Cervix Cancer Screening

IARC Press
2005
International Agency For Research On Cancer

The International Agency for Research on Cancer (IARC) was established in 1965 by the World Health Assembly, as an independently financed organization within the framework of the World Health Organization. The headquarters of the Agency are in Lyon, France.

The Agency conducts a programme of research concentrating particularly on the epidemiology of cancer and the study of potential carcinogens in the human environment. Its field studies are supplemented by biological and chemical research carried out in the Agency’s laboratories in Lyon and, through collaborative research agreements, in national research institutions in many countries. The Agency also conducts a programme for the education and training of personnel for cancer research.

The publications of the Agency contribute to the dissemination of authoritative information on different aspects of cancer research. Information about IARC publications, and how to order them, is available via the Internet at: http://www.iarc.fr/

This publication represents the views and opinions of an IARC Working Group on the Evaluation of Cancer Preventive Strategies which met in Lyon, France, Lyon, 20–27 April 2004
Contents

List of participants ........................................... ix

Preface ......................................................... xi

1. Cervical cancer and screening ............... 1
   Cervical cancer incidence and mortality worldwide ..................... 1
   Pathology of cervical neoplasia ................. 9
   Diagnosis and treatment of cervical preinvasive and invasive disease .......... 18
   The etiology of cervical cancer .............. 26
   Principles of screening ..................... 45
   Natural history of cervical cancer .......... 46
   Considerations for screening programmes .... 57

2. Screening tests ............................................ 59
   Cervical cytology ....................... 59
   Visual inspection ..................... 76
   Colposcopy ............................ 85
   Cervicography ....................... 90
   HPV DNA testing .................... 93
   Combinations of different modalities .... 109

3. Use of screening for cervical cancer ......... 117
   Delivery and uptake of screening .......... 117
   Behavioural considerations in screening .... 147

4. Efficacy of screening ...................... 163
   Methodology and analytical issues in assessment of efficacy ............. 163
   Cytological screening .................. 168
   Visual inspection ..................... 183
   Human papillomavirus testing .......... 186
   Other screening methods ............... 191
   Efficacy of screening among HIV-positive women .................. 196

5. Effectiveness of screening in populations .... 201
   Incidence and mortality trends in relation to screening ....................... 201
   Issues in the implementation of screening ........... 212
   Hazards of screening programmes .......... 214
   Performance evaluation .................. 217
   Economic evaluation and cost-effectiveness of cervical cancer screening .... 223

6. Summary of data ........................................ 227

7. Evaluation ............................................. 237

8. Recommendations for public health implementation and further research .... 239

References ............................................. 243

Glossary ................................................ 293

Working procedures ................................. 299
Note to the Reader

Anyone who is aware of published data that may influence any consideration in these Handbooks is encouraged to make the information available to the International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France

Although all efforts are made to prepare the Handbooks as accurately as possible, mistakes may occur. Readers are requested to communicate any errors to the IARC, so that corrections can be reported in future volumes.

Acknowledgements

We are very grateful for generous support received from the Bill and Melinda Gates Foundation, through the Alliance for Cervical Cancer Prevention (ACCP), that made this publication possible.
List of participants

A. Anttila
Finnish Cancer Registry
Institute for Statistical and Epidemiology Cancer Research
Liisankatu 21 B
00170 Helsinki
Finland

D. Aoki
Department of Obstetrics and Gynecology
Keio University
School of Medicine
35 Shinanomachi
Shinjuku-ku
Tokyo 160-8582
Japan

M. Arbyn
Scientific Institute of Public Health
European Network of Cervical Cancer Screening
J. Wytsmanstraat 14
1050 Brussels
Belgium

J. Austoker
Cancer Research UK
Primary Care Education Research Group
Division of Public Health
University of Oxford
Institute of Health Sciences
Old Road, Headington,
Oxford OX3 7LF
United Kingdom

X. Bosch**
Servei d'Epidemiologia
Institut Català d'Oncologia
Av. Gran Via s/n, km 2.7
08907 L'Hospitalet del Llobregat
Barcelona
Spain

Z.M. Chirenje
Department of Obstetrics and Gynaecology
University of Zimbabwe
PO Box A178
Avondale
Harare
Zimbabwe

J. Cuzick**
Cancer Research UK
Wolfson Institute of Preventive Medicine
Department of Epidemiology
Mathematics & Statistics
Charterhouse Square
London EC1M 6BQ
United Kingdom

N.E. Day (Chairman)
Institute of Public Health
Strangeways Research Laboratory
Wort's Causeway
Cambridge CB1 8RN
United Kingdom

L.A. Denny
Faculty of Health Sciences
Obstetrics and Gynaecology
Groote Schuur Hospital
Observatory 7925
Cape Town, Western Cape
South Africa

S. Fonn
School of Public Health
University of the Witwatersrand
7 York Road - Parktown
Braamfontein (PO Box 1038)
Johannesburg
South Africa

E. Franco**
Division of Cancer Epidemiology
McGill University
546 Pine Avenue West
Montreal QC, H2W 1S6
Canada

S. J. Goldie*
Department of Health Policy and Management
Harvard School of Public Health
718 Huntington Avenue, 2nd Floor
Boston MA 02115-5924
USA

T. Iftner**
Sektion Experimentelle Virology
Universitätsklinikum Tübingen
Elfriede-Aulhorn Strasse 6
72076 Tübingen
Germany

A. Kricker
Cancer Genes, Environment & Behaviour Program
School of Public Health
University of Sydney
Medical Foundation Building K25
92 Parramatta Road
Camperdown NSW
Australia

H. Lawson
Division of Cancer Prevention and Control
National Center for Chronic Disease Prevention
4770 Buford Highway N.E. MS-K57
Atlanta, GA 30341-3717
United States of America

* Unable to attend
**Invited specialist
E. Lynge
University of Copenhagen
Institute of Public Health
Blegdamsvej 3
2200 Copenhagen N
Denmark

L.D. Marrett
Division of Preventive Oncology
Cancer Care Ontario
620 University Avenue
Toronto, Ontario M5G 2L7
Canada

E. McGoogan**
University Medical School
Department of Pathology
Royal Infirmary of Edinburgh
51 Little France Crescent
Edinburgh EH164
United Kingdom

C.J. Meijer
Department of Pathology
Vrije Universiteit Medical Center
De Boelalaan 1117
POB 7057
1007 MB Amsterdam
The Netherlands

A.B. Miller (Vice-Chairman)
Department of Public Health Sciences
University of Toronto
Box 992
272 King Street
Niagara on the Lake
Ontario LOS 1JO
Canada

J. Patnick
NHS Cancer Screening Programme
The Manor House
260 Ecclesall Road South
S11 9PS Sheffield
United Kingdom

S.C. Robles
Pan American Health Organization
Program on Non-Communicable Diseases
525 23rd Street, N.W.
Washington, D.C. 20037
United States of America

G. Ronco**
Centro per la Prevenzione Oncol.
Piemonte
Azienda Sanitaria Ospedaliero S. G. Battista
Via S. Francesco da Paola 31
10123 Turin
Italy

M.H. Schiffman
Hormonal and Reproductive Epidemiology Branch
National Institutes of Health
Executive Plaza South
Room 7066
6120 Executive Blvd.
Rockville, MD 20892
United States of America

J.W. Sellers**
Program for Appropriate Technology
in Health
1455 NW Leary Way
Seattle WA 98107-5136
United States of America

A. Singer**
Whittington Hospital NHS Trust
Department of Women’s & Children’s Health
Highgate Hill
Jenner Building
London N19 SNF
United Kingdom

E.J. Suba
Kaiser Permanente Medical Center
1200 E1 Camino Real
South San Francisco, CA 904080
United States of America

T.C. Wright**
Department of Pathology
College of Physicians and Surgeons
Columbia University
Room 16-404, P&S Bldg
630 West 168th Street
New York, NY 10032
United States of America

Observers
N. Broutet
P. Claeyts
K. Shapiro
A. Ullrich
WHO, Geneva, Switzerland

Secretariat
S. Arrossi
F. Bianchini (Co-responsible officer)
F. Bray
J. Cheney (Editor)
V. Cogliano
G. Clifford
S. Franceschi
M. Hakama (Responsible officer)
A. Kreimer
A. Loos
C. Mahé
D.M. Parkin
P. Pisani
R. Sankaranarayanan
A. Sasco
B. Secretan
K. Straif
M. Tommasino
H. Vainio (Head of Programme)
S. Vaccarella

Post-meeting scientific assistance
F. Bianchini
M. Hakama

Technical assistance
J. Mitchell
A. Rivoire
J. Thévenoux

** Invited specialist
Preface
Why a Handbook on Cervix Cancer Screening?

Cervix cancer is an important public health problem. It is the third cancer in frequency in women worldwide and the most or second most common cancer among women in developing countries. The conventional screening modality for cervical squamous intraepithelial lesions is the cytological test, or Pap smear. This was introduced as a routine screening modality in much of Europe, North America and Oceania without formal evidence on efficacy from randomized trials. However, data on time trends in countries with centrally organized programmes that started in the 1960s and 1970s have provided convincing evidence that cervical cytology screening, by the identification and treatment of preinvasive lesions, can prevent a large proportion of invasive cervical cancers.

In 1985, the IARC, in collaboration with the UICC, published a monograph on cervical cancer screening, which included a detailed analysis of the effectiveness of different screening policies, including the frequency of screening and the age at which it should start. That volume has been widely used, particularly in Europe, to define national screening policy. Since 1985, there have been two notable advances. The most important is the identification of certain oncogenic types of human papillomavirus (HPV) as the major cause of cervical cancer; indeed it may be that the disease does not occur in the absence of HPV infection. With the development of vaccines against these oncogenic HPV types, it is becoming possible to envisage the primary prevention of most cases of cervical cancer. It will be several decades, however, before most women in the relevant age groups will benefit from such vaccines, since they will already have been at risk of exposure to the virus. Of more immediate potential benefit is the role of high-technology HPV testing in screening, either as an adjunct to cervical cytology or as the primary screening modality. The second advance has been the development of low-cost, low-technology cervical screening modalities. These may be appropriate as alternatives to cytology in many developing countries that have a high incidence of cervical cancer and limited infrastructure and health-care resources, as well as other competing health priorities. Furthermore, in the 20 years since the earlier monograph, the pattern of cervical cancer and its precursor lesions has changed in many countries, with rapidly increasing incidence in younger age groups and some evidence that the natural history may be age-dependent. Such age-dependence could have implications for screening policies.

The purpose of this Handbook is to consider the implications for cervical cancer screening of the advances that have been made over the past 20 years, and of the changing patterns of cervical cancer incidence. In particular, it gives an evidence-based critical evaluation of the efficacy and effectiveness of the modalities currently available for cervical cancer screening, and of their relative appropriateness depending on the resources available and competing priorities. Further aims are to provide recommendations for the public health implementation of screening, including the frequency of screening and the age groups that should constitute the target population, and to identify areas for further research.

Public health authorities in middle- and low-income countries are following closely the debate on the role of new screening technologies. Vaccination against HPV infection for primary prevention of cervical cancer opens a new avenue for control of cervix cancer. Between the fear of increased healthcare costs associated with the adoption of new technologies or boosting available efforts on the one hand and the promising results coming from the research front on HPV vaccines on the other, it is tempting to take a wait-and-see attitude concerning cervical cancer prevention. This posture could lead to decreased funding for cervical cancer screening in the false hope that HPV vaccines will be available soon to eradicate the disease. This scenario could prove disastrous by abolishing the hard-earned gains made in the last half century through cytological screening in reducing cervical cancer morbidity and mortality. Prophylactic vaccines offer great hope for future generations, but women who have initiated sexual intercourse will largely have to rely on screening for the prevention of cervical cancer for the foreseeable future.
Cervical cancer incidence and mortality worldwide

This section describes geographical patterns in cervical cancer incidence, survival and mortality, and the association of disease risk with classical demographic variables. Time trends in incidence and mortality, and the influence of screening programmes in determining them, are covered in Chapter 5.

In examining differences in risk of disease between populations and over time, it is best to use, whenever possible, data on cancer incidence. However, mortality data are, in general, more widely available and cover longer periods of time. The use of mortality data as a substitute for incidence data is based on the assumption that the ratio of incident cases to deaths—as expressed by survival—is more or less the same in the populations being considered (including, for study of trends, over time). The section below on survival gives an indication of the validity of this assumption.

In some studies, mortality, in terms of numbers of deaths or probability of death, may actually be the focus of interest, for example in comparing overall cancer burden or the combined result of all cancer control interventions (including early diagnosis and therapy). In this context, variables that take into account age at death (person-years of life lost) and the level of disability between diagnosis and death (quality-adjusted life years, disability-adjusted life years) have become more widely applied in health-care planning and evaluation.

Some methodological and data considerations

International comparative studies using the indicators summarized above depend upon assumptions about lack of bias arising from data-quality issues. Cancer incidence data, published in the Cancer Incidence in Five Continents series (Parkin et al., 2002) have been peer-reviewed for data quality and completeness of coverage of the population at risk. The mortality data available through the WHO statistical information system (http://www3.who.int/whosis/menu.cfm), based on information received from national statistical offices, may be biased by different practices in death certification, and, for some countries, may be quite incomplete, as far as population coverage is concerned. These sources of bias should be checked, using the tables showing estimated completeness of coverage (http://www3.who.int/whosis/mort/table3.cfm?path=whosis,inds,mort,mort_table3&language=english), before the rates are used for comparisons between populations or over time.

These caveats apply to all comparative studies, but two issues are specific to studies of cancer of the cervix. The principal one relates to the categories in the international classification of disease (ICD) which has, since its 7th edition (1955), provided for the coding of cancers of the uterus as ‘Cervix’, ‘Corpus’ or ‘Uterus, part unspecified’. The proportion of uterine cancer cases and deaths ascribed to the third of these categories, generally referred to as ‘not otherwise specified’ (NOS), varies widely both between countries and over time. The problem is much worse for mortality statistics than for incidence. The NOS category comprises more than 10% of uterine cancers in less than 10% of the cancer registries reporting in Cancer Incidence in Five Continents (Parkin et al., 2002). For mortality, in contrast, the proportion of deaths certified as due to cancer of ‘Uterus NOS’ can be very high—well over 50% in France and Italy, for example, in 1995 for women aged over 30 years (http://www-depdb.iarc.fr/who/menu.htm). As a result, comparative studies using data without correction for NOS may yield highly misleading assessments of geographical (Figure 1a) and temporal differences (Figure 1b) in mortality. For example, much of the apparent increase in the mortality rate from cancer of the cervix in Spain is due to a reduction in the rate for ‘Uterus NOS’ through better specification of cause of death (Figure 1b). Before comparative studies can be performed, therefore, some form of ‘reallocation’ of these NOS cases and deaths to the more specific categories is required. Several methods have been proposed (Arbyn & Geys, 2002; Bray et al., 2002). When the percentage of NOS cases is...
Figure 1 (a) Mortality from cancer of the cervix uteri in nine European countries, 1998; (b) Trends in mortality from cancer of the uterus, Spain
From http://www-depdb.iarc.fr/who/menu.htm
relatively small (< 25%, say), this reallocation can be according to the proportions of cases in the series with specified site, by age group. When a larger fraction of the cases have the precise site missing, it is preferable to use proportions from a different (reference) population which has data of better quality and is thought to be epidemiologically similar.

The second issue specific to cervical cancer epidemiology is that incidence and mortality rates are calculated using the entire female population as the ‘population at risk’, although women who have had a total hysterectomy for reasons other than the presence of cervical neoplasia are not at risk for the disease. Such women ought to be excluded from the population at risk, but the prevalence of hysterectomy is generally not known. Hysterectomized women may constitute a sizeable proportion of the population in some age groups and countries and this proportion may vary over time as well as by place and age. For example, in Ontario, Canada, the incidence of hysterectomy reached a peak in the early 1970s and then declined until 1990 (Snider & Beauvais, 1998). Rates were greatest in women aged 40–44 years. The self-reported prevalence of hysterectomy in 1994 varied from 13% to 28% of women aged 15 years and over by region of Canada; overall, 30% of women had had a hysterectomy by the time they attained age 65 (Snider & Beauvais, 1998). In England and Wales, the prevalence of hysterectomy was estimated as 21.3% at ages 55–59 in 1995 (Redburn & Murphy, 2001). Correction of the population at risk could therefore have a substantial impact on the estimated incidence and mortality rates, especially in older age groups, although little effect on the observed trends in Ontario (Marrett et al., 1999) and England and Wales (Redburn & Murphy, 2001) was seen.

Cervical cancer: world patterns
Cancer of the cervix uteri is the second most common cancer among women worldwide, with an estimated 471 000 new cases (and 233 000 deaths) in the year 2000 (Parkin et al., 2001). Almost 80% of the cases occur in developing countries, where, in many regions, it is the most common cancer among women, responsible for about 15% of all new cancers. The highest incidence rates are observed in Latin America and the Caribbean, sub-Saharan Africa, and south and south-east Asia (Figure 2). Cervical cancer is less common in the developed countries, where it was estimated to comprise about 4%

![Figure 2 Incidence of cancer of the cervix uteri](From Ferlay et al. (2001))
of cancers in women in the year 2000, ranking sixth in importance.

Figure 3 shows incidence rates recorded in cancer registries around 1995 (Parkin et al., 2002). These rates vary by at least 20-fold. The lowest (less than 14 per 100 000) are, in general, found in Europe (excluding some eastern European countries), North America and China. Incidence is generally higher in the developing countries of Latin America (average age-standardized incidence rates [ASR], 33.5 per 100 000) and the Caribbean (ASR, 33.5), sub-Saharan Africa (ASR, 31.0) and south-central (ASR, 26.5) and southeast Asia (ASR, 18.3) (Ferlay et al., 2001). Very low rates are observed in China and in western Asia (Figure 2); the lowest recorded rate is 0.4 per 100 000 in Ardabil, north-west Iran (Sadjadi et al., 2003).

This pattern is relatively recent, however; before the introduction of screening programmes in the 1960s and 1970s, the incidence in most of Europe, North America and Japan was similar to that seen in many developing countries today (Gustafsson et al., 1997b): for example, it was 38.0 per 100 000 in the Second National Cancer Survey of the USA (Dorn & Cutler, 1959), 37.8 per 100 000 in Hamburg, Germany, in 1960–62, 28.3 per 100 000 in Denmark in 1953–57 and 22.1 per 100 000 in Miyagi, Japan, in 1959–60 (Doll et al., 1966).

The majority of cases of cervical cancer are squamous-cell carcinomas.

---

**Figure 3** Age-standardized incidence rates (per 100 000) in selected cancer registry populations, around 1995, and the percentage of registered cases (of known histology) that are adenocarcinomas. Incidence rates by histological type were estimated by reallocating cases without specified histology (<10% of the total cases) to the three histological subtypes shown, according to observed proportions, by age group. From Parkin et al. (2002)
In general, the proportion of adenocarcinoma cases is higher in areas with a low incidence of cervical cancer (Figure 3). This probably relates to the presence of screening programmes, since cytological screening has, at least in the past, probably had little effect in reducing the risk of cervical adenocarcinoma (see Chapter 4). Adenocarcinomas occur within the cervical canal (from the glandular epithelium) and are not readily sampled by scraping the epithelium of the ectocervix (Fu et al., 1987; Sigurdsson, 1995); a case–control study (Mitchell et al., 1995) suggested that the risk of invasive adenocarcinoma was not reduced by screening.

There were an estimated 233 000 deaths from cervical cancer worldwide in 2000, 83% occurring in lower-resource areas, where this is the most common cause of cancer death (Parkin et al., 2001). While mortality rates are much lower than incidence rates (the worldwide ratio of mortality to incidence is 49%), they correlate rather well with incidence patterns.

**Demographic determinants of risk**

It was noted at an early date that cervical cancer has quite marked differences in incidence according to classical demographic variables (age, social class, marital status, ethnicity, religion, occupation). Later, epidemiological studies (mainly case–control studies) showed consistent associations between risk and early age at initiation of sexual activity, increasing number of sexual partners of females or of their sexual partners, and other indicators of sexual behaviour (Muñoz et al., 1992a,b). The part played by sexual behaviour of male partners in increasing risk was also the focus of interest in areas such as Latin America where cervical cancer is frequent, and where the median number of sexual partners of men is much greater than that of women, who are largely monogamous (Brinton et al., 1987, 1989a,b). These findings strongly suggested a causative role for a sexually transmitted agent. The development of methods for detecting the deoxyribonucleic acid (DNA) of HPV in tissues allowed the central role of this virus in the etiology of cervical cancer to be confirmed (IARC, 1995) (see section on Etiology in this chapter).

It is likely that the observed associations of the classical demographic variables with risk of cancer of the cervix are largely the result of differences in exposure (and possibly response) to HPV, as well as to differences in patterns of screening. This can be investigated in analytical studies, where the independent effects can be investigated. Although of little interest from an etiological point of view, these demographic variables remain useful in a health services context, for example in monitoring the use of screening programmes.

The general form of the curve of incidence versus age shows a rapid rise to a peak usually in the fifth or sixth decade, followed by a plateau and a variable decline thereafter (Figure 4). This pattern with an early
peak or plateau in risk is unique for an epithelial cancer, and reflects the natural history of infections with human papillomavirus (HPV) and the related carcinogenic mechanisms. The age profile is readily distorted by screening and also, when cross-sectional data (from a single time period) are examined, by birth-cohort-specific changes in risk (Ashley, 1966; Hakama & Penttinen, 1981). In an attempt to define the age-specific incidence patterns of cervical cancer in the absence of screening activity, Gustafsson et al. (1997b) compiled incidence data for 28 different populations for long periods of time between 1920 and 1989. After scaling (to permit direct comparison between countries with incidence rates of differing orders of magnitude), the rates for most populations fitted one of two reference curves used for descriptive purposes (Figure 5). The first group (green line), comprising Denmark, the former German Democratic Republic, the Federal Republic of Germany (before reunification), the Netherlands, Norway, Slovenia and Sweden, had an onset at about age 25, a rapid rise between 30 and 40 and a peak at ages 44–49 years. After the peak, the decline was fairly rapid, falling to half the maximum (the half peak value) at 70–75 years. The second group (blue line), comprising most American, Asian and African registries, plus Finland and Poland, had onset at about the same age but a slower rise to a peak at ages 50–65, followed by a decline similar to that in the first group. Data from the United Kingdom and China did not fit these curves. For the United Kingdom, this is almost certainly the result of long-term variation in risk by birth cohort (Hill & Adelstein, 1967; Osmond et al., 1982), while in China it is probably due to a low level, and late onset, of exposure to etiological factors, especially HPV (IARC, 1995). Analysis of temporal changes in the curves for the Nordic countries revealed shifts in the peak incidence with time towards earlier ages. This is also probably an effect of increasing risk among successive birth cohorts, since cross-sectional analysis of age-specific incidence showed that a 3% annual increase in successive birth cohorts would move the shape of the curves for the second group of countries towards the shape seen for the first group (Gustafsson et al., 1997b). This adds further weight to the other evidence that strong cohort effects exist that need to be taken into account in any analysis of incidence with respect to time.

One of the earliest observations in cancer epidemiology was the rarity of cancer of the cervix among (unmarried) nuns (Rigoni-Stern, 1842), an observation that has been confirmed more recently (Fraumeni et al., 1969). Risk is higher in women who are divorced or separated than in married women. The risk of cervical cancer is especially high among women marrying at young ages (Jones et al., 1958; Boyd & Doll, 1964). These associations are related to other aspects of sexual behaviour such as number of sexual partners and age at initiation of intercourse (Terris et al., 1967).

Women of lower socioeconomic status (defined by, for example, income, educational level or housing type) are at higher risk for cervical cancer (de Sanjose et al., 1997). HPV infection appears to be more prevalent in women of lower educational and income levels (Hildesheim et al., 1993; Varghese, 2000). Other correlates of social status such as nutrition, genital hygiene, parity, smoking, other genital infections and use of preventive ser-
vices (especially screening) may be responsible for the observed differences. Varghese (2000) found a significant association between social status and HPV infection in India, and social status remained a determinant of HPV infection even after adjustment for promiscuity.

In a review of data from the US Surveillance Epidemiology and End Results (SEER) programme for 1988–92, Miller et al. (1996) noted the highest incidence of cervical cancer among Vietnamese, with a rate some 7.4 times that in Japanese women (Figure 6). The incidence in black women was about 1.5 times that in whites. At least part of the racial differences is explicable by differences in terms of socioeconomic indicators, such as income and education; when adjustment is made for such factors, the black–white differences are greatly diminished (Devesa & Diamond, 1980). Other examples of striking differences between ethnic groups living in the same environment are the white and black populations of Harare, Zimbabwe (Bassett et al., 1995), and the Chinese, Indian and Malay populations of Singapore (Lee et al., 1988).

Certain religious groups in the USA, such as the Amish (Cross et al., 1968) and Mormons (Lyon et al., 1980), have been reported to have relatively low risks for cervical cancer compared with the general population. Jewish populations have also been noted to have lower risk than other religious groups among whom they reside (Boyd & Doll, 1964). Quite marked differences in incidence have been reported between different religious communities in Mumbai (Bombay), India (Figure 7) (Jussawalla & Yeole, 1984). The extent to which these different cancer risks reflect prevalence of HPV infection has not been studied.

Figure 6 Incidence rates of cancer of the cervix uteri in the US SEER programme for 1988–92
From Miller et al. (1996)

Figure 7 Incidence rates (per 100 000) of breast cancer and cancer of the cervix uteri among religious groups in Mumbai, India
From Jussawalla et al. (1981); Jussawalla & Yeole (1984)
High rates of cervical cancer have been reported among prostitutes (Rojel, 1952; Moghissi & Mack, 1968). Job/branch categories with excess relative risks for cervical cancer observed in studies using cancer registries or death certificates include hotel and restaurant personnel and waitresses (Williams et al., 1977; Kjaerheim & Andersen, 1994; Pukkala, 1995), maids, nurses’ aids (Sala et al., 1998), cleaners and cooks (Bulbulyan et al., 1992; Pukkala, 1995; Alterman et al., 1997) and woodworkers (Hall & Rosenman, 1991; Pukkala, 1995; Weiderpass et al., 2001). Exposure to various solvents has been found to be associated with increased risk (Blair et al., 1979; Berlin et al., 1995; Weiderpass et al., 2001). Women in agriculture seem to be at increased risk in some settings (Stubbs et al., 1984; Blair et al., 1993; McDuffie, 1994), but at decreased risk in others (Andersen et al., 1999). A twofold increase in risk for cervical cancer in workers exposed to multiple pesticidal agents has been reported (Wesseling et al., 1996). There are also associations with occupations of husbands and partners, specifically those necessitating prolonged absences from home (Beral, 1974).

### Survival and cancer control

Information on survival has long been recognized as an important indicator in monitoring cancer control activities (WHO, 2002), although it is not an adequate indicator of the effectiveness of cancer control on its own, but must be considered in context, together with incidence and mortality (Welch et al., 2000). Survival is usually studied to evaluate the effectiveness of treatment for cancer, and indeed, the availability and accessibility of high-quality treatment has a major influence on patient survival. It should be remembered, however, that population-based survival statistics from cancer registries reflect the outcome of the totality of cancer patients, including those who receive no treatment whatsoever. They are therefore the average result of the whole range of cancer-control activities, including screening and the organization of treatment services (Black et al., 1998). Estimates of survival in different populations may be influenced by a range of prognostic and other factors. Some prognostic factors, such as age and sex, are always available, and usually so too are tumour-related variables such as sub-site and histological type.

Stage of disease at diagnosis is generally the most important factor determining the survival of cancer patients, so that variations in the stage distribution of tumours among populations being compared are of particular concern. Table 1 shows a comparison of five-year relative survival, by stage, from several population-based series.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Country, period</th>
<th>Stage of cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ries et al. (2003)</td>
<td>USA: SEER (white), 1992–99</td>
<td>93  52  17</td>
</tr>
<tr>
<td>Dickman et al. (1999)</td>
<td>Finland, 1985–94</td>
<td>84  49  28</td>
</tr>
<tr>
<td>Carstensen (1993)</td>
<td>Denmark, 1978–87*</td>
<td>81  38  8</td>
</tr>
<tr>
<td>Yeole et al. (1998)</td>
<td>Mumbai, India, 1982–86</td>
<td>77  35  6</td>
</tr>
</tbody>
</table>

* Crude survival

Although an improvement in survival from the cancer of interest is considered to be a necessary but non-sufficient indicator of the success of a cancer screening programme, an effective cervical cancer screening programme may, paradoxically, have the opposite result. Thus, in Finland, Dickman et al. (1999) observed that, although survival improved over time between 1955 and 1994 for almost all cancers, cervical cancer was an exception; for this site, survival decreased slightly from 1965–74 to 1985–94. This is because, when overall incidence decreases, due to screening, a greater proportion of cases are advanced cancers in women who have not participated in the screening programme. It is possible, too, that interval cancers may represent a length-biased subset of more aggressive tumours that were not detected by screening in preinvasive or early invasive stages.

There are two related approaches to the estimation of survival: the Kaplan–Meier and actuarial, or life-table, methods (Berkson & Gage, 1950; Kaplan & Meier, 1958). The former is particularly useful when exact survival times are available, since
smooth estimates of survival as a function of time since diagnosis can be obtained. The actuarial method requires a life-table with survival times grouped usually into intervals that permit calculation of the cumulative probability of survival at time \( t \), from the conditional probabilities of survival during consecutive intervals of follow-up time up to and including \( t \). 'Observed survival' is influenced not only by mortality from the cancer of interest, but also by deaths from other causes. Relative survival takes into account the risk of death from causes other than the cancer under study (Ederer et al., 1961). For comparisons between different populations, a further standardization of survival by age is necessary.

**Factors influencing survival**

Survival of cervical cancer patients varies by age. In the EUROCARE-3 study (Sant et al., 2003), for example, relative survival of cases aged 15–44 years at diagnosis (74% at five years) was more than twice that of women who were aged 75 or more (34%), with a clear decreasing trend in survival with increasing age. The difference may be related to biological factors (tumour growth) or be the result of the higher prevalence of co-morbid disease such as hypertension and cardiovascular disease in the elderly, making the patient less likely to receive optimal treatment, or to have a favourable result from it.

Kogevinas and Porta (1997) summarized the results of ten studies that examined social class differences in survival from cancer of the cervix. In eight of these, patients of lower social class had poorer survival than those in high classes, although the differences were not great. The differences may relate to timing of diagnosis (patients of lower social class present later), in treatments applied, in the biological characteristics of the neoplasm, or in host factors. Staging procedures may be less intensive in patients of lower social class, so that there may be misclassification of more advanced cancer to earlier-stage disease. The life-tables (all-cause mortality) used to calculate relative survival only seldom allow for differences in competing causes of death between social classes. In general, however, it is thought that this is not an important source of error.

**International comparisons of survival**

Survival statistics for various periods from cancer registries in developed countries such as the USA, Canada, European countries, Japan and Australia have been published (Hakulinen et al., 1981; Berrino et al., 1995; Inoue et al., 1998; Berrino et al., 1999; Ries et al., 2003; Sant et al., 2003). Data on cancer survival from developing countries were sparse until 1995 (Nandakumar et al., 1995; Sriamporn et al., 1995). Sankaranarayanan et al. (1981) summarized survival data from several registries in developing countries, and more recently, the first data from Africa have become available (Wabinga et al., 2003; Chokunonga et al., 2004).

Five-year relative survival rates vary between regions, with quite good prognosis in low-risk regions, but even in developing countries, where many cases present at relatively advanced stage, survival rates are fair: 49% on average (Sankaranarayanan et al., 1998a).

**Time trends in survival from cancer of the cervix**

In the first half of the 20th century, there were major improvements in survival from cancer of the cervix, due in part to improving stage at diagnosis, and in part to better results of treatment within stage, particularly as a result of advances in radiotherapy (Pontén et al., 1995; Sparén et al., 1995). In most developed countries, there has, however, been little change in survival in recent decades. In Denmark, for example, five-year relative survival was 61.3% in 1958–62 and 63.9% in 1983–87 (Carstensen, 1993); in the USA, survival was 69.1% in 1974–76 and 71.3% in 1992–99 (Ries et al., 2003).

Figure 8 shows time trends in relative survival for nine populations (Chia et al., 2001). The series from Europe, the USA and Japan show little or no improvement in survival, while there has been a moderate improvement in Singapore, from 46% in 1968–72 to 63% in 1988–92. The relatively unfavourable trends in survival may be the result of a counterbalance between the effect of screening and improvements in treatment, as mentioned above. With the success of screening, the lesions that are diagnosed as invasive cancer between screenings will be those that are more aggressive and associated with poor survival.

---

**Pathology of cervical neoplasia**

The objective of cervical cancer screening programmes is to reduce the mortality from (and incidence of) the disease by identifying women with precancerous cervical lesions and early invasive cancers, and treating these women appropriately. Precancerous lesions are defined biologically as lesions that have, in principle, a capacity to progress potentially to invasive cervical cancer if left untreated. They are strongly associated with both morphological cellular changes and specific high-risk types of HPV, and continued expression of HPV-derived oncoproteins (e.g., E6 and E7) results in unregulated cellular proliferation. Phenotypically, precancers are characterized by intracellular high-risk HPV DNA, chromosomal instability with resulting aneuploidy, and monoclonality. Morphological appearances alone...
often do not allow distinction of precur-
sor lesions that have a substantial
capacity to progress from those
lesions that do not, contributing to
uncertainty for both clinicians and epi-
demiologists. Nevertheless, until more
precise methods are developed for use
in day-to-day settings, histological
appearance remains the basis for the
definition of both precancerous and
cancerous cervical lesions.

Intraepithelial squamous
lesions

Terminology

The uterine cervix is the cylindrically
shaped lower third of the uterus that
extends into the vagina. The cervix has
a central opening or external os that
opens into the endocervical canal
which communicates with the uterine
cavity (Figure 9). The cervical epithe-
lium is derived from two embryologi-
cally distinct sources. The part of the
cervix that projects into the vagina,
called the ectocervix or portio, is cov-
ered by non-keratinized stratified squa-
mous epithelium similar to that of the
vagina. This stratified squamous
epithelium is derived from the urogeni-
tal sinus. In contrast, the endocervical
canal is covered by tall, mucus-secret-
ing columnar cells that are of Müllerian
origin. The junction between these two
epithelia is termed the squamocolumn-
ar junction. The squamocolumnar
junction is not fixed anatomically, but
migrates throughout life. At the time of
puberty, it is usually positioned towards
the periphery of the ectocervix and
with age, it migrates inward towards
the external os (Figure 10). This migra-
tion occurs in large part by a process
termed squamous metaplasia, in
which the columnar endocervical-type
epithelium is replaced by a stratified
squamous epithelium. The area of the
cervix where this transformation from
columnar epithelium to stratified squa-
mous epithelium takes place is
referred to as the transformation zone
(Figure 10). The metaplastic area adja-
cent to the receding squamocolumnar
junction has, for unknown reasons, a
unique susceptibility to HPV-induced
neoplastic transformation, particularly
in the anterior and posterior areas.
These are the areas where most squa-
mous-cell carcinomas of the cervix
develop.
Cervical cancer and screening

Cervical cancer and intraepithelial lesions that develop in the transformation zone can be visualized by colposcopy and diagnosed by histological examination of colposcopy-directed biopsies of areas that appear abnormal.

It is now generally accepted that squamous and glandular neoplasms of the cervix are caused by infection of cervical epithelium by specific HPV types (Bosch et al., 1995; Muñoz et al., 2003). HPV infection is associated with a wide spectrum of histological appearances, some of which may be readily identified by routine light microscopy. Terminology used to classify these cellular changes has undergone periodic revision to incorporate advances in the scientific and clinical understanding of cervical neoplasia.

At least three separate, but for the most part interchangeable, histopathological classifications are currently in use (Table 2). All recognize that persistent HPV infection of cervical squamous epithelium leads to two categories of intraepithelial squamous lesions: productive, self-limited HPV infections, and those with the potential, if left untreated, to progress to invasive squamous-cell carcinoma (Wright et al., 2002b). Biopsies of productive HPV infections of the cervix have been classified as koilocytotic atypia, koilocytosis, condyloma, mild dysplasia, cervical intraepithelial neoplasia grade 1 (CIN 1) and low-grade squamous intraepithelial lesion (LSIL). CIN 1 lesions are heterogeneous with respect to their associated HPV types, clonality and ploidy status. The lesions can be associated with any of the anogenital HPV types, can be either monoclonal or polyclonal, and are aneuploid in only about one third of cases (Fu et al., 1983; Lungu et al., 1992; Park et al., 1996; Hering et al., 2000). They tend to be transient and are unlikely to act as cervical cancer precursors. Lesions more likely to represent cervical cancer precursors have

Figure 9 Gross anatomy of the uterine cervix
From Sellors & Sankaranarayanan (2003)
been classified as moderate dysplasia, severe dysplasia, CIN 2, CIN 3, carcinoma in situ, and high-grade squamous intraepithelial lesion (HSIL). CIN 2 and CIN 3 lesions are usually associated with high-risk types of HPV, are monoclonal and are usually aneuploid (Fu et al., 1983; Lungu et al., 1992; Park et al., 1996; Hering et al., 2000).

The designation carcinoma in situ was almost invariably used for full-thickness lesions of the uterine cervix by authors who adhered to the early WHO classification (Riotton et al., 1973). This was reflected in the early studies of the natural history of cervical cancer (see later in this chapter) and in the cases reported to cancer registries. Following Richart’s (1980) description of the cervical intraepithelial neoplasia (CIN) terminology, there was an increasing tendency to include cases referred to earlier as carcinoma in situ within the CIN 3 designation; this tendency accelerated when the Bethesda System was introduced (National Cancer Institute, 1989). Thus, while most authors continue to use the CIN 3 designation for histological diagnoses, the carcinoma in situ designation has now almost completely disappeared. Because CIN 3 combines severe dysplasia, which has a defined probability of regression, with carcinoma in situ, which regresses less, care is required in comparing the findings from earlier studies that used the term carcinoma in situ with more recent studies that have not.

The traditional dysplasia/carcinoma in situ and CIN classifications recognize that intraepithelial squamous lesions of low, intermediate and high risk for progression to invasive cervical cancer can be identified and attempt to stratify these lesions accordingly. However, inter-observer and intra-observer studies consistently document poor reproducibility of the distinction between CIN 2 and CIN 3.

Figure 10 Location of the squamocolumnar junction (SCJ) and transformation zone: (a) before menarche; (b) after puberty and at early reproductive age; (c) in a woman in her 30s; (d) in a perimenopausal woman; and (e) in a postmenopausal woman
From Sellors & Sankaranarayanan (2003)
(Ismail et al., 1989; Price et al., 2003). Many pathologists report histopathological diagnoses using more than one classification scheme. In this Handbook, the CIN terminology is used when referring to specific histopathological entities except when directly reporting studies that used different terminology.

Pathological findings
Intraepithelial squamous lesions are characterized by abnormal cellular proliferation and maturation together with nuclear atypia. Neither ultrastructural nor immunohistochemical studies currently contribute greatly to the routine diagnosis of intraepithelial squamous lesions. The microscopic alterations that comprise intraepithelial lesions are semi-quantitatively classified into three categories. The grading of CIN lesions is prone to high rates of intra-observer and inter-observer variability (Ismail et al., 1989, Robertson et al., 1989a; Stoler & Schiffman, 2001). Inter-observer agreement is higher among CIN 3 lesions and lower among lower-grade lesions (Stoler & Schiffman, 2001). However, despite the poor reproducibility of a diagnosis of a given grade of CIN, separation of CIN into three subcategories (e.g., CIN 1, CIN 2, CIN 3) correlates to a general extent with rates of progression and of regression of the lesion (Mitchell et al., 1996). With regard to microscopic morphological interpretation, poor reproducibility does not exclude accuracy (Renshaw, 2003).

**CIN 1 (flat condyloma; koilocytosis; mild dysplasia):** Neoplastic, basaloid cells and mitotic figures occupy the lower third of the epithelium in CIN 1 lesions. These lesions frequently show marked HPV cytopathic effects including perinuclear halos, multinucleation and nuclear membrane irregularities, and hyperchromasia (e.g., "koilocytosis") (Figure 11). Pathologists make frequent errors when attempting to distinguish reactive squamous proliferations from the HPV-induced lesions comprising this category. The most common error made in this category of lesions is ‘overcall’ of non-specific inflammatory or reactive lesions as productive HPV infections. In the National Cancer Institute's ASCUS-LSIL Triage Study (ALTS), 45% of biopsies initially classified as CIN 1 were downgraded to non-CIN when reviewed by a panel of expert gynaecological pathologists (Stoler & Schiffman, 2001). In particular, perinuclear haloes in the absence of significant nuclear atypia have been documented to be non-specific reactive features (Mittal et al., 1990).

**CIN 2 (moderate dysplasia):** In CIN 2, neoplastic basaloid cells and mitotic figures occupy the lower two thirds of the epithelium (Figure 12). Although CIN 2 lesions usually show somewhat less HPV cytopathic effect than do CIN 1 lesions, koilocytes are often still identified in the epithelium. Distinction between CIN 2 and both CIN 1 and CIN 3 in biopsy specimens is complicated by the fact that the thickness of the epithelium occupied by neoplastic basaloid cells and mitotic figures often varies greatly within any given cervical biopsy specimen, while variations in the angle at which the epithelium has been cut during histological sectioning can also have an effect (Wright et al., 2002b).

**CIN 3 (severe dysplasia; carcinoma in situ):** The characteristic histological feature of CIN 3 is the presence of neoplastic basaloid cells and mitotic

---

Table 2. Grading schemes for preinvasive histological abnormalities of uterine cervical squamous epithelium

<table>
<thead>
<tr>
<th>Dysplasia classification system</th>
<th>Cervical intraepithelial neoplasia (CIN)</th>
<th>Bethesda classification system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild dysplasia</td>
<td>CIN 1</td>
<td>LGSIL</td>
</tr>
<tr>
<td>Moderate dysplasia</td>
<td>CIN 2</td>
<td>HGSIL</td>
</tr>
<tr>
<td>Severe dysplasia</td>
<td>CIN 3</td>
<td>HGSIL</td>
</tr>
<tr>
<td>Carcinoma in situ</td>
<td>CIN 3</td>
<td>HGSIL</td>
</tr>
</tbody>
</table>

---

**Figure 11** Cervical intraepithelial neoplasia 1 (CIN 1).
The upper two thirds of the epithelium shows maturation and focal koilocytosis. There is a mild atypia throughout. From Tavassoli & Devilee (2003)

**Figure 12** Cervical intraepithelial neoplasia 2
Nuclear abnormalities are more striking than in CIN 1 and mitoses are seen (centre). The upper third of the epithelium shows maturation. From Tavassoli & Devilee (2003)
figures that occupy the full thickness of the epithelium. These cells have high nuclear:cytoplasmic ratios, with scant cytoplasm and dense, hyperchromatic nuclei having coarse clumped chromatin and irregular nuclear outlines (Figure 13). Although inter-observer variability among pathologists is moderate for histopathological diagnosis of CIN 2 and CIN 3 (Robertson et al., 1989a; Stoler & Schiffman, 2001), overcall and undercall errors are not uncommon. Immature metaplasia (Crum et al., 1983), atrophy and reparative processes are lesions without risk for progression to carcinoma that may be misinterpreted as CIN 2 and CIN 3. The distinction between CIN 2 or CIN 3 and atrophy in a postmenopausal patient can sometimes be established only after a repeat biopsy is taken after estrogen has been used to stimulate maturation of the cervical epithelium. Topical estrogen treatment induces maturation in atrophic cervical epithelium, but does not change the appearance of high-grade preinvasive lesions. In the future, immunohistochemical staining for various biomarkers such as p16 may be routinely usable to help distinguish CIN from its mimics. CIN 2 and CIN 3 lesions associated with extensive gland involvement may be confused with microinvasive squamous-cell carcinoma, resulting in overcall error.

**Intraepithelial glandular lesions**

**Terminology**

Adenocarcinoma *in situ* (AIS) is the only well characterized preinvasive glandular lesion of the uterine cervix; it is much less common than its squamous counterparts. The US SEER database recorded 72,357 *in situ* cervical cancers with histology records between 1973 and 2001 (National Cancer Institute, 2004), of which only 2% were AIS. Termination for intraepithelial glandular lesions with lower degrees of nuclear atypia and mitotic activity than AIS has been proposed; the proposed terms include *endocervical dysplasia, cervical intraepithelial glandular neoplasia* and *endocervical glandular atypia* (Bousfield et al., 1980; Gloor & Hurlimann, 1986; Ayer et al., 1987; Wright et al., 2002b). Because of the rarity of biopsy-documented non-AIS preinvasive glandular lesions, the utility of non-AIS terminology has not been established.

Nearly two thirds of cases of AIS have coexisting preinvasive squamous lesions or invasive squamous-cell carcinoma (Colgan & Lickrish, 1990; Denehy et al., 1997) and risk factors for AIS are similar to those for preinvasive squamous lesions (Ursin et al., 1996). Because no natural history studies of AIS have been published, the evidence that AIS is the precursor lesions for invasive endocervical adenocarcinoma remains circumstantial (Wright et al., 2002b). Like CIN 2 and CIN 3, AIS is associated with persistent infection with high-risk types of HPV (Tase et al., 1989; Duggan et al., 1994).

**Pathological findings and related errors**

AIS is characterized microscopically by replacement of the glandular cervical epithelium by cytologically malignant epithelial cells. The cells of AIS have enlarged hyperchromatic nuclei that tend to stratify, have frequent mitotic figures and can form epithelial tufts (Figure 14). Glands involved by AIS do not extend into the stroma beyond the depth of glands not involved by AIS, nor by definition do they produce stromal desmoplasia. Neither ultrastructural nor immunohistochemical studies contribute to the diagnosis of preinvasive glandular lesions. Endocervical, intestinal and endometrioid subtypes of AIS have been described; of these, the endocervical subtype is the most common (Jaworski et al., 1988). AIS must be distinguished from invasive adenocarcinoma, Arias–Stella reaction, glandular atypias due to inflammation and/or radiation, endometriosis, tubal metaplasia, microglandular hyperplasia and mesonephric remnants (Kurman et al., 1992).

**Invasive lesions**

The World Health Organization Classification for tumours of the uterine cervix recognizes three general categories of epithelial tumours: squamous tumours and precursors, glandular tumours and precursors, and

---

**Figure 13** Cervical intraepithelial neoplasia 3

Squamous epithelium consists entirely of atypical basaloid cells. Note the moderate nuclear polymorphism, coarse chromatin and mitotic figures in the upper half of the epithelium.

From Tavassoli & Devilee (2003)

**Figure 14** Adenocarcinoma *in situ*, coexisting with a normal endocervical epithelium (x 10)

From Sellors & Sankaranarayanan (2003)
Epithelial tumours
Squamous tumours and precursors
Squamous cell carcinoma, not otherwise specified 8070/3
  Keratinizing 8071/3
  Non-keratinizing 8072/3
  Basaloid 8083/3
  Verrucous 8051/3
  Warty 8051/3
  Papillary 8052/3
  Lymphoepithelioma-like 8082/3
  Squamotransitional 8120/3
Early invasive (microinvasive) squamous cell carcinoma 8076/3
Squamous intraepithelial neoplasia
  Cervical intraepithelial neoplasia (CIN) 3 / squamous-cell carcinoma in situ 8077/2
Benign squamous cell lesions
  Condyloma acuminatum 8052/0
  Squamous papilloma 8052/0
  Fibroepithelial polyp 8052/0
Glandular tumours and precursors
Adenocarcinoma 8140/3
  Mucinous adenocarcinoma 8480/3
  Endocervical 8482/3
  Intestinal 8144/3
  Signet-ring cell 8490/3
  Minimal deviation 8480/3
  Villoglandular 8262/3
  Endometrioid adenocarcinoma 8380/3
  Clear cell adenocarcinoma 8310/3
  Serous adenocarcinoma 8441/3
  Mesonephric adenocarcinoma 9110/3
  Early invasive adenocarcinoma 8140/3
Adenocarcinoma in situ 8140/2
Glandular dysplasia
Benign glandular lesions
  Müllerian papilloma 8098/3
  Endocervical polyp 8098/3
Other epithelial tumours
Adenosquamous carcinoma 8560/3
  Glassy cell carcinoma variant 8015/3
  Adenoid cystic carcinoma 8200/3
  Adenoid basal carcinoma 8200/3
Neuroendocrine tumours
  Carcinoid 8240/3
  Atypical carcinoid 8249/3
  Small cell carcinoma 8041/3
  Large cell neuroendocrine carcinoma 8013/3
Undifferentiated carcinoma 8020/3
Mesenchymal tumours and tumour-like conditions
Leiomyosarcoma 8890/3
Endometrioid stromal sarcoma, low grade 8931/3
Undifferentiated endocervical sarcoma 8805/3
Sarcoma botryoides 8910/3
Alveolar soft part sarcoma 9581/3

"other" epithelial tumours (Table 3). The staging system developed by the International Federation of Gynecology and Obstetrics (FIGO Committee on Gynecologic Oncology and IGCS Guidelines Committee, 2000) is widely accepted (Table 4).

Microinvasive squamous-cell carcinoma
Microinvasive squamous-cell carcinomas are clinicopathologically defined lesions with minimal stromal invasion, an excellent prognosis and an extremely low incidence of lymph node metastases (Benson & Norris, 1977; Fu & Berek, 1988; Creasman et al., 1998).

The FIGO defines microinvasive squamous-cell carcinomas as those demonstrating stromal invasion less than 5 mm, as measured from the point of origin of the invasive tumour elements. In addition, the surface extent of microinvasive squamous-cell carcinomas must be less than 7 mm, and there must be no histological evidence of vascular space invasion.

Invasive squamous-cell carcinoma
Three major pathological variants of invasive squamous-cell carcinomas in the uterine cervix are recognized: keratinizing carcinoma, large-cell non-keratinizing carcinoma and small-cell carcinoma. However, histopathological tumour type and tumour grade (well, moderately and poorly differentiated) are less predictive of patient outcome than the depth of invasion and presence (or absence) of lymphovascular tumour embolization (Crissman et al., 1987; Look et al., 1996) (Figure 15).

Unusual histological variants of cervical squamous-cell carcinoma include lymphoepithelial-like carcinoma (Mills et al., 1985), verrucous carcinoma (Tiltman & Atad, 1982), papillary carcinoma (Randall et al., 1986) and spindle-cell (sarcomatoid) carcinoma. Spindle-cell squamous carcinoma is rare in the uterine cervix (Steeper et
al. 1983) and this diagnosis should be established only after clinicoradiological evaluation of extrauterine sites, together with immunohistochemical analysis to exclude a diagnosis of melanoma.

The clinical behaviour of invasive squamous-cell carcinomas may be predicted by a variety of histopathological features and ancillary studies. Tumour size, depth of invasion, parametrial involvement and nodal status are significant prognostic factors (Zaino et al., 1992; Kristensen et al., 1999). The presence of HPV type 18 in invasive squamous lesions may be associated with worse clinical outcome (Burger et al., 1995; Rose et al., 1995; Nakagawa et al., 1996). However, neither tumour ploidy status (Atkin et al., 1990) nor cellular oncogene expression (Riou, 1988) has been established as an independent prognostic marker in invasive squamous lesions.

Invasive glandular lesions of the cervix
During the past several decades, many though not all countries have seen an appreciable increase in the proportion of endocervical adenocarcinoma (Parazzini & La Vecchia, 1990; Ursin et al., 1996; Vizcaino et al., 1998; Alfsen et al., 2000). Risk factors for invasive glandular lesions overlap with those for invasive squamous lesions, and invasive glandular lesions are associated with preinvasive squamous lesions in more than 50% of cases (Maier & Norris, 1980). Approximately 90% of invasive glandular lesions of the cervix are associated with high-risk HPV types, in particular HPV 18 (Lizano et al., 1997; Pirog et al., 2000). Use of immunohistochemical methods to distinguish endocervical glandular lesions...
Cervical cancer and screening

from endometrial glandular lesions revealed that 100% of AIS and 94% of endocervical adenocarcinomas were associated with high-risk HPV types (Zielinski et al., 2003). The term microinvasive endocervical adenocarcinoma has been applied to invasive tumours less than 5 mm in thickness (Lee & Flynn, 2000). However, this term may not refer to a reproducible and distinctive histopathological entity (Zaino, 2000).

Endocervical adenocarcinomas exhibit several histological patterns, and different patterns often coexist in the same lesion. Histological classification of endocervical adenocarcinoma is based on the predominant pattern. The most common subtypes are mucinous and endometrioid adenocarcinoma (Saigo et al., 1986; Kleine et al., 1989) (Figure 16). No significant prognostic differences have been detected between the more common histological subtypes of endocervical adenocarcinoma (Alfsen et al., 2001). Less common histological patterns include clear-cell adenocarcinoma (Noller et al., 1974), which may occur in young women with a history of in utero exposure to diethylstilbestrol (DES), minimal deviation adenocarcinoma (also referred to as adenoma malignum) (Steeper & Wick, 1986), papillary serous adenocarcinoma (Zhou et al., 1998), mesonephric

<table>
<thead>
<tr>
<th>Stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 0</td>
<td>Carcinoma in situ, preinvasive carcinoma</td>
</tr>
<tr>
<td>Stage I</td>
<td>Invasive carcinoma strictly confined to cervix</td>
</tr>
<tr>
<td>Stage IA</td>
<td>Invasive carcinoma identified microscopically (all microscopically visible lesions, even with superficial invasion, should be assigned to stage IB)</td>
</tr>
<tr>
<td>Stage IA1</td>
<td>Measured invasion of stroma 3.0 mm or less in depth and 7.0 mm or less in horizontal spread</td>
</tr>
<tr>
<td>Stage IA2</td>
<td>Measured invasion of stroma more than 3.0 mm but no greater than 5.0 mm in depth and 7.0 mm or less in horizontal spread</td>
</tr>
<tr>
<td>Stage IB</td>
<td>Clinically visible lesion confined to cervix or microscopic lesion greater than stage IA2</td>
</tr>
<tr>
<td>Stage IB1</td>
<td>Clinical lesions of 4.0 cm or less in size</td>
</tr>
<tr>
<td>Stage IB2</td>
<td>Clinical lesions more than 4.0 cm in size</td>
</tr>
<tr>
<td>Stage II</td>
<td>Carcinoma extending beyond cervix but not to pelvic sidewall; carcinoma involves vagina but not its lower third</td>
</tr>
<tr>
<td>Stage IIA</td>
<td>No parametral involvement</td>
</tr>
<tr>
<td>Stage IIB</td>
<td>Parametral involvement</td>
</tr>
<tr>
<td>Stage III</td>
<td>Carcinoma extending onto pelvic wall; the tumour involves lower third of the vagina. All patients with hydronephrosis or non-functioning kidney are included unless known to be the result of other causes</td>
</tr>
<tr>
<td>Stage IIIA</td>
<td>Involvement of lower third of the vagina; no extension of pelvic sidewall</td>
</tr>
<tr>
<td>Stage IIIB</td>
<td>Extension to pelvic sidewall and/or hydronephrosis or non-functioning kidney</td>
</tr>
<tr>
<td>Stage IV</td>
<td>Carcinoma extends beyond true pelvic or clinically involves mucosa of bladder or rectum. Bullous oedema does not allow a case to be designated as stage IV</td>
</tr>
<tr>
<td>Stage IVA</td>
<td>Spread of growth to adjacent organs</td>
</tr>
<tr>
<td>Stage IVB</td>
<td>Spread to distant organs</td>
</tr>
</tbody>
</table>

From FIGO Committee on Gynecologic Oncology and IGCS Guidelines Committee (2000)
carcinoma (Valente & Susin, 1987), villoglandular adenocarcinoma (Hopson et al., 1990), glassy-cell carcinoma (Maier & Norris, 1982), adenoid cystic carcinoma (Mazur & Battifora, 1982), adenoid basal carcinoma (Baggish & Woodruff, 1966), adenocarcinoma with carcinoid features (Albores-Saavedra et al., 1979) and adenosquamous carcinoma (Yazigi et al., 1990). The villoglandular and adenosquamous subtypes have been reported to confer a more favourable prognosis (Chen et al., 1998). The glassy-cell and papillary serous subtypes have been reported to be associated with less favourable prognosis.

Ancillary studies are generally required to distinguish primary endocervical adenocarcinoma from primary endometrial and metastatic adenocarcinomas. Immunohistochemical staining for carinoembryonic antigen (CEA) (Kudo et al., 1990), estrogen receptor protein (ER) (Staepler et al., 2002), progesterone receptor protein (PR) (Staepler et al., 2002), vimentin, 1C5 (Kudo et al., 1990) and mucus-specific antigens M1, M2, and M3 (Maes et al., 1988) have demonstrated utility in differentiating endocervical from endometrial origin for uterine adenocarcinomas. HPV in situ hybridization analysis has also been demonstrated to be useful in making this distinction (Staepler et al., 2002).

Other cervical neoplasms
Unusual cervical neoplasms include leiomyosarcoma (Abell & Ramirez, 1973), endocervical stromal sarcoma (Jaffe et al., 1985), embryonal rhabdomyosarcoma (sarcoma botryoides) (Brand et al., 1987; Daya & Scully, 1988), alveolar soft-part sarcoma (Flint et al., 1985), osteosarcoma (Gilmore et al., 1956), malignant schwannoma, liposarcoma, malignant fibrous histiocytoma (Clement, 1990), malignant mixed mesodermal tumour (Gersell et al., 1989), Wilms tumour (Bell et al., 1985) and primary melanoma (Hall et al., 1980).

Diagnosis and treatment of cervical preinvasive and invasive disease

Quality assurance is of critical importance to successful cervical cancer prevention (Miller, 2002a), and the pathologist is responsible for detecting and analysing errors that routinely occur in screening programmes by comparing histological findings with those from previous screening tests. The aim of this process of cytological/biopsy correlation (Table 5) is to ensure that women in target demographic groups receive appropriate clinical management.

Preinvasive cervical lesions

Diagnosis
Exfoliative cytology is the commonest way of diagnosing preinvasive disease; other methods, described later in this volume, include HPV DNA testing, screening colposcopy, visual inspection with acetic acid (VIA), visual inspection with Lugol’s iodine (VILI) and the newer and still experimental methods based on real-time imaging and tumour markers. A recent revision of the cytological classification criteria—the 2001 Bethesda system—(Solomon et al., 2002), although keeping the classification of LSIL and HSIL to refer to the low and high grades in respect of preinvasive precursor lesions, subdivides the lower degrees of abnormality of so-called atypical squamous cells (ASC) into two categories (i.e., atypical squamous cells of undetermined significance, ASCUS, and atypical squamous cells that cannot exclude HSIL, ASC-H). The latter term (ASC-H) recognizes the high risk associated with these apparently minor cytological abnormalities containing undiagnosed HSIL lesions.

In general, women with abnormal cytological findings are referred for colposcopic evaluation, the referral criteria varying from country to country. After colposcopic diagnosis, a punch biopsy may be taken to confirm the histological diagnosis or immediate treatment may be instituted without prior histological confirmation of disease, the so-called "see and treat" approach.

Management
The atypical squamous cell classification described above comprises two categories defined as ASCUS and ASC-H. The former is associated with a relatively low risk of discovering underlying high-grade disease (in 5–17%), while the latter group has a much higher risk (Wright et al., 1995; Manos et al., 1999; Solomon et al., 2001). The ASCUS-LSIL Triage Study (Solomon et al., 2001; Stoler & Schiffman, 2001) confirmed that HPV DNA testing was helpful in identifying women with underlying high-grade histological disease in both groups. The high negative predictive value of HPV DNA testing (99%) immediately excludes the risk of underlying high-grade lesions in women with a nega-
tive result. Those with a positive result would be candidates for referral to colposcopy. An alternative to employing HPV DNA testing is repeated cytology at six- to twelve-month intervals. However, it seems that all diagnostic modalities (i.e., cytology, HPV DNA analysis and colposcopy) are fairly equivalent in the management of women with atypical squamous cells (Wright et al., 2002a).

Histological confirmation of CIN 1, the most minor histological state of the three-stage CIN classification, presents problems in respect of consistency of diagnosis. In the ASCUS-LSIL Triage Study of low-grade disease, biopsies originally classified as CIN 1 were reclassified as CIN 1 in 43% of cases, downwards in 45% of cases and upgraded to CIN 2/3 in 12% (Stoler & Schiffman, 2001). Another problem is the presence of high-grade disease in up to 50% of excised specimens from women presenting with minor referral cytological abnormalities, i.e., mild dyskaryosis or LSIL (Massad et al., 1996).

Rates of regression of CIN 1 appear to be up to 60%, with progression rates of only 10% (Östör, 1993; Melnikow et al., 1998) and it is not clear whether these lesions should be treated or not. Two studies (Flannelly et al., 1994; Shafi et al., 1997) indicated that it is safe to monitor the women with cytological follow-up rather than the commonly employed practice of immediate excision when the colposcopist assumes that underlying high-grade intraepithelial disease (CIN 2 or 3) is present. Indeed, the rate of confirmed CIN 2 or 3 in excised specimens is low (Luesley et al., 1990). The Bethesda consensus conference (Wright et al., 2003) recommended that women could be followed by regular cytological surveillance with no increased risk of severe preinvasive lesions being undetected. Conservative management can also be supplemented by an HPV DNA test after 12 months, when a finding of a high-risk HPV type will indicate referral for colposcopy. Likewise, after two negative cytological tests over one year and a negative HPV DNA result, a woman can be safely returned to routine screening.

Lesions histologically confirmed as CIN 3 have been found to have an average likelihood of 12% of progression to cancer, while those with milder changes (CIN 2) had an average rate of progression to carcinoma in situ of 22% (Östör, 1993). Such estimates vary substantially (Mitchell et al., 1996; Melnikow et al., 1998), but treatment is indicated for both CIN 2 and CIN 3 lesions.

**CIN in pregnancy**

Regression of CIN 2 and 3 during pregnancy is minimal, but there is significant spontaneous regression post partum (Yost et al., 1999). A patient with a cytological diagnosis of high-grade disease should undergo colposcopy. Punch biopsy to confirm a high-grade preinvasive lesion is not contraindicated in pregnancy. Excisional procedures, however, should be performed only where there is a definite risk of microinvasion (Wright et al., 2003).

**CIN in immunosuppressed women**

Immunosuppressed women have an elevated incidence of persistent HPV infection, which in turn is associated with cervical cancer and its precursors. Women infected with human immunodeficiency virus (HIV) have increased rates of low- and high-grade epithelial lesions and atypical squamous cellular cytological abnormalities. (Massad et al., 2001). There seems to be a high rate of recurrence and persistence of CIN 2 and 3 after treatment in HIV-positive women, with failure rates approaching 25% (Holcomb et al., 1999) (see Chapter 4).

**Treatment techniques**

The object of treating histologically confirmed preinvasive CIN disease by any technique is to effectively eradicate the lesion, with minimal associated morbidity. Two categories of treatment techniques are available, namely destructive and excisional.

Techniques which involve destruction of the whole atypical transformation zone can be applied only if strict criteria are employed to ensure that no evidence of invasive cervical cancer lesions is present. These techniques, which include CO₂ vaporization, cryotherapy, electrocauterization and cold (thermo) coagulation, all have success rates of around 90%. Pre-treatment biopsy is mandatory and the level of destruction by these techniques extends to 7 mm depth; Anderson & Hartley (1980) showed that crypt involvement by CIN extends on average to about 3 mm. A meta-analysis has found that there is very little to choose between these techniques with regard to success or complications of treatment (Martin-Hirsch et al., 2004).

Excisional techniques involving surgical removal (followed by histological analysis) range from CO₂ laser excision through the cold knife technique to the rare application of hysterectomy. However, electrosurgical excision of the transformation zone using an electrosurgical unit which produces a constant low voltage, with the ability to blend the cutting and coagulation characteristics of the current, is now the most popular technique. It must be performed after a comprehensive colposcopic examination and the intention is to remove the entire transformation zone with an adequate margin of normal squamous epithelium surrounding the abnormal area, with minimal artifactual damage (Prendiville, 2003a). Various types of morphological excision have been popularized by Prendiville (2003b), who proposed three parameters of
<table>
<thead>
<tr>
<th>Previous Pap smear result</th>
<th>Current biopsy result</th>
<th>Correlation requirement</th>
<th>Post-correlation error category</th>
<th>Potential action</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIL</td>
<td>None/missing</td>
<td>None</td>
<td>Cytology and biopsy interpretations accurate; clinical sampling error possible</td>
<td>Correlate with smear collector ID and monitor</td>
</tr>
<tr>
<td></td>
<td>&lt;CIN 1</td>
<td>None</td>
<td>Cytology screening undercall</td>
<td>Correlate with cytotech ID and monitor</td>
</tr>
<tr>
<td></td>
<td>≥CIN 1</td>
<td>Review of previous Pap and current biopsy</td>
<td>Biopsy interpretive overcall</td>
<td>Correlate with pathologist ID and monitor</td>
</tr>
<tr>
<td>ASC-US</td>
<td>None/missing</td>
<td>Locality-specific</td>
<td>Cytology and biopsy interpretations accurate; clinical sampling error possible</td>
<td>Correlate with smear collector ID and monitor</td>
</tr>
<tr>
<td></td>
<td>&lt;CIN 2</td>
<td>None</td>
<td>Cytology interpretive undercall</td>
<td>Correlate with pathologist ID and monitor</td>
</tr>
<tr>
<td></td>
<td>≥CIN 2</td>
<td>Review of previous Pap and current biopsy</td>
<td>Biopsy interpretive overcall</td>
<td>Correlate with pathologist ID and monitor</td>
</tr>
<tr>
<td>LSIL</td>
<td>None/missing</td>
<td>Locality-specific</td>
<td>Cytology and biopsy interpretations accurate; clinical sampling error possible</td>
<td>Correlate with pathologist ID and monitor</td>
</tr>
<tr>
<td></td>
<td>&lt;CIN 1</td>
<td>None</td>
<td>Cytology interpretive overcall</td>
<td>Correlate with pathologist ID and monitor</td>
</tr>
<tr>
<td></td>
<td>≥CIN 1</td>
<td>None</td>
<td>Cytology interpretive undercall</td>
<td>Correlate with pathologist ID and monitor</td>
</tr>
<tr>
<td>AGC</td>
<td>None/missing</td>
<td>Alert programme manager</td>
<td>Clinical and/or systems errors</td>
<td>Clinical and/or process changes</td>
</tr>
<tr>
<td></td>
<td>&gt;CIN 1</td>
<td>Review of previous Pap and current biopsy</td>
<td>Cytology interpretive overcall</td>
<td>Correlate with pathologist ID and monitor</td>
</tr>
<tr>
<td></td>
<td>≥CIN 1</td>
<td>None</td>
<td>Cytology interpretive undercall</td>
<td>Correlate with pathologist ID and monitor</td>
</tr>
<tr>
<td>HSIL</td>
<td>None/missing</td>
<td>Alert programme manager</td>
<td>Clinical and/or systems errors</td>
<td>Clinical and/or process changes</td>
</tr>
<tr>
<td></td>
<td>&lt; CIN 1</td>
<td>Review of previous Pap and current biopsy</td>
<td>Cytology interpretive overcall</td>
<td>Correlate with pathologist and monitor</td>
</tr>
<tr>
<td></td>
<td>≥ CIN 1</td>
<td>None</td>
<td>Cytology and biopsy interpretations accurate; clinical sampling error possible</td>
<td>Alert colposcopist of possible incomplete clinical evaluation</td>
</tr>
</tbody>
</table>
transformation-zone morphology that should be considered before undertaking excisional treatment—these are the size of the transformation zone, the position of the upper limit of the transformation zone and the visibility of this upper limit. These characteristics then identify the transformation zone as being completely ectocervical, fully visible but with an endocervical component or not fully visible. Classification of transformation zones into these three types allows simple documentation and comparison.

Martin-Hirsch et al. (2004) conducted a meta-analysis of 28 individual trials of various treatments for CIN. Comparison of the ablative and excisional techniques revealed no technique superior to any other. However, the authors noted that the excisional technique using an electrosurgical unit for so-called large loop excision of the transformation zone (LLETZ) produced the least morbidity and the most favourable surgical specimen for histological analysis, and concluded that LLETZ seemed to be the ideal method for treating CIN.

The effectiveness of LLETZ was analysed in two studies by Flannelly et al. (1997, 2001), who categorized women into three distinct groups with respect to risk of recurrence. Women younger than 50 years and without margin involvement in the excised specimen had a 92% chance of a normal cytological result in subsequent follow-up, while those with margin involvement had an 86% chance of a normal result. In contrast, older women (aged 50 and over) with margin involvement had only a 57% chance of a normal result. Therefore the latter group, who comprised only 3% of the women treated, are at high risk of recurrence and should be more intensively followed up by cytology and colposcopy than the younger group. Other studies have confirmed the relatively low failure rate with the LLETZ procedure (Tables 6 and 7).

In the 'see-and-treat' procedure, a woman seen at a first visit can be treated on the basis of the colposcopic diagnosis, without histological confirmation. This approach obviates the need for a second visit but may result in overtreatment or unnecessary treatment of some women. In addition, some women require more time to consider their options with regard to treatment and potential complications.

New antiviral treatments using various therapeutic agents have been tested in an attempt to reduce the need for surgical intervention. Imiquimod, a non-specific immune response modulator, has been used in limited trials to treat low-grade lesions. Results suggest a variable clinical response.

---

### Table 5 (contd)

<table>
<thead>
<tr>
<th>Previous Pap smear result</th>
<th>Current biopsy result</th>
<th>Correlation requirement</th>
<th>Post-correlation error category</th>
<th>Potential action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malignant</td>
<td>None/missing</td>
<td>Alert programme manager</td>
<td>Clinical and/or systems errors</td>
<td>Clinical and/or process changes</td>
</tr>
<tr>
<td>&lt; Malignant</td>
<td>Review of previous Pap and current biopsy</td>
<td>Cytology interpretive overcall</td>
<td>Correlate with pathologist ID and monitor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Biopsy interpretive undercall</td>
<td>Correlate with pathologist ID and monitor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cytology and biopsy interpretations accurate; clinical sampling error possible</td>
<td>Alert colposcopist of possible incomplete clinical evaluation</td>
<td></td>
</tr>
<tr>
<td>Malignant</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*These are not practice guidelines. Rather, this table is intended to summarize the general principles of error detection and analysis through cytological/biopsy correlation studies. Previous Pap smear result(s), whenever and wherever available, should be reviewed before finalizing the histological analysis of any cervical biopsy specimen. When comparison of previous Pap smear results with current biopsy results indicates a discrepancy (correlation requirement), the previous Pap smear(s) should be retrieved and reviewed in conjunction with the biopsy. When possible, the previous Pap smear should be reviewed by a cytologist other than the one who initially examined the Pap smear.

**Abbreviations:** NIL, negative for intraepithelial lesion; ID, identity; CIN, cervical intraepithelial neoplasia; LSIL, low-grade squamous intraepithelial lesion; AGC, atypical glandular cells; HSIL, high-grade squamous intraepithelial lesion; ASC-US, atypical squamous cells of undetermined significance
but with associated troublesome systemic side-effects (Cruickshank, 2003). HPV vaccines have been tried as an alternate immunotherapy regime for treating low-grade lesions. These therapeutic vaccines to eliminate HPV from the basal cells have been used in limited trials. Difficulties in assessing end-points following treatment in conjunction with the high regression rate of low-grade lesions are among the problems that hamper the acceptance of such vaccines (Fiander & Man, 2003).

**Follow-up after treatment of CIN**

There is a well recognized risk of recurrence of CIN and rarely of invasive cancer following both ablative and excisional treatment of CIN (Table 6). Follow-up can be by colposcopy, cytology or HPV DNA testing, or by a combination of any of these. Colposcopy at the first post-operative visit is beneficial to evaluate the physical state of the cervix and to determine the mode of follow-up. For example, if there is cervical constriction, an endocervical brush for cytological specimen collection would be recommended. The efficacy of post-operative cytological surveillance will be determined by the structure of the cervix.

Combining HPV DNA with cytological testing has been evaluated in two large meta-analyses. Zielinski et al. (2004) showed that the combined tests had increased sensitivity to detect persistent or recurrent CIN and increased negative predictive value to identify women at little or no risk for persistence or recurrence. The combination proved more effective than either test alone. It was recommended that women treated for CIN 3 should have this combined test after six months. If positivity on either test is found, then colposcopy and close surveillance are indicated. Women with a double negative test can be safely seen at 24 months and, if negative at this stage, should be referred back to routine screening. Paraskevaidis et al. (2004) drew similar conclusions, but highlighted that a positive HPV test, even in the presence of normal cytological results, can pick up treatment failures more quickly and accurately. However, they noted that cytology and colposcopy may still need to be performed to rule out false positive or negative results.

### Table 6. Failure rates following loop excision (recurrence within one year)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of patients</th>
<th>Rate of residual disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prendiville et al. (1989)</td>
<td>102</td>
<td>3.0%</td>
</tr>
<tr>
<td>Murdoch et al. (1991)</td>
<td>721</td>
<td>4.6%</td>
</tr>
<tr>
<td>Bigrigg et al. (1994)</td>
<td>1000</td>
<td>5.0%</td>
</tr>
<tr>
<td>Flannelly et al. (1997)</td>
<td>1000</td>
<td>8.0%</td>
</tr>
<tr>
<td>Gardell et al. (1997)</td>
<td>225</td>
<td>8.5%</td>
</tr>
<tr>
<td>Baldauf et al. (1998)</td>
<td>288</td>
<td>6.9%</td>
</tr>
<tr>
<td>Dobbs et al. (2000)</td>
<td>322</td>
<td>4.3%</td>
</tr>
<tr>
<td>Narducci et al. (2000)</td>
<td>505</td>
<td>3.7%</td>
</tr>
<tr>
<td>Paraskevaidis et al. (2000)</td>
<td>635</td>
<td>4.9%</td>
</tr>
</tbody>
</table>

Reproduced with permission from Prendiville (2003b)

### Table 7. Findings of unexpected microinvasion or invasion in specimens after loop excision

<table>
<thead>
<tr>
<th>Series</th>
<th>Unexpected</th>
<th>Cytology or punch biopsy findings (when known)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prendiville et al. (1989)</td>
<td>1/102</td>
<td>(1%) CIN 3</td>
</tr>
<tr>
<td>Bigrigg et al. (1990)</td>
<td>5/1000</td>
<td>(0.5%) CIN 1 x 2, CIN 2 x 3</td>
</tr>
<tr>
<td>Gunesker et al. (1990)</td>
<td>1/91</td>
<td>(1%) CIN 3</td>
</tr>
<tr>
<td>Luesley et al. (1990)</td>
<td>4/616</td>
<td>(0.6%) –</td>
</tr>
<tr>
<td>Whiteley &amp; Olah (1990)</td>
<td>0/80</td>
<td>(0%) –</td>
</tr>
<tr>
<td>Halam et al. (1991)</td>
<td>8/1000</td>
<td>(0.8%) –</td>
</tr>
<tr>
<td>Wright et al. (1991)</td>
<td>1/157</td>
<td>(0.6%) CIN 2</td>
</tr>
<tr>
<td>Murdoch et al. (1992)</td>
<td>11/1143</td>
<td>(1%) –</td>
</tr>
</tbody>
</table>

Reproduced with permission from Prendiville (2003b)

---

#### Cervical glandular intraepithelial neoplasia and early invasive lesions

**Diagnosis**

Adenocarcinoma *in situ* (AIS) was first described in 1952. It is found in approximately 1% of all conizations performed for CIN. Coexisting CIN is found in approximately two thirds of AIS cases, while 10% of women with AIS have co-existing cancer of a glandular type. The Bethesda 2001 Conference (Solomon et al., 2002)
reclassified atypical glandular cells (AGC) as atypical endocervical or endometrial cells or not otherwise specified. The management of patients with glandular abnormalities depends on the primary site. The percentage of cases of AGC associated with underlying high-grade disease is higher than for ASCUS. A cytological finding of AGC, favouring neoplasia, or of an adenocarcinoma in situ itself dictates referral for colposcopy. An associated CIN lesion would most likely be colposcopically diagnosed; a glandular preinvasive lesion is extremely difficult to confirm by this technique (Cullimore, 2003). Multi-focal disease is found in approximately 16% of AIS cases, with the glandular lesions themselves extending for a variable distance into the endocervical canal, although the majority (95%) progress less than 25 mm from the anatomical external os. Because of the poor sensitivity of colposcopy and even of cytology, endocervical sampling in women with these abnormal glandular cytological findings is an important part of diagnosis.

**Treatment**

It is important to note that colposcopy cannot adequately evaluate the endocervix and where glandular lesions are suspected, a cone biopsy will usually be required (see above). The LLETZ technique is unsatisfactory in procuring a large enough specimen for diagnosis and is associated with a higher frequency of marginal involvement (Wolf et al., 1996) than with excision using the cold knife or CO₂ laser. A cylindrical base specimen with a width sufficient to encompass 5 mm on each side of the transformation zone is ideal (Cullimore, 2003). In women wishing to preserve fertility, the specimen taken usually extends for only 10 mm above the squamocolumnar junction of the cervix, but in older women it may include 20 mm of length of the endocervix.

There is a high risk of recurrence evidenced by finding residual AIS or even invasive adenocarcinoma in excisional samples. Soutter et al. (2001), in a study of 84 subjects from five hospitals, showed an incidence of residual disease at subsequent hysterectomy of 8/22 women in whom the margins of the initial excision specimen were involved by the glandular abnormality. In women in whom the excision margins are involved, a repeat conization is undertaken, especially in those wishing to maintain fertility, and hysterectomy in those who no longer desire fertility. Follow-up of these women by intense cytological screening and HPV DNA testing at six-monthly intervals with associated endocervical canal monitoring with either brush cytology or curettage is essential so as to check for recurrences.

**Cervical cancer**

**Diagnosis and staging**

A diagnosis of cervical cancer is uncommonly associated with an abnormal cytological result. More usually, the diagnosis is made when clinical symptoms develop. Confirmation is by biopsy taken from a suspicious lesion on the cervix or vaginal fornices. If colposcopy or a biopsy suggests early cancer in the form of microinvasion (Stage 1A), a mandatory and subsequent excisional biopsy (cone) must be taken which incorporates both epithelium and stroma, so that the depth and width of invasion below the basement membrane can be adequately assessed.

Staging for cervical cancer is based on clinical evaluation (bimanual digital pelvi-rectal examination), which should preferably be performed under general anaesthesia by an experienced examiner. The examination is designed to assess the extension of a tumour beyond the cervix into the parametria (transverse cervical ligaments), utero-sacral ligaments, the pelvic sidewall, the vagina and the bladder and/or rectum.

The clinical staging should not be changed in the light of findings of subsequent investigations. In case of doubt, the stage should be allocated to an earlier rather than a more advanced stage (FIGO Committee on Gynecologic Oncology and IGCS Guidelines Committee, 2000) (Table 4). Additional routine investigations to supplement the staging of clinically evident cervical cancer include liver and renal function tests, chest X-ray, intravenous pyelogram/urography or ultrasound of the ureters to diagnose hydronephrosis secondary to ureteral obstruction by a tumour, cystoscopy to diagnose occult bladder invasion and proctoscopy if rectal mucosa involvement is suspected. Other investigations, such as bone scans or skeletal X-radiography, should be performed according to the clinical presentation of the patient. Endocervical curettage, hysteroscopy and colposcopy may be required to assess early microinvasive disease. Attempts to document extension of tumour to the uterine corpus should not be made.

Investigations such as computerized tomography or magnetic resonance imaging (MRI) scanning should not be used as the basis for clinical staging, but can provide useful information for planning treatment or further investigation, such as fine-needle aspiration of suspected lymph node involvement with metastases.

Diagnosis of stage IA1 and IA2 cervical cancer is based on microscopic examination of cervical tissue, provided by a large excisional biopsy (cone biopsy). The two most important dimensions are the depth of invasion (< 5 mm from the base of the epithelium from which the lesion originates) and the width of the invasive lesion (horizontal spread should be less than 7 mm). Vascular or lymphatic space
invasion does not affect the staging, but should be recorded, as it may influence treatment options (Figure 17).

It should be noted that if hydronephrosis is present, the stage is automatically allotted to stage IIIB even if the clinical staging is less advanced. The presence of bullous oedema of the bladder does not imply allocation to stage IVA unless there is histological confirmation of invasion into the bladder.

**Treatment of different stages**

Stage IA1 disease (depth of invasion < 3 mm and < 7 mm wide) has a risk of metastasis to regional lymph nodes of 1.2%, with a death rate of less than 1% (Sevin et al., 1992; Benedet & Anderson, 1996). Where preservation of fertility is important, a cone biopsy may be considered a therapeutic procedure provided that (a) the woman is available for long-term follow-up, (b) the cervix is amenable to cytological and colposcopic evaluation, (c) the margins of the cone biopsy are free of both intraepithelial and invasive changes, and (d) there is no evidence of lymphatic or vascular invasion. Otherwise the optimal treatment is simple hysterectomy, which may be performed vaginally, abdominally or laparoscopically. Before hysterectomy, colposcopy should be performed to exclude occult extension to the vaginal fornices. If this is found, a cuff of vagina should be removed at the time of hysterectomy.

Stage IA2 (depth of invasion between 3 and 5 mm, width < 7 mm) has a risk of metastasis to regional lymph nodes of nearly 8% and a mortality rate of 2.4% (Sevin et al., 1992; Benedet & Anderson, 1996). The recommended treatment is modified radical hysterectomy and bilateral pelvic lymphadenectomy. In the absence of vascular or lymphatic invasion, a simple hysterectomy and bilateral node dissection is also considered adequate therapy (Elliott et al., 2000). If preservation of fertility is important, a large cone biopsy with nodal dissection or trachelectomy with nodal dissection (extraperitoneal or laparoscopic) may be considered (Dargent et al., 2000; Shepherd et al., 1998).

Microinvasive adenocarcinoma seems to be the counterpart to microinvasive squamous carcinoma (Ostör et al., 1997). Tumour involvement to at most 3 mm into the stroma is associated with minimal metastatic risk (Smith et al., 2002a). However, measurement of such involvement and the distinction between truly invasive and intraepithelial disease remain difficult (Kaspar et al., 1993). Treatment of early adenocarcinoma is contentious, but a recent meta-analysis reporting on 1170 cases treated by radical or simple hysterectomy and some even by conization showed that nodal metastasis was no more than 2.8% (Soutter, 2003). Disease-free survival rates for all treatment methods were just under 99%. Stage IA2 was not significantly different in prognosis from stage IA1 disease. The current recommendation for treatment of microinvasive adenocarcinoma is the same as for squamous-cell cancer.

Treatment strategies for stage IB and even early stage II invasive cancer include primary surgery with radical hysterectomy and pelvic lymphadenectomy or the option of primary radiation therapy with external beam radiation with either high-dose or low-dose rate brachytherapy. Published observational data indicate a five-year survival rate of 87–92% using either approach (Waggoner, 2003).

Radical hysterectomy and associated lymphadenectomy in younger patients involves ovarian conservation and avoids vaginal stenosis, which is a complication of radiotherapy. Surgical complications should be under 5% (Potter et al., 1990). Radiotherapy involves a combination of external irradiation (using 40–50 Gy administered...
over four to five weeks in daily portions) and intra-cavitary therapy (brachytherapy, with the intention of achieving a total dose of 80–85 Gy to point A and 50–55 Gy to point B), so as to treat the cervical disease, the parametrial sidewall tissues and pelvic nodes. Long-term complications involving the bladder and bowels may occur in up to 10% of cases.

Women with stage IB1 disease (under 4 cm in diameter) in whom fertility is important may be offered radical trachelectomy and laparoscopic lymphadenectomy, as described in relation to stage IA2 disease.

The treatment of stage IB1 cervical cancer (diameter < 4 cm) depends on the resources and the type of oncology services available and the age and general health of the woman. Multidisciplinary evaluation of women is recommended, to consider carefully the therapeutic options and toxicities. Dual modality treatments (surgery and radiotherapy) are more harmful, more expensive and associated with a higher rate of complications. Therefore, primary therapy should aim to use only one radical therapy, either surgery or radiation (with or without concurrent chemotherapy).

Recommended surgery involves radical hysterectomy (removal of the uterus, cervix, parametria, cuff of vagina, utero-sacral ligaments), bilateral pelvic lymph node dissection and ovarian suspension where appropriate.

If surgery is deemed inappropriate, primary radical chemo-radiation therapy is recommended. The standard radiation treatment is radical external beam radiation and brachytherapy. Concurrent chemotherapy is usually cisplatin given in a dose of 40 mg per m² weekly during external beam therapy.

When positive common iliac or para-aortic nodes have been identified as involved with metastatic disease, extended field radiation may be considered.

Five-year survival rates of generally 80–90% following either radical surgery or radical radiation as primary therapy for stage IB1 tumours have been reported (Hopkins & Morley, 1991; Landoni et al., 1997; Waggoner, 2003).

For stage IB2 disease (diameter > 4 cm), five-year survival rates are reduced to approximately 65–75% (Hopkins et al., 1991). Para-aortic nodes are commonly involved in this stage, as well as an increase in central and distant failures associated with recurrence.

Treatment options include (a) primary chemo-radiation therapy alone (Rose et al., 1999a); (b) primary radical hysterectomy with bilateral regional lymph node dissection, usually followed by radical adjuvant radiation (with or without concurrent chemotherapy), determined by pathological criteria such as disease-free margins, lymph-vascular space involvement and metastases to lymph nodes (Keys et al., 1999); and (c) neoadjuvant chemotherapy, followed by radical surgery as described above and the possible use of post-operative radiation (Sardi et al., 1993).

The standard primary treatment of advanced cervical cancer (stages IIA, IIB, IIIA, IIIB, IVA) is primary radical radiation with a combination of external beam and intracavitary brachytherapy, with concurrent chemo-radiation therapy (Keys et al., 1999; Rose et al., 1999b; Whitney et al., 1999). Results from five randomized trials on treatment of cervical cancer in the late 1990s prompted the US National Cancer Institute to recommend the incorporation of concurrent cisplatin-based chemotherapy in radiation therapy for cervical cancer treatment (National Cancer Institute, 2002).

If the surgical expertise and support services for post-operative care are available, pelvic exenteration may be considered for stage IVA disease, so long as there is no evidence of extension to pelvic sidewalls, in a patient with good general health.

Recurrent cervical cancer (stage IVB) may be in the pelvis, distant sites or both. The majority of recurrences occur within two years of diagnosis and the prognosis is poor, with most patients dying of their disease, with a mean survival time of seven months (van Nagell et al., 1979).

Management of women with distant metastases and advanced recurrent cervical cancer requires the efforts of a multidisciplinary team, and includes palliative use of anti-cancer therapies (chemotherapy, radiation therapy for treatment of symptoms, including surgery such as colostomy for relief of symptoms related to recto-vaginal fistulae), control of symptoms (pain, bleeding, discharge, symptoms related to specific metastases), emotional, psychological and spiritual support of the patient and her family.

**Invasive cervical cancer in pregnancy**

Approximately 3.5% of cervical cancers occur in pregnancy. Survival rates for those with stage I disease are around 85–95% (Hopkins et al., 1991). Individual management plans must be developed with proper consideration of the tumour size, the stage of the disease and the desire of women for continuation of the pregnancy.

Women with high-grade intraepithelial disease can deliver vaginally and be re-evaluated at four months post-partum. Those with early invasive disease (stage IA or B) may choose termination or continuation of the pregnancy until maturity of the fetus determines the date of delivery. The delay to accomplish fetal maturity, especially in patients with early invasive disease, leads to only a small degree of disease progression (Hopkins et al., 1991). The mode of delivery depends on the stage of the lesion and its volume; it is unclear whether vaginal delivery
influences progression. Most patients with stage I disease prefer caesarean section at the time of planned radical surgery, with vaginal delivery reserved for those with preinvasive or stage IA1 disease. Radical surgery and radiation offer similar cure rates, with the former being used for stages IA2, IB and IIA. Such pregnancies have, surprisingly, been associated with low morbidity and high survival rates when surgery is used (Goff et al., 2000). Retention of ovaries in young women is indicated. However, women with stage IIB or more advanced disease or those not medically fit are candidates for definitive radiation therapy, which should be initiated immediately after delivery. The actual application of radiation requires adaptation to the anatomical distortions created by the pregnancy and patients opting for primary radiation therapy who intend to have a primary termination should have this before external therapy begins.

Follow-up
Women who have had cervical cancer require four-monthly follow-up for the first two years and then twice yearly for the subsequent five years. An annual cytological examination of the vaginal vault and chest X-rays on a regular annual basis for five years have been recommended (American College of Obstetricians and Gynecologists, 2002). However, the interpretation of cytology and/or colposcopy following irradiation requires special expertise.

Immunosuppression
Women infected with HIV are of particular concern, particularly as cervical cancer is very common in areas where HIV infection is endemic, such as sub-Saharan Africa. Women who are immunosuppressed, with CD4+ counts below 200 cells/µL, are at particular risk if treated with either radiation or chemotherapy, both of which have immunosuppressive effects. Where possible or appropriate, surgical treatment is preferred for HIV-positive immunosuppressed women. Women with CD4+ counts over 350 cells/µL appear to tolerate either anti-cancer therapies as well as HIV-uninfected women (Lomalisa et al., 2000).

Cervical cancer in developing countries
In low-resource countries, many of the facilities for the treatment of cervical cancer do not exist, or if they do, the equipment is poorly maintained and will not provide either optimal or even suboptimal therapy. Chemotherapy may not be available, nor the resources or skills to provide radical surgical interventions. In these settings, even women with early cervical cancers will have a poor prognosis and the development of effective palliative care is essential.

In a low-resource environment, it will be difficult to apply the methods of diagnosis mentioned above to assist in defining the extent of malignant spread within the pelvis and allow accurate assessment of the stage of the cancer.

However, techniques such as cystoscopy, proctoscopy, radiography of pulmonary and renal systems with haematological and biochemical assessment can be performed within most settings. MRI and computerized tomography scanning may provide additional information but are not mandatory in assessing the FIGO staging. These sophisticated diagnostic techniques may provide valuable information for planning treatment but are extremely expensive. In low-resource settings, physical examination of the cervix, vagina, bladder and rectum sometimes offers the only feasible approach to staging.

Facilities for radical surgery, radiotherapy (in the form of external and intracavity radiation) and chemotherapy with platinum derivatives are available in many developing countries in Asia and Latin America (Sankaranarayanan et al., 1998a). However, this is not the case in large parts of sub-Saharan Africa.

Palliative care is inadequate in many regions of the world because of poor availability of medication, deficient health-care infrastructure, lack of training for health-care providers, associated with lack of counselling skills. Discomfort in discussing the diagnosis and management with patients and their families and a lack of awareness within the community of palliative care options are very common. Many health-care providers and policy makers lack awareness that there are inexpensive and effective ways to relieve advanced cancer symptoms (ACCP, 2003; WHO, 2001, 2002). WHO has developed a simple and inexpensive three-step analgesic ladder that can be easily incorporated into a treatment protocol (WHO, 2002).

The support of families and caregivers forms a most important part of the palliative management regime, and sometimes is the only source of social and psychological support. Patients and families must be informed and encouraged to believe that the cure of cervical cancer is feasible even in their sometimes deprived environment (Sankaranarayanan et al., 1998).

The etiology of cervical cancer

The epidemiological evidence relating HPV DNA to cervical cancer and its precursors includes a large and consistent body of data indicating, beyond any reasonable doubt, strong and specific associations between HPV infections and cervical cancer. The findings are consistent for all countries where investigations have taken place.
Once it was recognized that HPV represents a necessary cause of cervical cancer, reassessment of the role of co-factors in case–control studies was required by analysis restricted to HPV-positive women. Among persistently HPV-exposed women, some additional exposures further increase their risk of progression to advanced preinvasive lesions or invasive cancer. In the IARC studies, these co-factors were exposure to tobacco smoke, parity above five full-term pregnancies and use of oral contraceptives for five or more years. Presence of antibodies to *Chlamydia trachomatis* or to herpes simplex virus type 2 (HSV2) also modified the risk of progression significantly. Table 8 shows results pertaining to use of oral contraceptives, parity and cigarette smoking from the IARC multicentre case–control study (Castellsagué & Muñoz, 2003). The reported increases in risk for any of the co-factors were in general highly consistent within the range of 2.5- to 4-fold for the extreme categories of exposure.

### Human papillomaviruses

Natural history and follow-up studies have clearly shown that HPV infection precedes the development of cervical cancer by a number of years and confirmed that sexual transmission is the predominant mode of HPV acquisition. These studies provided biological support for the long-known clinical and epidemiological observations that cervical cancer displays the profile of a sexually transmitted disease (STD). Case–control studies, case series and prevalence surveys have unequivocally shown that HPV DNA can be detected in adequate specimens of cervical cancer in 95–100% of cases, compared with a prevalence of some 5–20% in cervical specimens from women identified as suitable epidemiological controls.

The association has been recognized as causal in nature by a number of international review parties since the early 1990s, and the claim has been made that this is the first necessary cause of a human cancer ever identified.

The implications of the recognition that, in the absence of viral DNA, cervical cancer does not develop, are of considerable public health relevance. On the one hand, the concept of risk groups comes into focus. High-risk women can be redefined as those who are persistent HPV carriers. Operationally this represents substantial progress from previous definitions of high-risk women according to their exposure to a constellation of ill-defined factors (low socioeconomic status, high number of sexual partners, smoking, use of oral contraceptives, history of STDs or any combination of the above). Most of these factors are now viewed either as surrogates of HPV exposure or as relevant co-factors given the presence of HPV DNA. On the other hand, if indeed HPV is a necessary cause of cervical cancer, the implication is that specific preventive practices targeting some putative non-HPV-related cervical cancer cases are no longer justified. Finally, methods are now available to screen the general population for HPV-DNA positivity.

On the basis of epidemiological surveys, the 40 HPV types infecting the genital area can be subdivided into low-risk types, which are mainly found in genital warts and CIN 1, and high-risk types, which are frequently associated with invasive cervical cancer. HPV DNA can be found in virtually all cervical carcinomas, with HPV types 16, 18, 31 and 45 being the most frequent ones (Bosch et al., 1995; Walboomers et al., 1999; Bosch et al.,)

### Table 8. Association of relevant co-factors among HPV-positive women

<table>
<thead>
<tr>
<th>Co-factor</th>
<th>HPV-positive women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases/controls</td>
</tr>
<tr>
<td><strong>OC use (status and years)</strong></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1071/163</td>
</tr>
<tr>
<td>Ever</td>
<td>605/92</td>
</tr>
<tr>
<td>1–4 years</td>
<td>274/64</td>
</tr>
<tr>
<td>5 years</td>
<td>331/28</td>
</tr>
<tr>
<td><strong>Full-term pregnancies (status and no.)</strong></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>57/24</td>
</tr>
<tr>
<td>Ever</td>
<td>1616/229</td>
</tr>
<tr>
<td>1–2</td>
<td>279/59</td>
</tr>
<tr>
<td>3–4</td>
<td>450/70</td>
</tr>
<tr>
<td>5</td>
<td>887/100</td>
</tr>
<tr>
<td><strong>Smoking (status and amount)</strong></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1265/218</td>
</tr>
<tr>
<td>Ever</td>
<td>409/36</td>
</tr>
<tr>
<td>1–5 cigarettes/day</td>
<td>181/17</td>
</tr>
<tr>
<td>6 cigarettes/day</td>
<td>211/18</td>
</tr>
</tbody>
</table>

OC, oral contraceptive

ORs adjusted for centre, age (<37, 37–45, 46–55, 56+), educational level (none, primary, secondary or higher), smoking amount (never, 1–5 cig./day, 6 cig./day+), age at first sexual intercourse (<17, 17–18, 19–22, 23+), lifetime number of sexual partners (1, 2–3, 4+), OC use (never, 1–4 years, 5–9 years, 10 years+), lifetime number of Pap smears (0, 1–5, 6+), and parity (0, 1–2, 3–4, 5–6, 7+). Adapted from Castellsagué & Muñoz (2003)
HPV and cervical neoplasia

Temporality

Cross-sectional studies have repeatedly found that sub-clinical HPV infections are highly prevalent in young individuals, whereas invasive cervical cancer typically develops in the third decade or later (Figure 18). The cross-sectional prevalence of HPV DNA decreases spontaneously to a background level of 2–8% in most populations in groups aged 35 years and above. In countries where intensive screening of young women takes place, part of the reduction in HPV prevalence may be attributable to aggressive treatment of HPV-related cervical lesions. Women who remain chronic HPV carriers are now considered the true high-risk group for cervical cancer. In some populations, a second peak of HPV DNA prevalence has been observed for older women (i.e., 50 years and above) and a second peak in the incidence of CIN 3 lesions and of invasive cervical cancer has also been reported (Herrero et al., 2000; Lazcano-Ponce et al., 2001). In all settings investigated, the point prevalence of HPV DNA in the young age groups is strongly related to the dominant sexual behaviour patterns in the population (Bauer et al., 1993; Melkert et al., 1993; Bosch et al., 1994; Kjaer et al., 1997; Jacobs et al., 2000; Schneider et al., 2000).

These population studies provide support for the concept that HPV infections precede the development of cervical cancer by some decades. Data from most cancer registries, including those in the USA, have established that the age-specific incidence of cervical cancer has a rising trend in the age interval 20–40 years, and shows a plateau or continues to increase smoothly after that age. Only occasionally do cases of invasive disease occur at earlier ages. Figure 18 shows the age-specific cross-sectional prevalence of high-risk HPV DNA in a screening programme in The Netherlands and the corresponding age-specific incidence rates of cervical cancer in that country. Very similar curves can be seen for other high- and low-risk countries (Bosch et al., 1992; Muñoz et al., 1992; Parkin et al., 1997; Jacobs et al., 2000).

Figure 18 Age-specific prevalence of high-risk HPV DNA in 3700 women entering a screening programme and age-specific incidence rate (ASIR) of cervical cancer in The Netherlands

Sources of data: Jacobs et al. (2000); Parkin et al. (1997)
From Bosch et al. (2002) (reproduced with permission from the BMJ Publishing Group)
Cervical cancer and screening

For the relationship between cervical cancer and HPV, compliance with the temporality criteria has been established by numerous cohort studies that monitored women from cytological normality to the stage of high-grade cervical intraepithelial neoplasia (HSIL or CIN 2 and 3). Continuing to monitor women to invasive disease is not acceptable on ethical grounds and thus information is not available.

Repeated sampling of women being followed for viral persistence and cervical abnormalities has shown that the median duration of infections detected at study enrolment is around eight months for high-risk HPV types (HPV16 in particular), compared with 4.8 months for the low-risk HPV types. In two unrelated studies, the time estimates were fairly consistent (Ho et al., 1998; Franco et al., 1999). In one study of incident HPV infections in Brazil, the mean duration of HPV detection was 13.5 months for high-risk HPV types and 8.2 months for the non-oncogenic types. HPV16 tended to persist longer than the average for other high-risk types (Franco et al., 1999). The results were remarkably similar in student populations in the USA and the United Kingdom (Ho et al., 1998; Woodman et al., 2001). The self-limiting course of most HPV infections is consistent with the cross-sectional profile displayed in Figure 18. However, the currently observed time intervals may still suffer from imprecision in the estimates of time at first exposure, from variability in the end-point definition and from censoring due to treatment of early lesions.

Follow-up studies of women with and without cervical abnormalities have indicated that the continuous presence of a high-risk HPV is necessary for the development, maintenance and progression of progressive CIN disease (Koutsky et al., 1992; Ho et al., 1995; Remmink et al., 1995; Ho et al., 1998; Nobbenhuis et al., 1999). A substantial fraction (15–30%) of women with high-risk HPV DNA who are cytologically normal at recruitment will develop CIN 2 or CIN 3 within the subsequent four years (Koutsky et al., 1992; Rozendaal et al., 1996, 2000). Conversely, among women found to be negative for high-risk HPV DNA and cytologically identified as having either ASCUS, borderline or mild dysplasia, CIN 2 or 3 is unlikely to develop during a follow-up of two years and their subsequent cytological results are likely to return to normal (Nobbenhuis et al., 2001a; Zielinski et al., 2001a). Women found positive for low-risk HPVs rarely become persistent carriers and their probability of progression to CIN 2 or 3 is extremely low (Manos et al., 1999; Zielinski et al., 2001a).

As existing cohorts extend their follow-up time, more precise estimates are being obtained on the predictive

<table>
<thead>
<tr>
<th>Table 9. Causality criteria and their fulfilment by the association of HPV DNA and cervical cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Criterion</strong></td>
</tr>
<tr>
<td>Time sequence</td>
</tr>
<tr>
<td>Experimental (prevention)</td>
</tr>
<tr>
<td>Strength and consistency</td>
</tr>
<tr>
<td>Biological plausibility and coherence</td>
</tr>
<tr>
<td>Dose–response</td>
</tr>
</tbody>
</table>

**Qualification of causality**

| Necessary | Exposure is present in all cases | Detailed investigation on ‘HPV-negative’ cervical cancer specimens. Exclusion of alternative explanations | ++ |
| Sufficient | Exposure always leads to disease | Natural history of transient infections | _ |

OR: Odds ratio; RR: Relative risk; CIN: cervical intraepithelial neoplasia

Table adapted from Bosch et al. (2002)
value of viral persistence as defined by repeated measurements of viral types and variants. In one such cohort in São Paulo, the incidence of cervical lesions in women who were HPV-negative twice was 0.73 per 1000 woman-months. The corresponding incidence among women who were twice HPV-positive for the same oncogenic types was 41.2 (95% CI 10.7–158.3) (Schlecht et al., 2001). Retrospective assessment of HPV DNA status using archival smears from cases of cervical cancer and controls has provided evidence that HPV infection preceded the development of invasive disease, and showed its value in signalling false negative cytological results (Zielinski et al., 2001a). It was also suggested that clearance of high-risk HPV in otherwise established cytological lesions is a marker associated with regression of CIN lesions (Nobbenhuis et al., 2001a; Zielinski et al., 2001b). Finally, persistence of HPV DNA detection after treatment for CIN 2 or 3 is an accurate predictor of relapse (Nobbenhuis et al., 2001b).

These results are useful in defining the clinical role of HPV testing. However, most observations on preinvasive disease have limitations for making inferences on cervical cancer causality. This is because even in controlled settings, observations cannot be allowed to continue beyond the stage of HSIL/CIN 3 or carcinoma in situ.

A valuable approach to conducting follow-up studies of invasive cancer (as opposed to studies of CIN 3) without ethical and time constraints is provided by nested case–control studies. These are studies initiated several years in the past that assembled and stored large banks of biological specimens from healthy individuals. Linkage of a serum sample bank to cancer registry data can then identify cases of cervical cancer (or any other condition) that have occurred in the interval; the original specimens can be analysed for the presence of HPV biomarkers. HPV DNA prevalence can then be compared with the corresponding prevalence in specimens from epidemiologically suitable controls (individuals from the same cohort who did not develop the condition under otherwise equivalent exposures). Such studies have documented the existence of HPV exposure years before the development of the disease, thus reproducing the conditions of a longitudinal study. With this approach, a relative risk estimate of 16.4 (95% CI 4.4–75.1) was observed for invasive cervical cancer in Sweden using DNA extracted from stored Pap smears (Wallin et al., 1999) and of 32 (95% CI 6.8–153) in The Netherlands (Zielinski et al., 2001a). In a study of similar design, an OR of 2.4 (95% CI 1.6–3.7) was obtained using serological markers of HPV exposure (Dillner et al., 1997).

**Strength of the association**

Figure 19 shows the HPV DNA prevalence in cervical cancer cases and controls in eight countries, from the IARC multicentric case–control study (Bosch et al., 2002). Table 10 shows the numbers of subjects in the study, the prevalence of HPV DNA in each relevant group, and the OR estimates. The first two studies, conducted in Spain and Colombia, used for HPV detection early versions of the MY09/11 PCR system that identified HPV DNA in some 75% of the cases. The rest of the studies were analysed using the GP5+/6+ PCR system and its modifications, which resulted in an almost 20% increase in the HPV DNA prevalence.
Cervical cancer and screening

detection rate (Muñoz et al., 2003). The ORs for squamous-cell carcinomas were statistically significant and very high. Restricting the analyses to studies that used the GP5+/6+ HPV detection system, the adjusted OR for HPV DNA detection (the factor by which the reference risk of cervical cancer is multiplied if HPV DNA is detected) was 158.2 for any single type (95% CI 113.2–220.6). The risk of adeno- or adenosquamous-cell carcinoma in eight countries (Algeria, Brazil, India, Morocco, Paraguay, Peru, The Philippines, Thailand) was estimated to be 77.2 (95% CI 41.2–144.8) (F.X. Bosch, personal communication).

The pool of IARC studies was large enough to provide type-specific risk estimates for 18 HPV types. Type-specific risk estimates and confidence limits are displayed in Figure 20 (Muñoz et al., 2003). These studies led to the conclusion that HPV types 16, 18, 31, 33, 35, 39, 45, rf51, 52, 56, 58, 59 and 68 should be considered high-risk carcinogenic types. Some evidence was also reported on a significant risk for HPV73 and 82. A second group of HPV types rarely found in cases was classified as low-risk, including HPV types 6, 11, 40, 42, 43, 44, 54, 61, 70, 72, 81 and CP6108.

Figure 20 shows that the risk for any given high-risk type was not statistically different from the risk reported for HPV16. Likewise, the risk related to the presence of multiple HPV types in the specimen was no different from the risk linked to a single HPV type. The standard estimates of the attributable fraction (AF %; the proportion of disease that is related to HPV DNA) derived from these and most other studies range from 90 to 98%.

Under extremely rare circumstances, HPV of the low-risk group (HPV6 or 11) was found as the only type in specimens of invasive cervical cancer. Although statistically the increases in risk are largely non-significant, it should be considered that low-risk types also show carcinogenic capacity under special but as yet unidentified conditions, at a very low level. It is plausible that a minute fraction of the population harbours a special susceptibility to HPV and even the presence of a low-risk type is capable of initiating a neoplastic process.

The results of the multicentre study are consistent with findings on invasive cervical cancer and preinvasive disease in Costa Rica (Herrero et al., 2000), Thailand (Thomas et al., 2001b), Norway (Olsen et al., 1995), Denmark (Kjaer et al., 1996) and virtually all other countries in which such studies have been conducted (Figure 21).

The proportion of specimens containing multiple HPV types varies across studies and particularly in relation to the HPV detection method used. Table 11 (adapted from Bosch et al., 2002) presents data on the proportion of specimens from cases and from the general population that showed multiple types. The table suggests that populations at high risk of cervical cancer and with high rates of HIV positivity tend to show higher proportions of multiple types as compared to populations not belonging to these groups. Longitudinal studies have suggested that the one-time, cross-sectional detection of type-specific HPV may underestimate the cumulative lifetime diversity of exposures to HPV (Woodman et al., 2001). However, in all studies of invasive carcinoma, the risk linked to multiple HPV types does not vary significantly from the risk linked to single HPV types.
Figure 22 shows the estimated percentages and numbers of cases of cervical cancer attributable to each HPV type, in all world regions combined. These were calculated by taking into account the estimated region-specific HPV genotype distribution and the number of incident cervical cancer cases (Munoz et al., 2004). Most of the cancer cases (70.7%) are accounted for by the two HPV types 16 and 18; the percentage rises to 87.4% when five other types (45, 31, 33, 52, 58) are considered. The 13 HPV types currently used for screening purposes seem to be responsible for 91.6% of all cancer cases.

**Exclusion of alternative explanations**

In the majority of studies of HPV and cervical cancer, a small fraction of cases is labelled as HPV-negative and these have been examined to assess whether HPV-negative cervical cancer is a true biological entity (Riou et al., 1990; Viladiu et al., 1997). The proportion of such cases tends to be greater in studies of preinvasive neoplasia (Burger et al., 1996; Tabrizi et al., 1999). In the IARC studies, it was clear that in broad terms, ‘HPV-negative’ cases retained the same epidemiological risk factor profile as the rest of the cases (i.e., similar age, high number of sexual partners, young age at first sexual intercourse, long-term use of contraceptives, high parity and similar prevalence of HPV16 antibodies). These results strongly suggested that the apparently HPV-negative cases were also STD-related; however, none of the sexually transmitted agents that had occasionally been associated with cervical cancer satisfied the causality criteria outlined in Table 9. In the evaluation of the putative HPV-negative cases, it was demonstrated that lack of identification of HPV DNA could be attributed primarily to the poor quality of the specimen (tumour necrosis, lack of cancer cells in the specimen, poor preservation) and to the quality of the amplification system used. Walboomers et al. (1999) showed that using histological verification of the specimen and the GP5+/6+ testing system, HPV DNA could be

<table>
<thead>
<tr>
<th>Table 10. Risk of squamous-cell cervical cancer associated with HPV DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Brazil</td>
</tr>
<tr>
<td>Mali</td>
</tr>
<tr>
<td>Morocco</td>
</tr>
<tr>
<td>Paraguay</td>
</tr>
<tr>
<td>Philippines</td>
</tr>
<tr>
<td>Thailand</td>
</tr>
<tr>
<td>Peru</td>
</tr>
<tr>
<td>Total invasive(^a)</td>
</tr>
<tr>
<td>Spain</td>
</tr>
<tr>
<td>Invasive</td>
</tr>
<tr>
<td>Carcinoma in situ</td>
</tr>
<tr>
<td>Colombia</td>
</tr>
<tr>
<td>Invasive</td>
</tr>
<tr>
<td>Carcinoma in situ</td>
</tr>
</tbody>
</table>

\(^a\) OR adjusted for age
\(^b\) OR adjusted for age and centre

From Muñoz et al. (2003)

<table>
<thead>
<tr>
<th>Table 11 Prevalence of multiple HPV types in cervical cancer cases and women without cervical cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Muñoz et al. (2000)</td>
</tr>
<tr>
<td>Herrero et al. (2000)</td>
</tr>
<tr>
<td>Castellsagué et al. (2001)</td>
</tr>
<tr>
<td>de Sanjosé et al. (2000)</td>
</tr>
<tr>
<td>Palefsky et al. (1999)</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Table adapted from: Bosch et al. (2002)
Cervical cancer and screening

detected in 99.6% of cases of cervical cancer worldwide, supporting the concept that HPV is indeed a necessary cause of the disease.

In the last decade there has not been any hypothesis supported by sound epidemiological or biological data indicating an HPV-independent etiology of cervical cancer, but such a hypothesis should be retained as a scientific and research option, for four reasons. (1) Epithelial cells are capable of developing into cancer cells in all human tissues regardless of a known, viral or non-viral, cause, so cells in the human cervix might too. (2) Cellular genes involved in HPV-related carcinogenesis should be liable to spontaneous or induced mutations that could lead to cancer in the absence of HPV, though available evidence suggests that this event is rare within the life expectation of the human population. (3) It is likely that any non-HPV-related cancers occur very rarely and probably cluster in very old women, but relatively few cases of cervical cancer in very old women have been investigated. (4) Non-epithelial cancers do occur in the cervix at a low frequency.

HPV transmission
Several groups of studies have clearly shown that HPV is predominantly and largely transmitted through sexual intercourse. Other forms of transmission, especially from mother to child, are briefly outlined below, but their implications in cervical cancer are likely to be no more than marginal. The evidence for non-sexual transmission of HPVs (reviewed by Cason, 1996; Mant et al., 2000), suggests that:

1. Genital HPV infections, including genital warts, may occur in populations without sexual experience, such as virgins, infants and children.
2. There is some evidence of non-genital transmission of low-risk HPVs.
3. Vertical and perinatal transmission of HPVs from mother to child does occur, although rates are generally very low.
4. High-risk genital HPVs have been detected in non-genital mucosa, such as that in the mouth, oropharynx and conjunctiva, and they have

Figure 21  Odds ratios and 95% confidence intervals for associations found in case–control studies after 2000
Sources of data: Bosch et al. (2000), Herrero et al. (2000), Josefsson et al. (2000), Muñoz et al. (2000), Thomas et al. (2001c), Ylitalo et al. (2000a)
From Bosch et al. (2002) (reproduced with permission from the BMJ Publishing Group).
been associated with a fraction of cancers of the oral cavity and oropharynx and with conjunctival squamous-cell carcinoma.

5. There is low concordance of HPV types and HPV16 genomic variants between heterosexual partners.

**HPV and sexual behaviour**

Epidemiological studies of risk factors for HPV infection have clearly and consistently shown that the key determinants among both women and men are related to their sexual behaviour. The best studied risk factors are their lifetime number of sexual partners, the age at which sexual intercourse was initiated and the likelihood that at least one of the sexual partners was an HPV carrier as measured by his sexual behaviour traits (Bauer et al., 1993; Hildesheim et al., 1993; Wheeler et al., 1993; Franco et al., 1995; Muñoz et al., 1996a; Kjaer et al., 1997; Rousseau et al., 2000; Silins et al., 2000; Kjaer et al., 2001). The role of males as possible vectors of HPV was measured in early epidemiological studies by questionnaires that addressed the sexual behaviour of the husbands or sexual partners of cervical cancer cases and controls. Later studies incorporated measurements of HPV DNA in exfoliated cells from the penile shaft, the coronal sulcus and the distal urethra (Barrasso et al., 1987; Kjaer et al., 1991; Bergman & Nalick, 1992; Bosch et al., 1996; Castellsagué et al., 1997). Figure 23 shows for both sexes the relationship between HPV DNA prevalence and either number of partners or age at first intercourse. In populations where female monogamy is dominant, the population of female sex workers plays an important role in the maintenance and transmission of HPV infections. Figure 24 shows the prevalence of HPV DNA in surveys of the general population in two European countries and in selected high-risk groups of sexual workers and imprisoned women from the same underlying populations. Figure 25 shows the correlation between the number of sexual partners and the prevalence of HPV DNA in the penis for husbands of monogamous and non-monogamous women.

The probability that a woman is an HPV carrier and her risk of developing cervical cancer have been consistently related to the presence of HPV DNA in the penis or the urethra of her husband or sexual partner (Kjaer et al., 1991; de Sanjosé et al., 1993; Bosch et al., 1994; Juarez-Figueroa et al., 2001; Thomas et al., 2001a). These and many other observations have consistently confirmed in terms of HPV infections the century-old observations and the hypothesis formulated 30 years ago that male sexual behaviour is a central determinant of the incidence of cervical cancer (Rigoni-Stern, 1842; Beral, 1974; Skegg et al., 1982).

Studies in couples have provided consistent evidence of the venereal nature of HPVs. Pridan and colleagues first showed an association between the number of sexual partners of the husband and the risk of cervical cancer among mostly monogamous Jewish women (Pridan & Lilienfeld, 1971). The male factor hypothesis was formulated soon after (Singer et al., 1976; Skegg et al., 1982). Geographical clustering of cervical and penile cancers (Smith et al., 1980; Franco et al., 1988; Bosch & Cardis, 1990) provides ecological support for the importance of men in the natural history of cervical cancer. Buckley et al. (1981) found that the risk of cervical cancer among monogamous women increased with the number of their husband’s sexual partners, and with the husband’s early age at first inter-

![Figure 22](https://example.com/figure22.png)
Cervical cancer and screening

The importance of the male role was also suggested by early studies of cancer clusters within couples. One study reported that subsequent wives of husbands whose previous wife developed cervical cancer had an increased risk of cervical neoplasia (Kessler, 1977), and other studies showed that wives of men with cancer of the penis had a high incidence and mortality due to cervical cancer (Martinez, 1969; Graham et al., 1979; Smith et al., 1980).

Data from the Swedish Family Cancer Database showed that husbands of women with in situ or invasive cervical cancer had an excess risk of anal cancer, a recognized HPV-related cancer (Hemminki & Dong, 2000). Anal cancer was also increased as a second primary cancer in women with cervical neoplasia (Hemminki et al., 2000a). Of special interest is the excess risk of both tonsilar cancer and cancer of the tongue in husbands of cervical cancer patients, supporting the evidence that HPV may be etiologically involved in a fraction of these tumours (Hemminki et al., 2000b, Herrero et al., 2003).

Quantitative evidence of the male role has been provided by formal case–control studies comparing either direct histories of sexual behaviour or clinical evidence of HPV-related lesions in male partners of women with and without cervical cancer (Zunzunegui et al., 1986; Brinton et al., 1989b; Kjaer et al., 1991; Bosch et al., 1996; Muñoz et al., 1996b). Huynh et al. (2004) showed an increased risk among women in Viet Nam in relation to the presence of their husbands in the army stationed in the south.

There is some evidence that HPV types differ in sexual transmissibility, with oncogenic (high-risk) types and non-oncogenic (low-risk) types having somewhat different risk factor profiles.

Figure 23  Prevalence of cervical and penile HPV DNA by lifetime number of sexual partners and age at first sexual intercourse in subjects without genital neoplasia (IARC studies)

Data from a series of 12 case–control studies of cervical cancer carried out by IARC in 10 countries. Data for females based on 2225 control women from studies in Algeria, Brazil, Colombia, India, Morocco, Paraguay, Peru, the Philippines, Spain and Thailand. Data for males based on 1140 men from studies in Brazil, Colombia, the Philippines, Spain and Thailand.

Adapted from Castellsagué & de Sanjosé (2003)
when considered as groups. The associations with number of sexual partners are stronger for the oncogenic types than for non-oncogenic types (Franco et al., 1995; Kjaer et al., 1997; Rousseau et al., 2000).

Most studies of concordance of genital HPVs in heterosexual couples, but not all (Baken et al., 1995), have found a relatively poor correlation of HPV-positivity and HPV type in cervical and penile samples (Hippelainen et al., 1994a; Kyo et al., 1994; Strand et al., 1995; Castellsagué et al., 1997; Franceschi et al., 2002). This is particularly important in case–control studies, in which a woman with cervical neoplasia is expected to be a long-term carrier of HPV DNA, whereas the husband is likely to have been a transient HPV DNA carrier (Hippelainen et al., 1994b). Moreover, in some couples, the current partner may not be the relevant one in determining the woman’s risk of HPV infection. Agreement in HPV findings, however, was also modest in couples where both the wife and husband reported only one lifetime sexual partner (Franceschi et al., 2002). Among women with cervical neoplasia, the relevant infection may have occurred years earlier. The relatively low prevalence of penile

**Figure 24** Prevalence of cervical HPV DNA in different risk groups in Denmark and Spain
Data for Denmark (adapted from Kjaer et al., 2000) include 182 female sex workers ( ), 187 female sexually transmitted disease clinic attendees ( ) and 1000 women from the general population ( ). Data for Spain (adapted from de Sanjosé et al., 2000, 2003 and Touze et al., 2001) include 187 female sex workers, 153 incarcerated women ( ) and 1101 women from the general population ( ) (Adapted from Castellsagué & de Sanjosé (2003))

**Figure 25** Penile HPV DNA prevalence by number of sexual partners of husbands of monogamous and non-monogamous women in Spain and Colombia.
Data (adapted from Castellsagué et al., 1997) include 595 men who were husbands or stable coital partners of women with and without cervical cancer.
HPV infection in their husbands suggests that viral shedding of advanced cervical lesions is limited. In addition, cross-sectional detection of penile HPV DNA may measure relatively recent exposures to HPVs that may be unrelated to the initiation of cervical neoplasia in the wife. Finally, the low agreement may be partly due to technical reasons, since a smaller amount of penile exfoliated cells is usually obtained from men compared with the cellular yield obtained from the cervix.

Male circumcision, penile HPV and cervical cancer
The IARC multicentric study on cervical cancer compared penile HPV DNA prevalence in circumcised and uncircumcised men and estimated the woman’s risk of cervical cancer according to the husband’s circumcision status. Circumcised men were about three times less likely to harbour HPV DNA in their penis than uncircumcised men. Male circumcision also reduced the risk of both genital HPV infections and cervical cancer in the female partner, particularly and most strongly in women whose male consorts had a promiscuous sexual history (Castellsagué et al., 2002). Other studies have, however, failed to report a lower prevalence of HPV DNA in the penis of circumcised males (Weaver et al., 2004).

Mother-to-child and perinatal transmission of HPV
Non-sexual transmission of HPVs was first suggested in 1956, in a case report of a male child that developed symptoms of laryngeal papillomatosis and penile warts at three and six months after birth to a mother with condyloma (Hajek, 1956). Since then a large body of epidemiological data on perinatal transmission of HPVs has accumulated (reviewed in Cason, 1996; Mant et al., 2000). A carefully conducted large study concluded that the risk of perinatal transmission of HPVs, although present, is probably very low (<3%) (Watts et al., 1998).

Perinatal HPV transmission has been unequivocally demonstrated for recurrent laryngeal papillomatosis, a rare, potentially life-threatening condition associated with HPV types 6 and 11, the types most commonly detected in genital warts.

Some evidence of intra-uterine infection with HPVs is available (Tseng et al., 1992; Armbuster-Moraes et al., 1994; Favre et al., 1998; Tseng et al., 1998).

Non-sexual transmission of HPV
Since Fleming et al. (1987) reported on a five-year-old boy with HPV2-positive warts on the anus and hand, a number of other case series have confirmed the possible non-sexual transmission of HPVs, particularly low-risk HPV types (reviewed in Lacey, 1996). These findings raise the possibility that patients with genital warts may transfer genital HPVs not only to their sexual partners by finger–genital contact but also horizontally to their children (Sonnex et al., 1999).

Finger–conjunctiva HPV transmission has been suggested by studies reporting the presence of HPV DNA, predominantly type 16, in human ocular surface squamous neoplasias, including conjunctival carcinomas (reviewed by Newton, 1996). A study in Uganda, a high-risk area for this tumour, found a statistically significant association between high titres of HPV16 antibodies and conjunctival squamous-cell carcinoma (Newton et al., 2002).

Transmission of HPV in blood, breast milk and sperm
It is very unlikely that HPVs are transmitted via blood, as HPVs do not have a known viraemic phase, and no case of HPV detection in blood has been documented.

Transmission of HPV to infants via the process of breast-feeding has not been documented.

Evidence on the possible role of sperm as a vector for HPVs suggests that semen may be a transmitter of cell-associated HPV during the process of ejaculation (Ostrow et al., 1986; Nieminen et al., 1991; Chan et al., 1994; Kyo et al., 1994; Lai et al., 1997; Olatunbosun et al., 2001).

Papillomavirus types
Early attempts to classify human papillomaviruses (HPV) were based on the rather strict tropism of certain HPV types for cornifying squamous epithelium (cutaneous types, e.g., HPV1, 4, 10) or mucosal epithelium (mucosal types, e.g. HPV6, 16, 31), with some types strongly linked to distinctive clinical presentations. However, this classification is over-simple and is incorrect in some cases, as demonstrated by the presence of the mucosal type HPV6 in cornifying genital warts. Another attempt to group papillomaviruses is the separation into skin types causing warts (e.g., HPV1) and genital types affecting primarily the anogenital area (e.g., HPV6, 16, 18). Again this classification is rather artificial, because HPV16 can also be found in nail-bed carcinomas on the hands. The modern classification into different HPV genotypes is based on DNA nucleotide sequence differences within the coding regions of the genes for the E6, E7 and L1 proteins, with different genotypes distinguished by having less than 90% sequence homology in these regions (de Villiers, 1994; Chan et al., 1995). By this definition, over 130 different HPV types have been described to date. Further subdivision into sub-types is based on sequence homology of 90–98% and variants with over 98% sequence homology.

HPV intratype variants are defined as having more than 98% nucleotide
sequence identity, determined over the E6, E7 and L1 open reading frames (ORFs), with the reference sequence (Van Ranst et al., 1992; Myers et al., 1996).

Because of the low prevalence of some genotypes, as detected by different genotyping methods, it has been difficult to categorize these according to risk, so that only 11 genotypes (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56 and 58) have been consistently classified as high risk (Lörincz et al., 1992; Bosch et al., 1995; Walboomers et al., 1999). In a recent analysis using pooled data from 11 studies (Muñoz et al., 2003), genotypes 59, 68, 73 and 82 were newly identified as high risk, while types 26, 53 and 66 were designated as probable high-risk types. Another possibility hindering precise risk classification is that the test sensitivities of genotyping methods may differ for distinct genotypes. For example, the most widely used genotyping systems based on polymerase chain reaction (PCR), such as the MY09/11 and GP5+/GP6+ systems (see Chapter 2), differ in their detection threshold levels for different HPV types. The sensitivity of the GP5+/6+ systems appears to be lower for HPV52, 53, 58 and 61 than that of the MY09/11 system, whereas GP5+/6+ has higher sensitivity for detection of HPV35 (Qu et al., 1997).

Papillomavirus biology
Papillomaviruses are widespread among higher vertebrates, but have strict species-specificity; transmission from non-primates to humans has not been reported. In general they cause local epithelial infections, although with animal fibropapillomaviruses (e.g., bovine papillomavirus; BPV), infection can also be found in the dermis. Viral spread to distant body sites does not occur.

Papillomaviruses are small icosahedral particles with a diameter of 55 nm, belonging to the family of Papillomaviridae. They have no envelope and consist of a capsid composed of 72 capsomeres, which accommodates the viral genome. The capsomeres are composed of two structural proteins: the 57 kDa late protein L1, which accounts for 80% of the viral particle and is considered to be a group-specific antigen, and the 43–53 kDa minor capsid protein L2 (Pfister & Fuchs, 1994). Because of the absence of an envelope, papillomaviruses are relatively stable and resistant to desiccation, retaining viability extracellularly for at least one week (Roden et al., 1997). They are also resistant to organic solvents and heat treatment to 56°C causes only a minor loss of infectivity.

Infections with papillomaviruses may cause local cell proliferation, which becomes apparent in the form of benign tumours such as common warts, condylomas and cervical intraepithelial neoplasia. The majority of benign tumours spontaneously regress in immunocompetent patients. Inherited or induced immune deficiencies lead to higher persistence of infections. In the case of the high-risk HPV types (see below), persistent infection confers a high risk for progression of the primary tumours into carcinomas (zur Hausen, 2000).

Structure of the HPV genome
The HPV genome consists of a double-stranded 8-kbp DNA molecule, which is associated with cell-derived histone proteins that produce a nucleosome-like superhelical twisted structure. The relative arrangement of the 9 to 10 ORFs within the genome is conserved within all papillomavirus types. One speciality of papillomaviruses is that the partly overlapping ORFs are arranged on only one DNA strand. To increase their coding capacities, HPVs make use of polycistronic transcripts and fusion proteins from different reading frames. The genome can be divided into three regions: the long control region (LCR), the region of early proteins (E1–E8) and the region of late proteins (L1 and L2). In accordance with this, two RNA poly-A addition sites, one for the early protein transcripts and one for the late protein transcripts, are always present (Pfister & Fuchs, 1994). A diagram of the genome organization of human papillomaviruses is given in Figure 26 and the functions of the different ORFs are summarized in Table 12.

The size of the long control region varies from 500 to 1000 bp in different HPVs. There are no ORFs in this area of the genome, but it does contain several control elements which regulate HPV DNA replication and gene expression (Iftner, 1990).

The proteins of papillomaviruses
Transcription of the genes E6 and E7 is a consistent feature in cervical carcinomas and was the first indication of an important role for these genes in HPV-associated tumorigenesis (Schwarz et al., 1985; Androphy et al., 1987; von Knebel Doeberitz et al., 1988; Sedman et al., 1991; Goodwin et al., 2000; zur Hausen, 2002). The E6 and E7 genes of HPV16 and HPV18 have been confirmed as potent viral oncogenes; the transforming and immortalizing abilities of their products have been demonstrated in numerous experiments in tissue culture and animal models (Münger & Howley, 2002).

The E6 ORF encodes several small proteins of approximately 150 amino acids, with molecular weights of about 16–18 kDa. Because of the presence of a splice donor and two splice acceptor sites within the E6 ORF of high-risk anogenital HPV types, smaller E6 proteins (E6*I and E6*II) are produced, which may auto-regulate the E6 promoter itself (p97) that is responsible for their expression. No enzymatic
Cervical cancer and screening

function of E6 proteins has been demonstrated, but physical interactions with several cellular factors resulting in the deregulation of the cell cycle or interference with DNA repair have been described (Mantovani & Banks, 2001). The key activity of high-risk E6 proteins is their ability to inhibit the function of p53 (Scheffner et al., 1990; Werness et al., 1990). p53 is a sequence-specific transcriptional transactivator with a growth arrest function and is regarded as a tumour-suppressor protein, which is stabilized post-translationally (increase in protein half-life) in case of DNA damage (El Deiry et al., 1993). The binding of E6 to p53 proteins leads to enhanced ubiquitin-dependent degradation of p53. This results in a shortening of its half-life from 3 hours to 20 minutes, with a corresponding loss of its biological function (Scheffner et al., 1990; Mantovani & Banks, 2001). For the ubiquitination of p53, E6 needs a cellular protein called E6-associated protein (E6AP), which acts as an E3-ubiquitin-protein ligase and links ubiquitin to a lysine side-chain, forming a stable isopeptide bond (Huibregtse et al., 1991). In non-infected eukaryotic cells, the ubiquitin-mediated proteolysis of p53 is triggered by the mdm-2 protein (Hengst et al., 2001). However, in cells infected with high-risk HPV, the formation of the E6-p53-E6AP-complex replaces the normal regulation of p53 by mdm-2.

Independently of the E6AP-dependent degradation of p53, high-risk E6 proteins lead to a down-regulation of p53-dependent transcription, which can be explained by the targeting of CBP/p300, a p53 co-activator. Further, E6 appears to be able to activate the cellular enzyme telomerase in differentiated cells (Klingelhofer et al., 1996; Steenbergen et al., 1996; Mantovani & Banks, 2001). This enzyme counteracts the continuous shortening of the chromosome’s telomeres which naturally occurs during replication of the cellular genome. Telomere shortening correlates with cell ageing and telomerase activity correlates with an increased life-span of the affected cell.

The E7 ORF encodes a small phosphoprotein of about 100 amino acids (10 kDa). E7 is a proliferation-inducing HPV oncogene and its activity is mediated through its ability to bind cellular proteins of the pRB family, which, in concert with the E2F family of transcription factors, control the transition of the cell cycle from the G1- to the S-phase (Dyson et al., 1989; Münger et al., 2001). Binding of E7 to the hypophosphorylated, active form of pRB leads to the activation of E2F transcription factors, permitting progression of the cell into the S-phase of the cell cycle with subsequent cell replication (Chellappan et al., 1992; Boyer et al., 1996). Apart from this proliferative capacity mediated by specific sequences within the N-terminus of E7, the E7 proteins of ‘low-risk’ types possess a tenfold lower efficiency of binding to pRB than ‘high-risk’ E7 proteins and are very inefficient in cell transformation assays together with activated ras oncogene (Gage et al., 1990; Münger et al., 2001). Chimeric substitution assays have attributed the difference in pRB-binding affinity and transforming capacity between low-risk and high-risk HPV types to the exchange of a single amino acid (Münger et al., 2001).

Forced entry into the S-phase is necessary for the virus to generate an environment that allows amplification of the viral DNA; this induces a number of cellular responses, like stabilization of the p53 protein which would lead to programmed cell death. To counteract
these cellular responses, high-risk papillomaviruses encode the E6 protein, which causes degradation of p53.

**Replication cycle in the infected epithelium**

The initial infection by HPV probably occurs in stem cells of the basal layer of stratified cervical epithelium or in associated hair follicles of the skin (Stanley, 1994; Schmitt et al., 1996). HPV genomes are then established as extrachromosomal elements in the nucleus. Upon cell division, one of the daughter cells stays in the basal layer and provides a reservoir of viral DNA, while the other cell migrates away from the basal layer and initiates a programme of differentiation. This leads to amplification of the viral DNA, expression of capsid proteins and finally the production of progeny virus. Since HPVs rely on cellular enzymes to replicate their genome, one major consequence of an HPV infection is a blockage of cell cycle exit. HPV-infected suprabasal cells undergo an incomplete S-phase to replicate HPV genomes to high levels (Laimins, 1996). With the high-risk HPV types, the blockage of cell cycle exit and induction of S-phase in differentiated suprabasal cells is mediated by the E6 and E7 proteins (Halbert et al., 1992; Cheng et al., 1995; Ruesch & Laimins, 1998).

HPVs maintain their genome at 10 to 100 virus copies per infected cell over long periods of time in vitro and this is thought to reflect the replication of viral DNA in basal cells in vivo (Laimins, 1996). Disturbances in the control of replication of high-risk HPV may have implications for the progression of high-risk HPV-induced lesions in vivo, as the viral DNA is extrachromosomal in precursor lesions, but is frequently found integrated into the host chromosomes in invasive cancers (Cullen et al., 1991). As no common integration site(s) have been identified, integration does not generally target proto-oncogenes or tumour-suppressor genes of the host cell. On the other hand, deletions and rearrangements of the integrated viral DNA occur. A model has been proposed which suggests that inactivation of the E2 gene releases E6/E7 oncogene expression from negative control (zur Hausen, 2002). However, no evidence has yet been presented that increased E6/E7 expression is indeed necessary for the progression of HPV-induced lesions. Viral DNA integration could simply be a consequence of an environment that does not support HPV

<table>
<thead>
<tr>
<th>Viral protein/genomic element</th>
<th>Molecular weight/size</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-coding elements</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long control region (LCR)</td>
<td>500–1000 bp</td>
<td>Origin of replication and regulation of HPV gene expression</td>
</tr>
<tr>
<td><strong>Early proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td>68–85 kDa</td>
<td>Helicase function; essential for viral replication and control of gene transcription</td>
</tr>
<tr>
<td>E2</td>
<td>48 kDa</td>
<td>Viral transcription factor; essential for viral replication and control of gene transcription; genome segregation and encapsidation</td>
</tr>
<tr>
<td>E3</td>
<td>Unknown</td>
<td>Function not known; present in only a few HPVs</td>
</tr>
<tr>
<td>E1–E4</td>
<td>10–44 kDa</td>
<td>Binding to cytoskeletal protein</td>
</tr>
<tr>
<td>E5</td>
<td>14 kDa</td>
<td>Interaction with EGF/PDGF receptors</td>
</tr>
<tr>
<td>E6</td>
<td>16–18 kDa</td>
<td>Interaction with several cellular proteins; high-risk HPV type E6 causes degradation of p53 and activate telomerase</td>
</tr>
<tr>
<td>E7</td>
<td>~ 10 kDa</td>
<td>Interaction with several cellular proteins, like with pRB and transactivation of E2F-dependent promoters</td>
</tr>
<tr>
<td>E8–E2C</td>
<td>20 kDa</td>
<td>Long-distance transcription and replication repressor protein</td>
</tr>
<tr>
<td><strong>Late proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>57 kDa</td>
<td>Major capsid protein</td>
</tr>
<tr>
<td>L2</td>
<td>43–53 kDa</td>
<td>Minor capsid protein</td>
</tr>
</tbody>
</table>
DNA replication. This is supported by fact that long-term extrachromosomal replication of high-risk HPV DNA has not been achieved in established HPV-positive or -negative tumour cell lines, but occurs almost exclusively in normal human keratinocytes (Frattini et al., 1996; Del Vecchio et al., 1992).

### Immunity to HPV

Numerous studies have found a positive association between the detection of HPV antibodies and the risk of cervical neoplasia, in line with the notion that HPV antibody detection is a marker of cumulative exposure to HPV. Although these antibodies, particularly those directed against the virion capsid proteins L1 and L2, might be effective in preventing infection, it is commonly accepted that antibodies are not important effectors of the regression of established HPV infections and related cervical lesions. Neutralizing antibodies are generated by a type-specific conformational epitope in the viral particle. In contrast, disrupted or partially disrupted viruses expose epitopes that are broadly cross-reactive or even group-specific (Cowsey et al., 1987; Christensen et al., 1996). HPV types 6 and 11 are an exception, having been shown to contain shared epitopes and type-specific epitopes on intact capsids (Christensen et al., 1994, 1996). Seroconversions against the HPV16 capsids are seen concomitantly with or within a few months after acquisition of HPV16 DNA (Andersson-Ellstrom et al., 1994, 1996; Wikstrom et al., 1995a, b; Carter et al., 1996), but in a subset of patients can be delayed many months after the detection of viral DNA. In Sweden, the risk for seroconversion to the major oncogenic HPV type, HPV16, was found to increase linearly by about 4% for each life-time sexual partner up to a plateau of about 32% among women with an average of eight lifetime sexual partners (Dillner et al., 1996). Low seroprevalence (2–7%) has often been found in monogamous women (Andersson-Ellstrom et al., 1994, 1996; Carter et al., 1996; Dillner et al., 1996; Wideroff et al., 1996; Viscidi et al., 1997; Kjellberg et al., 1999). HPV seropositivity in adult virginal women has not been reported, though the total number of virginal women tested is not large (Andersson-Ellstrom et al., 1994, 1996). Large-scale surveys among children aged under 13 years found HPV seroprevalences of the order of 2% (Mund et al., 1997; af Geijersstam et al., 1999).

The major isotypes of serum antibodies against HPV capsids are IgG1 and IgA (Wang et al., 2000). The serum IgA response is also HPV type-specific, as demonstrated by its correlation with the presence of type-specific HPV DNA (Wang et al., 2000). Secretory IgA antibodies to HPV capsids are detectable in cervical mucus. In contrast to serum IgG, however, serum IgA correlates with the number of recent sexual partners and with the life-time number of partners, mostly among young women (Wang et al., 2000), suggesting that the IgA response is less biologically stable over time than the IgG response. However, it is less clear whether antibodies against one HPV type protect against subsequent reinfection with the same or another closely related type and, if so, whether this protection is related to specific IgG and IgA subclasses. After the HPV infection has been cleared, serum antibody levels remain stable over time, even after 15 years of follow-up (Shah et al., 1997).

The cellular immune response is an important effector mechanism for the clearance of established HPV infections. The first line of defence is the immune response with natural killer (NK) cells inducing apoptosis in virus-infected cells and in tumour cells. The specific activity of NK cells requires so-called killer immunoglobulin-like receptors (KIRs) which enable them to distinguish normal from virally infected or tumour cells. A direct antiviral cellular immune response is mediated by cytotoxic T-cells that recognize and kill infected cells presenting viral peptides, with the help of human lymphocyte antigen (HLA) class I molecules on their surface.

### Viral and host risk factors

#### HLA haplotypes

The factors that determine whether an HPV infection is cleared or persists and that increase the risk for cervical cancer are not fully defined, but cellular immunity plays a major role. Altered HLA class I allele in cervical cancer has long been recognized and the presence of specific HLA class II alleles may be decisive for the risk for cervical cancer. In the case of HLA class I A2 (Montoya et al., 1998), B44 (Bontkes et al., 1998) and HLA-B7 (Duggan-Keen et al., 1996), negative associations have been described. The most likely underlying mechanism is allele-specific down-regulation of these antigens during cervical carcinogenesis. In addition, the existence of HPV16 variants with E6 mutations affecting HLA-A2 and -B7 binding motifs suggests that lack of CD8-restricted epitopes may enable the virus to escape the immune response (Ellis et al., 1995; Yamada et al., 1995). Many studies in humans have focused on the association of HLA class II with SIL or cervical cancer and several HLA class II haplotypes were found to be associated with disease, such as DQw3 increasing (Wank & Thomssen, 1991) and DRB1*1301 decreasing (Hildesheim & Wang, 2002) the risk for cervical cancer in general. Some associations were found to be type-specific; thus, DR15 increases the risk for HPV16-carrying cancer and DR7 may be either protective or increase the risk (Konya & Dillner, 2001).
**Cellular gene polymorphism**

Another seemingly important marker of risk is a single nucleotide polymorphism in codon 72 of the p53 gene. There are two structurally different forms of wild-type p53, containing either a proline (Pro) or an arginine (Arg) at amino acid residue 72 (Matlashewski et al., 1987). Storey et al. (1998) reported that the Arg p53 variant has higher susceptibility to HPV-E6-mediated degradation in tissue culture experiments, a discovery that could translate into a new cervical cancer risk factor. They also found that women with cervical cancer were more likely to be homozygous for Arg at position 72, rather than heterozygous or homozygous for Pro. It thus appeared that the Arg/Arg genotype at position 72 conferred susceptibility to HPV-induced cervical cancer and a few studies confirmed this association, but many more have failed to (reviewed by Koushik et al., 2004). Inter-laboratory variability in p53 genotyping has been proposed as a possible explanation for the null results (Makni et al., 2000). A recent meta-analysis using random-effects models that included 45 studies published since the original Storey et al. (1998) publication attempted to identify characteristics that significantly contributed to heterogeneity. For invasive cervical cancer with undefined histology, the Arg/Arg genotype was found not to affect risk (OR = 1.1; 95% CI 0.9–1.3). However, slightly increased risk was observed for squamous-cell carcinoma (OR = 1.5; 95% CI 1.2–1.9) and adenocarcinoma (OR = 1.7; 95% CI 1.0–2.7). Meta-regression analysis identified departures from Hardy–Weinberg equilibrium in the control group as the most important factor contributing to heterogeneity among results for invasive lesions. Summary ORs for studies in equilibrium were essentially null (Koushik et al., 2004). This study suggests a possible role of p53 codon 72 polymorphism at a late carcinogenic stage in cervical cancer. However, further investigations are required with appropriate attention to design, sample size and methodological issues.

**Loss of heterozygosity (LOH)**

Many studies have considered chromosomal abnormalities in cervical cancer. Chromosomal alterations have been consistently identified, such as LOH at chromosome 3p, 6, 11, 13, 16, 17 and 19, and chromosomal amplifications at 3q (Southern & Herrington, 1998; Lazo, 1999; Kaufmann et al., 2002); identification of the target genes (oncogenes or tumor-suppressor genes) affected in these areas is now required.

**Viral variants**

Although more than 100 HPV types have been identified, studies on variants in viral genes mainly relate to the E6 gene of HPV type 16 (Hecht et al., 1995; Lizano et al., 1997; Villa et al., 2000; Chan et al., 2002b). HPV16 variants with nucleotide alterations within the E6 gene, referred to as non-prototype-like variants, are reported to be more frequently associated with high-grade CIN and cervical cancer than wild-type genomes (Xi et al., 1997; Zehbe et al., 1998a), although this phenomenon could be population-dependent (Zehbe et al., 1998b; Nindl et al., 1999). On the basis of regional differences, HPV16 variants have been termed European (E), Asian-American (AA), African (Af1 and Af2) and North-American (NA) (Ho et al., 1991; Yamada et al., 1997).

Interestingly, a significant over-representation of G/T variants at position 350 (350G/T) of the HPV16 E6 gene was detected in cervical cancers of women with a p53 Arg/Arg polymorphism; a possible differentially oncogenic effect of HPV16 350G/T variants which is influenced by the p53 genotype was therefore suggested (van Duin et al., 2000). Another E6 variant (the 131G variant) was found to be present in 9.6% of cervical carcinoma patients (n = 94), of whom 78% had the HLA-B7 allele, already identified as a possible risk factor (Brady et al., 1999). Most of the studies performed did not consider other variations that may occur in the E6/E7 region or in other regions of the HPV genome. Therefore the current risk observed, which is associated with viral variants in general, might be an underestimate. Furthermore, this risk might be influenced by other genomic alterations.

**Viral load**

A number of cross-sectional epidemiological studies using the semiquantitative Hybrid Capture 2 (HC2) technique (Iftner & Villa, 2003) have demonstrated an association between viral load of high-risk HPV types and cervical cancer risk. However, estimates of viral copy numbers depend directly on the total input of cells and adjustment for cellular load is an absolute requirement that is frequently not fulfilled, as in the case of HC2. Using type-specific real-time quantitative PCR, Swan et al. (1999) and Ylitalo et al. (2000b) found that high viral load of HPV16 was associated with risk for progression. The number of HPV16 copies per µg of cellular DNA in patients with a normal cytological result (approximately 2 x 10^7) increased with the severity of the lesions, reaching a 100-fold increase in CIN 2/3 patients (Swan et al., 1999). Moreover, the risk of incident SIL increases with viral load and the progression or regression of a given cytological abnormality is correlated with higher or lower viral load, respectively (van Duin et al., 2002; Schlecht et al., 2003b); these associations have not been found for other HPV genotypes (Swan et al., 1997; Abba et al., 2003). Only few longitudinal data are available (Lörincz et al., 2002; Ylitalo et al.,...
While very low viral loads are associated with low subsequent risk, high viral loads are not necessarily associated with high risk of incident CIN 3 or cancer (Snijders et al., 2003). Moreover, the main determinant of viral load measurements from women with small CIN 3 lesions is the extent of surrounding, virus-producing CIN 1 and CIN 2 (Sherman et al., 2003c). Therefore, although HPV16 viral load appears to be correlated with cervical cancer, the utility of viral load for predicting progression from HPV infection to cancer remains unclear. In addition, little is known about the relationship of viral load of other types than HPV16 to cervical neoplasia.

**Viral DNA integration**
HPV DNA is maintained as an episome in benign infections, whereas integrated HPV genomes are frequently detected in CIN 3, cervical cancer and derived cell lines. It has been proposed that this integration event confers a certain growth advantage on the infected cells by activating the expression of the viral oncogenes (zur Hausen, 2000). The current model suggests that inactivation of the E2 gene as a consequence of integration releases E6/E7 oncogene expression from E2-mediated negative control. However, no evidence has yet been presented that increased E6/E7 expression is indeed necessary for the progression of HPV-induced lesions. In addition, a number of studies found exclusively episomal HPV16 DNA in 20–70% of cervical cancers (Cullen et al., 1991; Fuchs et al., 1989; Matsukura et al., 1989; Pirami et al., 1997) and in high percentages (75–97%) of CIN 3. Therefore it remains unclear whether HPV integration is simply a consequence of loss of normal epithelial cell differentiation capacity, without biologically conferring any further risk downstream, or whether the integration event does contribute to progression.

**Epigenetic events**
Epigenetic events are events that alter gene expression (e.g., phenotype) without a change in the DNA sequence; they include hypermethylation or hypomethylation of genes (e.g., the addition or the removal of a methyl group). The silencing of tumour-suppressor genes via promoter hypermethylation in HPV-infected host cells is a frequent human epigenetic event (Dong et al., 2001; Virmani et al., 2001). For example, silencing of the TSLC1 (tumour suppressor in lung cancer) gene is frequently seen in the progression from high-grade lesions to invasive cervical cancer (Steenbergen et al., 2004).

**Oral contraceptives and parity**
A recent meta-analysis showed a linear dose–response relationship of cervical cancer with hormonal contraceptives. The point estimate of the summary OR is between 2- and 3-fold; the duration of the effect after cessation of use of oral contraceptives remains to be determined (Smith et al., 2003).

High parity has been found consistently in most case–control studies to be associated with both cervical cancer and carcinoma in situ. Most of the major studies restricting analysis to HPV-positive women also report an increased risk of HSIL or cancer with increasing number of pregnancies. In the IARC multicentric study, the OR for cervical cancer in women with seven or more full-term pregnancies was fourfold higher than in nulliparous women, and the risk increased linearly with increasing number of full-term pregnancies (Muñoz et al., 2002). Risk of HSIL or cancer significantly increased with increasing number of live births in the Costa Rica study (Hildesheim et al., 2001). A borderline association with CIN 3 was found in the Manchester study (Deacon et al., 2000). The study in Denmark (Krüger-Kjaer et al., 1998) and the US prospective study (Castle et al., 2002a) did not find an association with the risk of HSIL and CIN 3 or cervical cancer, respectively. However, these results could be explained by the low parity of the study populations. In addition, oral contraceptives. The point estimate of the summary OR is between 2- and 3-fold; the duration or pack-years of smoking have shown an increased risk of cervical cancer with increasing exposure to tobacco smoking. A prospective study in the USA found a positive association with smoking status and smoking intensity (Castle et al., 2002a).

Malignant transformation of HPV-16-immortalized human endocervical cells by cigarette smoke condensate has been proven (Yang et al., 1996). The fact that nicotine and tobacco-specific carcinogens have been detected in the cervical mucus of smokers (Prokopczyk et al., 1997) strengthens the hypothesis of a synergistic action between cigarette smoking and HPV for the development of HSIL and cervical cancer. In a prospective study, smokers were found to maintain cervical HPV infections significantly longer and to have a lower probability of clearing an oncogenic infection than women who never smoked (Giulian et al., 2002). The significant association between the extent of smoking reduction and the reduction in lesion size found in an intervention study of
smoking cessation among women with minor-grade lesions further supports the role of tobacco smoking in HPV carcinogenesis (Szarewski et al., 1996).

**Herpes simplex virus type 2 and C. trachomatis**

The IARC multicentric case–control study investigated the presence of antibodies against the common sexually transmitted agents to assess their effect on cervical cancer risk in the presence of HPV DNA. The results show that among HPV-positive cases and controls there is a residual 1.5- to 2-fold increased risk linked to herpes simplex virus type 2 (HSV2) and C. trachomatis exposure, suggesting an interaction with the oncogenic capacity of HPV.

The pooled analysis of seven case–control studies included 1262 cases of invasive cancer and 1117 matched controls. Western blot analyses were used to detect type-specific antibodies to HSV types 1 and 2. As expected, seroprevalence was higher in cases than controls and the risk of squamous-cell carcinoma was significantly higher in analyses restricted to HPV DNA-positive cases and controls and adjusted for other possible confounders (OR = 2.19; 95% CI 1.41–3.40). The association was consistent in adeno- and adenosquamous-cell carcinoma (Smith et al., 2002b).

The prevalence of antibodies to C. trachomatis varies greatly by country and serum antibodies were associated with a 1.8-fold increased risk of squamous-cell invasive cervical cancer in all countries considered, except in Spain. The risk was higher in women with elevated C. trachomatis antibody titre and in women under 55 years of age. C. trachomatis and C. pneumoniae species-specific serum antibodies were differentiated using a microimmunofluorescence antibody assay. The increased risk of squamous-cell invasive cancer was found in women with C. trachomatis but not with C. pneumoniae antibodies. The study thus supports the possibility that C. trachomatis increases squamous-cell cervical cancer risk, after accounting for cervical HPV infection (Smith et al., 2004).

**Human immunodeficiency virus (HIV)**

The evidence for an interaction between HPV and HIV in the origin of cervical cancer has led to the recognition of cervical cancer as one of the criteria of acquired immunodeficiency syndrome (AIDS) among HIV-positive women. Subsequent studies largely confirmed this evidence, although some major confounders of the epidemiological association tend to obscure the results. In brief, these refer to the powerful impact of screening in some populations, the medical surveillance of HIV carriers in developed countries and the short survival time of AIDS patients in many populations at high risk of cervical cancer compared with the time interval between HPV infection and cervical cancer (Bayo et al., 2002; de Sanjosé & Palefsky, 2002; Gichangi et al., 2002).

Massad et al., (1999) reported on the baseline cervical cytology among 1713 HIV-positive and 482 HIV-negative women. Cervical cytology was abnormal in 38.3% of HIV-positive women versus 16.2% of the HIV-negative. High-grade lesions, low-grade lesions and ASCUS were all significantly more common among HIV-infected women. The risk factors for any abnormal cytological finding were a CD4+ count lower than 200 cells/µL (OR = 2.13; 95% CI 1.45–3.13), the presence of HPV DNA and a previous history of abnormal cytology.

Ahdieh et al., (2000) identified a higher baseline prevalence of cervical abnormalities among HIV-infected women (13.4%) compared with HIV-negative women (2.4%). In this follow-up study, 11 women were identified with CIN in subsequent visits, all of whom were HIV-positive and who had a median CD4+ count of 253 cells/µL. The risk for CIN was related to HPV persistence in all cases.

Thomas et al. (2001a) studied a group of 251 sex workers in Thailand. The HPV DNA prevalence was similar in HIV-positive and HIV-negative women. However, the risk of high-grade lesions was two-fold higher in women infected with both HPV and HIV than in the HIV-negative, HPV-positive women and 20 times higher than in HIV-negative, HPV-negative women.

Mandelblatt et al., (1999b) provided a pooled estimate of 15 studies on the association between HIV and CIN. HIV-infected women had an eight-fold increased risk of CIN (OR = 8.8; 95% CI 6.3–12.5). Sun et al., (1997) reported that, compared with HIV-negative, HPV-positive women, women co-infected with HPV and HIV had a lower regression rate of low-grade lesions and higher rates of progression from infection to CIN.

Ellerbrock et al., (2000) reported that HIV-positive women were 4.5-fold more likely than HIV-negative women to have or develop CIN within a 54-month follow-up interval. Among HIV carriers, transient HPV infections (RR = 7.4; 95% CI 1.0–57.4), persistent infections with HPV types other than 16 or 18 (RR = 8.9; 95% CI 1.2–66.2) and persistent infections with HPV type 16 or 18 (RR =11.0; 95% CI 1.4–88.7) were all significantly associated with CIN.

The International Collaboration on HIV and Cancer (2000) published cancer data from 23 prospective studies that included 47 936 HIV-infected subjects from North America, Europe and Australia for the period 1992–99. It was concluded that there had not been a significant change in the incidence of invasive cancer (rate ratio = 1.87; 99% CI 0.77–4.56) during this period.

An overview of early studies in Rwanda, South Africa and Uganda concluded that invasive cancer was not
related to exposure to HIV, with a summary OR of 0.8 (95% CI 0.5–1.4) (Newton et al., 1999). However, more recent data from a hospital-based case–control study in South Africa show increased risk for cervical cancer (OR = 1.6; 95% CI 1.1–2.3) and for vulvar cancer (OR = 4.8; 95% CI 1.9–12.2) among HIV-infected patients (Sitas et al., 2000).

Reports from the USA and Europe are generally consistent in detecting an increased risk for cervical cancer among HIV-infected women. Selik and Rabkin (1998) found a relative risk for cervical cancer of 5.5 among HIV-positive women in the USA. Frisch et al. (2000), using data from the US Cancer Match Registry for the period 1978–96, found an RR of 5.4 for invasive cervical cancer among HIV-positive women compared with the general US population. Similar increases in risk were observed for in situ cervical cancer (OR = 4.6), cancer of the vagina and vulva (OR = 5.8) and anal cancer (OR = 6.8). The risk showed no major change between the time before AIDS diagnosis and up to 60 months after diagnosis. Similar results are available for the population of New York (Gallagher et al., 2001).

In southern Europe, a strong association between invasive cervical cancer and AIDS has consistently been found. In Italy, the linkage of the National AIDS Registry and the population cancer registries showed a 15-fold increased cervical cancer risk in women with AIDS (Franceschi et al., 1998). The joint Italian–French follow-up study of HIV-positive women also showed a 13-fold increased rate of cervical cancer in HIV-positive women (Serraino et al., 1999). In Spain, the Catalonian AIDS surveillance system detected 58 cases of invasive cervical cancer among 823 HIV-positive women, an 18-fold increased risk compared with the general population (Mayans et al., 1999).

In summary, HPV and HIV share some behavioural traits that define a particularly vulnerable high-risk group. Progression of HPV infections to CIN lesions and cervical cancer in the context of limited or absent screening seems to be increased among HIV carriers and AIDS patients. Further-more, increasing evidence suggests that progression is related to the severity of immunosuppression, as indicated by CD4+ counts.

Dietary factors
Recent epidemiological evidence on the role of diet and nutrition on the risk of HPV persistence, SIL and invasive cervical squamous-cell carcinoma, taking into account HPV, has been systematically reviewed (Giuliano & Gapstur, 1998; Castle & Giuliano, 2003). Although there is some epidemiological support for a role of diet in cervical carcinogenesis, the evidence remains inconclusive. One of the most relevant new findings is a possible protective effect of fruit and vegetables on HPV persistence. In relation to nutrients, a protective effect against cervical neoplasia is considered probable for folate, retinol and vitamin E and possible for vitamins B12 and C and carotenoids. Conclusions for nutrients from studies taking HPV infection into account do not differ substantially from those that did not control for it. Overall, no clearly different patterns were observed for nutritional co-factors between low- and high-grade cervical lesions or between retrospective and prospective study designs.

No cohort studies on SIL and only few on HPV persistence have been reported which comprehensively assess suspected nutritional or dietary factors with control for HPV status.

Principles of screening
Screening was defined by the United States Commission on Chronic Illness (1957) as "the presumptive identification of unrecognized disease or defect by the application of tests, examina-

Cervical cancer and screening
3. Its purpose is to benefit the individuals being screened.

Those responsible for screening programmes also have an ethical responsibility to ensure that quality control of the screening tests is maintained and that the effectiveness of proven, beneficial programmes is continually monitored (Hakama et al., 1985).

Some other ethical issues are also important. The first is to reduce unnecessary anxiety to a minimum. This requires selecting screening methods that will provide the most attractive combination of negative predictive value (e.g., reassurance) and false positives that is attainable in a given setting. The second is to ensure that a useful remedy is available for all individuals identified as being true positive. There should be no one for whom either a definitive diagnostic test is not available or direct action cannot be offered. If this is not the case, screening will merely generate groups of anxious individuals for whom there is no benefit.

The process of screening should identify a test-negative group for whom no further action is warranted. For the test-positives, there should be a protocol which defines the diagnostic tests available and the treatment available for those with a true-positive test result. For cervical cancer, the protocol should define those referred immediately for treatment, and those for whom surveillance and repeat testing is recommended.

To attain these objectives, there is a need to ensure that those who are test-positive return for diagnosis and for treatment if found to be true-positives. A screening programme must ensure that there is sufficient contact with the individuals being screened to make them aware of the implications of a positive test result. Some provision should be made to ensure that they have somewhere to return to for further medical advice and if necessary, counselling. A screening programme that fails to take these considerations into account is failing in its duty of care.

Equity of access to screening services is another important consideration. All those who stand to gain from screening should have access to the procedure. A screening service should not be a service that relies on individuals seeking out particular tests or procedures that they have heard may be of value. Instead, those who organize the service have an obligation to ensure that those who have not heard of the test or procedure but who stand to benefit from it are adequately informed and located to enable them to be screened.

A final ethical issue concerns the extent to which the offer of screening in a community could divert resources from other, more important, health-care programmes. This is a particular problem for low-resource settings. There is an ethical responsibility to distribute limited resources equitably across the total community in order to obtain maximal health benefit. Under certain circumstances, the offer of screening could diminish the overall level of health in a community, if it resulted in fewer resources being available for other diseases. However, a well organized programme should promote equity through better use of resources.

**Natural history of cervical cancer**

For effective and efficient application of screening, “the natural history of the disease should be known” (Wilson & Junger, 1968). This is because screening is based on the expectation that the early detection of cancer, in the detectable preclinical phase (DPCP) (Cole & Morrison, 1980), will result in a reduction in mortality from the disease. If effective screening is directed primarily to the detection of precursors, the development of invasive cancer will be prevented.
Knowledge on the natural history of the disease will facilitate decisions on the appropriate ages to initiate and cease screening and on the optimal frequency of re-screening in those who test negative.

The concept that invasive squamous-cell carcinoma of the cervix arises from intraepithelial precursor lesions was first put forward over a century ago, on the basis of the histopathological identification of intraepithelial lesions immediately adjacent to frankly invasive cervical cancers that had morphological similarities to the invasive cancers. The recognition that there was a spatial relationship between certain intraepithelial squamous lesions and invasive cancers led to additional studies to define the temporal relationship between the intraepithelial and invasive cancers. A temporal relationship between carcinoma in situ and invasive cervical cancer was suggested by case reports of carcinoma in situ lesions of the cervix occurring several months to years before the subsequent development of invasive cervical cancer.

Putative precursor lesions can be identified through direct observation of the cervix with a colposcope after application of a solution of 3–5% acetic acid (see Chapter 2). This accessibility to direct observation using colposcopy and cytological/histological sampling has allowed the pathogenesis of cervical neoplasia to be intensively studied for the last 50 years. These investigations led to the identification of additional types of intraepithelial squamous lesions that are histologically less severe than carcinoma in situ. With our in-depth understanding of the central role of HPV in the pathogenesis of cervical cancer and the natural history of HPV infections (see above), we now have a coherent model for the natural history of cervical disease.

When cervical cancer screening programmes based on cervical cytology were being introduced, it was appreciated that the numbers of cases of presumed precursors were much larger than the numbers of invasive cancers occurring in the same population (Burns et al., 1968; Fidler et al., 1968; Coppleson & Brown, 1975). This led to a series of studies designed to clarify the natural history of the disease as identified by the abnormalities detected by cytology. The majority of these studies used post-biopsy histological diagnoses as their end-point. However, a few used cytological diagnoses, and these are specifically identified in the sections which follow. In interpreting the findings from these studies, it must be borne in mind that cytologically and histologically derived diagnoses are not necessarily identical, even if the same terminology is used (see earlier in this chapter). Because of the impossibility of directly observing the outcome of lesions treated surgically, their natural history has to be inferred by statistical techniques, usually by applying models of the presumed natural history, that are valid only to the extent that the assumptions that led to the model are valid.

In the 1960s, a variety of studies provided evidence interpreted as providing support for the existence of a spectrum of cytologically or histologically defined cervical cancer precursors. Studies using electron microscopy, time-lapse cinematography, ploidy analysis and DNA content led to a hypothesis that these intraepithelial lesions form a biological continuum from very early precursor lesions referred to as CIN 1 to more advanced lesions referred to as CIN 3. Although this morphology-based model of a continuum has now been supplanted by a more discrete theory of multi-stage carcinogenesis as described below, the CIN scale still merits consideration as the current basis of clinical management.

Findings in the 1970s that many CIN lesions presumed to be cervical cancer precursors had histological and cytological similarities to HPV-induced genital warts produced a profound change in the understanding of the pathogenesis of cervical neoplasia (Meisels & Fortin, 1976; Meisels et al., 1977; Purola & Savia, 1977). Subsequent studies have clearly shown that infection of the cervical epithelium with specific high-risk types of HPV plays a basic role in the pathogenesis of cervical cancer and its precursor lesions (see above).

Morphological appearances alone frequently do not permit the distinction of intraepithelial lesions that are associated with persistent high-risk HPV infections and have a substantial capacity to progress to invasive cervical cancer from those lesions that do not. Although the majority of histologically defined CIN 1 lesions demonstrating marked cytopathic effects represent transient lesions, it is important to recognize that some are associated with high-risk types of HPV and have biological features indicating that they may be true cancer precursors (Lungu et al., 1992). These features include monoclonality, aneuploidy and loss of heterozygosity (LOH) (Fu et al., 1988; Park et al., 1996; Chung et al., 2000). Similarly, it should be recognized that lesions that have the histological features of CIN 2, and a smaller subset of lesions with the histological features of CIN 3, are not aneuploid and do not show specific LOH. Therefore it is not possible to predict biological outcome solely on the basis of histopathological appearance.

Low-grade lesions
HPV infection of young women is frequent, and in the large majority of women transient. Repeated sampling of women being followed for viral persistence and cervical abnormalities has shown that the median duration of
a prevalently detected HPV infection is typically about eight months for high-risk types of HPV and 4.8 months for the low-risk types (see earlier in this chapter). If detected during follow-up following a negative cytological test, the median durations are doubled (Richardson et al., 2003).

Woodman et al. (2001) studied the natural history of HPV infections in a cohort of 1075 young women (15–29 years) in the United Kingdom, who agreed to undergo intensive cytology and HPV screening, and who on average had repeat tests every six months. This allowed accurate determination of the timing of HPV infections and the development of CIN that followed some of these infections. Somewhat surprisingly, of the 240 women who developed abnormal cytology and whose HPV status was known, 41% tested negative for HPV and another 34% only tested HPV-positive at the same visit as that in which the abnormal cytology was detected. Thus, for only 25% was a positive HPV test predictive of abnormal cytology, although this translated into a cumulative risk of 33% at three years.

Molano et al. (2003) showed in a cohort of women aged 13–85 years, based on 316 type-specific HPV infections, that HPV16 had a significantly lower clearance rate than low-risk types (RR = 0.47; 95% CI 0.32–0.72). HPV types related to HPV16 (types 31, 33, 35, 52 and 58) had intermediate clearance rates (RR = 0.62; 95% CI 0.47–0.94), while other high-risk types did not show evidence of slower clearance than low-risk types. Similar clearance rates were seen for infections with single and multiple HPV types and for women < 35 and ≥ 35 years of age.

Many women with transient HPV infections develop cytological abnormalities while they are actively shedding HPV particles (Figure 27). These occur because the viral life cycle is closely linked to the state of epithelial cell differentiation. When HPV is actively replicating in cells, it can produce characteristic cytopathic effects. These are nuclear enlargement, multinucleation, hyperchromasia and perinuclear cytoplasmic clearing or halos (e.g., “koilocytosis”) (see earlier in this chapter). When observed in cervical cytology specimens, these cellular changes are interpreted as either low-grade squamous intraepithelial lesion (LSIL) or atypical squamous cells (ASC), depending on the number of cells showing such changes and the severity of the changes.

The predominant age-curve of HPV prevalence shows a peak at young age and a steady decline in subsequent age groups (Jacobs et al., 2000; de Sanjose et al., 2003). Some population-based surveys of HPV prevalence (Herrero et al., 2000) and particularly the 12 areas included in the IARC Multi-Centre HPV Prevalence Survey have revealed, however, some variation in the age-pattern of HPV prevalence worldwide (Figure 28). Figure 28 is based on population-based surveys of several hundreds of cytologically normal women aged 15–74 years in areas where no organized screening programme and little or no opportunistic screening have existed. Furthermore, prevalence is reported only for high-risk HPV types.

In a few areas in Asia (Sukvirach et al., 2003; Shin et al., 2003) and Latin America (e.g., Concordia, Argentina, Matos et al., 2003; Bogotá, Colombia, Molano et al., 2003), the typical pattern again has a peak in women younger than 25 years of age, but a few additional types of age-curve are seen. In countries with very low HPV prevalence (e.g., Hanoi, Viet Nam, Anh et al., 2003), no peak in HPV prevalence is seen in young women. In the only sub-Saharan area included in the IARC survey (Ibadan, Nigeria, Thomas et al., 2004), there was very little decline in HPV prevalence across age decades between < 25 and > 64 years.

The interpretation of these different patterns is not entirely clear. The rapid rise of HPV prevalence in young women after start of sexual intercourse can be explained by the high prevalence of HPV in many populations and the high infectiousness of HPV.

The persistently high levels of HPV prevalence seen, for instance, in Nigeria point to the possibility that in countries where cervical cancer incidence is very high, HPV infection tends to persist and/or re-infection is substantial in every age-group.

In early studies, the end-point used was dysplasia and often the different stages of dysplasia were not distinguished. As these studies are uninterpretable with regard to the natural history of low-grade lesions, they are not summarized in this section. The authors’ terminology is retained in the summaries that follow.

Hakama & Rasanen-Virtanen (1976) used data emanating from the Finnish Cancer Registry and mass screening registries to estimate the probability of a woman aged 30–59 years being diagnosed with a cervical lesion after the first [negative] test in the national programme of five-yearly cytological screening. The probability of being diagnosed with dysplasia of low degree after a [negative] first result was estimated to be 0.006.

Campion et al. (1986) followed for 19–30 months 100 women under the age of 30 years who had cytological and colposcopic evidence of mild cervical atypia consistent with CIN 1 on three consecutive tests over a 16-week interval. They were reviewed every four months by cytology and colposcopy. If cytological evidence of severe dyskaryosis and colposcopic evidence of advanced CIN consistent with CIN 3 was obtained, a biopsy was performed and treatment given. Of the 100 women, 67 showed persistent mild
disease (of which three regressed to normal but recurred during the study period), 26 progressed to CIN 3 (one regressed to normal but disease recurred and progressed to CIN 3 during the study period) and seven regressed to normal, and did not recur.

Luthra et al. (1987) followed 428 women with cytologically diagnosed dysplasia for up to 84 months, the majority for less than 30 months. Those who progressed within six months were excluded. Of 268 women with mild dysplasia, two developed carcinoma in situ or worse (computed actuarial rate 28.9%). [The Working Group noted the substantial loss to follow-up in this study.]

Syrjänen et al. (1992) followed 528 women with HPV-associated genital lesions for up to 10 years. Of 270 with an HPV-associated lesion without CIN, 67% showed regression as defined by a negative cytological test, colposcopy and punch biopsy, 27% persistence and 6% progression defined by punch biopsy. Of 106 women with CIN 1, 56% regressed over a period of six years and 14% progressed, while 3% showed recurrence after treatment. When the lesions were reclassified using the Bethesda terminology, 376 were LSIL, of which 64% showed regression and 8% progression. [The Working Group noted the authors studied lesions associated with both low- and high-risk HPV types.]

Ostör (1993) reviewed reported studies that had considered the natural history of CIN 1. Only studies that employed cytology and/or cervical biopsy alone were reviewed (post-conization reports, and those where patients were treated by local destructive therapy, were excluded). Many were small studies (less than 60 subjects); 34 studies were included in the formal (summary) analysis. Only two of the studies of CIN 1 had over 1000 subjects. The total number of subjects with CIN 1 was 4504. After summarizing the data, Ostör concluded that 57% of CIN 1 regressed and 1% progressed to invasion. He suggested that, because of the small amount of cervical epithelium biopsied, it was probable that in many instances, more severe disease was not diagnosed. Therefore the proportion of cases of the various types that were estimated to have progressed was artificially increased. [The Working Group noted that the author did not conduct a formal meta-analysis, but summarized the reported percentages without taking note of the person-years of observation in the various studies. It is also likely that there was an effect of cervical biopsy in influencing the subsequent risk of progression. This "intervention effect" must have
had some impact upon the estimated progression rates.]

A cohort of 17,217 women, identified through one pathology laboratory in Toronto and whose records of cytological examinations spanned many years, was studied by Holowaty et al. (1999). The pathologists serving this laboratory had made a consistent attempt to identify the different degrees of dysplasia by cytology, and only referred women for further assessment if there was cytological evidence of progression. By linkage of the records from the laboratory with those of the Ontario Cancer Registry, women who were subsequently diagnosed with carcinoma in situ or invasive carcinoma of the cervix were identified. The maximum extent of regression occurred in those with cytological evidence of mild dysplasia. Progression to severe dysplasia or worse within 10 years occurred in only 10% of those with mild dysplasia. Most of these progressions occurred within five years. There was even less progression, even within ten years, when invasive cancer was used as the endpoint (Table 13).

In prospective follow-up studies which used HPV testing, it was found that cytological abnormalities were diagnosed in only a minority of women who were HPV-positive at study entry. For example, cytological abnormalities (LSIL or worse) occurred within three years in 25% of women who were HPV DNA-positive at enrolment in one study (Ho et al., 1998), in 28% (borderline dyskaryosis or worse) within three years in another (Woodman et al., 2001), in 22% (LSIL or worse) within 56 months in a study in young women and adolescents (Moscicki et al., 2001), in 17% (ASCUS or worse) over six years in another (Castle et al., 2002b) and 5.4% in a study of 2404 women followed by cytology and PCR-based HPV testing every 4–6 months over a period of eight years (Schlecht et al., 2003a). It was also noted in the latter study that the mean time for progression from ASCUS to LSIL or worse was less if the cytological specimen was positive for high-risk HPV (67 months), compared with 88 months in women with no HPV infection (Schlecht et al., 2003a). The reason why cytological abnormalities are not identified in a greater proportion of HPV-infected women is probably related to the transient nature of many productive infections and the fact that the cytological manifestations of a productive infection can be subtle. The risk of cytological abnormalities is consistently observed to be greatest during the six months immediately after HPV DNA detection and diminishes quickly thereafter. If such women undergo colposcopy, they are frequently found to have apparently low-grade acetowhite lesions that on biopsy often demonstrate the characteristic histopathological features of CIN 1. These CIN 1 lesions are heterogeneous with respect to a number of important biological parameters in addition to the associated HPV types. These include ploidy, clonality and LOH at specific chromosomal loci that may represent tumour-suppressor genes. The majority of CIN 1 are either diploid or polypoid, most are polyclonal, and LOH at specific chromosomal loci is most commonly absent.

Liaw et al. (1999) followed a cohort of 17,654 cytologically negative women using the records of Kaiser Permanente in California. Enrolment in the cohort commenced in 1989, and women were followed to the end of 1994. On average, each woman had 0.6 tests per year, and 20% had no repeat test. Of a total of 380 incident cases of cytological abnormality identified during follow-up, 154 were ASCUS and 179 LSIL. Cervical lavages had been collected and stored at enrolment, and on diagnosis another lavage specimen was collected. Similarly, specimens were collected from up to three matched controls without abnormality and all specimens were tested...
for HPV DNA using PCR. The data were analysed as a nested case–control study. Compared with women who were HPV DNA-negative on enrolment, those who were positive at enrolment had an OR of 3.8 (95% CI 2.6–5.5) for being diagnosed with LSIL. Infection with HPV16 was most likely to predict the development of SIL. The association was much stronger for HPV positivity detected at diagnosis.

Castle et al. (2002b) subsequently reported findings from a subcohort of 2020 women who were HPV-positive on enrolment but cytologically negative (from the same study as Liaw et al., 1999), who had at least one additional visit after enrolment and were followed for 57 months on average. The cumulative incidence of cytologically detected ASCUS or worse was 16.8% and of LSIL or worse 6.4%. The cumulative incidence of ASCUS or worse was 4.2% among cytologically negative women.

**High-grade lesions**

Once HPV infections clear, they have tended not to recur within the follow-up in the existing cohort studies. It is unclear whether there is an important chronic state of HPV latency related to subsequent cancer risk in immunocompetent women. Evidence for viral latency and re-emergence due to failed immune surveillance comes mainly from women with immunosuppression secondary to HIV infection or immunosuppressive medication used for organ transplantation (Sun et al., 1997).

A small proportion of women who become infected with HPV develop persistent HPV infections. The biological reasons why some women develop persistent infections are poorly understood, but probably include HPV type (with HPV16 persisting longer than other types) and differences in individual cell-mediated immune responses and other individual host factors, as well as environmental factors (e.g., possibly diet, smoking and co-infections with other sexually transmitted agents) (Sun et al., 1997; Kjaer et al., 2002b; Sedjo et al., 2002, Richardson et al., 2003). Persistent infections, for the most part, represent "abortive" viral infections. In abortive infections, no particles of HPV are produced within the infected cells and as a result, the cytopathic effects that are pathognomonic of productive HPV infections are greatly reduced. Follow-up studies have indicated that persistence of high-risk types of HPV is a prerequisite for the development of CIN 3 lesions and invasive cervical cancers (see earlier in this chapter) (Figure 27). For example, in a study of 1611 women with no cytological lesion on enrolment who were followed by HPV testing and cytology every 4–6 months for up to eight years, the incidence rate of SIL was 0.73 per 1000 women months (95% CI 0.5–0.9) among women free of HPV at the two initial visits and 8.68 (95% CI 2.3–15.1) among women with HPV type 16 or 18 infections persisting over both visits (Schlecht et al., 2001). Further, clearance of high-risk HPV in otherwise established cytological lesions is a marker associated with regression of CIN lesions (Nobbenhuis et al., 2001a; Zielinski et al., 2001a; Schiffman et al., 2002). Conversely, women with mild cytological abnormalities (e.g., ASCUS or LSIL) who lack identifiable high-risk types of HPV are at very low risk for developing CIN 2 or 3.

In the Liaw et al. (1999) cohort study of 17 654 cytologically negative women (see above), of a total of 380 incident cases of cytological abnormality identified, there were 47 with HSIL. Compared with women who were HPV DNA-negative on enrolment, those who were positive at enrolment had an OR of 12.7 (6.2–25.9) for HSIL. Infection with HPV16 was most likely to predict the development of SIL. The association was much stronger for HPV positivity detected at diagnosis.

Bory et al. (2002) reported results of follow-up of a sub-sample of 3091 women with normal cytological findings at first entry, among whom 659 (21%) had evidence of HPV infection. Of these 659 women, 241 (3% of the total studied) had a persistent HPV infection at 2–4 examinations, with a
final diagnosis of HSIL in 51 (0.7% of the total studied) within 4–36 months. The women who developed HSIL had a higher viral load than those with transient infections. All these women developed cytological abnormalities before or at the time of the colposcopy that led to the diagnosis. In contrast, of 2432 women who were negative for HPV and followed for a similar period as for the HPV-positive women, only two developed HSIL. Nobbenhuis et al. (1999) concentrated on CIN 3 as the end-point, using cytology and HPV testing by PCR to monitor for an average of 33 months 353 women who had been referred to gynaecologists with mild to moderate or severe dysplasia. Thirty-three women reached clinical progression, defined as CIN 3 covering three or more cervical quadrants, or a cytological result of suspected cervical cancer. All had persistent infection with a high-risk HPV type. The cumulative six-year incidence of clinical progression among these women was 40% (95% CI 21–59%).

Other data suggest that the intensity of an HPV infection (i.e., the viral load) is relevant to whether detectable disease develops. Thus, Cuzick et al. (1994) showed that in women with cytological abnormalities, a high viral load detected by a semi-quantitative PCR was strongly related to high-grade CIN. Subsequently, viral load determined with the Hybrid Capture technique has been reported to be directly related to the severity of the lesion (Clavel et al., 1999; George et al., 2000). Swan et al. (1999) reported a relationship between viral load and severity of CIN grade, which, however, appeared to be restricted to HPV16, and was absent for HPV18, 31 and 45. Ho et al. (1995) also suggested that women with SIL having a high viral load are more likely to have persistent SIL than those with a low level of HPV DNA. They conducted a study of 70 subjects with histopathologically confirmed cervical dysplasia, followed at three-month intervals for 15 months. Women with HPV type-specific infection and with a high viral load had the highest risk of persistent SIL compared with those having a low level of type-specific persistent infection or no type-specific infection (OR = 4.97; 95% CI 1.45–17.02). Josefsson et al. (2000) and Ylitalo et al. (2000b), in case–control studies nested within a cohort of screened women in Sweden, and using a fully quantitative PCR assay, demonstrated that cervical carcinoma in situ associated with HPV16 occurred mainly in HPV16-positive women who had consistently high long-term viral loads. These studies, based on tests of archived specimens, confirmed the long natural history of carcinoma in situ, as previously inferred from cytological observations. Thus, Ylitalo et al. (2000b) computed that the mean times from a first confirmed HPV infection before age 25 years to a diagnosis of carcinoma in situ in women with high and intermediate viral loads were 17 years and 19 years, respectively. Approximately 22% of women with evidence of a high viral load of HPV16 were eventually diagnosed with carcinoma in situ. Further, Josefsson et al. (2000) found that women in the highest 20% of the HPV16 DNA viral load distribution were at a 60-fold higher risk of being diagnosed with carcinoma in situ than women negative for HPV16. Subsequently, van Duin et al. (2002) found in a cohort study that in women with normal cytology, as well as those with abnormal cytology, an increased HPV16 viral load conferred an increased risk of developing a cervical lesion. This was confirmed by Schlecht et al. (2003b) in a cohort study of 473 women positive for HPV at the first two visits and followed by cytology and HPV testing for eight years. Compared with those having less than one viral copy per cell in specimens tested during the first two visits, the RR for incident SIL was 1.9 (95% CI 0.8–4.2) for those with 1–10 copies per cell and 4.5 (95% CI 1.9–10.7) for those with over 1000 copies per cell. The equivalent RR for HSIL for those with over 1000 copies per cell was 2.6 (95% CI 0.5–13.2).

The early studies frequently could not distinguish between different stages of CIN, or specifically between early lesions associated with transient HPV infections and those that represent persistent cancer precursors. Therefore, they are difficult to interpret in the light of our current understanding of the pathogenesis of cervical cancer. However, those that used histological diagnosis following biopsy and appropriate analytical methods contributed to our current understanding and these are summarized below.

In the early studies, the end-point used was carcinoma in situ. Subsequently, many authors used CIN 3 and more recently CIN 2/3 or HSIL (if cytology determined the end-point rather than histology). The authors’ terminology is retained in the summaries that follow.

Data derived from several screening programmes in the USA were used to derive estimates of the average duration of carcinoma in situ varying from five years (Dunn, 1960) to 8.1 years (Dunn & Martin, 1967) and 11 years (Kashgarian & Dunn, 1970). Dunn & Martin (1967) pointed out the discrepancy between some of these studies in the amount of disease that appeared to be occurring when using the curve of age-specific incidence, and suggested that some of the precursor lesions would terminate in regression. They reported that the maximum incidence of carcinoma in situ occurred in women aged 25–29, with a decrease to a low level after 35 years.

Kasper et al. (1970) analysed the prevalence and incidence of gynaecological cancer detected cytologically in
175 767 women screened in Alberta, Canada. They concluded that the generation of new cases of carcinoma in situ begins in the 20–24 age group at a very significant rate and that carcinoma in situ develops six times more frequently before age 45 years than in older women. They estimated that in the average case, detectable carcinoma in situ was present 8–10 years before the development of invasive cancer.

Theoretically, if all cases of carcinoma in situ progressed to invasive cancer, studies that could determine the cumulative incidence of carcinoma in situ (not prevalence) over a lifetime and compare this with the expected cumulative incidence of invasive cancer in the absence of screening should find almost or complete equivalence of these two cumulative rates. Allowance has to be made for those cases of carcinoma in situ that remain without progression and these are measured by the prevalence of disease. The difference between the cumulative incidence of carcinoma in situ by age less the prevalent (non-progressed) cases of carcinoma in situ and the cumulative incidence of invasive cancer gives an indication of non-progressive disease.

The British Columbia cohort study evaluated explanations other than non-progression for that difference (Boyes et al., 1982). The study utilized the records in the central cytology laboratory, the only one in the province. Records on all women born in 1914–18 (N = 52 452, Cohort 1) and 1929–33 (N = 66 701, Cohort 2) with follow-up to 1969 were extracted, and related to population data on invasive cancer, deaths, marriage and hysterectomy. The analysis concentrated mainly on corrected rates of prevalence and incidence of carcinoma in situ and of invasive cervical cancer. It was found that the gap persisted in spite of corrections for the false negative error and denominator error, and that the two cohorts 15 years of age apart had almost identical risks of carcinoma in situ at comparable ages. Thus the only remaining explanation for the gap was regression, which probably occurred in 40–60% of the detectable cases, especially at younger ages. Boyes et al. (1982) pointed out that when regression is part of the natural history, the ‘observed’ regression proportion is likely to be less than that which occurs. Therefore estimates of progression, made by the same process, are likely to be overestimates; however, they estimated that the proportion of carcinoma in situ that progressed ranged from 0.26 to 0.53.

The various models that had been applied earlier were reviewed by Prorok (1986). He concluded that the preclinical natural history of cervical cancer was complex, with a clear indication of regression of carcinoma in situ, as well as an age-dependence of transition probabilities and duration of disease status, that vary inversely by age. The mean duration of carcinoma in situ was model-dependent, with values in the 5–10 year range computed by authors who used early data, but values nearer 20–25 years were computed by authors who used longitudinal British Columbia data, which predated those subsequently analysed by Boyes et al. (1982).

In the Luthra et al. (1987) study (described in the section on low-grade disease above), of 138 women with cytologically diagnosed moderate dysplasia, 18 progressed (computed actuarial rate 29%) and of 22 with severe dysplasia two progressed (computed actuarial rate 36%). Of the total of 22 women who progressed during follow-up (including the two women originally diagnosed with mild dysplasia described above), 18 were diagnosed as carcinoma in situ and four early invasive cancer; there were no deaths.

[The Working Group noted the substantial loss to follow-up in this study.]

An extension of the Boyes et al. (1982) study, with follow-up to 1985, included over 75 000 women in Cohort 1, over 100 000 in Cohort 2 and a younger cohort (Cohort 3) of nearly 140 000 women born in 1944–48 (Miller et al., 1991b). Again there was little difference in incidence of carcinoma in situ at overlapping ages, confirming the absence of a cohort effect. The estimated proportions of cases of carcinoma in situ that regressed were 61% over ages 40–64 (Cohort 1), 70% over ages 25–54 (Cohort 2) and 77% over ages 15–39 (Cohort 3).

Syrjänen et al. (1992) defined regression by negative cytology, colposcopy and punch biopsy, and progression by punch biopsy. Of 68 women with CIN 2, 53% had regression over a period of six years, as did 14% of 42 with CIN 3. In contrast, 21% and 69%, respectively, showed progression, while 3% and 12% showed recurrence after treatment. When the lesions were reclassified using the Bethesda terminology, 376 were LSIL and 110 were HSIL. Of these HSIL lesions, 38% showed regression, 39% progression and 6% recurrence after treatment. [The Working Group noted that the authors studied lesions associated with both low- and high-risk HPV types.]

Koutsy et al. (1992) reported that after two years of follow-up of young college women, those initially HPV-positive had a cumulative risk of CIN 2 or 3 of 28%, compared with a cumulative risk of 3% for those initially HPV-negative. Risk was highest for those women infected with HPV16 or 18, the relative risk for those infected with type 16 or 18 being 11 (95% CI 4.6–26) for the development of any SIL compared with those without an HPV infection. Gaarenstroom et al. (1994) also showed that LSIL progressed to high-grade disease only if it contained high-risk HPV types.
In the Ostör (1993) review of the natural history of CIN described above for low-grade lesions, none of the studies of CIN 2 or 3 reviewed had over 1000 subjects. The total numbers of subjects in each of the categories were 2247 CIN 2 and 767 CIN 3. Ostör concluded that 43% of CIN 2 and 32% of CIN 3 regressed, and 5% and > 12% progressed to invasion.

Two more recent formal meta-analyses arrived at similar conclusions to those of Ostör (Mitchell et al., 1996; Melinkow et al., 1998).

Morrison et al. (1996) conducted a further update of the two cohorts originally studied by Boyes et al. (1982) to 1992, a follow-up period of over 40 years. The regression rate in the oldest cohort over this period was estimated to be 47% and in the younger cohort 72%.

In the Holowaty et al. (1999) cohort of 17,217 women, over 50% of those with cytological evidence of moderate dysplasia showed regression. Progression to severe dysplasia or worse within 10 years occurred in 32% of those with moderate dysplasia. Most of these progressions occurred within five years (Table 13).

In the Castle et al. (2002b) study of 20,810 women aged 16 years or more on enrolment who had satisfactory baseline cytology and samples suitable for HPV testing, they were screened annually by cytology for up to 122 months, during which period 171 women had CIN 3 or cancer diagnosed. The results of the HPV tests were not available for patient management, which was based on cytology according to standard practice. Of the women diagnosed with CIN 3 or worse, 123 had baseline cytology results of ASCUS or worse and/or a positive HPV test, 118 within the first 45 months of follow-up. During this period, the cumulative incidence of CIN 3 or worse was 4.5% among women with baseline cytology results of ASCUS or worse and/or a positive HPV test, compared with 0.16% among women negative on both tests.

Invasive cancer
Cervical cancer precursors can be defined in a variety of ways including virological measures, biological features and morphological terms. Persistence of infection with a high-risk type of HPV is now regarded as an absolute requirement for a lesion to be considered a precursor. Additional cellular events that occur in the majority of invasive cervical cancers and in many intraepithelial lesions include monoclonality, aneuploidy, genetic alterations which may result in the activation of oncogenes and inactivation of tumour-suppressor genes, and increases in telomerase activity. Monoclonality and aneuploidy are universally accepted features of malignancy at all tissue sites and are invariably found in invasive cervical cancers and in many CIN 3 lesions (Fu et al., 1988; Hering et al., 2000). Genetic alterations in cancers frequently appear to involve the inactivation of tumour-suppressor genes. Inactivation of tumour-suppressor genes is considered to be a two-hit process. One hit often involves a subtle mutation within one allele, whereas the other results in a gross deletion of the second allele (Weinberg, 1991) or methylation of that allele (Steenbergen et al., 2004). Molecular studies of invasive cervical cancers and CIN 2 and CIN 3 lesions have shown that a multitude of chromosomal loci are often lost in cervical carcinomas and that similar losses are occasionally observed in CIN 2 and CIN 3 lesions (Mitra et al., 1994a; Larson et al., 1997; Chung et al., 2000). High frequency of LOH, in more than 30% of cervical cancers studied, was found at 3p, 4q, 5p, 5q, 6p, 7q, 11p, 18p, and 18q (Mitra et al., 1994b; Mullokandow et al., 1996; Rader et al., 1996). Using comparative genomic hybridization (CGH) and LOH analysis, Steenbergen et al. (1996, 1998) demonstrated that HPV-mediated immortalization is associated with genetic gains at chromosomes 5p, 9q, 19 and 20 with LOH at 3p, 10p, 11p, 11q, 13q and 18q. In large part, these genetic alterations are thought to arise secondary to genetic instability induced in the cervical epithelium during the unregulated cell instability associated with expression of high-risk HPV-associated oncoproteins E6 and E7 in the basal layers, which is often associated with integration of the virus in the cellular genome.

In the study of Hakama & Rasanen-Virtanen (1976), the probability of being diagnosed with dysplasia
of high degree in the five-year interval after a negative cytological test was 0.010. For carcinoma in situ the probability was 0.012, while for microinvasive carcinoma it was 0.002. This compares with a probability of 0.010 for the development of frankly invasive carcinoma before the screening programme and 0.002 after the first [negative] test. Combining the probability of being diagnosed with in situ cancer with that for dysplasia of high degree, the authors estimated that 28–39% of such surgically treated preinvasive cases would otherwise have progressed to invasive cervical cancer, and that 21% of the frankly invasive cases are preceded by a preinvasive stage of shorter duration than the five-year interval between screenings, or have no preclinical stage.

Gustafsson & Adami (1989) used Swedish population-based incidence and mortality rates for cervical cancer to model the natural history of carcinoma of the cervix (including carcinoma in situ). They noted that the maximum incidence of carcinoma in situ occurred at 30 years. The proportion of cases of incident carcinoma in situ that progressed to invasive cancer was estimated to be 12.2%, with a mean duration of carcinoma in situ of 13.3 years, and the preclinical phase of invasive cancer without screening lasted on average for four years. They also estimated that 15–23% of prevalent carcinoma in situ progressed to invasive cancer. They did not estimate regression rates of carcinoma in situ, but concluded that the evidence was not compatible with two different types of cancer of the cervix with different natural histories. [The Working Group noted the differences between the estimates made by Gustafsson & Adami (1989) and those by other authors. One possible explanation is that a number of cases of CIN 3 were classified as carcinoma in situ in the data submitted to the cancer registry whose data these authors used, thus reducing the estimated proportion that progressed.]

The Holowaty et al. (1999) study of a cohort of 17 217 women found very little progression of moderate dysplasia, even within ten years, when invasive cancer was used as the end-point (Table 13).

**Discussion**

It is intuitively obvious that in order to gain knowledge of the natural history of the lesions discovered by screening, screening has to be performed and the detected abnormalities identified. However, as screening is performed in order to benefit the individual, it is rarely possible to avoid applying the currently best known therapy for the condition. For precursor lesions, this almost invariably involves surgical excision, resulting in complete, or almost complete, ablation of the cervical epithelium. Therefore, except in very unusual circumstances when observation rather than biopsy and treatment occurred (Barron & Richart 1968; Kinlen & Spriggs, 1978; Mclndoe et al., 1984), it is not possible to determine the natural history of the detected lesions by observing them. In that respect, the experience in New Zealand is salutary (Mclndoe et al., 1984). A total of 948 women with histologically confirmed carcinoma in situ diagnosed on punch biopsy but untreated were followed during 5–28 years. Of these, 817 had normal cytology on follow-up, 12 (1.5%) developed invasive carcinoma of the cervix. Of the remaining 131 women who continued to have abnormal cytology consistent with cervical neoplasia, 29% developed invasive carcinoma of the cervix or vaginal vault. Mclndoe et al. (1984) estimated that patients with continuing abnormal cytology after an initial diagnosis of carcinoma in situ are 24.8 times more likely to develop invasive carcinoma than women who have normal carcinoma on follow-up.

There are a number of inherent problems in almost all of the prospective studies of natural history. These include differing lengths of follow-up; differing definitions for what constitutes ‘disease’ (e.g., confirmed by colposcopy, cytology or histology), the fact that cervical biopsy, or perhaps even repeat cytological sampling, might affect natural history, and importantly that little effort has been made to examine the natural history of individual lesions on the cervix, as opposed to sampling the cervix as a whole. The latter may be particularly important when considering persistence of CIN 1 lesions. Unless location-specific information and in-depth HPV type-specific information is obtained during follow-up, it is impossible to determine whether a lesion that appears to be a persistent CIN 1 lesion is actually persistence of the same lesion or instead multiple sequential lesions associated with different HPV types. The inherent variability associated with a histo-pathological diagnosis of CIN is also critical when considering the outcomes of the natural history studies. Almost half of biopsy-confirmed CIN 1 lesions are reclassified as non-CIN when re-reviewed by expert pathologists, so that an apparent rate of approximately 50% regression of biopsy-confirmed CIN 1 could be observed through simply re-evaluating the original slides. The problems inherent in predicting natural history based on histological appearance alone are evident from the wide variations in observed outcomes when CIN lesions of various grades have been followed in prospective natural history studies.

A number of observations have indicated that high-risk HPV infection precedes the development of high-grade CIN lesions (see section above on Temporality). In studies involving
women who had CIN 3, Remmink et al. (1995) and Nobbenhuis et al. (1999) suggested that only women who were infected with high-risk HPV types were likely to have progressive disease, and only in women with persistent infection was their disease likely to progress. Further, Zielinski et al. (2001a) suggested that infection with HPV precedes the development of a cytological abnormality. However, cross-sectional surveys (conducted across ages at one point in time), such as those just described, do not provide much information on natural history. It is necessary to follow groups of women for some time, and conduct repeated tests, to obtain information about progression and regression of cervical lesions.

Among women with normal cytology who were positive for high-risk HPV, 8% developed CIN 3 within a four-year interval (Koutsy et al., 1992, Rozendaal et al., 1996, 2000). More recent prospective follow-up studies have confirmed that HPV infection precedes the development of low- and high-grade CIN lesions (Schlecht et al., 2001; Kjaer et al., 2002b). The distinction whether an intraepithelial lesion is associated with a transient HPV infection or whether it is associated with a persistent HPV infection now appears to be the single most important factor in determining the lesion's potential for progressing into an invasive cervical cancer. Interestingly, in studies by Nobbenhuis et al. (2001a) and Zielinski et al. (2001a), it was shown that clearance of high-risk HPV in otherwise established cytological lesions is a marker associated with the regression of CIN lesions (Londesborough et al., 1996).

In a retrospective study of women who developed cervical cancer, high-risk HPV was found to be present in cytologically normal smears taken more than 10 years before cervical cancer development (Zielinski et al., 2001a). That HPV infections precede the development of cervical cancer by at least 15 years in most cases is further supported by cross-sectional data (Bosch et al., 2002). In fact, in women over 30 years of age, the high-risk HPV prevalence declines to 2–4% (Jacobs et al., 2000; Clavel et al., 2001; Cuzick et al., 2003; Petry et al., 2003; Bulkmans et al., 2004).

Among HPV-positive women, very low viral loads correlate with cytological normality and benign natural history (Josefsson et al., 2000; van Duin et al., 2002; Schlecht et al., 2003b; Snijders et al., 2003; Giuliani et al., 2004; Hesselink et al., 2004). However, high viral loads are not always associated with high risk of progression to cancer precursor lesions. Because of the great variability in viral load and lesion severity, there is currently no consensus that viral load, as identified in routine clinical samples, presents a clinically useful parameter.

In most developed countries, the cumulative incidence to age 65 of invasive cervical cancer is approximately 1.5%, so that a cumulative incidence in excess of 4% of carcinoma in situ indicates the extent to which the majority of these lesions do not progress or regress. The cytology-based studies showed that at least 50% of women with detectable carcinoma in situ will not progress, while the proportion that regress is at least of a similar order to the proportion that progress. It is likely that this is also true for the lesions identified by HPV testing (Cuzick et al., 2003). However, the estimated progression rates in many studies may be overestimates, because what one observes (or can compute) in terms of regression is only a part of the regression that must be occurring, but is unobserved because of the intermittent nature of screening. Similar conclusions have also been reached concerning many of the dysplastic (SIL) abnormalities identified by cytology. The differences in extent of progression and/or regression reported in different studies may be due in part to differences in the application of similar terminology, or to differences in the way the classification was applied in different laboratories. Some authors have inferred that there can be direct progression to CIN 2 without passing through detectable CIN 1, but no data have been reported that directly addressed this issue. Although only one study has reported data using the Bethesda classification, it can be inferred that regression is an important part of the natural history of lesions classified both as HSIL (high-grade), as well as LSIL (low-grade) using this terminology.

The lessons from the HPV studies seem clear. Infection with HPV, especially the high-risk types, is predictive of subsequent development of squamous intraepithelial lesions, but such infection occurs far too frequently compared with the amount of cervical cancer expected in these populations. Thus not only are the large majority of infections in young women transient, but even the majority of those that appear persistent are of no concern. However, in older women, with documented evidence of persistent HPV infection with high-risk types, there is an appreciable probability that CIN 3 will develop.

There is, however, at least one other missing piece of information concerning the natural history of high-risk HPV infections. It is important to know what causes persistence of infection with oncogenic HPV viruses, and what causes precursors to progress and not regress. In this context, integration of the virus into the cellular genome, high viral load, diminished immunity of the host by HIV infection and additional genetic alterations such as LOH at 3p, 4, 5p, 10p and activation of hTert mRNA, which codes for the catalytic subunit
of telomerase, with increased levels of telomerase are likely to be important.

An ideal test would indicate that an oncogenic HPV has already enhanced genetic instability and rendered infected cells susceptible to transformation, thereby facilitating the development of cancer. In this respect, it should have the ability to detect those progressive cytological abnormalities that are caused by high-risk HPV infections and to discriminate them from transient low-grade lesions and those that only mimic morphological criteria of the onset of dysplasia or harbour HPV as an independent, but simultaneous event. Such a test should have greater true biological sensitivity and specificity than cytology and could possibly solve two problems inherent to conventional cytology. It could clarify how to consider the ASCUS and LSIL cytological abnormalities which, as already pointed out, represent mostly transient infections or in the case of ASCUS mainly diagnostic uncertainty. The other problem that contributes to low sensitivity of conventional cytology is overlooking and/or misinterpreting abnormal cells, a problem that also ideally should be overcome by a test that fulfils the criteria specified above and avoids sampling errors.

Figure 27 attempts to encapsulate our current understanding of the natural history of preclinical abnormalities of the cervix diagrammatically. It is important to recognize that although precise numbers cannot be given, as one moves from left to right across this figure, the probability of progression becomes higher, and as one moves in the opposite direction, the probability of regression increases.

### Considerations for screening programmes

The association of HPV and cervical cancer is unique in that a restricted number of HPV types have been identified as necessary for the development of cervical cancer worldwide. The implication is that in the absence of the viral infection (persistence is probably a requirement), cervical cancer is not expected to develop. Therefore, preventive strategies, either screening or vaccination, that target putative non-HPV-related cancers are no longer scientifically justified. The estimated risk linked to any of the high-risk HPV types for which sufficient evidence is available is, in statistical terms, equivalent to the risk of the most common ones, HPV16 and HPV18. Consequently, it is justified to use as screening tests cocktails of proven high-risk types. The role of the less common types in cervical cancer is not yet fully defined.

Some aspects of the epidemiological findings are relevant to screening recommendations:

(a) **Age to initiate screening.** Figure 18 shows a typical profile of age-specific HPV prevalence and of the incidence of cervical cancer in countries with established screening practices. The HPV prevalence is a function of the age at initiation of sexual activity of women and of the population-specific patterns of sexual exchange. Recommendations on the age to initiate HPV screening should aim to maximize detection of early cervical cancer cases while avoiding the bulk of transient HPV infections. It is thus important to carefully define country-specific HPV prevalence graphs as well as the age-specific incidence of cervical cancer and, if possible, of its closest precursors.

(b) **Age to terminate screening.** Most mortality from cervical cancer occurs in women aged 50 years and above. It is thus of interest to consider HPV screening along with cytology in women at the age of exit of their screening protocols, benefiting from the very strong negative predictive value of double negative results of these two screening tests.

(c) **Rescue of non-participants of screening programmes.** Deaths from cervical cancer in developed countries occur largely among women who escaped screening. There may be opportunities to offer some protection to these women linked to other health-related contacts, taking advantage of the longer protection associated with a normal cytological test result combined with HPV-negativity.

(d) **Women exposed to HIV are at higher risk of neoplastic progression.** Cervical cancer screening in women exposed to HIV may be enhanced by addition of HPV testing.

(e) **Other co-exposures of potential relevance.** Case–control studies have shown that at least three co-factors are likely to modulate the HPV-related carcinogenic process, namely high parity, long-term use of oral contraceptives and smoking. Although of little relevance in population-based screening programmes, individual preventive protocols at gynaecological clinics may take into account the presence of some of these exposures when making recommendations for HPV testing.
Chapter 2
Screening tests

Cervical cytology

Cytological testing involves collection of exfoliated cells from the cervix and microscopic examination of these cells after staining. The concept of utilizing exfoliative cytology to identify women with invasive cervical cancer was introduced by Papanicolaou and Babes in the 1920s (Papanicolaou, 1928; Papanicolaou & Traut, 1941). Subsequently, Papanicolaou refined the technique and demonstrated that conventional cytology could also be used to identify precancerous lesions of the cervix (Papanicolaou, 1954). The shift in emphasis from using cytology as a way to identify cases of invasive cervical cancer to using it to identify women with high-grade precursor lesions who are at risk for subsequently developing invasive cervical cancer was highly significant, as it meant that cervical cytology could be used to actually prevent the development of cervical cancer rather than simply identify cases at an early stage. In the 1960s, cervical cytology began to be widely used in many developed countries as a technique for cervical cancer prevention. Although the method was introduced over a half century ago, cytology-based screening programmes continue to be the mainstay of cervical cancer prevention.

Cytological terminology

Papanicolaou classes

The terminology developed by Papanicolaou separated cervical cytological findings into five categories or classes (Table 14) (Papanicolaou, 1954). At the time the classification was developed, there was only limited understanding of the relationship between cervical cancer precursor lesions and invasive cancers. Moreover, invasive cervical cancer was common and cervical cytology was initially viewed as a way of detecting early-stage, easily treated cancers. Therefore, the Papanicolaou classification system focused on how closely the exfoliated cells resembled those from an invasive cancer. Although the Papanicolaou classification was modified many times over the years, the problems inherent in this classification remain. For example, although it is clear how Class I and Class V translate into known histological entities, Classes II, III or IV correlate less clearly with standard histopathological lesions. For example, should a carcinoma in situ be classified as Class IV and all grades of dysplasia as Class III, or does mild dysplasia correspond to Class II? There was also no consensus as to what other non-neoplastic conditions were combined in Class II. Such ambiguity in the Papanicolaou classification resulted in its non-uniform use by different cytologists. Modifications of the Papanicolaou classifications are still used in some countries. In the Netherlands, a modified Papanicolaou system (CISOE-A) is used for classification. This redefined and subdivided the Papanicolaou classes in order to make the terminology correlate with histopathological terminology (Hanselaar, 2002).

World Health Organization terminology

In the 1950s, some cytologists began to promote a more scientifically accurate terminology that would allow cytological diagnoses to translate directly into histological diagnoses. This terminology (Table 15) was later adopted by the World Health Organization (WHO) (Riotton et al., 1973). The WHO terminology allows more precise correlation between cytological and histopathological findings, but is

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Absence of atypical or abnormal cells</td>
</tr>
<tr>
<td>II</td>
<td>Atypical cytology, but no evidence for malignancy</td>
</tr>
<tr>
<td>III</td>
<td>Cytology suggestive of, but not conclusive for, malignancy</td>
</tr>
<tr>
<td>IV</td>
<td>Cytology strongly suggestive of malignancy</td>
</tr>
<tr>
<td>V</td>
<td>Cytology conclusive for malignancy</td>
</tr>
</tbody>
</table>

From Papanicolaou, 1954
IARC Handbooks of Cancer Prevention Volume 10: Cervix cancer screening

difficult to use since it includes a number of different entities. These are mild dysplasia, moderate dysplasia, severe dysplasia, epidermoid carcinoma in situ, epidermoid carcinoma in situ with minimal stromal invasion, invasive epidermoid microcarcinoma and invasive epidermoid carcinoma. Studies have shown high rates of intra-observer and inter-observer variation with cervical cytology in general (Yobs et al., 1987; Klinkhamer et al., 1988; Selvaggi, 1999; Stoler & Schiffman, 2001). Classification systems that utilize more diagnostic categories have inherently higher rates of variability than do classification systems with fewer diagnostic categories (Yobs et al., 1987; Selvaggi, 1999; Stoler & Schiffman, 2001; Kundel & Polansky, 2003). Other limitations of the WHO terminology are that it does not adequately deal with non-neoplastic conditions nor with specimen adequacy. Despite its limitations, many cytologists around the world continue to utilize the WHO terminology.

Cervical intraepithelial neoplasia (CIN) terminology

As a result of advances in understanding of the pathogenesis of cervical cancer, the cervical intraepithelial neoplasia (CIN) terminology was introduced in the late 1960s (Richart, 1968, 1973). The CIN concept emphasized that dysplasia and carcinoma in situ represent different stages of the same biological process, rather than separate entities. It had a major impact on how precancerous lesions were treated, since all types of cervical cancer precursor were considered to form a biological and clinical continuum. In the CIN terminology, mild dysplasia is classified as CIN 1, moderate dysplasia as CIN 2 and severe dysplasia and carcinoma in situ are grouped together and classified as CIN 3 (Table 15). The CIN terminology is still widely used in many countries for reporting both histological and cytological diagnoses.

The Bethesda System terminology

By the late 1980s, advances in our understanding of the role of human papillomavirus (HPV) in the pathogenesis of cervical cancer needed to be incorporated into cytological terminology. Moreover, it was recognized that clinicians were often confused by the non-standard terminologies used to report cytological results and that this had a potential adverse impact on clinical care. Therefore, in 1988, the US National Institutes of Health held a conference in Bethesda, Maryland, to develop a new terminology that would ensure better standardization and accommodate current concepts of the pathogenesis of cervical disease, so that cytological findings could be transmitted to clinicians as accurately and concisely as possible. The terminology that resulted is known as the Bethesda System. In 1991 the Bethesda System was slightly modified on the basis of experience obtained during the first three years of its use and it was further modified in 2001 to take into account the results of new research and over a decade of experience with the terminology (Luff, 1992; Solomon et al., 2002).

The Bethesda system is viewed with caution in the United Kingdom, which retains its own British Society for Clinical Cytology (BSCC) 'dyskaryosis' terminology (British Society for Clinical Cytology, 1997). This can largely be mapped to the Bethesda system for comparison of data in a research setting, except for the borderline category, which may include koilocytes. Due largely to the robust nature of the 'severe dyskaryosis' category, fear of increasing the overtreatment inherent in cervical screening and the difficulty of achieving inter- and intra-observer agreement on 'low-grade' reports, the United Kingdom continues to use this terminology.

There are three distinct parts to each Bethesda System report: a state-

| Table 15. Comparison of different terminologies used for cytologic reporting |
|------------------|-----------------|-------------------|-------------------|
| Papanicolaou class system | World Health Organization | CIN | Bethesda System |
| Class I | Within normal limits | CIN1 | Low-grade SIL |
| Class II | Benign cellular changes ASC | CIN2 | High-grade SIL |
| Class III | Moderate dysplasia | CIN3 | Invasive carcinoma |
| Class IV | Carcinoma in situ | invasive carcinoma |
| Class V | Microinvasive carcinoma | Invasive carcinoma |

Abbreviations: CIN, Cervical intraepithelial neoplasia; ASC, Atypical squamous cells; SIL, Squamous intraepithelial lesions

ment of the specimen adequacy, a general categorization and a descriptive diagnosis (Table 16). These categories assist clinicians by providing answers to three basic questions: (1) Do I need to repeat the cervical cytology? (2) Was the cervical cytology normal? (3) If the specimen was not completely normal, what specifically was wrong?

Because cervical cytology is considered a screening, rather than diagnostic, test, the 2001 Bethesda System reports cytological findings as an 'interpretation' or 'result' rather than as a 'diagnosis'. This stresses the fact that cytological findings usually need to be interpreted in the light of clinical findings, and that the test is designed to reflect the underlying disease state but does not always do so.

In this Handbook, the Bethesda System (SIL) terminology is used for cytological interpretation of screening tests unless otherwise reported.

**Specimen adequacy**
The 2001 Bethesda System requires that every cervical cytology specimen be assessed with respect to its adequacy (Solomon et al., 2002). Specimens are classified into one of two categories: 'satisfactory for evaluation' or 'unsatisfactory for evaluation'. This represents a departure from the 1991 Bethesda System, which also included a third category for specimen adequacy that was called 'satisfactory for evaluation but limited by….' or SBLB. This 'satisfactory but limited by…' category was most frequently used when a specimen lacked either endocervical cells or squamous metaplasia from the transformation zone but was in all other aspects 'satisfactory'. With the 2001 Bethesda System, cytology specimens previously classified as ‘SBLB’ are classified as ‘satisfactory for evaluation’ and a quality indicator comment is made indicating what limiting features are present.

A 'satisfactory for evaluation' specimen must be appropriately labelled. To ensure proper identification, the woman’s name or identifying number should be written on, or affixed to, the slide before it is sent to the cytology laboratory. Cytology laboratories should not accept unlabelled slides and should return them to the submitting clinician. It is also critical that the smear-taker provide pertinent clinical information to the laboratory that will evaluate the specimen, including the woman’s age, date of last menstrual period, previous history of abnormal cervical cytology specimens or treatment for cervical disease, and the source of the specimen (e.g., vaginal or cervix). The minimal cellular requirements for a specimen to be considered ‘satisfactory for evaluation’ in the 2001 Bethesda System vary depending on whether the specimen is a conventional cytology specimen or a liquid-based cytology specimen. For classification of conventional cytology specimens as ‘satisfactory for evaluation’, an estimated 8000 to 12 000 well visualized squamous cells need to be present. For liquid-based cytology specimens, an estimated 5000 cells need to be present (Figure 29). Although the selection of these cut-offs is fairly arbitrary, the limit of 5000 cells for a liquid-based cytology specimen to be classified as ‘satisfactory for evaluation’ is based on a cell-counting study in which referent samples were diluted to produce preparations with defined numbers of squamous cells (Studeman et al., 2003). A clear demarcation in sensitivity was observed using the SurePath™ procedure (see below) between specimens with less than 5000 squamous cells and those with 5000 cells or more: the sensitivity for a reference diagnosis case of low-grade squamous intraepithelial lesion (LSIL) increased from 73% for specimens with less than 5000 squamous cells to 98% for preparations with over 5000 cells.

There is much controversy over the importance of identifying a transformation zone component (e.g., squamous metaplastic cells) or endocervical cells in a cervical cytology preparation. Because the majority of high-grade precursor lesions arise within the transformation zone, it was widely believed until recently that specimens lacking a transformation zone component (TZC) or endocervical cells (EC) should be considered somewhat less than ‘satisfactory for evaluation’. This view is supported by several studies that have shown the prevalence of SIL to be higher among cytology specimens that contain TZC/EC than among those that do not (Vooijs et al., 1985; Mitchell & Medley, 1992; Szarewski et al., 1993; Mintzer et al., 1999). However, other studies have failed to confirm this association and, perhaps more importantly, several retrospective longitudinal cohort studies have found that women lacking TZC/EC are no more likely on follow-up to be diagnosed with squamous lesions than are women whose specimens contain TZC/EC (Mitchell & Medley, 1991; Mitchell, 2001). One retrospective case–control study of true

A clear demarcation in sensitivity was observed using the SurePath™ procedure (see below) between specimens with less than 5000 squamous cells and those with 5000 cells or more: the sensitivity for a reference diagnosis case of low-grade squamous intraepithelial lesion (LSIL) increased from 73% for specimens with less than 5000 squamous cells to 98% for preparations with over 5000 cells.

There is much controversy over the importance of identifying a transformation zone component (e.g., squamous metaplastic cells) or endocervical cells in a cervical cytology preparation. Because the majority of high-grade precursor lesions arise within the transformation zone, it was widely believed until recently that specimens lacking a transformation zone component (TZC) or endocervical cells (EC) should be considered somewhat less than ‘satisfactory for evaluation’. This view is supported by several studies that have shown the prevalence of SIL to be higher among cytology specimens that contain TZC/EC than among those that do not (Vooijs et al., 1985; Mitchell & Medley, 1992; Szarewski et al., 1993; Mintzer et al., 1999). However, other studies have failed to confirm this association and, perhaps more importantly, several retrospective longitudinal cohort studies have found that women lacking TZC/EC are no more likely on follow-up to be diagnosed with squamous lesions than are women whose specimens contain TZC/EC (Mitchell & Medley, 1991; Mitchell, 2001). One retrospective case–control study of true

A clear demarcation in sensitivity was observed using the SurePath™ procedure (see below) between specimens with less than 5000 squamous cells and those with 5000 cells or more: the sensitivity for a reference diagnosis case of low-grade squamous intraepithelial lesion (LSIL) increased from 73% for specimens with less than 5000 squamous cells to 98% for preparations with over 5000 cells.

There is much controversy over the importance of identifying a transformation zone component (e.g., squamous metaplastic cells) or endocervical cells in a cervical cytology preparation. Because the majority of high-grade precursor lesions arise within the transformation zone, it was widely believed until recently that specimens lacking a transformation zone component (TZC) or endocervical cells (EC) should be considered somewhat less than ‘satisfactory for evaluation’. This view is supported by several studies that have shown the prevalence of SIL to be higher among cytology specimens that contain TZC/EC than among those that do not (Vooijs et al., 1985; Mitchell & Medley, 1992; Szarewski et al., 1993; Mintzer et al., 1999). However, other studies have failed to confirm this association and, perhaps more importantly, several retrospective longitudinal cohort studies have found that women lacking TZC/EC are no more likely on follow-up to be diagnosed with squamous lesions than are women whose specimens contain TZC/EC (Mitchell & Medley, 1991; Mitchell, 2001). One retrospective case–control study of true
positive and false negative cervical cytology specimens from women with CIN 3 found no difference in true positive rates between cases with or without TZC/EC (O’Sullivan et al., 1998). A prospective study of women with normal cytology at entry found that although specimens containing EC at the subsequent test were at significantly higher risk of both low- and high-grade squamous intraepithelial lesions than those without EC, the presence or absence of EC at entry had no significant effect (Mitchell, 2001). In another compelling study on the lack of importance of EC, all negative cervical cytology specimens obtained in the Netherlands between 1990 and 1991 were matched with results of subsequent cytological and histological examinations (Bos et al., 2001). There was no significant difference in the number of women subsequently diagnosed with CIN between women whose initial cytology specimens contained EC and those that did not. Moreover, the proportions of women diagnosed with cervical cancer were the same in both groups. It is also important to recognize that EC are less frequently found in cervical cytology specimens from women using oral contraceptives, who are pregnant or who are postmenopausal (Davey et al., 2002). It has therefore been argued that specimens lacking EC or a TZC should not be considered unsatisfactory and may not need to be repeated (Davey et al., 2002; Birdsong, 2001; Bos et al., 2001). The 2001 Bethesda System recommends that reports should state whether or not EC or a TZC are present. Specimens lacking endocervical cells or squamous metaplastic cells should be classified as ‘satisfactory for evaluation’ and the quality indicator comment should indicate that these components are not present. The numeric criterion for stating that such a component is present is 10 well preserved endocervical or squamous metaplastic cells. Specimens in which inflammation, blood or poor preservation cause 50–75% of the epithelial cells to be obscured should be classified as ‘satisfactory for evaluation’, but a quality indicator comment made indicating that there are partially obscuring factors.

A specimen is classified as ‘unsatisfactory for evaluation’ when either the minimal number of epithelial cells required for interpretation is not present or blood, inflammation or poor preservation obscures more than 75% of the epithelial cell component (Figure 30). Cases which the laboratory cannot process, such as those received unlabelled, are also classified as ‘unsatisfactory for evaluation’ and no interpretation is rendered.

General categorization
The ‘general categorization’ is included as an optional component of the Bethesda System to allow clinicians to readily determine whether any degree of abnormality is present. With the 2001 Bethesda System, all cytology specimens are classified into one of three general categories. These include ‘negative for intraepithelial lesion or malignancy’, ‘epithelial cell abnormalities’ and ‘other’. These categories are mutually exclusive and specimens should be categorized according to the most significant findings.

‘Negative for intraepithelial lesion or malignancy’ includes all specimens in which no intraepithelial lesion or malignancy is identified. This includes cases with common infections such as Trichomonas vaginalis, fungal organisms such as Candida species, Actinomyces or herpes simplex virus, a shift in bacterial flora consistent with bacterial vaginosis, reparative/reactive changes, changes associated with intrauterine devices, radiation reactions or atrophic changes.

The category ‘epithelial cell abnormalities’ includes both squamous and glandular cell abnormalities. This category is used whenever there are epithelial cell abnormalities, except for benign reactive or reparative changes.

The 2001 Bethesda System introduced a new general categorization, ‘other’. This category is used whenever there are no morphological abnormalities in the cells per se, but there are findings indicative that the woman is at some increased risk. An example is when benign-appearing endometrial cells are identified in a woman 40 years of age or older.

Squamous cell abnormalities
Atypical squamous cells (ASC): Epithelial cell abnormalities are subdivided into four categories (Table 16). ‘Atypical squamous cells’ (ASC) is used when cytological findings are considered suggestive but not diagnostic of a squamous intraepithelial lesion (SIL) (Figure 31). The term ASC was retained in the 2001 Bethesda System because of the wide recognition that these cells imply a significant risk for an underlying high-grade cervical intraepithelial lesion (SIL). In various studies, the prevalence of CIN 2 or 3 in women with ASC has varied between 10% and 20% (Wright et al., 2002a). The ASC category roughly correlates with the ‘borderline dyskaryosis’ category used in the
United Kingdom. However, neither the WHO terminology nor the CIN terminology incorporates a category similar to ASC. The 2001 Bethesda System also clearly separates ASC from reactive/reparative changes and an interpretation of ASC should not be made whenever a cytopathologist identifies minor cytological abnormalities. The term ASC should be used only when the cytological findings are suggestive, but not diagnostic, of SIL. Currently, approximately 4–5% of all cervical cytology specimens are classified as ASC in the USA (Jones & Davey, 2000).

The ‘atypical squamous cell’ category is formally subdivided into two subcategories: ‘atypical squamous cells – of undetermined significance’ (ASCUS or ASC-US) and ‘atypical squamous cells – cannot exclude a high-grade SIL’ (ASC-H). This subdivision was felt to be important because women with ASC-H (Figure 32) are at considerably higher risk for having CIN 2 or 3 and of being high-risk HPV DNA-positive than are women with ASCUS (Genest et al., 1998; Sherman et al., 1999, 2001; Selvaggi, 2003).

Information from the US National Cancer Institute ASCUS–LSIL Triage Study (ALTS) clinical trial indicates that the risk that a woman with ASC-H has CIN 2 or 3 is over twice that of a woman with ASCUS (Table 17) (Sherman et al., 2001). Moreover, the prevalence of high-risk HPV DNA-
positivity among women with ASC-H is almost as high as that of women with a high-grade squamous intraepithelial lesion (HSIL) cytological result. Therefore the recommended management of women with ASCUS and ASC-H differs (Wright et al., 2002a).

Low-grade squamous intraepithelial lesion: The LSIL category in the Bethesda System includes both HPV effects and CIN 1 (i.e. mild dysplasia). Most cytologists consider the cytopathic effects of HPV, including multinucleation, perinuclear halos and nuclear atypia with irregular nuclear outlines and hyperchromasia, to overlap the cytological features of CIN 1. These features are referred to as ‘koilocytosis’, a term derived from the Greek koilos, meaning hollow.

The classical studies of Reagan and others identified the key cytological features of CIN 1 (Table 18) (Reagan & Hamomic, 1956). The cells are of the superficial or intermediate-cell type. They are classically described as having nuclei 4–6 times the size of a normal intermediate-cell nucleus (Figure 33). However, nuclei may vary in size and, in many cases of LSIL that are characterized by marked HPV cytopathic effects, are only twice the size of a normal intermediate-cell nucleus. The nuclei are usually hyperchromatic, and multinucleation is common. The chromatin is finely granular and uniformly distributed. The cells typically occur as individual cells or as sheets of cells with well defined cell borders.

High-grade squamous intraepithelial lesion: Because the Bethesda System combines moderate and severe dysplasia together with carcinoma in situ in the HSIL category, there is wide variation in the cytological appearance of HSIL. When applying the 2001 Bethesda System, many cytopathologists utilize the option of subdividing HSIL into CIN 2 and CIN 3 lesions. As the severity of the lesion increases, the degree of differentiation and the amount of cytoplasm decreases, the nuclear:cytoplasmic ratio increases, and the degree of

### Table 17. Prevalence of high-risk HPV DNA and CIN 2 and CIN 3 in women with ASCUS and ASC-H in the ASCUS-LSIL triage study (ALTS)*

<table>
<thead>
<tr>
<th>Cytology result</th>
<th>No.</th>
<th>% high-risk HPV DNA-positive</th>
<th>% biopsy-confirmed CIN 2+</th>
<th>% biopsy-confirmed CIN 3+</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCUS</td>
<td>764</td>
<td>63.2%</td>
<td>11.6%</td>
<td>4.7%</td>
</tr>
<tr>
<td>ASC-H</td>
<td>116</td>
<td>85.6%</td>
<td>40.5%</td>
<td>24.1%</td>
</tr>
<tr>
<td>HSIL</td>
<td>213</td>
<td>98.7%</td>
<td>59.2%</td>
<td>37.6%</td>
</tr>
</tbody>
</table>

* Study provides the results for liquid-based cytology specimens that were tested for high-risk types of HPV using Hybrid Capture 2
From Sherman et al. (2001)

### Table 18. Cytological features of squamous intraepithelial lesions

<table>
<thead>
<tr>
<th>Bethesda system</th>
<th>LSIL</th>
<th>HSIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIN terminology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHO terminology</td>
<td>Mild dysplasia</td>
<td>Moderate dysplasia</td>
</tr>
<tr>
<td>Cell type</td>
<td>Superficial or intermediate</td>
<td>Parabasal</td>
</tr>
<tr>
<td>Cell arrangement</td>
<td>Singly or sheets</td>
<td>Singly or sheets</td>
</tr>
<tr>
<td>Number abnormal</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Koilocytosis</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Nuclear size</td>
<td>+++</td>
<td>++</td>
</tr>
<tr>
<td>Hyperchromasia</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Nuclear:cytoplasmic ratio</td>
<td>+</td>
<td>++</td>
</tr>
</tbody>
</table>

From Reagan & Hamomic, 1956

---

**Figure 32** Inflammatory smear with parabasal squamous cells with enlarged nuclei: ASC-H (ellipse) (obj. 20x)
nuclear atypia also increases. HSIL of the moderate dysplasia type typically contains cells similar to those seen in LSIL, as well as atypical immature cells of the parabasal type (Figure 34). The nuclei of these cells are more hyperchromatic and irregular than typically seen in LSIL. In severe dysplasia, the overall size of the cells is reduced compared to mild and moderate dysplasia, but because the cells demonstrate minimal differentiation, the nuclear:cytoplasmic ratio is greatly increased. In severe dysplasia, there are usually considerably greater numbers of neoplastic cells that are typically found individually. Carcinoma \textit{in situ} can be of the small-cell type, of the large-cell non-keratinizing type or of the large-cell keratinizing (pleomorphic) type. Although separation of carcinoma \textit{in situ} into these three different cytological types has little clinical significance, all three have quite different cytological appearances. Small-cell lesions consist of small basal-type cells similar to those seen in severe dysplasia but which demonstrate even less cytoplasm and higher nuclear:cytoplasmic ratios (Figure 35). Because of their small size, these cells can easily be overlooked during routine screening and such cases account for a disproportionate percentage of false negative cytological results. The cells of large-cell non-keratinizing lesions typically form syncytial-like cell sheets in which individual cell membranes are difficult to identify. These cells have enlarged, hyperchromatic nuclei and minimal amounts of cytoplasm. The keratinizing large-cell type of carcinoma \textit{in situ} is composed of pleomorphic, highly atypical cells, many of which have thick keratinized cytoplasm. These cells are often spindled or tadpole-shaped and have extremely dense nuclear chromatin (Figure 36).

\textit{Invasive squamous-cell carcinoma:} Cytologically, squamous-cell carcinomas of the cervix are subdivided into keratinizing and non-keratinizing types. Non-keratinizing carcinomas (Figure 37) typically have large numbers of malignant cells that form loose cell sheets and syncytial arrangements. The cells have enlarged nuclei with coarsely clumped chromatin, prominent macronucleoli and focal chromatin clearing. A key cytological feature is the presence of a ‘dirty’ background containing blood and necrotic material. This is often referred to as a tumour diathesis. This characteristic background is usually less prominent in liquid-based cytology specimens.

Cervical smears from women with \textit{keratinizing carcinomas} contain malignant cells of a variety of shapes and sizes (Figure 38). Some of the cells are pleomorphic or tadpole-shaped with nuclei that are irregular in shape and

---

**Figure 33** LSIL: typical eosinophilic and basophilic koilocytes associated with some parakeratosis and binucleated cells (obj. 20x)

**Figure 34** Parabasal cells arranged in a pile with nuclear enlargement, irregular nuclear outlines and coarse chromatin. HSIL (moderate dysplasia) (obj. 20x)

**Figure 35** HSIL (severe dysplasia): inflammatory smear containing many parabasal cells with enlarged nuclei with irregular chromatin (black arrow). Some cells with a mildly eosinophilic cytoplasm (ellipse) (obj. 20x)

**Figure 36** HSIL (severe dysplasia): basal cells with enlarged nuclei and irregular or very dense and opaque chromatin (arrow), accompanied by an atypical mature cell (obj. 40x)
quite hyperchromatic. Unlike non-keratinizing squamous-cell carcinoma, keratinizing squamous-cell carcinomas usually do not have a ‘dirty’ background or evidence of tumour diathesis.

**Glandular cell abnormalities**

Glandular cell abnormalities are categorized into four categories: *atypical glandular cells* (AGC), *atypical glandular cells – favour neoplasia*, adenocarcinoma in situ and adenocarcinoma. Whenever possible, atypical glandular cells are categorized as to whether they are endocervical or endometrial in origin.

*Atypical glandular cells (AGC):* Glandular cytological abnormalities are considerably less common than squamous abnormalities and most cytologists tend to be less comfortable recognizing and diagnosing them. In addition, the criteria used to differentiate reactive endocervical changes from neoplasia are less well established than those used for squamous lesions. Cytologists even have difficulty in differentiating atypical endocervical cells from cases of CIN 2 or CIN 3 that have extended into endocervical crypts. This accounts for the high prevalence of squamous abnormalities (approximately 30%) detected in women referred to colposcopy for AGC (Eddy *et al.*, 1997; Veljovich *et al.*, 1998; Ronnett *et al.*, 1999; Jones & Davey, 2000; Krane *et al.*, 2004).

The cytological features of atypical glandular cells vary depending on the degree of the underlying histopathological abnormality and whether or not the cells are endocervical or endometrial in origin. Atypical glandular cells of endocervical origin frequently form dense two- or three-dimensional aggregates that have minor degrees of nuclear overlapping. In some cases, the chromatin is somewhat granular and nuclear feathering can be seen at the periphery of the cellular aggregates (Figure 39). In cases interpreted as *atypical glandular cells – favour neoplasia*, there is more marked cytological abnormality and typically a greater number of abnormal cells.

*Adenocarcinoma in situ:* In cases of adenocarcinoma in situ, there are usually a larger number of atypical glandular cells that form crowded cellular clusters (Figure 40). The sheets are usually three-dimensional. The cells within these sheets occasionally form rosettes and have extensive feathering of the cells at the periphery. Individual endocervical cells are highly atypical with enlarged round, oval or elongated nuclei that vary in size from cell to cell. In most cases, the chromatin is coarsely clumped and multiple mitoses are seen.
Adenocarcinoma: Invasive adenocarcinomas should be subclassified into the endocervical or endometrial type whenever possible. The cytological diagnosis of invasive adenocarcinoma is relatively straightforward. Adenocarcinoma cells from either an endocervical or an endometrial primary type have enlarged nuclei, high nuclear:cytoplasmic ratios, coarsely clumped chromatin and prominent nucleoli (Figure 41). They can occur singly or in clusters.

Other terminologies
Although the 2001 Bethesda System classification is applied in many countries, other classification systems are also widely used. As mentioned previously, many countries prefer to subclassify high-grade intraepithelial lesions into at least two categories. This is the approach used in the United Kingdom, where squamous intraepithelial abnormalities are divided into five categories (borderline changes; mild, moderate, severe dyskaryosis and severe dyskaryosis or possibly invasive cancer) (British Society for Clinical Cytology, 1997).

Conventional cervical cytology
The importance of proper specimen collection cannot be overemphasized. Although no formal studies have demonstrated that educating clinicians on the optimal technique of obtaining cervical cytology samples improves specimen quality, there is considerable anecdotal evidence that this is important (Krieger et al., 1998). One half to two thirds of false negative cervical cytology results are attributable to either poor patient conditions at the time the cervical specimen is collected or the manner in which it is collected (Morell et al., 1982; Gay et al., 1985; Vooijis et al., 1985; Agency for Health Care Policy and Research, 1999).

Therefore it is important that clinicians and nurses obtaining specimens be adequately trained in specimen collection and that they avoid situations that may reduce the performance of the test (McGoogan et al., 1998). This is especially important in low-resource settings, where women may undergo screening only once or twice in their lifetime.

Preparing the woman
Whenever possible, appointments for a cervical cytology examination should be scheduled approximately two weeks after the first day of the last menstrual period. Patients should be instructed to avoid sexual intercourse and douching for 24 to 48 hours before having the cytology specimen collected. In addition, women should not use any intravaginal products or medicine for several days before the smear is taken. Women using an intravaginal estrogen product should discontinue its use several days before the examination.

Circumstances that may interfere with the interpretation of a cervical cytology test include active menstruation, significant cervical or vulvovaginal infections and a timing less than eight weeks post-partum. When a woman is actively menstruating, blood and cellular debris from the endometrium tend to obscure the cells on the smear, particularly during the first few days. Similarly, a cytology specimen should not be obtained when an abnormal vaginal or cervical discharge is observed. Women with a discharge should be evaluated for cervicitis and vaginitis using appropriate tests and be treated before the cytology specimen is taken, otherwise the specimen may be compromised by the inflammatory exudates or mildly reactive cells may be misinterpreted as a significant cytological abnormality.

There is controversy as to the ideal timing of post-partum smears. Smears obtained less than eight weeks post-partum are often difficult to interpret because of marked inflammation and reparative changes, so a high rate of mild cytological abnormalities may be diagnosed. Another factor that can adversely affect the interpretation of cervical cytology specimens is severe atrophy.

Although one should strive to collect specimens under ideal conditions, failure to comply with suggested screening intervals presents a greater risk to women. For previously non-compliant women, particularly those at risk for cervical neoplasia, a smear obtained under less than ideal conditions is preferable to no smear at all.

Equipment
To collect a conventional cervical cytological specimen, the equipment required is a speculum, a light source, a collection device, a glass slide and fixative. Since most cervical cancer precursors and invasive cancers occur in the transformation zone, the use of specially designed devices that sample this area is recommended. The most common is a wooden or plastic spatula that conforms to the curvature of the portio. It is critical that the endocervical canal be sampled in order to obtain
reasonable sensitivity (Martin-Hirsch et al., 1999) and many spatula-type devices have extended tips designed to collect cells from this area. Either a moistened cotton swab or a brush-type endocervical sampler device (e.g., cytobrush) can be used to collect a second sample directly from the endocervical canal after the portio has been sampled (Koonings et al., 1992; Kohlberger et al., 1999). Recently developed collection devices that sample the endocervix and exocervix simultaneously do not provide a significantly lower false negative rate than the combination of spatula and a conical cervical brush (Szarewski et al., 1993).

There is no consensus as to whether a single-slide technique, with both samples of the ectocervix and endocervix placed on the same slide, or a technique in which the two samples are put on two separate slides is preferable. Comparative studies of the two techniques have reported similar results (Saitas et al., 1995; Quackenbush, 1999). The single-slide approach has the advantage of reducing screening time and laboratory workload and it decreases the storage space required for archiving slides. When a single-slide technique is utilized, there also is no consensus on whether the specimens from the ectocervix and endocervix should be mixed together on the slide or kept separate as in the V (vagina) C (ectocervix) E (endocervix) technique.

**Collecting the sample**

A conventional cytology specimen is typically obtained using a spatula and conical cervical brush. The slide must first be labelled with the woman’s name or number. Laboratories should have a written protocol specifying what is considered adequate labelling and should not accept inadequately labelled specimens. The person collecting the specimen should ensure that a test requisition is accurately and legibly filled out before collecting the specimen. The information most commonly requested by laboratories includes:

- **Woman’s name and indication if there has been a name change in the last five years. Some laboratories also use unique patient identifier numbers**
- **Date of birth or age**
- **Menstrual status (date of last menstrual period, whether the woman is pregnant, post-partum, on hormone replacement therapy, or has had a hysterectomy)**
- **Previous history of abnormal cervical cytology, or treatment for CIN or cancer**
- **Whether the clinician considers the woman to be at high risk for developing CIN or cancer. Possible risk factors include smoking, infection with HIV, lack of previous screening and multiple partners.**
- **Specimen source – vaginal or cervical**

Good visualization of the cervix is important for obtaining an adequate specimen. Cervical cytology specimens are generally collected with the woman in the dorsolithotomy position. A sterilized or single-use bivalve speculum of appropriate size is inserted into the vagina in such a manner as to allow complete visualization of the cervical os and as much of the transformation zone as possible. The cervix should not be contaminated with lubricant or water-soluble gel that may obscure the smear. Therefore the smear must be obtained before any bimanual examination. Gentle removal of excess mucus and discharge from the cervix with a large cotton-tipped applicator can produce a better-quality smear (Kotaska & Matisic, 2003), but vigorous cleansing may remove many of the most easily exfoliated cells. Saline should not be used to help clear debris from the surface of the cervix. It is also preferable not to apply 3–5% acetic acid to the cervix before taking the cytology specimen, as this can reduce the cellularity of the smear and produce poor staining (Griffiths et al., 1989; Cronje et al., 1997).

Before the specimen is collected, the cervix should be carefully inspected with the naked eye for grossly visible masses or ulcerations that may indicate an invasive cervical cancer. If a grossly visible lesion is identified, the woman should be referred for further confirmation. In many cases, the lesion can be directly sampled and the cellular sample obtained can be submitted separately for cytological assessment. The procedure for collecting cells from the cervix varies depending on the type of device used and the number of slides to be prepared. If a spatula and conical cervical brush are utilized, the first step is to place the spatula firmly against the ectocervix with the long projection extending into the endocervical canal. The spatula is then rotated several times 360° around the portio and removed. It is important to ensure that the entire squamocolumnar junction is sampled, since this is the site where most CIN lesions develop. In most women, the spatula will come into contact with the squamocolumnar junction if the pointed end is placed in the os, but in young women with a large ectopy, the spatula may need to be moved laterally to sample a peripherally positioned squamocolumnar junction. When rotating the spatula, it is easy to miss part of the cervix; this can be alleviated by directly visualizing the cervix while sampling. Transfer is best performed by using the spatula to thinly...
spread the cells onto the glass slide. It is important to ensure that as much cellular material as possible is transferred from both sides of the spatula.

The endocervical canal is then sampled, using a conical cervical brush, which is placed in the endocervical canal so that the last few bristles remain visible and then gently rotated 90° to 180° once. One such rotation will adequately sample the endocervical canal and generally does not produce bleeding. Material from both sides of the spatula should be spread onto the slide.

If collection devices that simultaneously sample both the endocervix and the ectocervix are used, the manufacturers' directions should be followed for each type of device.

Cell fixation must be performed within a few seconds of specimen collection in order to prevent air-drying, which obscures cellular detail and hinders interpretation (Somrak et al., 1990). Immersing the slide in alcohol or spraying it with a specially formulated spray fixative can prevent air-drying. With immersion fixation, the slide is either immersed in alcohol and transferred to the laboratory in the container of alcohol or allowed to fix for 20 to 30 minutes in the alcohol, removed and allowed to air-dry. Various different spray fixatives are available. Only spray fixatives specifically designed for cytological specimens should be used and the manufacturer's instructions for a given product must be followed. The fixative should be liberally applied such that the slide appears moist over its entire surface. In order to prevent disruption of the cellular layer on the slide, the container of spray fixative should generally be held 15–25 cm from the slide during application.

**Performance of conventional cytology**

Despite the proven effectiveness of cervical cytological screening in reducing the incidence of cervical cancer, over the last decade the accuracy of cervical cytology has been questioned. Two factors need to be considered when assessing the accuracy of any screening or diagnostic test. One is whether the test is specific in detecting a given condition; the other is the sensitivity of the test for detecting the condition. Several large meta-analyses have indicated that both the sensitivity and specificity of cervical cytology are lower than previously thought (Fahey et al., 1995; McCrory et al., 1999; Nanda et al., 2000). [The Working Group considered the estimates of cytology test performance obtained through these meta-analyses to be of concern, given current cytology practices. In particular, it felt that it is very unlikely that specificities as low as 60–70% would be observed in a modern cytological screening practice.]

Table 19 presents the sensitivities and specificities of conventional cervical cytology observed in a number of recent large cervical cancer screening studies. Even within the confines of research studies, a wide range of performance has been reported.

**Liquid-based cervical cytology**

Liquid-based cytology (LBC) was introduced in the mid-1990s as a way to improve the performance of the test. Rather than having the clinician prepare the cytological specimen at the bedside by spreading the exfoliated cells onto a glass slide, the cells are transferred to a liquid preservative solution that is transported to the laboratory, where the slide is prepared.

### Table 19. Performance of conventional cytology in various large research studies

<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>Ages</th>
<th>Study size</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Histological cut-off</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cuzick et al. (1999a)</td>
<td>United Kingdom</td>
<td>34+</td>
<td>2988</td>
<td>86</td>
<td>98</td>
<td>CIN 2+</td>
</tr>
<tr>
<td>Hutchinson et al. (1999)</td>
<td>Costa Rica</td>
<td>18+</td>
<td>8636</td>
<td>55</td>
<td>98</td>
<td>CIN 2+</td>
</tr>
<tr>
<td>Ratnam et al. (2000)</td>
<td>Canada</td>
<td>18–69</td>
<td>2098</td>
<td>56</td>
<td>62</td>
<td>CIN 2+</td>
</tr>
<tr>
<td>Denny et al. (2000a)*</td>
<td>South Africa</td>
<td>35–65</td>
<td>2944</td>
<td>70</td>
<td>85</td>
<td>CIN 2+</td>
</tr>
<tr>
<td>Denny et al. (2002)*</td>
<td>South Africa</td>
<td>35–65</td>
<td>2754</td>
<td>40</td>
<td>96</td>
<td>CIN 1+</td>
</tr>
<tr>
<td>Cuzick et al. (2003)</td>
<td>United Kingdom</td>
<td>30–60</td>
<td>11 085</td>
<td>77</td>
<td>96</td>
<td>CIN 1+</td>
</tr>
<tr>
<td>Petry et al. (2003)</td>
<td>Germany</td>
<td>30+</td>
<td>8466</td>
<td>44</td>
<td>98</td>
<td>CIN 2+</td>
</tr>
<tr>
<td>Salmerón et al. (2003)</td>
<td>Mexico</td>
<td>15–85</td>
<td>7868</td>
<td>59</td>
<td>98</td>
<td>CIN 1+</td>
</tr>
<tr>
<td>Sankaranarayan et al. (2004b)</td>
<td>India</td>
<td>25–65</td>
<td>10 591</td>
<td>65</td>
<td>92</td>
<td>CIN 2+</td>
</tr>
</tbody>
</table>

Cytological cut-off for referral for all studies is ASC or greater except for those studies marked by asterisk, where a cut-off of LSIL or greater was used.

Sensitivity and specificity are estimated cross-sectionally (see Chapter 4).
A number of different LBC techniques are in use worldwide. These include ThinPrep®, SurePath™, Cytoscreen™, Cytoseasy®, Labonord Easy Prep, Cytoslide, SpinThin and PapSpin. The first two of these are approved for use in the USA by the Food and Drug Administration (FDA) and are the most widely used methods worldwide. They are therefore the best characterized in terms of performance. With the ThinPrep method, clumps of cells and mucus are broken up by mechanical agitation and then the liquid preservative solution is filtered through a membrane filter with a pore size specifically designed to trap epithelial cells while allowing contaminating red blood cells and inflammatory cells to pass through. The epithelial cells collected on the membrane filter are then transferred onto a glass slide and stained. This produces a relatively thin, monolayer-type preparation. The ThinPrep-2000 processor allows one specimen to be processed at a time, whereas the newer ThinPrep-3000 processor is more fully automated and allows up to 80 samples to be processed at a time. In contrast, with the SurePath method, clumps of cells and mucus are broken up by aspiration through a syringe. The cell suspension is then layered on top of a density gradient and the red blood cells and inflammatory cells are separated from the epithelial cells by density gradient centrifugation. The resulting cell pellet containing predominantly epithelial cells is then inserted into a robotic workstation, where it is resuspended and transferred to a glass microscope slide. The SurePath method allows up to 48 samples to be processed at a time.

LBC is purported to have a number of advantages over conventional cervical cytology. These include a more representative transfer of cells from the collection device to the glass slide, a reduction in the number of unsatisfactory cytology specimens, the availability of residual cellular material for subsequent molecular testing or for making additional glass slides, and possibly increased detection of HSIL.

**Performance of liquid-based cytology methods**

Numerous studies have evaluated the comparative performance of the two most commonly used LBC methods (ThinPrep and SurePath) and conventional cytology with respect to test positivity, their sensitivity and specificity for identification of CIN, the time required for evaluation of the specimens, and specimen adequacy. Although there is reasonable agreement that LBC improves specimen adequacy and reduces screening time compared to conventional cytology, there is considerable controversy surrounding the relative sensitivity and specificity of the two approaches, largely due to a lack of well designed comparative studies.

Most comparative studies have utilized one of two types of study design: split-sample studies and historical control studies. Split-sample studies collect cells from the cervix using a single collection device and a conventional cervical cytology specimen is prepared first. Residual cells remaining on the device are then transferred to a liquid-based cytology preservative. Therefore each woman acts as her own control and detection rates in conventional and LBC specimens are compared. The other widely used study design, known as ‘direct to vial’, compares the performance of LBC collected in the routine manner (direct transfer to the preservative solution) during a given time period with historic control data obtained using conventional cytology. Both study designs have significant limitations. With split-sample studies, it is difficult to ensure that the two cytology specimens are comparable. Since the conventional cytology slide is prepared first and the LBC specimen is prepared second, this design would seem to lead inherently to bias against LBC. Therefore it has been argued that split-sample studies do not demonstrate the full benefit that could be obtained when LBC is utilized in routine clinical practice. Studies utilizing historical controls avoid the need to prepare several cytology specimens from a single woman, but introduce other potential biases, including the comparability of the populations being compared.

Other significant limitations found in many of the studies evaluating LBC include failure to compare test performance with a reference standard of ‘blinded’ colposcopy/biopsy and a study population of women followed up for a prior cytological abnormality rather than women undergoing routine screening. A review of new cervical cytology methods conducted in 2001 for the US Preventive Services Task Force and the Agency for Healthcare Research and Quality found that out of 962 potentially relevant studies, not one met their predefined inclusion criteria (Hartmann et al., 2001). This was commonly due to lack of an adequate reference standard, but most studies were excluded for more than one reason. At the time the review was conducted, only one study, from Costa Rica, had applied a definitive clinical reference standard to a random sample of women with normal screening test results and allowed the relative sensitivity and specificity for LBC and conventional cervical cytology to be calculated (Hutchinson et al., 1999). The Costa Rica study was a split-sample study rather than a direct-to-vial study. The other studies that were reviewed used various types of clinical reference standard, including a combination of histological follow-up and conventional cytology follow-up with incomplete data, a consensus expert panel diagnosis of the index specimens, histological follow-up or...
consensus expert panel diagnosis in cases of missing follow-up, and histological follow-up of HSIL combined with a balanced follow-up diagnosis of all other available follow-up data. The most common reference standard used in studies of LBC performance has been the expert panel review of selected cytology specimens. Unfortunately, with expert panel review, the screening test findings are not related to the true disease status of the cervix, making determination of false negatives and false positives, and hence sensitivity and specificity, impossible. Cervical biopsy diagnoses obtained as part of routine follow-up of women with abnormal cervical cytology results are another commonly used reference standard in studies of LBC. However, unless the pathologist is blinded to the original cytological findings, it is quite possible that the interpretation of cervical biopsy specimens will be biased. Large, randomized controlled clinical trials comparing the performance of LBC and conventional cytology need to be conducted by laboratories in which the techniques are well established. Although the results of no such studies are yet available, one large randomized trial is currently under way in the Netherlands (M.A. Arbyn, personal communication).

Several systematic, evidence-based reviews of the published literature on LBC have been published (Nanda et al., 2000; Payne et al., 2000; Bernstein et al., 2001; Hartmann et al., 2001; Sulik et al., 2001; Abulafia et al., 2003; Klinkhamer et al., 2003; Arbyn et al., 2004a). These reviews are based on test positivity ratios or detection rates, i.e., relative sensitivities and specificities of histologically confirmed lesions. They have come to somewhat conflicting conclusions (Table 22). It is important to note that the comparative utility of LBC relative to conventional cervical cytology will vary from one setting to another. The National Health Service of the United Kingdom recently agreed to introduce LBC throughout the country, in view of the reduction of inadequate specimens from 9% with conventional cervical cytology to 1–2% with LBC (National Institute for Clinical Excellence, 2003).

Table 20 presents data from a number of ‘direct-to-vial’ studies. Although there is considerable variation between the studies in the prevalence of HSIL identified using either conventional cytology or LBC, on average the use of LBC increased the rate of detection of HSIL in these stud-

<table>
<thead>
<tr>
<th>Reference</th>
<th>LBC test</th>
<th>Population</th>
<th>Conventional No.</th>
<th>Conventional LSIL</th>
<th>Conventional HSIL</th>
<th>Liquid-based cytology No.</th>
<th>Liquid-based LSIL</th>
<th>Liquid-based HSIL</th>
<th>Increase in HSIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bolick &amp; Hellman (1998)</td>
<td>TP</td>
<td>Screening</td>
<td>39 408</td>
<td>0.8%</td>
<td>0.3%</td>
<td>10 694</td>
<td>2.3%</td>
<td>0.8%</td>
<td>173%</td>
</tr>
<tr>
<td>Dupree et al. (1998)</td>
<td>TP</td>
<td>Screening</td>
<td>22 323</td>
<td>0.9%</td>
<td>0.2%</td>
<td>19 351</td>
<td>1.4%</td>
<td>0.3%</td>
<td>50%</td>
</tr>
<tr>
<td>Papillo et al. (1998)</td>
<td>TP</td>
<td>Screening</td>
<td>18 569</td>
<td>0.9%</td>
<td>0.5%</td>
<td>8541</td>
<td>1.6%</td>
<td>0.7%</td>
<td>55%</td>
</tr>
<tr>
<td>Carpenter &amp; Davey (1999)</td>
<td>TP</td>
<td>High-risk</td>
<td>5000</td>
<td>4.4%</td>
<td>1.9%</td>
<td>2727</td>
<td>6.9%</td>
<td>2.4%</td>
<td>26%</td>
</tr>
<tr>
<td>Diaz-Rosario &amp; Kabawat (1999)</td>
<td>TP</td>
<td>Screening</td>
<td>74 756</td>
<td>1.6%</td>
<td>0.26%</td>
<td>56 339</td>
<td>2.7%</td>
<td>0.52%</td>
<td>102%</td>
</tr>
<tr>
<td>Guidos &amp; Selvaggi (1999)</td>
<td>TP</td>
<td>Screening</td>
<td>5423</td>
<td>1.0%</td>
<td>0.3%</td>
<td>9583</td>
<td>3.6%</td>
<td>1.0%</td>
<td>233%</td>
</tr>
<tr>
<td>Vassilakos et al. (1999)</td>
<td>SP</td>
<td>Screening</td>
<td>88 569</td>
<td>1.6%</td>
<td>0.4%</td>
<td>111 358</td>
<td>2.5%</td>
<td>0.7%</td>
<td>79%</td>
</tr>
<tr>
<td>Hatch (2000)</td>
<td>TP</td>
<td>High-risk</td>
<td>16 260</td>
<td>2.9%</td>
<td>1.5%</td>
<td>7934</td>
<td>6.1%</td>
<td>3.2%</td>
<td>116%</td>
</tr>
<tr>
<td>Tench (2000)</td>
<td>SP</td>
<td>Screening</td>
<td>10 367</td>
<td>0.6%</td>
<td>0.5%</td>
<td>2231</td>
<td>1.0%</td>
<td>0.7%</td>
<td>46%</td>
</tr>
<tr>
<td>Weintraub &amp; Morabia (2000)</td>
<td>TP</td>
<td>Screening</td>
<td>126 619</td>
<td>0.5%</td>
<td>0.1%</td>
<td>39 455</td>
<td>1.8%</td>
<td>0.5%</td>
<td>400%</td>
</tr>
<tr>
<td>Obwegeser &amp; Brack (2001)</td>
<td>TP</td>
<td>Screening</td>
<td>1002</td>
<td>3.7%</td>
<td>1.8%</td>
<td>997</td>
<td>4.7%</td>
<td>1.6%</td>
<td>– 11%</td>
</tr>
<tr>
<td>Baker (2002)</td>
<td>TP</td>
<td>Screening</td>
<td>4872</td>
<td>2.8%</td>
<td>0.7%</td>
<td>3286</td>
<td>4.1%</td>
<td>1.0%</td>
<td>43%</td>
</tr>
<tr>
<td>Cheung et al. (2003)</td>
<td>TP</td>
<td>Screening</td>
<td>191 581</td>
<td>1.0%</td>
<td>0.25%</td>
<td>190 667</td>
<td>1.7%</td>
<td>0.24%</td>
<td>– 4%</td>
</tr>
<tr>
<td>Moss et al. (2003)</td>
<td>TP</td>
<td>Screening</td>
<td>67 856</td>
<td>2.3%</td>
<td>1.4%</td>
<td>34 128</td>
<td>2.6%</td>
<td>1.7%</td>
<td>21%</td>
</tr>
<tr>
<td>SP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colgan et al. (2004)</td>
<td>SP</td>
<td>Screening</td>
<td>445 225</td>
<td>1.4%</td>
<td>0.40%</td>
<td>445 011</td>
<td>1.8%</td>
<td>0.35%</td>
<td>–</td>
</tr>
</tbody>
</table>

Abbreviations: TP, ThinPrep; SP, SurePath
ies. The wide variety of study populations makes comparisons difficult. This is because estimates of performance are influenced by outlying results of a few studies. In a comprehensive formal meta-analysis of all published ‘direct-to-vial’ studies that adjusted for outlying results, Arbyn et al. (2004a) found a pooled ratio for detection rate of HSIL in ThinPrep specimens versus conventional cytology of 1.72 (95% CI 1.42–2.08) and for SurePath specimens versus conventional cytology of 1.47 (95% CI 1.14–1.89). It is important to bear in mind the limitations to interpretation of these studies, as described above, and that the actual number of additional cases classified as HSIL using LBC is quite small—only about three cases per 1000 women screened.

**Specimen adequacy**

The effect of LBC on specimen adequacy rates has been evaluated in a number of the ‘direct-to-vial’ studies (Table 21). Both the ThinPrep and SurePath methods appear to produce fewer specimens classified as either ‘limited by obscuring factors’ such as blood, inflammation or poor preservation than does conventional cytology. In addition, in many studies both methods have reduced the number of specimens classified as ‘unsatisfactory for evaluation’. In a recent pilot study in the United Kingdom (Moss et al., 2003), the use of ThinPrep reduced the ‘inadequate’ rate from 9.7% to 2.0%. The use of SurePath reduced the ‘inadequate’ rate from 9.1% to 0.9%. For all study sites combined, there was an 82.7% reduction (rate ratio 0.173, 95% CI 0.17–0.19). The reduction was significant in each of three age groups: 20–34, 35–49 and 50–64 years. In a meta-analysis of the comparative performance of LBC, Arbyn et al. (2004a) estimated the ratio of the inadequacy rate versus conventional cytology of ThinPrep in ‘direct-to-vial’ studies to be 0.70 (95% CI 0.39–1.27) and of SurePath to be 0.13 (95% CI 0.07–0.26).

**Specimen interpretation time**

A few studies have evaluated the impact of LBC on specimen interpretation time. LBC seems to be associated with shorter interpretation times than required for conventional cytology specimens. Payne et al. (2000), in their systematic review for the United Kingdom National Health Service, provided estimates of three minutes for LBC compared with 4–6 minutes for conventional cytology. This is not surprising given that the total surface area that needs to be screened is considerably less for both ThinPrep and SurePath than for conventional cytology specimens. The need for continuous adjustment to focus is also reduced using LBC, since the cells tend to be in the same plane of focus. With conventional cytology specimens, the screener needs to continually adjust the focus to evaluate clusters of cells. Payne et al. (2000) reported, however, that cytologists in Edinburgh found screening monolayers to require more intense concentration than screening conventional cytology specimens, making it more tiring. In part, this reflects the fact that occasionally only one or two HSIL cells are present.

---

**Table 21. Comparison of specimen adequacy in conventional cytology with LBC in "direct-to-vial" studies**

<table>
<thead>
<tr>
<th>Reference</th>
<th>LBC test</th>
<th>Population</th>
<th>Conventional</th>
<th>Liquid-based cytology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>No. Limited (%)</td>
<td>Unsatisf. (%)</td>
</tr>
<tr>
<td>Bolick and Hellman (1998)</td>
<td>TP</td>
<td>Screening</td>
<td>39 408 17.8</td>
<td>1.0</td>
</tr>
<tr>
<td>Dupree et al. (1998)</td>
<td>TP</td>
<td>Screening</td>
<td>22 323 2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Diaz-Rosario and Kabawat (1999)</td>
<td>TP</td>
<td>Screening</td>
<td>74 756 22.0</td>
<td>0.2</td>
</tr>
<tr>
<td>Carpenter and Davey (1999)</td>
<td>TP</td>
<td>High-risk</td>
<td>5000 19.4</td>
<td>0.6</td>
</tr>
<tr>
<td>Guidos and Selvaggi (1999)</td>
<td>TP</td>
<td>Screening</td>
<td>5423 21.4</td>
<td>1.2</td>
</tr>
<tr>
<td>Vassilakos et al. (1999)</td>
<td>SP</td>
<td>Screening</td>
<td>88 569 4.7</td>
<td>1.5</td>
</tr>
<tr>
<td>Tench (2000)</td>
<td>SP</td>
<td>Screening</td>
<td>10 367 31.0</td>
<td>2.9</td>
</tr>
<tr>
<td>Weintraub and Morabia (2000)</td>
<td>TP</td>
<td>Screening</td>
<td>130 050 27.8</td>
<td>0.3</td>
</tr>
<tr>
<td>Obwegeser and Brack (2001)</td>
<td>TP</td>
<td>Screening</td>
<td>1002 2.5</td>
<td>0.0</td>
</tr>
<tr>
<td>Baker (2002)</td>
<td>TP</td>
<td>Screening</td>
<td>4872 18.2</td>
<td>0.7</td>
</tr>
<tr>
<td>Cheung et al. (2003)</td>
<td>TP</td>
<td>Screening</td>
<td>191 581 2.6</td>
<td>0.48</td>
</tr>
<tr>
<td>Moss et al. (2003)</td>
<td>SP</td>
<td>Screening</td>
<td>47 632 9.1</td>
<td>0.9</td>
</tr>
</tbody>
</table>
### Table 22. Systematic reviews of comparative performance of LBC and conventional cytology

<table>
<thead>
<tr>
<th>Author</th>
<th>Subgroup</th>
<th>Indicator</th>
<th>Key conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nanda et al. (2000)</td>
<td>ThinPrep</td>
<td>Histologically confirmed lesion</td>
<td>Higher sensitivity of LBC, but only three studies were evaluated</td>
</tr>
<tr>
<td>Payne et al. (2000)</td>
<td>Test positivity or histologically confirmed lesion</td>
<td></td>
<td>Some evidence that LBC offers an improvement in sensitivity</td>
</tr>
<tr>
<td>Bernstein et al. (2001)</td>
<td>ThinPrep</td>
<td>Test positivity</td>
<td>ThinPrep is as good as, or superior to, conventional cytology for diagnosing CIN</td>
</tr>
<tr>
<td>Hartmann et al. (2001)</td>
<td>All studies</td>
<td>Histologically confirmed lesion</td>
<td>Current evidence is inadequate to gauge whether LBC is “better” than conventional cytology</td>
</tr>
<tr>
<td>Sulik et al. (2001)</td>
<td>Histologically confirmed lesion</td>
<td>LBC demonstrated higher sensitivity (90%; 95% CI 77–96%) than conventional cytology (79%; 95% CI: 59–91%) for CIN 2 or more severe</td>
<td></td>
</tr>
<tr>
<td>Abulafia et al. (2003)</td>
<td>ThinPrep only</td>
<td>Test positivity</td>
<td>ThinPrep tends to be more sensitive than conventional smears in detecting CIN</td>
</tr>
<tr>
<td>Klinkhamer et al. (2003)</td>
<td>All studies</td>
<td>Histologically confirmed lesion</td>
<td>Indications that SurePath has lower sensitivity than conventional for ASC or greater No definitive statement can be made for detection of LSIL or higher or HSIL or higher for SurePath because of conflicting results Indications that ThinPrep has higher sensitivity than conventional for ASC or greater Likely that ThinPrep has higher sensitivity than conventional for LSIL or higher Likely that ThinPrep has higher positivity rate and greater absolute sensitivity than conventional for HSIL</td>
</tr>
<tr>
<td>Arbyn et al. (2004a)</td>
<td>Split-sample studies</td>
<td>Test positivity</td>
<td>More LSIL in LBC than in conventional cytology Positivity rates for ASC and HSIL not statistically different More LSIL detected by LBC 80% (95% CI 52–112%) ThinPrep; 54% (95% CI 25–90%) SurePath More HSIL detected by LBC 72% (95% CI 42–108%) ThinPrep; 47% (95% CI 14–89%) SurePath Positivity rates for ASC were the same There was no reduction in positive predictive value for CIN 2 and CIN 3 of LBC versus conventional cytology</td>
</tr>
<tr>
<td></td>
<td>Direct-to-vial studies</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
in an LBC specimen, necessitating careful scrutiny of every individual cell.

**Availability of residual cellular material for molecular testing**

One of the major benefits of LBC in many settings is the availability of residual cellular material for molecular testing for agents such as *Chlamydia trachomatis* or HPV. In the USA, the 2001 Consensus Guidelines for the Management of Women with Cytologic Abnormalities considered HPV DNA testing of residual LBC fluid to be the preferred approach to managing women with ASCUS cytological results (Wright *et al.*, 2002a), on the grounds that such reflex HPV DNA testing offers the advantage that women do not need to return to the office or clinic for an additional clinical examination. In addition, the 40–60% of women who are high-risk HPV DNA-negative will be spared a colposcopic examination and can be rapidly assured that they do not have a significant cervical lesion. A comprehensive study of triage methods for women with ASCUS indicated that reflex HPV DNA testing provides the same or greater life expectancy benefits and is more cost-effective than either a programme of repeat cytology or immediate colposcopy (Kim *et al.*, 2002).

**Quality assurance and quality control issues**

An advantage of cervical cytology over screening methods such as visual screening is that even though quality assurance and quality control programmes can be developed for both, the availability of archival glass slides facilities such programmes. Various definitions for quality control and quality assurance are used by laboratories. In general, *quality control* can be thought of in terms of the actual assessments that are done to ensure high quality and *quality assurance* can be thought of in terms of the entire process of maintaining minimum standards and continually striving for excellence. Quality assurance should be a coordinated effort that is designed to control, detect and prevent the occurrence of errors and hopefully to improve patient care. In general, there are three stages to the process of quality control (Bozzo, 1991):

- Setting standards for what one wishes to control and defining the benchmarks;
- Developing a mechanism for assessing what one wishes to control;
- Defining the response to be taken when deficiencies are identified.

For cervical cytology screening, quality assurance programmes can include a number of types of activity and should take into account country- and location-specific needs. What may be considered acceptable or even mandatory in one setting may serve simply to limit the availability of screening in other settings. It is critical, however, that any cervical cytology laboratory or programme have an established quality assurance programme. In general, it is preferable for cytology services to be centralized as much as possible, to facilitate quality assurance. The use of computerized data collection systems that can integrate cytological findings, histological findings and follow-up information is highly desirable (Miller *et al.*, 2000).

**Preanalytic quality control**

Preanalytic quality control measures include the records that laboratories should maintain relating to specimen receipt, preparation of specimens, staining of specimens and upkeep of equipment and microscopes, as well as records of personnel and their training and education.

**Training**

Training of both the cytotechnicians who perform the initial screening in the laboratory and the pathologists who provide the final interpretation is critical to obtaining optimal performance of a cervical cytology programme. Cytopathologists should either receive formal training in an established academic programme or be trained in an established national centre for cervical cytology for at least six months (Miller *et al.*, 2000). This training should typically include not only the interpretation of cervical cytology specimens, but also cervical histopathology. A cytopathologist who will run a laboratory is generally selected for leadership potential and ability to organize, run and manage a successful cytopathology laboratory, and will require training in laboratory management skills.

In most developed countries, cytotechnicians undergo 1–2 years of formal didactic training in order to develop a high level of competence in evaluating all types of cytological specimens, including gynaecological cytology. However, in some countries consideration is now being given to intensive six-month training programmes focusing only on gynaecological cytology. In addition, cytotechnicians should periodically participate in competence-based education programmes. Unfortunately, cytotechnology training programmes are not available in many developing countries and extended formal training programmes are not an option. In these settings, cytotechnicians are often ‘bench-trained’, being tutored by a person with some level of training in interpreting gynaecological cytology specimens. Training in this manner should be avoided unless the laboratory where it occurs processes at least 15 000 specimens annually and training should last for at least six months (Miller *et al.*, 2000). Although there is little evidence that cytotechnicians who are trained in such a ‘hands-
on’ fashion perform less well than those who receive formal training, the variability in training inherent in this approach is a cause for concern. Whenever possible, cytotechnicians should receive formal, structured, competence-based training in interpreting cervical cytology specimens. The International Federation of Cytology has an international qualification for cytotechnicians, which can be used to ensure that competence has been obtained.

Workload limits – maximum and minimum
It is now widely accepted that, because of the repetitive nature of screening cytology specimens, there should be workload limits on the number of specimens that a cytotechnician can screen in any given period. In the USA, federal regulations require that anyone performing primary screening of cervical cytology specimens should evaluate no more than 100 cytology specimens per 24-hour period and in not less than eight hours (Federal Register, 1992). In addition, every laboratory must establish individual workload limits for each cytotechnician, based on their experience and skill. This must be reassessed every six months using laboratory-defined performance standards. In many European countries, this workload limit is considered too high and other limits are used. In the United Kingdom, for example, time limits rather than slide limits are used. Cytotechnicians are restricted to screening for only four hours per day, regardless of whether they are screening conventional or LBC specimens. Since LBC specimens can be screened more rapidly than conventional cytology specimens, this means that greater numbers of LBC specimens can be screened by each cytotechnician. A recent consensus panel recommended that a daily workload limit of 60 cases was preferable (Miller et al., 2000).

It is also important that a laboratory process a minimum number of specimens per year in order to maintain an adequate level of competence (Krieger et al., 1998). In reviews of US laboratories by the College of American Pathologists, screening error rates were found to be greatest in laboratories processing less than 5000 specimens per year and having no dedicated screening cytotechnicians (College of American Pathologists, 1997). In the United Kingdom, laboratories are now required to process at least 15,000 specimens per year. Evaluation of a minimum annual number of specimens is also to be considered desirable in low-resource settings. The Peruvian Society of Cytopathology does not certify laboratories that process under 5000 specimens annually (Salvetto & Sandiford, 2004). A recent World Health Organization consensus panel recommended that each laboratory should process at least 20,000 specimens yearly in order to maintain acceptable skills (Miller et al., 2000).

Review of abnormal cases
It is generally accepted that a pathologist should review all specimens deemed by the screening cytotechnician to have any degree of cytological abnormality (American Society of Cytopathology, 2001). Identification of discordant cases provides an element of quality control for the screening process and allows identification of specific cytotechnicians and specific areas of cytology requiring additional education. It is important for quality monitoring that all reviews be documented.

Rescreening of negative cases
Some form of rescreening of specimens initially considered negative is important for quality control. In the USA, federal regulations stipulate that at least 10% of all samples interpreted as negative by each cytotechnician must be reassessed by either a pathologist or a supervising cytotechnologist before the result is reported (Federal Register, 1992). This regulation is controversial for a variety of reasons. One is the level of discrepancy that is considered significant. It has been argued that negative specimens classified as atypical (e.g., ASC) upon review should not be considered errors, because of the inherent subjectivity of this diagnosis (Krieger et al., 1998). Another problem with performing 10% rescreening of negative cases is that significant lesions are quite uncommon in the reviews. Given an underlying rate of SIL of only 2–3% in the screened population and assuming that even a poorly performing cytotechnician will be able to identify 75% of specimens containing SIL, large numbers of specimens must be rescreened to determine which cytotechnicians or laboratories are performing poorly. This lack of statistical power greatly hampers its use as a quality control measure (Hutchinson, 1996).

The technique of rapid rescreening of all negative specimens has been the subject of a number of studies and appears to present an attractive alternative to the 10% rescreening approach. Using this technique all, or most, of the specimens classified as negative by a laboratory undergo a second, more rapid evaluation by a different screener. This is the approach to rescreening adopted in the United Kingdom by the National Health Service (NHSCSP, 2000). Another approach is referred to as ‘prescreening’, in which all specimens undergo a rapid review before the intensive screening. In a recent meta-analysis of published data, Arbyn et al. (2003) found that the pooled estimated sensitivity of rapid prescreening for HSIL or more severe lesions was 86% and that the technique showed diagnostic properties that support its use as a quality...
control measure. The same group previously demonstrated that rapid pre-screening was superior to 10% random rescreening in identifying cases that were missed (Arbyn & Schenck, 2000).

**Cytology–histology correlations and clinical follow-up**

If a laboratory has access to histological specimens obtained at the time of colposcopy for an abnormal cytological finding, it should compare all premalignant and malignant cytological results with the histopathological observations. This allows the laboratory to refine its cytological criteria. If histological specimens are not available, the laboratory may attempt to obtain referral and follow-up information. It is important that the laboratory obtain follow-up information on women with HSIL to ensure that they have not been lost to follow-up.

**Measuring the performance of the laboratory**

Laboratories need to carefully and continuously monitor their performance as a whole, as well as that of individual cytotechnicians. Information that can be useful for a given laboratory includes the percentages of specimens classified as having a given result (e.g., ASC, LSIIL, HSIL, etc.), the rate of unsatisfactory specimens, the ASC:LSIIL ratio, the laboratory turnaround time, etc. One of the most important measures is screening sensitivity (Krieger et al., 1998; NHSCSP, 2000; American Society of Cytopathology, 2001). However, it is very difficult to determine the sensitivity of screening in a real-world laboratory setting. One approach that has been proposed to estimate screening sensitivity in a laboratory is to calculate the ‘false negative proportion’, which is essentially the number of false negative LSIL or greater specimens identified through a 100% rapid rescreen programme divided by the total number of LSIL or greater specimens identified through regular screening and the rapid rescreen process combined (Krieger et al., 1998).

**Proficiency testing**

Proficiency testing programmes provide laboratories, cytopathologists and individual cytotechnicians with sets of stained cytology specimens on which the interpretation has been agreed to according to a set procedure. The slide sets are then evaluated by the person being tested and their interpretation is compared with that of the panel or with their peers (Coleman & Evans, 1999; NHSCSP, 2000). This allows the performance of both whole laboratories and individual cytotechnicians or cytopathologists to be compared against others in an unbiased manner. Periodic retesting should be conducted every 6–12 months (Miller et al., 2000). Individuals who perform poorly on proficiency testing should receive additional training to improve their skills and any who continue to perform poorly after retraining should be reassigned to non-screening tasks.

Recent evidence suggests that performance on proficiency testing provides some evidence of the real-world performance of cytotechnicians (Keenlyside et al., 1999). A recent report from Peru and Nicaragua has shown that proficiency testing can be implemented successfully in developing countries (Salvietto & Sandiford, 2004).

**Visual inspection**

The use of visual inspection methods to screen for cervical neoplasia began with the use of Schiller’s test in the 1930s (Schiller, 1933). In the 1980s, the idea of looking at the cervix with the naked eye for early detection of disease (known as ‘down-staging’) in low-resource settings was promoted (Stjernswärd, 1987). Over the last ten years, the use of dilute (3–5%) acetic acid applied to the cervix before inspection (visual inspection with acetic acid, VIA) has been investigated. More recently, the application of Lugol’s solution has been used and is referred to as visual inspection with Lugol’s iodine (VILI).

Visual tests are inherently subjective. Published studies of the test performance characteristics vary with regard to important methodological aspects that result in biases and other difficulties in generalizing the findings to other populations. For example, studies may use different definitions of test positivity. Differences in training of personnel and in the light sources used also generate variability in test performance characteristics across different study settings. Varying abilities of colposcopists to detect lesions and of pathologists to interpret histology accurately also affect the assessment of test performance.

Colposcopy with directed biopsy is the usual reference standard by which the performance of visual tests is assessed, but biases may impair the validity of the assessment. Verification bias arises if colposcopy is not applied equally to all women because of the study design. Blinding between those performing the visual test under evaluation and those performing the reference colposcopy is crucial to avoid information or expectation bias. Test performance for detection of CIN 2 or worse lesions and potential biases of all studies reviewed are summarized below.

**Unaided visual inspection (VI)**

Visual inspection (VI) (also called ‘down-staging’ or ‘unaided visual inspection’) consists of a clinical examination of the cervix using only a speculum and a light source. Test positivity is defined by the presence or
absence of specific characteristics, usually with low and high thresholds of positivity (Table 23). Only one of the six published studies reporting test characteristics of VI (Table 24) did not suffer from obvious verification bias (see Glossary and Chapter 4) (Basu et al., 2002); it found sensitivity to be low (<50%) irrespective of the threshold used to define test positivity. The other five studies used cytology as the reference standard. In all studies, the high-threshold definition of test positivity (corresponding to a 5–10% positivity rate) was associated with rather low sensitivity (30–60%). Gains in sensitivity using the low-threshold definition of test positivity led to concomitant decreases in specificity. Thus it is clear that VI lacks sufficient sensitivity for use as a primary screening test.

Visual inspection using acetic acid (VIA)
VIA involves naked-eye inspection of the cervix one minute after application of a 3–5% solution of acetic acid using a cotton swab or a spray. Test positivity is based on the appearance of acetowhite areas in the transformation zone, close to the squamocolumnar junction or the os. The cervix is examined using a bright light source such as a torch or halogen focus lamp. VIA is also known as direct visual inspection (DVI), acetic acid test (AAT) and cervicography.

Dilute acetic acid causes what is thought to be a reversible coagulation of intracellular proteins, resulting in noticeable opacity and a decrease in the usual reddish hue imparted by the subepithelial vascularature. This effect, called acetowhiteing, is not specific to cervical neoplasia and may also occur in immature squamous metaplasia and in inflamed, regenerating cervical epithelium. The degree of opacity due to the acetowhite reaction varies according to the thickness of the neoplastic change present in the epithelium and thus according to the grade of intraepithelial neoplasia.

The most common features observed using VIA are summarized in Table 25. VIA results are reported using negative and positive categories. VIA-positive cervices are illustrated in Figure 42.

In 17 published studies, test positivity rates ranged from 3% to 53% (Table 26). Seven studies were designed to minimize verification bias. In two other studies (Denny et al., 2000a, 2002), only women negative by cytology, VIA, HPV DNA testing and cervicography were not subjected to colposcopy, reducing susceptibility to bias. Seven of these nine studies (Londhe et al., 1997; University of Zimbabwe/JHPIEGO Cervical Cancer Project, 1999; Denny et al., 2000a; Belinson et al., 2001; Denny et al., 2002; Cronjé et al., 2003; Sankaranarayanan et al., 2004a), accounting for more than 95% of the total sample size, reported sensitivities of approximately 75%.

One study used a gold standard for enhanced disease ascertainment that was based on directed biopsy of any abnormal area(s), four-quadrant biopsies and endocervical curettage (ECC) in all women (Belinson et al., 2001). The other study that had enhanced design features was based on 55 000 women enrolled at 11 sites in six West African countries and India (Sankaranarayanan et al., 2004a). Each site followed a common testing protocol that included VIA, VILI and colposcopy with directed biopsy, as required, performed by separate individuals. Although similar training methods and test result definitions were used, there was substantial variation in the reported positivity rates (7–27%), sensitivity (56–94%) and specificity (74–94%).

Numerous studies have shown VIA to have sensitivity similar to that of cervical cytology for identifying women with HSIL, but much lower specificity (Table 27). Only two studies compared the accuracy of VIA and HPV DNA testing (Table 28); these showed the two tests to have similar accuracy. The reproducibility of VIA has been documented to be equivalent to that of histology, cytology and colposcopy (Sellors et al., 2002). In the multicentre study in Africa and India, the agreement between master trainers and local providers using 36 cervical photographs was fair (raw agreement, 64.5%; kappa, 0.38) (Sankaranarayanan et al., 2004a).

Visual inspection using acetic acid with low-level magnification (VIAM)
VIAM is VIA with low-level magnification (2–4 x), using a hand-held device to inspect the cervix one minute after application of acetic acid. Table 29 presents test results from four studies comparing VIA and VIAM. None of these studies documented any significant difference in test performance characteristics between VIAM and VIA.

Table 23. Test definition for visual inspection

<table>
<thead>
<tr>
<th>Test definition</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Normal-looking cervix, nabothian cysts</td>
</tr>
<tr>
<td>Positive (low threshold)</td>
<td>Cervicitis, erosion, polyp, wart, unhealthy cervix, reddish-looking cervix</td>
</tr>
<tr>
<td>Positive (high threshold)</td>
<td>Low-threshold features plus bleeding on touch, bleeding erosion, hypertrophied elongated cervix, growth, ulcer</td>
</tr>
</tbody>
</table>
A correlation study, with a sample size of 2080 previously screened women and a positivity rate of approximately 5% for VIAM, reported poor associations between VIAM and HPV test positivity and between VIAM and cytology test results (Rodriguez et al., 2004).

**Visual inspection using Lugol’s iodine (VILI)**

The use of Lugol’s solution to aid inspection of the cervix with the naked eye was described in 1933 by Schiller, but fell into disuse as cytological testing became available (Schiller, 1933; Wright, 2003). Several decades later, research on visual inspection methods led to the observation that nurses and midwives recognized non-staining areas on the cervix after application of Lugol’s solution more readily than acetowhite areas (Sankaranarayanan & Wesley, 2003), which led to renewed interest in this technique (referred to in the past as ‘Lugol’s iodine test’ and ‘Schiller’s iodine test’).

Lugol’s iodine stains glycogen stored in cervical epithelial cells. Mature squamous epithelium stores more glycogen than either columnar epithelium or immature squamous metaplastic epithelium. The application of iodine solution to the cervix thus results in black or dark brown staining of mature squamous epithelium. Columnar epithelium does not stain and retains its reddish hue. Areas of immature metaplasia stain a very light brownish hue, if at all. Neoplastic squamous epithelium

---

**Table 24. Studies of visual inspection**

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample size</th>
<th>Population (age, recruitment, location)</th>
<th>Provider</th>
<th>Reference diagnosis</th>
<th>Positivity rate (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singh et al. b</td>
<td>44 970</td>
<td>Opportunistic India</td>
<td>Cytology/ histology</td>
<td>(H) 11 (L) 69</td>
<td>63</td>
<td>89</td>
<td>Verification bias</td>
<td></td>
</tr>
<tr>
<td>Bhargava et al. b</td>
<td>3608</td>
<td>Opportunistic India</td>
<td>Midwife</td>
<td>Cytology/ histology</td>
<td>(H) 5 (L) 65</td>
<td>25</td>
<td>96</td>
<td>Verification bias</td>
</tr>
<tr>
<td>Sujathon et al.</td>
<td>3602</td>
<td>Opportunistic India and referred India</td>
<td>Cyto-pathologist</td>
<td>Cytology</td>
<td>(L) 63</td>
<td>89</td>
<td>50</td>
<td>Not designed for accuracy estimation</td>
</tr>
<tr>
<td>Nene et al. c</td>
<td>2135</td>
<td>Community-based</td>
<td>Health worker</td>
<td>Cytology/ histology</td>
<td>(H) 6 (L) 57</td>
<td>60</td>
<td>43</td>
<td>Verification bias</td>
</tr>
<tr>
<td>Wesley et al.</td>
<td>2843</td>
<td>Opportunistic India</td>
<td>Health worker</td>
<td>Cytology/ histology</td>
<td>(H) 6 (L) 45</td>
<td>29</td>
<td>94</td>
<td>Verification bias</td>
</tr>
<tr>
<td>Basu et al.</td>
<td>6399</td>
<td>Community-based</td>
<td>Health worker</td>
<td>Colposcopy/ histology</td>
<td>(H) 7 (L) 25</td>
<td>32</td>
<td>93</td>
<td>Verification bias</td>
</tr>
</tbody>
</table>

*a* Some test characteristics of the table are not exactly those reported in corresponding publications. Estimates have been computed when they were not provided or have been corrected to achieve comparability between studies. This correction was performed to take into account differences in study design or analysis, due to various factors: different threshold of test positivity, different disease definition, only a subset of the population used for estimation of characteristics, improper computation method, etc.

*b* Detection of any lesions

*c* Detection of cancer

Test positivity was defined at high-threshold (H) and at low-threshold (L), the sensitivity was estimated with the threshold CIN2–3 unless otherwise specified.

The study with no verification bias is highlighted.
contains little or no glycogen and does not stain with Lugol's iodine, taking a bright mustard or saffron yellow colour. Atrophic epithelium stains partially with Lugol's iodine, which makes interpretation difficult in postmenopausal women. A condylomatous lesion may not stain or only partially stain with Lugol's iodine. Areas of leukoplakia (hyperkeratosis) and areas partially denuded of squamous epithelium do not stain with iodine and remain colourless in a surrounding black or dark brown background. Results of VILI are categorized in Table 30. Images of VILI-positive cervices are shown in Figure 43.

The single published report of VILI test characteristics (Sankaranarayanan et al., 2004a) involved 54 981 women aged 25–65 years. The reference standard was colposcopy-directed biopsy. VIA and VILI were performed independently by blinded individuals in order to minimize information (expectation) bias. In this setting, VILI was more sensitive than VIA and equally specific (Figure 44). The reproducibility of

<table>
<thead>
<tr>
<th>Table 25. Test definition of visual inspection with acetic acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denomination</td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>Normal</td>
</tr>
<tr>
<td>Indeterminate</td>
</tr>
<tr>
<td>Ill-definite lesion</td>
</tr>
<tr>
<td>Definite lesion</td>
</tr>
<tr>
<td>Suspicious cancer</td>
</tr>
</tbody>
</table>

Figure 42 Example of VIA-positive lesions
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample size</th>
<th>Population (age, recruitment, location)</th>
<th>Provider</th>
<th>Reference diagnosis</th>
<th>Positivity rate (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slawson et al. (1992)</td>
<td>2690</td>
<td>15–45 Family practice USA</td>
<td>Clinician</td>
<td>Colposcopy/histology</td>
<td>3</td>
<td>29</td>
<td>97</td>
<td>Not designed for test accuracy estimation</td>
</tr>
<tr>
<td>Cecchini et al. (1993)</td>
<td>2036</td>
<td>17–83 Opportunistic Italy</td>
<td>Smear-taker</td>
<td>Colposcopy/histology</td>
<td>25</td>
<td>88</td>
<td>75</td>
<td>Not designed for test accuracy estimation</td>
</tr>
<tr>
<td>Megevand et al. (1996)</td>
<td>2426</td>
<td>20–83 Opportunistic South Africa</td>
<td>Nurse</td>
<td>Colposcopy/histology</td>
<td>3</td>
<td>65</td>
<td>98</td>
<td>Not designed for test accuracy estimation</td>
</tr>
<tr>
<td>Londhe et al. (1997)</td>
<td>372</td>
<td>Opportunistic India</td>
<td>Clinician</td>
<td>Colposcopy</td>
<td>53</td>
<td>78</td>
<td>49</td>
<td>Only 74% of the enrolled patient underwent colposcopy</td>
</tr>
<tr>
<td>Sankaranarayanan et al. (1998)</td>
<td>2935</td>
<td>20+ Opportunistic India</td>
<td>Cytotech.</td>
<td>Colposcopy/histology</td>
<td>10</td>
<td>87</td>
<td>91</td>
<td>Not designed for test accuracy estimation</td>
</tr>
<tr>
<td>University of Zimbabwe/ JHPIEGO (1999)</td>
<td>2148</td>
<td>25–55 Opportunistic Zimbabwe</td>
<td>Midwife</td>
<td>Colposcopy/histology</td>
<td>40</td>
<td>77</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Sankaranarayanan et al. (1999)</td>
<td>1268</td>
<td>22–70 Opportunistic India</td>
<td>Nurse</td>
<td>Colposcopy/histology</td>
<td>36</td>
<td>95</td>
<td>68</td>
<td>Not designed for test accuracy estimation</td>
</tr>
<tr>
<td>Denny et al. (2000)</td>
<td>2944</td>
<td>Opportunistic South Africa</td>
<td>Nurse</td>
<td>Colposcopy/histology</td>
<td>18</td>
<td>67</td>
<td>83</td>
<td>Minimal verification bias</td>
</tr>
<tr>
<td>Cronjé et al. (2001)</td>
<td>6298</td>
<td>Mean age: 34 Opportunistic South Africa</td>
<td>Nurse</td>
<td>Histology for VIA+ and 20% of VIA-</td>
<td>18</td>
<td>50</td>
<td>84</td>
<td>Unbiased estimation not possible from the published data</td>
</tr>
<tr>
<td>Study</td>
<td>Sample size</td>
<td>Population (age, recruitment, location)</td>
<td>Provider</td>
<td>Reference diagnosis</td>
<td>Positivity rate (%)</td>
<td>Sensitivity (%)</td>
<td>Specificity (%)</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>Belinson et al. (2001)</td>
<td>1997</td>
<td>35–45 Opportunistic China</td>
<td>Gynaecol.</td>
<td>4 quadrant histology + ECC</td>
<td>28</td>
<td>71</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Singh et al. (2001)</td>
<td>402</td>
<td>Mean age: 37 Referred Symptomatic India</td>
<td>Gynaecol.</td>
<td>Colposcopy/histology</td>
<td>42</td>
<td>87</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Denny et al. (2002)</td>
<td>2754</td>
<td>35–65 Opportunistic South Africa</td>
<td>Nurse</td>
<td>Colposcopy/histology for Pap+, HPV-, VIA+ or cervicography+</td>
<td>25</td>
<td>70</td>
<td>79</td>
<td>Minimal verification bias</td>
</tr>
<tr>
<td>Rodriguez-Reyes et al. (2002)</td>
<td>376</td>
<td>19–45 Opportunistic Mexico</td>
<td>Histology</td>
<td></td>
<td>48</td>
<td>92</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Cronjé et al. (2003)</td>
<td>1093</td>
<td>21–65 Opportunistic South Africa</td>
<td>Nurse</td>
<td>Histology</td>
<td>53</td>
<td>79</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>Ngelangel et al. (2003)</td>
<td>3316</td>
<td>25–65 Opportunistic Philippine</td>
<td>Nurse</td>
<td>Colposcopy/histology</td>
<td>10</td>
<td>37</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>Tayyeb et al. (2003)</td>
<td>501</td>
<td>30–60 Opportunistic Pakistan</td>
<td>Colposcopy/histology for Pap+ or VIA+</td>
<td>31</td>
<td>94</td>
<td>78</td>
<td>Verification bias</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Some test characteristics of the table are not those reported in the corresponding publications. Estimates have been computed when they were not provided or have been corrected to achieve comparability between studies. This correction was performed to take into account differences in study design or analysis, due to various factors: different threshold of test positivity, different disease definition, only a subset of the population used for estimation of characteristics, improper computation method, etc.

\(^b\) R stands for the range within the studies reported

Outcome threshold CIN 2–3
The studies with no verification bias are highlighted
### Table 27. Comparison of VIA and cytology accuracy in published studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample size</th>
<th>Positivity rate (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>VIA</td>
<td>Cytology</td>
<td>VIA</td>
</tr>
<tr>
<td>Slawson <em>et al.</em> (1992)</td>
<td>2690</td>
<td>3</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td>Cecchini <em>et al.</em> (1993)</td>
<td>2036</td>
<td>25</td>
<td>4</td>
<td>88</td>
</tr>
<tr>
<td>Megevand <em>et al.</em> (1996)</td>
<td>2426</td>
<td>3</td>
<td>13</td>
<td>65</td>
</tr>
<tr>
<td>Londhe <em>et al.</em> (1997)</td>
<td>372</td>
<td>53</td>
<td>6</td>
<td>78</td>
</tr>
<tr>
<td>Sankaranarayanan <em>et al.</em> (1998)</td>
<td>2935</td>
<td>10</td>
<td>10</td>
<td>87</td>
</tr>
<tr>
<td>University of Zimbabwe/JHPIEGO (1999)#</td>
<td>2148</td>
<td>40</td>
<td>13</td>
<td>77</td>
</tr>
<tr>
<td>Sankaranarayanan <em>et al.</em> (1999)</td>
<td>1268</td>
<td>36</td>
<td>16</td>
<td>95</td>
</tr>
<tr>
<td>Denny <em>et al.</em> (2000)</td>
<td>2944</td>
<td>18</td>
<td>15</td>
<td>67</td>
</tr>
<tr>
<td>Cronjé <em>et al.</em> (2001)#</td>
<td>6298</td>
<td>18</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>Singh <em>et al.</em> (2001)</td>
<td>402</td>
<td>42</td>
<td>42</td>
<td>87</td>
</tr>
<tr>
<td>Denny <em>et al.</em> (2002)#</td>
<td>2754</td>
<td>25</td>
<td>70</td>
<td>57</td>
</tr>
<tr>
<td>Ngelangel <em>et al.</em> (2003)</td>
<td>3316 (VIA)</td>
<td>10</td>
<td>2</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>3195 (Cytology)</td>
<td>10</td>
<td>2</td>
<td>37</td>
</tr>
<tr>
<td>Tayyeb <em>et al.</em> (2003)</td>
<td>501</td>
<td>31</td>
<td>16</td>
<td>94</td>
</tr>
<tr>
<td>Cronje <em>et al.</em> (2003)#</td>
<td>1093</td>
<td>53</td>
<td>9</td>
<td>79</td>
</tr>
<tr>
<td>Sankaranarayanan *et al.*b (2004b)</td>
<td>22 663</td>
<td>17</td>
<td>9</td>
<td>72</td>
</tr>
</tbody>
</table>

*a* Some test characteristics of the table are not the ones reported in corresponding publications. Estimates have been computed when they were not provided or have been corrected to achieve comparability between studies. This correction was performed to take into account differences in study design or analysis, due to various factors: different threshold of test positivity, different disease definition, only a subset of the population used for estimation of characteristics, improper computation method, etc.

*b* Subset of five Indian studies from Sankaranarayanan *et al.*(2004a)

*c* R stands for the range within the studies reported

Cytology threshold: ASCUS+, unless otherwise indicated (# , LSIL+)

Outcome threshold: CIN 2–3

---

**Figure 43** Example of VILI-positive lesions
VILI appears to be greater than that of VIA.

**Quality control for visual inspection tests**

The substantial variability in test performance characteristics of visual inspection tests reflects, at least in part, the subjective nature of visual inspection. Definitions of test result categories should be standardized to improve reproducibility. Due to the subjective nature of visual inspection methods, it is difficult to maintain the quality of assessment among trained staff. Adequate training, routine process measurements (e.g., test positivity rates, histological confirmation rates) and on-site supervision are critical to support high-quality visual inspection-based screening services. Although reliable methods of correlating daily competence with proficiency testing results have not been

---

### Table 28. Comparison of VIA and HPV testing accuracy in published studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample size</th>
<th>VIA</th>
<th>HPV testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positivity rate (%)</td>
<td>Sensitivity (%)</td>
</tr>
<tr>
<td>Denny et al. (2000)</td>
<td>2944</td>
<td>18</td>
<td>67</td>
</tr>
<tr>
<td>Sankaranarayanan et al. (2004c)a</td>
<td>18 085</td>
<td>11</td>
<td>[Rb: 54–79]</td>
</tr>
</tbody>
</table>

Outcome threshold: CIN 2–3

a Subset of five Indian studies from Sankaranarayanan et al. (2004a)

b R stands for the range within the studies reported

---

### Table 29. Comparison of accuracy of visual inspection with acetic acid, with or without magnification (VIA and VIAM), in published studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample size</th>
<th>VIA</th>
<th>VIAM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positivity rate (%)</td>
<td>Sensitivity (%)</td>
</tr>
<tr>
<td>Denny et al. (2000)</td>
<td>2944</td>
<td>18</td>
<td>67</td>
</tr>
<tr>
<td>Denny et al. (2002)</td>
<td>2754</td>
<td>25</td>
<td>70</td>
</tr>
<tr>
<td>Ngelangel et al. (2003)</td>
<td>3316 (VIA)</td>
<td>10</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>3447 (VIAM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sankaranarayanan et al. (2004d)</td>
<td>16 900</td>
<td>14</td>
<td>[Ra: 11–19]</td>
</tr>
</tbody>
</table>

Outcome threshold: CIN 2–3

a R stands for the range within the studies reported
established for cytological screening programmes (Vooijs et al., 1998), visual inspection-based screening programmes may utilize periodic assessments of practitioners’ skills using collections of VIA and VILI cervical photographs.

There is no consensus on the number of visual inspections that should be performed correctly by an individual before he/she is deemed competent, nor on the minimum daily rate that is required to maintain skills. The periodic computation of test positivity rates for visual inspections performed by each individual may be useful for monitoring visual inspection. However, because there is no permanent record when visual assessments are made by VIA or VILI, unless a photographic image is taken for subsequent review by a supervisor (Sellors et al., 2002; Wright, 2003), these positivity rates are not verifiable by audit. Visual inspection

---

**Table 30. Categorization of visual inspection with Lugol’s iodine (VILI) test results**

<table>
<thead>
<tr>
<th>VILI-negative</th>
<th>Patterns include the following:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• A normal pattern of dark brown or black staining of the squamous epithelium and no change in colour of the columnar epithelium, or</td>
</tr>
<tr>
<td></td>
<td>• patchy, indistinct, ill-defined, colourless or partially brown areas, or</td>
</tr>
<tr>
<td></td>
<td>• pale areas of no or partial iodine uptake present on polyps, or</td>
</tr>
<tr>
<td></td>
<td>• a leopard-skin appearance (associated with <em>T. vaginalis</em> infection), or</td>
</tr>
<tr>
<td></td>
<td>• pepper-like, non-iodine uptake areas seen in the squamous epithelium, far away from the squamocolumnar junction, or satellite, thin, yellow, non-iodine uptake areas with angular, or digitating margins, resembling geographical areas seen far away from the squamocolumnar junction</td>
</tr>
</tbody>
</table>

| VILI-positive | Presence of a dense, thick, bright, mustard-yellow or saffron-yellow iodine non-uptake area seen in the transformation zone close to or abutting the squamocolumnar junction or, in the absence of a visible squamocolumnar junction, close to the os, or when the entire cervix turns bright yellow |

| VILI-positive, invasive cancer | Frank, nodular, irregular, ulceroproliferative growth visible on the cervix, which turns densely yellow on application of iodine |

Adapted from Sankaranarayanan & Wesley (2003)

---

**Figure 44** Plot of sensitivity and specificity for VIA and VILI for each of the studies. The size of the bubble reflects the precision of the estimates. The bigger the bubble, the lower the variance of the sensitivity and specificity and the higher the precision due to a larger study sample size. The bubbles with thick borders represent the pooled estimates. (Adapted from Sankaranarayanan et al., 2004a)

Note: bubble size = \( k/(d_{se}^2 + d_{sp}^2) \), where \( k \) is a constant, \( d \) is the difference between the lower and upper confidence limits, \( se \) is sensitivity and \( sp \) is specificity.


methods coupled with immediate cryosurgical treatment for test-positive cases ('see and treat') have also been suggested for cervical screening in low-resource settings (Denny et al., 2002, Gaffikin et al., 2003). Because biopsies to exclude invasive carcinoma are not performed before ablative treatment, the 'see and treat' approaches carry the risk of undertreatment of invasive carcinoma and reduced opportunities to diagnose potentially curable invasive disease. It has been argued that, outside of research settings, it may not be feasible to monitor the safety or confirm the effectiveness of 'see and treat' programmes (Suba et al., 2004).

Advantages and potential hazards of visual inspection methods

Concerns about personal modesty and discomfort caused by the vaginal speculum are common to all screening techniques. Local irritation of tissues and allergic reactions to iodine or vinegar are rare. In both VIA and VILI, staining of the epithelium is temporary, although iodine staining lasts longer (up to 45 minutes) than acetowhiteness (Sankaranarayanan et al., 2004a).

Application of visual inspection methods should probably be restricted to women under the age of 50 years. With increasing age, the squamocolumnar junction migrates inward from the readily visible portion of the ectocervix towards the endocervical canal, so lesions probably become more difficult to identify with visual methods in older women. In addition, the accuracy of visual inspection may be highly dependent on the underlying prevalence of sexually transmitted diseases, which may increase the level of inflammation and render visual inspection difficult to assess.

Visual inspection-based tests are simple, safe and well accepted. They require a very low level of infrastructure and can be performed by a wide range of personnel, such as doctors, nurses, midwives, paramedical workers and trained non-medical personnel, after a short period of training (1–3 weeks). In addition, results are available without delay, allowing immediate referral for confirmatory testing. However, if immediate treatment is performed, high rates of overtreatment may result, given the relatively low test specificity of VIA and VILI.

Colposcopy

Colposcopy is a procedure that allows illuminated stereoscopic and magnified (typically 6–40 x) viewing of the cervix and the vagina. For colposcopy, the woman is placed in the lithotomy position, with the cervix exposed with a bivalve speculum in place, and various solutions (normal saline, 3–5% dilute acetic acid and Lugol's iodine) are applied to the cervical epithelium in sequence. A green filter is rarely used except when the subepithelial vascular pattern is examined (Jordan, 1985; Sellors & Sankaranarayanan, 2003). The aim is to examine the transformation zone, an area bounded laterally by the original squamocolumnar junction, in which metaplastic squamous epithelium develops, the medial or internal border being defined by the new squamocolumnar junction. This latter junction defines the upper limit of the squamous metaplastic process, which in certain conditions may become abnormal. When such abnormal areas develop within this zone, they are graded according to morphological features, namely, acetowhiteness, margins, blood vessels and iodine uptake.

Hinselmann (1925) first described colposcopy. The modern coloscope is a binocular microscope with a variable-intensity light source providing a stereoscopic view of the cervix, with a field of view and depth of focus that vary inversely with the magnification selected.

 Provision of a high-quality colposcopic service requires the availability of a trained colposcopist and access to a competent histopathologist able to perform assessment of removed biopsy material. The findings from video colposcopy seem to agree with those obtained with traditional optical colposcopy (Ferris et al., 2000b).

Colposcopic findings

Terminology to describe the morphological findings in a standard fashion has evolved over the years (Dexeus et al., 1977; Jordan, 1985; Sellors & Sankaranarayanan, 2003; Walker et al., 2003) (Table 31). Many of the qualitative descriptions have been quantified as to the degree of abnormality and have been combined into a scoring system (Table 32) that is used by many colposcopists to grade abnormal squamous epithelial areas (Reid & Scalzi, 1985; Reid, 1993). The uncommon glandular epithelial lesions tend to be more difficult to diagnose and appear as strikingly dense acetowhite or milky-white areas compared to the surrounding villi of columnar epithelium. Microinvasive and frankly invasive squamous cancers are densely acetowhite with markedly atypical blood vessels (bizarre, irregular branching and gross fluctuations in calibre and course). The surface configuration gradually changes from small protuberances, ex crescences or microconvolutions in microinvasive cancer to frankly invasive cancer with strikingly raised edges, irregular surface contour, and strikingly bizarre blood vessels that bleed spontaneously or on touch.

Histological confirmation

Biopsies are obtained under colposcopic visualization from the locations with the most severe changes, in order to histologically confirm the degree of severity of the neoplastic process. Since it is essential to rule out the presence of cancer, it is standard practice in some settings to obtain a
The colposcope can also be used to assess the remainder of the lower genital tract (vagina, vulva and perianal skin), especially if no cervical lesion is found in a woman with abnormal cytology. Women who are HIV-positive tend to have multifocal disease involving the vagina, vulva and perianal areas, and therefore these regions need to be examined (Abercrombie & Korn, 1998).

Since colposcopy was introduced in the 1960s to the English-speaking world, it has been selectively applied for diagnosis in women who are referred because of an abnormal cytological test. Current indications for colposcopy are listed in Table 33.

### Biases and caveats in the assessment of colposcopy

Most assessments of the sensitivity and specificity of colposcopy and directed biopsy are susceptible to bias. The colposcopic impression confounds the reference standard of diagnosis (histology) since it dictates where the histological specimen is obtained from, leading to an inflated estimate of the accuracy of colposcopy. In contrast to

### Table 31. International terminology for colposcopy from the International Federation for Cervical Pathology and Colposcopy

| I | Normal colposcopic findings |
|   | Original squamous epithelium |
|   | Columnar epithelium |
|   | Transformation zone |
| II | Abnormal colposcopic findings |
|   | Flat acetowhite epithelium |
|   | Dense acetowhite epitheliuma |
|   | Fine mosaica |
|   | Fine punctuation |
|   | Coarse punctuationa |
|   | Iodine partial positivity |
|   | Iodine negativitya |
|   | Atypical vesselsa |
| III | Colposcopic features suggestive of invasive cancer |
| IV | Unsatisfactory colposcopy |
|   | Squamocolumnar junction not visible |
|   | Severe inflammation, severe atrophy, trauma |
|   | Cervix not visible |
| V  | Miscellaneous findings |
|   | Condylomata |
|   | Keratosis |
|   | Erosion |
|   | Inflammation |
|   | Atrophy |
|   | Deciduosis |
|   | Polyps |

From Walker et al. (2003)

a Major changes

The rationale behind this combined testing approach is that it decreases false negative and false positive rates associated with cytology alone and also reduces the need for call-back for repeat cytology, the coloscope being used as a guide to collection of the cytology specimens (Van Niekerk et al., 1998). Within Germany at least, there is some reluctance to support the continued use of colposcopy as a screening tool to assist in the taking of cytological specimens, since there is no evidence that the quality of smears is improved (Hilgarth & Menton, 1996). Furthermore, constraints limiting the application of colposcopy to universal screening include its high cost relative to cytology, the availability and accessibility of adequately trained colposcopists, and the lower ability of colposcopy to detect endocervical lesions (Van Niekerk et al., 1998; Belinson et al., 2001).

Primary screening and diagnosis

Colposcopy continues to be used routinely as part of a standard gynecological examination by many clinicians in some European and Latin American countries, probably as a result of the long-standing tradition rooted in German medical teaching from the time of Hinselmann in Hamburg (Jordan, 1985; Dexeus et al., 2002). When colposcopy has been evaluated for primary screening, it has been usually accompanied by simultaneous cytology (Dexeus et al., 2002).
studies in which colposcopy is used for primary screening (with or without cytology), studies assessing colposcopy as a diagnostic procedure are conducted on women referred with abnormal screening cytology and having, therefore, a higher probability and possibly a more severe spectrum of cervical pathology. Since women with more pronounced findings and disease may be selected by screening, the performance of colposcopy in a diagnostic capacity may exceed its accuracy and reproducibility when it is used as a screening tool. If possible, all women evaluated with a test under assessment should have the reference standard applied to avoid verification bias and where this is not possible, statistical correction should be made. When colposcopic findings are compared with the pathological diagnosis, the colposcopist and the pathologist should be blind to corresponding information from the other test.

In relation to a large multidisciplinary study of precancerous lesions in China, Belinson et al. (2001) observed that increased use of technology alone does not guarantee that detection improves. Important factors are whether the quality of light used optimizes perception, the adequacy of training of the personnel, and the attributes of the population studied, such as prevalence of cervical inflammation. The definition of abnormality and certainty thresholds used by colposcopists in a study is important, since these determine the replicability

---

**Table 32. Combined colposcopic index, commonly used to score and document abnormal areas seen on colposcopic examination**

<table>
<thead>
<tr>
<th>Colposcopic sign</th>
<th>Zero points</th>
<th>One point</th>
<th>Two points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colour</td>
<td>Less intense acetowhite (not completely opaque); indistinct, semi-transparent acetowhiteing. Acetowhiteing beyond the margin of the transformation zone Snow-white colour with with intense surface shine</td>
<td>Intermediate, shiny, grey-white shade</td>
<td>Dull, oyster-white</td>
</tr>
<tr>
<td>Lesion margin and surface configuration</td>
<td>Feathery, indistinct, or finely scalloped edges Angular, irregularly shaped, geographic margins Satellite lesions with margins well removed from the new squamocolumnar junction Lesion with a condylomatous or micropapillary contour</td>
<td>Regularly-shaped lesion with sharp, straight edges</td>
<td>Rolled, peeled edges Internal margins separating lesions with differing scores, the more central one with the higher score tending to be nearest the new squamocolumnar junction</td>
</tr>
<tr>
<td>Blood vessels</td>
<td>Fine punctuation or mosaic pattern</td>
<td>Absent vessels (after application of acetic acid)</td>
<td>Coarse punctuation or mosaic pattern</td>
</tr>
<tr>
<td>Iodine staining</td>
<td>Positive iodine staining (mahogany-brown colour) Negative iodine staining in an area that scores 3 or less on the first 3 criteria</td>
<td>Partial iodine uptake giving a veriegated pattern</td>
<td>Negative for uptake giving a mustard yellow appearance in area that is significant (4 or more points) by the other 3 criteria</td>
</tr>
</tbody>
</table>

From Reid & Scalzi (1985); Reid (1993)
A score of 0–2 is compatible with CIN 1; 3–5 with CIN 1 or 2; and 6–8 with CIN 2 or 3.
of findings and the test cut-off for what are the minimal criteria for abnormality.

Studies of diagnostic colposcopy

Two meta-analyses have been performed on the accuracy of diagnostic colposcopy applied to women referred with abnormal cytology. Mitchell et al. (1998a) performed a systematic review of 86 articles published between 1960 and 1996, nine of which met the inclusion criteria and eight were eligible for meta-analysis. At the cut-off level of normal versus abnormal on colposcopy, the average weighted sensitivity, specificity and area under the receiver operating characteristic (ROC) curve of histological CIN 2 or more were 96%, 48% and 80%, respectively. At the cut-off level of normal and LSIL versus HSIL and cancer on colposcopy, the corresponding results were 85%, 69%, and 82%. This suggests that, independent of prevalence and compared with low-grade lesions, high-grade lesions and cancer are diagnosed with higher sensitivity. Olaniyan (2002) reviewed publications from 1966 to 2000 and the results of his meta-analysis, based on eight studies, seven of which were included in the previous meta-analysis, were similar.

A recent study of the diagnostic accuracy of colposcopy in China (Belinson et al., 2001) included methodological features intended to reduce selection bias and to assess the degree to which colposcopically directed biopsy is confounded with the colposcopic impression and the reference standard. In this study, vaginal and cervical specimens from 8,497 women (aged 27 to 56 years) were screened for 13 oncogenic types of HPV (Hybrid Capture 2 assay) and by liquid-based cytology (AutoCyte, TriPath, Burlington, NC) (Pretorius et al., 2004). Colposcopy was performed on 3,063 women who had an abnormality on screening and a directed biopsy was obtained from any abnormality. If colposcopy showed no lesion in a quadrant of the transformation zone, a biopsy was obtained in the original squamocolumnar junction in that quadrant. An ECC was then performed after biopsies had been obtained. Based on all of the women who had colposcopy (including 11 with unsatisfactory colposcopy), the sensitivity and specificity of colposcopy for detection of CIN 2 or worse lesions were 62.4% (234/375; 95% CI 57.3–67.3%) and 93.7% (7612/8122; 95% CI 93.2–94.2%), respectively (Pretorius et al., 2004).

Among the women with satisfactory colposcopy in the same study, directed biopsy detected 57.1% of high-grade lesions and cancers, while four-quadrant biopsy and ECC detected 37.4% and 5.5%, respectively. Among women referred for a cytological abnormality, directed biopsies were 4.8 times more likely to show a high-grade lesion or cancer than four-quadrant biopsies (26.5% versus 5.5%). The yields of CIN 2 or higher from four-quadrant biopsies for women referred because of HSIL, LSIL or ASCUS with a positive HPV test were 17.6%, 3.6% and 1.7%, respectively. One of 20 women in whom CIN 2 or worse was detected only by ECC had cancer despite satisfactory colposcopy.

A cohort study of 255 colposcopically negative women with abnormal cytology and 726 controls with normal cytology were followed for five years to assess the probability of false-negative colposcopy (Milne et al., 1999). Subsequent neoplasia was found in 19% versus 3% of controls ($p < 0.0001$).

Studies of screening colposcopy

In a cross-sectional study, 1997 unscreened Chinese women (aged 35–45 years) first were assessed by VIA performed by a gynaecologist and then a second gynaecologist (blinded to the VIA results) performed colposcopy with directed biopsies being taken from abnormal areas (Belinson et al., 2001). All women also had a biopsy taken from each of the four quadrants (and all had an ECC) in order to estimate the performance of colposcopy in a screening setting. Sensitivity and specificity of colposcopy and directed biopsy for high-
grade CIN or cancer were 81% (95% CI 72–89%) and 77% (95% CI 75–78%) compared with the combined histological findings from the directed, four-quadrant and ECC specimens.

A similar study in Germany enrolled 4761 women 18–70 years of age who were screened by conventional cytology (obtained under colposcopic vision), HPV testing of cervicovaginal samples by PCR and probing for 13 high-risk types and colposcopy when they visited one of ten gynaecologists for standard care (Schneider et al., 2000). Biopsy and EEC were performed where appropriate and if colposcopy was normal, biopsies at 6 and 12 o’clock and ECC were obtained. The sensitivity and specificity of screening colposcopy for detecting at least CIN 2, by histological confirmation, were 13.3% (95% CI 7.0–20.5) and 99.3% (95% CI 99.0–99.6), respectively.

Five studies of the simultaneous use of colposcopy and cytology to detect cervical cancer, performed more than 30 years ago, showed that the combined sensitivity of the two methods for cervical cancer varied from 95.0% to 99.4% (Dexeux et al., 1977). A recent case series from a German university using colposcopy and cytology for primary screening showed that the sensitivity of colposcopy for detecting at least CIN 2 was 90.8% (148/163) based on directed biopsy (Hilgarth & Menton, 1996). A similar study in the USA, based on 196 women who were screened opportunistically in a gynaecologist’s practice, gave estimated sensitivities of screening cytology, colposcopy and their combination of 48%, 76% and 91% (Davison & Marty, 1994). The estimated specificities were 100%, 96% and 96%, respectively.

**Validity of visual signs**

Reid and Scalzi (1985) published a scoring system which quantified the degrees of difference within certain morphological parameters (Table 32). These included reference to the colour of the cervical epithelial, blood vessel structure and the surface configuration of the epithelium of the transformation zone, as well as the degree of iodine staining. However, few major studies have studied the incorporation of this scoring system within a colposcopic management regime. One retrospective study of 134 women with biopsy-proven lesions using the modified Reid index score showed that it gave more accurate prediction of low-grade versus high-grade disease than when the 1976 International Nomenclature for Colposcopic Classification was employed (Carriero et al., 1991).

Prospective research on the predictive validity of visual signs in 425 women with abnormal cytology referred to a Canadian colposcopy clinic has shown that among three morphological characteristics routinely evaluated within the abnormal transformation zone (borders, degree of acetowhiteness, abnormal blood vessels), performance based on acetowhiteness was as good as all three signs combined (Shaw et al., 2003). A prospective study of 2112 women referred to the Cook County Hospital Dysplasia Clinic in Chicago did not use standardized grading criteria, but did show an association between histology and colposcopic impression \((p < 0.001)\), although agreement was poor (kappa, 0.20) (Massad & Collins, 2003).

The size of a lesion (categorized as the number of quadrants with positive histology) affects the sensitivity of colposcopy for detecting at least CIN 2 when the lesion size on referral cytology or histology is controlled (Pretorius et al., 2001). Colposcopy had a sensitivity of 65% (95% CI 47–79%) if the lesion involved only one quadrant of the cervical surface and 100% if more surface was involved (Belinson et al., 2001). Shafi et al. (1991) excised the entire transformation zone by loop electrosurgical excision procedure (LEEP) and confirmed an association between lesion area and histological grade. A study that estimated lesion size from cervigrams concluded that lesion size affects the sensitivity of cytology (Barton et al., 1989). Colposcopically inapparent high-grade lesions, remaining after a directed biopsy was taken, were evenly distributed among the four quadrants at 2, 4, 8, and 10 o’clock (Pretorius et al., 2004).

While most studies of colposcopically directed biopsy have shown less than perfect sensitivity for detecting the presence of a higher-grade lesion found on a subsequent LEEP specimen (Howe & Vincenti, 1991; Barker et al., 2001), the rate of underestimation among HIV-positive women may be substantially higher (Del Priore et al., 1996).

**Reproducibility of colposcopy**

Observer agreement studies of visual methods have been conducted using cervical photographs taken after the application of dilute acetic acid. Between three expert colposcopists, intra-observer and inter-observer agreement was poor to good when assessing border characteristics (range of inter-observer kappa, 0.13–0.41; of intra-observer kappa, 0.26–0.58) and the colour of acetowhiteness (range of inter-observer kappa, 0.21–0.47; of intra-observer kappa, 0.34–0.75). There was excellent agreement as to the site of the lesion from which a biopsy should have been obtained (raw agreement, 95.3%, 143/150) (Sellors et al., 1990). Ferris et al. (2000b) studied the inter-observer agreement within pairs of colposcopists using optical and video colposcopes and found that colposcopic impression agreement with histopathology (kappa, 0.60; 95% CI 0.53–0.68), biopsy intent agreement (79.9%) and biopsy site selection
agreement by quadrant (A, 78.3%; B, 81.3%; C, 85.3%; D, 82.7%) were not significantly different ($p \geq 0.3$), despite the use of different coloscopes. Similar agreement was obtained when telecolposcopy (using a video colposcope) was viewed by an expert colposcopist at a remote location and compared with the video colposcopy performed by an expert on-site. The kappa values for colposcopic impression and histopathology agreement varied between 0.16 and 0.31 ($p$ values not given) and for biopsy intent, kappa was 0.32 ($p = 0.002$) (Ferris et al. 2002).

Assuming that colposcopists use the same definitions, reproducibility of colposcopic assessment depends in part on colposcopists using similar ‘thresholds of certainty’ for categorizing findings as to normal versus abnormality and grade.

**Quality control**

Like other medical services, colposcopy services can be audited and compared with national standards, such as those established for the English National Health Service, for process and outcome (Ferris et al., 2002). Indicators recommended for periodic audit include waiting time for colposcopy by grade of referral smear; adequacy of communication between primary and secondary level; frequency of procedures; agreement between colposcopic diagnosis, referral cytology and histology; treatment method by histological diagnosis; efficacy of treatment (e.g., whether histological evidence of CIN is present in over 90% of women undergoing ‘see and treat’); and follow-up rates at one year (Luesley, 1996).

Cervical imaging using colpophotography, video colpography, and telecolposcopy has been studied. All methods give a true representation of what is seen at colposcopy and have been recommended for teaching and audit, as well as for patient care (Sellors et al., 1990; Etherington et al., 1997; Milne et al., 1999; Harper et al., 2000; Li et al., 2003). Harper et al. (2000) showed that a telescopy network that allows transmission and sharing of static colposcopic images for consultation and teaching purposes on a regular basis was technically feasible, acceptable to women and health care providers living in remote areas, and gave good inter-observer agreement between the on-site colposcopists and the off-site review colposcopists as to degree of abnormality (kappa = 0.68; 95% CI 0.54–0.82). Ferris et al. (2003) showed that network telecolposcopy using high-speed telecommunications lines and computer telecolposcopy using modems and telephone lines to transmit static images was superior to cervicography as measured by the number of confirmed CIN lesions detected and timeliness of results. On-site colposcopy had the highest sensitivity to detect CIN because of the stereoscopic vision, the ability to manipulate the cervix and view the acetowhite reaction as it occurs, and the ability to resolve vascular and epithelial features. Compared with telecolposcopy, the ability to assess whether a colposcopic examination is satisfactory appears to be better with on-site colposcopy (Sellors et al., 1990).

Documentation of colposcopic images and data using the latest digital photographic and information systems allows not only recording and comparison of colposcopic findings with subsequent examinations, but also the retention of data for audit, post-treatment follow-up and comparison of data between units.

**Potential side-effects of colposcopy**

A routine colposcopic examination involves some discomfort due to the insertion of the vaginal speculum and more when a tissue specimen is obtained by punch biopsy. Psychological morbidity should be appreciated and counselling considered (Howard et al., 2002a). Studies using measures of anxiety such as the State-Trait Anxiety Inventory have consistently shown that anxiety scores before colposcopy are markedly elevated to levels seen in patients awaiting surgery, and fall immediately after colposcopy is completed. The fears that women have before colposcopy relate to cancer, fertility, danger to partner, social stigma and pain or embarrassment during the procedure. Other women may have a significant level of anxiety about the examination because of a possible history of sexual abuse. Educational booklets and counselling are effective in reducing anxiety (Ferris et al., 2003). Colposcopy service providers need to be sensitive and responsive to women’s needs in order to provide an acceptable service and to optimize adherence to appointments.

**Cervicography**

Cervigrams are replicate photographs of the cervix taken after application of 5% acetic acid, using a camera with a fixed focal length and internal light source. The images are projected onto a screen at a fixed distance to simulate magnification and are interpreted by a trained evaluator.

It is now possible to achieve equal visual resolution with digital cameras, producing images that can be immediately downloaded and transmitted for expert review, the images being evaluated with computer-generated magnification as needed. Future efforts related to cervicography will depend on digital techniques capable of generating images as good as those using high-quality film, with the advantages of ‘telemedicine’-based screening and centralized image analysis (Wright, 2003).
Cervigrams are interpreted using the categories presented in Table 34. With these criteria, initial studies showed poor reproducibility because of differing distinction of the very subtle acetowhite lesions that represent either immature squamous metaplasia or HPV changes.

In a large inter-observer study among 3637 women, a comparison of dichotomous results (positive versus not) assigned by the initial versus the second evaluator yielded a kappa statistic of 0.5, indicating only moderate agreement beyond that expected by chance (Schneider et al., 2002).

The deficiencies and inconclusive results of several small-scale studies of cervicography were summarized by Nuovo et al. (1997). In later large-scale evaluations, summarized in Table 35, cervicography proved insufficiently accurate to serve as a stand-alone screening test. In summary, cervicography appeared somewhat less accurate than cytology, primarily because of inferior specificity. This was a result of the overcalling of acetowhite epithelial changes.

The percentage of technically inadequate cervigrams varies widely by study; satisfactory results depend on the experience of the evaluator (De Sutter et al., 1998). Adjudicated cervigram reviews and histological re-confirmation of CIN 2, CIN 3 or cancer did improve performance over a single interpretation, but suggested the upper limit of sensitivity (Schneider et al., 2002). The sensitivity and specificity depended on the cut-point of positivity and targeted disease end-point, but no choice of cut-point generated excellent overall accuracy of detection of CIN 3 and cancer.

A major limitation of cervicography (and possibly, by extension, other static visual techniques) is the poor sensitivity among older women, whose transformation zones are often beyond the field of vision (Schneider et al., 1999). Women aged 40–60 years would be expected to represent a sizeable proportion of women being screened in the low-resource settings where a non-cytological technique such as cervicography might be particularly helpful (Wright, 2003). However, in this group, the targeted precancerous lesions can be small enough to be easily missed. In one evaluation, the sensitivity of cervicography for detection of CIN 2, CIN 3 and cancer was only 30.0% among 2196 post-menopausal women, compared with 54.7% among 6264 pre-menopausal women, using a positive cut-point, and findings were similar using an atypical cut-point (see Table 34 for definitions) (Schneider et al., 1999).

The relative accuracy of direct visual inspection compared with distant, expert review of a static visual image is not clear. There appears to be a trade-off between colposcopic

### Table 34. Cervigram classificationa

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Not referred for colposcopy:</strong></td>
<td></td>
</tr>
<tr>
<td>Negative:</td>
<td>No definite lesion is visible</td>
</tr>
<tr>
<td>Atypical 1 (A1):</td>
<td>A lesion inside the transformation zone is visible; based on the lesion's site and morphology, the lesion is presently considered to be of doubtful significance</td>
</tr>
<tr>
<td>Atypical 2 (A2):</td>
<td>A lesion outside the transformation zone is visible; based on the lesion's site and morphology, the lesion is presently considered to be of doubtful significance</td>
</tr>
<tr>
<td>Technically defective:</td>
<td>The cervigram slide is not adequate for evaluation</td>
</tr>
<tr>
<td><strong>Referred for colposcopy:</strong></td>
<td></td>
</tr>
<tr>
<td>Positive (all categories below):</td>
<td>A lesion is visible and colposcopy is recommended because of the lesion's site and morphology, or no definite lesion is visible, but the appearance warrants colposcopy to exclude significant disease.</td>
</tr>
<tr>
<td>Positive 0 (P0):</td>
<td>Probably normal variant; appearance warrants colposcopy to exclude significant disease</td>
</tr>
<tr>
<td>Positive 1A (P1A):</td>
<td>A lesion extending into the canal, the visible portion of which is presently considered to be of doubtful significance</td>
</tr>
<tr>
<td>Positive 1B (P1B):</td>
<td>Compatible with a low-grade lesion</td>
</tr>
<tr>
<td>Positive 2 (P2):</td>
<td>Compatible with a high grade lesion</td>
</tr>
<tr>
<td>Positive 3 (P3):</td>
<td>Compatible with cancer</td>
</tr>
</tbody>
</table>

a As of 1 January 1995, National Testing Laboratories worldwide revised the atypical category. Previously, atypical 1 referred to trivial lesions outside the transformation zone and atypical 2 referred to trivial lesions inside the transformation zone. Modified from Schneider et al. (1999)
expertise and the loss in visual discrimination inherent in examining an image compared to real-time examination. In one cross-sectional screening study with limited statistical power due to small numbers of precancerous outcomes, cervicography was apparently more accurate than direct visual inspection by nurses, due to increased specificity (Rodriguez et al., 2004). However, in a statistically more powerful study, distant review by experts of digitized, static colposcopy images was significantly less sensitive (but more specific) than colposcopy performed by local gynaecologists and nurses with varied training (Ferris & Litaker, 2004).

Since it has been concluded, on the basis of accumulated data, that cervicography is inadequate as a stand-alone screening technique, research has shifted to evaluation of combining cervicography with

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Disease target threshold</th>
<th>Cervicography</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coibion et al. (1994)</td>
<td>Belgium, n = 4015</td>
<td>&gt; CIN 1, n = 123</td>
<td>Old Atyp.</td>
<td>86</td>
<td>99</td>
<td>76</td>
<td>99.0</td>
</tr>
<tr>
<td>Schneider et al. (1996)</td>
<td>Germany, n = 967</td>
<td>&gt; CIN 2, n = 38</td>
<td>Atyp. or Pos.</td>
<td>45</td>
<td>91</td>
<td>17</td>
<td>97.6</td>
</tr>
<tr>
<td>Baldauf et al. (1997)</td>
<td>France, n = 1351</td>
<td>&gt; CIN 1, n = 168</td>
<td>Positive</td>
<td>51</td>
<td>96</td>
<td>44</td>
<td>97.1</td>
</tr>
<tr>
<td>De Sutter et al. (1998)</td>
<td>Belgium, n = 5192</td>
<td>&gt; CIN 2, n = 33</td>
<td>Positive</td>
<td>55</td>
<td>97</td>
<td>11</td>
<td>99.7</td>
</tr>
<tr>
<td>Schneider et al. (1999)</td>
<td>Costa Rica, n = 8460b</td>
<td>&gt; CIN 2, n =136</td>
<td>Atyp. or Pos.</td>
<td>63</td>
<td>85</td>
<td>6</td>
<td>99.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Positive</td>
<td>49</td>
<td>95</td>
<td>14</td>
<td>99.1</td>
</tr>
<tr>
<td>Denny et al. (2000)</td>
<td>South Africa, n = 2611</td>
<td>&gt; CIN 2, n = 79</td>
<td>Positive</td>
<td>58</td>
<td>91</td>
<td>58</td>
<td>93.4</td>
</tr>
<tr>
<td>Costa et al. (2000)</td>
<td>Italy, n = 992</td>
<td>&gt; CIN 2, n = 90c</td>
<td>Atyp. or Pos.</td>
<td>76</td>
<td>91</td>
<td>51</td>
<td>97.4</td>
</tr>
<tr>
<td>Cronjé et al. (2001)</td>
<td>South Africa, n = 1747a</td>
<td>&gt; CIN 1, n = 342</td>
<td>Positive</td>
<td>42</td>
<td>79</td>
<td>32</td>
<td>84.8</td>
</tr>
<tr>
<td>Cronjé et al. (2003)</td>
<td>South Africa, n = 1093</td>
<td>&gt; CIN 2, n = 90</td>
<td>Positive (not P0)</td>
<td>49</td>
<td>88</td>
<td>26</td>
<td>95.0</td>
</tr>
<tr>
<td>Ferreccio et al. (2003)</td>
<td>Costa Rica, n = 8457</td>
<td>&gt; CIN 3, n = 110 including follow-up</td>
<td>Atyp. or Pos.</td>
<td>62</td>
<td>85</td>
<td>5</td>
<td>99.4</td>
</tr>
</tbody>
</table>

*a* Evaluation not performed according to National Testing Laboratory criteria.

*b* Population-based screening of a high-risk province, where attempts were made to vary cervicography cut-point and disease end-point to explore performance.

*c* Women with negative colposcopy presumed to be disease-negative.

*d* Analysis of subgroup of large group of screened women. Subgroup included those with biopsied acetowhite lesions, as well as 1/5 of women with seemingly normal cervix. Predictive values not adjusted for sampling.

Sensitivity and specificity are estimated cross-sectionally (see Chapter 4)
cytology or HPV for screening, and to a possible role for cervicography in the triage of women with equivocal cytology. These topics are considered below in the section on combined techniques.

**HPV DNA testing**

Research on the use of HPV DNA assays as a potential cervical cancer screening tool began in the late 1980s, as a reflection of the emerging evidence that these viruses played a causal role in the genesis of cervical neoplasia (zur Hausen, 1976; Deligeorgi-Politi et al., 1986). Although much of that research began with a focus on viral detection as an end in itself (reviewed by Schiffman, 1992), attention soon turned to the potential clinical utility of HPV testing for identifying cervical cancer precursors (Lörincz et al., 1990; Wilbur & Stoler, 1991; Lörincz, 1992). The basic assumption was that standardized molecular testing of exfoliated cervical cells for the putative causal agent of cervical cancer could have acceptable diagnostic performance, while being more reproducible and more easily adapted for automated, high-volume testing in clinical practice than conventional cytological testing. Concerns in the USA about the quality of smears processed in cytopathology laboratories added pressure to study the potential use of HPV testing as an adjunct to cytology (Reid et al., 1991; Reid & Lörincz, 1991), despite some opposing views (Nuovo & Nuovo, 1991; Beral & Day, 1992). More recently, cytology has been characterized not only as a sufficient screening test, but also as a likely necessary component of future screening programmes based on HPV or visual testing, due to the low relative specificity of non-cytological methods (Suba & Raab, 2004).

Techniques to detect the presence of HPV in cervical cell specimens have evolved considerably in the last 25 years, from (i) simple scoring of karyocytes (a type of cytopathic effect taken to indicate the presence of HPV in the host epithelial cells) in cervical smears (Komorowski & Clowry, 1976) to (ii) immunocytochemical staining (Syrränen & Pyrhonen, 1982); non-amplified nucleic acid hybridization methods, such as (iii) dot blot (Parkkinen et al., 1986), (iv) Southern blot (Okagaki et al., 1983) and (v) filter in-situ hybridization (Schneider et al., 1985); signal-amplified, immunoassay-based nucleic acid hybridization techniques such as (vi) the Hybrid Capture™ (HC) assay (Farthing et al., 1994); and (vii) a variety of type-specific (Dallas et al., 1989) and general or consensus-primer (Gregoire et al., 1989; Manos et al., 1989; Snijders et al., 1990; Roda Husman et al., 1995; Kleter et al., 1998; Gravitt et al., 2000) polymerase chain reaction (PCR) techniques. In addition, adaptations of the solution-based, non-amplified hybridization and PCR protocols have been used to detect HPV DNA in histological sections or smears, to allow confirmation of the presence of the virus in particular target cells. Such in situ techniques (Gupta et al., 1985; Nuovo et al., 1991) have been useful in molecular pathology studies, but have found little interest as potential screening tools for cervical cancer and its precursors. Serological assays to detect antibodies to HPV capsid or functional protein antigens have also received attention as investigational tools in epidemiological and clinical studies (Jochmus-Kudielka et al., 1989; Galloway, 1992). However, as with in situ assays, they have not been considered as candidate methods for screening cervical cancer precursors. Serology detects humoral immune response to HPV antigens, which may reflect cumulative exposure to HPV infection acquired in mucosal sites other than genital, and thus is not suitable, in principle, as a screening tool.

Early clinical studies used non-amplified DNA hybridization methods (without signal amplification) to gauge the screening utility of HPV testing to identify and manage cervical lesions. Such methods are no longer used, however, because of their insufficient sensitivity and specificity for epidemiological and clinical studies (Franco, 1992; Schiffman & Schatzkin, 1994). Only the commercially available HC assay and a few PCR protocols have been the focus of investigations conducted in the last 10 years. Although a number of biotechnology companies are currently developing HPV DNA diagnostic systems for clinical use, few are yet available commercially or have reached the stage of large-scale clinical studies. For this reason, this overview focuses primarily on the HC assay and on the more popular PCR protocols that have been used in screening studies.

**Hybrid Capture™ assay**

Most clinical investigations of HPV testing have used first- or second-generation Hybrid Capture™ (HC) systems (Digene, Inc., Gaithersburg, MD), the only HPV test currently approved by the US FDA. The HC system is a nucleic acid hybridization assay with signal amplification for the qualitative detection of DNA of high-risk, cancer-associated HPV types in cervical specimens. It cannot determine the specific HPV type present, since detection is performed with a combined probe mix. The first HC assay (HC1) was a tube-based detection system and probed for only nine of the high-risk HPV types: 16, 18, 31, 33, 35, 45, 51, 52 and 56. The second-generation HC system (HC2) has improved reagents and is based on a microplate assay lay-out that targets 13 high-risk HPV types: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56,
58, 59 and 68. A probe set for a few non-oncogenic HPV types (6, 11, 42, 43, 44) has been available for both the HC1 and HC2 assays but its utility has not been sufficiently investigated in clinical or epidemiological studies. It is often designated as probe A, whereas probes for high-risk HPV types are referred to as probe B.

HC2 is an entire system that can be used with a dedicated cervical sampler kit containing a special cervical conical brush and a vial with specimen transport medium (STM). The brush is designed for optimal collection of cells from both the ectocervix and endocervix. The brush is shaped as a cone (Christmas-tree-like) that fits the cervical canal and samples the endo- and ecto-cervix. This brush is inserted gently into the cervical canal and fully rotated three times. It is then retrieved without touching the vaginal wall and inserted into the collection tube containing STM. The tip is broken and the tube is closed. According to the manufacturer, specimens in STM can be held at room temperature for up to two weeks and can be stored for an additional week at 4°C. If not tested in the first three weeks after collection, they can be stored at –20°C for up to three months.

HC2 is a solution hybridization assay that uses long synthetic RNA probes that are complementary to the DNA sequence of the 13 high-risk HPV types (or to the probe A types) listed above. The initial reaction step denatures the exfoliated cells in STM, thus releasing host and any existing HPV DNA molecules to the solution. HPV DNA molecules then bind (i.e., hybridize) with the respective RNA probe, resulting in the formation of DNA–RNA hybrids reflecting the composition of HPV types present in the mixture. This hybridization step occurs in solution inside the wells of a specially treated 96-well plastic microtiter plate previously coated with polyclonal IgG antibodies that are specific for RNA–DNA hybrids, regardless of sequence homology. Any such hybrids will then be captured by the solid-phase-bound antibodies, hence the name 'hybrid capture'. After washing steps to remove unbound molecules, a solution of a conjugate reagent consisting of the same anti-RNA–DNA hybrid antibody covalently linked with the enzyme alkaline phosphatase is added to the wells. Conjugate antibody molecules will then bind to any solid-phase-bound hybrids. After further washing to remove unbound molecules, a solution containing a chemiluminescent diode-tane substrate is added to the wells. Cleavage of the substrate by the alkaline phosphatase releases a luminescent reaction product into the solution. The intensity of the light emitted is proportional to the amount of HPV DNA originally present in the specimen and is measured in a luminometer provided with the system. The reaction signal of each specimen is expressed on a scale (relative light units or RLU) relative to the average reactivity measured in triplicate wells with a positive control containing 1.0 pg of HPV16 DNA per ml. Specimens yielding RLUs greater than or equal to 1.0 are considered positive; some studies have assessed the validity of this cut-point using ROC curve analysis (Schiffman et al., 2000). In most clinical settings, the manufacturer (Digene) certifies the laboratory that intends to perform HC2 testing, thus ensuring quality control.

Because the RLU signal is proportional to the amount of HPV DNA present in the specimen, the HC2 assay has occasionally been used to infer viral load, on a semi-quantitative basis (Clavel et al., 1998; Sun et al., 2001; Cuzick et al., 2003). The assay is easy to perform in clinical practice and amenable to automation, which makes it attractive for high-volume screening use. To this end, a robotic assay workstation named Rapid Capture System™ (Digene) is available, which performs specimen transfer, all pipetting operations, incubations, shakings and washings. However, the denaturation of specimens in the sample device tubes still has to be performed by hand. This automatic station increases the accuracy of the test and allows a single user to test 352 specimens within four hours.

Since it is based on signal, rather than target amplification (as in the case of PCR protocols), HC2 is less prone to cross-specimen contamination, thus obviating the need for special laboratory facilities to avoid cross-contamination (Coutlee et al., 1997). In practice, only the high-risk probe mix (probe B) is used for cervical lesion screening, which reduces the time and cost to perform the test. At the standard FDA-approved cut-off of 1 pg/ml (RLU ≥ 1.0) and even at higher discriminant levels, there is cross-reactivity between certain HPV types not present as targets in the probe B set (e.g., 53, 66, 67, 73) and the RNA probes used in that set (Peyton et al., 2002b, 2004). Cross-reactivity with non-cancer-causing types would have an adverse impact on test specificity in settings with high prevalence of the low-risk types. On the other hand, cross-reactivity with other high-risk types not represented in the probe B set may be beneficial for test sensitivity (Castle et al., 2003).

**Polymerase chain reaction**

PCR is based on the repetitive replication of a target sequence of DNA flanked at each end by a pair of specific oligonucleotide primers, which initiate the polymerase-catalysed chain reaction. Because of the exponential increase in the amount of target DNA sequence after a few reaction cycles of denaturation, annealing and extension, PCR has very high levels of molecular
sensitivity and permits the detection of less than 10 copies of HPV DNA in a mixture. Therefore, PCR has a lower threshold of molecular detection for HPV DNA than the HC assay. PCR is based on target amplification with type-specific or consensus or general primers. The latter are able to amplify sequences from several different HPV types because they target conserved DNA regions in the HPV genome. The amplified DNA products can be revealed by ethidium bromide staining following agarose or acrylamide gel electrophoresis, which permits presumptive verification of the expected molecular weight of the amplified target, thus confirming positivity. Verification can also be done by methods that further probe the post-amplification products for their sequence homology with the target. Dot blot, Southern blot or line strip hybridization are used to this end and generally result in improved molecular sensitivity and specificity as compared with electrophoresis and staining (Gravitt & Manos, 1992; Gravitt et al., 1998).

Finally, use of restriction enzymes to analyse the fragment length signatures in combination with probe hybridization (Bernard et al., 1994) and direct DNA sequencing provide the highest possible resolution to distinguish the HPV types present in a biological specimen.

The very high sensitivity of PCR is its very limiting factor in terms of clinical applicability. Molecular threshold does not correlate directly with clinical sensitivity and specificity (Snijders et al., 2003). Because millions of copies of the DNA target can be produced from a single molecule, there is a high probability of contamination of other specimens and control samples with HPV sequences in airborne droplets and aerosolized reaction mixtures. In fact, cross-contamination was a major problem in some early applications of PCR in HPV testing. Extreme care is needed in PCR testing laboratories. Several procedures are well established to minimize the potential for contamination, the most important of which is the separation of pre-amplification and post-amplification areas.

Judicious analysis of sequence homology among different genes of distinct HPV types using software that aligns DNA sequences will reveal countless segments that could serve as candidates for PCR primer design. In fact, many type-specific and consensus HPV testing PCR protocols have been published in the last 15 years. However, because of the requirements for validation, reproducibility, and general acceptability, relatively few have become established to the point of being widely used in clinical and epidemiological studies. Primer systems targeting sequences in the L1, E1, E6 and E7 genes have been most commonly used. Because of their well conserved sequences, L1 and E1 have been targeted by the consensus primer protocols. E6 and E7, on the other hand, have more sequence variation among HPV genotypes, making them less suitable as targets for amplification of a broad spectrum of HPV types (Gravitt & Manos, 1992).

The most widely used PCR protocols are of the consensus or general primer (degenerate or non-degenerate) type, i.e., they can potentially amplify sequences of multiple HPV types with one primer set in one reaction pass. The size of the amplified product is the same irrespective of the HPV type present in the starting mixture, and thus electrophoresis cannot reveal the actual type present in the sample. The post-amplification hybridization or sequencing techniques described above must be used to identify the HPV type or types originally present. Three consensus primer systems (and their technical variations) based on L1 sequence detection have become well established. They can detect essentially all types of HPV that infect the mucosal areas of the lower genital and upper aerodigestive tracts. Two of these, the MY09/11 (Manos et al., 1989) and the GP5/6 (Snijders et al., 1990; Van den Brule et al., 1990) systems have evolved into variants with better primer composition and internal oligonucleotide probing, such as the PGMY09/11 (Gravitt et al., 1998, 2000) and the GP5+/6+ (Roda Husman et al., 1995; Jacobs et al., 1995, 1997; Van den Brule et al., 2002) protocols. Over the years, the original radioactively labelled hybridization probes have gradually been abandoned in favour of biotinylated probes and enzyme immunoassay formats. The third protocol is designated SPF10 LiPA, for line probe assay based on the SPF10 primer set (Kleter et al., 1998; Quint et al., 2001). Although these three consensus protocols amplify targets within the L1 gene of HPV, they do so for segments of considerably different sizes: 450 base pairs (bp) for MY09/11, 140 bp for GP5/6, and 65 bp for SPF10 LiPA. The size of the amplified product is not a trivial matter. Although discrimination of sequence homology is better for longer gene segments and thus would in theory permit improved HPV type resolution, shorter fragments tend to yield better sensitivity with severely degraded specimens, such as paraffin-embedded, archival tumour tissue. Damage is often pronounced in DNA extracted from such archival specimens, resulting in DNA fragments of less than 200 bp. In these circumstances, a protocol targeting a short fragment, such as GP5+/6+ or SPF10 LiPA, tends to yield fewer false negative results (Gravitt & Manos, 1992).

The newly developed Roche prototype Microwell plate assay (Roche MWP) employs an oligonucleotide set which amplifies a short fragment of the L1 gene of high-risk HPV types (170 bp, compared with 450 bp with
PGMY09/11). This amplicon is immobilized using a pool of capture molecules bound to the wells of a microtitre well plate and visualized by colorimetric detection. The new test was developed to employ the TaqGold DNA polymerase, which minimizes the amount of non-specific amplification and increases the sensitivity of the test. Since it amplifies a shorter fragment, it is considered to be more sensitive than PGMY09/11 PCR and also suitable for less well preserved specimens; it has been reported that these primers detect about 13% more HPV in cervical smears than the PGMY primers (Ifnner & Villa, 2003). However, because these primers were designed for high-risk types only (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68), this test is not truly generic, but rather comparable to the HC2 test. In addition, the use of PCR assays aiming at maximum sensitivity for detection of HPV in a screening setting irrespective of concomitant disease may be inappropriate with regard to clinical usefulness.

The reproducibility and agreement of HPV testing results among the three most popular PCR protocols, as well as between them and the FDA-approved HC2 assay, for overall HPV detection have been extensively studied. While agreement at the overall positivity level may be considered adequate in clinical settings, concordance at the level of type detection leaves much to be desired (Qu et al., 1997; Kletter et al., 1998, 1999; Peyton et al., 1998; Swan et al., 1999; Gravitt et al., 2000; Castle et al., 2002c; Van Doorn et al., 2002; Castle et al., 2003).

Biomedlab Co. (Republic of Korea) has developed a HPV oligonucleotide microarray-based system for detection of HPV types that currently allows detection of 22 HPV types, by immobilizing HPV type-specific oligonucleotide probes and a control (beta-globin probe) on an aldehyde-derivatized glass slide. Target DNA is subjected to a standard PCR in the presence of fluoresceinated nucleotides (labelled with Cy5 or Cy3) employing primers for both the beta-globin (PC03/04) and for the L1 region (modified GP5/6 primers) of several HPV types. Randomly labelled PCR products are then hybridized to specific oligonucleotides on the chip, which is afterwards scanned by laser fluorescence. In the case of multiple infections, multiple hybridization signals can be seen. Because signal detection in microarrays is subject to variation, additional levels of control would be desirable. These should include quality control of the efficiency of the PCR reaction and the hybridization conditions, include a measurement of the homogeneity of the probes on the chips and allow some sort of quantification. In addition, the read-out requires expensive equipment for signal detection and would need to be performed with the help of special software that allows threshold settings.

**Sensitivity and specificity of HPV assays**

Dozens of studies have provided data on the diagnostic performance of HPV DNA testing methods. However, only some of them provided direct comparisons with cytological testing in detecting high-grade precancerous lesions and clearly specified the type of population, i.e., whether it was a primary screening or secondary triage study. The vast majority of studies either did not provide data on cytology or presented data on mixed series of subjects and could not be unequivocally designated as screening or triage settings. Table 36 summarizes the main features of selected published studies that provided data on the comparison of HPV testing with cytology in primary screening for cervical cancer and its precursors; it also gives estimates of sensitivity and specificity for HPV testing and for cytology in the same studies. Data on the performance of HPV testing in triage studies are presented later in this chapter. For all studies, specificity estimates are based on women free of histologically demonstrable squamous lesions.

The studies vary considerably in terms of investigational design, choice of population and methods, which, as expected, leads to enormous variability in the results observed. Most studies assessed HPV test performance on the basis of prevalent lesions using simple cross-sectional designs or retrospective case series, whereas some assessed both prevalent and seemingly short-term incident lesions based on cross-sectional investigations with extended follow-up (ASCUS-LSIL Triage Study (ALTS) Group, 2003a, b). Lesion definition varied across studies and included either CIN 1 or CIN 2/3 or worse lesions, diagnosed by histology on specimens obtained by colposcopy-guided biopsy. Sometimes the SIL terminolgy was used for these histologically diagnosable squamous lesions.

In some studies, the colposcopic result was used if no biopsy was taken. Some studies used direct community recruitment, but usually the study population was clinic-based. None of the studies was based on long-term follow-up for more relevant end-points, such as incidence of CIN 2 or 3 or cancer or mortality from invasive cervical cancer (see Chapter 4).

For many of the studies, the purpose was to compare HPV testing with other screening technologies for cervical cancer (primarily cytology). None of the investigations was a randomized controlled trial; all were based on concomitant testing for HPV and cytology alone or with additional tests. Such investigations are known as split-sample studies because the cervical specimen, collected in single or multiple exfoliative procedures using a swab, a cytobrush or other collection device, is split into several sub-
### Table 36. Selected studies that provided data on the comparison of HPV testing with cytological testing in primary screening for cervical cancer and its precursor lesions: characteristics and estimates of screening performance

<table>
<thead>
<tr>
<th>Study, country</th>
<th>Study size</th>
<th>Age (years)</th>
<th>HPV test</th>
<th>Study features</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cuzick <em>et al.</em> (1995), UK</td>
<td>2009</td>
<td>20–45</td>
<td>Type-specific PCR (16, 18, 31, 33)</td>
<td>Women free of cytological abnormalities at enrolment</td>
<td>75</td>
<td>46</td>
<td>96 98</td>
</tr>
<tr>
<td>Cuzick <em>et al.</em> (1999), UK</td>
<td>2988</td>
<td>34+</td>
<td>HC1, HC2, MYO9/11 PCR</td>
<td>Women free of cytological abnormalities at enrolment</td>
<td>95</td>
<td>79</td>
<td>95 99</td>
</tr>
<tr>
<td>Kuhn <em>et al.</em> (2000), South Africa</td>
<td>2944</td>
<td>35–65</td>
<td>HC1, HC2</td>
<td>Unscreened population, community recruitment</td>
<td>88</td>
<td>78</td>
<td>82 97</td>
</tr>
<tr>
<td>Ratnam <em>et al.</em> (2000), Canada</td>
<td>2098</td>
<td>18–69</td>
<td>HC1, HC2</td>
<td>Multiple screening practices, 10% random sample of Pap-/HPV- women referred for colposcopy</td>
<td>68</td>
<td>27</td>
<td>91 96</td>
</tr>
<tr>
<td>Schiffman <em>et al.</em> (2000), Costa Rica</td>
<td>8554</td>
<td>18+</td>
<td>HC1, HC2</td>
<td>Population-based, HPV positivity not a criterion for colposcopy referral</td>
<td>88</td>
<td>78</td>
<td>89 94</td>
</tr>
<tr>
<td>Schneider <em>et al.</em> (2000), Germany</td>
<td>4761</td>
<td>18–70</td>
<td>GP5+/6+ PCR</td>
<td>Multiple screening practices, cross-sectional plus 8 months follow-up testing</td>
<td>89</td>
<td>20</td>
<td>94 99</td>
</tr>
<tr>
<td>Belinson <em>et al.</em> (2001), China</td>
<td>1997</td>
<td>35–45</td>
<td>HC2</td>
<td>Unscreened population, community recruitment, all women underwent colposcopy</td>
<td>95</td>
<td>87</td>
<td>85 94</td>
</tr>
<tr>
<td>Blumenthal <em>et al.</em> (2001), Zimbabwe</td>
<td>2073</td>
<td>25–55</td>
<td>HC2</td>
<td>Completed recruitment (irrespective of HIV sero-status) of study in Womack <em>et al.</em> (2000), all women underwent colposcopy</td>
<td>80</td>
<td>44</td>
<td>61 91</td>
</tr>
<tr>
<td>Clavel <em>et al.</em> (2001), France</td>
<td>7932</td>
<td>15–76</td>
<td>HC2</td>
<td>Women free of cytological abnormalities at enrolment, cross-sectional plus 15 months follow-up testing, HPV positivity 100 alone not a criterion for immediate colposcopy referral</td>
<td>100</td>
<td>68</td>
<td>87 95</td>
</tr>
</tbody>
</table>

*a* Conventional cytology and HC2, all ages

*b* Bias-controlled
<table>
<thead>
<tr>
<th>Study, country</th>
<th>Study size</th>
<th>Age (years)</th>
<th>HPV test</th>
<th>Study features</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HPV</td>
<td>Pap</td>
<td>HPV</td>
</tr>
<tr>
<td>Kulasingam et al. (2002), UK</td>
<td>4075</td>
<td>18–50</td>
<td>HC2, MYO9/11 PCR</td>
<td>Family planning clinic recruitment, ThinPrep cytology, 41% random sample of Pap-/HPV- women referred for colposcopy</td>
<td>100</td>
<td>58</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HPV</td>
<td>Pap</td>
<td>HPV</td>
</tr>
<tr>
<td>Petry et al. (2003), Germany</td>
<td>8466</td>
<td>30+</td>
<td>HC2 and PCR</td>
<td>Primary screening network of clinics and gynaecological practices: patients attending routine cervical screening stratified to be representative of all of Germany. Bias controlled</td>
<td>98</td>
<td>37</td>
<td>95</td>
</tr>
<tr>
<td>Salmeron et al. (2003), Mexico</td>
<td>7868</td>
<td>15–85</td>
<td>HC2</td>
<td>Women attending opportunistic screening. Self-collected vaginal samples compared with clinician-collected cervical samples. Those positive for HPV or by cytology were referred for colposcopy and biopsy</td>
<td>93</td>
<td>59</td>
<td>92</td>
</tr>
<tr>
<td>Sankaranarayanan et al. (2004), India</td>
<td>18 085</td>
<td>25–65</td>
<td>HC2</td>
<td>Women attending primary screening in 3 different sites in India. All subjects investigated by colposcopy and, when necessary, underwent a biopsy. Averted verification bias in the design.</td>
<td>46</td>
<td>37</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HPV</td>
<td>Pap</td>
<td>HPV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HPV</td>
<td>Pap</td>
<td>HPV</td>
</tr>
</tbody>
</table>

* LSIL threshold (majority of studies) or LSIL or persistent ASCUS (Kulasingam et al., 2002). Unless otherwise indicated, results are based on conventional cytology. Unless otherwise stated, estimates shown are for CIN 2/3 lesions or worse as disease outcome.

* Verification bias: bias-controlled denotes verification of disease status among all participants and bias-adjusted denotes correction of estimates based on the verification of disease status in a random sample of test-negative women.
samples for testing. Studies varied in terms of timing of collection, collection method, or whether or not visual methods for cervical inspection were used as adjunct screening techniques.

One advantage of HPV DNA testing is that it is suitable for self-sampling, as in many of the studies shown in Table 36. Self-sampling is likely to improve compliance, and is particularly appealing in populations with social or religious limitations on the acceptability of vaginal examinations.

A few large randomized controlled trials of HPV testing in primary cervical cancer screening are currently in progress. Of note are the UK "HPV in Addition to Routine Testing" (HART) investigation (Cuzick et al., 2003), the Dutch POBASCAM trial (Bulkmans et al., 2004), the UK "A Randomized Trial in Screening to Improve Cytology" (ARTISTIC) (H. Kitchener, personal communication), the Osmanabad trial in India (R. Sankaranarayanan, personal communication), the Italian trial (G. Ronco, personal communication), the Canadian Cervical Cancer Screening Trial (CCCaST) (E. Franco, personal communication) and a trial in Finland (Nieminen et al., 2003).

The majority of the estimates in Table 36 must be interpreted with caution because of selection biases and other issues that affect computation of screening performance indices. For instance, the sensitivity and specificity estimates of most studies shown in Table 36 are relative, not absolute, because they are not based on interval cancer incidence and are subject to verification bias (see Chapter 4). The latter occurs whenever the probability of disease verification via the gold standard is dependent on the screening test result. In general, such studies used a design in which only women with one or more positive screening tests were referred for colposcopy and biopsy, which prevented the unbiased estimation of absolute sensitivity and specificity (their estimates should be considered relative). These studies relied on the fact that with two or more tests, there were always combinations of either cytology-negative or HPV-negative women with verified disease status available for analysis. However, the biasing effects of the unequal verification of disease status can be strong and may lead to estimates of screening efficacy that cannot be generalized for cost considerations and other public health uses (Franco, 2000). Such verification bias was averted (by applying the gold standard of disease verification to all women) or corrected (by extrapolating the screening results from a random fraction of women with negative screening tests to those without colposcopic verification) in a few studies, as indicated in Table 36.

An important assumption in dealing with the issue of verification bias is the expectation that the gold standard of colposcopy-guided biopsy provides perfect ascertainment of disease. Studies that either avoided or corrected for the putative bias assumed that a colposcopy-guided biopsy accurately reveals the existence of cervical lesional tissue, which was then used to ascertain the distribution of diseased and non-diseased women, allowing the computation of adjusted estimates of screening validity. While the approach is correct for its intended purpose, i.e., to obtain an improved estimate of the distribution of disease conditional on test results, it should be recognized that a simple colposcopy or even a colposcopy-guided biopsy cannot guarantee that a lesion will be detected. In many test-negative women, the colposcopist cannot visualize lesional tissue and may decide that the colposcopic impression of no disease alone serves as definitive diagnosis. However, a lesion could be hidden in the endocervical canal and not visible. Although this pitfall could be minimized by adopting a colposcopy protocol in which blind biopsies are collected, it is still possible that a fraction of the existing lesions will remain undetected, because of either their location or size. Therefore, in any cross-sectional survey of screening efficacy, the ethically acceptable gold standard for cervical lesions (colposcopy-guided biopsies) is an imperfect one because of inadequate sampling of the entire cervical tissue that is at risk for squamous-cell malignancy.

Even if tissue sampling could be done optimally with respect to lesion site and time of development, one needs to consider also the misclassification of lesion outcome status that exists even with histopathological ascertainment. Studies that involve multiple expert pathologists indicate that the reproducibility in grading histopathology specimens is not high, even with large specimens, such as LEEP-obtained tissue samples. Therefore, a study that is simply based on lesion ascertainment by a single expert pathologist will be more prone to lesion misclassification than one employing a panel of readers that reaches a consensus diagnosis in every case.

Furthermore, as the design of screening efficacy studies evolves from the traditional single-opportunity sampling, cross-sectional layout to long-term, repeated sampling investigations over many years, disease case definition becomes a more dynamic process, requiring the juxtapositioning of screening and diagnostic test results.
obtained from multiple samples collected over time. This process involves combining the results from different diagnostic approaches, which may be differentially triggered by the severity of the lesion grade presumed by the test (HPV or cytology), e.g., colposcopy with simple biopsy for equivocal or low-grade lesions, LEEP for high-grade lesions, etc. Natural history investigations of HPV and cervical neoplasia are examples of studies that have to grapple with this added complexity by having to differentiate between prevalent and incident lesions, progression and regression, and relating them to screening test performance. Calculation of sensitivity and specificity in such studies involves the combination of diagnostic information over multiple samples, which greatly reduces the chance that any lesions are missed through the pitfalls described above for a cross-sectional study relying on colposcopy-guided biopsies alone. On the other hand, the repeated sampling layout of these investigations obviates the need for invasive diagnostic procedures among women testing consistently negative for both HPV and cytology over many visits. The longitudinal nature of the investigation ends up providing the test and diagnostic data that approaches the true distribution of disease dynamics, conditional on study duration. Therefore, correction for verification bias is not a critical issue in these longitudinal studies with intensive follow-up of test-negative women and repeated histological sampling of test-positive cases. However, such studies do have to contend with the issue of distinguishing between prevalent and incident lesions to properly assign the distribution of disease for the purposes of gauging screening test efficacy.

As described in Chapter 4, the ideal estimation of sensitivity assumes follow-up and clinical surveillance, via a cancer registry or otherwise, for invasive cancer that is diagnosed between the screening tests.

Another issue that affects screening performance of HPV tests is the type of specimen. In theory, clinician-collected cervical specimens are ideal in terms of sampling exfoliated cells from the target tissue. Therefore, clinical correlations between lesion severity and presence of HPV should be optimal with such specimens. However, in public-health practice, convenience for the patient and cost-saving considerations have led some to propose self-sampling as a viable alternative to collection by a clinician; the assumption is that the loss in screening accuracy would not be substantial to the point of offsetting the benefits of simplifying specimen collection (Sellors et al., 2000). Self-sampling of genital specimens remains an attractive option in developing countries and in remote regions where health-care providers cannot be available at point-of-care settings. However, issues of validity, acceptability and training present obstacles to wider application of self-sampling.

Costs and potential hazards

The most important obstacles to more widespread acceptance of HPV testing in cervical cancer screening are its high unit cost and the fact that the technology is not in the public domain, as it is for cervical cytology. The cost-effectiveness of HPV testing is heavily dependent on assumptions related to the intrinsic cost of the test, the infrastructure available in the setting where the screening will be implemented, the length of interval between screening visits, and the existing expenditures incurred by quality assurance imposed by local legislation.

There are no additional physical hazards associated with the application of HPV testing technology for the purpose of cervical cancer screening, as the specimen used in the test is the same as that collected for a traditional cytological test. Only minor discomfort and very minimal risk are associated with obtaining exfoliated cervical cell samples. On the other hand, little is known about the psychological and emotional impact of communicating positive HPV test results to women. As knowledge about HPV has become more widespread, there has been a gradual shift in how the medical and public-health communities consider cervical cancer prevention; the perspective has moved from an oncological one to a model in which a sexually transmitted infection is the target (Franco, 2003). Implementation of testing for HPV in primary screening for cervical cancer would lead to a large proportion of women having to be told that they harbour a sexually transmitted viral infection that can ultimately cause cancer. There is a dearth of research on the merits and consequences of conveying this information. The vast majority of such women will not be required to change their lifestyle or to be referred for a more aggressive diagnostic procedure on the basis of this information, since their infection will be found to be transient. Therefore, it is debatable whether conveying this information would bring any real benefit to a screening participant. Practically nothing is known on the potential negative impact, including social and legal implications, of imparting this information. Also the dynamics of between-gender transmission of HPV infection are poorly understood. Such information is important in screening contexts, e.g. for health providers to convey meaningful information on risk to couples.

Another concern with the use of HPV testing in cancer screening is the potential for a breakdown in quality-control safeguards if too many commercial test suppliers enter the market without a certain level of regulatory control of performance standards by health-care or government agencies. At present there are only a
few commercial suppliers of HPV diagnostic systems. The two major ones can afford to keep up strict quality-control standards in reagent batch production and performance characteristics by passing on the costs to the consumers or their health insurers, private or public. However, increased competition resulting in diminishing market shares and reductions in the cost of testing might lead test manufacturers to relax their standards of quality. Such a scenario could prove disastrous in many respects, since there are theoretically many more variables that can affect the performance of HPV testing than there are for cytology-based screening. It is imperative, therefore, that early performance and proficiency standards be agreed upon by public health agencies involved with quality assurance of cervical cancer screening.

**Other emerging techniques**

This section describes three new developments in screening methods: (1) computer-assisted cytological interpretation of cervical smears, (2) use of physical real-time devices and (3) detection of molecular surrogate markers of cancer progression.

**Computer-assisted reading of cervical smears**

The aim of automation-assisted screening is to increase the sensitivity of cytological testing by finding, for instance, small abnormal squamous and glandular cells, known to be very difficult to detect in conventional screening; it should also increase specificity by selecting only lesions corresponding to objective reproducible criteria. Automated screening should also increase productivity by excluding normal slides or part of the slides from manual screening by selecting most atypical images from a slide to be checked by the cytologist, so as to allow more slides to be screened without increasing the number of staff.

The PAPNET system, that is no longer commercially available, included neural network software and traditional imaging technology. It selected 128 of the most suspect fields in conventional cervical cytological specimens and presented these images on a video review screen. The cytotechnician then interpreted the images on the screen and decided whether to carry out manual screening. Another system that was introduced in the 1990s was AutoPap. This computerized scanning device was originally designed for algorithmic classification of conventional cervical cytology specimens, but was later approved for liquid-based cytology specimens. The device was initially approved in the USA by the FDA as a method to be used for quality control. In the quality control mode, only those specimens classified as normal (within normal limits) were reviewed through the device. Subsequently AutoPap was approved by the FDA for primary screening (Dunton, 2000). In the primary screening mode, all slides are processed through the device and then, on the basis of an ‘abnormality index’ assigned to the slide by the algorithmic processing feature, each slide is either filed without manual review by a cytotechnician (up to 25% of all specimens) or reviewed manually in the normal manner.

Both of these devices have served as prototypes for newer devices that are being developed to help automate the evaluation of cervical cytology specimens. Recently, the ThinPrep Imaging System (Cytyc, Boxborough, MA, USA) has received FDA approval for use in primary screening of liquid-based cytology specimens in the USA. This system uses image analysis and algorithmic processing to identify a fixed number of the worst microscopic fields on a given slide; a motorized computer-controlled microscope stage then takes the cytotechnician directly to these specific microscope fields. TriPath Imaging has published results using a similar type of device, referred to as the Focal Point location-guided screening device, that is not yet FDA-approved (Wilbur et al., 2002). This device is based on the earlier AutoPap device.

Several studies have reported the test accuracy of automation-assisted screening (Kok & Boon, 1996; Wilbur et al., 1996; Koss et al., 1997; Michelow et al., 1997; Bartels et al., 1998; Doorneward et al., 1999; Halford et al., 1999; PRISMATIC Project Management Team, 1999; Bergeron et al., 2000a; Duggan, 2000; Kok et al., 2000). They show generally a better test sensitivity with at least the same specificity as conventional screening. Most studies were retrospective (quality control) and/or involved rather small numbers of smears. One larger prospective study, conducted by the PRISMATIC Project Management Team (1999), including 21 700 smears, also showed equal sensitivity but better specificity for automated screening, as well as higher productivity. Results of only two randomized prospective public health trials in a primary screening setting have been reported. One of these studies found clearly higher detection rates of *in situ* and invasive carcinoma (Kok & Boon, 1996). However, the second study, integrated in the Finnish organized screening programme and involving several cytological laboratories, did not clearly confirm this result (Nieminen et al., 2003) (Table 37), showing sensitivity and specificity nearly equal to those of traditional cytological screening (Table 38).

The few randomized prospective studies and other performance studies have shown that automation-assisted screening may be feasible as a part of...
Table 37. Comparison of histologically verified cervical lesions between the PAPNET® arm and the conventional screening arm: number (N) and proportion (per 1000) of screenees, odds ratios (OR), with 95% confidence intervals (CI) (logistic regression)

<table>
<thead>
<tr>
<th>Histological diagnosis</th>
<th>PAPNET arm</th>
<th>Conventional arm</th>
<th>OR</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total 65 527</td>
<td>Total 25 767</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>per 1000</td>
<td>N</td>
<td>per 1000</td>
</tr>
<tr>
<td>Invasive cancer</td>
<td>44</td>
<td>0.67</td>
<td>8</td>
<td>0.31</td>
</tr>
<tr>
<td>In situ carcinoma</td>
<td>79</td>
<td>1.20</td>
<td>18</td>
<td>0.68</td>
</tr>
<tr>
<td>CIN 3</td>
<td>124</td>
<td>1.89</td>
<td>44</td>
<td>1.70</td>
</tr>
</tbody>
</table>

NS, not significant
From Kok & Boon (1996)

Table 38. Specificity of the PAPNET and conventional Pap-smear test with cut-off levels of ASCUS+ and LSIL+ for invasive cancer and for an outcome of CIN 2+ or invasive cancer in primary screening setting

<table>
<thead>
<tr>
<th>Cytological threshold: ASCUS+</th>
<th>Negative histology</th>
<th>Negative Pap smear</th>
<th>Specificity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome: invasive cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAPNET</td>
<td>36 222</td>
<td>33 447</td>
<td>92.3</td>
</tr>
<tr>
<td>Conventional</td>
<td>72 453</td>
<td>67 241</td>
<td>92.8</td>
</tr>
</tbody>
</table>

| Outcome: CIN2+                |                    |                    |               |
| PAPNET                         | 36 171             | 33 447             | 92.5          |
| Conventional                  | 72 353             | 67 240             | 92.9          |

<table>
<thead>
<tr>
<th>Cytological threshold: LSIL+</th>
<th>Negative histology</th>
<th>Negative Pap smear</th>
<th>Specificity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome: invasive cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAPNET</td>
<td>36 222</td>
<td>35 972</td>
<td>99.3</td>
</tr>
<tr>
<td>Conventional</td>
<td>72 453</td>
<td>71 890</td>
<td>99.2</td>
</tr>
</tbody>
</table>

| Outcome: CIN2+                |                    |                    |               |
| PAPNET                         | 36 171             | 35 970             | 99.4          |
| Conventional                  | 72 353             | 71 887             | 99.4          |

From Nieminen et al. (2003)
routine primary screening and with the devices tested it seems to perform at least as well as conventional screening in an organized well functioning programme. Automation-assisted screening may improve the results of a sub-optimal screening organization, but may have no advantage over a well organized, high-quality screening programme other than possibly handling more samples with same quality.

A new generation of automated devices for use with liquid-based cytology is now being launched, the performance of which has not yet been evaluated in randomized trials

**Physical real-time devices**

Advantages of physical real-time devices would be to allow non-invasive on-spot diagnosis, an ability to acquire an objective machine-generated result, applicability in primary health-care settings, the non-requirement for highly trained colposcopists, and high acceptability by women (Soler & Blumenthal, 2000; Basen-Engquist et al., 2003; Wright, 2003).

Normal and neoplastic cervical epithelia have different physical and biochemical properties, yielding distinct patterns in conductance of electrical pulses and reflectance of light waves (Mahadevan et al., 1993; Ramanujam et al., 1994; Richards-Kortum et al., 1994; Wright et al., 2002c). These differences have been applied in fluorescent spectroscopic devices which capture electro-physical signals from the stimulated cervix and analyse the patterns using algorithms to discriminate between normal and neoplastic tissue (Burke et al., 1999; Follen Mitchell et al., 1999). The TruScreen (formerly Polarprobe, Polartechnics Limited, Sydney, Australia) is a portable device that measures the response of the cervical surface to low-voltage electric stimuli and light waves of four different wavelengths. The sensor captures the emitted signals and computer software integrates the information and provides a diagnosis of CIN or the absence of CIN (Mould & Singer, 1997; Singer et al., 2003). In a study of 651 women in ten international centres, the relative sensitivity for histologically confirmed CIN 2 or worse lesions as diagnosed by TruScreen was 70%; the corresponding sensitivity for cytology was 69% and for a combination of TruScreen and cytology was 93% (Singer et al., 2003).

Fluorescence spectroscopy is based on the measurement of autofluorescence from tissue molecules such as FAD, NADPH and collagen that emit light after excitation with low-power laser light of certain wavelengths (Burke et al., 1999; Ferris et al., 2001a; Follen Mitchell et al., 1999). Multi-modal spectroscopy integrates several types of spectroscopy such as intrinsic fluorescence, diffuse reflectance and light scattering. These integrated techniques allow the examination of both biochemical characteristics and morphological features (nuclear size, blood perfusion, cellular changes), so as to optimize the distinction between normal and abnormal tissue (Nordstrom et al., 2001; Georgakoudi et al., 2002). Spectroscopic instruments are continually being improved (Drzek et al., 2003).

The performance of fluorescence spectroscopic devices has been found promising in several small trials, usually conducted by the manufacturer and most often on selected groups of women. In a series of 111 women, accuracy to detect CIN 2 or worse lesions was higher for multimodal spectroscopy than for cytology (area under the ROC curve (AUC): 95% for spectroscopy and 78% for cytology) (Ferris et al., 2001a). Larger multicentre trials are needed to confirm these preliminary results.

**Molecular surrogate markers**

Because both HPV DNA testing and cytological screening yield considerable numbers of women to be referred for colposcopy and biopsy, who are subsequently found not to harbour high-grade disease, it seems worthwhile to look for markers which, at one test occasion, might identify women susceptible to progression with a high predictive value.

Biomolecular pathways leading from HPV infection to the development of cervical dysplasia and cancer are becoming well understood (zur Hausen, 2000, 2002). Continued expression of the viral early oncogenes E6 and E7 appears to be an essential factor in the neoplastic transformation and maintenance of immortalized cell lines, apparently by inactivation of the tumour-suppressor proteins p53 and pRb, respectively (zur Hausen, 1994, 2000) (see Chapter 1). Certain key DNA, RNA or protein markers arising during the neoplastic transformation process might be measured to predict the progressive character of disease in screening, diagnosis and prognosis. However, it is not clear that one common molecular carcinogenetic pathway is involved and it is therefore possible that no single molecular marker will ever on its own allow distinction between progressive and non-progressive disease.

Potential markers of progression include messenger RNA for the E6 or E7 proteins, HPV DNA sequences integrated into the human genome, over-expression of cell cycle regulator proteins or proliferation protein markers, and determination of certain genetic or immunological profiles (Sotlar et al., 1998; Arias-Pulido et al., 2002; Bibbo et al., 2002; Kadish et al., 2002; von Knebel Doeberitz, 2002; Altiov, 2003; Sherman, 2003; Solomon, 2003; Wang & Hildesheim, 2003) (see Table 39).
### Table 39. Potential markers for HPV-induced cervical intraepithelial neoplasia or cancer

<table>
<thead>
<tr>
<th>Marker</th>
<th>Change in expression</th>
<th>Family</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protein markers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p16(^{INK4A})</td>
<td>Increased</td>
<td>Cyclin-dependent kinase (CDK) inhibitor</td>
<td>E7-mediated degradation of the Rb gene gene yields enhanced transcription of the gene coding for p16.</td>
</tr>
<tr>
<td>p53</td>
<td>Decreased</td>
<td>Anti-tumour regulating protein, involved in apoptosis</td>
<td>Viral E6 protein from oncogenic HPV types binds p53, facilitating its degradation through the ubiquitation pathway.</td>
</tr>
<tr>
<td>Ki-67</td>
<td>Increased</td>
<td>Cell proliferation marker</td>
<td>Abnormal cell proliferation beyond basal cell layers.</td>
</tr>
<tr>
<td>PCNA (proliferating cell nuclear antigen)</td>
<td>Increased</td>
<td>Cell proliferation marker</td>
<td>Abnormal cell proliferation beyond basal cell layers.</td>
</tr>
<tr>
<td>Cyclin E</td>
<td>Increased</td>
<td>Protein associating with CDK2</td>
<td>Cyclin E associated with CDK2 drives cells from G1 to S phase through phosphorylation of pRb and other targets.</td>
</tr>
<tr>
<td>Mcm5, Cdc6 c-myc</td>
<td>Increased</td>
<td>Proliferation markers</td>
<td>Abnormal cell proliferation beyond basal cell layers.</td>
</tr>
<tr>
<td>Telomerase</td>
<td>Increased</td>
<td>Nucleoprotein consisting of hTR (RNA) and hTERT (enzyme)</td>
<td>Controls length of telomeres and plays a role in cell immortalization</td>
</tr>
<tr>
<td><strong>RNA markers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E7 or E6 mRNA</td>
<td>Presence</td>
<td>Viral mRNA, transcripts of E6 or E7 gene</td>
<td>Presence of mRNA for E6 or E7 indicates active expression of oncogenes.</td>
</tr>
<tr>
<td><strong>DNA markers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decrease in the ratio E2/E6 and E2/E7 viral DNA</td>
<td></td>
<td>Viral DNA sequences integrated in host genome</td>
<td>Viral integration often occurs at the E2 gene of the HPV genome. Disruption of the E2 gene yields a more intensive transcription of the oncogenes E6 and E7. In the episomal state E2 and E6 DNA are present in equal amounts, while in the integrated form, less or no intact E2 is present.</td>
</tr>
<tr>
<td>Markers of genetic host – polymorphism of p53 gene</td>
<td></td>
<td></td>
<td>p53 with arginine at position 72 of p53 should have a higher affinity for the E6 oncogene. Arg/Arg homozygotes should therefore have a higher risk of cervical neoplasia than Arg/Pro heterozygotes or Pro/Pro homozygotes(^a)</td>
</tr>
</tbody>
</table>

\(^a\) Increased risk for CIN or cervical cancer among Arg/Arg homozygotes is an inconsistent finding in the literature (see Chapter 1).
**p16**

Cyclin-dependent kinases (CDKs), cyclins and CDK inhibitors are key molecules that control the cell cycle and coordinate DNA synthesis, chromosome separation and cell division (Morgan, 1997). Viral oncoproteins interfere directly or indirectly with several of these CDKs (Cho et al., 2002). p16INK4A inhibits the CDK4/6 interaction with cyclin D1, preventing progression through the G1/S checkpoint of the cell cycle (Keating et al., 2001). Accumulation of p16INK4a mRNA and protein has been reported in response to inactivation of the retinoblastoma gene product (pRb) through binding with viral E7 (Xiong et al., 1993; Serrano, 1997). However, this overexpressed p16 is inert since pRB-function is neutralized by E7 (Medema et al., 1995; Khleif et al., 1996). Overexpression of p16 protein is considered to be a marker for progression from HPV infection to cervical cancer (von Knebel-Doeberitz, 2002).

Immuno-detection of p16 using monoclonal antibodies in histological material was described by Klaes et al. (2001). Other studies of p16 immunoreactivity in histological material with different levels of abnormality and HPV infection status (Table 40) have used various types of primary and secondary antibody and chromogen. The sensitivity of p16-immunohistochemical detection of CIN 2 or worse lesions in histological preparations varied between 70% and 100%, while the specificity ranged from 34% to 100% (Keating et al., 2001; Bibbo et al., 2003; Murphy et al., 2003; Negri et al., 2003; Zielinski et al., 2003; Wang et al., 2004b).

Klaes et al. (2002) showed improved inter-observer concordance in histological interpretation of p16-immunostained material (group kappa = 0.94; 95% CI 0.84–0.99) compared with haematoxylin–eosin stained material (group kappa = 0.71; 95% CI 0.65–0.78).

The potential use of p16 immunostaining in cytological smears, and the correlation with cytological and histological results, has been examined recently (Table 41). p16 immunostaining has been found to facilitate the retrieval of dysplastic cervical cells on a slide (Bibbo et al., 2002, 2003; Sahebali et al., 2004).

Strong nuclear and cytoplasmic p16 staining in conventional or liquid-based cervical smears gave a sensitivity for detection of HSIL or worse of about 90–100%; the specificity varied between 36% and 100% (Klaes et al., 2001; Saqi et al., 2002; Murphy et al., 2003; Nassar et al., 2003). Sporadic immunoreactivity in normal squamous metaplastic, inflammatory cells and more systematic staining of endometrial and tubal metaplastic cells and of bacteria has been reported (Bibbo et al., 2002, 2003; Saqi et al., 2002; Riethdorf et al., 2002).

These accuracy measures have been computed from very small and highly selective series and cannot be considered representative for real screening or clinical situations, but the results are promising.

Possible advantages of immunostaining of protein markers include the higher reproducibility of microscopic interpretation, quicker detection of stained lesions and appropriateness for automated detection. A disadvantage is the presence of background staining and positive staining of endometrial or tubal cells, requiring the determination of criteria to define positivity, balancing the sensitivity against specificity.

Biochemical detection of p16 protein in lysates of cervical swab samples using a sandwich ELISA assay is a potentially simple approach for resource-poor settings (Herkert et al., 2004).

**Ki-67**

Expression of the Ki-67 protein occurs in proliferating cells and its presence is normally confined to the basal or suprabasal epithelial cell layers. Expression of Ki-67 allows distinction of atrophic cells (negative for Ki-67) from neoplastic cells (positive for Ki-67) in menopausal women (Ejersbo et al., 1999; Mittal et al., 1999; Bulten et al., 2000). Expression beyond the inner third of the cervical epithelium is observed in case of CIN and cancer (Bulten et al., 1996; Keating et al., 2001). Several authors have found a significant correlation between the presence or intensity of Ki-67 and the severity of cytological abnormality in cytological preparations (Dunton et al., 1997; Sahebali et al., 2003). Dunton et al. (1997) found a sensitivity of 89% for Ki-67 immunostaining in a set of selected abnormal smears for detection of histologically confirmed CIN 2+ lesions, whereas the specificity was 65%.

**Other proliferation or cell cycle regulating markers**

Several other proteins are over-expressed in proliferating cells and certain cell progression regulators have been proposed as potential markers for cervical neoplasia, such as proliferating cell nuclear antigen (PCNA) (Demeter et al., 1994; Mittal et al., 1993), Mcm5 and Cdc6 (Williams et al., 1998) and cyclin E (Altiok, 2003).

Proliferation markers are physiologically present in basal or para-basal epithelial cells, and are an objective indicator of neoplasia when observed beyond the lower cell layers. In cervical smears lacking architectural information, the presence of proliferation markers is less informative and can easily yield false positive results.

**Telomerase**

Telomeres are repeated arrays of six nucleotides (TTAGGG) at the chromosome ends that protect chromosomes against degradation and aberrant fusion or recombination (Collins &
# Overview of p16 immunoreactivity in histological material by severity of lesion and by HPV status

<table>
<thead>
<tr>
<th>Reference</th>
<th>Detection system: primary and antibody and chromogen</th>
<th>p16 positivity</th>
<th>Lesion, HPV status</th>
<th>N</th>
<th>% p16+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sano et al. (1998)</td>
<td>-Mouse monoclonal antibody (JC8)</td>
<td>Diffuse</td>
<td>HPV 16</td>
<td>60</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td>-Biotinylated horse anti-mouse antibody</td>
<td></td>
<td>Other hrHPV+</td>
<td>28</td>
<td>96.4%</td>
</tr>
<tr>
<td></td>
<td>-Chromogen: DAB</td>
<td></td>
<td>HPV 6 or 11</td>
<td>34</td>
<td>0.0%</td>
</tr>
<tr>
<td>Keating et al. (2001)</td>
<td>-Clone G175-405 (Pharmigen, -Biotinylated goat anti-mouse antibody</td>
<td>Diffuse</td>
<td>Normal</td>
<td>24</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>-Chromogen: DAB</td>
<td></td>
<td>LSIL</td>
<td>24</td>
<td>37.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HSIL</td>
<td>37</td>
<td>70.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hrHPV+ and CIN</td>
<td>40</td>
<td>70.0%</td>
</tr>
<tr>
<td>Klaes et al. (2001)</td>
<td>-Clone E6H4 (MTM Lab., Heidelberg)</td>
<td>Diffuse</td>
<td>Normal, hrHPV–</td>
<td>32</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>-Biotinylated goat anti-mouse antibody</td>
<td></td>
<td>Normal, hrHPV+</td>
<td>10</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>-Chromogen: aminoethylcarbazole with hydrogen peroxide in acetate buffer</td>
<td></td>
<td>Inflammation, hrHPV–</td>
<td>30</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Inflammation, hrHPV+</td>
<td>18</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Reserve cell hyperplasia, hrHPV–</td>
<td>13</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Reserve cell hyperplasia, hrHPV+</td>
<td>8</td>
<td>12.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN1, hrHPV–</td>
<td>32</td>
<td>46.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN1, hrHPV+</td>
<td>15</td>
<td>86.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN2, hrHPV–</td>
<td>14</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN2, hrHPV+</td>
<td>18</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN3, hrHPV–</td>
<td>9</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN3, hrHPV+</td>
<td>51</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Invasive cancer, hrHPV–</td>
<td>5</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Invasive cancer, hrHPV+</td>
<td>55</td>
<td>96.4%</td>
</tr>
<tr>
<td>Bibbo et al. (2002)</td>
<td>-Clone E6H4 (MTM Lab., Heidelberg)</td>
<td>Focal or diffuse</td>
<td>Normal</td>
<td>3</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>-Mouse non-avidin-biotin Envision+ polymer (Dako)</td>
<td></td>
<td>CIN 1</td>
<td>19</td>
<td>73.7%</td>
</tr>
<tr>
<td></td>
<td>-Chromogen: DAB</td>
<td></td>
<td>CIN 2</td>
<td>11</td>
<td>90.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN 3</td>
<td>14</td>
<td>100.0%</td>
</tr>
<tr>
<td>Bibbo et al. (2003)</td>
<td>-Clone E6H4 (MTM Lab., Heidelberg)</td>
<td>Focal or diffuse</td>
<td>Chronic cervicitis</td>
<td>5</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>-Mouse non-avidin-biotin Envision+ polymer (Dako)</td>
<td></td>
<td>Squamous metaplasia</td>
<td>2</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>-Chromogen: DAB</td>
<td></td>
<td>CIN 1</td>
<td>5</td>
<td>40.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN 2</td>
<td>4</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN 3</td>
<td>11</td>
<td>100.0%</td>
</tr>
<tr>
<td>Murphy et al. (2003)</td>
<td>-Clone G175-405 (Pharmingen, San Diego)</td>
<td>&gt;10% positive staining</td>
<td>Normal</td>
<td>21</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>-Biotinylated universal antibody, avidin–biotin complex (Vector Laboratories, Burlingame)</td>
<td></td>
<td>cGIN</td>
<td>5</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td>-Chromogen: DAB</td>
<td></td>
<td>CIN 1</td>
<td>38</td>
<td>92.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN 2</td>
<td>33</td>
<td>72.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIN 3</td>
<td>46</td>
<td>91.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Squamous-cell carcinoma</td>
<td>8</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Adenocarcinoma</td>
<td>2</td>
<td>100.0%</td>
</tr>
<tr>
<td>Negri et al. (2003)</td>
<td>-Clone E6H4 (MTM Lab., Heidelberg)</td>
<td>Diffuse</td>
<td>Reactive cells</td>
<td>15</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>-Avidin–biotin kit (Lab. Vision Corp., Fremont)</td>
<td></td>
<td>Endocervical glandular atypia</td>
<td>4</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>-Chromogen: aminoethylcarbazole</td>
<td></td>
<td>Adenocarcinoma in situ</td>
<td>8</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Adenocarcinoma</td>
<td>18</td>
<td>94.4%</td>
</tr>
<tr>
<td>Zielinski et al. (2003)</td>
<td>-Clone E6H4 (MTM Lab., Heidelberg)</td>
<td>Diffuse &amp; strong</td>
<td>Adenocarcinoma, hrHPV–</td>
<td>5</td>
<td>20.0%</td>
</tr>
<tr>
<td></td>
<td>-Biotinylated rabbit anti-mouse antibody</td>
<td></td>
<td>Adenocarcinoma, hrHPV+</td>
<td>20</td>
<td>95.0%</td>
</tr>
<tr>
<td></td>
<td>-Chromogen: DAB or aminoethylcarbazole</td>
<td></td>
<td>Adenocarcinoma endom., hrHPV-</td>
<td>15</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Abbreviations: DAB, 3,3’-diaminobenzidine; hr, high risk; cGIN, cervical glandular intraepithelial neoplasia
<table>
<thead>
<tr>
<th>Reference</th>
<th>Detection system: primary antibody and chromogen</th>
<th>Preparation</th>
<th>Cytological lesion/histological lesion in corresponding biopsy</th>
<th>N</th>
<th>% p16+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klaes et al. (2001)</td>
<td>-Clone E6H4 (MTM Lab., Heidelberg) -Biotinylated horse anti-mouse antibody -Chromogen: aminoethylcarbazole</td>
<td>Conventional smears</td>
<td>Pap I/II Pap IIID+ (LSIL+)</td>
<td>36</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>100.0%</td>
</tr>
<tr>
<td>Bibbo et al. (2001)</td>
<td>-Clone E6H4 (MTM Lab., Heidelberg) -Mouse non-avidin-biotin Envision+ polymer (Dako) Chromogen: DAB</td>
<td>ThinPrep LBC</td>
<td>Within normal limits LSIL HSIL</td>
<td>26</td>
<td>96.2%</td>
</tr>
<tr>
<td>Saqi et al. (2002)</td>
<td>-p16 antibody (Neomarkers, Fremont) -Envision + system (Dako)</td>
<td>SurePath LBC</td>
<td>Within normal limits AGUS LSIL HSIL Squamous-cell carcinoma Adenocarcinoma</td>
<td>25</td>
<td>4.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>60.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>80.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>90.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>100.0%</td>
</tr>
<tr>
<td>Bibbo et al. (2003)</td>
<td>-Clone E6H4 (MTM Lab., Heidelberg) -Mouse non-avidin-biotin Envision+ polymer (Dako) -Chromogen: DAB</td>
<td>ThinPrep LBC</td>
<td>Chronic cervicitis Squamous metaplasia CIN 1 CIN 2 CIN 3</td>
<td>5</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>40.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>83.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>100.0%</td>
</tr>
<tr>
<td>Murphy et al. (2003)</td>
<td>-Clone G175-405 Pharmingen, San Diego -Biotinylated universal antibody, avidin-biotin complex (Vector Laboratories, Burlingame) -Chromogen: DAB</td>
<td>ThinPrep LBC</td>
<td>Normal cGIN CIN 1–3</td>
<td>12</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td>100.0%</td>
</tr>
<tr>
<td>Nassar et al. (2003)</td>
<td>-Monoclonal antibody (Neomarkers) Mouse non-avidin-biotin Envision+ polymer (Dako)</td>
<td>Surepath LBC</td>
<td>Not neoplastic Benign cell changes ASCUS LSIL HSIL</td>
<td>10</td>
<td>50.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td>11.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td>14.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>50.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>100.0%</td>
</tr>
<tr>
<td>Negri et al. (2003)</td>
<td>-Clone E6H4 (MTM Lab, Heidelberg) -Avidin-biotin kit (Lab Vision Corp., Fremont) -Chromogen: aminoethylcarbazole</td>
<td>ThinPrep AGUS</td>
<td>AGUS</td>
<td>10</td>
<td>100.0%</td>
</tr>
<tr>
<td>Nieh et al. (2003)</td>
<td>-Clone E6H4 (MTM Lab., Heidelberg) -Mouse non-avidin-biotin Envision+ polymer (Dako) -Chromogen: DAB</td>
<td>ASCUS Pap smears</td>
<td>Reactive CIN 1 CIN 2/3 Squamous carcinoma Adenocarcinoma in situ</td>
<td>21</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td>8.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
<td>94.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Abbreviations: DAB, 3,3′-diaminobenzidine; LBC, liquid-based cytology; cGIN, cervical glandular intraepithelial neoplasia; AGUS, atypical glandular cells of undetermined significance.
Mitchell, 2002). They become progressively shorter as cells multiply, resulting in chromosomal instability and senescence when a critical short length is reached (Counter et al., 1992). The enzyme telomerase is a ribo-nucleoprotein composed of an RNA part (hRT) and a catalytic part (hTERT), which controls telomere length and is believed to play a role in immortalization of cells (Mathon & Lloyd, 2001; Blasco, 2002). Its activity is increased in CIN and cancer. The intensity of telomerase activity is reported to be correlated with the severity of the abnormality in biopsies and in cervical scrapings, but reliable detection of hTR, hTERT and telomerase activity is still limited by analytical deficiencies (Oh et al., 2001; Jarboe et al., 2002; Fu et al., 2003).

Detection of viral oncogene transcripts
Viral mRNA can be detected using (nested) real-time PCR or nucleic acid sequence-based amplification assay (NASBA) (Smits et al., 1995; Sotlar et al., 1998; Deiman et al., 2002). Presence of viral mRNA transcripts coding for the E6 and E7 proteins from high-risk HPV might be a more specific predictor of progressive infection than simple presence of HPV DNA (Nakagawa et al., 2000; Cuschieri et al., 2004). A commercial kit exists (PreTect HPV-Proofer, NorChip AS, Klokkarstua, Norway) which detects E6 mRNA from HPV16 and E7 mRNA from HPV types 18, 31, 33 and 45. The presence of E6 and E7 mRNA and absence of viral L1 DNA (negative test result on consensus PCR) indicate integration of viral DNA in the human genome, yielding enhanced transcription of the E6–E7 sequence. Molden et al. (2004) found that rates of HPV-Proofer positivity and presence of HPV DNA (measured with GP5+/6+ consensus PCR and type-specific PCR) increased with the severity of cytological or histological cervical abnormality. Nevertheless, lower proportions of mRNA-positive results were observed in normal cases, ASCUS and LSIL (see Table 42).

Viral DNA integration markers
Testing for HPV integration appears to increase the predictive value that an HPV-positive sample is derived from tissue containing progressive CIN or cervical cancer (Klaes et al., 1999). Viral integration often occurs at the E2 gene of the HPV genome. Disruption of the E2 gene is believed to result in more intensive transcription of the oncogenes E6 and E7. In the episomal state, E2 and E6 DNA are present in equal amounts, while in the integrated form, less intact E2 is present (zur Hausen, 2002). A decrease in E2/E6 DNA ratio assessed with real-time PCR is another potential progression marker. However, other authors have reported exclusively episomal HPV DNA in tumours (Cullen et al., 1991; Pirami et al., 1997).

Micro-array technology
It is believed that profiles of multiple host–virus interaction factors will reveal possibilities for accurate individualized risk assessment and prognosis prediction. By the use of DNA micro-array technology or DNA chips, expression of many genes can be analysed at once using only a small amount of sample (Hughes & Shoemaker, 2001). The first step consists in extraction of mRNA from a tissue sample. Using reverse transcriptase, complementary DNA (cDNA) is synthesized, which is labelled with a fluorescent molecule. This labelled cDNA is subsequently divided over a slide or membrane where hundreds or thousands of known target DNA sequences are fixed. Hybridization of the labelled cDNA with target DNA is detected as a coloured light signal at a particular locus on the array, which indicates expression of a particular gene.

Post-translational changes also play an important role in pathogenesis, and can be studied using protein arrays or proteomics techniques (Wulfkuhle et al., 2003; Lee et al., 2004).

Table 42. Positivity rate for HPV DNA and HPV mRNA in a series of 4136 women presenting at an outpatient gynaecological service, Oslo, Norway

<table>
<thead>
<tr>
<th></th>
<th>Normal 3950</th>
<th>ASCUS 57</th>
<th>LSIL 20</th>
<th>CIN 2 5</th>
<th>CIN 3 12</th>
<th>Squamous cancer 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV mRNA+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>95</td>
<td>12</td>
<td>6</td>
<td>2</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.0001</td>
<td>0.08</td>
<td>0.009</td>
<td>1.00</td>
<td>0.62</td>
<td>1.00</td>
</tr>
<tr>
<td>HPV DNA+</td>
<td>368</td>
<td>27</td>
<td>15</td>
<td>2</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>p-value</td>
<td>2.4%</td>
<td>21.1%</td>
<td>30.0%</td>
<td>40.0%</td>
<td>75.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

From Molden et al. (2004)
General comments
Research on molecular markers has so far been largely restricted to correlation studies documenting the presence or absence or the intensity of the considered marker in cytological or histological material from selected women. These test accuracy measures can be assessed for detection of CIN 2+ but are not representative for real screening, triage or follow-up settings.

Potential advantages from the use of molecular markers in future clinical practice include: triage of women with minor cytological abnormalities (ASCUS and LSIL) with higher specificity than HPV DNA detection; improvement of the accuracy of histology as gold standard for screening test assessment, by more accurate and reproducible classification of histological squamous and glandular cervical lesions and clearer distinction between cervical and endometrial glandular lesions; selection of best treatment procedures; prognosis prediction; and last but not least, more accurate primary screening for cervical progressive cancer precursors.

Combinations of different modalities
As the previous sections have demonstrated, no single currently available screening test for cervical cancer provides an optimal trade-off between sensitivity and specificity. Because various screening techniques are available, applying them in combinations might be advantageous. Although combinations of tests necessarily require extra resources, the added testing accuracy might increase the detection of treatable disease and allow lengthened screening intervals. Research is in progress to find the combinations that are most complementary, to determine how these tests should be combined, and to clarify how to interpret and manage the increased complexity of results. The results so far are promising but far from complete.

Screening with more than one technique
The main classes of available screening techniques are cytology, visual inspection and HPV DNA testing. It is possible to consider combinations of two techniques within a class (e.g., conventional and liquid-based cytology), but most interest has focused on combining two techniques of different classes in the hope of gaining benefit from complementarity. Thus, researchers have examined cytology plus visual techniques, cytology plus HPV DNA testing and, to a limited degree, HPV DNA testing plus visual techniques. Because of the practical limitations of resources, there has been only occasional interest in combining more than two techniques (Reid et al., 1991).

Whenever two screening techniques are combined, with abnormal results from either test taken to constitute an overall positive result, the sensitivity will be higher than that of either test alone (Franco & Ferenczy, 1999). However, the key question is whether the increase in sensitivity is sufficiently greater than random to merit consideration. Increased sensitivity will typically lead to an offsetting decrease in specificity and the trade-off must be examined to determine the overall effect of the combination on screening accuracy. Various statistical methods for evaluating the added value of adding a second test have been suggested, but none has been fully accepted. The best statistical methods generate roughly equivalent conclusions (Macaskill et al., 2002; Ferreccio et al., 2003), although the interpretations depend on varying regional standards of acceptable safety and cost.

The studies on combination of different modalities have been run with designs that provide estimates of sensitivity and specificity that are not totally comparable. Most of the designs were cross-sectional without correction for verification bias (see Chapter 4). Therefore, the results and conclusions depend on the number of tests applied. If two tests only are considered, the cross-sectional sensitivity estimate for the test combination is bound to be 100%. The same applies, albeit not as a logical consequence, if too few women with dual negative tests are subjected to a commonly accepted reference standard such as colposcopic examination with guided biopsy. Colposcopy itself is not sufficiently sensitive to rule out missed disease and therefore its own errors must be recognized when considering results that depend upon it. In general, the sensitivity estimate of any combination of two tests is smaller if more tests are used for detection of disease and those who tested negative on every test are not subjected to the reference standard (colposcopy). Only one study (Sherman et al., 2003b) has been based on interval cancer incidence, the ideal to estimate the true sensitivity (see Chapter 4). In the absence of data on the expected incidence, the risk in screen positives versus that in screen negatives was used as an indicator of sensitivity.

Cytology plus HPV testing
The residual cytology specimen from liquid-based cytology or a co-collected specimen can be tested for oncogenic HPV types. There is much evidence that screening of women with both cytology and HPV DNA tests increases sensitivity for detection of prevalent CIN 3 or cancer sufficiently to permit longer screening intervals than with cytology alone. After consideration of the accumulated evidence regarding increased sensitivity, decreased specificity and the possibility of lengthened screening intervals using the combination, the US FDA
approved HC2 for HPV DNA testing as an adjunct to cytological testing for women aged 30 years and older. The supporting evidence has been summarized by Franco (2003), with additional recent support (Cuzick et al., 2003; Ferreccio et al., 2003). However, in two other recent studies (de Cremoux et al., 2003; Coste et al., 2003), HC2 performed worse than in other published studies and conventional cytology performed considerably better than is usually reported. Several of the supportive studies are summarized in Table 43. The increase in sensitivity from adding HPV testing was generally greater than the decrease in specificity and, in some studies, cytology actually added little to the performance of HPV testing.

In practice, the introduction of HPV testing into routine screening in combination with cytology produces multiple risk strata ranging from very low to very high absolute risk (positive predictive value) of prevalent or incipient CIN 3 or cancer. A woman with HSIL cytology (especially with a positive oncogenic HPV test) has a high risk of underlying CIN 3. In contrast, negative cytology (whether conventional or liquid-based) with a negative HPV test is associated with a risk of CIN 3 within two years that approaches zero (Ferreccio et al., 2003). These extremes of risk stratification are easily managed. However, strategies to manage the very large numbers of women who are HPV-positive and cytology-negative need to be developed and evaluated. Repeating viral and cytological tests between six and twelve months has been proposed (Wright et al., 2004) as an interim measure until more data are available to develop truly evidence-based guidelines.

The negative predictive value of adding an HPV test to cytology is its major utility. HPV infection is so common that a positive test conveys only a moderate positive predictive value for prevalent or incipient CIN 3 and cancer (Sherman et al., 2003b). However, because persistent infection with oncogenic types of HPV is the necessary cause of virtually all cases of cervical cancer, a negative test for oncogenic HPV has unusually high negative predictive value, or reassurance, in the context of negative or even mildly abnormal cytology. The uncommon combination of a negative HPV test and an HSIL cytological result merits further evaluation because of its rarity and the possibility that one of the two results is in error.

**Cytology plus visual techniques**

The interest in combining cytology with some kind of visual inspection is natural, since cytology and colposcopy have comprised the strategy responsible for a half century of successful screening for cervical cancer. Population screening using cytology and colposcopy concurrently has been explored (Olatunbosun et al., 1991), but is much too expensive and demanding of limited expertise for use in most regions. The search for an easier alternative to colposcopy has led to the combination of cytology and cervicography and of cytology and direct visual inspection. Two representative studies to examine these combinations are summarized in Table 44. Cervicography is no longer available commercially, but it is worth noting that studies of cervicography as an adjunct have suggested that a visual technique might complement conventional cytology. The two kinds of technique tend to detect different groups of women with CIN 2 or 3 without an obviously unacceptable loss in specificity (Ferreccio et al., 2003), resulting in overall increased accuracy. The cost-effectiveness of combining cytology with some kind of visual inspection (Shastri et al., 2004) should be compared with alternative strategies and evaluated more fully and formally on a regional basis.

Real-time cervical scanning based optico-electrical devices might improve the sensitivity of cytology (Singer et al., 2003), but the influence on specificity of adding such new techniques is not yet clear.

**HPV plus visual inspection**

In many developing countries, approaches that do not rely on an extensive infrastructure of highly trained personnel must be considered. Because first-rate cytological screening programmes are difficult to create and maintain, there is interest in establishing programmes that rely on more easily performed and standardized techniques. It may be feasible to combine an HPV test for primary screening with triage modalities other than cytology, such as direct visual assessment by non-physician health-care providers. The use of HPV testing plus visual inspection is best considered as a sequential strategy, to maximize sensitivity with acceptable specificity (see below).

**Combining two techniques from the same class**

The combination of two cytological techniques can be seen as a logical extension of re-screening of slides or of computer-assisted screening, as discussed in the first section of this chapter. The combination of conventional and liquid-based cytology is not particularly complementary (Ferreccio et al., 2003) and holds little promise because of the expense of conducting the two tests for each woman screened. Similarly, there is probably no reason to combine two visual techniques (e.g., cervicography and direct inspection) because of the correlated nature of the results and limitations of the techniques (Shastri et al., 2004). HPV testing with multiple techniques could reveal some testing errors, but not enough to justify the high cost.
<table>
<thead>
<tr>
<th>Study</th>
<th>No. of tests</th>
<th>Case identification</th>
<th>Colposcopy for negatives</th>
<th>No. of women</th>
<th>No. of CIN</th>
<th>CIN threshold</th>
<th>Test</th>
<th>Cut-points</th>
<th>Sens.</th>
<th>Spec.</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratnam et al. (2000)</td>
<td>2 P</td>
<td>8</td>
<td>2098</td>
<td>30</td>
<td>CIN2/3</td>
<td></td>
<td>Cytol.</td>
<td>ASCUS</td>
<td>40</td>
<td>92</td>
<td>11</td>
<td>98.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HC1, HC2</td>
<td>10,1 pg/ml</td>
<td>68</td>
<td>91</td>
<td>15</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol.+HC</td>
<td>Either +</td>
<td>76</td>
<td>86</td>
<td>12</td>
<td>99.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC</td>
<td>ASCUS</td>
<td>84</td>
<td>85</td>
<td>17</td>
<td>98.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC+HC2</td>
<td>Cytol.</td>
<td>86</td>
<td>76</td>
<td>13</td>
<td>98.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HC2</td>
<td>ASCUS</td>
<td>94</td>
<td>78</td>
<td>16</td>
<td>99.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol.+HC2</td>
<td>1 pg/ml</td>
<td>95</td>
<td>85</td>
<td>23</td>
<td>99.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC</td>
<td>Either +</td>
<td>92</td>
<td>93</td>
<td>39</td>
<td>99.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC+HC2</td>
<td>ASCUS</td>
<td>59</td>
<td>98</td>
<td>36</td>
<td>99.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol.</td>
<td>1 pg/ml</td>
<td>93</td>
<td>92</td>
<td>15</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol.+HC2</td>
<td>Either +</td>
<td>98</td>
<td>91</td>
<td>12</td>
<td>99.9</td>
</tr>
<tr>
<td>Belinson et al. (2001)</td>
<td>5 P</td>
<td>No</td>
<td>1997</td>
<td>86</td>
<td>CIN2/3</td>
<td></td>
<td>Cytol.</td>
<td>ASCUS</td>
<td>46</td>
<td>98</td>
<td>10</td>
<td>99.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC</td>
<td>1 pg/ml</td>
<td>97</td>
<td>95</td>
<td>9</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC+HC2</td>
<td>Either +</td>
<td>100</td>
<td>95</td>
<td>8</td>
<td>100</td>
</tr>
<tr>
<td>Salmeron et al. (2003)</td>
<td>3 P</td>
<td>No</td>
<td>7732</td>
<td>93</td>
<td>CIN3</td>
<td></td>
<td>Cytol.</td>
<td>ASCUS</td>
<td>34</td>
<td>97</td>
<td>10</td>
<td>98.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HC2</td>
<td>1 pg/ml</td>
<td>64</td>
<td>86</td>
<td>7</td>
<td>99.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol.+HC2</td>
<td>Either +</td>
<td>72</td>
<td>85</td>
<td>7</td>
<td>99.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC</td>
<td>ASCUS</td>
<td>63</td>
<td>94</td>
<td>12</td>
<td>99.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC+PCR</td>
<td>Either +</td>
<td>85</td>
<td>88</td>
<td>9</td>
<td>99.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC+PCR</td>
<td>Either +</td>
<td>91</td>
<td>88</td>
<td>9</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol.</td>
<td>Borderline</td>
<td>77</td>
<td>96</td>
<td>16</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HC2</td>
<td>Mild</td>
<td>70</td>
<td>99</td>
<td>34</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol.+HC2</td>
<td>1 pg/ml</td>
<td>97</td>
<td>93</td>
<td>13</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol.+HC2</td>
<td>2 pg/ml</td>
<td>96</td>
<td>94</td>
<td>15</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol.+HC2</td>
<td>Mild/2 pg</td>
<td>100</td>
<td>94</td>
<td>14</td>
<td>99.9</td>
</tr>
</tbody>
</table>

P, detected at screening; I, interval cancers; Cytol, conventional cytology; LBC liquid-based cytology; PCR, polymerase chain reaction; HC, hybrid capture; PPV, positive predictive value; NPV, negative predictive value.
Sequential tests (triage)
The classical scheme of secondary cancer prevention consists of three steps: sensitive screening of asymptomatic individuals to identify those at risk of disease, specific diagnosis of the disease state and treatment of those with cancer or a cancer precursor (see Figure 45). Triage is an additional step interposed between screening and diagnosis to further stratify individuals with positive primary screening results according to their risk for the disease state. In other words, a second test is performed only if the first test is neither completely normal nor definitely indicative of need for treatment. In this respect, triage is conceptually related to the consideration of residual error in stepwise prediction models.

The utility of a triage test in the context of a cervical cancer screening programme will depend not only on the performance characteristics of the test itself in relation to the primary test, but also on the target screening population, the prevalence of disease, the cost of follow-up, the available resources (logistic and monetary) and patient compliance (Solomon, 2003). Triage is of most value when the screening test lacks specificity and/or the diagnostic procedure is expensive or a limited resource. An efficient triage test should reduce overtreatment, patient anxiety and inconvenience, as well as overall management costs, usually by reducing the number of diagnostic procedures performed—all without excessively sacrificing sensitivity for detection of disease. However, the sequential use of imperfect tests tends invariably to reduce sensitivity somewhat.

Triage of cytology
In cervical cancer screening, the primary test has traditionally been a programme of repeated cytological tests, which generally succeeds because of the typically long natural history of HPV persistence leading to cervical carcinogenesis. A single cytological test is not sufficiently sensitive to serve as an adequate screening test. In an effort to maximize sensitivity and negative predictive value, a test finding of ASCUS or above is used as the threshold for referral for additional follow-up in the USA. ASCUS is a common cytological interpretation, applied in approximately 5% of screening cytological tests. Similarly, in the United Kingdom, approximately 4% of all smears show borderline or mildly dyskaryotic changes. The threshold of test positivity at equivocal cytology substantially increases sensitivity for identifying histological CIN 2/3 (Kinney et al., 1998), but at the cost of repeated cytology or referring millions of women for colposcopy and biopsy. Many of these women are not infected with oncogenic HPV types, and some 90% do not have prevalent CIN 2 or CIN 3 and are not destined to develop it in the immediate future. In this setting of lower specificity, a triage test that further stratifies women according to cancer risk is appealing.

A multicentre, randomized clinical trial was conducted by the US National Cancer Institute to compare different strategies for managing the 2–3 million women with ASCUS and 1.25 million women with LSIL cytological results in the USA each year (Schiffman & Adrianza, 2000). About 40–50% of women with ASCUS are HPV-positive (Manos et al., 1999; Solomon et al., 2001), the actual proportion depending on the patient population and the

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of cases</th>
<th>No. of CIN</th>
<th>CIN</th>
<th>Test</th>
<th>Cut-points</th>
<th>Sensit. (%)</th>
<th>Specif. (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferreccio et al. (2003)</td>
<td>8551</td>
<td>110</td>
<td>CIN3</td>
<td>Cytol.</td>
<td>ASCUS</td>
<td>63</td>
<td>94</td>
<td>12</td>
<td>99.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC</td>
<td>ASCUS</td>
<td>86</td>
<td>88</td>
<td>9</td>
<td>99.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cervicog.</td>
<td>A</td>
<td>62</td>
<td>85</td>
<td>5</td>
<td>99.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol. + Cervicog.</td>
<td>LSIL or P</td>
<td>75</td>
<td>91</td>
<td>10</td>
<td>99.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LBC + Cervicog.</td>
<td>ASCUS or P</td>
<td>93</td>
<td>84</td>
<td>7</td>
<td>99.9</td>
</tr>
<tr>
<td>Shastri et al. (2004)</td>
<td>4039</td>
<td>57</td>
<td>CIN2/3</td>
<td>Cytol.</td>
<td>LSIL</td>
<td>57</td>
<td>99</td>
<td>38</td>
<td>99.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VIA</td>
<td>P</td>
<td>60</td>
<td>88</td>
<td>7</td>
<td>99.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VILI</td>
<td>P</td>
<td>75</td>
<td>84</td>
<td>7</td>
<td>99.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol. + VIA</td>
<td>LSIL or P</td>
<td>83</td>
<td>87</td>
<td>9</td>
<td>99.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytol. + VILI</td>
<td>LSIL or P</td>
<td>89</td>
<td>83</td>
<td>7</td>
<td>99.8</td>
</tr>
</tbody>
</table>

Abbreviations: Cytol. = conventional cytology; LBC = liquid-based cytology; A = equivocal cervigram; P = positive cervigram; PPV, positive predictive value; NPV, negative predictive value; VIA, visual inspection with acetic acid; VILI, visual inspection with Lugol’s iodine
cytomorphological threshold utilized. Importantly, virtually all of the occult CIN 2 or 3 associated with ASCUS is found in the HPV-positive fraction. Therefore, in the context of ASCUS cytology, triage by HPV testing can save approximately 50% of women from unnecessary colposcopy without compromising sensitivity.

This result has been supported in a recent meta-analysis of 15 comparable studies of HPV DNA testing as an alternative to repeat cytology in women who had equivocal results on a previous cytological test (Arbyn et al., 2004b). The pooled sensitivity and specificity to detect histologically confirmed CIN 2 or worse were 84.4% (95% CI 77.6–91.1%) and 72.9% (95% CI 62.5–83.3%), respectively, for overall HPV testing. Restriction to the nine studies (Table 45) where the HC2 assay was used yielded a pooled sensitivity of 94.8% (95% CI 92.7–96.9%) and a pooled specificity of 67.3% (95% CI 58.2–76.4%).

Consensus management guidelines for follow-up of an ASCUS cytology result, based on the accumulated evidence, were developed for the USA under the sponsorship of the American Society for Colposcopy and Cervical Pathology (ASCCP). Acceptable options following an ASCUS cytological interpretation include repeat cytology, immediate colposcopy or HPV testing (Wright et al., 2002c; American College of Obstetricians and Gynecologists, 2003).

Triage by HPV DNA testing of women with ASCUS is now very common in the USA, where, if initial liquid-based cytology is used, ‘reflex’ HPV testing (see Glossary) is considered the preferred triage approach, as it obviates the need for a repeat visit (Wright et al., 2002c). HPV testing has also been recommended to be introduced in the United Kingdom for borderline cytological cases on a pilot basis (Cuzick et al., 1999a, b) and the HART study in women over the age of 30 years has confirmed the validity of this approach (Cuzick et al., 2003). With conventional cytology smears, there is no residual sample available for HPV testing, but in the USA, an additional specimen is now often col-lected to be used for triage if an ASCUS interpretation is obtained (o-th-erwise the co-collected specimen is discarded). When oncogenic HPV is detected in conjunction with an ASCUS cytological interpretation, the tendency at present is to report both findings, rather than to upgrade the cytological interpretation to SIL (Levi et al., 2003).

The ASCUS-LSIL Triage Study (ALTS) and other studies have shown that cytologically identified LSIL, when interpreted stringently (as in, for example, France, Sweden and the USA) is so highly associated with HPV that an HPV triage test (as a sequential test) is not useful, due to low specificity (ASCUS-LSIL Triage Study (ALTS) Group, 2000, 2003a; Arbyn et al., 2002; Scott et al., 2002). In a meta-analysis of studies based on HC2 for detection of CIN 3 (Arbyn et al., 2002), the estimates of relative sensitivity, specificity, positive predictive value and negative predictive value were 95.7% (95% CI 91.2–100), 32.9% (95% CI 17.8–48.0), 32.4% (95% CI 13.4–51.3) and 98.8% (95% CI 97.1–100), respectively. Cuzick et al. (2003) suggested that HPV testing might be useful for triage of mild dyskaryosis based on high negative predictive value, despite the high HPV prevalence associated with this cytological interpretation.

Atypical glandular endocervical cells (AGC), the glandular counterpart of ASC, is a much less common cytological interpretation with a higher risk for underlying precancerous lesions or cancer than ASCUS. Current ASCCP guidelines recommend colposcopic evaluation with endocervical sampling for all women with AGC or ‘atypical endocervical cells’ (Wright et al., 2002c). One study has suggested the possible utility of HPV triage following an AGC result (Ronnett et al., 1999). Finally, colposcopic referral is recommended for another relatively uncommon equivocal interpretation, ASC-H, because of the high risk of underlying CIN 2 or 3 (Wright et al., 2002c).

International variation in cytological terminology, compounded by the use of different morphological criteria for similarly termed diagnoses, might imply that results could not be general-ized between countries (Scott et al., 2002). However, use of an atlas of cytology images, with known HPV status and disease outcome, should allow the performance of HPV triage to be transferred between classification systems and screening programmes without the need for costly repetition of trials (Solomon, 2003).

In regions where expert colposcopic services are limited or expensive, the possibility of triage with another visual technique is attractive. However, several evaluations of triage by cervicography or visual inspection after cytological testing (Costa et al., 2000; Denny et al., 2000b; Mould et al., 2000; Blumenthal et al., 2001; Ferris et al., 2001b) have indicated lower accuracy than HPV DNA testing, with inad-equate sensitivity.

**HPV first, then triage by cytology or visual inspection**

The combination of HPV as an adjunct to cytology may be an interim strategy in an evolution that ultimately leads to primary screening by HPV with triage by cytology. In fact, HPV testing followed by cytology is a rational approach for older women, given the higher sensitivity of HPV testing and the greater specificity of cytology (Sasieni & Cuzick, 2002).

The performance of cytology as a triage test might be very different from...
its characteristics as a screening test (Solomon, 2003). If used as a triage test, there would be a dramatic reduction in the number of tests overall and a marked increase in the yield of positive results, altering the ratio of negative to abnormal specimens. It is unclear how this would affect the sensitivity and specificity of cytological testing.

**HPV then visual inspection**

It might be possible to combine an inexpensive, rapid HPV test of the kinds now under development with simplified visual inspection to produce screening strategies with good characteristics. Evidence supporting this possibility comes from a very few studies where HPV testing and cervicography were both analysed (Ferreccio et al., 2003; Jeronimo et al., 2003). Screening, triage and even treatment services could be combined in the same visit and thereby reduce loss to follow-up in areas remote from health clinics. Simple visual assessment categories (e.g., normal versus lesion treatable by cryotherapy versus lesion requiring a gynaecologist) could be calibrated to maximize reliability and sensitivity if screening was done only once or twice in a woman’s lifetime.

The use of this approach depends upon the development of a widely applicable inexpensive and rapid HPV test. For such a combined approach, revised (more sensitive) criteria might be needed for visual assessment, to prevent a serious loss of overall sensitivity (Denny et al., 2000b). A disadvantage would be the loss of the limited ability of cervical cytology to detect non-cervical neoplasia, particularly endometrial cancer, in a strategy restricted to HPV and visual inspection.

**Triage following visual inspection**

Some investigators have considered two-stage cervical cancer screening in which visual inspection would be followed by a second test (Denny et al., 2000b). However, in a simple application requiring both results to be positive before treatment, the overall sensitivity would be limited by the numbers of cases missed by either test, regardless of the order in which they are applied.

**Follow-up of positive test results**

The conventional confirmatory test following an abnormal primary screening result has been colposcopically directed biopsy. Certain procedures could appear in several of the categories depicted in Figure 45. Thus colposcopy is used in some settings, particularly in Europe, as an adjunctive screening test, but it can also be categorized as a triage modality.

Although colposcopically directed biopsy has been used as the gold standard for diagnosis, recent findings suggest that it misses about a quarter or more of prevalent CIN 3 (ASCUS-LSIL Triage Study (ALTS) Group, 2003b). This implies that women with an apparently negative diagnosis on colposcopy remain at increased cancer risk, possibly requiring more than resumption of routine screening (Viikki et al., 2000). Moreover, the three-stage strategy of screening, requiring a return for histological diagnosis and a third visit for treatment, leads in many regions to unacceptable loss to follow-up. Consequently, there is reason to explore other ways of combining the first test, triage, diagnostic test and management. Women diagnosed with less than CIN 2 by colposcopy are at approximately 10% risk of CIN 2 or CIN 3 within two years; this risk is similar regardless of whether the colposcopically directed biopsy result was ‘negative’ or ‘CIN 1’ (Cox et al., 2003). There is insufficient evidence regard-

### Table 45. Triage of ASCUS cytology by HPV DNA testing (Hybrid Capture 2) for detection of histologically confirmed CIN 2+

<table>
<thead>
<tr>
<th>Study</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>Test positivity</th>
<th>Prevalence of CIN 2+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manos et al. (1999)</td>
<td>0.892</td>
<td>0.641</td>
<td>0.151</td>
<td>0.988</td>
<td>0.395</td>
<td>0.067</td>
</tr>
<tr>
<td>Bergeron et al. (2000b)</td>
<td>0.833</td>
<td>0.616</td>
<td>0.208</td>
<td>0.968</td>
<td>0.432</td>
<td>0.108</td>
</tr>
<tr>
<td>Fait et al. (2000)</td>
<td>0.857</td>
<td>0.971</td>
<td>0.906</td>
<td>0.954</td>
<td>0.235</td>
<td>0.248</td>
</tr>
<tr>
<td>Lin et al. (2000)</td>
<td>1.000</td>
<td>0.745</td>
<td>0.692</td>
<td>1.000</td>
<td>0.527</td>
<td>0.365</td>
</tr>
<tr>
<td>Shlay et al. (2000)</td>
<td>0.933</td>
<td>0.739</td>
<td>0.230</td>
<td>0.993</td>
<td>0.313</td>
<td>0.077</td>
</tr>
<tr>
<td>Morin et al. (2001)</td>
<td>0.895</td>
<td>0.742</td>
<td>0.162</td>
<td>0.992</td>
<td>0.292</td>
<td>0.053</td>
</tr>
<tr>
<td>Rebello et al. (2001)</td>
<td>0.857</td>
<td>0.759</td>
<td>0.581</td>
<td>0.932</td>
<td>0.413</td>
<td>0.280</td>
</tr>
<tr>
<td>Solomon et al. (2001)</td>
<td>0.959</td>
<td>0.484</td>
<td>0.196</td>
<td>0.989</td>
<td>0.568</td>
<td>0.116</td>
</tr>
<tr>
<td>Zielinski et al. (2001)</td>
<td>0.917</td>
<td>0.687</td>
<td>0.149</td>
<td>0.993</td>
<td>0.347</td>
<td>0.056</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value
Modified from Arbyn et al. (2004)
ing the optimal management of women diagnosed with less than CIN 2 by colposcopically directed biopsy. In ALTS, various re-triage strategies combining follow-up cytology and HPV testing were compared (Guido et al., 2003). A single HPV test at 12 months gave the best trade-off of sensitivity and referral percentage. As an alternative, semi-annual cytological sampling appeared to be useful. Further studies are needed to find assays or strategies that more efficiently identify women with occult CIN 3 and permit the majority of women to safely return to routine screening.

More sensitive screening and triage strategies that translate into increased detection of ‘early’ and often very small CIN 3 lesions may lead to earlier treatment, but the impact on cancer outcomes has not been established. Many small high-grade lesions might regress, and others could be detected later, when larger but still intraepithelial (Sherman et al., 2002). Very early detection leads to a greater likelihood of overtreatment of lesions, particularly CIN 2, that might otherwise regress. Identifying markers of risk of progression to cancer is a priority in order to reduce unnecessary treatment and attendant complications and costs associated with treating all cases of CIN 2 or 3. Novel approaches were considered in the previous section of this chapter.
Chapter 3
Use of screening for cervical cancer

Delivery and uptake of screening

Cervical cancer screening started with the introduction of the Papanicolaou test into clinical practice. In many countries, this occurred as part of family-planning services, so that the target group was younger women. Because such services are frequently not well integrated with secondary levels of care, it was not always possible to ensure adequate diagnosis and treatment of women with a positive test result. It has now become clear that organized screening programmes have a greater impact than opportunistic screening because they have the potential to achieve greater participation and this can improve equity of access and the likelihood of reaching women at higher risk.

Cervical cancer screening comprises various types of care or services, ranging from provision of the screening test to diagnosis and treatment, as shown in Figure 46.

Implementation of a national programme requires that there be a national policy that defines the screening age and interval and what method of screening will be used, as well as sufficient political and financial investment. The major issues that have to be considered are:

- The budget to run the programme
- Training of health-care providers in: the logic of the screening policy; carrying out the screening test; patient counselling; and collection and interpretation of monitoring data (participation and follow-up rates)
- Setting up equipment supply systems for the clinic or health centre
- Ensuring that high-quality laboratory services are available
- Establishing a referral pathway for treatment of patients (which may involve training of people at local level and referral for more advanced cases needing specialized treatment)
- Developing the capacity to offer treatment (for in situ disease, definitive treatment and palliative care)
- Setting up national monitoring systems
- Education of the population to ensure participation in the screening programme

Overall, a screening programme should be an integrated system in which, as seamlessly as possible, women are recruited, screened, receive and understand the results, are referred for treatment as required, return for repeat screening as determined by the policy and become advocates for others to participate. This means that all staff must know, understand and give the same message to patients, that services be accessible, equipped and welcoming, and that transport and communications mechanisms with institutions for reading of results and treatment are functional. In other words, a functional health system must operate with sufficient coverage, so that all women in the target group have satisfactory access to services.

The organization and financing of the overall health-care system of a

Organized screening programme

- an explicit policy with specified age categories, method and interval for screening;
- a defined target population;
- a management team responsible for implementation;
- a health-care team for decisions and care;
- a quality assurance structure; and
- a method for identifying cancer occurrence in the target population
Figure 46  The cervical cancer screening process
From Zapka et al. (2003)
country affects the potential effectiveness of a cervical screening programme, in particular if only part of the service is free of charge or covered by insurance (state or other). Further influences on programme effectiveness include the accessibility of services in poorly developed health-care systems and the way that information about the programme is conveyed to the target population.

**Europe**

Among 38 European countries, 25 are member states of the European Union, which includes all western European countries except Iceland, Norway and Switzerland, and (as of 2004) also several eastern European countries. The European Union includes 450 million inhabitants. Most data on the use of cervical cancer screening are available from western European countries. Despite its relatively good level of resources, Europe has rather few national well organized and documented programmes. In most European countries, cervical cancer screening started as an opportunistic activity, performed on the initiative of women or doctors. This opportunistic screening activity is still predominant in most European countries.

The European Union has recommended cervical cancer screening since the start of the Europe against Cancer programme in 1987. European guidelines for quality assurance in cervical cancer screening were issued in 1993 (Coleman et al., 1993). A Council recommendation in 2003 stressed the need for the adoption of organized screening programmes with personal invitations and quality assurance (Boyle et al., 2003; European Commission, 2003).

Table 46 lists European countries that have organized screening programmes. Nationwide programmes with personal invitation started in the 1960s and 1970s in Iceland, Finland, Sweden, Denmark and Latvia. However, participation in the Latvian programme decreased after 1987. In the United Kingdom, a computerized call/recall system was established in 1988. This is a system that invites women who are registered with a GP, keeps track of any follow-up investigation and, if all is well, recalls the woman for screening in three or five years time. National coordination and quality assurance was adopted in 1995. In the Netherlands, local programmes existed from the 1970s and a national organized programme was set up in 1996. A national programme started in Norway in 1995. A programme for the Flemish Region of Belgium including about 60% of the national population started in 1994. In Italy, a few local programmes started in the late 1980s or early 1990s, and national guidelines recommended organized programmes on a regional basis in 1996. In 2002, 12 regions out of 20 had programmes targeting in total about half of Italian women. Nationwide programmes were very recently started in Hungary and Slovenia. In the other countries listed, organized programmes are mainly pilot programmes and cover a small percentage of the national population. In Germany, a committee is currently considering the possibilities for establishing organized screening.

**Screening test**

Cytological (Pap smear) testing is generally used. A combination Ayre’s spatula and brush or an extended-tip spatula is commonly used for sampling, although in Germany most smears are reported to be taken with a cotton-swab (Schenk & von Karsa, 2000). In Germany, a gynaecological examination is also a mandatory part of screening, while colposcopy is left to the discretion of the physician (Schenk & von Karsa, 2000). Colposcopy, although not recommended, is still quite common in opportunistic screening in Italy (Segnan et al., 2000).

Most smears are fixed on glass by the smear-taker. However, in the United Kingdom (National Institute for Clinical Excellence, 2003) and Denmark (Hoelund, 2003; Patologifadellen, Hvidovre Hospital, 2003), the screening programmes are changing to liquid-based cytology. The smears are read by cytotechnicians. In Finland, a trial of the use of neural-network-assisted screening (Papnet) is in progress (Nieminenn et al., 2003). Since 1999, the largest screening programme in Denmark has used automated reading (Focal Point; Patologifadellen, Hvidovre Hospital, 2003). In Denmark, national guidelines have been issued on the use of HPV testing in assessment of women with atypical cytological results (Kjaer et al., 2002a). Trials on the use of HPV testing for primary screening are under way in Finland (Nieminenn et al., 2003), Italy (Ronco et al., 2004), the Netherlands (Bulkmans et al., 2004), Sweden (Dillner, 2000) and the United Kingdom (Cuzick et al., 2003).

**Screening interval and age**

European guidelines (European Commission, 2003) recommend three- to five-year screening intervals, depending on the resources available, but there is wide variation in actual recommendations at the national level across Europe. The most commonly recommended interval between normal cytological tests is three years (Table 47). Five-year intervals are recommended in Finland, Ireland and the Netherlands. A three- to five-year interval used to be recommended in the United Kingdom, but recently a three-year interval was recommended for women aged 25–49 and five years for women 50–64 years old. These recommendations were based on an audit of screening histories (Sasieni et al., 2003). The recommended interval is
## Table 46. Organized programmes in Europe

<table>
<thead>
<tr>
<th>Country</th>
<th>Population included in organized cervical screening</th>
<th>Start</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Regional programme in Vorarlberg (120 000 women, 4% of Austrian women 20+)</td>
<td>1970</td>
<td>Breitenecker <em>et al.</em> (2000)</td>
</tr>
<tr>
<td>Belgium</td>
<td>Flemish Region (about 60% of Belgian population)</td>
<td>1994</td>
<td>Arbyn &amp; Van Oyen (2000)</td>
</tr>
</tbody>
</table>
| Denmark | Regional programmes  
| France | Pilot programmes in 4 departments: Bas-Rhin, Isère and Doubs (500 000 women) and Martinique | 1990 (Isère)  
1991 (Martinique)  
| Greece | Pilot programmes in Ormylia (< 20 000 women)  
and Ilia and Messinia Region | | Linos & Riza (2000) |
| Hungary | Nationwide | 2003  | Döbrössy & Bodo (personal communication) |
| Iceland | National programme | 1964 Reykjavik  
| Ireland | Pilot programme in Mid Western Health Board Region (67 000 women) | 2000  | O’Neill (2000) |
| Italy | Regional programmes in 12 of 20 regions targeting 52% of women aged 25–64 | Most after 1996 (where 13% of population targeted). | Segnan *et al.* (2000); Ronco *et al.* (2003a) |
| Latvia | National programme | 1972. After 1987 the programme gradually disappeared due to economic and political factors | V. Gruunberg (personal communication) |
| Netherlands | Regional programmes  
Nationwide with national co-ordination and policy | 1970 local programmes and opportunistic screening  
| Norway | National programme  
National coordination and policy | 1959 first pilot  
1980s - 1990s many opportunistic smears  
1995 national programme | Messelt & Høeg (1967); Nygård *et al.* (2002) |
| Portugal | Regional programme in central Portugal (300 000 women) | 1990  | Real *et al.* (2000) |
Use of screening for cervical cancer

The European guidelines recommend screening for cervical abnormalities "starting at the latest by the age of 30 and definitely not before the age of 20" (European Commission, 2003). In the updated European Code against Cancer, this has been phrased as "Women from 25 years of age should participate in cervical cancer screening" (Boyle et al., 2003). However, wide variation is also seen here in what is recommended at the national level. Most European countries recommend screening from age 25 up to age 64 or 65. The organized programmes in Finland and the Netherlands target women aged 25 to 60 years. In Germany and Austria, women aged 20 or older are eligible for annual cytology, and in Luxembourg those aged 15 or older are eligible.

The combination of differences in the recommended age group and in screening interval results in dramatic differences in the number of recommended lifetime smears, from 6–8 in Finland, Ireland and the Netherlands, 12–18 in most European countries, up to 50 or more in Austria, Germany and Luxembourg.

**Invitations**
A call/recall system based on personal invitations is considered to be a key element of an organized programme in Europe. For this purpose, an accurate list of the target population with names and addresses is needed. Sources of such lists vary between countries and include population registries, health service registers, general practitioners’ (GPs) medical files, electoral registers and others.

Usually, only women who are not registered as having had a cytological test within the recommended interval are invited. This ‘integrated’ approach is applied with the intention of saving resources by avoiding re-screening of recently tested women (Coleman et al., 1993). It requires comprehensive registration of cytological testing, including opportunistic tests, at the population level. In some Italian programmes, all women are invited independently of their screening history (Ronco et al., 1998). This approach may be used if cytology registration is incomplete or if it is hoped to modify the spontaneous frequency of screening. In Finland, the organized programme invites all women (Niemenen et al., 1999); until the 1990s, all smears from the organized programme were analysed in laboratories run by the Cancer Society of Finland (Niemenen et al., 2002).

The nature of the invitation may vary from a suggestion to contact a smear-taker to a pre-assignment of a modifiable place and date. In randomized trials (Wilson & Leeming, 1987; Pierce et al., 1989; Segnan et al., 1998), compliance rates were significantly higher with letters offering pre-allocated appointments than with open-ended invitations.

Outside organized programmes, no systematic active personal invitation is sent. In Germany until 1995, statutory insurers used to issue yearly vouchers for reimbursement to all eligible women, which also served as a reminder.

Information campaigns via mass media are implemented both in areas covered by invitational programmes and in areas not covered.
<table>
<thead>
<tr>
<th>Country</th>
<th>Age group</th>
<th>Interval between normal tests (years)</th>
<th>Lifetime number of recommended smears</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>20+</td>
<td>1</td>
<td>50 (to age 70)</td>
<td>Breitenecker et al. (2000)</td>
</tr>
<tr>
<td>Denmark</td>
<td>23–59</td>
<td>3</td>
<td>13</td>
<td>Sundhedsstyrelsen (1986)</td>
</tr>
<tr>
<td>Finland</td>
<td>30–60</td>
<td>5</td>
<td>6</td>
<td>Anttila &amp; Nieminen (2000)</td>
</tr>
<tr>
<td>France</td>
<td>25–65</td>
<td>3</td>
<td>14</td>
<td>Schaffer et al. (2000)</td>
</tr>
<tr>
<td>Germany</td>
<td>20+</td>
<td>1</td>
<td>50 (to age 70)</td>
<td>Schenk &amp; von Karsa (2000)</td>
</tr>
<tr>
<td>Greece</td>
<td>25–64 (in pilot)</td>
<td>3 after 2 negative smears (Omylia), 2 (Ilia/Messinia)</td>
<td>15 (Omylia) 21 (Ilia/Messinia)</td>
<td>Riza et al. (2000)</td>
</tr>
<tr>
<td>Hungary</td>
<td>25–65 (previously 18+)</td>
<td>3 after 2 negative smears (from 2003, before 1)</td>
<td>15</td>
<td>Döbrössy &amp; Bodo, (personal communication)</td>
</tr>
<tr>
<td></td>
<td>25–69 (1970–87)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20–69 (1988–)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>25–64</td>
<td>3</td>
<td>14</td>
<td>Segnan et al. (2000)</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>15+</td>
<td>1</td>
<td>55 (to age 70)</td>
<td>Scheiden et al. (2000)</td>
</tr>
<tr>
<td>Norway</td>
<td>25–69</td>
<td>3</td>
<td>15</td>
<td>Nygård et al. (2002)</td>
</tr>
<tr>
<td>Portugal</td>
<td>20–64</td>
<td>3 after 2 negative smears</td>
<td>17</td>
<td>Real et al. (2000)</td>
</tr>
<tr>
<td>Romania</td>
<td>25–65</td>
<td>3</td>
<td>14</td>
<td>Suteu et al. (2003)</td>
</tr>
<tr>
<td>Slovenia</td>
<td>20–64 (formerly 20+)</td>
<td>3 after 2 negative smears (from 2003, before 1)</td>
<td>17</td>
<td>Primic Zakelj (personal communication)</td>
</tr>
<tr>
<td>Spain</td>
<td>35–64, below 35 if with risk factors. Most regions women aged 25–65</td>
<td>5 after 2 negative smears, but 3 in organized programme</td>
<td>14 (age 25–65, 3 year)</td>
<td>Asuncze Elizaga et al. (1993); AETS (2002); Fernandez Calvo et al. (2000)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>20 (or 25 from 2004)–64</td>
<td>3 for age 25–49 5 for age 50–64 (From 2004, before 3–5)</td>
<td>12</td>
<td>Patrick (2000) NHS (2003a)</td>
</tr>
</tbody>
</table>
GPs frequently act as advisers both in the presence and in the absence of personal invitations. Attendance following invitation was higher with letters signed by the GP than with letters signed by programme staff (Segnan et al., 1998; Palm et al., 1993). In the United Kingdom, GPs are paid for screening based on the coverage among their patients. This was introduced to increase coverage (Rudiman et al., 1995).

The facilities for testing and the professionals involved vary widely between countries and between organized and opportunistic activity (Table 48). In the Netherlands and the United Kingdom, smears are most commonly taken by GPs or their assistants, while gynaecologists play a major role in health nurses and midwives. In the Finnish organized programme, the smears were taken at maternity and child health centres invariably by public health nurses and midwives. In the Italian organized programmes, smears are most often taken by midwives in family-planning clinics, and midwives also participate in Finland and Sweden.

Coverage and participation
Table 49 shows published estimates of participation by cervical screening at the national level. This measure has been provided at a national level only in Finland, Iceland, Norway, England and the Netherlands. For France, an estimate has been made on the basis of individual linkage on a sample of women, using insurance data. For a number of other countries, estimates are based on interviews, with the possibility of recall bias, although some kind of validation was frequently performed. Comparability of the findings is also limited by differences in the age groups considered, and by the fact that hysterectomized women were excluded in some case (e.g., England) but not in others (frequently not mentioned).

Participation over 80% is seen in England and Iceland and in rural areas of Sweden and Denmark. Participation of 70–80% is found in the Flemish part of Belgium, Finland, the Netherlands, Norway and Copenhagen, Denmark. Participation is around or below 60% in Austria, France, Italy and Spain. In Germany, where women are eligible for yearly screening, the number of tests in 1996 was about 50% of the number of women (Schenck & von Karsa, 2000). A three-year participation of 65% was estimated for the European Union (women aged 25–54 years) on the basis of an interview survey in 1991 (Coleman et al., 1993).

Participation also varies between areas within countries. This variation was quite low in England, with participation ranging from 76% in London to 85% in the East Midlands in 2002–03 (NHS, 2003a). However, in Spain, the reported participation (women 40–70 years) ranged from 25% in Castilla-La Mancha to 61% in Madrid (AETS, 2002). In Italy, there was a strong gradient in participation from northern-central (53–61%) to southern Italy (26%) (Ronco et al., 2003a; Mancini et al., 2004). In France, the annual number of tests ranged from 17 to 39 per 100 women, with a north–south and west–east increasing trend (Rousseau et al., 2002).

There is also a difference in activity by age, with a common pattern of lower activity at the highest ages. In England in 2002–03, participation was over 80% among women aged 30–59 years, 74% among those aged 25–29, and 77% among those aged 60–64 (NHS, 2003a). In Spain, participation decreased from 61% in the 40–45-year age group to 31% in the 61–65-year age group (AETS, 2002). In Flanders (Belgium), participation remained high up to 40 years of age, and decreased thereafter (Arbyn et al., 1997). In Italy, it was 27% at age 25–34, over 50% at age 35–44, and 43% at age 55–64 (Mancini et al., 2004). In France, the rate of activity increased slightly up to age 50–54, and then decreased rapidly (Rousseau et al., 2002).

In England, the introduction of the computerized call/recall system in 1988 and the target payments for GPs in 1990 increased the five-year participation for women aged 25–64 years from 40% in 1988 (Havelock et al., 1988; Shroff et al., 1988; Robertson et al., 1989b) to persistently over 80% between 1992 and 2003 (NHS, 2003b). In Norway, opportunistic screening has been very common, but when organized screening with personal invitations was introduced in 1995 the participation increased from 65% to 71% (Nygård et al. 2002). In France, a three-year participation of 69% was found in the organized programme of Bas-Rhin after four years of activity (Fender et al., 2000), compared with the national estimate of 54% (Rousseau et al., 2002). In Castilla y Leon, a region of Spain with an organized programme, the estimated three-year participation of 41% was similar to the 44% for Spain (AETS, 2002). In Italy, national participation data for 1999–2000 were only marginally influenced by the recently started organized programmes, although in one of the latter, the three-year participation was estimated to be 74%, compared with 43% before (Ronco et al., 1997). A strong reduction in variability by age, education and marital status was also observed.

Excess use of Pap testing
Pap testing is unevenly distributed among women in many countries, with many women not screened at all or not screened within the recommended interval, and other women screened more frequently than recommended.

In general, over-testing is assumed to be high in opportunistic screening. The level of testing is high in Germany, where women are eligible for yearly
<table>
<thead>
<tr>
<th>Country</th>
<th>Smear-taker</th>
<th>Communication of cytology results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Gynaecologists</td>
<td>Mail or phone to the smear taker</td>
</tr>
<tr>
<td>Belgium</td>
<td>Gynaecologists/GPs</td>
<td>Report to the smear taker</td>
</tr>
<tr>
<td>Denmark</td>
<td>GPs</td>
<td>Mostly: woman asked to call GP</td>
</tr>
<tr>
<td>England</td>
<td>GPs or general practice nurses</td>
<td>Report to the smear taker</td>
</tr>
<tr>
<td>Finland</td>
<td>Midwives or public health nurses</td>
<td>Letter to woman</td>
</tr>
<tr>
<td>France</td>
<td>Gynaecologists/GPs</td>
<td>Not specified</td>
</tr>
<tr>
<td>Germany</td>
<td>Office-based gynaecologists (90%) and GPs (10%)</td>
<td>By the smear taker</td>
</tr>
<tr>
<td>Greece</td>
<td>Organized: Gynaecologists (Ormylia) gynaecologists, trained rural doctors and midwives (Ilia/Messinia) Opportunistic: gynaecologists (Riza et al., 2000)</td>
<td>Organized: letter directly to the woman</td>
</tr>
<tr>
<td>Iceland</td>
<td>Gynaecologists/GPs (Sigurdsson et al., 1991)</td>
<td>Not specified</td>
</tr>
<tr>
<td>Ireland</td>
<td>GPs, family planning and community clinics, hospitals</td>
<td>Letter to woman</td>
</tr>
<tr>
<td>Italy</td>
<td>Organized: mainly midwives in family planning clinics Opportunistic: mainly gynaecologists</td>
<td>Organized: mostly letter directly to the woman</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>GPs and/or gynaecologists</td>
<td>Not specified</td>
</tr>
<tr>
<td>Netherlands</td>
<td>GPs and their practice assistants</td>
<td>Via the GP</td>
</tr>
<tr>
<td>Norway</td>
<td>Primary physicians (Krogh &amp; Malterud, 1995)</td>
<td>Not specified</td>
</tr>
<tr>
<td>Portugal</td>
<td>GPs (organized)</td>
<td>Organized: letter via the GP</td>
</tr>
<tr>
<td>Spain</td>
<td>Organized: family doctors Opportunistic: mainly gynaecologists (&gt;80%) and family clinics of primary care centres (about 20%) (AETS, 2000)</td>
<td>Organized: letter via the primary care physician</td>
</tr>
<tr>
<td>Sweden</td>
<td>Nurse-midwives (Sarkadi et al., 2004)</td>
<td>Letter to woman</td>
</tr>
</tbody>
</table>

Modified from Linos & Riza (2000)
screening; the number of smears in 1996 was about 50% of the number of women in the target population. In Italy, 52% of screened women reported having a test every year (Mancini et al., 2004).

Over-testing in countries and regions with organized screening programmes depends on the organization of the programme. The total level is often high in countries where the organized programme runs independently of opportunistic activity. In Finland, 200 000 smears are taken annually within the organized programme and 400 000 smears are taken outside (Finnish Cancer Registry, 2003). In an Italian programme in which all women are invited independently of previous testing, 20–25% of those who joined the programme also had tests outside the protocol (Ronco et al., 1997). Sweden, where there were regional differences in the organizational set-up, over-testing was heavy in 1994, with 292 000 smears taken in the organized programme and 656 000 taken outside. Opportunistic testing was free, whereas a small fee had to be paid for the organized screening (Dillner, 2000).

<table>
<thead>
<tr>
<th>Country</th>
<th>Estimated coveragea</th>
<th>Source</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Lifetime:</td>
<td>Interviews with sample of women</td>
<td>Vutuc et al. (1999) (reported in Breitenecker et al., 2000)</td>
</tr>
<tr>
<td></td>
<td>60% 2+ smears</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10% 1 smear</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30% never</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>74% in 3 years, Flemish region, 64% in 3 years, Walloon region</td>
<td>National health interview survey</td>
<td>Arbyn &amp; Van Oyen (2000)</td>
</tr>
<tr>
<td>Denmark</td>
<td>85% County of Funen</td>
<td>Register linkage</td>
<td>Hoelund (2003); Patologiøndelungen, Hvidovre Hospital (2003)</td>
</tr>
<tr>
<td></td>
<td>73% Copenhagen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>England</td>
<td>81% in 5 years, 71% in 3 years, 2003</td>
<td>Register linkage</td>
<td>NHS (2003b)</td>
</tr>
<tr>
<td>Finland</td>
<td>70% (organized screening)</td>
<td>93% (all smears)</td>
<td>Finnish Cancer Registry (2003)</td>
</tr>
<tr>
<td></td>
<td>54% in 3 years, 1998–2000</td>
<td>Registration of smears in two health insurance systems and linkage for sample of 9374 women</td>
<td>Rousseau et al. (2002)</td>
</tr>
<tr>
<td>Germany</td>
<td>42–47% in 1997 for women aged 25–54</td>
<td>Personal interviews</td>
<td>Kahl et al. (1999)</td>
</tr>
<tr>
<td>Iceland</td>
<td>83% in 1990–92</td>
<td>Register linkage</td>
<td>Sigurdsson (1999)</td>
</tr>
<tr>
<td>Italy</td>
<td>50% reporting usual frequency of 3 years, 1999–2000</td>
<td>National periodic survey on health (44 433 women)</td>
<td>Mancini et al. (2004)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>About 80% in 2 years, 1996–97</td>
<td>Register linkage</td>
<td>Van Ballegooijen &amp; Hermens (2000)</td>
</tr>
<tr>
<td>Spain</td>
<td>44% reporting one or more tests in 3 years for preventive reasons</td>
<td>Personal interview, random sample (2409 women)</td>
<td>AETS (2002)</td>
</tr>
<tr>
<td>Sweden</td>
<td>&gt;80% northern Sweden, 20–30% Malmö, 50–70% most common</td>
<td>Register linkage</td>
<td>Dillner (2000)</td>
</tr>
</tbody>
</table>

a Proportion of the target population having had at least one test in the defined interval
2000). Since a government report examined this issue (Socialstyrelsen, 1998), all counties have changed to the ‘integrated approach’ (see above) for invitations.

There can be a lower level of over-testing in integrated programmes, although this is not always the case. In England, GPs are paid based on coverage of their patients and not on the number of smears taken. Data on over-testing are not published, but the level seems to be small. In 2002–03, a total of 4.1 million smears were taken. There were 13.8 million women in the target 25–64-year age group, of whom 66.1% were screened within the last three years, giving 3.0 million annual smears used for screening. An additional 1.0 million women were recalled more often than every third year for surveillance etc. This leaves very few smears corresponding to possible ‘over-use’ (NHS, 2003b). In Norway, the average number of smears per woman aged 25–69 in a three-year period decreased from 1.68 to 1.52 when an integrated programme was introduced, at the same time increasing participation (Nygård et al., 2002). Denmark runs integrated programmes, and the level of over-testing was 28% in Copenhagen in 1999–2001 (Patologifaelingen, Hvidovre Hospital, 2003), but only 4% in Fyn in 1999 (Hoelund, 2003). In the Netherlands, out of about one million smears taken in 1996, 450 000 were in the organized programme, 300 000 were other ‘primary’ smears, and 250 000 ‘secondary’ (follow-up or repeat) smears (van Ballegooijen & Hermens, 2000). In Bas-Rhin, France, 63% of women had a second test before the recommended interval (Fender et al., 2000).

The difficulties in limiting over-testing are illustrated by the fact that 27–29% of female primary physicians in Norway in 1995 recommended screening more often that the three years stated in the national guidelines (Krohg & Malterud 1995). In Stockholm, a common practice among private gynaecologists seems to be to have an annual appointment with private patients and a Pap test is often part of the consultation (Sarkadi et al., 2004).

Cytological interpretation and management of abnormal results

There is no unique European system of classification. National classification systems are applied in the Netherlands and in the United Kingdom. In Germany (where standardized national reporting forms exist) and in Austria, the Munich classification is used (Schenk & von Karsa, 2000; Breitenecker et al., 2000). In Italy, the Bethesda system is widely applied, although with many local adaptations (Ronco et al., 1998). A standard reporting protocol, related to the Bethesda system, is applied in the Flemish Region of Belgium (Arbyn & Van Oyen, 2000). Tables of ‘equivalent terminology’ between different classifications have been published in the European Guidelines (Coleman et al., 1993). Data on comparability of the criteria actually used in different countries are limited. In a study of agreement in cytological interpretation conducted in six Italian laboratories and one Danish using the Bethesda 1991 classification, agreement between the Danish and the Italian results was similar to that within Italians (Ronco et al., 2003b).

Table 48 summarizes the methods of communication of test results to the woman. In some cases the report is sent directly to the woman (e.g., Italian organized programmes), while more frequently the laboratory reports to the smear-taker. The practice of sending negative test results directly to the women was abandoned in one Danish county after a survey showed that women preferred to have the results reported via their GP (Andreasen et al., 1998). In most countries, the decision on the action to be taken following a non-negative smear is left to the smear-taker (Belgium, Germany), while in others (England) the recommendation is given by the laboratory and it is the responsibility of the smear-taker to ensure that the woman receives the result.

Criteria for management of women based on cytology results vary widely, also depending on the cost and availability of colposcopy facilities. National guidelines with implementation policies are applied in England and the Netherlands. National guidelines for the management of abnormal smears exist in France (ANAES, 1998), Austria (stated in Breitenecker et al., 2000) and Germany (Bundesärztekammer, 1994; Schenk & von Karsa, 2000). In Italy the national guidelines recommend development of detailed local protocols for the management of abnormal results (Ronco et al., 1998). Recommendations have also been prepared in the Flemish programme (Arbyn & Van Oyen, 2000).

In the United Kingdom, women with moderate and severe dyskaryosis (equivalent to HSIL in the Bethesda system) are referred for colposcopy, while those with mild dyskaryosis (Bethesda: LSIL) and borderline cytology (Bethesda: ASCUS) are advised to repeat testing and referred for colposcopy only in case of persistence, although some of these cases are in fact directly referred for colposcopy. The same policy is applied in Belgium and the Netherlands. In France, a choice between colposcopy and repeat cytology is left for borderline and low-grade lesions. In Portugal, women with ASCUS are advised to undergo repeat testing, while those with LSIL or worse are referred for colposcopy. In Italy, most organized programmes refer all women with ASCUS or more severe cytology for colposcopy, although some programmes perform a repeat cytology in the case
of ASCUS (but not LSIL). Referral of all ASCUS-positive women for colposcopy is also the usual practice in opportunistic activity in Italy. In Finland, treatment is given to women with low-grade dysplasia, whereas in Norway a more conservative stance is taken and treatment is given only after three repeated low-grade abnormal results (Nygård et al. 2002). In Denmark, follow-up of women with atypia varies from direct referral to colposcopy to a repeat test within 6–12 months (Bigaard et al., 2000).

An issue relevant to screening effectiveness is that all women needing further action (repeat or colposcopy) should actually have it. Fail-safe methods are implemented in most organized programmes. Monitoring of follow-up is applied in French organized programmes (Schaffer et al., 2000). In the Netherlands, it is the responsibility of the GP to inform the woman of the results and to ensure completion of follow-up. It has been planned that laboratories should provide GPs with lists of women with incomplete follow-up (van Ballegooijen & Hermens, 2000). A similar system is implemented in Denmark (Patologi-afdelingen, Hvidovre Hospital, 2003). In Finland and in most Italian organized programmes, women needing colposcopy receive a pre-arranged appointment to a reference centre and are reminded in case of default.

Quality assurance
Several factors contribute to the impact of cervical screening, including coverage of the targeted population, the actual participation, the quality of smear-taking and interpretation, the follow-up of women with abnormal results, the quality of diagnostic procedures and initial treatment. Several initiatives have been taken to develop and promote quality assurance systems, and guidelines have been published both at European and national levels.

Quality assurance addresses the following issues: rules concerning structural requirements (e.g., number of smears interpreted, qualification and training of staff), procedures to be followed (e.g., methods of smear preparation, re-interpretation of smears, including guidelines for the management of women); and monitoring of performance for taking action in response to such information.

National guidelines concerning cytological interpretation exist in Austria (reported in Breitenecker et al., 2000), England (Johnson & Patnick, 2000), France (Marsan & Cochand-Priollet, 1993), Germany (Bundésärztekammer, 1994; reported in Schenk & von Karsa, 2000) and the Netherlands (van Ballegooijen & Hermens, 2000). Guidelines for quality assurance in colposcopy were published for England (Luesley, 1996) and the Netherlands (Helmerhorst & Wijnen, 1998). Poor compliance with guidelines in Austria has been reported (Breitenecker et al., 2000).

Rules or guidelines concerning the number of smears to be read exist in many countries both as a maximum number per cytologist (Austria, England, Germany, the Netherlands) and as a minimum (Denmark, England, Italy). Laboratories vary greatly in size and in some countries many are small. In the Flemish Region of Belgium, a total of 620 000 smears were processed in 1993 in over 100 laboratories (Arbyn & van Oyen, 2000). In Austria, the annual number of smears per laboratory varied from 3000 to 150 000, with an average of 25 000 (Breitenecker et al., 2000). In Germany, an annual total of 17 000 000 smears are interpreted by some 2000 laboratories (Schenk & von Karsa, 2000). In the Netherlands, the annual number varies from 5000 to over 50 000 per laboratory (van Ballegooijen & Hermens, 2000). Among Italian laboratories involved in organized programmes in 1997, the workload varied from 3000 to over 50 000 smears per year (Ronco et al., 1998).

Proficiency testing for cytological interpretation is compulsory for all laboratory staff in the United Kingdom, and for cytopathologists (but not for cytotechnicians) in Germany. In many other countries, proficiency testing is encouraged but not compulsory.

Re-screening of a sample of negative smears, which is mandatory in the USA, is not compulsory in most European countries. A rapid review (Faraker & Boxer, 1996) of all negative smears is mandatory in England (NHSCSP, 2000) and this policy is supported by a meta-analysis (Arbyn & Schenk, 2000). Suspicious smears are usually reviewed by a supervisor. In Italy, where quality assurance programmes are decided on a regional basis, circulation and discussion of sets of smears is becoming widely applied on both a local and national basis in order to improve consistency between laboratories (Branca et al., 1998; Ronco et al., 2003b).

Data registration
Monitoring of screening performance requires comprehensive registration of all events related to screening and the recovery and linkage of the data at an individual level in order to allow reconstruction of the screening histories and their results. European guidelines (Coleman et al., 1993) recommend comprehensive registration of all cytological and histological findings.

Registration of events related to screening, particularly cytological results, exists in the organized programmes listed in Table 46 and in other areas. Individual linkage of data takes place in Denmark, Finland, Iceland, the Netherlands, Norway, Sweden, the United Kingdom, and in the organized programmes in Italy. In Germany, individual linkage is prohibited due to regulations on privacy.
Where results are registered, the level of comprehensiveness varies. Comprehensive computerized registration of all tests is performed in the Netherlands through a national database covering all pathology units (the PALGA system). In the United Kingdom, registration of cytology is comprehensive (data are registered locally and forwarded periodically for central analysis) and highly complete, leaving out only a small percentage of privately performed tests.

Screening registers with comprehensive registration of cytology have also been set up in the French organized programmes through agreements with the involved parties (laboratories, GPs, gynaecologists) and in the Flemish Region in Belgium. In the latter case, personal identification codes are encrypted. In other areas, however, complete registration only of cytology and histology taken within the organized programme is possible, while registration of opportunistic test results and of histology performed outside the reference centres is incomplete or absent. This is the case for some Italian programmes, where completeness also depends on the number of laboratories in the area and on the amount of private activity.

In France, outside organized programmes, computerized registration of cytological testing is performed at the national level by the social security system, although mainly for administrative purposes. Individual linkage has been performed experimentally on a sample of women in order to estimate participation (Rousseau et al., 2002). In Germany, cytology reports are registered on paper, transmitted to regional insurance billing offices, and finally registered on a central computerized database. Only results of initially abnormal test results and of a random sample of normal ones are registered. Follow-up and histology data after an initially abnormal result are reported on the same sheet. However, problems of quality and completeness have been reported. No special registration of colposcopies or biopsies exists (Schenk & von Karsa, 2000). Colposcopies performed in referral centres are recorded by most Italian programmes and partially in England.

Performance indicators
The European Guidelines proposed a number of standard tabulations and parameters (coverage, interval to reporting, proportion of unsatisfactory smears, treatment compliance, sensitivity and specificity, distribution of invasive cancers, interval cancers) for ‘short-term monitoring’ of programmes (Coleman et al., 1993). In England, a national system of measurements and of reference standards for them, each related to an objective, was adopted (NHSCSP, 1996). Annual reports, produced in both a synthetic (NHS, 2003b) and detailed (NHS, 2003a) format, are available on the NHS screening web site (http://www.cancerscreening.nhs.uk). In Italy, a system of process indicators (partly with reference standards) has been published (Ronco et al., 1999) and included in official guidelines. Annual surveys of the activity of organized screening programmes have been conducted from 1998 and a report published annually from 2002 (Ronco et al., 2002). In the Netherlands, regular reports are produced on the outcome of the organized programme. Regular reports have been published from the Icelandic (Sigurdsson, 1995), Norwegian (Nygård et al., 2002) and Finnish (Finnish Cancer Registry, 2003) programmes. Process measures have been published for French organized programmes (Fender et al., 2000; Schaffer et al., 2000) and the distribution of cytology results for the Flemish Region has been reported (Arbyn & van Oyen, 2000).

Participation is a key indicator and has been described in a previous section. Another important indicator is the proportion of unsatisfactory cytological tests. The proportion is high in England, at 9.4% of smears (NHS, 2003b), plausibly as a result of strict criteria for adequacy. In Norway, 4.7% of smears were considered unsatisfactory in 1998–2000 (Nygård et al., 2002). The Netherlands has 1% (van Ballegooijen & Hermens, 2000), Finland 0.004% (Finnish Cancer Registry, 2003), Flemish programmes 0.6–1.0% (Arbyn & van Oyen, 2000), the French programmes 0.12–2% (Schaffer et al., 2000) and the Italian organized programmes 3.8% (Ronco et al., 2003a). In Copenhagen in 1999–2001, 2.5% of smears were unsatisfactory, but 8.5% were normal without endocervical cells. In the latter case, the GP decides whether testing should be repeated (Patologiatdellingen, Hvidovre Hospital, 2003). In Funen in 1999, 7.5% of smears were unsatisfactory, including smears without endocervical cells. This percentage decreased to 2.5% after introduction of liquid-based cytology (Hoelund, 2003). In many organized programmes, reports on unsatisfactory smears are sent to smear-takers.

A measure of the proportion of women referred for further action (repeat cytology or colposcopy) is obviously useful as an indication of the human and economic cost of screening. More frequently, the proportion of abnormal cytological results (or of screened women with abnormal results) is reported. However, this does not translate immediately to referral or repeat action, both because guidelines leave choice for some diagnoses (e.g., LSIL/ASCUS in France) and because of referral for clinical reasons. Cyto-histological correlation data are frequently reported, sometimes in terms of positive predictive value. However, comparison is difficult because of the
variability of criteria for inclusion, both in relation to the cytological diagnoses considered (frequently only certain cytological categories among those of referred women are included) and to the presence of histology (in some cases but not others, women examined colposcopically but without histology are included, assuming that no biopsy was done because no suspect area was identified at colposcopy). It is nevertheless clear that the proportion of screened women immediately referred for colposcopy varies between countries from 0.8% in Finland (Finnish Cancer Registry, 2003) to 2.9% in Italian organized programmes (Ronco et al., 2003a).

Some measure of completeness of follow-up is also reported or can be computed from available data, although sometimes the data relate to colposcopy only (Italy) and sometimes to either colposcopy or repeat testing (France). Statistics on the time between referral and attendance for colposcopy are computed in England.

United States and Canada

The type of organization of cervical cancer screening services in the USA and Canada covers the range from opportunistic screening, with access based on availability of individual or third-party financial resources, to organized screening at the local, regional and national level funded by work-based groups or government agencies.

Organization and financing

United States

In the USA, cervical cancer screening is provided in various settings: private practices, public health clinics, community health centres, sexually transmitted disease clinics, family planning clinics and prenatal clinics. This screening is offered on an entirely opportunistic basis. Financing for cervical cancer screening and other preventive services depends on a woman’s personal resources and/or health insurance coverage. Most insurance plans cover cervical cancer screening services, but if follow-up testing is necessary, there may be cover only for part of the expenses.

In the Government-sponsored Medicare and Medicaid, the proportions paid by individuals, if any, are limited by law and are often related to income levels. Medicare provides reimbursement for screening services of individuals aged 65 years or more and some younger disabled individuals. Medicaid, administered by states, provides reimbursement for very low-income families with highly limited resources. Persons covered by some type of insurance, public or private, represent a median of 84% of the state populations (Mansley et al., 2002) and there is some geographical variation in the availability of insurance. For example, the State of Wisconsin has coverage for 91% of its constituents whereas the State of Texas has coverage for only 76% of its residents (Mansley et al., 2002). Populations with lower socioeconomic status are more likely to have no or insufficient insurance coverage (Henson et al., 1996).

Special programmes like the National Breast and Cervical Cancer Early Detection Program (NBCCEDP), administered by the Centers for Disease Control and Prevention (CDC), make cancer screening services available to uninsured or underinsured women who meet certain income and family size criteria but are not eligible for other government reimbursement programmes. Funding is available to provide screening and certain follow-up services for only 6% of women eligible for this programme, aged 18–64 years. In the absence of a structured national health care registration system, women are informed about and recruited into the NBC-CEDP by a variety of means including the media, community- and religion-based organizational health fairs and public health announcements. However, the first contact with the programme is initiated voluntarily by the woman when she applies for eligibility. Since 2001, many professional organizations and government agencies (CDC, the Institute of Medicine–US Preventive Services Task Force (USPSTF) and the National Institutes of Health (NIH)) have deliberated on the features of cervical cancer screening in the USA (see Table 50). After extensive review of published evidence and consensus building among the various groups, updated recommendations have been published for cytological testing and follow-up of women with abnormal results. New screening recommendations include changes in the age to begin screening (previously 18 years, now 21), frequency of repeat screening if results are negative (previously annual, now up to every three years if over age 30), and the age to consider ending screening (previously no recommendation, now age 70 if recent screening results are negative) (Saslow et al., 2002; American College of Obstetricians and Gynecologists, 2003; US Preventive Services Task Force, 2003).

Canada

In Canada, organization and provision of health care is the responsibility of the provincial and territorial governments. The universal coverage includes cancer screening and follow-up activities. Furthermore, most provinces have cancer agencies that are usually responsible for planning, coordinating and monitoring cancer screening programmes.

Since the introduction of cytological testing in Canada, opportunistic cervical cancer screening has been the most frequently used method to screen women. In recognition of the fact that
Table 50. Recommendations for cervical cancer screening, United States, 2003

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>When to start cervical screening</td>
<td>Age 21 or within 3 y of start of sexual activity</td>
<td>Age 21 or within 3 y of start of sexual activity</td>
<td>Age 21 or within 3 y of start of sexual activity</td>
</tr>
<tr>
<td>Interval</td>
<td>Annually with conventional or every 2 y using liquid-based cytology; age &gt; 30, women with 3 negative may be screened every 2–3 y</td>
<td>Every year for women &lt; 30 or every 2–3 y for women &gt; 30 (except women with HIV, immunosuppressed or DES exposure)</td>
<td>HPV-negative, Pap-negative: every 3 y</td>
</tr>
<tr>
<td>Thin Prep</td>
<td>Recommend</td>
<td>Option</td>
<td>Insufficient evidence</td>
</tr>
<tr>
<td>HPV testing with ASCUS</td>
<td></td>
<td>Option</td>
<td>Insufficient evidence</td>
</tr>
<tr>
<td>HPV testing &gt; 30</td>
<td>Guidelines not out before FDA approval; preliminary recommend</td>
<td>Option</td>
<td>Insufficient evidence</td>
</tr>
<tr>
<td>Post-hysterectomy</td>
<td>Discontinue if for benign reasons</td>
<td>Discontinue except in special circumstances</td>
<td>Discontinue</td>
</tr>
<tr>
<td>When to stop cervical screening</td>
<td>Age 70 or 3 or more negative tests within 10-year period</td>
<td>No upper limit</td>
<td>&gt; 65</td>
</tr>
</tbody>
</table>

- Cytological screening should start at age 18 years or at initiation of sexual activity.
- A second test should, in general, be performed after one year, especially for women who begin screening after age 20.
- If the first two tests are satisfactory and show no significant epithelial abnormality, women should, in general, be advised to be re-screened every three years up to age 69.
- Screening should occur at this frequency in areas where a population-based information system exists for identifying women and allowing notification and recall. In the absence of such a system, it is advisable to repeat tests annually.
- Women over the age of 69 who have had at least two satisfactory tests and no significant epithelial abnormality in the last nine years and who have never had biopsy-confirmed severe dysplasia or carcinoma in situ can leave the cervical cytology screening programme.
- If mild dysplasia is found, a cytological test is to be repeated every six months for two years.
- If the lesion persists or progresses to moderate or severe dysplasia, the patient must be referred for colposcopy.
- More frequent testing may be considered for women at high risk (first intercourse at less than 18 years of age, multiple sexual partners, partner who has had multiple sexual partners, smoking, low socioeconomic status).
- Women do not need to be screened if they have never had sexual intercourse or have had a

---

effective organization not only reduces the cost of screening programmes but improves their effectiveness, recommendations have been made on several occasions over the years for the development of organized screening programmes that incorporate a computerized information system, population-based recruitment and effective quality management.

Summarized below are highlights of the recommendations from the 1989 National Workshop on Cervical Cancer Screening (Miller et al., 1991). These recommendations have been variably accepted and updated by provincial agencies responsible for screening.

- Cytological screening should start at age 18 years or at initiation of sexual activity.
- A second test should, in general, be performed after one year, especially for women who begin screening after age 20.
- If the first two tests are satisfactory and show no significant epithelial abnormality, women should, in general, be advised to be re-screened every three years up to age 69.
- Screening should occur at this frequency in areas where a population-based information system exists for identifying women and allowing notification and recall. In the absence of such a system, it is advisable to repeat tests annually.
- Women over the age of 69 who have had at least two satisfactory tests and no significant epithelial abnormality in the last nine years and who have never had biopsy-confirmed severe dysplasia or carcinoma in situ can leave the cervical cytology screening programme.
- If mild dysplasia is found, a cytological test is to be repeated every six months for two years.
- If the lesion persists or progresses to moderate or severe dysplasia, the patient must be referred for colposcopy.
- More frequent testing may be considered for women at high risk (first intercourse at less than 18 years of age, multiple sexual partners, partner who has had multiple sexual partners, smoking, low socioeconomic status).
- Women do not need to be screened if they have never had sexual intercourse or have had a
hysterectomy for a benign condition with adequate pathological documentation that the cervical epithelium has been totally removed and previous screening tests have been normal.

Two Canadian provinces, British Columbia and Nova Scotia, have well established, organized programmes for cervical cancer screening and in recent years other provinces such as Alberta, Manitoba, Ontario and Prince Edward Island have launched new programmes. These programmes target all women in the provincial population in a specified age range (usually 18–69 years), but no province yet has population-based recruitment. Variation between provinces in the implementation of screening programme components reflects the maturity of their programme development (Table 51) (Health Canada, 2002).

Most women who develop cervical cancer have either not been screened or have been screened too infrequently (Quality Management Working Group, 1998). About 60% of cervical cancers occur in women who have not been screened in the previous three years. Lack of organization has contributed to this failure, including an inability to reach high-risk women, inadequate quality control, or ineffective follow-up procedures.

The cost of cervical cancer screening and follow-up of abnormal findings is covered through universal provincial health funding.

Extent of use and access

United States

The extent of screening in the USA is affected by the proportion of women with some reimbursement for primary care and preventive health, and approximately 50 million cytological tests are performed annually (Kurman et al., 1994). The NBCCEDP provides approximately 250 000 of those tests. About 7% of all tests are reported to reveal an abnormality requiring further testing (Jones & Davey, 2000).

The Behavioral Risk Factor Surveillance System (BRFSS) and the National Health Interview Survey (NHIS), both administered by the CDC, show that more than 85% of women in the USA have had a previous cervical cancer screening test, and that approximately 80% have had one in the past two years (Blackman et al., 1999).

Overall, screening tends to occur more frequently among younger women (every 1–2 years under age 40) than among older women (40 years or older), who present themselves for screening services less often as they have less need for reproductive health services. According to the NHIS 2000 data, over 82% women aged at least 25 years reported having a test within the last three years; the numbers are slightly different from those from the BRFSS because of the time interval included. The groups with the lowest proportions of women who had a test within the previous three years were women without a usual source of health care (58.3%; 95% CI 55.3–61.3), women without health insurance (62.4%; 95% CI 58.1–66.8) and women who immigrated to the USA within the last 10 years (61.0%, 95% CI 55.2–66.8). Women with lower levels of education, women with limited income and women with chronic disabilities had lower levels of screening compared with other groups (Swan et al., 2003).

The burden of cervical cancer remains highest among women who are rarely or never screened, who account for an estimated 60% of newly diagnosed invasive cancers (Sung, 2001).

Canada

In Canada, access to cervical cancer screening is available to all women who meet the criteria for screening either through national or provincial programmes. Table 52 shows the frequency of self-reported cytological tests in Canada by province and age group for the period 1998–99.

More recently, the Canada Community Health Survey (CCHS), a national, biennial, cross-sectional survey, provided information on cervical cancer screening (Statistics Canada, 2002). From the survey cycle of 2000/2001, an estimated 89% of Canadian women aged 20–69 years answered "Yes" to the question "Have you ever had a Pap test?", with the highest percentage in the Atlantic Provinces (95%) and the lowest in Quebec (83%). Nationally, 53% and 73% of women aged 18–69 years reported having had a test within the last year and last three years, respectively, with the highest percentages in Nova Scotia (60% for one year; 80% for three years) and the lowest in Quebec (50% for one year and 67% for three years).

Methods of assuring quality

United States

In the 1980s, intensive media coverage of poor cytology laboratory practices and charges of lax enforcement of federal regulations contributed to the passage of the Clinical Laboratory Improvement Amendments (CLIA) in 1988 and the regulations that now define standards of cytology laboratory practice in the USA. CLIA and the related regulations serve as a baseline, through biennial inspections and certification, for assessing the quality of laboratory work including cervical cytology (Lawson et al., 1997). The regulations allow for enforcement of CLIA standards and for corrective measures when laboratories fail to meet these standards. The Centers for Medicare and Medicaid Services (CMS) in the US Department of Health and Human Services and the CDC are responsible for establishing and implementing the CLIA regulations. CMS is responsible
for enforcing the regulations, and CDC provides technical and scientific support to CMS. The CMS central office in Baltimore, Maryland, establishes CLIA programme policies and oversees and coordinates the work of 10 CMS regional offices. The regional offices are responsible for enforcing the CLIA regulations among the cytology laboratories in their jurisdictions. These regulations were amended in 1991 and have been further streamlined over the last decade. Some issues still exist regarding interpretation and review of screenings and review for false negative tests, but some improvements in cervical cancer screening technology such as thin-layer cytology, liquid-based preparations and HPV DNA testing may increase the sensitivity and specificity of testing.

To ensure the reliability of cytological tests, the following steps must be performed and monitored correctly and adequately (Lawson et al., 1997):

<table>
<thead>
<tr>
<th>Province</th>
<th>Programme</th>
<th>Year of inception</th>
<th>Computerized information system</th>
<th>Target age group</th>
<th>Screening frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newfoundland</td>
<td>No</td>
<td>–</td>
<td>✓</td>
<td>18+</td>
<td>Annual</td>
</tr>
<tr>
<td>Nova Scotia</td>
<td>Yes</td>
<td>1991</td>
<td>✓</td>
<td>18+</td>
<td>Annual</td>
</tr>
<tr>
<td>Prince Edward Island</td>
<td>Yes</td>
<td>2001</td>
<td>✓</td>
<td>20–69</td>
<td>After three normal annual tests, screening should be continued at least every 2 y</td>
</tr>
<tr>
<td>New Brunswick</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Quebec</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>18–69</td>
<td>Annual</td>
</tr>
<tr>
<td>Ontario</td>
<td>Yes</td>
<td>2000</td>
<td>✓</td>
<td>20–69</td>
<td>After three normal annual tests, screening should be continued every 2 y</td>
</tr>
<tr>
<td>Manitoba</td>
<td>Yes</td>
<td>1999</td>
<td>✓</td>
<td>18–69</td>
<td>After three normal annual tests, screening should be continued every 2 y</td>
</tr>
<tr>
<td>Saskatchewan</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Alberta</td>
<td>Yes</td>
<td>2000</td>
<td>Under development</td>
<td>18–69</td>
<td>Annual (to be reviewed when all components of programme in place</td>
</tr>
<tr>
<td>British Columbia</td>
<td>Yes</td>
<td>1960</td>
<td>✓</td>
<td>18–69</td>
<td>After three normal annual tests, screening should be continued every 2 y. If high risk, continue annually</td>
</tr>
<tr>
<td>Northwest Territories</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>18+</td>
<td>After three normal annual tests, screening should be continued every 2 y</td>
</tr>
<tr>
<td>Yukon</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>18+</td>
<td>After three normal annual tests, screening should be continued every 2 y</td>
</tr>
<tr>
<td>Nunavut</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>18+</td>
<td>After three normal annual tests, screening should be continued every 2 y. If high risk, continue annually</td>
</tr>
</tbody>
</table>

*a Has a provincial computerized information system for cytology, which may have been implemented before inception of full programme. From Health Canada (2002)
Patients must be properly examined and cervical cells from the transformation zone must be sampled.

Specimens must be properly collected and labelled.

Laboratory requisition forms must be complete and contain sufficient information.

Cytological tests must be evaluated in a CLIA-certified laboratory.

Laboratory reports must be reviewed to identify patients who require follow-up.

Health-care providers and their patients must be notified of the screening results and any follow-up indicated.

Appropriate follow-up must be taken.

Any substantive discrepancy between clinical, cytological and histological findings must be resolved by the referring clinician and an anatomic pathologist.

In 1988, the first Bethesda System Conference was organized to streamline and update the use of cervical cancer screening terminology. This conference established the Bethesda System, to simplify and improve communication of findings between cytopathologists, clinicians and patients (see Chapter 2). The Bethesda system, with refinements made in 1991 and 2001, is widely accepted in the USA and Canada (Solomon et al., 2002).

Canada

Cervical cancer screening programmes have adopted a system closely resembling the 2001 Bethesda System to classify cytological specimens on the basis of their perceived adequacy for interpretation: satisfactory for interpretation or unsatisfactory. The "unsatisfactory" category is used when the smear quality is inadequate for interpretation (Health Canada, 2002).

Three provinces reported on specimen adequacy for smears taken in 1998, before the changes resulting from the 2001 Bethesda System conference. The percentage of unsatisfactory smears varied from 0.3% to 3.8% (Health Canada, 2002).

Performance indicators

United States

While estimates of screening participation in the USA are readily available, they come primarily from self-reported data collected in the BRFSS and NHIS national surveillance programmes. These figures are not validated and may represent either over- or under-estimates. Few if any measures of adequacy of interpretation other than those required by CLIA exist to monitor accuracy. Little information is available at the national level on such performance indicators as proportion of unsatisfactory tests, proportion of women not receiving indicated follow-up and the proportion of women diagnosed with cancer who were never or rarely screened. Information that is not population-based has come primarily from case series and from the quality control experience of cytology reference laboratories (Sung, 2001; Krieger & Naryshkin, 1994; Tabbara & Sidawy, 1996).

In the NBCCEDP, performance indicators measure adequacy and timeliness of screening and follow-up, the proportion of unsatisfactory tests and the proportion of women never or rarely screened who enter the programme each year. A target has been set of 20% of women screened annually who have rarely or never been screened. In addition, to monitor over-utilization of screening, a target of 75% has been established as the proportion of women who are moved to a triennial test interval if they have had three suc-

<table>
<thead>
<tr>
<th>Age group</th>
<th>Canada %</th>
<th>BC %</th>
<th>ALTA %</th>
<th>SASK %</th>
<th>MAN %</th>
<th>ON %</th>
<th>QUE %</th>
<th>NB %</th>
<th>NS %</th>
<th>PEI %</th>
<th>NFLD %</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–29</td>
<td>80</td>
<td>76</td>
<td>80</td>
<td>91</td>
<td>86</td>
<td>77</td>
<td>81</td>
<td>84</td>
<td>92</td>
<td>84</td>
<td>92</td>
</tr>
<tr>
<td>20–39</td>
<td>86</td>
<td>83</td>
<td>85</td>
<td>85</td>
<td>88</td>
<td>87</td>
<td>82</td>
<td>96</td>
<td>92</td>
<td>89</td>
<td>91</td>
</tr>
<tr>
<td>40–49</td>
<td>82</td>
<td>79</td>
<td>89</td>
<td>82</td>
<td>83</td>
<td>80</td>
<td>83</td>
<td>80</td>
<td>90</td>
<td>83</td>
<td>85</td>
</tr>
<tr>
<td>50–59</td>
<td>77</td>
<td>79</td>
<td>70</td>
<td>74</td>
<td>90</td>
<td>77</td>
<td>78</td>
<td>73</td>
<td>69</td>
<td>77</td>
<td>67</td>
</tr>
<tr>
<td>60–69</td>
<td>60</td>
<td>58</td>
<td>72</td>
<td>64</td>
<td>73</td>
<td>60</td>
<td>56</td>
<td>54</td>
<td>59</td>
<td>68</td>
<td>53</td>
</tr>
<tr>
<td>Total (20–69)</td>
<td>79</td>
<td>77</td>
<td>81</td>
<td>80</td>
<td>85</td>
<td>78</td>
<td>78</td>
<td>80</td>
<td>83</td>
<td>82</td>
<td>80</td>
</tr>
</tbody>
</table>

From Health Canada, 2002
ALTA, Alberta, BC, British Columbia; MAN, Manitoba; NB, New Brunswick; NFLD, Newfoundland; NS, Nova Scotia; ON, Ontario; PEI, Prince Edward Island; QUE, Quebec, SASK, Saskatchewan
cessive confirmed negative programme tests over a five-year period.

**Canada**

Objective information on cervical cancer screening in Canada is not readily available in most provinces on a routine basis. Furthermore, administrative databases at the provincial level, such as physician billing data, cannot be used for the most part as they may not include separate billing codes for Pap tests and, even if they do, will not allow distinction between tests done specifically for screening in asymptomatic women and those done for follow-up. Self-reports from women in national and provincial health surveys and small-area studies are thus the only sources of information.

The only cervical cytology performance indicators currently available relate directly to cytology laboratories. The indicators collected are:

- Cyto-histological correlation rates for each grade of squamous intraepithelial lesion and for carcinomas measured against follow-up surgical material or clinical outcome.
- False-negative rates. The false-negative rate of the laboratory and of individual cytotechnologists should be separately measured. A false-negative result is defined as a screening miss, in a satisfactory smear, of an abnormality graded as ASCUS or worse, or an equivalent if an alternative terminology is in use (Canadian Society of Cytology, 1996).
- The rate of satisfactory and unsatisfactory specimens at the laboratory and slide-taker levels.
- The total number and rates of abnormal gynaecological diagnoses and specific diagnostic categories for the laboratory.
- Turnaround time. Clinicians and laboratories should establish a mutually agreed turnaround time from the date the specimen is received in the laboratory to the date of the final report; an optimal time could be approximately one month.

**Latin America and the Caribbean**

This region includes 28 countries, ranging from the small island states in the Caribbean to Brazil.

**Organization**

Health services in Latin America and the Caribbean began offering screening services with cervical cytology in the early 1960s, through family-planning services and later within primary health care. Table 53 provides information about the current status of cervical cancer screening programmes in the region. There is considerable variation with regard to the age range of the target population, but most of the guidelines recommend screening every three years. To what extent these guidelines are followed by health-care providers is unknown.

Several countries have attempted to set up organized screening programmes, often achieving partial organization. Chile and Colombia have had a national organized programme for at least 15 years, with documented improvements in quality and coverage, as well as decreased mortality in Santiago, Chile and decreased incidence in Cali, Colombia.

In Chile, the programme was reorganized in 1987, focusing on improving follow-up of women screened positive and involving all public laboratories that served the programme in a quality assurance system. No data are available on rates of follow-up, but the country has a large health infrastructure, a public subsidy for health insurance for 70% of the population, with 30% paying for private insurance. In both these systems, services for treatment of pre-malignant lesions are widely offered and access to cancer treatment is guaranteed. In 1990, 40% of women aged 25–64 years had been screened in the past three years, according to a national survey, and participation increased to 66% in 1997, remaining at a similar level in 2000 (68%). The programme is centrally supervised by the Ministry of Health, but managed by each health region. A management agreement for budget allocation is signed between regions and the Ministry of Health, in which specific services and outcomes are agreed upon.

In Colombia, organization of a national programme started in 1989 and guidelines were approved in 1990. These guidelines placed emphasis on diagnosis and treatment for women with a positive screening result. Health sector reform in the 1990s privatized and decentralized the delivery of health care, but special legislation ensured preventive care including cervical cancer screening.

**Extent of use**

In the absence of country-wide population registries, participation in cytological screening has been assessed through surveys. One source of information is a series of Health and Fertility Surveys sponsored by the US Agency for International Development (USAID) in countries in which they have a reproductive health programme. A probabilistic sample of women aged 15–49 years are interviewed in their homes to investigate various aspects of reproductive health. Several countries collect data on screening for cervical lesions by asking women if they had a Pap test in the last 12 months. Table 54 shows participation in cytological screening in countries for which data are available. The lowest participations are observed in Jamaica (15.3%) and Nicaragua (20.5%). Ecuador and Costa Rica report high participation (72.2% and
<table>
<thead>
<tr>
<th>Country</th>
<th>State of cervical cancer screening programme</th>
<th>Screening policy</th>
<th>Lifetime no. of smears</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico</td>
<td>In 1996, a national guideline was approved by consensus with the participation of all major health-care providers. Various studies have shown deficiencies in quality services, including cytology.</td>
<td>Age 25+ of all</td>
<td>13+</td>
</tr>
<tr>
<td>Guatemala</td>
<td>No organized cervical cancer screening programme exists in these countries. Although cytology is performed through family-planning programmes, no follow-up of women screened positive is ensured. Initial attempts to pilot and evaluate visual inspection have been carried out in Nicaragua and El Salvador. Health-care systems in these countries are fragmented and coverage for women is restricted to maternal health.</td>
<td>30–45, annual 30–59, every 2 y</td>
<td>15</td>
</tr>
<tr>
<td>El Salvador</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Honduras</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nicaragua</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominican Republic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Several attempts have been made to organize a cervical cancer screening programme. They have been able to centralize the cytology laboratory and maintain good quality standards. Nearly 90% of the population is covered by insurance and high coverage of cervical cytology has been reported, but no significant reduction in incidence or mortality from cervical cancer.</td>
<td>20–59, every 2 y</td>
<td>17</td>
</tr>
<tr>
<td>Panama</td>
<td>Screening has been available, but no organized screening programme is in place</td>
<td>15+, every 3 y</td>
<td>17+</td>
</tr>
<tr>
<td>Cuba</td>
<td>A programme is in place; coverage has been estimated at 70% using records reported from provincial health departments. No reduction in mortality has been observed.</td>
<td>25–59, every 3 y</td>
<td>11</td>
</tr>
<tr>
<td>Colombia</td>
<td>Has had a cervical cancer screening programme since 1989. With health-sector reform in which multiple health-care providers emerged, major efforts were made to maintain the programme and improve coverage. Decreased incidence has been observed in data from the Cali cancer registry.</td>
<td>25–64, every 3 y</td>
<td>13</td>
</tr>
<tr>
<td>Venezuela</td>
<td>Efforts were made at the state level to improve follow-up of women screened positive by promoting out-patient care and use of colposcopy and LEEP.</td>
<td>25–64, every 3 y</td>
<td>13</td>
</tr>
<tr>
<td>Ecuador</td>
<td>In the city of Quito, a cytology quality assurance programme was implemented through a large NGO in charge of cancer care (SOLCA); more recently coverage has been extended and efforts directed to screen women aged 35–59 years every five years.</td>
<td>30–59, every 5 y</td>
<td>7</td>
</tr>
<tr>
<td>Peru</td>
<td>National guidelines were issued in 2000. Cervical cytology is available. A large project in the province of San Martin has improved capacity and is leading the way for other provinces to improve their programmes.</td>
<td>25–59, every 2 y</td>
<td>17</td>
</tr>
<tr>
<td>Bolivia</td>
<td>The Ministry of Health has recognized the importance of the problem, but cervical cancer screening, diagnosis and treatment were recently excluded from the maternal insurance package.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paraguay</td>
<td>A recent needs assessment was conducted, but no cervical cancer screening programme has been organized.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uruguay</td>
<td>Presents among the lowest cervical cancer mortality rates in Latin America; opportunistic screening is offered in health care services.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>A large cervical cancer screening programme was initiated in 1996 through the National Cancer Institute. In 1998 a large campaign to reach women who had never been screened was launched. Efforts are now being made to incorporate the screening programme into the primary care family health programme.</td>
<td>25–60 every 3 y</td>
<td>11</td>
</tr>
<tr>
<td>Argentina</td>
<td>Cervical cytology is widely offered throughout the country. Cervical cytology programmes have been organized at the provincial level. No quality assurance programme is in place.</td>
<td>35–64, every 3 y</td>
<td>10</td>
</tr>
<tr>
<td>Chile</td>
<td>A cervical cancer programme was reoriented in 1987, with clear attention to improving quality of cytology and follow-up of women screened positive before increasing coverage. Mortality from cervical cancer has begun to decrease, particularly in Santiago where the programme started.</td>
<td>25–64, every 3 y</td>
<td>13</td>
</tr>
<tr>
<td>Haiti</td>
<td>No programme</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch/ English-speaking Caribbean</td>
<td>No organized programme exists in any of the 14 countries, although over the past 15 years several attempts have been made. In 2003, the Ministers responsible for health of the Caribbean Community (CARICOM) placed cervical cancer high on their agenda and 10 countries have assigned new resources to this area.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
66.9%, respectively), largely due to family-planning programmes in their primary health-care infrastructure.

In other Latin American countries not covered by Health and Fertility Surveys, various studies have assessed screening participation in the population. The proportions of women ever screened by cytology were reported as 68.9% among women aged 15–69 years in São Paulo, Brazil (Nascimento et al., 1996); 65% among women aged 20–69 years in Pelotas, Brazil (Dias-Da-Costa et al., 1996); 63.3% among women aged 15–49 years in Morelos, Mexico (Lazcano-Ponce et al., 1996); 81.6% in Guadalajara, Mexico (Jiménez-Pérez & Thomas, 1999) and 45% in Mexico City (Aguilar-Pérez et al., 2003). Consistent with the 1993 fertility survey, in a sample of women aged 18 years and older in Guanacaste province, Costa Rica, 87.8% reported having had a cytological test in their lifetime (Herrero et al., 1997); and in 2000–03, 44.3% of women aged 25–59 years in the province of San Martin, Peru, declared having had a test in the last three years (Ferreccio et al., 2004), consistent with the figure reported in the 1996 survey. The programme in Chile, using a periodic household survey to ask women aged 25–64 whether they had a test in the last three years, reported participation of 68% in the year 2000. Other countries collect similar data from periodic surveys, but these are seldom published. The quality of such information is difficult to assess. In Latin America, higher participations have been found in women who are aware of the benefits of screening (Lazcano-Ponce et al., 1999b; Aguilar-Pérez et al., 2003), have high socioeconomic status, as measured by schooling or housing conditions (Torres-Mejia et al., 2002;
Dias-da-Costa et al., 2003) and who report satisfaction with previous care (Alvarez, 1996; Brenna et al., 2001).

**Quality of cytology**
Several cytology laboratories in Latin America and the Caribbean perform quality assurance procedures, by studying reproducibility, performance evaluation or cyto-histological correlation. In Mexico, two studies on reproducibility of cytological testing, as measured by weighted kappa, concluded that intra-class agreement was low for dysplasia and carcinoma in situ but higher for invasive carcinoma (Alonso de Ruiz et al., 1996; Lazcano-Ponce et al., 1997a). Kappa coefficients ranged from –0.02 to 0.17 for moderate and severe dysplasia between the two studies and were 0.29 and 0.36 for invasive carcinoma. One study (Lazcano-Ponce et al., 1997a) also assessed reproducibility for histology of cervical lesions among 30 pathologists; kappa coefficients were 0.07 for severe dysplasia, 0.23 for moderate dysplasia and 0.64 for invasive carcinoma. Following these studies, a major initiative was undertaken to improve cytology in the country.

An initial evaluation of all cytology laboratories of the Mexican Ministry of Health found that only 70% of microscopes were in satisfactory condition; deficiencies in the supply of reagents and in laboratory facilities were reported. An evaluation of the performance of cytotechnicians was conducted by reviewing sets of slides previously diagnosed by an expert panel. In this exercise only 16% of cytotechnicians achieved a good or excellent score; the percentage increased to 69% after the course. At a third assessment six months to one year later, the proportion with good or excellent scores decreased to 56%, still much better than the situation found at baseline.

Seven countries in Latin America participate in a joint quality assurance programme (RedPAC). The programme includes on-site evaluation and technical support, as well as periodic performance evaluation, measured as agreement between the cyto-histologists and an expert panel. Improvements were observed in Costa Rica, Ecuador and Mexico, where in three consecutive years (1999–2001) weighted kappas rose from 0.43 to 0.65, from 0.44 to 0.63 and 0.52 to 0.63 respectively (Luciani et al., 2003). The improvement was attributed mostly to the reduction in the proportion of slides that were undercalled. Other areas of concern have also been addressed, such as the proportion of inadequate slides, which was reduced after training programmes from 3.71% to 1.98% in Ecuador and from 4.47% to 2.0% in Mexico. In three other participating countries (Bolivia, Peru and Venezuela), no significant change has been observed and agreements (weighted kappas) were under 0.40. The only data from Brazil are for the city of São Paulo, where an analysis of cytology and the corresponding histology in a sample of 157 cases from a major laboratory found agreement in 75.8% of the cases; cytology underestimated 17.2% of the lesions (Di Loreto et al., 1997).

**Performance indicators**
Major efforts and investments to increase participation have been made in many Latin American countries, but other aspects of the programme have not received equal attention. In Mexico, the cytology-based cervical cancer screening programme has been able to avert only 13% of potentially preventable deaths, as estimated by Lazcano-Ponce et al. (1996). This situation has been attributed to low accuracy of the test or quality of cytology in these settings (Alonso de Ruiz et al., 1996). A basic element of a screening programme is to provide diagnostic and treatment services for those screened positive, but this has been neglected in many countries. In a study in Peru, only 25% of women who were screened positive received appropriate follow-up (Gage et al., 2003); a similar situation was found in El Salvador (Robles, 2004). In a municipality of São Paulo, Brazil, 78.3% of women with a positive cytological result received adequate treatment in primary care services; but when they were referred to secondary care, only 37.4% had adequate follow-up (Santiago & Andrade, 2003).

Overall, cytological testing has been offered throughout Latin America and the Caribbean, but most countries have failed to organize screening programmes. Research is under way to assess the potential effectiveness of low-cost technology such as visual inspection methods and of HPV DNA testing in various populations of Latin America and the Caribbean.

**Africa**
Screening is difficult to achieve in Africa partly because of significant competing, urgent health-care needs, in particular the HIV epidemic, and partly because of poorly functioning health-care delivery systems.

A review on the functioning of health-care systems undertaken by the World Health Organization (2000) rated and ranked these systems worldwide, revealing that among the 191 WHO Member States, most of the least functional health-care delivery systems are in Africa.

To make an impact on the epidemiology of cervical cancer, as with any
screening programme, national organized programmes that achieve high participation are required. Research in African settings confirms that integration of screening services into the existing health-care systems is the only way that high participation rates could be achieved (JHPIEGO, 1997). Other research is attempting to find technological solutions (screening methods that are cheaper, that do not require laboratory back-up, or that allow immediate treatment, with low-technology and low-cost treatment options) to address some of the health system inadequacies that are prevalent in Africa and some research is looking at other treatment options (Adewole et al., 1998; Darwish & Gadallah, 1998). In addition, such efforts may provide pilot sites from which national programmes can evolve.

**Organization**

Only South Africa has an official national cervical screening policy, which recommends three cytological tests for women over the age of 30 years at ten-year intervals. However, the need for national policy has been articulated in many countries in Africa (Adanu, 2002). Absence of policy has resulted in lack of action and in screening of women at inappropriate ages (Ngwalle et al., 2001; Konje et al., 1991); in addition, facilities for treatment of precancerous lesions are often inadequate (for example, Tanzania: Ngwalle et al., 2001).

There are indications, however, that even in the absence of formal policy, many countries have plans to implement programmes. In Malawi, a programme for cervical cancer screening and early treatment, as a partnership between the Ministry of Health and Population and an international non-governmental organization, has been initiated and a nationwide campaign is planned which will start in the southern part of country (Anonymous, 2002). In Zimbabwe, screening based on visual inspection with acetic acid (VIA) has been under consideration (Chirenje et al., 2000), but it is not clear if this is a national programme and whether this is a priority for Zimbabwe (Rutgers & Verkuyl, 2000). The current state of the Zimbabwe health service suggests that rapid implementation is highly unlikely. Cameroon has had an operational cervical screening programme since 1992, but services are provided only in two major cities and the cost is such that most women do not have access to the service (Robyr et al., 2002).

It is clear that access to screening is poor for most women in Africa (Kenya and Sierra Leone; Burkina Faso, Senegal, Ghana, Nigeria: Brown & Morgan, 1998) and this results in late presentation of women at treatment centres (Gharoro et al., 1999; Were & Buziba, 2001).

Although policy frameworks are useful and their absence can lead to irrational programmes and confusion (Moyo et al., 1997), they are not sufficient to ensure implementation; processes to ensure that policy is understood are essential (Anonymous, 2001). In South Africa, there is now a national initiative to implement a national cervical screening policy, which is probably the best developed in the region and might provide an example for other countries to follow. South Africa is, however, different from many other countries in the region in having a relatively well developed laboratory-based capacity to provide cytology services.

**South Africa**

The South African National Cancer Control Programme was adopted as official government policy in 1997. Policy development is a national function, but implementation and delivery is a provincial (and more recently a local government) function. While policy indicates what is required, it often does not explain how services should be developed. This leads to unrealistic expectations at the policy level and frustration for implementers. After various attempts to implement screening by promoting the policy, a national strategy spearheaded by the national Department of Health was initiated. The main objectives of this strategy are to strengthen the existing health-management systems to implement, monitor and sustain the cervical cancer screening programme; to ensure maximum coverage of the target population; to ensure provision of facilities for screening and treatment of precancerous lesions and develop referral links between screening and treatment services; to increase awareness of cervical cancer and its prevention; and to ensure monitoring and evaluation of the programme.

Despite this policy framework, progress in cervical cancer screening has been slow and hard to achieve. The nine provinces in South Africa are at different levels of development and there are resulting differences in implementation. Two provinces (Limpopo and Eastern Cape) are predominantly rural in nature and poor, with health-care delivery systems that function poorly. Others (Gauteng and Western Cape) are urban and have good resources. The other provinces are located between these two extremes.

Study of the South African situation shows that where policy does exist, it is more likely that resources will be dedicated to cervical screening services. Thus in the Western Cape province, cervical screening has been identified as a priority service and development of services has been included in the key performance indicators or in performance agreements of health-system managers. Similarly in KwaZuluNatal, cervical screening is a performance target for districts.
These two provinces have the greatest access to screening services. In three provinces budgetary allocation for cervical screening exists. While in provinces such as Limpopo and Mpumalanga, poor transport systems are considered a barrier to implementation, as this hampers the delivery of slides to laboratories. However, other possibly poorer provinces have found creative ways of overcoming this problem, for example by linking cervical screening with the tuberculosis programme and using the same laboratory transport system. This illustrates the value of integrating cervical screening into existing health-care services. In another case, private taxis transport the slides as part of their routine runs. While in all provinces there are colposcopes at the district level, there may be no trained staff available to use them. In one rural province a partnership with a non-governmental organization has been set up to improve access to screening services and three of the five districts comprising the province report offering services. Nonetheless, in some instances women in the wrong age group are being screened. However, it is becoming increasingly clear what kinds of intervention are required in order to assist provinces to implement a programme. Pilot programmes have been set up and manuals for managers have been developed.

The South African experience has shown the need to provide health service managers with tools to assist them with practical aspects of implementation. For example, one tool indicates how to work out the target population and thus the number of smears to be taken in a year and the workload and equipment needs at each service site. Another tool indicates what clinic-based data to collect and how to estimate the participation rate and the follow-up in each service delivery area.

The rest of the region
No data on cervical cancer screening are available from northern Africa.

In a review of cervical cancer diagnosis and treatment in countries of east, central and southern Africa, a lack of policy guidelines, infrequent supply of basic materials and absence of suitably qualified staff were the common reasons reported for the low percentage of women actually screened. The review found that 95% of the facilities at primary care level had the basic infrastructure to offer cervical screening (Chirenje et al., 2001). However, once slides were taken, there was no way to send them for reading and limited access to referral pathways to treat patients. In Botswana, for example, cervical screening is limited because few facilities have easy access to laboratories to read cervical smears (Baakile et al., 1996).

Given the competing demands on health-care services, the only way in which cervical screening programmes will gain the required political support is if they are developed in such a way as to benefit the overall functioning of the health-care system. The imperative of high coverage requires that services be decentralized, thus integrating cervical screening into the existing health-care services offers the best approach to reducing cervical cancer mortality. However, such integration is not simple, as much of the organization of health services in Africa is donor-driven, often resulting in single-service facilities such as family-planning services. In a study in Kenya, integration of cervical screening into family planning clinics was reported to be feasible and acceptable to both providers and patients, and would benefit the patients screened. However, only a small percentage of women utilize these services and in the Kenyan study, 43.5% of women were less than 30 years of age, so that the potential reduction in cervical cancer mortality is limited (Claeys et al., 2003).

It has been suggested that many African countries are not able to implement screening and that cervical cancer is not a priority (McCoy & Barron, 1996; Rutgers & Verkuyl, 2000) or that it is important but impossible to achieve (Wilkinson, 1997). However, the benefits that accrue from setting up a cervical screening service can be extended to other services, so that referral pathways, laboratory services, equipment supply systems and monitoring systems, once operational for cervical screening, can be extended to other health-care needs or be developed in tandem with and complement existing systems (Fonn, 1997).

There are neither the resources nor the human capacity in Africa to develop vertical programmes (single-service programmes with staff working only in that programme, frequently with unique conditions of service and unlinked independent supply and monitoring systems, unrelated to other health-care services often provided in adjacent sites). Yet existing resources can be marshalled and applied to cervical screening. An approach that recognizes and builds health-systems capacity to deliver cervical cancer screening can improve the overall functioning of health services in general.

Asia
Data on cervical cancer screening from western and south-central Asian countries were available to the Working Group only from Israel and India. Data are also available for a number of south-eastern and far eastern Asian countries.

Israel
As the incidence rate of cervical cancer in Israel is very low, the official policy is not to screen average-risk women. However, the National Insurance Plan reimburses cytological testing for women aged 35–54 years once every three years. In practice, many women are screened, usually at
shorter intervals than recommended, with generally low-quality cytology. In addition, most women attending screening are of high socioeconomic status and probably are not the women at highest risk. About 150 000 tests are performed every year.

**India**

India has a National Cancer Control Programme that supports the principle of early diagnosis and treatment of cancer of the cervix. Although cytological testing is available to a limited population of mainly urban women, there are no screening programmes (Dinshaw & Shastri, 2001; Shanta, 2001; Varghese et al., 1999). India is a high-risk country for cervical cancer (Shanta et al., 2000; Sen et al., 2002). Women at highest risk of cervical cancer are those over the age of 35 years, in low socioeconomic strata and with little or no education. Given that over 80% of the population lives in rural areas, screening programmes need to work within this sector (Dinshaw & Shastri, 2001; Sankaranarayanan et al., 2001; Shanta, 2004). Cytology-based screening is not regarded as practical or achievable in India (Dinshaw & Shastri, 2001).

Visual inspection-based approaches to cervical cancer screening such as VIA have been extensively investigated in India, although their long-term efficacy in reducing the burden of cervical cancer has not yet been demonstrated (Sankaranarayanan et al., 2001; Basu et al., 2003). Visual inspection is now regarded as the best option for proposed cancer control programmes, with training curricula and courses developed by international organizations such as IARC and JHPIEGO (Shanta, 2001). The Bill and Melinda Gates Foundation funds the Alliance for Cervical Cancer Prevention which supports projects in India, including through IARC (http://www.alliance-cxca.org/index.html).

The advantages of visual inspection methods are the lower costs than cytology and the short training period required for health workers, including the ability to train nursing and nonmedical workers (Basu et al., 2003; Sankaranarayanan et al., 2003). Some 457 000 women (0.25% of all eligible women at risk) have participated in screening studies of visual inspection methods (Shanta, 2004). The results indicate that the women accepted screening by visual inspection with acetic acid or magnification after application of acetic acid (and colposcopy and cryotherapy) by nurses, that a moderate level of compliance with screening and treatment was reached, and that these methods have higher sensitivity and lower specificity than cytology in the Indian setting (Basu et al., 2003). The low specificity, however, that causes high rates of referral and treatment, was a major limitation. Nevertheless, visual inspection has been recommended as the immediate option for cervical cancer control initiatives in 54 districts of India (Shanta, 2001; Sankaranarayanan et al., 2001).

**Viet Nam**

Viet Nam does not have a national screening programme. Research activity relates mainly to HPV prevalence. A substantial difference in the prevalence of cervical cancer and of HPV infection between the north and south of the country are regarded as due to the greater isolation of north during the decades of war and socialist economy (Pham et al., 2003). A survey carried out in 1997, within the framework of an IARC multicentre study, found that HPV infection was rare in Hanoi and five-fold higher in Ho Chi Min City (Pham et al., 2003).

In November 1998, the Western Pacific Regional Office of WHO collaborated in strengthening cervical screening programmes with a series of training sessions on cytological screening, and pilot projects were established in Hanoi and Ho Chi Minh City. In co-operation with the Viet/American Cervical Cancer Prevention Project, the feasibility of cytological screening in Viet Nam was established by a formal cost-effectiveness analysis, and population-based cytological screening services were established in 1999 in Ho Chi Minh City (Suba et al., 2001; Le Van et al., 2004). Pilot-scale screening is continuing to assess whether cervical cancer constitutes a public health problem of sufficient magnitude in northern Viet Nam to warrant the initiation of population-based screening.

**Thailand**

Thailand has attempted to establish a cervical cancer prevention programme for 30 years, with activity in selected districts through maternal and child health or family-planning services (Gaffikin et al., 2003). In 1997, a national policy for cervical cancer proposed that screening be offered to women aged 35–54 years with a Pap test every five years and, in the north-east of Thailand, using visual inspection methods. However, surveys have found that few women know about the Pap test and few have ever been screened (Kritpetcharat et al., 2003; Tinker, 2004). National annual participation is estimated to be no more than 5% (Gaffikin et al., 2003). A mobile unit programme offering cytological testing was established in 1993 (Swaddiwudhipong et al., 1999) and a demonstration programme of visual inspection methods in 2000 (Gaffikin et al., 2003). The latter programme concluded that a single-visit approach with VIA and cryotherapy by nurses was safe, acceptable and feasible and could be considered in areas where setting up cytological screening is unlikely (Gaffikin et al., 2003). Concerns have been raised, however, about potential
Use of screening for cervical cancer

over-treatment due to the low specificity of VIA (Walraven, 2003). Other impediments to the use of VIA in a population-based programme include the need for standardized initial training and continuing education, and that no study has shown that VIA screening reduces cervical cancer incidence or mortality (Walraven, 2003).

**Philippines**

Cervical cancer is the second most common cancer in women in the Philippines. The Department of Health in the Philippines has proposed an organized cervical cancer screening programme, with recommendations for regular cytological tests every three years, although a recent policy shift has recommended visual inspection methods (Ngelangel & Wang, 2002; Ngelangel et al., 2003). Changes in public health policy, including aspects related to education of screening personnel, strategies for ensuring compliance with screening and health insurance coverage for preventive services, have been mentioned as barriers to the development and implementation of a screening programme (Ngelangel et al., 2003). The lack of a skilled workforce is also an issue. The Philippines Cancer Society is involved in cytological testing, although this is not widely available (http://www.kanser.com.ph).

**Republic of Korea**

In Korea, a national screening programme for cervical cancer started in 1999. Cytological screening is recommended every two years for women 30 years or older, of whom 33% had a test in 1999–2000. The provision of services is insurance-based, administered by the Ministry of Health and Welfare. This insurance covers the costs to individual women of screening for cervical cancer selectively for lower-income women. Discussion and planning are continuing in order to define the screening interval, upper age limit, the test and quality control procedures.

**China**

Occurrence of cervical cancer is largely unrecorded in China, but it is known to be higher in rural areas; mortality from cervical cancer is around the level in the USA (Belinson et al., 1999). Mortality has decreased over the past 25 years, maybe as a result of the major social changes and the health programme set up by the Chinese government after the founding of the People's Republic in 1949, in particular, the outlawing of prostitution and closure of brothels, and establishment of health facilities in factories and other work units, and specific public health programmes to screen for and treat sexually transmitted diseases (Li et al., 2000). Cervical cancer accounted for only 1% of cancers in women in 1995 (Wang, 2001). Cancer prevention is not a high priority and lacks government funding. There is no national screening programme and cytology-based services are patient- or employer-initiated. Women who have insurance can attend a hospital for screening. While government agencies cover employees for insurance, private employees may or may not be covered. Screening is now less common than 10 years ago and is becoming more of a personal activity, organized at the level of individual companies or groups and subject to market forces. This change has led to fewer individuals being screened (Wang, 2001).

Small centres offer a screening service, mainly associated with universities or small studies. In the past, the work unit was the sole channel through which screening could be offered and was responsible for organizing any screening that took place. Health professionals hope that the new medical insurance system will cover cancer screening and prevention (Wang, 2001).

In Shandong Province, a cytological screening programme started in 1970–72 and covered 1.5% of the female population, before government funding was withdrawn; screening then continued in only a few areas (Li et al., 2000).

More recently, a pilot study in a high-risk province (Shanxi province) conducted a trial of cytological testing and HPV direct and self-testing in 1997 among women aged 35–45 years (Belinson et al., 2003).

Three national demonstration centres will be set up for screening, organized through field stations in women's and children's health centres or village clinics.

**Hong Kong**

Cytological screening has been offered for 10 years in Hong Kong by the Department of Health through 34 gynaecological clinics and health centres, including maternal and child health centres, social hygiene clinics and others such as family planning, for which the service statistics indicate around 100 000 tests per year in 1997–2001 (http://www.famplan.org.hk/fpahk/en/template1.asp?style=template1.asp&content=info/statistics.asp&type=3). It has been estimated that the Department of Health and Family Planning was responsible for 24% of Pap tests, the Hospital Authority for 15% and private hospitals and medical practitioners for 37% in 2003 (Asian HPV Summit, 2003). The existing cytological testing has not been part of an organized programme and has no target population, no screening register and no formal quality-control process.

Several surveys conducted in Hong Kong showed that around half of all women were not receiving cervical screening (Yeung & Cheung, 2003). The Shatin Community Clinic for the Prevention of Cervical Cancer was set up in 1995 to reduce the incidence of cervical cancer. It provides a free
service to women who have never been tested, without age restriction, and tested 20,000 women in 1995–99. In 1998, it received funding from the Hong Kong Cancer Society for an automated screening instrument, and has accreditation from the Australian laboratory accreditation authority. The clinic provides training in diagnostic cytology to the Chinese PLA General Hospital in Beijing.

A population-based cervical screening programme was to be launched by the Department of Health in collaboration with other health services providers in late 2003 or early 2004 (Yeung & Cheung, 2003), targeting women aged 25–64 years with three-yearly screening, recruitment being through invitation letters, publicity campaigns and community outreach. The programme will provide training for smear-takers, have a central register and establish quality assurance indicators.

Taiwan
In 1993, an estimated 40% of women in Taiwan had never been screened (Wang & Lin, 1996) and a programme of free mass screening was established as part of the national health insurance in 1995 (Chen et al., 2002). An educational and cervical screening programme at 12 public health centres in Taipei was extended (Pair & Ruey, 1996), using outreach clinics in areas with inadequate medical facilities and offering training courses (Chen et al., 2002). The goal of the programme is to achieve a screening rate of 40% in women 30 years or older, who are offered free three-yearly screening. The Central Department of Health monitors and evaluates the programme, which is delivered by the City and County Health Bureau and local health stations with follow-up of HSIL grades by local public health nurses. The programme has a central registry and a process for laboratory quality control.

Singapore
The 1998 National Health Survey reported that about two thirds of Singapore women had ever had a cytological test (Yian, 2000). Screening of sexually active women aged 20–69 years is carried out at 16 polyclinics, 1,900 private medical clinics and hospitals and the Singapore Cancer Society, although no annual figures are available (Thamboo et al., 2003). Women pay for the test but older age groups are offered a subsidy. In January 1999, the Ministry of Health launched a Cervical Cancer Education Program, with a recommendation for an annual test in the first two years and three-yearly tests thereafter (http://www.hpb.gov.sg/hpb/haz/haz01123.asp).

A coordinated national programme has been set up to start in 2004 (Asian HPV Summit, 2003, Dr Quek), with training programmes for smear-takers, a smear reporting system as in Australia, a system of audit and management guidelines for abnormal results. The full programme is expected to commence after a one-year pilot programme at selected primary health care clinics.

Japan
Organization and financing
Cytological screening for cervical cancer was introduced in selected regions of Japan in 1961. These early programmes were established and organized voluntarily by gynaecology practitioners in cooperation with local government officials. National government funding, initiated in 1967, led to implementation of screening programmes at a nationwide level. In 1983, the Health and Medical Service Law for the Aged was virtually inactivated, leaving implementation to be decided by each regional government, the continuation of cancer screening programmes has been strongly advocated by the national government. Cervical cancer screening is now funded by each regional government. Women have to make an out-of-pocket payment of 10–30% of the total cost of the test, the proportion differing between regions.

In addition to the mass-screening offered by the regional governments, many women have the opportunity to participate in company-based cancer screening, often offered by employers as part of a health insurance and benefits package, or personal health examinations, usually including cervical cancer screening, at private clinics and institutions. Cervical cancer screening offered by these programmes is implemented under almost the same criteria as programmes sponsored by regional governments. In a questionnaire survey, from 216 completed questionnaires, 147 companies (68%) stated that they offered cervical cancer screening as part of their employee health check-up (Nagai et al., 1998). Thus, a variety of screening programmes are available to most women.

Extent of use and method of screening
Since the passage in 1983 of the Health and Medical Service Law for the Aged, the screening protocol recommended by the national government has been offered to residents of all prefectures. The target population includes all women aged 30 years or above, with a screening interval of one year. Women are individually invited to participate. The test is performed by obstetricians/gynaecologists under a speculum examination using a cotton swab, spatula, scraper or brush for...
sampling. Interpretation of cytological specimens is carried out by certified cytotechnologists and cytopathologists; their certification is carried out under the auspices of the Japanese Society of Clinical Cytopathology. All cytotechnologists and cytopathologists are members of this society and in order to assure the quality of cytology screening, the society offers regular training and education courses in addition to renewal of certifications every four years, based upon stipulated conditions.

The results of screening are reported in a five-tier evaluation designated as Classes I to V, based upon a modified Papanicolaou classification. Although the tiered evaluation system was not incorporated in the Bethesda System of 2001, the use of only a written evaluation report is still not widely accepted among clinical practitioners in Japan and the five-tier cytology classification is still used to avoid misunderstanding and to facilitate reporting of the screening results. The five-tier evaluation consists of Class I as normal, Class II as inflammatory change, Class III as dysplasia, Class IV as carcinoma in situ, and Class V as invasive carcinoma, as deduced from cytological features. All Class III results and above are interpreted as screen-positive and a repeat cytological examination as well as a colposcopically guided cervical biopsy are recommended as secondary testing.

The accuracy of cervical cancer screening carried out in Miyagi prefecture as part of the nationwide programme showed sensitivity of 94.7%, specificity of 98.9% and a false negative rate of 5.3% by linkage analysis, when all women screened were compared with patients having invasive cervical cancer or carcinoma in situ who were registered in a regional cancer registry (Table 55; Yoshida et al., 2001). False negative cases were defined as those diagnosed as having cancer, including carcinoma in situ, within one year after the negative screening result.

Quality assurance control for cervical screening programmes is administered by Management Control Committees for Lifestyle-related Disease established in each prefecture. These committees monitor information regarding the total number of participants, participation rate, secondary screening rate and individual participation history for cervical cancer screening programmes in each city or town in all prefectures (Ministry of Health, Labor, and Welfare, 1998).

The only data on total participation numbers and rates, as well as screening results for the programmes provided by regional governments, are integrated at a nationwide level and published annually in a Report on Elderly Health Care by the Statistics and Information Department, Minister's Secretariat of the Ministry of Health, Labor, and Welfare. A participation rate of about 14–15% is reported. No similar comprehensive analysis is available for non-government-sponsored cervical cancer screening programmes, although the estimated overall participation, combining both government and non-government-sponsored programmes, has been estimated to be 24%. Few quality assurance controls exist for non-government-sponsored cervical cancer screening programmes.

**Future perspectives**

The low participation of the cervical cancer screening programmes has been a matter of concern. In 2000, the national government presented a ‘National Health Promotion Movement in the 21st Century (Healthy Japan 21)’, which included the goal of increasing the number of participants by more than 50%.

Another concern is to broaden the target population, taking into account an observed increase in the incidence of carcinoma in situ and invasive cancer in younger women (Research Group for Population-based Cancer Registration in Japan, 2003). In April 2004, the Ministry of Health, Labor, and Welfare recommended that screening should be initiated at 20 years of age with an interval of two years.

**Oceania**

**Australia**

Cytological screening was readily available to Australian women from the 1960s but largely on an opportunistic basis. In 1988, however, the Australian Health Ministers’ Advisory Council set up the Cervical Cancer Screening Evaluation Steering Committee, which recommended an organized programme that was established in 1991 and renamed the National Cervical Screening Program in 1995. The organized programme is a joint initiative of the commonwealth, state and territory governments to provide cervical screening by coordinating the local programmes in individual states and territories, each of which has adopted or endorsed the goals and priorities of the national programme and uses the same performance indicators and targets. States and territories are responsible for regional coordination and delivery of screening.

Cervical screening is available to all sexually active women between 18 and 69 years of age, with a two-year screening interval. Women are asked to give their consent to their details being entered in the local cervical cytology register, from which a reminder is sent two years after their last screen. If the test result is abnormal, the register sends a letter to the woman and her medical practitioner to help ensure that appropriate follow-up action is taken. Coordination and programme administration are funded jointly through a national and state government agreement that covers...
Extent of use and access
The screening programme offers a test every two years to all sexually active women from around age 18–20, or younger if appropriate, and up to age 69 at which time screening may cease after two normal test results within five years. Women 70 years and older who have never had a test, or who request one, are eligible for screening. Women who have an intact uterus and have no symptoms or history suggestive of cervical pathology are eligible. There are separate national guidelines for management outside of the screening programme for women with a history of high-grade cervical lesions or who are being followed up for a previous abnormal test result.

Education campaigns encourage eligible women and under-screened groups to attend. General medical practitioners take most Pap smears (80%); gynaecologists, women’s health nurses, Indigenous health groups and sexual health and other clinics are other providers. The scarcity of medical practitioners, particularly women practitioners, in remote and rural parts of Australia limits access to screening test services, although women’s health nurses may be available. Accredited cervical cytology laboratories read the slides and send the results to the state or territory screening register and also to the health-care provider who took the smear and to the woman.

The costs of a visit to a medical practitioner and the laboratory costs for reading the slide are reimbursed by Medicare, the national health insurance scheme funded by the Commonwealth government. The same is true for treatment. The woman’s contribution varies, since some medical practitioners charge more than the Medicare reimbursement and the women themselves must fund the difference in the fee. For women who are eligible, there is no charge for visits to providers funded through Health Program Grants or by state governments.

In 1991, the year before the national programme began, 52% of women nationally had a screening test. In the programme in 2000–01, 61.8% of women had a test. Participation varied by state: 58% in Queensland, 60–63% in New South Wales, Western Australia, the Northern Territory, and the Australian Capital Territory, and 66–67% in South Australia and Tasmania (Australian Institute of Health and Welfare, 2003a). Nationally, 32% of women registered with the programme were re-screened within less than the recommended two-year interval in 1999–2000 and 2000–01 (Australian Institute of Health and Welfare, 2003b). Women known to be under-screened are those of low socioeconomic status or with indigenous or other culturally and linguistically diverse backgrounds, and women 60 years and older or from rural and remote areas (Department of Health and Aged Care, 2000). No identifier by indigenous status is available in screening registers. Published four-yearly mortality rates, however, show that death rates from cervical cancer are much higher in indigenous (11.4 per 100 000 in 1998–2001) than non-indigenous (2.5 per 100 000) women (Australian Institute of Health and Welfare, 2003b).

The computerized cytology registers set up in each Australian state and territory in 1989–99 record contact details of consenting women and the smear-takers forwarded by health-care practitioners; results of tests and identification of the reporting laboratory are sent directly by laboratories. All information is confidential. Around 1–3% of women choose not to be registered by name and de-identified demographic details and smear results only are recorded for them. All registers have a protocol to ensure that women with test abnormalities have appropriate follow-up.

Registry data are collated nationally at the Australian Institute of Health and Welfare, which has produced five annual reports on the performance indicators endorsed by the national screening programme, beginning with data for 1996–97 and continuing up to 2000–01. Data standards in place ensure consistent and reliable data for performance indicators.

Government expenditure on screening with cervical cytology in 1994–95 was mainly (61%) through Medicare reimbursement for private medical services and some pathology. 23% was a direct national government contribution to the screening programme and 16% came from the local...
resources of the states and territories (AusAID, 1999). In 1999–2000, Medicare expenditure on cytological tests and pathology was $84.2 million, while national recruiting activities were allocated $4.8 million over the three years. Funding to increase participation was introduced in November 2001 and offers incentives to medical practitioners for increasing participation by women who were not tested in the last four years.

Methods for assuring quality
The National Advisory Committee to the National Cervical Screening Program (Advisory Committee on Cancer Prevention, 2000) has five working groups with members from all programmes, health professionals (pathology, general practice, health economics, epidemiology, gynaecology), and a consumer and indigenous representative. A working group responsible for quality assurance monitors specified programme outcomes and identifies areas for improvement in laboratory adherence to performance standards and methods to improve quality of Pap smears.

Australia has required formal accreditation of all pathology laboratories since 1987 with three-yearly inspections to ascertain compliance with national standards set by industry and professional authorities (National Pathology Accreditation Advisory Council (NPAAC), 2003). In addition, the screening programme, in consultation with pathology accreditation authorities, has formulated performance standards for technical competence in gynaecological cytology in laboratories; these were voluntary from 1996 and mandatory since 1999. The formal accreditation process requires laboratories to submit standard data, which are compiled in a national report distributed to all laboratories with their own performance data. No individual laboratory is identified in the national report. The information in these reports is verified against data supplied by the cervical cytology registers. Laboratory performance is self-assessed using in-house quality systems, which set up corrective measures as necessary and report on the process to the assessment authorities. Reimbursement through the national health-care system is unavailable for services in unaccredited laboratories.

The pathology performance standards for Australian laboratories reporting cervical cytology were reviewed in 2003 by NPAAC. Performance measures include the proportion of unsatisfactory and satisfactory specimens, the positive predictive value of a cytology report of a high-grade intraepithelial lesion, the false negative rate among women with histologically confirmed carcinoma in situ, and the turnaround time in processing slides. All states have a feedback loop whereby the register supplies laboratories with a performance report on test reporting and laboratories feed information back to the register to allow monitoring of data integrity. Performance indicators, endorsed by the national advisory committee, are reported annually for screening programmes: these include the proportion of women participating by age, the percentage of women with re-screening in the 21 months following a negative screen, the ratio of low- to high-grade abnormalities verified on histology for women aged 20–69 years, the incidence of micro-invasive and all invasive cervical cancers, and mortality. Incidence and mortality by location and mortality in indigenous women is reported every four years (Australian Institute of Health and Welfare, 2003a, b).

New Zealand
A national screening programme came into operation in 1990–91 following recommendations in 1985 for routine cervical screening and the Cartwright Inquiry that recommended a nationwide population-based programme with central coordination (Skegg et al. 1985; Ministry of Health, 1997; National Cervical Screening Programme (NCSP), 2002).

In 1993, enrolment in the programme increased after a legislative change from registration only of those women who consent (‘opt-on’) to non-registration only of those who specifically refuse (‘opt-off’), and histology reporting became compulsory. Maori women’s data was protected in 1995 under the Kaitiaki regulations and in 1997, the Ministry of Health established a data management group for Pacific women. Statistical reports were produced in 1993, 1995 and 1998. In 1996–97, a national cervical cytology screening register was centralized in Wellington and in 1998, responsibility for national coordination of the screening programme passed from the Ministry of Health to the Health Funding Authority. Following a review, an additional $1.4 million was injected into the NCSP during the 1999/2000 year to improve quality standards, set up new independent monitoring processes, improve coordination between providers and improve information for women and training for health educators. The National Screening Team transferred to the Ministry of Health as a separate unit, the National Screening Unit (NSU), within the Public Health Directorate in January 2001.

The NSU funds the screening programme by contracting four independent service providers to provide health promotion to Maori, Pacific and other women in their regions and 12 laboratories (2 public and 10 private) to provide cervical cytology services. Regional screening services are contracted through 13 District Health Boards that provide health promotion and cytology to priority-group women and regional coordination; eight of the
regional services are responsible for local management and data entry of laboratory results to the screening register. In addition, the NSU also funds some low-cost or free cytological and colposcopic services and treatment provided by District Health Boards.

An inquiry into the apparent under-reporting of abnormal results in the Gisborne region found that during the 1990s the NCSP lacked the necessary organization, coordination and some of the constituent parts required for safe and effective screening programmes. A key finding of the inquiry was the need for the Health Act 1956 to be amended to enhance the capacity to monitor, audit and evaluate the NCSP. Appropriate legislation is now before Parliament.

The NCSP Operational Policy and Quality Standards were introduced from November 2000 across the programme. These set standards for laboratory and publicly funded colposcopy services. Since mid 2003, 758 585 women have had one or more screening tests in the NCSP in accordance with the new standards. The national programme sets minimum standards for health services; providers contracted by the Ministry of Health are monitored against these standards. A complete copy of the standards is available on the National Screening Programme website (http://www.healthywomen.org.nz).

The screening programme targets all women aged 20–69 who have ever had intercourse for a three-yearly cytological test. In particular, the programme targets women who have never been tested or whose previous test was more than five years ago; these women have a second test after one year. Women who have had a hysterectomy for a benign condition, with complete removal of histologically normal cervical epithelium and a normal cytological history, do not require further screening. Women aged over 40 years and Maori and Pacific women are priority groups in the NCSP.

Pap smears are taken by general medical practitioners (70%), specialists (5%), nurse smear-takers (25%), and two lay smear-takers without a professional background. Cytology reading is funded by the government in 10 community-based and two public hospital-based laboratories. The cost to the woman is only the normal consultation fee, except when the screening unit is directly funding the service, and an additional fee when a liquid-based cytology specimen is used, and private colposcopy services.

Test results are forwarded to the register unless the woman has opted off and does not want the result sent to the programme; results are also sent to the smear-taker. The register sends the woman a reminder when a test is overdue and also supplies the women’s cervical screening histories to smear-takers and laboratory cytologists to assist in management decisions. The register also makes sure that the woman is informed of an abnormal result. It holds the name, address, date of birth, smear-taker and their details, cytological and histological history, and a record of letters sent to the woman.

**Extent of use and access**

The programme aims to cover 85% (adjusted for hysterectomy) of all 20–69-year-old women recorded on the screening register in the previous 36 months. In 1998, 76% of eligible women (84% after adjustment for hysterectomy) were tested (Independent Monitoring Group of the National Cervical Screening Programme, 2003). By May 2003, 99.14% of the eligible population (1 084 592 women) were enrolled on the Register (http://www.csi.org.nz/other_reports/N CSPQuestionsnAnswers.htm). Extent of use and access are currently estimated using census data, as the lack of a population register precludes ready identification of women for targeted recruitment.

The rate of cervical cancer for Maori women in 1997 was nearly three times that of non-Maori (Ministry of Health, 1997). An important function of the screening programme, therefore, is to address the disparities in health outcomes for Maori women. The issue of choice of service provider is important to Maori women. A Maori Women’s Cytology Working Group was established and funded. Although Maori smear-takers are not available in all areas, Maori community health initiatives are offered in most areas and a National Kaitiaki Group was formed to act as guardian of registry data (Ministry of Health, 1997).

**Performance indicators**

National indicators for quantitative monitoring have been developed as part of the process of improving overall quality assurance in the NCSP (National Screening Team, 2000).

An independent monitoring group at the University of Otago is contracted to evaluate the programme against national indicators and targets (Independent Monitoring Group of the National Cervical Screening Programme, 2003). Performance indicators are reported quarterly, six-monthly or annually. The indicators reported quarterly are short-interval re-screening, delayed re-screening for women with a high-grade abnormality, follow-up of women with HSIL cytology, laboratory test reporting, including cytology and histology turnaround time, satisfactory but limited and unsatisfactory smears by laboratory and smear-taker, and the positive predictive value of cytological reports of HSIL.

Annual reporting is required for the numbers of women enrolled, participation, coverage of women having a screening test recorded on the registry in the 36 months preceding the end of
the reporting period, non-participation, re-participation, incidence of cervical cancer and incidence by stage, cervical cancer mortality, rates of cervical abnormality and histology abnormality reporting on the register, interval cancers, programme sensitivity, the opt-off rate, the accuracy of negative cytology reports and residual high-grade disease after treatment (Independent Monitoring Group of the National Cervical Screening Programme, 2003).

Behavioral considerations in screening

Success in delivering a screening programme requires a good understanding of, and attention to, behavioral factors. These factors include communication about cervical cancer and the screening process, the psychological consequences of participating in screening and issues that affect participation. Most research in this area has focused on predictors of attendance at screening and the evaluation of strategies designed to increase participation.

Information and understanding

The cancer screening process can have substantial negative consequences for an individual in terms of anxiety and, if screening results are positive, additional tests and treatment. The psychological consequences of cervical screening are discussed further in Chapter 5.

The fact that screening usually targets individuals who do not have symptoms enhances the significance of potential negative consequences. Women should receive accurate, evidence-based information about both the hazards and the benefits of a screening programme, so that they can make informed decisions about whether to take part (see, for example, General Medical Council, 1998). While research has identified barriers and deficiencies in information provision, little is known about effective ways of enabling women to make informed decisions (Cockburn et al., 1995; Raffle, 1997; Coulter, 1998; Anderson & Nottingham, 1999). It can be difficult to reconcile the aim of promoting effective forms of health care with that of promoting patient choice and the rights of women who may decide not to be screened (Austoker, 1999).

Knowledge and understanding of cervical cancer

When making a decision about participation in screening, women should ideally take into account their own understanding of cervical cancer and perception of their risk of it, in addition to their understanding of the risks and benefits of screening.

Survey data about knowledge of risk factors for cervical cancer in diverse groups of women has shown a low prevalence of information. HPV testing is increasingly being incorporated in some cervical screening programmes. Knowledge about HPV is, in general, poor. In a survey of female university staff, 70% of respondents (280/400) reported that they had never heard of HPV infection (Pitts & Clarke, 2002). A survey of university students (of whom only 18% had undergone screening due to the age distribution) reported a very similar percentage of 69% (Philips et al., 2003). Among these students, 51% thought that HPV might increase the risk of cervical cancer. In a series of groups of low-income and minority women in the USA, just over half of the participants had heard of HPV, but they greatly overestimated the risk of developing cancer if infected with the virus (Anhang et al., 2004).

Implementation of HPV testing in primary screening for cervical cancer would result in a large proportion of women having to be told that they harbour a sexually transmitted viral infection that can ultimately cause cancer (see Chapter 2). The potential negative impact of imparting this information to the screening public has not been well assessed from a psychological standpoint.

Among Vietnamese migrants to the USA, 52% identified many sexual partners as a risk factor, 49% identified sexual intercourse at an early age and 59%, incorrectly, thought that cervical cancer was familial (Schulmeister & Lifsey, 1999). Another survey reported that 35% of mainly black South African women, all cancer patients and approximately 90% of medical students and student nurses from the same catchment area had some basic knowledge about cervical cancer (Wellensiek et al., 2002). Similar low proportions of women with limited or no knowledge of cervical cancer and risk factors have been reported in Ghana (Adanu, 2002), Botswana (McFarland, 2003), Kenya (Gichangi et al., 2003) and Nigeria (Ayinde & Omigbodun, 2003).

Knowledge and understanding of cervical screening

While some earlier surveys reported poor general knowledge of cervical screening, (Kennedy, 1989; Schwartz et al., 1989; Nicoll et al., 1991; Nugent & Tamlyn-Leaman, 1992; McKie, 1993a, b; Greimel et al., 1997), more recent research indicates an improvement (Eiser & Cole, 2000; Slater, 2000; Eaker et al., 2001a; Marteau et al., 2001; Idestrom et al., 2002; Pitts & Clarke, 2002; Philips et al., 2003). Additional explanations may help communication: the proportion of UK women who understood the implications of a normal result increased from 52% to 70% after an explanation that ‘normal’ meant that risk was low and not that there was no risk of cancer (Marteau et al., 2001). In a high-resource country (Sweden) where screening is well established, 92% of
survey respondents were aware that cytological testing detects abnormalities in asymptomatic women (Eaker et al., 2001a) and 95% of respondents knew the purpose of screening, although fewer (62%) knew the type of cancer detected by the screening test (Ideström et al., 2002).

Knowledge of the implications of an abnormal test result and the reasons why colposcopy is needed are also not well understood (Nugent & Tamlyn-Leaman, 1992; Onyeka & Martin Hirsch, 2003). As may be expected, understanding was greater among women who had received an abnormal result in the past than among women who had not undergone colposcopy (Pitts & Clarke, 2002); once an abnormality was detected, the perception of personal risk increased (Kavanagh & Broom, 1998).

Bankhead et al. (2003) reviewed 49 studies on beliefs and behaviour related to cervical cancer screening, all observational, and found that adherence to follow-up recommendations after a positive test result ranged from 53 to 75%. Several intervention studies have looked at the provision of information to try to reduce anxiety, with the assumption that lower anxiety would result in better adherence to follow-up recommendations, but have found that provision of information increased knowledge without reducing anxiety or increasing attendance at follow-up.

Informed choice about whether or not to attend for cancer screening was explicitly introduced in England in 2001. New mandatory leaflets explaining the benefits and limitations of screening were to accompany every invitation (http://www.cancerscreening.nhs.uk/news/001.html). Since this policy and new leaflet were introduced, despite concerns in some areas, no change has been observed in acceptance of screening (Department of Health, Statistical Bulletins for 2001/2, 2002/3, 2003/4, available on http://www.cancerscreening.nhs.uk/news/001.html).

Bankhead et al. (2003) commented on the relatively poor quality of studies into health behaviour, attitudes and beliefs with regard to cervical cancer screening, although the trend in most observational studies is towards a beneficial effect.

**Predictors of attendance for screening**

Obtaining the high levels of attendance for screening that are essential to reduce the incidence of cervical cancer has been a major problem in many countries with and without organized screening programmes (Ponten et al., 1995, Lazcano-Ponce et al., 1999a). Besides a high screening coverage of the population at risk, a comprehensive cervical screening programme must also assure maximum return rates among women with abnormal screening results and ensure appropriate care for women requiring follow-up treatment.

Establishing the main determinants of participation is essential to devise effective strategies to increase attendance. These include factors such as the health-care system organization, the socioeconomic level of the population, the costs involved, women’s perceptions of vulnerability, anxiety and fear about cervical cancer, beliefs about the relevance of the test, concurrent family difficulties, and the priority accorded to cervical screening (Austoker, 1994). The relative importance of each of these factors will depend on the specific setting.

Studies of predictors of participation published in the last 10 years are presented in Table 56. Studies carried out among specific social or ethnic groups and qualitative studies were not included. Most studies were carried out in developed countries, with only five from developing countries, and most analysed predictors of participation in cytology-based screening; one analysed determinants of participation in visual inspection-based screening (Sankaranarayan et al., 2003).

Only one of the studies included in Table 56 analysed perceived barriers to screening (Eaker et al., 2001a); it found that time-consuming barriers and economic barriers were associated with non-attendance. The association of emotional barriers was found to be non-significant.

**Socio-demographic factors**

**Age**

Most studies have found that younger women are more likely to attend for screening than older women (Calle et al., 1993; Perez-Stable et al., 1995; Mandelblatt et al., 1999a; Hsia et al., 2000; Chan et al., 2002a; Sankaranarayan et al., 2003).

**Socioeconomic status**

Participation in cervical cancer screening was associated with higher income and educational level in many studies (Calle et al., 1993; Katz & Hofer, 1994; Perez-Stable et al., 1995; Nascimento et al., 1996; Lazcano-Ponce et al., 1997b, 2002; Borras et al., 1999; Hsia et al., 2000; Maxwell et al., 2001; Chan et al., 2002a; Hewitt et al., 2002; O’Malley et al., 2002, Siahpush & Singh, 2002; Selvin & Brett, 2003). For example, Calle et al. (1993) found that 19% of women under the poverty line had never had a screening test compared with only 5.8% of women whose incomes were at least 300% of the poverty level.

Although socioeconomic level may be an important determinant of the ability to pay for preventive services, Katz & Hofer (1994) found that women with higher income and education in the USA and Canada were more likely to have been tested, and that there was no difference between countries, despite Canada’s universal health coverage. The authors suggested that
Use of screening for cervical cancer

Table 56. Predictors of attending for cervical cancer screening

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type</th>
<th>Country</th>
<th>Study population</th>
<th>Key findings</th>
</tr>
</thead>
</table>
| Calle et al. (1993) | Population-based cross-sectional study; reported screening | USA              | 12 252 women aged 18 and older who participated in the National Health Interview Survey in 1987 | Ever having had a test associated with:  
  Demographic: having been married; white/black background; younger than 65; education > 12 years; income above the poverty line. No difference according to rural residence. |
| Lurie et al. (1993) | Cross-sectional. Women enrolled in a large Mid-western health plan | USA              | 24 713 women aged 18 to 75 years old                | Having had a test associated with:  
  Health care: having a female physician. |
| Katz & Hofer (1994) | Population-based cross-sectional survey          | USA/Canada       | 23 521 and 23 932 women aged 18 and older participating in the 1990 Ontario Health Survey and in the 1990 US National Health Interview Survey respectively | Ever having had a test associated with:  
  Demographic: college degree, higher income, no differences between countries. Disparities persisted when the effect of health insurance was controlled. |
| Majeed et al. (1994) | Cross-sectional/correlation study. General practice | England          | 174 724 women aged 25 to 64                          | Having had a test associated with:  
  Demographic: percentage of the practice population under 5 years of age. Overcrowding, age 35–44, and change of address negatively associated with attendance.  
  Health care: female partner. Size of practice and computerization not significant predictors of screening uptake |
| Ronco et al. (1994) | Survey of attenders and non-attenders to a pilot programme in Turin (invitation by GPs) | Italy            | 374 (372 analysed) compliers and 513 (398 analysed) non-compliers aged 25–64 years | Attendance increased with:  
  Demographic: older age, lower education level  
  Health care: having had a test more than three years ago. Receiving an invitation with a pre-fixed appointment.  
  Cognitive: Anxiety |
| Bowman et al. (1995) | Prospective study after invitation              | Australia        | 504 women aged 18–70 from intervention groups       | Attendance associated with:  
  Demographic: Younger age,  
  Health care: oral contraception, and receiving a GP letter.  
  Cognitive: perceive screening as necessary |
| Perez-Stable et al. (1995) | Population-based cross-sectional study; reported screening | USA              | 1242 Latino and Anglo women aged 35–74 years         | Test in the last three years associated with:  
  Demographic: age 35–49; education level > 12 years  
  Ethnicity was not a significant predictor for use of screening in the previous three years |
| Lantz et al. (1997) | Population-based telephone survey               | USA              | 1168 rural Wisconsin women aged 40 years and older  | Having a test in the last three years associated with:  
  Demographic: being married  
  Health care: having health insurance, having a regular physician, having seen a health practitioner in the past year, and having a physician make a recent recommendation for a test  
  The association with income and education was non-significant. Having a hysterectomy and perceiving screening tests to involve physical discomfort were negatively associated. |
### Table 56 (contd)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type</th>
<th>Country</th>
<th>Study population</th>
<th>Key findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Malley et al. (1997)</td>
<td>Population-based cross-sectional study; reported screening</td>
<td>USA</td>
<td>1420 women aged 18 years and older from four Hispanic groups and three black groups</td>
<td>Health care: Ever and recent screening associated with usual site of care</td>
</tr>
<tr>
<td>Potosky et al. (1998)</td>
<td>Population-based cross-sectional survey. Reported screening</td>
<td>USA</td>
<td>9455 women 18 years and older participating in the National Health Interview Survey in 1992</td>
<td>Screening in the past three years was associated with: <strong>Health care:</strong> having health insurance.</td>
</tr>
<tr>
<td>Segnan et al. (1998)</td>
<td>Prospective study after intervention</td>
<td>Italy</td>
<td>8385 women aged 25–64 years</td>
<td>Higher attendance when age 45–54, married, receiving an invitation from GP, higher education level</td>
</tr>
<tr>
<td>Borras et al. (1999)</td>
<td>Population-based cross-sectional survey, reported screening</td>
<td>Spain</td>
<td>5865 women aged 20 years and older participating in the Catalan Health Survey in 1994</td>
<td>Ever having a test associated with: <strong>Demographic:</strong> younger age, higher educational level <strong>Health care:</strong> enrolled in voluntary health insurance</td>
</tr>
<tr>
<td>Mandelblatt et al. (1999a)</td>
<td>Population-based telephone survey; reported screening</td>
<td>USA</td>
<td>1420 women aged 18–74 from four Hispanic groups and three black groups in New York city</td>
<td>Ever and recent test associated with: <strong>Demographic:</strong> younger age, higher education <strong>Health care:</strong> regular source of care, health insurance <strong>Cognitive:</strong> positive cancer attitudes</td>
</tr>
<tr>
<td>Hsia et al. (2000)</td>
<td>Cross-sectional study; reported screening</td>
<td>USA</td>
<td>55 278 women aged 50–79 participating in the Women’s Health Initiative Observational Study in 1994–97</td>
<td>Test in the last three years associated with: <strong>Demographic:</strong> younger age, higher education; higher income, married status, ethnic background <strong>Health care:</strong> usual care provider, medical visit in the past year; health insurance <strong>Health status:</strong> Presence of some chronic medical conditions, smoking.</td>
</tr>
<tr>
<td>Eaker et al. (2001a)</td>
<td>Population-based telephone survey; validated reported screening</td>
<td>Sweden</td>
<td>430 non-attenders and 514 attenders aged 25–59 years.</td>
<td>Attendance associated with (5 years for women 30–59 and 3 years for 25–29 years): <strong>Cognitive:</strong> perceived satisfactory benefits. <strong>Access:</strong> time-consuming barriers and economic barriers associated with non-attendance Association between perceived severity of cancer, social support and anxiety non-significant</td>
</tr>
<tr>
<td>Eaker et al. (2001b)</td>
<td>Population-based telephone survey; validated reported screening</td>
<td>Sweden</td>
<td>430 non-attenders and 514 attenders aged 25–59 years</td>
<td>Attendance associated with (5 years for women 30–59 and 3 years for 25–29 years): <strong>Demographic:</strong> living in urban areas. Socioeconomic status was not determinant of participation <strong>Health care:</strong> oral contraception, visits to a physician 1–5 times/year or less than once a year; seeing the same gynaecologist; having a test on their own initiative <strong>Cognitive:</strong> Knowing the recommended screening interval</td>
</tr>
<tr>
<td>Maxwell et al. (2001)</td>
<td>Population-based cross-sectional survey</td>
<td>Canada</td>
<td>23 287 women aged 25–64 years participating in the National Population Health Survey 1996–1997</td>
<td>Screening associated with: <strong>Demographic:</strong> being non-migrant, ever married, higher education, spoken language: English, younger age <strong>Health behaviour:</strong> having a recent blood pressure check and frequent physical activity <strong>Health care:</strong> regular physician</td>
</tr>
</tbody>
</table>
Use of screening for cervical cancer

**Table 56 (contd)**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type</th>
<th>Country</th>
<th>Study population</th>
<th>Key findings</th>
</tr>
</thead>
</table>
| Coughlin et al. (2002) | Population-based cross-sectional study; reported screening | USA     | 131,813 women aged 18 years and older who participated in the Behavioral Risk Factor Surveillance System in 1998–99 | Having had a test in the last three years associated with:  
Demographic: higher education, higher income, increased level of urbanization, fewer than three persons in the household, being currently married and White/Hispanic/black background. Women aged 30–39 years were more likely to have had a recent test.  
Health care: health insurance coverage.  
Health status: having seen a physician in the past year, good or excellent health status, non-smoking and no alcohol consumption.                                                                                           |
| Hewitt et al. (2002) | Population-based cross-sectional study; reported screening | USA     | 10,847 women aged 15–44 years who participated in the National Survey of Family Growth in 1995 | Having had a test in the last year associated with:  
Demographic: family incomes at or above 300% of the poverty level, at least a college degree; not having a non-Hispanic black background.  
Health care: health insurance.  
Cervical cancer risk: Having at least one of five risk factors for cervical cancer                                                                                                                                                                                                                       |
| O’Malley et al. (2002) | Population-based telephone survey; reported screening | USA     | 1205 women 40 years and older | Reported receipt of the last two screening tests within the recommended intervals for age:  
Demographic: higher income  
Health care: Women under 65: continuity of care, primary care more comprehensive, counselling, patient–physician relationship, trust  
Older than 64: knowledge about cervical cancer, ever married, continuity of care, counselling, patient–physician relationship                                                                                                                                                                                                                       |
| Siahpush & Singh (2002) | Population-based cross-sectional study | Australia | 7572 women aged 18–69 years | Ever having a test associated with:  
Demographic: age 30–49, being married, being born in Australia or New Zealand and high education level. Socio-economic level was not a predictor of participation.                                                                                                                                                                                                                                 |
| Selvin & Brett (2003) | Population-based cross-sectional study; reported screening | USA     | 5509 women aged 40–64 years who participated in the 1998 National Health Interview Survey | Test within three years before interview associated with:  
Demographic: Income above or at 200% of poverty level (except for black women), bachelor’s degree or higher  
Health care: Having a usual source of care; private health insurance  
Health behaviour: non-smoking except for non-Hispanic black  
Residential status, self-reported health and marital status not significant predictors of attendance                                                                                                                                                                                                                     |
| **Developing countries** | | | |                                                                                                                                                                                                                                                                |
| Nascimento et al. (1996) | Population-based cross-sectional survey; reported screening | Brazil   | 967 women aged 15–59 years | Ever having had a test associated with:  
Demographic: older than 24, higher education, higher income, ever married  
Health care: consulted a physician in the year preceding the survey, oral contraceptive use, and performed breast self-examination                                                                                                                                                                                                                     |
| Lazcano-Ponce et al. (1997) | Population-based cross-sectional survey; reported screening | Mexico   | 4208 women aged 15–49 years from Mexico City or in selected rural areas of the state of Oaxaca | Ever having had a test associated with:  
Demographic: higher education level; higher socioeconomic level; living in urban areas  
Health care: access to social security health care  
Cognitive: knowledge of what the Pap test is used for  
Health status: having seen a physician in the past year, good or excellent health status, non-smoking and no alcohol consumption.                                                                                                                                                                                                                     |
Table 56 (contd)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type</th>
<th>Country/Study population</th>
<th>Key findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chan et al. (2002)</td>
<td>Population-based cross-sectional survey; reported screening</td>
<td>Hong Kong 2067 women aged 44–55 years</td>
<td>Test in the previous 12 months associated with: Demographic: Younger age, education level higher than primary level. Health care: receiving hormone treatment; breast self-examination performed. Health status: premenopausal status; chronic disease.</td>
</tr>
<tr>
<td>Lazcano-Ponce et al. (2002)</td>
<td>Population-based cross-sectional survey</td>
<td>Mexico 2094 women aged 15–49 years with a CCSP (National Cervical Cancer Screening Programme) Pap test history</td>
<td>Increased screening associated with: Demographic: higher educational level of the head of the family. Health care: history of using two or more family planning methods. Cognitive: knowing why the screening test is employed; good experience of screening quality.</td>
</tr>
<tr>
<td>Sankaranarayanan et al. (2003b)</td>
<td>Prospective study of women invited for screening; objective measure of screening</td>
<td>India (rural) 48 225 women aged 30–59 years</td>
<td>VIA-based screening associated with: Demographic: Younger age, higher educational level, being married, multiparous status and low-income level. Health care: having had tubal sterilization for birth control.</td>
</tr>
</tbody>
</table>

Factors linked to recruitment and service delivery should also be taken into account in explaining socioeconomic differences. For example, in rural India, women with higher income levels were less likely to participate in the visual-inspection-based screening (Sankaranarayanan et al., 2003). The authors considered that the fact that the screening clinics were organized in public institutions such as health centres or schools might have deterred high-income women from attending.

Marital status
Married, divorced and widowed women were more likely than single women to attend for screening (Calle et al., 1993; Nascimento et al., 1996; Lantz et al., 1997; Segnan et al., 1998; Hsia et al., 2000; Maxwell et al., 2001; Siahpush & Singh, 2002).

Rural residence
Most research suggests that women living in urban areas are more likely to attend for screening (Eaker et al., 2001b; Coughlin et al., 2002). However, in some European countries, higher participation is seen in rural areas than in large urban centres.

Ethnicity
The evidence on the influence of ethnicity in screening attendance is not conclusive. In the USA, for example, Calle et al. (1993) found that women with white or black ethnic background were more likely to be screened and Coughlin et al. (2002) found that Hispanic women were as likely as white women to be screened. Perez-Stable et al. (1995), in their study comparing Latino and ‘Anglo’ (non-Latino white) women found no significant effect of ethnicity and Hsia et al. (2000) found that only Asian/Pacific islander women were less likely to participate. It is important to note that ethnicity is often a proxy of socioeconomic status, particularly in the USA.

Health status
Some studies have shown that attendance was higher among women with good or excellent health status (Coughlin et al., 2002), while others have found that women with chronic diseases were more likely to be screened (Chan et al., 2002a). In the USA, Mandelblatt et al. (1999a) found that both younger women in poor health and elderly women in good health were more likely to have ever had or to recently have had a test.

Interactions with the health system
Most studies have shown that contacts with the health-care system increase the likelihood of a woman being screened (Lantz et al., 1997; O’Malley et al., 1997; Mandelblatt et al., 1999a; Eaker et al., 2001b; Hsia et al., 2000; Maxwell et al., 2001; Coughlin et al., 2002; Lazcano-Ponce et al., 2002; O’Malley et al., 2002). For example, Maxwell et al. (2001) reported that, in Canada, having had a medical consultation in the past year and having a last blood pressure check less than two years ago were important predictors of cervical cancer screening. A contact with the health-care system seems to be one of the main determinants of attendance also in developing
countries (Nascimento et al., 1996; Sankaranarayanan et al., 2003b). For example, in India, both a greater number of children and use of family planning methods were associated with higher participation (Sankaranarayanan et al., 2003b). A previous contact with gynaecological and maternal services may increase awareness about gynaecological procedures and encourage further contacts with health-care services, making women more responsive to screening. Having a regular source of care was also linked to higher attendance in many studies (Lantz et al., 1997; Mandelblatt et al., 1999a; Eaker et al., 2001b; Maxwell et al., 2001; O’Malley et al., 2002). Indeed, Lantz et al. (1997) found that having a regular source of care was a main predictor of screening, even after controlling for other health-care access factors such as health insurance.

Aspects related to the patient–physician relationship and contact with the health-care system appear to be important determinants of attendance. The probability of attending screening was higher in women who reported good experiences with the health system. For example, O’Malley et al. (2002) reported that women who evaluated primary care as more comprehensive, and their relation with their physician based on trust, were also more likely to be screened. In Mexico, women with good previous screening experiences were four times more likely to be re-screened (Lazcano-Ponce et al., 2002). In El Salvador, a pilot study that implemented a quality-improvement process resulted in screening 25% of women aged 30–59 years who had never been screened. Another qualitative study summarizing the experiences of research projects in Bolivia, Peru, Kenya, South Africa, and Mexico carried out by the Alliance for Cervical Cancer Prevention (ACCP) reported that women expressed the need for confidentiality and privacy. Women commonly report feeling ashamed, especially when privacy is lacking or when male providers perform the examination (Bingham et al., 2003).

The effect of the gender of the physician was examined in two studies and in both it appeared that women having a female physician were more likely to be screened (Lurie et al. 1993; Majeed et al., 1994). A recommendation by a doctor to attend screening also appeared to have an important influence on women’s decision to be screened (Bowman et al., 1995; Lantz et al., 1997; Segnan et al., 1998). For example, in the USA, women who received a physician’s recommendation for screening were 2.3 times more likely to be screened (Lantz et al., 1997). Ronco et al. (1994) reported that if the recommendation included a fixed appointment, women were more likely to attend than if the invitation did not.

Among the factors related to access to health care, health insurance appears to be one of the most important predictors of screening, as most studies have found that having health insurance increased the likelihood of participation (Katz & Hofer, 1994; Lantz et al., 1997; Potosky et al., 1998; Borras et al., 1999; Lazcano-Ponce et al., 1997b; Hsia et al., 2000; Selvin & Brett, 2003; Hewitt et al., 2002). There is limited evidence related to other access factors such as distance to a health-care centre and cost of transport. The introduction of mobile clinics in rural Thailand increased participation in a cytology-based screening from 21 to 57% (Swaddiwudhipong et al., 1995, 1999).

Knowledge and attitudes as predictors of attendance

Knowledge of screening

All of the studies included in Table 56 that analysed the effect of knowledge (Bowman et al., 1995; Mandelblatt et al., 1999a; Eaker et al. 2001b; Lazcano-Ponce et al., 1997b, 2002; O’Malley et al., 2002) found that knowledge about the screening test increased the probability of screening. For example, in Mexico, women who knew why the test was given were three times more likely to be screened (Lazcano-Ponce et al., 2002). In Sweden, knowing the recommended screening interval increased the probability of attendance (Eaker et al., 2001b). Also, women are more likely to attend screening if they perceive screening as necessary or beneficial (Bowman et al., 1995; Eaker et al., 2001a). There is no evidence that awareness of risk influences women’s decisions on whether to be screened.

Fear/anxiety

In most studies, increased anxiety was associated with lower probability of women attending for screening. In Italy, Ronco et al. (1994) found that anxiety caused by periodic controls was an important negative determinant of compliance. In the USA, positive attitudes to cancer (less anxiety and hopelessness and a lower level of denial) were key determinants of participation (Mandelblatt et al., 1999a). Extensive qualitative research on reasons for non-attendance indicates that both fear of cancer and anxiety and knowledge of screening are key barriers to screening. A recent in-depth qualitative analysis in five Latin American countries indicated ‘fear of cancer’ as an underlying reason for women not to seek screening services (Agurto et al., 2004). The authors suggest that this aspect is articulated by women in different forms, such as poor knowledge and understanding or not having time, depending on the questions of the survey; but when women are prompted to explain further they consistently alluded to ‘fear of cancer’.
Strategies to encourage participation in cervical cancer screening programmes

Many approaches to increase participation in screening have been analysed (Table 57). Studies differ in the populations addressed, the settings where they were implemented and the methods used to test the approach. Even when two studies test the same strategy, for example, an invitation by letter, it is not possible to control for all the variables that may have an influence on women’s participation, such as the style and tone of the letter, the actual service being offered, who signs the letter and so on. In general, limited information is available about these methodological differences, but it is clear that apparently similar approaches can differ in various aspects that might affect participation.

The effectiveness of a strategy will depend on how the health system is organized; thus, an invitation letter may be effective if the service is provided free of charge, but may not be in certain settings if women have to pay for the test. The types of intervention that have been investigated include strategies targeting individual women, healthcare providers and communities.

Strategies targeting individual women

Invitation letters

Of the 29 studies presented in Table 57, 15 evaluated the effectiveness of invitation letters compared with a control group with no intervention (McDowell et al., 1989; Pierce et al., 1989; Clementz et al., 1990; Ornstein et al., 1991; Lancaster & Elton, 1992; Bowman et al., 1995; Lantz et al., 1995; Pritchard et al., 1995; Binstock et al., 1997; Buehler & Parsons, 1997; Somkin et al., 1997; Burack et al., 1998; Del Mar et al., 1998; Torres Mejia et al., 2000; Vogt et al., 2003). Invitation letters significantly increased screening uptake in 10 of these studies. Two studies found a non-significantly better participation in the control group (Clementz et al., 1990; Del Mar et al., 1998). In the one by Clementz et al. (1990), subjects were offered up to seven other screening tests in the same letter (including faecal occult blood test, digital rectal examination, sigmoidoscopy, pelvic bimanual examination, breast examination and mammography), so it is difficult to draw conclusions related only to cervical cancer screening, as the reasons for low attendance may be related to how patients react when invited for several screening tests. Additional evidence from programme evaluations (not included in Table 57) confirms the importance of the invitation letter. For example, a screening programme in Denmark issued personal invitations to target women over a period of 15 years. The participation among women aged 30–50 was 91% (Lynge et al., 1992). After that, personal invitations were stopped and participation in the same age group dropped to 66%.

The context of the trial and the characteristics of the letter varied across the studies in Table 57. For example, in two studies (Burack et al., 1998; Vogt et al., 2003), the letters were accompanied by educational brochures. In six studies, letters were followed by mailed or phone reminders (McDowell et al., 1989; Ornstein et al., 1991; Lantz et al., 1995; Buehler & Parsons, 1997; Torres Mejia et al., 2000; Vogt et al., 2003). In the study by Vogt et al. (2003), the increase in attendance was significant only when the letter was followed by a phone reminder; a letter followed by a mailed reminder led to no significant increase.

The two studies that analysed the effect of sending a letter to minority group women found no effect of the intervention (Del Mar et al., 1998; Hunt et al., 1998). However, in one of these (Del Mar et al., 1998), the letter followed a media campaign introduced two months before in the whole region, so an effect of contamination between the intervention and control groups cannot be eliminated.

The only study conducted in a developing country found significantly higher participation among women who received an invitation letter (Torres Mejia et al., 2000). The project was carried out in the context of the Mexican Social Security Institute, which provides medical services for nearly 60% of the Mexican population.

In the total group of women invited, the effectiveness of the intervention was 20.1% versus 3.3% in the control group. However, it may be difficult to adopt this strategy in other developing countries due to the lack of mailing lists, ineffective or nonexistent postal systems and women’s difficulties in reading or understanding letters.

Two studies used invitation letters signed by different authority sources. Significantly better participation was observed in the intervention groups receiving letters signed by GPs versus letters signed by female nurse practitioners (Bowman et al., 1995) or by programme coordinators (Segnan et al., 1998).

Three studies explored the effect of including a fixed appointment in the letter versus an open invitation to make an appointment (Wilson & Leeming, 1987; Pritchard et al., 1995, Segnan et al., 1998); all found a favourable effect of a fixed appointment. For example, in Italy, the overall compliance with screening was 36.1% when the letter signed by the GP included a fixed appointment, but 22.7% when it only included a prompting to contact the screening centre to make an appointment (Segnan et al., 1998).

In three studies, invitation letters were compared with phone invitations. Binstock et al. (1997) found that telephone invitations were more effective, whereas in the study by McDowell et al. (1989) invitation letters were more effective. However, the latter study included a reminder 21 days later if the
### Table 57. Studies evaluating interventions to increase cervical cancer screening attendance

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type</th>
<th>Country</th>
<th>Study population</th>
<th>Key findings</th>
</tr>
</thead>
</table>
| Wilson & Leeming (1987), UK | Screening uptake      | National Screening Programme; women aged 45–65 years with no record of having a previous smear | a. Letter of invitation to make an appointment + two reminders, \( N = 125 \) (122 analysed)  
b. Letter with an appointment date + two reminders. \( N = 125 \) (118 analysed) | a. 32%  
b. 47%  
Significant differences |
| McDowell et al. (1989), USA | Test during study year | Hospital; no test in past year | a. GP letter and reminder letter after 21 days. \( N = 367 \)  
b. Physician reminder. \( N = 332 \)  
c. Telephone call. \( N = 377 \)  
d. Control group. \( N = 330 \) | a. 25.9%  
b. 16.1%  
c. 20%  
d. 13.7%  
Significant increase only for the GP letter |
| Pierce et al. (1989), UK   | Test during study year | General practice; women registered with a general practice eligible for a test | a. Letter asking women to have a smear. \( N = 140 \)  
b. Physician reminder. \( N = 142 \)  
c. Control group. \( N = 134 \) | a. 32%*  
b. 27%*  
c. 15%  
*p \( <0.01 \)  
Differences between intervention groups NS |
| Robson et al. (1989), UK   | Test within preceding three years | General practice | a. Patients had open access to health promotion nurse and had their risk factors assessed and followed up by both their GP and the nurse. \( N = 799 \)  
b. Control, usual care (i.e. managed by GP alone). \( N = 806 \) | a. 76%  
b. 49%  
p \( <0.001 \) |
| Clementz et al. (1990), USA | Test up to four months after the intervention | Female patients attending ambulatory clinic; aged 50–69 years | a. Personalized GP’s letter, one month before due date of tests with an educational component. \( N = 102 \)  
b. Control group received usual care (not described). \( N = 76 \) | a. 20.6%  
b. 30.3%  
NS differences |
| McAvoy & Raza (1991), UK   | Test within four months after the intervention | National screening programme; Asian women resident in Leicester, aged 18–52 years with no record of having had a test | a. Home visit and a multilingual video. \( N = 263 \)  
b. Home visit, multilingual leaflet and fact sheet. \( N = 219 \)  
c. Posted multilingual leaflet and fact sheet. \( N = 131 \)  
d. Control group received no intervention. \( N = 124 \) | a. 30%*  
b. 26%*  
c. 11%#  
d. 5%  
*Difference with control group significant  
#Difference with control group NS |
| Ornstein et al. (1991), USA | Screening uptake      | Family medicine clinic; women aged 18 years and over; not screened in previous 2 years; active patient of the family medicine centre (i.e. had visited clinic in previous 2 years) | a. Physicians received computerized reminders. \( N = 1988 \) participants; 14 physicians  
b. Participants were sent and invitation to attend followed by another personalized reminder letter (6 months later). \( N = 1925 \) participants, 12 physicians  
c. Both physicians and participant reminders. \( N = 1908 \) participants, 13 physicians  
d. Control group, no intervention. \( N = 1576 \) participants, 10 physicians | Slight decline in intervention group. |
<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type</th>
<th>Country</th>
<th>Study population</th>
<th>Key findings</th>
</tr>
</thead>
</table>
| Ward *et al.* (1991), Australia   | Test up to one month after the initial consultation | General practice; women aged 20–65 years; provided consent               | a. Minimal intervention: GP advised eligible women of need for test and offered to perform it immediately. Those not consenting advised to make appointment for test within a week. *N = 99*  
  b. Maximal intervention: GP advised women of need for test and offered to perform it immediately; GP attempted to persuade those not consenting during that consultation by exploring barriers and reasons for self-exclusions. If still did not consent, GP advised making an appointment for test within a week. *N = 103*  |
| Lancaster & Elton (1992), UK      | Test during a six week period                   | General practice; women aged 50–64 years; resident in study area         | a. Cervical screening invitation sent with breast screening invitation. *N = 474*  
  b. Breast screening invitation only sent (control). *N = 483*  |                                                                                                                                         |
| Bowman *et al.* (1995), Australia | Screening uptake at six months after the intervention | Community and general practice. Women aged 18–70 years who reported ever having sexual intercourse and who had not had a test in the previous three years if a complete set of information was available | a. GP reminder letter. *N = 220* (178 analysed)  
  b. Women’s health clinic invitation for screening by a female nurse practitioner. *N = 220* (164 analysed)  
  c. Mailed educational pamphlet personally addressed to women. *N = 219* (162 analysed)  
  d. Control group (not stated) *N = 219* (155 analysed).  
  Note: analysis carried out on women who responded to the follow-up survey  | a. 36.9%  
  b. 22.6%  
  c. 25.9%  
  d. 24.5%  
  Letter from the general practitioner significantly increased pap smear uptake. Differences between other groups and control group NS. |
| Lantz *et al.* (1995), USA         | Screening uptake 6 months after the intervention | Community health centre; women aged 40–79 years, enrolled in benefit scheme* no claim for Pap test in past 3 years | a. Reminder letter from primary-care physician for test(s) required. Follow-up phone call/letter from a health educator (nurse or social work intern) 7–10 days later, to offer barrier counselling and/or assistance with appointment making. *N = 337*  
  b. Control group received ‘usual care’ (not described). *N = 322*  | a. 19%  
  b. 6%  
  [values recalculated from original data]  
  Significant difference |
| Pritchard *et al.* (1995), Australia | Test within 12 months of entry into the study   | General practice; women aged 36–69 years at a university general practice in a socio-economically disadvantaged area of Perth | a. Physician reminder (tagged notes). *N = 198*  
  b. Letter with invitation to make an appointment. *N = 206*  
  c. Letter with fixed appointment. *N = 168*  
  d. Control group (usual care). *N = 185*  | a. 21.2%*  
  b. 25.7%#  
  c. 30.4%#  
  d. 16.8%  
  *NS difference with the control group  
  # Difference with the control group statistically significant. |
| Yancey *et al.* (1995), USA        | Test within 5 months after the intervention     | Health clinic; women attending one of the two study clinics.              | a. Culturally sensitive health education videos dealing with breast and cervical cancer played in waiting room. *N = 868*  
  b. Control, no intervention. *N = 876*  | a. 19.4%  
  b. 13.7%  
  *NS difference with the control group  
  p < 0.05  
  .  |
### Table 57 (contd)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type</th>
<th>Country</th>
<th>Study population</th>
<th>Key findings</th>
</tr>
</thead>
</table>
| Binstock *et al.* (1997), USA | Screening uptake at 12 months after intervention | Health maintenance organization. Women aged 25–49 years, enrolled for three years at the Kaiser Permanente Health Plan who were likely to seek out-patient care at one of the three medical centres | a. Telephone call. *N* = 1526  
  b. Letter. *N* = 1526  
  c. Memo to woman's primary provider. *N* = 1526  
  d. Chart reminder affixed to outside of woman's medical record. *N* = 1526  
  e. No intervention (control group). *N* = 1526 | Differences with control group significant for all groups. *p* < 0.001  
  No significant differences between b, c and d |
| Buehler & Parsons (1997), Canada | Screening uptake at 6 months after intervention | Family medicine clinic. Patients of the clinics aged 18–69 years who had not had a test in the previous three years | a. Personal letter and reminder letter 4 weeks later (letter head of the provincial cytology registry signed by co-investigators). *N* = 178  
  b. Control group not received letter. *N* = 208 | a. 10.7%  
  b. 6.3%  
  NS differences |
| Somkin *et al.* (1997), USA | Test in the six months following the intervention | Health maintenance organization (HMO); women aged 20–64 years, no test in previous 36 months; residents of study area; continuously enrolled as a member of HMO for the previous 36 months | a. Letter signed by a physician inviting women to make an appointment. *N* = 1188  
  b. Letter signed by a physician inviting women to make an appointment and chart. Providers encouraged by presentations by researchers and memoranda describing the project. *N* = 1188  
  c. Usual care (required a referral from a physician). *N* = 1188. | a. 19.4%*  
  b. 22.8%*  
  c. 9.1%  
  Difference with control group significant. *p* < 0.01 |
| Sung *et al.* (1997) | Test up to six months after completion of the intervention | Community; African American women; aged 18 years and older | a. Lay health workers visited women three times to provide a culturally sensitive educational programme emphasizing need for screening through printed material and video. *N* = 163  
  b. Control group received educational information on completion of follow-up | No significant increase |
| Burack *et al.* (1998), USA | Screening uptake during the study year | Health maintenance organization (HMO) Age 18–40 years; HMO member; visited one of the primary care study sites in the previous years; had not had a test in the previous year. | a. Patient reminder (invitation letter signed by the HMO director) + National Cancer Institute educational brochure. *N* = 964  
  b. Reminder for physician. *N* = 960  
  c. Reminders for both physician and participants. *N* = 960  
  d. Control (no reminder). *N* = 964 | a. 29%  
  b. 29%  
  c. 32%  
  d. 28% Differences with control group NS |
| Margolis *et al.* (1998), USA | Screening uptake 1 year after the intervention | Community health centre; women aged 40 years and over attending appointments in the clinics; had not had a test in the previous year. | a. Lay health workers assessed screening status and offered women screening with a female nurse practitioner. *N* = 566 (470 analysed)  
  b. Usual care group. *N* = 536 (437 analysed) | a. 63.2%  
  b. 50.3%  
  *p* < 0.002 |

Reference: Binstock *et al.* (1997), USA

Buehler & Parsons (1997), Canada

Somkin *et al.* (1997), USA

Sung *et al.* (1997)

Burack *et al.* (1998), USA

Margolis *et al.* (1998), USA
<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type</th>
<th>Country</th>
<th>Study population</th>
<th>Key findings</th>
</tr>
</thead>
</table>
| Segnan et al. (1998), Italy| Screening uptake at 12 months after the initial invitation | General practice in national screening programme; women aged 25–64 years; resident of Turin | a. Personal letter signed by GP with prefixed appointment (control). N = 2100  
 b. Personal letter, signed GP with open-ended appointment. N = 2093  
 c. Personal letter signed by programme coordinator with prefixed appointment. N = 2094  
 d. Personal letter with extended text signed by GP with prefixed appointment. N = 2098 | a. 36.1%  
 b. 22.7%*  
 c. 30.9%*  
 d. 36.7%  
 Personal invitation letters signed by the GP with prefixed appointments induced a significant increase in compliance with screening, p value not provided  
 * Significant difference with group (a) |
| Rimer et al. (1999), USA   | Screening uptake 16 months after the intervention | Community health centre, women aged 18 years or over; client of medical centre who had visited centre in previous 18 months (care for black, low income, low education population), had not had a test in the last year | a. Provider prompting intervention only. N = 202  
 b. Provider prompting and tailored educational print communications (Healthy Birthday cards) N = 204  
 c. Provider prompting, tailored educational print communications and tailored telephone counselling. N = 213 | a. 56%  
 b. 52%  
 c. 64%  
 Provider prompting, tailored educational print communications and tailored telephone counselling induced increased compliance, p = 0.05 |
| Allen et al. (2001), USA   | Screening uptake within past three years        | Worksie; women 40 years and older who were employed for more than 15 hours per week on a permanent basis | a. Intervention sites: Voluntary advisory boards with worker participation; peer health advisors (PHA); group sessions led by PHAs; one-to one outreach activities, worksite-wide health education; other events. N = 1489  
 b. Control sites: no specific activities. N = 1308 | a. 89.9%  
 b. 87.7%  
 Significant difference |
| Vogt et al. (2003), USA    | Test within 12 weeks after the intervention     | Health-care organization; women aged 18–70 who had at least 3 years of continuous membership before the study who had not received a test during the same 3-year period | a. Letter from the programme and brochure followed by a second letter to women who had not had an examination six weeks later after the first contact. N = 206  
 b. Letter as in the first group followed by a phone call to all those not screened in the first six weeks. N = 113  
 c. Phone call followed by a second phone call to all those not screened in the first six weeks. N = 88  
 d. Usual care. N = 280 | a. 22%*  
 b. 54%#  
 c. 50%#  
 d. 17%  
 #p <0.0001  
 *p = 0.16  
 Interventions were more effective than usual care except for the letter/letter only intervention. Letter with phone was as effective as phone/phone outreach |
| Del Mar et al. (1998), Australia | Screening uptake one year after the intervention | Community; women aged 18–67 years; Vietnamese | a. Personal letter in Vietnamese informing them about screening and its benefits. N = 359  
 b. Control group did not receive a letter. N = 330 | a. 10%  
 b. 12%  
 NS differences |
**Use of screening for cervical cancer**

### Table 57 (contd)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type</th>
<th>Country</th>
<th>Study population</th>
<th>Key findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Minority groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| **Navarro et al. (1998), USA** | Screening uptake 3 months after the intervention | Latino community, 18 years and older | a. Por La Vida (PLV) programme with community workers (consejeras) taking 12 weekly educational sessions with the groups of women. N = 274; analysed 199  
   b. Control, no PLV programme; instead consejeras participated in a community living skills programme. N = 238; analysed 162 | a. 65.3%  
   b. 61.1%  
   NS differences  
   Differences in increase between the control and the intervention NS                                                                 |
| **Hunt et al. (1998), Australia** | Screening uptake three months after the intervention | Community health centre; aboriginal women seen more than twice in the past three years and had no record of hysterectomy, had not had a test in the last 3 years | a. Personal approach. Women approached by aboriginal health workers and invited for screening. N = 119  
   b. Letter. Designed by aboriginal workers stating individual overdue for test and inviting them to attend. N = 125  
   c. Control. Usual care with reminder tags for clinic staff attached to medical records. N = 122 | a. 6.7%  
   b. 2.4%  
   c. 0  
   NS differences.                                                                 |
| **Taylor et al. (2002a), USA** | Screening uptake                  | Community; Chinese women 20–69 years of age; spoke Cantonese, Mandarin or English, had no history of invasive cervical cancer and were under-utilizers of screening | a. Culturally appropriate outreach worker intervention. Health education, video, motivational pamphlet, educational brochure and a fact sheet. Home visits, tailored counselling by outreach workers. N = 161; analysed = 129  
   b. Direct mail with a cover letter, the education video, motivational pamphlet, educational brochure and fact sheet. N = 161; analysed = 139  
   c. Control. Usual care with reminder tags for clinic staff attached to medical records. N = 160; analysed = 134 | a. 39%  
   b. 25%  
   c. 15%  
   a versus c, p <0.001  
   b versus c, p = 0.03  
   a versus b, p = 0.02                                                                 |
| **Taylor et al. (2002), USA** | Screening uptake in the last twelve months. | Community; Cambodian refugees in Seattle; women aged 18 and older | a. Control: usual care. N = 160; analysed = 134  
   b. Introductory mailing, home visit including educational video and tailored counselling; group meetings. N = 144  
   b. Control: usual care. N = 145 | a. 61%  
   b. 62%  
   NS differences                                                                 |
| **Developing countries**   |                                 |                                                    |                                                                                 |                                                                                             |
| **Torres Mejia et al. (2000), Mexico** | Screening uptake up to 8.5 weeks after the intervention | Mexican Social Security Clinics (Morelos). Women 20–64 years who had not had a test in the previous 12 months | a. Intervention: Letter of invitation and reminder. N = 2119  
   b. Control group not received letter. N = 2100 | a. 20.1%  
   b. 3.3%  
   Significant differences                                                                 |
women did not respond to the first letter, whereas in the former no reminder was sent after the phone call. When a phone call was followed by a second phone call to all women not screened in the first six weeks, it was more effective than a letter followed by a mailed reminder (Vogt et al., 2003). These three studies also compared an invitation phone call and no intervention. In all three studies, the uptake of screening was higher in the phone call group, but not significantly in the study by Binstock et al. (1997). In the study by Vogt et al. (2003) the phone–phone approach was as effective as a letter–phone approach. However, the letter followed by a phone call was the most cost-effective approach. It was estimated that the phone–phone approach produced one additional screening for $305 versus $185 with the letter–phone approach.

**Personal approach**

The efficacy of a personal contact was evaluated in six studies (McAvoy & Raza, 1991; Sung et al., 1997; Hunt et al., 1998; Margolis et al., 1998; Taylor et al., 2002a, b). The face-to-face approach significantly increased screening uptake in three of them (McAvoy & Raza, 1991; Margolis et al., 1998; Taylor et al., 2002a). However, the conditions under which the personal contact took place varied enormously from study to study. In the study by Margolis et al. (1998), lay health workers approached women attending a community health centre and offered screening with a female nurse practitioner. The effectiveness of this approach was evaluated in relation to usual care; 63.2% of the invited women complied with screening versus 50.3% in the control group. However, 36% of randomized women cancelled or missed appointments, so they were not contacted at all and were not included in the final analysis. With a strategy based on approaching women attending a health centre, non-users will never be reached. Taylor et al. (2002a) evaluated home visits that included delivering educational material (a health education video, a motivational pamphlet, an educational brochure and fact sheet) and providing tailored counselling, in comparison with only mailing the educational material and with a control group receiving usual care. The first approach was the most effective in increasing participation rates. However, no data on costs were provided; as up to ten attempts were made to contact each woman in the first group, it would presumably have been the most expensive.

**Educational interventions**

Ten studies evaluated different types of educational intervention including printed material (Rimer et al., 1999; Bowman et al., 1995), video/slide presentations (Yancey et al., 1995), face-to-face contacts (Navarro et al., 1998; Allen et al., 2001) or combinations of various educational approaches (Taylor et al., 2002a, b; McAvoy & Raza, 1991; Sung et al., 1997). Mailed printed materials do not appear to increase uptake of screening. For example, in Australia, Bowman et al. (1995) found that mailed educational pamphlets personally addressed to women did not increase uptake compared with a letter from the GP without the educational pamphlet or a control group (no intervention). Educational video tapes were found to be effective compared with no intervention, both when mailed (Taylor et al., 2002a) and when played in a health-care setting. For example, video presentations played in the waiting room of health clinics increased uptake of screening among women attending the clinics in the USA by around 30% (Yancey et al., 1995). The effectiveness of face-to-face educational interventions seems to be low, as no (Navarro et al., 1998) or modest (Allen et al., 2001) effect was found in the two studies included. However, the face-to-face approaches differed greatly. Navarro et al. (1998) compared the effect of 12 educational sessions on cervical cancer screening with the effect of 12 educational sessions about living in the community. Allen et al. (2001) invited women to a group meeting carried out in their workplace. Multi-component interventions seemed to be the most effective approach (Allen et al., 2001; Taylor et al., 2002a). For example, Allen et al. (2001) found that health education provided in workplaces together with worker participation, group sessions and one-to-one outreach activities significantly increased the uptake of screening.

Counselling, with exploration of possible barriers to screening and reasons for self-exclusion, to persuade the woman to have screening was analysed in three studies. In the USA, Rimer et al. (1999) found that complementing the physician’s reminder and mailed tailored educational print communication with tailored telephone counselling increased participation. Attendance was increased in a general practice setting in the United Kingdom if patients had access to a health promotion nurse and had their risk factors assessed and followed up by both the GP and the nurse (Robson et al., 1989). In contrast, Ward et al. (1991) found no significant difference in Australia between a minimal intervention by which GPs advised eligible women to be screened and a maximal intervention during which they also attempted to provide counselling. However, since the average time spent on counselling was only 91 seconds (range 6 seconds to 3 minutes and 44 seconds), it is impossible to determine whether counselling was really ineffective or if insufficient time was allocated to perform it effectively.

**Strategies targeting health-care providers**

Despite the recognized influential role of doctors in promoting screening,
many women still report that their doctor failed to recommend screening. Seven of the studies included in Table 57 evaluated the effect of several types of physician reminder (including a flag reminder affixed to the woman’s medical record) versus a control group with no intervention (Binstock et al., 1997; Burack et al., 1998; McDowell et al., 1989; Ornstein et al., 1991; Pritchard et al., 1995; Pierce et al., 1989; Somkin et al., 1997). Only two of these studies found a significant increase in screening uptake compared with no intervention (Binstock et al., 1997; Pierce et al., 1989). However, no differences were found between the physician reminders and other types of intervention measured in these two studies (i.e., telephone or mailed invitations sent to women).

In the United Kingdom, target payments for GPs have been linked to the level of coverage achieved, with the payment for coverage of 80% or over being almost four times that for 50% to 79%. Introduction of such payments led to a dramatic improvement in coverage, from less than 40% to over 80% (Patnick, 2000; NHS 2003a, b).

Community strategies

Mass media campaigns

The impact of mass media campaigns in increasing attendance in cervical cancer screening was summarized in a review of different strategies by Black et al. (2002). Studies included in this review showed that mass media campaigns combined with other strategies were effective at increasing either screening rates or early cancer detection. Of the four studies reviewed that used mass media campaigns alone (Suarez et al., 1993a, b; Mitchell et al., 1997; Suarez et al., 1997), only one was effective, in a specific subpopulation targeted with language-specific materiel (Mitchell et al., 1997). Shelley et al., (1991) reported an increased attendance in New South Wales, Australia, after a mass media campaign including television and radio commercials, advertisements in two women’s magazines and posters and pamphlets distributed to GPs.

Involving family and community members

For many women, particularly those of ethnic or minority groups in developed countries and women in developing countries, their decision about cervical cancer screening will be greatly influenced by the husband or other key family and community members (Lazcano-Ponce et al., 2002). Involving family and community members has been proposed as an important strategy to increase attendance in screening programmes. However, there is limited evidence on the effectiveness of this approach. In a randomized controlled trial carried out in rural India to evaluate the effectiveness of VIA (Sankaranarayanan et al., 2003b), the main components of the project included health education activities, personal invitations, mobile clinics and involving key members and leaders of the community. Attendance reached 63.4%, which represents a reasonable level considering that no women had ever been tested previously in the region.

Strategies to improve follow-up after an abnormal test result

Obtaining good levels of attendance for screening is a necessary but insufficient condition for effectiveness of a cervical cancer screening programme. Screened women with abnormal tests must also receive follow-up and appropriate treatment. Rates of incomplete follow-up vary enormously across settings and populations; between 7 and 49% of women with abnormal test results fail to receive adequate follow-up (Yabroff et al., 2000) and a study in the Amazonian region of Peru found that only 25% of women with abnormal cytology received appropriate follow-up care (Gage et al., 2003). Despite the importance of assuring good levels of compliance with follow-up, most efforts have focused on increasing attendance in screening programmes. For example, in a Cochrane Review of strategies to increase attendance at cervical cancer screening (Forbes et al., 2004), only three of the selected 35 studies were about compliance with follow-up.

A notification letter including some educational material was evaluated in two studies (Paskett et al., 1990; Marcus et al., 1992). Adding educational materials to the notification letter did not have a significant impact on the uptake of follow-up visits in any of these studies except in the group that also received an educational slide-tape programme (Marcus et al., 1992). In this study, providing transportation incentives increased the odds of follow-up compared with women receiving usual care, but the effect was lower than that obtained with the letter plus slide-tape programme. A combination of computerized tracking of follow-up, transportation and financial incentives yielded only a limited increase in the intervention group in relation to the control group (Kaplan et al., 2000). An invitation to consult a nurse who presented educational information about abnormal tests did not result in a significant difference between the groups (Peters et al., 1999).

Although educational interventions have been shown to improve women’s knowledge about the meaning of an abnormal test result, whether this improved knowledge correlates with lower anxiety or improved adherence for follow-up is unclear (Zeisler et al., 1997; Fylan, 1998).

Studies using alternative screening approaches

Alternative methods are being evaluated to provide simple and low-cost screening in developing countries where organizing a cytology-based programme is not feasible (Sankaranarayanan et al., 2001). Combining testing with an immediate offer of treat-
ment for screen-positive women has been shown to be a feasible option in low-resource settings (Gaffikin et al., 2003). One advantage of the "see-and-treat" or "screen-and-treat" approaches is that they reduce the need for follow-up visits and thus decrease the probability of loss to follow-up. Several recent studies have evaluated compliance with treatment using this approach. For example, in a VIA-based demonstration project in rural Thailand, nearly 93% of screen-positive women received cryotherapy in a "screen-and-treat" scheme (Gaffikin et al., 2003). In India, compliance with treatment for high-grade lesions was 74%; in this study, cryotherapy was offered on a "see-and-treat" basis and LEEP was provided through referral to a hospital (Sankaranarayanan et al., 2003b).
Chapter 4
Efficacy of screening

Methodology and analytical issues in assessment of efficacy

The core concept of screening is that detection of early disease offers the opportunity to change its prognosis. Earlier diagnosis may improve prospects for survival because early intervention permits treatment at a more tractable stage (Morrison, 1992). However, as experience with screening has accumulated and understanding of cancer biology has evolved, it has become apparent that there is much heterogeneity among cancers at particular sites, and that this heterogeneity can influence the impact of screening. Models of screening should take account of this heterogeneity.

General definitions
A simplified model of screening is presented in Figure 47. Several definitions are needed to understand this model. First, the model assumes that there is a period in which there is no detectable disease, but early malignant changes may have taken place and a clone of cells is dividing and de-differentiating until it attains a size that can be detected by screening. The point at which a lesion can be found by screening is the beginning of the sojourn time (Zelen & Feinleib, 1969) or ‘detectable preclinical phase’ (DPCP) (Cole & Morrison, 1980). For cervical cancer screening, lesions during the DPCP are mainly preinvasive, but also include some early invasive and microinvasive lesions. ‘Lead time’ refers to the period between the moment a lesion is found by screening and the time of diagnosis of the invasive cancer that would have developed (Morrison, 1992). Sojourn time is a combined function of the lesions and of the screening test. Lead time will in addition usually be affected by the frequency of screening, depending on the distribution of the sojourn time. Both sojourn time and lead time will vary widely in a population. Neither is directly observable for an individual, unless a screening test is repeated at frequent intervals, the results of a positive screening test are ignored and the woman is observed until she becomes symptomatic. Such a situation is clearly not tenable. However, in a population that has undergone screening, the distribution of lead time and sojourn time can be estimated (Walter & Day, 1983).

Sensitivity of the screening test for early detection of invasive lesions
In addition to sojourn time and lead time, two parameters traditionally of importance in screening are sensitivity and specificity. For a condition which either exists or does not, such as Tay–Sachs disease, these two parameters are defined in terms of a $2 \times 2$ table:

<table>
<thead>
<tr>
<th>Result of screening test</th>
<th>‘True’ disease state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Positive</td>
<td>$a$</td>
</tr>
<tr>
<td>Negative</td>
<td>$c$</td>
</tr>
</tbody>
</table>

Sensitivity $= a/(a + c)$, specificity $= d/(b + d)$

![Figure 47](image-url) Scheme of progression of a chronic disease, with the intervention of an early-detection screening test
The situation is more complex for screening for cervical cancer, because it is a progressive condition. At the time at which screening is performed, there is no ‘gold standard’ diagnostic test for the disease: the condition being screened for is a future invasive disease. The ‘true’ disease state being sought at the time of screening is a lesion that will progress into an invasive cancer. This state can be determined for an individual only by following her forward in time. Since, however, a positive result at screening should lead to an intervention to prevent the development of a clinical cancer, much of the information required for direct estimation of sensitivity and specificity will be missing, so that there is no direct measure of the quantity \(a\) in the table. If one follows forward in time a group of individuals who showed no lesion on the screening test, some will develop invasive disease. There is a tradition to estimate the unknown quantity \(c\) by follow-up after screening: the women who develop invasive disease after a negative screening result thus constitute the cell entry \(c\) in the table above.

The length of time after screening that is used to define this group of ‘screen-negative’ and ‘disease-positive’ individuals is variable, and is a somewhat arbitrary interval. The interval between screening rounds is a natural choice if an organized rescreening programme is being evaluated.

In actual screening programmes, a variety of screening intervals are used and a relatively long interval may be used to define sensitivity. This has the advantage that it is less subject to statistical variation due to small numbers and less dependent on the exact date of diagnosis, although more affected by bias due to new cancers. Clearly, the longer the interval used to define sensitivity, the lower will be the resulting estimate (as follows from the discussions below and Figure 48).

To estimate sensitivity, one must then identify the individuals, or indirectly estimate their number, who constitute the cell entry \(a\). The ‘true’ disease state is agreed to be invasive cancer appearing after a positive screening episode. Thus one needs to estimate the number of lesions, detected at screening and treated, which in the absence of screening would have progressed to an invasive cancer. This group forms the screen-positive, disease-positive group. The quantity \(a + c\) is the number of cancers that would have presented as frankly invasive cases in the screened group if no screening had taken place. If one has a directly comparable unscreened population, as in a randomized trial, the quantity \(a + c\) is observable. In the absence of a comparison group strictly defined by randomization, other approaches are needed, but for any general population sample, estimates based on age-adjusted cancer incidence data from a comparable population or a time when screening was not practised should provide a good approximation, if used judiciously. The quantity \(a\) is then obtained by subtraction, and the episode sensitivity estimate is given as before (Day, 1985):

\[
a/(a + c).
\]

This approach to the estimation of sensitivity, called the ‘Incidence method’, can be expressed graphically as in Figure 48 and can be used to estimate sensitivity by means of the proportionate incidence of interval cancers (see below).

Since screening for cervical cancer is aimed at detection of preclinical lesions, most of which are preinvasive, the estimate of \(a/(a + c)\) (for invasive disease) is to be interpreted as the proportion of screen-detected preinvasive lesions that would have progressed to invasion among all lesions in the DPCP that would have progressed to invasion. This is equal to one minus the proportion of observed invasive cancer post-screening to the expected invasive cancer rate in the absence of screening.

An additional issue is the categorization of microinvasive (stage IA1) disease. This is almost always screen-detected and can be effectively treated with minimal morbidity. While this category is usually included among the invasive cancers, for the evaluation of screening, it should be considered as a success and is best included with the carcinoma in situ or CIN 3 cases as a less than fully invasive lesion.

It should be realized that diagnostic issues related to disease verification as discussed in Chapter 2 are fundamental to this discussion. The frequency of disease that is identified in a screening...
programme is also dependent on the diagnostic methods that are used to augment colposcopy and biopsy (e.g., LEEP investigation) and on how aggressively one looks for disease.

**Positive predictive value and specificity**

A similar approach can be taken to the definition of specificity and positive predictive value, as, if one has estimates of the values of \(a\) and \(c\), and \(a + b\) and \(c + d\) are known from the results of screening, then clearly one has estimates of \(b\) and \(d\). For positive predictive value, for example, it is of special interest to estimate the proportion of lesions detected at screening that would have progressed to clinical cancers (i.e., \(a/(a + b)\)) before the next round in a periodic screening programme. For specificity, or rather its complement, one might be interested in the proportion of individuals who had a positive screening result among those who would not have developed a clinical cancer in the interval between screening tests (i.e., \(b/(b + d)\)). The positive tests that were confirmed negative should be included for the test specificity. The test validity indicators correspond to each other like those of episode validity. In particular, it is deficient to report (only) episode specificity and only test sensitivity.

Attention should also be paid to the definition of the screening test. Screening for cervix cancer is essentially a multiple-step process, with the initial screening test leading, if positive, to more detailed investigations, culminating in a biopsy for a definitive diagnosis. The definitions of sensitivity and specificity discussed in this section refer to the complete screening episode, the final assessment of positivity or negativity being based on the results of the screening test and all further assessment. It is a common experience that women with a positive test but classified as negative (i.e., disease-free) on further assessment are at higher risk of subsequent disease than the general population. The implication is that if only the screening test is considered, it will be more sensitive than the overall screening episode, although of course with less specificity. The sensitivity of screening tests could be estimated in analogous fashion to the sensitivity of the complete screening episode, but in practice such estimation is rarely attempted.

**Relative sensitivity of different screening tests**

The preceding paragraph considered sensitivity as the capacity of a screening episode and test to detect future invasive disease. This measure might be termed absolute sensitivity. Often, however, one requires a more rapid and direct method of comparing the sensitivity of two screening tests, probably based on cross-sectional rather than follow-up data. In this situation, it is useful to define the relative sensitivity of the two screening tests, In the context of cervical screening, this surrogate is usually taken as histological diagnosis of the screen-detected lesions.

It should be noted that not all lesions diagnosed histologically as malignant would have progressed to invasive cancers. It is known that many preinvasive lesions, including CIN 3 and carcinoma in situ, will regress. Furthermore, colposcopically directed biopsies are known to miss significant lesions. It is therefore clear that relative sensitivity and absolute sensitivity are measures of different quantities. Relative sensitivity would normally be larger than absolute sensitivity, due to overdiagnosis at histology and the length bias of cytologically detected lesions. Hence, if relative and absolute sensitivities differ, due to inherent bias in the relative sensitivity, the absolute sensitivity will give the more correct estimate.

To determine relative sensitivity, a sample of women would undergo both screening tests (e.g., split-sample studies; see Chapter 2). In some studies, all women then undergo colposcopy as well, and the histological diagnosis is obtained on the entire sample. The results would then be summarized as in the table below.

<table>
<thead>
<tr>
<th>Histological diagnosis</th>
<th>Screening test 1 Positive</th>
<th>Screening test 1 Negative</th>
<th>Screening test 2 Positive</th>
<th>Screening test 2 Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive x1</td>
<td>y1</td>
<td>x1</td>
<td>y2</td>
<td>x2</td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Detection rate of Test 1 is \(w_1/(w_1 + x_1 + y_1 + z_1)\)

Test positivity rate of Test 1 is \((w_1 + y_1)/(w_1 + x_1 + y_1 + z_1)\)
histological diagnosis available on those positive on one of the screening tests to be compared. These studies give an unbiased estimate of the ratio of their respective relative sensitivities, provided all those positive on one of the two screening tests have a histological diagnosis available. However, they give a biased overestimate of the sensitivity relative to histology, as they exclude from the denominator women positive by histology but negative on the screening tests under consideration.

Estimates of specificity can also be obtained from such studies, since there will be women negative on one of the screening tests among the women who are positive on at least one of the tests. This estimate will be severely biased (underestimated), since a major component of a correct specificity estimate will be women negative on both screening tests and by histology.

This bias in sensitivity and specificity is known as the verification bias (Franco, 2000, 2003). Some studies have chosen a sub-sample of women negative on the screening test to undergo a histological diagnosis, in order to attempt to correct for verification bias. This attempt may be more or less successful, depending on the size of the sample, the comparability of the colposcopy and perhaps other factors in reducing verification bias. An example of correction for verification bias is given in Table 58 (taken from Ratnam et al., 2000). The effects are clearly large. [The ratio of the uncorrected sensitivities differs slightly from the ratio of the corrected sensitivities since not all women positive on one or both of the two tests were included in the uncorrected estimates, although included in the corrected estimates.]

One must also be alert to potential bias in the interpretation of the net efficacy of screening when two tests are used in series or in parallel in the same women (Macaskill et al., 2002). A nominal increase in relative sensitivity always occurs by chance whenever an adjunct test (e.g., HPV DNA testing) is used in parallel with a conventional one (e.g., cytology), even if the new test gives totally random results with respect to the disease being evaluated. This increase in sensitivity can be misleading, even if deemed significant by a statistical test. Combined testing prevents a loss in specificity but may offer no real sensitivity gain in certain conditions (Franco & Ferenczy, 1999). An empirically valuable adjunct test, such as the HPV assay, should complement cytological testing so that the net combined sensitivity and specificity will be truly superior to those of cytology alone. In practice, we can compute the "expected null values" for sensitivity and specificity using the following formulae (Franco, 2000):

\[
S_{\text{exp}} = S_{\text{cyt}} + P \left(1 - S_{\text{cyt}}\right)
\]

for the expected null sensitivity, and

\[
W_{\text{exp}} = W_{\text{cyt}} - P \left(W_{\text{cyt}}\right)
\]

for the expected null specificity, where \(S = \text{sensitivity}, W = \text{specificity}, S_{\text{cyt}} \text{ and } W_{\text{cyt}} \text{ represent, respectively, the sensitivity and specificity for the original cytological test, } S_{\text{exp}} \text{ and } W_{\text{exp}} \text{ denote the adjusted (for the addition of the new test) sensitivity and specificity, and } P \text{ is the expected positivity rate for HPV testing or any other test used as an adjunct to cytology in the same population. These expected values should then be used in a comparison with the equivalent indices from the combined testing approach.}

An analogous issue to the latter bias that stems from the application of adjunctive testing may also appear in randomized controlled trials as an asymmetry problem. Figure 49 illustrates this problem using a generic example with two screening tests, one (test A) that serves as the paradigm (e.g., cytology) and the other (test B) that serves as the adjunctive, experimental technology (e.g., HPV DNA testing), whose benefit is to be evaluated. If a trial is designed without long-term follow-up, it will probably misinterpret the difference between arms in detection rates of preinvasive lesions as being indicative of the putative efficacy attributable to the intervention. One needs to exercise caution and avoid inadvertently claiming that the combination of cytological and HPV DNA testing in such trials is superior simply because it detected more prevalent (or short-term incident) high-grade lesions than the cytology-only arm, as a measure of the greater sensitivity of the combined testing approach. The asymmetry bias virtually guarantees that this would happen even if the HPV test performed randomly with respect to

<table>
<thead>
<tr>
<th>Screening</th>
<th>Definition of</th>
<th>Uncorrected</th>
<th>Corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity</td>
<td>Specificity</td>
<td>Sensitivity</td>
</tr>
<tr>
<td>Cytology</td>
<td>LSIL or worse</td>
<td>38.2</td>
<td>80.5</td>
</tr>
<tr>
<td>HPV</td>
<td>Positive</td>
<td>85.3</td>
<td>58.0</td>
</tr>
<tr>
<td>Combination</td>
<td>HPV+ or ≥ LSIL</td>
<td>97.1</td>
<td>51.3</td>
</tr>
</tbody>
</table>

From Ratnam et al. (2000)
lesions and cytology. This is because there will be more women selected for triage, which will increase the probability of detecting incipient lesions that would never be found were it not for the contribution of the adjunctive test, however inadequate it may be (Franco, 2004).

**A more complex view of cancer**

The model shown in Figure 47 describes the operational process of screening, incorporating no information on the biology of the carcinogenic process. Current knowledge of the neoplastic process allows us to distinguish a number of steps, which may begin with mutation at specific genetic loci and other cellular events and continue until cells divide and disseminate throughout the organism. Cancer development is a long process, and not all the steps are necessarily irreversible. In the future, screening modalities may be developed to target these early molecular changes. In that case, more complex models of the screening process will be required.

If all women do not undergo each test, \( w_1 + x_1 \) is no longer equal to \( w_2 + x_2 \) (as in the direct-to-vial studies mentioned in Chapter 2). While cross-sectional analysis mimics the relationship between relative sensitivities, follow-up may allow estimation of the absolute sensitivities, depending on the validity of controls.

**Methods used to evaluate the efficacy of screening**

Reductions in mortality and/or incidence of invasive disease are fundamental measures of the efficacy of screening. A reduction in the incidence of invasive disease as a consequence of the treatment of disease precursors is a predictor of a reduction in mortality from cervical cancer.

Because screening for cervical cancer results in the detection and treatment of precursors, reduction in the incidence of the disease is an appropriate outcome measure. Reduction in mortality was used in some early studies to evaluate cervix cancer screening. It is accepted that case survival is not an appropriate outcome, because of lead time, length bias, selection bias and overdiagnosis bias.

Randomized controlled trials are the most valid method to assess reductions in mortality or incidence consequent on screening (Prorok et al., 1984), but until recently have not been employed in the evaluation of screening for cervical cancer.

**Use of observational studies in evaluation of screening**

Observational studies can be used to evaluate the efficacy of screening, provided that the programme was introduced sufficiently long before the study that an effect can be expected to have occurred. The major bias that potentially affects observational studies in the evaluation of screening is selection bias, as the health-conscious may select themselves for screening and are likely to have different underlying (lower) risks of developing cancer of the cervix from those who refuse screening. Thus any comparison that essentially compares the incidence of cancer of the cervix in those who accept invitations to attend for screening with those who decline such invitations is potentially affected by selection bias.

In the cohort study design, the incidence of cancer of the cervix in an individually identified and followed screened group (the cohort) is compared with the incidence in a control population, often derived from the general population, sometimes using data from before screening started, or from another study population in which screening has not been used. The incidence in the controls should be adjusted for the incidence in the refusers (Cuzick et al., 1997) for a valid estimate of efficacy. An estimate of efficiency is obtained by comparing the incidence in those offered screening (attenders and refusers) with that among controls. A historical comparison could be biased if changes in risk factors in the population are affecting the incidence of the disease. If mortality...
were to be used as the end-point, it must be recognized that those recruited into a screening programme are initially free of the disease of interest, so that it is not appropriate to apply population mortality rates for the disease to the person-years experience of the study cohort. Rather, as is required in estimating the sample size required for a controlled trial of screening, it is necessary first to determine the expected incidence of the cases of interest, then apply to that expectation the expected case-fatality rate from the disease to derive the expectation for the deaths (Moss et al., 1987). In practice, it is difficult in a cohort study of screening to adjust for selection, so the results have to be interpreted with caution.

A case–control study of screening is another approach that can be used to evaluate screening. Case–control studies should be designed to mimic randomized controlled trials as far as possible, especially in terms of the cases used (ideally, in this instance, cases of invasive cancer of the cervix). They depend on comparing the screening histories of the cases with the histories of comparable controls drawn from the population from which the cases arose. Individuals with early-stage disease if sampled would be eligible as a control, providing the date of diagnosis was not earlier than that of the case, as diagnosis of disease truncates the screening history. However, a bias would arise if advanced disease were compared only with early-stage disease, as the latter is likely to be screen-detected, though this is just a function of the screening process, not its efficacy (Weiss, 1983). Cases have to reflect the end-points used to evaluate screening, i.e., those that would be expected to be reduced by screening.

Selection bias may be difficult to adjust for in the analysis, though this should be attempted if relevant data on risk factors for the disease (confounders) are available. Such bias may not be a problem, however, if it can be demonstrated that the incidence of cancer in those who declined the invitation to the screening programme is similar to that expected in an unscreened population. This is only seldom true, however (see Chapter 3).

Even if data are available on risk factors for disease, control for them may not result in avoiding the effect of selection bias. Experience in breast cancer screening studies in Sweden and the United Kingdom, where case–control studies were performed within trials, shows that although breast cancer incidence among those who refuse invitations for screening is similar to that of controls, their breast cancer mortality experience is worse than that of controls. This means that the estimate of the effect of screening in such case–control studies will be greater than could be expected in the total population (Miller et al., 1990). This differential could also arise because the case–control analysis directly measures the effect in those screened.

There are other problems with case–control studies of screening, notably two relating to the exposure measure (Weiss, 1994, 1998). One is the issue of excluding tests that are done because the disease is present. These are so-called diagnostic tests, often performed in women with symptoms, or suspected to be at high risk. The second is that of bias due to counting only negative tests as exposure to screening. This will bias the result by omitting positive tests, although these are the only tests that can possibly influence risk of disease.

### Assessing the efficacy of a new screening test

When a new screening test becomes available as a supplement or even a replacement for a screening test that is known to be effective in reducing incidence of invasive cancer, a comparative assessment can be made in terms of absolute sensitivity. Since absolute sensitivity, as defined earlier, quantifies the reduction in the incidence of invasive disease following screening, a comparison between two tests of their absolute sensitivity should be predictive of the comparative efficacy of the two tests. Often, however, it will initially be easier to compare relative sensitivity, and specificity, in cross-sectional studies using histological diagnosis of preinvasive lesions as a surrogate. These results need to be interpreted with caution, since a proportion of such preinvasive lesions will have little potential to progress to invasion. However, if the new test identifies both the lesions detected by the existing test (or the great majority of such lesions) together with additional lesions, the problem of differential regression of lesions identified by the new test is of less concern. Nevertheless, confirmation, preferably by follow-up studies to establish absolute sensitivity, would normally be required before the new test can be considered of at least equal efficacy to the existing test.

The Working Group concluded that at present no surrogate marker suitable for evaluating the efficacy of a new screening test for cervical cancer has been fully established.

### Cytological screening

#### Randomized trials

Efficacy of cervical cancer screening programmes using cytological testing was never tested in a randomized trial. Evidence has therefore been derived from observational studies (cohort and case–control). Studies relating trends in cervical cancer incidence or mortality to screening have also provided very convincing evidence in support of the effectiveness of these programmes, as described in Chapter 5.
Cohort studies

In the evaluation of cervical cancer screening, the ‘exposure’ has been defined in cohort studies either by invitation to screening or by participation in screening. In efficacy trials (randomized), there is no bias, but in observational studies several biases must be assumed. In the early studies, the rates expected in the absence of screening were usually calculated from the population rates during the latest period before screening was implemented or from concurrent rates in areas without a screening programme. More recently, reference rates were obtained either from those among non-screened or non-participants, or from the average rates in the general population during the screening period. In studies that followed incidence among participants and compared it with that among non-participants, there is potential for selection of a more healthy group to participate compared with non-participants (see Chapter 3), which would introduce a bias towards overestimation of the impact of screening. In studies using the general population rates as the reference, a large part of the general population had been screened in the programme. Therefore, only the early studies with a non-targeted population (not intended to be screened) as controls will in principle give estimates on efficiency with less bias, i.e., on the effect if screened albeit not in ideal but in routine conditions. This, however, assumes that the risk in the control population was adjusted for the risk among non-responders to screening (Cuzick et al., 1997). Such adjustment was done by the Working Group and is indicated in brackets in the following and called an efficacy estimate. Not all cohort studies provided the necessary information for such a correction, however, and none of the studies with non-responders as controls or incidence in the total population as reference value can yield such an estimate.

Several of the cohort studies reviewed below, which use invasive cervical cancer as the main outcome parameter, were included in a review by Lynge (2000), upon which this chapter is based. Table 59 summarizes the characteristics and main findings of the cohort studies.

British Columbia, Canada

The British Columbia cervical cancer screening project started in 1949. In 1959, about 8% of women above the age of 20 years were screened, and about 44% in 1971. The first cohort study was based on the population incidence data and individual screening records from 1958 to 1966. In 1965, 13 clinical invasive carcinomas were detected among the screened women, with 81.2 cases expected based on the 1955–57 incidence rates (standardized incidence ratio (SIR) = 0.16). Among the unscreened women, the numbers were 67 and 62.1 respectively (SIR = 1.08). The effectiveness of the programme was estimated for the total population at 80 observed versus 143.4 expected (SIR = 0.56). [The efficacy estimate of the SIR corrected for selective participation (Cuzick et al., 1997) was 0.17.] Pre-clinical invasive cancers were not included among the observed cases (Fidler et al., 1968). A later cohort study from the same programme reported cancer incidence rates after negative tests (van Oortmassen & Habbema, 1986) (included in the IARC study (IARC, 1986), see below).

Finland

One of the first indications of the magnitude of the effect of screening for cervical cancer in the Nordic countries was provided by a cohort study (Hakama & Räsänen-Virtanen, 1976). In Finland, an organized screening programme started in 1963 and it gradually developed to become nationwide. A cytological test was offered every fifth year to all women aged 30–55 years. Data on 407 000 women screened at least once during 1963–71 and followed up from their first screening until the end of 1972 revealed that the (relative) risks of invasive cervical cancer were 0.2 after a negative test and 1.6 among non-attenders, in terms of the unit risk of reference from the overall Finnish incidence in years preceding the start of the programme. There were 1.4 million woman-years in the follow-up of invasive carcinoma among screened women, and their average follow-up time was thus 3.5 years. The effect of the screening test applied was 80% (100 x (1.0 – 0.2)), without correction for the selective attendance. The attendance rate was 85%, and the effectiveness of the public health service was estimated at 60%, indicating the result of outcome evaluation of the programme among the whole population in that early follow-up phase. [The efficacy estimate for SIR corrected for selective attendance was 0.22].

A further cohort follow-up study in Finland was based on the follow-up of a sample of 45 572 women with a Papanicolaou group I test result in the mass screening programme during 1971–76, when up to the third invitational round was in action (Viikki et al., 1999). The follow-up was performed using the files of the cancer registry up to 1994, and the reference risk was obtained from expected numbers calculated from the general population rates, including women screened in the programme. Overall, 48 invasive cancers were observed (SIR = 0.5; 95% CI 0.4–0.7). In the five-year follow-up since screening, the SIR estimates were, respectively, 0.3 and 6.5 among those with Papanicolaou group I or II–V results but no malignancy confirmed in the screening episode. Follow-up of those with a positive test result at entry was also reported by Viikki et al. (2000). The SIR estimate
<table>
<thead>
<tr>
<th>Location (reference)</th>
<th>Cohort description: Numbers of women, screening period, source of screening data, follow-up period and source of follow-up data, screening recommendation</th>
<th>Cervical cancer end-point</th>
<th>Screening assessment</th>
<th>Observed (expected) rate per 100 000</th>
<th>No. of cases/ deaths observed (expected)</th>
<th>Relative risk</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>British Columbia, Canada (Fidler et al., 1968)</td>
<td>310 000 screened and 233 000 unscreened women, screening in 1958–65, screening laboratory database, incidence follow-up to 1965, laboratory database and pathological files in the province plus a mortality registry, screening recommendation in 20+ years old women once a year</td>
<td>Incidence, clinical squamous carcinoma (invasive) in 1965</td>
<td>Screened previously at least once in the programme</td>
<td>4.2 (26.2)</td>
<td>13 (81.2)</td>
<td>0.16</td>
<td>Pre-clinical 'occult' invasive carcinoma not included</td>
</tr>
<tr>
<td>Finland (Hakama &amp; Räsänen-Virtanen, 1976)</td>
<td>406 358 screened and 35 279 unscreened women, screening 1963–71, screening registry, incidence follow-up in 1963–72, cancer registry, screening 30–55-year women every 5 years</td>
<td>Incidence, invasive carcinoma after the first programme test</td>
<td>Screened at least once in the programme</td>
<td>7.7 [38.5]</td>
<td>109 [545]</td>
<td>0.2</td>
<td>Average follow-up time 3.5 years among screened women (1.4 million person-years). Expectation was drawn from time before screening.</td>
</tr>
<tr>
<td>Iceland (Johannesson et al., 1978)</td>
<td>Not reported [percentage of women in 1974 who had ever been screened was approx. 89% in women aged 30–54], screening 1964–74, screening registry, mortality follow-up in 1965–74, cancer and mortality registry, screening women of 25–59 (25–70 from 1969) every 2–3 years</td>
<td>Mortality from cervical cancer</td>
<td>Women with an initial negative screening</td>
<td>2.6 (NR)</td>
<td>29.5 (NR)</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Maribo, Denmark (Berget, 1979; Mellemgaard et al., 1990; Lyng, 2000)</td>
<td>16 187 women invited to the first screening round, screening 1967–70, not described, incidence and mortality follow-up to the end of 1984, not described, not described</td>
<td>Incidence of cervical cancer</td>
<td>Participated in screening (after invitation)</td>
<td>115 (217)</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Østfjold, Norway (Magnus et al., 1987)</td>
<td>45 960 women without previous invasive or preinvasive lesions of the cervix and invited to screening, screening 1959–77, not described, incidence and mortality follow-up in 1959–82, cancer registry, screening women of 25–59 every 2–4 years</td>
<td>Incidence and mortality, invasive carcinoma [after invitation]</td>
<td>Whole study population</td>
<td>267 (341.5)</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| | | | Screened | 103 (123.9) | 0.83 |
| | | | 178 (286.1) | 0.62 |
| | | | Incidence | 55 (102.6) | 0.54 |
| | | | Mortality | 89 (55.4) | 1.61 |
| | | | Incidence | 48 (21.3) | 2.25 |

| | | | Unscreened | |

| | | | NR | |

| | | | NR | |

Comments: Numbers of women, screening period, source of screening data, follow-up period and source of follow-up data, screening recommendation
Efficacy of screening

<table>
<thead>
<tr>
<th>Location (reference)</th>
<th>Cohort description: Numbers of women, screening period, source of screening data, follow-up period and source of follow-up data, screening recommendation</th>
<th>Cervical cancer end-point</th>
<th>Screening assessment</th>
<th>Observed (expected)</th>
<th>No of cases/deaths observed (expected)</th>
<th>Relative risk</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden (Sparén, 1996)</td>
<td>386 990 women, screening and population registries, incidence follow-up in 1968–92, cancer registry, any screenings</td>
<td>Screened ever vs never</td>
<td>Women with any negative tests in the programme</td>
<td>125 (259.4) incidence</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Women with 5 negative tests in the programme</td>
<td>41 (92.6) mortality 6 (34.2) incidence</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Women referred to gynaecologist with other diagnosis than severe dysplasia, ca. in situ or invasive carcinoma</td>
<td>11 (26.6) incidence</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Women referred to gynaecologist with other diagnosis than severe dysplasia, ca. in situ or invasive carcinoma</td>
<td>5 (10.0) mortality</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland (Viikki et al., 1999)</td>
<td>A sample of 45 572 women screened negative, screening in 1971–76, screening registry, incidence follow-up in 1971–94, cancer registry, screening 30–55 y old women every 5 y</td>
<td>Incidence, invasive carcinoma after 1st, 2nd or 3rd programme smear</td>
<td>Screened negative</td>
<td>48 (94)</td>
<td>0.5</td>
<td></td>
<td>SIR estimate from population rates during the screening period (including the screened population). Follow-up of women with positive results reported also in Viikki et al. (2000)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Screened negative, follow-up 5 years</td>
<td></td>
<td></td>
<td>0.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Screened positive, follow-up 5 years</td>
<td></td>
<td></td>
<td>6.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>All attenders, follow-up 5 years</td>
<td></td>
<td></td>
<td>0.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

* Vital status or losses from follow-up were not reported.
* NR, not reported

If not available in the original publication, confidence intervals were estimated based on an assumption that the observed number of cases followed a Poisson distribution (indicated in square brackets).
for all attenders combined was 0.7. Among non-attenders, the estimated SIR was 1.6. The approximate relative risk (RSIR) between attenders and non-attenders was 0.7/1.6 = 0.44; and between test negatives and non-attenders 0.3/1.6 = 0.19 (confidence intervals not available). The long-term protection provided by screening scheduled every five years was evaluated. The SIR for invasive cancer remained less than unity (with a 3% annual increase) during all the 23-year follow-up. The SIR of preinvasive lesions exceeded unity at follow-up year 10, i.e., at the second screening round. The results confirmed the appropriateness of the five-year screening interval used in Finland.

**Iceland**

Cervical cancer screening started in Iceland in 1964 and became nationwide in 1969. Women aged 25–59 were invited every 2–3 years (later extended to women aged 25–70). In 1974, approximately 89% of women aged 30–54 had had at least one test; the age-group-specific coverage proportions varied from 81% to 95%. Among women aged 25–29 and 55–59 years, the proportions were lower (47% and 77%). In all Icelandic women aged 25–59, the mortality from cervical cancer changed from 20 per 100 000 aged 25–59, to 21 in 1960–64, 32 in 1965–69 and 15 in 1970–74 (Johannesson et al., 1978). The rates in never-screened women 25–59 years old were 30 in 1965–69 and 23 in 1970–74. The average mortality rate among women with an initial negative screening result was 2.6 per 100 000 in ten years of follow-up. The population mortality rate had decreased in a later study by 60% between 1959–70 and 1975–78, and the mortality rates among the unscreened population were more than ten-fold greater than among the screened (Johannesson et al., 1982).

**Maribo, Denmark**

An organized screening programme was started in 1967 in Maribo County, Denmark, among women aged 30–49 years. The 16 187 women who were invited to the first round in 1967–70 were followed up for incidence of cervical cancer to the end of 1984 and the observed numbers were compared with the expected number based on rates for all Danish women. In the 87% of the invited women who participated, 115 cervical cancer cases were observed compared with 217 expected (SIR 0.53), whereas the numbers were 63 and 35.96 (SIR 1.75) among the 13% of invited women who did not participate (Berget, 1979; Møllergaard et al., 1990; Lynge, 2000). The effectiveness of the programme estimated by the SIR for the total population was 0.70 (178/253). The comparison group was the total population of Denmark, where there was extensive spontaneous screening, and some other organized programmes were operating. The Maribo cohort was later extended to include all women screened in the area in 1967–82; the results on the follow-up for cervical cancer incidence after negative tests were included in the IARC study (see below; IARC, 1986). [The efficacy estimate of SIR corrected for selective attendance was 0.60, but, as mentioned above, it was affected by contamination in the control population].

**Manitoba, Canada**

A province-wide cervical cytology screening programme was initiated in Manitoba in 1963, and included a screening registry. Cases of cervical cancer were recorded at the Manitoba Cancer Registry (Choi & Nelson, 1986). The data on cancer incidence after a negative screening test were included in the IARC study (see below; IARC, 1986).

**Sweden**

From 1964 onwards, several counties in Sweden gradually introduced organized cytological screening programmes for cervical cancer. Women aged 30–49 years were targeted, with a four-year screening interval. The programme covered all of Sweden except the municipality of Gothenburg by 1973. All tests within the organized programme were reported to the National Board of Health and Welfare, where 930 127 women were registered with at least one test during the period 1967–75. This cohort was followed up for incidence of invasive cervical cancer to the end of 1980 (Pettersson et al., 1986). The data for women with a negative result at entry were included in the IARC study (see below; IARC, 1986).

In a later study, a cohort of 386 990 women resident in Uppsala and Gävleborg counties was followed up for invasive squamous-cell cancer of the cervix (Sparén, 1996). The screening histories were derived from computerized registers including any cytological tests performed in the area. Record linkage allowed complete follow-up with regard to cancer incidence, migration and deaths during 1968–92. The relative risk of squamous-cell cervical cancer incidence among ever- versus never-screened women was 0.55 (95% CI 0.51–0.61). The lowest age-specific relative risks among the screened women were in the age group 40–59 years (RRRs from 0.27 to 0.38).

**Østfold, Norway**

A regional cervical cancer screening programme was organized in Østfold County, Norway. The first round took place in 1959–65 and the last (fifth) round in 1974–77. A cohort follow-up study included all 45 960 women invited to the first screening in the age-group 25–59 years and not previously diagnosed with cervical cancer. The cohort was followed up to the end of 1982, and the observed incidence and mortality were compared with those of women in five neighbouring counties.
which did not offer organized screening. During the period 1959–82, 267 new cases of invasive cervical cancer were observed in the Østfold cohort compared with 341.5 expected cases (SIR = 0.78), and 103 deaths from cervical cancer were observed and 124 expected (SMR = 0.83) (Magnus et al., 1987). Women not participating in the screening programme had a 61% higher incidence of cervical cancer than that observed in the reference population and a more than two-fold excess in the mortality rate. Women with any negative test in the programme had an incidence of cervical cancer of 48% and those with five negative tests 18%, compared with the incidence expected from the counties without organized screening. [The efficacy estimate of SIR corrected for selective attendance was 0.77.]

The quantitative results from the cohort studies are illustrated in Figure 50. Four of the cohort studies (in Canada, Denmark, Finland and Norway) allow an estimate of efficacy with presumably only small bias. The results show large variation in effect, with RRs from 0.17 in the British Columbia study to 0.77 in the Østfold county study. Only part of the variation can be accounted for by bias or random variation; most of it is likely to be true. The screening programmes have an effect that varies from close to eradication of invasive disease to a marginal one, which further emphasizes the need for organization and quality assurance, as outlined in other chapters of this volume.

Case–control studies

Case–control studies do not measure the impact of screening in relation to the situation that would be expected in the absence of screening in the population subjected to the screening programme. Instead they compare the risks among the screened to that among non-screened or never-screened groups. As the absolute risks remain unknown, it is not possible to adjust for selective attendance. Despite the inherent biases in their design, resulting in overestimation of efficacy, case–control studies, like cohort studies, have been crucial in the assessment of the efficacy of screening. In 1986, the conclusions of the IARC Working Group on Cervical Cancer Screening were based on a review of studies performed in a number of countries with widely different approaches (Hakama et al., 1986). Among these, the results of five case–control studies were analysed; two (Macgregor et al., 1985; Geirsson et al., 1986) were nested case–control studies within organized programmes (in Aberdeen, Scotland, and in Iceland) and three (Clarke & Anderson, 1979; Berrino et al., 1986; Raymond et al., 1984) in areas where screening was not centrally organized (Toronto, Canada; Milan, Italy; Geneva, Switzerland).

The studies in Iceland and Aberdeen were designed to determine the reduction in risk of invasive cervical cancer among women with a previous negative test, in terms of time elapsed since the smear was taken (Macgregor et al., 1985; Geirsson et al., 1986). A combined analysis of the two studies showed a relative protection (RP) of more than ten-fold (RP = 11.1; 95% CI 2.4–52.2) for women with their last negative test performed 0–11 months before the diagnosis of the case, compared with women who had had their last negative test ten or more years before (Moss, 1986). The other three case–control studies were designed to evaluate the effects of cytological screening for invasive cervical cancer. They differed in the criteria used for case and control selection and in the definition of screening history, but the odds ratios (OR) observed for ever versus never screened were similar, ranging from 0.26 to 0.37. Furthermore, the effects were similar to those observed in the other studies when time elapsed since the last screening and number of previous tests were taken into account.

Since the publication of the IARC monograph (Hakama et al., 1986) on the efficacy of cytological screening, several case–control studies have been carried out to evaluate screening programmes and activities. Fourteen studies published between 1986 and 1998 were reviewed by Zappa and
Ciatto (2000): three of these had been carried out in North America, two in Central America, four in Asia and five in Europe. Since then, another four case–control studies have reported effects of cytological screening by screening status (screened versus non-screened), in Mexico (Jiménez-Pérez & Thomas, 1999), Finland (Nieminen et al., 1999), Sweden (Andersson-Ellström et al., 2000) and South Africa (Hoffman et al., 2003). Table 60 summarizes the main characteristics and findings of these studies.

About half of the studies were carried out within organized programmes or with active invitation of women. Exposure was usually defined by the screening (attendance) status, not by screening invitation. It is also possible that there had been some screening activity among the ‘non-screened’ groups, diluting the exposure contrast and affecting the rates. This might also have been the case in studies using smear archives and registers. In all studies, cervical cancer incidence was used as the outcome. Two studies also considered cervical cancer mortality in separate analyses. Many studies limited cervical cancers to the squamous subtype, whereas two studies made separate analyses for squamous and adenocarcinomas (Herrero et al., 1992; Sato et al., 1997), while in others the histological type of cervical cancer was not specified. In certain studies, attention was paid to the stage of cervical cancer and/or only advanced stages were considered (van der Graaf et al., 1988; Zhang et al., 1989) or separate analyses were carried out for different grades of invasion (Herrero et al., 1992). About half of the studies were population-based and the others were hospital-based. In some studies, controls were selected from subjects with a negative test at the date of diagnosis of matched case (Macgregor et al., 1994); in another (Sobue et al., 1990), the controls for screen-detected cases were selected among subjects with a negative test in the same year as the diagnosis of the respective case. Two of the studies were nested within cohorts of women invited to be screened (Zhang et al., 1989; Sato et al., 1997).

The proportion of controls ‘ever screened’ may give an idea of the coverage of cytological testing in the general population, although in most studies controls were matched to cases for several co-variables, so that the actual coverage cannot be directly estimated. Proportions tested differ substantially from one study to another. The proportion of ever-screened controls ranged from 20% in Osaka and 37% in Bangkok to 88% in Miyagi and 93% in Maryland. Most studies tried to identify and exclude tests performed because of symptoms by excluding those performed within 6 or 12 months before the index date.

In spite of the differences mentioned in relation to eligibility criteria for cases and/or controls, methods of collection of screening history and adjustment for confounding variables, the results from the review by Zappa and Ciatto were quite similar, ORs ranging from 0.27 in the Danish study to 0.43 in the Canadian study. Some results fall outside this range, with lower risks in the studies in Miyagi, Japan and Jingan, China (OR = 0.16) and a higher risk observed in Mexico City (OR = 0.76). In the latter study, the OR fell to 0.38 (95% CI 0.28–0.52) when only tests performed in the absence of gynaecological symptoms were considered.

Two fairly recent studies show smaller impact of screening, with ORs in the range 0.5–0.8 (Nieminen et al., 1999 for opportunistic screening; Andersson-Ellström et al., 2000). In the Finnish study (Nieminen et al., 1999), there was a clear difference between self-reported screening in the organized programme and that in opportunistic screening. The OR of cervical cancer was 0.25 (95% CI 0.13–0.48; all ages included) among those who participated in the organized screening only (but who were not screened in the spontaneous screening modality) in comparison with the non-screened; the corresponding OR was 0.57 (95% CI 0.30–1.06) for those who had at least one test in the spontaneous modality only (and did not participate in the organized programme). Most women had had tests in both screening modalities (organized and spontaneous) versus those with spontaneous screening alone was 0.47 (95% CI 0.29–0.75). This difference in effect was obtained with less resources in the organized screening (see Chapter 3).

In the Swedish study (Andersson-Ellström et al., 2000), the effect of screening was estimated in comparison with women who had not been tested during the last six years before the index date. Some of these might have been tested earlier, diminishing the screening contrast. There was a large proportion of cases, compared with controls, in whom carcinoma in situ or another (milder) lesion had been previously treated (see Table 60). In only 18 cases (16%) had all the previous tests been negative (the corresponding figure for controls was not given). [These findings suggest that inadequacies in the management of treatment may have, at least partly, accounted for the rather modest effect of screening; or that a fraction of the case women had been tested in the course of management follow-up activity.]

Figure 51 presents a summary of results from the case–control studies on incidence included in Table 60. For the study by Hernandez-Avila et al. (1998), the results excluding tests performed on account of symptoms were considered most relevant for the purpose of estimating efficacy in this
diagram. From the study by Nieminen et al. (1999), only the results on organized screening were included. The study by Andersson-Ellstrom et al. (2000) was not included, as symptomatic women may have been included. There was a strong indication of heterogeneity in the results over all studies reported in the table, reflecting the many differences in studies mentioned above, whereas in the results selected for Figure 51, the heterogeneity is very much less \((p = 0.288, \text{calculated after Der Simonian \\& Laird, 1986})\). There was indication of publication bias, however \((p = 0.023, \text{calculated after Begg \\& Mazumdar, 1994})\). Most studies reported large decreases in the cervical cancer risk attributable to screening, although because of the limitations of and selection in the individual studies, one needs to be cautious when interpreting the pooled point estimate \((0.34)\) of the impact from the case–control studies. A crude attempt to adjust for bias can be made by assuming that selection was the most important source of bias and that it was relatively constant over the studies. The relatively high homogeneity over the case–control studies in the estimated ORs is another justification for such an assumption. The ORs in the cohort studies in risk between non-responders and controls (not non-responders) was about 1.5, which indicates that the non-responders may have an inherent risk up to two times that among responders. Such an adjustment would imply a true protective effect in the populations subjected to case–control studies of screening of about 0.7. If the estimates of ORs in the case–control studies were 0.6 or larger, there is a possibility that the programme was practically without effect.

When death from cervical cancer is taken as the end-point, the protective effect of screening tends to be slightly higher than estimated from the incidence studies. In the Scottish study (Macgregor et al., 1994), the OR for mortality was 0.25 (95% CI 0.11–0.48) and for incidence 0.35 (95% CI 0.25–0.50), while in the Osaka study (Sobue et al., 1990) the corresponding figures were 0.22 (95% CI 0.03–1.95) and 0.41 (95% CI 0.13–1.29). Most of the studies did not report impact on mortality.

**Cervical cancer incidence after screening negative**

The IARC joint study on the incidence of invasive cervical cancer by number of previous negative tests (Day, 1986; IARC, 1986) was based on data from ten centres worldwide from which individual screening histories were available and could be linked to cancer registry data. Five were cohort studies, two nested case–control studies and three population-based case–control studies. The cohort studies were those listed above from British Columbia (van Oortmarssen & Habbema, 1986), Manitoba (Choi & Nelson, 1986), Sweden (Pettersson et al., 1986), Norway (Magnus & Langmark, 1986) and Denmark (Lynge & Poll, 1986a, b). The nested case–control studies (Macgregor et al., 1985; Geirsson et al., 1986) were carried out within organized screening programmes in Aberdeen and Iceland. The population-based case–control studies (Clarke \\& Anderson, 1979; Berrino et al., 1986; Raymond et al., 1984) were from areas where screening was not centrally organized (in Toronto, Milan and Geneva, respectively). In the case–control studies and in one of the cohort studies (British Columbia), the reference population was the potentially selected group of unscreened women; in the other cohort studies, the expected incidence in the absence of screening was derived from corresponding population incidence rates in a period before mass screening was started. A negative result was defined in the IARC study as either a Papanicolaou group I result or one or two suspicious (group II) results followed by a group I result. The relative risk of squamous-cell carcinoma of the cervix among women aged 35–64 years, whose second negative test occurred at age 35, by time since the index negative smear is given in Table 61. The risk estimates were 0.07 (95% CI 0.04–0.10) during the first year (12 months), 0.08 (95% CI 0.05–0.13) during the second year, 0.13 (95% CI 0.08–0.19) during the third year and 0.36 (95% CI 0.25–0.53) during the fifth year since screening negative. These risks were used to calculate the cumulative percentage reduction in risk of squamous-cell carcinoma of the cervix assuming different screening intervals (see below).

One recent cohort follow-up study (Van den Akker-van Marle et al., 2003a) and two case–control studies (Miller et al., 2003; Sasieni et al., 2003) have also reported cervical cancer incidence rates after negative screening results. These studies collected data on screening history from archive sources and followed cervical cancer incidence since time from the index negative smear.

Van den Akker-van Marle et al. (2003a) followed invasive cervical cancer incidence among women who tested negative in the Dutch screening programme during 1975–97. Data on screening were derived from a national pathological archive, and information on cervical cancers was obtained from the same source for the period 1994–97. Incidence rates were calculated for women aged 35–64 years with one and with two previous negative tests. A negative screen was defined as an episode consisting of a cytological or histological examination with a negative result, or a cytological examination with a positive result but without histological confirmation of invasive cervical cancer or a precursor.
Table 60. Main characteristics and results of case–control studies on cervical cancer screening published after 1986 (modified and updated from Zappa & Ciatto, 2000)

<table>
<thead>
<tr>
<th>Country (reference)</th>
<th>Outcome, period of observation, number and source of cases and controls</th>
<th>Screening modality</th>
<th>Proportions of cases/controls ever screened (%)</th>
<th>OR ever vs never screened</th>
<th>95% CI</th>
<th>Data source for screening information. Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangkok, Thailand (Wangsuphachart et al., 1987)</td>
<td>Incidence (all histological types, ages 15–54 y), 1979–83, 189/1023, hospital records</td>
<td>Not invitational</td>
<td>30/37</td>
<td>0.39 (screened every 2–5 y vs never)</td>
<td>0.21–0.74</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>Denmark (Olesen, 1988)</td>
<td>Incidence (all histological types, mean age 52.6 y), 1983, 428/428, cancer registry</td>
<td>Invitational</td>
<td>45/67</td>
<td>0.27</td>
<td>0.18–0.42</td>
<td>Questionnaire to general practitioners</td>
</tr>
<tr>
<td>Nijmegen, Netherlands (van der Graaf et al., 1988)</td>
<td>Incidence (FIGO &gt;1A, age &lt;70 y), 1979–85, 36/120, cancer registry and registrar's office</td>
<td>Invitational</td>
<td>47/68</td>
<td>0.22</td>
<td>0.1–0.81</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>Maryland, USA (Celentano et al., 1988)</td>
<td>Incidence (age 22–84 y), 1982–84, 153/153, hospital admission records</td>
<td>Not invitational</td>
<td>72/93</td>
<td>0.29 (screened within 3 y vs never)</td>
<td>0.15–0.58</td>
<td>Interview</td>
</tr>
<tr>
<td>Washington, USA (Shy et al., 1989)</td>
<td>Incidence (FIGO &gt;1B-occult, ages 31–75 y), 1979–83, 92/178, cancer registry</td>
<td>Not invitational</td>
<td>85/93</td>
<td>0.21</td>
<td>0.09–0.50</td>
<td>Telephone interview. Smears collected in the follow-up of an abnormal test or at the cancer diagnosis were excluded. OR estimate a re-calculation by Zappa &amp; Ciatto (2000).</td>
</tr>
<tr>
<td>Jingan, China (Zhang et al., 1989)</td>
<td>Incidence (FIGO &gt;1A, squamous), 1965–74, 109/545, screening archive</td>
<td>Invitational</td>
<td>Not available</td>
<td>0.16 (smears performed within last 2 y vs smears performed 6 or more y earlier)</td>
<td>0.05–0.58</td>
<td>Archive</td>
</tr>
<tr>
<td>Osaka, Japan (Sobue et al., 1990)</td>
<td>Incidence (ages 30–79 y), 1965–87, 28/272, cancer registry and dwelling history</td>
<td>Invitational</td>
<td>25/39 (within 10 years)</td>
<td>0.41 (screened within 10 y vs not screened within 10 y)</td>
<td>0.13–1.29</td>
<td>Archive. Only negative tests included.</td>
</tr>
<tr>
<td>Osaka, Japan (Sobue et al., 1990)</td>
<td>Mortality (age &lt;80 y), 1965–87, 15/150, cancer registry and dwelling history</td>
<td>Invitational</td>
<td>7/20 (within 10 years)</td>
<td>0.22 (screened within 10 y vs not screened within 10 y)</td>
<td>0.03–1.95</td>
<td>Archive. Includes diagnostic smears</td>
</tr>
<tr>
<td>Florence, Italy (Palli et al., 1990)</td>
<td>Incidence (age &lt;75 y), 1982–85, 191/540, cancer registry and residents list</td>
<td>Invitational</td>
<td>19/48</td>
<td>0.29</td>
<td>0.15–0.55</td>
<td>Archive</td>
</tr>
<tr>
<td>Bogota, Mexico City, Panama, Costa Rica (Herrero et al., 1992)</td>
<td>Incidence (age &lt;70 y), 1986–87, 759/1433, cancer treatment centres, hospital admission list and partly from census list.</td>
<td>Not invitational</td>
<td>50/72</td>
<td>0.40</td>
<td>0.31–0.48</td>
<td>Interview</td>
</tr>
<tr>
<td>Country (reference)</td>
<td>Outcome, period of observation, number and source of cases and controls</td>
<td>Screening modality</td>
<td>Proportions of cases/controls ever screened (%)</td>
<td>OR ever vs never screened</td>
<td>95% CI</td>
<td>Data source for screening information. Notes</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------------------------------------</td>
<td>-------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------</td>
<td>-------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Manitoba, Canada (Cohen, 1993)</td>
<td>Incidence (ages 25–64 y), 1981–84, 415/29269, cancer registry and residents list</td>
<td>Not invitational</td>
<td>76/87 (within 10 y)</td>
<td>0.43</td>
<td>0.32–0.57</td>
<td>Health care files</td>
</tr>
<tr>
<td>South-east Scotland, (Macgregor et al., 1994)</td>
<td>Incidence (squamous CC), 1982–91, 282/564, screening records</td>
<td>Invitational</td>
<td>45/73</td>
<td>0.35</td>
<td>0.25–0.50</td>
<td>Cytopathology database</td>
</tr>
<tr>
<td>South-east Scotland, (Macgregor et al., 1994)</td>
<td>Mortality (squamous CC), 1982–91, 108/216, screening records</td>
<td>Invitational</td>
<td>35/73</td>
<td>0.25</td>
<td>0.11–0.48</td>
<td>Cytopathology database</td>
</tr>
<tr>
<td>UK (Sasieni et al., 1996)</td>
<td>Incidence (age &gt;20 y), 1992, 348/677, pathology laboratories and registry of local health authority</td>
<td>Invitational</td>
<td>73/85</td>
<td>0.26 (tests performed 24–35 months before, vs not screened or screened &gt;66 months before)</td>
<td>0.14–0.47</td>
<td>Archive. These data were included in the later study (Sasieni et al., 2003)</td>
</tr>
<tr>
<td>Miyagi, Japan (Sato et al., 1997)</td>
<td>Incidence (ages 35–79 y), 1984–89, 119/218, screening archive</td>
<td>Invitational</td>
<td>55/88</td>
<td>0.16 (screened within 5 y vs not screened within 5 y)</td>
<td>0.09–0.28</td>
<td>Interview and archive</td>
</tr>
<tr>
<td>Mexico City, Mexico (Hernandez-Avila et al., 1998)</td>
<td>Incidence, 1990–92, 397/1005, hospital admissions records and sample of residents</td>
<td>Not invitational</td>
<td>42/51</td>
<td>0.76</td>
<td>0.59–0.98</td>
<td>Interview. OR = 0.38 (95% CI 0.28–0.52) when tests due to gynaecological symptoms were excluded.</td>
</tr>
<tr>
<td>Guadalajara, Mexico (Jiménez-Pérez &amp; Thomas, 1999)</td>
<td>Incidence (age &lt;70 y), 1991–94, 143/311, hospital records</td>
<td>Not invitational</td>
<td>54/82</td>
<td>0.3</td>
<td>0.2–0.4</td>
<td>Interview</td>
</tr>
<tr>
<td>Finland (Nieminen et al., 1999)</td>
<td>Incidence, 1987–94, 147/1098, hospital records and population files</td>
<td>Invitational</td>
<td>56/72 all ages, 68/88 ages 30–59, 64/66 all ages, 80/80 ages 30–59</td>
<td>0.36</td>
<td>0.25–0.53</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>Värmland, Sweden (Andersson-Eilstrom et al., 2000)</td>
<td>Incidence (ages 20+ y), 1990–97, 112/112, pathology and population files</td>
<td>Any smears (about 50% of the tests were after invitation)</td>
<td>83/88 in ages 20–59 (within 6 y)</td>
<td>0.83</td>
<td>0.19–0.57</td>
<td>Pathology database. 16 cases (14%) and 4 (4%) controls had been previously treated for carcinoma in situ of the cervix (p &lt; 0.01); 32 cases (29%) and 6 controls (5%) had former atypia (p &lt; 0.001)</td>
</tr>
<tr>
<td>Western Cape, South Africa (Hoffman et al., 2003)</td>
<td>Incidence (stage &gt;IA), 524/1540, hospital records</td>
<td>Not invitational</td>
<td>50/73</td>
<td>0.3</td>
<td>0.3–0.4</td>
<td>Interview. OR 0.2 among those with at least 3 tests; and 0.3 among those with &lt;10 y since the last screen</td>
</tr>
</tbody>
</table>
The incidence expected in the absence of screening was estimated using incidence data from three regions during 1965–69 (the latest period before the screening programme was started) covering 8% of women in the whole country. In addition, age-period-cohort (APC) modelling was used to refine the expected incidence without screening, using the same pre-screening period incidence data as the input. The relative risk of invasive cervical cancer increased from 0.13 in the first year after screening to 0.24 after more than six years from screening for women with one previous negative screening (confidence intervals not available). These figures decreased to 0.12 (95% CI 0.08–0.17) and 0.06 (95% CI 0.03–0.10) for 0–6 and 7–12 months since the last negative screening and 0.18 (95% CI 0.11–0.30), respectively, for more than six years among women with two or more previous negative screening results. The identification and linking method used in the pathology register was not perfect (i.e., the identification code consisted of the same characters for two or more women); this was considered to have produced an upward bias in the incidence rate after a negative test. As a consequence, the true reduction in relative risk might have been somewhat larger than reported. On the other hand, the analysis using APC modelling suggested overestimation of the background risk.

Miller et al. (2003) analysed negative cytological histories within a selected group of women having continuous participation in the Kaiser Permanente medical care programme in northern California, USA, for at least 30 months before the diagnosis date of the cervical cancer cases. The cases (N = 482), diagnosed between 1983 and 1995, were drawn from the files of the medical care programme, SEER and the California cancer registry; controls (N = 934) were matched for age, length of membership and race. About 92% of women aged 20 years or more

\[ \text{Odds ratio} \]

\[ (95\%) \]

Figure 51 Forest plot of results from case–control studies with invitational and non-invitational screening, including a pooled odds ratio estimate (incidence) using a random effects model (Der Simonian & Laird, 1986)
had been screened at least once and 89% within the last three years. A test was defined as negative if the cytological result did not require a change in the follow-up interval (i.e., no referral or control test was required). In addition, there was a group of 'other smears' including those for which the result was missing or unrelated to invasive cancer (e.g., atypical or dysplastic endometrial cells, atrophic changes, Trichomonas infection); and a group of 'abnormal' results. For 32% of the cases and 10% of the controls, no negative results were available (indicating either that the women had been screened elsewhere or had not been screened at all, or that they had had positive or other smears). In the follow-up of the last negative test (irrespective of the other two groups in earlier screenings), the OR for a two-year (19–30 months) versus one-year (0–18 months) follow-up interval was 1.72 (95% CI 1.12–2.64) and for a three-year (31–42 months) versus one-year follow-up interval 2.06 (95% CI 1.21–3.50). Adjustment for ever having had an abnormal result before the index test and for having at least one previous consecutive negative result within 36 months before the index test did not essentially change the results. For the sub-sample of women with at least two consecutive negative results, the OR was 2.15 (95% CI 1.12–4.11) for the two-year follow-up and 3.60 (95% CI 1.50–8.68) for the three-year follow-up, as compared with the one-year follow-up. [The study did not quantify the overall reductions in cervical cancer attributable to screening. The Working Group noted that the baseline risk with one-year follow-up was difficult to estimate and could be subject to bias; therefore the results of this study were not included in the table.]

A study in the United Kingdom (Sasieni et al., 2003) used screening data on women registered within a group practice drawn from a computerized database of the screening programme; information on cervical cancer cases was obtained from pathology laboratories. There were 1305 women aged 20–69 years, diagnosed between 1990 and 2001 with frankly invasive cervical cancer, and 2532 age-matched controls. It was not possible to identify which cancers were screen-detected, because some 50% of the women screened in England in the mid-1990s did not attend in response to an invitation to the group practice. In all analyses, the date of

<table>
<thead>
<tr>
<th>Time interval since screening years (months)</th>
<th>Relative risk (95% confidence interval)</th>
</tr>
</thead>
</table>
| IARC (1986)  
 Ages 35–64                        |                                        |
| 1 (0–11)                                    | 0.07 (0.04–0.10)                        |
| 2 (12–23)                                   | 0.08 (0.05–0.13)                        |
| 3 (24–35)                                   | 0.13 (0.08–0.19)                        |
| 4 (36–47)                                   | 0.19 (0.13–0.28)                        |
| 5 (48–59)                                   | 0.36 (0.25–0.53)                        |
| 6 (60–71)                                   | 0.28 (0.17–0.48)                        |
| 7–10 (72–119)                               | 0.63 (0.30–1.67)                        |
| Van den Akker-van Marle et al. (2003a)  
 Ages 35–64 |                                        |
| 1 (0–6)                                     | 0.12 (0.08–0.17)                        |
| 1 (7–12)                                    | 0.06 (0.03–0.10)                        |
| 1–2                                        | 0.08 (0.06–0.12)                        |
| 2–4                                        | 0.15 (0.11–0.19)                        |
| 4–6                                        | 0.20 (0.14–0.29)                        |
| 6–10                                       | 0.18 (0.11–0.30)                        |
| Sasieni et al. (2003)  
 Ages 20–39  
 Ages 40–59  
 Ages 55–69 |                                        |
| 1 (0–18)                                    | 0.24 (0.16–0.37) 0.12 (0.08–0.18) 0.13 (0.08–0.22) |
| 2 (18–30)                                   | 0.33 (0.21–0.51) 0.14 (0.08–0.22) 0.13 (0.07–0.23) |
| 3 (30–42)                                   | 0.67 (0.43–1.04) 0.25 (0.16–0.40) 0.15 (0.08–0.26) |
| 4 (42–54)                                   | 1.06 (0.65–1.72) 0.30 (0.18–0.50) 0.18 (0.09–0.34) |
| 5 (54–66)                                   | 1.40 (0.75–2.62) 0.61 (0.34–1.09) 0.28 (0.14–0.57) |
| 6 (66–78)                                   | 1.86 (0.88–3.93) 0.72 (0.36–1.43) 0.33 (0.14–0.79) |
| >6 (>78)                                    | 2.37 (1.16–4.85) 0.69 (0.36–1.34) 0.55 (0.27–1.10) |

a Including invasive squamous-cell carcinoma of the cervix uteri, since the last negative test at ages 35–64 years, in comparison with expectation in the absence of screening. Assuming that a woman is screened negative at age 35 and that she had at least one negative screen previously.

b Invasive cervical cancer for 35–64-year old women since two or more previous negative screenings, in comparison with expectation without screening.

c Invasive cervical cancer in various age groups since the last operationally negative smear.
Other countries have taken a different view. They have noted that although young women below the age of 25 or 30 have much higher rates of cervical abnormality than older women, the rise in cervical cancer incidence does not take place until the next decade. While treatments are very successful and have very low rates of complications, the consequences for a young woman can be much greater than for an older woman, for whom preservation of fertility is not an issue. For younger women, the risk of harm may be greater than the risk of benefit.

In Europe, the age to start screening varies widely, with women in Finland and the Netherlands invited to the organized programmes from the age of 30 years, while some countries start screening at much younger ages (Miller, 2002b). Sasieni et al. (2003), on the basis of a study principally designed to determine the frequency of re-screening, found that the effectiveness of cytological screening was relatively low in young women, but rose in older women (lower part of Table 61). This led to the decision in England to move from a recommended age of 20 years for starting screening to the age of 25 years.

For developing and some middle-income countries, in order to maximize use of resources, and given the infrequency of cervical cancer below the age of 35 years, it is generally recommended to start screening at 35 years and only extend screening to younger ages when resources permit (WHO, 1986). It has been pointed out that age is the most important risk factor for cervical cancer and that screening should aim to target high-risk women. A good guide would be to take the age at the beginning of the rise in incidence of cervical cancer and begin screening five years before this age. In most countries, this would be at about 30–35 years of age (Miller et al., 2000).

Frequency of re-screening

Screening programmes seek to maximize the reduction in incidence of and mortality from disease, for a given level of resources. The optimal screening interval is one that provides the most favourable ratio between degree of disease control and cost of screening. The design of a screening programme defines two key parameters for achieving these objectives, the target population and the screening interval. Compliance with these parameters is crucial in maintaining the effectiveness of the programme and in measuring its cost-effectiveness in order that resources can be used to increase population coverage (see Chapter 3). Significant deviation from the recommended screening interval or target population may reduce the programme efficiency either by using excessive resources, as in the case of annual re-screening for cervical cancer, or by allowing the disease to 'escape' the period at which early intervention can lead to treatment and/or cure. Models can facilitate decisions on the optimal periodicity of screening.

Determining the frequency of screening is helped by understanding the natural history of the condition to be screened for, especially the duration of the asymptomatic (latent) phase. A high frequency of screening will result in a low number of cases per screen and thus a low predictive value. The reason for this is that the prevalence of asymptomatic disease will be low in the population if the screening frequency is high. On the other hand, screening too infrequently will leave much of the disease uncontrolled (Cole & Morrison, 1980).

An early evaluation of cervical cancer screening in British Columbia using a Markov–Chain model supported a prolonged natural history of carcinoma in situ (an average sojourn time of at least nine years) and suggested that cytologically negative women should be rescreened every five years (Shun-
The IARC Working Group on Cervical Cancer Screening Programmes (IARC, 1986) established a proper approach to re-screening. This study showed that there was very little evidence to support annual screening and largely provided the basis for international recommendations for three-yearly or even less frequent screening. It underlined the importance of concentrating screening between the ages of 35 and 64 years, with almost as much benefit expected from three-yearly screening as from annual re-screening (Tables 61 and 62). These findings have been reinforced by a study in the Netherlands (Van den Akker-van Marle et al., 2003a; see Table 61).

It is important to note that the greatest percentage reduction in cumulative incidence can be obtained only if a high proportion of the population complies with screening. However, even in the best of circumstances, experience in highly efficient cytology screening programmes of many countries shows that no realistic screening schedule results in the abolition of invasive cervical cancer.

Part of the reason for the imperfect outcome of screening programmes is failure of an essential component of the programme, which can occur at the level of the woman, her physician or the laboratory examining the cytological smears (Miller, 1995). However, another reason is likely to be the variability of the natural history in different women. The models are based on averages of transition probabilities, each with a different distribution or range of time periods during which some lesions progress from one stage to the next, while others regress to normal, and still others remain stable for long periods of time. Some lesions may progress so rapidly that they cannot be found in a curable stage even with annual screening, and it seems unlikely that the majority of such lesions would be detected by more frequent screening. This does not mean there are different types of cancer of the cervix, as suggested many years ago (Ashley, 1966), just that the fast-growing lesions represent one extreme of the distribution of progression (sojourn) times.

Celentano et al. (1989) conducted a case–control study of 153 cases of invasive cancer, 153 case-nominated controls and 392 randomly selected controls. The results were largely congruent for the two sets of controls. The relative protection after a self-reported negative test was significant for 2–3 years from the last negative test (OR = 8.28; 95% CI 3.44–19.9 for case-nominated controls, OR 4.62; 95% CI 2.04–10.5 for randomly selected controls) and some degree of protection was seen for 4–6 years (OR = 4.30; 95% CI 1.46–12.7; 3.63; 95% CI 1.38–9.57, respectively) after adjustment for a variety of confounders.

Herbert et al. (1996) studied the incidence of cervical cancer in a region of the UK after the introduction of the UK computerized call-and-recall system. The incidence of invasive cancer was significantly higher in women who had not been screened in the previous five years than in those who had (RR = 2.6; 95% CI 1.6–4.3); the incidence was higher in those with an interval of 3.5–5.5 years compared to 0.5–3.5 years (RR = 2.2; 95% CI 1.3–3.8). The RRs were higher when screen-detected cancers were excluded. The authors concluded that a five-year interval is too long.

Vikki et al. (1999) studied the risk of cervical cancer after a negative test in the context of the five-yearly organized screening programme in Finland (see above). They found that the SIR was low initially after screening and increased gradually until the time the next test was due. There was an estimated 3% annual increase in risk of invasive disease. The risk did not reach the national average within the more than 20 years of follow-up. They also found that the relative risk of a preinvasive lesion after an initial negative result was decreased up to the second rescreening round at 10 years and concluded that the five-year screening interval applied in Finland was appropriate.

Goldie et al. (2001) modelled the natural history of cervical cancer using published data on transition and regression rates, and data from a study in Cape Town, South Africa. They concluded that in developing countries, if the limited resources are such as to allow three screenings in a lifetime, it may be more cost-effective to give these tests every five years from the age of 35 or 40 years, rather than every 10 years from the age of 35, as had

<table>
<thead>
<tr>
<th>Screening frequency</th>
<th>% reduction in the cumulative rate*</th>
<th>Number of tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year</td>
<td>93.5</td>
<td>30</td>
</tr>
<tr>
<td>2 years</td>
<td>92.5</td>
<td>15</td>
</tr>
<tr>
<td>3 years</td>
<td>90.8</td>
<td>10</td>
</tr>
<tr>
<td>5 years</td>
<td>83.6</td>
<td>6</td>
</tr>
<tr>
<td>10 years</td>
<td>64.1</td>
<td>3</td>
</tr>
</tbody>
</table>

* Assuming a screen occurs at age 35 years, and that a previous negative screen had been performed

From IARC (1986)
been modelled by the IARC (1986) study, and subsequently adopted as a suggested policy for South Africa (Provincial Administration Western Cape: Department of Health, 1995).

Miller et al. (2003) conducted a case–control study of cases of invasive cancer diagnosed between 1983 and 1995 within the Kaiser Permanente medical care programme (see above). The ORs for various intervals between screens, with a one-year interval as the referent, adjusted for ever having had an abnormal cytological finding before the last negative result and for having at least one negative result within 36 months before the last negative one, increased to 2.24 (95% CI 1.28–3.92) at 2 years, 3.37 (1.97–5.76) at 3–5 years and 5.72 (3.48–9.41) at 5–10 years, although there was a low absolute risk of developing invasive cervical cancer within three years of a previous negative result. [The Working Group noted that the cited odds ratios were entirely dependent on the validity of the estimated referent risk level.]

Sawaya et al. (2003) studied the prevalence of biopsy-proven cervical neoplasia among 938 576 women under 65 years of age. The prevalence of all grades of CIN was highest in women aged less than 30 years and much higher in those with no previous negative cytological result than in those with one or more. No invasive cervical cancers were detected in those who had had three or more previous negative tests. Using a Markov model, and various rates of progression and regression from the literature, they estimated that for women aged 30–64 years who had had three or more consecutive negative tests, extending the re-screening interval from one year to every three years would result in an average excess risk of about 3 per 100 000.

Sasieni et al. (2003) conducted a case–control study in the United Kingdom based on the screening histories of 1305 women age 20–69 with stage 1B cancer of the cervix and 2532 age-matched controls from the records of the screening programme (Table 61). The OR for occurrence of cervical cancer increased with time from last negative result; it reached 1.0 (no protection) at three years for women aged 20–39, approached 1.0 at six years for women aged 40–54 and was still ~0.5 at six years for women aged 55–69 years. The authors estimated the proportion of cervical cancer that would be prevented by different schedules of re-screening. These proportions varied by age. For women aged 20–39, 30% would have been prevented by five-yearly, 61% by three-yearly and 76% by annual screening. The corresponding percent-ages for women aged 40–54 were 73%, 84% and 88%, and for women aged 55–69 were 83%, 87% and 87%, respectively. On the basis of these results, the authors recommended three-yearly screening for women aged 25–49, five-yearly screening for women aged 50–64 and, for women aged 65 or over, screening only of those who had not been screened since age 50.

Age to stop screening
Many countries recommend stopping screening or inviting women at around age 60 or 65 years, for a number of reasons. For example, older women have tended to be poor attenders for screening, and good-quality smears are difficult to obtain in women so far past the menopause. In addition, if they have had regular tests with a normal outcome in the past, they are considered to be at low risk of developing cervical cancer. However, in view of the relatively high age-specific incidence rates of invasive cervix cancer in all countries in older women, there is a consensus in developed countries that women over the age of 60 years who have never been screened or have not been screened for many years should be encouraged to have at least two tests, and only if both are negative should they stop screening (e.g., Miller et al., 1991a; Sasieni et al., 2003). However, in developing countries where resources are limited but available for some screening in older women, it has been recommended that women who have never previously been screened and are older than 60 years of age should have one test only (Miller et al., 2000).

Cecchini et al. (1996) reviewed data for women aged 60–70 years from the Florence screening programme and the Tuscany Cancer Registry. Only five of 242 women with invasive cervical cancer had had two or more negative results between 50 and 60 years of age. However, of 11 342 women aged 58–60 who had a negative test between 1980 and 1987 and were followed to December 1990, only one invasive cancer was diagnosed, compared with 13.95 expected from age-specific incidence data from the cancer registry (OR = 0.07; 95% CI 0.002–0.39). The authors recommended reconsideration of continuing screening after 60–64 years of age.

In North America, it is generally recommended that women who have been actively screened and always been negative should cease screening at 69–70 (Miller et al., 1991a; Saslow et al., 2002). In Europe, the guidelines recommend 64 years as the upper age limit for active invitation for screening (Coleman et al., 1993).

There have been suggestions that women who have been active participants in screening but never had a cytological abnormality could stop screening at younger ages (e.g., 55 or even 50 years) (Cruickshank et al., 1997). Flannelly et al. (2004) analysed screening data for women aged 50 years and over who had had a satisfactory result between 1988 and 1996 (N = 36 512) from five regions in
England and Scotland. Women with prior dyskaryosis or borderline nuclear abnormalities had RRs for a positive test after the age of 50 of 4.39 and 3.08, respectively, compared with women whose screening history before the age of 50 was negative. However, 1.8% of women with a negative screen history before the age of 50 had dyskaryosis detected after the age of 50 during a median duration of follow-up of 33.2 months.

Visual inspection

Four screening techniques based on visual inspection have been assessed for early detection of cervical neoplasia, mostly in low-resource settings:

- Unaided visual inspection (also known as downstaging)
- Visual inspection with 3–5% acetic acid (VIA)
- Visual inspection with acetic acid using low-level magnification (VIAM)
- Visual inspection with Lugol’s iodine (VILI)

Unaided visual inspection involves naked-eye visualization of the cervix, without application of acetic acid, to identify abnormal tissue harbouring cervical neoplasia, particularly invasive cancer. Cross-sectional studies in India have shown low sensitivity (30–50%) for unaided visual inspection to detect cervical cancer precursors and it is no longer considered a suitable screening test (Sankaranarayanan et al., 1997; Basu et al., 2002). VIAM involves the use of low-level magnification (2–4 x) in visualizing acetowhite lesions after application of acetic acid. The test characteristics of VIA and VIAM have been evaluated in cross-sectional studies in India and South Africa (Denny et al., 2000a, 2002; Sankaranarayanan et al., 2004e). The results from these studies indicate that magnification did not improve the test performance over and above that of naked-eye visualization. Low-level magnification is no longer widely used for visualization after application of acetic acid.

VIA

VIA involves naked-eye inspection of the cervix one minute after application of 3–5% dilute acetic acid. VIA has been widely investigated for its test characteristics in detecting CIN 2–3 lesions and invasive cancer in several cross-sectional studies, mostly in developing countries (Slawson et al., 1992; Cecchini et al., 1993; Megevand et al., 1996; Londhe et al., 1997; Sankaranarayanan et al., 1998b, 1999; University of Zimbabwe/JHPIEGO Cervical Cancer Project, 1999; Denny et al., 2000a; Cronje et al., 2001; Belinson et al., 2001; Denny et al., 2002; Rodriguez-Reyes et al., 2002; Ngelangel et al., 2003; Tayyeb et al., 2003; Cronje et al., 2003; Sankaranarayanan et al., 2004a). The relative sensitivity of VIA to detect high-grade precancerous lesions and invasive cervical cancer varied from 29% to 95% and the specificity varied from 68% to 98% in cross-sectional studies suffering from verification bias (Slawson et al., 1992; Cecchini et al., 1993; Megevand et al., 1996; Londhe et al., 1997; Sankaranarayanan et al., 1998b, 1999; Cronje et al., 2001; Tayyeb et al., 2003) (see Chapter 2, Table 26). In cross-sectional studies with minimal verification bias, the sensitivity of VIA to detect CIN 2–3 lesions varied from 37% to 92% and the specificity from 49% to 91% (University of Zimbabwe/JHPIEGO Cervical Cancer Project, 1999; Denny et al., 2000a; Belinson et al., 2001; Singh et al., 2001; Denny et al., 2002; Rodriguez-Reyes et al., 2002; Ngelangel et al., 2003; Cronje et al., 2003; Sankaranarayanan et al., 2004a) (see Chapter 2, Table 26). Conventional cytology was concurrently evaluated in most of the above studies and the sensitivity of VIA was found to be similar to or higher than that of cytology as provided in the respective study settings, but the specificity of VIA was consistently lower than that of cytology (Londhe et al., 1997; Cecchini et al., 1993; Sankaranarayanan et al., 1998b, 1999; University of Zimbabwe/JHPIEGO Cervical Cancer Project, 1999; Denny et al., 2000a; Cronje et al., 2001; Denny et al., 2002; Cronje et al., 2003; Sankaranarayanan et al., 2004f) (see Chapter 2, Table 27). HPV testing was concurrently evaluated in cross-sectional studies in India, South Africa and Zimbabwe and was found to have sensitivity similar to that of VIA (Denny et al., 2000a; Womack et al., 2000; Sankaranarayanan et al., 2004b) but similar (Womack et al., 2000) or higher specificity than VIA (Denny et al., 2000a; Sankaranarayanan et al., 2004b).

VIA is being evaluated in three randomized intervention trials in India, to assess the reduction in incidence of and mortality from cervical cancer as compared to a control group with no screening (Sankaranarayanan et al., 2003a, b, 2004c, d). Early results in terms of participation, detection rates of cervical neoplasia and stage distribution of invasive cancers detected have been reported from two of these studies.

The impact of a single round of screening with VIA provided by trained nurses on cervical cancer incidence and mortality as compared to a control group with no screening is being investigated in a cluster-randomized trial in Dindigul district, south India(Sankaranarayanan et al., 2003a, 2004d). Women aged 30–59 years living in 507 villages grouped into 113 clusters were randomized to VIA screening (57 clusters, 48 225 women) by nurses and to a control group (56 clusters, 30 167 women). The early results from the study are given in Table 63. All the screen-positive women were investigated with colposcopy by nurses and most had biopsies taken. The detection
rates of lesions among screened women were 5.8% for CIN 1, 0.7% for CIN 2–3 and 0.2 for invasive cancer. 71% of women with CIN 1 and 80% of those with CIN 2–3 lesions accepted excisional treatment by mid-level cryotherapy provided by nurses and those with CIN 2–3 lesions accepted 71% of women with CIN 1 and 80% of CIN 2–3 and 0.2 for invasive cancer. women were 5.8% for CIN 1, 0.7% for rates of lesions among screened women were 5.8% for CIN 1, 0.7% for CIN 2–3 and 0.2 for invasive cancer. 71% of women with CIN 1 and 80% of those with CIN 2–3 lesions accepted excisional treatment by mid-level cryotherapy provided by nurses and those with CIN 2–3 lesions accepted 71% of women with CIN 1 and 80% of women were 5.8% for CIN 1, 0.7% for CIN 2–3 and 0.2 for invasive cancer. women were 5.8% for CIN 1, 0.7% for CIN 2–3 and 0.2 for invasive cancer. women were 5.8% for CIN 1, 0.7% for

The impact of screening by VIA, cervical cytology or HPV testing (using the Hybrid Capture® 2 (HC 2) probe B assay; see Chapter 2) on cervical cancer incidence and mortality, as compared to a control group, is being investigated in a cluster-randomized controlled trial in Osmanabad district, India (Sankaranarayanan et al., 2003a; 2004c). Women aged 30–59 years living in 52 clusters of 497 villages in rural Osmanabad District, were randomized to a single round of screening by either VIA (13 clusters, 34 149 women) or cytology (13 clusters, 32 136 women) or HPV testing (13 clusters, 34 515 women) or to a control group (13 clusters, 30 378 women). The early results are given in Table 64. Participation of eligible women in screening was 78.4% in the VIA group, 79.5% in the cytology group and 78.7% in the HPV group. The test-positive rates were 14.0% for VIA, 7.0% for cytology and 10.4% for HPV testing. Test-positive women were investigated with colposcopy and biopsy based on colposcopy findings. Biopsies were taken from 9.9% of screened women in the VIA, 3.5% in the cytology and 4.3% in the HPV groups.

Low-grade lesions were detected in 1068 (4.0%) screened women in the VIA group, 304 (1.2%) in the cytology group and 327 (1.2%) in the HPV group. VIA had a significantly higher rate for detection of low-grade lesions than cytology or HPV testing (p < 0.001). The detection rates of CIN 2–3 were 0.7% in the VIA, 1.0% in the cytology and 0.9% in the HPV groups. The detection rates of CIN 2–3 were significantly different between arms (p < 0.001); after adjustment for socio-economic factors affecting detection rates, the detection rate of CIN 2–3 in the VIA arm was significantly lower than in the cytology arm (OR = 0.7, p = 0.005). During 2000–03, 121 women in the VIA group, 131 in the cytology group, 100 in the HPV group and 59 in the control group were diagnosed with invasive cancer. In the intervention groups, 70–74% of the cancers were screen-detected, and 48–60% were diagnosed in stage I as opposed to 24% in the control group. The preliminary findings from this study indicate satisfactory participation rates for screening, diagnosis and treatment. VIA detected significantly fewer CIN 2–3 cases than did cytology. The trial participants are being followed up to document cervical cancer incidence and mortality in the four groups.

The preliminary findings from the above trials indicate that a VIA-based screening programme is feasible, safe and acceptable for a population in rural settings, and that it results in early detection of cervical neoplasia. VIA is associated with high detection of low-grade CIN. The detection rates of CIN 2–3 lesions by VIA were similar in both the trials. While the detection rate of CIN 2–3 lesions for VIA was constant in the Dindigul trial throughout recruitment, it declined from 1.0% at the beginning to 0.5% at the end of recruit-

<table>
<thead>
<tr>
<th>VIA-screened group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of women</td>
<td>48 225</td>
</tr>
<tr>
<td>Received screening</td>
<td>30 577</td>
</tr>
<tr>
<td>Screened positive</td>
<td>2939 (9.6%)a</td>
</tr>
<tr>
<td>Number of screen-positive women who had colposcopy</td>
<td>2939</td>
</tr>
<tr>
<td>Number of women who received biopsy</td>
<td>2777</td>
</tr>
<tr>
<td>CIN 1</td>
<td>1778 (5.8%)b</td>
</tr>
<tr>
<td>CIN 2–3</td>
<td>222 (0.7%)b</td>
</tr>
<tr>
<td>Number with invasive cancer</td>
<td>97</td>
</tr>
<tr>
<td>Number with stage I cancer</td>
<td>34 (35.0%)c</td>
</tr>
<tr>
<td>Number with stage II cancer</td>
<td>18 (18.6%)c</td>
</tr>
<tr>
<td>Number with stage III cancer</td>
<td>45 (46.4%)c</td>
</tr>
<tr>
<td>Number with stage IV cancer</td>
<td>0 (0.0%)</td>
</tr>
</tbody>
</table>

a Percentage of screened women
b Indicates detection rate of CIN per 100 screened women
c Percentage of all cancers

e From Sankaranarayanan et al. (2004d)
ment in the Osmanabad trial; with cytology, the rate remained constant in the latter study. A high proportion of invasive cancers were diagnosed in stage I in women screened with VIA. The ultimate efficacy of VIA in reducing cancer incidence and mortality will become clearer with follow-up for cancer incidence and mortality in these studies.

An innovative option taking advantage of the immediate availability of test results with VIA is the screen-and-treat or single-visit approach, to ensure high treatment compliance among screen-positive women. This approach is based on the following premises: studies have reported high sensitivity for VIA to identify precancerous lesions; the lack of or inadequate infrastructure and resources for diagnostic facilities such as colposcopy and histopathology in many low-resource settings; the possibility of high rates of loss to follow-up associated with multiple visits; and the potential for protection against cervical neoplasia among women who had ablation (with electrocoagulation or cryotherapy) of the ectopic cervical epithelium and the transformation zone (Vonka et al., 1984).

In the screen-and-treat approach, screen-positive women without clinical evidence of invasive cancer and satisfying the criteria for ablative therapy are immediately treated by cryotherapy, without confirmatory colposcopic or histological investigations. The safety, acceptability and feasibility of such a single-visit approach combining VIA and cryotherapy was assessed in a recent study in rural Thailand (RTCOC/JHPIEGO, 2003). Trained nurses tested 5999 women with VIA and 798 (13.3%) women were VIA-positive. Overall, 756 women received cryotherapy (either immediately or postponed). No major complications were recorded following cryotherapy; only 33 women (4.4%) of treated women returned for a perceived problem. At a one-year follow-up visit, the VIA test negative rate among treated women was 94.3%.

The efficacy of the screen-and-treat approach with VIA as compared to HPV testing and treatment in reducing the frequency of high-grade CIN is being assessed in a randomized clinical trial in South Africa, which has not yet published any results.

In order to assess how screen-and-treat with VIA will perform in a routine health service setting, a large demonstration project to screen women aged 30–49 years has been launched in the St Martin province of Peru. This programme aims to cover 80 000 women in three years; no results have yet been published.

A greater proportion of cervical cancers detected by VIA were in stage I than

<p>| Table 64. Initial results after the screening phase of the cluster-randomized controlled trial of visual inspection for cervical cancer with acetic acid, cytology and HPV DNA testing in Osmanabad district, India |</p>
<table>
<thead>
<tr>
<th>Group screened with VIA</th>
<th>Group screened with cytology</th>
<th>Group screened with HPV testing</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of women</td>
<td>34 149</td>
<td>32 136</td>
<td>34 515</td>
</tr>
<tr>
<td>Received screening</td>
<td>26 755</td>
<td>25 535</td>
<td>27 159</td>
</tr>
<tr>
<td>Screened positive</td>
<td>3731 (14.0%)</td>
<td>1790 (7.0%)</td>
<td>2812 (10.4%)</td>
</tr>
<tr>
<td>Number of screen positive</td>
<td>3682</td>
<td>1559</td>
<td>2475</td>
</tr>
<tr>
<td>who had colposcopy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of women who</td>
<td>2528</td>
<td>828</td>
<td>1114</td>
</tr>
<tr>
<td>received biopsy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-grade lesions</td>
<td>1068 (4.0%)</td>
<td>304 (1.2%)</td>
<td>327 (1.2%)</td>
</tr>
<tr>
<td>CIN 2 lesions</td>
<td>84 (0.3%)</td>
<td>103 (0.4%)</td>
<td>105 (0.4%)</td>
</tr>
<tr>
<td>CIN 3 lesions</td>
<td>112 (0.4%)</td>
<td>162 (0.6%)</td>
<td>138 (0.5%)</td>
</tr>
<tr>
<td>Number with invasive cancer</td>
<td>121</td>
<td>131</td>
<td>100</td>
</tr>
<tr>
<td>Number with stage I cancer</td>
<td>58 (47.9%)</td>
<td>67 (51.2%)</td>
<td>60 (60.0%)</td>
</tr>
<tr>
<td>Number with stage II cancer</td>
<td>18 (14.8%)</td>
<td>12 (9.2%)</td>
<td>9 (0.0%)</td>
</tr>
<tr>
<td>Number with stage III cancer</td>
<td>30 (24.8%)</td>
<td>29 (22.1%)</td>
<td>10 (10.0%)</td>
</tr>
<tr>
<td>Number with stage IV cancer</td>
<td>5 (4.1%)</td>
<td>1 (1.0%)</td>
<td>1 (1.0%)</td>
</tr>
</tbody>
</table>

a Percentage of screened women  
b Indicates detection rate of CIN per 100 screened women  
c Percentage of all cancers

From Sankaranarayanan et al. (2004c)
among cancers occurring in unscreened controls. The long-term impact of VIA screening in reducing cervical cancer incidence remains to be established.

**VILI**

Visual inspection with Lugol's iodine (VILI) involves naked-eye examination of the cervix to identify mustard-yellow iodine-non-uptake areas after application of Lugol's iodine. The test characteristics of VILI provided by nurses, midwives and trained non-medical workers have been studied in a set of ten cross-sectional studies, with a similar protocol, involving 49,080 women aged 25–65 years conducted in Burkina Faso, Republic of Congo, Guinea, India, Mali and Niger (Sankaranarayanan et al., 2004a). VILI was also simultaneously evaluated in all 10 studies, conventional cytology in five and HPV testing in three studies. No untoward reaction to iodine was observed. VILI had a significantly higher pooled sensitivity than VIA (91.7% versus 76.8%) but similar specificity (85.4% versus 85.5%) in detecting CIN 2–3 lesions. VILI had significantly higher pooled sensitivity than cytology (83.9% versus 45.4%) but lower specificity (82.5% versus 99.2%) in detecting histologically confirmed CIN 2–3 lesions (Sankaranarayanan et al., 2004f). VILI had similar sensitivity to that of HPV testing, but lower specificity, to detect histologically confirmed CIN 3 lesions in the pooled analysis of three cross-sectional studies (Sankaranarayanan et al., 2004b). There are no randomized trials evaluating the efficacy of VILI in reducing cervical cancer incidence and mortality.

**Human papillomavirus testing**

Almost all of the studies on HPV testing have focused on the sensitivity and specificity of the test under various conditions. No studies have prospectively investigated its impact on subsequent cancer rates, but a few have retrospectively studied the detectability of HPV in archival smears which were negative on cytology some years before a diagnosis of cancer. Issues of the persistence of HPV in high-grade lesions and the length of protection following a negative HPV test have been addressed in a few studies. A comprehensive review of the role of HPV testing in cervical screening appeared in 1999 (Cuzick et al., 1999b) and several important studies have been reported since then.

**Relative sensitivity for CIN 2 or 3 compared with cytology**

Most of the recent screening studies have used the Hybrid Capture™ 2 test for high-risk HPV types, which is the only test now commercially available. It is clear that this test is more sensitive than cytology for CIN 2 or 3 and for CIN 3, but it also has lower specificity (Table 65). The specificity improves if testing is restricted to women over the age of 30 years.

Typically, HPV testing has a sensitivity of 95% for detecting CIN 2 or worse lesions compared with 75% for cytology at the borderline (ASCUS) or above level and 70% for cytology at the mild dyskaryosis (LSIL) level (i.e., when the cytology threshold (or cut-off) is ASCUS or worse, or when it is LSIL or worse). Thus, virtually all of the lesions detected by cytology were HPV-positive, as were an additional 25% which were negative on cytology. In women over the age of 30 years, specificity is about 93%, compared with 95% for cytology at the borderline level and 98% at the mild level. For younger women, both tests have poorer specificity. For example, in the English screening programme, for cytology at the borderline cut-off, specificity is about 89% in women aged less than 30 and 96% for older women (NHS, 2003a). For HPV, the specificity is about 85% for women aged less than 30 and 93% for older women. Where available, the studies show even greater sensitivity for detecting CIN 3 (Table 65b).

**Lower sensitivities of both HPV and cytological testing are seen in developing countries.** In the three large European studies (Clavel et al., 2001; Petry et al., 2003; Cuzick et al., 2003), the sensitivity of HPV testing was uniformly high (97% or higher), whereas the sensitivity of cytology was lower and highly variable between countries.

**Retrospective studies of HPV evaluation**

Eleven published studies have evaluated HPV infection in stored material (Table 66). Seven used archival smears (de Roda Husman et al., 1995; Walboomers et al., 1995; Chua & Hjerpe, 1996; Wallin et al., 1999; Carozzi et al., 2000; Ylitalo et al., 2000a; Zielinski et al., 2001a), one used previous biopsy specimens (Konno et al., 1992) and three tested for HPV antibodies in stored serum samples (Chua et al., 1996; Lehtinen et al., 1996; Dillner et al., 1997).

Two of these studies (Konno et al., 1992; de Roda Husman et al., 1995) did not include controls. They looked at a total of 15 women with invasive cervical cancer and five with CIN 3, and examined smears and biopsies taken up to 10 years previously. All stored specimens tested positive for HPV16, 18 or an unknown type. Chua and Hjerpe (1996) analysing archival smears, with two matched controls per case, obtained odds ratios of 16, 11 and 18 for invasive squamous, adenocarcinoma and carcinoma in situ of the cervix based on 18, 12 and 58 cases, respectively. Walboomers et al. (1995) used as controls women from a gynaecological clinic, some of whom were being treated for CIN. They used general
### Table 65. Relative sensitivity and specificity of Hybrid Capture (HC) 2 compared with cytology on biopsy in cross-sectional screening studies

#### (a) for CIN2+

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cytology ≥ LSIL</td>
<td>HPV</td>
<td>Cytology &lt; LSIL</td>
</tr>
<tr>
<td>Clavel et al. (1999)</td>
<td>85</td>
<td>100</td>
<td>95</td>
</tr>
<tr>
<td>Cuzick et al. (1999a)</td>
<td>79</td>
<td>95</td>
<td>99</td>
</tr>
<tr>
<td>Schiffman et al. (2000)</td>
<td>78</td>
<td>88</td>
<td>94</td>
</tr>
<tr>
<td>Ratnam et al. (2000)</td>
<td>40</td>
<td>90</td>
<td>77</td>
</tr>
<tr>
<td>Denny et al. (2000a)</td>
<td>78</td>
<td>73</td>
<td>85</td>
</tr>
<tr>
<td>Denny et al. (2000b)</td>
<td>82</td>
<td>72</td>
<td>93</td>
</tr>
<tr>
<td>Schneider et al. (2000)</td>
<td>20</td>
<td>89</td>
<td>99</td>
</tr>
<tr>
<td>Womack et al. (2000)</td>
<td>44</td>
<td>81</td>
<td>91</td>
</tr>
<tr>
<td>Clavel et al. (2001)</td>
<td>68</td>
<td>100</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>88</td>
<td></td>
<td>93</td>
</tr>
<tr>
<td>Petry et al. (2003)</td>
<td>37</td>
<td>98</td>
<td>99</td>
</tr>
<tr>
<td>Kulasingam et al. (2002)</td>
<td>36</td>
<td>63</td>
<td>96</td>
</tr>
<tr>
<td>Cuzick et al. (2003)</td>
<td>70</td>
<td>97</td>
<td>99</td>
</tr>
<tr>
<td>Sankaranarayanan et al. (2004b)</td>
<td>37–72</td>
<td>46–81</td>
<td>87–98</td>
</tr>
<tr>
<td>Salmeron et al. (2003)</td>
<td>59</td>
<td>93</td>
<td>98</td>
</tr>
<tr>
<td>Nieminen et al. (2004)</td>
<td>83</td>
<td>98</td>
<td>94</td>
</tr>
</tbody>
</table>

#### (b) for CIN3+

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cytology ≥ LSIL</td>
<td>HPV</td>
<td>Cytology &lt; LSIL</td>
</tr>
<tr>
<td>Cuzick et al. (1999a)</td>
<td>79</td>
<td>100</td>
<td>99</td>
</tr>
<tr>
<td>Ferreccio et al. (2003)</td>
<td>63</td>
<td>85</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>86</td>
<td></td>
<td>88</td>
</tr>
<tr>
<td>Salmeron et al. (2003)</td>
<td>60</td>
<td>94</td>
<td>98</td>
</tr>
<tr>
<td>Petry et al. (2003)</td>
<td>40</td>
<td>97</td>
<td>99</td>
</tr>
<tr>
<td>Cuzick et al. (2003)</td>
<td>77</td>
<td>98</td>
<td>99</td>
</tr>
<tr>
<td>Kulasingam et al. (2002)</td>
<td>57</td>
<td>91</td>
<td>90</td>
</tr>
<tr>
<td>Sankaranarayanan et al. (2004b)</td>
<td>80</td>
<td>77–89</td>
<td>95</td>
</tr>
<tr>
<td>Nieminen et al. (2004)</td>
<td>86</td>
<td>100</td>
<td>93</td>
</tr>
</tbody>
</table>

Primers to probe archival smears and, consistent with other studies from this group, found a very strong association with high-risk HPV types. Sixteen of the 17 women with invasive carcinoma had HPV in archival smears compared with seven of the 50 controls, giving an odds ratio of 49. Further, all nine cases with two archival smears had the same type of HPV detected on both. The smears were taken between two months and six years before cancer diagnosis (median 1 year). By design, all smears were originally classed as normal. On reanalysis, four of the 26 archival smears from the cases were deemed inadequate, and the rest showed severe dyskaryosis or worse. Wallin et al. (1999) compared archival smears, all of which had normal cytology, from 118 women with subsequent cervical cancer with those from 118 controls. The average duration between smears and cancer was 5.6 years (range, 0–26 years). HPV was detected in 30% of the cases but only 3% of the controls (OR = 16.4; 95% CI 4.4–75). The PCR in this study used both MY09/MY11 and GP5/6 consensus primers.

Ylitalo et al. (2000a) reviewed all previous smears in 484 cases of carcinoma in situ and 619 matched controls.
in Uppsala, Sweden. Smears were available for up to 26 years before diagnosis. Only HPV16 was tested for. The case smears from 16–18 years before diagnosis were HPV-positive for about 10% of women, which was similar to controls, but the proportion rose linearly to 56% 2.3 years before diagnosis, which was highly significant. A positive HPV16 result on the two last smears before diagnosis was associated with an odds ratio of 31.2 (95% CI 10.6–91.8) for carcinoma in situ. The mean time from HPV positivity to diagnosis was estimated to be between 7 and 12 years.

Carozzi et al. (2000) assessed archival smears classified as normal from 79 cases of CIN 2 or worse and matched controls. They used a consensus system which detected HPV types 16, 18, 31, 33, 52 and 58. An odds ratio of 64 (95% CI 31–133) for CIN 2 or worse was found for HPV positivity in all smears, which rose to 103 (95% CI 43–251) in smears taken less than four years before the last cytological test; overall 77% of case smears were HPV-positive compared with 5.1% of control smears.

Zielinski et al. (2001a) examined the last normal archival smear from 57 women who later developed cancer and 114 control women, using GP5+/6+ primers. HPV was detected in 65% (37) of the smears from cases compared with 6% (7) of the control smears (OR = 28; 95% CI 11–72). Positivity was only slightly higher in smears taken within three years of diagnosis (76%) than in smears taken 4–20 years previously (65%).

Three studies (Chua et al., 1996; Lehtinen et al., 1996; Dillner et al., 1997) looked for HPV16 (or HPV16 and 18) antibodies in stored serum samples, using a nested case–control design. All three found increased risk of cancer or carcinoma in situ in women with prior seropositivity to HPV16. The odds ratios associated with HPV antibodies in these studies ranged from 3 to 13. A longer lag time from sampling of sera to diagnosis was associated with greater relative risk. Chua et al. (1996) estimated the progression rates to cancer or carcinoma in situ in women of different ages with and without HPV16 antibodies. The incidence of cancer or carcinoma in situ decreased with age, as did the relative risk associated with HPV16 antibodies, whereas seropositivity increased with age in the controls. A possible explanation of this finding is that CIN is associated with an active HPV infection and women who developed antibodies some years earlier no longer necessarily carry the virus. The study also looked at antibodies for HPV18 and 33, but these were not significantly associated with disease. The largest of the studies (Dillner et al., 1997) combined cohorts from Finland, Norway and Sweden and included 182 cases of invasive carcinoma. Overall, the relative risks were 2.7 for HPV16 antibodies and 2.2 for HPV16, 18 or 33 antibodies. The relative risk associated with HPV16 antibodies increased to 3.9 in those women with a lag time of over five years. The third study measuring antibodies (Lehtinen et al., 1996) included 27 cases of invasive cancer and 25 of carcinoma in situ. Overall, the odds ratio for risk of developing cervical carcinoma according to the presence of HPV antibodies was 13.2. It was greater for invasive cancer (OR = infinite; 95% CI 2.0–infinite) than for carcinoma in situ (OR = 6.0; 95% CI 1.2–29.7) and for lag times of over five years (OR = 18; 95% CI 2.3–142) compared with under five years (OR = 8.6; 95% CI 1.0–75). [Due to the lack of sensitivity of serological tests (50% at best for some high-risk HPV types), odds ratios from these studies are inevitably much lower than for direct DNA testing in cervical specimens.]

**Duration of protection**

The higher relative sensitivity of HPV testing for CIN 2 or 3 compared with cytology suggests that it might be safe to lengthen screening intervals if HPV testing were used. Two studies of this issue have been reported (Bory et al., 2002; Sherman et al., 2003a) and others are in progress. Bory et al. (2002) found that among 2432 women who were negative for high-risk HPV, only two (0.08%) developed high-grade CIN after a median follow-up of 27 months. Both cases were HPV-positive at the time of diagnosis (after 18 and 24 months). This was compared with 21.2% developing CIN 2 or 3 among women who were initially HPV-positive but cytologically normal.

Sherman et al. (2003a) reported a 10-year follow-up of 20 810 women screened by cytology and HPV DNA testing at Kaiser Permanente in Portland, Oregon. Cervical lavage specimens were used for HPV testing. A total of 171 women were diagnosed with CIN 3 or cancer on follow-up. HPV positivity was more sensitive for detecting CIN 3 on follow-up than cytology (89 versus 58 in the first 45 months and 110 cases versus 59 cases overall). Conversely, there were fewer cases in HPV-negative women (29 versus 60 in the first 45 months and 61 versus 112 overall). Detection rates were similar between tests in the first nine months (15 cases in HPV-positive women, 15 cases in HPV-negative women) but HPV negativity was much more protective than cytology after that (14 versus 45 cases in months 10–45; 46 versus 97 in months 10–112).

In addition, modelling studies have suggested that the use of HPV testing could safely allow the screening interval to be lengthened. Goldie et al. (2004) modelled the US data and recommended that women whose results are negative by both HPV DNA testing and cytology should not be
**Table 66. Studies retrospectively analysing stored samples for HPV DNA or antibodies**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Assay</th>
<th>Follow-up</th>
<th>Setting</th>
<th>Age (years)</th>
<th>Case</th>
<th>Control</th>
<th>Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>Konno et al. (1992)</td>
<td>In situ hybridization PCR of negatives</td>
<td>&lt; 10 y</td>
<td>Japan</td>
<td>CIN 3 (5)</td>
<td>Microinvasive (2) Invasive carcinoma (1)</td>
<td>All with HPV16/18 on hysterectomy section</td>
<td>Hysterectomy sections</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HPV 16 found in all previous biopsy specimens</td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Roda Husman et al. (1995)</td>
<td>General primer PCR</td>
<td>2–9 y; mean, 5.8 y</td>
<td>Netherlands Screening programme</td>
<td>Cervical cancer (12) None</td>
<td>50 controls from gynaecology clinic including women with CIN</td>
<td>Archival smears</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HPV found in archival smears of all 12 cases. Same type found in smear and biopsy of tumour.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walboomers et al. (1995)</td>
<td>GP5/6</td>
<td>2 mo–6 y; median, 1 y</td>
<td>Netherlands 3-yearly screening programme</td>
<td>17 cancers with normal archival smears</td>
<td>50 controls from gynaecology clinic including women with CIN</td>
<td>Archival smears</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Archival HPVhr Total</td>
<td>Women Smears</td>
<td></td>
</tr>
<tr>
<td>Chua &amp; PCR (nested):</td>
<td>i MY09/11</td>
<td>2–7 y; mean 3 y</td>
<td>Sweden</td>
<td>Adenocarcinoma (12) No carcinoma in situ for 5 y post-smear (but some history smears)</td>
<td>Carcinoma in situ (58) of abnormaly (age matched)</td>
<td>Archival smears</td>
<td></td>
</tr>
<tr>
<td>Hjerpe (1996)</td>
<td>ii GP5+/6+</td>
<td></td>
<td></td>
<td></td>
<td>Case</td>
<td>Control</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No. HPV+ (%)</td>
<td>No. HPV+ (%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Carcinoma in situ</td>
<td>58</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Invasive squamous-cell</td>
<td>18</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Adenocarcinoma</td>
<td>12</td>
<td>58</td>
</tr>
<tr>
<td>Chua et al. (1996)</td>
<td>Serology L1 and L2 capsids for HPV16 antibodies</td>
<td>&lt; 6.5 y; mean 3 y</td>
<td>Sweden</td>
<td>CIN 2/3 (41) CIN 1 (10) CIN not otherwise specified (23)</td>
<td>Population-based Match on age, date of blood Probability of CIN within 3 months given HPV seropositivity</td>
<td>Sera</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No. Seropositive (%) Age (years) + – Case</td>
<td>Control</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CIN 2–3</td>
<td>41</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CIN 1</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CIN not otherwise specified</td>
<td>23</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Controls</td>
<td>148</td>
<td>16</td>
</tr>
</tbody>
</table>
### Table 66 (contd)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Assay</th>
<th>Follow-up</th>
<th>Setting</th>
<th>Age (years)</th>
<th>Case</th>
<th>Control</th>
<th>Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lehtinen et al. (1996)</td>
<td>HPV16 L1 and L2 capsid</td>
<td>0.7–22.8 y; mean, 10 y</td>
<td>Finland</td>
<td>Mean, 39 at</td>
<td>27 invasive carcinomas</td>
<td>143 individually matched</td>
<td>Stored sera</td>
</tr>
<tr>
<td></td>
<td>antibodies</td>
<td></td>
<td>18814 women flagged to cancer registry</td>
<td>baseline, 49</td>
<td>at diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dillner et al. (1997)</td>
<td>L1 and L2 capsids to HPV16, 18</td>
<td>14% under 1 y; 33% over 5 y</td>
<td>Finland and Norway</td>
<td>&lt; 40</td>
<td>182 invasive carcinomas</td>
<td>~ 3 matched controls per case</td>
<td>Stored sera</td>
</tr>
<tr>
<td></td>
<td>and 33 antibodies</td>
<td></td>
<td>and Sweden population-based serum</td>
<td>(45%)</td>
<td>(total, 538)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&gt; 45 (26%)</td>
<td>Carcinomas</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wallin et al. (1999)</td>
<td>MY09/11 and GP5/6</td>
<td>0–26 y, median, 5.6 y</td>
<td>Sweden</td>
<td>19–74, mean</td>
<td>Cancers with prior negative</td>
<td>118</td>
<td>Archival</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>44</td>
<td>smears (118)</td>
<td></td>
<td>smears (and biopsies)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ylitalo et al. (2000a)</td>
<td>PCR HPV 16</td>
<td>0.26 y; median, 8 y</td>
<td>Sweden</td>
<td>20–70 median; 35</td>
<td>484 CIS (2228 smears)</td>
<td>619 matched controls (1806 smears)</td>
<td>Archival smears</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carozzi et al. (2000)</td>
<td>PCR (consensus HR primers HPV)</td>
<td>0–6 y; mean, 2.5 y</td>
<td>Florence, Italy</td>
<td>25–64</td>
<td>92 total smears</td>
<td>3 controls per case; total number</td>
<td>Archival</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15 CIN 2</td>
<td>Archival of control smears, 332</td>
<td>smears</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>59 CIN 3</td>
<td>All smears cytologically negative</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5 invasive</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zielinski et al. (2001a)</td>
<td>GP5+/6+ PCR</td>
<td>0–18 y</td>
<td>Netherlands 3-yearly screening programme</td>
<td></td>
<td>57 cancers with normal smear</td>
<td>114 age matched controls</td>
<td>Archival</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>smears</td>
</tr>
<tr>
<td>For HPV16, relative risk increased with increasing lag time: 3.9 (95% CI 1.6–9.6) over 5 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CIS, carcinoma in situ; HR, high-risk; OR, odds ratio. Modified and updated from Cuzick et al. (1999a)
re-screened before three years. Women with normal cytological results, but who are positive for high-risk HPV DNA, are at relatively low risk of having high-grade cervical neoplasia, and colposcopy should not be performed routinely in this setting. Instead, HPV DNA testing along with cervical cytology should be repeated in these women after 6–12 months. If test results of either are abnormal, colposcopy should then be performed. As a result of these recommendations, the American Cancer Society and the American College of Obstetricians and Gynecologists have approved the extension of screening intervals from annually to three-yearly in women aged over 30 years, when HPV is added to cytology. [The Working Group noted that modelling results are no substitute for direct evaluation of this question.]

Cuzick et al. (2003) showed that women who originally had negative or borderline cytology but were subsequently found to have CIN 2 or worse lesions remained HPV-positive. [The Working Group noted the importance of persistence in determining the likelihood of underlying high-grade disease, but that currently this could only be assessed by two tests less than one year apart.]

### Other screening methods

#### Liquid-based cytology systems

**Introduction**

Several literature reviews and meta-analyses have been published comparing the relative sensitivity and specificity of liquid-based cytology systems with those of conventional cervical testing (see Chapter 2).

The comprehensive systematic review and meta-analysis by Arbyn et al. (2004a) includes split-sample and more recent direct-to-vial studies. The authors noted that early studies often yielded favourable results for LBC when comparing test positivity rates for low-grade abnormalities on cytology only, whereas in studies using detection rates for biopsy-confirmed CIN 2 or 3, no significant differences between conventional and liquid-based cytology were found.

This meta-analysis showed that test positivity was higher in direct-to-vial studies than in traditional testing, suggesting a bias in previous split-sample studies to the disadvantage of LBC. In direct-to-vial studies, more LSIL and HSIL lesions were identified in both ThinPrep and SurePath LBC systems (Table 67) (Arbyn et al., 2004a). Identification of equivocal samples (ASCUS) was similar. The authors were concerned that increased identification of cytological abnormalities (ratios >1) alone provides insufficient evidence for improved sensitivity of LBC in a screening programme and that verification with a valid gold standard is needed. Figures 52 and 53 show the test positivity ratios from studies available to the Working Group.

Only a few studies verified all cases (test positives and negatives) with a gold standard, allowing evaluation of sensitivity and specificity of both LBC and conventional cytology without verification bias (see Table 68). All these studies used the split sample design and none showed a statistically significant difference. In fact, the relative sensitivity was somewhat lower for LBC than for conventional testing.

Consideration of the positive predictive value (PPV) would allow determination of whether higher cytological positivity rates with LBC are due to an increase in false-positive tests. As positive predictive values for presence of CIN 2+ pooled from studies with at least an 80% gold standard verification of test-positives by colposcopically directed biopsy did not significantly differ from that of conventional cytology, this was probably true.

#### Evidence of efficacy from LBC in routine cervical screening programmes

Experience with LBC as the routine test in screening programmes is relatively recent and there is no long-term follow-up in terms of effects on incidence and mortality in the populations served. Thus the evaluation of efficacy has been made in terms of short-term surrogate markers such as relative sensitivity and specificity of LBC compared with those of high-quality conventional cytology. Short-term evaluations of LBC as the routine test in cervical screening in England, Scotland and Canada are described below.

A systematic review of the literature and modelling of cost-effectiveness, commissioned in the United Kingdom by the National Institute for Clinical Excellence (NICE) (Payne et al., 2000), concluded that, despite the lack of published studies providing direct evidence regarding cost-effectiveness of LBC for cervical screening, it was likely that LBC would reduce the number of inadequate samples, reduce the number of false negative results and decrease the time required for examination of specimens by cytologists. NICE immediately commissioned a full cost-effectiveness trial of LBC in a low-prevalence population for routine screening. LBC was introduced, after a learning transition period of 3 to 6 months in three selected laboratories, as part of a 12-month pilot project aiming at 100 000 routine screening tests. In two laboratories, the ThinPrep system was introduced and SurePath in another. The cytological results of the first six months of the pilot period were compared with the four previous years where exclusively conventional cytology was used (Moss et al., 2003). It was noted that different sampling devices were used before (wooden Aylebury spatula) and after (Cervex broom) the introduction of LBC, but there is published evidence that there is no statistical difference in sensitivity.
Table 67. Meta-analysis: pooled ratios of test positivity rates for liquid-based cytology (direct-to-vial studies) versus conventional testing

<table>
<thead>
<tr>
<th>Test threshold</th>
<th>ThinPrep Pooled estimate</th>
<th>95% CI</th>
<th>No. of studies</th>
<th>SurePath Pooled estimate</th>
<th>95% CI</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSIL+</td>
<td>1.72</td>
<td>1.42–2.08</td>
<td>21</td>
<td>1.47</td>
<td>1.14–1.89</td>
<td>7</td>
</tr>
<tr>
<td>LSIL+</td>
<td>1.74</td>
<td>1.47–2.06</td>
<td>21</td>
<td>1.52</td>
<td>1.24–1.86</td>
<td>7</td>
</tr>
<tr>
<td>LSIL</td>
<td>1.80</td>
<td>1.52–2.12</td>
<td>21</td>
<td>1.54</td>
<td>1.25–1.90</td>
<td>7</td>
</tr>
<tr>
<td>ASC+</td>
<td>1.23</td>
<td>1.07–1.40</td>
<td>19</td>
<td>1.19</td>
<td>0.96–1.46</td>
<td>7</td>
</tr>
<tr>
<td>ASC</td>
<td>0.95</td>
<td>0.84–1.09</td>
<td>19</td>
<td>0.93</td>
<td>0.67–1.31</td>
<td>7</td>
</tr>
</tbody>
</table>

From Arbyn et al. (2004a)

**Figure 52** Ratios of test positivity for HSIL+ in direct-to-vial studies with ThinPrep LBC, compared with conventional Pap cytology
< > indicates the combined effect estimated by random-effects model

between these two devices (Buntinx & Brouwers, 1996).

The proportions of inadequate tests before and after conversion to LBC are shown in Table 69. The evaluation showed an 80% reduction in the rate of inadequate samples in all laboratories and in all age groups after introduction of LBC. The rate was lowest in the SurePath laboratory. The quality of conventional smears increased with age, but no age differential was observed with LBC. The inadequate rate in English laboratories is higher than in many other countries (average 9%).

The relative changes in the identification of cytological abnormalities with LBC versus conventional testing by
The age group and by laboratory are shown in Table 70. There was no significant increase in the rates of HSIL when averaged across the three sites. However, in the ThinPrep laboratories, significantly more SIL and HSIL lesions were found, and one of them (lab C) also found more borderline lesions. In the laboratory where SurePath was used, less HSIL and borderline lesions were detected. The reason for this difference is not known. The increased identification of LSIL or worse lesions by SurePath was concentrated in the 20–34-year age group.

The reduction in inadequate rate should lead to fewer tests being performed, with a resulting decrease in workload for laboratories and primary care as well as recall systems. Referrals to colposcopy are likely to be affected only if the overall reporting of high-grade lesions increases. Comparing the running costs of LBC with those of conventional testing was complex since they utilize different amounts of laboratory resources.

There was some debate in the United Kingdom on the extent to which published differences between LBC and conventional cytology represented a true improvement (Herbert & Johnson, 2001, Moseley & Paget, 2002); a more recent article (Coste et al., 2003) contradicted the National Health Service pilot findings.

In another LBC pilot study, conducted by the Scottish Cervical Screening Programme (SCSP, 2002), four regional groups of smear-takers were randomized into two groups, collecting respectively conventional cervical smears and ThinPrep preparations from women attending for routine or follow-up screening tests. ThinPrep alone was chosen as the LBC system studies, since Cytyc had a more established infrastructure to support the laboratories at the time and the numbers in the study (30,000 LBC) were insufficient for evaluation of LBC with two systems. Smears and ThinPrep LBC vials were sent to four selected laboratories where cytotechnologists had received training in the interpretation of ThinPrep slides. The LBC samples were collected using the plastic cervix broom, while conventional smears continued to be collected by the wooden Aylesbury spatula. The results, summarized in Table 71, showed a sharp reduction in the rate of unsatisfactory.
factory samples and a significant improvement in the identification of high-grade lesions (between 3 and 9 women per 1000 tested). Reduced workload and increased productivity were also demonstrated in laboratories. In Ontario, Canada, SurePath was adopted for routine cervical screening in large screening laboratories in 2001, after training of large numbers of cytotechnologists. Almost one million routine cervical screening results were reviewed using the Ontario Provincial database. The results for 445,011 SurePath samples reported between January and June 2002 were compared with 445,225 conventional smear results from the same period in 2001 (Colgan et al., 2004; McLachlin et al., 2004). All slides had been screened manually. The SurePath cases showed 21% higher reporting of LSIL+ but a 15% decrease in HSIL detection. Assessment with the addition of colposcopic diagnostic rates is in progress to determine the relative sensitivity and specificity of the SurePath technology as a routine screening test.

### Table 68. Ratio of sensitivity and specificity for CIN 2+ of two liquid-based cytology systems (LBC) relative to the conventional Pap smear (CP), pooled from studies with complete verification by colposcopy and/or biopsy

<table>
<thead>
<tr>
<th>Preparation system</th>
<th>Ratio of sensitivities (LBC/CP)</th>
<th>95% CI</th>
<th>No. of studies</th>
<th>Ratio of specificities (LBC/CP)</th>
<th>95% CI</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>SurePath LBC</td>
<td>0.95</td>
<td>0.81–1.11</td>
<td>1</td>
<td>0.94</td>
<td>0.87–1.01</td>
<td>1</td>
</tr>
<tr>
<td>ThinPrep LBC</td>
<td>1.08</td>
<td>0.90–1.30</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional Pap smear</td>
<td>0.95</td>
<td>0.88–1.03</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Test threshold HSIL+
Adapted from Arbyn et al. (2004a)

### Table 69. Prevalence of inadequate specimens by preparation system

<table>
<thead>
<tr>
<th>Preparation system</th>
<th>% inadequate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional Pap smear</td>
<td>9.7</td>
<td>9.4–10.0</td>
</tr>
<tr>
<td>ThinPrep LBC</td>
<td>2.0</td>
<td>1.8–2.2</td>
</tr>
<tr>
<td>SurePath LBC</td>
<td>0.9</td>
<td>0.8–1.1</td>
</tr>
</tbody>
</table>

Computed over four years of conventional cytology and the first six months’ use of LBC
From Arbyn et al. (2004a), adapted from Moss et al. (2003)

### Table 70. Relative change in test positivity for abnormal results, SIL and HSIL in liquid-based cytology in comparison with conventional cytology, by laboratory and by age (crude and weighted)

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>Abnormal (≥ borderline)</th>
<th>RR</th>
<th>95% CI</th>
<th>SIL (≥ mild dyskaryosis)</th>
<th>RR</th>
<th>95% CI</th>
<th>HSIL (≥ moderate dyskaryosis)</th>
<th>RR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.99</td>
<td>0.94–1.04</td>
<td>1.25</td>
<td>1.19–1.31</td>
<td>1.10</td>
<td>0.98–1.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.94</td>
<td>0.90–0.99</td>
<td>1.00</td>
<td>0.93–1.08</td>
<td>0.85</td>
<td>0.75–0.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1.37</td>
<td>1.30–1.44</td>
<td>1.73</td>
<td>1.60–1.86</td>
<td>1.55</td>
<td>1.36–1.76</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Age**

<table>
<thead>
<tr>
<th>RR</th>
<th>95% CI</th>
<th>RR</th>
<th>95% CI</th>
<th>RR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–34</td>
<td>1.15</td>
<td>1.11–1.19</td>
<td>1.25</td>
<td>1.19–1.31</td>
<td>1.18</td>
</tr>
<tr>
<td>35–49</td>
<td>0.96</td>
<td>0.91–1.02</td>
<td>1.03</td>
<td>0.94–1.13</td>
<td>1.01</td>
</tr>
<tr>
<td>50–64</td>
<td>0.93</td>
<td>0.85–1.01</td>
<td>1.03</td>
<td>0.87–1.23</td>
<td>0.75</td>
</tr>
<tr>
<td>20–64 (MH)</td>
<td>1.06</td>
<td>1.03–1.09</td>
<td>1.18</td>
<td>1.13–1.23</td>
<td>1.10</td>
</tr>
</tbody>
</table>

MH: overall Mantel–Haenzel-adjusted relative risk
Laboratories A and C used Thin Prep; B used Sure Path
From Arbyn et al. (2004a), adapted from Moss et al. (2003)
Use of liquid-based cytology systems in cervical screening programmes

In May 1996, the ThinPrep T2000 processor was approved by the US Food and Drug Administration (FDA) for use in cervical screening on the basis of lower inadequacy rates and higher identification of LSIL and HSIL in comparison with conventional cytology and this was extended to the fully automated ThinPrep T3000 processor in May 2000. The SurePath approval by FDA in June 1999 indicated improved specimen quality and equivalent identification of cytological abnormalities to conventional cytology and in May 2003 a claim for increased HSIL identification was approved by the FDA. These two LBC systems now account for over 80% of cervical screening tests in the USA.

The Scottish Department of Health decided in 2002 that LBC should be implemented as the routine screening test throughout Scotland, following the Scottish LBC pilot project (see above). Scottish laboratories opted to use the ThinPrep system and all laboratories are now fully converted. In England and Wales, the NICE (2003) recommended to the National Health Service that LBC be introduced as the primary means of processing samples in its cervical screening programme. Health technology assessments in several other countries have not yet led to approval of LBC, although it is already widely used in the private sector in other parts of the world.

Other liquid-based systems are marketed, but there is little or no evidence in peer-reviewed literature for their efficacy (Johnson et al., 2000; Bergeron & Fagnani, 2003; Alves et al., 2004).

Automation-assisted devices

Automation-assisted screening is aimed at enhancing the performance of manual microscopic screening by excluding some of the normal slides from manual screening or by relocating the most suspicious cells down the microscope or collecting images into a gallery for review on computer screens. These technologies have the potential to decrease the fatigue of the user, allow 50% more slides to be reviewed per day, decrease the screening false negative rate due to human error, with appropriate decision support, identify morphological features that are not apparent in routine human review and, in particular, identifying small numbers of small abnormal cells, known to be very difficult to find in conventional screening.

Most of the automated scanning devices are capable of processing either conventional or liquid-based smears, potentially allowing their use in different kinds of screening programmes. Since LBC systems deposit cells onto a thin layer on a microscope slide, problems of cell overlap and obscuring debris are mitigated and thus LBC facilitates the performance of computer-assisted imaging (see Chapter 2)

Technological developments are very rapid in this area and several new approaches are emerging. A development of the FDA-approved AutoPap® system is the FocalPoint™ system, which is designed for use with SurePath (LBC) slides. The few published studies on the FocalPoint system suggest a performance equivalent to that of AutoPap (Cengel et al., 2003, Parker et al., 2004).

Another system that has become commercially available recently is the Cytyc ThinPrep Imager. Again, only a few studies on the performance of this system have been reported, but these show statistically significant improvement in sensitivity of the Imager review method over the conventional manual review for HSIL+. Specimen adequacy can be determined with the Imager

### Table 71. Proportion of inadequate specimens and cytological abnormalities and ratios observed in two randomized groups of women from four Scottish areas

<table>
<thead>
<tr>
<th>Cytological category</th>
<th>Proportion</th>
<th>Ratio*&lt;sub&gt;(ThinPrep/ conventional)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory</td>
<td>1.86%</td>
<td>0.23</td>
<td>0.20–0.26</td>
</tr>
<tr>
<td>Borderline</td>
<td>3.67%</td>
<td>0.84</td>
<td>0.76–0.94</td>
</tr>
<tr>
<td>Mild dyskaryosis</td>
<td>2.10%</td>
<td>1.93</td>
<td>1.60–2.32</td>
</tr>
<tr>
<td>Moderate dyskaryosis</td>
<td>0.97%</td>
<td>2.01</td>
<td>1.52–2.66</td>
</tr>
<tr>
<td>Severe dyskaryosis</td>
<td>1.09%</td>
<td>1.84</td>
<td>1.42–2.38</td>
</tr>
</tbody>
</table>

*Confidence intervals are approximate. The SCSP report stated that in half of the total of 30 288 women a conventional Pap smear was taken and in the other half a ThinPrep, so the calculation of confidence intervals was made assuming exactly 15 144 individuals in each group.

From Arbyn et al. (2004a), adapted from SCSP (2002)
review method. Cytotechnicians were able to double their daily work output while maintaining the same quality (Biscotti et al., 2003; McKee et al., 2003).

A few randomized prospective studies using the obsolete Papnet system (Nieminen et al., 2003) (see Chapter 2) show that automation-assisted screening is feasible in routine primary screening and that it performs in organized screening programmes at least as well as conventional manual microscopy. It is suggested that automation-assisted screening would not improve the outcome of an optimal cervical cytology service.

### Efficacy of screening among HIV-positive women

There are no data specifically on the efficacy of screening in HIV-positive women. In considering this group of women, three aspects of screening need to be taken into account.

- The accuracy of cytology as a screening test in HIV-positive versus HIV-negative women;
- The natural history of preinvasive disease of the cervix in HIV-positive versus HIV-negative women;
- The impact of anti-retroviral treatment on the natural history of preinvasive lesions of the cervix.

### Accuracy of cytology in HIV-positive women

Several studies have addressed the accuracy of the cytological test in HIV-positive women. Maiman et al. (1991) reported that the false negative rate of cytology was significantly higher in HIV-infected women than in HIV-uninfected women and recommended that routine colposcopy and histological sampling be performed in these women. A follow-up study confirmed the earlier findings (Maiman et al., 1998), based on an evaluation of 285 HIV-infected and 685 HIV-negative women, among whom 255 of the HIV-infected women underwent colposcopy and biopsy. Abnormal cytology detected 62% of all biopsy-confirmed CIN and 83% of all high-grade CIN in the HIV-infected women, but 38% of all CIN diagnosed would have been missed had colposcopy not been performed. The false negative rate of cytology, however, was not significantly different from that recorded in the HIV-negative group, reflecting the limitations of cervical cytology as a screening test rather than a poorer performance of cytology in HIV-infected women.

Fink et al. (1994), in a cross-sectional analysis of 51 HIV-positive women examined by cytology, colposcopy and biopsy, showed that there was good reproducibility of cytological results in HIV-infected women, but a high false negative rate for cytology compared to colposcopy. For instance, of 29 women who had normal cytological findings, 24% had CIN on biopsy. The authors recommended that this high false negative rate of cytology and the high prevalence of CIN in HIV-infected women warranted the inclusion of routine colposcopy for all HIV-infected women as a part of primary screening.

Korn et al. (1994) tested 52 HIV-positive women by cytology, colposcopy and histology, and a group of 85 women who self-reported HIV-negative status. The prevalence of CIN was 50% in the HIV-infected group and the sensitivity of cytology was 63% with a specificity of 84%. The performance of cytology in the control group was similar, however, and it was concluded that the accuracy of cytology was not significantly lower in HIV-positive women. The authors noted a high rate of loss to follow-up in HIV-positive women, as well as a significant incidence of concurrent lower genital tract pathology, which may justify initial colposcopic evaluation in this group.

Spinillo et al. (1998) reported on a cross-sectional study of 241 HIV-positive women and 991 controls (404 known HIV-negative and 587 of unknown HIV status). Among HIV-positive women, the sensitivity of cytology was 73% and the specificity 97%. The corresponding figures for the control group were 84% and 99% and the differences were not statistically significant. However, the negative predictive value of cytology was significantly lower in the HIV-positive group. The authors suggested more frequent screening of HIV-positive women rather than primary screening with cytology, colposcopy and biopsy.

Goodman et al. (2000) undertook a prospective study of cytology and concurrent colposcopically directed biopsies in HIV-positive and -negative women to determine the accuracy of cytology in the two groups. Among 82 HIV-positive women, the prevalence of CIN was 37%, compared with 17% in the HIV-negative group; the false negative rates of cytology in the HIV-positive and -negative women were 37% and 21% respectively, if ASCUS findings on cytology were included among the negative results. These false negative rates fell to 10% and 14%, respectively, if ASCUS was counted as a positive result. The authors concluded that ASCUS diagnosis comprised the majority of false negative calls in HIV-positive women and they too recommended an initial screening colposcopy to detect cases of CIN missed by cytology. Thereafter, they recommended six-monthly cytological screening.

Boardman et al. (1994), who compared 41 HIV-positive women with 228 HIV-negative and 409 women with unknown HIV status, also found no difference in the performance of cytological testing between HIV-positive women and women of negative or unknown HIV status. With HIV-negative women as the reference group, the
relative risk of cytology–histology discrepancy was 1.1 for HIV-positive women compared with 1.5 for women whose HIV status was unknown.

Adachi et al. (1993) performed colposcopy on 48 women with a cytological diagnosis of SIL, of whom 95% had colposcopic or histological findings that were no more severe than the cytological result. They concluded that the positive predictive value was high. Del Priore et al. (1995) also reported a high PPV of abnormal cytology to predict disease in a series of 52 HIV-positive women. The PPV of cytology was 96% for HIV-positive women versus 78% for HIV-negative women. The sensitivity of cytology among HIV-positive women was only 57%, with a specificity of 92%. The authors concluded, however, that prediction of the presence and degree of an intraepithelial lesion by abnormal cytology was no worse in HIV-positive than in HIV-negative women.

Where colposcopy services are readily available and accessible, an initial colposcopy may be warranted, particularly in women at greatest risk of disease, such as those with significant immune compromise or borderline abnormal cytology. Where such a service is lacking, more frequent cytological surveillance, e.g., six-monthly, has been recommended, although with little supporting evidence.

**Natural history of preinvasive cervical disease in HIV-positive versus HIV-negative women**

It is now clear that women infected with HIV have a higher prevalence of infection with HPV and are more likely to develop persistent infection with multiple types of HPV, as well as having a higher incidence and prevalence of preinvasive lesions of cervix, possibly a more rapid progression to cervical cancer and a higher incidence of cervical cancer (Schafer et al., 1991; Klein et al., 1994; Wright et al., 1994; Sun et al., 1997; Palefsky et al., 1999; Ellerbrock et al., 2000).

In 1992, the Centers for Disease Control and Prevention (CDC) included cervical cancer as an AIDS-defining disease, on the basis of extrapolation of data on the higher frequency of CIN in HIV-positive than in HIV-negative women (Centers for Disease Control and Prevention, 1992). A number of studies have supported this view and shown that HIV-positive women generally present with more advanced lesions and have a poorer prognosis than HIV-negative women (Maiman et al., 1993, 1997).

In a large population-based study by the AIDS-Cancer Match Registry Study Group in the USA, Frisch et al. (2000) found that, compared with the general population, HIV-infected individuals were at considerably increased risk for all types of anogenital HPV-associated cancers and their precursor lesions. This elevated risk spanned the decade from five years before the onset to five years after the diagnosis of AIDS. The relative risks of in situ cancer of the cervix (N = 722) (4.6; 95% CI 4.3–5.0) and invasive cancer (N = 44) (5.4; 95% CI 3.9–7.2) were similar in women with HIV infection or AIDS.

In the population-based Cancer and AIDS Registry Linkage Study in Italy (Dal Maso et al., 2003), women with HIV infection or AIDS had a relative risk of invasive cervical cancer (N = 18) of 21.8 (95% CI 12.9–34.6).

Lomalisa et al. (2000) presented data from South Africa on 60 HIV-seropositive and 776 HIV-seronegative women with newly diagnosed invasive cervical cancer. HIV-positive women presented with cervical cancer almost 10 years earlier than HIV-negative women (mean age 44 years versus 53 years), although the stage distribution was not different in the two groups, a finding of particular importance for screening programmes. In addition, severely immunocompromised women (e.g., CD4+ counts below 200 cells/µL) were significantly more likely to have advanced-stage disease at initial diagnosis than HIV-negative women.

A study in Senegal (Hawes et al., 2003) provided support for these conclusions, showing that HIV infection was associated with increased rates of cervical infection with high-risk types of HPV and that high-grade cervical cancer precursors and invasive cervical cancer were significantly more common in HIV-positive than in HIV-negative women (OR = 8.0; 95% CI 2.0–31.5). The degree of cervical abnormality was related to increased HIV viral load and increased immunosuppression, as expressed by low CD4+ cell counts.

Sitas et al. (2000) identified 167 cases of invasive cervical cancer among HIV-positive South African women versus 1323 among HIV-negative women (OR = 1.6; 95% CI 1.1–2.3), suggesting an increased cervical cancer risk among HIV-positive women in a country with high rates of cervical cancer and HIV seropositivity.

Numerous studies have indicated an increased prevalence of preinvasive lesions of the cervix in HIV-positive women. Mandelblatt et al. (1992) reviewed 21 studies from 1986 to 1990, and found five studies with sufficient data and a comparison group. All five studies showed a significant association between HIV infection and CIN, with an odds ratio for HIV-positive women of 4.9 (95% CI 3.0–8.2) compared with HIV-negative women.

Wright et al. (1994) conducted a cross-sectional study of 398 HIV-positive and 357 HIV-negative women; 20% of HIV-positive women had colposcopically confirmed CIN, compared with 4% of the HIV-negative women. The sensitivity and specificity of cytological testing did not differ significantly between the two groups and the authors concluded that cytology was
an effective screening test in HIV-positive women. In addition, by multiple logistic regression analysis, CIN was found to be associated with HPV infection (OR = 9.8), HIV infection (OR = 3.5), CD4+ T-lymphocyte count of less than 200 <mu>L (OR = 2.7) and age greater than 34 years (OR = 2.0).

Delmas et al. (2000) reported on the effect of immunodeficiency on the prevalence and incidence of SIL in 485 HIV-positive women. Compared with prevalence and incidence of SIL in 485 women with CD4+ counts of over 500 HIV-positive women. Compared with prevalence and incidence of SIL in 485 the effect of immunodeficiency on the prevalence and incidence of SIL in 485 women. Risk factors for any abnormal SIL were CD4+ counts lower than 200 <mu>L (OR = 2.7) and age greater than 34 years (OR = 2.0).

In the Women's Interagency HIV study (WHIS) (Massad et al., 1999), baseline cytology in 1713 HIV-positive women and 482 high-risk HIV-negative women was abnormal in 38% of HIV-positive women compared with 16% of HIV-negative women. Risk factors for any abnormal cytology were CD4+ counts lower than 200 <mu>L (OR = 2.13; 95% CI 1.45–3.13), presence of HPV DNA and history of abnormal cytology.

Ellerbrock et al. (2000) showed that the prevalence of abnormal cytological findings was 4.3-fold higher in HIV-infected than in uninfected women, confirming the findings of other studies showing abnormal cytology rates of 23–60% (Provencher et al., 1988). HIV-positive women were 4.5 times more likely to have histologically confirmed CIN at 54 months of follow-up than HIV-negative women.

Ahdieh et al. (2000) followed 84 HIV-negative and 184 HIV-positive injection drug users with six-monthly visits. Of the HIV-positive women, 70% were HPV DNA-positive at baseline compared with 26% of HIV-negative women. Cervical abnormalities were found in 13% of HIV-infected women versus 2% of HIV-negative women.

Following treatment, HIV-positive women have generally shown high recurrence rates ranging from 38% to 62%, compared with 15–18% in HIV-negative women (Petry et al., 1994; Maiman et al., 1999; Chirenje et al., 2002).

**Impact of anti-retroviral therapy on the natural history of preinvasive cervical lesions**

The use of anti-retroviral therapy (ART) for treatment of HIV-infected individuals in developed countries has substantially reduced the associated morbidity and mortality. The increase in life expectancy may affect the burden of cervical cancer in either direction, depending on the degree to which the immune reconstitution allowed by ART is sufficient to diminish the risk of cervical cancer. The International Collaboration on HIV and Cancer (2000) found no change in the incidence of cervical cancer between 1992–96 and 1997–99 (i.e., after the use of ART had become widespread) in a reanalysis of cancer risk in 23 cohort studies in developed countries. This contrasted with marked reductions in the incidence of Kaposi sarcoma and non-Hodgkin lymphoma.

Recent studies have shown a beneficial impact of ART, with greater regression of HPV-associated lesions in treated women. In a cohort of French women, the prevalence of cervical HPV infection among 34 HIV-positive women remained unchanged at 81% five months after the initiation of ART. However the prevalence of CIN decreased from 69% to 53% after a median of five months (p = 0.04) and the mean CD4+ cell count was higher among women who regressed, suggesting that the loss of immune response and ART-induced reconstitution of immunity may have played a role in protecting against CIN (Heard et al., 1998).

In a long-term follow-up study (Heard et al., 2002) (median follow-up 17.7 months) of 168 HIV-positive women, 96 of whom were receiving ART, regression of CIN (defined as a regression to normality or to a lower grade of CIN) was seen in 40% of the 168 women. In a multivariate analysis, the grade of the lesion and the use of ART were independently associated with regression of CIN, after adjustment for CD4+ cell count. The relative hazard of regression of CIN in women receiving ART was 1.93 (95% CI 1.14–3.29; p = 0.01) compared with untreated HIV-positive women. In addition, a trend for a greater increase in CD4+ cell counts after six months of ART was observed in women who regressed compared with those who did not.

Minkoff et al. (2001) reported that women on ART were 40% more likely to show regression of cervical cytological abnormalities towards normality or lower-grade disease and less likely to show progression (OR = 0.68; 95% CI 0.52–0.88), after control for stage of HIV disease and severity of cytological abnormality. Among HIV-infected women, persistence of HPV infection and high HIV viral load were associated with cytological progression. Conversely, low HIV viral load and high CD4+ cell counts were associated with regression.

Moore et al. (2002) reported on 71 HIV-positive women who were examined by cytology, colposcopy and biopsy before starting ART and had at least one similar assessment six months after starting ART. The baseline prevalence of cervical disease was 55%, and at six months after starting ART 13% of the women showed regression without treatment of the cervix. No individual factor (e.g., smoking, HIV viral load, stage of HIV disease or CD4+ cell count) was significantly associated with regression, although a greater increase in CD4+ cell count in women on ART was most strongly associated with regression.
(OR = 1.66 per 50 cell increase, \( p = 0.08 \)).

These data suggest that the best responses of cervical disease in women on ART are seen among women with higher CD4+ cell counts or that women who respond to ART are those with the largest increase in CD4+ cell counts. It is important to note, however, that progression from CIN to cervical cancer occurs over many years. Before ART was introduced, HIV-positive women most often died of other HIV-related diseases and there was insufficient time for the development of cervical cancer. If ART leads to prolongation of life, women with cervical disease may be at greater risk of developing invasive cervical cancer if they do not enter a screening programme. Until a clear impact of ART on the regression and progression of HPV-associated lesions of the cervix is confirmed, HIV-positive women being treated with ART should undergo cervical screening and be actively treated where appropriate.
Chapter 5
Effectiveness of screening in populations

This chapter deals with population measures of effectiveness of cervical cancer screening. At the population level, effectiveness is ultimately assessed by the reduction in mortality due to cancer of the cervix. There are two reasons to expect mortality to decrease as a result of screening: removal of incident cases through detection and treatment of premalignant lesions, and diagnosis of invasive lesions at earlier, more curable stages. Because screening for cervical cancer can detect precursor lesions that can be treated to prevent progression to invasive disease, reduction of cervical cancer incidence can be used as a measure of effectiveness as well.

Four methods have been used to assess the effectiveness of screening: individual-based studies using case–control or cohort designs (see Chapter 4); ecological analyses (correlating screening activity with changes in mortality or incidence rates across time, place or age group); modelling of screening policy and practice to estimate effectiveness; and evaluation of operational parameters of screening. The latter includes screening performance indicators such as participation, quality and adequacy of follow-up of positive test results. This chapter is concerned with the last three of these.

**Incidence and mortality trends in relation to screening**

Time trends are of considerable interest, in part for the light that may be shed on changes in exposure to etiological factors (especially between women of different generations) and in part as a means of evaluating the success, or otherwise, of screening programmes. Because of their comprehensive coverage and availability, mortality data are often used in studies of time trends; however, care is needed in doing so, on account of the changing proportions of deaths assigned to ‘Uterus, unspecified’ (NOS) (see Chapter 1), and possible changes in treatment-induced survival, which may be quite large if long time series are studied (Pontén et al., 1995).

A reduction over time in the incidence of invasive cervical cancer, especially in those age groups where screening is mostly targeted, is another long-term indicator of effectiveness. However, high-quality population-based incidence data, as provided by population-based cancer registries, are available in relatively few regions of the world (Parkin et al., 2002) and fewer still have incidence data covering extended periods of time.

**Time trends by region**

Until quite recently, most studies focused on the overall cervical cancer trends rather than looking separately at adenocarcinoma and squamous-cell cancer. However, cytological screening identifies mainly the latter. Since most cervical cancers are squamous-cell carcinomas, studies of overall cancer incidence and mortality largely reflect trends in this histological type.

**Developed countries**

**Europe**

Trends in cervical cancer incidence and mortality have been intensively studied in the Nordic countries, where it has also been possible to compare the trends in the different countries in relation to the intensity of screening undertaken (Hakama, 1982; Hakulinen et al., 1986, Läärää et al., 1987; Engeland et al., 1993; Sigurdsson, 1993, 1995; Hristova & Hakama, 1997; Anttila & Läärää, 2000; Moller et al., 2002). In these countries, national incidence and mortality data are available from before and after the times that screening programmes were implemented. Towards the end of the 1960s, Finland, Sweden and Iceland had nationwide, organized screening programmes, and the same was true for several Danish counties. Norway, in contrast, had organized screening only in a single county covering about 5% of the population. Throughout the Nordic countries, opportunistic testing also increased at the same time.
From the late 1960s, a decrease was seen in both the incidence of and mortality from cervical cancer in Finland, Sweden, Iceland and Denmark (Figure 54). The decrease, relative to the levels before screening, was largest in Finland, where the age-standardized mortality rate decreased more than 80% from 6.6 deaths per 100 000 in early 1960s to 1.2 deaths per 100 000 in the early 1990s (decrease 82%) (rates adjusted for age to the world standard population). In the earlier period, women 30–55 years of age were invited with a five-year screening interval and it was only in the early 1990s that the maximum age for invitation was raised to 60 years in Finland. The decreases in the mortality rates were 65% and 55%, respectively, in Sweden and Denmark, with partial coverage by organized programmes. A reduction in cervical cancer incidence was also observed in the Danish counties with organized screening compared to those without (Lynge et al., 1989). In Norway, the incidence increased until the mid-1970s, and the decrease in mortality was considerably less (41% from the early 1960s to the early 1990s) than in the other Nordic countries. At that time, opportunistic screening had become frequent also in Norway. A national organized programme of cervical cancer screening started in Norway in 1995 (Nygard et al., 2002). The trend in incidence was quite similar to the trend in mortality within each country up to the mid-1990s in terms of percentage reduction in the age-standardized rate. Also the incidence to mortality ratios were quite stable.

In general, incidence and mortality have also declined in the last 20–40 years in many other European countries (Coleman et al., 1993; Beral et al., 1994), but in some populations increases have been observed among younger women (aged under 35 years), particularly during the 1970s and 1980s (Figure 55). This was first noted in England and Wales, where generations of women born since about 1935 were observed to be at increasingly high risk (Hill & Adelstein, 1967; Cook & Draper, 1984; Parkin et al., 1985). Similar phenomena have been seen in Belgium (Vyslouzilova et al., 1997), Slovenia (Krn et al., 1992), Slovakia (Vlasak et al., 1991), Spain (Llorca et al., 1999) and in several other countries of eastern Europe (Beral et al., 1994).

Since the early 1990s, the incidence rate has started to increase in Finland among women below 55 years of age (Figure 56) (Anttila et al., 1999). This trend is probably due to a combination of changes in sexual lifestyles and increased transmission of papillomaviruses in younger generations of women, as well as inadequacies in the screening programme such as changes in laboratory procedures during this time (Nieminen et al., 2002). Because the effect of increasing incidence has been partly obscured by the protective effect of screening, in some countries, there has been little or no increase in
risk in young women (for example, Sweden (Figure 57); Bergström et al., 1999).

In the United Kingdom, cytological screening was introduced in the 1960s, but an organized programme, including a call/recall system and quality assurance, was implemented only from the 1988 onwards, leading to increased coverage within the targeted population. A sharp decrease in cervical cancer incidence and mortality rates since 1990 has been attributed to this organized programme (Sasieni et al., 1995, Gibson et al., 1997; Quinn et al., 1999; Sasieni & Adams, 1999) (Figure 58). The average drop in the age-adjusted mortality rate was estimated as 1–2% per year during
1960–88 and 7% since then (Sasieni et al., 1995).

Coding of deaths as due to cancer of the uterus NOS has been common in many countries and this affects comparability over time. In Belgium, an attempt has been made to estimate the proportions of deaths ascribed to cancer of the uterus NOS that should be redistributed to cervix and other uterine cancer (Arbyn & Geys, 2002). The corrected age-standardized mortality rates decreased from 14 per 100 000 in the 1950s to 4.5 in the 1990s (68% decrease), while the certified rates decreased from 6.3 to 3 (52% decrease).

In a number of eastern European countries such as Bulgaria, Romania and the Russian Federation, where little or no screening has taken place, cervical cancer mortality rates are rapidly rising (Figure 59), notably among recently born generations, as seen for Bulgarian women. In more affluent eastern European countries such as the Czech Republic, Hungary and Poland, there is some evidence of very recent declines (Figure 59), and in terms of birth cohort, mortality may have peaked among women born between 1945 and 1960 and then decreased, as observed in Hungary.

### North America

Overall, cervical incidence and mortality in the USA have declined for many decades in both black and white populations (Figure 60); this has been attributed to the effect of cytological screening programmes countering any increase due to changes in risk factors (Devesa et al., 1989). Increases at younger ages have not been observed in white or black women (Devesa et al., 1989; Wang et al., 2004).

In British Columbia, Canada, the age-adjusted incidence rate of squamous-cell cervical cancer was 28.4 cases per 100 000 woman-years in 1955, before the large-scale population-based centrally organized screening programme, and decreased to 6.4 in 1985 (a 78% decrease) (Boyes et al., 1981; Anderson et al., 1988). The corresponding mortality rate decreased from 11.4 deaths per 100 000 woman-years in 1958 to 3.1 deaths in 1985 (a 72% decrease). The lifetime coverage of cytological testing was estimated at 85% from 1970 onwards.

Although screening for cervical cancer commenced in North America towards the middle of the 20th century (in British Columbia, Canada, in 1949), there was a concomitant decline in mortality from the disease that initially did not seem to be associated with screening (Kinlen & Doll, 1973). Therefore, studies were initiated in the USA and Canada which attempted to evaluate the association between the extent of the decline in mortality from cancer of the cervix and the intensity of screening (Cramer, 1974; Miller et al., 1976). In both countries, a strong association was found when regional declines were analysed in relation to screening data from various sources. In Canada, the cytology data were derived from a national survey and the mortality data were rates among women aged 30–64 years, the ages at which mortality was expected to be most strongly associated with screening (Miller et al., 1976). Mortality from cancer of all parts of the uterus was used, as the extent to which deaths were attributed to cancer of the uterus NOS varied across the country and with time. The association between reduction in mortality and screening was strong at the national level for declines from 1960–62 to 1970–72, and at the census district level, and it was demonstrated that the decline was not explained by censusederived risk factors. In a further analysis, Miller et al. (1981) showed that the decline was not explained by changes in hysterectomy rates.

Subsequently, Miller (1986) re-examined the correlation between mortality rate and screening intensity in various parts of Canada for later time periods. Although he found consistent negative correlations between screening intensity and mortality rate at different points in time, he did not find consistent correlations with mortality reduction during time periods after the 1960s. Problems the author noted with this approach were possible changes in the underlying incidence of cervical cancer and the marginal effect expected with marginal increases in screening activity over time once a certain level of activity had been established.

### Australia and New Zealand

Although cervical cancer incidence rates in Australia (Figure 61, New South Wales) and New Zealand have not greatly decreased (Coleman et al., 1993), the mortality rates have been clearly declining in Australia for many decades; in women aged under 35 years, decreases have been seen since the mid-1980s in incidence and from around 1990 in mortality. Some increases in mortality rates during the 1970s were noted, notably in younger women (Armstrong & Holman, 1981), and a more recent study observed continuing period-specific declines in incidence and mortality from 1972 to 1996, alongside increasing rates in successive generations (Taylor et al., 2001). Increasing cohort-specific risks in women born in the late 1930s in New Zealand were reported (Cox & Skegg, 1986), but not confirmed later (Cox & Borman, 1994).

### Japan

Although incidence and mortality from cervical cancer in Japan have been reported to be falling for many decades (Figure 62) (Coleman et al., 1993), there is evidence of some increase (particularly in mortality) during the 1980s and 1990s in women aged under 35 years. In a study of cervical cancer incidence in Miyagi Prefecture during 1959–87, an age-period-
Figure 59 Age-standardized mortality rates of cervical cancer in Bulgaria, the Czech Republic, Hungary, Poland, Romania and the Russian Federation, ages 0–85+
cohort model showed that risk had decreased in recent periods and in younger generations of women (Minami et al., 1996).

Time trends in developing countries
There is limited information on time trends in cervical cancer in developing countries. In general terms, rates of incidence and mortality have been relatively stable or shown rather modest declines (Sankaranarayanan et al., 2001). The absence of the declines in incidence and mortality that have been observed in high-resource populations probably reflects the lack of screening programmes, or, where they exist, the low population coverage and poor quality of cytology (Lazcano-Ponce et al., 1998).

Figure 60 Cervical cancer incidence (----, black, ○ white) and mortality (––) trends in the USA, all ages

Figure 61 Cervical cancer incidence (----) and mortality (—) trends in Australia, New South Wales, all ages

Figure 62 Cervical cancer incidence (----) and mortality (—) trends in Japan, all ages

Figure 63 Annual cervical cancer mortality rates (per 100 000) in selected Latin American countries, age-adjusted to the world population, 1960–95. Five-year moving averages
Latin America

In contrast to most developed countries, mortality due to cervical cancer in Latin America increased between 1975 and 1985 (Restrepo et al., 1993). A later analysis (Robles et al., 1996) showed almost no significant downward change in mortality in Latin American countries between 1960 and 1993.

Figure 63 shows trends in age-adjusted cervical cancer mortality in eight Latin American countries between 1960 and 1994. In Puerto Rico, with rates similar to those of Mexico, Venezuela and Uruguay at the beginning of the period, a persistent decline has been observed, that gave it, by the end of the 1990s, the lowest risk in the region. This decline parallels the introduction of a screening programme (Robles et al., 1996), the effect of which can be seen in the progressive decline in age-specific rates, especially in the middle of the age range (30–69), where screening should have the highest effect (Figure 64a). In Cali, Colombia, a decline in the incidence of invasive carcinoma was accompanied by an increase in registrations of carcinoma in situ following the introduction of a screening programme in 1967 (Figure 64b) (Aristizabal et al., 1984). In Chile, Costa Rica, Cuba and Mexico, very limited changes in mortality from cervical cancer appear to have followed the introduction of screening. Mortality increased from 1965 onward in Mexico, where a national cervical cancer screening programme was initiated in 1974; although a slight decreasing trend has been observed since the 1990s, the risk remains among the highest in the region. In Costa Rica, cytology testing has been available nationwide to women aged over 15 years since 1970, but mortality and incidence have remained almost unchanged (Herrero et al., 1992). In Cuba likewise, the national screening programme was judged to have had no impact on either incidence or mortality in the period 1980–94 (Fernandez Garrote et al., 1996). In Chile, mortality rates increased steadily between 1960 and 1975, and then began to decrease, although rather slowly. This decline has been modest despite the operation of an organized screening programme since the early 1970s.

Figure 64 Age-specific incidence rates of cervical cancer in successive time periods a. Puerto Rico; b. Cali, Colombia
(Taucher et al., 1996; Sankaranarayanan et al., 2001). However, the proportion of cancer of the uterus NOS has steadily decreased from almost 50% at the beginning of the 1960s to around 10% in the 1990s, and this would have masked some of the decline in mortality from cancer of the cervix, as described above.

**Asia**

Figure 65 shows trends in cervix cancer incidence reported by the cancer registries of Mumbai (India) and Singapore. Declines in incidence are relatively modest (except for the Indian population of Singapore, among which the age-standardized rate declined from 29.8 per 100 000 in 1968–72 to 8.2 in 1993–97). In contrast, dramatic declines in cervix cancer in China have been reported. The age-adjusted incidence in Shanghai fell from 26.7 to 2.5 per 100 000 between 1972–74 and 1993–94 (Jin et al., 1999) and mortality rates have fallen dramatically, especially in urban populations, although the trend has reversed recently in younger women (Yang et al., 2003). The declines have been attributed to cytological screening, treatment programmes and improved female genital hygiene, while the increased rates among younger women may reflect changing economic circumstances and sexual habits leading to a greater prevalence of infection with HPV and other agents (Li et al., 2000).

**Africa**

There are very few data on time trends from Africa. In Bulawayo, Zimbabwe, the frequency of cervical cancer increased significantly during the period 1963–77 (Parkin et al., 1994). Mortality data from South Africa suggested some increase in rates for the ‘coloured’ population between 1949 and 1979, but little change in the black population from 1964 to 1977 (Bradshaw & Harington, 1985). After about 1980, the mortality in the ‘coloured’ population remained more or less constant, while in the white population, mortality declined from the mid-1960s (Bailie et al., 1996). The difference was ascribed to the availability of screening services, particularly for older women.

In some registry series, recent incidence rates appear to be higher than earlier ones. In Kampala, Uganda, for example, there has been a significant increase since the 1960s (Wabinga et al., 2000). On the other hand, there seems to have been little change in the recorded rate in Nigeria; it was 20.9 in 1960–69 and 19.9 in 1998–99 (Parkin et al., 2003).

**Caveats in the evaluation of time trends in relation to intensity of screening**

Trends in cancer incidence and mortality are a complex phenomenon to study, having substantial limitations and potential errors associated with them (Saxen, 1982; Muir et al., 1994). In addition, there are specific issues that concern the interpretation of time trends of cervical cancer, including changes over time in the proportions of deaths certified as uterus NOS and in the prevalence of hysterectomy (see Chapter 1).

The effect of screening is difficult to separate from the effects of other factors influencing rates of cancer diagnosis or death. For example, cervical cancer mortality rates were declining in North America before widespread screening was introduced, and the rate of decline changed little over the period 1946–74 despite considerably increased screening activity (Gardner & Lyon, 1977). This has also been noted in other parts of the world.
(e.g., Miller, 1999; Arbyn & Geys, 2002). Some studies have found underlying risk to be greater for recent-born cohorts (e.g., in Belgium, Arbyn & Geys, 2002; England and Wales, Parkin et al., 1985). The increased incidence in young Finnish women in the 1990s may also indicate changing risk, although some inadequacies in screening might also be responsible (Anttila et al., 1999). Because screening has been in use for decades in many parts of the world, especially in developed countries, it is difficult, if not impossible, to accurately estimate the risk of cervix cancer in its absence, in order to estimate the true impact of screening (van Ballegooijen et al., 2000). Further difficulties in interpreting trends result from a lack of information on the extent and quality of screening, particularly when a large proportion of tests are outside organized programmes.

**Age, period and cohort effects for squamous-cell carcinoma**

Trend studies generally fail to distinguish adenocarcinomas from squamous-cell carcinomas, although their etiology may be rather different, and their susceptibility to detection by cytological screening certainly is (Mitchell et al., 1995b, 2003). In an attempt to further evaluate the effectiveness of screening, the trends in incidence of the squamous-cell carcinoma by age at diagnosis, period of diagnosis, and birth cohort have been examined (Bray et al., 2004) using an age–period–cohort model (Case, 1956; Holford, 1983; Clayton & Schifflers, 1987a,b). An examination of cancer rates according to birth cohort may provide insight into the nature and intensity of disease-correlated exposures that may vary across successive generations. Cohort effects may relate to birth itself, or may appear to be related to birth only as a result of influences that are shared in the same group as they age together. Temporal changes in environmental risk factors tend to affect particular generations of individuals in the same way as they age together, and are more likely to exert particular influence on earlier stages of carcinogenesis.

Cancer rates by time, on the other hand, may act as surrogate measures of events that quickly change incidence or mortality with the same order of magnitude in all age groups under study. These effects may be the result of planned interventions that act at later stages of carcinogenesis, such as new therapies that improve survival in all age groups. More frequently, they are due to influences that artificially raise or lower the number of observed events (e.g., changes in classification or improvements in diagnostic procedures).

It is likely that any major effect of a general screening policy will be more visible in the period than in the cohort parameters. Such an interpretation is crude and obviously subject to uncertainties; thus, a screening policy may focus on a narrow window of ages and be of short duration, corresponding to a cohort.

Age is a powerful determinant of cancer risk, since it parallels the cumulative exposure to carcinogens over time and the accumulation of the series of mutations necessary for the unregulated cell proliferation that leads to cancer (Peto et al., 1985).

In presenting the age, period and cohort effects for squamous-cell cervical cancer incidence, the effect of age was fixed *a priori* as a biological constant. Two characteristic age curves that related the time before screening distorted the age–incidence pattern. Here the Gustafsson et al. (1997a) proposal was applied and the final choice for each population took into account the credibility of the curves from a biological point of view and empirical evidence that the subsequent period and cohort effects were in reasonable agreement with the observed trends.

Figure 66 provides estimates of squamous-cell carcinoma trends from age–period–cohort models for women aged 30–64 years. In Finland, the declines observed since screening was introduced in the 1960s have recently reversed, rates having steadily increased in cohorts of women born after 1945 (Figure 66a) and diagnosed in the 1990s. These model-based estimates are consistent with the observed overall time trends (Figure 56) (Anttila et al., 1999). Similar declines in period trend and fluctuation in cohort parameters are seen in Sweden (Figure 66b), although notable changes in rates in younger generations are not clear in the observed trend (Figure 57) (Bergstrom et al., 1999). In England, increasing rates are seen in generations born after 1935 (Figure 66c). In the observed trends, a deceleration in the rise has taken place among very recent generations (Vizcaino et al., 2000). The period parameters for England have reversed since the late 1980s, a finding which is consistent with the overhaul of the screening programme from 1988 (Walker et al., 1998; Quinn et al., 1999). In Estonia, where little screening has taken place (Aareleid et al., 1993), the period parameters have no trend, and there have been clear cohort-driven rises in women born since the mid-1930s (Figure 66d). In the USA, there are clear and uniform cross-sectional declines in period parameter, observed in the rates from the 1970s in both black and white women (Figure 60). Rates were relatively stable among successive generations of white women (Figure 66e) and steadily decreasing cohort trends among black women (Figure 66f).

In conclusion, the age–period–cohort modelling seems to confirm what is known on screening activities, effectively summarizing the data, as well as shedding additional light on the effects of etiological exposures.
Figure 66 Estimates of squamous-cell carcinoma trends from age–period–cohort models for women aged 30–64 years. Left-hand curve: cohort parameters. Right-hand curve: period parameters.
Use of modelling in the design and evaluation of screening

Statistical models have been developed to explore the effect of screening test, policy and programme characteristics on the expected reductions in incidence and mortality (and derivative quantities such as years of life saved). These have led to improved understanding of the relative importance of various screening parameters, which in turn has made it possible to infer what changes in screening programmes might be most effective. The quality of the models has improved over time as the underlying parameters (natural history, test sensitivity, etc.) have become better understood. The models have also become more widely used, as the contribution of the sophisticated methodology has become better appreciated and the statistical techniques more widely disseminated. As with any model, they depend on the availability of data and on the accuracy of the assumptions.

The pooling of several case–control and cohort studies (IARC Working Group on Cervical Cancer Screening, 1986) was used to estimate the reduction in incidence in a cohort of women as they age from 20 to 64 years under different assumptions of the ages of testing and its frequency (WHO, 1986). The results have been widely quoted and used as input data in various models.

In the absence of direct observations, models can examine the influence of variation in the underlying risk of disease over time or between population subgroups, and in the quality of screening procedures, on the effectiveness of screening as measured by incidence. Simulation models have been developed. These use observed data on the natural history of the disease, screening test performance and effectiveness of different options for treatment of precancerous lesions, and can allow for variation in risk, accessibility, compliance and feasibility. The early models of this type were used to examine the relative effectiveness of different programmes and were relatively simple computer simulations (Knox, 1976; Eddy 1980). More complex models use Monte Carlo simulation methods to provide greater flexibility and more realistic simulation of disease natural history (van Oortmarssen et al., 1981; Goldie, 2002). Gustafsson and Adami (1992) developed a differential equation describing natural history based on the use of computerized identification techniques.

The main findings derived from these models were that, with increasing numbers of tests, the marginal gains become smaller with each additional test (or unit cost). With few tests, the optimal age to start screening is around the age of 35 years, and as more tests are added to the schedule, the optimum age at start diminishes but less than the addition of years for examination at older ages. Attendance, test sensitivity and completeness of follow-up, at moderate levels of screening intensity, improve effectiveness more than increasing numbers of tests.

Van Ballegooijen et al. (2000) used the more general MISCAN simulation model to compare the impact of policies and characteristics of screening programmes across Europe on the modelled reduction in life-years lost due to cervical cancer. They did not take into account possible regional variations in, for example, natural history, underlying risk, prognosis or test sensitivity. They estimated reductions from 21% to almost 100% for different screening policies and coverage rates operative in European countries, under the most conservative assumption about round-to-round participation in screening (Table 72). They also estimated the reductions in incidence and mortality for each country in the light of their screening policy and assuming complete coverage. This work is continuing.

Using modelling techniques without simulation, Goel et al. (1998) estimated the impact of various potential improvements to screening in Canada on the incidence of cervical cancer. The results suggested that the number of

| Table 72. Percentage reduction in life-years lost according to policy and coverage |
|----------------------------------------|---------------------------------|-----------------|
|                                       | Netherlands, Finland | Belgium, France, Greece, Italy, Spain | Germany |
| Starting age                          | 30 y                | 25 y              | 20 y        |
| Interval                              | 5 y                 | 3 y               | 1 y         |
| Ending age                            | 60 y                | 64 y<sup>a</sup>  | 72 y        |
| Lifetime number of tests              | 7                   | 14                | 53          |
| Interval coverage (%)                 | % Reduction in life-years lost |
|                                      | 25                  | 21                | 24          |
|                                      | 50                  | 42                | 47          |
|                                      | 75                  | 63                | 71          |
|                                      | 100                 | 84                | 94          |

<sup>a</sup> For France, the stopping age is 65 years

Adapted from Van Ballegooijen et al. (2000)
cervical cancer cases could be reduced by 15% if Canada achieved full coverage with three-yearly screening. If, instead, smear quality were improved so that all smears were satisfactory for evaluation (versus 5% inadequacy assumed by the model), cases would decline by about half this amount.

Hakama and Hristova (1997) used age-period-cohort modelling to project mortality to 2017 across the Nordic countries under three scenarios: with no screening, with present levels of screening (and projections based on recent trends) and with improved screening as for Finland, that was termed optimal screening. They estimated that 91% of cervical cancer deaths were prevented in Finland by screening, i.e., could be prevented with optimal screening. Further, they noted that greater declines than observed could have occurred in the Nordic countries other than Finland had the Finnish (optimal) screening practices been adopted. It was predicted that screening would prevent the loss of 10 000 woman years in the Nordic countries in 2010 and that the costs of the health services were less with screening than without, assuming the organization practised in Finland.

### Issues in the implementation of screening

The Papanicolaou test was never rigorously evaluated in a randomized controlled trial such as those to which new screening techniques are subject today. Observational data have been used to demonstrate the efficacy and effectiveness of screening in controlling cervical cancer (see Chapter 4 and above) and it would now be unethical to conduct a randomized study in the presence of existing cytology-based screening. Where screening has failed to work, the blame can be laid on the design or delivery of the screening service (Zapka et al., 2003).

### Reaching the women at risk

The many organized screening programmes around the world are described in Chapter 3, but much cervical screening is also undertaken spontaneously. In the USA, where screening is opportunistic or spontaneous, about 80% of women over the age of 25 years report having had a test in the last three years (Breen et al., 2001; Swan et al., 2003). In England, where screening is organized, 81.5% of women aged 25–64 are reported as having had a test in the last five years (Statistical Bulletin, 2003). These are clearly very comparable rates. Gustafsson et al. (1995) suggested that a screening test can be equally effective whether performed in an organized or opportunistic setting. [The Working Group noted that Gustafsson et al. inappropriately considered the prevalence of carcinoma in situ and microinvasive cancer in estimating the effectiveness of screening, rather than the incidence of clinical invasive cancer.] However, while opportunistic screening may avoid the costs of the central call/recall bureaucracy, it can be considerably more expensive (Schaffer et al., 1995).

A further issue is whether screening reaches the population at high risk. In the United Kingdom, before organized call/recall was introduced in 1988, only around a quarter of women had had a recent test and these were largely lower-risk women (Farmery & Gray, 1994). In the context of a generally organized and centrally funded health service, opportunistic screening was failing a large proportion of the population and, in addition, cervical cancer rates were beginning to rise, particularly among younger women (Beral & Booth, 1986). Organization of the screening programme in England and Wales raised the coverage rates and led directly to a 42% drop in cervical cancers between 1988 and 1999, after an initial increase in the number of cases diagnosed (Quinn et al., 2001). Following a case–control study, Nieminen et al. (1999) concluded that the substantial decrease in the incidence of and mortality from cervical cancer in Finland was due mainly to the organized mass screening that had taken place rather than to any opportunistic testing and that opportunistic testing was far less efficient.

The part of a population that is hardest to reach generally includes many of the high-risk women (Davey-Smith et al., 1994), for reasons that are surprisingly similar despite the different health systems observed. These include socioeconomic deprivation, cultural and language barriers, often being from a minority ethnic group, being highly mobile in residence and not having a ‘usual care provider’ (Lawson et al., 2000). Many strategies have been employed to attract such women for screening, with varying degrees of success. It has been found that use of nurses to take smears improves acceptance, particularly among deprived women (Baker & Middleton, 2003) and non-medical smear-takers may also be used (National Cervical Screening Programme (NZ), 1998). Even when the initial test and any follow-up required is provided without charge, difficulties in reaching deprived women can persist (Chiou, 2003). Organized call/recall systems are more likely to reach these women, although the accuracy of the register is a key factor in the degree of success (Baker & Middleton, 2003).

Screening is best organized on a population basis as a public health programme. Evidence from the Netherlands illustrates the difficulties in organizing cervical screening within a general practice setting (Hermens et al., 1998). Even when professional thinking on policies is clear, external assistance is required to bridge the gap between policy and effective delivery of the programme (Hanselaar, 2002).
Most cervical cancer screening now takes place in the developed world, although 80% of the cases are found in developing countries. The scarcity of the skills and resources required to report cervical cytology in developing countries together with the difficulty of finding and treating women have led to interest in investigating alternative techniques for cervical screening in these areas, such as visual inspection of the cervix (see Chapter 4).

Identifying abnormalities
The process of performing and reporting the original test has a number of distinct phases. The first of these is obtaining the sample. When cervical screening was first implemented in a structured way in British Columbia, Canada, in 1949, the objective was to demonstrate the effectiveness of the technique first reported by Papanicolaou in the 1940s, and the majority of smears were taken by general practitioners during examinations. Most cervical screening still takes place in the primary-care setting, but the nature of the individual who actually performs the test varies from one country to another (Boyes & Worth, 1976).

The sampling device used in the original Canadian system was the Ayre’s spatula. This has remained in use to the present day, often in combination with an endocervical brush to ensure sampling of the endocervical canal. Cotton swabs have also sometimes been used [the Working Group noted that this is not an efficient sampling technique]. Extended-tip spatulae, such as the British ‘Aylesbury’ spatula, have come into use over the last 15 years and more recently plastic brooms. These are almost always used where a liquid-based specimen is to be taken. Buntinx and Brouwers (1996) conducted a meta-analysis looking for any relationship between sampling device and detection of abnormality and concluded that either the extended-tip spatula, a combination of any spatula plus the Cytobrush or cotton swab, or the plastic broom should be used for cervical screening.

The most common screening test remained the conventional Papanicolaou smear until relatively recently, when liquid-based cytology (LBC) was introduced. LBC testing is now used for the majority of cervical screening in the USA (Noller et al., 2003) and this is spreading elsewhere. The United Kingdom is now converting its entire programme (NICE, 2003), as a result of improvements in efficiency due to the dramatic drop in the number of tests reported as inadequate for diagnosis and improved laboratory productivity. LBC may increase the detection rate of cervical screening (see Chapter 4), although studies are generally based on findings at one test, rather than in a population over time, so there is a lack of long-term data (Payne et al., 2000). In the United Kingdom, the introduction of LBC was modelled to be cost-effective, as discussed below (Moss et al., 2003).

Changing to LBC facilitates a move to using automated devices to assist in reporting, although few studies have yet provided data obtained with currently available equipment.

The place of testing for high-risk HPV DNA in a cervical cancer control programme is not yet defined. Testing women for high-risk HPV as triage for borderline or ASCUS cytological results is becoming common following the publication of data from studies conducted primarily in the USA (Manos et al., 1999; Soloman et al., 2001; ANAES, 2002; Arbyn et al., 2004). Studies are also being conducted into the possibility of using HPV DNA testing as a primary screening test (see Chapters 2 and 4). Adding HPV testing to cervical cytology allows the interval to be increased for HPV-negative women with normal cytological results (van den Akker-van Marle et al., 2003).

Quality assurance is an integral part of most screening programmes, although in practice it varies from the comprehensive quality assurance programme seen in the United Kingdom to systems which cover the laboratory only, such as the Clinical Laboratory Improvement Amendments (CLIA) in the USA. European guidelines, originally produced in 1993 (Coleman et al., 1993), are currently being revised. Evaluation of screening programmes in the longer term requires monitoring of cervical cancer incidence and mor-
tality rates and comparison of data in the screened population with what might have been seen in unscreened populations (Day, 1986).

Follow-up and treatment of abnormalities
Sasieni et al. (1996) calculated that in England, 21% of the cervical cancers with inadequate screening history in 1992 were due to failure to follow up abnormalities according to the then current guidelines. Failure to investigate and treat women with cytological abnormalities and loss to follow-up after treatment are well documented pitfalls in the operation of cervical screening programmes (see Chapter 3). The majority of preinvasive cervical lesions can today be treated under colposcopic guidance and with no or only a local anaesthetic. This has considerably lessened the harm caused by cervical screening compared with the early days when radical surgical techniques were the treatment of choice (Boyce & Worth, 1976). Treatment is now extremely successful, with a complication rate of less than 2% (Luesley & Leeson, 2004). However, due to the risk of recurrent disease, women who have been treated for cervical intraepithelial neoplasia (CIN) are generally recommended to have annual cytological screening for around 5–10 years before returning to a longer cycle.

A systematic review on HPV DNA testing in the follow-up after treatment of CIN indicates that a positive HPV test can pick up treatment failure more quickly than cytology (Paraskevaidis et al., 2004).

Demonstration projects
Each population to which screening will be applied has different characteristics, priorities and health systems. In order to determine the optimal service design and delivery for a given population, a demonstration project should be undertaken (Miller et al., 2000). This should consider the feasibility of the proposed arrangements and testing of those arrangements in practice.

Hazards of screening programmes
Screening is an unusual medical intervention in that it is an "investigation, which does not arise from a patient's request for advice for a specific complaint" (McKeown, 1968). While there are excellent data supporting the implementation of mass cervical cancer screening programmes, there are also negative consequences of screening large numbers of healthy women in order to prevent significant disease in a few. These include:

- Psychological consequences of a positive screening result, with increased anxiety and fear among women;
- Misunderstanding by women and health-care providers of the meaning of a positive screening test, such that a positive result is interpreted as a ‘cancer diagnosis’;
- Misunderstanding by women and health-care providers of the meaning of a negative test as implying no risk rather than low risk for cervical cancer, which may lead to underinvestigation of symptoms;
- False positive screening results leading to unnecessary interventions, with both human and financial cost implications;
- False negative screening results giving false reassurance;
- Over-treatment of preinvasive lesions that left alone would neither progress nor cause any clinically significant disease, particularly as there are still no reliable markers to determine which high-grade cervical cancer precursors will progress to cancer or will remain clinically insignificant;
- Complications of treatment such as cervical stenosis, cervical incompetence and infertility, as well as the results of more radical therapies, such as hysterectomy, with a range of potentially negative sequelae related to the surgical intervention;
- Opportunity costs to the healthcare system of introducing a screening programme;
- Impact of incidental findings during screening.

Psychological consequences of screening
There have been few studies directed specifically at evaluation of the psychological impact of participation in a cervical cancer screening programme. In such programmes, women who are well and asymptomatic are required to undergo a gynaecological examination, which for many women is uncomfortable and experienced as a relatively invasive procedure in a private and intimate part of their bodies. After a delay of varying intervals, depending on the quality of the screening service, the woman receives her result. Approximately 1–10% of all smears are considered abnormal. In most screening services, low-grade cervical abnormalities are managed by repeated testing at defined intervals, while for high-grade abnormalities, women are referred for colposcopic evaluation, another uncomfortable and invasive procedure. Both approaches may cause significant anxiety in women (Marteau et al., 1990).

The notification of an abnormal result may cause anxiety and fear among women. For many, the concept of a ‘precancerous’ lesion is difficult to grasp and the assumption may be that they either have or are at great risk of having an established cancer (Posner & Vessey, 1988). Despite the markedly improved
outcome after treatment for cancer in the past 20 years, many women still equate a 'cancer diagnosis' with a 'death sentence' (Greer, 1985). The word precancer itself causes anxiety, because what women hear is the word 'cancer' without the qualification of the medical meaning that it is a precursor lesion that may never develop into an invasive lesion (Kavanagh & Broom, 1998).

In addition, in many colposcopy services there is considerable delay for an appointment and the waiting period may be associated with acute anxiety, particularly if the woman believes that she has cancer, even though, due to the long latent period in the natural history of cervical cancer, there may be no clinically significant consequence of this delay.

Posner and Vessey (1988) used a semi-standardized interview to study 153 women from the time they were referred to the clinic for colposcopy to after their final check-up. Of these women, 65% described feeling 'worried or alarmed' after receiving notification of their abnormal test and 27% used words such as 'shocked', 'stunned' and 'devastated'. Women's anxiety was related principally to the belief that the positive result implied cancer and death. Further, even after appropriate treatment, 35% of the women still felt afraid of the possibility of cancer and 43% worried about recurrence of disease. Women also reported having a different view of their bodies and a different attitude to sex after a positive test.

There is very little good information on whether women who attend colposcopy clinics after an abnormal test result differ in perception of cancer risk from those who fail to attend, as suggested by Posner and Vessey. Funke and Nicholson (1993) found no difference in such risk perception between women who attended for colposcopy and non-attenders. Lerman et al. (1990) found that women who had been screened in the past three years had less fear of cancer than those who had not been screened in the same period. McKee et al. (1999) found no difference between women compliant with colposcopy attendance with regard to fear of cancer compared with those who were non-compliant. They also found no difference in the perception of the gynaecological examination as embarrassing between attenders and non-attenders at colposcopy clinics.

In evaluating women's responses to an abnormal test result, factors other than fears associated with cancer and death also need to be taken into account. For instance, McKie (1993) found that women made negative links between cervical cancer and sexual promiscuity. The epidemiological findings that having multiple sexual partners or a partner with multiple partners increases the risk of cervical cancer may be interpreted by women with abnormal tests as implying that they or their partners have been promiscuous; the comment has been made that 'their character, as well as their cervix, is smeared' (McCormick, 1989).

In most screening programmes, information given about screening is aimed at achieving high uptake of screening, in keeping with the well documented benefit of wide coverage on reduction in cervical cancer. However, giving information that emphasizes only the positive aspects of screening may have negative consequences for some women who feel let down by the screening process, particularly women who receive false negative or false positive results. A belief that screening gives full protection may in itself have negative consequences and it is important that women understand that screening will not prevent all deaths related to cervical cancer. In addition, many screen-positive women will be treated for an asymptomatic condition that, if left undetected, would never have progressed to a clinically significant lesion.

Concern has been raised that giving realistic information to the public, including explaining that certain individuals can suffer adverse outcomes, can have a negative impact on uptake of screening, and brings up the issue of the 'public good' versus 'individual autonomy'. Wardle and Pope (1992), commented that attention to the psychological costs of screening had lagged far behind the technical and organizational aspects of screening services. Research into this qualitative aspect of screening has suggested a substantial toll of emotional turmoil, but most studies have been uncontrolled and involved subjective evaluation.

**Unnecessary treatment, overtreatment and adverse consequences**

The appropriateness of screening for the prevention of cervical cancer should be seen as a balance between the beneficial health effects and the adverse effects and costs of intervention. Unnecessary referrals and diagnostic and therapeutic procedures are often cited as the major adverse effect of cervical cancer screening, although few studies have attempted to quantify them. The Dutch Evaluation Commission (Evaluation Commission Cervical Cancer Screening, 1988, reported in Van Ballegooijen et al., 1990) evaluated the amount of diagnostic and treatment procedures induced by cervical cancer screening prospectively and in relation to mortality reduction using data from the Dutch screening programme. A model-based analysis led to the following estimates: (1) a mean duration of preinvasive disease of 17 years, with the shortest at older ages; (2) a regression rate of preinvasive disease of 60% on average, with the highest at young ages;
and (3) a sensitivity of cytology of around 70% for CIN 3. The false positive rate of cytology was assumed to be 0.4%. The group calculated that for five invitations for screening among women aged 37–70 years at eight-yearly intervals, 13 deaths were avoided per million women per screening year. Each death avoided is balanced by 2800 preventive tests, nine women referred for a gynaecological assessment and four for minor treatment procedures (e.g., conization of the cervix). Increasing the invitations to 25 from five would avoid 27 deaths per million women per screening year, but would require 7300 preventive tests, 22 referrals to gynaecology and eight minor treatment procedures. These data clearly showed that more intensive screening greatly increases the need for intervention with diminishing returns for the extra efforts required.

CIN lesions have been treated using a variety of ablative and excisional techniques over the past 40 years (Martin-Hirsch et al., 2004), with each method having its own range of complications and consequences (although the same efficacy). Ablative techniques rely on a histological diagnosis provided by colposcopically directed punch biopsy, which may both undercall (leading to missed diagnosis of microinvasive cancers and under-treatment of these conditions) or overcall (leading to over or unnecessary treatment of the cervix). The most widely used ablative techniques include cryotherapy and laser therapy (see Chapter 1).

In addition to the potential for unnecessary interventions, the widespread use of excisional procedures to treat preinvasive lesions of the cervix may have led to considerable overtreatment. Data on negative histological findings in tissue obtained during loop electrosurgical excision procedure (LEEP) have been reported in a number of studies using a ‘see-and-treat’ approach, i.e., treatment of CIN after a colposcopic diagnosis without prior histological confirmation. Murdoch et al. (1991) reported an overall 41% rate of negative histology after LEEP in a highly selected group of women attending a colposcopy clinic because of abnormal cytology. In women who had had prior histological sampling, negative LEEP histology was found in 43%, compared with 38% of women treated with LEEP on a see-and-treat basis. Of the women who had negative LEEP histology and who were treated on a see-and-treat basis, the majority (53%) had index cytology of CIN 1. As a consequence, the authors cautioned against the see-and-treat approach in women with low-grade referral cytology.

In a retrospective analysis of LEEP performed at a colposcopy clinic in South Africa (Denny et al., 1995), 18% of LEEPs performed after prior histological sampling (21 out of 116) yielded histologically negative findings compared with 14% of those treated on a see-and-treat basis (16 out of 114). The women were referred to this colposcopy clinic as a result of persistent LSIL (2–3 LSIL cytological results over 12–18 months) or one result of HSIL or suspicious of malignancy. An additional finding in this study was that 25% of punch biopsies were falsely negative; the authors emphasized that a punch biopsy is only as reliable as the colposcopist’s ability to identify the most abnormal area for biopsy.

Rates of negative LEEP histology ranging from 5 to 41% have been reported. All of these studies were performed in colposcopy clinics where women had been referred with abnormal cytology (Prendiville et al., 1989; Luesley et al., 1990; Whiteley & Olah, 1990; Bigrigg et al., 1991; Hallam et al., 1993; Denny et al., 1995). False positive cytology or colposcopy, false negative histology in the LEEP specimen and possible complete excision or spontaneous resolution of the lesion after prior biopsy are possible explanations for negative LEEP histology. In addition, some series have reported negative LEEP histology where there has been extensive thermocoagulation preventing a histological diagnosis.

While excisional procedures of the cervix are performed under local anaesthetic in an outpatient setting and complications are considered relatively benign, this is not always the case. In a randomized trial of LEEP, cryotherapy and laser vaporization of the cervix, Mitchell et al. (1998) reported complications from LEEP in 7.6% (10/130) of women treated. The majority of the complications related to bleeding (70%); there was one case of infection, one woman complained of severe pain and required treatment and one woman subsequently developed cervical stenosis.

Ferenczy et al. (1996) reported on 1070 women who underwent LEEP and who returned for post-LEEP follow-up. Complications were recorded in 71 women (7%); 37 had significant intra- or post-operative bleeding, and one required admission to hospital to control bleeding. Seven women developed a purulent vaginal discharge and pelvic pain within one week of treatment. A further 13 women (1.2%) developed cervical stenosis, of whom 11 were over the age of 45 and not on hormone replacement therapy.

Cervical stenosis, while rather rarely reported, is a significant complication of treatment of the cervix and may result in infertility or menstrual complications such as haematometria, making follow-up with cytology and/or colposcopy difficult if not impossible. These complications may necessitate hysterectomy, with all the potential sequelae associated with this more radical surgical intervention. Hysterec-
tomy may also be the result of clinical uncertainty as to the meaning of persistently low-grade cytology or persistently inadequate tests, neither of which can be resolved satisfactorily.

In some women, if the excisional procedure removes large amounts of cervical stroma, a complication of treatment may be cervical incompetence. This has been associated with pre-term delivery due to premature rupture of membranes (Sadler et al., 2004).

In addition to the complications associated with treatment, the risk of persistence or recurrence of lesions after treatment makes long-term follow-up of treated women essential. Reports from both non-randomized and randomized trials of treatment for cervical cancer precursors using a variety of treatment modalities indicate that most treatments have about a 90% success rate (Martin-Hirsch et al., 2004).

Opportunity costs to the health system
Setting up a screening programme designed to detect disease in healthy individuals necessitates diversion of human and financial resources from other health interventions, in particular, treatment of already existing or apparent disease. Thus a screening programme and its hazards and benefits need to be evaluated in the context of the competing health needs of the specific country to ensure that resources are used to the maximum benefit of the entire population.

One feature of cervical cancer screening programmes in low-resource countries with endemic HIV infection is the diagnosis of CIN lesions in as many as 20–30% of HIV-infected women. This may lead to the consumption of scarce health resources for treatment of CIN. However, in countries such as Zimbabwe where HIV infection has become pandemic, HIV-infected women succumb to opportunistic infections long before invasive cervical cancer arises (Chokunonga et al., 1999). In situations where it is not possible to provide any form of treatment for HIV-related conditions (e.g., treatment of opportunistic infections and/or antiretroviral therapy), cervical cancer screening may not be a priority.

Incidental findings
While screening is an activity designed for healthy, asymptomatic women, there can be unexpected incidental findings at the time of screening that may cause harm as well as benefit to women and the health system. For instance, the discovery of underlying diabetes due to the diagnosis of diabetic vulvitis at the time of performing the screening test may be of benefit to the woman, but the discovery of a lethal co-existent cancer may only increase suffering without offering any improvement in quality of life, if there is no effective treatment for that cancer.

In low-resource countries where women generally have little access to health care, screening may identify a significant number of women with co-morbid health conditions which it is impossible to manage with the available resources.

One possible incidental finding of particular significance is identification of a woman as HIV-positive. The prevalence of CIN among women infected with HIV is nearly five times that in HIV-negative controls (Chirenje et al., 2002). In low-resource countries where cervical cancer screening is mainly opportunistic, there is a large burden of women harbouring CIN lesions with concomitant HIV infection that has not been identified.

Cervical cancer screening in HIV-positive women
It is estimated that there are now up to 42 million people worldwide living with HIV infection or AIDS. About 70% of these individuals live in sub-Saharan Africa (UNAIDS, 2003) and more than half of the infected people are women; cervical cancer screening has been widely unavailable in this area up to now.

In many HIV-endemic countries, CIN lesions may be detected at the same time that HIV infection is first diagnosed. The discovery of CIN lesions in an HIV-infected woman may create major psychological and morale challenges, not only to the woman, but also to health workers.

The diagnosis of HIV infection still carries a high level of stigmatization and fear of disclosure of an incurable sexually transmissible disease, with some women being exposed to violent response from their male partners and permanent psychological isolation from the community. In the absence of antiretroviral treatment, some women become suicidal and in this group an added diagnosis of a CIN lesion through cervical cancer screening causes special concern. Linkage of a cervical cancer screening programme to HIV testing may present a new barrier to participation in cervical screening.

Another concern is how to interpret a positive cervical screening test result in the presence of HIV-positivity, bearing in mind that natural progression of both conditions is dependent on availability of effective treatment. Treatment of CIN lesions in HIV-positive women by ablative or excision procedures results in epithelial disruption which can theoretically enhance viral acquisition or transmission.

Health workers treating HIV-positive women for cervical disease should take universal precautions as for all health interventions.

Performance evaluation
The minimal essential elements of a cervical screening programme are: a defined population to screen; invitation to this population to participate;
assessment of coverage; a quality control system; and treatment for test-positive women. The means of achieving these, e.g., population registers or geographical location to define the population; personal invitation letters (call/recall) or mass education to invite women to participate; population-linked cervical cancer registry or sample surveys to assess coverage) will depend on the local circumstances. The recognition that the effectiveness of screening "is determined by the proportion of progressive lesions that are successfully detected and treated" (Pontén et al., 1995), which depends in turn on screening policy and its implementation, has led to the development of indicators of screening programme performance for routine monitoring. The determinants of this proportion represent the essential elements of a good screening programme (Hakama et al., 1985). They include coverage of the target population (participation); attendance for rescreening; adequacy of smear-taking; quality of interpretation of smears; and follow-up of abnormal results. Specific indicators for these programme performance areas have been developed in the context of cytology and more specifically Pap smears.

The Council of Europe recently recommended all Member States to offer organized screening for three cancers including cervical cancer, and stressed that this should be managed in such a way that the performance can be evaluated fully (Council of the European Union, 2003). Organized screening requires adequate data collection systems to be set up concerning invitation and participation of the target population, registration of screen test results and follow-up of screen positives (Arbyn et al., 1999; Advisory Committee on Cancer Prevention, 2000). Screening databases, including personal records, should be linkable with cancer and mortality registers, in order to allow full evaluation of the programme. This needs to be done with full respect for national legislation. Opportunistic screening systems are in general less cost-effective and do not allow evaluation (Advisory Committee on Cancer Prevention, 2000).

**Screening policy**

Unlike breast cancer screening, where optimal screening policy (in terms of age range, frequency and modality) has been determined by randomized trials and most programmes follow similar policies, results of observational studies of protection offered by cytological screening have produced estimates of effectiveness for a variety of alternative policies. Thus, the maximal effectiveness of a programme will depend on what policy is adopted. The impact of policy on effectiveness has been modelled recently for European countries by van Ballegooijen et al. (2000) (see Table 72).

Most screening programmes recommend the same screening interval for the entire target age range. However, a recent audit of screening in the United Kingdom suggested that the policy of one screening interval across the entire target age range may not yield optimal effectiveness and recommended screening women aged 25–49 more frequently than older women (Sasieni et al., 2003). This further illustrates how effectiveness can be determined in part by policy.

Many factors influence what policy is adopted. For example, a national workshop in Canada recommended that screening intervals not be raised to three years without the security provided by an information system (Miller et al., 1991). Since most Canadian jurisdictions have not in the past had information systems, this almost certainly resulted in over-screening. Medical legal issues may also influence screening policy (and therefore effectiveness) in some countries. For example, a shorter screening interval is safer than a longer one and a broader age range safer than a narrower one. However, the marginal utility of each extra test diminishes rapidly. This is of particular concern in developing countries.

**Screening delivery**

It is generally accepted that, for optimal effectiveness, cervical cancer screening should be offered within an organized programme (see Chapter 3). The programme should further include the specific elements as described by Hakama et al. (1985) (see above). Hakama and others have concluded that organized screening programmes, as practised to varying degrees in the Nordic countries and particularly in Finland, are more effective than opportunistic screening activities. This conclusion was based largely on comparison of incidence and mortality trends (e.g., Hakama, 1982; Läärä et al., 1987) and cohort studies, as discussed previously in this volume. A case–control study by Nieminen et al. (1999) suggested that tests in an organized programme may be more effective than those delivered outside the programme, within the same jurisdiction. However, since countries differ greatly in their screening practices as well as in the level of organization, it is difficult to attribute higher effectiveness definitively to a more organized programme.

The premise that a totally organized programme is significantly more effective than a programme based on largely opportunistic screening with some central coordination and elements of organized screening, such as is practised in many parts of the developed world, has been challenged (Madlensky et al., 2003). Finland may be one of the very few countries in the world whose screening programme...
meets all the criteria outlined by Hakama et al. (1985). Although it may not be feasible to adopt all these elements of organized screening, many jurisdictions have incorporated some of them. For example, recruitment strategies may be targeted to specific population groups in the absence of a population register that allows personal invitations to screening and this can lead to achievement of low rates of cervical cancer. For example, incidence rates for cervical cancer (uncorrected for hysterectomy prevalence) and trends over the past 30 years are almost identical in different provinces of Canada: British Columbia, which has had centralized cytological screening and a cytology information system for several decades (see, for example, Morrison et al., 1996); Ontario, which has a recent information system including about 80% of the province’s screening tests and a programme that sets policy and standards but where smears are taken and read in a totally decentralized system; and Quebec which has no organized programme and no information system but opportunistic screening (see Figure 67). However, unless historical patterns of screening and cancer incidence rates are taken into account, inferences from such data regarding effectiveness are uncertain (see Chapter 4).

On the other hand, poor organization of a screening programme can produce poor outcome. In the United Kingdom during the 1970s and 1980s, population coverage was low; low-risk groups were over-screened and the technical quality of screening process parameters was moderate. Implementation of call/recall systems and targeted rewards for primary care providers, achieving high levels of coverage among eligible women (Patnick, 2000), resulted in a substantial decline in incidence and mortality in all age groups of the target population (Sasieni et al., 1995; Quinn et al., 1999; Sasieni & Adams, 1999, 2000).

There are, however, several common problems with opportunistic screening versus organized screening, and opportunistic screening should therefore be discouraged:

1. It is less cost-effective (see below)
2. Hard-to-reach women are less likely to be adequately screened
3. In many settings, especially in developing countries, there is disproportionate representation of women who are in contact with the health-care system for other health interventions such as reproductive care, so that those in some age groups are inadequately screened (Were & Buziba, 2001)
4. It can create sporadic work flow, which can lead to reduced proficiency, etc.
5. It can result in greater chances of harm due to over-screening.
6. It may be difficult to ensure quality.
7. Screen-positive women may not have easy access to diagnostic and treatment services.

Performance indicators
An integrated information system or a set of systems that can be linked as required is recommended as the ideal support for performance monitoring; such a system can also support programme operation (Miller, 1992). These systems should permit identification of each woman as well as each test and link them. A model for a comprehensive information system is shown in Figure 68.

For performance monitoring, the system should ideally contain a screening database including results of cytology and follow-up (colposcopy, histopathology, treatment) with periodic linkage to a population register, tumour registry, mortality file and hysterectomy data. However, even in areas where population registers and/or other files do not exist or are not accessible, information systems can be developed that permit estimation of many indicators. Others can be estimated periodically by special studies. A number of indicators are based on negative test results. This generally assumes exclusion of negative results for women who are under special surveillance (e.g., following colposcopy, previous positive history, etc.).

Table 73 outlines performance indicators coinciding with the determinants identified by Pontén et al. (1995) and Hakama et al. (1985), along with the programme area they are designed to specifically evaluate and required data where relevant. These have been adapted from Coleman et al. (1993), who proposed a menu of specific indicators with targets for the Europe Against Cancer Programme.

All indicators should be evaluated, reviewed and published annually. Some of the indicators require multiple sources of information, sometimes linked at the individual level. Where this is not possible, alternative methods, such as periodic special studies (e.g., the health and fertility surveys noted in Chapter 3), should be used.

Participation (or coverage)
The participation rate is the proportion of eligible women in the target population who participate in screening within the time interval specified by local screening policy. It can alternatively be defined in terms of some chosen period of time (e.g., tested in the last three years). Its estimation requires, at a minimum, counts of screened women in the target age range and of the target population. If estimates of the prevalence of hysterectomy are available, they should be used to reduce population counts to reflect more accurately the population at risk. Women with only inadequate smears in the interval should not be counted as having been screened in that
interval. Participation should be evaluated according to age group, geography and other locally relevant indicators (e.g., health-care provider, ethnicity).

The effect of participation in reducing mortality and incidence has been demonstrated descriptively. For example, in the United Kingdom, Quinn et al. (1999) showed that incidence declined dramatically starting in the late 1980s after the introduction of a call/recall programme which resulted in greatly increased coverage. Miller et al. (2000) stated that “… the programme must focus on achieving the highest possible coverage rate. To support this, indicators such as number of women...

Figure 67 Cervical cancer incidence in three large provinces of Canada, with different intensity of programmatic components, 1970–99

(——) Quebec, (—) British Columbia, (—) Ontario

Table 73. Short-term performance indicators for assessing programme effectiveness and efficiency

<table>
<thead>
<tr>
<th>Programme component</th>
<th>Measure</th>
<th>Definition</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance for screening</td>
<td>Participation (or compliance) rate&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Percentage of women in the target population with at least one test within the recommended interval</td>
<td>Adjust denominator for prevalence of hysterectomy. Estimate for age, region and other risk indicators</td>
</tr>
<tr>
<td>Adequacy of referral and treatment systems</td>
<td>Compliance with recommendations for follow-up of unsatisfactory smears and positive test results</td>
<td>Percentage of women (or tests) with specific positive (or unsatisfactory) results with follow-up action according to recommendations</td>
<td>‘Recommendations’ are those set by the programme and include follow-up action (e.g., repeat test, colposcopy) and time to follow-up. Report according to reason for follow up (e.g., inadequate, ASCUS, etc.) and person/institution responsible for ensuring follow-up. Reasons for non-compliance should be noted.</td>
</tr>
<tr>
<td>Compliance with treatment recommendations</td>
<td>Compliance with treatment recommendations</td>
<td>Percentage of women requiring treatment who receive it according to recommendations</td>
<td>‘Recommendations’ include type of treatment and time to treatment. Report and investigate as for follow-up compliance.</td>
</tr>
<tr>
<td>Overall stage of invasive cancers</td>
<td>Stage of invasive cancers</td>
<td>Percentage distribution of stage at diagnosis for all invasive cervical cancers in region</td>
<td>Requires population-based cancer registry.</td>
</tr>
<tr>
<td>Interval cancers</td>
<td>Number of invasive cancers diagnosed following negative test result and before next expected screen per 100 000 person years at risk&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Person years calculated from date of test to date of diagnosis, date of next expected screen or date removed from population at risk. Reasons for individual cases should be investigated as a kind of audit, along with other cancers that are not interval cancers</td>
<td></td>
</tr>
<tr>
<td>Efficiency</td>
<td>Over-screening</td>
<td>Percentage of women with negative test who are screened again before end of screening interval&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Also generally referred to as ‘coverage’. However, this is an ambiguous term as it can also refer to the percentage of population targeted for screening.

<sup>b</sup> May also be expressed per 100 000 negatively screened women
Modelling has also shown that participation rate is the most important programmatic determinant of effectiveness. Van Ballegooijen et al. (2000) modelled the expected reduction in life-years lost for a number of European countries as a function of screening policy and coverage. They found that differences in coverage resulted in more or less proportional differences in effectiveness, essentially independent of screening policy (see Table 72).

High rates of participation are, however, not sufficient to ensure high effectiveness, if other essential elements of a screening programme are suboptimal. This is evident from the situation in some Latin American countries (see Chapter 3 and this chapter).

In populations where the prevalence of prior hysterectomy is substantial, participation rates will be underestimated, especially in women aged over 50 years, if adjustment is not made. Because rates of hysterectomy vary across time and place, lack of

---

**Figure 68** Model for a comprehensive cervical screening information and reporting system

1 Periodic linkages would uptake addresses, identify unscreened women, ascertain ‘failures’ (cancers) and remove women who no longer need screening

From Marrett et al. (2002)
such adjustment may invalidate comparisons of participation rates over time and between populations.

It is most appropriate to estimate participation based on all screening tests, whether from an organized programme or opportunistic.

**Quality indicators**

For programmes to be effective, all parts should be quality assured, with indicators for each part of the programme. Furthermore, the programme needs to have access to such quality assurance information and to ensure that it is fed back, along with standards or comparisons, to those providing the service (e.g., primary care providers, laboratories, etc.). Quality indicators should cover test-taking (e.g., inadequacy rates, interpretation (e.g., positive predictive value), treatment and follow-up and programme organization. Acceptable ranges for any such indicators should be specified and values that are outside the range should be investigated. Many of these are discussed in more detail in Chapter 3. Programmes should regularly audit cases of cervical cancer to identify possible shortcomings of the programme.

Modelling of the impact of quality improvements suggests that they have less impact than improvements in coverage. However, in a programme with high coverage, improvements in quality can increase effectiveness. Goel et al. (1998) estimated that the impact of reducing the false negative rate from 0.25 to 0.10 (i.e., increasing sensitivity from 75% to 90%), while leaving screening otherwise unchanged, might result in 25% fewer incident cases. As noted by Fahey et al. (1995), however, even the lower of these sensitivity values may be higher than is generally achieved. Modelling of effectiveness of screening programmes must use realistic estimates of test quality if it is to provide valid estimates of effectiveness.

**Follow-up**

Preinvasive (or early invasive) lesions identified via screening must be treated if development of invasive (lethal) disease is to be avoided. Thus, referral for and presentation at follow-up of a positive test result as well as post-treatment follow-up of confirmed precancerous lesions, in accordance with treatment policy, are important. It is also essential that test results be provided to the health-care provider and woman, including specification of the need for follow-up, in a timely fashion.

**Follow-up consistent with recommendations**

A screening programme should have a policy for follow-up of unsatisfactory tests and positive test results. Policy should specify the action required and the time frame within which this action should take place for specific test results and patient history. Actual follow-up should then be monitored in relation to programme policy. An unsatisfactory test means that a woman has not been adequately screened and the appropriate follow-up recommendation in this situation is for a repeat test. Compliance with this recommendation should also be monitored.

The proportions of women who are followed up according to programme policy (in terms of both action and time frame) should be calculated according to reason for follow-up. Reasons for follow-up failure should be documented.

Computation of indicators of compliance with follow-up for positive test results requires linkage between these tests and follow-up data. This often includes information on colposcopy visits.

**Treatment consistent with recommendations**

Screening programmes should also have policies regarding treatment of confirmed abnormalities, again specifying both action and time frame.

As with abnormal or inadequate screening test results, the proportions of women receiving adequate treatment should be calculated according to reason for treatment. Reasons for non-compliance should be documented.

Goel et al. (1998) estimated that improving efficacy of follow-up and treatment following a positive test result from 0.8 to 0.9 might reduce the number of cancer cases by 2%. The overall impact is relatively small because this improvement affects only women with positive results, which represent a very small fraction of all tests. Pontén et al. (1995) estimated that in a programme with fewer lifetime screenings than in the Canadian context, about half of the ultimate reduction in mortality might result from detection and treatment of early invasive disease and half from removal of screen-detected precursors. In such situations, an improvement in the efficacy of follow-up might have greater impact on mortality.

**Overall short-term indicators**

Because cervical screening can detect asymptomatic invasive disease, an effective programme would be expected to lead to a shift towards more microinvasive and early invasive disease. Thus the major changes in incidence that occurred in Iceland following introduction of screening were among advanced cancers (Johannesson et al., 1982). Evidence for the effectiveness of screening due to early detection of disease can also be obtained by examining trends in survival (see, for example, Adami et al., 1994). However, observed increases in survival can be due to improvements in treatment. In general, a trend in survival is not a good indicator of an effect of screening (see Chapter 1).
**Stage distribution of incident cervical cancers**

The distribution of all newly diagnosed cervical cancers in the programme area according to stage at diagnosis should be calculated annually. This requires a population-based cancer registry that includes standardized staging information on all or at least a high proportion of newly diagnosed cases. If such data are not routinely available, periodic special studies can be conducted. This indicator is particularly important in areas where screening is not yet well established and screening intervals are relatively long. In such situations, a relatively large part of the effect of screening on mortality will be due to earlier stage at diagnosis.

**Interval cancers**

Interval cancers are those that arise following a negative test and before the next scheduled screen. These can arise for two reasons: either the previous result was a false negative or the (pre)cancer was not detectable at the time of the previous screen (for example, because it was fast-growing or did not go through detectable preclinical stages). A test repeated every three years on women aged 35–64 years has been estimated to ‘prevent’ 84–91% of invasive cancers (Day, 1989; Sasieni *et al.*, 2003) if all abnormalities are effectively treated. In the United Kingdom, testing at three-year intervals was estimated to be capable of preventing about 70% of all cancers occurring in women aged 40–69 years, after allowing for cancers that develop in women with positive test results (Sasieni *et al.*, 2003).

The interval cancer rate is calculated as the number of interval cancers per 100 000 person years at risk. Person-years at risk are estimated by summing time from the date of the last negative test to the end of the recommended screening interval or to diagnosis of cancer (or until a woman becomes ineligible due to emigration, death, etc.) for all women having a negative test. Use of this rate to assess programme effectiveness is difficult in regions where screening is well established, because the expected rate of cervical cancer in the absence of screening is unknown. However, the interval cancer rate can be followed over time and compared across programmes with similar screening policies. It may be closely paralleled by the ratio of the number of cancers diagnosed in screen-negative women during the screening interval divided by the number of screen-negative women, expressed per 100 000 women.

Calculation of the interval cancer rate requires knowledge of cancers occurring in negatively screened women. In general, this requires linkage between a population-based cancer registry and the screening test data. In many regions, no cancer registry exists or linkage between screening data and the cancer registry, at least on a routine basis, is not permitted. In such cases, audit studies of interval cancers should be performed.

Screening histories (preceding test and other follow-up results) of all invasive cancers, whether they are true interval cancers or not, should be examined routinely to identify areas where programme improvements may be required.

**Indicators of efficiency**

There are a variety of parameters that indicate the efficiency of a screening programme. The most important of these relates to over-screening. This can be assessed by the proportion of women with a negative result having a subsequent test before the end of the screening interval. Retention of screened women for rescreening can be assessed by the proportion of screen-negative women returning for rescreen at about the right time. Both of these can be examined in relation to length of interval between tests, region, smear taker, etc. The average number of tests per woman during the recommended screening interval is another indicator of the extent of over-screening.

**Process quality indicators**

A number of process indicators should be monitored to ensure that screening is operating as it should. Those that are chosen will depend on the issues that are important in a particular setting. These might include elapsed times (e.g., between test-taking and reporting; between reporting of a positive result and follow-up colposcopy), results of laboratory proficiency testing, etc. Acceptable ranges for any such indicators should be specified and values that are outside the range should be investigated.

**Economic evaluation and cost-effectiveness of cervical cancer screening**

The basic principle of a decision analytic approach is that all consequences of decisions (e.g., individual clinical outcomes, population-based outcomes and costs) should be identified, measured and valued. When a decision analysis formally compares the relationship between the health and economic consequences associated with different public health care interventions, it is considered a cost-effectiveness analysis. The application of economics to public policy does not necessarily mean that less money should be spent, but rather that the use of resources might be more efficient.

Different types of economic evaluation are commonly confused. For example, there are distinct differences between cost-minimization analysis (how much money can be saved?) and
cost-effectiveness analysis (how much health improvement can be gained, per unit expenditure?). The results of a cost-effectiveness analysis are summarized using an incremental cost-effectiveness ratio. In this ratio, all health outcomes associated with a particular strategy (compared with an alternative) are included in the denominator, and all costs or changes in resource use with a particular strategy (compared with an alternative) are included in the numerator. This type of analysis defines the 'opportunity cost' of choosing one clinical or public health approach over another.

Advances in the various technologies and preventive modalities for cervical cancer screening mean that policymakers in national and international agencies are confronted with various strategies from which to choose. However, there are important differences between developed and developing countries in the policy questions that are most relevant to cervical cancer control. Scarce resources, limited infrastructure and competing health priorities have prevented most low-resource countries from implementing successful cervical cancer screening programmes.

In countries classified as low-income economies (gross national income per capita equal to or less than US $755 in 2000), the key problem is how to implement a sustainable screening programme in the setting of competing health priorities and limited resources. If a cytology-based screening programme is to be introduced, cost-effectiveness modelling using locally derived information about costs and the age–incidence curve for cancer in the population in question, together with internationally accepted data on efficacy, will assist in deciding on the number of screening rounds and the age group to be targeted.

For a developed country where a conventional cytology screening programme already exists, information on effectiveness may be available, but whether it is cost-effective may not have been fully established. However, it is possible to assess the likely cost-effectiveness of a new technology in detecting precursor lesions of cervical cancer relative to conventional cytology. For example, England recently used cost-effectiveness modelling as a major consideration in deciding to move to using liquid-based cytology in its programme. This modelling began with the assumption that the effectiveness of new and conventional cytology was the same (Payne et al., 2000). The evaluation of the new technology was then based on costs derived locally from pilot implementation of liquid-based cytology (Moss et al., 2003). The Payne et al. (2000) conclusion of equivalence of effectiveness between conventional and liquid-based cytology was based on a surrogate measure of effectiveness, the identification of cervical abnormalities, rather than long-term follow-up of outcome, that is reductions in incidence and mortality from cervical cancer.

Data sources
Cost-effectiveness measures require data on natural history of cervical cancer, the overall effectiveness of the policy or intervention, survival rates associated with cancer, test characteristics, and quality of life.

Data sources could include: randomized trials, observational studies, meta-analyses; other published literature, expert opinion and health systems statistics. However, as implied above, surrogate measures of efficacy may have to be used to support assumptions relating to the likely effectiveness of new technology. To the extent that these assumptions are uncertain, the result of the modelling will also be uncertain, even if sensitivity analyses are performed to attempt to encompass the extent of uncertainty.

Sources of cost data could include the costs of:
- Training of staff
- The screening test
- Administration of the screening test
- Laboratory procedures
- Reporting and referral of women with abnormalities
- Diagnostic tests
- Treatment of precursors
- Treatment of clinically invasive cancer
- Patient time for all aspects of screening
- Transportation of specimens
- Programme organization

These data must be collected locally or estimated according to local conditions. The eventual judgement as to whether a particular strategy is cost-effective or not will depend on local health circumstances.

Cost-effectiveness studies
All models consistently support the messages that organized screening is more cost-effective than opportunistic screening and that the most pressing question in all settings is how to reach the highest proportion of women at greatest risk for cervical cancer. Increasing coverage is always more cost-effective than using resources in any other area of the programme. In the United Kingdom, altering the payment system in 1990 as an incentive to primary-care physicians to maximize coverage rather than as a fee for services greatly facilitated the needed increase in coverage, which rose from around 40% of women in 1989 to over 80% of women by 1993 (NHS 2003a, b; Patnick, 2000).

Goldie et al. (2001) evaluated the cost-effectiveness of visual inspection, cytology and HPV testing, largely using surrogate measures of efficacy, within a developing country situation. A very important determinant of cost-effectiveness in this setting was the
requirement for three visits for women with abnormal cytological results, whereas for screening by visual inspection a one-visit strategy was modelled, and for HPV testing two visits. A substantial loss when women are required to return after the initial screening visit is observed in many developing countries, and this has an important effect on the cost-effectiveness of cytology, and to a lesser extent of HPV testing. When the authors modelled the effect of a limited number of tests in a lifetime, three tests at five-year intervals from the age of 35 years proved to be more cost-effective than a 10-year schedule commencing at the same age. It was concluded that cervical cancer screening strategies that incorporate DVI or HPV DNA testing and eliminate colposcopy may offer attractive alternatives to cytology-based screening programs in low-resource settings.

Goldie et al. (2004) reported the results of an analysis comparing the cost-effectiveness of HPV testing with that of conventional cytology in women aged 30 years or more. This was set in the US context of annual conventional cytological testing, which was compared with three-year screening using liquid-based cytology and three-year screening using HPV testing. [Although the latter strategy proved to be more cost-effective in the analysis, it is unclear what the results would have been if three-yearly conventional cytology had been incorporated in the analysis.] Goldie et al. (2004) considered that for women aged 30 years and more, a strategy of screening every two or three years with either HPV DNA testing in combination with cytology for primary screening or cytology with reflex HPV DNA testing for equivocal results will provide a greater reduction in cancer and be less costly than annual conventional cytology.
Chapter 6
Summary of data

Cervical cancer and screening

Incidence and mortality worldwide
The majority of cervical cancer cases today occur in the developing world. However, before the introduction of screening, the rates of cervical cancer in most of Europe, North America and Japan were very similar to those now seen in developing countries.

The reported incidence and mortality rates for different populations have different degrees of reliability. Even when cases of and deaths from cervical cancer are reported, they may be reported or recorded as ‘uterus not otherwise specified (NOS)’ rather than as cancer in the uterine cervix specifically. The rates of death from and cases of cancer of the ‘uterus NOS’ must be borne in mind when considering incidence and mortality rates of cervical cancer. A further influence on incidence and mortality rates in a population is the hysterectomy rate, since this affects the denominator used in the calculations.

Cancer of the cervix uteri is the second most common cancer among women worldwide, with an estimated 471,000 new cases and 233,000 deaths in the year 2000. Almost 80% of the cases occur in developing countries, where, in many regions, it is the most common cancer among women, responsible for about 15% of all new cancers. The highest incidence rates are observed in Latin America and the Caribbean, sub-Saharan Africa, and south and south-east Asia. Cervical cancer is less common in economically developed countries, where in the year 2000, it was estimated to comprise about 4% of cancers in women, ranking sixth in importance.

The demographic determinants of risk include age, marital status, socioeconomic status and ethnic and religious groupings.

Survival has not been shown to vary between populations when the data are corrected for clinical stage at presentation, provided that adequate and equivalent treatments are available, and that co-morbidities (e.g., HIV status) are taken into account. After major improvements in survival of cervical cancer patients in the first half of the twentieth century, there has been little additional progress in recent years. Indeed, it is an apparent paradox that, often, when screening has become established in a population, no improvement in survival is seen. This is chiefly because there usually remains a proportion of late-stage tumours which are diagnosed in women who were inadequately screened or not screened at all. These become important when calculating survival rates if the majority of the population is well screened and may avoid developing cervical cancer altogether.

Pathology of cervical neoplasia

Intraepithelial squamous lesions
HPV infection of cervical squamous epithelium leads to two categories of intraepithelial squamous lesions: productive, self-limited HPV infections, and those with potential to progress to invasive squamous-cell carcinoma. Biopsies of productive HPV infections of the cervix have been classified variously as koilocytic atypia, koilocytosis, condyloma, mild dysplasia, cervical intraepithelial neoplasia grade 1 (CIN 1), and low-grade squamous intraepithelial lesion (LGSIL or LSIL). Lesions more likely to represent cervical cancer precursors have been classified as moderate dysplasia, severe dysplasia, CIN 2, CIN 3, including carcinoma in situ, and high-grade squamous intraepithelial lesion (HGSIL or HSIL). Many pathologists report histopathological diagnoses using more than one classification scheme.

Intraepithelial glandular lesions
Adenocarcinoma in situ (AIS) is the only well characterized intraepithelial glandular lesion of the uterine cervix. These lesions are less common than their squamous counterparts and are associated with persistent infection by high-risk types of HPV. The utility of diagnostic terms for intraepithelial glandular lesions with lower degrees of atypia than AIS, including endocervical dysplasia, cervical intraepithelial glandular neoplasia, and endocervical glandular atypia, has not been established.
Invasive squamous and glandular lesions

Diagnosis and treatment
Pre-invasive cervical lesions
By diagnosing and treating pre-invasive (pre-cancerous) lesions, the rate of invasive cancer can be reduced. Women with the minor intraepithelial abnormality of CIN 1 may be managed conservatively with cytology augmented by HPV testing or repeated cytology according to established protocols. Those with a cytological diagnosis of atypical squamous cells, an equivocal epithelial abnormality, have a small increase in risk of underlying high-grade disease which may be detected by using HPV DNA testing, or repeated cytology and referral to colposcopy. Women without CIN 2 or 3 may resume routine screening according to protocols.

Treatment of CIN 2 or 3 lesions may involve destructive or excisional techniques. Meta-analysis indicates little difference in success or morbidity in relation to the different techniques, which include cryotherapy, loop excision, and laser vaporization. The loop electrosurgical technique (LEEP) or LLETZ for transformation zone removal has become the standard way of treating CIN in developed countries; cryotherapy has been shown to be effective and safe in developing countries. The ‘see-and-treat’ method, where the excision is performed at the first diagnostic colposcopy visit, depends on expert assessment of the atypical transformation zone, usually by colposcopy in developed countries in the presence of a report showing high-grade cytological abnormality, but by the trained examiner in developing countries. Treatment failures (overall rates of 3–7%) are most likely to occur in women aged over 50 years with involved resection margins in the excised histological specimen.

Women with immunosuppressive states, for example due to HIV infection, are at increased risk of having high-grade CIN and need intensive follow-up after treatment in view of a higher recurrence rate.

After a check-up at six months with colposcopy and cytology, follow-up can be based on cytology alone at yearly intervals up to 10 years, with colposcopy in case of any cytological abnormality to judge the need to take biopsies or for re-treatment. Adding HPV testing gives more sensitivity and is quicker than follow-up cytology in detecting recurrent disease.

Cervical cancer
The rates of cure of cervical cancer depend on the stage at which diagnosis occurs. Diagnosis is by examination of a woman with suspicious symptoms or uncommonly as a result of abnormal cytology. The earliest pre-clinical stages of invasive disease (stage IA or micro-invasion) can be managed conservatively by excisional procedures with nodal sampling. At a more advanced stage (IA2), radical excision of the cervix and associated para-cervical tissue (radical trachelectomy) in women with fertility aspirations can be applied, but when fertility is not important, hysterectomy with associated lymph node sampling is recommended.

Staging of clinically invasive cancer involves a multi-disciplinary approach including radiological examination to assess the extent of invasion within the cervix and its surroundings. Early clinical invasive disease (stage IB1) may be considered for radical cervical removal as described for stage IA2. The usual procedure for a tumour of diameter above 4 cm (stage IB2) is radical hysterectomy combined with radiotherapy with or without simultaneous chemotherapy. Nodal sampling is mandatory.

For more advanced disease (IIB–B) with cervical and parametrial-vaginal extension, radiotherapy with concurrent chemotherapy with cisplatin derivatives is indicated; this regime has been associated with improved survival.

Stage IV cancer with recurrent or refractory disease with associated high mortality rates represents a challenge to radical pelvic surgery in limited cases but chemo-radiation is the usual treatment modality.

Palliative care is a basic need in any cervical cancer screening programme. No matter what the availability and accessibility of cancer treatment, palliative care services and medications should be provided according to standards recommended by WHO.

Etiology
Cervical cancer is an uncommon outcome of a common sexually transmitted infection. The causal association is restricted to certain genotypes of the human papillomavirus (HPV) family, denoted as high-risk types. Infections of the epithelium with HPV are usually of transient nature and may lead to the generation of an immune response, of which cellular immunity seems to be the most important for regression. Any event inhibiting normal differentiation of the epithelium or prevention of the normal sequence of viral replication may lead to the development of persistent infections, which can remain
clinically latent or become active due to a compromised immune status or other factors.

Infection of the cervix with HPV occurs during sexual intercourse with an HPV-infected male. Other forms of HPV transmission are of little relevance to genital tract infections.

The age at first exposure to HPV and the age-specific HPV DNA prevalence are strongly related to the patterns of sexual behaviour and are therefore population-specific. The risk of HPV infection and the risk of cervical cancer in a woman is directly related to the number of lifetime additional sexual partners of her sexual partner and to the number of sexual contacts with prostitutes. Male circumcision offers some protection from both HPV infection and cervical cancer in the spouse.

The development of a long-term persistent infection is required for progression towards cervical cancer. Factors possibly affecting persistence include HLA class I antigens, HLA class II haplotypes, polymorphisms in certain human genes (such as p53), partly in combination with viral variants, loss of heterozygosity and epigenetic events leading to the loss of cellular protein expression. Events compromising the immune system increase the frequency of persistent infections and consequently the risk for malignant progression.

The association of high-risk HPV types and cervical cancer is causal in nature and, under optimal testing conditions, HPV DNA can be identified in all specimens of invasive cervical cancer. The association is consistent worldwide and includes the squamous-cell carcinomas, the adenocarcinomas and the vast majority (>85%) of the high-grade precursors of cancer (CIN 2 and 3).

The recognition that HPV infections are a necessary cause of cervical cancer has several profound implications for cancer prevention. Firstly, in the absence of persistent viral infection, cervical cancer is not expected to develop. Consequently, preventive strategies based on HPV screening or prophylactic vaccination should be viewed as targeting virtually all cervical cancer cases. Secondly, the distribution of HPV types in cervical cancer cases shows a strong predominance of HPV 16 and 18. The 13 most common types account for an estimated 98% of the cancers worldwide. Thirdly, the risk estimate for any of these 13 types is not statistically different from the risk linked to the most common types, HPV 16 or 18. Therefore, the use of a probe set including high-risk HPV types in screening and patient management is justified.

Co-factors that further increase the risk of invasive cancer among HPV DNA-positive women include increasing age, the long-term use of oral contraceptives (five or more years), high parity (five or more full-term pregnancies), smoking and HIV. Co-factors that possibly increase cancer risk include previous exposure to Chlamydia trachomatis and herpes virus type 2.

The age at first exposure to HPV and the age-specific HPV DNA prevalence are strongly related to the patterns of sexual behaviour and are therefore population-specific. To make efficient and effective age-specific recommendations for HPV screening, the HPV attack rate and the age-specific incidence of invasive cancer should be described.

Principles of screening

When cancer precursors can be detected by screening tests, as for cancer of the cervix, the aim of screening is to reduce the incidence, and as a consequence also the mortality from the disease. Screening programmes are directed to populations, but dependent on individuals accepting the invitation to be screened and administration of a high-quality screening test. Subsequent diagnostic tests and treatment are required for those found to be test positive. Ethical issues are important. Women should be aware that screening cannot prevent all cases of invasive cancer occurring and should be informed of the processes and consequences of screening.

Natural history of cervical cancer

Studies of the natural history of cancer of the cervix based on cervical cytology have been of two types, one based upon the invasive cases of the disease and reconstruction of the natural history retrospectively, often using a case–control design, the other on cohort studies following up women who had been screened. Studies incorporating HPV testing have been predominantly cross-sectional or cohort in type. The earlier cytology-based studies had histologically confirmed carcinoma in situ or invasive cancer as the end-point; more recent studies have concentrated on various degrees of CIN, some based only on cytological diagnoses. Several studies have modelled the data from other reported studies.

Most precursor lesions arise within a specific region of the cervix that is referred to as the transformation zone. The transformation zone appears to be particularly susceptible to neoplasia induced by high-risk types of HPV.

Current theories of the pathogenesis of cervical cancer consider infection of the cervical epithelium with specific high-risk types of HPV to play an integral role in the pathogenesis of cervical cancer and its precursor lesions. There is good evidence to support a model of cervical cancer pathogenesis involving a multistep process. Infection with high-risk types of HPV is the first stage in this process. HPV infection of young women is frequent, and transient in the large majority of...
women. The median duration of a prevalently-detected HPV infection is typically about a year for high-risk types of HPV and shorter for the low-risk types. The greatest determinant of clearance of HPV infection is age (maximal in young women) and HPV type (lowest in those infected with HPV type 16). Many women with transient HPV infections will develop cytological abnormalities. When HPV is actively replicating in cells, it can produce characteristic cytopathic effects and borderline or mild cytological abnormalities. Borderline or mild cytological abnormalities are most commonly identified within the first six months of initial infection. In some women with transient infections, colposcopy reveals CIN 1 lesions. A small proportion of women who become infected with HPV develop persistent HPV infections. The biological reasons why some women develop persistent infections are poorly understood.

The great majority of invasive cancers develop after a pre-invasive stage of sufficient length to allow their detection by screening programmes. Persistence of high-risk types of HPV is a prerequisite for the development of CIN 3 lesions and invasive cervical cancers. A variety of natural history follow-up studies based either on cytology alone or using a combination of colposcopy and cervical biopsy have demonstrated that CIN 1 lesions have a relatively high rate of spontaneous regression in the absence of treatment and low rates of progression to higher-grade CIN or invasive cervical cancer. In contrast, CIN 3 lesions and carcinoma in situ lesions have much lower rates of spontaneous regression and higher rates of progression to invasive lesions. The biological behaviour of CIN 2 lesions taken in aggregate appears to be intermediate between that of CIN 1 and CIN 3 lesions in terms or rates of progression, regression and persistence.

Cervical cancer precursors can be defined in a variety of ways including virological measures, biological features and morphological terms. Cellular properties of the majority of invasive cervical cancers and many precursor lesions include monoclonality and aneuploidy. Other cellular events include genetic alterations which may result in the activation of oncogenes and inactivation of tumour-suppressor genes, and increases in telomerase activity. There are a number of inherent problems in many prospective studies of the natural history of cervical cancer precursors. Therefore critical parameters including the rate of progression from precursor to invasion and the proportion of precursors that will progress if left untreated may be poorly characterized. However, it is clear that progression occurs only in HPV-positive women, with very low probabilities of progression in women under the age of 30 years. Estimated progression rates in studies of older women have varied depending on the end-point used. For CIN 3 or carcinoma in situ, the progression rates approximate to 50%, though lower rates have been reported, and rates are higher in studies of prevalent disease, and in older than younger women. For CIN 2 or moderate dysplasia or less, all studies have estimated progression rates of 20% or less.

Regression is an important part of the natural history of both carcinoma in situ and dysplasia (CIN), though the estimated extent of regression has varied, from about 30% of cases of carcinoma in situ or CIN 3 at ages above 50 to 70% at younger ages, while for moderate or slight dysplasia (CIN 2 or less), the majority of lesions regress within five years.

## Screening tests

### Cervical cytology

Cytological testing involves collection of exfoliated cells from the cervix and microscopic examination of these cells after staining. This allows abnormal cells to be detected and an estimation of whether there is an underlying cervical cancer precursor, so as to determine whether the woman needs further follow-up.

Cytology-based screening programmes continue to be the mainstay of cervical cancer prevention worldwide. Over the last two decades there have been major changes in the terminology used for reporting results. In most areas of the world, either the World Health Organization terminology, the cervical intraepithelial neoplasia (CIN) terminology, or the 2001 Bethesda System are used to classify intraepithelial squamous lesions. It is important that cervical cytology specimens be assessed with respect to their adequacy and that a uniform terminology be utilized for borderline cytological changes that are not diagnostic of an intraepithelial lesion. In many countries these borderline specimens account for 3–5% of all specimens and for up to 50% of all women with biopsy-confirmed CIN 2 or CIN 3.

With newer liquid-based cytology (LBC) methods, epithelial cells scraped from the cervix are transferred to a liquid fixative and transported to the cytology laboratory for processing. The unit cost for LBC is considerably higher than that of conventional cytology. Purported advantages of LBC over conventional cervical cytology include a more representative transfer of cells from the collection device to the glass slide, a reduction in the number of unsatisfactory specimens, the possibility of using residual cellular
material for additional molecular testing, and a statistically significant increase in detection of HSIL. However, LBC systems differ in their test characteristics, and data obtained from earlier systems cannot necessarily be extrapolated to new systems.

The importance of training to ensure proper specimen collection cannot be overemphasized. One-half to two-thirds of false negative cervical cytology results are a result of either poor patient conditions (such as active menstruation or severe cervicovaginal infections) at the time the specimen was collected or the manner in which the specimen was collected. It is critical that the entire transformation zone be sampled during specimen collection, since this is the area where almost all CIN 2 and CIN 3 lesions develop.

Adequate quality control and quality assurance programmes are critical to maintaining a high level of performance in a cytology service. These need to provide continuous monitoring of record-keeping, review of abnormal cases by a cytopathologist, review of negative cases either by a 10% rescreening programme or by use of a rapid prescreening or rescreening of all samples, correlation of cytological and histological results from abnormal specimens whenever possible, and proficiency testing programmes. It is also important to set upper and lower workload limits.

**Visual inspection**

Sensitivity of visual inspection with application of acetic acid (VIA) or with Lugol's iodine (VILI) has been found similar to that of cytology for detecting CIN 2–3 or invasive cancer in some developing countries, but specificity is lower. In one large study, VILI showed better sensitivity and potentially better reproducibility. VIA with magnification has not shown any advantage over VIA using various low magnification devices. The low sensitivity of unaided visual inspection precludes its use as a screening test.

Visual inspection tests are inexpensive, safe and acceptable techniques, and require a lower level of infrastructure than laboratory-based tests. They can be performed by a wide range of personnel after a short period of training, and test results are available immediately.

There are no universally accepted definitions of test results for VIA and VILI. Visual inspection methods are subjective and present challenges to maintain the quality of testing. Adequate training and supervision are critical to implement visual inspection-based screening.

**Colposcopy**

Despite the extensive reliance on colposcopy, understanding of how to optimize its performance as a diagnostic or screening test is still incomplete. Recent studies show that colposcopy-directed biopsy has sensitivity for detecting CIN 2 or worse lesions as low as 57%. Colposcopists are well able to differentiate high-grade lesions from other conditions, but differentiation of low-grade changes from normal tissue is more problematic. Relatively few studies have been performed to assess the accuracy of the Reid scoring system, but it appears that there is scope for improvement.

**HPV DNA testing**

Research on the use of HPV DNA testing as a potential cervical cancer screening and management tool began in the late 1980s in response to the evidence that these viruses play a causal role in cervical carcinogenesis and that HPV testing of cervical cells could have acceptable diagnostic performance, while being more reproducible and more easily adapted for clinical practice than conventional cytology.

Techniques to detect the presence of HPV in cervical cell specimens have evolved considerably in the last 25 years and have included methods based on cytological, immunocytochemical and nucleic acid hybridization principles. The Hybrid Capture™ (HC) assay and polymerase chain reaction (PCR) techniques are among the most common and represent signal and target-amplified DNA hybridization approaches, respectively. The former has become an approved technique for screening and triage of equivocal cervical abnormalities in many developed countries. Acceptable standards of testing formats are constantly evolving but in essence, they are based on the principle of detecting, either individually or jointly, the main types of HPV that are associated with cervical cancer. Research is continuing on reproducibility and on agreement between test formats; such performance characteristics must be calibrated with respect to detecting cervical cancer precursors and not merely the presence of HPV in cervical specimens.

For primary screening of women older than 30 years of age, HPV testing yields on average about 10–20% greater sensitivity and 10% lower specificity than cytology (either conventional or liquid-based). In some studies, the combination of cytology and HPV testing (as independent or reflex testing) attained very high sensitivity and negative predictive values (approaching 100%). A testing combination with such a high negative predictive value could potentially allow screening intervals to be increased, e.g., from the minimum of three years up to five years or longer, depending on the population and risk profile. The drawback of this approach is the loss in specificity with respect to either test in isolation due to the excessive number of patients who would need to be referred for colposcopy.
The high unit cost of HPV testing and the fact that it is not a public domain technology, like cervical cytology, remain important impediments to its wider acceptance in cervical cancer prevention. The cost-effectiveness of HPV testing is heavily dependent on assumptions related to the cost of the test, the infrastructure available in the setting where the screening will be implemented, the length of the interval between screening visits, and the existing expenditures incurred by quality assurance imposed by local legislation. More studies are needed in low-, middle- and high-income countries to assess effectivity as a function of these variables.

Although there are no additional physical hazards associated with HPV test in cervical cancer screening, reservations are nonetheless noted. Little is known about the psychological and emotional impact of communicating positive HPV test results to women. If it were eventually implemented in primary screening for cervical cancer, testing for HPV would result in a large proportion of women having to be told that they harbour a sexually transmitted viral infection that can ultimately cause cancer. There is a dearth of research on the merits and consequences of conveying this information. Understanding of the dynamics of sexual transmission of HPV infection is insufficient for health providers to convey meaningful information on risk to men and women.

Other emerging techniques

Computer-assisted reading of cervical smears

Automation-assisted screening is aimed at increasing the sensitivity of cytological testing by finding, for instance, small abnormal squamous and glandular cells which are very difficult to find in conventional screening; it should also increase specificity by selecting only lesions corresponding to objective and reproducible criteria. Automated screening is designed also to increase productivity by excluding normal slides or part of the slides from manual screening by selecting most the atypical images from a slide to be checked by the cytologist.

The few randomized prospective studies and other performance studies have shown that automation-assisted screening may be applicable as a part of routine primary screening and can perform at least as well as conventional screening.

A new generation of automated devices for liquid-based cytology is now being launched, the performance of which has not yet been evaluated in randomized trials.

Physical real-time devices allowing an instant machine-generated result without requiring highly trained personnel hold promise, but have been insufficiently evaluated.

Molecular surrogate markers

Certain DNA, RNA or protein markers associated with the neoplastic transformation process subsequent to HPV infection may be applicable in screening, diagnosis and prognosis. Potential advantages from the use of such markers in clinical practice include: triage of women with minor cytological abnormalities (atypical squamous cells of undetermined significance (ASCUS) and LSIL) with higher specificity than HPV DNA detection; selection of women with lesions with high potential of progression needing treatment; prognosis prediction; improvement of the accuracy of histology as the gold standard for screening test assessment, by more accurate and reproducible classification of histological squamous and glandular cervical lesions and clearer distinction between cervical and endometrial glandular lesions; and last but not least, more accurate primary screening for cervical progressive cancer precursors.

Correlation studies have documented the presence or absence of certain molecular markers in cytological or histological material from selected patients. Test accuracy measures can be assessed for detection of CIN 2, CIN 3 or cancer, but are not representative for real screening, triage or follow-up settings. No such markers have yet been validated for use.

Combination of different modalities

Adding a sensitive detection method for high-risk HPV DNA to cytology yields a substantial increase in test sensitivity and negative predictive value for CIN 3 or cancer and probably allows an increase in screening intervals. The concomitant decrease in specificity is of particular concern for large populations. Cost-effectiveness analyses are still hampered by lack of long-term outcome data and information on psycho-social effects of mass HPV screening. The additional gain in sensitivity from adding cytology to HPV testing often is minor or negligible. Adding an HPV test to visual inspection in resource-poor settings is impractical because of the cost.

A second test can be used sequentially as a triage method where the purpose of the second test is to restrict the number of screen-positive women requiring referral. Among women showing equivocal cytology results, HPV DNA testing is more accurate for detecting underlying CIN 3 or worse disease than repeat cytology. HPV DNA testing is not useful for triage of mild cytological lesions with very high HPV positivity (e.g., LSIL as interpreted in the USA).

Colposcopy finds its place in the follow-up of cytological screen-positive women to decide the need for biopsy and to orient biopsy or excision for diagnosis or treatment. Negative colposcopic findings are not conclusive of
absence of high-grade cervical disease because of the intrinsic false-negative rate and the difficulty of observing localized endocervical disease.

### Use of cervical cancer screening

#### Delivery and uptake of screening Europe

Europe has several well organized and documented cervical cancer screening programmes at national and regional levels, but opportunistic screening activity is still predominant in most countries. Most countries recommend three-year screening intervals from age 25 to 64 or 65 years, but a few countries recommend five-year intervals and some one-year intervals. This results in large differences in the lifetime number of recommended screening tests across countries.

The reported statistics on screening activity and performance, available for a few countries or regions, mostly for organized programmes, are not always comparable. Most screening activity is not documented and there is an almost complete lack of information for some countries, especially in eastern Europe. Attendance above 80% or in the 70–80% range is found in a few countries, especially in the presence of organized programmes and in northern Europe, while levels of 60% or below are found in other countries, especially in southern Europe. Excess testing is substantial in most countries, with the exception of England and a few other areas.

The organization of screening delivery, cytological classification and management of women with abnormal test results varies between countries. Guidelines or regulations related to quality assurance, mainly of cytology, exist in different countries. Systems to ensure compliance with follow-up recommendations are adopted in organized programmes.

#### USA and Canada

In the USA, screening is mainly opportunistic, with only about 1% of the population being covered by a national cervical cancer screening programme administered by the Centers for Disease Control and Prevention (CDC). In the opportunistic service, payment is largely the responsibility of the individual woman. New cervical cancer screening guidelines mostly recommend annual cervical cancer screening from age 21 until age 30 for conventional testing or every two years for liquid-based testing, and every 2–3 years between age 30 and 65–70 if previous tests were normal. Figures from the National Health Interview Survey show that screening coverage in the past three years is 82% in women aged 25 and older. Quality of cytology testing is regulated at the national level. Recommendations for follow-up of abnormal cytological test results now also include HPV DNA testing.

In Canada, screening is mostly opportunistic, with the services being reimbursed by provincial health plans. Some provinces have organized cervical cancer screening programmes for women aged 18–69 years, to which, once enrolled, women are invited at regular intervals. Population-based recruitment plans have not yet been implemented. Canada has issued screening and follow-up recommendations. New recommendations from 2003 include use of liquid-based testing and HPV DNA testing. Survey data show the three-year coverage to be 79% for women aged 20–69 years.

#### Latin America and the Caribbean

Attempts to organize screening programmes have failed in most Latin American countries, in spite of a coverage of over 60% reported in many countries. Long-standing organized screening programmes are in operation in Colombia and Chile, and are beginning to show reduction in incidence and mortality. In the rest of the region, there has been a lack of attention to quality assurance of cytology and, in many instances, lack of follow-up, diagnosis and treatment for women screened positive.

#### Africa

Many studies have concluded that providing cervical cancer screening services in sub-Saharan Africa is essential as this region carries a high burden of disease. However, access to early diagnosis and treatment is limited by severe resource constraints, competing health and development needs and dysfunctional health-care systems. The challenge for Africa is to develop screening services integrated into existing health services in such a way as to improve the overall functioning of health-care systems.

#### Oceania and Asia

Australia and New Zealand have established organized national cervical cancer screening programmes, while China (Taiwan) and the Republic of Korea have recently initiated screening programmes. All programmes use cytology and screen women every 2–3 years, targeting women aged 30 years and over in China (Taiwan) and the Republic of Korea and aged 20 years and over in Australia and New Zealand. These screening programmes have management structures for administration and quality assurance, and collect data to monitor performance indicators. Visual inspection with acetic acid is under trial in India and Thailand for possible use in screening. Outside of the few organized screening programmes, most women do not yet have access to screening for cervical cancer.

In Japan, an organized nationwide cervical cancer screening programme was initiated in 1983. Although government financing was phased out in
1998, the organized screening programme is still offered by each regional government. Annual screening by cytological testing is recommended to begin at age 30. There is no recommended age to end screening. One-year coverage in Japan is estimated to be about 25%. In April 2004, the Ministry of Health, Labor, and Welfare issued new recommendations stating that screening should be initiated at 20 years of age with an interval of two years.

This screening programme also stipulates the establishment of Management Control Committees that organize screening delivery and quality assurance and monitor performance indicators.

**Behavioural considerations in screening participation**

Factors associated with participation in cervical cancer screening include: having a contact with the health system, a good patient–physician relationship, a female screening provider, and providing the screening test in a setting where privacy is assured. In many countries, an invitation letter increased attendance. However, this procedure may be difficult to implement in low-resource settings.

Multi-component interventions that include education, home visits and involving family and key community members appear to be effective in increasing uptake among hard-to-reach women.

**Efficacy of screening**

In evaluating the efficacy of screening, it is preferable to have data from randomized screening trials. However, no such data are available with the incidence of clinical invasive cancer of the cervix as the end-point. The data available come from observational studies of screening in defined populations: cohort and case–control studies, which may not be free of bias, and studies of incidence and mortality trends which could be affected by changes in the impact of risk factors for the disease.

For new screening modalities, the data available are generally from studies that compared the sensitivity and specificity of different screening tests, allowing estimation of relative sensitivity of the tests in detecting cervical cancer precursors.

For conventional cervical cytology, studies using cohort, case–control or geographical correlation (before/after analysis) designs indicate substantial effects in reducing the cervical cancer incidence and mortality rates, the impact exceeding 80% among women screened in various organized settings. Studies in Scandinavia, the United Kingdom and British Columbia, Canada, have been most informative. There is evidence that the screening impact is particularly large in the organized screening programmes. Opportunistic screening has also been found to reduce cervical cancer incidence, although generally to a smaller extent than in organized programmes, and requires far more resources.

There is evidence that the duration of the low risk after a negative cytology screening test is less in women under the age of 35 years than in older women. However, the incidence of invasive cancer of the cervix is extremely low in women aged less than 25, while in women aged 25–34 there is a low absolute risk of invasive cancer of the cervix after a negative screening test during the following three years. In women over the age of 35, and especially over the age of 50, the risk of invasive cancer of the cervix after a negative test is low for the next five years.

The evidence does not support annual screening at any age, a repeat test one year after the first test, nor screening after the age of 65 in cytologically negative women.

Meta-analyses have shown that liquid-based cytology is at least equivalent to conventional cytology in terms of relative sensitivity and specificity. There are insufficient data to evaluate the efficacy of currently available automation-assisted cytological screening systems, although data using a system that is no longer commercially available suggest that automated systems can be as sensitive and specific as high-quality conventional cytology. There are no long-term studies of impact on incidence or mortality using either of these new technologies.

Over a dozen studies have shown that testing for high-risk HPV is substantially more sensitive (around 95%) for detecting CIN 3 than conventional cytology (around 70%). Two studies have found a lower rate of CIN 3 on follow-up of HPV-negative compared to cytology-negative women, suggesting that the screening interval can be safely lengthened following a negative HPV test. Studies looking at archival smears and antibodies in stored sera indicate that HPV is present several years before the diagnosis of cancer or CIN 2/3 in cytologically negative women. The data suggest that testing for HPV infection, the necessary cause of cervical cancer, possibly at a longer interval than for cytology, may lead to lower invasive cervical cancer rates, but there are no data on cancer incidence or mortality rates after HPV screening.

Unaided visual inspection (also referred to as downstaging) is associated with low sensitivity (30–50%) to detect cervical cancer precursors and is no longer considered to be a suitable screening test. Naked-eye visual inspection after application of 3–5% acetic acid (VIA) and VIA with low-level magnification (VIAM) have similar test characteristics and VIAM has been shown to have no advantage over VIA. VIA has been evaluated in several cross-sectional studies, mostly in developing countries. In most cross-
sectional studies comparing VIA with conventional cytology, similar or higher sensitivity but lower specificity was seen. VIA is also being investigated in three randomized controlled trials in India for which early results are available from two cluster randomized trials. Rates of detection of CIN 2/3 lesions by VIA were similar in both trials (7/1000). A significantly higher frequency of stage I cancers (35–48%) was observed in the VIA screening group compared to the control group (0–24%). However, the detection rate of CIN 3 by VIA was significantly lower than by cytology in one trial.

Pooled results from 10 cross-sectional studies of visual inspection with Lugol’s iodine (VILI) screening in India and Africa indicate higher relative sensitivity and similar specificity to VIA.

No significant difference in the performance of cytology in HIV-positive and HIV-negative women has been found. There are increased incidence and prevalence of HPV infection, intraepithelial lesions and cervical cancer among HIV-positive women compared to uninfected women. However, the impact of these data on the efficacy of screening in HIV-positive women is not clear.

**Effectiveness of screening in populations**

**Incidence and mortality trends in relation to screening**

Time trends in the incidence and mortality rates of cervical cancer are of considerable interest, as they may shed light on changes in exposure to etiological factors and provide a means of evaluating the success, or otherwise, of screening programmes. Comparisons of trends in the Nordic countries have been particularly informative. Decreases in incidence and mortality since the late 1960s were greatest in Finland, Sweden and Iceland, which had the most extensive screening programmes, and least in Norway, which had organized screening in only a single county.

Cervical cancer incidence and mortality rates have generally declined in the last few decades in many other populations in Europe, the USA, Japan and Oceania. However, there have been periods of increase, particularly in women under 35 years of age, in some areas, although these have occurred at different times in different countries. Observation of these trends has sometimes (e.g., in the United Kingdom) resulted in changes in screening practices, in order to attempt to reverse the upward trend. Cervical cancer mortality rates have been rapidly rising in a number of eastern European countries where there is little screening; this trend may have been reversed recently in the more affluent parts of eastern Europe.

There is less information on time trends in cervical cancer in developing countries. Rates of incidence and mortality have generally been stable, or shown modest declines. This probably reflects the lack of screening programmes, or, where they exist, their low coverage and/or poor quality.

The overall trends in cervical cancer incidence largely reflect trends in squamous-cell carcinoma, the dominant morphological type, and the type for which screening techniques have historically been more effective. Increases in cervical adenocarcinoma have occurred in many countries, especially in young women.

**Use of modelling in the design and evaluation of screening**

Statistical models can be used to explore the relationships between screening test, policy and programme characteristics and the expected reduction in incidence and mortality (and derivatives such as years of life saved). Simulation models have been developed that use observed data on disease natural history, screening test performance and effectiveness of different treatment options for precancerous lesions and can allow for heterogeneity of risk, accessibility, compliance and feasibility.

The quality of the models has improved over time as the underlying parameters (natural history, test sensitivity, etc.) have become better understood. The models have also become more widely used, as the contribution of the sophisticated methodology has become better appreciated and the statistical techniques more widely disseminated. As with any model, they are subject to the accuracy of the assumptions.

Use of such models has led to an improved understanding of the relative importance of various screening parameters and the relative gains to be expected; this can help screening programmes to infer where changes might be most effective.

**Issues in the implementation of screening**

The cytology test has been shown to be effective when well applied. Where cytology screening has failed to work, blame can be laid on the design or delivery of the screening service.

Demonstration projects can and should be used to ensure that screening of proven efficacy is implemented in an optimal manner for a given population.

**Hazards of screening**

The hazards of screening include anxiety and fear among women related to the difficulty of understanding the meaning of both negative and positive screening results, as well as the difficulty of understanding the concept of precancer. Further, there are problems related to the test itself, with both false positive and false negative
results, which may lead to either overtreatment or unnecessary medical interventions, or to undertreatment of significant lesions. Other hazards include the complications of treatment (cervical stenosis and incompetence leading to infertility) and the use of more medical therapies (e.g., hysterectomy) due to complications of treatment. Finally, in HIV-endemic settings and where HIV-positive women are not treated for HIV, cervical cancer screening may result in an inappropriate diversion of resources.

**Performance evaluation**

Cervical cancer screening should be implemented within the context of an established programme policy regarding age range and between-screen interval. A programme’s policies will determine its maximal effectiveness. However, programme implementation and delivery will determine the actual effectiveness.

The essential elements for optimal effectiveness include identification of women in the target population; measures to guarantee high coverage and attendance; high-quality smear taking, reading and reporting; and methods to ensure follow-up of women identified by screening as having lesions needing further assessment. Indicators have been established to assess performance on these elements. The most important of these, in terms of effectiveness, is coverage or participation (proportion of women being screened). However, high coverage alone is not sufficient. If indicators in other areas, such as quality, are suboptimal, effectiveness will be compromised.

Performance indicators should be monitored regularly, preferably through routine systems of information. Where these do not exist, special studies such as population surveys and programme audits should be conducted periodically.

**Cost-effectiveness**

Cost-effectiveness models require data on natural history of cervical cancer, the overall effectiveness of the policy or intervention, survival rates associated with cancer, test characteristics, and quality of life. Sources could include: trials, observational studies, meta-analyses; other published literature, expert opinion and health systems statistics. Cost data must be collected locally or estimated according to local conditions. The eventual judgement as to whether a particular strategy is cost-effective or not will be dependent on local circumstances.

Common points are that expending resources on improving coverage and on achieving and maintaining adequate quality is almost always more cost-effective than any other potential service improvement and that opportunistic screening is less cost-effective than organized screening.
In reaching its evaluation, the Working Group distinguished evidence of two types. The strongest evidence derives from historical or prospective data on efficacy, currently available only for cervix cancer from observational studies or time trends in populations. However, evidence based upon surrogate markers of reduction in cancer incidence was utilized when derived from a comparison with comparable data following screening with a test shown to reduce cancer incidence by the first type of evidence. In the evaluations that follow, an evaluation based on the first type of evidence is expressed by the words "has reduced cervical cancer incidence and mortality rates", and those based on the second type of evidence by the words "can reduce cervical cancer incidence and mortality rates".

There is sufficient evidence that screening by conventional cytology has reduced cervical cancer incidence and mortality rates.

There is sufficient evidence that testing for human papillomavirus infection as the primary screening modality can reduce cervical cancer incidence and mortality rates.

There is limited evidence that screening by visual inspection with application of acetic acid can reduce cervical cancer incidence and mortality rates.

There is limited evidence that screening by visual inspection with application of Lugol’s iodine can reduce cervical cancer incidence and mortality rates.

Overall evaluation
There is sufficient evidence that screening for cervical cancer precursors every 3–5 years between the ages of 35 and 64 years by conventional cytology in a high-quality programme reduces the incidence of invasive cervical cancer by 80% or more among the women screened. In women aged 25–34 years, screening at intervals of three years or less may have smaller impact. There is no evidence that screening annually in either age group results in much greater efficacy. Other forms of cytology screening using a validated system at the same ages and frequency can be expected to be as effective as conventional cytology.

Efficacy of conventional cytology has been demonstrated only for squamous-cell carcinoma.

Screening in well organized programmes is more cost-effective, with less harm due to overscreening and overtreatment, than opportunistic screening.

Data for analysing cost-effectiveness must be gathered locally and any conclusions drawn must be appropriate to the context. Investing in obtaining high rates of population coverage is critically important in achieving a cost-effective intervention.

There is sufficient evidence, based on surrogate markers, that the efficacy of HPV testing, using a validated system, as the primary screening modality can be expected to be at least as good as that of conventional cytology.

For visual inspection with application of acetic acid (VIA) or with Lugol’s iodine (VILI), there is limited evidence of efficacy. Cross-sectional studies in low-resource settings have shown VIA and VILI to be similar to conventional cytology in detecting CIN 2–3. In view of the current uncertainties in the definition of test results, reproducibility and quality assurance, long-term results from randomized trials that are in progress are essential for further evaluation of visual inspection.
Chapter 8
Recommendations for public health implementation and further research

Introduction
Much of the evidence to be generated on the long-term effectiveness of modified or new screening modalities, in terms of reduction in the incidence of invasive disease, will come from evaluation of the results of organized population-based programmes. Modifications of screening modalities in existing screening programmes therefore need to be introduced in a way that will facilitate rigorous evaluation of long-term effectiveness. This is best achieved by incorporating randomization.

This requirement is particularly relevant for cervical cancer screening, where much of the evidence for programme modifications comes from short-term end-points (e.g., cross-sectional rates of histological diagnosis of CIN 3) and cost considerations.

For the reasons above, this section presents recommendations for public health implementation and further research combined.

General
1. Two major determinants of the effectiveness of public health screening programmes are high coverage of the target population and quality of the total screening process including the primary screening test and follow-up of those positive. Research is needed on methods (i) to improve coverage, especially among deprived populations, and (ii) for continuing quality assurance, whatever the screening modality in operation.
2. Once an organized system is in place, opportunistic (or unscheduled) screening should be discouraged.
3. New developments in screening technology can be evaluated in short-term or cross-sectional studies using surrogate markers of efficacy such as sensitivity and specificity for a histological diagnosis of CIN 3, compared with screening tests known to reduce cervical cancer incidence such as high-quality conventional cytology. The design of such short-term studies is most efficient if the same women undergo both the new and the established test.
4. For longer-term assessment of efficacy in terms of absolute sensitivity or incidence of invasive cancer, a design in which women are allocated to different modalities is required. For example, comparison might be between HPV testing only versus cytology, cytology versus HPV testing plus cytology or visual inspection with acetic acid (VIA) versus cytology. Collection of material for retrospective analysis of another modality is an acceptable study design.
5. The adoption of a new screening modality in a population-based screening programme should depend upon the local cost environment, expertise and facilities. Considerations include the capacity both for the primary screening test and for management of screen-detected lesions. Any such implementation should be based on population-based studies.
6. All screening will have associated negative effects. These include psychosocial, biological and economic effects of the screening episode. Research is needed to minimize the impact of each of these components. In particular for cervical screening, research is needed into the possible negative effects of overtreatment of screen-detected lesions. In all comprehensive assessments and comparisons, full account needs to be taken of the potential harmful consequences of screening.
7. HIV-positive women have a higher prevalence of HPV infection and cervical cancer and its precursors of all degrees of severity. These women may therefore benefit from more frequent screening than HIV-negative women. Whether
screening should begin at a younger age in HIV-positive women is unclear and requires study.

**Conventional cytology**

The issues identified by the Working Group for implementation of conventional cytology were the ages at which screening should start and might stop, and the possibility of adopting different screening frequencies at different ages.

1. There is minimal benefit and substantial harm in screening below age 25. Organized programmes should not include women aged less than 25 years in their target populations.

2. Women who have always tested negative in an organized screening programme should cease screening once they attain the age of 65, as there is little benefit of screening to women over the age of 65 who have had at least two negative tests in the last 10 years. Research is needed to determine whether screening can cease earlier.

3. For women over age 50, a five-year screening interval is considered appropriate. For women aged 25–49, a three-year rather than a five-year interval might be considered in countries with the necessary resources. Annual screening is not recommended at any age.

4. Implementation of the preceding three recommendations needs continuing long-term evaluation in terms of invasive cervical cancer incidence and mortality.

5. In countries with limited resources, solutions other than those recommended above are likely to be required. However, screening should always be introduced after an informed strategic analysis within the context of the national (or regional) cancer control programme, and only after the necessary resources and facilities to permit high-quality screening, efficient diagnosis and management of detected abnormalities are secured.

6. New screening policies developed in accordance with the previous recommendation should usually be piloted in a feasibility (demonstration) project in a defined area with a defined population, preferably with a population-based cancer registry. In the absence of such a registry, an information system that includes data on the cases of clinically invasive cancer that occur in the population can be used.

**New developments in cytological screening**

1. Large, randomized controlled clinical trials comparing the performance of LBC and conventional cytology need to be conducted by laboratories in which the techniques are well established.

2. Implementation of liquid-based cytology and automation-assisted screening in organized screening programmes needs to be based on cost and local feasibility. It is imperative that the introduction of each new modality is accompanied by long-term evaluation of impact on invasive cancer and continuing quality assurance and monitoring. The age and screening interval for conventional cytology should also apply here.

3. New modifications to these modalities are frequently proposed. Each such modification needs rigorous evaluation in short-term assessments of relative sensitivity and specificity for histologically diagnosed CIN 3 compared to the current standard, as well as economic and logistic evaluations before implementation.

**HPV testing**

If a country, on reviewing the available evidence, decides to introduce HPV testing as a primary screening modality, it must consider local circumstances, including the acceptability of the test. Introduction would be facilitated by the availability of low-cost public-domain HPV tests. Implementation should be preceded by demonstration projects. Large-scale implementation needs to be designed so as to allow rigorous long-term evaluation.

1. It is likely that the same reduction in incidence of invasive disease could be achieved with a longer screening interval using HPV testing as a screening test than the intervals recommended above for cytological screening. It is expected that evidence supporting a longer interval may emerge from properly designed public health screening programmes in which HPV testing has been incorporated.

2. The optimal ages for starting and stopping HPV screening require further research. Age-specific population rates of HPV positivity and cancer incidence should be based on samples for immediate (reflex) cytology testing. Cohort studies and demonstration projects are appropriate for this research. At present, commencing HPV-based screening at ages below 30 years is not recommended.

3. The management of women who are HPV-positive but negative on cytology is of vital importance to avoid overtreatment, particularly in younger women, in whom transient infections are common. Research is required to identify secondary biomarkers, whether cellular or viral, which are accurate predictors of persistence of viral infection and/or progression of cervical lesions.
4. In countries that lack expertise for high-quality conventional cytology which adopt HPV testing as a primary screening test, long-term follow-up studies should be conducted of the effectiveness of HPV testing followed by colposcopy without cytology in women who test positive.

5. Efficient implementation of HPV screening requires research into HPV as a viral infection as well as a screening test. Aspects to be studied include: (i) factors, both individual and social, determining transmission and susceptibility; (ii) factors influencing age-specific rates of infection, reinfection, duration and natural history of infection in the relevant populations (these studies need to be by type and variant of HPV and, in particular, need to consider older women, for whom relatively little information currently exists); (iii) the consequences of using a test for a sexually transmitted agent as a primary screening test in terms of behavioural and psychosocial impact; and (iv) the natural history of HPV infection in males, including risk factors for persistence and transmission as well as the social impact of knowledge of HPV infection status in men and women.

6. Health professionals and the population at large must be educated on HPV and its connections with cervical cancer and screening programmes.

7. HPV testing systems need to be standardized and specification requirements for test performance need to be defined.

8. New commercial testing systems need rigorous evaluation and validation before being adopted by the public health system.

Visual inspection
1. The evidence on visual inspection with application of acetic acid (VIA) or with Lugol's iodine (VILI) was considered not yet sufficient for it to be recommended for adoption as the primary screening test in a public health programme. The size and quality of studies currently in progress should provide definitive evidence on the impact of a single test on the cumulative incidence of both invasive cancer and advanced disease. Future research will be required to assess the impact of repeated testing using a variety of intervals. The validity of the test is highly dependent on the training and skills of those performing the test. Research is needed to establish quality assurance markers for visual inspection.

2. Studies in other settings are needed to confirm the performance of VILI.

3. Studies are required to improve the specificity of criteria in order to increase the positive predictive value of "screen and treat" policies, and so reduce the false positive rate.

Colposcopy
1. Colposcopy is less accurate than is commonly believed and improving colposcopy depends upon improved training and assuring quality control of existing services. Basic clinical research in relation to improved visual diagnosis and use of tissue sampling techniques should be encouraged.

2. Studies are required to improve the specificity of colposcopic criteria to increase the positive predictive value of "see and treat" policies, and so reduce the false positive rate.


Anderson, C.M. & Nottingham, J. (1999) Bridging the knowledge gap and communicating uncertainties for informed consent in cervical cytology screening; we need unbiased information and a culture change. *Cytopathology, 10*, 221–228


Asian HPV Summit, Leura, Australia 2003, CD supplied by Diagnostic Technologies


Australian Institute of Health and Welfare (2003a) Cervical Screening in Australia 1999-2000 (Cancer Series No. 21, AIHW Cat. No. 16), Canberra, AIHW


References


References


shah, k.v. (1995) prevalence of human papillomavirus in cervical cancer: a worldwide per-

schiffman, m.h., moreno, v., kurman, r. & Gonzalez, l.c. (1992) risk factors for cervical cancer in colombia and spain. int. j. cancer, 52, 750–758

bosch, f.x., muñoz, n., de sanjose, s., guerrero, e., ghaffari, a.m., kaldor, j., castellsagüé, x., shah, k.v. & gaffari, a.m. (1994) importance of human papillomavirus endemi city in the incidence of cervical cancer: an extension of the hypothesis on sexual behavior. cancer epidemiol. biomarkers prev., 3, 375–379

bosch, f.x., manos, m.m., muñoz, n., sherman, m., jansen, a.m., peto, j., schiffman, m.h., moreno, v., kúrman, r. & shah, k.v. (1995) prevalence of human papillomavirus in cervical cancer: a worldwide perspective. international biological study on cervical cancer (ibscs) study group. j. natl cancer inst., 87, 796–802

bosch, f.x., castellsagüé, x., muñoz, n., de sanjose, s., ghaffari, a.m., gonzalez, l.c., gili, m., izarzugaza, i., viladiu, p., navarro, c., vergara, a., ascunce, n., guerrero, e. & shah, k.v. (1996) male sexual behavior and human papillomavirus dna: key risk factors for cervical cancer in spain. j. natl cancer inst., 88, 1060–1067

bosch, f.x., muñoz, n., chichareon, s., ngelangel, c., caceres, e., eluf-neto, j., el gueddari, b., rolón, p.a., diaz, m., meijer, c.j.l.m. & waltomers, j.m.m. (2000) hpv and cervical adenocarcinoma: an iarc based multicentric case-control study. in: castellsagüé, x., bosch, f.x., de sanjose, s., moreno, v. & ribes, j., eds, 18th international papillomavirus conference—program and abstracts book. barcelona, thau l., p. 131 (available online: http://www.hpv2000.com)

bosch, f.x., lörincz, a., muñoz, n., meijer, c.j. & shah, k.v. (2002) the causal relation between human papillomavirus and cervical cancer. j. clin. pathol., 55, 244–265


bowa, j., sanson-fisher, r., boyle, c., pope, s. & redman, s. (1995) a randomised controlled trial of strategies to prompt attendance for a pap smear. j. med. screen., 2, 211–218

boyd, j.t. & doll, r. (1964) a study of the aetiology of carcinoma of the cervix uteri. br. j. cancer, 13, 419–434

boyer, s.n., wazer, d.e. & band, v. (1996) e7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. cancer res., 56, 4620–4624

boyes, d.a. & worth, a.j. (1976) cytological screening for cervical carcinoma. in: jordan, j.a. & singer, a., eds, the cervix, london, w.b. saunders, pp. 404–411

boyes, d.a., worth, a.j. & anderson, g.h. (1981) experience with cervical screening in british columbia. gynecol. oncol., 12, s143–s155

boyes, d.a., morrison, b., knox, e.g., draper, g.j. & miller, a.b. (1982) a cohort study of cervical cancer screening in british columbia. clin. invest. med., 5, 1–29

boyle, p., autier, p., bartelink, h., baselga, j., boffetta, p., burn, j., burns, h.j., christensen, l., denis, l., dicato, m., diehl, v., doll, r., franceschi, s., gillis, c.r., gray, n., gricute, l., hackshaw, a., kasler, m., kogevinas, m., kivinsland, s., la vecchia, c., levi, f., mcvie, j.g., maisonneuve, p., martin-moreno, j.m., bishop, j.n., oleari, f., perrin, p., quinn, m., richards, m., ringborg, u., scully, c., siracka, e., storm, h., tubiana, m., tursz, t., veronesi, u., wald, n., weber, w., zaridze, d.g., zatonski, w. & zur hausen, h. (2003) european code against cancer and scientific justification: third version (2003). ann. oncol., 14, 973–1005


brady, c.s., duggan-keen, m.f., davidson, j.a., varley, j.m. & stern, p.l. (1999) human papillomavirus type 16 e6 variants in cervical carcinoma: relationship to host genetic factors and clinical parameters. j. gen. virol., 80 (pt 12), 3233–3240

branca, m., morosini, p., duca, p., verderio, p., giovagnoli, m.r., riti, m.g. & leoncini, l. (1998) reliability and accuracy in reporting cin in 14 laboratories. developing new indices of diagnostic variability in an interlaboratory study. the working group for external quality control in cervical cytopathology. acta cytol., 42, 1370–1376

brand, e., berek, j.s., nieberg, r.k. & hacker, n.f. (1987) rhabdomyosarcoma of the uterine cervix. sarcoma botryoides. cancer, 60, 1552–1560


brenna, s.m., hardy, e., zeferino, l.c. & namura, i. (2001) conhecimento, atitude e pratica do exame de papanicolaou em mulheres com cancer de colo uterino. [knowledge, attitudes, and practices related to the pap smear among women with cervical cancer]. cad. saude publica, 17, 909–914

brinton, l.a., hamman, r.f., huggins, g.r., lehman, h.f., levine, r.s., mallin, k. & fraumeni, j.f., jr (1987) sexual and reproductive risk factors for invasive squamous cell cervical cancer. j. natl cancer inst., 79, 23–30

brinton, l.a., reeves, w.c., brenes, m.m., herrero, r., de brutton, r.c., gaitan, e., tenorio, f., garcia, m. & rawls, w.e. (1989a) parity as a risk factor for cervical cancer. am. j. epidemiol., 130, 486–496

brinton, l.a., reeves, w.c., brenes, m.m., herrero, r., de brutton, r.c., gaitan, e., tenorio, f., garcia, m. & rawls, w.e. (1989b) the male factor in the etiology of cervical cancer among sexually monogamous women. int. j. cancer, 44, 199–203

british society for clinical cytology (1997) recommended code of practice for laboratories providing a cytopathology service 1997. cytopathology, 8 suppl. 1, 1–25
References


Cancer, DNA-positive, cytologically negative women.

Absolute risk of a subsequent abnormal Pap


College of American Pathologists (1997) Interlaboratory Comparison Program in Cervicovaginal Cytology: 1996 Year-end Summary; Northfield, IL


IARC Handbooks of Cancer Prevention Volume 10: Cervix Cancer Screening


References


General Medical Council (1998) Seeking Patients' Consent: The Ethical Considerations, London

References


Suppl. 4

cytology.
Med.,
Am. Stat. Assoc.,
Kaplan, E.L. & Meier, P . (1958) Nonparametric
Gynecol.,
papilloma virus-related cervical neoplasia.

Kaufmann, A.M., Backsch, C., Schneider, A. &
Durst, M. (2002) HPV induced cervical
carcinogenesis: molecular basis and vaccine
development. Zentralbl. Gynakol., 124,
511–424
screening by means of questionnaire and
Suppl. 4, 1–224
Med., 46, 437–444
Keating, J.T., Cviko, A., Riethdorf, S.,
Riethdorf, L., Quade, B.J., Sun, D., Duensing,
S., Sheets, E.E., Munger, K. & Crum, C.P.
(2001) Ki-67, cyclin E, and p16INK4 are com-
plimentary surrogate biomarkers for human
papilloma virus-related cervical neoplasia.
Am. J. Surg. Pathol., 25, 884–891
Keenlyside, R.A., Collins, C.L., Hancock, J.S.,
Gagnon, M.C., Cohn, R.D., Menoff, A.L.,
Dodd, L.G., Kurytcz, D.F., Hearn, T.L. & Baker,
E.L., Jr (1999) Do proficiency test results cor-
relate with the work performance of screeners
who screen Papanicolaou smears? Am. J.
Clin. Pathol., 112, 769–776
Kenny, F.D. (1989) Patients’ knowledge of
cervical screening and risk factors for cervical
cervical cancer: evidence from marital clus-
ters. Cancer, 39, 1912–1919
Keys, H.M., Bundy, B.N., Stehman, F.B.,
Muderspach, L.I., Chafe, W.E., Suggs, C.L.,
radiation, and adjuvant hysterectomy compa-
red with radiation and adjuvant hysterecto-
my for bulky stage IB cervical carcinoma. New
Khleif, S.N., DeGregori, J., Yee, C.L.,
Otterson, G.A., Kaye, F.J., Nevins, J.R. &
Howley, P.M. (1996) Inhibition of cyclin D-
CDK4/CDK6 activity is associated with an
E2F-mediated induction of cyclin kinase
inhibitor activity. Proc. Natl Acad. Sci. USA,
93, 4350–4354
Kierkegaard, O., Byralsen, C., Hansen, K.C.,
Association between colposcopic findings and
histology in cervical lesions: the significa-
tance of the size of the lesion. Gynecol. Oncol.,
57, 66–71
Cost-effectiveness of alternative triage strate-
gies for atypical squamous cells of under-
determined significance. JAMA, 287, 2382–2390
Kilien, L.J. & Doll, R. (1973) Trends in mortal-
ty from cancer of the uterus in Canada and in
27, 146–149
with positive cervical smears but without surgical
intervention. A follow-up study. Lancet, 2,
463–465
Kinnunen, W.K., Manos, M.M., Hurley, L.B. &
Ransley, J.E. (1998) Where’s the high-grade
cervical neoplasia? The importance of mini-
mally abnormal Papanicolaou diagnoses.
Obstet. Gynecol., 91, 973–976
Epidemiological evaluation of cervical cancer
Gynaecol. Oncol., 13, 75–82
Kjaer, S.K., de Villiers, E.M., Dahl, C.,
Engholm, G., Bock, J.E., Vestergaard, B.F.,
study of risk factors for cervical neoplasia in
Denmark. I: Role of the “male factor” in women
with one lifetime sexual partner. Int. J. Cancer,
48, 39–44
Kjaer, S.K., Van den Brule, A.J., Bock, J.E.,
Poll, P.A., Engholm, G., Sherman, M.E.,
Human papillomavirus – the most significant
risk determinant of cervical intraepithelial
Kjaer, S.K., Van den Brule, A.J., Bock, J.E.,
Poll, P.A., Engholm, G., Sherman, M.E.,
Determinants for genital human papillo-
mavirus (HPV) infection in 1000 randomly
chosen young Danish women with normal
Pap smear: are there different risk profiles for
oncogenic and nononcogenic HPV types?
Cancer Epidemiol. Biomarkers Prev., 6,
799–805
Kjaer, S.K., Svre, E.I., Worm, A.M.,
Walboomers, J.M., Meijer, C.J. & van Den
infection in Danish female sex workers.
Decreasing prevalence with age despite con-
tinuously high sexual activity. Sex. Transm.
Dis., 27, 438–445
Kjaer, S.K., Chackerian, B., van den Brule,
A.J., Svre, E.I., Paull, G., Walboomers, J.M.,
Schiller, J.T., Bock, J.E., Sherman, M.E.,
human papillomavirus is sexually transmitted:
evidence from a follow-up study of virgins
starting sexual activity (intercourse). Cancer
Kjær, S.K., Svre, E.I., Munk, C., Jensen, M.L.
Hoelund, B., Norrild, B., Nielsen, L.P, Westh,
H., Andersen, E. S., Lundh, J., Olesen, F. &
Holten I. (2002a) [Guidelines for use of HPV-
testing in Denmark]. Ugeskr. Laeger, Status
rapport 8 (in Danish)
Kjær, S.K., Van den Brule, A.J., Paull, G.,
Svre, E.I., Sherman, M.E., Thomsen, B.L.,
Suntum, M., Bock, J.E., Poll, P.A. & Meijer,
C.J. (2002b) Type specific persistence of high
risk human papillomavirus (HPV) as indicator
of high grade cervical squamous intraepithelial
lesions in young women: population based
prospective follow up study. BMJ, 325, 572

References


References


References


References
The effect of hysterectomies and screening on mortality from cancer of the uterus in Canada. 


National Cervical Screening Programme (NZ) (1998) *Cervical Screening: Information for Health Professionals*, Wellington, Health Funding Authority


References


Information leaflets, verbal information and 23

274

probes.

with molecularly cloned virus-specific DNA cervical and vaginal intraepithelial neoplasia


References


Role of the pap test in cervical cancer screening.


for cervical intraepithelial neoplasia and risk of preterm delivery. JAMA, 291, 2100–2106


References


References


Slater, D.N. (2000) Are women sufficiently well informed to provide valid consent for the cervical smear test? Cytopathology, 11, 166–170


Sparén, P. (1996) Early Detection and Screening for Cancer of the Cervix in Sweden during the 20th Century, (Comprehensive summaries of Uppsala Dissertations from the Faculty of Medicine 599), Uppsala, Acta Universitatis Upsaliensis


References


Statistics Canada (2002) The Canadian Community Health Survey Cycle 1.1, Ottawa


Sundhedsstyrelsen (1986) [Vejledende retningslinier for screeningen mod livmoderhalskraeft], Copenhagen


Tavassoli, F.A. & Stratton, M.R., eds (2003) Pathology and Genetics of Tumours of the Breast and Female Genital Organs (WHO Classification of Tumours), Lyon, IARCPress


References


References
IARC Handbooks of Cancer Prevention Volume 10: Cervix Cancer Screening


The presence of high-risk HPV combined with specific p53 and p16INK4a expression patterns points to high-risk HPV as the main causative agent for adenocarcinoma in situ and adenocarcinoma of the cervix. J. Pathol., 201, 535–543


### Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atypical squamous cell</td>
<td>Cells that are considered suggestive but not diagnostic of a squamous intraepithelial lesion, at cytology.</td>
</tr>
<tr>
<td>Background cervical cancer incidence rate</td>
<td>The cervical cancer incidence rate expected in the absence of screening. It is not directly observable but estimated from the incidence in the target population before screening started (and adjusted for trend) or incidence at about the same time in an unscreened referent population, or in unscreened controls in the case of a randomized trial.</td>
</tr>
<tr>
<td>Biopsy</td>
<td>Tissue specimen for morphological or immunohistochemical diagnosis</td>
</tr>
<tr>
<td>Cancer registry</td>
<td>System of ongoing reporting of cancer patients in a defined population. More broadly a research institute that utilizes a cancer register and other information for epidemiological research.</td>
</tr>
<tr>
<td>Cervical cancer incidence rate</td>
<td>The rate at which new cases of cervical cancer occur in a population. The numerator is the number of newly diagnosed cases of cervical cancer that occur in a defined period. The denominator is the population at risk of a diagnosis of cervical cancer during this defined period multiplied by the length of this period, sometimes expressed as person-time.</td>
</tr>
<tr>
<td>Cervical cancer mortality rate</td>
<td>The rate at which deaths from cervical cancer occur in a population. The numerator is the number of cervical cancer deaths that occur in a defined time period. The denominator is the population at risk of dying from cervical cancer during this defined period multiplied by the length of the period, sometimes expressed as person-time.</td>
</tr>
<tr>
<td>Cervical cancer register</td>
<td>Recording of information on all new cases of and deaths from cervical cancer occurring in a defined population.</td>
</tr>
<tr>
<td>Cervicography</td>
<td>Photography of the cervix taken after the application of 5% acetic acid, using a camera with a fixed focal length and internal light source. The images are projected on a screen at a fixed distance to simulate magnification and are interpreted as to grade of neoplasia by a specially trained evaluator.</td>
</tr>
<tr>
<td>Cohort effect</td>
<td>Effect of an etiological exposure or medical or societal intervention that affects differently persons born in successive birth cohorts.</td>
</tr>
<tr>
<td>Colposcopy</td>
<td>Magnified visual examination of the cervix using a low-power stereoscopic binocular field microscope with a powerful light source.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Cost-effectiveness</td>
<td>An analysis of the costs relative to the effectiveness of a procedure or activity, or comparisons of similar activities to determine the relative degree they will achieve similar effectiveness.</td>
</tr>
<tr>
<td>Coverage</td>
<td>Number of women invited as a proportion of target population. Also the number of women who have a screening test within the recommended interval as a proportion of all women who are eligible to attend for screening. In the second meaning, this term is equivalent to attendance or participation rate.</td>
</tr>
<tr>
<td>Delay time</td>
<td>The time between when a lesion destined to became cancer could be detected by screening and when it is actually detected by screening. Not directly observable. Cf. lead time</td>
</tr>
<tr>
<td>Demonstration project</td>
<td>A health-care project with built-in provision for measuring cost, performance and outcome of a model service.</td>
</tr>
<tr>
<td>Detectable preclinical phase</td>
<td>The time between that at which a tumour could be found by screening and that at which it would become clinically recognized (not directly observable). Length of DPCP is sojourn time and it is composed of delay time and lead time.</td>
</tr>
<tr>
<td>(DPCP)</td>
<td></td>
</tr>
<tr>
<td>Detection method for</td>
<td>To estimate sensitivity by detection rate and interval cancer incidence.</td>
</tr>
<tr>
<td>sensitivity</td>
<td></td>
</tr>
<tr>
<td>Detection rate</td>
<td>Proportion of cancers (preinvasive lesions) confirmed during the screening episode among those screened or in the target population.</td>
</tr>
<tr>
<td>Direct-to-vial</td>
<td>Where liquid-based cytology is used and cells exfoliated from the cervix are placed directly and completely in the vial of preservative liquid.</td>
</tr>
<tr>
<td>Down-staging</td>
<td>Screening with identification of invasive disease in asymptomatic women at an earlier clinical stage than those detected clinically.</td>
</tr>
<tr>
<td>Effect</td>
<td>The result of screening. Effect measures are changes in incidence of and/or mortality from cervical cancer.</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>The reduction in incidence of and/or mortality from invasive cervical cancer due to screening practice, under real conditions and among those in the target population.</td>
</tr>
<tr>
<td>Efficacy</td>
<td>The reduction in incidence of and/or mortality from invasive cervical cancer under ideal conditions (in randomized trials), and among those screened compared to the incidence or mortality in those randomized not to be screened but compliant if invited to be screened.</td>
</tr>
<tr>
<td>Efficiency</td>
<td>The effects or end results achieved in relation to the effort expended in terms of money, resources and time.</td>
</tr>
<tr>
<td>Episode</td>
<td>The period from the time of test (taking the smear) to the end of time of further assessment, i.e., the time of decision to intervene or not.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>False (change) gain in sensitivity</td>
<td>When a second, adjunct test is added to a conventional, primary test and positive results by the second test are used to supplement the positivity of the primary test, the estimate of sensitivity will always be greater than that of the first test used alone, even if the second test were totally random with respect to the disease or to the first test. The increased combined sensitivity may or may not be greater than that contributed by an unrelated adjunct test in the same screening setting. Ideally, studies should consider sensitivity gains of combined testing only after taking into account this chance increase in sensitivity.</td>
</tr>
<tr>
<td>Further assessment</td>
<td>Additional diagnostic steps (either non-invasive or invasive) performed to clarify the nature of an abnormality detected by the screening test, either at the time of screening or on recall or as a result of referral.</td>
</tr>
<tr>
<td>Gold standard</td>
<td>A diagnostic method that is considered to have the best sensitivity and specificity among all methods available.</td>
</tr>
<tr>
<td>Incidence method for sensitivity</td>
<td>To estimate sensitivity as 1 – ratio of interval cancer incidence rate between two screens to that expected if there was no screening.</td>
</tr>
<tr>
<td>Incidence (annual) of preinvasive lesions</td>
<td>Detection rate of the lesion at given subsequent screen divided by the screening interval. Alternatively, the number of new cases of preinvasive lesions divided by the person time, which equals number of women screened multiplied by screening interval.</td>
</tr>
<tr>
<td>Informed choice</td>
<td>Decision about whether or not to participate, based on the provision of information about the benefits and limitations of screening.</td>
</tr>
<tr>
<td>Informed consent</td>
<td>Voluntary consent given by a subject for participation, after being informed about the purpose, procedures, benefits and risks.</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Material and human resources and their interrelationships.</td>
</tr>
<tr>
<td>Interval cancer</td>
<td>An invasive cervical cancer diagnosed in an attender, after a negative screen, either:</td>
</tr>
<tr>
<td></td>
<td>• before the next invitation to screening was due or</td>
</tr>
<tr>
<td></td>
<td>• within a period equal to a screening interval for a woman who has reached the upper age limit for screening.</td>
</tr>
<tr>
<td>Interval cancer (incidence) rate</td>
<td>Interval cancers divided by person years in the period the cancers are derived from. The rate is different for test, episode and programme.</td>
</tr>
<tr>
<td>Lead time</td>
<td>Period between when a lesion destined to become cancer is found by screening and when it would have been clinically recognized if no screening took place (cf. delay time).</td>
</tr>
<tr>
<td>Length bias</td>
<td>The bias towards detection by screening of cancers with longer sojourn times and therefore better prognosis.</td>
</tr>
</tbody>
</table>
Loop electrosurgical excision procedure (LEEP)  
LEEP uses a thin wire loop electrode attached to an electrosurgical generator as a precise and rapid surgical tool. The generator transmits a painless electrical current that quickly cuts away affected cervical tissue in the immediate area of the loop wire.

Microinvasive cancer  
Cancers that have invaded no more than 5 mm deep and 7 mm wide into the underlying cervical stroma.

Organized screening  
Screening programmes organized at national or regional level, with an explicit policy, that includes several essential elements from target population to treatment.

Opportunistic screening  
Screening outside an organized or population-based screening programme, as a result of, for example, a recommendation made during a routine medical consultation for the woman, consultation for an unrelated condition, on the basis of a possibly increased risk for developing cervical cancer or by self-referral.

Outcome  
Event related to objective of screening (death from cervical cancer), sometimes also to the performance of screening.

Overcall  
Recall or referral with poor specificity

Overdiagnosis  
Detection of cervical cancers or preinvasive lesions that would never have progressed to be clinically recognized during a woman's life.

Overtreatment  
Treatment of lesions that would never have progressed to be clinically recognized during a woman's life.

Participation rate  
Proportion of those screened among those invited according to the scheduled policy (organized screening). In a programme not based on invitations, participation has the same meaning as coverage.

Performance  
Quality of screening activities mainly related to the laboratory, sometimes to all the screening process rather than outcome.

Performance indicators  
Quantitative measures of the process of screening. Generally, targets are set of the quantity which is required for good quality process.

Period effect  
Effect of an etiological exposure or medical or societal intervention that affects differently in time.

Pilot study  
A demonstration project that provides information on performance but not on outcome and is based on a limited population.

Population access  
Proportion of the national population of eligible women who have access to a screening programme (cf. coverage).

Positive predictive value  
Proportion of diagnoses of cancer in all positive results of the screening test. A process measure

Positivity rate of test  
Proportion of diagnoses of cancer in all positive results of the screening test. A process measure.
### Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary screening</td>
<td>Detection of cases of cervical cancer or of its precursor lesions among asymptomatic women without a referral diagnosis, i.e., as true population screening, either opportunistic or systematic.</td>
</tr>
<tr>
<td>Quality assurance</td>
<td>Maintenance of minimum standards and continual striving for excellence.</td>
</tr>
<tr>
<td>Quality control</td>
<td>The supervision and control of all operations involved in a process, usually involving sampling and inspection, in order to detect and correct systematic or excessively random variations in quality.</td>
</tr>
<tr>
<td>Recall</td>
<td>Clarification of a perceived abnormality detected at screening, by performance of an additional procedure.</td>
</tr>
<tr>
<td>Recall rate</td>
<td>The number of women recalled for further assessment as a proportion of all women who were screened (test positivity rate).</td>
</tr>
<tr>
<td>Referral</td>
<td>Physical referral of women to a clinical facility as a consequence of the screening test for diagnostic confirmation, e.g., by histology.</td>
</tr>
<tr>
<td>Reflex HPV testing</td>
<td>A protocol for routine triage of equivocal cervical cytological interpretations, by HPV testing either the residual liquid cytology specimen or an additional specimen collected at the same time as the original sample.</td>
</tr>
<tr>
<td>Relative sensitivity</td>
<td>Ratio of detection rate of malignancy after test A to the detection rate of malignancy after test B. Also sensitivity of test A relative to histology. See verification bias.</td>
</tr>
<tr>
<td>Relative survival</td>
<td>Survival if cervical cancer were the only cause of death among cervical cancer patients.</td>
</tr>
<tr>
<td>Screen and treat</td>
<td>A procedure where testing, confirmation and treatment take place during the same episode.</td>
</tr>
<tr>
<td>Screening interval</td>
<td>Fixed interval between routine screenings decided upon in each programme, depending on screening policy.</td>
</tr>
<tr>
<td>Screening policy</td>
<td>Specific policy of a screening programme which dictates the targeted age group, the geographical area, the screening interval, etc. Opportunistic systems may also have policies.</td>
</tr>
<tr>
<td>Screening test</td>
<td>Test applied to all women in a programme that results in discrimination between those who test positive from those who test negative (e.g., Pap smear). Those who test positive will be recalled or referred for further assessment or diagnostic confirmation.</td>
</tr>
<tr>
<td>See and treat</td>
<td>A procedure where the cervix is treated at first attendance for colposcopy and no histology information is available.</td>
</tr>
</tbody>
</table>
Sensitivity

Capacity of screening to identify unrecognized disease, i.e., future invasive cervix cancer in a population or disease in the DPCP.

- sensitivity of test is the proportion of those with a positive test among those with disease in the DPCP
- sensitivity of the episode is the proportion of those with disease detected by screening among those with the disease in the DPCP among those screened
- programme sensitivity is the proportion of those with disease detected by the screening organization among those with disease in the DPCP among total target population
- sensitivity by incidence method is estimated with interval cancers and background incidence

Sojourn time

Detectable preclinical phase; time between that at which a tumour could be found by screening and that at which it would be clinically recognized if the woman was not screened (not directly observable).

Specificity

Capacity of screening to identify those who remain healthy in a population.

Split sample

Sample of the exfoliated cervical cells where liquid-based cytology sample is split between preparation of a conventional Pap smear and the balance of cells being deposited in a vial of liquid preservative.

Target population

The population eligible for screening, i.e., all women recommended to undergo screening according to the policy adopted.

Triage

Detection of cases of cervical cancer or of its precursor lesions among women who were initially found to have an abnormal screening test that requires further evaluation.

Undercall

Recall or referral with poor sensitivity

Verification bias

A bias in the relative sensitivity and specificity estimates that occurs if the probability of disease verification via the gold standard (e.g., colposcopy and biopsy) is dependent on the screening test result. It may also occur when there are two screening tests whose results the investigator uses to decide who will be referred for the gold standard. In that case, bias will ensue if the positivity of the second test is evaluated conditionally on the positivity of the first test.
Prevention of cancer is one of the key objectives of IARC. Secondary prevention by early diagnosis and screening is a fundamental component of any cancer control programme. The aim of secondary prevention is to reduce mortality and suffering from the disease. When screening is planned as part of a cancer control programme, only strategies proved to be effective should be proposed to the general population. Screening usually requires repeated interactions between ‘healthy’ individuals and health care providers, which can be inconvenient and costly. Furthermore, screening requires an ongoing commitment between the public and health care providers.

Scope
Cochrane (1972) first discussed the concepts of efficacy and effectiveness in the context of health interventions. Efficacy was later defined by Last (1995) as "the extent to which a specific intervention, procedure or service produces a beneficial result under ideal circumstances". In contrast, the related term "effectiveness" is defined by the same author as "... a measure of the extent to which a specific intervention, procedure, regimen or service, when deployed in the field in routine circumstances, does what it is intended to do for a specific population." The distinction between efficacy as measured in experimental studies and the effectiveness of a mass population intervention is a crucial one for public health decision-making. In particular, the fact that the effectiveness of a screening procedure may be different in different populations is often overlooked. A mass programme of screening must satisfy certain minimal requirements (e.g. acceptability, availability of relevant personnel, facilities for screening and access to pertinent health services) if it is to achieve the results that have been documented in epidemiological studies. The acceptance and use of screening services may vary from one population to another, implying that a given screening procedure is not universally effective. Even when a screening procedure is effective as a mass intervention, other outcomes such as harm and costs and the potential for other interventions to achieve equivalent benefits must be considered.

Efficacy is a necessary but not a sufficient basis for recommending screening. The efficacy of a screening procedure can be inferred if effectiveness can be proven. Screening has sometimes been implemented by a given procedure on the assumption that ‘earlier is better’, even when no evidence of efficacy was available. If such interventions result in a significant reduction in mortality that cannot otherwise be explained, it can be inferred that the procedure is effective. However, uncontrolled interventions in which individuals are exposed to unknown risks and benefits should be avoided.

Objectives
The objectives of the Working Group are:
(1) to evaluate the strength of the evidence for the efficacy of a screening procedure;
(2) to assess the effectiveness of defined screening interventions in defined populations;
(3) to assess the balance of benefit and harm in target populations; and
(4) to formulate recommendations for further research and for public health action.

The conclusions of the Working Group are published as a volume in the series of the IARC Handbooks of Cancer Prevention.

Working groups
An international working group of experts is convened by the IARC. The tasks of the group are:
(1) to ascertain that all appropriate data have been retrieved;
(2) to select the data relevant for evaluation on the basis of scientific merit;
(3) to prepare accurate reviews of data to allow the reader to follow the reasoning of the working group;
(4) to evaluate the efficacy and effectiveness of the screening procedure;
(5) to summarize the potential adverse consequences of screening;
(6) to prepare recommendations for research and for public health action; and
(7) to prepare an overall evaluation of the screening procedure at the population level.

Approximately 13 months before a working group meets, the topics of the Handbook are announced, and prospective participants are selected by IARC staff in consultation with other experts. Working group participants who contributed to the considerations and evaluations within a particular handbook are listed, with their
addresses, at the beginning of each publication. Each participant serves as an independent scientist and not as a representative of any organization, government or industry. They are expected to put aside any stake they may have in a particular outcome and to evaluate the evidence objectively and with scientific rigour. All participants are required to complete a form before the meeting on which they declare any potential conflict of interest, due for example to recent links with commercial or industrial bodies that have a stake in the outcome of the meeting. Participants who declare any such potential conflict of interest are excluded from chairing the meeting or any of its subgroups, from drafting evaluations and from any voting that may be involved in reaching the final conclusions. They may otherwise participate fully in the meeting, and are designated in the list of participants (pages vi-viii) as ‘invited experts’.

Scientists nominated by national and international agencies, industrial associations and consumer and/or environmental organizations may be invited as observers. IARC staff members involved in the preparation of the handbook are listed as secretariat.

Subsequently, relevant data are collected by the IARC from all available sources of published information. About eight months before the meeting, the material collected is sent to meeting participants who are asked to prepare sections for the first drafts of the handbook. These drafts are then compiled by IARC staff and sent, before the meeting, to all participants of the working group for review.

**Data for handbooks**

The handbooks do not necessarily cite all of the literature on the agent or strategy being evaluated. Only those data considered by the working group to be relevant to making the evaluation are included. Meeting abstracts and other reports that do not provide sufficient detail upon which to base an assessment of their quality are generally not considered.

With regard to reports of basic scientific research, epidemiological studies and clinical trials, only those that have been published or accepted for publication in the openly available scientific literature are reviewed by the working group. In certain instances, government agency reports that have undergone peer review and are widely available are considered. Exceptions may be made *ad hoc* to include unpublished reports that are in their final form and publicly available, if their inclusion is considered pertinent to making a final evaluation.

The available studies are summarized by the working group. In general, numerical findings are indicated as they appear in the original report; units are converted when necessary for easier comparison. The working group may conduct additional analyses of the published data and use them in their assessment of the evidence. These analyses are described in the handbook. Important aspects of a study, directly impinging on its interpretation, are brought to the attention of the reader.

**Evaluation of screening**

The framework of a handbook on screening includes the following eight chapters:

**Chapter 1. Disease characteristics, global burden and rationale for screening**

**Descriptive epidemiology**

The purpose of this section is to document the importance of the disease in the context of the general health status of different populations. The worldwide burden of the cancer is described (mortality, incidence, prevalence and survival rates) and integrated with measures of the occurrence of cancers at other sites, of mortality from all causes and life expectancy. Expected trends in the absence of screening are a relevant component of this section.

**Natural history of the disease as relevant to screening**

In this section, the natural history of the disease of interest and the relevance and potential of screening for early detection and for reducing mortality are described. Evolving concepts and principles pertinent to screening are also discussed.

There is now a wealth of evidence (both direct and indirect) to support the principle that screening and detection of certain cancers in appropriate target populations are associated with a lower probability of dying from the disease. The scheme (on the next page) illustrates the temporal framework commonly subscribed to in modern screening models.

It should be noted that early diagnosis, due to greater awareness and improved access to appropriate medical services, has resulted in many countries in a reduction in diagnostic delay, probably reducing mortality. As a consequence, symptomatic cancers are frequently diagnosed and treated early after the onset of symptoms in many developed nations. In such instances, screening for the disease will improve outcomes (for example, reducing mortality) only if treatment of the disease at an even earlier phase in its development provides additional benefit. The rapid evolution of molecular or genetic markers of pre-malignant conditions or individuals at high risk has modified the concepts of ‘disease onset’ and ‘lead time’. Hence, the model outlined above may require adaptation or development to allow for detection of pre-clinical conditions of undetermined significance (including serological and molecular markers and genetic predisposition), if they are relevant for screening for the cancer in question.

**Chapter 2. Screening tests**

It is important to distinguish between
screening tests and screening procedures, i.e. the test itself and the way in which it is administered. The two merit separate, detailed evaluation. Each of the screening tests to be considered is described. The ability of each test to detect cancer and to distinguish cancer from non-cancer conditions will be assessed as:

- the validity of the test, expressed as its sensitivity and specificity under various conditions;
- all known or potential side-effects; and
- the cost of the test when implemented in mass screening programmes.

Chapter 3. Delivery and uptake of screening

Information on how screening is delivered in different countries is reviewed in this section, with emphasis on the following aspects:

- infrastructure for diagnosis and treatment: the nature of standard diagnostic procedures and treatment regimens and their availability to the target population;
- extent of population coverage and participation rates;
- equity, as defined by the extent to which access to the procedure (including diagnostic investigation and treatment) is ensured for all eligible individuals, irrespective of any personal characteristics;
- informed decision and informed consent: the extent to which individual values are respected when information on potential benefit and harm is conveyed; and
- behavioural and demographic considerations that affect participation in screening.

Chapter 4. Efficacy of screening tests

In this section, evidence from epidemiological studies is reviewed, and aspects of study design and analysis are critically discussed. The handbooks are not intended to summarize all published studies. The working group considers the following aspects:

(1) the relevance of the study;
(2) the appropriateness of the design and analysis to the question being asked;
(3) the adequacy and completeness of the presentation of the data; and
(4) the degree to which chance, bias and confounding may have affected the results.

Studies that are judged to be inadequate or irrelevant to the evaluation are generally omitted. They may be mentioned briefly (i) when the information is considered to be a useful supplement to that in other reports, (ii) if they provide the only data available or (iii) in exceptional cases, if they have been widely perceived as being pertinent but are deemed otherwise by the working group. Their inclusion does not imply acceptance of the adequacy of the study design nor of the analysis and interpretation of the results, and their limitations are outlined.

The appropriate outcomes) (mortality or incidence) of a given procedure, e.g. the detectable phases) of the natural history of the disease, are also defined. Aspects that are particularly important in evaluating experimental studies are: the selection of participants, the nature and adequacy of the randomization procedure, evidence that randomization achieved an adequate balance between the groups, the exclusion criteria used before and after randomization, compliance with the intervention in the screened group and ‘contamination’ with the intervention in the control group. Other considerations are the means by which the end-point was determined and validated (either by screening or by other means of detection of the disease), the length and completeness of follow-up of the groups and the adequacy of the analysis.

Whenever possible, similar criteria should be used to evaluate non-experimental comparative studies.

In the Working Group’s analysis of the efficacy of the screening procedure, a meta-analysis may be used, when applicable.

In evaluating case–control and cohort studies, particular attention is paid to the definition of cases, controls and exposure and, for cohort studies, the length and completeness of follow-up. Potential bias, especially selection bias, is carefully examined in all observational studies.

Chapter 5. Effectiveness of population-based screening

The impact of the screening procedure when implemented in defined populations is examined in this section. Indicators used to monitor effectiveness, such as positive and negative predictive values, detection rate, rates of interval cancers and the number of tests performed, are reported. Time trends before and after implementation of screening as well as geographical comparisons of the occurrence of the disease and death from the disease in populations exposed and not exposed to screening are reviewed and interpreted. In doing this, the Working Group takes into account differences in
screening procedures (e.g. frequency and the age of the target population) and of participation rates.

An integral component of this section is an evaluation of the benefits and harms of the screening procedure to the population. Reductions in mortality and/or incidence of invasive disease are fundamental measures of benefit. An additional benefit is that more cases can be treated by less aggressive, less invasive procedures, thus improving the quality of life.

The spectrum of health care is dynamic, and a screening procedure should not be viewed in isolation. Greater awareness of the disease, brought about by publicity about screening that may result in early diagnosis, could be regarded as another benefit of a screening programme. This section should also consider the possibility that there might have been a change in treatment of the cancer, which even in the absence of screening would have resulted in a substantial decrease in mortality. As far as possible, an evaluation should be made of the extent to which improved treatment has been responsible for any changes seen in mortality from the specific disease.

Estimates of the rates of false-positive and false-negative findings in screened individuals and their consequences (false sense of security with false-negatives and false alarm with false-positives) are an integral part of this section. The rates of short- and long-term side-effects and the possibility of unnecessary treatment of borderline or indolent cases detected at screening are discussed. Management procedures for lesions detected at screening are reviewed. Psychological factors, such as anxiety induced by undergoing the test procedure, are also considered.

Finally, the cost-effectiveness of various modalities of test administration in various settings is considered. The discussion takes into account the costs per case detected and per death prevented.

Chapter 6. Summary of data
In this section, the relevant data are summarized. Inadequate studies identified in the preceding text are generally not included.

Chapter 7. Evaluation
Evaluation of the efficacy of the screening procedure
An evaluation of the degree of evidence for the efficacy of a screening procedure is formulated according to the following definitions:

Sufficient evidence of the efficacy of cancer-preventive activity will apply when screening interventions by a defined procedure are consistently associated with a reduction in mortality from the cancer and/or a reduction in the incidence of invasive cancer, and chance and bias can be ruled out with reasonable confidence.

Limited evidence of the efficacy of cancer-preventive activity will apply when screening interventions by a defined procedure are associated with a reduction in mortality from the cancer and/or a reduction in the incidence of invasive cancer or a reduction in the incidence of clinically advanced cancer, but bias or confounding cannot be ruled out with reasonable confidence as alternative explanations for these associations.

Inadequate evidence of the efficacy of cancer-preventive activity will apply when data are lacking or when the available information is insufficient or too heterogeneous to allow an evaluation.

Sufficient evidence that the screening procedure is not efficacious in cancer prevention will apply when any of the following cases hold:

• the test does not result in earlier diagnosis than with standard tests already in use;
• the survival of cases detected at screening is no better than that of cases diagnosed routinely;
• the screening interventions are consistently associated with no reduction in mortality from or incidence of invasive cancer, and bias can be ruled out with reasonable confidence.

In the case of limited or inadequate evidence, the Working Group should highlight those aspects of the procedure for which information is lacking and which led to the uncertainty in evaluation. This will provide indications of research priorities.

Overall evaluation
Finally, the body of evidence is considered as a whole, and summary statements are made about the cancer-preventive effects of the screening intervention in humans and other beneficial or adverse effects, as appropriate. The overall evaluation is usually in the form of a narrative. The data on the effectiveness of the screening intervention are summarized, including the factors that determine its success and failure under routine conditions. Finally, the balance between expected benefit and harm is described.

Chapter 8. Recommendations
After its review of the data and its deliberations, the working group formulates recommendations, where applicable, for:

• further research and
• public health action.

References