Effects and Costs of Cervical Cancer Screening

Marjolein van Ballegooijen

CIP gegevens Koninklijke Bibliotheek, Den Haag
Ballegooijen, Marjolein van

Effects and Costs of Cervical Cancer Screening:
Marjolein van Ballegooijen.

Proefschrift Erasmus Universiteit Rotterdam.
Met lit. opg. – met samenvatting in het Engels en in het Nederlands.

ISBN 90-72245-86-5

Trefwoorden: bevolkingsonderzoek / baarmoederhalskanker

Printing: Universiteitsdrukkerij TUE

Effects and Costs of Cervical Cancer Screening

Proefschrift

Ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof dr P.W.C. Akkermans M.A.
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
woensdag 15 april 1998 om 15:45

door
Marjolein van Ballegooijen
geboren te Etterbeek (België)

Promotie-commissie

Promotor:	Prof. dr ir J.D.F. Habbema
Overige leden:	Prof. dr T. Helmerhorst
	Prof. dr S. Thomas
	Prof. dr J.G. van der Tweel
Co-promotor:	Dr ir G.J. van Oortmarssen

Contents

Chapter 2–5 and 7–9 are based on the following papers:

2 Ballegooijen M van, Koopmanschap MA, Oortmarssen GJ van, Habbema JDF, Lubbe JThN, Agt HMA van, Diagnostic and treatment procedures induced by cervical cancer screening → Eur J Cancer, 26, 941–945 (1990)

3 Ballegooijen M van, Habbema JDF, Oortmarssen GJ van, Koopmanschap MA, Lubbe JThN, Agt HHA van, Preventive Pap-smears: Balancing costs, risks and benefits → Br J Cancer, 65, 930–933 (1992)

4 Ballegooijen M van, Koopmanschap MA, Tjokrobardojo AJS, Oortmarssen GJ van, Care and costs for advanced cervical cancer → Eur J Cancer, 28A, 1703–1708 (1992)

5 Ballegooijen M van, Koopmanschap MA, Habbema JDF The management of cervical intraepithelial neoplasia: extensiveness and costs → Eur J Cancer, 13A, 1672–1676 (1995)

7 Ballegooijen M van, Boer R, Oortmarssen GJ van, Koopmanschap MA, Lubbe JThN, Habbema JDF Cervical cancer screening: Age-ranges and intervals. An updated cost-effectiveness analysis → Erasmus University Rotterdam (1993)

8 Ballegooijen M van, Beck S, Boon ME, Boer R, Habbema JDF The rescreen effect in conventional and Papnet screening observed in a study using material enriched with positive smears → Acta Cytologica (in press)

9 Ballegooijen M van, Akker ME van den, Warmerdam PG, Meijer CJLM, Walboomers JMM, Habbema JDF Present evidence on the value of HPV testing for cervical cancer screening: a model-based exploration of the (cost-)effectiveness → Br J Cancer, 76(5) 651–657 (1997)

1

Introduction

1.1 Cervical cancer mortality in the Netherlands

In recent years in the Netherlands, around 235 women have died annually from cervical cancer. This is 0.35% of total female mortality [38]. Cervical cancer mortality stabilised between 1950 and 1965. Between 1965 and 1970 mortality started to decrease. This decrease occurred before an effect from screening could be expected *(see Figure 1.1)*. The decrease continued after the start of screening, and currently the mortality rate is less than 40% of the level in the period 1950–1959. A decreasing mortality before the introduction of screening can be seen in other countries as well [82, 126, 96, 16]. Since randomized cervical cancer screening trials have never taken place, it is not possible to estimate to what extent, if any, the decrease in mortality is due to screening.

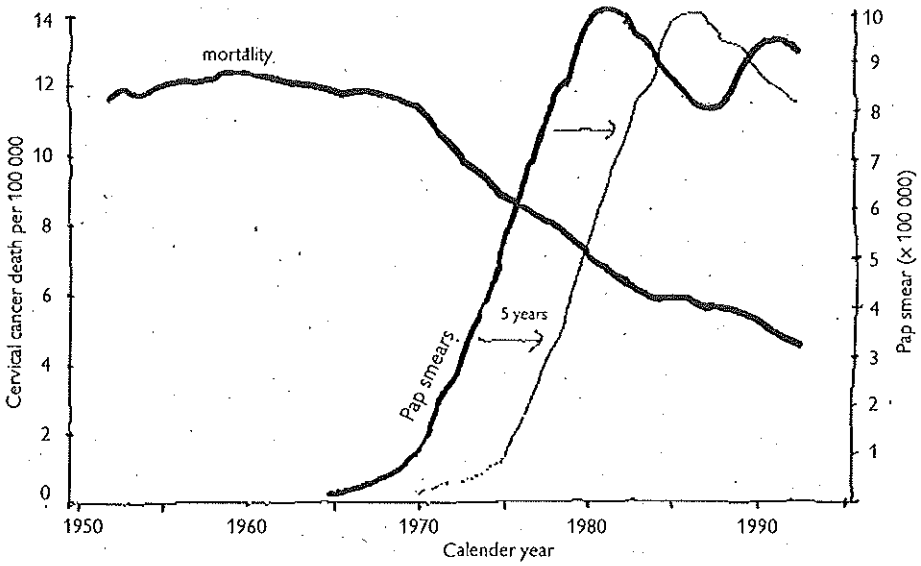
1.2 Cervical cancer worldwide

The incidence and mortality of cervical cancer varies between regions in the world [168, 171]. In developed countries, the incidence and mortality rates are lower (2.5 times for mortality) than in developing countries. A lower mortality is partly due to a better survival in developed countries (around 60% 5-year survival for Europe [19]) than in developing countries (e.g. a 38% 5-year survival in Bangalore, India [102]). The Netherlands is amongst the countries with low incidence with 9 new cases per 100 000 women annually (1993) [156]. Mortality rates in the EU-countries are low in the Netherlands (2.4 per 100 000 women, world standard rate (WSR) [154]) and highest in Denmark (5.0 per 100 000 women, WSR [1]).

1.3 The pap smear

Cervical cancer can be detected at an early stage by screening women who have not yet developed signs or symptoms. The screening test that has been

Figure 1.1 *Mortality from cervical cancer by calendar year in the Netherlands [39](moving 5-yearly average per 100 000 women aged 20 years and over), age adjusted to the Dutch population in 1995, and annual numbers of smears by true calendar year and shifted 5 years later, after which time interval a notable effect from screening on mortality is expected to start*



used up to now is known as the PAP smear. The technique of this cytological test is as follows: cells are scraped from the cervix uteri e.g. by the general practitioner, and evaluated microscopically by cytotechnicians and pathologists. If the test is positive (abnormalities are observed) a follow-up takes place. This can result in the detection of cervical cancer or its precursors at a stage in which symptoms are absent i.e. pre-clinical. In over 90% of cases, the detected neoplasia is intraepethelial (without invasion of underlying tissues), e.g. dysplasia and carcinoma in situ, but early invasive cancers are also detected by screening. When left untreated, cervical intraepithelial neoplasia (CIN) will, at least in some women, develop into invasive cervical cancer. After treatment, prognosis is excellent for CIN, with well over a 99% cure [114, 175, 72, 106, 62, 198]. Prognosis is also quite good for early invasive cervical cancer, with a 97% 5-year survival in FIGO stage IA and around 80% in stage IB [80, 20, 54]. A new test that might be useful for the early detection of cervical cancer is the test for human papilloma-virus (HPV) on a cervical scrape. The HPV-test is currently being investigated for its possible use in combination with or, as an alternative to, cytological evaluation.

1.4 The effectiveness of screening

Because of the good prognosis of screen-detected and treated cases, PAP smear testing is expected to reduce the mortality from cervical cancer. Mortality reduction, the main goal of cancer screening, is ideally estimated from randomized controlled trials, but such trials have never been carried out for cervical cancer screening. The reason is that cervical cancer screening started, in parts of the US and Canada, around 1950 in an era in which the idea of evidence-based medicine was not yet established [147]. In fact there was such a conviction that screening would be effective, a randomized trial seemed unethical. Evidence for mortality reduction from cervical cancer screening and the assessment of its magnitude consequently had to emanate from less powerful methods, e.g. by comparing regions or periods or individuals with different screening intensities in non-randomized cohort studies [126]and in case-control studies [43, 100, 81].

1.5 Balancing benefits, negative side-effects and costs

The main benefit of effective screening apart from the reduction of mortality is the prevention of invasive cancers. But screening will inevitably also have negative side-effects for some women. Currently, more than 10% of women screened will have positive test-results [14], a large majority of whom will have no benefit at all from screening. The financial costs of screening are also high and depend largely on the number of PAP smears carried out.

In cervical cancer screening, the balance between benefits, harms and costs is vulnerable. The background cumulative risk for cervical cancer in the Netherlands is only around 1.5% (Chapter 6), which means that the incidence and mortality reduction will affect at most 1.5% of the population, whereas the negative side-effects may apply to a much higher proportion of the screened women (for example 10% per round), and the costs are dominated by the costs of the PAP smears and therefore apply to the total screened population. The quality of screening, in terms of this balance, depends highly on the screening strategy followed. Important components of this strategy are the age range and interval between subsequent PAP smears, and the follow-up strategy in cases of abnormal smear test results.

1.6 Age range and interval

Nowadays, in West-European and North-American countries, many women have at least one PAP smear during their life. Most of these women have several and some have many smears. Screening was introduced generally in the sixties,

Table 1.1 *Number of smears and screening interval used in countries with cervical cancer screening programmes or recommended in national guidelines*

Country	Number of smears per woman	Screening interval in years
The Netherlands until 1996	7	3
The Netherlands from 1996 onwards	7	5
Denmark [200]	13	3
Sweden [135]	12	3
Iceland [195]	17	2–3
Finland [96]	7	5
United Kingdom [158]	10–16	3–5
USA (guidelines) [203]	±17	<3
Australia [61]	27	2

as opportunistic (non-invitational) screening by gynaecologists and general practitioners and partly also by implementing invitational programmes.

Screening strategies in developed countries differ. In Europe, women are invited 7 to 16 times during a life-time, with an interval of 2 to 5 years *(see Table 1.1)*. In Northern America, annual screening was recommended in the sixties and early seventies [120]. Guidelines proposed to start screening at age 20, implying that up to 40 smears would be taken per woman during her lifetime [194]. Later, guidelines tended to advise intervals of three years, still implying at least 17 smears per woman because of the advice to start screening at a young age [203]. In Australia, programme screening has started targeting women between the ages of 18 and 70 every two years, thus inviting women 27 times during their life [61]. The differences in the number of smears per woman between countries are often not a reflection of differences in the risk of mortality from cervical cancer. They are probably caused by differences in the process of decision making, e.g. differences in the criteria, in the methods used, and in the influence of interest groups.

Apart from programme screening (i.e. with a personal invitation to have a PAP smear at set ages), there is also a large amount of opportunistic screening taking place. This occurs when the opportunity for taking a PAP smear is at hand, or when the woman and/or her physician think it is time to take a smear. In general, opportunistic screening tends to start earlier in life, to concentrate more in young age, and to be applied more frequently than programme screening [100, 164, 85]. The proportion of women screened (coverage), a crucial parameter for the public health impact of screening, is lower in a situation where only opportunistic screening exists [77, 5]. In situations where an invitational programme is in operation, opportunistic screening usually runs in parallel. Opportunistic screening reaches those women missed by programme

screening, but especially adds to screening outside (particularly under) the target age-range and increases the frequency of screening within the target age-range. Opportunistic screening has often been criticised for its lack of efficiency and the amount of unfavourable effects that it generates. But some investigators have been more supportive. [85, 22].

1.7 The follow-up of abnormal pap smears

As far as follow-up is concerned, a PAP smear has three possible outcomes: negative (i.e., no particular follow-up is advised and the normal screening schedule is to be followed), borderline (i.e., repeat smears are advised) and positive (i.e. referral to the gynaecologist-colposcopist is required). How to define normal, borderline and positive abnormalities has always been and still is highly debated. Between the sixties and nineties, there has been an ongoing shift towards more intensive follow up *(see Table 10.3)*. This trend has led to the present practice of very intensive follow-up. The tension between the stress on physicians (sometimes manifested in the form of lawsuits) to do 'everything possible' to prevent a potentially life threatening disease, and the desire for a medical intervention with as little unnecessary harm as possible, is very prominent in cervical cancer screening. In follow-up too the balance between benefits, harms and costs should be of major concern.

1.8 Screening in the Netherlands: a historical overview

The work presented in this thesis has been performed as part of the evaluation of programme screening in the eighties and nineties in the Netherlands. Cervical cancer screening was introduced in the late sixties in the Netherlands. After some small scale feasibility projects in the early seventies, a pilot project for programme screening was started in 1976. It was centrally organized and financed by the Government, and took place in three pilot regions: Nijmegen, Rotterdam and Utrecht, covering 24% of the Dutch female population. Women from age 35 to 53 were invited every 3 years to attend screening. The PAP smears were taken by specially trained women at local public health care units. This programme was meant to serve as a regional trial and was planned to compare mortality trends in the screening regions with trends in the other parts of the country. Under political pressure however, the screening programme was soon extended to other regions, reaching almost nationwide coverage around 1980. Meanwhile, gynaecologists and general practitioners increasingly took smears, particularly in women who were too young for programme screening. The total number of smears increased from an estimated

Table 1.2 *Important factors influencing the effects and costs of cervical cancer*

- Incidence of disease by age and year of birth
- Regression rate of pre-invasive disease by age
- Duration and its variability of preclinical disease
- Test characteristics (sensitivity and specificity)
- Prognosis for the situation without screening and the improvement in prognosis resulting from early detection
- Attendance pattern and relationship between attendance and risk
- Demography, deaths from causes other than cervical cancer, hysterectomies for reasons other than cervical cancer

100 000 in 1970 to 500 000 in 1975 and 1 million in 1982 [7]. In 1982, the Government policy changed and programme screening directly financed by the government was stopped. By 1985, organized screening had ended in most parts of the country, and the number of smears also dropped.

During the late eighties programme screening was started again, covering 85% of the municipalities in 1990 [123]. The number of smears was 750 000 in 1987 and 940 000 in 1990 after which it remained at about the same level [14]. Programme screening during this period was strongly decentralized, the municipalities were responsible for inviting women and for monitoring and evaluating the programme. Women were invited to make an appointment with their general practitioner to have a smear. The costs for the municipalities had to be covered by general preventive funds (Wet Collectieve Voorzieningen). The screening age range and intervals remained unchanged. This decentral way of conducting a cervical cancer screening programme was strongly criticized. Of particular concern was the quality of the smears being taken by the general practitioners, the coverage of the programme and the amount of inefficient screening (e.g. in too young women and too frequently).

In 1990, the Ministry of Public Health instructed that a study be undertaken to look at where the 'bottlenecks' of the (decentrally organized) cervical cancer screening programme were occurring. An important conclusion of this study was that due to a lack of directive powers and earmarked financing, effect-evaluation, and hence the opportunity for adjustments, was insufficient [123]. The Government decided that programme screening should be continued, but only if it could be better organized [192]. The Dutch Health Insurance Council was requested to advise the Government on an improved way of organizing and financing the programme. A total reconsideration followed, including the screening age range and interval. It resulted in a newly organized screening programme, which was implemented nationwide in 1996. The age-range was extended from 35–53 to 30–60 years, and the interval from 3 to 5 years, leaving the number of seven invitations a life-time unchanged. The choice of age-range

and interval was amongst others based on the work presented and discussed in this thesis, in particular on the results presented in Chapter 7. According to the new recommendations, the follow-up for borderline abnormalities was intensified, so that more women will be referred for colposcipical/histological evaluation. This decision to intensify follow-up was not evidence-based (see chapter 10). On the other hand, it was decided that the percentage of borderline smear ('Pap 2') should decrease. The organization of this new programme remains largely decentralized. However, the Government has taken (financial) responsibility for quality control and for evaluation of effects and costs of the programme. Regional organizations are responsible for reporting the data that are necessary for quality control and evaluation. These data are arranged in a set of evaluation tables which are compatible with those issued by the Europe Against Cancer Programme for quality assurance in cervical cancer screening [45]. One of the objectives of this reorganisation and of the evaluation system is that the costs of improving the screening programme should be compensated by the decrease in opportunistic PAP smears outside the age-range and frequency guidelines.

1.9 New developments

New developments in screening need thorough evaluation before implementation in a mass-screening can be considered. The first important new development explored in this thesis is computer aided cyto-morphologic evaluation of PAP smears, as a possible method of improving sensitivity and/or specificity [122]. Computer aided screeing is expected to affect the costs of the laboratory evaluation, a substantial part of the total costs of cervical cancer screening [118]. The second development that will be addressed is the increasing knowledge of the relationship between HPV (human papillomavirus) infections and the development of cervical cancer. HPV is found in more than 90% of invasive cervical cancers [26], while the observed prevalence of oncogene HPV types in the general female population over age 30 years, at least in developed countries, is around 3–6% [15, 149, 53, 180]. The first longitudinal studies seem to confirm that HPV infections are a strong risk factor for the future development of cervical cancer. Hence, HPV detection possibly could be useful in cervical cancer screening. And if eventually an effective preventive vaccine against HPV infections is developed [201, 73], primary prevention of cervical cancer could be an efficient alternative, gradually substituting cervical cancer screening.

1.10 The objective of this thesis

The objective of this thesis is to evaluate cervical cancer screening by assessing its various effects and costs. Questions to be addressed are:

1. Is cervical cancer screening effective in reducing mortality? And if so, how large are the beneficial effects (reduction in incidence and mortality)?
2. How large are the unfavourable health effects of cervical cancer screening? Are they outweighed by the beneficial health effects?
3. Under what condition is cervical cancer screening cost-effective?
4. What screening age range and intervals should be chosen from the viewpoint of cost-effectiveness?
5. What follow-up strategy after non-negative smears should be chosen from the viewpoint of cost-effectiveness?
6. Should cervical cancer screening be continued? And what changes should be recommended in order to improve its efficiency?
7. Should new techniques, like automated cytological evaluation of cervical smears or HPV-detection in cervical scrapes, be added to or replace conventional PAP smear screening?

In Chapter 2–10 the findings regarding these questions will be reported and discussed. The conclusions for each of the seven questions will be presented in Chapter 11.

1.11 The use of a model

Prediction of the effects and costs of cervical cancer screening for different screening strategies is a complex matter that involves many factors *(see Table 1.2)*. Therefore, a computerized simulation model is indispensable. The work presented in this thesis is based on the use of a cervical cancer screening version of MISCAN, a micro-simulation programme that was designed for the evaluation of cancer screening [87, 118]. This model has also been used for evaluation of breast cancer screening [117]. In the model, all the factors listed in *Table 1.2* are described and quantified (see Chapter 6). *Table 1.3* gives an overview of the data used for quantification and validation. After this quantification and validation ([159, 160, 161] and Chapter 6) the model has been used to predict health effects and costs of cervical cancer screening policies.

Table 1.3 *Overview of the data used in estimating model parameter values. For details and references see also chapters 6 and 7*

- Screening data (detection and incidence rates by age, rank and interval) from British Columbia, Dutch pilot regions, and the IARC study
- Cervical cancer incidence, mortality and survival data
- Data on the screening pattern in the Dutch female population
- Cost data (concerning screening, diagnosis and treatment)
- Hysterectomy data
- Dutch national demographic data

1.12 Reading guidance

Evaluation is a cyclic process of analyzing new data, evaluation of costs and effects of screening, explorative analysis of new options, etc. This means that the more or less chronological order of the manuscripts included in this thesis does not always represent a logical ordering of steps in the evaluation of a screening programme.

A logical ordering of the chapters of this thesis is:

0. Analyses of the data for model quantification regarding the effects of screening
   — see the thesis of Gerrit van Oortmarssen [162]
1. Testing the model predictions against observed data
   — Chapter 6
2. Performing cost-analyses to assess the costs of screening and link the effects to costs and savings
   — Chapters 4 and 5
3. Predicting the benefits, the negative side-effects and the costs of different screening strategies
   — Chapters 2 (comparing efficient policies with different numbers of smears per woman), 3 (comparing opportunistic with efficient screening) and 7 (optimizing screening age-ranges and intervals)
4. Investigating the potential value of two new developments
   — Chapters 8 (computer-assisted PAP smear evaluation) and 9 (HPV-screening)

2

Diagnostic and treatment procedures induced by cervical cancer screening

This chapter is based on M. van Ballegooijen et al
Eur J Cancer, Vol.26, no.9, pp.941–945. 1990

Abstract

The amount of diagnostic and treatment procedures induced by cervical cancer screening has been assessed prospectively and related to mortality reduction. Assumptions are based on data from Dutch screening programmes and on a scenario for future developments. With 5 invitations for screening, between ages 37–70 every eight years, 13 deaths are avoided per million women per screening year. Each death avoided is balanced by 2800 preventive smears, 9 women referred to a gyneacology department and 4 minor treatment procedures (conserving treatment or exconisation). 25 invitations in a life-time avoids 27 deaths per million women per screening year but with per death avoided 7300 preventive smears, 22 referrals and 8 small treatment procedures. Thus intensifying screening will not only result in diminishing returns of extra screening efforts, but also in increasing risk for women to undergo unnecessary (no invasive disease or death avoided) diagnostic and treatment procedures. The balance between beneficial and adverse effects deteriorates strongly when hysterectomies play an important part in the management of cervical intraepithelial neoplasia.

Acknowledgement This study was financed by the Prevention Fund.

2.1 Introduction

The appropriateness of screening for cervical cancer should be balanced between beneficial health effects and adverse effects and costs. As the avoidance of death is the principal aim, most evaluations have concentrated on the relation between life-years saved and costs, which shows a diminishing return for the extra efforts involved in screening women more frequently [64, 132, 167]. Although

Table 2.1 *Primary treatment procedures found (a) in eight Dutch hospitals*, in 1982–1986, and (b) in the pilot region of Nijmegen in 1981–1984*

Treatment	<CIN III† (a)	<CIN III† (b)	CIN III (a)	CIN III (b)	IA (a)	IA (b)
None	ND	61%	—	—	—	—
Conserving treatment	ND	10%	25%	28%	—	—
Exconisation	ND	6%	64%	15%	38%	0%
Exconisation + hysterectomy	ND	—	4%	—	46%	50%
Hysterectomy	ND	23%	7%	48%	16%	50%
Unknown	ND	—	—	9%	—	—
Total (n) 100%		66	663	46	24	2

* Academisch Medisch Centrum (Amsterdam), Medisch Centrum Alkmaar (Alkmaar), Sint Franciscus Gasthuis Rotterdam (Rotterdam), Ziekenhuis Leyenburg (Den Haag), Ikazia Ziekenhuis Rotterdam (Rotterdam), RK Ziekenhuis Groningen (Groningen), Catherina Ziekenhuis Eindhoven (Eindhoven), Westeinde Ziekenhuis Den Haag (Den Haag). Annual reports 1982–1986

† Women referred to a gynaecology department for cervical cancer assessment, in whom no CIN or only CIN I/II is found

ND No data

unnecessary referrals and diagnostic and local therapeutic procedures are often discussed as the major adverse effect of cervical cancer screening, reports on this effect are scarce. Treatments of advanced cancer will be reduced by early detection. But due to the follow-up of false positive smears and of lesions that would have regressed spontaneously, there will be a considerable increase in the number of diagnostic and minor therapeutic procedures. We have assessed this increase and related it to the number of life-years gained.

2.2 Methods

2.2.1 Yield of early detection

The effects of cervical cancer screening have been predicted by a mathematical model [90, 88, 118]. The model has been validated by analysing cervical cancer screening data from the British Columbia Cohort Study [27] and from pilot cervical cancer screening projects (1976–1985) in the Netherlands [70]. Both analyses led to similar conclusions [88]. Model-based analysis of the Dutch data led to the following estimates: (a) mean duration of the pre-clinical stages, 17 years and shortest at old ages; (b) regression rate of pre-invasive disease, 60% on average and highest at young ages; and (c) sensitivity of cervical cytology, 70% for cervical intraepithelial neoplastic (CIN III). With other assumptions (e.g. shorter duration, less regression or higher sensitivity) the observed Dutch

screening results and incidence and mortality data could not be explained. The false positive rate, defined as the percentage of women in whom no CIN III or invasive cancer was found after an abnormal (at least moderate dysplasia) cervical smear, was 0.4%. This number has been used in our calculations.

The validated model was used to simulate different screening policies in Dutch women during the period 1988–2015. Effects of screening occurring after 2015 have been taken into account. The population dynamics including deaths by causes other than cervical cancer have been incorporated into the model. The impact of cervical cancer screening before 1988 and of hysterectomies performed for reasons other than cervical cancer have also been taken into account. The population attendance rate has been assumed to be 65% on average, as observed in the Dutch pilot regions. Attendance decreases gradually for women over 50, to 55% for women aged 70.

The model predicts the yearly number of smears, the yearly number of women diagnosed with CIN III, micro-invasive or cervical cancer by stage, the yearly number of women detected in whom no CIN III or worse is found during follow up and the yearly number of women dying from cervical cancer.

Effects were calculated compared with no early detection, which was also simulated and in which women are only detected by symptoms. In our model detection by symptoms occurs in stage IB at the earliest.

2.2.2 Diagnosis and treatment

Assessment of the number of women referred to a gyneacological department because of an abnormal cervical smear is of major importance when estimating the extent of diagnosis induced by screening. Referred women are supposed to undergo colposcopy at least once; most will have ectocervical biopsy and in some endocervical curettage will be done. We calculated the number of referred women from the predicted numbers of detected women. We assumed that all women with a histologically confirmed CIN III+ and 65% of the women followed up without a confirmed CIN III+ have been referred to a gynaecologist. The other 35% of the latter group is assumed to have had only repeat smears done by the general practitioner, as the data from the pilot regions indicate. In the pilot regions, a PAP smear with moderate dysplasia was followed by a repeat smear within a few weeks. A smear with severe dysplasia or more led to immediate referral to a gynaecologist.

To calculate the numbers of primary treatment procedures a schedule has been devised a realistic frequency of treatments by stage of disease. This schedule was applied on the predicted numbers of detected women by stage.

The possible treatment procedures in CIN and microinvasive cervical cancer are summarised in *Table 2.1.* Conserving treatments (cryocoagulation, electrocoagulation, laser-evaporation and diathermic ablation) and exconisations

Table 2.2 *Assumptions used in the prospective calculation of numbers of primary treatment procedures: percentages of primary treatment procedures by cervical cancer stage*

Treatment	<CIN III*	CIN III	IA	IB	Detected by Screening II+	Detected by Symptoms II+
None (absence of CIN III)	81%	—	—	—	—	—
Conserving	12%	24%	—	—	—	—
Conserving + exconisation†	1%	1%	—	—	—	—
Exconisation	6%	64%	38%	—	—	—
Econisation + hysterectomy	—	3%	62%	—	—	—
Hysterectomy	—	8%	0%	90%	40%	16%
Primary radiation	—	—	—	10%	60%	79%
No primary treatment (advanced disease)	—	—	—	—	—	5%

* Women referred to a gynaecology department for cervical cancer assessment, in whom no CIN or only CIN I/II is found

† Because of exconisations as a retreatment procedure assumed in 5% of the women who had conserving treatment

are considered as minor treatment procedures. *Table 2.1* also shows the treatment procedures used in eight Dutch hospitals of different size and in different regions during 1982–1986 and in the screening programme in the pilot region of Nijmegen during 1981–1984 (office of Evaluation and Registration of Cervical Cancer).

The difference between the two data sets stresses the variability in medical practice. Hysterectomies were more frequent in Nijmegen. The reasons for this were not investigated. Current developments towards less aggressive treatments are the further introduction of colposcopy in the assessment of suspected cervical abnormalities and the decreasing rates for hysterectomies in general (there has been a more than 30% fall in the age group 35–50 in the Netherlands since the late 1970s) [186]. Hospital departments that reported data on their treatment procedures are probably departments that follow these developments more closely than average. The percentage for less aggressive procedures in these hospitals, however, may be a good approximation for future practice. So we used the hospital data for the assumptions on the average stage-specific treatment *(see Table 2.2)*.

As no detailed data were available on treatment procedures in women with less than CIN III we based our assumptions for this group on reported information on interviews with expert gynaecologists. Of the women followed up in whom no CIN III or invasive disease is found 70% are women without any neoplasia and 30% are with CIN I/II (Office of Evaluation and Registration of Cervical

Cancer and [125]). We assumed no treatment in the first group. The percentages for conserving treatment and exconisations in the group with CIN I or CIN II are based on the assumption that about 60% of these women are treated and 40% are strictly followed up by cytology. From the treated group 30% must (at least) have exconisation because in about 30% of the women with CIN the transformation zone and/or the lesion cannot entirely be seen with colposcopy. So the percentage of conserving treatments in the women referred with less than CIN III is assumed to be 13% (30% × 60% × 70%) and the percentage of exconisations to be 6% (30% × 60% × 30% [rounded]). Moreover, 5% of the women who had conserving treatment are assumed to have exconisation as a re-treatment.

In the management of women with cervical cancer stage IA, at least an exconisation is assumed. In the Netherlands there is consensus to avoid diagnosing stage IA without doing exconisation first. The two possible primary treatment procedures in invasive cervical cancer stages IB or higher are radical surgery (with or without additional radiotherapy) and radiotherapy (alone). Data were studied from the eight hospitals, the Nijmegen region and from several other Dutch studies [107, 202]. All results were similar, and since no changes are expected in the near future, they were adopted. The difference in *Table 2.2* in the assumptions between stage II+ detected by early detection and stage II+ detected by symptoms is based on the more favourable distribution over the stages IIA, IIB, III and IV in women detected by screening compared with these detected by symptoms [69].

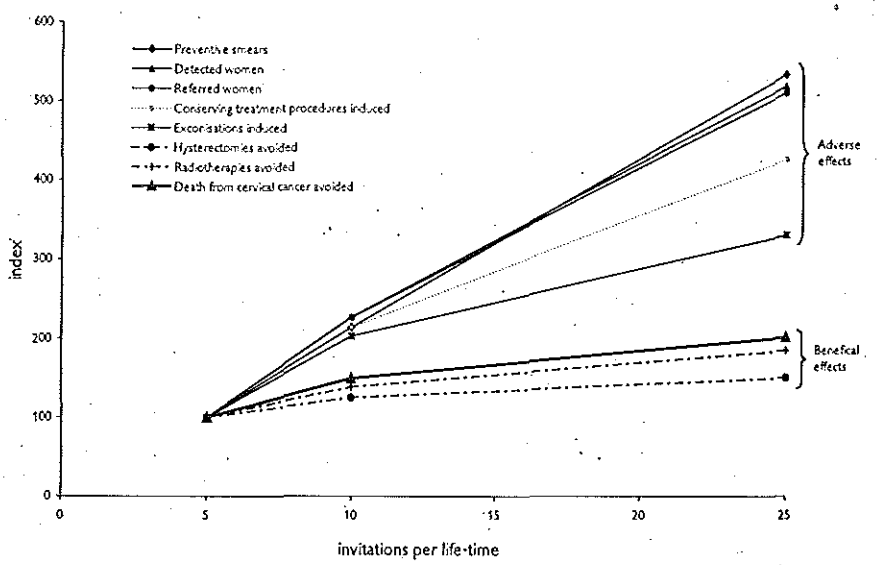
2.2.3 Screening policies

We will discuss the results for three screening policies, with an increasing number of invitations for screening during a life-time: 5, 10 and 25. More invitations implies a shorter interval between successive invitations. These intervals are 8, 5 and 2 years. The age-ranges are, respectively, 37–70, 27–72 and 26–74 years. This is a broad age range that does not correspond with most current practice. All three are so-called efficient screening policies. Policies are efficient when their distribution of screenings over the ages leads to the largest number of life-years gained at a certain level of costs, or conversely to the least costs for a certain number of life-years gained. Our study on the costs of cervical cancer screening is presented elsewhere [118].

2.3 Results

The stage distribution, number of smears, number of women followed up and mortality from cervical cancer are shown in *Table 2.3* for the three detection strategies. All results are presented per screening year, which means that

Figure 2.1 *Relationship between various beneficial and adverse effects of screening and intensity of screening, as represented by efficient cervical screening policies with 5, 10 and 25 invitations. Level of effects at 5 invitations is set to 100 (all lines go through origin)*



the total outcomes and effects of screening performed in 1988–2015 have been divided by the number of screening years: 27 (i.e. 'by year').

With more intensive screening a shift occurs in the detected precancers and cancers towards earlier stages. The total number of women followed-up increases significantly: from 349 (5 invitations) to 1044 (25 invitations) per million women per year, mainly due to the increase in follow-up of women with no histological abnormalities or only CIN I or CIN II. With intensification of screening from 5 to 10 invitations, 184 additional life-years are gained, while a further increase from 10 to 25 invitations adds only 134 life-years more per million women per year.

Intensifying the policy from 5 to 25 invitations per life-time causes a five-fold increase in the number of women referred *(see Table 2.4)*. This induces additional diagnostic and treatment procedures. More specifically, an increase from 5 to 25 invitations makes the number of conserving treatments rise from 25 to 106 and exconisations rise from 47 to 157 per million women per year.

The effect on the number of hysterectomies (total plus radical) of intensifying early detection efforts is modest. The reason is that the rise in the number of total hysterectomies caused by the increase in the number of women detected with CIN III *(see Table 2.3)* is neutralised by a decrease in radical hysterec-

Table 2.3 *Outcomes of three efficient cervical cancer early detection strategies. Prospective calculations. Numbers are per year per million women (all ages). Results without early detection are given in comparison*

	Number of invitations*			No early detection
	5	10	25	
< CIN III†	140	299	745	0
CIN III	53	113	174	0
IA	5	6	5	0
IB	49	42	36	57
II+scr	1	1	1	0
II+sympt	101	92	82	124
Total	349	554	1 044	182
Number of smears	36 800	78 700	196 100	0
Deaths from cervical cancer	67	60	54	80
(difference)‡	(-13)	(-20)	(-27)	
Lifeyears lost	1172	989	854	1434
(difference)‡	(-262)	(-446)	(-580)	

* Per woman during her lifetime
† Women with abnormal smears in whom no CIN or only CIN I/II is found
‡ Difference with the situation with no early detection

tomies caused by the fall in the number of women diagnosed with cervical cancer stage IB or higher. The number of primary radiotherapies falls as a result of the decrease in the number of diagnosed stages IB and higher.

Tables 2.3 and *2.4* can be used to calculate the balance between the efforts and risks needed to prevent more deaths from cervical cancers by more intensive screening. With 5 invitations, for every death avoided, 2800 preventive PAP smears, 9 referrals and 4 minor treatment procedures are needed. With 25 invitations, the corresponding figures are 7300, 22 and 8.

One reason for the deteriorating balance between beneficial and adverse effects is the strongly increased detection of women with no abnormalities or with lesions which are less advanced than CIN III *(see Table 2.3)*. The second reason is the increased detection of women with CIN III that would have regressed spontaneously. According to our predictions, this number would go from 32 to 77 to 121 per million women per year when intensifying the policy from 5 to 10 to 25 invitations, respectively. The resulting additional treatments, a negative side-effect inherent to early detection of cervical cancer, do not result in avoided invasive disease or life-years gained. Adverse effects all increase at last two to three times as fast as the beneficial effects when intensifying the early detection policy *(see Figure 2.1)*.

Overall the diagnostic and treatment procedures induced by cervical cancer screening precede the mortality reduction by many years. Discounting of

Table 2.4 *Effect on diagnostic and primary treatment procedures of three efficient cervical cancer early detection strategies (prospective calculations). Per year per million women (all ages), difference compared with no screening*

	Number of invitations*			No early
	5	10	25	detection
Women referred	+117	+266	+598	182
Conserving treatments	+25	+53	+106	0
Exconisations	+47	+96	+157	0
Hysterectomies	-2	-2	-3	72
Primary radiotherapies	-19	-26	-35	104
Total treatments	+52	+121	+226	175

* Per woman during her lifetime

effects (when not considering the costs) is disputable. Nevertheless, when all effects are discounted with a rate of 5%, the ratios between unfavourable and favourable effects rise between 160% and 200%. There is not only a diminishing return for the extra efforts involved in screening women more frequently, but there is also an increasing risk for women to be referred and treated without having the benefit of invasive disease being presented.

2.3.1 More aggressive early treatment

The results presented are very sensitive to the treatment procedures assumed in pre-invasive and micro-invasive stages, because of the large number of women concerned. Therefore we also made calculations for a more aggressive schedule, with a shift from minor treatment procedures to hysterectomies and less untreated women *(see Table 2.5)*. The results *(see Table 2.6)* should be compared with those of *Table 2.4*. There is a sharp increase in the number of hysterectomies: for a limited early detection policy (5 invitations), a net decrease of 2 hysterectomies has now turned into a net increase of 48. For more intensive screening policies, this number increases nearly linearly, contrary to the number of deaths avoided. This is an important finding, as intensive policies are often advocated, for instance by the National Cancer Institute [152].

The beneficial effects of screening are unaffected by the new assumptions, as survival is reported to remain excellent when more local treatment procedures are used in CIN and cervical cancer stage IA, provided the right selection criteria are used [72, 106, 175].

The numbers and types of treatments without early detection remain unchanged, as no women with pre-invasive or micro-invasive disease are detected by symptoms in our calculations.

Table 2.5 *More aggressive treatment schedule in* CIN *and micro-invasive cervical cancer (percentages of primary treatment procedures by stage)*

Treatment	< CIN III*	CIN III	IA
None (absence of CIN III)	55%	—	—
Conserving treatment	10%	20%	—
Exconisation	10%	20%	5%
Exconisation + hysterectomy	—	—	45%
Hysterectomy	25%	60%	50%

* Women referred to a gynaecology department for cervical cancer assessment, in whom no CIN or only CIN I/II is found

2.4 Discussion

Complete data on treatments used in women with CIN I and CIN II are scarce. About 10% of women with CIN III in British clinics [78, 170, 193] have total hysterectomies, and 60–70% have conserving treatment and 20–30% have exconisations. This is the opposite of what is indicated by Dutch data. The possibility of a quality-bias in the British data, cannot be ruled out. But if the average treatment schedule was as suggested by the British data, the numbers of conserving treatments and exconisations in *Table 2.4* should be interchanged. The total level of adverse effects would be lower, but the pattern of increasing risks for women with increasingly intensive policies would remain unchanged.

Results have been calculated for an average duration of preclinical disease of 17 years, a regression rate of 60% and a sensitivity of 70%. The calculations have been repeated for other values for duration, regression and sensitivity, which could still explain the results of screening in Canada and the Netherlands. The results were not affected substantially. In a sensitivity analysis we also studied the influence of possible changes in incidence and natural history on our results. In case of a higher cervical cancer incidence (in the absence of screening) in the future, all beneficial effects and some adverse effects increase proportionally. A smaller (or larger) number of spontaneously regressing lesions makes adverse effects decrease (or increase), leaving the number of deaths avoided unaltered.

In efficient policies screening starts at about age 30. It is often advocated to start screening much younger, perhaps around age 20 [51][162]. Indeed, present screening by general practitioners and gynaecologists concentrates between the ages 20 and 35. Our calculations indicate that such early detection policies lead to many more diagnostic and treatment procedures per death avoided. For instance a screening policy that invites women every year between the age of 20 and 35, and every 5 years between 35 and 60 [51] has the following effects. Per death avoided, 11 800 smears, 37 referrals and 16 minor treatment procedures are needed. The number of deaths avoided, however, is only 17 per million women per screening year. We found that the ratio between beneficial

Table 2.6 *Prospective calculations with a more aggressive treatment schedule: effects on the quantity of diagnosis and primary treatment resulting from three efficient early detection policies (differences with the situation with no screening), and from the situation without screening (absolute numbers), per million women, per screening year*

	Number of invitations*			No early
	5	10	25	detection
Women referred	+117	+266	+598	182
Conserving treatments	+20	+42	+83	—
Exconisations	+20	+45	+86	—
Hysterectomies	+48	+103	+204	72
Primary radiotherapies	-19	-26	-35	104
Total treatments	+71	+164	+338	175

* Per woman during her life-time

and adverse effects is dramatically unfavourable compared with those from efficient policies. This is mainly caused by the regression rate of CIN III lesions being especially high in young women and if screening starts from age 30 or 35, most of the women with progressive CIN III at early age are still detected in this pre-invasive stage, the mean duration of progressive CIN III being about 15 years.

Our predictions were made for the situation in which follow-up is advised only after a smear with at least a cytologically moderate dysplasia. Colposcopy after smears with mild dysplasia (or even less severe abnormalities) is also advocated. For quantification of the effects of such a strategy, too little is known about the incidence, regression rate and duration of CIN I and CIN II, and about the sensitivity of the PAP smear for these abnormalities. Furthermore it is uncertain which protocols for treatment will be used in the future, and how medical practice will be. But the number of referrals and diagnostic and conserving treatment procedures per radical treatment or death avoided will undoubtedly increase. The reasons for this are a relative increase in the number of false positive results (the predictive value will be lower) and the larger number of treated women with regressive lesions, making the plausible assumption that regression is also possible from CIN I and CIN II.

Because we deal with symptom-free women, there is a close relation between balancing risks and benefits in individual women, and decision making in health care. What counts in both views is an acceptable ratio between beneficial and adverse effects of early detection of cervical cancer. The question is how many referrals for diagnosis, conserving treatment procedures, exconisations and total hysterectomies are acceptable per radical treatment or death avoided. Our aim was not to provide an absolute answer to this question, but to provide the material required to compare different policies on these criteria.

3

Preventive pap-smears: balancing costs, risks and benefits

This chapter is based on M. van Ballegooijen et al
Br J Cancer (1992), 65, 930–933

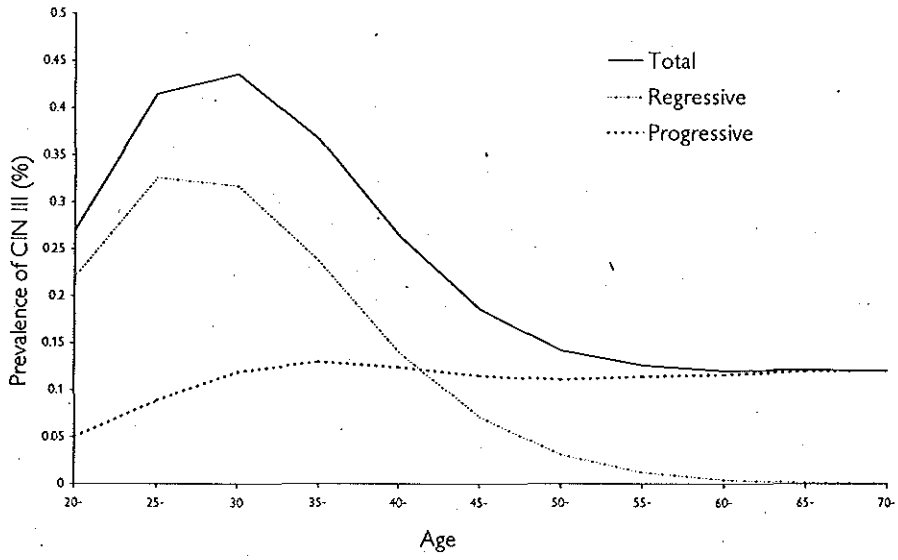
Abstract

The pattern of spontaneous screening for cervical cancer by general practitioners and gynaecologists in the Netherlands is compared with an efficient screening policy resulting from a cost-effective study. Spontaneous screening tends to start and stop too early in a woman's life, and leaves too many women overscreened or unprotected. The combination in young age of a low incidence of invasive cancer and a high incidence of regressive lesions explains relative ineffectiveness and harmfulness of present screening practice. When screening would take place between ages 30 and at least 60, with intervals of about 5 years, as many lives could be saved for half the costs and with only 60% of the unnecessary referrals and treatments. Much attention should be paid to the coverage of the target population. Therapeutic follow-up policies for dysplastic lesions should be restrained.

3.1 Introduction

Screening has contributed to the decrease in cervical cancer mortality in several countries [58, 92, 126, 56, 81]. There is still debate on the age to start screening and on the interval. Some screening recommendations call for intensive screening at a young age [2, 51] but studies which analyse the health effects of screening conclude that screening efforts should be directed to middle aged and older women [126, 112, 143, 57, 167]. The advocated interval has been lengthening the last few years but in practice the interval tends to be still short.

Figure 3.1 *Age-specific prevalence of* CIN III *(histologically confirmed severe dysplasia or carcinoma in situ) in the unscreened population. Estimates which are based on observed data from cervical cancer screening programmes in the Netherlands (see text). Speculative under age 30 (few data available)*



3.2 Results

The pros and cons of screening policies critically depend on the duration and detectability of the preclinical stages of the disease. Knowledge of these important parameters can be derived from the results of existing screening programmes. Therefore, a detailed analysis was made of data from the early detection programmes in British Columbia and in the Netherlands. Both analyses led to very similar conclusions [88]. The first one has been published recently in this journal [169].

In this article we study the consequences of the results on duration and regression for balanced PAP smear taking. We compare spontaneous screening with optimised screening, studying the costs, risks and benefits.

3.3 Methods and materials

3.3.1 The natural history

For the Netherlands, the following estimates were derived:

- a smear will detect 70% of the cases of CIN III (cervical intraepithelial neoplasia) (sensitivity, that pertains to the situation in which women with at

Table 3.1 *Results: number of smears and the major effects of two different approaches to cervical cancer screening. All numbers are per million women per year*

Screening patterns	Smears[c]	Life-years gained	Deaths avoided	Women referred	Unnecessarily treated women[d]
Spontaneous[a]	120 000	400	14	370	135
Efficient[b]	65 000	400	18	210	80

[a] Spontaneous screening pattern by general practitioners and gynaecologists
[b] Efficient pattern, age 33 to 68, every 5 years, attendance 65%
[c] See *Figure 3.2* for the age distribution of the smears
[d] At least local treatment (e.g. cryocoagulation or laser-evaporation)

least (cytologically) moderate dysplasia twice or severe dysplasia once are referred for colposcopy)

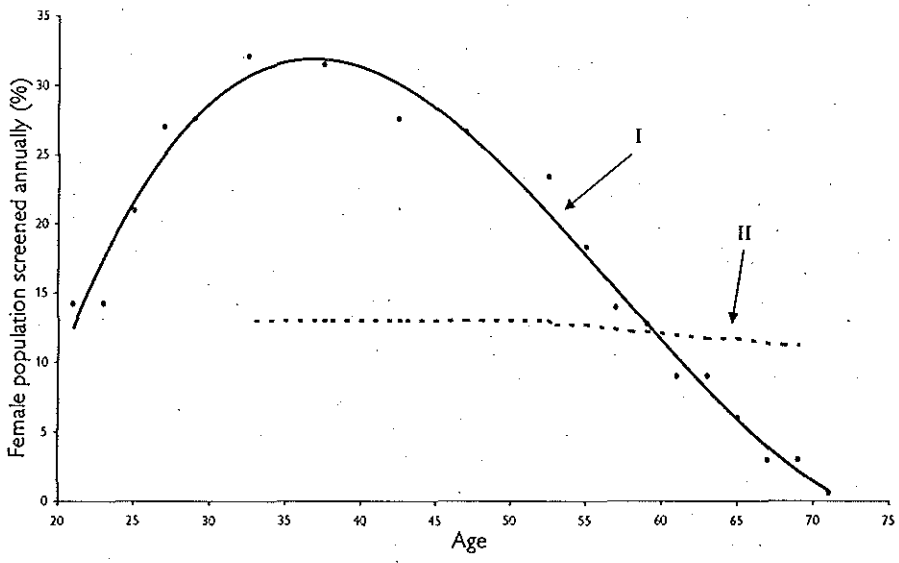
- 0.4% of the smears will be false-positive (no CIN, or at the most CIN II will be found histologically)
- the mean duration of CIN III is 15 years
- on average 60% of the cases of CIN III will regress spontaneously, this percentage is highest at younger age *(see Figure 3.1)*
- a higher incidence of cervical cancer in non-attenders to screening than in attenders

3.3.2 Predictive calculations

The assumptions on natural history have been implemented in a computerised epidemiometric model, which uses also assumptions on demography, age- specific incidence and stage-specific survival (see Habbema et al [89], for a full description of the model). Screening policies were assumed to be operational in the Netherlands in the period 1988–2015. Health effects and changes in number of women referred and treated after the termination of the programme have also been taken into account.

Outcomes are effectiveness (number of life years gained), costs (number of screenings) and risks (the number of women unnecessarily referred and treated because of false positive test results or regressive lesions). All these results have been calculated as differences with the (hypothetical) situation in which there is no early detection of cervical cancer. As we emphasise the ratio between positive and negative effects, for which discounting is disputable, undiscounted results are presented. The comparison between different policies is only very little affected by discounting.

Figure 3.2 *Two screening patterns: annually percentage of the female population screened by age. I. Spontaneous screening pattern by general practitioners and gynaecologists (see text). II. Efficient screening pattern (see text): age 33 to 68 every 5 years, attendance 65%*



3.3.3 Spontaneous screening

Spontaneous screening has been defined as screening in the situation without any invitational programme, resulting from the existing diversity of initiatives among the women and the doctors involved. We studied data on screening by general practitioners and gynaecologists in the Netherlands during the period 1985–1988, during which there were almost no invitational screening programmes running. We found *(see Figure 3.2)* that it starts at very young ages, declines in intensity after age 35 and stops nearly entirely at age 55–60. Population coverage is rather poor at older ages. This pattern corresponds with reports of other European and North American countries [110, 94, 41, 3, 165]. Detailed data on individual screening patterns in spontaneous screening were not available. We assumed that 50% of the screened women have a smear every 2 years, the others being screened less often. The spontaneous screening pattern was incorporated in our model and the costs, risks and benefits were calculated.

3.3.4 Efficient and spontaneous screening compared

We identified the efficient (with the lowest costs) screening policy with 65% attendance that results in the same number of life-years gained as the spon-

taneous screening pattern described. We assumed a 65% attendance level (percentage of the women screened) because this was reached in centrally organised screening with a population based invitation system in Dutch pilot regions [71]. The efficient policy differs from spontaneous screening in four ways *(see Figure 3.2)*:

- there is no screening in very young women: starting age is 33 years;
- women are screened until later in life: ending age is 68 years;
- the interval is longer: 5 years;
- coverage is higher, especially in older women.

Costs, risks and benefits of both screening patterns are presented in *Table 3.1*. The efficient policy requires half the number of smears to reach the same number of life-years gained as spontaneous screening, and the adverse effects will be cut down by more than 40%. In order to explore the reasons for these large differences in risks and benefits, we will now have a detailed look at the four characteristics of efficient screening mentioned.

3.3.5 Screening at a young age

The isolated effect of screening at young age vs screening later in life is demonstrated for the case of a single screening *(see Table 3.2)*. With a single invitation at age 40, the number of women unnecessarily referred for CIN III or lesser abnormalities and unnecessarily treated for each death avoided are seven and five times lower than with a single screening invitation at age 20. The chance that a first screened woman has a CIN III is highest at young age (continuous line in *Figure 3.1*). As women with diagnosed CIN III are nearly always treated, regression (discontinuous line in *Figure 3.1*) can not be observed.

The long duration of progressive CIN III (about 15 years on average) results in timely detection in the large majority of the cases when screening starts at age 30. Thus, only a few deaths will be avoided by additional screening under 30 years, at the expense of a very large number of screenings and a considerable risk of treatment of regressive lesions.

We basically assumed a stable incidence of cervical cancer for the birth cohorts from 1948 onwards. Even when we assumed an increase in the incidence for women born after 1960 with 50%, the starting age of the efficient policies still did not fall much under 30.

3.3.6 Screening in old age

To study the difference in results with and without screening women between 50 and 70, we compared two screening policies that both start at age 33, the one (already presented in *Table 3.1*) ending at age 68, the other at age 51 *(see Table 3.2)*. The latter policy is certainly not efficient: 15% more life-years can

Table 3.2 *Results: number of smears and the major effects of different cervical cancer screening patterns. All numbers are per million women per year*

Screening patterns	Smears	Life-years gained	Deaths avoided	Women referred	Unnecessarily treated women[e]
Young ages[a]					
1 smear at 20	9 000	20	0.4	30	10
1 smear at 40	10 500	110	4	40	20
Old ages[b]					
until age 68	65 000	400	18	210	80
until age 51	67 500	340	12	220	90
Intervals[c]					
every 8 years	37 000	260	13	120	45
every 2 years	196 000	580	27	600	210
Attendance[d]					
100%, 5×	51 000	450	23	170	65
50%, 25×	129 000	440	20	400	140

[a] Single screening at age 20, attendance 75% respectively single screening at age 40, attendance 75%

[b] Efficient pattern, age 33 to 68, every 5 years, attendance 65% respectively screening from age 33 to 51, every 3 years, attendance 65%

[c] Efficient pattern, age 39 to 71, every 8 years, attendance 65% respectively efficient pattern, age 26 to 74, every 2 years, attendance 65%

[d] Efficient pattern, age 39 to 71, every 8 years, attendance 100% respectively efficient pattern, age 26 to 74, every years, attendance 50%

[e] At least local treatment (e.g. cryocoagulation or laser-evaporation)

be gained with even less (5%) screenings when the policy is extended to the age-group 51–68 by increasing the interval from 3 to 5 years.

Is the chance that a women will develop cervical cancer later negligible when she reached the age of 50 without developing a precursor of cervical cancer? When this would be true, the high death rate in old age could only be caused by poor screening under 50 years. Available epidemiologic data suggest otherwise. The detection rate for preinvasive plus invasive cancer in women who were first screened between 50 and 55 years in Nijmegen and Utrecht [46] was 4.1–7.6 per 1000. This is clearly less than the cumulative incidence of invasive cancer of 11.8 per 1000 women of age 55–84 in 1975, i.e. before screening became widespread [197]. The gap between detection rate and cumulative incidence can only partly be explained by a sensitivity of the PAP smear of e.g. 70%.

The poor screening history in women over age 50 is in itself reason enough to screen until at least age 65 during the forthcoming decade [75, 148]. Meanwhile, new evidence could be collected on incidence in older women and on the need for further screening in women who received adequate screening until age 50–55.

3.3.7 The interval between successive screenings

The effect of screening frequency is quantified by comparing intervals of 2 and 8 years *(see Table 3.2)*. With an interval of 8 years 2800 smears are needed per death avoided. With an interval of 2 years this number rises to 7300 smears. The reason is that the chance of getting invasive cancer decreases substantially by a screening in the previous 2–3 years *(see Figure 3.3)*. As pointed out in the report of the IARC working group [58], this decrease can be seen in data from screening programs even 10 years after a negative screening. This is not surprising with a mean duration of CIN III of 15 years.

The balance between risks and benefits also gets worse. With an interval of 8 years, nine women are referred and three women are treated per death avoided. With an interval of 2 years, these numbers increase to 22 women referred and eight women treated.

3.3.8 The coverage of the target population

As shown in *Table 3.2*, cervical cancer mortality would be lower when all women would have a PAP smear five times in their life, than when 50% of the women would be screened 25 times.

Most cases of invasive cervical cancer nowadays occur in unscreened or poorly screened women [128]. Incidence in non-attenders appears to be higher than in the total population. This conclusion of our analysis of the Canadian and Dutch screening data is supported by data from Denmark and Norway [18, 134]. A further reduction in mortality can primarily be achieved by screening the as yet unscreened women. The use of a shorter screening interval would mainly result in a more frequent screening of those who are already being screened.

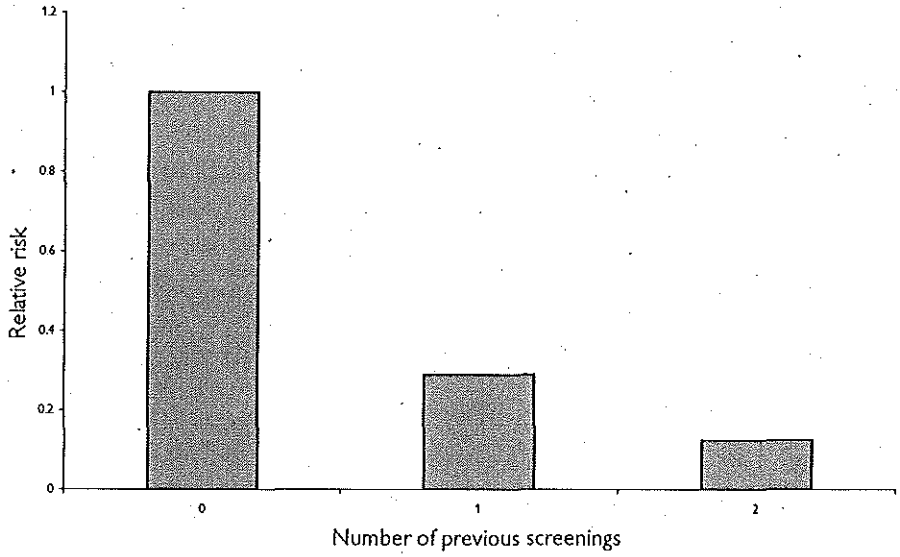
3.4 Discussion

A comparable study has been performed by Eddy [65]. Although his outcomes show a very small difference in effectiveness when lengthening the interval from 1 to 4 years, he surprisingly recommends screening at least every 3 years. Eddy recommends to start screening in the early 20s, without studying adverse effects and assuming an age-independent regression rate. In our view high regression rates at young age cause extra risks of screening for young women.

3.4.1 Follow-up and treatment

Cervical cancer screening will always induce unnecessary treatment, because of the partly regressive nature of CIN. The seriousness of this adverse effect

Figure 3.3 *Relative risk of invasive cervical cancer in screened women with a most recent screening 2–3 years ago compared to unscreened women. Calculated from Day* [58]



depends on the treatment applied. We found that in some Dutch gynaecological centers nearly 50% of the women with CIN III were treated with hysterectomy and in other centers 10% [9]. From the USA, hysterectomy rates in women with cervical carcinoma in situ are reported to be 50% [79]. In a screening programme with excellent gynaecological follow-up, the number of hysterectomies for cervical cancer in the population should fall because of the decreasing number of invasive cancers. But with an excessively aggressive treatment of preinvasive lesions, the number of hysterectomies can increase 3-fold when an intensive screening programme is carried out.

3.5 Conclusions

Our analysis clearly shows the consequences of screening efforts still starting and stopping too early in life, and being performed too frequently. The importance of a high coverage cannot be overemphasised.

4

The management of cervical intraepithelial neoplasia: extensiveness and costs in the Netherlands

This chapter is based on M. van Ballegooijen et al
European Journal of Cancer, Vol.31A, No. 10, pp. 1672–1676, 1995.

Abstract

In order to provide greater insight into both the extensiveness and the medical costs of the diagnosis and treatment of screen-detected cervical intraepithelial neoplasia (CIN) in general medical practice in the Netherlands, data from national registries and gynaecology departments were retrieved, and experts were interviewed. Of the 5060 women diagnosed with CIN in 1988, more than 50% were treated in hospital with conisation or hysterectomy, which on average took 5.5 days stay per admission. The assessed average duration of the total pre- and posttreatment period is 4.6 years. The average total medical costs in women with detected CIN III are Dfl 3700 per woman. The diagnosis of CIN I and II involves more medical procedures and time than CIN III, but fewer women have conisation or hysterectomy, resulting in lower total medical costs (Dfl 2572). The overall extent and costs of the management of CIN should be accounted for when balancing the benefits, unfavourable effects and costs of cervical cancer screening.

Acknowledgement This study was supported by the Ministry of Welfare, Public Health and Cultural Affairs of the Netherlands, and by the Dutch Health Insurance Council.

4.1 Introduction

The aim of this study was to provide greater insight into the extent and the medical costs of the diagnosis and treatment of screen-detected cervical intra- epithelial neoplasia (CIN) in general medical practice. Such knowledge

is most important for assessing the balance between costs and favourable and unfavourable effects of screening for cervical cancer. Part of the unfavourable effects is caused by detection of non-progressive lesions. The number of women involved is increasing because the cut-off point in cytology for colposcopic follow-up has been shifting to lower grade abnormalities. It is often stressed that the management of CIN is quick, safe and cheap. Even if this were true in an ideal situation, such a statement should be verified for general medical practice. Available literature refers to data from one or two selected treatment centres per study [105, 78, 193]. We studied the number of the various diagnostic and treatment procedures from national data, and when this was not possible, from a number of gynaecological departments covering 25% of the population. Data from the Dutch national hospital admission registration (coverage over 99%) were combined with centrally collected data from at least 85% of the cytopathology laboratories (PALGA, Dutch Network and National Database for Pathology). Currently the Netherlands is one of the few countries for which these data are available on a national scale.

4.2 Methods and materials

The study concerns women with histologically confirmed CIN: the episode studied starts with the first visit to the gynaecology department for follow-up after an abnormal PAP smear and ends after the woman is referred back to normal screening practice. The situation for women having invasive cervical cancer has been considered elsewhere [90]. This study concentrates on pre-invasive disease, cervical intra-epithelial neoplasia: CIN I/II and III.

The annual number of women in whom CIN I, II and III has been diagnosed in the Netherlands was retrieved from PALGA, which is connected to the cytopathology laboratories. All women with histologically confirmed CIN in the period 1987–1990 were selected [163]. For each of these women, the maximal diagnosis within this period was established. The population coverage rate of PALGA was 70%, 85%, 98% and 99.6% in 1987, 1988, 1989 and 1990 respectively. The annual total numbers (corrected for incomplete coverage) of women detected with CIN I/II and CIN III for 1988 are 2590 and 2570, respectively, and for 1990 3410 and 3420, respectively.

The annual number of conisations (2180), hysterectomies (425) and days of hospitalization (15780) associated with the diagnosis and treatment of CIN *(see Table 4.1)*, were retrieved from national data on hospital admissions, selecting admissions with dysplasia or carcinoma in situ of the cervix uteri (ICD-9 codes 622.1 and 233.1 respectively) as the main diagnosis [187]. Some of the cases with hysterectomy also had a secondary diagnosis, but there were also hysterectomy cases with dysplasia or carcinoma in situ in the secondary diagnoses, which

were not included in our figures. Together with the retrieved annual number of women with CIN, the hospital data enabled us to compute an average number of medical procedures per woman with CIN *(see Table 4.1)*.

In the search for CIN grade-specific data and data on out-patient procedures (particularly conserving treatments), all gynaecology departments in the Netherlands were sent a request for annual figures on diagnostic and treatment procedures in the management of CIN. Gynaecologists from 10% of the hospitals responded with relevant data, mostly by sending annual reports, data on treatment of women with CIN III (there are hardly any data on CIN I and II), annual number of conisations, conserving treatments and colposcopies, covering the period 1984–1989. These hospitals cover 25% [190] of the Dutch population. In these data, 33% of the women with CIN III had conservative treatment (cryocoagulation, laser evaporation and hot loop diathermy are widely used in the Netherlands), 56% had conisation and 11% hysterectomy. These fractions were accepted as our estimates. In the 50% responding hospitals with the highest frequency of conservative treatment, these percentages summated to 58%, 35% and 7%, respectively. In addition, we assumed that 15% of all women treated will need retreatment at least once [211], 75% of whom will receive conservative treatment and 25% of whom will receive conisation, amounting to 0.44 ($0.33 + 0.75 \times 0.15$) conservative treatments and 0.60 ($0.56 + 0.25 \times 0.15$) conisations.

In the next step, the fraction for conisations was multiplied by the total number of women diagnosed with CIN III, resulting in the number of conisations which should be attributed to women with CIN III ($0.60 \times 2570 = 1542$). The number of conisations which should be attributed to women with CIN I or II was calculated by subtraction ($2180 - 1542 = 638$), and the fraction of women with CIN I or II having conisation was determined ($638/2590 = 0.25$). The same procedure was followed to calculate the number of hysterectomies and conservative treatments in women with CIN I/II. The total annual number of conservative treatments was calculated from the ratio between conservative treatments and conisations in the reporting gynaecology departments and the total annual number of conisations ($1.273 \times 2180 = 2775$). The fraction — for which no data were available — of women with CIN I/II without a conservative treatment, conisation, or hysterectomy was calculated as the residue after subtracting the total fraction of treatment procedures from 1 plus the fraction of retreatments ($1.15 - 0.63 - 0.25 - 0.05$, *see Table 4.1*). The annual number of colposcopies reported from the gynaecology departments was extrapolated to the national level (15 466 colposcopies).

To collect information on aspects for which no detailed large scale data were available - number of consultations at the gynaecology department, Pap smears, colposcopies and (ecto- or endocervical) biopsies, and the duration of the total period studied — four gynaecologist-colposcopists from different hos-

Table 4.1 *Estimated average number of diagnostic and treatment procedures, and days in hospital in women with screen detected* CIN. *The Netherlands, women diagnosed in 1988*

	CIN I/II number per woman a	CIN III number per woman b	All CIN annual number c
Treatment procedures			
No treatment	0.21	0.00	544
Conserving treatment	0.63	0.44	2775
Conisation	0.25	0.60	2180
Hysterectomy	0.05	0.11	425
Hospital days*			
Conisation	1.1	2.7	9810
Hysterectomy	0.7	1.3	5100
Other	0.1	0.2	870
Total	1.9	4.3	1578
Hospital admissions			
Conisation	0.25	0.60	2180
Hysterectomy	0.05	0.11	425
Other	0.03	0.08	285
Total	0.33	0.79	2890
Assessment procedures			
Pap smears	8.7	7.3	41383
Colposcopy	6.0	4.6	27200
Biopsy	2.4	1.8	10799
Number of consultations and duration of total period studied			
Consultations	9.0	7.6	42680
Years	4.9	4.3	23851

$c = a \times 2590 + b \times 2570$, in which 2590 and 2570 are the annual number of women with CIN I/II and CIN III respectively

* Example: 1.1 days for conisations per women with CIN I/II = 0.25 conisations × 4.5 days *(see Table 4.2)*

pitals (three university hospitals and one regional hospital) were interviewed in a standardised way. From these interviews, we could assess which aspects show variation in practice, and consequently, since large scale data are missing, remain uncertain. The possible influences of these uncertainties were studied in a sensitivity analysis.

A cost-effectiveness analysis was the goal of the cost part of this study and the costs were assessed from the viewpoint of society. An analysis of the true resource costs of all relevant procedures was, however, beyond the scope of this study. Apart from the cost per colposcopy and per hospital day, which reflect an assessment of the true resource costs, the costs of diagnosis and treat-

Table 4.2 *Costs (in Dfl) per medical procedure in the management of* CIN

	Cost per procedure	Days in hospital	Total costs*
Visit	67	0	67
Pap-smear	52	0	52
Prim. colposcopy	145	0	145
Sec. colposcopy	106	0	106
Biopsy	87	0	87
Cryocoagulation etc.	123	0	123
Conisation	401†	4.5	2622
Hysterectomy	2536†	12	8458

* Cost per procedure plus Dfl 494 per day in hospital
† Including Dfl 165 for pre-surgery laboratory procedures and chest X-ray

ment have been approximated by tariffs charged (1993) in the Netherlands *(see Table 4.2)*. The cost per colposcopy was assessed by interviewing colposcopists for time investment, by reviewing financial accounts of gynaecology departments, and by cost analysis of the equipment. The cost per hospital day is an estimate of the average cost per day (weighted average for the general and teaching hospitals, and of general and intensive care), including "hotel" costs, nursing and medical staff, standard medical equipment, medication and overhead costs. We performed a sensitivity analysis on a lower cost per hospital day (see Discussion).

The costs are presented in Dutch guilders (Dfl) in *Table 4.3*. In 1993, the exchange rates for the British pound was Dfl 2.8 and for the U.S. dollar Dfl 1.85.

4.3 Results

The resulting numbers of treatments and days of hospitalisation are summarised in *Table 4.1*. As expected, the fraction of women treated with hysterectomy or conisation was much lower in CIN I/II than in CIN III. Important differences in practice occurred in the management of low grade lesions. The gynaecologists interviewed confirmed that in some of the gynaecology departments in the Netherlands, women with CIN I and, to a lesser extent, CIN II were treated only if there was persistence or progression in the first 2 years after diagnosis. In our calculation, the fraction of women with diagnosed CIN I/II who have no treatment was 21%. Assuming that half the women with CIN I/II show regression within 2 years, this would mean that approximately 40% of the women diagnosed with this condition were initially followed-up with cytology and colposcopy. In a sensitivity analysis, we examined the effect of all women

with CIN I/II being treated immediately after diagnosis, other than by conisation or hysterectomy. We also examined the effect of all these women being initially followed up with cytology and colposcopy *(see Table 4.4)*.

The required numbers of PAP smears, colposcopies, and biopsies per women reported by the colposcopists interviewed showed little variation with primary diagnosis. We assumed 1 PAP smear, 1.2 colposcopies and 1.4 biopsies per women during 1.2 consultations over a period of 0.3 years. In women initially followed up with cytology and colposcopy only, but treated eventually (in our calculations 21% of the women with CIN I/II), these numbers were doubled because a second diagnosis is required. In treated women, the numbers were increased by 30% for (suspected) recurrence. During follow up without treatment, one colposcopy and PAP smear per 6 months was added.

Finally, for all women with CIN, we accounted for consultations, PAP smears and colposcopies during the follow-up after primary management. This follow up ends when a woman is referred for routine screening. The different kinds of schedules used in this period are:

Procedure at consultations	Number of consultations
5 years of follow-up with both cytology and colposcopy	8
5 years of follow up with only cytology	8
3 years of colposcopy followed by 2 years of cytology	8
1 year with one colposcopic and two cytologic evaluations	2

We basically assumed an intermediate schedule with six consultations, six PAP smears, three colposcopies and a duration of 4 years. In a sensitivity analysis we account for more and less intensive schedules *(see Table 4.4)*. The resulting numbers of diagnostic procedures during the total period studied are shown in *Table 4.1*.

Table 4.2 presents the cost per procedure. In *Table 4.3*, the estimated costs for diagnosis and treatment of CIN are presented. The costs of the medical procedures were approximately Dfl 1600 for all grades of CIN. The total costs in CIN III were 45% higher than in CIN I/II, due to the larger number of hospital days. The costs of hospitalisation accounted for more than half of the total costs of Dfl 3727 of the diagnosis, treatment and after treatment follow-up per woman with CIN III.

4.4 Discussion

Multiplication of the estimated numbers of colposcopies per women with the estimated annual number of cases results in the total number of colposcopies (27 200) *(see Table 4.1)* which seems to be (80%) too high compared with

Table 4.3 *Estimated average medical costs (in Dfl) per woman with screen detected* CIN

	CIN I/II		CIN III	
Consultation*	201	7.8%	201	5.4%
Pap smear	454	17.7%	380	10.2%
Colposcopy†	673	26.2%	522	14.0%
Biopsy	206	8.0%	158	4.2%
Conserving treatment	78	3.0%	54	1.4%
Conisation	646	25.1%	1573	42.1%
Hysterectomy	465	18.1%	930	24.9%
Other days in hospital	51	2.0%	109	3.1%
Total costs	2572	100.0%	3727	100.0%
Of which:				
Costs for procedures	1649	64.1%	1634	43.8%
Costs of hospitalization‡	923	35.9%	2093	56.2%

* Only consultations without colposcopy are charged
† One 'Primary colposcopy', all others 'Secondary colposcopy' *(see Table 4.2)*
‡ Dfl 494 per day in hospital

the annual number of 15 466 estimated from the data from the gynaecology department. This could partly be due to the cervical cancer screening programme reintroduced in the Netherlands since 1987. The total number of smears in 1988 was 20% higher than in 1987 [163]. A substantial number of the colposcopies induced by screening in 1988 will only be incorporated in later years and are not present in the reports of 1988. Moreover, the gynaecologists interviewed suggested that there had been substantial underregistration of colposcopies in their departments.

As can be seen from *Table 4.4*, the length of the follow-up period after treatment is important both for the extent of the management for the women involved and for the costs. It also affects the number of colposcopies, which is important in terms of the capacity of the gynaecology departments. The considerable variety in practice shows the need for more evidence on how long a woman treated for CIN should be followed-up colposcopically. How much information colposcopy after treatment adds to cytology is questionable [130]. When only cytology is required, the women can be referred to the general practitioner.

Other important cost factors are the number of women with CIN having conisation and hysterectomy respectively. For the Dutch situation, we found that the number of hysterectomies in women with CIN III differs considerably (from 10% to almost 50%) between hospitals [9]. The use of hysterectomy for the treatment of CIN has been decreasing over recent decades. In a literature search for data from 1980 onwards, we found three studies from British

Table 4.4 *Sensitivity analysis on average medical practice. Percental changes in annual total number of colposcopies, consultations, length of follow-up period and costs (in Dfl of the management of screendetected* CIN *in the Netherlands, calculations for 1990*

Assumptions	Colposcopies	Consultations	Number of women in follow up	Total costs (MLN)
Baseline assumptions	35 977	56 467	31 559	21.5
Initial management of women with CIN I and CIN II				
Immediate treatment in all cases	−14%	−9%	−7%	−4.2%
Cyto-/colposcopic follow-up initially in all cases*	+19%	+12%	+9%	+5.6%
Follow-up schedule after treatment				
One year, colposcopy	−38%	−48%	−65%	−13%
5 Years, colposcopy during the whole period	+95%	+24%	+22%	+20%

* Where we assume that 50% of the followed up women are treated after all, because regression is not noticed within two years

hospitals, in which 15–40% of the women with CIN had conisation and 1–7% hysterectomy [105, 78, 193]. This compares favourably with the 41% and 8%, respectively, (retreatments included) recorded in this study. Each of these studies, however, came from one treatment centre, and quality bias (towards more conservative treatment) cannot be ruled out in data from publishing (= selected) centres. Our estimates on the use of conisation and hysterectomy are based on national data. Furthermore, in two of the studies from the United Kingdom [105, 78] conservative treatment was performed as an in-hospital treatment and under general anaesthesia. Conservative treatment in the Netherlands is performed as an out-patient treatment, under local (if any) anaesthesia.

Current Dutch data, only the distribution of the conisations (and hysterectomies) over the different grades of CIN was based on less complete data (although coverage was much higher than in one-centre studies). The responding 10% of the gynaecology departments represent the larger hospitals, possibly causing an overestimation of the use of conservative treatment in CIN III.

Correction of such an overestimation in our assessment would shift part of the conisations and hysterectomies from women with CIN I of II to women with CIN III, without affecting the total number of these treatment modalities used.

Data from Loizzi and associates [129], also obtained from one treatment centre, show 73% conisations in women with CIN. Goodwin and colleagues [79], who

performed a population based study on all women from New Mexico with diagnosed carcinoma in situ of the cervix in 1982–1985, found 44% hysterectomies. (In the Netherlands, 40% of the CIN III diagnosis concerns carcinoma in situ).

These figures reflect more aggressive treatment strategies and probably large differences in gynaecological practice between regions or countries.

In the Netherlands, a conisation on average takes 4.5 days in hospital. If we hypothetically assume only one hospital day per conisation, this would reduce the total number of hospital days by 29%, and the total costs for the management of screendetected CIN by 14% (calculations for 1990). As far as the cost per hospital day is concerned, women hospitalised for the treatment of CIN are probably more healthy than the average hospital population, and consequently, even having general anaesthesia, they might need less general care. If we arbitrarily decrease the cost per hospital day from Dfl 494 to Dfl 400, the total costs of the management of CIN would decrease by 9%.

The evidence as to whether women with CIN I and II should be followed up initially, to give regression a chance, or should be treated immediately, is inconclusive. Immediate treatment is less costly than the "wait and see" management, but the difference is rather small *(see Table 4.4)*. Treatment without delay saves 2 years of follow-up maximum, at the expense of treating more women. The difference in the psychological burden to the women involved might be important, but this has hardly been quantified: on the one hand, anxiety may occur as long as no treatment has been given, on the other hand women in whom CIN regresses have the psychological advantage of not needing treatment for a condition associated with cancer. The gynaecologist's counselling strategy and attitude probably plays a major role in these psychological effects.

Total costs for the management of woman with detected CIN amounted to 21.5 million guilders in the Netherlands in 1990 *(see Table 4.4)*. This is almost as much as the sum of the assessed costs for the diagnosis and primary treatment of women with invasive cervical cancer (757 woman in 1990 [154]) of 14 million, and for advanced disease (288 women died from cervical cancer in 1990 [37]) of 9 million [11].

In Conclusion, the extent of the overall management of CIN is greater and the costs higher than might be expected. Although there has been a trend towards more conservative treatment, in the Netherlands half of the cases are treated with conisation or hysterectomy. The costs are substantial and the effect on the well being of the women may be considerable. It is this overall practice that should be accounted for when balancing the benefits, unfavourable effects and costs of cervical cancer screening. A further trend towards less aggressive treatment in the near future from LEEP (loop electrosectional excision procedures) combined with 'see and treat' strategies, and specific treatments for

human oncogene, papillomavirus positive women, is controversial. Expectations will have to be reconciled not only with data from excellent treatment centres, but also with data reflecting overall practice.

5

Care and costs for advanced cervical cancer

This chapter is based on M. van Ballegooijen et al
Eur J Cancer, Vol. 28A, No. 10, pp. 1703–1708, 1992.

Abstract

The types, amounts and costs of hospital and home care in patients who died from cervical cancer are investigated, using both national data sources and hospital files. Our goal has been assessment of the savings on treatment and care of advanced cervical cancer resulting from cervical cancer screening.

Hospital costs account for 70% of the total cost per patient of Dfl 29 200. The amount of hospital care decreases significantly with increasing age. The average number of days of hospitalisation per patient with advanced disease decreases from 62 days below age 50 to less than 10 days at age 70 and older. In-hospital medical procedures, home care and nursing home care account for 24, 22, and 8% of the costs, respectively. Mass screening programmes for cervical cancer will result in a reduction in both advanced disease and mortality. The potential savings compensate approximately 10% of the costs of screening.

Acknowledgement This study was supported by the Ministry of Welfare, Health and Cultural Affairs of the Netherlands.

5.1 Introduction

Knowledge of the total amount of care and related medical costs in patients with advanced cancer is significant both in itself and for studies on the cost-effectiveness of cancer prevention. Our goal has been assessment of the savings on treatment and care of advanced cervical cancer resulting from cervical cancer screening. In an earlier study on the costs and savings of cervical cancer screening [118, 9] we roughly estimated the costs of treating advanced disease. This topic is now being dealt with in a more detailed way, accounting for hospitalisation days, diagnostic and (palliative) treatment procedures and home care.

Table 5.1 *Hospital days for invasive cervical cancer*

Year	Hospital days: Invasive cervical cancer (a)*	Advanced cervical cancer (b)=(a)-(i)	Advanced cervical cancer, per death (c)=(b)/(d)	Deaths from cervical cancer (absolute number) (d)
1987	17 654	7392	24.6	301
1988	19 531	8496	25.4	334

Year	Hospital days: Primary diagnosis and surgical treatment (e)†	Primary (brachy-) radiotherapy (f)‡	Complications of primary treatment (g)§	Treatment in survivors of recurrent cervical cancer (h)‖	Total (i)=(e)+(f)+(g)+(h)
1987	6615	3217	245	185	10 262
1988	8033	2560	242	201	11 035

* The total number of hospital days from admissions with main diagnosis invasive cervical cancer (ICD-9 code 180) or with main diagnosis metastasis (ICD-9 codes 160 to 179) and invasive cervical cancer as second diagnosis [188]

† The number of hospital days for primary diagnosis and surgical treatment of cervical cancer, assessed by summing up the hospital days from admissions with invasive cervical carcinoma (ICD-9 code 180) as the main diagnosis and one of the possible surgical diagnostic or primary treatment procedures (including cervical or uteral biopsies, curettages, exconisations and hysterectomies) as one of the operations coded [188]

‡ The number of hospital days (BrDays) for primary (brachy)radiotherapy (megavolt radiotherapy is given in an outpatient setting) is calculated using the formula: $BrDays = (700 - Hys) \times 1.6 \times 5$, in which 700 is the estimated yearly incidence of invasive cancer, *Hys* is the yearly number of hysterectomies performed for invasive cervical cancer [188], 1.6 is the average number of brachytherapy sessions for a woman who is receiving primary radiotherapy, 5 is the estimated number of hospital days per session of brachytherapy [188]

§ The number of hospital days for complications after primary therapy $(Hys \times 0.03 \times 10) + [(700 - Hys) \times (0.04 \times 9)]$, in which 700 is the estimated yearly incidence of invasive cancer, *Hys* is the yearly number of hysterectomies performed for invasive cervical cancer [188], 0.03 and 10 are the risk of a complication and the number of hospital days for a complication requiring a supplementary hospital admission after radical hysterectomy [8], 0.04 and 9 days are the risk of a complication and number of hospital days for a complication requiring a supplementary hospital admission after primary radiotherapy [8]

‖ The number of hospital days for treatment of recurrent invasive cervical cancer in women that survive is: $deaths \times 5\% \times 12 days$, in which *deaths* is the number of deaths from cervical cancer, 5% is the estimated cure rate in women with advanced disease [8], 12 is the estimated number of hospital days per admission for curative treatment of recurrence [8]

5.2 Material and methods

The relevant episode of care starts at the first diagnosis of recurrence or metastasis after primary therapy and always ends with death. If women receive no primary curative treatment, because of poor prognosis at the time of primary diagnosis, the episode studied starts with primary diagnosis. From here on, we will use the term 'advanced cervical cancer' for all cases.

5.2.1 In-hospital care

We analysed national data on the number of admissions and length of stay in Dutch hospitals for patients with advanced cervical cancer during the years 1987–1988 [187] (see *Table 5.1* for details).

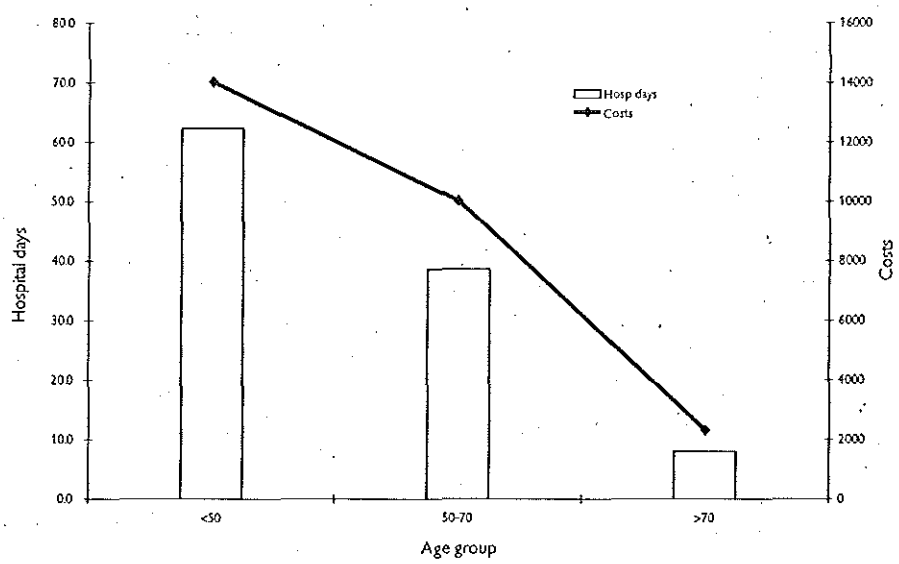
For additional information, especially on number and type of medical procedures, a file study was performed in the Dr Daniel den Hoed Clinic in Rotterdam, which is one of the two cancer centres in the Netherlands, with an associated radiotherapy center. From the medical registration, the files of 40 women who died from cervical carcinoma during the period 1 June 1987 to 31 December 1988 were collected. This registration also includes outpatients never admitted. Five cases have been excluded from further analysis because of incomplete information on stays in other hospitals. In the remaining 35 cases, primary therapy had been radical hysterectomy (8 women) or radiotherapy (22 women), or no primary curative therapy had been applied because of advanced disease with poor prognosis (5 women).

For the major cost categories, we used estimates of the actual resource costs. This is the case for nursing costs per hospital day (excluding costs of medical procedures) [204], and for radiotherapy. The costs of radiotherapy are based on a special study on the costs of radiotherapy in breast cancer [115], but taking differences between breast cancer and cervical cancer treatment into account. The costs of the other medical procedures were approximated by the tariffs charged.

5.2.2 Nursing home care

The average length of stay in nursing homes has been estimated from national data concerning nursing homes admissions [189], covering 80% of all Dutch nursing homes (1986–1988). The data studied concerned a selection of women dying in a nursing home and having neoplasm of the female urogenital tract (NFUGT) and metastasis registered as first two diagnoses (in any order).

Figure 5.1 *Average number of hospital days and average costs Dfl) of in-hospital medical procedures per patient in different age groups*



5.2.3 Home care

Patients who stay at home can receive informal and professional care. Informal care is defined as care provided by relatives, neighbours, friends and volunteers. Professional care in the Netherlands is mainly provided by district nurses, home helps, general practitioners and private nurses.

The available data enabled us to discern two successive phases to assess the total amount of home care provided: a low-care phase and a terminal phase in which the level of care increases rapidly. To estimate the total amount and cost of care in the two phases, information is required concerning the length of the phase, the proportion of patients receiving different types of home care, the intensity of care, and the costs per unit of care.

Low care phase. Information about the amount of care during the low-care phase has been based on data collected in 1986, 1987 and 1988 by the National District Nursing Association [157]. The data are based on a 4% sample of the Dutch population, which is representative with respect to age and sex composition. From this registry, we selected data on patients with cancer of the female genital tract as main diagnosis. To estimate only the care for patients in the low-care phase, we excluded the patients recorded as terminal patients by district nurses.

National data about home help only provide information about the average level of care by age-group [39]. Diagnosis-specific information on home help was available from one organisation [98].

Costs per hour of care were calculated using average wages per type of health care worker, derived from financial reports of the Intensive Homecare Project (IHP) [127] and from the national associations of district nursing, home help and general practitioners [157, 39, 145, 151]. These financial reports also provided the material costs.

Terminal phase. The IHP, whose services will be called 'intensive home care', in operation in an experimental setting since 1987 in three regions in the Netherlands, provided data concerning intensive home care [144, 111]. This project concerns mainly cancer patients who die within a relatively short period and who would need hospital or nursing home care in the absence of this project. This type of care is available for 24 h per day. The amount of care, delivered by district nurses, home helps, general practitioners, private nurses, and informal help was registered daily. We used data from 1988 about care for all 32 patients with cancer of the female genital organs (cervical cancer was not classified separately).

Next to this a significant part of the terminal patients receive 'regular home care', which is restricted to a maximum of 2.5 h of district nursing and 8 h of home help per day. Data on this group of patients were obtained from 2 sources: a study of district nursing care in Amsterdam and data about the level of care for patients in the last week before entering the IHP [127].

5.3 Results

5.3.1 In-hospital care

National data on the number of days of hospitalisation for the years 1987–1988 are presented in *Table 5.1.* The average number of days of hospitalisation for advanced cervical cancer estimated for the years 1987 and 1988 is 25 (see column (c)).

Average age at death for the 35 women in the file study was 53 years. This is lower than the average age at death from cervical cancer on a national level: 65.3 years in 1987 and 66.1 years in 1988 [36]. Of the 35 patients, 15 (43%) died at home, which is not significantly different from the percentage (50%) reported from Dutch national data on cervical cancer deaths [35]. The average duration of the episode of advanced disease was 11 months.

The hospital data from the 35 cases in the file study are presented in *Table 5.2.* The average number of hospital admissions per patient was 4.2 and the average length of stay was 10.3 days. Chemotherapy cycles required relatively short admissions of 2–5 days. On average, the length of stay in hospital was 43 days per woman. This number decreases markedly with increasing age *(see Figure 5.1).* Below age 50, the average number of days is 62, but the 7 women older than 70 years had a mean total duration of stay of only 8 days. This

Table 5.2 *Medical record study of advanced cervical cancer**

Procedure	Mean no. per woman	Costs per procedure (Dfl)	Costs per patient (Dfl)	% Of total costs
Outpatient visits	13.1	17	225	2.2
Radiotherapy†	0.5	5 100	2 625	26.1
Surgery‡	0.7	2 080	1 486	14.8
Chemotherapy cures	1.8	1 075	1 935	19.2
CT-scans	1.9	435	821	8.2
X-rays (IVP excluded)§	4.8	57	272	2.2
IVP	0.7	110	66	0.7
Echoscopies	1.2	141	165	1.6
ECG	0.9	45	41	0.4
Other diagnostics	0.9	325	279	2.8
Intercollegial consultations	1.3	66	83	0.8
diagnostic punctions	1.3	199	251	2.5
Biopsies	1.2	95	117	1.2
Exam under narcosis	0.7	380	250	2.5
Laboratory	53.0	27	1 456	14.5
Total			10 070	100.0
Hospital days	43.4			
of which in intensive care	0.5			

* Outpatient visits, diagnostic and treatment procedures, and hospital days in 35 women who died from cervical cancer (Daniel den Hoed Clinic)

† Referring to a whole course of treatment; 18 out of the 35 women had radiotherapy

‡ Including the urinary tract (13), the digestive tract (7), laparotomies (2), radical hysterectomy (1), brain surgery (1) and cordotomy (1)

§ 65% Of the thorax, 18% of the abdomen, pelvis and lumbar spinal column

age trend should be kept in mind in comparing the hospital data to national data, since the women in the file study were relatively young. Standardisation of the mean duration of stay on the age distribution of cervical cancer deaths in the Netherlands results in a much lower estimate of 29 days for the average duration of stay in hospital (and an extra 0.3 days in intensive care), which is in reasonable agreement with the study on national admission data, although it is slightly higher (29 vs 25 days).

For each patient in the file study, the total costs of all in-hospital medical procedures (excluding the costs of hospitalisation) have been calculated *(see Table 5.2)*. The average costs are Dfl 10 070 per patient. The main constituents are costs of radiotherapy (26%), chemotherapy (19%), surgery (15%) and laboratory costs (15%).

In the costs of hospital procedures, again, a marked age-trend was found, very similar to the trend in number of days of hospitalisation *(see Figure 5.1)*. For women who died before age 50, the average costs amount to Dfl 14 000, but for women over age 70 the costs are only Dfl 2300. Standardisation for the age dis-

tribution of all cervical cancer deaths in the Netherlands results in an estimated average of Dfl 7100. After primary treatment, 4 out of the 35 women received only radiotherapy, 4 only surgery, 1 only chemotherapy, 4 radiotherapy and chemotherapy, 6 radiotherapy and surgery, 2 chemotherapy, 4 all three treatment modalities and 10 none of them. For the group of 25 women who had radiotherapy, surgery and/or chemotherapy, the average costs of in-hospital procedures are Dfl 13 700 and the average number of hospital days is 39. Radiotherapy, surgery and chemotherapy then account for 27%, 15% and 20% of these costs, respectively.

The total in-hospital cost for advanced cervical cancer is based on the sum of the age-standardised costs of procedures (Dfl 7100) and the costs of hospitalisation. Assuming 27 hospital days (an intermediate between the 2 described estimates), a cost per hospital day of Dfl 470 for normal care (medical procedures excluded), and Dfl 1880 for intensive care [204], the costs for hospitalisation amount to Dfl 13 250 per patient *(see Table 5.3)*.

5.3.2 Nursing home care

The average length of stay in a nursing home for the 95 women who died from advanced cancer of the urogenital tract during the period 1986–1989 was 230 days. 61% of these women came from hospitals. From the data on hospital admissions, it appears that each year between 4 and 16 women with advanced cervical cancer enter a nursing home after leaving the hospital. On the basis of these figures we assumed that 5% of the approximately 300 women who die from cervical cancer in a year in the Netherlands are admitted to a nursing home because of advanced disease for an average 230 days per patient.

Accounting for the costs per nursing home day (Dfl 200), these estimates mount up to the average costs of nursing home care of Dfl 2400 per patient with advanced cervical cancer. This is 8% of the total costs of care for these patients *(see Table 5.3)*.

5.3.3 Home care

The estimates concerning the four relevant variables (phase length, participation in different types of home care, intensity of care, and costs per hour) are presented in *Table 5.4*.

Phase length. The 32 women with metastases of the female genital tract in the IHP received on average 3 weeks of intensive home care, which we accepted as the length of the terminal phase. The average duration of the low-care phase should therefore be 10 months: the total period of advanced disease according to the hospital patient files (11 months), minus the terminal 3 weeks.

Participation. In the data from the regions covered by the IHP, we found that 27% of the terminal patients received intensive home care. We extrapo-

Table 5.3 *Estimated costs of diagnosis and treatment of advanced cervical cancer*

	Costs per patient (Dfl)		% Of total costs	
In-hospital care	20 350		70	
hospital days		12 690		43
intensive care		560		2
in-hospital procedures		7 100		24
Nursing home care	2 400		8	
Home care	6 440		22	
Total care	29 200		100	

lated this percentage to all women dying from cervical cancer, as nationwide coverage of intensive home care is to be implemented in the near future. The percentage of women dying at home with regular care (23%) follows as the remainder from the total number of deceased minus the number of patients dying in hospitals and nursing homes (approximately 50%) [35], and at home with intensive terminal care (27%).

Intensity. No data were available with respect to private nursing during the terminal period in patients who did not participate in the IHP. The amount of informal care in these patients has been estimated at 43 h per week, similar to patients receiving intensive home care. In view of the much lower amount of professional care, this may well be an underestimation.

Total costs of home care. Combining the estimates presented in *Table 5.4* results in average costs per patient, shown in *Table 5.5*. The average total cost of home care per patient with advanced cervical cancer is Dfl 6440. More than 60% of these costs are incurred by the intensive terminal home care.

5.3.4 Total costs of advanced disease

The total costs of (professional) care for a patient with advanced cervical cancer are Dfl 29 000 *(see Table 5.3)*. From the hospital costs, the hospital days account for 45% of the total costs, and in-hospital procedures for 24%, which amounts to Dfl 20 350, over two-thirds of the total costs. Professional home care accounts for Dfl 6440 per patient, or 22% of the total costs.

5.3.5 Savings by cervical cancer screening

As stated before, our concern has been to study the costs and effects of cervical cancer screening, which demands an assessment of the savings resulting from deaths avoided. We calculated the impact of these savings on the total costs of screening using our model for cervical cancer [118, 9]. In these calculations

Table 5.4 *Average values for episode length, participation in different types of home care, intensity of care, and costs per hour, for cervical cancer patients having advanced disease*

	Low care episode	Terminal episode	
		Intensive care	Regular care
Episode length	10 months	3 weeks	
Participation			
District nursing	22% each month	]	]
Home help	22% each month		
Private nursing	NA	27%	23%
General practitioner	NA		
Informal care	NA	]	]
Care intensity (hours/week)			
District nursing	0.6	29	8
Home help	5	20	10
Private nursing	NA	19	NA
General practitioner	NA	1.5	1
Informal care	NA	43	43
Costs per hour (Dfl)			
District nursing	75	81	86
Home help	36	47	42
Private nursing	61	73	69
General practitioner	110	110	110

NA No data available

we used detailed cost estimates on screening and primary diagnosis and treatment described elsewhere [118, 9], and the cost estimates of care for advanced disease presented in this article. The results are presented in *Table 5.6.* In the Netherlands, a hypothetical screening policy in which women are invited 7 times between ages 37 and 73, assuming a participation rate of 65%, will cost 325 million guilders in the period 1988–2015. The effectiveness of screening is reflected in a decrease of the costs for care of advanced disease from Dfl 164 (128 + 36) million to 127 (99 + 28) million. The saving of Dfl 37 million represents 10% of the costs of screening plus the incremental costs of diagnostic and treatment procedures for primary disease. The costs per life- year gained are Dfl 21 700.

5.3.6 Sensitivity analyses

The main uncertainty in our analysis of national hospital data on the number of hospital days is the number of admission days per brachytherapy session in primary therapy *(see Table 5.1).* Assuming an average stay of 3 days, which is a minimum per session, this would increase the number of hospital days per

Table 5.5 *Average costs for home care per patient with advanced cervical cancer**

Type of care	Low-care episode	Terminal episode		Total period	Total costs
		Intensive	Regular		
District nursing	500	1910	480	—	2890
Home help	770	760	290	—	1820
General practitioner	NA	130	80	—	210
Private nursing	NA	1120	NA	—	1120
Material costs				400	400
Total	1270	3920	850	400	6440
Informal care (h)	NA	35	30		65

* Costs per type of care and episode in Dfl
NA No data available

women dying from cervical cancer from 25 to 27 days, which would even be more close to the number found in the file study (29 days).

We accepted the (age-corrected) outcomes for the costs of treatment procedures of the cancer clinic as representative for the current and future practice on a national level. However, the role of chemotherapy in advanced cervical cancer is still the subject of randomised trials. We calculated the results in case chemotherapy could be abandoned, accounting for the resulting savings on admission days. The total costs of treatment and care of advanced cervical cancer would then decrease from Dfl 29 200 to Dfl 25 100 (-14%). The proportion of women receiving district nursing and home help in the low- care phase is very uncertain. Halving the propoïtion of district nursing results in Dfl 6120 costs of home care, a decline of only 5%. Our estimate of the proportion of women receiving home help was quite conservative. Doubling this proportion, a quite dramatic change, gives the total costs for home care in advanced cervical cancer of Dfl 7060 and only a 2% rise in the total costs. Appraisal of the informal care would probably have more impact. It is possible by use of shadow prices [213]. However, we think that the bulk of informal care is given in the low-care phase, for which no data are available.

5.4 Discussion

For estimating the amount of in-hospital care one should collect empirical data from different settings: cancer treatment centres, universities and other major hospitals, and smaller local hospitals. This was not possible because only cancer treatment centres systematically follow the vital status and cause of death of all patients. Without such a complete follow-up, selecting patients

Table 5.6 *Cost-effectiveness of an efficient policy of cervical cancer screening, assuming total costs of advanced disease per woman of Dfl 29 200. 5% discount rate*

Costs (million Dfl)	Screening policy*	No early detection	Difference in costs
Screening, and diagnostic and treatment Procedures for primary disease	556	230	326
Care of advanced disease			
In-hospital care	99	128	-29
Home care	28	36	-8
Total costs	682	393	289
Life-years gained	13 321		
Costs (Dfl) per life-year gained	21 713		

* Screening women 7 times between 37 and 73 (every 6 years) during the period 1988–2015 in the Netherlands, attendance rate 65%

who died would retrieve mainly patients who died in hospitals. Next to this, in most hospitals out patients are not recorded in the medical registration. Patients without any hospital admissions would be missed when analysing data from these hospitals.

For the total length of stay in hospitals, by far the most important cost component, we analysed two independent sets of data, one on a national basis, one from a specialised centre. The difference between the results was small (25 vs 29 hospital days), suggesting that average medical practice in respect to treatment of advanced cervical cancer differs little from the practice in specialised cancer treatment centres.

We found a clear association between age and number of hospital days, and also between age and treatment costs. Treatment costs are probably lower at higher age because of the risks and adverse effects of treatment that will often not be counterbalanced by the low chance of remission.

Riley et al [176]compared medical costs in different age groups of elderly patients in the last year of life with cause of death. For all sites they report lower costs for patients with cancer in age-group 85+ than in age-group 65–74. For patients with cancer of the genital organs this decrease was 40%, from US$ 9126 to US$ 5440. Our data show that this trend starts in younger age groups.

We had to assess the costs of home care for advanced cervical cancer by data on patients with neoplasma of the female genital tract. However, the difference in the amount of intensive home care between different cancer sites is very small [144], therefore this is not a major shortcoming.

We only found one study in which the costs of advanced cervical cancer are assessed from empirical data: an analysis of the Medicare file in the USA [6]. The different cost components are not presented in this study. The total costs of care for the period following primary therapy until death are US$ 17 700, which is higher than in our study (US$ 1 is approximately Dfl 2). The US tariff per hospital-day however, is approximately twice the Dutch tariff, which could already level off the difference. We could not find any study in which the amount of care in advanced cervical cancer patients is presented explicitly. In the present study, only the medical costs have been assessed. Advanced cancer, leading to death, also causes a psychological burden and social costs (loss of productivity and costs of informal care). Effective mass screening programmes for cervical cancer will reduce both. On the other hand, cervical cancer screening also affects a larger number of women in a negative way because of false-positive test results and detection of regressive intraepithelial neoplasias.

The costs of hospital and home care in advanced cervical cancer are substantial. Nevertheless, in cervical cancer screening the savings on the treatment of advanced cervical cancer cannot compensate the costs of screening. To answer the question whether cervical cancer screening is expensive, the cost-effectiveness should be compared with that of other medical interventions. Such a comparison shows that the costs per life-year gained are intermediate between breast cancer screening, which is cheaper per life-year gained, and hypertension treatment in men and heart transplantation, which is more expensive [181].

6

Comparing the predictions of a cervical cancer natural history model with incidence and mortality trends after the introduction of screening

M. van Ballegooijen et al, submitted for publication

Abstract

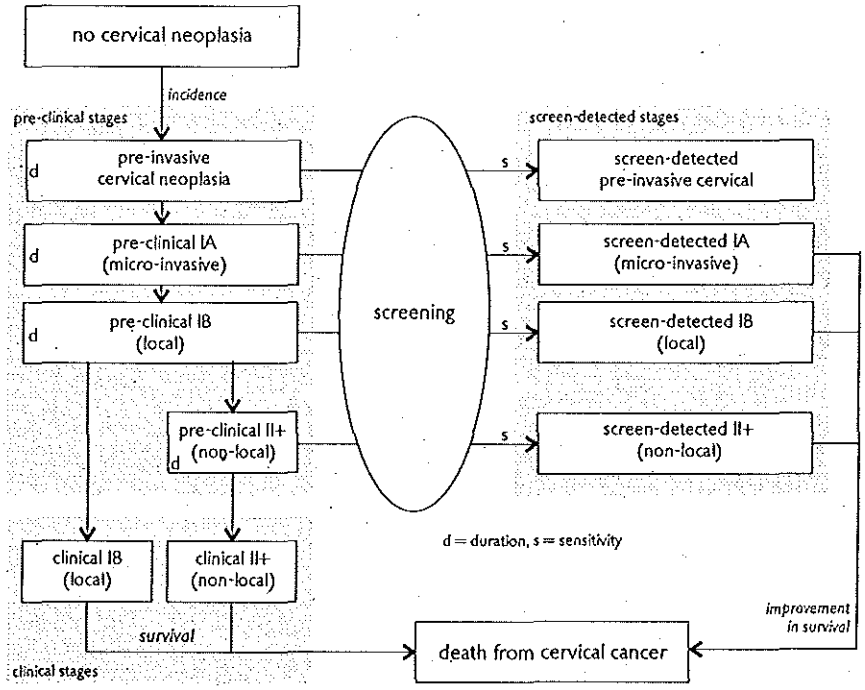
Background: Knowledge about the natural history of cervical cancer and its precursors is essential for the planning and evaluation of cervical cancer screening policies. One way of studying the natural history has been the use of quantitative models for estimating the duration of preclinical stages and the sensitivity of the PAP smear from population data on cervical cancer and screening. The validity question addressed in this paper is how compatible these estimates, combined with an assumed full cure after treatment for pre-invasive stages, are with the observed trends in cervical cancer incidence and mortality in the Netherlands after the introduction of screening.

Material and Methods: The MISCAN simulation model has been used to predict age-specific incidence and mortality trends for the period 1965–1992 in the Netherlands on the basis of the observed screening pattern. In the simulations, we accounted for autonomous trends in cervical cancer risk estimated from Dutch pre-screening mortality data. Estimates of the duration of preclinical disease (mean 15.7 years) and the sensitivity of the screening test (80%) were based on screening data from British Columbia, Canada. The predicted incidence and mortality were compared with observed trends.

Results: Predicted incidence and mortality rates were found to correspond reasonably well with observed trends in cohorts for which the risk of cercical cancer could be estimated from pre-screening data. Predictions are too high if no effect from screening on incidence and mortality is assumed.

Conclusion: The agreement between predictions based on previously derived natural history assumptions and the incidence and mortality trends since the start of

Figure 6.1 *Schematic presentation of the disease model. Pre-clinical stages, clinical stages, screen detected stages and disease specific mortality. Birth and mortality from other causes are not represented*



screening in the Netherlands, provides further support for an estimated duration for detectable preclinical progressive cervical cancer of approximately 16 years, an estimated sensitivity of the PAP smear of about 80%, and a full cure after treatment for pre-invasive stages.

6.1 Introduction

The natural history of cervical cancer has been studied extensively, not only to attain a better understanding of the disease itself but also to learn more about the way in which screening works. Cervical cancer screening leads to detection and treatment of intraepithelial neoplasia. A reduction in incidence and mortality is expected because intraepithelial neoplasia is considered to be a pre-invasive precursor of invasive cancer and appropriate treatment is assumed to result in almost a 100% cure ([106, 62]). An additional mortality reduction is expected from the detection of early invasive cases which have a relatively good prognosis.

Screening for cervical cancer has been shown to reduce the incidence of invasive cancer in screened women (see for example [43, 100, 81]) and the nationwide cervical cancer mortality [142, 126]. However, serious doubt has been expressed over whether mass screening does reduce mortality in the UK and New Zealand [173][82].

The objective of this paper is to test whether earlier derived natural history, sensitivity and cure-rate assumptions are compatible with incidence and mortality trends in The Netherlands over the past 30 years. In order to predict these trends, the assumptions were combined in an integrated model for cervical cancer screening which includes the sensitivity of the PAP smear, the incidence and duration of pre-invasive and early invasive lesions, and the screening pattern in the Netherlands since the start of mass screening in the early 1970's. The assumptions made on duration of pre-clinical detectable disease and test sensitivity stem from an analysis of screening from British Columbia [159]. The model was calibrated to the incidence and mortality in the Netherlands before the start of screening.

We used the model for predicting trends in cervical cancer incidence and mortality after screening started and compared these predictions with observed recent incidence data and with mortality in 1978–1982 and 1988–1992. In addition, we used the model to predict mortality for the hypothetical situation where screening would have had no effect at all in the Netherlands. This made it possible to compare predictions with and without screening effect for their agreement with observations.

It was considered that a good fit between the observed and predicted incidence and mortality trends would improve the credibility of the model. This is important, because we use these models to assess the percentage reduction in mortality effectuated by screening (see Section 6.4), and to compare different screening policies with regard to the public health effects and costs, see for example Koopmanschap et al [118]and Chapter 7.

6.2 Methods and materials

We used the micro-simulation program MISCAN to model cervical cancer screening and predict its effects [89]. In micro-simulation, fictitious individual life-histories are generated on the basis of probability distributions and other parameters specified in the model. Each life-history is characterized by a date of birth and the age of death from other causes. Some life-histories may include an age at which a hysterectomy is performed for reasons other than cervical cancer, after which the woman is no longer at risk. In case the woman develops cervical cancer, the life-history includes clinical diagnosis and survival. For each life-history, screening tests are carried out at specified ages which in case

Table 6.1 *Overview of the model parameters and the data sets used to estimate them. The upper part concerns the disease model* (see Figure 6.1), *the lower part describes the situation in the Netherlands*

Model parameters Data used for estimation
The disease model
Duration of pre-clinical cervical neoplasia (1) *Sensitivity* of PAP smear for pre-clinical cervical neoplasia British Columbia data
Survival (2) and *improvement in prognosis* Pre-screening ratio of incidence to mortality, stage specific survival data
Age- and cohortspecific *incidence* of disease Age- and cohortspecific pre-screening mortality data and accounting for (1)+(2)+(3)+(4)
Age- and cohort-specific cervical cancer *risks* (4) Pre-screening mortality data from the Netherlands (1950–1975)
The situation in The Netherlands
Births per cohort *Deaths* from other causes National demographic data
Hysterectomies for other reasons (3) National hospital registry
Screening pattern in the Netherlands in 1970–1992 Various data on the screening uptake in the Netherlands (see text)

of a (true) positive test, result in detection and subsequent treatment, which in turn may lead to a change of the disease history. The predicted effects of screening for cervical cancer depend on the quantification of the different components of the model. *Figure 6.1* gives an overview of the complete disease model. *Table 6.1* shows how the model assumptions were based on previous analyses. *Table 6.2* gives an overview of the main parameter values. The assumptions and parameter values will be explained below.

6.2.1 Year of birth, death from other causes, hysterectomy

The distribution of the year of birth follows from the size of the birth cohorts in The Netherlands. Death rates for the Dutch female population were directly adapted from demographic data [37]. Annual age-specific rates for total hysterectomies were obtained from the national Hospital Admission Registration [186].

6.2.2 Incidence of progressive pre-invasive cancer

The life-time risk of developing progressive cervical cancer differs between 10-year birth-cohorts *(see Figure 6.2 a)*. The estimates result from a log-linear (APC) analysis of pre-screening mortality rates available from 1950 onwards, in which age, period and cohort factors are assumed to operate independently [44]. For women born after 1915, mortality has been included in the APC analysis until 1974 because of substancial participation in screening after it started around 1970. Mortality has been included up to 1989 for older cohorts. The APC model adequately explains pre-screening mortality rates ($p = 0.8$). Linear trends can in the APC model be explained by both cohort effects and period effects. We chose to attribute them to cohort effects, assuming no trend in the period factors, i.e. we assumed no major improvements in stage distribution or treatment of clinical cervical cancer over the period considered. Data to investigate this assumption are not available for the fifties and sixties in the Netherlands. In the south of Sweden, where pre-screening data on stage distribution and stage-specific survival have been available since 1930, there was improvement in stage distribution and stage specific survival before 1950, but not after 1950 (Sparen 1995). In the national data of Sweden however, the improvements in survival occurred 10 to 15 years later. Thus, survival improved by approximately 10% after 1950 [172]. In the discussion, we will address the possible impact of such an improvement on our results.

The cervical cancer risk as estimated from the APC-analysis is more or less constant for cohorts born between 1903 and 1927, and then decreases sharply for younger cohorts. The figure suggests stabilization for the cohorts born after 1940, but for these cohorts, the estimates are based on very small numbers of cases *(see Figure 6.2 a)*.

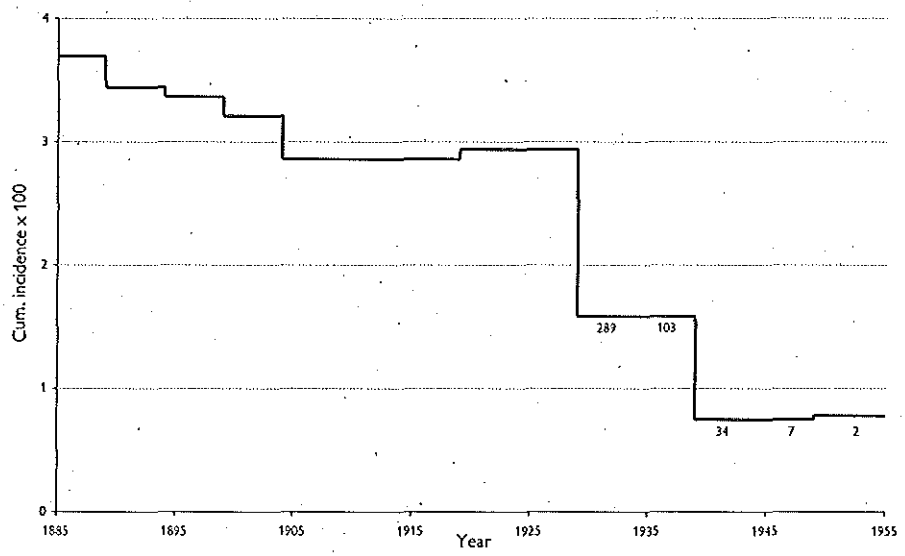
The age-distribution of incidence of progressive pre-invasive neoplasia was back-calculated from the age-components of the mortality derived in the APC-analysis. In this back-calculation procedure we used the distribution of the duration between onset of pre-invasive lesions and death from cervical cancer. This duration includes the total duration of the preclinical stages, and the duration between clinical diagnosis and death combined with the probability of death from cervical cancer (see below). The resulting incidence has a peak in age-group 30–34 years *(see Figure 6.2 b)*.

6.2.3 Pre-clinical disease stages and their duration

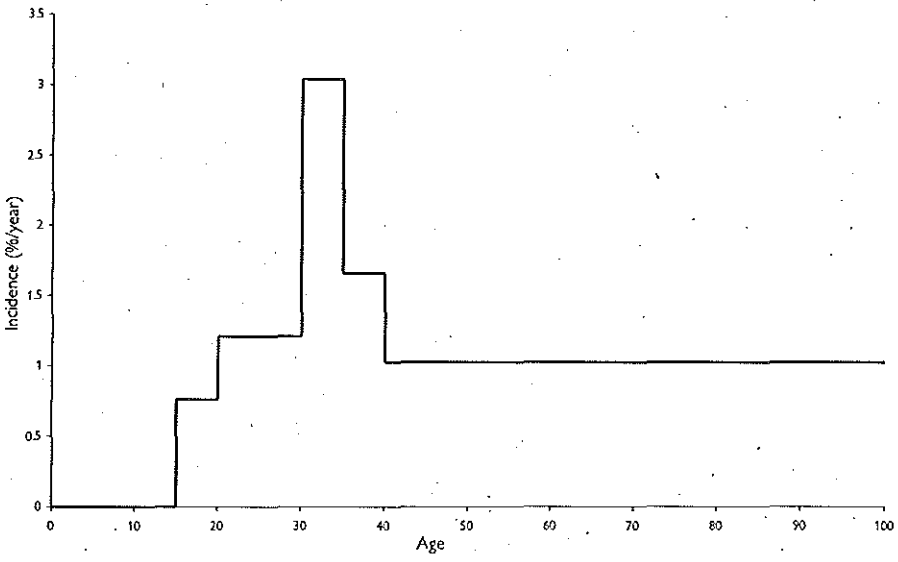
The stages and possible transitions in preclinical disease are presented in *Figure 6.1*. The first stage is screen-detectable pre-invasive neoplasia (corresponding with CIN: cervical intraepithelial neoplasia), with a mean duration of 11.8 years [159]. Regression is possible from this stage. Regression is an important aspect of the disease process when the model is used for evaluation of

Figure 6.2 *Incidence of progressive pre-invasive cervical neoplasia by birth cohort and by age. The figures correspond to the number of cases on which the estimates were based*

(a) The cumulative incidence of progressive pre-invasive cervical neoplasia by birth-cohort



(b) The percentage of the cumulative incidence that occurs at each year of age



adverse effects and costs of screening. However, detection of regressive disease does not affect incidence of invasive cervical cancer nor subsequent mortality. Therefore, only progressive lesions have been included in the present analysis.

In invasive cancer, the screen-detectable stages IA, IB and II+ (corresponding with FIGO definitions) are distinguished. Micro-invasive cancer stage IA is assumed never to give symptoms (except for accidental findings, micro-invasive stage IA cancer is only detected by screening), hence in the situation without screening it is always followed by a preclinical stage IB. Stage IB sometimes becomes clinical, and sometimes progresses to stage II+ before symptoms develop (see the next paragraph on clinically diagnosed cancers). To account for variability, the duration in the pre-invasive and invasive preclinical stages is assumed to follow a Weibull probability distribution (with shape parameter 1.9). The total mean duration of 3.9 years for invasive preclinical disease *(see Table 6.2)* has been obtained from the ratio of detection rates at the first smear in the data from British Colombia and clinical incidence before screening started [27, 159]. In the Dutch pilot screening programme a similar ratio was found between incidence and detection rates of invasive cancer at the first screening, and 54% of the invasive cancers detected at first smears were in stage IA[71], suggesting that preclinical stage IA and stage IB+ have approximately the same average duration.

6.2.4 Survival of clinically diagnosed cancers

In clinical disease we use a subdivision into stages IB and II+. In the absence of population based data on clinical stage-distribution in the Netherlands, we used data from Dutch hospital registries [107]and Norwegian population registries [50]. In the pre-screening period, both data sets showed a clearly more favourable stage distribution for younger age groups. We assumed that the proportion of stage IB among clinically diagnosed cancers decreases linearly from 58% at age 30 to 26% at age 70.

In agreement with the Norwegian survival data [50], we assumed that lethality is highest in the first years following diagnosis, and that after 5–7 years mortality is already relatively small *(see Table 6.2)*. We adjusted the age-specific long-term survival to the ratio of cervical cancer mortality to cervical cancer incidence, using Dutch mortality and incidence figures from the pre-screening period in the Netherlands (1968–1972 and 1965–1969 respectively) *(see Table 6.2)*.

6.2.5 Survival of screen detected cases

Cure is assumed to be complete for screen-detected pre-invasive lesions. Screen-detection of invasive cancer is also assumed to improve survival. This improvement depends on the stage at detection *(see Table 6.2)*, and is modeled as a

Table 6.2 *Assumptions about the natural history of cervical cancer and its precursors (a–c) and about the screening effect (d–e)*

(a) Duration of preclinical stages

Stage	Mean duration (years)
Pre-invasive neoplasia	11.8
Micro-invasive IA	2.0
Pre-clinical IB → clinical IB	1.9
Pre-clinical IB → pre-clinical II+	1.0
Pre-clinical II+	0.9
Total pre-clinical	15.7

(b) Survival: long-term relative survival by clinical stage

Age	IB	II+
<25	0.699	0.200
30	0.812	0.500
50	0.812	0.500
>65	0.624	0.000

(c) Survival: duration of survival of women who die from cervical cancer

Time since diagnosis	Probability of surviving
1.5 years	0.616
4 years	0.153
7 years	0.077
99 years	0.000

Screening effect:

(d) Survival in screen detected cases: reduction in risk of dying of cervical cancer by stage in which (pre)cancer is detected

Pre-invasive neoplasia	100%
Micro-invasive IA	80%
Pre-clinical IB	40%
Pre-clinical II+	20%

Screening effect:

(e) Sensitivity of the Pap smear by stage

Pre-invasive neoplasia	80%
Micro-invasive IA	85%
Pre-clinical IB	85%
Pre-clinical II+	90%

reduction of the risk for cervical cancer patients of dying from screen-detected cervical cancer compared to their risk of dying from the cancer in the situation without screening. For example, if the probability of dying from cervical cancer for patients of a particular age is 60% without screening, then detection by screening in stage IA will reduce this probability by 80% *(see Table 6.2)* to only

12%. This 80% improvement for stage IA was found to reproduce the reported 97% 5-years relative survival for this stage [80]. For the screen detected stage II+, the stage-distribution within stage II+ in a period with little screening (1970–1975) was compared to the stage-distribution in screen detected II+ cases [71]. Accounting for survival at stages II, III and IV, this resulted in an estimated 20% improvement of prognosis. For screen detected stage IB, an intermediate improvement of prognosis of 40% was assumed.

6.2.6 Sensitivity

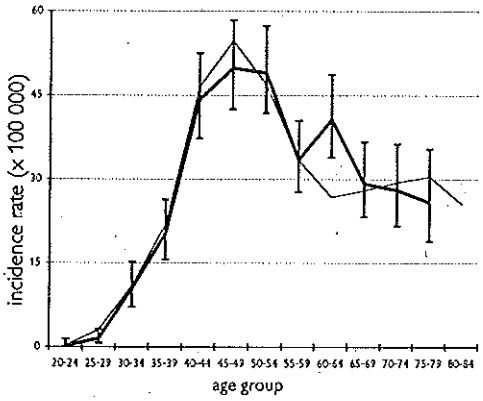
When a woman with disease is screened, the test result can be either true positive or false negative. The estimated sensitivity for the pre-invasive detectable stage is 80% [159]. Sensitivity for invasive pre-clinical stages was assumed to be somewhat higher *(see Table 6.2)*. False positive results, which are important in a complete evaluation of screening, are not considered in this analysis because they do not affect clinical incidence or mortality.

6.2.7 The screening pattern

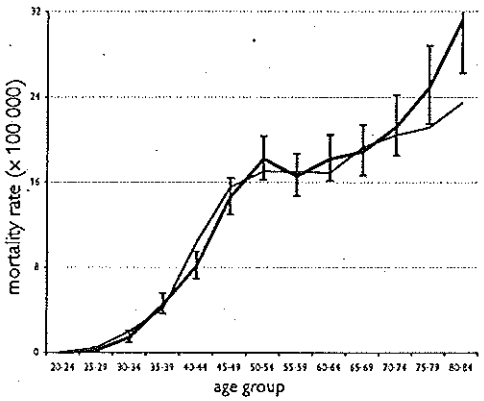
In the simulation of the Dutch situation, screening is assumed to start in 1970. PAP smears are assumed not to have been taken below age 20. In order to describe the screening uptake in the female population, data on the total number of smears are of limited value. Data on the distribution of smears over the female population and on individual patterns in frequency of screening are also needed. But such information on the pattern of individual screening in the Dutch population is scarce. For predicting the impact of screening on incidence of invasive cancer and on mortality, the annual number of first smears is especially important, because first smears are expected to have the highest impact. The assumed proportion of women who had at least one smear over the period 1970–1991 is given in *Table 6.3*. As indicated, the proportions for the years 1976 and 1991 were based on empirical data [67, 29]. We decreased the fraction of ever screened women as reported by the 1991 national health survey by a factor 10%, assuming an association between non-response to this survey (44%) and non-participation to screening. After this correction, the highest coverage of 83% is found in the 1939–48 birth cohort. After accounting for the hysterectomies for reasons other than cervical cancer, this corresponds with a coverage of approximately 90%. We assumed that 10% of women never participate in screening, and therefore 90% is a maximum coverage. In the Dutch cervical screening pilot regions (1976–1986) the coverage per 3-year round never exceeded 80% of the target population [68]. Literature on screening coverage rarely addresses the question of 'ever' participating at PAP smear screening [150]. Usually data concern smear utilisation in e.g. the last 5 years [97, 177, 14]. In these population based studies, the reported coverage rates

Figure 6.3 *Observed and simulated incidence and mortality rates in the Netherlands before (a and b) and after (c, d and e) screening started. Observed data: CCR 1973 [40] (a), NCR 1992 [153] and 1993 [154] (d), CBS 1994 [37] (b, c and e)*

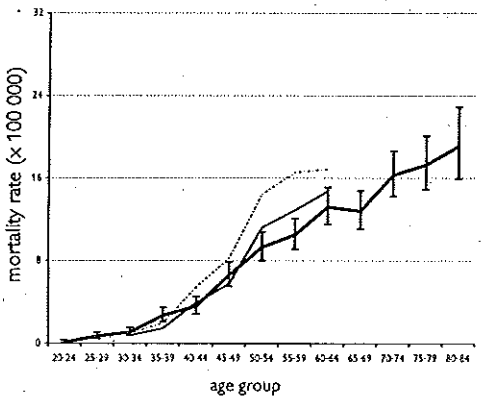
(a) pre-screening incidence 1965–1969.



(b) pre-screening mortality 1968–1972.

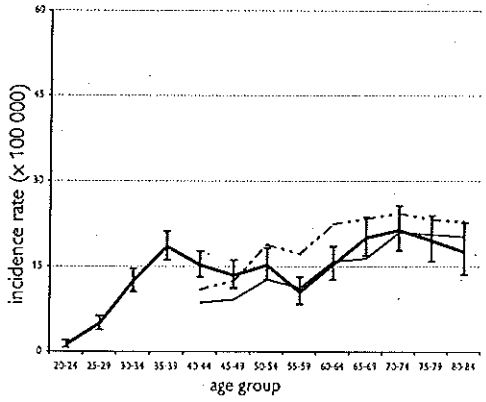


(c) mortality 1978–1982

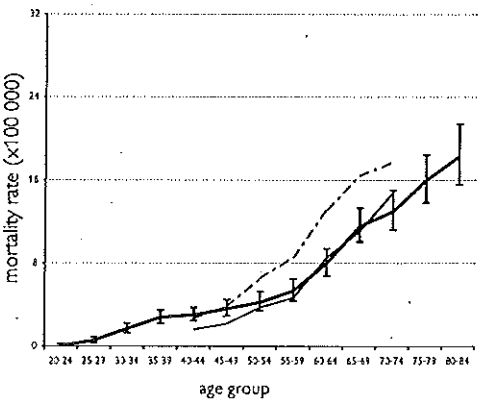


observed + 95% CI
predicted, no screening effect
predicted, with screening effect

(d) incidence 1989–1990.



(e) mortality (1988–1992)



vary between 65 and 80%. We also assumed that the 10% never attending women are at a three times higher risk for cervical cancer than attenders. A high risk for non-participants was reported in Denmark [18], Norway [134] and British Columbia [27]. In British Columbia the risk for participants was estimated at 0.74 of the average risk [159].

In the simulation, the probability of a second smear varies from 90% in women under age 40 at their first smear, to 10% in women over age 60. Of these second smears, 93% are taken within 3 years. Among women who had at least two smears, 90% will have another smear within 3 years. These assumptions are based on the following data. In the Netherlands, invitational screening was combined with opportunistic screening throughout the period under consideration (1970–1992). The majority of the women who had PAP smears attended the organized screening program, in which women between age 35 and 53 were invited at 3 year intervals. Of the women who attended the 3- yearly screening program, 90% had either attended the preceding round or reported to have had a PAP smear outside the program recently [68]. This was measured around 1980, when coverage was increasing.

6.2.8 Reference data

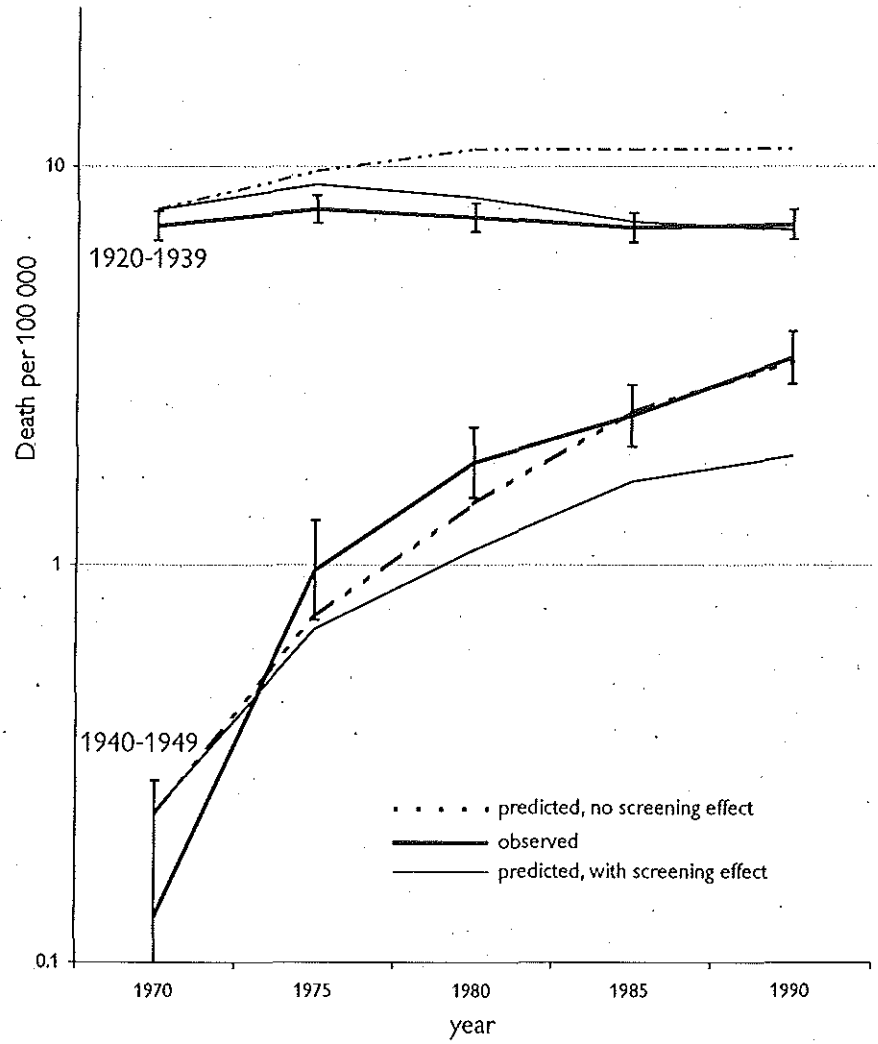
In the Netherlands, a nationwide cancer incidence registry was only in operation from 1989 onwards. Data from this registry were used for the years 1989–1990 [154, 155], when incidence was affected by screening. For the period 1960–1971, which represents the situation before screening, incidence data are available from three local registries in the (rural) province of Friesland and in the cities of Rotterdam and The Hague [40]. National incidence figures have been calculated from these regional incidences. Regional differences have been accounted for by using differences between the age-standardized mortality rate for these three regions and for the total country for the period 1968–1978 [34] as a proxy.

Nationwide age-specific cervical cancer mortality rates have been available since 1950, and are used for the periods 1968–1972, 1978–1982 and 1988–1992 [37]. The last two periods represent test data, and the first (pre-screening) period was part of the data used to calibrate the model.

6.3 Results

Using the model described above, we made predictions of the cervical cancer incidence and mortality in the Netherlands during the period 1965–1992, starting during a situation without screening and taking the effects of screens caried out since 1970 into account (thin solid lines in *Figures 6.3* and *6.4*). For comparison, we also made predictions assuming no effect of screening at all (dashed lines in *Figures 6.3* and *6.4*).

Figure 6.4 *Observed and predicted cervical cancer mortality in two birth cohorts (1920–39, and 1940–49). Observed data: CBS 1994* [37]



6.3.1 Pre-screening incidence and mortality in 1965–1972

Pre-screening incidence and mortality rates are reproduced in *Figure 6.3a* and *6.3b*. Notice that we used the observed rates of this period to calibrate the mopdel (see Methods and Materials).

Because the ratio of mortality to incidence showed a dip in age-group 60–64 (not reproduced in the smoothed model assumptions), predicted incidence in age group 50–69 is 10% lower than observed. In age group 75 and over, observed mortality is equal to or higher than the observed incidence, which indicates that there is a registration problem in the oldest age groups. For the other age-groups, there are no systematic or statistically significant differences between simulated and observed mortality in the period 1968–1972, except for a too high simulated mortality in age-group 40–44. The higher simulated rate below age 35, corresponding with the birth cohort 1940–49, is based on a very small number of observed cases.

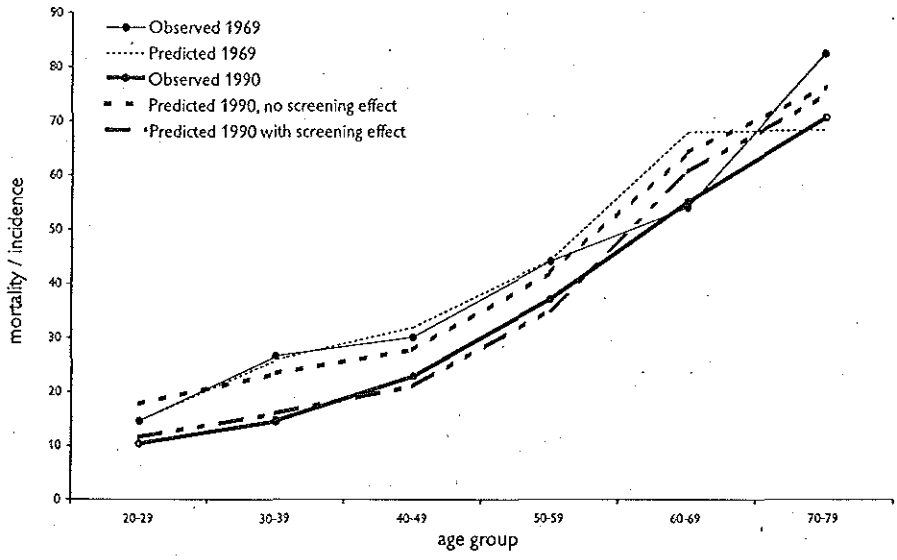
The pre-screening ratio between mortality and incidence is shown in *Figure 6.5* (the thin '1969' lines). This ratio reflects the lethality from cervical cancer and the figure shows the strong age-dependency of lethality. Up to age 60, the simulated ratio is close to the observed data. The deviation in age groups 60–69 and 70–79 have been discussed above. These higher age-groups had hardly any screening in the subsequent years, and are of minor importance in evaluating incidence and mortality in 1988–92.

6.3.2 Post-screening incidence and mortality: women aged 50+ in 1990

For women aged 50+, the predicted incidence and mortality around 1990 are fairly good (*Figure 6.3d* and *6.3e*). They are not statistically different from observed incidence and mortality, except for a too low simulated incidence in age-group 65–69.

In the intermediate period (1978–1982) *(see Figure 6.3 c)*, mortality is predicted correctly in age-group 40–49, but too high in age group 50–69. Since there are no incidence data from this period, we cannot discern between too little incidence reduction or too little improvement in prognosis. A possible explanation is that the simulated screening activity in the 1970s was too low for women born before 1930. Correction of the divergence would need 50% instead of 30% of women with a first smear by 1976 in women born between 1919 and 1928 *(see Table 6.3)*, and 21% instead of 9% in women born before 1918. Another explanation would be that the simulated short term screening effect is too small. From the parameters we investigated in this analysis, i.e. the duration of the detectable pre-clinical phase and the sensitivity of the test, the duration is of little help in this respect. A longer duration hardly affects the short term effects of screening on mortality. A shorter duration decreases the effect on mortality, particularly the long term effects, but also the short

Figure 6.5 *Observed and simulated cervical cancer mortality/incidence ratio for the Netherlands in a prescreening period and in a recent period. Pre-screening period: 1968–1972/1965–1969 (mort/inc). Recent period: 1988–1992/1989–1990 (mort/inc). Observed data: CBS 1994 [37] (mortality), CCR 1973 [40] and NCR 1992 [153]and 1993 [154](incidence)*



term effects. A higher sensitivity on the other hand, will increase the short term effect on mortality (and also the long term effects, but to a smaller extent because of repeated screening). To explain the total difference between predicted and observed mortality around 1980 in this way, an unrealistic 100% sensitivity is required.

In *Figure 6.4*, mortality trends are shown by cohort. The solid lines again denote the observed mortality (bold line with markers) and the model-predicted mortality (thin line). The cohort on top (cohort 1920–39) corresponds with the women aged 50–69 in 1990. This cohort is especially interesting: it has the highest screening exposure in the 1970's *(see Table 6.3)*, and it also has enough pre-screening mortality to assess its background risk accurately. For this cohort, the predictions are close to the observations, except for a slightly higher initial level and a somewhat more distinct decrease over time in the simulation. The predicted trend without screening effect (dashed line) clearly differs from the observed trend; its initial increase reflects the age-pattern of cervical cancer mortality. The model predicts a 37% reduction in mortality in 1990 in this cohort.

6.3.3 Post-screening incidence and mortality: women aged <50 in 1990

In women under age 50, the predicted incidence and mortality around 1990 are clearly too low *(see Figure 6.3 d* and *6.3e)*. One could suppose that this deviation in the predictions is due to a too strong reduction in incidence and mortality caused by screening. However, even rates predicted without screening effect (dashed lines) are too low.

The same discrepancies appear in the cohort-wise representation of mortality in the <50 women in (*see Figure 6.4*, cohort 1940–49). This cohort corresponds to women age 40–49 in 1990, and is the youngest cohort for which pre-screening mortality data were available. From 1973–77 onwards, the simulated mortality rates are clearly too low and even the simulated rates without screening effect are below or at most equal to the observed level. In 1968–1972, the predicted mortality is too high, but the observed rate is only based on 5 cases. After this initial period, there is a remarkable parallelism between the observed and predicted (with screening effect) lines. Together with the fact that the discrepancies between observed and predicted incidence rates follow the same pattern, it corroborates the hypothesis that the discrepancies could be most parsimoneously explained by a higher risk for the cohort concerned than estimated.

The predicted and observed mortality to incidence ratios for 1990 *(see Figure 6.5)* reflect the effect of screening on lethality. The agreement between prediction and observations is good for ages under 60. The model predicted lethality is too high in age-groups 60–69 (as it was already in the pre-screening period) and 70–79 (where screening has had little effect). The mortality to incidence ratio in the predictions without screening, effect where there is no improvement in prognosis in early detected invasive cancers, is too high.

6.4 Discussion

The analysis of the British Columbia data that resulted in the parameter estimates describing the natural history, was based on clinical incidence before screening started and on detection rates by rank and interval since last screening. Gustafsson et al deduced their assumptions on natural history from trends in incidence and mortality, accounting for screening practice (using incidence of carcinoma in situ as a proxy measure for screening intensity) [83]. The results of these two analyses are consistent. The parameter values derived by Gustafsson (duration of the pre-invasive detectable stage 13.3 years, duration of invasive pre-clinical stage 4 years) are similar to ours (11.8 years and 4 years respectively). In an earlier analysis [161], we found these quantifications for duration of and test sensitivity for pre-invasive disease to be consistent with international data on interval cancers after negative smears [100].

In the present analysis, these assumptions on duration and sensitivity have been tested against screen-affected incidence and mortality trends in the Netherlands. Such an independent test of a cervical cancer screening model describing the natural history of cervical cancer has not been published before.

In estimating age- and cohort-factors in the risk of dying from cervical cancer, we assumed that no period trend occurred in this risk during the period considered (from 1950 until the start of screening). We already mentioned that in Sweden, although not in all regions, there was some improvement in survival after 1950, suggesting a downward period trend [199, 172]. There was no change in stage distribution, but the stage specific survival improved, presumably by an improvement in treatment. We could have explained part of the downward trend in pre-screening mortality in the Netherlands after 1950 by a period trend. When this trend is assumed to have continued until the end of the period for which we made predictions (1992), predictions will not be affected. Only if we assume that the improvement in treatment of clinical cancer stopped after e.g. 1970, our predictions are affected. The predicted incidence and mortality level will become higher in 1980 and more pronounced so in 1990. Hence, the difference between the observed rates and predicted rates without screening effect will be larger. The predictions with screening effect will still be closer to the observations than the predictions without screening effect, but they also will be too high.

The baseline estimates of the cohort risks in women born between 1940 and 1949 were based on a reasonable number (41 cases) of cervical cancer deaths in the pre-screening period up to 1975. The marked difference between predicted and observed time trend for this cohort as shown in *Figure 6.4* strongly suggest that the underlying cervical cancer risk increases over time. It could be that in this period and in this cohort, period and cohort factors have not acted independently. Tentatively, knowing the role of sexual behaviour and genital HPV infection as a risk factor for cervical cancer, it could be argued that this increase is related to the marked increase in STD's in the late 1960s and in the 1970s (following introduction of oral anticonceptives) [33, 48]. Transmission of STD's is much lower in older age groups, and it seems reasonable to assume that this sexual revolution has had only limited impact on the cohorts born before 1940.

Estimation of cohort factors for women born after 1950 is impossible because of insufficient baseline data from the pre-screening era (only 2 cervical cancer deaths before 1975). Cohort trends from the turbulent pattern in the pre-1950 cohorts can be extrapolated to more recent cohorts in many different directions, and has a strong influence on the predictions for younger age groups, up to age 40 in the more recent periods considered. Indeed, the predictions differ considerably from observed data and the comparison for 1988–1992 mortality suggest that in these cohorts the underlying risk of cervical cancer was also higher than the very low level at young age in women born between 1940 and

Table 6.3 *Assumed fractions of the Dutch female population that had a first smear by birth cohort and calendar year. All the presented fractions have been incorporated in the model (between the calendar years presented, linear intrapolation is used)*

Year of birth	1976[1] % (mid-age)	1981[3] % (mid-age)	1988[3] % (mid-age)	1991[2] % (mid-age)
1889–1918	9% (71)[4]	14% (73)[4]	16% (77)[4]	17% (78)[4]
1919–1928	30% (53)	44% (58)	52% (65)	56% (68)
1929–1938	39% (43)	52% (48)	63% (55)	76% (58)
1939–1948	28% (33)	48% (38)	66% (45)	83% (48)
1949–1958	10% (23)	33% (28)	57% (35)	78% (38)
1959–1968	0% (13)	5% (18)	18% (25)	53% (28)
1969–1978	0% (3)	0% (8)	0% (15)	9% (18)

[1] Based on survey data from a starting mass screening program in three regions of the Netherlands (24% of the Dutch population) (EVAC [67])

[2] Based on data from the Dutch National Health Survey, in which 2700 women over age 16 reported on their PAP smear consumption (de Bruin e.a. [29]) (see text)

[3] Intrapolated between [1] and [2], accounting for the varying intensity of programme screening over the calender years (see text)

[4] Calculated with 84 years as upper age limit for getting a first smear

1949. For the birth cohort 1940–1954, calibration of the risk to an exact fit of cervical cancer mortality in 1988–1992 leads to a clearly higher level than shown in *Figure 6.2*. This higher level is comparable to the risk of women born in 1930–1939, but still lower than those born in the twenties.

When not accounting for any incidence and mortality reduction caused by screening, the fit between observed and predicted incidence and mortality is much poorer than when accounting for the assessed screening effects. The mortality reduction from screening is estimated to be 32% for the Netherlands in 1992. The reduction is primarily limited to this figure because of a lack of coverage. Many women, especially at older age, did not have previous screening at all or only very infrequently. Moreover, in the model, it is assumed that 10% of women will never participate in screening and these women have a threefold risk compared to participants. Coverage has increased in more recent years, but it will take more time before this can be fully expressed in mortality reduction. Using the model described in this article, we have also predicted mortality reduction in a cohort of women who all participate in screening from age 30 to 60 every 5 years. The predicted reduction in cervical cancer mortality is 75%. The 25% residual mortality is mainly due to the fact that there is no screening over 60 (17%) years of age, and the remaining eight percent is devided between no screening under age 30 (3%), imperfect test sensitivity (3%) and screening every 5 years instead of annually from age 30 to 60 (2%).

Table 6.4 *Observed and simulated incidence and mortality rates per 105 women for various periods between 1965–1992. Simulated with the model as described in the Methods and Materials and accounting for screening effect. The figures correspond with* Figure 6.3

Age groups	20–24	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69	70–74	75–79	80–84
Prescreening period													
INCIDENCE													
Period 1965–1969													
Observed rates[1]	0.20	1.49	10.4	20.3	44.3	49.9	49.0	33.5	40.7	29.2	28.0	25.9	na
Model rates	0.27	3.03	10.9	21.9	46.5	54.7	47.0	33.4	26.8	28.0	29.4	30.5	25.5
Observed cases	1	7	27	57	133	153	154	108	118	75	58	39	na
MORTALITY													
Period 1968–1972													
Observed rates	0.01	0.25	1.4	4.4	8.1	14.6	18.2	16.6	18.2	18.9	21.2	24.9	31.2
Model rates	0.05	0.45	2.0	4.1	10.4	15.5	17.0	17.0	16.9	19.3	20.4	21.1	23.5
Observed cases	0	6	27	84	153	277	307	271	275	247	219	180	129
Period after the introduction of screening													
INCIDENCE													
Period 1989–1990													
Observed rates	1.21	4.91	12.5	18.5	15.3	13.4	15.2	10.5	15.3	20.5	21.4	19.5	17.6
Model rates	np	np	np	np	8.6	9.2	12.7	11.2	15.9	16.4	21.0	ni	ni
Observed cases	15	62	148	208	175	117	117	77	108	137	117	92	59
MORTALITY													
Period 1978–1982													
Observed rates	0.14	0.71	1.1	2.7	3.6	6.6	9.3	10.5	13.2	12.8	16.3	17.3	19.1
Model rates	np	np	0.8	1.5	4.0	5.7	11.2	12.9	14.7	ni	ni	ni	ni
Observed cases	4	20	32	61	72	125	174	193	212	191	210	171	118
Period 1988–1992													
Observed rates	0.06	0.57	1.7	2.8	3.0	3.6	4.2	5.3	8.0	11.6	13.0	15.0	18.3
Model rates	np	np	np	np	1.6	2.2	3.7	4.7	8.5	11.0	14.8	ni	ni
Observed cases	2	18	51	79	86	81	82	98	141	196	181	188	154

[1] Standardized, see text

[na] Not available

[np] No prediction, cohorts without pre-screening mortality data

[ni] No independant data: the observed data are used for the calibration of the model (see the description of the APC-analysis in the Methods and Materials section)

In conclusion, this analysis provides further support for the estimated duration of detectable pre-clinical cervical cancer of around 16 years, for the estimated sensitivity of the PAP smear for this pre-clinical phase of around 80% and a cure rate for pre-invasive disease of close to 100%. The present model can therefore, in our opinion, be used for assessment and comparison of the effects and costs of different screning strategies. For cos-effectiveness calculations, also regressive lesions should be added. For application in other countries, the Dutch data on incidence level and cohort-specific risks, demography, hysterectomies for other reasons and screeing pattern have to be replaced by the pertinent ones for the region concerned.

7

The effects and costs of cervical cancer screening: the scientific background to lengthening the pap smear screening interval from 3 to 5 years in the Netherlands

This chapter is based on the results presented in: *Cervical cancer screening: Age-ranges and intervals. An updated cost-effectiveness analysis* (in Dutch). Ballegooijen M. van et al. 1993. Report Erasmus University Rotterdam.

Abstract

Objective: As part of a total re-evaluation of the Dutch cervical cancer screening programme, new guidelines were developed for the age-range and intervals for PAP smears. To this end, a cost-effectiveness analysis was conducted. Although the Government stated that the existing number of 7 screens per woman should be maintained, other numbers have also been evaluated in order to be able to assess incremental cost-effectiveness.

Methods: A validated cervical cancer screening model was implemented in the MISCAN micro-simulation programme for cancer screening evaluation. Predictions of the effects and costs of different combinations of age-ranges and intervals were made for policies with 5 to 10 PAP smears per woman and the policy implications considered.

Results: For 7 screenings per woman, the optimal interval was estimated at seven years and the optimal age-range at 27–69 years. The incremental costs per life year gained compared with an optimal policy of 5 screenings were approximately Dfl 35 000. Even for 10 screenings, the optimal interval did not fall below five years because a broad age range appears to be more important for effectiveness than a short interval between screenings. At a Dutch national consensus meeting, a 5-year interval was considered the maximum acceptable length for PAP smear screening. Within this constraint, the optimal age-range for 7 screenings is from ±30 tot ±60 years.

Conclusions: When seven PAP smears are offered per woman during her life, the optimal screening interval is estimated at 7 years. The interval in the new national guidelines has been reduced to five years because of the expected lack of acceptability of longer screening intervals.

7.1 Introduction

The European Committee's guidelines [45]advise that women should be offered PAP smears for cervical cancer screening every 3 to 5 years. The guidelines of the US preventive Task Force [203] recommend an interval of 3 years. Until recently, in North and West Europe, only Finland [96] and parts of the United Kingdom [158] have a 5-year interval. In the Netherlands, the screening interval has been 3 years since programme screening started in 1976. Recently, it was decided to change this to a 5-year interval, starting from 1996 [59]. In this chapter, the analysis that corroborated this decision and its role in the discussion will be described.

Before 1996, the cervical cancer screening policy in the Netherlands was based on offering 7 PAP smears per woman in the age range 35–53 years. The change in both the interval and age range was part of a total re-evaluation of the management of the national cervical cancer screening programme, including follow-up strategies, organisation, evaluation and financing. This re-evaluation began in 1991 and resulted in new guidelines in 1993 [59]which were ratified by the Minister of Health in 1994.

There were several reasons to reconsider the age-range and interval. For example, many clinicians argued that the starting age of 35 years should be lowered. Since they were generally not in favour of lengthening the 3-year interval period, this would have implied an increased number of screenings per woman and an increase in costs of around 30%. Cost effectiveness calculations in 1988 estimated the costs per life year of efficient policies incorporating 7 screening invitations at approximately Dfl 25 000 [118]. The incremental costs-effectiveness ratio of moving to a policy with more screenings would be much higher. This compared unfavourable with the estimated Dfl 8 000 per life-year gained in breast cancer screening as performed in the Netherlands (from 50–70 years, every 2 years) [116]. The Government did not agree that the cervical cancer screening programme should be intensified and proposed that a programme with at most 7 smears per woman should be financed. In practice, this meant that exactly 7 smears per woman would be performed as the parties involved were not in favour reducing this number.

Although the Government had imposed a policy constraint of 7 smears per woman, the question remained as to what were good combinations of age-range and interval from a cost-effectiveness perspective. Using newly available data

we estimated the incremental cost-effectiveness of policies using other numbers of smears and also considered policies which incorporated 5–10 screenings.

We used the micro-simulation programme MISCAN to model cervical cancer screening and predict the costs and effects for different age-range and interval combinations [88]. The cervical cancer screening model is an extension of the model used for analyzing cervical cancer screening data from British Columbia [159] and was adapted to the situation in the Netherlands (see [12] and Chapter 6).

A problem with the evaluation of screening according to an agreed policy is that in the Netherlands, as in many other European countries, 'opportunistic' preventive smears are being carried out in addition to the organized, invitation-based, screening programme. In the cost-effectiveness calculations, we have chosen to consider the situation without opportunistic screening. Quality of life considerations are also not accounted for in the analysis. The possible consequences of these simplifying assumptions for the results and conclusions will be discussed.

7.2 Materials and methods

7.2.1 Cost-effectiveness calculations

The costs and effects were estimated for a screening programme which started in 1993 and is intended to run until 2020 (27 years) in the Dutch population. We assumed a situation without opportunistic screening and with an attendance rate of 75% in women at risk (with cervix uteri) until age 50, decreasing with 0.5% per year after age 50. This is in accordance with coverage of the total PAP smear activities observed in the Netherlands in 1994 [14]. The simulated effects have been accounted for until all women who could have benefited from the programme died. Both costs and effects were discounted to 1993 (the year in which the predictive calculations start) at a rate of 5% per year.

7.2.2 Assumptions determining the effects of screening

A description of the cervical cancer screening model and how it was tested against observed trends in cervical cancer incidence and mortality in the Netherlands between 1965 and 1992, is given in Chapter 6. There are some differences between the model used in the present chapter and the model described in Chapter 6. These differences are described in Appendix A.

Important assumptions in the model are that the average duration intraepithelial neoplasia) is 11.8 years, that the average duration of pre-clinical invasive cervical cancer is 4 years, and that the sensitivity for pre-invasive cancer is 80% [159].

Table 7.1 *Costs of assessment and treatment of screen detected and clinically diagnosed cervical neoplasias per case by stage, in Dfl*

Stage	Costs
CIN	3 500
FIGO IA (micro-invasive)	9 550
FIGO IB (local)	20 250
FIGO II+ (non-local) screendetected	19 100[1]
FIGO II+ (non-local) clinically diagnosed	17 450[1]

[1] The difference in costs for screen detected and clinically diagnosed stage II+ cervical cancers is caused by the less favourable stage-distribution in clinically diagnosed cases, generating a lower number of radical hysterectomies and a higher number of radiotherapeutic treatments. The costs of the latter treatment are about 35% lower

The risk in the youngest cohorts (born after 1938) could not be assessed from pre-screening mortality data (Chapter 6). Using the model, the cohort risks were calibrated to the observed mortality in these cohorts between 1988 and 1992.

The main unfavourable side effects of cervical cancer screening are caused by false positive test-results and detection of regressive disease. Regarding false-positive test-results we assumed that 5% of attending women on average need 1 repeat smear before they return to the regular screening schedule, and that 1.6 per thousand attending women are referred for a colposcopy and a biopsy after which no cervical neoplasia is found. We assumed that 70% of the CIN lesions that start before age 35 will regress spontaneously, and that this percentage is lower (40%) between age 35 and 54. All neoplasias starting after age 54 are assumed to be progressive. All these assumptions on false positive rates and regression are based on positive rates, follow-up practice and detection rates for histologically confirmed neoplasias from the period before 1985 (see Appendix A).

We assumed an attendance rate of 75%; 10% of the population never attends and has a risk 3 times higher than the attenders.

The effects of screening policies are predicted for screening in the period 1993–2019, and related to the situation without screening from 1993 onwards. The impact of screening activities in the Netherlands before 1993 was taken into account. The number of life-years gained was considered as the effect measure. The number of smears was used as a proxy for the costs of a screening policy. This simplified approach could be adopted because we had previously found an almost linear relation between total costs (including changes in costs of diagnosis and treatment of cervical cancer) and number of smears [90]. Predictions have been made for about 80 policies with different age-ranges, intervals and numbers of smears per woman *(see Table 7.9)*. The evaluated policies have

been compared to identify efficient policies. A policy is considered efficient when there is no alternative policy offering the same or lower costs and more life-years gained, or with lower costs and the same amount of life-years gained. This means that in *Figure 7.1*, for a policy to be efficient there should be no other policies in the upper-left quadrant of the cost-effectiveness point. For the efficient policies, detailed cost calculations have been made on the basis of the following cost assumptions.

7.2.3 The costs

Resource costs were assessed for PAP smears, colposcopic evaluations and radiotherapeutical treatments. For the other relevant medical procedures (e.g. biopsies, local treatments including cryocoagulation, conisations, and total and radical hysterectomies), tariffs charged in the Netherlands were used [12]. The costs are presented in Dutch Guilders (price level 1993).

The fixed costs for coordinating and evaluating the cervical cancer screening programme at a national, regional and local level were assumed to be Dfl 6 mln annually. The other costs of the screening programme are variable costs and were estimated at Dfl 58 per smear. This includes the costs for the personal invitation (Dfl 1.80), registration in the PALGA (Dutch Network and National Database for Pathology) (Dfl 2.60), material (Dfl 1.00), taking of the smear (Dfl 16.00), time (Dfl 5.30) and travelling (Dfl 2.50) for the woman, the cytologic evaluation (Dfl 27.30), and the costs of 2% of smears that are repeated because of insufficient quality. The costs of the cytologic evaluation depend on the number of smears evaluated annually in the laboratory. The weighted average cost per smear for the 63 laboratories based on approximately 500 000 smears annually was Dfl 27.30.

The corresponding figure based on 250 000 smears was Dfl 30.70, and for 750 000 smears Dfl 26.50.

Other costs accounted for were Dfl 88 per repeat smear after a positive smear, and Dfl 870 per referral to the gynaecologist in false positive cases. The costs of assessment, treatment and follow-up procedures after treatment per woman with CIN or invasive cervical cancer are presented in *Table 7.1*. The costs of treatment of recurrence and palliative care per woman who dies from cervical cancer were assessed at Dfl 30 660 [11]. Indirect costs (e.g. loss of productivity) other than for the screening itself (see the above mentioned Dfl 5.30) have not been taken into account.

Table 7.2 *Efficient combinations of age-range and interval with 5 to 10* PAP *smears, and of the new and the old policy of the Dutch screening programme*

Annual number of Pap smears*	Number of screenings	Starting age	Last age	Age-range	Interval	Average screening age	Label in Fig.7.1
Efficient policies							
316 000	5	27	67	41	10	47	=
382 000	6	27	67	41	8	47	}
438 000	7	27	69	43	7	48	{
502 000	8	27	69	43	6	48	]
609 000	10	27	72	46	5	49.5	[
Dutch policy implemented in 1996							
478 000	7	30	60	31	5	46	G
Dutch policy used until 1996							
505 000	7	35	53	19	3	35	U

* Assuming 75% attendance rate

7.3 Results

The number of primary smears (as a proxy for the costs) and life-years gained for the respective combinations of age-ranges and intervals *(see Table 7.9)* are shown in *Figure 7.1*. In this figure, the efficient policies are indicated by an arrow. These efficient policies are described in *Table 7.2*. Their distinguishing characteristic is a broad age-range. Compared to less intensive efficient screening policies (with 5 or 6 screening-ages), the main difference of more intensive efficient screening policies is a shorter interval. In all efficient policies, screening starts at age 27, and the age of the last screening is hardly affected by the number of screenings per woman. The fact that the starting age of 27 did not change at all might have been influenced by the fact that we did not make predictions for screening in women older than 69 years (except for policies with 10 screenings), thus also limiting the starting age for the respective combinations of intervals and number of smears.

The costs of the efficient policies and of the old and new Dutch policy (compared to the situation without screening) are given in *Table 7.3*. The cost of screen-induced follow-up in false positive and non-invasive cases is largely compensated for by the savings from prevented treatment in invasive cases and care for advanced disease. The total extra costs compared to the situation without screening mainly consist of the costs of screening. In *Tables 7.4* (no discounting) and *7.7* (5% discounting), the effects of the policies are presented. The predicted nondiscounted number of deaths from cervical cancer prevented per screening year is 150, 180 and 210 respectively for the three policies pre-

sented. Discounting at 5% will have a much larger impact on health effects than on costs. It will reduce the costs with 20–25%, but the lives saved with 66% and the life-years gained even with 82%. The cost-effectiveness ratios are given in *Table 7.5*. The costs per life-year gained increase gradually from Dfl 22 000 per life year gained in the case of 5 screens to Dfl 28 000 in the case of 10 screens. The incremental costs per life-year gained when moving to a more intensive policy, increase much more steeply, from approximately Dfl 35 000 (from 5 to 7 smears) to approximately Dfl 56 000 (from 7 to 10 smears).

7.3.1 Sensitivity analysis

For the policies with 7 screenings, we explored the influence of the following key assumptions on the optimal combination of age-range and interval: i.e. the average duration of detectable preinvasive disease, the false negative rate of the PAP smear for preinvasive disease, attendance, the risk in the youngest cohorts, and the discount rate. For each parameter, high and low values as listed in *Table 7.6* were explored. For policies with 7 screenings, the age-range and interval of efficient combinations were only slightly affected, with the following exceptions.

The length of the optimal interval is influenced by the average duration of the pre-invasive stage and by the attendance rate (both a shorter duration and a lower attendance make shorter intervals relatively more attractive). When the assumed average duration of pre-invasive disease is shortened by 5 years (6.8 instead of 11.8), the cost-effectiveness of policies with 5 and 7-year intervals become similar. With an attendance as low as 40%, the difference in cost-effectiveness between 5 and 7-year interval policies also decreases. The optimal starting age is influenced by the average duration of the pre- invasive stage, by the assumed risk in the youngest cohorts, and by the discount rate. Assuming a shorter pre-invasive duration favours higher starting ages, shifting the optimal age-range slightly upwards. The starting age is about 5 years younger when women born after 1939 are taken into account as they have a risk three times higher than the base-line assumptions. Such a high risk is not very likely in the Netherlands (see Chapter 10). With no discounting, the optimal starting age dropped by about 3 years.

7.3.2 The role the results played in the decision on the new guidelines

The decision on the new guidelines for age-ranges and intervals was made in a national consensus meeting where all professionals and institutions involved (general practitioners, pathologists, gynaecologists, epidemiologists, municipal and regional health care centres, and the government) were represented [59]. The results reported above were presented at the meeting, pointing out that the efficient interval for a policy with seven screenings is estimated at seven

Table 7.3 *Costs in Mln Dfl of efficient combinations of age-ranges and intervals with increasing number of smears, and for the new (labelled G in* Figure 7.1*) and old (labelled U in* Figure 7.1*) nationally recommended age-range and interval combination, for a 27 year cervical cancer screening programme in the period 1993–2020. Further assumptions: attendance 75%, no opportunistic screening, 5% discount rate for effects and costs*

	..... Efficient			..... Used		
Label Figure 7.1	=	{	[	G	U	
Number of smears	5	7	10	7	7	No screening
Ages[1]	27(10)67	27(7)69	27(5)72	30(5)60	35(3)53	
Annual number of smears (0% disc.)	316 000	438 000	609 000	478 000	505 000	0
Costs						
Screening	412	527	683	564	590	0
Repeat smears	24	34	45	36	38	0
Referred, no CIN +[2)]	7	10	10	10	10	0
CIN[2]	56	68	83	77	62	0
Invasive carcinoma[2)]	202	187	177	184	196	264
Advanced disease[3)]	150	141	133	147	154	193
Total costs	851	968	1131	1018	1050	457
Differences with the situation without screening						
Screening	412	527	683	564	590	
Repeat smears	24	33	45	36	38	
Referred, no CIN +[2)]	7	10	14	11	12	
CIN[2]	56	69	79	66	51	
Invasive carcinoma[2)]	-63	-76	-87	-71	-58	
Advanced disease[3)]	-43	-52	-60	-45	-39	
Total costs	394	511	674	561	594	

[1] Starting age (interval) last age

[2] Costs of assessment and treatment

[3] Costs of treatment of recurrence and palliative care in women who die from cervical cancer

years and the optimal age range at 27–69 years. However, most attenders to the meeting considered 5 years to be the maximum acceptable length for the interval between subsequent PAP smears. It was envisaged that a longer screening interval would lead to many extra smears between programme screenings. Because the policy should incorporate 7 smears, discussion focused on the optimal starting and finishing age between first and last screening of 30 years (5 years interval × 6 intervals per woman). It appears that the starting age goes up and the last age is down relative to the optimal age-range of 27–69 years. The cost-effectiveness results are in favour of policies starting at around age 30 to age 40 *(see Figure 7.1)*. Over this range, the cost-effectiveness ratio seems relatively stable.

Table 7.4 *Effects and costs of efficient combinations of age-ranges and intervals with increasing number of screenings, and for the new (labelled G in* Figure 7.1*) and old (labelled U in* Figure 7.1*) nationally recommended age-range and interval combination, for a 27 year cervical cancer screening programme in the period 1993–2020. Costs in Dfl. Further assumptions: attendance 75%, no opportunistic screening, 0% discount rate*

	Efficient			Used		
Label Figure 7.1	=	{	[	G	U	
Number of smears	5	7	10	7	7	No screening
Ages[1]	27(10)67	27(7)69	27(5)72	30(5)60	35(3)53	
Annual number of smears (0% disc.)	316 000	438 000	609 000	478 000	505 000	0
Costs						
Total costs (× Mln)	1 939	2 100	2 339	2 182	2 258	1 442
Effects						
Deaths from cervical cancer	16 611	15 799	15 057	16 169	16 831	20 731
Lost life-years from cervical cancer	310 211	291 917	277 840	294 179	308 247	413 003
Differences with the situation without early detection						
Total costs (× Mln)	497	659	898	740	816	
Deaths prevented	4 120	4 932	5 674	4 563	3 900	
Life-years gained	102 792	121 086	135 163	118 824	104 756	

[1] Starting age (interval) last age

More opportunistic screening takes place in young age groups when programme screening starts at age 35 compared to age 30. Therefore, the policy offering PAP smears every 5 years for women between ages 30 and 60 was chosen for the revised national cervical cancer screening programme.

After the consensus meeting, we reconsidered policies incorporating 7 screenings per woman and an interval of 5 years, and calculated the cost-effectiveness when accounting for the true costs instead of the number of smears *(see Table 7.8)*. This re-analysis confirmed that the comparison between the policies had been only slightly influenced by taking number of smears as a proxy for the total net costs: the difference in costs per smear between the 'youngest' policy (starting at age 20 years) and the 'oldest' policy (starting at age 40 years) was less than 2%. The random error of the incremental cost-effectiveness ratio's of policies with different starting ages is rather large (see the bottom row of *Table 7.8*). Nevertheless it seems that changing from the policy that starts at age 30 to one starting at a younger age is relatively unfavourable. The cost-effectiveness ratio of policies that start at age 30, 35 and 40 are not very different. Of these policies, if one accepts this level of cost-effectiveness, the logical policy is the one with the largest effects, i.e. the policy that starts at age 30 years.

Table 7.5 *Cost-effectiveness ratios (CERs) of efficient combinations of age-ranges and intervals with increasing number of screenings, and for the new (labelled G in* Figure 7.1*) and old (labelled U in* Figure 7.1*) nationally recommended age-range and interval combination, for a 27 year cervical cancer screening programme in the period 1993–2020. Costs in Dfl. Further assumptions: attendance 75%, no opportunistic screening, 5% discount rate*

	Efficient			Used	
Label Figure 7.1	=	{	[	G	U
Number of smears	5	7	10	7	7
Ages[1]	27(10)67	27(7)69	27(5)72	30(5)60	35(3)53
Annual number of smears (0% disc.)	316 000	438 000	609 000	478 000	505 000
CERs compared with the situation without screening					
Costs per death prevented	287 274	308 624	348 888	383 329	477 361
Costs per life-year gained	21 854	23 840	27 679	27 602	33 276
CERs compared with the efficient policy with less smears					
Incremental costs (× Mln)	n.a.	117	163		
Incremental deaths prevented	n.a.	285	276		
Incremental life-years gained	n.a.	3 416	2 914		
Incremental costs per l.y.g.	n.a.	34 250	55 937	Not efficient	
95% CI because of of stochastic output	–	32 000 to 37 000	51 000 to 62 000		

[1] Starting age (interval) last age
n.a. Not available

7.3.3 The new policy compared to the old policy and the optimal policy

In comparison with the previous policy, it is predicted that the new policy, using a 5% discount rate, will result in 18% more deaths being prevented and 14% more life years gained *(see Table 7.7)*. In addition costs will be 6% lower. Compared to the optimal policy of 7 PAP smears per woman, the number of deaths prevented for the new policy is 12% lower, the number of years gained 5% lower and the costs 10% higher. According to our predictions, this suggests that while the new policy is not optimal it marks a considerable improvement over the original policy.

7.4 Discussion

In practice, opportunistic screening will certainly occur. One reason why we did not account for this in our predictions, was that it is instructive to explore the cost-effectiveness of pure policies. But the main reason was that it is

Table 7.6 *Sensitivity analysis: baseline, low and high assumptions for selected parameters*

Parameter	Baseline assumptions	Low assumption	High assumption
Average duration of pre-invasive detectable neoplasia	11.8 years	6.8 years*	16.8 years*
Sensitivity of the screening test (pre-invasive neoplasia)	80%	60%	90%
Attendance	75%	40%	90%
Cumulative incidence of progressive neoplasia in participating women born after 1938	1.2%	–	Three times baseline
Discount rate	5%	0%	–

* The disease onset is shifted to five years older and younger respectively to compensate the influence of the changes in duration on the age of incidence of invasive cancer and mortality

very difficult to predict the interaction between an organized policy and the frequency and age-distribution of opportunistic screening. In the decision process as described, the likely impact of opportunistic screening has tentatively been taken into account. A longer interval than 5 years was expected to generate much more opportunistic 'in between' screening, and a starting age over 30 to generate a large amount of opportunistic screening in young age.

When the number of PAP smears per woman is given, a screening policy is characterized by the starting age and the length of the interval between subsequent screenings. We will discuss possible influences of the methods we used and of remaining uncertainties about parameter quantifications on the conclusions reached on starting age and interval.

7.4.1 The length of the screening interval

The optimal screening interval, for a given number of screenings, depends to a large extent on the assumed duration of detectable pre-invasive disease (CIN) before it becomes invasive. The estimate of 11.8 years that we used for the average duration of CIN is based on analysis of data from British Columbia [161] and tested against data on interval cancers from the IARCstudy [161, 100]. Our estimate is intermediate compared to estimates by Brookmeyer and Day in 1987 [28] (6.4 to 28 years including invasive pre-clinical duration), Gustafsson in 1989 [83] (13.3 years for CIS) and Bos in 1997 [24] (16 years including invasive pre-clinical duration). When the average pre-invasive duration was shortened to 6.8 years, the optimal interval between seven screenings still did not drop under 5 years.

Table 7.7 *Effects and costs of efficient combinations of age-ranges and intervals with increasing number of screenings, and for the new (labelled G in* Figure 7.1*) and old (labelled U in* Figure 7.1*) nationally recommended age-range and interval combination, for a 27 year cervical cancer screening programme in the period 1993–2020. Costs in Dfl. Further assumptions: attendance 75%, no opportunistic screening, 5% discount rate*

	..... Efficient			..... Used		
Label Figure 7.1	=	{	[	G	U	
Number of smears	5	7	10	7	7	No screening
Ages[1]	27(10)67	27(7)69	27(5)72	30(5)60	35(3)53	
Annual number of smears (0% disc.)	316 000	438 000	609 000	478 000	505 000	0
Costs						
Total costs (× Mln)	851	968	1 131	1 018	1 050	457
Effects						
Deaths from cervical cancer	4 847	4 562	4 286	4 754	4 975	6 219
Lost life-years from cervical cancer	51 833	48 417	45 503	49 534	52 029	69 868
Differences with the situation without early detection						
Total costs (× Mln)	394	511	674	561	594	
Deaths prevented	1 372	1 657	1 933	1 464	1 244	
Life-years gained	18 035	21 451	24 365	20 334	17 839	

[1] Starting age (interval) last age

7.4.2 The starting age in screening

The optimal starting age is influenced by the discount rate chosen. Without any discounting, the estimated optimal starting age of 27 years *(see Table 7.2)* would drop approximately 3 years. Discounting means that effects and costs that occur in the future have less impact on the calculated cost-effectiveness. Discounting therefore decreases the advantage of presenting a cervical cancer death in a younger than in an older woman, because the life-years gained far in the future count less. Moreover, the time lag between the detection of pre-invasive disease and the time the woman otherwise would have died from the neoplasia, is on average longer in younger age groups than in older ones. Discounting is generally advised for cost-effectiveness analyses, although recently there has been a tendency to recommend a lower rate of 3% instead of the usual 5% discount rate [210]. This again supports the decision to chose the 30 to 60 years policy instead of a policy with 7 smears and a 7 year interval that starts at age 35 or 40 years.

7.4.3 Quality of life

Inclusion of quality of life aspects, which have not been considered in this analysis, would influence the estimated optimal age-range. Compared to screening at older age, screening at a younger age would, according to our analysis, generate more negative as well as more beneficial side effects (other than mortality reduction). One important factor in this respect is the assumed higher occurrence of regressive disease at younger age. This assumption was based on the analysis of the British Columbia screening data [159], but was not confirmed by other analyses [83, 24]. More onset of regressive lesions at younger age means more detection, diagnosis and treatment of lesions without preventing invasive cancer or death, and therefore more loss in quality of life (and more costs). On the other hand, the high overall survival at young age is favourable for extra gains in quality of life in younger age: relative to the mortality reduction, the reduction in incidence is higher than in older age. Assessment of the quality of life impact on the optimal starting age requires health status measurement and valuation of the burden for women caused by the detection of pre-invasive disease and by the diagnosis of invasive cervical cancer (see also Chapter 10).

If the number of screenings per woman is fixed, accounting for quality of life impacts of screening would hardly have any impact on the optimal length of the interval.

7.4.4 Cost-effectiveness level

For assessing whether cervical cancer screening is a good investment, the (incremental) costs per life year gained need to be compared with other health care interventions. Therefore it is important to know how sensitive the estimates for the cost-effectiveness of cervical cancer screening is for the assumptions made. A very important assumption in this respect is the discount rate applied. Since there is a long duration between costs and effects in cervical cancer screening, the costs per life-year gained increase with an increased discount rate. For instance, at 5% discounting, the costs per life-year gained of the new Dutch policy is estimated at Dfl 28 000. This compares with Dfl 16 000 at 3% discounting and only Dfl 6 000 at a zero discount rate. In general, time preference is unfavourable for early detection when it is compared to curative medical interventions, where the time lag between the intervention and its benefits is usually much shorter (e.g. treatment with antibiotics in acute conditions). The unfavourable influence of time preference is especially strong in cervical cancer screening, with its long detectable pre-clinical phase of about 15 years on average.

The estimated cost-effectiveness level is also influenced by the assumptions regarding attendance. This in part is due to the fixed costs of running the programme and economies of scale. The association between attendance and

Table 7.8 *Effects, costs and cost-effectiveness ratios (CERs) of 5 policies with 7 screening-ages and an interval of 5 years, for a 27 year cervical cancer screening programme in the period 1993–2020. Costs in Dfl. Further assumptions: attendance 75%, no opportunistic screening, 5% discount rate*

Label Figure 7.1	.	9	G	e	a
Starting age	40	35	30	25	20
Number of annual smears (0% disc.)	404 000	446 000	478 000	499 000	510 000
Effects and costs: differences with the situation without early detection					
Total costs (× Mln)	463	515	561	595	616
Deaths prevented	1 496	1 485	1 464	1 402	1 275
Life-years gained	17 564	19 323	20 334	20 511	19 352
CERs compared with the situation without screening					
Costs per death prevented	309 269	347 157	338 329	424 530	482 867
Costs per life-year gained	26 334	26 673	27 602	29 024	31 811
CERs compared with the policy starting at older age					
Incremental costs per l.y.g.	–	30 045	45 371	192 393	Less effects more costs
95% CI because of stochastic output	–	26 000 to 34 000	36 000 to 60 000	79 000 to no l.y.g.	No l.y.g. to no l.y.g.

risk also plays a role: i.e. the higher the attendance, the more high risk women are expected to be covered by screens [119]. We assumed 75% attendance in our base-line calculations, which is much higher than the attendance rates of 35 to 50% reported from regions in the Netherlands over the last 5 years. However, 75% of the women of the target age-group had a PAP smear in the five years preceding 1995 [14]. Many of these women had opportunistic PAP smears outside the screening programme. Since our predictive calculations apply to the situation without opportunistic screening, an assumed attendance rate of 75% would be a high estimate. On the other hand, the better organization of the new screening programme is expected to increase the coverage of screening.

Large amounts of opportunistic smears result in higher costs per life-year gained with screening. In order to prevent opportunistic screening, the Dutch Health Insurance Council has decided no longer to reimburse the costs of these smears.

The costs of cervical cancer screening mainly consist of the costs of running the programme and the costs of taking and evaluating the PAP smears *(see Table 7.3)*. Recently, new tariffs have been determined for these aspects, which were intended to reflect the true costs of the newly organised screening programme. The 1996 tariffs result in about 30% higher costs for a screening programme with seven PAP smears than previously estimated by us. Future

budget analyses will clarify how realistic the new tariffs are. If we have underestimated the variable costs of screening, the case for the decision not to increase the number of screenings will become stronger.

In Chapter 10 of this thesis, a comparison is made between our cost-effectiveness estimates for cervical cancer screening and other estimates from literature.

7.4.5 Varying assumptions on the follow-up strategy

The assumptions on follow up practice have not been subjected to a sensitivity-analysis. They may have been too conservative, assuming too little follow-up compared to the recent practice (see Chapter 10 and Appendix A). Ideally, follow-up should have been varied to estimate the optimal combination of follow up strategy and the screening policy. More intensive follow-up for instance would possibly allow for longer screening intervals. This is an important (but difficult) subject for further analysis, particularly in the present situation where, according to the follow-up guidelines, over 10% of the women are referred for follow-up.

7.4.6 International perspective

Several criteria can be used when choosing between cervical cancer screening policies: i.e. a target for the health effects, budget constraints, or a target incremental cost-effectiveness level. Setting any of these targets is necessary and sufficient for identifying a unique optimal policy. In the Netherlands, by setting the number of smears per women, a quasi budget constraint was set. This constraint was amongst others inspired by earlier estimates of the cost-effectiveness for cervical cancer screening policies at different cost-levels. Apparently, in other countries other criteria than estimated cost-effectiveness have sometimes played an important role. The respective national guidelines on the number of screenings per woman differs widely, e.g. from 27 in Australia [61] to six or seven in Finland [96] and the Netherlands. A large number of smears does not always reflect a higher risk level or a higher available budget for health care. For instance in Australia, the risk for cervical cancer is similar to the risk in the Netherlands.

Large differences in risk level will influence the optimal policy. In Denmark for instance, the risk of cervical cancer is about twice the risk in the Netherlands. If all other circumstances are assumed to be equal, a doubled risk would roughly halve the estimated (incremental) costs per life year gained. If an incremental cost-effectiveness of about Dfl 35 000 is accepted (which according to our calculations for the Netherlands corresponds about with moving from 6 to 7 screenings), this would result in at least ten screenings in the Danish situation, and consequently in an optimal screening interval of 5 years or even somewhat shorter. In most Danish counties, women are invited for PAP smear screening 13 times during their life, with an interval of 3 years [200].

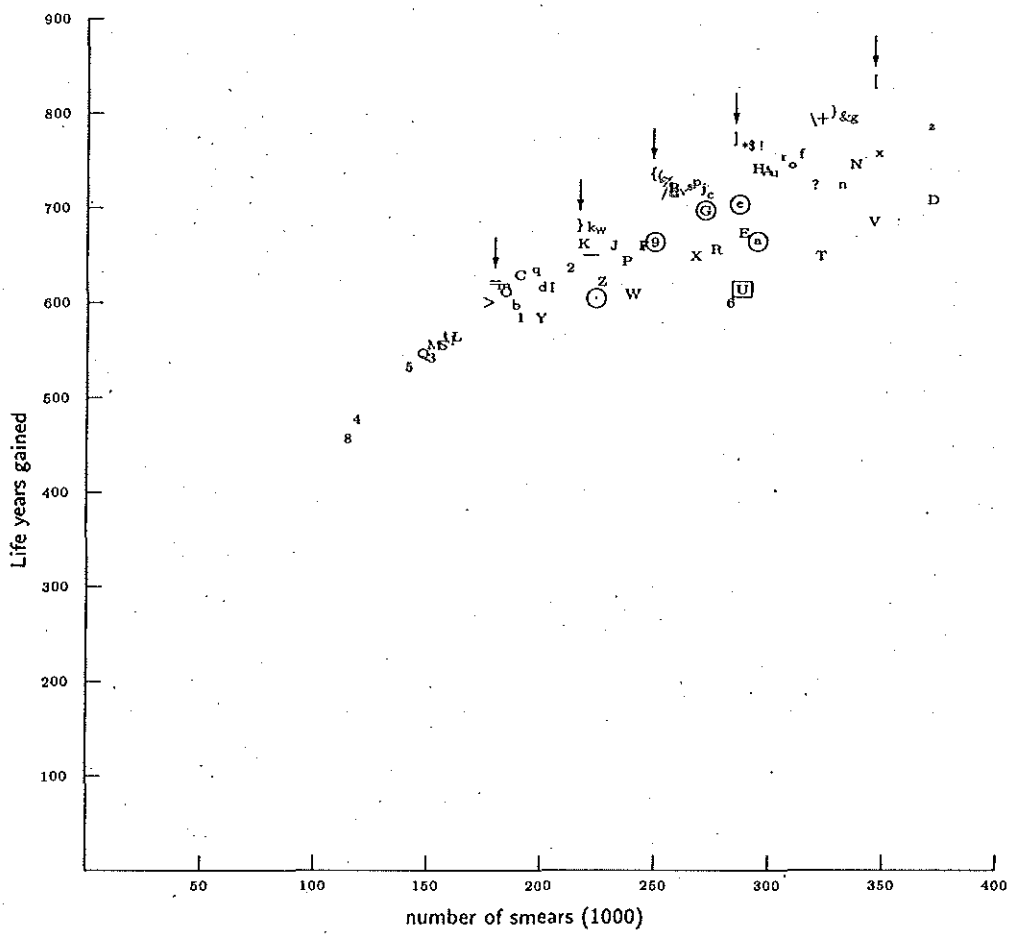
7.4.7 Conclusion

For the situation in the Netherlands but without opportunistic screening, and for a policy with 7 screenings per woman, a PAP smear screening interval of 7 years is optimal from a cost-effectiveness point of view. A shorter screening interval covers a too narrow age-range, missing the opportunity to save more life-years using the same amount of resources with a 7-year interval. In a Dutch national consensus meeting, 5 years was considered the largest acceptable interval, which is an improvement over a 3-year interval, but falls somewhat short compared to the optimal interval in an ideal situation without opportunistic screening. A 5-year interval might be a reasonable choice in reality where an amount of opportunistic screening can be expected to increase strongly with increasing screening interval.

Table 7.9 *Characteristics of the policies for which predictions have been made, and their label in* Figure 7.1. *Starting age (SA), last age (LA), interval (I), number of smears per woman (#), mean age (MA)*

Label	#	SA	I	LA	MA	Label	#	SA	I	LA	MA
4	3	35	10	55	45	p	7	26	6	62	44
8	3	37	10	57	47	(	7	26	7	68	47
1	4	25	10	55	40	s	7	27	6	63	45
t	4	27	10	57	42	{	7	27	7	69	48
L	4	30	8	54	42	v	7	28	6	64	46
M	4	30	10	60	45	B	7	29	6	65	47
Q	4	32	10	62	47	E	7	30	4	54	42
S	4	33	8	57	45	G	7	30	5	60	45
3	4	35	8	59	47	/	7	30	6	66	48
5	4	35	10	65	50	R	7	33	4	57	45
b	5	22	10	62	42	U	7	35	3	53	44
d	5	24	8	56	40	X	7	35	4	59	47
m	5	25	10	65	45	9	7	35	5	65	50
q	5	26	8	58	42	6	7	36	3	54	45
=	5	27	10	67	47	|	7	37	6	73	55
C	5	29	8	61	45	.	7	40	5	70	55
I	5	30	6	54	42	!	8	24	6	66	45
O	5	31	8	63	47	f	8	25	5	60	43
>	5	34	8	66	52	$	8	25	6	67	46
Y	5	35	5	55	45	n	8	26	4	54	40
1	5	35	6	59	47	o	8	26	5	61	44
,	5	39	8	71	55	*	8	26	6	68	47
i	6	25	6	55	40	r	8	27	5	62	45
k	6	25	8	65	45	]	8	27	6	69	48
}	6	27	8	67	47	u	8	28	5	63	46
w	6	28	7	63	46	A	8	29	5	64	47
F	6	30	5	55	43	H	8	30	5	65	48
J	6	30	6	60	45	T	8	34	3	55	45
K	6	30	7	65	48	g	9	25	5	65	45
P	6	32	5	57	45	&	9	26	5	66	46
W	6	35	4	55	45	)	9	27	5	67	47
Z	6	35	5	60	48	+	9	28	5	68	48
2	6	35	6	65	50	x	9	29	4	61	45
a	7	20	5	50	35	\	9	29	5	69	49
c	7	24	6	60	42	D	9	30	3	54	42
@	7	24	7	66	45	N	9	31	4	63	47
e	7	25	5	55	40	?	9	34	4	66	52
j	7	25	6	61	43	V	9	35	3	59	47
%	7	25	7	67	46	[	10	27	5	72	50

Figure 7.1 *The number of smears and the number of life years gained for different combinations of age ranges and intervals. The symbols refer to the policies described in* Table 7.9. *The numbers are per screening year in the Netherlands. Further assumptions: 75% attendance rate, no opportunistic screening, and 5% discount rate*



↓ Policies with an efficient combination of age-range and interval
○ Policies with 7 PAP smears per woman and a 5 years interval, with the following starting ages:
○ 40 years, ⑨ 35 years, ⓖ 30 years, ⓔ 25 years, ⓐ 20 years.
G New Dutch policy
U Old Dutch policy

8

The rescreen effect in conventional and Papnet screening observed in a study using material enriched with positive smears

This chapter is based on M. van Ballegooijen et al, in press in Acta Cytologica

Abstract

Objective: PAPNET-computer assisted cervical smear screening is used routinely in the LCPL, showing encouraging results. The purpose of this paper is to study the rescreen effect and the PAPNET effect on enriched material derived from smears screened routinely using PAPNET and conventional microsopy.

Study design: A series of 432 smears (containing 122 atypical [ASCUS] and 44 at least dysplastic [squamous intraepithelial lesions, SIL+] ones) screened routinely with the conventional method, were rescreened using the PAPNET system. An other series of 461 smears (containing 140 ASCUS and 52 SIL+ ones) screened routinely with PAPNET, were rescreened conventionally. The rescreen-effect, defined as the effect of differences between the rescreen and the routine screening situation, was investigated by comparing the rescreen results in both series of smears with the routine results in both series. The effect of using either method of screening was studied by comparing the PAPNET results in both series with the results of conventional screening in both series.

Results: The rescreen effect was statistically significant both for a higher number of smears classified as negative (less than ASCUS) and a higher number of smears classified as high grade SIL or more. PAPNET-assisted screening resulted in a significantly higher number of smears classified as high grade SIL+; although for this latter finding there is an unexplained significant difference between conventional and PAPNET screened cases in the changes made by the cytopathologist in the cytotechnologists' diagnoses.

Conclusion: The rescreen effect should not be ignored when enriched material is used.

8.1 Introduction

Ever since the introduction of cervical cytology as a diagnostic tool for routine screening to detect premalignant and malignant conditions, various techniques have been sought to improve sensitivity and/or specificity. One of these techniques is the recently developed computer-assisted screening method using the PAPNET system [122]. In this computer assisted system, smears are processed by a neural network computer before the human cytotechnologist is involved. This computer processing produces a digital tape with 128 images, which are selected as having the greatest likelihood of containing abnormal cells.

Results of a study that compared the use of this new method in a routine setting to conventional screening showed a decrease in false negative rates [113]. However, as always when a new screening method emerges, studies to begin with concerned PAPNET in a rescreen situation, using series of smears enriched with positive cases [23, 108, 122, 136, 137, 179, 182, 185, 196, 206]. The rescreen situation is completely different from the routine screening situation, resulting in a change in diagnostic pattern [25], the so-called rescreen effect. The impact of the rescreen-effect is a well recognized pitfall in comparing different modalities of laboratory screening tests. The purpose of this paper is to study it quantitatively.

The present study concerns a rescreen study on enriched material derived from both PAPNET-assisted and conventionally screened smears. This was possible because screening with the PAPNET system was used routinely at the Leiden laboratory alongside conventional screening. Moreover, for rescreening, we also could use both screening modalities. This design enabled the rescreen effect to be evaluated independently of the technique used.

8.2 Materials and methods

8.2.1 Papnet-assisted screening

In the PAPNET-assisted diagnostic procedure, the video images of the 128 cell groups selected by the computer were examined and the images of the 16 most abnormal ones were marked using the available computer technology. The 16 images were brought together on the video summary screen to further enhance the diagnostic information. The XY co-ordinates of each of these 16 images were utilized to localize under the microscope the strands of material in the smear that contained abnormal fragments and/or abnormal cells. As a final step, the cytological diagnosis was made, taking into account the diagnostic information presented on the video summary screen plus the light microscopic evaluation of the abnormal cells under a high power objective (40×).

8.2.2 Cytological classification

The smears were classified as negative, ASCUS/AGUS (including atypical squamous cells of uncertain significance and atypical glandular cells of uncertain significance, and smears showing only condyloma, excluding benign cellular changes), low grade SIL(including light and moderate dysplasia) and high grade SIL+ (including severe dysplasia, carcinoma in situ and invasive carcinoma). Since in the Dutch system smears with light and moderate dysplasia were classified into one category, we had to include moderate dysplasia into the low grade SIL.

8.2.3 Screening and rescreening procedure

Both in the routine and the rescreen situation, the cytotechnologists were allowed to sign out negative and ASCUS/AGUS cases. However smears signed out by cytotechnologists as at least low grade SIL, were also evaluated by the cytopathologist, who gave the final diagnosis as presented in the tables. All selected and available smears were rescreened by one of the five cytotechnologists who also performed the initial screening. The cytotechnologists and the cytopathologist were blinded for the initial result.

8.2.4 Initial screening material

In a five-month period in 1993, 28 337 samples were sent to the laboratory. About 500 smears per week were screened using the PAPNET-assisted method. This amounted to 11 117 smears or 39% of all the samples received. The other 17 220 smears were screened conventionally. We call this PAPNET-assisted and conventional evaluation of the smears the initial screening; it was performed in a routine screening situation (as opposed to the study situation).

8.2.5 Selection of material for rescreening

From the total of 28 337 smears screened, 95.5% were classified as negative, 2.7% as ASCUS, 1.4% as low grade SIL and 0.4% as high grade SIL or higher abnormality. 1000 cases were selected for rescreening on the basis of the initial cytological diagnosis; 500 were chosen from the smears initially screened conventionally while 500 were selected from the material initially screened by the PAPNET method. The initial diagnoses in the two selected series are summarized in *Table 8.1.* Abnormal smears were deliberately overrepresented in the study samples (such that the material selected contained enough positive cases to compare outcomes in these cases). All the smears with a diagnosis of at least SIL were selected from both series. For ASCUS and negative smears, 120 and 180 smears respectively were randomly selected. To compensate for eight smears with SIL fewer among the smears initially screened with the PAPNET-assisted

Table 8.1 *Initial cytological results of the smears selected and ( ) available for the study*[1]

Cytological result	Negative	ASCUS	Low grade SIL	High grade SIL+	Total
Smears initially screened conventionally	180 (157)	120 (109)	142 (122)	58 (44)	500 (432)
Smears initially screened with PAPNET	188 (163)	120 (118)	140 (137)	52 (43)	500 (461)
Total	368 (320)	240 (227)	282 (259)	110 (87)	1000 (893)

[1] The analyses have been also performed after having corrected the figures for the differences in number of available smears per class of initial diagnosis, see results

method, eight extra negative smears were selected for this series. However, 68 of the conventional smears and 39 of the PAPNET-smears were not available for rescreening *(see Table 8.1)* due to technical problems, such as poor coverslipping of the smear, slide breakages, additional immunostaining and other irremediable circumstances.

8.2.6 Data analysis

Two factors play a role in the cytological results: the method used (PAPNET vs conventional) and the rank of the cytologic evaluation (rescreening vs initial). To start with, the two series of smears were considered separately, by cross-classification analysis of the initial and the rescreen results (*see Table 8.2* and Table 8.3 for conventional initial screening and PAPNET initial screening respectively). In this analysis, the two factors are confounded.

In the next two analyses, the two aspects are put forward separately. In the first one *(see Table 8.4)*, the effect associated to the rescreen situation versus the initial situation is visualized. To this end, the results of the two series of smears were combined: all rescreen results (either from conventional or from PAPNET rescreening) were compared with all initial screening results (either from conventional or from PAPNET screening).

In *Table 8.5*, PAPNET screening is compared with conventional screening, independently of rank of screening. The results of the two series of smears again are combined. However, this time all the conventional diagnoses (from either initial screening or rescreening) were compared to all the PAPNET diagnoses (from either initial or rescreening).

For the calculation of the p values for the differences in the number of diagnoses (*Tables 8.2–8.5*), we used contingency tables and log-linear models [60].

Table 8.2 *Cross-classification of cytological results: conventional initial versus PAPNET rescreen*

Papnet rescreen	Conventional initial				
	Negative	ASCUS	Low grade SIL	High grade SIL+	ALL
Negative	132	71	2	0	205
ASCUS	19	35	12	0	66
Low grade SIL	6	3	83	1	93
High grade SIL+	0	0	25	43	68
ALL	157	109	122	44	432

With rescreen:
48 Negative more (+31%), or 48 ASCUS+ less (-17%); $p < 0.0001$
43 ASCUS less (-39%)
29 Low grade SIL less (-24%)
24 High grade SIL+ more (+55%); $p < 0.0001$

8.3 Results

The results of the rescreening with PAPNET of smears initially screened with the conventional method are presented in *Table 8.2*, and those of rescreening conventionally smears initially screened with PAPNET in *Table 8.3*. The tables show that both upgrading and downgrading of the initial diagnosis occurred. Upgrading from initially negative smears and smears with ASCUS to high grade SIL+ and downgrading from initially high grade SIL+ to negative smears never occurred; downgrading from initially high grade SIL+ to ASCUS occurred only twice.

A summary of the statistical significance of the upgrading and downgrading is provided in the bottom of *Table 8.2* and *Table 8.3*. These outcomes depend highly on the sample composition in the study. They were only considered in this aggregated way to compare the results of the different analyses. In the PAPNET rescreen run *(see Table 8.2)*, significantly fewer smears were classified as positive (ASCUS and higher) and significantly more as highly abnormal (HSIL+) than in the same smears in the conventional initial run. In the conventional rescreen run however *(see Table 8.3)*, again the technique used in the rescreen method (which is the conventional one this time) classified significantly fewer smears as positive (ASCUS+). The number of smears classified as highly abnormal now is almost unchanged.

In *Table 8.4*, where the effect of rescreening becomes visible independent of the methods used, the same pattern arises as in *Table 8.2*: during rescreening significantly fewer (-19%) smears were classified as abnormal (ASCUS and higher) and significantly more (+25%) as highly abnormal (HSIL+) than in the initial screening.

Table 8.3 *Cross-classification of cytological results:* PAPNET *initial versus conventional rescreen*

Conventional rescreen	Papnet initial				
	Negative	ASCUS	Low grade SIL	High grade SIL+	ALL
Negative	144	75	6	0	225
ASCUS	18	33	10	2	63
Low grade SIL	1	10	114	7	132
High grade SIL+	0	0	7	34	41
ALL	163	118	137	43	461

With rescreen:
62 Negative more (+38%), or 62 ASCUS+ less (-21%); $p < 0.0001$
55 ASCUS less (-47%)
5 Low grade SIL less (-4%)
2 High grade SIL+ less (-5%); $p = 0.62$

In comparing PAPNET with conventional screening *(see Table 8.5)*, PAPNET screening classified 2.7% (not significant) more smears as positive (ASCUS+), but 30% (significant) more smears as highly abnormal.

We repeated all the analyses after having corrected the figures for the differences in number of available smears per class of initial diagnosis *(see Table 8.1)*. The results were not affected, neither in direction nor in statistical significance.

The rescreen-effect on the number of smears classified as negative was much larger than the PAPNET-effect. (compare *Table 8.4* with *Table 8.5*). For high grade SIL+ we see an increase in the number of smears classified as such, both in rescreening and in using PAPNET. In other words, the rescreening effect on this enriched material points in the same direction as the PAPNET effect.

8.4 Discussion

This study shows a significant rescreen effect both on the number of smears classified as negative (less than ASCUS) and on the number of smears classified as higly abnormal (SIL+). In our view, based on observations from the workfloor, this rescreen effect was mainly caused by three factors. Firstly, the diagnostic responsibility is completely different. The diagnosis in the rescreen situation will have no clinical consequence, resulting in more completely negative and more completely positive cases. As a consequence there are less 'in between' cases. Secondly, hightened alertness of the cytologist, through which smears with only few abnormal cells will be more easily spotted, is expected in a study situation. Thirdly, the material selected for this study contained 60% positive smears, whereas in the initial routine screening situation only 4% of

Table 8.4 *Cross-classification of cytological results: initial versus rescreen*

Rescreen (conventional + Papnet)	Initial (conventional + Papnet)				
	Negative	ASCUS	Low grade SIL	High grade SIL+	ALL
Negative	276	146	8	0	430
ASCUS	37	68	22	2	129
Low grade SIL	7	13	197	8	225
High grade SIL+	0	0	32	77	109
ALL	320	227	259	87	893

With rescreen:
110 Negative more (+34%), or 110 ASCUS+ less (-19%); $p < 0.0001$
98 ASCUS less (-43%)
34 Low grade SIL less (-13%)
22 High grade SIL+ more (+25%); $p < 0.0001$

the diagnoses were positive. As a consequence, the cytotechnician suffers less from habituation. Habituation has roughly the opposite effect as increased allertness. For those of us who have routinely screened 200 000 smears with the PAPNET method, it is not surprising that apart from the rescreen effect there is also a PAPNET effect in finding more high grade SIL+ cases. When PAPNET is used, there is a higher alertness that the visual information of the 128 tiles might be diagnostically important, otherwise they would not have been selected out of the 300 000 cells of the smear. In addition, there is less habituation because the cytotechnologist sees less benign cells.

In this rescreen study, like in the routine screening situation, smears classified by cytotechnologists as at least low grade SIL, were also evaluated by the cytopathologist, who gave the final diagnosis. It was interesting to see that in the smears rescreened with PAPNET, compared to the cytopathologist, the cytotechnologists had a tendency to diagnose low grade versus high grade SIL+ (this occurred 8 times versus 0 times the other way around). In the smears rescreened conventionally the opposit is seen: cytotechnologists tended to diagnose high grade versus low grade SIL (this occurred 11 times versus 2 times the other way around). This difference is statistically highly significant ($p < 0.001$). As a consequence, the PAPNET-effect which was found at the cytopathologist level was not significantly present at the level of cytotechnologists. We have no explanation for this finding, and have not investigated it in greater detail, but it emphasizes the importance of well recorded laboratory screening procedures. The observed rescreen-effect was equal at the cytopathologist and cytotechnologist level.

The problem with using enriched series for rescreening studies, as was also done by us, is that it contains a too small number of negative smears to cal-

Table 8.5 *Cross-classification of cytological results: conventional versus* PAPNET

Papnet (initial + rescreen)	Conventional (initial + rescreen)				
	Negative	ASCUS	Low grade SIL	High grade SIL+	ALL
Negative	276	89	3	0	368
ASCUS	94	68	22	0	184
Low grade SIL	12	13	197	8	230
High grade SIL+	0	2	32	77	111
ALL	382	172	254	85	893

With Papnet:
14 Negative less (-4%), or 14 ASCUS+ more (+2.7%); $p = 0.32$
12 ASCUS more (+7%)
24 Low grade SIL less (-9%)
26 High grade SIL+ more (+30%); $p = 0.0005$

culate reliably specificity and sensitivity. Since in a routine situation, more then 90% of the smears are conventionally diagnosed as negative, this category may carry a heavy weight on these calculations. The impact of small numbers of negative smears can be illustrated by rescaling the outcomes of this study to the distribution of diagnoses in the screening situation (95.5% negative smears, 2.7% ASCUS, 1.4% low grade SIL and 0.4% high grade SIL+). After such rescaling of the results of PAPNET versus conventional screening (presented in *Table 8.5*), the 2.7% increase in smears classified as positive (ASCUS+) becomes a 85% increase. The cause for this shift is that the 4% decrease in smears initially classified as negative in PAPNET screening now applies to 96% of the smears classified as negative in the screening situation (versus 40% in the study material). Adding follow-up data would determine the histologic yield of this higher number of positive cytologic diagnoses. But the expected prevalence of e.g. high grade SIL (as defined according to the Dutch system where moderate dysplasia counts as not high grade) in smears initially called negative is only about one per thousand (assuming a prevalence of high grade SIL of five per 1000 women and a sensitivity for high grade SIL of 80%). Hence, many thousands of smears initially classified as negative have to be screened with the new method to asses a change in sensitivity for high grade lesions. This not only holds for the assessment of the sensitivity of PAPNET in routine screening, but also in rescreening negative smears for quality assurance purposes in cytology laboratories, using any of a number of techniques including the 'automated' ones.

This study clearly shows the necessity not to ignore the rescreen-effect when enriched material is used.

9

Present evidence on the value of HPV testing for cervical cancer screening: a model-based exploration of the (cost-)effectiveness

This chapter is based on M. van Ballegooijen et al
British Journal of Cancer 76(5), 651–657, 1997

Abstract

Human papillomavirus (HPV) is the main risk factor for invasive cervical cancer. High risks ratios are found in cross-sectional data on HPV prevalence. The question raised is whether this present evidence is sufficient for making firm recommendations on HPV-screening.

A validated cervical cancer screening model was extended by adding HPV infection as a possible precursor of cervical intraepithelial neoplasia (CIN). Two widely different model quantifications were constructed so that both were compatible with the observed HPV risk ratios. One model assumed a much longer duration of HPV infection before progressing to CIN and a higher sensitivity of the HPV test than the other. In one version of the model, the calculated mortality reduction from HPV-screening was higher and the (cost-)effectiveness was much better than for PAP smear screening. In the other version, outcomes were the opposite, although the cost-effectiveness of the combined HPV+ cytology test was close to that of PAP smear screening.

Although small follow-up studies and studies with limited strength of design suggest that HPV testing may well improve cervical cancer screening, only large longitudinal screening studies on the association between HPV infection and the development of neoplasias can give outcomes that would enable a firm conclusion to be made on the (cost-)effectiveness of HPV screening. Prospective studies should address women aged 30–60 years.

Acknowledgement This study was financed by the Dutch Health Insurance Council

9.1 Introduction

Molecular and epidemiological studies have clearly demonstrated that HPV is the main risk factor for cervical cancer [101, 214]. These epidemiological studies are case-control studies that consistently show a very high-risk ratio for HPV in women with (precursors of) cervical cancer compared with controls with negative cytology [146, 149, 66, 183]. The association between CIN and high-risk HPV infection is stronger in high-grade than in low-grade abnormalities [30, 17, 131, 76, 109] and is well over 90% in invasive cancers [30, 26]. A few small follow-up studies also corroborate the crucial role of HPV infections: progression is found almost only in women with (persistent) high-risk HPV genotypes both in normal [180] and in dysplastic cases [95, 174]. In a small retrospective study on archived false-negative smears from women with subsequent invasive cervical cancer, the high-risk HPV types found in the cancers were detected in nearly 100% of the preceding smears [207].

On the other hand, test-positive rates for high-risk HPV types in women over 30 years of age with normal cytology in North American and western European countries vary from 3% to 6% [149, 15, 53, 180]. This is much higher than can be explained by the life-time risk of developing cervical cancer in these countries. For example, in the Netherlands, the rate for high-risk HPV types in woman aged 30+ with normal cytology is around 4%, while the cumulative risk for invasive cervical cancer is around 1.5%; the risk in women aged 30+ with normal cytology is again much smaller. Therefore, only a fraction of the infections with high-risk HPV types will progress to cervical cancer.

The goal of this study was to incorporate the very high observed HPV-associated risks ratios in a cervical cancer screening model and to investigate the consequences for HPV screening as expressed in predicted mortality reduction, negative side-effects and costs. The outcome of the follow-up studies carried out to date have been incorporated in the model in so far that HPV infections were assumed to preceed HPV-infected neoplasias. They were not used for the quantification of the model as these studies were small or interpretation in quantitative epidemiological terms was limited by their design. The possible impact, however, will be discussed. The present study focusses on the question of whether recommendations about HPV screening can already be made on the basis of the available data and, if not, what type of data will be required to decrease uncertainty.

9.2 Materials and methods

9.2.1 The data

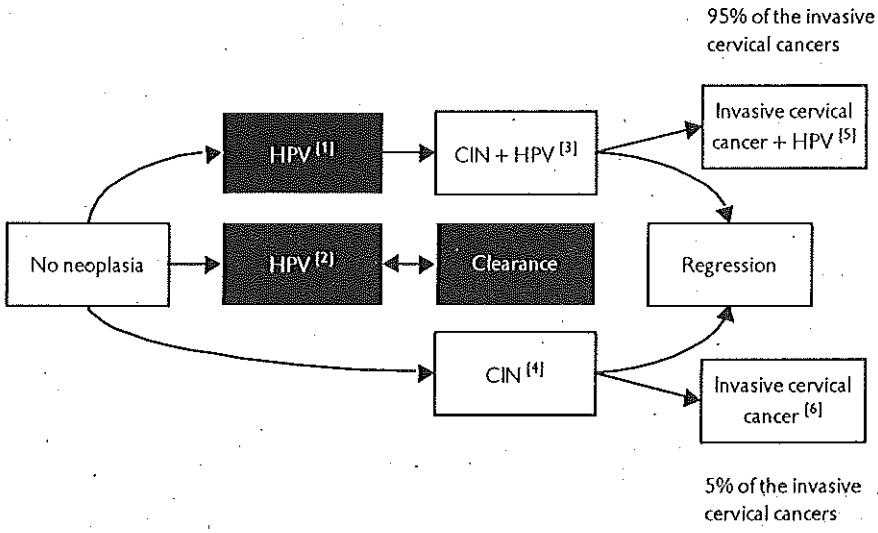
Test-positive rates for high-risk HPV types in women between the ages of 30 and 60 years were estimated on the basis of empirical data. Polymerase chain reaction (PCR-)based HPV-positive rates on cytological material of the cervix from women with negative cytology are 4% in the Netherlands [180], 5.7% in Portland, Oregon, USA [15], and 4.6% in Spain [149]. PCR-based HPV-positive rates on cytological material of women with a histologically confirmed diagnosis of CIN are 71% in Spain and 54% in Colombia [183], 75% in the USA [146], 72% in the UK [62] and 59% in the Netherlands [76]. HPV rates are higher in high-grade than in low-grade lesions. Noting that the reported results are of the same order of magnitude, we summarized them by assuming 4% HPV positiveness in cytologically negative women and 67% in women with CIN. On the basis of the worldwide study on histological material of Bosch et al [26], we assumed that 95% of the invasive carcinomas were HPV infected, i.e. only 5% of invasive cervical cancers developed without being preceded by an HPV infection. In accordance with the results of the Dutch study [141], HPV-positive rates are assumed to be constant between 30 and 60 years of age.

9.2.2 The model

Here, the relationship between HPV and cervical cancer in a stochastic microsimulation screening model is described. HPV in the model represents high-risk HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 54, 56, 58, 59, 66, 68). As shown in *Figure 9.1*, the model is based on the hypothesis that the onset of HPV infections found in invasive cervical cancer and in CIN has preceded these neoplastic stages. Women who go through an HPV infection either become clear from the infection or develop HPV-infected CIN, which either regresses or progresses into HPV-positive invasive cervical cancer. Women can also develop CIN without an HPV infection, and this CIN again can regress or progress (only sometimes, see later) into invasive cancer. Allowing for the possibility that women can develop CIN (with or without HPV) after having become clear from HPV infection would cause a shift between the several arms in the model, without affecting the model outcomes presented in this article; therefore we did not complicate the model in this manner. This model is an extension of a validated cervical cancer screening PAP smear model [118, 119, 10].

According to this model, the average duration of CIN is 11.8 years and preclinical invasive cancer is 3.9 years *(see Table 9.1)*. The sensitivity of the PAP smear is 80% in CIN and 87.5% in preclinical invasive carcinoma. These estimates on duration and sensitivity were derived from the British Columbia (Canada) screening data [159] and were compatible with data on interval cancers collected by the IARC[160, 161]. The incidence of progressive CIN was cho-

Figure 9.1 *The stages and possible transitions in the* HPV *to* CIN *to invasive cervical cancer model. The disease stages that describe non-neoplastic conditions, and that have been added to the validated* CIN *to cervical cancer model, have been shaded*



sen to reproduce cervical cancer incidence and mortality in the Netherlands between 1965 and 1992. The regression rate was 72% of disease onset under 35 years, 40% between the age of 35 and 54 and very low in women aged 54 and over. These estimates resulted from subtracting progressive CIN from the age-specific CIN detection rates observed in the Dutch population [163]. When adding HPV infection to the model, the part describing CIN and invasive cervical cancer was kept unchanged; the predicted CIN and cervical cancer incidences and prevalences were not affected. Consequently, previous validations are still valid. The incidence in the Dutch population accounted for is lower than incidences, for example, in the UK and the USA (7.8 and 12.9 per 100 000 for the Netherlands and the UK respectively, in 1978–1982 [104] and 9.9 in the USA in 1985 [168]). The incidence level however did not influence the comparison of screening strategies.

9.2.3 Two model versions

Because only cross-sectional HPV data were available for the quantification of the model, there was an identification problem for the parameters describing HPV infections. Test-positive rates in women screened for the first time are a result of incidence × duration × sensitivity. In view of this non-identifiability, we decided to construct two model quantifications that were contrasting in

Table 9.1 *Parameter values in model version A and B on the duration of detectable preclinical stages and the sensitivity of the HPV test for these stages*

	Model version A	Model version B
Duration of stages (years)		
HPV[1] that will develop into CIN+HPV	10	1
HPV[2] that will be cleared	1	10
CIN (with or without HPV)[3][4]	11.8	11.8
Invasive cancer (with or without HPV)[5][6]	3.9	3.9
Sensitivity of HPV test (%)		
HPV[1][2]	100	50
CIN+HPV[3]	100	80
Invasive cancer+HPV[5]	100	87.5

[ ] Refers to the numbering of the disease stages in *Figure 9.1*

HPV-screening outcomes. We varied duration and sensitivity and adjusted the incidence level to the observed test-positive rates for HPV. The longer the duration of progressive (to CIN) HPV infections (stage HPV[1] in *Figure 9.1*) and the higher the sensitivity of the HPV test, the more effective HPV screening will be in reducing cervical cancer mortality. In order to minimize the negative side-effects (i.e. follow-up of HPV-positive women who will not develop cervical neoplasia), it is favourable to assume a short duration of harmless (non-progressive) HPV infections (stage HPV[2] in *Figure 9.1*).

In model quantification A *(see Table 9.1)*, the extra duration of the detectable preclinical phase because of HPV detection was assumed to be 10 years. The assumed sensitivity for HPV was 100% at all stages. Long duration and high sensitivity made model version A very favourable for HPV screening. In version B of the model, the detectable preclinical phase was only 1 year longer than in PAP smear screening, and sensitivity for high-risk HPV types was considerably lower than in version A. In HPV-infected neoplasia stages, sensitivity of the HPV test was equal to the sensitivity of the PAP smear (80% in HPV-positive CIN and 87.5% in HPV-positive invasive cancer), and sensitivity was only 50% in HPV infections without neoplasia. Compared with model A, model B was very unfavourable for HPV screening. The consequences of the two sets of assumptions for the sensitivity of the test (or combination of tests) are given in *Table 9.2*.

As a result of differences in sensitivity of the HPV test, the HPV test-positive rate of scrapes in invasive cervical cancer cases was (100% sensitivity × 95% invasive cervical cancers with preceding HPV infections =) 95% in model A and (87.5% × 95% =) 83% in model B. A high rate is in accordance with some PCR studies on cytological material of women with invasive cervical cancer (up to 100% [30]), but a lower rate has been found in other studies (e.g. 84% [66]).

Table 9.2 *Sensitivity by test (or combination of tests), stage and model version resulting from the values for sensitivity of the HPV test given in* Table 9.1

Stages	Any model version	Model version A		Model version B	
	Cytology only	Cytology + HPV	HPV only	Cytology + HPV	HPV only
HPV[1][2]	0	100	100	50	50
CIN + HPV[3]	80	100	100	96	80
CIN[4]	80	80	0	80	0
Invasive cancer + HPV[5]	87.5	100	100	98.4	87.5
Invasive cancer[6]	87.5	87.5	0	87.5	0

[ ] Refers to the numbering of the disease stages in *Figure 9.1*

9.2.4 Simulated compared with observed HPV test-positive rates

In both model versions, predicted HPV test-positive rates in the age group 30–60 years was 4.01% in women with negative cytology and 67% in women with CIN.

9.2.5 Consequences of true-positive test results

In the simulation, women with only negative tests at screening had a future screening after the regular screening interval. Women with positive cytology were followed up and in true-positive cases this led to the detection of neoplasia. Women with a negative PAP smear and a positive HPV test were assumed to be followed up with HPV tests and PAP smears every six months. This follow-up stopped either when the HPV infection was cleared (after which women go back to screening) or when there was a transition of the HPV infection to HPV infected CIN (the neoplasia is detected). Detected CIN was assumed to be managed so that no invasive cancer would develop. For the management of CIN (diagnosis, treatment and after treatment check-ups), we accounted for 4 years of follow-up. This was in accordance with current practice in the management of CIN, at least in the Netherlands [13].

9.2.6 Consequences of false-positive test results

Women with borderline (ASCUS) or low-grade abnormalities in their PAP smears in the Netherlands, and also in many other countries, are followed up with repeat smears. Some of these women have negative repeat smears and are referred back for routine screening. Women with high-grade abnormalities in their PAP smears are referred to the gynaecologist. In a proportion of these women, no neoplasia is found. As the model was adjusted for histologically confirmed detection rates, these so called 'false-positive' cytological outcomes have to be accounted for separately. We made the following assumptions:

Table 9.3 *Assumptions on the costs by type of procedure, in Dfl*

Procedure	Costs	Costs in the sensitivity analyses
Screening PAP smear[a]	70	
Repeat PAP smear[a]	100	
HPV test[a]	90	45/155
PAP smear and HPV test in one screening session[a]	135	90/200
Follow-up session in HPV-positive women with negative cytology	140	280
Diagnostic work-up of the referral when no neoplasia is found	800	
Management of CIN[b] [13]	3 100[c]	
Curative primary treatment		
microinvasive carcinoma	9 500	
IB invasive carcinoma	20 200	
II+ invasive carcinoma	19 100	
Care for advanced disease [11]	30 700	

[a] Including Dfl 25 in total for costs for carrying out the smear/scrape and the costs for the women (time and transport, [118]

[b] CIN with or without HPV infection

[c] Including the costs of 15% recurrence of disease after primary treatment of CIN

- Five per cent of the screening smears generated two repeat smears in women that did not have neoplasia
- Five per 10 000 screened women without CIN were referred to the gynaecologist [164]

9.2.7 The costs of screening

In order to account for the costs and savings of early detection, the costs of screening, follow-up, diagnosis and treatment were considered *(see Table 9.3)*. The true resource costs were assessed for the screening PAP smear, the HPV test, colposcopy and radiotherapy. Costs charged in the Netherlands for the other medical procedures were used. The costs are presented in Dutch Guilders, for which the US$ exchange rate during 1995 was, on average 1.61.

9.2.8 Screening strategies

In both model versions, the effects and costs have been calculated for several screening strategies for women between the ages of 30 and 60 years. We made predictive calculations for 3-yearly cytology and for six alternative strategies. Within these alternative strategies, we considered two screening test (or combination of tests) and three screening schedules. The screening tests were: cytology plus HPV test and HPV test only. In the three screening schedules, women were screened between 30 and 60 years of age: every 3 years (11 screenings per woman), every 5 years (seven screenings per woman) and every 10 years (four screenings per woman).

9.2.9 The cost-effectiveness calculations

Calculations were made for a cohort of women who attended all screenings. Effects, costs and savings of the screenings were accounted for from birth to death. Outcomes were presented per 1000 women and have not been discounted.

9.3 Results

9.3.1 Mortality reduction, years in follow-up and cost-effectiveness

The model predictions of the main effects and costs of the different combinations of frequency and types of screening tests are summarized in *Table 9.4*. For each of the two model versions and for each of the two alternative screening tests (cytology plus HPV test and HPV test alone), only the policy with the lowest screening frequency that had the same or higher mortality reduction compared with 3-yearly PAP smear screening is presented.

According to the model version A, which was favourable for HPV-screening, the combined test (cytology plus HPV test), even if performed only once every 10 years, reduced mortality more (91% vs 79%) than 3-yearly PAP smears. Costs were 37% lower, mainly because of the less frequent screening, and costs per life-year gained decreased by 41%. The number of years in follow-up was 26% lower, and the years in follow-up per life-year gained decreased by 27%. For 10-yearly screening with the HPV test only, mortality reduction was also higher than for 3-yearly cytology and only a little lower (89% vs 91%) than for the combined test. The costs for HPV only were very low, only 31% of the costs of 3-yearly PAP smear screening. Costs per life-year gained were 69% lower. The number of life-years spent in follow-up was less than half (because the repeat smears of the borderline cytology do not occur in screening for HPV), and this also counts for the number of life-years in follow-up per life-year gained.

The results of model version B, which was unfavourable for HPV screening, were quite different. Combined screening performed every 5 years yielded a slightly higher mortality reduction (80% vs 79%, it was predicted at 77% with 10-yearly combined screening) than screening with cytology every 3 years, and was 63% more costly, resulting in 60% higher costs per life-year gained. The number of years in follow-up were 2.5 times higher, as were the number of years in follow-up per life-year gained. In the predictions for screening with the HPV test alone, even a 3-yearly interval did not result in a mortality reduction as high as with 3-yearly PAP smear screening (the 1 year extra detectable phase for which sensitivity is 50% is outbalanced by the 5% progressive lesions that are not detectable because they are HPV negative). Costs per life-year gained and years in follow-up per life-year gained were 1.8 and 2.6 times as high respectively.

Table 9.4 *Model outcomes: effects and costs of different screening policies in women between 30 and 60 years of age, two model versions. Only the least frequent* HPV *screening strategies with the same or higher mortality reduction compared to 3-yearly* PAP *smear screening are presented. All figures are per 1000 women screened, except for percentages (in brackets)*

	Any model version	A model version		B model version	
	Cytology only 3-yearly[a]	Cytology + HPV 10-yearly[a]	HPV only 10-yearly[a]	Cytology + HPV 5-yearly[a]	HPV only 3-yearly[a]
Favourable effects					
Mortality reduction (%)	(79)	(91)	(89)	(80)	(76)[b]
Life-years gained (%)	65 (88)	68 (93)	66 (90)	66 (89)	62 (85)
Unfavourable effects					
Years in follow up	700	520	290	1760	1790
Costs (in Dfl ×1000)					
Screening	650	460	300	800	830
Follow up of HPV-positive cases	–	60	60	361	470
Follow up of false positive cytology[c]	95	35	0.2	65	1.5
Diagnosis and treatment					
CIN	180	120	80	170	140
Invasive and advanced cancer	-190	-220	-210	-195	-185
Total costs	740	460	230	1 200	1 250
Ratios (per life-year gained)					
Years in follow-up	11	8	4	27	29
Costs	11 400	6 800	3 500	18 300	20 100

[a] At primary screening interval
[b] Using the HPV test only, according to model B, one would have to screen more frequently than 3-yearly to result in at least the same mortality reduction as 3-yearly cytology
[c] At screening and during follow-up of HPV-positive cases

Based on the model version A calculations, a decision might be made to replace PAP smear screening with HPV screening with a longer interval. This would lead to a greater mortality reduction at lower costs in terms of resources and negative side-effects. However, the model version B calculations suggest that PAP smear screening should not be replaced by any of the studied HPV screening strategies; costs and negative side-effects increased, while prevention of mortality did not improve.

9.3.2 Sensitivity analyses

We also calculated the costs of HPV screening assuming that HPV-positive women with negative cytology would be followed up every 3 years instead of every 6 months. The resulting total costs of HPV screening were lower, in

Table 9.5 *Sensitivity analysis: costs per life-year gained with alternative cost assumptions, as percentage difference with the costs per life-year gained of 3-yearly cytology*

	Any model version	A model version		B model version	
	Cytology only 3-yearly	Cytology + HPV 10-yearly	HPV only 10-yearly	Cytology + HPV 5-yearly	HPV only 3-yearly
Baseline cost assumptions[a]	11 400	6 800	3 500	18 300	20 100
		-40	-70	+60	+80
Alternative cost assumptions[b]					
HPV test, Dfl 45	-60	-90	+25	+20	
HPV test, Dfl 155	-10	-40	+110	+160	
HPV follow up, Dfl 280	-30	-60	+110	+140	

[a] HPV test Dfl 90; HPV follow-up, Dfl 140

[b] These changes in assumptions do not affect the costs per life-year gained of 11 400 of 3-yearly cytology

particular according to model B in which cost-effectiveness of the combined test was close to the cost-effectiveness of PAP smear screening. However, less intensive follow-up in HPV-positive women would, with current knowledge, not be an acceptable option.

Economies of scale play an important role in the costs of an HPV test. Our estimate was based on a situation with, on average, 12 000 PCRs per year per laboratory. If the testing was concentrated in fewer laboratories, the test would become cheaper. Moreover, new developments can cause an increase or decrease in the costs of routine HPV tests. Therefore, calculations were repeated under the assumption that the laboratory costs per HPV test of Dfl 65 were less than one-third, i.e. Dfl 20, or doubled to Dfl 130. The total costs per test, including the Dfl 25 for carrying out the smear/scrape consequently will be Dfl 45 and Dfl 155, respectively, for the HPV test and Dfl 90 and Dfl 200 for the combined test (PAP smear + HPV test). In our basic calculations, a follow-up session for HPV-positive women was restricted to an HPV test and a PAP smear. We repeated the calculations with twice the costs per follow-up session (Dfl 280 instead of Dfl 140). This would be approximately the costs incurred when a colposcopy is added.

The results are summarized in *Table 9.5.* Options that were more cost-effective than 3-yearly PAP smear screening remained more cost-effective and those that were less cost-effective also remained less cost-effective. The conclusions were, therefore, not affected by considerable changes in the assumptions about the costs of HPV screening.

9.4 Discussion

We produced two model versions that both explained the high observed risk ratios for high-risk HPV types in women with cervical neoplasia compared with women with normal cytology. In addition, they were both compatible with the 'clearance' rates in repeated HPV tests observed in women with normal cytology. In model A, this clearance resulted from a short duration of harmless HPV infections. In model B, the low sensitivity of the HPV test explained why woman that were HPV positive at a first screening will often be HPV negative at the next one. The effects of HPV screening predicted by the two model versions widely differed. Hence, the high-risk ratios alone were inconclusive for the outcomes expected from HPV screening.

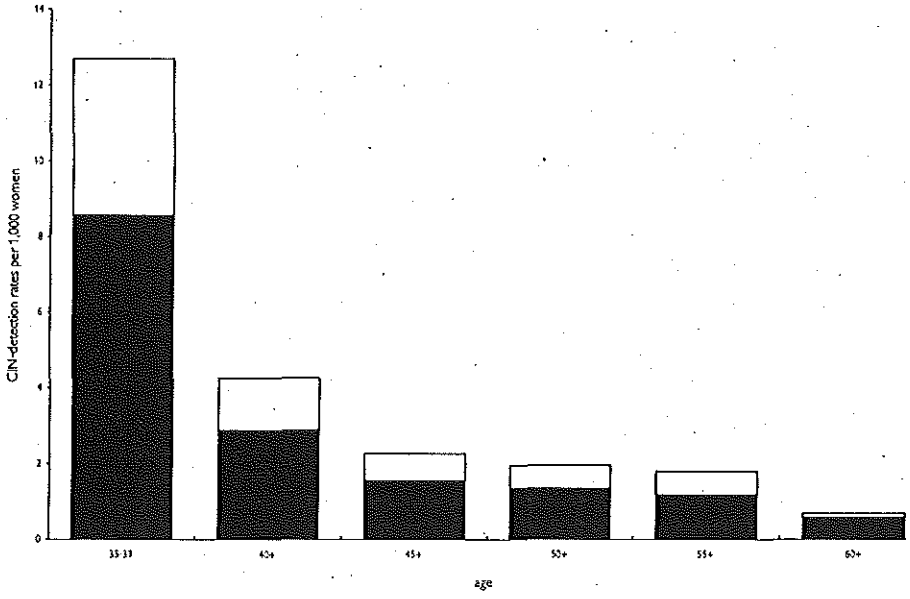
The first non-cross-sectional evidence for the crucial role of high-risk HPV infections for the development of cervical cancer has been found in observational follow-up studies. These studies show only progression to high- grade neoplasias in the presence of (persistent) HPV infections. This concerns women with normal [180]and abnormal [95, 174] cytology. Although these studies are very important for showing that HPV infection precedes the (progression of) neoplasia, they are too small [180] or have an inadequate design [95, 174] for assessing the duration between HPV infection and the development of CIN, and the sensitivity of the HPV test. Nevertheless, they suggest that the sensitivity for progressive HPV infections is high and, in that respect, they support our favourable model version A more than the unfavourable model B. This support emphasizes how worthwhile it is to carry out the required large prospective studies on the association between HPV and cervical neoplasia that hopefully will confirm the 'preliminary' findings.

The presented disease model has a number of simplifications. It does, for example, not discern low-grade on high-grade pre-invasive lesions, while HPV-negative CIN cannot become HPV positive. These simplifications, however, are not important for the results, and model refinements will be of little help as long as adequate longitudinal data on HPV detection are not available.

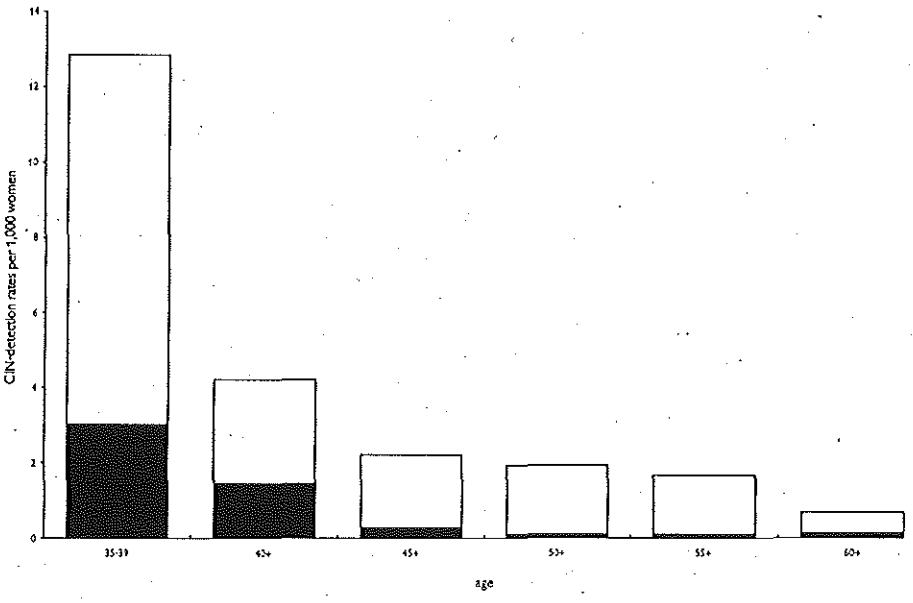
The results of the cost-effectiveness calculations concerning the policies that combine HPV testing and PAP smear screening are complex and their outcomes could not have been predicted easily. For the calculations concerning policies using only the HPV test, it is not surprising that when it takes 10 years for HPV infections to produce CIN, HPV screening can improve PAP smear screening. This is clearly not the case when HPV infection precedes CIN changes only by 1 year. But it is important to realize that these widely different assumptions are both compatible with the observed very strong association between HPV infection and cervical cancer, even if it is accepted that the HPV infection preceded the neoplastic changes that led to the invasive carcinomas. The work of Jenkins et al [103], who also assessed the effectiveness of HPV testing as a

Figure 9.2 *Simulated results of a hypothetically observational study using model A and B: age-specific histologically confirmed* CIN *detection rates at* PAP *smear screening in women who 5 years previously had had a negative* PAP *smear, by* HPV *status 5 years previously and age group at present screening*

Model A



Model B



• 5 years previously: cytology -/HPV+
○ 5 years previously: cytology -/HPV-

primary screening tool by using a stochastic model, illustrates this issue. The authors did not vary the parameters that are crucial for the outcomes. They used assumptions on the sensitivity of the HPV test that were very similar to those in our model version A. In the sensitivity analysis, the simulated screening situation was further improved (by assuming that 100% of the cancers develop in the presence of high-grade HPV), but lower sensitivity was not tested. As far as duration is concerned, Jenkins' assumptions are intermediate to ours. Although the authors agreed that selection of the progression parameters (which determine the duration of stages) was not unique, they did not vary the progression rate of HPV infection and therefore did not describe the complete range of possible (cost-)effectiveness of HPV screening.

To explore the impact of longitudinal data, we simulated an observational cohort study with the two model versions A and B. In the simulation, women who entered the study with negative cytology have a PAP smear 5 years later. Predicted CIN detection rates in women who at entry were HPV negative and those who were HPV positive were discerned *(see Figure 9.2)*. As the description of cervical neoplasia (CIN and invasive cervical cancer) of the model was the same in both model versions, the detection rate for CIN at PAP smear screening 5 years after negative cytology was the same. In version A, however, almost 70% of the women with histologically confirmed CIN (low and high grade) came from previously HPV-positive women, whereas in model version B this was only 20%. This reflects a higher predictive value for future CIN of a positive HPV test in version A. The fact that longitudinal outcomes clearly differ in both models means that different longitudinal outcomes can be consistent with present cross-sectional data, and that, once such longitudinal data are available, at least one (and probably both) of models A and B can be rejected. The range of combinations of parameter values on duration of HPV infections and sensitivity of the HPV test that are compatible with observed data will strongly decrease, and better predictions can be made of results expected from HPV screening.

Although the cross-sectional data show a strong association between HPV and cervical neoplasia, the results are insufficient to arrive at recommendations on screening. The discussion, therefore, on the representativeness of the test-positive rates that we aimed at in our simulation (4% in cytologically negative women, 67% in women with CIN and from 83% to 95% in women with invasive cervical cancers) is premature. Nonetheless, it is interesting to assess the influence of lower or higher observed HPV test-positive rates. In women with invasive cancer, the higher HPV positiveness, the better this will be for the effectiveness of HPV screening. Higher test-positive rates in women with normal cytology and in women with CIN, however, can only mean that more women who do not develop cervical cancer will be HPV positive (all women that will develop HPV-positive cervical cancer are already assumed to be HPV

positive before the development of the cancer). These women will unnecessarily be detected and followed up, and the negative side-effects and cost of follow-up will increase. In other words, given that HPV infection precedes, for example, 95% of the progressive neoplasias, lower HPV prevalence in the cytologically negative women and in women with CIN implies less harmless and less costly HPV screening.

A modelling approach, as presented in this paper, is useful for a joint analysis of cross-sectional, longitudinal and other relevant epidemiological data. We will adjust our model as soon as new evidence becomes available.

Data from large PCR-based cohort studies will accumulate in the forthcoming years. The fact that many of them are solely focussed on young women should be of major concern. The Copenhagen study [109] is restricted to women under 30 years of age, and the median age of the women in the Portland study is 34 years [15]. Screening for HPV in very young women would cause many women to be followed-up (because of the high prevalence in this age group of HPV infections that will clear) and is therefore not advisable. Moreover, the fact that prevalence is so much higher in younger age groups is also an expression of a different natural history of the HPV infections (at least a higher clearance rate) in this age group. Follow-up results from these women are obviously not transferable to the older age groups. Hence, further cohort studies should aim at women aged 30–60 years.

10

Discussion

In this chapter, I will first integrate and complete some important discussion points raised in the previous chapters. Subsequently I will discuss priorities for further research.

10.1 Effects, side-effects and cost-effectiveness

10.1.1 The effectiveness of cervical cancer screening

The implementation of routine cervical cancer screening was not preceded by randomized controlled trials. Hence, there is no unbiased direct evidence indicating that screening leads to a reduction in mortality. Effect assessment has to be based on evidence from non-experimental data, implying a relatively broad range of uncertainty. Within their limitations however, some non-experimental studies have provided convincing evidence that screening has reduced the level of incidence [100] and mortality [126] of cervical cancer.

We predicted the effects of screening by using a simulation model comprising all relevant aspects of cervical cancer screening. The natural history and sensitivity parameters used in the predictions resulted from an analysis of screening data from British Colombia [159] and have been validated with the IARC data on cancers after negative PAP smears [161]. We studied whether these estimated parameters were compatible with the incidence and mortality trend in the Netherlands (Chapter 6). In this analysis, we corrected for the fact that in the Netherlands, as in other countries [82, 126, 16], mortality had already decreased before screening(effect) started, mainly because of a decrease in risk in successive cohorts of women born between 1920 and 1940. Extrapolation of the effects of this decrease only partly explained the observed downward trend in mortality in the period after screening started. When the reduction from screening as predicted with our model was added, the predicted downward trends in incidence and mortality were similar to the observed trend.

Sceptics about the effectiveness of cervical cancer screening often refer to the situation in the United Kingdom especially in the seventies and eighties. However, there are two specific explanations for non-decreasing or even increasing mortality rates in the United Kingdom. Firstly, the UK cervical cancer screening programme experienced serious organizational problems at least until the late eighties [5]. Secondly, the risk for cervical cancer had been rising in women born after 1935 [166, 47]. Mortality in women born around 1951 was about twice as high as mortality in those born around 1941. Such a marked mortality increase in young cohorts was mainly seen in England and Wales and in some allied countries (Scotland, Ireland, New Zealand and Australia). It is much less pronounced in many other countries, e.g. the former Federal Republic of Germany, Finland, and the Netherlands [16]. In contrast with the situation in the UK, the highest estimate for achieved mortality reduction (60%) was made for Finland, where women were screened between the ages of 30 to 55 every 5 years since the mid-sixties [96].

It is questionable whether we can expect stronger evidence in the future on the effectiveness of cervical cancer screening than has been available to date. The current evidence is strong enough to make randomized trials comparing screening no screening unethical. After more than 25 years of widespread screening, it will become more and more difficult to disentangle screening effects from other effects.

10.1.2 The starting age for cervical cancer screening

It has been argued that screening in the Netherlands should start at an earlier age than the recently officially recommended 30 years. Arguments put forward to support this view include the following:

The relatively high histologically confirmed detection rates for high grade CIN under age 30 years [163]. However, detection rates are of limited value: it is quite possible that higher detection rates in young age are caused by higher prevalence of regressive lesions [159]. Moreover, since cervical cancer has a long pre-invasive phase, most of the cases will also be detected in time if screening starts at age 30.

The current incidence peak in age-group 35–39 *(see Figure 6.3d).* Here, the question is to what extent the current incidence peak in the age group 35–39 results from an increased underlying risk of cervical cancer in young birth cohorts, or whether it results from a decreased incidence (caused by screening or other factors) in middle-aged (40–69 years) women. From comparing pre-screening incidence data with the recent incidence data *(see Figure 6.3a* and *6.3d)*, it can be concluded that a decrease in incidence in the age-group 40–69 has been much more important than an increased incidence in women under age 40.

Moreover, it should be taken into account that in the Netherlands, mortality from cervical cancer is still much higher in older than in younger age and does, contrary to the incidence, not show a local peak around age 40 years *(see Figure 6.3e)*. The lower ratio of mortality to incidence in younger than in older age, also in the prescreening period, is an expression of the fact that overall clinical survival in young age is better than in older age. This means that mortality reduction resulting from prevention and early detection of invasive cervical cancers will be larger in older age than in younger age.

The increasing incidence and mortality rates in young birth cohorts in other countries [16, 47, 49, 55]. However, as mentioned in the previous section, the magnitude of the increase observed in the countries referred to (UK, New Zealand, Australia) is not generally observed in West European and North American countries, and also not in the Netherlands [16].

10.1.3 Side-effects of screening

Whenever medical procedures are applied, there will be side-effects, some of which are unfavourable. A full evaluation of health effects considers both the aimed-at effects and the side-effects. In cervical cancer, the primary goal is mortality reduction. That is why life-years gained were taken as the effect measure for cost-effectiveness in this thesis. The frequency of side-effects has been quantified and discussed (see Chapter 2 and 3). We concluded that the ratio of unfavourable to favourable effects becomes worse when the frequency of screening is increased and when screening is concentrated in young age. But until now, we have used costs per life year gained as the cost-effectiveness ratio, and this measure does not include side-effects. I will tentatively now include other effect measures in the cost-effectiveness. These other effect measures are related to quality of life. The net health effect of screening can be obtained by correcting the life years gained for changes in the health-related quality of life. Therefore, all types of effect should be quantified and expressed in a common measure. One method to achieve a common measure is the quality-adjusted life-years analysis. In this type of analysis, health states are valued after these states have been measured (described). The values or utilities should reflect the average willingness of individuals to trade off quality against quantity of life. There are no reports of such measurement and valuation in literature concerning health states produced or prevented by cervical cancer screening. I have therefore used outcomes of quality of life measurements for health states related to breast cancer screening performed by de Haes at al. in 1991 [91] to explore the favourable and adverse effects in quality of life of cervical cancer screening.

10.1.4 Side-effects of screening: tentative calculations

The main unfavourable effects of cervical cancer screening occur in women who require repeat smears, women referred for colposcopy and other diagnostic procedures, and women treated for pre-invasive conditions. The current number of women referred and treated is higher than the figures in Chapters 2, 3 and 5 which were based on data from the seventies and eighties, when follow-up was less intensive (see the Appendix of this thesis). For a more up to date assessment, we calculated the expected number of women followed up and treated from recent (1994 [141]) cytological and histological detection rates, accounting for the newly revised follow up guidelines in the Netherlands. We combined these figures with the incidence and mortality reduction predicted with our model (see Chapter 7) for the recently revised screening age range (30 to 60 years) and screening interval (5 years), assumed 75% attendance, no opportunistic screening, and 5% borderline screening results (after which women are advised to have repeat smears). The resulting ratios between the expected unfavourable effects and the expected beneficial effects are as follows: per prevented death, 2800 women have a screening exam; 140 women have a period of repeat PAP smears, 70 women are referred for colposcopical/histological evaluation, and 30 women are treated for preinvasive cervical neoplasia. But an additional beneficial effect of two prevented cases of invasive cervical cancer can also be expected for each death prevented. Because the estimated average gain in life expectancy is 26 years per death prevented, the numbers per life year gained are 26 times lower (see leftmost column of *Table 10.1*).

In the calculations presented here *(see Table 10.1* and *Table 10.2)*, the loss in utility resulting from cervical cancer screening attendance was assumed to be the same as that assessed for in breast cancer screening attendance, but it was applied for a period of two weeks (instead of one week for breast cancer screening), since this is the usual time interval before receiving the PAP smear result. This same disutility was applied to a period of one year after a positive smear in which women have (half) yearly repeat smears. This is in accordance with the guidelines that imply that it takes on average one year before women either return to regular screening or are referred for colposcopy.

The disutility (and its duration) caused by referral for colposcopical/ histological evaluation, were assumed equal to those for the diagnostic phase after a positive mammography. For the treatment of preinvasive cervical neoplasia, the loss in utility was assumed to be the same as for primary surgical treatment of breast cancer, but the duration of this loss was adapted: 2 days in the case of an out-patient treatment procedure (cryocoagulation or loop excision), 1 month for conisation and 2 months for total hysterectomy, that resulted in a weighted average of 2 weeks after accounting for the observed fractions for either treatment as presented in Chapter 4.

For the primary and palliative treatment and for the terminal phase of invasive cervical cancer, the disutility and the duration of corresponding breast cancer states were used *(see Table 10.1)*. The calculations are expressed in terms of how much quality loss and quality gain respectively is obtained per life year gained by screening.

The resulting tentative assessment of the quality of life impact shows a net loss in quality of life which corresponds to 16% of the gain in quantity of life. After discounting at 3% per year, the net loss increases to 26%, corresponding to a 35% increase in the cost-effectiveness ratio. Using a 5% discount rate, the loss in quality of life equals 44% of the gain in life years. This stronger impact of quality of life correction at the higher discount rate is due to the fact that unfavourable effects occur in the period immediately after screening, and therefore are much less affected by discounting than the beneficial effects.

According to these tentative calculations, the disutilities with the largest impact for the quality of life adjustment are the disutility of having biennial repeat smears after a mildly abnormal PAP smear, the long term disutility after a hysterectomy for the treatment of CIN, and the long term gain in utility because of less life-years with (treated) invasive cervical cancer. Applying 0.6% disutility throughout a one year period with one repeat smear half way and one more at the end might be an overestimation of the anxiety. The disease-free years after a hysterectomy for CIN and those after primary treatment of invasive cervical cancer have both been weighted equally to the disease-free years after mastectomy for treating breast cancer. Factors that play a role here are anxiety over possible recurrence, the impact for women of having lost the uterus or a breast respectively, and possible long term physical problems after surgery and/or radiotherapy. Accounting for all these factors, one would expect the long term burden of hysterectomy for CIN treatment to be on average less severe than the long term burden of mastectomy for breast cancer treatment, and also less than the long term burden of radical treatment of cervical cancer. To weight hysterectomy for CIN treatment as equal to radical cervical cancer treatment therefore overestimates the loss in quality of life.

Current PAP smear screening in the Netherlands differs unfavourably from the assumptions made in the above calculations in two respects. Firstly, although the coverage in 1994 was indeed approximately 75%, the total number of preventive (as opposed to follow-up) smears was around 30% higher than in our calculations because many smears additional were taken in relation to the targeted age-range and interval [14]. Extra smears in symptomless women will increase the negative side-effects of screening considerably more than the benefits (see Chapter 2). Secondly, repeat smears (taken for reasons other than because of inadequate quality) were advised after 10% of the screening tests instead of the assumed 5%.

Table 10.1 *A tentative correction for the quality of life impact per life year gained in cervical cancer screening (from age 30 to 60 every 5 years), using quality of life values for health states induced and prevented by breast cancer screening by de Haes et al* [91](see Table 10.2). *No discounting*

State	Number per life year gained*	Duration	I-Utility	Quality adjustment in years per life year gained
Screening attendance	106	2 weeks	0.006	-0.024
Period with repeat smears	5.3	1 year	0.006	-0.032
Colposcopical/ histological evaluation				
	0.47[1]	5 weeks	0.105	-0.005
	2.12[2]			-0.021
Primary treatment for pre-invasive cervical neoplasia				
	0.42[1]	2 weeks	0.133	-0.002
	0.64[2]			-0.003
First year after primary treatment for pre-invasive cervical neoplasia				
	0.42[1]	1 year	0.006	-0.003
	0.64[2]			-0.004
Period after total hysterectomy for pre-invasive cervical neoplasia				
	2.31[3]	All life-years after hysterectomy	0.053	-0.123
Primary treatment of invasive cervical cancer				
Micro invasive (IA)	0.01	2 months	0.133	-0.000
Local (IB)	-0.04	2 months	0.133	0.001
Non-local (II+)	-0.05	2 months	0.197	0.002
3 Months – 1 year after primary treatment of invasive cervical cancer				
	-0.09	10 months	0.156	0.011
> 1 Year after primary treatment of invasive cervical cancer				
	-0.59[3]	All extra life years with cervical cancer	0.053	0.031
Palliative treatment				
Radiotherapy course	-0.02	1 month	0.419	0.001
Surgery	-0.03	5 weeks	0.383	0.001
Chemotherapy cure	-0.07	4 months	0.469	0.011
Terminal phase	-0.04	1 month	0.712	0.002+
Total unfavourable adjustment				-0.217
Total favourable adjustment				0.060+
Total quality of life adjustment				-0.157

* The average number of life-years gained per death prevented is 26

[1] Direct referral after a highly positive (PAP3a mild dysplasia +, KOPAC P $\geq$ S, S $\geq$ 6 A $\geq$ 5, High grade SIL) screening cytology

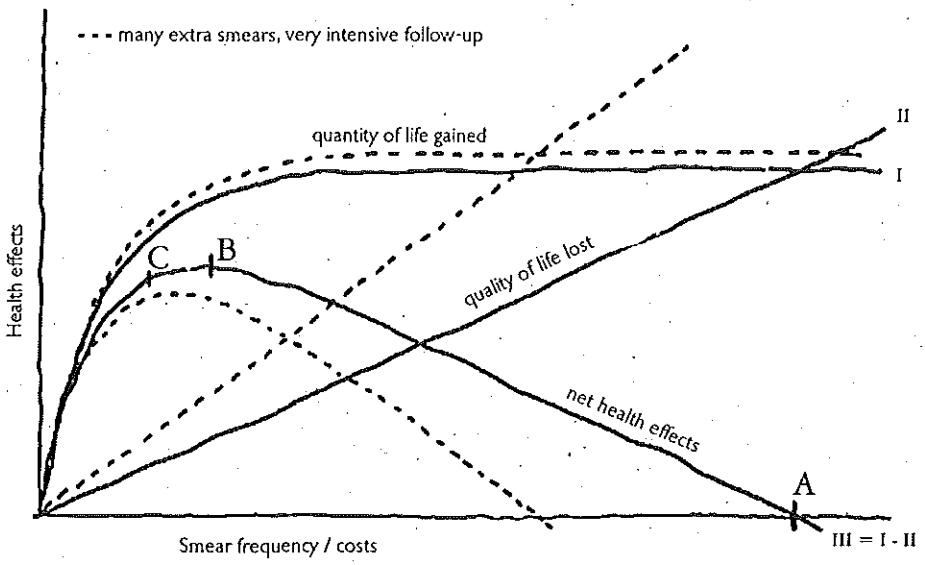
[2] Women with screening cytology results that initially generate a recommendation for repeat smears (PAP2 and PAP3a mild dysplasia, KOPAC-P234 G345, A34, ASCUS, low grade SIL)

[3] (Extra) life-years with hysterectomy / invasive cervical cancer respectively per life-year gained

Table 10.2 *Relationship between the Disutility and duration of states used for cervical cancer screening in* Table 10.1 *and those applied for breast cancer screening by de Haes et al [91]*

State	Relationship
Screening attendance	Disutility (0.006) and duration (2 weeks) as assessed for breast cancer screening attendance
Period with repeat smears	Disutility (0.006) as assessed for breast cancer screening attendance applied to 1 year after which the women either go back to screening, or are referred
Colposcopical/ histological evaluation	Disutility (0.105) and duration (5 weeks) as assessed for the diagnostic phase after a positive mammography
Primary treatment for pre-invasive cervical neoplasia	Disutility (0.133) as assessed for the primary surgical treatment of breast cancer, duration (2 weeks) accounting for 2 days in an out-patient setting, 1 month for conisation and 2 months for total hysterectomy
First year after primary treatment for pre-invasive cervical neoplasia	Disutility (0.006) as assessed for breast cancer screening attendance applied to one year because most recurrence/residue occur within one year
Period after total hysterectomy for pre-invasive cervical neoplasia	Disutility (0.053) and duration (all life-years after hysterectomy) as assessed for > 1 year after mastectomy
Primary treatment of invasive cervical cancer	Disutility and duration as assessed for breast cancer for primary surgical treatment (cervical cancer stages IA and IB) and primary radiotherapy (cervical cancer stage II+)
3 Months – 1 year after primary treatment of invasive cervical cancer	Disutility (0.156) and duration (10 months) as assessed for mastectomy
> 1 Year after primary treatment of invasive cervical cancer	Disutility (0.053) and duration (all extra life years with cervical cancer) as assessed for mastectomy
Palliative treatment	Disutility and duration as assessed for respective therapies in palliative treatment in advanced breast cancer
Terminal phase	Disutility and duration as assessed for the terminal phase of breast cancer

Figure 10.1 *Schematic representation of the health effects of screening by increasing smear frequency: the quantity of life gained, the quality of life lost and the net health effect for two screening situations (represented by the solid and dashed lines respectively)*



Accounting for the extra negative side-effects of both factors, the tentative quality of life correction on the life years gained in the calculations presented in *Table 10.1* would become 26% instead of 16%. After discounting at 3%, this would correspond to around 45%, which would almost double the cost-effectiveness ratio. This is an underestimation because we did not account here for the extra costs. On the other hand, we did not account for the extra benefits either, but these are relatively small (see Chapter 3 for the discussion on the extra benefits of preventive smears added to the programme smears). Therefore, these figures stress the importance of the reduction of extra preventive smears and of balanced follow-up guidelines for beneficial and efficient screening.

The tentative calculations presented here suggest that for a restricted screening situation with seven smears per women, little extra preventive smears and no more than 5% borderline screen results, the total health benefits of cervical cancer screening in the Netherlands would clearly exceed the total harm. They also show that for a less restricted practice as was carried out in the first half of the nineties, the harm in relation to the benefits might be substantial.

Figure 10.1 is a graphical representation of the relationship between health effects of screening and the frequency of screening. Line I represents the quantity of life gained (e.g. the number of life years gained). This line corresponds

with the line through the efficient policies in *Figure 7.1* in Chapter 7. Line II represents the relationship between the loss in quality of life and the frequency of screening. Line III represents the net health effects, and results from subtracting line II from line I. Since the unfavourable quality of life effects increase more or less linearly with increasing frequency of screening, while the gain in length of life levels off, the unfavourable effects will exceed the beneficial ones at some frequency of screening (point A). The level at which the incremental net health effects from increasing the frequency of screening becomes negative (point B) then has already been passed. All policies to the right of point B are inadmissible. The crucial level from a cost-effectiveness point of view is the one where the extra costs of increasing the screening frequency will become too high in relation to the extra net health effects (for example at point C). The tentative calculations presented above suggest that for the Netherlands, point A will not be reached by screening women 7 times. The dashed lines in *Figure 10.1* represent the influence of many extra smears and/or very intensive follow-up on the health effects. These lines show that in trying to maximize incidence and mortality reduction beyond a certain level, health effects in fact are diminished.

In order to make a better estimate of the position of line III, a better estimate of the position of line II is required. To this end, cervical cancer screening impacts on quality of life have to be systematically investigated, with the aim of expressing these impacts in quality adjusted life years.

10.1.5 Other recent cost-effectiveness studies on cervical cancer screening

In order to trace possible gabs or flaws in our cost-effectiveness analysis, we looked critically at differences with other published cost-effectiveness studies on cervical cancer screening. The most important studies and their main results are listed and compared to our MISCAN results in *Tables 10.3* and *10.4*. We will discuss the respective studies seperately below. A summary of the most important differences is given in *Table 10.4*. The studies differ in perspective (e.g. some predictions concern a cohort of women, others a population), in methodology (e.g. the discount rate) and in local characteristics (e.g. the background risk for cervical cancer and the costs of medical procedures in the area under consideration).

The background risk of cervical cancer that we derived for the Netherlands is considerably lower than in any of the other studies *(see Table 10.4)*. This is partly due to differences in incidence and mortality before screening was introduced, which were higher in e.g. Denmark and the UK than in the Netherlands. Another part of the difference is that we found that part of the decrease in cervical cancer incidence and mortality observed in the Netherlands since screening started is not due to screening and would have occurred anyway

Table 10.3 *Results of other cost-effectiveness studies on cervical cancer screening compared to our results. The discount rate (%) is indicated between brackets*

Study	Policy[1]	Results	
Costs per life year gained (Dfl)			
MISCAN 1993 (Chapter 7)	7:30-60 [5]	27 000 (5%)	6 000 (0%)
Hristova and Hakama 1997	7:30-60 [5]		-5 000 (0%)
Gyrd-Hansen et al 1995	7:30-60 [5]	7 000 (5%)	
Eddy 1990	16:20-74 [4]	20 000 (5%)	
Number of smears per invasive cervical cancer case prevented			
MISCAN 1993 (Chapter 7)	7:30-60 [5]	250 (5%)	1300 (0%)
Sherlaw-Johnson et al 1995	16:18-64 [3]		600 (0%)
Mortality reduction with 7 smears[2]			
MISCAN 1993 (Chapter 7)	7:30-60 [5]	50%	
Gustafsson and Adami 1992	7:28-66 [var]	76%	

[1] Number of smears: starting age-ending age [interval]
[2] Independent of discounting

(compare the dashed lines in *Figures 6.3 a/b* with *Figures 6.3 d/e* in Chapter 6). The last explanation for the difference in background risk, is that we accounted for the amount of screening that took place in the Netherlands before 1993, the first year of the predictions of effects and costs of screening. The baseline risk in a screended population is ofcourse lower then in an unscreened population.

In *Table 10.5*, we present cost-effectiveness ratios for the Dutch situation, when we adjust the MISCAN model for the differences listed in *Table 10.4*. This was carried out to explore to what extend these differences could explain differences in predicted cost-effectiveness ratios. A detailed overview of the adjustments is given in Appendix B. Note that remaining differences in the adjusted screening policies (age range and intervals) may to some extend explain or hide remaining differences in cost-effectiveness.

10.1.6 Systematic discussion of the studies

Hristova and Hakama 1997

Hristova and Hakama [96] estimated the costs per life year gained for screening as performed in Finland (between age 30 and 60 every 5 years) for the Nordic countries. They used the mortality reduction estimated for Finland on basis of the observed mortality trends. The number of invasive cancers was assessed at approximately twice the number of deaths from cervical cancer prevented. The number of detected CIS was assumed to be three times the number of prevented invasive cancers.

Hristova and Hakama arrived at negative cost-effectiveness ratio, because the savings from preventing cancer cases and death cases were larger than the costs of screening. In our analysis (and in all other analyses discussed here) the screening costs were larger than the savings. The difference is not explained by differences in screening costs, but by a combination of a higher number of (advanced) cancer cases prevented and higher treatment costs per case of (advanced) cancer in the Nordic study. For a more detailed comparison, we considered the figures the authors presented for Denmark. The higher effects are explained by a higher background incidence (rate ratio 2.4) and a higher mortality reduction (90% versus 50%). Also, the savings per invasive cancer and per advanced case prevented are 60% higher.

The estimated 90% mortality reduction results from a combination of a high estimate of the background risk of dying from cervical cancer and a low estimate of the future mortality in screened women. The estimate for the background risk is based on pre-screening mortality data from Denmark, and results from the fact that mortality increased from 1955 to 1960 in most age-groups. Such an age-independent increase is suspect for a flaw in the data. The estimate of future mortality rates in screened women are based on Finnish mortality trends after screening started and results from the assumption that an increase in mortality reduction in a screened cohort of women goes on with age, whereas one would expect that the mortality reduction levels off.

The higher the diagnosis and treatment costs per detected pre-invasive case for the Nordic countries (Dfl 8000 versus Dfl 3500) combined with a higher detection rates for pre-invasive disease results in higher costs for the management of detected pre-invasive disease. However, these high costs are almost entirely compensated for by the lower unit costs for smears (Dfl 20 for the Nordic countries versus Dfl 70 for the smear plus Dfl 6 for false positives for the Netherlands) *(see Table 10.4)*. If we account for the differences mentioned, the MISCAN model will also predict savings for cervical cancer screening instead of costs *(see Table 10.5)*.

Gyrd-Hansen et al 1995

Gyrd-Hansen et al [86] estimated the cost-effectiveness of cervical cancer screening for the Danish situation. The disease model used by Gyrd-Hansen incorporated a pre-clinical detectable phase which precedes clinical cancer, which has a certain duration and for which screening has a certain sensitivity. This model was very similar to the model used in our study.

The four times lower costs per life year gained (Dfl 7,000 per life-year gained) can to a high extent be explained by the higher background incidence (factor 3.4) and a higher assumed attendance (80% vs 75%) combined with no difference in risk between attenders and non attenders. If we account for these differences, our MISCAN model produces a CER of Dfl 4000 per life year gained.

Table 10.4 *Methodology and parameter values used in the respective cervical cancer cost-effectiveness analyses*

Study Country	MISCAN Netherlands	Hristova Denmark	Gyrd-Hanson Denmark	Eddy USA	Waugh UK	Sherlaw-Johnson UK	Gustafsson (Sweden)
Model/approach	Population	Population	Population	Cohort	Population	Cohort	Cohort
Discounting	5%	0%	5%	5%	0%	0%	0%
Policy[1]	7:30–60 (5)	7:30–60 (5)	7:30–60 (5)	16:20–74 (4)	14:20–59 (3) versus 9:20–60 (5)	16:18–64 (3)	7:28–66 (var)
Background risk ratio	1	2.4	3.4	1.8	± 2	3.5	NA
Previous screening	Yes	No	No	No	Yes	No	No
Attendance	75%	]	80%	100%	]	70%	]
Relative risk attenders	0.74	NA	1	1	NA	1	75%
Sensitivity	80%	]	?	Modelled differently	]	?	]
Regression	Yes	Yes	Yes	Yes	NA	Yes	Yes
Mortality reduction	50%	90%	?	88%	—[3]	—[4]	75%
Costs per screening smear (Dfl)							
Screening smears	70	20	71	152	]	Costs are not included in the analyses	
False positives	6	]	?	]	76		
Management of CIN	8	80	?	?	]		
Savings per invasive cervical cancer case prevented (Dfl)							
Initial therapy	19 000	53 000[2]	?	34 000	24 000[2]		
Costs per advanced cancer prevented (Dfl)							
Advanced cancer	31 000	]	?	44 000	]		

[1] Number of smears: starting age – ending age (interval in years)
[2] Medical costs per invasive case, including treatment for advanced cancer in some of the women. For the Netherlands we predicted one case of advanced cancer per 2.2 cases of invasive cancer, and the corresponding total costs are Dfl 33 000
[3] 8% Incremental mortality reduction of 3-yearly vs 5-yearly screening
[4] 64% Incidence reduction
? Lacking information in the publications consulted
NA No Assumptions

A more detailed comparison between this Danish model and ours was not possible because of incomplete information on this model in the English literature. The assumed duration of pre-clinical disease and the sensitivity of the smear for this disease phase were not given. However, the model predicted the IARC rates for interval carcinomas fairly well and so does our model [161]. Hence we do not expect important differences in the prevention of invasive cancers. The total clinical survival was similar to ours. The survival in screen-detected invasive cancers seems to be higher in the Danish assumptions, but since the number of screen-detected cancers is low in cervical cancer screening, this will

Table 10.5 *The cost-effectiveness ratio calculated from the* MISCAN *predictions adjusted for the differences* (see Table 10.4) *between the respective studies and our analysis. For* MISCAN *we present predictions for the policy* (see Table 7.8) *that was closest to the policy in the other study. Parameters adjusted for are the background risk, the mortality reduction and the costs. The discount rate (%) is indicated in brackets. The details of the adjustment are given in Appendix B of this thesis*

	.... Other studies		..... MISCAN		
Study	Policy[1]	Results	Policy[1]	Unadjusted results	Adjusted results
Costs per life year gained (Dfl)					
Hristova and Hakama 1997	7:30–60 [5]	-5 000 (0%)	7:30–60 [5]	6 000 (0%)	-2 000 (0%)
Gyrd-Hansen et al 1995	7:30–60 [5]	7 000 (5%)	7:30–60 [5]	27 000 (0%)	4 000 (5%)
Eddy 1990	16:20–74 [4]	20 000 (5%)	10:27–72 [5]	28 000 (5%)	21 000 (5%)
Incremental costs per life year gained 3-yearly versus 5-yearly screening (Dfl)					
Waugh et al 1996	14:20–59 [3] versus 9:20–60 [5]	24 500 (0%)	9:30–54 [3] versus 6:30–55 [5]	33 000 (0%)	20 000 (0%)
Number of smears per prevented invasive cancer case					
Sherlaw-Johnson et al 1995	16:18–64 [3]	600 (0%)	10:27–72 [5]	1300 (0%)	350 (0%)

[1] Number of smears: starting age – ending age [interval]

not have a very big impact on the (cost-)effectiveness. The costs per smear were the same (Dfl 71 and Dfl 70 per smear), but the costs of false positive test results (in our predictions Dfl 6 per smear) were not available for the Danish analysis (there are 0.62% false positive results, which was less than in our analysis, but the associated costs were not given). The regression rate assumed by the Danish authors is somewhat higher than in our model, but the costs per detected CIN/CIS were not given. Information on the savings per invasive case prevented were also lacking.

Eddy 1990

Eddy [65] estimated the cost-effectiveness of cervical cancer screening in the United States. The author also used a disease model with a pre-clinical disease - clinical cancer sequence, similar to our model.

The policy with the lowest number of smears in Eddy's analysis was 16, for which the author predicted a cost-effectiveness ratio of Dfl 20 000 per life year gained. In our analysis, the highest number of smears was 10, with an estimated Dfl 28 000 per life year gained. The cost-effectiveness ratio increases with higher screening frequency, indicating that for similar screening policies the difference in cost-effectiveness will be much more then 40%. We

will use our predictions for this policy with 10 smears to make a comparison with Eddy's predictions in more detail. When we compare the effects and the costs, it becomes apparent that part of the differences are masked in the cost-effectiveness ratio: both the effects and the costs are higher in Eddy's predictions. Higher effects can be explained by an assumed higher background incidence for the USA compared to the Netherlands (factor 1.8) and by a higher mortality reduction (88% for Eddy's model versus 58% for MISCAN). This higher mortality reduction is almost completely explained by the fact that Eddy considered women who attend all screening rounds and assumed that these women are at average risk. The higher costs are explained by higher costs per smear (Dfl 152 vs Dfl 76, both including the diagnostic costs of false positives). These higher costs are only partially compensated by the higher savings per prevented case with invasive or advanced cancer. The costs for diagnosis and treatment of detected pre-invasive disease were also considerably higher in Eddy's calculations (Dfl 11 000 vs Dfl 3 500). However, although Eddy assumed regression it was not clear how much. Therefore the total impact of the difference in costs per detected case with a pre-invasive lesion is unknown. If we account for the differences mentioned (using our incidence of regressive CIN), the costs per life-year gained calculated with MISCAN for the policy with 10 smears per women are lowered to Dfl 21 000. This is higher than the Dfl 20 000 estimated by Eddy for 16 smears per women. A further difference between the two models is that we accounted for the screening that took place in the Netherlands in the two decades before the start (in 1993) of the period for which we made our predictions. Without his history of screening, our model would probably indeed predict a clearly lower cost-effectiveness ratio than Eddy's.

A difficulty in comparing Eddy's results with ours in more detail is that Eddy's predictions concern a cohort of women, and not — and more realistic — a population during a defined period of time as in MISCAN. In particular the discounting poses problems, because it is not clear from Eddy's publication to what point in time effects and costs have been discounted.

Waugh et al 1996

Waugh et al [208, 209] estimated the incremental costs per life-year saved from 3-yearly screening compared to 5-yearly screening at Dfl 24 500. This was based on an analysis of the screening history of 24 invasive cases, concluding that 2 cases would probably have been prevented if the interval would have been 3 instead of 5 years. The autors assumed that this corresponded to one case of death prevented. Accounting for the fact that the effectiveness estimate was based on two cases only, the authors calculated a 95% confidence interval for the CER of Dfl 5 000–50 000 per life year gained.

Waugh et al compared two policies with the same age-range, and consequently the policy with a 3-year interval had 55% more (14 vs 9) smears per woman than the policy with a 5-year interval. For purposes of comparison, we could use the same approach in MISCAN, by comparing a policy with age-range 30–54 years and an interval of 3 years (9 smears per woman) to a policy with age-range 30–55 years and an interval of 5 years (6 smears per woman). Since the authors from the UK did not include discounting, we also looked at undiscounted results from MISCAN. Comparing the two policies mentioned, the incremental costs per life-years gained were estimated at Dfl 33 000 using MISCAN for the Dutch situation. The mortality reduction increases by 8%, which is equal to the 8% incremental mortality reduotion estimated by Waugh et al. The total number of lifetime smears is much larger in Waugh's study, and therefore one would expect a much higher incremental mortality reduction than for the policies compared by the UK study. However, the effects in Waugh's study were based on two cases only. Therefore, although the point estimate for effectiveness was more favourable, the confidence interval of the authors results easily include our estimate.

The background incidence in Waugh's study population was approximately a factor 2 higher than we assumed in our calculations. As far as the costs are concerned, the costs per smear (including costs of false positive smear results and the management of CIN) were Dfl 74 in Waugh's analysis versus Dfl 84 in ours. On the other hand, the savings per invasive cancer case prevented were also lower (24 000 vs 33 000, both including the costs of advanced disease). After adjusting our model for the differences mentioned, the incremental CER with at zero % discounting of 9 smears with a 3-year interval versus 6 smears with a 5-year interval becomes Dfl 20 000 *(see Table 10.5)*. Because of decreasing incremental effectiveness, this incremental CER will become higher when we compare exactly the same two policies (with 14 and 9 smears) that have been considered in Waugh's paper.

Sherlaw-Johnson et al 1994

Sherlaw-Johnson [184] and colleagues estimated the incidence reduction and the number of smears needed for screening women aged between 18 to 64 years every 3 years (16 smears per woman) for the United Kingdom. To this end, the authors used a disease model with a pre-clinical disease — clinical cancer sequence.

Although we think that incidence reduction is not the most appropriate effect measure for cost-effectiveness analyses, we are able to compare these outcomes with ours. From the predictions the authors present (zero discounting), it can be calculated that around 600 smears are needed per invasive case prevented. Without discounting, the number of smears per invasive cancer prevented for the Dutch situation and for the policy of 10 smears per women is assessed

at 1300. The assumed background incidence in the UK analysis was 59 per 100 000 women aged 18 and older, which was 3.5 times higher than the background incidence in our analysis. The estimated incidence reduction was 64% versus 50% in our analysis. But if we calculate the incidence reduction for 70% attendance in women of average risk in the Dutch situation, the mortality reduction increases to 63% as well. If we adjust MISCAN for these differences, we predict approximately 350 PAP smears per prevented case of invasive cancer, which is lower than the 600 smears predicted by Sherlaw-Johnson et al. A lower number is expected because of decreasing incremental effectiveness, since our predictions concern a policy of 10 smears per woman, and Sherlaw's for 16 smears per woman.

Gustafsson and Adami 1992

The study of Gustafsson and Adami [84] is not a cost-effectiveness analysis, but is nevertheless interesting to compare with our analysis. The authors estimated what are efficient policies with different numbers of smears per woman. To this end, they used a disease model with a pre-clinical disease — clinical cancer sequence. The predictions did not focus on an existing population although important parameter value estimates were based on Swedish data. The analysis did not include costs or total numbers of smears. The outcomes which we can compare with ours are the mortality reduction for the efficient policy with 7 smears per women, and the characteristics of this policy. The estimated mortality reduction of a the optimal policy of 7 smears per woman in Gustafsson's analysis was around 75% versus 50% in our analysis. The difference is partly explained by the difference in the assumption of 'smear-efficiency', defined by Gustafsson as the proportion of the prevalent cases of pre-clinical disease that is eliminated in an invited cohort of women, and assumed to be 75%. In our model this corresponds with *sensitivity* × *attendance rate* × *the risk in attenders compared to the average risk*, that is 80% × 75% × 0.74 = 44% for first smears in a womans life, and even lower in smears with a higher rank. The mortality reduction for an efficiency of 50% and 25% according to Gustafsson would be around 60% and 40% respectively.

In identifying an efficient policy of 7 smears, Gustafsson also considered policies with variable intervals between smears. The authors argue that since the prevalence curve for pre-invasive disease shows a peak in young age (around 35 years) in the situation without screening, it will be more efficient to screen with shorter intervals in younger than in older age. When they optimized for the number of cancers prevented, the efficient policy of 7 smears was to screen women at age 28, 32.5, 37, 42.5, 49, 57, and 66 years. The respective intervals are 4.5, 4.5, 5.5, 6.5, 8, and 9 years. The optimal starting and ending ages are similar to ours (28 and 66 years versus 27 and 69 years). The authors report that optimizing on life-years gained instead of incidence reduction did not make much difference for the characteristics of optimal policies. We also made

predictions with a policy with smaller intervals in younger than in older age, similar to the one suggested by Gustafsson and Adami. Undiscounted, this resulted in more life years gained than the corresponding policy with a fixed interval of 7 years, but with a 5% discount rate, the advantage of using variable intervals disappeared. Gustafsson et al optimized without discounting.

10.1.7 Conclusions

The comparison with other studies did not reveal gaps or flaws in our cost-effectiveness analysis. In fact it confirmed the comprehensiveness and soundness of our study. Three important features of our analysis were not present in most of the other analyses. Firstly, we accounted for an autonomous trend (independant of screening effect) towards a lower risk for cervical cancer that continued after screening started. Secondly, the already relatively low background risk for cervical cancer in the Netherlands was further reduced in our calculations because we assumed that only 75% of the population would attend, and that the non-attending women as a group were at a higher risk than attenders. Thirdly, we accounted for the fact that when the predictions start the population is already screened to a certain extent, resulting in a lower prevalence of pre-clinical disease and thus a lower base-line risk for cervical cancer. As a result, the mortality reduction in the total population and the number of life years gained in our predictions are modest, and the cost-effectiveness of cervical cancer screening estimated for the Dutch situation was less favourable than the other estimates for other countries.

The comparison made shows that cost-effectiveness results should be interpreted with caution. Large differences may occur not only because of true differences in epidemiology of cervical cancer, the screening policy evaluated and the cost of screening and treatment of cervical cancer, but also because of methodological differences between studies.

10.2 Priorities for further research

10.2.1 Optimization of follow-up strategies

The ongoing trend towards more active follow up of (very) mild cytological abnormalities is of major concern. In the Netherlands in the seventies, cytological PAP class 2 was considered to be a negative outcome, not needing any follow up [4, 178]. In the eighties, it was stated that a PAP 2 result should be followed by a repeat smear after one year [205]. Recently, it was decided at a national level that the repeat smear should be made at an interval of 6 months [59]. Similar trends have occurred in other countries [169, 121, 124, 63, 31, 74]. Consequently, the proportion of women referred for at least cytological follow-up (and after further cytology sometimes also for colposcopy) has increased

Table 10.6 *Rates of positive smears (i.e. smears with a follow-up advice other than regular screening) in different periods (excluding smears of inadequate quality).*

Country	In the early days of screening		Recently	
Norway	Ostfold, 1960–1968 [169]	0.6%	National data, 1972–1994 [21]	5.5%
USA	Massachusetts, 1975 [63]	0.2%	Mc Neil, 1995 [139]	3%–10% or higher
Netherlands	Utrecht, 1982 [178]	0.8%	National data, 1994 [14]	10%

strongly over the last decades from less than 1% in the sixties to 5%–10% in recent years *(see Table 10.6).*

Nowadays in medicine in general, and certainly also in (cancer) screening, major changes in diagnostic or treatment procedures must be supported by solid evidence for its (cost-)effectiveness, preferably from randomized controlled trials. No such evidence has driven the intensification of follow-up strategies. The unfavourable side-effects of follow-up have not been described well, although this could be remedied by collecting long term and complete follow- up data. Most probably, unfavourable side-effects increase with more intensive follow-up. Assessment of the extra benefits of more active follow-up is less easy. Ideally, randomized trials could measure differences in mortality, or at least in incidence. But the incidence rate of invasive cancer in women with low grade cytological abnormalities in their PAP smears is very low for any of the alternative follow-up strategies. Study populations of hundred thousands of women with abnormal smears would be required. Such large studies would not be considered. In the USA, a randomized trial (the ALTS-trial) is currently taking place, investigating the effects of different follow-up strategies for smears with ASCUS (atypia of squamous cells of undefined significance) and low grade SIL (squamous intraepithelial lesions) [139]. However, for financial and other reasons, it was not possible to use decrease in clinical cervical cancer or mortality as effect measures. Instead, a surrogate endpoint has been chosen, missed detection of high grade SIL, which makes the interpretation of the results very problematic. Some strategies are more likely to miss less high grade SIL than others. Even if one assumes that this necessarily incurs a reduction in cancer incidence and mortality, it will remain uncertain as to how many (deaths from) cervical cancer will be prevented. Possibly, after extrapolation of the extra detection of high grade SIL to extra prevention of cancer and death, even the upper-hand of the confidence interval for the incremental benefits of intensifying the follow-up may result in an unfavourable cost-effectiveness ratio, thus rejecting the intensive follow-up strategy. Such

an analysis would be important to stop further intensification of the follow-up. It will not however provide an estimate of what could be an optimal follow-up strategy.

In the Netherlands, we will use the data retrieved from the PALGA (with a high national coverage of cervical cytology and histology over the past 10 years) to perform a descriptive study of the follow up practice and its outcomes. This will, at least for the Dutch situation, show the long-term consequences for women (in terms of numbers of repeat smears and biopsies etc., numbers of years in follow-up) of the follow-up practice. These non-experimental data are not appropriate for answering the question on the differences in effectiveness of different follow-up strategies: outcomes (e.g. cancer rates during follow-up) may be (and probably are) biased because at least part of the variation in intensity of follow-up is not random with respect to outcome. Women with more intensive follow up, by self selection and by the physicians selection, probably as a group do not have the same risk for cervical cancer compared to women with less intensive follow up. Again, it cannot be excluded that an analysis would provide a minimum for the incremental costs per life-year gained by intensifying the follow-up that is so high that the appropriateness of intensive follow-up should be seriously questioned.

Meanwhile, studies have begun started to investigate to what extent HPV-detection performed on a cervical scrape could solve the problem of the follow-up of 'borderline' smears (see the paragraph 'human papilloma virus (HPV)' below in this chapter). It is hoped that women with borderline cytology but a negative HPV-test would not require follow-up other than further screening after the regular screening interval, and that only women with persistent positive tests for high risk HPV-types need referral for colposcopy and biopsies. We await the results of these studies.

10.2.2 The influence of screening on quality of life

For estimating the total effects of screening, it is necessary to assess the influence of the various positive and negative side-effects of screening on the quality of life. This influence is not well known. Some work has been published [138, 212, 32], but in order to assess the costs per quality adjusted life year (QALY) gained, which is a more complete measure for the balance between benefits, negative side effects and costs than the cost per life year gained, more thorough quantification is required. In a previous section ('Side-effects of screening') of this chapter, I used crude approximations on the basis of utilities assessed for health states produced and saved by breast cancer screening. It is important to determine values for quality of life specially for health states related to cervical cancer screening.

10.2.3 Computer-aided screening

Improving the test-characteristics of the cyto-morphologic evaluation of the smear could lead to a higher sensitivity without increasing the false positive rate or to a higher specificity without losing sensitivity. Automated pattern recognition has become a serious option over the last decade, and several systems for computer aided evaluation of PAP smears have been developed and tested [99]. A new method for cytological evaluation is usually initially tested in rescreening studies on series of smears enriched with positive cases. The results of such studies on e.g. PAPNET, one of the computer aided cervical smear evaluation systems, are encouraging, but of limited value when routine screening is concerned, because of biases which are inherent in these experiments (Chapter 8). One study on larger series of smears with a more or less routine mixture of negative and positive smears has been published [113] and others are in progress. The results will provide a better estimate of the performance (both in terms of false negatives and false positives) of PAPNET and other computer-aided screening techniques.

Along with estimates for the decrease in costs because of time-saving for the cytotechnicians, and estimates for the increase of costs because of the technology used, this should provide insight into the cost-effectiveness of these new techniques.

10.2.4 Human Papilloma Virus HPV

After HPV infections have been found to be strongly associated with the presence of CIN and invasive cervical cancer in cross-sectional studies [149, 183, 146, 66], attention has been focused on HPV-detection as a possible discriminator between harmless and pre-cancerous neoplastic morphological changes [124]. It has also been suggested that HPV-screening could be used in primary screening. Almost all (95% or more) invasive cervical cancers are HPV-infected [141, 26]. The hypothesis is that only a small fraction of the invasive cancers develop without having been preceded by an HPV infection. This is corroborated by longitudinal epidemiologic studies on pre-invasive disease [180, 95, 174]. Cross sectional data also show that only a small fraction of the infections with oncogene HPV genotypes are followed by the development of invasive cervical cancer. In the Netherlands, the prevalence of oncogene HPV is relatively high (at least 10%) in young (age <30 years) women, and much lower (around 4%) after the age of 35 years [180]. This decrease in prevalence, and the fact that the life-long cumulative risk for having cervical cancer diagnosed in the situation without screening is around 1.5%, shows that most of the HPV infections are cleared and not followed by progression to cancer. The current hypothesis is that an HPV infection is a necessary but not sufficient step in the cascade of events that leads to invasive cancer, in other words that cervical cancer is a

Table 10.7 *Possible designs for longitudinal studies on the effects of* HPV-*screening in combination with* PAP *smear screening*

..... Type of study			..... Effect-measure .....
1a	Observational	retrospective	high grade CIN
1b	Observational	retrospective	invasive cervical cancer
2a	Observational	prospective	high grade CIN
2b	Observational	prospective	invasive cervical cancer, only possibly in meta-analysis, power probably not sufficient
3a	Intervention, randomized	prospective	high grade CIN
3b	Intervention, randomized	prospective	invasive cervical cancer, only possible in meta-analysis, power probably not sufficient

rare complication of HPV infection [140]. The exploration of the possible role of HPV-detection in population based cervical cancer screening as presented in Chapter 9 is based on this hypothesis. In this scenario, HPV-positive women are followed-up until the HPV infection is cleared or cervical intraepithelial neoplasia has developed (and is then treated). The exploration shows that, from a cost-effectiveness point of view, the current data do not give clear evidence for or against HPV-screening. Only further longitudinal data can give the information required: how long beforehand cervical cancer is preceded by an HPV infection and what is the sensitivity of an HPV-test in a cervical scrape during this period.

In *Table 10.7*, various alternative designs for longitudinal studies for further evaluation of HPV-screening are displayed. One 2a type study in women over 30 years of age has been published [180]. In this study, the relative risk of developing CIN 3 within 3 years after a negative PAP smear was significantly higher in high risk HPV positive women than in negative women, with a point estimate of 116 times (95% CI 13–990). The total number of cases however (7 CIN 3 cases) is too small to allow for a (cost-)effectiveness analysis of HPV-screening, so larger studies are required. Conclusive information about the effectiveness of HPV-screening can only be attained if invasive cervical cancer is the effect-measure. In a prospective study (2b and 3b type studies), this would require a study population of hundred thousand of women and a follow-up period of several years. Therefore, research is currently focused on the testing of archived smears from cases with cervical cancer (and smears of controls) for HPV (1b type studies). This is another way of investigating to what extent

invasive carcinomas have been preceded by detectable HPV infections. The results of two retrospective studies so far have been published. Chua and Hjerpe found 60% of smears to be HPV-positive on average 4 years before diagnosis of invasive cervical cancer in 30 cases versus 11% in the age-matched controls [42]. Walboomers at al. reported 96%–100% of smears to be HPV-positive in 24 false negative smears made on average 2.4 years before the diagnosis of invasive cervical cancer in 16 cases [207](no matched controls).

Before implementing HPV-testing on a regular basis in a population screening programme, further information is also needed concerning the acceptability, the feasibility and the outcomes in a real life situation. This can only be accomplished in an intervention study. For these questions, a 3a type study design is sufficient, and a number of this type of studies is under way: studies are (planned to be) undertaken in the Nordic countries and in the Netherlands, in which women with negative cytology and women with 'borderline' smear results are followed-up by HPV-tests in the intervention arm and not (negative smears) or by repeat smears (borderline smears) in the non-intervention arm.

Studies type 1a are also underway, and may provide a link between type 1 and 3 studies: to compare the expected yield of high grade CIN, on the basis of the findings in the retrospective observational 1a type studies, with the observed yield in 3a type intervention studies.

Meanwhile, there is hope that an effective HPV vaccine will be developed [73, 201]. If indeed HPV infections have a conditional role in the etiology of cervical cancer, and if a vaccine at low cost becomes available, vaccination could replace early detection. In countries where already mass-screening for cervical cancer is practised, starting with vaccination would possibly mean that screening can be phased out when cohorts of women that became sexually active before the vaccination started have become older than 60 years of age.

10.3 International and national perspective

10.3.1 International perspective

The conclusions of this thesis, which are summarized in the next chapter, are based on an analysis of important international datasets and extrapolation of the results of this analysis to the Dutch situation. In principle, this can also be done and to some extent has been done for other countries (see paragraph 'Other recent cost-effectiveness studies on cervical cancer' above). The following general conclusions however can be made: the coverage of the screening is the most important factor for the (cost-)effectiveness of screening; more than seven smears per woman in a cervical cancer screening programme is only reasonable from a cost-effectiveness point of view if the background risk for

cervical cancer is much higher than in the Netherlands or if financial considerations are different. When the number of smears is indeed restricted to 10 or less, a broad age-range is more important than a short (<5–7 years) interval; extra (opportunistic) smears outside the target age-range and frequency have an adverse effect on the cost-effectiveness of screening.

Even if the national authorities are aware of these conclusions and their implications, the problem is implementation in daily practice. This particularly holds for the intensity of screening. As far as the interval between subsequent screenings is concerned, it seems that there is a favourable international trend to less frequent screening. As far as the number of women followed up is concerned, this number until now has only increased.

In a situation with a centrally organized screening programme with an invitational system, it is already difficult to control the screening practice. It will be even more difficult in a situation with only opportunistic screening. Part of the problem in both situations is that many professionals (in particular general practicioners, gynaecologists and pathologists involved) are afraid, also because of possible law suits, to 'miss' cases. Good management of cervical cancer screening requires a consensus between these professionals, and official quality assurance guidelines to back up the physicians. These guidelines should have legal status in order to prevent law suits threatening the quality of screening.

10.3.2 Ongoing evaluation of the Dutch national programme

Cervical cancer screening needs close evaluation, because screening practice can easily become ineffective or inefficient. In 1991, the Dutch Government decided that complete evaluation of the national cervical cancer screening programma should be conditional for the programme to be continued [192]. Since then, major efforts have been put forward to ameliorate the data collection and registration. As a result, a much more complete evaluation is already currently possible. The main remaining data problems are the lack of reliable information on the reason why the smear was taken (i.e. was the smear taken as a screening test or because of symptoms?), and the fact that the identification of individuals in PALGA (date of birth, the first four caracters of the name at birth, and gender) is not 100% unique and sometimes combines two or more women in one identification code in the registration process.

Main evaluation issues that have to be addressed are coverage of the women (most) at risk, the amount of screening outside the targeted age-range and interval, the intensity and completeness of follow-up, and the quality of the smear taking and of the cytologic evaluation of smears.

Sice its inception, cervical cancer screening has been more intensive than prescribed in official guidelines. The relatively high prevalence of pre-invasive

neoplasia in young women and the occurrence of invasive cancers at young age was an incentive to start screening earlier in a woman's life. The occurrence of interval cancers pushed clinicians to shorten the screening interval and/or to intensify the follow up strategy. These mechanisms, that can result in more harm than good, make close monitoring particularly important.

11

Conclusions

In the introduction (Chapter 1), seven crucial questions have been put forward. The results of the analyses dealing with these questions have been presented and discussed in the respective Chapters. The answers will be summarized here.

1 **Is cervical cancer screening effective in reducing mortality? And if so, how large are the beneficial effects (reduction in incidence and mortality)?**

For the Netherlands in the early nineties, mortality reduction as a result of screening activities since the early 1970's is estimated at about one third of total cervical cancer mortality (Chapter 6), thus preventing around 150 cervical cancer deaths and 330 cases of invasive cancer annually. For the new programme with 7 screenings between the ages 30 and 60, and assuming 75% attendance, the mortality reduction is predicted at approximately 50%. For women attending all 7 screens, the risk of dying from cervical cancer is reduced by 75%. The main reason for not attaining 100% in regularly attending women is the incidence after age 60, followed by false negative test-results, incidence under age 30, and fast growing tumors.

2 **How large are the unfavourable health effects of cervical cancer screening? Are they outweighed by the beneficial health effects?**

Before considering cost-effectiveness, it must be clear that the net (beneficial minus unfavourable) health effects are positive. For a programme in the Netherlands with limited screening every 5 years between age 30 and 60, the unfavourable effects can be estimated as follows: per prevented death and two prevented cases of invasive cervical cancer, 2800 women have to attend a screening; 140 women will have a period of repeat PAP smears, 70 women will be referred for colposcopical/histological evaluation, and 30 women will be treated for preinvasive cervical neoplasia. Using tentative weighing factors

for the beneficial and unfavourable effects, the quality of life effects represent a net unfavourable effect that values about 20% of the life years gained (see Chapter 10).
Limited screening here means that no preventive smears that are additional to the official age range and interval schedule are taken, and that of all screened women only 5% will have border line smear results (requiring repeat smears). When screening is less limited, as in the early nineties, and under a wide range of plausible weighing factors, the beneficial health effects remain larger than the unfavourable ones, but the impact of the unfavourable effects on the net health effects may be considerable.

3 Under what conditions is cervical cancer screening cost-effective?

For screening women seven times between age 30 and 60 every 5 years in the Netherlands, the estimated costs per life-year gained are Dfl 28 000. For the optimal policy with 7 screenings per women, screening every 7 years from age 27 to 69, the estimated costs per life-year gained are Dfl 25 000. In a situation with limited screening (7 smears per women, no opportunistic screening, a limited amount of follow-up), the costs per quality of life adjusted life year (QALY) gained, tentatively accounting for the unfavourable side-effects, are up to 35% higher. These costs per life-year or per QALY gained are still within the limits of what generally is considered acceptable in the Dutch health care system. The incremental cost-effectiveness rapidely becomes more unfavourable with every next smear per woman added to the seven screens policy (Chapter 7), and with an increasing amount of opportunistic screening and more intensive follow-up of borderline smears (Chapters 3 and 10).

4 What screening age range and intervals should be chosen from the viewpoint of cost-effectiveness?

If seven smears are to be offered per woman, optimal screening (in terms of costs per life year gained) starts between 25 and 30 years, ends between 65 and 70 years, with an interval between subsequent screens of about seven years. If more than seven smears are offered, the interval can be shortened accordingly. For any number of screns it is very important to maintain a broad age-range, and adapt the screening interval accordingly (Chapter 7).

5 What follow-up strategy after non-negative smears should be chosen from the viewpoint of cost-effectiveness?

Since the introduction of screening three decades ago, there has been an important trend towards more follow-up after non-negative smears. The differences in costs and unfavourable side-effects between follow-up strategies can be estimated from long term and complete follow-up data from daily practice, but

more research still needs to be done. The differences in beneficial effects are as yet unknown because they are very difficult to estimate. Therefore, the question on optimal follow-up strategies is unresolved, and a priority for further research.

6 Should cervical cancer screening be continued? And what changes should be recommended in order to improve its efficiency?

In the Dutch situation, PAP smear cervical cancer screening could potentially prevent 50% of mortality from cervical cancer at reasonable costs and with acceptable unfavourable side effects. Sometimes however, physicians and attending women reason and act as if the aim of screening is to minimize the risk of interval cancers at all costs. This results in an inefficient and probably harmful practice of screening women under the target starting age, screening women more frequently than according to the target interval, and following up a high proportion of the screened women. Restriction of these practices is necessary. Furthermore, increasing the coverage of the target population would increase effectiveness and improve the balance between benefits, negative side effects, and costs (Chapter 2).

7 Should new techniques, like automated cytological evaluation of cervical smears or HPV-detection in cervical scrapes, be added to or replace conventional PAP smear screening?

Drawbacks of PAP smear screening are the high costs of screening all women e.g. seven times during their life-time, and the unfavourable side-effects caused by follow-up of false-positive test-results and by the detection of non- progressive conditions. The effectiveness in women who do attend regularly screening cannot be much improved. Thus, in order to be useful, new techniques should make cervical cancer screening less costly and/or improve its specificity for progressive pre-cancerous conditions. The potential effectiveness of automated and HPV-screening may be promising, but the costs and unfavourable effects in a routine screening situation have not been well established so far. Thus, although these techniques are interesting and further research is recommended, implementation in routine practice at this moment is not yet warranted.

A

Overview of differences between model versions

Because of newly available data and further analyses, the input parameters of the simulation model changed over the years. The '1988 version' of the model, described by Habbema et al in 1988 [90], was used in chapters 2, 3 and 5, the '1993 version', described by van Ballegooijen et al in 1993 [12] in chapter 7, the '1996a version' is described in Chapter 6, and '1996b' is used in chapter 9.

I will describe the differences (the main ones are summarized in *Table A.2*) and discuss their impact on the conclusions.

A.1 The 1993 version of the model compared to the 1988 version

For the 1993 analysis we had better age-specific invasive cervical cancer incidence data at our disposal, both for the recent period and for the pre-screening period. We used these incidence data in modelling the disease course between age-specific incidence of pre-invasive disease and age-specific mortality, where in the 1988 analysis only mortality and stage-specific survival data were used [90, 12]. Inclusion of these additional data led to three closely linked changes in model assumptions:

1 The stage-specific clinical survival in the 1988 version did not depend on age. But in the 1993 version it decreases from age 30 onwards. This was a direct consequence of fitting simultaneously age-specific incidence and age-specific mortality data.

2 The incidence of progressive dysplasia, which was fairly constant over ages in the 1988 version of the model, has a distinct peak in age-group 20 to 30 years in the 1993 version. This change was made to compensate for the influence of step 1. on mortality.

Table A.1 *Model versions and corresponding chapters*

Chapter	2, 3 and 5	6	7	9
Model version	1988	1996a	1993	1996b

3 For women under age 35, 83% of all new cases of CIN was assumed to be regressive in the 1988 model. This percentage was decreased to 72% in the 1993 model, to compensate for the influence of step 1 on detection rates.

In other words, in the 1993 model version, a higher proportion of the onset of CIN at young age was progressive, generating more invasive cancers at young age but with a higher survival. The effect was that screening in young age became relatively more effective in preventing incidence from cervical cancer, but not in preventing mortality from cervical cancer.

Moreover, in the 1993 version the risk in the women born after 1938 was increased with 10%, and the evaluation concerned 1993–2019 instead of 1988–2014, thus decreasing the impact of the older cohorts with a high relative risk.

For the optimal policies, the result was a slight decrease in the optimal age to start and stop screening.

A.2 The 1996a version of the model compared to the 1993 version

In the 1996a model used and described in chapter 6, the risk in women born after 1940 was estimated from the Dutch cervical cancer mortality in these birth cohorts only before screening started. In the 1993 version of the model, which was used for predictions in Chapter 7, we already calibrated the risk in women born after 1940 to the mortality from cervical cancer in recent years *(see Figure A.1)*. This resulted in a higher risk in the young cohorts in the 1993 model than in the 1996a model. The reason is that in 1993, we already anticipated to the conclusion from chapter 6 about the risk of the women born after 1940. When we will use the 1996a model for predictions to support health policy decisions, the risk in the youngest cohorts will again be calibrated to mortality from cervical cancer observed in recent years.

Furthermore, the refined analysis reported in chapter 6 resulted in a 5-year survival of clinically diagnosed cancers in middle ages (50–64 years) that is around 15% higher than in the 1993 version. This change is an adaption of the age-dependency of survival introduced in the 1993 model: survival in age-

Table A.2 *Main changes in successive versions of the model*

Model version	Main changes
1993	Clinical survival age-dependent
1996a	Cohort risks fitted to the pre-screening mortality only. Clinical survival fitted to pre-screening mortality/incidence ratio only
1996b	Incidence of regressive lesions increased to fit more recent data on detection rates of histologically confirmed pre-invasive disease

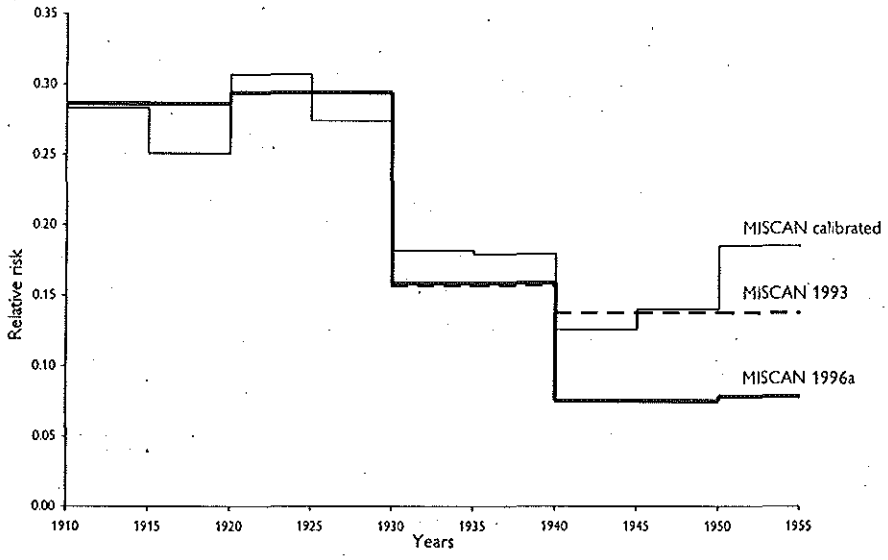
group 50+ was decreased in the 1993 analysis, and this decrease has been slightly reversed in 1996. We recognized that to test the model against incidence and mortality trends after screening started, it is appropriate to use the ratio between mortality and incidence before screening started to derive survival of clinically diagnosed cancers. In the 1993 analysis used for predictions into the future, we derived survival from recent mortality and incidence data. Consequently, screening in young age in the 1996a model will become slightly less cost-effective compared to the 1993 model.

A.3 The 1996b version of the model compared to the 1993 version

For exploring the (cost-)effectiveness of HPV-screening (model 1996b, Chapter 9), the 1993 model version was used, except for regressive CIN. We increased the incidence of regressive CIN up to a level where the detection rates (of progressive plus negative CIN) corresponded with higher observed levels in the early 1990's in the Netherlands.

In the 1988 and 1993 model versions, these assumptions were based on detection rates from the three pilot cervical cancer screening projects in the Netherlands between 1976 and 1984 [90], and from British Columbia from 1949–1969 [169]. The use of these data from a period with low referral rates explains why the predicted numbers of women referred and treated in the 1988 model (see Chapters 2 and 3) and in the 1993 model (results not shown) are relatively low. The consequence for the results presented in Chapters 2 and 3 is that — compared to the 1996b version— the unfavourable effects are underestimated for the present situation and the foreseenth future situation, and for the results presented in Chapter 7 it means that the costs for the management of CIN are underestimated.

Figure A.1 *The relative cervical cancer cohort risks assumed in the successive model versions* (see Table A.1) *used in this thesis, and the calibrated risk, using the 1996a model, to the mortality around 1990*



A.4 The predictions concerning negative side-effects in chapter 10 compared to the 1996b and previous model versions

The predicted numbers of women with repeat smears and women referred and treated for pre-invasive neoplasia as presented in Chapter 10.1.4 are not only based on recent detection rates, as in the 1996b version, but account for the influence of a further intensification of the follow-up practice to be expected from the revised follow-up guidelines that are currently being implemented as well. These referral and treatment rates are roughly two times higher than in the 1996b model and roughly sixfold higher than in the 1988 model (see Chapters 2 and 3) and in the 1993 model (results not shown).

B

Adjusted cost-effectiveness ratios (CERs) to compare studies

I will here describe how we accounted for differences between other cost-effectiveness studies and our MISCAN assumptions to calculate adjusted CERs (cost-effectiveness ratios) in chapter 10. The results of the adjustment was presented in Table 10.4. The CER in the MISCAN prediction for the Dutch cervical cancer screening situation is:

$$\frac{\text{Costs}}{\text{Effects}} = \frac{(A + B - C - D)}{E}$$

In which:

A = # smears × average costs per smear including costs for false positive cases

B = # detected CIN cases × costs per CIN case

C = # prevented invasive cervical cancer cases × the average costs per case

D = # prevented advanced cancer cases × the costs per advanced cancer case

E = # life years gained

This CER was adjusted for differences between the situation as described in the respective other studies considered in chapter 10 and the situation in the Netherlands as described by us. The differences accounted for are listed in Table 10.3 in chapter 10. The adjustment was performed as follows:

$$\text{AdjustedCER} = \frac{A \times a^{(1)} + B \times b^{(2)} * r - C \times c \times r \times e - D \times d \times r \times e}{E \times r \times e}$$

(1) including the costs for follow-up of false-positive screening test results, and in some studies the costs of detection of pre-invasive disease

(2) as far as not included in the costs per smear.

In which:

a, b, c, d are the respective costs used in the other studies relative to the corresponding costs assumed in MISCAN for the Dutch situation.
For example for Eddy's study $a = \frac{152}{76}$

r is the background risk for cervical cancer used in the other studies relative to the risk assumed in MISCAN for the Dutch situation.
For example for Eddy's study $r = 1.8$

e is the relative mortality reduction
For example for Eddy's study $e = \frac{88}{58}$ (see text on Eddy's study in Chapter 10)

Summary

In recent years in the Netherlands around 235 women have died annually from cervical cancer. The current mortality rate is less than 40% of the level around 1960. Mortality was already decreasing before it could have been affected by PAP smear screening, which started around 1970. The PAP smear can detect cervical cancer and its non-invasive precursors before symptoms occur. Because randomized cervical cancer screening trials have never taken place, a direct estimate of the contribution of screening to the decrease in mortality is not available.

The objective of this thesis is to evaluate cervical cancer screening by assessing its various effects and costs. Questions to be addressed concern the reduction in incidence and mortality, the unfavourable effects and the cost-effectiveness of different screening policies (age range and interval combinations). To this end, we used a cervical cancer version of MISCAN, a simulation model that was designed and programmed to evaluate cancer screening. This cervical cancer screening model had already been tested against screening data sets from British Columbia (Canada) and from Dutch pilot screening projects in 1976–1985. The potential impact of new developments in cervical cancer screening, such as automated cytological evaluation and HPV-testing, has been investigated. The ultimate questions are: should cervical cancer screening be continued, and what changes should be recommended in order to improve its efficiency?

The relationship between benefits and unfavourable effects was studied for three efficient cervical cancer screening strategies incorporating 5, 10 and 25 smears per woman (Chapter 2). It was concluded that when the screening frequency was increased, the ratio of favourable to unfavourable effects became worse.

Compared to the opportunistic screening in the Netherlands, efficient screening may yield the same mortality reduction with only 50% of the smears, the costs thus being also around 50% lower (Chapter 3). The reason is that opportunistic screening smears were taken at relatively young age and with shorter intervals and that the resulting coverage of the population at risk was lower.

Screening will lead to changes in medical practice concerning cervical cancer and its precursors, and accordingly in the costs of this practice. Positive screening results will produce diagnostic and therapeutic procedures for cervical intraepithelial neoplasia (CIN) and early invasive cancer. The management of CIN will on average cost Dfl 3500 per woman (Chapters 4 and 7). In half of the women, treatment by conisation or hysterectomy was performed. Effective screening will prevent death from cervical cancer and therefore also incur financial savings. The costs of medical care of advanced incurable cervical cancer were estimated at Dfl 30 000 per woman (Chapter 5).

A previous analysis of detection rates of pre-clinical disease and pre-screening incidence in British Columbia resulted in an estimated average duration of 16 years for pre-clinical detectable disease, and an estimated sensitivity of the PAP smear of 80%. These estimates are crucial in predicting the impact and cost-effectiveness of screening. We therefore checked them against trends in cervical cancer incidence and mortality in the period 1965–1992, which encompasses some 10 years before screening became widespread and could have had a marked impact. To this end, we imbedded the estimates for duration and sensitivity in the MISCAN population cervical cancer screening simulation model. This model was adapted to Dutch pre-screening cervical cancer incidence and mortality together with the screening pattern since its inception around 1970. The model-predicted incidence and mortality trends in the Netherlands after the start of screening were compared with observed incidence and mortality trends (Chapter 6). The similarity of observed and simulated trends supports the estimated natural history, sensitivity and cure rate parameters.

Using MISCAN, we predicted the life years gained and costs of screening for policies with 5 to 10 screenings per woman during her life, also varying the age-range and the screening interval. From an effectiveness point of view, a broad age-range is more important than a short interval. With seven screenings per women, efficient policies are those which screen women between 25–30 and 65–70 every 7 years. If less screenings per woman are planned, the broad age-range should be maintained and the interval must be increased accordingly. Given that in the Netherlands 5 years was considered the maximum acceptable length for the interval between successive screenings, screening women from age 30 to age 60 is viewed as a good policy. Furthermore, this policy will prevent opportunistic screening which will take place in young age groups when programme screening would start at later age. The expected mortality reduction from this policy (with an attendance rate of 75% and no opportunistic screening) is 50%, corresponding with around 200 fewer deaths from cervical cancer per year.

We investigated the potential usefulness of two recent developments, computer-aided cytological evaluation and HPV(human papillomavirus)-screening. At a Dutch cyto-pathological laboratory, a series of enriched (with positive cases)

smears initially screened conventionally was rescreened with a computer-aided method (PAPNET). A similar series initially screened with the computer aided method was rescreened conventionally (Chapter 8). Results showed a significantly lower number of non-negative smears (with borderline and higher abnormalities) in rescreening than in initial screening. In comparing conventional to computer aided screening, we corrected for the effects on the cytological evaluation that might have resulted from differences between the rescreening situation and the initial routine screening situation. After correction, the difference in the number of non-negative smears between conventional and computer aided screening was small and not significant. We concluded that rescreening of enriched (with positive smears) series of smears, as usually performed in the first instance to investigating a new PAP smear evaluation method, was of limited value for investigating the possibilities of these methods for routine screening because of a strong rescreen effect. Before a decision can be made about computer aided screening, further testing is needed in larger studies with a close to routine mixture of negative and positive smears.

We explored the cost-effectiveness of HPV-screening, using an extended version of the MISCAN cervical cancer disease model which is based on the hypothesis that HPV-infections found in women with CIN and invasive cervical cancer generally precede the detectable neoplasia (Chapter 9). Two different models were constructed, one extremely favourable and one extremely unfavourable for HPV-screening, but both in concordance with observed HPV-positive rates in women with and without cervical neoplasia. According to the favourable model, compared with 3-yearly PAP smear screening, 10-yearly HPV-screening was more effective in reducing mortality. In addition, costs were lower. According to the unfavourable model, 3-yearly HPV-screening was less effective than 3-yearly PAP smear screening and costs were higher. This is true both for HPV-testing alone and in combination with cytology. The crucial uncertainty concerns the natural history and the detectability of HPV-infections before neoplasia is present. Longitudinal data are required to decrease this uncertainty to such an extent that HPV-screening can either be rejected or recommended. Meanwhile it is possible that an effective, harmless and cheap HPV-vaccine will be developed in the future.

The effect measure used up to Chapter 10 in the estimated cost-effectiveness ratio's were life years gained. This means that effects on quality of life, and therefore the unfavourable effects, were not accounted for. In Chapter 10, we explored the magnitude and potential impact in cost-effectiveness evaluation of PAP smear screening of these unfavourable effects, that mainly involve women being advised to have repeat smears and being referred for colposcopy. Tentative weighing of the effects on quality of life as expressed in quality adjusted life years (QALY's) showed that the unfavourable effects in a very restricted screening situation are probably acceptable. But in the current screening sit-

uation, with a high frequency of smears and very intensive follow up after borderline screen results, the unfavourable effects may well have considerable negative impact on the total health effects and thus on the cost-effectiveness of cervical cancer screening.

In conclusion, the cost-effectiveness of a well-planned and organized cervical cancer mass-screening programme in the Netherlands is expected to be acceptable providing the following conditions are taken into account:

- The coverage of screening must be maintained or, even better, improved.
- Women and physicians must be discouraged to have preventive smears taken outside the target age range of 30 to 60 years and with an interval shorter than the targeted 5 years.

The following issues must also be taken under consideration:

- The guidelines for follow-up after borderline results require further evaluation because they might incur too intensive a follow-up.
- Computer aided cytology and HPV-screening are important developments but need further research before any decision can be made on their role in the screeningprocess.

Samenvatting

In de afgelopen jaren stierven er in Nederland rond de 235 vrouwen per jaar aan baarmoederhalskanker. Het huidige sterfte niveau is nog geen 40% van dat rond 1960. De sterfte was al aan het dalen voordat vroege opsporing van baarmoederhalskanker, die middels het nemen van uitstrijkjes bij symptoomloze vrouwen rond 1970 was begonnen, daar een rol in kon spelen. Men verwacht dat het maken van uitstrijkjes sterfte aan baarmoederhalskanker zal doen afnemen, omdat met het uitstrijkje baarmoederhalskanker en haar niet invasieve voorstadia ontdekt kan worden nog voor er symptomen optreden en voordat de ziekte onbehandelbaar wordt. Echter, omdat gerandomiseerde trials nooit hebben plaatsgevonden is een directe schatting van de bijdrage van screening op de daling in de mortaliteit niet mogelijk.

Het doel van dit proefschrift is vroege opsporing van baarmoederhalskanker te evalueren door de verschillende effecten en kosten te schatten. Het gaat om de reductie van incidentie en sterfte, de ongunstige effecten en de kosteneffectiviteit van verschillende uitnodigings schema's (combinaties van begin- en eindleeftijd en interval tussen opeenvolgende uitstrijkjes). Om deze te schatten gebruikten we een baarmoederhalskanker versie van MISCAN, een simulatiemodel dat is gemaakt om screening op kanker te evalueren. Dit baarmoederhalskanker model was al in overeenstemming gebracht met de screenings gegevens uit British Columbia (Canada) en uit de Nederlandse proefregio's in 1976–1985. Er is tevens onderzoek gedaan naar de mogelijkheden van nieuwe ontwikkelingen, zoals geautomatiseerde cytologische beoordeling van uitstrijkjes en het testen van schraapsels van de baarmoedermond op humaan papillomavirus (HPV). De uiteindelijke vragen zijn: moet vroege opsporing van baarmoederhalskanker worden voortgezet, en welke veranderingen moeten er komen om het efficiënter te maken.

De relatie tussen gunstige en ongunstige effecten is onderzocht voor drie efficiënte screenings schema's: één met 5, één met 10 en één met 25 uitnodigingen per vrouw (Hoofdstuk 2). De conclusie was dat bij een toenemend aantal uitstrijkjes, de ongunstige effecten meer toenemen dan de gunstige.

Vergeleken met opportunistische ('wilde') screening, kan efficiënte screening dezelfde sterfte reductie bereiken met half zo veel uitstrijkjes, waarbij ook de

kosten ongeveer 50% lager liggen (Hoofdstuk 3). De reden hiervoor is dat opportunistische uitstrijkjes gemiddeld op jongere leeftijd en met een korter screenings interval worden gemaakt en dat het bereik van de doel-leeftijdsgroep lager is.

Vroege opsporing zal veranderingen in de medische praktijk ten aanzien van baarmoederhalskanker tot gevolg hebben. Positieve uitstrijkjes zullen diagnostische en therapeutische ingrepen voor cervicale intraepitheliale neoplasie (CIN) en vroeg invasieve cervixcarcinomen induceren. De diagnostiek, behandeling en nazorg van CIN kost gemiddeld Dfl 3500 per geval (Hoofdstukken 4 en 7). In de helft van de gevallen wordt er een conisatie of een uterusextirpatie verricht. Aan de andere kant zal effectieve screening sterfte aan baarmoederhalskanker voorkomen, en daarmee besparingen met zich mee brengen. De medische kosten van vergevorderde baarmoederhalskanker werden op Dfl 30 000 per vrouw geschat (Hoofdstuk 5).

Een eerdere analyse van detectie cijfers van pre-klinische ziekte bij screening en de incidentie voor de aanvang van screening in British Columbia, leverde een schatting op voor de gemiddelde duur van pre-klinische detecteerbare ziekte van 16 jaar, en voor de sensitiviteit van het uitstrijkje van 80%. Deze schattingen zijn cruciaal voor de te verwachten resultaten en kosten-effectiviteit van cervix screening. Daarom hebben we ze getoetst aan de trend in incidentie en sterfte van baarmoederhalskanker in de periode 1965–1992, waarvan de eerste ca. 10 jaren nog nauwelijks door screening beïnvloed zijn. Daartoe werden de geschatte duur en sensitiviteit in het MISCAN populatie model voor baarmoederhalskanker screening ingebed. Dit model werd aangepast aan de Nederlandse incidentie van en sterfte aan baarmoederhalskanker voor aanvang van screening en aan het screenings patroon zoals het sinds 1970 bestond. De model voorspellingen voor de trends in incidentie en sterfte na aanvang van screening werden vergeleken met de waargenomen trends (Hoofdstuk 6). De overeenkomst tussen waargenomen en voorspelde trends ondersteunt de eerder gemaakte schattingen van de parameters die het natuurlijk beloop en de sensitiviteit beschrijven, en de veronderstelde uitstekende behandelbaarheid van pre-invasive stadia.

Eveneens gebruik makend van het MISCAN model, deden we voorspellingen voor de gewonnen levensjaren en de kosten van screening voor screenings schema's met 5 tot 10 uitstrijkjes per vrouw, waarbij de leeftijds range en het screenings interval werden gevarieerd. Voor de effectiviteit is een breed leeftijds bereik belangrijker dan een kort screenings interval. Met zeven uitstrijkjes per vrouw, is het het meest efficiënt om vrouwen tussen de 25 à 30 en 65 à 70 jaar te screenen om de zeven jaar. Als er met minder uitstrijkjes wordt gescreend, moet het brede leeftijds bereik worden gehandhaafd, en moet het interval overeenkomstig worden verlengd. Gegeven echter het feit dat in Nederland 5 jaar als de maximaal haalbare lengte voor het screenings inter-

val wordt beschouwd, is het screenen van vrouwen tussen de 30 en 60 jaar een goed beleid. De verwachte mortaliteits reductie van het screening van vrouwen tussen de 30 en 60 jaar om de vijf jaar met een deelname-percentage van 75% en zonder opportunistische screening, is 50%, hetgeen overeenkomt met ongeveer 200 voorkomen sterf-gevallen aan baarmoederhalskanker per jaar.

We hebben onderzoek gedaan naar het potentiële nut van twee recente ontwikkelingen, computer ondersteunde cytologische evaluatie van uitstrijkjes en screening op humaan papillomavirus (HPV). Door een Nederlands cytologisch laboratorium werden een serie uitstrijkjes die oorspronkelijk in de routine praktijk op de conventionele wijze was gescreend, met een computer ondersteunde methode (PAPNET) opnieuw gescreend. Tevens werd een serie uitstrijkjes die oorspronkelijk in de normale routine met PAPNET was gescreend op de conventionele wijze opnieuw gescreend. Beide series waren in gelijke mate verrijkt met positieve uitstrijkjes (Hoofdstuk 8). De resultaten laten zien dat bij het opnieuw screenen een significant kleiner aantal uitstrijkjes als niet-negatief (borderline of sterk afwijkend) werden geclassificeerd dan in de routine praktijk. Door de studie opzet kon conventionele screening met PAPNET screening vergeleken worden onafhankelijk van het gevonden rescreen effect. Tussen conventionele screening en screening met PAPNET waren de verschillen in aantallen niet-negatieve uitstrijkjes klein en niet significant. We concludeerden dat het opnieuw screenen van verrijkte series van uitstrijkjes, het eerste wat gedaan wordt om een nieuwe screenings methode te testen, door het optreden van sterke rescreen effecten van beperkte waarde is voor het schatten van de performance van zo'n nieuwe methode in routine screening. Om voor routine screening tot een beslissing te kunnen komen zijn grotere studies met een meer realistische mix van (veel) negatieve (weinig) positieve uitstrijkjes nodig.

Voor een exploratie van de kosten-effectiviteit van HPV-screening hebben we gebruik gemaakt van een versie van het MISCAN model waarin het ziekte model was uitgebreid met HPV-infecties, onder de hypothese dat HPV-infecties die worden gevonden in vrouwen met CIN en (invasief) cervixcarcinoom, aan de nieuwvorming zijn voorafgegaan (Hoofdstuk 9). We lieten zien dat het mogelijk is om twee verschillende modellen te construeren, één uitermate gunstig voor HPV-screening en één uitermate ongunstig, die beide in overeenstemming zijn met de gemeten HPV-positiviteit in de algemene populatie en in vrouwen met cervix neoplasie. Volgens het eerste model is 10-jaarlijkse HPV-screening effectiever in het voorkomen van sterfte aan baarmoederhalskanker dan 3-jaarlijkse cytologische screening, en bovendien goedkoper. Volgens het tweede model is 3-jaarlijkse HPV-screening minder effectief dan 3-jaarlijkse cytologie en duurder. De cruciale onzekerheid betreft het natuurlijk beloop van HPV-infecties voordat er neoplasie ontstaat en de detecteerbaarheid van HPV-infecties tijdens die periode. Om een uitspraak te kunnen doen over de (kosten-)effectiviteit van HPV-screening zijn longitudinale data nodig. Onder-

tussen is het niet uitgesloten dat er in de toekomst een effectief, ongevaarlijk en goedkoop vaccin tegen HPV-infecties komt.

De effectmaat die tot Hoofdstuk 10 is gebruikt voor het berekenen van kosten-effectiviteits ratio's is het aantal gewonnen levensjaren. Dat houdt in dat effecten op de kwaliteit van leven, die per saldo negatief zijn, niet in deze ratio tot uitdrukking komen. In Hoofdstuk 10 hebben we de mogelijke grootte onderzocht van de invloed wanneer wel wordt rekening gehouden met negatieve effecten. Het gaat bij die negatieve effecten met name om vrouwen die vanwege (licht) positieve uitstrijkjes wordt aangeraden op relatief korte termijn nogmaals een uitstrijkje te laten maken, en vrouwen die wegens een afwijkend uitstrijkje voor verder onderzoek worden verwezen naar de gynaecoloog. Op tentatieve wijze werden effecten op de kwaliteit van leven uitgedrukt in kwantiteit van leven, zodat voor kwaliteit gecorrigeerde gewonnen levensjaren (gewonnen QALY's) konden worden berekend. De exploratieve berekeningen lieten zien dat de ongunstige effecten van screening met een zeer beperkte intensiteit waarschijnlijk acceptabel zijn. Maar het werd ook duidelijk dat de ongunstige effecten van screening zoals die in afgelopen jaren werd toegepast, met relatief frequente uitstrijkjes en intensieve follow-up ook al bij zeer lichte afwijkingen, een aanzienlijke negatieve impact kunnen hebben op de netto gezondheidseffecten, die overblijven na aftrek van deze negatieve effecten.

Concluderend is te verwachten dat de kosten-effectiviteits verhouding van een goed gepland en georganiseerd bevolkingsonderzoek naar baarmoederhalskanker in Nederland acceptabel zal zijn, mits aan een aantal voorwaarden wordt voldaan:

- Het bereik van de screening moet op zijn minst gehandhaafd worden maar liever nog verhoogd worden
- Uitstrijkjes bij symptoomloze vrouwen buiten de gestelde leeftijds range van 30 tot 60 jaar en met een hogere frequentie dan om de vijf jaar moeten tot een minimum beperkt worden.

Tevens zijn de volgende zaken van belang:

- De geldende richtlijnen voor vervolgonderzoek na (licht) positieve uitstrijkjes moeten opnieuw geëvalueerd worden want induceren mogelijk een te intensieve follow-up praktijk
- Computer ondersteunde cytologie en HPV-screening zijn belangwekkende ontwikkelingen die zeker verder onderzoek behoeven. Echter, hun nut voor bevolkingsonderzoek op baarmoederhalskanker is nog niet aangetoond.

Acknowledgement

The evaluation of cervical cancer screening was financed by The Prevention Fund, the Dutch Ministry of Health, and the Public Health Insurance Council.

Chapter 2–5 and 8–9 have been reused with permission:

2 Ballegooijen M van, Koopmanschap MA, Oortmarssen GJ van, Habbema JDF, Lubbe JThN, Agt HMA van, Diagnostic and treatment procedures induced by cervical cancer screening → Eur J Cancer, 26, 941–945 (1990).
copyright ©1990 Pergamon Press Ltd

3 Ballegooijen M van, Habbema JDF, Oortmarssen GJ van, Koopmanschap MA, Lubbe JThN, Agt HHA van, Preventive Pap-smears: Balancing costs, risks and benefits → Br J Cancer, 65, 930–933 (1992)
copyright ©1997 Cancer Research Campaign

4 Ballegooijen M van, Koopmanschap MA, Tjokrobardojo AJS, Oortmarssen GJ van, Care and costs for advanced cervical cancer → Eur J Cancer, 28A, 1703–1708 (1992)
copyright ©1992 Pergamon Press Ltd

5 Ballegooijen M van, Koopmanschap MA, Habbema JDF The management of cervical intraepithelial neoplasia: extensiveness and costs → Eur J Cancer, 13A, 1672–1676 (1995)
copyright ©1992 Pergamon Press Ltd

8 Ballegooijen M van, Beck S, Boon ME, Boer R, Habbema JDF The rescreen effect in conventional and Papnet screening observed in a study using material enriched with positive smears
→ Acta Cytologica (in press)

9 Ballegooijen M van, Akker ME van den, Warmerdam PG, Meijer CJLM, Walboomers JMM, Habbema JDF Present evidence on the value of HPV testing for cervical cancer screening: a model-based exploration of the (cost-)effectiveness → Br J Cancer, 76(5) 651–657 (1997)
copyright ©1997 Cancer Research Campaign

Dankwoord

Dit proefschrift heeft alleen door toedoen van vele anderen tot stand kunnen komen.

Sommigen hebben een belangrijke inhoudelijke inbreng gehad. Dit geldt in hoge mate voor Gerrit van Oortmarssen, mijn co-promotor, en Rob Boer. Gerrit kan ik niet genoeg bedanken voor zijn inbreng, hulp en steun. Vele malen was hij het geheel met me oneens, en altijd op prettige wijze. De discussies met Rob, daartoe altijd bereid, gaven niet zelden stof tot nadenken. En natuurlijk was daar Dik Habbema, mijn promotor. Behalve voor het vele dat ik van hem heb geleerd dank ik hem voor de ruimte die ik kreeg om de dingen op mijn eigen wijze te doen, en voor het altijd opbouwende commentaar dat hij onvermoeibaar op mijn stukken leverde.

Het geleverde werk is het resultaat van langdurige multi-disciplinaire samenwerking. In de eerste jaren heeft Koos Lubbe me de wereld van de baarmoederhalskanker binnengeleid. Bij hem kon ik bovendien altijd mijn frustraties kwijt. Verder dank ik Harry de Koning en Mark Koopmanschap (de discussies met hem heb ik, nadat hij naar het iMTA was vertrokken, gemist). Heleen van Agt, en in de latere jaren Anita Bos, Elske van den Akker, Willem-Jan Meerding, Peter Warmerdam en Joke Ter Harmsel hebben als projectgenoten een direct aandeel gehad.

iMGZ bood het kader waarin dit alles kon plaatsvinden. De keuzes die destijds door Paul van der Maas en Dik Habbema zijn gemaakt ten aanzien van het modelleren van vroege opsporing van kanker staan aan de basis van mijn werk. En ik dank het aan hun leiding en aan de overige iMGZetters dat ik enerzijds zelfstandig onderzoek kon doen, anderzijds deel uitmaakte van een stimulerende wetenschappelijke gemeenschap.

Ilse Philips dank ik voor de blijmoedige volharding waarmee ze wijzigingen bleef invoeren, en voor allerlei andere zaken die ze voor me heeft gedaan (incluis het afstaan van boterhammen!). Daarnaast hebben Saskia Drent, Atie Slikkerveer, Tineke Kurtz, Daniëlle Romijn en Arry de Bruyn mij allerlei werk uit handen genomen.

Ook aan mensen buiten iMGZ ben ik veel dank verschuldigd. I thank Elsebeth Lynge [134] and Nereo Segnan for the many inspiring discussions we had. In de

beginjaren was Hans Wijnen een steunpunt. Dokter Subandono maakte het mij mogelijk statusonderzoek te doen zonder dat hierdoor de borstvoeding in gevaar kwam. Later waren Théo Helmerhorst, Chris Meijer, Jan Walboomers, Rence Rozendaal, en Mathé Burger geïnteresseerd, kritisch en behulpzaam. En niet te vergeten Mathilde Boon.

Als onderzoeker ben ik in de gelukkige omstandigheid geweest in nauw contact met de beleidsmakers te staan. In de eerste jaren was dat Wim van Veen die kritisch meekeek naar onze werkzaamheden. De discussies later met Bert Boer van de Ziekenfondsraad waren verhelderend. Ik heb veel gehad aan de gedachtenwisselingen met leden van de Coördinatie Commissie Bevolkingsonderzoek Baarmoederhalskanker en de werkgroep 'KIE', die ik via mijn lidmaatschap veelvuldig heb mogen hebben.

Het SIG/PALGA was van groot belang voor de recente gegevens. De open samenwerking met Renso Kamps maakte veel mogelijk.

Het Praeventiefonds, het Ministerie en de Ziekenfondsraad dank ik voor de financiering die de continuering van ons werk ten aanzien van de evaluatie van de vroege opsporing van baarmoederhalskanker steeds mogelijk maakte en nog steeds maakt.

De steun van het thuisfront was onmisbaar en hartverwarmend.
Ik ben Janna Marijn en Jan Joris meer dan dankbaar. En niet te vergeten Djuna. Jullie zijn echte lieverds.

Tenslotte was mij dit alles nooit gelukt zonder de steun, het vertrouwen en de aanmoedigingen van mijn ouders.

Curriculum vitae

Marjolein van Ballegooijen is in 1954 geboren in Etterbeek (België) en opgegroeid in Brussel. Haar schoolopleiding genoot ze aldaar aan de Europese School van Uccle. Na het 'bac' wiskunde–moderne talen (HBS-b) verhuisde zij in 1972 naar vader- en moederland Nederland om geneeskunde te studeren aan de (Gemeentelijke) Universiteit van Amsterdam. Na een 6-jarige onderbreking van haar studie geneeskunde, waarin ze ondermeer het propaedeuse economie aan de Universiteit van Amsterdam behaalde, deed ze in 1986 artsexamen. In 1987 werd ze aangesteld aan het Instituut Maatschappelijke Gezondheidszorg aan de Erasmus Universiteit Rotterdam, om met andere onderzoekers onder leiding van Dik Habbema beleidsondersteunend onderzoek te doen naar de effecten en kosten van vroege opsporing van baarmoederhalskanker. In 1995 leidde dit tot een universitair docentschap Medical Technology Assessment. Tegenwoordig houdt zij zich tevens bezig met evaluatie van screening op dikke darmkanker.

Ze is getrouwd en heeft twee kinderen, Janna Marijn en Jan Joris.

References

[1] Abel-Smith B, Figueras J, Holland W, McKee M, Mosrialos E (1995) Choices in health policy. An agenda for the European Union → Luxemburg, Office for official publications of the European Communities

[2] American College of Obstetricians and Gynecologists (ACOG) (1988) Report of Task Force on routine cancer screening → The American College of Obstetricians and Gynecologists: Washington DC

[3] Anderson GH, Boyes DA, Benedet JL, Le Riche JC, Matisic JP, Sven KC, Worth AJ, Millner A, Bennett OM (1988) Organization and results of the cervical cytology screening programme in British Columbia → Brit Med J, 296, 975–978

[4] Appelman CIM, Pal RS. (1986) Follow-up recomendations after Pap IIIa mildly and moderately displastic smears in cervical cancer mass-screening (in Dutch) → Ned Tijdschr Geneeskd, 130(41), 1844–1847

[5] Austoker J (1994) Screening for cervical cancer → Br Med J, 309, 241–248

[6] Baker MS, Kessler LG, Smucker RC (1989) Site-specific Treatment costs for cancer: an analysis of the Medicare continuous history sample file → In Scheffer RM, Andrews NC (eds): Cancer cost issues. Michigan, Health Adm Press Perspectives, Ann Arbor, 127–138

[7] Ballegooijen M van, Lubbe JThN, Oortmarssen GJ (1988) Cervical cancer epidemiology, natural histology and early detection → In: Habbema JDF et al. The costs and effects of mass-screening for cervical cancer (in Dutch). Rotterdam, Erasmus University Rotterdam

[8] Ballegooijen M van (1988) Diagnosis and treatment of cervical cancer → In Habbema JDF, Lubbe JThN, Agt HME van, Ballegooijen M van, Koopmanschap MA, Oortmarssen GJ van. Costs and effects of mass screening for cervical cancer (in Dutch). Rotterdam, Erasmus University Rotterdam

[9] Ballegooijen M van, Koopmanschap MA, Oortmarssen GJ van, Habbema JDF, Lubbe JThN, van Agt HMA (1990) Diagnostic and treatment procedures induced by cervical cancer screening → Eur J Cancer, 26, 941–945 (Chapter 2 of this thesis).

[10] Ballegooijen M Van, Habbema Jdf, Oortmarssen Gj Van, Koopmanschap Ma, Lubbe Jthn, Agt Hma Van (1992) → Preventive Pap-smears: balancing costs, risks and benefits Br J Cancer, 65, 930–933 (Chapter 3 of this thesis)

[11] Ballegooijen M van *et al.* (1992) Care and costs for advanced cervical cancer → Eur J Cancer, 28A, 1703–1708 (Chatper 5 of this thesis)

[12] Ballegooijen M van, Boer R, Oortmarssen GJ van, Koopmanschap MA, Lubbe JThN, Habbema JDF (1993) Cervical cancer screening: Age-ranges and intervals. An updated cost-effectiveness analysis (in Dutch) Rotterdam, Erasmus University Rotterdam →

[13] Ballegooijen M Van, Koopmanschap Ma, Habbema Jdf (1995) The management of cervical intraepithelial neoplasia: extensiveness and costs → Eur J Cancer, 13a, 1672–1676 (Chapter 4 of this thesis)

[14] Ballegooijen M van, Bos AB, Marle ME van, Oortmarssen GJ van, Boer R, Habbema JDF (1997) A first description of the screening practice in The Netherlands in 1994 based on data from PALGA (in Dutch) → Rotterdam, Erasmus University Rotterdam

[15] Bauer HM, Hildesheim A, Schiffman MH, Glass AG, Rush BB, Scott DR, Cadell DM, Kurman RJ, Manos MM (1993) Determinants of genital human papillomavirus infection in low-Risk women in Portland, Oregon → Sexually Transmitted Diseases, 20(5), 274–278

[16] Beral V, Hermon C, Muñoz N, Devesa SS (1994) Trends in cancer incidence and mortality → Cancer Surveys, 19/20, 265–285

[17] Bergeron C, Barrasso R, Beaudenon S, Flamant P, Croissant O, Orth G (1992) Human Papillomaviruses Associated With Cervical Intra-epithelial Neoplasia → Am J Surg Pathol, 16,641–649

[18] Berget A (1979) Influence of population screening on morbidity and mortality of cancer of the uterine cervix in Maribo Amt → Danish Med Bull, 26, 91–100

[19] Berrino F, Sant M, Verdecchia A, Capocaccia R, Hakulinen T, Estève J (editors) (1995) Survival of cancer patients in Europe. The Eurocare study → IARC Scientific Publications no. 132. Lyon, Evans AS,

[20] Bjorge T, Thoresen SO, Skare GB (1993) Incidence, survival and mortality in cervical cancer in Norway, 1956–1990 → Eur J Cancer, Vol.29A, N016. 2291–2297

[21] Bjørge T, Gunbjørud AB, Hangen OA, Skare GB, Trope C, Thoresen SØ (1995) Mass screening for cervical cancer in Norway: evaluation of the pilot project → Cancer Causes Control, 6, 477–484

[22] Boon ME, Graaff Guilloud JC de (1987) Dutch general practitioners have higher scores for positive cervical smears than programme screening (in Dutch) → Medisch Contact, 13, 407–408

[23] Boon ME, Kok LP (1993) Neural network processing can provide means to catch errors that slip through human screening of Pap-smears → Diagn Cytopath, 9, 411–416

[24] Bos AB, Ballegooijen M van, Oortmarssen GJ, Marle ME van, Habbema JDF, Lynge E (1997) Non-progression of cervical intraepithelial neoplasia estimated from population-screening data → Br J Cancer 75(1), 124–130

[25] Bosch MMC, Rietveld-Scheffers PEM, Boon ME (1992) Characteristics of false-negative smears in the normal screening situation → Acta Cytol, 711–716

[26] Bosch FX, Manos MM, Munoz N, Sherman M, Jansen AM, Peto J, Schiffman M, Moreno V, Kurman R, Shah KV (International Biological Study on Cervical Cancer (IBSCC) Study Group) (1995) Prevalence of human papillomavirus in cervical cancer: a world wide perspective → J Nat Cancer Inst, 87, 796–802

[27] Boyes DA, Morrison B, Knox EG, Draper GJ, Miller AB (1982) A cohort study of cervical cancer screening in British Columbia → Clin Invest Med, 5, 1–29

[28] Brookmeyer R and Day NE (1987) Two-stage models for the analyses of cancer screening data → Biometrics 43, 657–669

[29] Bruin A de, Koning HJ de, Ballegooijen M van (1993) Pap smears and mammographies, Dutch National Health Interview Surveys 1991 → Monthly Bulletin of Health Statistics, 12–5

[30] Brule Ajc Van Den, Walboomers Jmm, Maine M Du, Kenemans P, Meijer Cjlm (1991) Difference in prevalence of human papillomavirus genotypes in cytomorphological normal cervical smears is associated with a history of cervical intraepithelial neoplasia → Int J Cancer, 48, 404–408

[31] Buckley CH, Herbert A, Johnson J, et al. (1994) Borderline nuclear changes in cervical smears: guidelines on their recognition and management. National Coordinating Network (National Cervical Screening Programme), British Society for Clin-

ical Cytology, and Royal College of Pathologists' Working Party. → J Clin Pathol, 47, 481–492

[32] Campion MJ, Brown JR, McCance DJ, Atia W, Eduards R, Cuzick J, Singer A (1988) Psychosexual trauma of an abnormal cervical smear → Br J Obstet Gynaecol, 95, 175–181

[33] Caterall RD, Nicol CS, ed (1976) Sexually transmitted diseases → London, Academic Press

[34] Netherlands Central Bureau of Statistics (CBS) (1980) Atlas of cancer mortality in The Netherlands 1968–1979 → Voorburg, CBS

[35] Netherlands Central Bureau of Statistics (CBS) (1987) Death by cause of death, age and sex → Series B1, Voorburg, CBS

[36] Netherlands Central Bureau of Statistics (CBS) (1989) Death by cause of death, age and sex → Series A1, Voorburg, CBS

[37] Netherlands Central Bureau of Statistics (CBS) (1994) Death by cause of death, age and sex. Series A1, 1950–1992 → Voorburg, CBS

[38] Netherlands Central Bureau of Statistics (CBS) (1995) Death by cause of death, age and sex. Series A1, 1950–1993 → Voorburg, CBS

[39] Central Council for Home help (CCH) (1988) Data concerning volume and costs of home help in june 1988 (in Dutch) → Information bulletin nr. 28, Driebergen-Rijsenburg

[40] Netherlands Central Cancer Registration (CCR) (1973) Incidence of carcinoma of the cervix uteri in Friesland, Rotterdam and The Hague, 1960–1970 → Dutch Cancer Society

[41] Choi NW, Nelson NA (1986) Results from a cervical cancer screening programme in Manitoba, Canada → In: Hakama, M., Miller, A.B., Day, N.E. (eds). Screening for cancer of the uterine cervix, pp. 61–67. Lyon, International Agency for Research on Cancer

[42] Chua KL and Hjerpe A (1996) Persistence of human papillomavirus (HPV) infections preceding cervical carcinoma → Cancer, 77, 121–127

[43] Clarke EA, Anderson TW (1979) Does screening by 'Pap' smears help prevent cervical cancer? A case-control study → Lancet, 2(8132), 1–4

[44] Clayton D, Schifflers E (1987) Models for temporal variation in cancer rates. I and II: Age-period and cohort models → Stat Med, 6, 449–481

[45] Coleman D, Day N, Douglas G, Farmery E, Lynge E, Philip J, Segnan N. (Europe Against Cancer) (1993) European guidelines for quality assurance in cervical cancer screening Eur J → Cancer, 29A (suppl. 4), S1-S38

[46] Collette HJA, Linthorst G, de Waard F (1974) Report of the activities of 'Cyt-U-Universitair' in early detection of cervical cancer (1970–1973) → Utrecht, Laboratorium Cyt-U-Universitair

[47] Cook GA, Draper GJ (1984) Trends in cervical cancer and carcinoma in situ in Great Britain. → Br J Cancer, 50, 367–375

[48] Coutinho RA (1980) The increase in the number of patients with STD's in Amsterdam → Ned T Geneesk, 124(8), 284–286

[49] Cox B, Skegg DCG (1986) Trends in cervical cancer in New Zealand → NZ Med J, 99,795–798

[50] Cancer Registry in Norway (CRN) (1980) Survival of cancer patients. Cases diagnosed in Norway 1968–1975 → Oslo, CRN

[51] Canadian Task Force (1982) Cervical cancer screening programs. Summary of the 1982 Canadian Task Force report → Can Med Ass J, 127, 581

[52] Cuzick J, Terry G, Ho L, Hollingworth T, Anderson M (1994) Type-specific human papillomavirus dna in abnormal smears as a predictor of high-grade cervical intraepithelial neoplasia → Br J Cancer, 69(1), 167–171

[53] Cuzick J, Szarewski A, Terry G, Ho L, Hanby A, Maddox P, Anderson M, Kocjan G, Steele ST, Guillebaud J (1995) Human papillomavirus testing in primary cervical cancer screening → Lancet, 345, 1533–1536

[54] Damhuis RAM (editor) (1994) Gynaecologic oncology. Treatment and survival, 1984–1991 (in Dutch) → Rotterdam, Integraal Kankercentrum Rotterdam

[55] D'Arcy C, Holman J, Armstrong BK (1987) Cervical cancer mortality trends in Australia - an update → Med J Aust, 146, 410–412

[56] Day NE (1984) Effect of cervical cancer screening in Scandinavia → Obstet Gynecol, 63, 714–718

[57] Day NE (1986) The epidemiological basis for evaluating different screening policies. → In: Hakama M, Miller AB, Day NE, eds, Screening for cancer of the uterine cervix, pp. 199–209. Lyon, International Agency for Research on Cancer

[58] Day NE (1986) Screening for squamous cervical cancer: duration of low risk after negative results of cervical cytology and its implication for screening policies. IARC working group on evaluation of cervical cancer screening programmes → Br Med J, 293, 659–664

[59] Dutch Health Insurance Council (1993) Starting points restructuring cervical cancer screening programme (in Dutch) → Decision Dutch Health Insurance Council, 592

[60] Dobson AJ (1983) An introduction to statistical modelling → London, Chapman and Hall

[61] Department of Health, Housing and Community Services (1991) Screening for the prevention of cervical cancer → Camberra, AGPS, 3

[62] Dræby-Kristiansen J, Garsaae M, Bruun M, Hansen K (1991) Ten years after cryosurgical treatment of cervical intraepithelial neoplasia → Am J Obstet Gynecol, 165(1), 43–45

[63] Duffy BJ (1976) Papanicolao testing — Are we screening the wrong women? → New Engl J Med, 294(4), 223

[64] Eddy DM (1980) Screening for cancer: theory, analysis and design → Englewood Cliffs, Prentice Hall

[65] Eddy DM (1990) Screening for cervical cancer → Ann Intern Med 113(3), 214–226

[66] Eluf-neto J, Booth M, Muñoz N, Bosch Fx, Meijer Cjlm, Walboomers Jmm (1994) Human Papillomavirus And Invasive Cervical Cancer In Brazil → Br J Cancer, 69(1), 114–119

[67] Evaluation Committee (EVAC) (1980) Population screening for cervical cancer in the pilot regons Nijmegen, Rotterdam and Utrecht. A report by the Evaluation Committee → First and second interim report (in Dutch). Leidschendam, Ministry of Welfare, Public Health and Cultural Affairs

[68] Evaluation Committee (EVAC) (1984) Report from the Evaluation Committee concerning the early detection of cervical cancer. About the second round of the screening programme in the pilot regions Nijmegen, Rotterdam and Utrecht → (in Dutch) Leidschendam, Ministry of Welfare, Public Health and Cultural Affairs

[69] Evaluation Committee Cervical Cancer Screening (EVAC) (1986) Registration of cervical cancer in the pilot regions Nijmegen, Rotterdam and Utrecht (in Dutch) → Rijswijk, Government Department of Public Health

[70] Evaluation Committee Cervical Cancer Screening (EVAC) (1988) Final Report of the Evaluationcommission concerning the early detection of cervical cancer (in Dutch) → Rijswijk, Department of Public Health

[71] Evaluation Committee Cervical Cancer Screening (EVAC) (1989) Population screening for cervical cancer in The Netherlands. A report by the Evaluation Committee → Int J Epidemiol, 18(4), 775–781

[72] Evans AS, Monaghan JM (1983) The treatment of cervical intraepithelial neoplasia using carbon dioxide laser → Br J Obstet Gynaecol, 90, 553–556

[73] Feltkamp MC, Smits HL, Vierboom MP (1993) Vaccination with cytotoxic Tlymphocyte epitope containing peptide protects against a tumour induced by human papillomavirus type 16 – transformed cells → Eur J Immunol, 23, 2242–2249

[74] Flannelly G, Anderson D, Kitchener HC, Mann EM, Campbell M, Fisher P, Walker F, Templeton AA (1994) Management of women with mild and moderate dyskaryosis → Br Med J, 308, 1399–1403

[75] Fletcher A (1990) Screening for cancer of the cervix in elderly women → Lancet, 335, 97–99

[76] Gaarenstroom Kn, Melkert P, Walboomers Mm, Brule Ajc Van Den, Bommel Pfj Van, Meyer Cjlm, Voorhorst Fj, Kenemans P, Helmerhorst Thjm (1994) Human Papillomavirus Dna And Genotypes: Prognostic Factors For Progression Of Cervical Intraepithelial Neoplasia → Int J Gynaecol Cancer, 4, 73–78

[77] Gessel-Dabekousen AAMW van, Ballegooijen M van (1994) Participation in mass screening for cervical cancer (in Dutch) → Mndber Gezondheid (CBS) 12

[78] Giles JA, Walker PG, Chalk PAE (1987) Treatment of cervical intraepithelial neoplasia (CIN) by radical electrocoagulation diathermy: 5 years experience → Br J Obstet Gynaecol, 94, 1089–1093

[79] Goodwin JS, Hunt CH, Key CK, Samet JM (1990) Changes in surgical treatment: the example of hysterctomy versus conization for cervial carcinoma in situ → J Clin Epidemiol, 43, 977–982

[80] Graaf Y van der (1987) Screening for cervical cancer. The Nijmegen project → Dissertation, Nijmegen

[81] Graaf Y van der, Zielhuis GA, Peer PGM, Vooijs GP (1988) The effectiveness of cervical cancer screening. A population based case-control study → J Clin Epidemiol, 41(1), 21–26

[82] Green GH (1979) Rising cervical cancer mortality in young New Zealand women → New Zealand Med J, 89, 89–91

[83] Gustaffson L, Adami HO (1989) Natural History of cervical neoplasia: consistent results obtained by an identification technique → Br J Cancer, 60, 132–141

[84] Gustafsson L, Adami HO (1992) Optimization of cervical cancer screening → Cancer Causes Control 3, 125–136

[85] Gustafsson L, Sparen P, Gustafsson M, Wilander E, Bergström R, Adami HO (1995) Efficiency of organised and opportunistic cytological screening for cancer in situ of the cervix → Br J Cancer, 72, 498–505

[86] Gyrd-Hansen D, Hølund B, Andersen P (1995) A cost-effectiveness analysis of cervical cancer screening: health policy implications → Health Policy, 34, 35–51

[87] Habbema JDF, Oortmarssen GJ van, Lubbe JThN, Maas PJ van der (1984) The MISCAN simulation program for the evaluation of screening for disease → Comput Methods Programs in Biomed, 20, 79–93

[88] Habbema JDF, van Oortmarssen GJ, Lubbe JThN, van der Maas PJ (1985) Model building on the basis of Dutch cervical cancer screening data → Maturitas, 5, 11–20

[89] Habbema JDF, Lubbe JThN, van Oortmarssen GJ, van der Maas PJ (1987) A simulation approach to cost-effectiveness and cost-benefit calculations of screening for the early detection of diseases → Eur J Oper Res, 29, 159–166

[90] Habbema JDF, Lubbe JThN, Agt HME van, Ballegooijen M van, Koopmanschap MA, Oortmarssen GJ van (1988) Costs and effects of mass screening on cervical cancer (in Dutch) → Rotterdam, Erasmus University Rotterdam

[91] Haes H de, Koning HJ de, Oortmarssen GJ van, Agt HME van, Bruyn AE de, Maas PJ van (1991) The impact of a breast cancer screening programme on quality-adjusted life-years → Int J Cancer 49, 538–544

[92] Hakama M (1985) Effect of population screening for carcinoma of the uterine cervix in Finland → Maturitas, 7, 3–10

[93] Hakama M (1993) To screen or not to screen for cervical cancer → Eur J Cancer, 29A(16), 2218–2220

[94] Hakulinen T, Hakama M (1985) The effect of screening on the incidence and mortality of cervical cancer in Finland → Nowotwory, 35, 290–296

[95] Ho Gyf, Burk Rd, Klein S, Kadish As, Chang Cj, Palan P, Basu J, Tachezy R, Lewis R, Romney S (1995) Persistant genital human papillomavirus infection as a risk factor for persistent cervical dysplasia → J Nat Cancer Inst, 87, 1365–1371

[96] Hristova L (1995) Effect of screening for cancer on mortality, costs and quality of life in Finland → University of Tampere, Dissertation

[97] Hristova L, Hakama M (1997) Effect of screening for cancer in the Nordic countries on deaths, cost and quality of life up to the year 2017 → Acta Oncol 36, suppl. 9, 1–60

[98] Humanitas (home help agency) (1990) Data concerning home help for women having advanced cancer, 1987–1989 → Amsterdam

[99] Husain OAN, Butler EB, Nayagam M, Mango L, Alonzo A (1994) An analysis of the variation of human interpretation, Papnet? — a mini-challenge → Analytical Cellular Pathology, 6, 157–163

[100] International Agency for Research on Cancer (IARC) Working group on evaluation of cervical cancer screening programmes (1986) Screening for squamous cervical cancer: duration of low risk after negative results of cervical cytology and its implication for screening policies → Br Med J, 293, 659–664

[101] International Agency for Research on Cancer (IARC) Working Group (1995) IARC Monographs On The Evaluation Of Carcinogenic Risks To Humans. Human Papilloma Viruses → Lyon, IARC Scientific Publication no. 64, IARC

[102] Nandakumar A, Anantha N, Venugopal TC (1995) Incidence, Mortality and survival in cancer of the cervix in Bangalore, India → Br J Cancer, 71(6), 1348–1352

[103] Jenkins D, Sherlaw-johnson C, Gallivan S (1996) Can Papilloma Virus Testing Be Used To Improve Cervical Cancer Screening? → Int J Cancer, 65, 768–773

[104] Jensen OM, Estève J, Møller H, Renard H (1990) Cancer in the European Community and its Member States → Eur J Cancer, 26, 1167–1256

[105] Johnson DB, Rowlands CJ (1990) Diagnosis and treatment of cervical intraepithelial neoplasia in general practice → Br Med J, 299, 1083–1086

[106] Jordan JA, Woodman CBJ, Mylotte MJ, Emens JM, Williams DR, Macalary M (1985) The treatment of cervical intraepithelial neoplasia by laser vaporization → Br J Obstet Gynaecol, 92, 394–398

[107] Ketting BW (1981) Surgical treatment of invasive carcinoma of the uterine cervix → Dissertation. University of Amsterdam

[108] Kharazi M, Keyhani-Rofagha S, O'Toole R (1993) Automated screening of cervical cytology using PAPNET: Comparative cytological findings with biopsy → Am J Clin Path, 99, 337–338

[109] Kjaer SK, Brule AJC Van Den, Bock JE, Poll PA, Engholm G, Sherman Me, Walboomers Jmm, Meijer Cjlm (1996) Human Papillomavirus - The Most Significant Risk Determinant Of Cervical Intraepithelial Neoplasia → Int J Cancer, 65, 601–606

[110] Kjellgren O (1986) Mass screening in Sweden for cancer of the uterine cervix: effect on incidence and mortality → Gynecol Obstet Invest, 22, 57–63

[111] Knapen M, Mensink J; Ramakers C (1988) Home care experiments, final report period september 1987–may 1988 → (in Dutch), Institute for applied social sciences. Nijmegen

[112] Knox EG (1976) Ages and frequencies for cervical cancer screening → Br J Cancer, 34, 444

[113] Kok MR, Boon ME (1996) Consequences of neural network technology for cervical screening: increase in diagnostic consistency and positive scores → Cancer, 78, 112–117

[114] Kolstad P, Klem V (1976) Long-Term Follow-up of 1121 cases of carcinoma in situ → Obstet Gynaecol, 48(2), 125–129

[115] Koning HJ de, Ineveld BM van, Oortmarssen GJ van *et al.* (1990) Costs and effects of breast cancer screening (in Dutch). Final report. → Rotterdam, Erasmus University Rotterdam

[116] Koning HJ de, Ineveld BM van, Oortmarssen GJ van, Haes JCJM de, Collette HJA, Hendriks JHCL, Maas PJ van der (1991) Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the posible impact of uncertain factors → Int J Cancer 49, 531–537

[117] Koning HJ de (1993) The effects and costs of breast cancer screening → Dissertaton, Erasmus University Rotterdam

[118] Koopmanschap MA, Lubbe JThN, Oortmarssen GJ van, Agt HME van, Ballegooijen M van, Habbema JDF (1990) Economic aspects of cervical cancer screening → Soc Sci Med, 30, 1081–1087.

[119] Koopmanschap M, Oortmarssen Gj Van, Agt Hma Van, Ballegooijen M Van, Habbema Jdf, Lubbe Jthn (1990b) Cervical-cancer Screening: Attendance And Cost-effectiveness → Int J Cancer, 45, 410–415

[120] Koss LG, Phillips AJ (1974) Summary and recommendations of the workshop on uterine-cervical cancer → Cancer, 33(Suppl.), 1753–1754

[121] Koss LG (1978) Dysplasia: A real concept or a misnomer? → Obstetrics and Gynaecology, 51(3), 374–379

[122] Koss LG, Lin E, Schreiber K, Elgert P, Mango L (1994) Evaluation of the PAP-NET cytologic screening system for quality control of cervical smears → Am J Clin Pathol, 101, 220–229

[123] KPMG Klijnveld Management Consultants (1991) Study of the bottlenecks of the cervical cancer mass-screening programme (in Dutch) → Rapport, KPMG

[124] Kurman RJ, Henson DE, Herbst AL, Noller KL, Schiffman MH for the 1992 National Cancer Institute Workshop (1994) Interim guidelines for management of abnormal cervical cytology → JAMA, 271(23), 1866–1869

[125] Kwikkel HJ (1985) Diagnosis and Treatment of Cervical Intraepithelial Neoplasia (CIN) → Dissertation. Vrije Universiteit of Amsterdam

[126] Läärä E, Day NE, Hakama M (1987) Trends in mortality from cervical cancer in the Nordic countries: Association with organized screening programmes → Lancet, i, 1247–1249

[127] Lapré F, van den Bos G (1987) Evening- and night care by district nursing in Amsterdam (in Dutch). → Amsterdam, District nursing Amsterdam

[128] La Vecchia C, Decarli A, Gallus G (1987) Epidemiological data on cervical carcinoma relevant to cytopathology → Appl Pathol, 5, 25–32

[129] Loizzi P, Mor-Yosef S, Pearson S, Ireland D, Monaghan JM (1992) Rational use of cryosurgerey and cold knife conization for treatment of cervical intraepithelial neoplasia → Eur J Gynaecol Oncol, 13(6), 507–513

[130] Lopez A, Carbirriero C, Gesn'A di, Resta L, Nappi R (1990) Is routine colposcopic assessment necessary following laser ablation of cervical intraepithelial neoplasia? → Br J Obstet Gynaec, 97, 175–177

[131] Lorincz At, Reid R, Jenson Ab, Greenberg Md, Lancaster Wd, Kurman Rj (1992) Human Papillomavirus Infection Of The Cervix: Relative Risk Associations Of Fifteen Common Anogenital Types → Obstet Gynecol, 79,328-3-37

[132] Luce BR (1981) The implications of cost-effectiveness analysis of medical technology. An application to cervical cancer screening → USA Washington, Office Technology Assessment (OTA) congress

[133] Lynge E, Arffmann E, Hansen KC, Henriksen B, Holten IW, Holund B, Jakobson AB, Jespersen NB, Olesen F et al (1994) Status of preventive screening for uterine cervix cancer in Denmark (in Danish) → Ugesk Laeger 156(4), 471–473

[134] Lynge E (1998) Personal communications → Copenhagen

[134] Magnus K, Langmark F, Andersen A (1987) Mass screening for cervical cancer in Ostfold country of Norway 1959-77 → Int J Cancer, 39, 311–316

[135] Mählek CG, Jonsson H, Lenner P (1994) Pap smears screening and changes in cervical cancer mortality in Sweden → Int J Gynecol Obstet, 44, 267–272

[136] Malberger E (1992) Determination of the suitability if the PAPNET automated cytologic screening system to rapidly implement wide-scale screening for cervical cancer in a population that is currently largely unscreened → Analyt Quant Cytol Histol, 14, 272

[137] Mango LJ, Holund B, Holm K, Koss LG, Schreiber K, Alonzo A, Decker D (1993) Rescreening of 89 'negative' cervical smears preceding invasive carcinoma by the PAPNET cytological screening system → Acta Cytol, 37, 780

[138] Marteau TM, Walker P, Giles J (1990) Anxieties in women undergoing colposcopy Br J → Obstet Gynaecol, 97, 859–861

[139] McNeil C (1995) Getting a handle on ASCUS: a new clinical trial could show how → J Nat Cancer Inst, 87, 788–791

[140] Meijer CJLM (1995) → Personal communications (in Dutch)

[141] Melkert Pwj, Hopman E, Brule Ajc van den, Risse EK, Dienst PJ van, Bleker OP, Helmerhorst T, Schipper ME, Meijer CJLM, Walboomers JM (1993) Prevalence of HPV in cytomorphological normal cervical scrapes, as determined by the polymerase chain reaction, is age-dependent → Int J Cancer, 53, 919–923

[142] Miller AB, Lindsay J, Hill GB (1976) Mortality from cancer of the uterus in Canada and its relationship to screening for cancer of the cervix → Int J Cancer, 17, 602–612

[143] Miller AB (1985) Screening for cancer of the cervix → In: Miller, A.B. (ed) Screening for cancer, pp. 295-323. New York: Academic Press

[144] Miltenburg Th, Zwan H van der, Hoorn R van (1989) Prices of home care (in Dutch) → Nijmegen, Institute for applied social sciences

[145] Miltenburg Th, Mensink J, Tunissen C (1989) Data on care for cancer patients in the home care experiments may-december 1988 (in Dutch). → Nijmegen, Institute for applied social sciences

[146] Morrison Eab, Ho Gyf, Vermund Sh, Goldberg Gl, Kadish As, Kelley Kf, Burk Rd (1991) Human Papillomavirus Infection And Other Risk Factors For Cervical Neoplasia: A Case-control Study → Int J Cancer, 49(1), 6–13

[147] Morrison AS (1992) Screening in chronic disease. Second edition → Oxford, Oxford University Press

[148] Muller C, Mandelblatt J, Schechter CB, Power EJ, Duffy BM, Wagner JL (1990) Costs and effectiveness of cervical cancer screening in elderly women → Background paper U.S. Congress, Office of Technology Assessment, OTA-BP-65. Washington DC, U.S. Government Printing Office

[149] Muñoz N, Bosch FX, Sanjosè S de, Tafur L, Izarzugaza I, Gili M, Viladiu P, Navarro C, Martos C, Ascunce N, Gonzalez LC, Kaldor JM, Guerrero E, Lörincz A, Santamaria M, Alfonso de Ruiz P, Aristizabal N, Shah K (1992) The causal link between human papillomavirus and invasive cervical cancer: a population-based case-control study in Colombia and Spain → Int J Cancer, 52(5), 743–749

[150] Murray M, Mc Millan C (1993) Social and behavioural predictors of women's cancer screening practices in Northern Ireland → J Publ Health Med 15(2), 147–53

[151] Netherlands National Association of General Practitioners (NAG) (1990) → Personal communication

[152] USA National Cancer Institute (1988) Working guidelines for early cancer detection → International Cancer News, 11

[153] Netherlands Cancer Registry (NCR) (1992) Incidence of cancer in the Netherlands, 1989 → Utrecht, SIG Health Care Information

[154] Netherlands Cancer Registry (NCR) (1993) Incidence of cancer in the Netherlands, 1990 → Utrecht, SIG Health Care Information

[155] Netherlands Cancer Registry (NCR) (1994) Incidence of cancer in the Netherlands, 1991 → Utrecht, SIG Health Care Information

[156] Netherlands Cancer Registry (NCR) (1996) Incidence of cancer in the Netherlands, 1993 → Utrecht, SIG Health Care Information

[157] Netherlands National District Nursing Association, NDNA (1989) Data about district nursing for women with cancer, sampleproject 1986–1988 → Bunnik

[158] NHS cervical screening programme (1996) Cervical screening. A pocket guide → NHS cervical screening programme, United Kingdom

[159] Oortmarssen GJ van, Habbema JDF (1991) Epidemiologic evidence for age-dependent regression of pre-invasive cervical cancer → Br J Cancer, 64, 559–565

[160] Oortmarssen GJ van, Habbema JDF, Ballegooijen M van (1992) Predicting mortality from cervical cancer after negative smear test results → Br Med J, 305, 449–451

[161] Oortmarssen GJ van, Habbema JDF (1995) The duration of pre-clinical cervical cancer and the reduction in incidence of invasive cancer following negative Pap smears → Int J Epidemiol, 24, 300–307

[162] Oortmarssen GJ (1995) Evaluation of mass screening for cancer. A model-based approach → Dissertation, Erasmus University Rotterdam

[163] PALGA – Dutch Network And National Database For Pathology (1992) Results of retrieval action on cervical cytology from 1987–1990 → Utrecht, SIG/PALGA

[164] PALGA – Dutch Network And National Database For Pathology (1995) Results of retrieval action on cervical cytology from 1977–1994 → Utrecht, SIG/PALGA

[165] Parkin DM, Leach K, Cobb P, Clayden AD (1982) Cervical cytology screening in two Yorkshire areas: results of testing → London, Public Health, 96, 3–14

[166] Parkin DM, Nguyen-Dink X, Day NE (1985) The impact of screening on the incidence of cervical cancer in England and Wales → Br J Obstet Gynaecol, 92, 150–157

[167] Parkin DM, Moss SM (1986) An evaluation of screening policies for cervical cancer in England and Wales using a computer simulation model → J Epidemiol Comm Health, 40, 143–153 .

[168] Parkin Dm, Pisani P, Ferlay J (1993) Estimates Of The Worldwide Incidence Of Eighteen Major Cancers In 1985 → Int J Cancer, 54, 954–606

[169] Pedersen E, Hoeg K, Kolsted P. (1971) Mass screening for cancer for the uterine cervix in Ostfold Country, Norway: an experiment. Second report of the Norwegian Cancer Society. → Acta Obstet Gynecol Scand Suppl, Suppl 11, 1–18

[170] Pill CF, Letchworth AT, Noble AD (1984) Effect of introduction of colposcopy into a district general hospital → Postgrad Med J, 60, 461–463

[171] Pisani P, Parkin DM, Ferlay J (1993) Estimates of the worldwide mortality from eigtheen major cancers in 1985. Implications for prevention and projections of future burden → Int J Cancer, 55, 891–903

[172] Pontèn J, Adami HO, Bergström R, Dillner J, Friberg LG, Gustafsson L, Miller AB, Parkin DM, Sparèn P, Trichopoulos D. (1995) Strategies for global control of cervical cancer. → Int J Cancer, 60, 1–26

[173] Reaffle AE, Alden B, Mackenzie EFD (1995) Detection rates for abnormal cervical smears: what are we screening for? → Lancet, 345, 1469–1473

[174] Remmink Aj, Walboomers Jmm, Helmerhorst Tjm, Voorhorst Fj, Rozendaal L, Risse Ekj, Meijer Cjlm, Kenemans P (1995) The Presence Of Persistent High Risk HPV Genotypes In Dysplastic Cervical Lesions Is Associated With Progressive Disease: Natural History Up To 36 Months → Int J Cancer, 61, 306–311

[175] Richart RM, Townsend DE, Crisp W, DePetrello A (1980) An Analysis of long-term follow-up results in patients with cervical intraepithelial neoplasia treated by cryotherapy → Am J Obstet Gynecol, 137, 823–826

[176] Riley G, Lubitz J, Prihoda R, Rabey E (1987) The use and costs of Medicare services by cause of death → Inquiry, 24, 233–244

[177] Rolnick S, LaFerla JJ, Wehrle D, Trygstad E, Okagaki T (1996) Pap smear screening in a health maintenance organization: 1986–1990 → Prev Med 25(2), 156–161

[178] Rombach JJ (1984) Cyt-U-Universitair: Anouncements concerning early detection of cervical cancer → Utrecht, Laboratorium Cyt-U-Universitair

[179] Rosenthal DL, Mango LJ, Acosta DA, Peters RD (1993) 'Negative' Pap smears preceding carcinoma of the cervix: Rescreening with the PAPNET system → Am J Clin Pathol, 100, 331

[180] Rozendaal L, Walboomers JMM, Linden JC van der, Voorhorst FJ, Kenemans P, Helmerhorst THJM, Ballegooijen M van, Meijer CJLM (1996) The PCR-based high risk HPV test in cervical cancer screening gives an objective risk assessment of women with cytomorphological normal cervical smears → Int J Cancer, 68, 766–769

[181] Rutten F, Haan G (1990) Cost-effective use of medical technology: regulatory instruments and economic incentives → In Jönsson B, Rutten F and Vang J (eds): Policy making in health care. Linköping, Linköping Collaborating Centre

[182] Ryan MR, Stasny JF, Remmers R, Pedigo MA, Cahill LA, Frable WJ (1996) PAPNET-directed rescreening of cervicobaginal smears. A study of 101 cases of atypical squamous cells of undetermined significance → Am J Clin Pathol, 105, 711–718

[183] Sanjosè S De, Muñoz N, Bosch Fx, Reimann K, Pedersen Ns, Orfila J, Ascune N, Gonzalez Lc, Tafur L, Gili M, Lette I, Viladiu P, Tormo Mj, Moreo P, Shah K, Wahren B (1994) Sexually Transmitted Agents And Cervical Neoplasia In Colombia And Spain → Int J Cancer, 56(3), 358–363

[184] Sherlaw-Johnson C, Gallivan S, Jenkins D, Jones MH (1994) Cytological screening and management of abnormalities in prevention of cervical cancer: an overview with stochastic modelling → J Clin Pathol 47, 430–435

[185] Sherman ME, Mango LJ, Kelly D et al (1993) PAPNET review of smears preceding the diagnosis of a high grade squamous intraepithelial lesion or carcinoma → Acta Cytol, 37, 781

[186] SIG – Information centre for Health Care (1985) Hospital Diagnose Statistics 1963–1985 (in Dutch) → Utrecht, SIG

[187] Information Centre for Health Care (SIG) (1989) LMR data on hospital admissions, results of retrieval action on dysplasia, carcinoma in situ and cancer of the cervix uteri: 1985–1988 → Utrecht, SIG

[188] SIG – Information Centre for Health Care (1989) LMR data on hospital admissions, 1985–1988 → Utrecht, SIG

[189] SIG – Information Center for Health Care (1989) SIVIS data on nursing home admissions, 1986–1988 → Utrecht, SIG

[190] SIG – Information Centre for Health Care (1991) LMR data on hospital adherence to the dutch population → Personal communication. Utrecht, SIG

[191] SIG – Information Centre for Health Care (1992) LMR data on hospital admissions, 1975–1990 → Utrecht, SIG

[192] Simons HJ (1991) Letter of the Ministry of Welfare, Public Health and Cultural Affairs to the chairman of the Dutch Health Insurance Council → PEP/GZ 001027, from April 18, 1991

[193] Singer A, Walker P, Tay SK, Dyson J (1984) Impact of introduction of colposcopy to a district general hospital → Br Med J, 289, 1049–1051

[194] Singer A (1995) Cervical cancer screening: state of the art → Baillière's Clin Obstet Gynaecol, 9(1), 39–64

[195] Sirgurdsson K, Hrafukelsson J, Geirsson G, Gudmundsson J, Salvarsolottiv A (1991) Screening as a prognostic factor in cervical cancer: analysis of survival and prognostic factors based on Icelandic population data, 1964–1988 → Gynaecol Oncol 43, 64–70

[196] Slagel DD, Zaleski S, Mango LJ, Cohen MB (1993) Evaluation of PAPNET automated cytological screening system for gynecologic smears → Modern Path, 6, 32a

[197] Smid T (1983) Cervical cancer in The Netherlands in the period 1969–1980, morbidity and mortality in the population at risk → Wageningen, University of Agriculture (Department of Health)

[198] Souter WP,Barros Lopes A de, Fletcher A, Monaghan JM, Duncan ID, Parakevaidis E, Kitchener HC (1997) Invasive cervical cancer after conservative therapy for cervical intraepithelial neoplasia → Lancet 349, 878–80

[199] Sparèn P, Gustafsson L, Friberg LG, Pontèn J, Bergström R, Adami HO. (1995) Improved control of invasive cervical cancer in Sweden over six decades by earlier clinical detection and beter treatment. → J Clin Oncol, 13,715–725

[200] Sundhedsstyrelsen (1996) Screening for cervical cancer (in Danish) → Copenhagen, Sundhedsstyrelsen

[201] Suzich JA, Ghim S, Palmer-Hill FJ, White WI, Tamura JK, Bell JA et al (1995) Systematic immunization with papillomavirus L1 protein completely prevents the development of viral mucosal papillomas → Proceedings of the National Academy of Sciences, USA, 92, 11553–11557

[202] Timmer PR (1982) The treatment of cervical cancer in the Academic Hospital of Groningen from 1970–1979 → Dissertation. University of Groningen

[203] US Preventive Services Task Force (1996) Guide to Clinical Preventive Services 2nd ed → Baltimore, Williams and Williams

[204] Dutch Association of Sickness Funds (VNZ) (1990) → Personal communication

[205] Vooijs GP (1987) Advices after abnormal findings in cytology of the cervix uteri (in Dutch) → Ned Tijdschr Geneeskd, 131, 1662–1663

[206] Vuong PN, Vacher-Lavenu MC, Marsan C, Baviera E (1995) Computer-assisted rescreening of Papanicolaou-stained cervical/vaginal smears. Evaluation of the PAPNET system based on a series of 225 cases. (In French) → Arch Anat Cytol Path, 43, 147–153

[207] Walboomers JMM, Roda Husman AM de, Snijders PJF, Stel HV, Risse EKJ, Helmerhorst TJM, Voorhorst FJ, Meijer CJLM (1995) Human papillomavirus in false negative archival cervical smears: implications for screening for cervical cancer → J Clin Pathol 1995, 48(8), 728–732

[208] Waugh N and Robertson A (1996) Costs and benefits of cervical cancer screening. II. Is it worthwhile reducing the screening interval from 5 to 3 years → Cytopathol 7, 241–248

[209] Waugh N, Smith I, Robertson A, Reid GS, Halkerston R, Grants A (1996) Costs and benefits of cervical cancer screening. I. The costs of the cervical screening programme → Cytopathol 7, 231–240

[210] Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russel LB; for the Panel on Cost-Effectiveness in Health and Medicine (1996) Recommendations of the panel on cost-effectiveness in health and medicine → JAMA, 276(15), 1253–1258

[211] Wetchler SJ (1984) Treatment of cervical intraepithelial neoplasia with CO2 laser: laser versus cryotherapy → A review of effectiveness and cost. Obstet Gynaecol Survey, 38, 469

[212] Wilkinson C, Jones MJ, McBride J (1990) Anxiety caused by abnormal results of cervical smear test: a controlled trial → Br Med J, 300, 440

[213] Wright K (1987) The economics of informal care of the elderly → Centre for health economics, discussion paper 23, York, University of York

[214] Zur Hausen H (1994) Molecular pathogenesis of cancer of the cervix and its causation by specific human papillomavirus types → In: Zur Hausen H (Ed) Human Pathogenic Papillomaviruses. Berlin, Springer-verlag