Kirsten Rozemeijer

The Effects of New Screening Tests in the Dutch Cervical Cancer Screening Programme

ISBN: 978-94-6169-780-6

Cover illustration, lay-out and print: Optima Grafische Communicatie, Rotterdam, The Netherlands

Most studies reported in this thesis were funded by the Dutch National Institute for Public Health and the Environment.

This thesis is partly realised due to the financial support of the Department of Public Health Erasmus MC Rotterdam and the Erasmus University Rotterdam.

The Effects of New Screening Tests in the Dutch Cervical Cancer Screening Programme

De effecten van nieuwe screeningstesten in het
Nederlandse Bevolkingsonderzoek naar Baarmoederhalskanker

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

donderdag 21 januari 2016 om 13.30 uur

door

Kirsten Rozemeijer
geboren te Rijnsburg

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotor:
Prof.dr. H.J. de Koning

Overige leden:
Prof.dr. F.J. van Kemenade
Prof.dr. P.J.E. Bindels
Prof.dr. L.F.A.G. Massuger

Copromotor:
Dr. I.M.C.M. de Kok

CONTENTS

CHAPTER 1

General introduction

EPIDEMIOLOGY OF CERVICAL CANCER

Worldwide, cervical cancer is the fourth most common cancer in women. It was estimated that 528,000 women were diagnosed with cervical cancer in 2012 while 266,000 women died as a result of this disease. Incidence and mortality rates vary widely between developed and less developed regions. In 2012, 85% percent of the new cervical cancer cases occurred within the less developed regions, just as 87% of the deaths arising from cervical cancer. Mortality rates ranged from less than 2 per 100,000 women in Western Asia, Western Europe, and Australia/New Zealand, to almost 28 in Eastern Africa(1). The disease mainly affects young women, with a peak incidence at ages 40 to 55 years in unscreened women(2).

In the Netherlands, the cervical cancer incidence and mortality have decreased since 1990 and 1970, although an increase in incidence was observed between 2004 and 2012 (Figure 1-1). In 2013, the incidence and mortality rate were 5.9 and 1.4 per 100,000 women-years, both standardized to the world's age distribution(3).

A prerequisite for the development of cervical cancer is infection with a high-risk type of the human papillomavirus (HPV)(4, 5). There are more than 100 HPV types of which approximately 40 are able to infect the epithelium of the genital tract(6, 7). Approximately 15 of these HPV types are oncogenic and thus considered as high-risk HPV types(7, 8). As high-risk HPV is sexually transmitted, the risk of infection is strongly related with sexual behaviour. It increases with an earlier sexual debut and an increasing number of sexual partners, both for the woman and her male partner(9-11). The prevalence is age-dependent and reaches a maximum before the age of 30(7, 11-14) where it can be as high as 27.4%(15). However, as most high-risk HPV infections are transient, only 15 to 30% of them will progress to cervical intraepithelial neoplasia (CIN) (16-18). CIN lesions are considered as asymptomatic prestadia of cervical cancer and they are ranked to the severity of the lesion (CIN I, II, or III). Only a small part of these precursor stadia (approximately 1, 5, and 12% of CIN I, CIN II, and CIN III lesions,

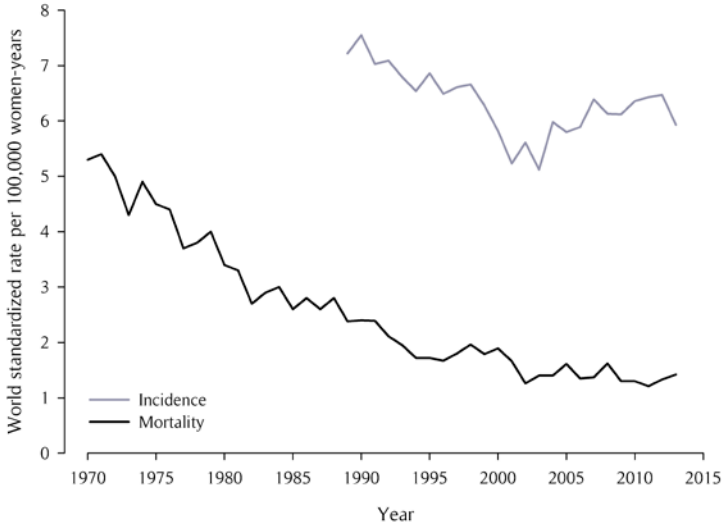

Figure 1-1. Cervical cancer incidence and mortality in the Netherlands, standardized to the world's age distribution(3).

respectively(19)) will progress to cervical cancer in the absence of screening and treatment(20). Co-factors which are associated with an increased risk of cervical cancer are smoking, long-term oral contraceptive use, p53 polymorphisms, and infections with other sexually transmitted diseases(9).

CERVICAL CANCER SCREENING

The preclinical detectable phase of cervical cancer (ie. CIN I until CIN III) is estimated to last on average more than ten years(21-23), which makes it ideal for screening. In developed countries, cervical abnormalities are detected by using a Pap smear. Cervix uteri exfoliated cells are collected by a nurse, general practitioner, or gynaecologist and examined by a cytologist

to identify the presence of cellular abnormalities. According to the severity of these abnormalities, women will either be invited for a triage test a couple of months later or they will be referred directly to a gynaecologist for a colposcopy. For the use of different cytological classification systems, and the correspondence between them and histological outcomes of biopsies taken at the gynaecologist, see Table 1-1.

A CIN II^{+} lesion can be treated by local destructive therapy or by complete excision of the transformation zone, of which the latter is preferred according to European guidelines(24). Treatment diminishes the risk of cervical cancer with 95%(25), although it increases the risk of preterm delivery, extremely low birth weight infants and perinatal mortality(26). As the majority of the lesions would not have progressed to cervical cancer(19, 20, 22), it is important to find a balance between prevention of cervical cancer on the one hand and overdiagnosis and overtreatment on the other. Since the percentage of CIN I lesions progressing to cervical cancer is extremely low, Dutch guidelines advice to treat them only in exceptional cases. For instance, when the CIN I lesion is persistent or if adequate follow-up cannot be guaranteed(27).

Table 1-1. Agreement between different cytological classification systems and histological outcomes(28, 29). In the Netherlands, the Pap classification was used before 1996. Thereafter, the CISOE-A classification was used. Another classification system, which is internationally often used, is the Bethesda classification(30).

Histological outcomes	Cytological outcomes according to Pap	Cytological outcomes according to CISOE-A	Cytological outcomes according to Bethesda
Normal	1	S1 O1 E1/E2	Negative
CIN I	2/3a1	S2/S3/S4 O3/O4 E3/4	ASC-US/ LSIL / ASC-H*
CIN II	3a2	S5 O5 E5	HSIL
CIN III	3b/4	S6 O6/O7 E6/E7	HSIL
Carcinoma	Carcinoma	S8/S9 O7/O8 E9	Carcinoma

*ASC-H has no official equivalent at the Pap and CISOE-A classification systems. Usually, it is classified in the same group as ASC-US and LSIL.

History of screening in the Netherlands

Cervical cancer screening exists in the Netherlands since the 1970s. At that time no centrally organized programme was in place, but screening was implemented and organized per region. Since the mid-80s decentralized screening was offered to women aged between 35 and 53 years with a 3-year interval. As coverage was low and false-positivity rate was high(31-33), screening was reorganized in 1996: the age range was extended to ages 30 to 60, the interval was extended from three to five years, all non-attendees received a reminder, financing and managing was centrally organized, opportunistic screening smears were discouraged and no longer reimbursed(34), and the description of cervical smears by Pap terminology was replaced by using the more detailed CISOE-classification(35) (Table 1-1). Moreover, the definition of a negative smear was extended. Sole inflammation of the epithelium was no longer defined as a Pap 2 but as a Pap 1, which increased the specificity while keeping the sensitivity at a similar level(36). In addition, the triage scheme was adjusted. Women with a primary Pap 2/3a1 were invited for cytology triage six and 18 months later and women with a ≥Pap3a2 were referred directly to the gynaecologist. In 2002, the definition of a negative primary smear was again altered. The absence of endocervical cells was no longer a reason for a woman to be referred to triage(34).

This reorganization and subsequent changes mainly resulted in a decreased percentage of women sent to follow-up, an increased compliance with follow-up, and a decreased average time in follow-up(34). Overall, it has proven that the introduction of cervical cancer screening has effectively reduced the incidence and mortality of cervical cancer(2, 37).

Liquid-based cytology testing and HPV triage

Approximately since 2005, the use of cytology triage testing has gradually been replaced by the use of cytology combined with HPV triage testing, which is accepted by the Dutch national guidelines for Pathology(38). With cytology triage testing alone, women are tested both at six and 18 months, independent of the result at six months. When HPV triage testing at six

months is added, women are referred back to the routine screening programme if no cytological abnormalities are found and no high-risk HPV virus is detected (Figure 1-2). This resulted in an increased specificity as fewer women underwent secondary repeat testing. In addition, women will no longer be referred to the gynaecologist in case the cytology triage test at six months is classified as a Pap 2/3a1, and no HPV virus is detected, and the triage cytology at 18 months is classified as Pap 1 (Figure 1-2). Despite this change in protocol, no decrease in the number of women being referred to the gynaecologist was observed. Moreover, an unexpected effect of the addition of HPV triage testing was an increase in the rate of CIN I and CIN II lesions detected via triage. An explanation for this phenomenon could be that prior knowledge of the HPV result affected cytological classifications(39).

Liquid-based cytology (LBC) has been developed as an alternative for the use of the conventional cytology Pap test as it has certain advantages, such as: ease of processing, reduction of unsatisfactory slides, reduction of time needed to read the slides and the ability of co-testing for the HPV virus(40-48). Conventional cytology and LBC share the same method of sampling cells from the cervix (ie. scraping off cells with a brush or similar device from the histological transition zone). LBC differs from conventional cytology with respect to the transfer of cells from brush to slide. With conventional cytology, cells are directly smeared on a slide, while with LBC, the brush is first rinsed into a vial with a preservative fluid and then transported to the laboratory(49). In the laboratory, a uniform layer of cells is prepared onto the slide(50, 51). The rationale behind this is that this method of cell transfer results in a better representation of the entire sample as compared to conventional cytology where a selective proportion of the sample is transferred to the slide(52). In addition, this uniform monolayer is easier to interpret for a cytologist as compared to conventional cytology smears. The latter can be obscured by blood or inflammation, have a bad cell fixation and/or an inhomogeneous distribution of cells(49). Earlier studies have shown that the sensitivity to detect CIN II^{+} lesions is similar between con-

Table 1-2. Preparation and protocol differences between LBC tests SurePath and ThinPrep(53, 54).

Protocol concerning:	SurePath	ThinPrep
Preservative liquid	Ethanol based	Methanol based
Collecting device	The cervical specimen is collected using a broom like device with detachable head. The detachable head is placed in a vial with preservative fluid. Then, the vial is transported to the laboratory	The cervical specimen is collected using a Cervix Brush and the brush is rinsed in a vial with preservative fluid. Cells are released by pushing the brush to the bottom, forcing the bristles apart and swirling the brush into the fluid. Subsequently, the brush is discarded and the vial is transported to the laboratory.
Method of cell transfer	At the laboratory, the brush is discarded followed by centrifugation of the fluid and cells to isolate the cells from the fluid. The cells are resuspended in a sucrose density gradient followed by slide transfer using gravity for adherence. A manual imprint system is used to prepare slides in advance.	At the laboratory, cells are isolated from the fluid by vacuum filtration and are transferred to the slide using air pressure for adherence.

ventional cytology and LBC testing(49, 52, 55-60). Because of this supposed similarity in CIN II$^+$ sensitivity and the above-mentioned advantages of LBC over conventional cytology, the Dutch society for Pathology permitted the use of LBC tests SurePath and ThinPrep(38). As a result, the majority of Dutch laboratories processing primary screening tests switched from using conventional cytology to either SurePath or ThinPrep since 2000. However, differences in techniques are apparent between these types of LBC tests (Table 1-2), while the effects of these differences on their performances were unknown at the time.

Recently, some of the laboratories processing primary LBC screening tests have implemented automated reading of these tests. This means that the automated reader assists the cytologist in examining the slide by pointing

out areas which need a closer look (ie. were cell abnormalities are likely to be present). The effect of automated reading on CIN II^{+} sensitivity of LBC tests is currently unknown as study results are heterogeneous(61-64).

Altogether, these innovations have resulted in a mixture of using three different types of primary cytology tests (ie. conventional cytology, SurePath or ThinPrep) and a mixture of using two different triage schemes (ie. cytology or cytology combined with HPV triage testing) within the current Dutch cervical cancer screening programme (Figure 1-2). In addition, examination of LBC tests SurePath and ThinPrep can be done with or without the assistance of automated reading.

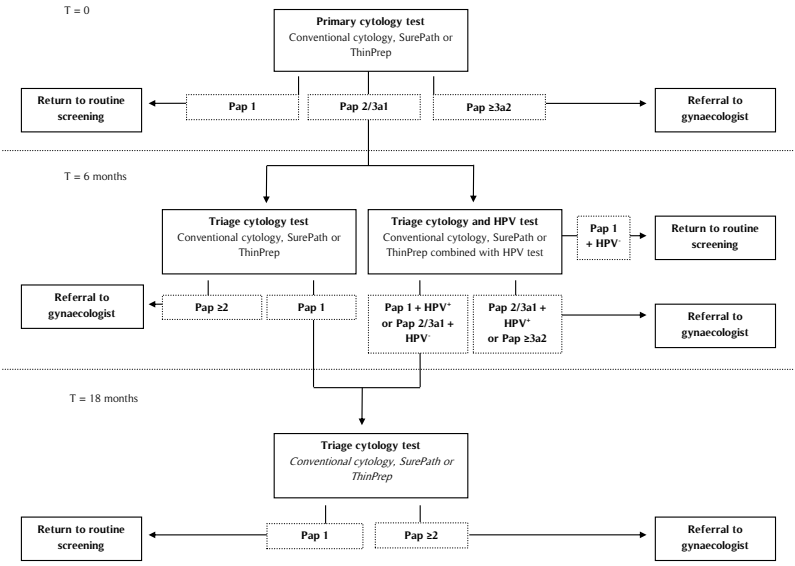

Figure 1-2. Screening scheme as currently used in the Netherlands. The examination of LBC tests SurePath and ThinPrep can be done with or without the assistance of automated reading. See Table 1-1 for the agreement between the Pap and Bethesda classification system.

Primary HPV and self-testing

The causality between HPV and cervical cancer makes primary HPV screening a good alternative to primary cytology screening. Multiple studies have shown that the sensitivity to detect CIN II$^+$ or CIN III$^+$ lesions is higher when using primary HPV as compared to primary cytology testing(65, 66). Furthermore, as the negative predictive value of a primary HPV test is higher than that of a primary cytology test, the interval between screening rounds could be extended when replacing primary cytology by primary HPV screening without affecting its effectiveness (65, 66). It is expected that from 2016 onward, primary high-risk HPV testing will be implemented in the Dutch cervical cancer screening programme and the screening interval will be extended from five to ten years starting from the age of 40(67, 68). Consequently, women will be invited for screening at ages 30, 35, 40, 50, and 60 years. If they do not attend screening or had a positive HPV test in the previous screening round, the ten-year interval will be lowered to five years. Triage will consist of immediate cytology co-testing and cytology testing six months later. If cell abnormalities are found (ie. ≥Pap 2) besides the HPV virus, the woman will be referred to the gynaecologist (Figure 1-3).

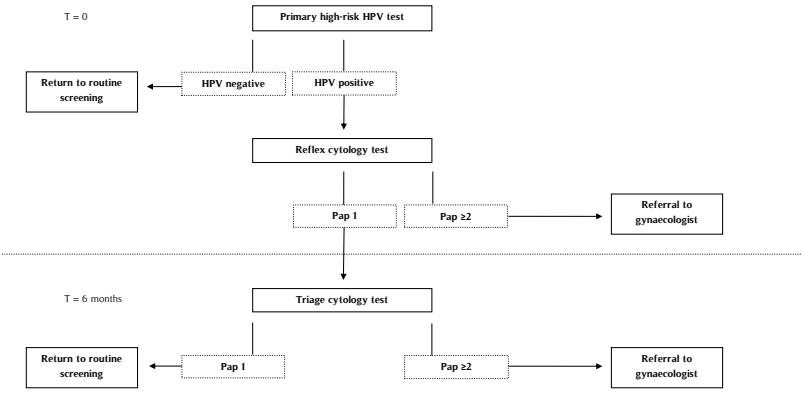

Figure 1-3. Future screening scheme which is expected to be used from 2016 onward.

The possibility of self-sampling is another important advantage of HPV testing over cytology testing(69-71). This is important as self-sampling may increase the attendance when offered to non-attendees(72, 73). As the majority of the cervical cancers in the Netherlands are estimated to occur in non-attendees(74), it might be a promising method to decrease the cervical cancer incidence and mortality. Therefore, the Dutch Ministry of Health has decided that HPV self-sampling will be added to the primary HPV cervical cancer screening programme by offering it to non-attendees(67, 68).

HPV vaccination

Since 2009, HPV vaccination has been implemented in the Dutch National Immunisation Programme. Twelve year old girls are vaccinated with the bivalent Cervarix vaccine against high-risk HPV types 16 and 18(75), which are responsible for more than 70% of the cervical cancer cases in Europe and North America(8). In addition, a catch-up campaign was organized for thirteen to sixteen year old girls born between 1993 and 1996)(75). Thus, the first vaccinated cohort will reach the screening age of 30 in 2023. As the coverage is approximately 60%(76) and vaccinated women are still at risk for cervical cancer caused by HPV types other than 16 and 18, it is expected that cervical cancer screening will still be necessary. The composition of this programme (eg. separate screening policies for vaccinated and unvaccinated women) is currently under investigation(77).

MONITORING OF THE DUTCH CERVICAL CANCER SCREENING PROGRAMME

Since 1996, monitoring and evaluation is part of the cervical cancer screening programme. In that way obstacles which endanger the effectiveness of the programme can be identified as early as possible, using short-term (eg. attendance rates, the number of women referred to triage, and CIN detection rates) and long-term outcomes (eg. interval cancer rates: defined as a cervical cancer diagnosed after a negative primary screening smear). Sub-

sequently, interventions can take place in order to ensure the effectiveness of the programme. This is of importance as the screening strategies (and screening tests) are based on microsimulations and assumptions which may differ from real-life settings. Moreover, long-term outcomes such as interval cancers are too rare to be examined in most clinical trials, while they can be assessed using population-based data.

The national monitoring and evaluation has been performed by the Department of Public Health of the Erasmus MC in Rotterdam. It comprises of publishing periodic reports (eg. in 2009, 2011, and 2013(78-80)) and cost-effectiveness analyses examining the effects and costs of the programme after the reorganisation of 1996(81).

The nationwide registry of histo- and cytopathology (PALGA)

This database contains information on all cytological and histological examinations of the cervix uteri taken in the Netherlands, covering all Dutch pathology laboratories. Every woman has her own identification code, while date, reasoning (ie. screening or not), and results are also registered. This makes it possible to follow individual screening histories. Therefore, for most of our analyses described in this thesis, data were retrieved from the PALGA database. As these data contain millions of records and dozens of variables, SQL queries were designed to select the data needed for the analyses we were interested in.

AIM OF THE THESIS

The aim of this thesis is to give an overview of the effects of using new screening tests in the Dutch cervical cancer screening programme. In the first part (questions 1-4, corresponding with Chapters 2-5), we focus on the effects of using LBC tests SurePath and ThinPrep in the current programme. In the second part (questions 5-7, corresponding with Chapters 6-8), we

focus on the effects of non-attendance and offering HPV self-sampling to non-attendees in the future primary HPV programme.

1. What is the effect of using SurePath and ThinPrep as primary test method on CIN II$^+$ detection rates?

Although multiple studies compared the ability to detect CIN II$^+$ between ThinPrep and conventional cytology(45, 49, 52, 57, 59, 60), only two have compared the ability between SurePath and conventional cytology(42, 58). In addition, none have compared the performance of both SurePath and ThinPrep with conventional cytology, while differences in protocol are present. Therefore, we examined the effect of using LBC tests SurePath and ThinPrep versus conventional cytology as primary test method within the current Dutch cervical cancer screening programme. All primary smears taken within this programme from 2000 to 2011 were analyzed using PALGA. We performed logistic regression analyses to examine whether CIN and cervical cancer detection rates differed between the types of cytological tests, adjusted for confounding factors.

2. Are increasing CIN detection rates caused by implementation of LBC tests?

During the last decade, CIN I, II, and III detection rates have increased in the Dutch screening programme (source: PALGA). As LBC testing was recently implemented in the programme, we quantified the increase in CIN detection rates and assessed whether the increase was still present when adjusting for differences in types of cytology tests used over time.

3. Is there a difference in sensitivity to detect progressive CIN lesions between SurePath, ThinPrep, and conventional cytology?

As in the absence of screening (and associated treatment) only a fraction of CIN lesions would progress to cervical cancer, detecting an equal or increased CIN rate is not necessarily equivalent to preventing equally or increased numbers of cervical cancers. To assess whether

the ability to detect progressive CIN lesions differs between SurePath, ThinPrep, and conventional cytology, we compared interval cancer rates between the three types of cytology tests by performing Cox regression analyses, adjusted for confounding factors.

4. What is the amount of overdiagnosis in the Dutch cervical cancer screening programme?

Cervical cancer screening is aimed at finding and treating CIN lesions in order to prevent the development of cervical cancer, which has resulted in a decrease of the cervical cancer incidence rate. Therefore, cervical cancer overdiagnosis is not really an issue within cervical cancer screening, although the detection of CIN lesions did increase and it is estimated that only a few of these lesions would have progressed to cervical cancer without treatment(19). In addition, the detection of a CIN lesion is associated with a decreased quality of life because of the psychosocial burden(82), and because treatment results in increased risks on adverse pregnancy outcomes(26). Therefore, we calculated the amount of overdiagnosis in cervical cancer upon inclusion of pre-invasive lesions using the MISCAN-Cervix model.

5. How many cervical cancer cases in young women can potentially be prevented using a more sensitive screening test at age 30?

The CIN II$^+$ sensitivity of primary high-risk HPV testing is higher as compared to that of conventional cytology(65, 83), which probably leads to an increased prevention of cervical cancers cases and deaths. Notwithstanding, the CIN II$^+$ specificity is lower(65, 83), which probably leads to increased numbers of triage testing and referrals to gynaecologist, both associated with an inevitable loss in quality of life(82). Nevertheless, it is expected that in general the benefits of primary HPV testing will outweigh the harms(84, 85). However, it is questionable whether this is also the case within young women aged below 35 years as (i) HPV infections are more often transient within younger women(65) and (ii) the prevalence of high-risk HPV infections peaks before the age of 35. To get a better understanding of the

potential benefit of switching from conventional cytology to a more sensitive screening test at young age, we estimated the proportion of cancer cases within 30 to 35 year old women that could have been prevented by using a more sensitive screening test at first screening at age 30. We analyzed the screening history of 30 to 35 year old women diagnosed with cervical cancer between 2004 and March 2009 and assessed the percentage of cervical cancer cases that were preceded by a negative cytology test under the age of 35. Also, we assessed the percentage of cervical cancer cases without a history of cervical cancer screening in order to get a better understanding of the potential impact of increasing attendance rates as compared to switching to a more sensitive screening test.

6. When is it effective to offer self-sampling to non-attendees of organized primary HPV screening?

Offering HPV self-sampling to non-attendees of cervical cancer screening could be used to increase participation rates and thereby decrease cervical cancer incidence and mortality. Its effect on the effectiveness of the programme probably not only depends on the increase in attendance, but also on the test characteristics of HPV self-sampling and on the ability to target higher underlying risk non-attendees. Moreover, "switching" of regular attendees from office-based to self-sampling could, given a loss in detection (ie. more loss to follow-up and a possible lower sensitivity), result in a decrease of the effectiveness of the programme. In other words, the quality-adjusted life-years (QALYs) gained by attracting non-attendees could be annulled by the QALYs lost by switching of regular attendees. However, it is unclear at which level of switching this will happen. Therefore, we examined effects of parameters, such as: the relative CIN II^{+} sensitivity and specificity (self-sampling versus regular sampling), the extra attendance via self-sampling, the underlying risk of extra attendees, and the percentage of women switching. We used the MISCAN-Cervix model to estimate under which conditions adding HPV self-sampling to the primary HPV screening programme would be more effective (ie. QALYs would be gained).

7. Is lower socioeconomic status associated with an increased underlying cervical cancer risk?

Lower socioeconomic status (SES) is considered to be a risk factor for developing cervical cancer(86, 87), which might be attributed solely to their lower screening participation rate(88-90), but could also be partly due to other factors. Therefore, we included all first-time attendees without any history of cervical examinations to eliminate influences of differences in screening history and uptake. We compared the cervical cancer risk between women living in low and intermediate versus high SES neighbourhoods who attended the organized cervical cancer screening programme between 2000 and 2007. As the SES of a neighbourhood is indicative of the SES of its inhabitants, we extrapolated our results from the ecological to the individual level.

PART I

PRIMARY LIQUID-BASED CYTOLOGY TESTING

CHAPTER 2

Comparing SurePath, ThinPrep and conventional cytology as primary test method: SurePath is associated with increased CIN II^{+} detection rates

Kirsten Rozemeijer[1], Corine Penning[1], Albert G Siebers[2, 3], Steffie K Naber[1], Suzette M Matthijsse[1], Marjolein van Ballegooijen[1], Folkert J van Kemenade[4], and Inge MCM de Kok[1]

1. Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, the Netherlands
2. Radboud University Medical Center, Department of Pathology, Nijmegen, the Netherlands
3. PALGA, the nationwide network and registry of histo- and cytopathology in the Netherlands, Houten, the Netherlands
4. Erasmus MC, University Medical Center, Department of Pathology, Rotterdam, the Netherlands

Cancer Causes Control. 2015 Oct 12. [Epub ahead of print]

ABSTRACT

Background. Within the last decade, SurePath and ThinPrep [both liquid-based cytology (LBC) tests] have replaced conventional cytology as primary test method in cervical cancer screening programmes of multiple countries. The aim of our study was to examine the effect in the Dutch screening programme.

Methods. All primary smears taken within this programme from 2000 to 2011 were analyzed using the nationwide registry of histo- and cytopathology (PALGA) with a follow-up until March 2013. The percentage of smears classified as borderline/mildly dyskaryotic (BMD) and >BMD as well as CIN and cervical cancer detection rates were compared between SurePath and ThinPrep versus conventional cytology by logistic regression analyses (adjusted for age, screening region, socioeconomic status, and calendar time).

Results. We included 3,118,685 conventional cytology, 1,313,731 SurePath and 1,584,587 ThinPrep smears. Using SurePath resulted in an increased rate of primary smears classified as >BMD [odds ratio (OR) = 1.12 (95% confidence interval (CI): 1.09, 1.16)]. CIN I and II$^+$ detection rates increased by 14% [OR = 1.14 (95% CI: 1.08, 1.20)] and 8% [OR = 1.08 (95% CI: 1.05, 1.12)]. Cervical cancer detection rates were unaffected. Implementing ThinPrep did not result in major alterations of the cytological classification of smears, and it did not affect CIN detection rates. While not significant, cervical cancer detection rates were lower [OR = 0.87 (95% CI: 0.75, 1.01)].

Conclusions. The impact of replacing conventional cytology by LBC as primary test method depends on the type of LBC test used. Only the use of SurePath was associated with increased CIN II$^+$ detection, although it simultaneously increased the detection of CIN I.

Keywords. Cervical intraepithelial neoplasia; Liquid-based cytology; SurePath; ThinPrep; Conventional cytology; Screening

BACKGROUND

Since the 1980s, a national cervical cancer screening programme exists in the Netherlands. From 1996 onward, women are invited every five years from ages 30 to 60 years. The screening strategy consists of primary cytology screening with triage by repeat cytology or triage by a combination of repeat cytology and human papillomavirus (HPV) testing (Figure 2-1). Despite its limited sensitivity(65), the conventional cytology test has long been used as primary test method.

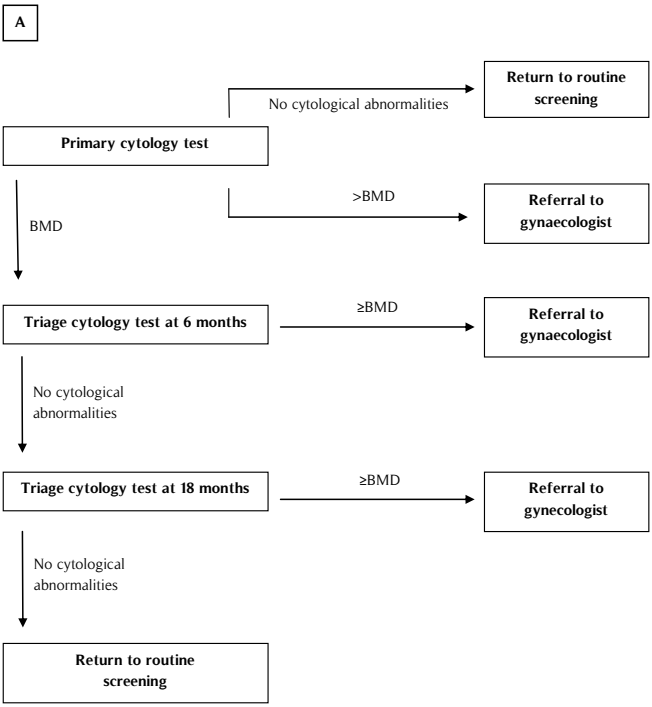

Figure 2-1. Triage protocol consisting of triage cytology (A) without HPV testing, and (B) with HPV testing.
HPV = Human papillomavirus; BMD = Borderline and mildly dyskaryotic smears.

Within the last 10 to 15 years, conventional cytology has been replaced by liquid-based cytology (LBC) tests SurePath or ThinPrep in most of Dutch laboratories processing primary screening tests. Conventional cytology and both LBC systems share the same method of sampling cells from the cervix (ie. scraping off cells with a brush or similar device from the histological transition zone). LBC differs from conventional cytology with respect to the transfer of cells from brush to slide: with conventional cytology, cells are directly smeared on a slide, while with LBC, the brush is first rinsed into a vial with a preservative fluid and then transported to a laboratory(49). In the laboratory, a uniform layer of cells is prepared on the slide(50, 51). It is thought that this method of cell transfer (which differs between SurePath and ThinPrep) results in a better representation of the entire sample as compared to conventional cytology(52). A review which evaluated the applicability of LBC in the Dutch cervical cancer screening programme concluded that further research was needed to determine the applicability of SurePath. Furthermore, they recommended to further analyse the costs and benefits of ThinPrep before deciding whether or not to implement this method(91). Yet, public health authorities in the Netherlands permitted use of both LBC systems based on perceived advantages such as: ease of processing, reduction in unsatisfactory slides(40, 42, 46, 48), and time needed to read the slides(42, 43, 45, 47). Finally, the use of LBC allowed for easier application of HPV co-testing(41, 44).

The use of conventional cytology as primary test method has also been replaced by the use of SurePath and/or ThinPrep in many other countries with and without organized cervical cancer screening programmes, such as Denmark, the UK, and the USA(92, 93). It is believed that the sensitivity of LBC for detecting cervical intraepithelial neoplasia (CIN) II^+ lesions is similar to that of conventional cytology(55, 56). However, when stratifying for the type of LBC test used, many studies have been published comparing CIN detection between ThinPrep and conventional cytology(45, 49, 52, 57, 59, 60), while only two studies have compared CIN detection between SurePath and conventional cytology(42, 58). Moreover, no studies have

been published comparing CIN detection rates between the three types of cytology tests. As the outcome of all cervix uteri cytological and histological tests taken within the Dutch screening programme were available [ie. are registered in the Dutch Pathology Register (PALGA)(94)] and we were able to deduce which type of primary cytology test had been used, we assessed whether differences in CIN detection rates were present when screened by SurePath or ThinPrep as compared to conventional cytology. In addition, we assessed the effect on cervical cancer detection rates and on the classification of smears.

METHODS

Information on all cervix uteri cytological and histological tests in the Netherlands registered from January 2000 until March 2013 was retrieved from PALGA. Women are identified through their birth date and the first eight letters of their (maiden) family name. This identification code enables linkage of tests belonging to the same woman, allowing us to follow individual screening histories. We identified primary smears (ie. first smear of an episode) taken within the national cervical cancer screening programme between January 2000 and December 2011. A minimum duration of 15 month follow-up was ensured as data until March 2013 were available. Histological confirmed CIN lesions and cervical cancer cases were identified by selecting all PALGA records that included corresponding pathology codes. Detection of these conditions was assigned to the type of cytology test used. Age was defined as the woman's age at the time of the primary smear and was categorised as: 29-33, 34-38, 39-43, 44-48, 49-53, 54-58, and 59-63 years. As women are invited every five years in the year they turn 30, 35, .., 60 years, these age categories reflect different screening rounds. The cervical cancer screening programme is organized by five different screening organizations, each accounting for a geographical region (ie. screening region) (North, South-West, Middle-West, South and East). Screening regions were coded corresponding with the place of residence at the time

of the primary smear. Socioeconomic status (SES) was defined (low, middle, high) according to the status score, which is an ecological variable based on the four-digit postal code of the woman's place of residence at the time of the primary test(95). Status scores per four-digit postal code were provided by the Netherlands Institute for Social Research based on 1) mean income, 2) percentage of households with a low income, 3) percentage of households with, on average, a low education, and 4) unemployment rate in 2010(96). Low SES corresponded with a status score lower than -1 (ie. average status score minus standard deviation), intermediate SES with a score of ≥-1 and ≤1, and high SES with a score higher than 1 (ie. average status score plus standard deviation).

In PALGA, the type of cytology testing is not routinely registered. Therefore, the date of conversion was retrieved from the laboratories fixed to one of the quarters per year, since most labs had a phase in-phase out transition period of 2-4 months. This information was linked to PALGA as a proxy for which type of primary cytology test was used (ie. in the Netherlands, laboratories supply the tools for cytology and thus determine the type of cytology test that is used by the general practitioner).

Type of cytology testing

With conventional cytology, cervical specimen is collected (ie. no data were available on the type of device or brush used), and cells are directly smeared from the sampling device on the slide. With SurePath, cervical specimen is collected using a broom like device with detachable head. The detachable head is placed in a vial with an ethanol based preservative fluid. At the laboratory, the fluid and cells are centrifuged to isolate the cells from the fluid. The cells are resuspended in a sucrose density gradient followed by slide transfer using gravity for adherence. With ThinPrep, cervical specimen is collected using a Cervix Brush, and the brush is rinsed in a vial with a methanol based preservative fluid. Cells are released by pushing the brush to the bottom, forcing the bristles apart, and swirling the brush into the fluid. Subsequently, the brush is discarded. At the laboratory, cells are isolated

from the fluid by vacuum filtration and are transferred to the slide using air pressure for adherence(53).

Statistical analyses

Since LBC was implemented per laboratory at different points in time, calendar time is expected to differ between the three types of cytology tests. The demographic characteristics of attending women (ie. age, screening region, and SES) also differ between laboratories; hence, we expected that they also differ between the cytology tests. As age, SES, screening region, and calendar time are all associated with CIN and/or cervical cancer(97, 98), they are all potential confounding factors. We used a Pearson's chi-squared test to assess whether their distributions differed between the types of cytological tests. Thus, we tested whether they were confounders or not. A p value of less than 0.05 was considered to be statistically significant.

We performed logistic regression analyses to examine whether CIN and cervical cancer detection rates differed between the types of cytological tests, adjusted for confounding factors. Moreover, we assessed how these overall changes in CIN and cervical cancer detection rates, if present, were composed. First, we examined whether the rate of primary smears classified as borderline/mildly dyskaryotic (BMD) differed between the types of cytological tests. Second, we assessed whether CIN and cervical cancer detection rates in women with a BMD smear were different between the types of tests, which would indicate that the positive predictive value (PPV) of a primary BMD smear differed. Third, we combined these two steps to examine whether the tests differed in the fraction of primary smears both classified as BMD and resulting in the detection of a CIN or cervical cancer. By performing the same analyses for having a >BMD smear, we could assess whether potential differences in CIN and cervical cancer rates were (mainly) caused by differences in the triage (ie. those with a primary BMD smear) or direct referral pathway (ie. those with a primary >BMD smear). Finally, we assessed the overall difference in CIN and cervical cancer detection rates, regardless of the cytological result.

Missing values were imputed with 10 multiple imputations for confounding factors. The odds ratio (OR) was interpreted as relative risk if the prevalence of the outcome (ie. BMD, >BMD, CIN I, CIN II, CIN III, cervical cancer or CIN II^{+}) was <10% in the respective logistic regression model analysis(99). The software programme SPSS (version 20) was used to perform the statistical analyses.

RESULTS

We included 3,118,685 primary conventional cytology smears, 1,313,731 primary SurePath smears, and 1,584,587 primary ThinPrep smears in our analyses. The distribution of calendar time significantly differed between

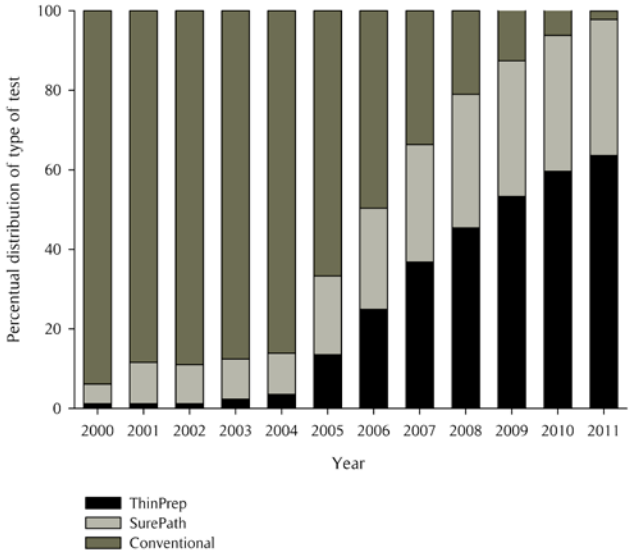

Figure 2-2. Distribution of the types of cytological tests used within the Dutch screening programme. The total number of primary smears where the type of cytological test was known varied from 441,663 in 2000 to 541,587 in 2007.

Table 2-1. Population characteristics.

	Conventional	SurePath	ThinPrep	*P* value
N	3,118,685	1,313,731	1,584,587	
Screening region				<0.001
1, n (%)	430,548 (13.8)	503,967 (38.4)	352,790 (22.3)	
2, n (%)	822,189 (26.4)	178,844 (13.6)	538,890 (34.0)	
3, n (%)	482,137 (15.5)	311,276 (23.7)	296,609 (18.7)	
4, n (%)	872,931 (28.0)	294,939 (22.5)	206,098 (13.0)	
5, n (%)	501,852 (16.1)	24,471 (1.9)	187,279 (11.8)	
Unknown, n (%)	9,028 (0.3)	234 (0.0)	2,921 (0.2)	
SES				<0.001
Low, n (%)	257,544 (8.3)	156,058 (11.9)	107,983 (6.8)	
Middle, n (%)	2,574,027 (82.5)	1,045,158 (79.6)	1,331,613 (84.0)	
High, n (%)	239,623 (7.7)	87,591 (6.7)	132,439 (8.4)	
Unknown, n (%)	47,491 (1.5)	24,924 (1.9)	12,552 (0.8)	
Age				<0.001
29-33 years, n (%)	428,600 (13.7)	170,699 (13.0)	195,935 (12.4)	
34-38 years, n (%)	522,173 (16.7)	191,193 (14.6)	220,462 (13.9)	
39-43 years, n (%)	533,438 (17.1)	222,906 (17.0)	271,924 (17.2)	
44-48 years, n (%)	496,856 (15.9)	219,118 (16.7)	265,672 (16.8)	
49-53 years, n (%)	446,596 (14.3)	195,127 (14.9)	243,354 (15.4)	
54-58 years, n (%)	388,637 (12.5)	171,194 (13.0)	207,405 (13.1)	
59-63 years, n (%)	302,385 (9.7)	143,494 (10.9)	179,835 (11.3)	

The distributions of factors associated with CIN and/or cervical cancer detection rates between the three primary test methods are given. If a distribution significantly differs between the primary tests (which is tested with a Pearson's Chi-Square test), the variable is considered to be a confounding factor.

the methods of cytology testing (p <0.001). In 2000, 94% of the primary cytology tests performed within the Dutch screening programme consisted of conventional cytology, while in 2011 this percentage has dropped to 2% (Figure 2-2). The distribution of age, SES, and screening region also

Table 2-2. Logistic regression analyses on the classification of smears and histological outcomes when tested by SurePath or Thinprep versus conventional cytology, adjusted for age, SES, screening region, and calendar time.

Outcome	SurePath	ThinPrep
*BMD	**0.96 (0.94, 0.97)**	**1.02 (1.01, 1.04)**
¥PPV of a primary BMD smear on histological outcomes		
CIN I	**1.26 (1.18, 1.34)**	1.03 (0.97, 1.10)
CIN II	**1.16 (1.08, 1.25)**	1.04 (0.97, 1.12)
CIN III	0.95 (0.88, 1.02)	**0.87 (0.81, 0.94)**
Cervical Cancer	0.74 (0.49, 1.12)	**0.62 (0.41, 0.92)**
#Fraction of primary smears both classified as BMD and resulting in the detection of the following histological outcomes		
*CIN I	**1.20 (1.13, 1.27)**	1.06 (0.99, 1.12)
*CIN II	**1.14 (1.07, 1.22)**	**1.08 (1.00, 1.15)**
*CIN III	0.99 (0.92, 1.06)	0.93 (0.87, 1.00)
*Cervical Cancer	0.77 (0.51, 1.15)	0.66 (0.43, 1.00)
*>BMD	**1.12 (1.09, 1.16)**	**0.96 (0.93, 0.99)**
¥PPV of a primary >BMD smear on histological outcomes		
CIN I	0.92 (0.83, 1.03)	**0.86 (0.77, 0.97)**
CIN II	1.06 (0.98, 1.15)	1.08 (0.99, 1.17)
CIN III	0.97 (0.91, 1.03)	1.06 (0.99, 1.13)
Cervical Cancer	0.94 (0.80 , 1.10)	0.98 (0.83, 1.15)
$Fraction of primary smears both classified as >BMD and resulting in the detection of the following histological outcomes		
*CIN I	1.05 (0.94, 1.16)	**0.83 (0.74 , 0.92)**
*CIN II	**1.17 (1.09, 1.27)**	1.02 (0.94, 1.10)
*CIN III	**1.10 (1.06, 1.15)**	1.00 (0.96, 1.04)
*Cervical Cancer	1.07 (0.91, 1.24)	0.93 (0.79, 1.09)
Overall histological outcomes		
*CIN I	**1.14 (1.08, 1.20)**	0.98 (0.93, 1.04)
*CIN II	**1.14 (1.09, 1.20)**	1.04 (0.99, 1.10)
*CIN III	**1.06 (1.02, 1.10)**	0.98 (0.94, 1.01)
*Cervical Cancer	0.99 (0.86, 1.14)	0.87 (0.75, 1.01)

Odds ratios with a 95% confidence interval are given. This table shows how the overall changes in CIN and cervical cancer detection rates, if present, are composed. The differ-

ences in the odds of primary smears classified as BMD combined with the differences in the odds of the PPV of a BMD smear led to differences in the fraction of primary smears both classified as BMD and resulting in the detection of a CIN or cervical cancer. By performing the same analyses for having a >BMD smear, we could assess whether potential differences in CIN and cervical cancer rates were (mainly) caused by differences in the triage (ie. those with a primary BMD smear) or direct referral pathway (ie. those with a primary >BMD smear). Altogether, this led to differences in odds of overall CIN and cervical cancer detection.

Bold = Significant. A *p* value of <0.05 was considered to be statistically significant.

BMD = Borderline and mildly dyskaryotic smears; PPV = Positive predictive value

*Odds ratio could be interpreted as detection rate ratio because the prevalence of the outcome was <10%.

¥This can be interpreted as: Does a BMD or >BMD smear more often lead to the following histological outcomes when using SurePath or ThinPrep as compared to conventional cytology.

#Histological outcomes detected via triage.

$Histological outcomes detected via direct colposcopy.

significantly differed between the methods of cytology testing (Table 2-1). For instance, most conventional cytology tests were performed in screening region 4 (28%), while most SurePath and ThinPrep tests were performed in screening regions 1 (38.4%) and 2 (34.0%). Thus, calendar time, age, SES, and screening region were all considered confounding factors and missing values were imputed for 1.6 % of the primary smears.

The effect of SurePath versus conventional cytology, adjusted for confounding factors

When comparing using SurePath with using conventional cytology as primary test method, 4% fewer primary smears were classified as BMD [OR of 0.96 (95% confidence interval (CI): 0.94 – 0.97)], while a BMD smear more often led to a CIN I [OR of 1.26 (95% CI: 1.18, 1.34)] or CIN II diagnosis [OR of 1.16 (95% CI: 1.08, 1.25)]. Combined this led to a 20% [OR of 1.20 (95% CI: 1.13, 1.27)] and 14% [OR of 1.14 (95% CI: 1.07, 1.22)] increase in the fraction of primary smears both classified as BMD and resulting in the detection of a CIN I or CIN II lesion (Table 2-2, for the unadjusted results see the Supplementary Information).

The rate of primary smears classified as >BMD increased by 12% [OR of 1.12 (95% CI: 1.09, 1.16)], whereas a smear classified as >BMD led to a similar number of CIN I, CIN II, CIN III and cervical cancer diagnoses. As a result, the fraction of primary smears both classified as >BMD and resulting in the detection of a CIN II or CIN III lesion increased by 17% [OR of 1.17 (95% CI: 1.09, 1.27)] and 10% [OR of 1.10 (95% CI: 1.06, 1.15)].

Overall, CIN I, CIN II, and CIN III detection rates increased by 14% [OR of 1.14 (95% CI: 1.08, 1.20)], 14% [OR of 1.14 (95% CI: 1.09, 1.20)], and 6% [OR of 1.06 (95% CI: 1.02, 1.10)], respectively, when using SurePath as compared to using conventional cytology as primary test method. Cervical cancer detection rates were equivocal between both tests [OR of 0.99 (95% CI: 0.86, 1.14)]. CIN II^{+} detection rates increased by 8% [OR of 1.08 (95% CI: 1.05, 1.12)].

The effect of ThinPrep versus conventional cytology, adjusted for confounding factors

When using ThinPrep as compared to using conventional cytology as primary test method, the rate of primary smears classified as BMD increased by 2% [OR of 1.02 (95% CI: 1.01, 1.04)], although a primary smear classified as BMD less often resulted in a CIN III [OR of 0.87 (95% CI: 0.81, 0.94)] or cervical cancer diagnosis [OR of 0.62 (95% CI 0.41, 0.92)]. Combined this led to a marginally significant 8 % increase [OR of 1.08 (95% CI 1.00, 1.15)] in the fraction of primary smears both classified as BMD and resulting in the detection of a CIN II lesion. The fraction of primary smears both classified as BMD and resulting in the detection of a CIN I lesion nonsignificantly increased [OR of 1.06 (95% CI: 0.99, 1.12)], while the fraction both classified as BMD and resulting in the detection of a CIN III or cervical cancer nonsignificantly decreased [ORs of 0.93 (95% CI: 0.87, 1.00) and 0.66 (95% CI: 0.43, 1.00), respectively] (Table 2-2, for the unadjusted results see the Supplementary Information).

The rate of primary smears classified as >BMD decreased by 4% [OR of 0.96 (95% CI: 0.93, 0.99)]. A primary smear classified as >BMD less often resulted in a CIN I diagnosis [OR of 0.86 (95% CI: 0.77, 0.97)], although it nonsignificantly resulted in more CIN II and CIN III diagnoses [ORs of 1.08 (95% CI: 0.99, 1.17) and 1.06 (95% CI: 0.99, 1.13), respectively]. As a result, fraction of primary smears both classified as >BMD and resulting in the detection of a CIN I lesion decreased by 17% [OR of 0.83 (95% CI: 0.74, 0.92)].

Overall, using ThinPrep as primary test method did not have a significant effect on CIN I [OR of 0.98 (95% CI: 0.93, 1.04)], CIN II [OR of 1.04 (95% CI: 0.99, 1.10)], CIN III [OR of 0.98 (95% CI: 0.94, 1.01)] (Table 2-2) or CIN II$^+$ detection rates [OR of 0.99 (95% CI: 0.96, 1.02)]. Cervical cancer detection rates were nonsignificantly lower [OR of 0.87 (95% CI: 0.75, 1.01)].

DISCUSSION

Using SurePath versus conventional cytology as primary test method resulted in a 12% increase in the rate of primary smears classified as >BMD. The rate of primary smears classified as BMD decreased by 4% and women with a primary BMD smear were more often diagnosed with CIN I or II. Combined this led to increased fractions of primary smears both classified as BMD and resulting in the detection of a CIN I or CIN II lesion, and to increased fractions of primary smears both classified as >BMD and resulting in the detection of a CIN II or CIN III lesion. Altogether, the detection of CIN II$^+$ increased by 8% accompanied by a 14% increase in the detection of CIN I. Cervical cancer rates were unaffected. The comparison of using ThinPrep versus conventional cytology did not result in such findings, although the sensitivity to detect cervical cancers might be lower.

Given the differences in preparation between both LBC methods, it is possible that the sensitivity for CIN II$^+$ differs between them as well. For

instance, it was shown that the cell yield is larger when the collecting device was retained instead of discarded from the vial with preservative fluid(100, 101), meaning that if the protocol is followed the cell yield is larger when using SurePath (ie. collecting device is retained) than when using ThinPrep (ie. collecting device is discarded). Therefore, the probability of transferring abnormal cells from the cervical specimen (if present) to the slide is probably larger when using SurePath. The study of Rask et al. seems to confirm this, since they found that replacing conventional cytology by SurePath resulted in a significant 31% increase in cytological abnormalities within 23-29 aged women, while replacing conventional cytology by ThinPrep resulted in a nonsignificant 11% decrease(92).

Our research demonstrated that CIN II$^+$ detection rates are similar between ThinPrep and conventional cytology, which is compatible with results of previous studies. For instance, the observed CIN II$^+$ detection rate ratio of 0.99 (95% CI: 0.96, 1.02) fits with the pooled relative CIN II$^+$ sensitivity of 1.03 (95% CI: 0.97, 1.09) as reported in the meta-analysis of Arbyn et al. (ie. our point estimate lies within the 95% CI)(55). However, that ratio was based on seven studies comparing LBC with conventional cytology of which two did not use ThinPrep as LBC test method. When only focusing on the included ThinPrep studies, we found and calculated (ie. using data provided in the study) CIN II$^+$ detection rate ratios of 1.17 (95% CI: 0.87, 1.56)(57), 0.97 (95% CI: 0.61, 1.55)(60), 0.95 (95% CI: 0.62, 1.48)(52), and 1.09 (95% CI: 0.80, 1.48)(59), which were all compatible to the detection rates observed in the present study. The CIN II$^+$ detection rate ratio of the fifth included ThinPrep study was not provided nor could be calculated(45). Furthermore, the largest randomized controlled trial performed so far, including almost 90,000 participants, found a CIN II$^+$ detection rate ratio of 1.00 (95% CI: 0.84, 1.20)(49) which also fits our data. When focusing on studies comparing SurePath with conventional cytology, only one previous study matched our criteria (ie. providing a CIN II$^+$ detection rate ratio, or data needed to calculate it, at a cut-off of ASCUS or BMD)(58). Again, their data [ie. a CIN II$^+$ detection rate of 1.01 (95% CI: 0.76, 1.33)] fitted with ours

[ie. a ratio of 1.08 (95% CI: 1.05, 1.12)] (ie. our point estimate lies within the reported 95% CI).

It is expected that from 2016 onward, primary cytology screening will be replaced by primary HPV screening in the Dutch cervical cancer screening programme. If high-risk HPV is present, a reflex cytology triage test will be carried out on the same sample followed by another triage test six months later, if the reflex cytology triage test shows no abnormalities. If one of these smears is classified as ≥BMD, the woman will be referred to the gynaecologist for colposcopy, otherwise she will be referred to routine screening. Whether our results can be extended from a primary screening to a triage population depends on the performance of the cytology tests on (i) fluid remnant after primary HPV testing in HPV-positive women (in case of reflex triage testing), and (ii) directly taken material in (previously) HPV-positive women (in case of triage testing at six months). Although prior knowledge of the HPV status influences the interpretation of cytological smears(102, 103), we assume this effect to be similar for the three types of cytology tests. If true, we expect the differences in sensitivity between Surepath, ThinPrep, and conventional cytology in a triage population to be equivalent to the differences in a primary screening population. However, this assumption has not been tested yet. In addition, because conventional cytology cannot be performed on fluid remnant after primary HPV testing(50, 55), our results of comparing SurePath and ThinPrep with conventional cytology cannot be extended to reflex triage testing. As data of Cuzick et al. suggested that the performance of HPV assays depends on the type of LBC test used(104), it is also possible that the performance of LBC tests on fluid remaining after HPV testing depends on the type of HPV assay used. Thus, more research is needed to assess which combination of primary HPV test and secondary reflex LBC test has the highest CIN II$^+$ sensitivity.

We were the first who compared CIN and cervical cancer detection rates between Surepath, ThinPrep and conventional cytology. Furthermore, we

included more than six million primary smears, and we showed its effect in real practice instead of in a strictly controlled setting.

At the same time, the lack of a more controlled setting is one of the limitations of our study. As ThinPrep and SurePath were used in different women, differences in demographic factors were inevitable. Although we were able to correct for confounders age, screening region, SES, and calendar time, we were not able to correct for other potential confounding factors such as screening history or compliance with the given advice. Both could have resulted in biased effect estimates if their distribution differed between the types of cytology tests. In addition, no data are present whether cytology triage testing at six months was combined with HPV testing. Because of the possibility of co-testing, it is likely that the use of HPV triage is correlated with the use of primary LBC testing. As it is known that more CIN I and CIN II lesions are detected when cytology triage is combined with HPV(39), it is probable that the increased sensitivity of SurePath to detect CIN I and CIN II was partly caused by the simultaneous use of HPV testing. However, the entire increase in CIN III detection rates when comparing Surepath with conventional cytology, and for a large part also the increase in CIN II detection rates, is caused by an increase of primary smears being classified as >BMD. Therefore, we still believe that SurePath results in increased CIN II+ detection rates although it might be accompanied by a smaller increase in CIN I detection than estimated. Also, we did not have individual data on which type of primary test was used. Therefore, we combined the date of the primary smear and the quarter of the year within which the laboratory introduced the LBC test as proxy for the type of cytology test that was used. This means that primary screening smears taken during this quarter could have been misclassified, resulting in slightly underestimated effects. Another shortcoming of the study was that we were not able to correct for the use of automated reading, although this has only been introduced in relatively few Dutch laboratories. As study results on the effect of automated screening are heterogeneous, it is unknown how this affected our effect estimates. If automated reading does not affect the sensitivity for CIN II$^+$, as shown by Klug and Palmer et al.(61, 62), our estimates are not biased. If

automated reading results in a decreased sensitivity for CIN II+, as shown in the MAVARIC study(63), we might have underestimated the effect of using SurePath and ThinPrep on CIN II+ detection rates. If it results in an increased sensitivity, we might have overestimated the effects. At last, we did not correct for possible learning curve effects, as the aim of our study was to examine the effect of using SurePath and ThinPrep in routine-practice, which also includes a possible learning effect.

Our results indicate that the widespread use of SurePath as primary test method has led to an increased probability to detect both CIN I and CIN II+ lesions. As only a small fraction of CIN I lesions progress to cancer, increased CIN I detection is often regarded as increased overdiagnosis. In contrast, CIN II+ lesions are associated with a substantial cancer risk and are therefore often considered as clinically relevant. However, whether the increased probability to detect CIN II+ lesions indeed corresponds with an increased sensitivity for progressive lesions remains to be investigated. If this is the case, using SurePath would in due time result in a decrease of the incidence and mortality of cervical cancer, thereby increasing the health benefits of the screening programme. If not, it would only lead to increased burden and harms through overdiagnosis (and treatment) of regressive CIN lesions. The widespread use of ThinPrep as primary test method did not lead to changes in CIN II+ detection rates, although cervical cancer detection was nonsignificantly lower. Whether these results imply a decreased sensitivity for progressive CIN II+ lesions is unknown. For evidence as to whether the detection of progressive CIN II+ lesions is higher with any of the LBC tests than with conventional cytology, cervical interval cancer rates have to be compared.

ACKNOWLEDGEMENTS

This study was funded by the Dutch National Institute for Public Health and the Environment (007/12 V&Z NvdV/EM). We thank C. Looman for statis-

tical advice, the regional coordinating pathologists (RCP) for their help in collecting data, and L.I.H. Overbeek for linking data of the laboratories to the PALGA database.

Table 2-S1. Logistic regression analyses on the classification of smears and histological outcomes when tested by SurePath or Thinprep versus conventional cytology, unadjusted and adjusted for confounding factors. Odds ratios with a 95% confidence interval are given. This table shows the impact of adjusting for confounding factors.

	Effect of using either SurePath or ThinPrep instead of conventional cytology as primary test method				
	Unadjusted	Adjusted for year	Adjusted for year and age	Adjusted for year, age, and screening region	Adjusted for year, age, screening region and SES
Fraction of primary smears classified as BMD					
SurePath	**1.23 (1.21, 1.24)**	0.99 (0.97, 1.01)	0.99 (0.97, 1.00)	**0.97 (0.95, 0.98)**	**0.96 (0.94, 0.97)**
ThinPrep	**1.40 (1.39, 1.42)**	**1.05 (1.03, 1.07)**	**1.05 (1.03, 1.06)**	**1.02 (1.00, 1.04)**	**1.02 (1.01, 1.04)**
[¥]PPV of a primary BMD smear on histological outcomes					
CIN I					
SurePath	**1.37 (1.30, 1.43)**	**1.23 (1.17, 1.31)**	**1.23 (1.16, 1.30)**	**1.26 (1.19, 1.34)**	**1.26 (1.18, 1.34)**
ThinPrep	**1.16 (1.11, 1.21)**	1.03 (0.97, 1.09)	1.02 (0.96, 1.08)	1.04 (0.97, 1.10)	1.03 (0.97, 1.10)
CIN II					
SurePath	**1.29 (1.22, 1.36)**	**1.17 (1.10, 1.25)**	**1.15 (1.07, 1.23)**	**1.17 (1.09, 1.26)**	**1.16 (1.08, 1.25)**
ThinPrep	**1.18 (1.12, 1.24)**	1.07 (0.99, 1.14)	1.04 (0.97, 1.12)	1.04 (0.97, 1.12)	1.04 (0.97, 1.12)
CIN III					
SurePath	0.94 (0.89, 1.00)	0.99 (0.92, 1.06)	0.95 (0.89, 1.02)	0.94 (0.88, 1.02)	0.95 (0.88, 1.02)
ThinPrep	**0.87 (0.82, 0.92)**	0.93 (0.87, 1.00)	**0.91 (0.84, 0.98)**	**0.87 (0.81, 0.94)**	**0.87 (0.81, 0.94)**
Cervical cancer					
SurePath	**0.70 (0.51, 0.95)**	0.76 (0.52, 1.11)	0.76 (0.53, 1.11)	**0.74 (0.60, 0.91)**	0.74 (0.49, 1.12)
ThinPrep	**0.58 (0.43, 0.77)**	**0.65 (0.43, 0.96)**	**0.64 (0.43, 0.96)**	**0.62 (0.50, 0.77)**	**0.62 (0.41, 0.92)**
[#]Fraction of primary smears both classified as BMD and resulting in the detection of the following histological outcomes					
CIN I					

Table 2-S1. (continued)

	Effect of using either SurePath or ThinPrep instead of conventional cytology as primary test method				
	Unadjusted	Adjusted for year	Adjusted for year and age	Adjusted for year, age, and screening region	Adjusted for year, age, screening region and SES
SurePath	**1.62 (1.55, 1.69)**	**1.20 (1.14, 1.27)**	**1.20 (1.14, 1.27)**	**1.22 (1.15, 1.29)**	**1.20 (1.13, 1.27)**
ThinPrep	**1.60 (1.53, 1.67)**	**1.07 (1.01, 1.13)**	**1.07 (1.01, 1.14)**	1.05 (0.99, 1.12)	1.06 (0.99, 1.12)
CIN II					
SurePath	**1.55 (1.47, 1.64)**	**1.15 (1.08, 1.22)**	**1.15 (1.08, 1.22)**	**1.16 (1.09, 1.24)**	**1.14 (1.07, 1.22)**
ThinPrep	**1.63 (1.55, 1.71)**	**1.11 (1.04, 1.18)**	**1.11 (1.04, 1.19)**	**1.07 (1.00, 1.15)**	**1.08 (1.00, 1.15)**
CIN III					
SurePath	**1.16 (1.10, 1.22)**	0.97 (0.91, 1.04)	0.98 (0.92, 1.05)	0.99 (0.93, 1.07)	0.99 (0.92, 1.06)
ThinPrep	**1.22 (1.16, 1.28)**	0.98 (0.91, 1.05)	0.99 (0.92, 1.06)	0.93 (0.87, 1.00)	0.93 (0.87, 1.00)
Cervical cancer					
SurePath	0.85 (0.63, 1.16)	0.75 (0.52, 1.09)	0.75 (0.52, 1.09)	**0.77 (0.63, 0.95)**	0.77 (0.51, 1.15)
ThinPrep	0.81 (0.60, 1.08)	0.68 (0.46, 1.01)	0.68 (0.46, 1.01)	**0.66 (0.53, 0.81)**	0.66 (0.43, 1.00)
Fraction of primary smears classified as >BMD					
SurePath	**1.21 (1.18, 1.24)**	**1.06 (1.03, 1.09)**	**1.07 (1.04, 1.10)**	**1.15 (1.11, 1.18)**	**1.12 (1.09, 1.16)**
ThinPrep	**1.11 (1.09, 1.14)**	**0.93 (0.90, 0.96)**	**0.93 (0.90, 0.96)**	**0.95 (0.92, 0.99)**	**0.96 (0.93, 0.99)**
[¥]PPV of a primary >BMD smear on histological outcomes					
CIN I					
SurePath	1.03 (0.95, 1.12)	0.94 (0.85, 1.05)	0.93 (0.84, 1.04)	0.92 (0.83, 1.03)	0.92 (0.83, 1.03)
ThinPrep	0.98 (0.90, 1.06)	**0.87 (0.77, 0.97)**	**0.87 (0.78, 0.98)**	**0.86 (0.77, 0.97)**	**0.86 (0.77, 0.97)**
CIN II					
SurePath	**1.14 (1.07, 1.21)**	1.04 (0.96, 1.12)	1.04 (0.97, 1.13)	1.06 (0.98, 1.15)	1.06 (0.98, 1.15)

Table 2-S1. (continued)

	Effect of using either SurePath or ThinPrep instead of conventional cytology as primary test method				
	Unadjusted	Adjusted for year	Adjusted for year and age	Adjusted for year, age, and screening region	Adjusted for year, age, screening region and SES
ThinPrep	**1.23 (1.16, 1.30)**	**1.09 (1.00, 1.18)**	1.08 (1.00, 1.17)	1.08 (0.99, 1.18)	1.08 (0.99, 1.17)
CIN III					
SurePath	0.99 (0.94, 1.03)	0.97 (0.92, 1.03)	0.99 (0.94, 1.05)	0.97 (0.91, 1.03)	0.97 (0.91, 1.03)
ThinPrep	**1.13 (1.08, 1.19)**	**1.12 (1.05, 1.19)**	**1.09 (1.03, 1.17)**	1.06 (0.99, 1.13)	1.06 (0.99, 1.13)
Cervical cancer					
SurePath	0.96 (0.85, 1.08)	0.99 (0.86, 1.15)	0.98 (0.85, 1.13)	0.93 (0.80, 1.09)	0.94 (0.80, 1.10)
ThinPrep	0.95 (0.84, 1.06)	0.99 (0.85, 1.16)	1.00 (0.85, 1.17)	0.98 (0.83, 1.15)	0.98 (0.83, 1.15)
[$]Fraction of screened women having both a >BMD smear and the following histological outcomes					
CIN I					
SurePath	**1.25 (1.15, 1.35)**	1.01 (0.91, 1.11)	1.01 (0.91. 1.11)	1.07 (0.96, 1.19)	1.05 (0.94, 1.16)
ThinPrep	**1.09 (1.01, 1.18)**	**0.81 (0.73, 0.90)**	**0.82 (0.73, 0.91)**	**0.82 (0.74, 0.92)**	**0.83 (0.74, 0.92)**
CIN II					
SurePath	**1.35 (1.27, 1.43)**	**1.10 (1.02, 1.18)**	**1.10 (1.02, 1.18)**	**1.20 (1.11, 1.29)**	**1.17 (1.09, 1.27)**
ThinPrep	**1.32 (1.25, 1.39)**	0.99 (0.92, 1.07)	1.00 (0.93, 1.08)	1.02 (0.94, 1.10)	1.02 (0.94, 1.10)
CIN III					
SurePath	**1.20 (1.16, 1.24)**	**1.05 (1.01, 1.09)**	**1.05 (1.01, 1.09)**	**1.12 (1.08, 1.17)**	**1.10 (1.06, 1.15)**
ThinPrep	**1.17 (1.14, 1.21)**	0.97 (0.93, 1.01)	0.98 (0.94, 1.02)	1.00 (0.95, 1.04)	1.00 (0.96, 1.04)
Cervical cancer					
SurePath	**1.16 (1.03, 1.31)**	1.05 (0.91, 1.21)	1.05 (0.91, 1.21)	1.08 (0.93, 1.26)	1.07 (0.91, 1.24)
ThinPrep	1.05 (0.94, 1.18)	0.92 (0.79, 1.08)	0.92 (0.79, 1.07)	0.93 (0.79, 1.09)	0.93 (0.79, 1.09)

Table 2-S1. (continued)

	Effect of using either SurePath or ThinPrep instead of conventional cytology as primary test method				
	Unadjusted	Adjusted for year	Adjusted for year and age	Adjusted for year, age, and screening region	Adjusted for year, age, screening region and SES
Overall histological outcomes					
CIN I					
SurePath	**1.50 (1.45, 1.56)**	**1.13 (1.08, 1.19)**	**1.13 (1.08, 1.19)**	**1.16 (1.10, 1.21)**	**1.14 (1.08, 1.20)**
ThinPrep	**1.46 (1.40, 1.51)**	0.99 (0.94, 1.04)	0.99 (0.95, 1.05)	0.98 (0.93, 1.03)	0.98 (0.93, 1.04)
CIN II					
SurePath	**1.44 (1.39, 1.50)**	**1.11 (1.06, 1.16)**	**1.11 (1.06, 1.16)**	**1.16 (1.11, 1.22)**	**1.14 (1.09, 1.20)**
ThinPrep	**1.47 (1.42, 1.53)**	1.05 (1.00, 1.10)	**1.05 (1.00, 1.11)**	1.04 (0.99, 1.10)	1.04 (0.99, 1.10)
CIN III					
SurePath	**1.18 (1.15, 1.21)**	1.02 (0.98, 1.05)	1.02 (0.99, 1.06)	**1.08 (1.04, 1.12)**	**1.06 (1.02, 1.10)**
ThinPrep	**1.18 (1.15, 1.21)**	0.97 (0.93, 1.00)	0.98 (0.94, 1.01)	0.97 (0.94, 1.01)	0.98 (0.94, 1.01)
Cervical cancer					
SurePath	1.08 (0.97, 1.21)	0.98 (0.86, 1.11)	0.97 (0.85, 1.11)	1.01 (0.87, 1.16)	0.99 (0.86, 1.14)
ThinPrep	1.00 (0.90, 1.11)	0.87 (0.76, 1.01)	0.87 (0.75, 1.00)	0.87 (0.75, 1.01)	0.87 (0.75, 1.01)
CIN II+					
SurePath	**1.25 (1.22, 1.28)**	**1.04 (1.02, 1.07)**	**1.05 (1.02, 1.08)**	**1.10 (1.07, 1.13)**	**1.08 (1.05, 1.12)**
ThinPrep	**1.26 (1.23, 1.29)**	0.99 (0.96, 1.02)	1.00 (0.97, 1.03)	0.99 (0.96, 1.02)	0.99 (0.96, 1.02)

Bold = Significant. A *p* value of <0.05 was considered to be statistically significant.

[¥]This can be interpreted as: Does a BMD or >BMD smear more often lead to the following histological outcomes when using SurePath or ThinPrep as compared to conventional cytology.

[#]Histological outcomes detected via triage.

[$]Histological outcomes detected via direct colposcopy.

SES = Socioeconomic status; BMD = Borderline/mildly dyskaryotic; PPV = Positive predictive value.

CHAPTER 3

Exploring the trend of increased cervical intraepithelial neoplasia detection rates in the Netherlands

Kirsten Rozemeijer[1], Folkert J van Kemenade[2], Corine Penning[1], Suzette M Matthijsse[1], Steffie K Naber[1], Joost van Rosmalen[1,3], Marjolein van Ballegooijen[1], and Inge MCM de Kok[1]

1. Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, the Netherlands
2. Erasmus MC, University Medical Center, Department of Pathology, Rotterdam, the Netherlands
3. Erasmus MC, University Medical Center, Department of Biostatistics, Rotterdam, the Netherlands

J Med Screen. 2015 Sep;22(3):144-50.

ABSTRACT

Background. Over the last decade, cervical intraepithelial neoplasia (CIN) detection has increased in the Netherlands. We investigated the underlying mechanism by quantifying the increase, and analyzing patterns of CIN and cervical cancer detection over time.

Methods. We observed annual CIN and cervical cancer detection rates per 10,000 primary smears within the Dutch screening programme for 2000 to 2011. Joinpoint analyses were performed to determine changes in time trends, logistic regression analyses to assess the relative risk of calendar time on histological outcomes, adjusted for demographic factors and type of primary cytology test used.

Results. Trends of increased detection occurred for all CIN grades (ie. detection rates increased from 17.8 to 36.1, from 21.0 to 35.5, and from 43.4 to 64.6 for CIN I, II, and III from 2003 to 2009). After adjusting for demographic factors, detection rates were still 2.11 [95% confidence interval (CI): 1.95, 2.29], 1.79 (95% CI: 1.66, 1.92), and 1.59 (95% CI: 1.50, 1.67) times higher in 2009. When also adjusting for the type of cytology test, detection rates were 1.90 (95% CI: 1.62, 2.22), 1.48 (95% CI: 1.22, 1.79), and 1.55 (95% CI: 1.39, 1.73) times higher. No trends in cervical cancer detection rates were found.

Conclusions. The implementation of liquid-based cytology contributed to the CIN increase. If some of these extra detected CIN are regressive this leads to overdiagnosis. Other factors, such as an increased cervical cancer risk, and implementation of imaging-assisted reading, could also have contributed.

Keywords. Cervical cancer; Screening; Trend; Cervical intraepithelial neoplasia

BACKGROUND

Cervical intraepithelial neoplasia (CIN) is considered a preclinical precursor of cervical cancer and is graded (CIN I, II, and III) by the severity of the dysplasia(105, 106). The long total duration of the detectable preclinical dysplastic stage [ie. more than 10 years on average(21-23)] means that screening is an appropriate approach for cervical cancer prevention. A national cervical cancer screening programme has existed in the Netherlands since the 1980s, with women aged 30–60 invited every five years since 1996. Despite its limited sensitivity(65), cytology is still utilized as the primary screening and triage test. After a borderline or mildly dyskaryotic (BMD) result (ie. corresponding to ASC-H of the Bethesda classification), women are invited for triage cytology tests six [with or without high-risk human papillomavirus (HPV) testing] and 18 months later (see the Supplementary Information for a detailed triage protocol). For a worse result (ie. >BMD), women are immediately referred to a gynaecologist. According to the guidelines, CIN II and III are always treated, while CIN I is only treated when persistent(107). The Dutch screening programme achieves a cervical smear coverage of almost 80%(80).

During the last decade, CIN detection rates have rapidly increased in the Dutch screening programme(80). It is important to reveal the underlying mechanism, so that possible adjustments to the screening programme can be made in order to ensure its effectiveness. We therefore quantified the increase in CIN detection rates, unadjusted and adjusted for differences in demographic factors and types of cytology tests used over time. In addition, we analyzed patterns of CIN detection by age, screening region, and the cytological classification of smears. Furthermore, we studied cervical cancer detection rates and the cytological classification of smears over time.

METHODS

Data collection

Information on all cervix uteri cytological and histological tests in the Netherlands registered from January 2000 until March 2013 were retrieved from the nationwide network and registry of histology and cytopathology PALGA(94). Women were invited every five years in the year they turn 30, 35, .., 60 so that attendance occurs in age groups 29-33, 34-38, …, 59-63. These women are identified through their birth date and the first eight letters of their (maiden) family name. This identification code enables linkage of the woman's tests, allowing us to follow individual screening histories. We identified primary smears (ie. first smear of an episode) taken within the national screening programme between January 2000 and December 2011, which ensures a minimum follow-up of 15 months after a primary smear. Histologically confirmed CIN lesions or cervical cancer cases following primary smears were identified by selecting all PALGA records that included corresponding pathology codes. Detection of these conditions was assigned to the calendar year of the primary smear. Age was defined as the age at the primary smear, and was categorized as: 29– 33, 34–38,..., 59–63. The cervical cancer screening programme is organized by five different screening organizations, each accounting for a geographical region (ie. screening region) (North, South-West, Middle-West, South, and East). Screening regions were coded corresponding with the place of residence at the time of the primary smear. Women were stratified to low, intermediate or high socioeconomic status (SES) according to their status score, which is an ecological variable based on the four-digit postal code of the woman's place of residence at the time of the primary test(95). Status scores per four-digit postal code were provided by the Netherlands Institute for Social Research(96) based on (i) mean income, (ii) percentage of households with a low income, (iii) percentage of households with, on average, a low education, and (iiii) unemployment rate in 2010. Low SES corresponded with a status score lower than -1 (ie. average status score minus standard deviation), intermediate

SES with a score of ≥-1 and ≤1, and high SES with a score higher than 1 (ie. average status score plus standard deviation).

The type of primary cytology test used (ie. conventional cytology, SurePath, or ThinPrep) was based on the date of the primary smear, the laboratory involved, and the date of conversion from the laboratory. This conversion date was retrieved from the laboratories and fixed to one of the quarters per year, as most laboratories had a phase in-phase out transition period of 2-4 months, and this information was linked to PALGA as a proxy for which type of primary cytology test was used.

Statistical analyses

Annual CIN and cervical cancer detection rates were calculated as the number of diagnoses per 10,000 primary smears, allocated to the calendar year of the primary smear, unstratified and stratified by cytological classification (ie. BMD or >BMD). CIN was defined as all histological neoplastic lesions in the cervix similar to CIN I, CIN II, or CIN III, including glandular neoplasia (such as adenocarcinoma in situ). The annual rate of smears classified as BMD and >BMD were also calculated.

Joinpoint analyses were performed to identify time trends, using the Joinpoint Regression Programme (version 4.0.4.) from the Statistical Research and Applications Branch of the US National Cancer Institute(108). Annual percent changes (APCs) with their corresponding 95% confidence intervals (CIs) were calculated by fitting a regression line to the natural logarithm of the rates using calendar time as independent variable [ie. y = ax + b where y = ln(rate) and x = calendar year, then APC = 100 * (e^a-1)].

We performed logistic regression analyses to examine whether the rate of smears classified as BMD or >BMD increased over time (ie. for the period depicted by joinpoint analysis), unadjusted and adjusted for differences in demographic factors (ie. age, screening region, and SES) and differences in types of cytology tests used over time. As the effect of the type of primary

cytology used could differ between age groups(92) and laboratories, we included two-way interaction terms between the type of primary cytology test and age, and between the type of primary cytology test and screening region (ie. as proxy for the laboratories involved).

Similar analyses were performed to assess whether CIN and cervical cancer detection rates increased over time. Missing values were imputed with 10 multiple imputations for the type of primary cytology test used, screening region, and SES (ie. 15.2% of the cases had one or more missing values). The odds ratio (OR) was interpreted as relative risk if the prevalence of the outcome was <10%(99). The software programme SPSS (version 20) was used to perform statistical analyses.

RESULTS

Out of 6,470,400 primary cytology tests in the screening programme from 2000 to 2011, there were 16,837 CIN I (26.0 per 10,000 primary smears), 17,193 CIN II (26.6 per 10,000 primary smears), 34,380 CIN III (53.1 per 10,000 primary smears), and 2,180 cervical cancer (3.4 per 10,000 primary smears) diagnoses.

The average CIN I detection rate decreased from 2000 to 2004 by 6.3% (95% CI: -10.7, -1.6) per year (Figure 3-1), followed by an average increase of 15.7% (95% CI: 10.7, 20.9) per year from 2004 to 2009. CIN II detection rates increased from 2000 to 2005 by 3.8% (95% CI: 1.4, 6.3) annually, and from 2005 to 2009 by 11.8% (95% CI: 6.9, 16.9). Average CIN III detection rates decreased by 5.6% (95% CI: -10.6, -0.3) from 2000 to 2003, then increased by 6.6% (95% CI: 4.2, 9.1) per year from 2003 to 2009. From 2009 onward, no significant annual changes were detected for all CIN grades. There were no significant trends for cervical cancer detection over the studied period [APC of 1.2 (95% CI: -0.3, 2.8)].

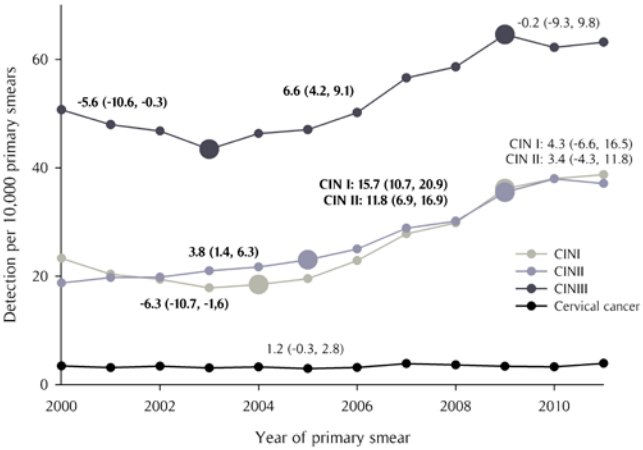

Figure 3-1. Trends in crude CIN and cervical cancer detection rates as observed within the national screening programme. Joinpoints are depicted by larger symbols. The annual percent changes are given with their 95% confidence interval for the periods depicted by joinpoint analyses (eg. CIN I detection decreased by 6.3% per year from 2000-2004 and increased by 15.7% from 2004-2009). Bold estimates are statistically significant

Table 3-1. Relative risk of CIN I, CIN II, CIN III, and cervical cancer in 2009 compared with 2003, unadjusted and adjusted for confounding factors.

	CIN I	CIN II	CIN III	Cervical cancer
Model 1	**2.03 (1.88, 2.19)**	**1.69 (1.57, 1.82)**	**1.49 (1.41, 1.57)**	1.10 (0.89, 1.35)
Model 2	**2.11 (1.95, 2.29)**	**1.79 (1.66, 1.92)**	**1.59 (1.50, 1.67)**	**1.12 (1.01, 1.25)**
Model 3	**1.90 (1.62, 2.22)**	**1.48 (1.22, 1.79)**	**1.55 (1.39, 1.73)**	1.05 (0.70, 1.59)

Odds ratio given can be interpreted as relative risk as the prevalence of the outcomes is <10%. The 95% confidence interval is given in brackets. Bold = Significant. A *p* value of <0.05 was considered to be statistically significant.
Model 1 = The effect of calendar time.
Model 2 = The effect of calendar time adjusted for demographic factors age, screening region, and SES.
Model 3 = The effect of calendar time adjusted for demographic factors age, screening region, and SES + type of primary cytology test used. Two-way interaction terms between screening region and the type of primary cytology test used, and age and the type of primary cytology test used were also included.

Between 2003 and 2009 the detection rate increased for CIN I from 17.8 to 36.1 diagnoses per 10,000 primary smears (Figure 3-1), for CIN II from 21.0 to 35.5, and for CIN III from 43.4 to 64.6. This corresponds with a 2.11 (95% CI: 1.95, 2.29), 1.79 (95% CI: 1.66, 1.92), and 1.59 (95% CI: 1.50, 1.67) times increased probability of being diagnosed with CIN I, CIN II, or CIN III, respectively, when adjusted for differences in demographic factors (Table 3-1). When also adjusting for the type of primary cytology test used, these relative risks were 1.90 (95% CI: 1.62, 2.22), 1.48 (95% CI: 1.22, 1.79), and 1.55 (95% CI: 1.39, 1.73), respectively. The probability of being diagnosed with cervical cancer did not change over time.

Table 3-2. Relative risk of CIN I, CIN II, and CIN III in 2009 compared with 2003, per age group and screening region.

		CIN I	CIN II	CIN III
Age group	29-33	**2.28 (1.72, 3.02)**	**1.50 (1.15, 1.96)**	**1.74 (1.46, 2.07)**
	34-38	**2.10 (1.54, 2.86)**	1.13 (0.79, 1.62)	**1.38 (1.10, 1.72)**
	39-43	**1.59 (1.05, 2.39)**	**1.57 (1.15, 2.15)**	**1.60 (1.22, 2.10)**
	44-48	**1.74 (1.21, 2.50)**	**1.93 (1.31, 2.84)**	**1.55 (1.17, 2.06)**
	49-53	**1.83 (1.24, 2.70)**	1.43 (0.87, 2.37)	**1.60 (1.09, 2.35)**
	54-58	**1.78 (1.09, 2.93)**	**2.33 (1.18, 4.61)**	1.07 (0.59, 1.97)
	59-63	1.04 (0.49, 2.23)	1.06 (0.41, 2.76)	0.74 (0.35, 1.57)
Screening region	1	**1.83 (1.49, 2.25)**	**1.34 (1.09, 1.64)**	**1.26 (1.08, 1.45)**
	2	**1.95 (1.42, 2.67)**	**1.54 (1.13, 2.09)**	**1.67 (1.36, 2.04)**
	3	**2.46 (1.29, 4.68)**	1.81 (0.68, 4.82)	1.79 (0.97, 3.33)
	4	**1.84 (1.43, 2.37)**	1.21 (0.92, 1.60)	**1.92 (1.60, 2.29)**
	5	**1.53 (1.01, 2.33)**	**2.06 (1.45, 2.92)**	**1.44 (1.12, 1.86)**

Adjusted for age (if applicable), screening region (if applicable), SES, and the type of primary cytology test used. Two-way interaction terms between screening region and the type of primary cytology test used (if applicable), and age and the type of primary cytology test used (if applicable) were also included. Odds ratio given can be interpreted as relative risk as the prevalence of the outcomes is <10%. The 95% confidence interval is given in brackets. Bold = Significant. A *p* value of <0.05 was considered to be statistically significant.

Trends seemed similar between most age groups for all CIN grades (Supplementary Information). When adjusting for demographic factors and the type of primary cytology test used, no significant differences in increased CIN detection rates seemed were found between age groups (ie. 95% CIs did overlap) (Table 3-2). Between screening regions, the trends differed slightly (Supplementary Information). When adjusting for demographic factors and the type of primary cytology test used, the increase in CIN III detection rates seemed to differ between the screening regions (ie. 95% CI did not overlap) (Table 3-2).

When restricting analyses to CIN detected via triage, trends of increased CIN I, CIN II, and CIN III detection occurred over time (Figure 3-2a). After adjustment for confounders, the probability of a CIN I, II, or III diagnosis was 1.94 (95% CI: 1.62, 2.32), 1.56 (95% CI: 1.25, 1.96), and 1.76 (95% CI: 1.44, 2.15) times higher in 2009 compared with 2003 (data not shown). The detection of cervical cancer via triage did not increase over time [OR of 0.69 (95% CI: 0.21, 2.27)]. The increase in CIN detected via triage was mainly explained by a 77% increased probability of primary smears being classified as BMD [OR of 1.77 (95% CI: 1.59, 1.98)] (data not shown-see Figure 3-3 for BMD detection rates per year and results of the joinpoint analysis). A BMD smear did not lead to a significant change in CIN I [OR of 1.11 (95% CI: 0.95, 1.30)], CIN II [OR of 0.84 (95% CI: 0.69, 1.03)], or CIN III detection rates [OR of 0.99 (95% CI: 0.80, 1.23)] (data not shown). Restricting analyses to CIN detected via direct referral yielded similar results. Trends of increased CIN I, II, and III detection occurred (Figure 3-2b). The probability of being diagnosed with CIN I, II, or III increased from 2003 to 2009 [ORs of 1.66 (95% CI: 1.26, 2.20), 1.33 (95% CI: 1.04, 1.71), and 1.47 (95% CI: 1.30, 1.67), respectively], but the probability of being diagnosed with cervical cancer was unaffected [OR of 1.06 (95% CI: 0.68, 1.65)] (data not shown). The increase in CIN detected via direct referral was mainly explained by a 41% increased probability of primary smears being classified as >BMD [OR of 1.41 (95% CI: 1.27, 1.58)] (data not shown – see Figure 3-3 for >BMD detection rates per year and results of the joinpoint analysis).

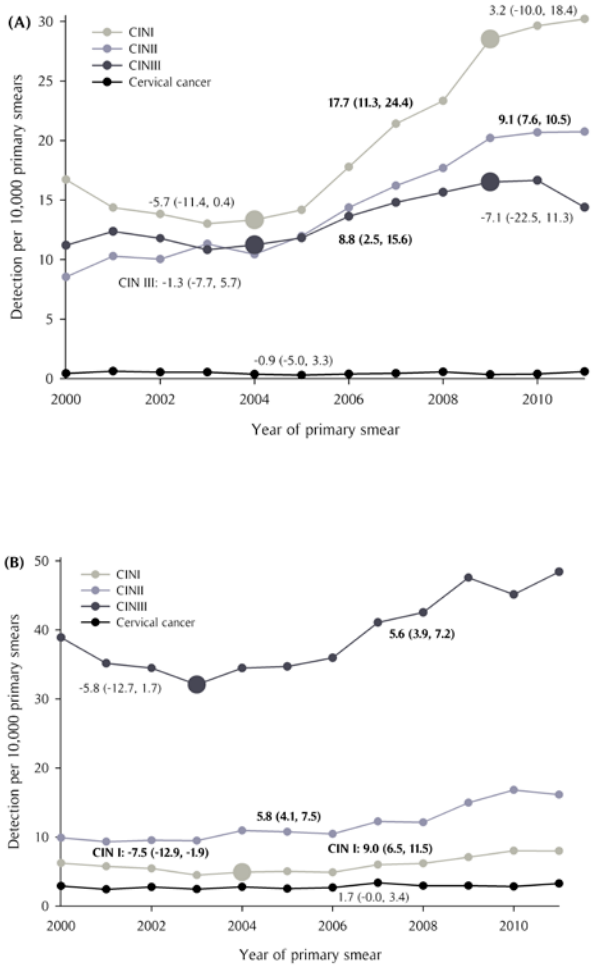

Figure 3-2. Trends in crude CIN and cervical cancer detection rates as observed within the national screening programme via (A) triage (ie. indirect referral to the gynaecologist) or via (B) direct referral to the gynaecologist. Joinpoints are depicted by larger symbols. The annual percent changes are given with their 95% confidence interval for the periods depicted by joinpoint analyses (eg. when detected via triage, CIN I detection increased by 17.7% per year from 2004-2009 while CIN III detection increased by 8.8% per year). Bold estimates are statistically significant.

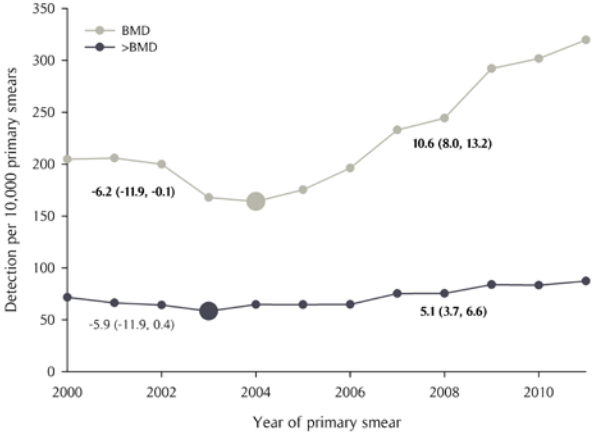

Figure 3-3. Trends in crude abnormal cytological detection rates (BMD and >BMD) as observed within the national screening programme. Joinpoints are depicted by larger symbols. The annual percent changes are given with their 95% confidence intervals for the periods depicted by joinpoint analyses (eg. BMD decreased by 6.2% per year from 2000-2004 and increased by 10.6% from 2004-2011). Women with a BMD outcome receive triage advice; women with a >BMD outcome receive direct referral advice. Bold estimates are statistically significant.

A >BMD smear did not lead to increased CIN I [OR of 1.19 (95% CI: 0.89, 1.58)], CIN II [OR of 0.93 (95% CI: 0.74, 1.17)], or CIN III detection rates [OR of 1.08 (95% CI: 0.91, 1.27)] (data not shown).

DISCUSSION

In the Dutch screening programme trends of increased detection occurred for all CIN grades from 2003–2005 to 2009. When adjusted for differences in the distribution of demographic factors, the probability of having a CIN I [ie. OR of 2.11 (95% CI: 1.95, 2.29)], CIN II [ie. OR of 1.79 (95% CI: 1.66, 1.92)], or CIN III diagnosis [ie. OR of 1.59 (95% CI: 1.50, 1.67)] was still considerably higher in 2009 as compared with 2003. When also adjusting for the type

of primary cytology test used, these relative risks were 1.90 (95% CI: 1.62, 2.22), 1.48 (95% CI: 1.22, 1.79), and 1.55 (95% CI: 1.39, 1.73), respectively. Overall, the increase in CIN detection did not seem to significantly differ between age groups and screening regions, except for the increase in CIN III detection rates among the screening regions. Trends of increased CIN detection still existed when restricting analyses to CIN detected via triage or via direct referral, mainly explained by increased probabilities of smears being classified as BMD or >BMD (ie. increased probabilities of women receiving a triage or direct referral advice). No trends in CIN detection rates were found from 2009 to 2011, and no significant trends have yet been detected for cervical cancer over time.

Although the increasing trend of CIN detection rates seems to have been temporary (ie. until 2009), CIN detection rates remained at this increased level afterwards. As a consequence, the number of treated CIN lesions increased from approximately 3500 in 2000 (70 per 10,000 primary smears) to 5000 in 2011 (100 per 10,000 primary smears) (ie. assuming that all CIN II and III lesions are treated, and 500,000 women attend screening each year).

The total increase in CIN detection is probably caused by multiple factors, including an increased risk of developing cervical cancer. However, it is unlikely that the strong increase in CIN detection rates could have been caused solely by an increased prevalence of risk factors (eg. changes in sexual behaviour, smoking, or long-term oral contraceptive use). The gradual implementation of liquid-based cytology contributed to a small extent to the CIN increase, although this effect could be underestimated if the effect of switching to liquid-based cytology differed between laboratories. The interaction between screening region (ie. proxy for the laboratories involved) and the type of primary cytology test used seems to confirm this, and could partly explain the increase in the number of cytological smears classified as BMD or >BMD, although the implementation of imaging-assisted reading could also have contributed(92). Increased attendance of previously unscreened women may also have contributed to the CIN increase, although coverage rates of the screening population did not increase from 1998 to

2006(29). Increased CIN detection rates may have been a compensation for the observed previous decrease in CIN detection. This decrease was probably the effect of a period with increased screening intensity from 1996 to 1998, when the age range eligible for screening was extended from 35-53 years to 30-60 years and several extra birth cohorts were invited(34, 109). Finally, the screening protocol may occasionally be violated by performing co-testing (ie. primary cytology combined with primary high-risk HPV-testing), which could explain a small part of the CIN increase(110).

Changes in the distribution of demographic factors (age, screening region, and SES) did not explain the increase in CIN detection. Changes in registration method or completeness are also unlikely explanations, as definitions and methods of data collection in PALGA are unchanged over the past decade. Although the interobserver agreement among pathologists is lower for CIN I and CIN II than for CIN III(111, 112), there are currently no data that suggest any changes over time. In addition, the positive predictive value of a direct referral did not significantly increase. It is unlikely that a late effect of a higher cytology cut-off (ie. due to the introduction of the CISOE-A classification in 1996(28)) has contributed to the CIN increase, as the increase was also observed in new participants of the screening programme (ie. aged 29-33). In addition, no significant differences in increased CIN detection rates have been observed between age groups.

Contrary to CIN trends, trends in cervical cancer detection rates have not yet been observed. Future trends of cervical cancer incidence may help to understand the underlying mechanism causing increased CIN detection rates. An increased sensitivity to detect, and therefore treat, CIN will eventually lead to lower cervical cancer incidence and mortality, assuming that a proportion of the extra detected CIN is clinically relevant. If not, it would only lead to increased burden and harms through overdiagnosis (and treatment) of regressive CIN lesions. An increased underlying risk of cervical cancer and, therefore, progressive CIN lesions, will probable lead to increased cervical cancer incidence and mortality.

Our results agree with those of a Danish study which found a strong increase in the number of CIN I, II, and III diagnoses(113). The Danish authors concluded that the increase was too sudden and strong to be caused by biological factors, such as an increase in HPV prevalence. In Finland the CIN detection rate decreased from 37.8 per 10,000 primary smears in 2006 to 32.5 in 2009(114). In the UK the percentage of abnormal smears in women aged 25-64 increased from 5.2% in 2004-2005 to 6.7% in 2009-2010 and decreased to 5.8% in 2011-2012(115). These variations in trends among European countries do not explain the observed trend in the Netherlands, and may indicate regional and national changes in the screening programme (eg. a new primary screening test), rather than changes in biological factors.

A limitation of our study was that we could not correct for differences in follow-up time after a positive primary smear. The individual follow-up period varied from 48 (ie. primary smear taken before April 2009) to 15 months (ie. primary smear taken in December 2011). The latter follow-up might have been too short for women with a primary BMD smear who were invited for multiple repeat cytology testing six and 18 months later. As the number of CIN I, II, and III lesions detected after these multiple repeat tests were 27.2, 16.5, and 7.4% of the total CIN I, II, and III detection (source: PALGA), low-grade CIN rates might have been somewhat underestimated in 2011. Also, we could not correct for difference in lost to follow-up (ie. women who did not comply with the given advice) over time, but its effect on the CIN increase would be negligible as the increase in CIN detection rates is mainly explained by an altered distribution of cytological classifications.

In the Netherlands, trends of increased detection were present for all CIN grades from 2003-2005 to 2009, although they were not (yet) detected for cervical cancer. The gradual implementation of liquid-based cytology caused some of the increase in CIN detection rates. This could lead to lower incidence and mortality of cervical cancer, if a proportion of the extra detected CIN are clinically relevant. If they are not, it would only lead

to increased burden and harms through overdiagnosis (and treatment) of regressive CIN lesions. Other factors, such as an increased risk of developing cervical cancer and implementation of imaging-assisted reading, could also have contributed to the increased CIN detection rates.

ACKNOWLEDGEMENTS

This study was funded by the Dutch National Institute for Public Health and the Environment (007/12 V&Z NvdV/EM). We thank C. Looman for statistical advice, the regional coordinating pathologists (RCP) for their help in collecting data, and both L.I.H. Overbeek and A.G. Siebers for linking data of the laboratories to the PALGA database.

SUPPLEMENTARY INFORMATION

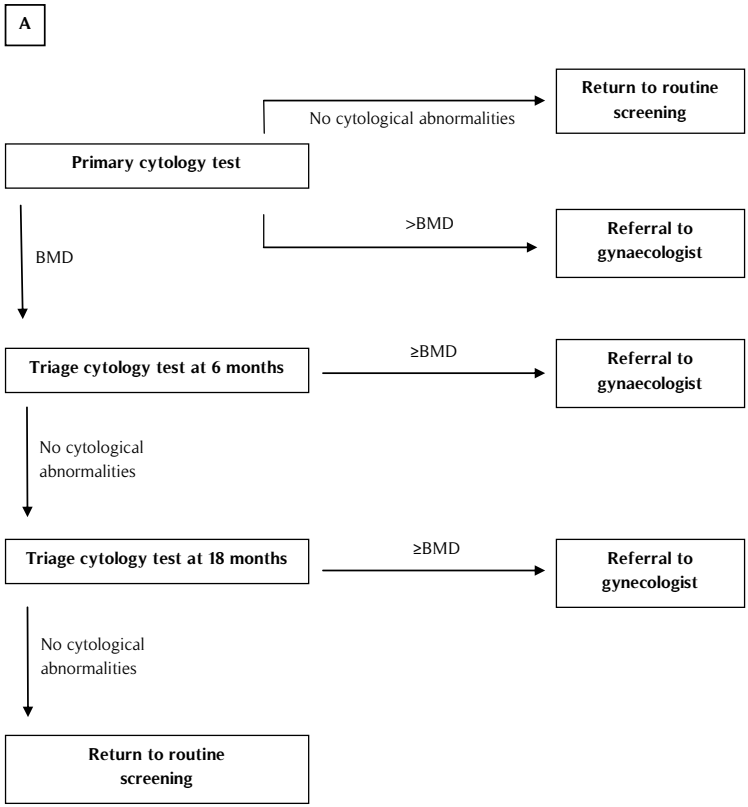

Figure 3-S1. Triage protocol consisting of triage cytology (A) without HPV testing, and (B) with HPV testing. HPV = Human papillomavirus; BMD = Borderline and mildly dyskaryotic smears.

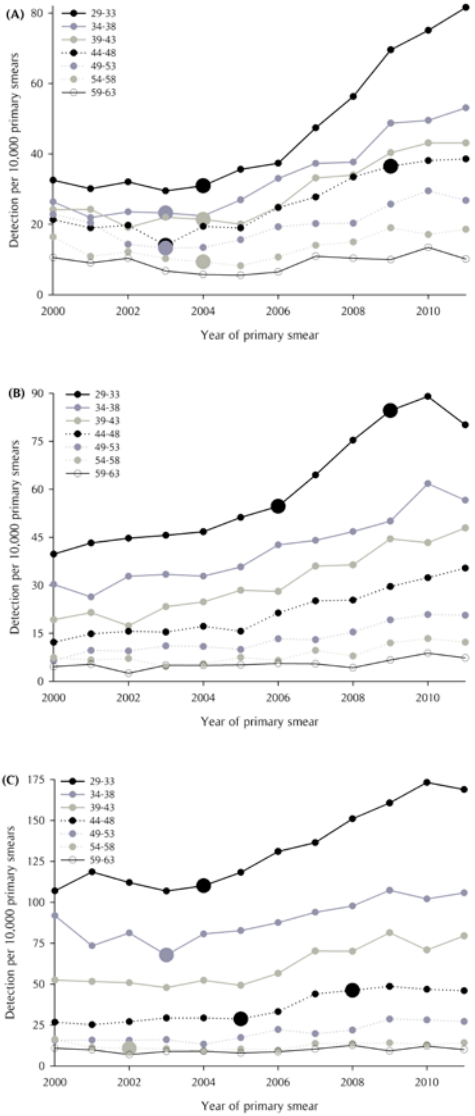

Figure 3-S2. Trends in crude (A) CIN I, (B) CIN II, and (C) CIN III detection rates as observed within the national screening programme, per age group. Joinpoints are depicted by larger symbols.

Table 3-S1. Trends of CIN I, II and III detection rates per age group as given by joinpoint analyses.

Outcome	Age group	Period	APC (95% Confidence Interval)
CIN I	29-33	2000-2004	-1.1 (-8.3, 6.8)
		2004-2011	**16.1 (13.2, 19.1)**
	34-38	2000-2003	-4.2 (-16.9, 10.5)
		2003-2011	**12.4 (9.3, 15.5)**
	39-43	2000-2004	-3.5 (-14.3, 8.8)
		2004-2011	**12.6 (7.7, 17.7)**
	44-48	2000-2003	-10.2 (-21.9, 3.3)
		2003-2009	**15.4 (9.3, 21.9)**
		2009-2011	1.9 (-16.4, 24.1)
	49-53	**2000-2003**	**-17.6 (-30.3, -2.5)**
		2003-2011	**11.0 (7.3, 14.8)**
	54-58	2000-2004	-11.8 (-24.6, 3.1)
		2004-2011	**12.3 (5.5, 19.6)**
	59-63	2000-2011	3.0 (-1.8, 7.9)
CIN II	29-33	**2000-2006**	**5.1 (3.4, 6.8)**
		2006-2009	**17.5 (8.6, 27.2)**
		2009-2011	-3.1 (-10.0, 4.3)
	34-38	**2000-2011**	**7.4 (5.9, 8.9)**
	39-43	**2000-2011**	**9.6 (8.0, 11.3)**
	44-48	**2000-2011**	**10.2 (8.6, 11.8)**
	49-53	**2000-2011**	**10.0 (7.9, 12.1)**
	54-58	**2000-2011**	**7.7 (3.8, 11.7)**
	59-63	**2000-2011**	**5.7 (1.7, 9.9)**
CIN III	29-33	2000-2004	0.1 (-5.0, 5.4)
		2004-2011	**7.0 (4.9, 9.1)**
	34-38	2000-2003	-5.4 (-14.0, 4.1)
		2003-2011	**5.3 (3.1, 7.6)**
	39-43	**2000-2011**	**5.0 (3.1, 6.9)**
	44-48	2000-2005	2.5 (-0.8, 5.9)
		2005-2008	**18.0 (4.6, 33.1)**
		2008-2011	-1.5 (-6.7, 3.9)
	49-53	**2000-2011**	**6.9 (4.5, 9.3)**
	54-58	2002-2011	-21.9 (-45.8, 12.6)
		2002-2011	**4.7 (1.3, 8.3)**
	59-63	2000-2011	2.0 (-1.0, 5.2)

Bold = Significant annual percent changes were present within that period of time.
A *p* value of <0.05 was considered to be statistically significant.
APC = Annual percent change.

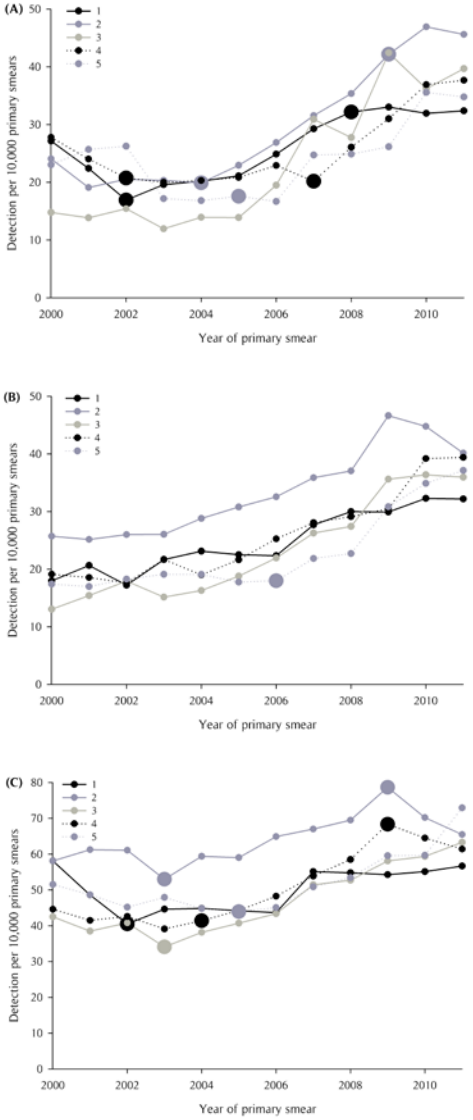

Figure 3-S3. Trends in crude (A) CIN I, (B) CIN II, and (C) CIN III detection rates as observed within the national screening programme, per screening region. Joinpoints are depicted by larger symbols.

Table 3-S2. Trends of CIN I, II and III detection rates per screening region as given by joinpoint analyses.

Outcome	Screening region	Period	APC (95% Confidence Interval)
CIN I	1	**2000-2002**	**-22.2 (-36.5, -4.6)**
		2002-2008	**11.3 (6.3, 16.6)**
		2008-2011	0.7 (-7.8, 9.9)
	2	2000-2004	-3.1 (-11.2, 5.8)
		2004-2009	**17.0 (8.7, 25.9)**
		2009-2011	4.6 (-13.0, 25.8)
	3	**2000-2011**	**13.0 (8.6, 17.5)**
	4	2000-2002	-15.1 (-30.8, 4.3)
		2002-2007	1.9 (-5.1, 9.4)
		2007-2011	**16.1 (9.5, 23.2)**
	5	2000-2005	-8.4 (-18.2, 2.6)
		2005-2011	**13.9 (5.6, 22.8)**
CIN II	1	**2000-2011**	**5.9 (4.5, 7.3)**
	2	**2000-2011**	**6.2 (4.5, 7.8)**
	3	**2000-2011**	**10.6 (8.5, 12.7)**
	4	**2000-2011**	**8.1 (6.3, 9.9)**
	5	2000-2006	0.9 (-3.0, 5.0)
		2006-2011	**15.8 (11.1, 20.8)**
CIN III	1	2000-2002	-15.5 (-30.9, 3.4)
		2002-2011	**3.9 (1.9, 5.9)**
	2	2000-2003	-3.4 (-8.0, 1.4)
		2003-2009	**5.3 (3.3, 7.5)**
		2009-2011	-6.9 (-14.8, 1.6)
	3	2000-2003	-5.4 (-12.4, 2.1)
		2003-2011	**7.8 (6.1, 9.4)**
	4	2000-2004	-2.4 (-5.3, 0.6)
		2004-2009	**10.9 (7.9, 14.0)**
		2009-2011	-3.9 (-11.3, 4.0)
	5	**2000-2005**	**-3.4 (-6.6, -0.0)**
		2005-2011	**8.4 (5.8, 11.0)**

Bold = Significant annual percent changes were present within that period of time. A *p* value of <0.05 was considered to be statistically significant.
APC = Annual percent change.

CHAPTER 4

Cervical cancer incidence after a negative cytological smear in routine screening: Comparing SurePath, ThinPrep and conventional cytology

Kirsten Rozemeijer[1], Steffie K Naber[1], Corine Penning[1], Lucy IH Overbeek[2], Caspar WN Looman[1],Inge MCM de Kok[1], Suzette M Matthijsse[1], Matejka Rebolj[3], Folkert J van Kemenade[4], and Marjolein van Ballegooijen[1]

1. Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, the Netherlands
2. PALGA, the nationwide network and registry of histo- and cytopathology in the Netherlands, Houten, the Netherlands
3. Clinical Research Centre and Department of Pathology, Copenhagen University Hospital, Hvidovre, Denmark
4. Erasmus MC, University Medical Center, Department of Pathology, Rotterdam, the Netherlands

Submitted

ABSTRACT

Background. Several studies have compared cervical intraepithelial neoplasia (CIN) II$^+$ detection rates between conventional cytology and liquid-based cytology tests SurePath and/or ThinPrep. As detecting more CIN does not necessarily mean preventing more cervical cancers, we examined the incidence of interval cancers (ie. cervical cancers diagnosed after a negative primary smear) in the Dutch cervical cancer screening programme.

Methods. All primary negative screening smears taken within this programme from 2000 to March 2012 were analyzed using the nationwide registry of histo- and cytopathology (PALGA), stratified by the type of cytology test used (ie. SurePath, ThinPrep, or conventional cytology), with a follow-up until March 2013. The 72-month cumulative incidence of interval cancers was calculated for each screening method. Cox regression analyses were performed to assess the hazard ratio (HR), adjusted for calendar time, age, screening region, and socioeconomic status. In addition, we stratified for the reason of the cervical examination prior that led to the interval cancer diagnosis (ie. clinically or screen-detected).

Results. We included 5,924,474 primary negative screening smears, resulting in 23,833,123 women-years. The 72-month cumulative interval cancer incidence was 66.8 [95% confidence interval (CI): 56.7, 78.7], 58.5 (95% CI: 54.6, 62.7), and 44.6 (95% CI: 37.8, 52.6) per 100,000 negative ThinPrep, conventional cytology, and SurePath smears, respectively. When compared to conventional cytology, the overall hazard of interval cancer was 17% lower [HR = 0.83 (95% CI: 0.70, 0.99)] for SurePath, caused by a 26% lower hazard [HR = 0.74 (95% CI: 0.59, 0.91)] of a clinically detected interval cancer. For ThinPrep, the overall hazard was on average 20% higher [HR = 1.20 (95% CI: 1.01, 1.41)], caused by a 58% higher hazard of a screen-detected interval cancer [HR = 1.58 (95% CI: 1.19, 2.08)].

Conclusions. The interval cancer rate was lowest after a negative SurePath smear and highest after a negative ThinPrep smear. This strongly suggests

that the sensitivity to detect progressive CIN lesions is highest when using SurePath and lowest when using ThinPrep as primary test method.

Keywords. Cervical cancer; Interval cancer; False-negative smear; LBC; SurePath; ThinPrep; Conventional cytology; Screening

BACKGROUND

The use of conventional cytology as primary test method has been replaced by the use of liquid-based cytology (LBC) in many countries with organized cervical cancer screening programmes, such as the UK, the Netherlands, and Denmark(92, 93). The main advantages of using LBC instead of conventional cytology are facilitating co-testing [the residual material can be tested for the presence of the human papillomavirus (HPV)](41, 44) and reducing the number of slides of unsatisfactory quality(40, 42, 46, 48, 116). In addition, it is believed that the sensitivity of LBC for detecting cervical intraepithelial neoplasia (CIN) II^+ lesions is similar to that of conventional cytology(55, 56). However, although many studies have been published comparing CIN detection between ThinPrep and conventional cytology(45, 49, 52, 57, 59, 60), only two studies have compared CIN detection between SurePath and conventional cytology(42, 58). Therefore, we compared CIN II^+ detection rates between these three types of cytology tests in our previous study, including more than six million smears taken within the Dutch cervical cancer screening programme(117). While the use of SurePath led to an 8% increased detection of CIN II^+ as compared to conventional cytology, the use of ThinPrep did not affect CIN II^+ detection rates. These results were compatible with results of other studies. In our previous study we showed that our point estimates of the CIN II^+ detection rate ratios lied within the 95% confidence interval (CI) of those from other studies comparing CIN II^+ detection rates between SurePath or ThinPrep and conventional cytology(117). Furthermore, a recently published study by Rebolj et al. confirmed our findings(64). When applying the same reading technology, they found

increased CIN II$^+$ detection rates when using SurePath, while these were unaffected when using ThinPrep.

As in the absence of screening (and associated treatment) only a fraction of CIN would progress to cervical cancer, detecting more CIN lesions is not necessarily equivalent to preventing more cervical cancers. If the increase of detected CIN lesions would be mainly regressive, this increase would not translate into altered carcinoma incidence. To assess whether the ability to detect progressive CIN lesions differs between different types of LBC tests and conventional cytology, the probability of a cervical cancer diagnosis shortly after a negative primary screening smear (ie. interval cancer) has to be compared. Whereas detecting more progressive CIN lesions will lead to fewer interval cancers, detecting more regressive CIN lesions will not. As the incidence of interval cancers is rare [6-year cumulative incidence rate of 48 per 100.000 negative smears (95% CI: 43, 54)](36), such a comparison can only be feasible if performed by an observational population-based study where a large number of smears can be evaluated.

The Netherlands is one of many countries with an organized cervical cancer screening programme where primary conventional cytology testing has been replaced by primary LBC testing. Here, organized cervical cancer screening exists since the 1980s and women aged between 30 and 60 years have been invited every five years since 1996. The screening strategy consists of primary cytology screening with cytology triage, the latter either alone or in combination with HPV testing. All cervix uteri cytological and histological tests taken inside and outside the Dutch screening programme are registered in the Dutch Pathology Register (PALGA)(94) and women can be traced even when they moved around the country. By using these data, we were able to assess any differences in interval cancer risk between different types of LBC tests (ie. SurePath and ThinPrep) and conventional cytology, thereby indicating whether there is a possible difference in sensitivity to detected progressive CIN lesions. In addition, we stratified for the reason of cervical examination prior to the interval cancer diagnosis (ie. clinically or

screen-detected). Moreover, possible differences in overdiagnosis rates were assessed by comparing CIN detection rates.

METHODS

Information on all cytological and histological examinations of the cervix uteri taken in the Netherlands between January 2000 and March 2013 were available and retrieved from PALGA. Multiple quality checks ensured the reliability of the retrieved data(29, 118). Women were identified through their birth date and the first eight letters of their (maiden) family name. This identification code enables linkage of multiple tests belonging to the same woman, allowing us to follow individual screening histories. We identified and selected episodes starting with a negative primary screening smear taken within the Dutch screening programme between January 2000 and March 2012. Women with a primary smear of unsatisfactory quality followed by a negative smear within the same episode were also selected. An episode was defined as starting with a primary test followed by one or more secondary tests in case the result was abnormal (ie. at least borderline mild dyskaryosis) or of unsatisfactory quality. Unless the follow-up of a primary test had already been completed according to guidelines, tests taken within four years following a primary test were considered as follow-up or secondary tests(38). All other tests were seen as primary tests.

Negative primary screening smears were stratified by the type of cytology test used (ie. SurePath, ThinPrep, or conventional cytology). Since PALGA does not register this routinely, regional coordinating pathologists obtained conversion dates (ie. fixed to the first date of the quarter) from individual laboratories part of their region (ie. five regions covering 44 laboratories). This information was linked to (i) the pathology laboratory involved, and (ii) the examination date as a proxy for which type of primary cytology test was used (ie. in the Netherlands, laboratories supply the general practitioners with cytology kits and thus determine the type of cytology used).

Follow-up lasted for a period of six years, or until the next episode started, or the end of the database was reached (March 2013), whichever came first. We chose for a period of six years because it covers the next screening round which is scheduled to take place five years after a negative screening smear. Histologically confirmed cervical cancer cases were identified by selecting all PALGA records that included pathology codes describing invasive cancers originating in the cervix uteri. These codes were manually checked to avoid over-counting both of non-invasive lesions and primary cancers originating elsewhere.

Since women in the Netherlands are invited for screening in the year they turn 30, 35, …, and 60, age was categorized as: 29-33, 34-38, …, and 59-63 years at the time of the negative primary cytological smear. Calendar year was also defined at the time of the negative cytological smear. The Dutch screening programme is organized by five screening organizations, each covering a geographical region (ie. screening region; North, South-West, Middle-West, South and East). Screening region was determined by a woman's place of residence at the time of the negative smear. Socioeconomic status (SES), categorized as low, middle, or high, was defined by the status score. This is an ecological variable based on the household characteristics of the four-digit postcode area where the woman was living at the time of the primary test(95). Status scores per four-digit postal code were provided by the Netherlands Institute for Social Research(96) based on 1) mean income, 2) percentage of households with a low income, 3) percentage of households with, on average, a low education, and 4) unemployment rate in 2010. Low SES corresponded with a status score lower than -1 (ie. average status score minus standard deviation), intermediate SES with a score of ≥-1 and ≤1, and high SES with a score higher than 1 (ie. average status score plus standard deviation).

Statistical analyses

Laboratories implemented LBC testing at different points in time. Therefore, follow-up (FU) and calendar time were expected to differ between the three types of cytology tests. As demographic characteristics of screened women (ie. age, screening region, and SES) probably differed between laboratories, we expected them to differ between the types of cytology tests as well. Since age, SES, screening region and calendar time were all associated with CIN and/or cervical cancer detection rates(86, 87, 97, 98), they were all potential confounding factors. We used a Pearson's chi-squared test to test whether their distributions differed between the types of cytological tests. Thus, we tested whether they were confounders or not. A p value of less than 0.05 was considered to be statistically significant.

Cumulative incidence and hazard ratio

For each type of test, the cumulative interval cancer incidence per 100,000 negative cytological screening smears was calculated. Differences in FU time were taken into account and the 95% CIs were estimated by non-parametric Kaplan-Meier product-limit estimator for log(hazard)(36) (119). Cox regression analyses were performed to compare the hazard of an interval cancer between the types of cytology tests, also taking differences in FU time into account and adjusting for confounding factors. In addition, we stratified for the reason (i.e. screen-detected when programme smear, or clinically in all other cases, which includes opportunistic screening as well as direct biopsies) of the cervical examination that led to the interval cancer diagnosis. Missing values were imputed with 10 multiple imputations for confounding factors. Time dependencies of the hazard ratios (HRs) were statistically tested by splitting the total follow-up time in two periods with a roughly equal number of cases. Subsequently, HRs were assessed for each time period. If the sum of the deviance of both sub-models was significantly lower than the deviance of the original model, the hazard ratio was time-dependent as it differed significantly between the time periods.

Difference in CIN detection rates and 72-month cumulative interval cancer incidence per 100,000 (negative) primary screening smears

We assessed the difference in CIN detection rates per 100,000 SurePath and 100,000 ThinPrep smears (ie. as compared to the CIN detection rates per 100,000 conventional cytology smears) and we compared it with the difference in the 72-month cumulative interval cancer incidence per 100,000 SurePath and ThinPrep negative smears. Information on the calculation of the difference in detection rates per 100,000 primary screening samples can be found in the Supplementary Material. The 72-month cumulative incidence rates for SurePath and ThinPrep were calculated by multiplying the distribution of the 72-month cumulative incidence rate for conventional cytology with the distribution of the adjusted HRs for SurePath and ThinPrep versus conventional cytology, as obtained by Cox regression.

RESULTS

Within the follow-up period, 1,042 interval cancers were diagnosed after 3,028,865 negative conventional cytology smears, 231 interval cancers were diagnosed after 1,303,817 negative SurePath smears, and 328 interval cancers were diagnosed after 1,591,792 negative ThinPrep smears (Table 4-1). This corresponds with the diagnoses of 7.6, 4.8, and 6.3 cervical cancers per 100,000 women-years, respectively.

Crude cumulative incidence, taking differences in the duration of FU time into account

As compared with conventional cytology, the 12-, 24-, 36-, 48-, 60- and 72-month cumulative interval cancer incidences were significantly lower for SurePath smears (Figure 4-1). When SurePath was compared with ThinPrep, all but the 24-month cumulative incidences were significantly lower for SurePath. No significant difference was detected between ThinPrep and conventional cytology. The 72-month cumulative incidence was 44.6 (95% CI: 37.8, 52.6) after 100,000 negative SurePath smears, 58.5 (95% CI: 54.6,

Table 4-1. Baseline characteristics.

	Conventional	SurePath	ThinPrep	*P* value
Negative primary smears, n	3,028,865	1,303,817	1,591,792	
Women-years at risk	13,796,018	4,835,917	5,201,188	
Interval cancers, n	1,042	231	328	<0.001
FU time				<0.001
0-1 years, n (%)	208,668 (6.9)	73,905 (5.7)	95,563 (6.0)	
1-2 years, n (%)	105,945 (3.5)	191,027 (14.7)	321,784 (20.2)	
2-3 years, n (%)	129,165 (4.3)	187,410 (14.4)	311,295 (19.6)	
3-4 years, n (%)	203,768 (6.7)	189,063 (14.5)	284,262 (17.9)	
4-5 years, n (%)	920,825 (30.4)	334,677 (25.7)	339,590 (21.3)	
5-6 years, n (%)	1,460,494 (48.2)	327,735 (25.1)	239,298 (15.0)	
Age				<0.001
29-33, n (%)	411,873 (13.6)	167,015 (12.8)	193,998 (12.2)	
34-38, n (%)	503,889 (16.6)	187,179 (14.4)	217,213 (13.6)	
39-43, n (%)	516,728 (17.1)	218,559 (16.8)	267,194 (16.8)	
44-48, n (%)	482,822 (15.9)	218,476 (16.8)	267,585 (16.8)	
49-53, n (%)	434,620 (14.3)	192,594 (14.8)	240,801 (15.1)	
54-58, n (%)	381,312 (12.6)	173,572 (13.3)	219,277 (13.8)	
59-63, n (%)	297,621 (9.8)	146,422 (11.2)	185,724 (11.7)	
Screening region				<0.001
1, n (%)	417,594 (13.8)	498,554 (38.2)	353,757 (22.2)	
2, n (%)	797,228 (26.3)	176,577 (13.5)	537,844 (33.8)	
3, n (%)	472,708 (15.6)	306,500 (23.5)	293,197 (18.4)	
4, n (%)	845,727 (27.9)	297,858 (22.8)	210,042 (13.2)	
5, n (%)	486,976 (16.1)	24,085 (1.8)	193,971 (12.2)	
Unknown, n (%)	8,632 (0.3)	243 (0.0)	2,981 (0.2)	
SES				<0.001
Low, n (%)	248,097 (8.2)	153,494 (11.8)	108,492 (6.8)	
Middle, n (%)	2,501,696 (82.6)	1,038,602 (79.7)	1,337,521 (84.0)	
High, n (%)	232,658 (7.7)	87,193 (6.7)	132,863 (8.3)	
Unknown, n (%)	46,414 (1.5)	24,528 (1.9)	12,916 (0.8)	
Calendar time				<0.001
2000-2003, n (%)	163,1520 (53.9)	162,014 (12.4)	26,499 (1.7)	
2004-2007, n (%)	118,6547 (39.2)	435,108 (33.4)	404,603 (25.4)	
2008-2012*, n (%)	210,798 (7.0)	706,695 (54.2)	116,0690 (72.9)	

*Until 31 March 2012.

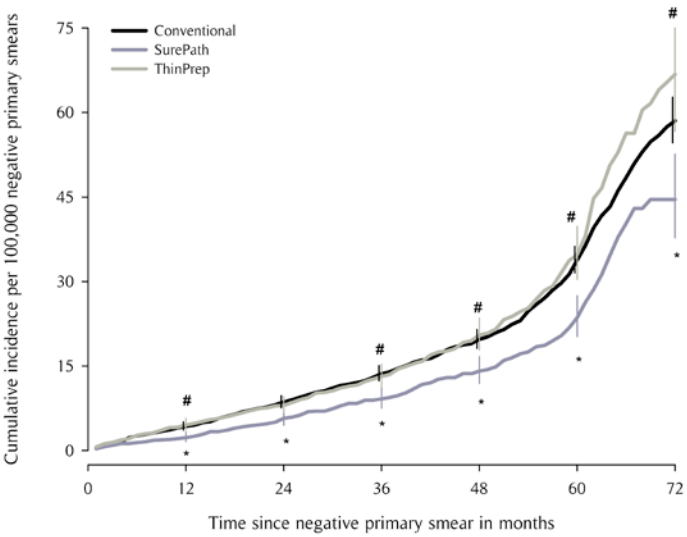

Figure 4-1. Comparing the crude cumulative interval cancer incidence per 100,000 negative primary screening smears between conventional cytology, SurePath, and ThinPrep. The 95% confidence intervals are depicted by vertical lines. A p value of < 0.05 was considered to be statistically significant.

*Significant difference between SurePath and conventional cytology.
#Significant difference between SurePath and ThinPrep.
No significant differences between ThinPrep and conventional cytology were detected.

62.7) after 100,000 negative conventional cytology smears, and 66.8 (95% CI: 56.7, 78.7) after 100,000 negative ThinPrep smears.

Confounding factors

Large and significant differences in the distributions of FU time, screening region, and calendar time were observed. For instance, almost 80% of the negative conventional cytology smears had a FU time of at least four years, while for SurePath and ThinPrep this was the case for slightly more than 50 and 35% of the negative smears. Small but significant differences were also present in the distributions of SES and age (Table 4-1). Thus, FU time,

screening region, calendar time, SES, and age were all considered confounding factors when comparing the occurrence of interval cancers between the three testing methods. Missing values were imputed for 1.6% of the primary negative smears.

Cox regression analyses of interval cancers, taking differences in the duration of FU time into account and adjusted for confounding factors

When comparing SurePath with conventional cytology, adjusted for confounding factors, the hazard of an overall interval cancer was significantly lower [HR of 0.83 (95% CI: 0.70, 0.99)] (Table 4-2). This decreased hazard was mainly caused by a decreased hazard of a clinically detected interval cancer [HR of 0.74 (95% CI: 0.59, 0.91)], the hazard of a screen-detected interval cancer was similar to that of conventional cytology [HR of 0.98

Table 4-2. Cox regression analyses of interval cancer, overall and stratified by reason of the cervical examination that led to the interval cancer diagnosis. Differences in FU were taken into account and hazard ratios are shown unadjusted and adjusted for age, SES, screening region, and calendar time.

	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
Overall		
SurePath versus CC	**0.74 (0.64, 0.85)**	**0.83 (0.70, 0.99)**
ThinPrep versus CC	1.07 (0.95, 1.22)	**1.20 (1.01, 1.41)**
SurePath versus ThinPrep	**0.68 (0.57, 0.81)**	**0.72 (0.60, 0.86)**
Clinically detected		
SurePath versus CC	**0.70 (0.59, 0.84)**	**0.74 (0.59, 0.91)**
ThinPrep versus CC	1.03 (0.88, 1.20)	1.03 (0.92, 1.14)
SurePath versus ThinPrep	**0.68 (0.55, 0.84)**	**0.75 (0.59, 0.94)**
Screen-detected		
SurePath versus CC	0.80 (0.64, 1.02)	0.98 (0.74, 1.29)
ThinPrep versus CC	1.16 (0.93, 1.44)	**1.58 (1.19, 2.08)**
SurePath versus ThinPrep	**0.68 (0.51, 0.91)**	**0.65 (0.47, 0.89)**

Bold = Significant. A p value of <0.05 was considered to be statistically significant.
CC = Conventional cytology

(95% CI: 0.74, 1.29)]. All HRs were not time-dependent (ie. the HRs did not differ within the six year time period; p = 0.449 for HR of overall interval cancer; p = 0.590 for HR of clinically detected interval cancer; p = 0.448 for HR of screen-detected interval cancer).

When comparing SurePath with ThinPrep, adjusted for confounding factors, the hazard of an overall interval cancer was significantly lower [HR of 0.72 (95% CI: 0.60, 0.86)]. This decreased hazard was both caused by a decreased hazard of a clinically detected interval cancer [HR of 0.75 (95% CI: 0.59, 0.94)] and a decreased hazard of a screen-detected interval cancer [HR of 0.65 (95% CI: 0.47, 0.89)]. All HRs were not time-dependent (p = 0.781 for HR of overall interval cancer; p = 0.661 for HR of clinically detected interval cancer; p = 0.853 for HR of screen-detected interval cancer).

When comparing ThinPrep with conventional cytology, adjusted for confounding factors, the hazard of an overall interval cancer was on average significantly higher [HR of 1.20 (95% CI: 1.01, 1.41)]. This effect seemed to differ over time (p = 0.051), with a HR of 1.03 (95% CI: 0.82, 1.29) in the first 44 months after the negative screening smear and a HR of 1.43 (95% CI: 1.12, 1.83) thereafter. This overall increased hazard was caused by an increased hazard of a screen-detected interval cancer [HR of 1.58 (95% CI: 1.19, 2.08)], the hazard of a clinically detected interval cancer was unaffected [HR of 1.03 (95% CI: 0.92, 1.14)]. These HRs were not time-dependent (p = 0.210 for HR of clinically detected interval cancer; p = 0.401 for HR of screen-detected interval cancer).

Difference in CIN detection rates and 72-month cumulative interval cancer incidence per 100,000 (negative) primary screening smears, adjusted for confounding factors

The use of SurePath versus conventional cytology as primary test method resulted in 94.4 (95% CI: +68.9, +120.6) extra CIN diagnoses per 100,000 screening smears, while the 72-month cumulative interval cancer incidence decreased by 9.7 (95% CI: -13.6, -4.8) (Table 4-3). The use of ThinPrep ver-

Table 4-3. The increase (+) or decrease (-) in CIN detection rates and 72-month cumulative interval cancer incidence per 100,000 (negative) SurePath or ThinPrep smears versus 100,000 (negative) conventional cytology smears. These numbers were corrected for differences in the distribution of follow-up time (in case of interval cancers), age, screening region, SES, and calendar time. The 95% confidence intervals are given.

	Base-case: CC*	SurePath versus CC	ThinPrep versus CC
CIN I	216.1	**+30.1** **(+18.1, +42.8)**	-3.5 (-14.3, +7.9)
CIN II	220.0	**+31.2** **(+19.0, +44.1)**	+9.4 (-2.1, +21.5)
CIN III	495.0	**+30.3** **(+12.0, +49.3)**	-12.2 (-29.6, +5.9)
Total CIN	931.0	**+94.4** **(+68.9, +120.6)**	-6.8 (-30.6, +17.6)
Interval cancer	58.5 (54.6 to 62.7)	**-9.7** **(-13.6, -4.8)**	**+11.1** **(+4.1, +20.9)**

Bold = Significant. A *p* value of <0.05 was considered to be statistically significant.
CC = Conventional cytology; CIN = Cervical intraepithelial neoplasia
*Observed CIN detection rate and 72-month cumulative interval cancer incidence per 100,000 (negative) conventional cytology smears.

sus conventional cytology showed quite different results. While the number of CIN diagnoses did not change significantly, the 72-month cumulative interval cancer incidence increased by 11.1 (95% CI: +4.1, +20.9) per 100,000 negative screening tests.

DISCUSSION

The interval cancer incidence rate among the three cytology modalities was highest for conventional cytology and lowest for SurePath. The 72-month cumulative interval cancer incidence, when also taking differences in FU time into account, was again lowest for SurePath but highest for ThinPrep. When we also adjusted for confounding factors, the overall interval cancer risk was 17% lower for SurePath in comparison to conventional cytology, which was caused by a 26% lower risk for a clinically detected interval can-

cer. The use of SurePath resulted in ten fewer interval cancers per 100,000 primary smears, while the number of CIN lesions increased by 94. The overall interval cancer risk was 20% higher for ThinPrep in comparison to conventional cytology, but it differed over time. Within the first 44 months after the negative screening test, the risks were comparable. Thereafter, the risk was 43% higher when using ThinPrep. Both the overall increased risk and the difference over time is due to a 58% higher risk for a screen-detected interval cancer when using ThinPrep (ie. the recommended Dutch screening interval is five years and thus 60 months). The use of ThinPrep resulted in 11 additional interval cancers per 100,000 negative primary smears.

In our previous study, using the same data as in our current study, we showed that the detection of CIN II$^+$ was increased by using SurePath, while it was unaffected by using ThinPrep(120). As the use of SurePath resulted in decreased interval cancer rates, this indicates that at least part of the extra detected CIN lesions were progressive. As the use of ThinPrep resulted in increased interval cancers, this indicates that although similar numbers of CIN lesions were detected fewer of them were progressive. These differences in sensitivity to detect progressive CIN lesions are most likely caused by differences between the techniques of the LBC tests, such as: the extent of fixation, the technique of taking a representative sample from the vial, and the retention of the brush (ie. the collecting device) in the fluid(53, 54). Studies have shown that retaining the brush, as is done when using SurePath, is associated with an increased cell yield as compared to rinsing and discarding the brush, as is done when using ThinPrep(100, 101). Therefore, the ability to transfer abnormal cells may differ between SurePath and ThinPrep, possibly resulting in a difference in sensitivity. Moreover, as the thoroughness of rinsing the brush in the vial with preservative fluid, and therefore the cell yield(101), might differ between clinicians, it is possible for the interobserver agreement of ThinPrep to be lower than that of SurePath. Thus, the differences in sensitivity to detect progressive lesions might differ more between laboratories using ThinPrep than SurePath. In addition, the interobserver agreement of conventional cytology is prob-

ably lower than that of SurePath and ThinPrep as the quality of the cell transfer, and therefore the quality of the conventional cytology smear, can differ between clinicians. Thus, how much there is to gain when replacing conventional cytology by another primary test method could differ between clinicians and therefore, between laboratories. In conclusion, the difference in sensitivity to detect progressive lesions between SurePath and ThinPrep versus conventional cytology might differ between clinicians and therefore, between laboratories.

If we take the most recent distribution of cytology testing into account (ie. obtained in the first quarter of 2012), each year 325,000 women are screened by ThinPrep and therefore prone to 36 extra interval cancers while 175,000 women are screened by SurePath and therefore protected against 17 of them. However, as the risk for a clinically detected interval cancer was unaffected by ThinPrep, the extra found interval cancers are probably diagnosed in an early stage where clinical symptoms are absent or rarely present. Therefore, the negative effects of using ThinPrep are probably less pronounced on the interval cancer mortality than on the incidence. On the other hand, the protective effect of SurePath on interval cancer mortality is probably more pronounced than on the incidence as we found that SurePath was primarily protective for clinically detected interval cancers. As no data on mortality were available, we were not able to estimate the effects of LBC implementation on interval cancer mortality. While our results may be less relevant for the future of the Dutch cervical cancer screening programme (ie. it is expected that from 2016 onward, primary cytology screening will be replaced by primary HPV screening(67)), they certainly can be relevant to other countries with organized primary cytology screening programmes who have switched to using SurePath and/or ThinPrep or will switch in the near future.

An important drawback of cervical cancer screening is the diagnosis and treatment of CIN lesions that would never have progressed to clinical cervical cancer in the absence of screening. With the use of SurePath, the

prevention of ten extra interval cancers is accompanied by the detection of 94 extra CIN lesions. Therefore, when including CIN diagnoses in the definition, the increased sensitivity to detect progressive CIN lesions also led to increased overdiagnosis rates according to the following definition: number of extra diagnoses with screening divided by total number of diagnoses in a population with screening. As both the number of excess diagnoses and the number of diagnoses in the population increased by 84, the ratio increased and therefore the overdiagnosis rate. How the use of SurePath will affect the number of QALYs gained, and therefore the effectiveness of the Dutch cervical cancer screening programme, is beyond the scope of this article.

This study is the first that compared interval cancer rates between two different types of LBC tests and conventional cytology, thereby examining differences in sensitivity to detect progressive CIN lesions. In addition, we examined the drawbacks of LBC implementation by comparing indicators of overdiagnosis.

Our study has some limitations, many of which are linked to the fact that this was not a randomized trial. First, differences in the distribution of demographic factors were present since the study did not take place in a controlled setting. Although we were able to correct for FU time, age, screening region, SES, and calendar time, we were not able to correct for other potential confounders such as a woman's screening history. This may have resulted in biased effect estimates if the distribution differed between the types of cytology tests. The fact that no large differences in age and SES were found was reassuring. Second, as the risk of a screen-detected interval cancer was increased for ThinPrep and the FU time was too short for most women with a negative primary ThinPrep smear to be invited for the next screening round, this would mean that the overall risk of an interval cancer was underestimated for ThinPrep. However, as we corrected for differences in the distribution of calendar time, which is linked to FU time, we believe this underestimation would be minimal. Indeed, when we only selected women attending screening within six years after a negative primary ThinPrep smear (data not shown), we found a slightly higher risk of

a screen-detected interval cancer than before, which seems to confirm our statement that we minimally underestimated our screen-detected and overall risk for ThinPrep. As no difference in screen-detected interval cancer risk was found between SurePath and conventional cytology, (shortage in) FU time had no influence on the overall interval cancer risk for SurePath. Third, we were not able to correct for the use of automated reading, although the possible influence would be small given that automated reading has only been introduced in relatively few Dutch laboratories. Moreover, as multiple studies demonstrated that CIN II$^+$ detection were unaffected(61, 62, 64) or slightly decreased(63) by adding automated assisted reading to the use of ThinPrep or SurePath, we do not believe our estimated were significantly biased. Fourth, as we did not have an unique identification code [ie. identification code was based on the first eight letters of the (maiden) family name and birth date], tests belonging to different women may have been allocated to a single woman (so-called fusions). However, we think it is unlikely that these fusions would be correlated with the type of cytology test used. Fifth, we did not have individual data on which type of primary test was used. Therefore, we used date of the primary cytological smear and conversion date (ie. fixed to the first date of the quarter) of the laboratory examining the smear to deduce which type of cytology test was used. This means that negative primary screening smears taken during this quarter may have been misclassified, leading to a slight underestimation of the effects. Sixth, we were not able to censor follow-up for death and migration. However, since mortality rates are relatively low at screening ages(121), and demography was relatively similar between the groups, we do not expect this has biased our results. Seventh, we did not correct for possible learning curve effects, since the aim of our study was to examine the effect of using SurePath and ThinPrep in routine practice, which also includes a possible learning effect.

In conclusion, the use of SurePath versus the use of conventional cytology as primary test method was associated with lower interval cancer rates, strongly suggesting that the sensitivity to detect progressive CIN lesions is higher. The use of ThinPrep versus the use of conventional cytology was

associated with higher interval cancer rates, strongly suggesting that the sensitivity to detect progressive CIN lesions is lower. Our findings should urge reconsideration for the assumed lack of difference in test characteristics between LBC and conventional cytology.

ACKNOWLEDGEMENTS

This study was funded by the Dutch National Institute for Public Health and the Environment (007/12 V&Z NvdV/EM). We thank the regional coordinating pathologists (RCP) for their help in collecting data and A.G. Siebers for linking data of the laboratories to the PALGA database.

SUPPLEMENTARY INFORMATION

In this Supplementary Material, we describe the material and methods that were used to determine the difference in CIN detection rates per 100,000 primary screening smears.

Selecting data from PALGA: CIN lesions

We identified primary smears taken within the national cervical cancer screening programme between January 2000 and December 2011. As data until March 2013 were available to us, a minimum duration of 15 months follow-up was ensured. Histologically confirmed CIN lesions were identified by selecting all PALGA records that included corresponding pathology codes. Subsequently, lesions were linked to the type of cytology test used. Age, screening region, SES, and calendar year at the time of the primary smear were assessed in similar ways as in the main analysis.

Statistical analyses: CIN lesions

We compared CIN detection rates per 100,000 SurePath and 100,000 ThinPrep smears with CIN detection rates per 100,000 conventional cytology smears. As confounding factors are present, comparing observed CIN detection rates was not sufficient. Therefore, we calculated CIN detection rates

Table 4-S1. Factors to calculate the adjusted CIN detection rates for SurePath and ThinPrep. Given factors are odds ratios comparing SurePath and ThinPrep with conventional cytology, adjusted for age, screening region, SES and calendar time.

	SurePath versus CC (95% CI)	ThinPrep versus CC (95% CI)
CIN I	**1.14 (1.08, 1.20)**	0.98 (0.93, 1.04)
CIN II	**1.14 (1.09, 1.20)**	1.04 (0.99, 1.10)
CIN III	**1.06 (1.02, 1.10)**	0.98 (0.94, 1.01)
Total CIN	**1.10 (1.07, 1.13)**	0.99 (0.97, 1.02)

Bold = Significant. A *p* value of <0.05 was considered to be statistically significant.
CC = Conventional cytology; CIN = Cervical intraepithelial neoplasia.

per 100,000 SurePath and ThinPrep smears by multiplying the observed CIN detection rates per 100,000 conventional cytology smears with the adjusted odds ratios for SurePath and ThinPrep versus conventional cytology, as obtained in our previous study (Table 4-S1(117)). These odds ratios were adjusted for differences in the distribution of age, screening region, SES, and calendar time between the three cytology tests.

CHAPTER 5

The role of pre-invasive disease in overdiagnosis: a microsimulation study comparing mass screening for breast cancer and cervical cancer

Paula A van Luijt[1], Kirsten Rozemeijer[1], Steffie K Naber[1], Eveline AM Heijnsdijk[1], Joost van Rosmalen[2], Marjolein van Ballegooijen[1], and Harry J de Koning[1]

1. Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, the Netherlands
2. Erasmus MC, University Medical Center, Department of Biostatistics, Rotterdam, the Netherlands

Submitted

ABSTRACT

Background. Early detection of cancer prevents cancer deaths if an effective treatment is available for the early stage at detection. A drawback of mass screening is overdiagnosis. The potential harm of overdiagnosis depends on its frequency and the consequences of diagnosis and treatment. There is much debate on the topic of overdiagnosis in screening for breast cancer, but less so on overdiagnosis in screening for cervical cancer.

Methods. We estimated overdiagnosis rates by microsimulation for breast cancer screening and for cervical cancer screening, using a cohort run of women born in 1982 with lifelong follow-up. Overdiagnosis estimates were made analogous to two definitions formed by the UK 2012 breast screening review. Pre-invasive disease was included in both definitions.

Results. Whereas breast cancer screening averted 1.3% of invasive cancers and 21% of related deaths, cervical cancer screening averted 55% of cervical cancers and 59% of related deaths. Breast cancer overdiagnosis rate was estimated at 2.5%, when including pre-invasive disease. Cervical cancer overdiagnosis rate was 74.8%, when including pre-invasive disease. For women of all ages in breast cancer screening, an excess of 207 diagnoses per 100,000 women was found with screening, compared to an excess of 3,999 diagnoses per 100,000 women in cervical cancer screening.

Conclusions. For breast cancer, the frequency of overdiagnosis in screening is relatively low, but its consequences are evident. For cervical cancer, the frequency of overdiagnosis in screening is high, because of detection of pre-invasive disease, but the consequences per case are relatively small due to less invasive treatment. This illustrates that it is necessary to present overdiagnosis in relation to disease stage and consequences.

Key words. Breast cancer screening; Cervical cancer screening; Pre-invasive disease; Overdiagnosis; Microsimulation

BACKGROUND

The purpose of cancer screening is to prevent cancer death by detecting a (pre)cancerous lesion early, when treatment is still a viable option and more effective(122). Screening advances the diagnosis of disease to an earlier age, resulting in a higher incidence just after the initiation of screening. After the upper age limit of screening, the incidence rate will drop(123).

Breast cancer screening detects invasive breast cancer and ductal carcinoma in situ (DCIS), which are both considered a cancer diagnosis(124). The number of breast cancer diagnoses has increased since the introduction of screening, due to both lead time and changes in underlying risk. In a mature cervical cancer screening programme, the screen detection of invasive cancer is rare due to the higher frequency of detection of precursor lesions, thus altering the natural history of those lesions that are progressive. Screening for cervical cancer mostly detects cervical intraepithelial neoplasia (CIN), which is not regarded as a cancer diagnosis. The incidence rate of cervical cancer had been decreasing prior to the introduction of screening and has continued to decrease since the introduction of screening(125, 126). Given the fact that screening for colorectal cancer will also focus on detecting precancerous lesions, it is expected that the incidence of invasive colorectal cancer will decrease after screening is introduced(126).

The downside of early detection is the possibility of detecting abnormalities that would never have become clinically apparent in the absence of screening(127). This may occur because abnormalities spontaneously regress, as is described for cervical cancer(128-130), or because they remain indolent, as is described for breast cancer(131, 132). The detection of such an abnormality is called overdiagnosis, and most overdiagnoses lead to overtreatment. Overdiagnosis has been the topic of a fierce debate in breast cancer screening(127). In cervical cancer screening, overdiagnosis is usually quantified as a lack of specificity for clinically significant disease.

The impact of overdiagnosis depends on its frequency and its consequences. In breast cancer screening, the overdiagnosis rate is relatively low(127). In cervical cancer screening, the overdiagnosis rate is usually not established. The consequence of overdiagnosis is unnecessary treatment which is inherently harmful. The consequences of overdiagnosis in breast cancer screening are more severe than those in overdiagnosed non-progressive CIN lesions in cervical cancer screening. For an individual patient the name of the disease carries weight as well.

We aimed to exemplify the impact of overdiagnosis by comparing these two screening programmes, which have been implemented for several decades in the Netherlands; for cervical cancer since 1985 (ie. currently, women aged 30-60 are invited every five years), and for breast cancer since 1990 (ie. currently, women aged 50-74 are invited every two years)(133, 134).

In literature, estimates of overdiagnosis in breast cancer screening vary from 4-54%(127, 135-138). The proper estimate of overdiagnosis has been the topic of many debates and the cause of many misunderstandings. We chose to use the definitions put forward by the UK independent review panel(139).

This is the first study aiming to compare different screening programmes by addressing the potential amount and composition of overdiagnosed cases in the same overdiagnosis framework. As more types of cancer will become eligible for screening, we hope that in the future balanced reports will elucidate the impact of any cancer screening on the advanced cancer rate and disease-specific mortality, while also publishing the properly estimated extent of overdiagnosis.

METHODS

The MISCAN model is a microsimulation model. This means the model simulates all individual life histories in a population. We have a model for

breast cancer screening (MISCAN-Breast) and a model for cervical cancer screening (MISCAN-Cervix)(127, 140). In order to obtain a representative population, the models are fitted with a birth table and a life table. Each life history has its own probability of developing a (pre)cancerous lesion. In MISCAN-Breast, this probability is determined by fitting model parameters hazard, onset, and incidence, to data on incidence without screening from the Dutch Cancer Registry(141). In MISCAN-Cervix, the model is fitted to incidence data from the Dutch Cancer Registry and data on detection obtained from PALGA(141). From each state the disease may progress to the next stage by a semi-Markov progression model (Figure 5-1). In MISCAN-Breast, screening is implemented in the model using data on gradual roll-out, attendance rate, and re-attendance rate in the Dutch screening programme. Sensitivity, stage distribution, and distribution of sojourn-time were estimated by fitting these parameters to data on incidence and stage distribution in situations with (1991-2010) and without screening (1990).

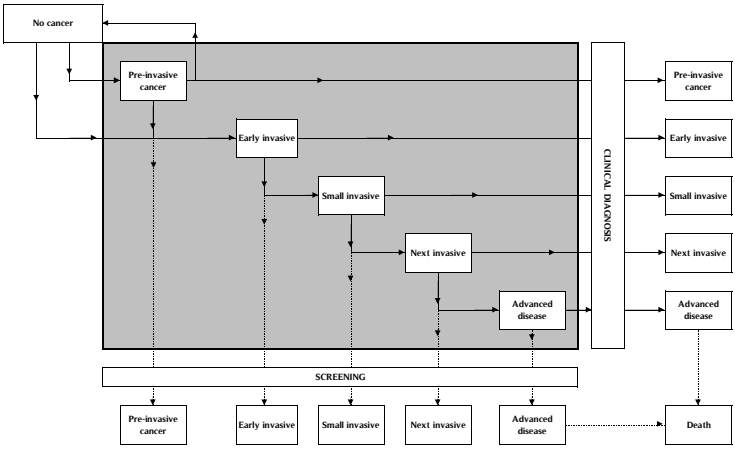

Figure 5-1. Progression in the MISCAN model. Every woman starts at the top left, where she has no cancer. From there she may progress through the different stages of cancer. If the cancer is detected by screening, the woman moves to the bottom of the graph (screen-detected). If the cancer is clinically detected she moves to the far right of the graph (clinically detected).

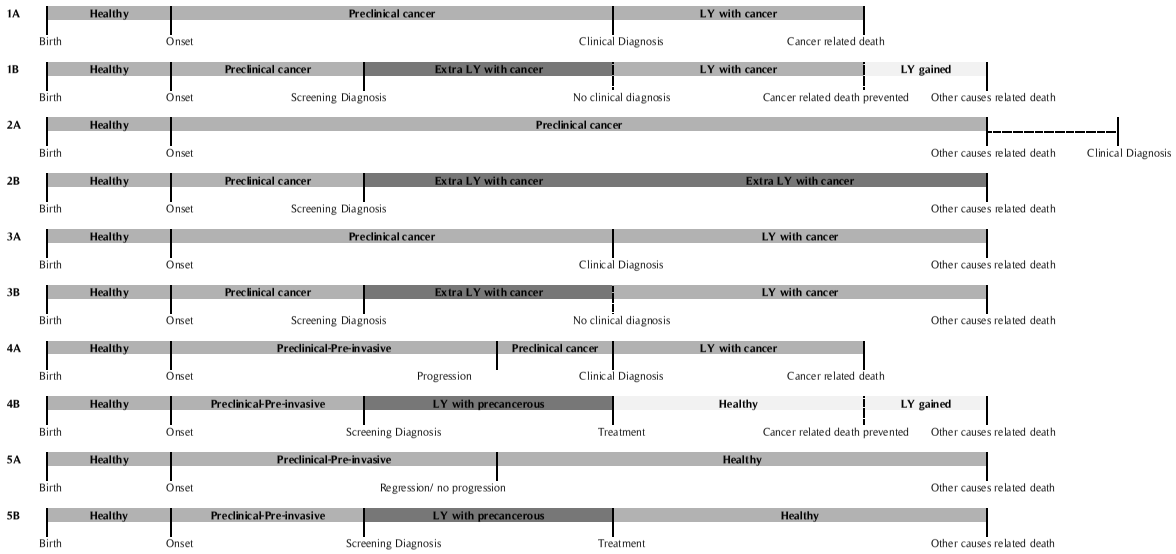

Figure 5-2. Life histories of women affected differently by screening. The numbers indicate different women, each of them having a life history (A) without screening, and (B) with screening. Dark grey cells represent the negative effects of screening, light grey cells represent the positive effects of screening. LY = Life-years.

MISCAN-Breast assumes a 1.4% annual percentage change in underlying incidence(142). Mortality reduction in the breast cancer model is based on the results of the Swedish trials(143). The mortality reduction in the cervical cancer model is based on observational data, provided by the Dutch Cancer Registry and PALGA over the period 1998 to 2007.

The impact of screening on an individual life history is illustrated by Figure 5-2, in which there are five different women, and each has a scenario with and without screening. Dark grey cells represent the negative effects of screening, while light grey cells represent the positive effects of screening. Woman number 1 will benefit from screening. In situation 1A, there is no mass screening. She will have an onset of cancer; this cancer will grow and develop up to the point when she develops symptoms. The cancer will be clinically diagnosed and she will die from this cancer. In situation 1B, there is mass screening. The woman will have the same onset and the same pre-clinical disease phase, but now mass screening will detect her cancer before she develops symptoms. Therefore, the disease is in a less advanced state and treatment is successful. She has gained life-years and will die of other causes than cancer. Woman number 2 does not benefit from screening. Like woman number 1, she has an onset of cancer, followed by a preclinical disease phase. This phase however, would extend beyond her lifespan. She will never be diagnosed with cancer in the situation without screening (2A). In the situation with screening (2B) the cancer will be detected by screening and she will be treated accordingly. She will still die at the same time, but now she has lost several quality-adjusted life-years (QALYs) due to the fact that she had a cancer diagnosed. Woman number 3 develops a pre-invasive disease that will progress to a clinically detected cancer, but she will not die from this cancer (3A). She will also not gain any life-years by screening (3B). Woman number 4 has a type of cancer with an obvious pre-invasive precursor state (ie. CIN in cervical cancer). In this case the preclinical phase is divided into two phases, one with preclinical pre-invasive disease and one with preclinical cancer. The preclinical-pre-invasive state will progress to preclinical cancer, which becomes clinically detected and leads to cancer-

related death in the situation without screening (4A). When this woman is screened (4B) while her disease is in the pre-invasive phase and her condition is detected, she may be cured completely. Thus, cancer was prevented and she benefits from screening. Woman number 5 does not benefit from screening; she has a preclinical-pre-invasive disease that will not progress, or may even regress back to normal without screening (5A). Screening (5B) will give her a diagnosis of pre-invasive disease, but she will not gain any life-years.

MISCAN-Breast assumes a regression rate of 2%, and a progression rate of 11%, for DCIS(144). MISCAN-Cervix has six different disease paths, five assume regression, and assumes progression from onset to invasive disease. Each woman has an age-dependent probability of ending up in one of the disease paths.

We performed a cohort run using our breast cancer and cervical cancer models. The cohort consisted of 10,000,000 women, all born in 1982. The year 1982 was chosen so all women were 30 years and invited for cervical cancer screening in 2012, the most recent year with complete data. The number of simulated women alive in 2012 was also chosen as the denominator to convert raw data to rates. Between 2012 and 2032 (the year all women are invited to breast cancer screening for the first time) approximately 2% of the simulated women die of all-cause mortality (including cancer). Follow-up was completed for ages 30-100 years. Output measures were: number of diagnoses during entire follow-up in the situation without screening and in the situation with screening, and the number of diagnoses during the screening ages in the situation without screening and in the situation with screening. All results are presented per 100,000 women aged 30 in 2012 and stratified by precancer (DCIS for breast cancer and CIN grades I, II, and III for cervical cancer) and invasive cancer.

To estimate overdiagnosis we used the definitions set forward by the UK Independent review panel, which are: (i) "from the population perspective,

the proportion of all cancers ever diagnosed in women invited to screening that are overdiagnosed", and (ii) "from the perspective of a woman invited to screening, the probability that a cancer diagnosed during the screening period represents overdiagnosis"(139). To be able to address all diagnoses in the programme, we extend the definitions above to include pre-invasive lesions, such as CIN I, II, and III.

These definitions translate into the following calculations:

1. From the population perspective: Number of extra diagnoses with screening/Total number of diagnoses in a population with screening. For the purpose of comparison we used ages 30-100 years. No significant amount of cancers occur before the age of 30.
2. From an individual perspective: Number of extra diagnoses with screening/Total number of diagnoses in women of screening age. For breast cancer screening this age range is 49-75 years. For cervical cancer screening this age range is 29-60 years, but we used 29-64 years because the diagnostic process in cervical cancer screening may take some time due to follow-up.

The number of extra diagnoses with screening is the difference between the total number of diagnoses in women aged 0-100 without screening and the total number of diagnoses in women aged 0-100 with screening. When we consider overdiagnosis, we included pre-invasive disease. If we had not included pre-invasive disease, overdiagnosis measures would not have applied.

RESULTS

All results are given per 100,000 women aged 30 in 2012. The model predicted 1,669 cervical neoplasia diagnoses (Table 5-1) and 13,210 breast cancer diagnoses per 100,000 women without screening (Table 5-2). Screening added 3,999 cervical neoplasia diagnoses and 207 breast cancer diagnoses. The extra cervical diagnoses were 4,920 extra CIN lesions, which cannot be

Table 5-1. Number of cervical cancer cases per 100,000 women aged 30 years in 2012 in the situations with versus without screening.

	Without screening		With screening	
	Clinically detected	Clinically and screen-detected	Screen-detected	Clinically detected
Diagnoses during entire life (ages 30-100 years)				
CIN I	0	1,138	1,138	0
CIN II	0	1,189	1,189	0
CIN III	0	2,593	2,593	0
Cervical cancer	1,669	748*	117	632
Total diagnoses	1,669	5,668	5,037	632
Cervical cancer death	644	266		
Diagnoses during screening (ages 30-64 years)				
CIN I	0	1,138	1,138	0
CIN II	0	1,189	1,189	0
CIN III	0	2,593	2,593	0
Cervical cancer	1,138	424	117	307
Total diagnoses	1,138	5,344	5,037	307

CIN = Cervical intraepithelial neoplasia
*The total of screen-detected and clinically detected cervical cancers do not add up as a result of rounding

clinically detected, and 921 (-55.2%) fewer cervical cancer diagnoses. The extra breast cancer diagnoses were the result of 376 extra DCIS diagnoses (+61.7%), and 169 fewer invasive cancers (-1.3%).

From a population perspective, the breast cancer overdiagnosis rate was estimated to be 1.5%. The cervical cancer overdiagnosis rate varied from 70.6%, when including all CIN and invasive diagnoses, to 50.0%, when including only CIN III and invasive disease. From the individual perspective, the breast cancer overdiagnosis rate was 2.5%. Cervical cancer overdiagnosis rate varied from 74.8%, when including all CIN and invasive diagnoses,

Table 5-2. Number of breast cancer cases per 100,000 women aged 30 years in 2012 in the situations with versus without screening.

	Without screening	With screening		
	Clinically detected	Clinically and screen-detected	Screen-detected	Clinically detected
Diagnoses during entire life (ages 30-100 years)				
DCIS	610	985	531	454
Breast cancer	12,600	12,431	3,523	8,908
Total diagnoses	13,210	13,417	4,055	9,362
Breast cancer death	4,637	3,668		
Diagnoses during screening (ages 49-75 years)				
DCIS	364	746	531	215
Breast cancer	7,286	7,447	3,523	3,924
Total diagnoses	7,650	8,194	4,055	4,139

DCIS = Ductal carcinoma in situ

Table 5-3. Cervical cancer and breast cancer overdiagnosis rates including different (pre-) stadia and stratified by perspective.

(Pre)stadia included as overdiagnosis	Population perspective: Excess diagnoses / Lifetime diagnoses	Individual perspective: Excess diagnoses / Screening age diagnoses
CIN I, II, and III + cervical cancer	70.6%	74.8%
CIN II, and III + cervical cancer	63.2%	68.0%
CIN III + cervical cancer	50.0%	55.4%
DCIS + breast cancer	1.5%	2.5%

Overdiagnosis rates were calculated by using the numbers of cases per 100,000 women given in Tables 5-1 and 5-2. Excess diagnoses were calculated by subtracting all diagnoses in women aged 30-100 in the situation without screening from all diagnoses in women aged 30-100 in the situation with screening.
Diagnoses are considered to be lifetime diagnoses when detected in women aged 30-100. Diagnoses are considered to be screening age diagnoses when detected in women aged 30-64 for cervical cancer, and in women aged 49-75 for breast cancer.
CIN = Cervical intraepithelial neoplasia; DCIS = Ductal carcinoma in situ.

to 55.4%, when including only CIN III and cervical cancer (Table 5-3). The number of cervical cancer deaths reduced by 59% in the situation with screening (ie. from 644 to 266 deaths per 100,000 women). The number of breast cancer deaths reduced by 21% (ie. from 4,637 to 3,668 deaths per 100,000 women).

DISCUSSION

Both breast cancer screening and cervical cancer screening prevent cancer-specific mortality at the expense of overdiagnosis, when the detected pre-invasive lesions are included in its definition.

The burden of overdiagnosis depends on its frequency and its consequences. Although the overdiagnosis frequency is high in cervical cancer screening relative to breast cancer screening, the impact is limited because treatment is minimally invasive. For CIN I most often no treatment is necessary, and for CIN II or CIN III a loop excision or conisation may be done in an out-patient setting(145). These procedures have relatively limited risks, and no apparent cosmetic impact. However, cold knife conisation and large loop excision may be associated with preterm delivery, low birth weight, caesarean section, and preterm rupture of the membranes in future pregnancies(146-148). For breast cancer screening, the frequency is low relative to cervical cancer screening, but the impact is higher due to more invasive treatment. The treatment of DCIS consists of lumpectomy or even mastectomy, in some cases followed by radiation therapy(149, 150). The risks of these treatments include (rare) standard operation risks (haemorrhage or infection) and the risk of generalized anaesthesia. Additionally, the cosmetic result of these procedures has significant impact(150). The perception of the individual also needs to be taken into account. The information provided with each diagnosis, whether it is cancer or pre-invasive disease, is crucial to the impact of this event. The decision to count a diagnosis as overdiagnosis has to be related to its severity, treatment warranted, and on the impact of the information provided at diagnosis.

Our estimates for overdiagnosis of breast cancer were different from those previously published using the MISCAN model. This is a direct result of using cohort runs instead of simulating a realistic population. If we run our model with a population aged 0-100, we obtain an overdiagnosis rate directly comparable to that of De Gelder et al.(127). This rate is: from a population perspective, for all diagnoses 4.6%; and from an individual perspective, for all diagnoses 8.1%(127, 135-138). For cervical cancer no comparable numbers were published.

Our analysis for cervical cancer screening was performed on the current situation (ie. primary conventional cytology testing with cytology triage) in the Netherlands. However, over the last years most laboratories have added a test to detect human papillomavirus (HPV) infections in the triage phase which slightly increases CIN I and CIN II detection(39). In addition, most laboratories processing primary screening tests have switched from using conventional cytology to liquid-based cytology tests SurePath and ThinPrep. Rozemeijer et al. have shown that CIN II^{+} detection rates increase by using SurePath, while they are unaffected by using ThinPrep(117). This means that overdiagnosis rates are probably somewhat higher in the current Dutch situation than estimated in our study. Also, it is expected that from 2016 onward, cervical cancer screening will be further modified in the Dutch programme. Primary cytology will then be replaced by primary HPV screening with cytology triage. Furthermore, women will be invited for screening five times in their lifetime(151). On the one hand there is a risk of increasing overdiagnosis by detecting disease at yet an earlier stage, on the other hand overdiagnosis may decrease due to less screening examinations in a lifetime.

Internationally breast cancer screening programmes vary. In the USA many women are annually screened for breast cancer from the age of 40, despite the recommendation made by the United States Preventive Services Task Force (USPSTF)(152), while in the UK women are invited from ages 50-70 every three years (ie. the programme is currently extending to include

women aged 47-73)(153). In addition, cervical cancer screening programmes also vary. In some countries, such as Finland, the cervical cancer screening programme is comparable to that of the Netherlands and women are invited seven times a lifetime, while in the UK, Sweden and Denmark women are invited for screening 12, 13, and 13 times a lifetime starting at the ages of 25, 23, and 23, respectively(154-156). Therefore, our estimated overdiagnosis rates for both breast and cervical cancer screening may be different in other countries. They are expected to increase for screening programmes with increasing number of screening examinations and for screening programmes with an earlier onset of screening. With each added screening round overdiagnosis may increase, also for younger women. In fact more non-progressive CIN is found in younger women than in older women(21).

Looking towards the future, if we were to analyse the data for colorectal cancer screening we would expect results in between those of breast cancer and cervical cancer screening, depending on the screening test being used. Faecal occult blood tests, especially the older guaiac tests but also the newer immunochemical tests, have a lower sensitivity for early, pre-invasive disease than endoscopy. The most sensitive test will find more pre-invasive disease, which will need less invasive treatment but also more often would not have developed into clinical disease. Thus, the frequency of overdiagnosis would be high but the per case consequences would be low.

In order to compare the two programmes, which offer screening at different ages, we performed a cohort run. Although this results in a lifetime estimate of harms and benefits, it remains hypothetical as the homogeneity of a cohort never resembles a real population. Mathematical modelling requires assumptions made in the model on natural history of cancer. The mean duration of sojourn time and the probability of progression are interchangeable in the model, the assumptions used have influenced the overdiagnosis estimate(157).

We have compared the burden of screening for two of the population cancer screening programmes currently in use in the Netherlands. For breast cancer, overdiagnosis estimates are relatively low, but the consequences for overdiagnosed women are significant. On a population level, these consequences are, however, quite small. For cervical cancer, overdiagnosis estimates of pre-invasive disease are high, but the consequences are relatively small due to less invasive treatment. Screening eligible women should not only be informed about the potential benefits from screening, but also about the probability, and its potential harms, from being overdiagnosed.

ACKNOWLEDGEMENTS

This study was partly funded by the Dutch National Institute for Public Health and the Environment (007/12 V&Z NvdV/EM).

PART II

INCREASING ATTENDANCE BY OFFERING HPV SELF-SAMPLING TO NON-ATTENDEES

CHAPTER 6

How many cervical cancer cases can potentially be prevented using a more sensitive screening test at young age?

Inge MCM de Kok[1], Joost van Rosmalen[1,2], Kirsten Rozemeijer[1], Corine Penning[1], and Marjolein van Ballegooijen[1]

1. Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, the Netherlands
2. Erasmus MC, University Medical Center, Department of Biostatistics, Rotterdam, the Netherlands

Int J Cancer. 2014 Jan 15;134(2):460-6.
Reprinted with kind permission from John Wiley and Sons.

ABSTRACT

Background. The human papillomavirus (HPV) DNA test has higher sensitivity than cytology for cervical cancer screening. Therefore, cervical cancer cases that are missed by cytology could potentially be identified if we use primary HPV testing. Studies showed that HPV screening is the preferred primary test at age 35 and over. Given the high prevalence of harmless HPV infections, the use of HPV testing at younger age is less obvious. The number of cancers in young age is often mentioned to indicate the possible benefits of a more sensitive test. We actually estimated the proportion of those cases that is potentially preventable in the Netherlands by the use of a more sensitive screening test at first screening age 30, given that the more sensitive test is used at age 35 and over.

Methods. We analyzed the screening history of women diagnosed with cervical cancer in the period 2004 to March 2009, using data from the Dutch Pathology Register.

Results. Only 15-30% (two to four cases per 100,000 women) of the cases was preceded by negative cytology under age 35 and therefore could have been prevented by a more sensitive test at age 30.

Conclusions. The lower the screening coverage and the shorter the screening interval in those screened at young age, the smaller the gain of a more sensitive test. So, as long as the current screening pattern is not changed, the majority of the cervical cancer cases at young age would still occur even when applying a more sensitive test at the younger ages.

Keywords. Cervical cancer; Screening; Human papillomavirus DNA test; Cytology; Young women

BACKGROUND

Mass screening for cervical cancer has been operational for women from age 30 years onward in the Netherlands since 1996(34). Nevertheless, still approximately 700 women are diagnosed with cervical cancer annually in the Netherlands, of which 160 cases in the age group 30-39(158). Some of these cases were diagnosed based on symptoms while unscreened or were detected by their first screening moment(74). In these women, lack of screening test sensitivity did not play a role. However, some cases had screening but tested negative(74), which might well be due to lack of sensitivity of, in the current situation, cytological screening. Meta-analyses and pooled analyses have established that human papillomavirus (HPV) DNA tests have higher sensitivity than cytology for detecting high-grade, clinically relevant, cervical intraepithelial (CIN) lesions(83, 159). Therefore, cervical cancer cases that are missed by cytology could have been identified if we had used primary HPV DNA testing as the primary screening test. Several studies showed that HPV screening is the preferred primary test at age 35 and over(84, 85, 160-164). HPV testing is less specific than cytology because it can detect harmless HPV infections. As these are considerably more prevalent in young women, the net benefits of using HPV testing under age 35 is less obvious(165, 166). Frequent screening at a young age detects many transient infections, cytological and histological abnormalities and every screening round adds to unnecessary triage, overdiagnosis, and overtreatment as a consequence. Conversely, the number of life-years gained per extra death prevented in young women is high, and the prevention of cervical cancer incidence and mortality at young age is valuable.

The goal of this study is to estimate how many cases are potentially preventable in the Netherlands by the use of a more sensitive screening test before the age of 35, assuming that this more sensitive test is (already) offered after the age of 35 years. To this end, we analyzed the screening history of women with cervical cancer diagnosed in the period January 2004 to March 2009,

using data from the Dutch national pathology file that includes cervical cytology and histological results(94).

METHODS

To analyse the screening history of women diagnosed with cervical cancer in the period January 2004 to March 2009, we used data from the Dutch nationwide network and registry of histo- and cytopathology (PALGA). The registration began in the late 1970s and achieved practically complete coverage of pathology laboratories in 1990. The network registers cervical smears and biopsies taken in all settings: primary smears within the screening programme, opportunistic screening, smears and biopsies taken because of medical complaints, and secondary (diagnostic and follow-up) tests, regardless of whether they are taken or read by public or private healthcare providers and laboratories.

In the PALGA network, women are identified through their birth date and the first four letters of their (maiden) family name. This identification code enabled linkage of multiple tests belonging to the same woman, allowing us to follow the individual screening and disease histories. The problem of false identity matches was avoided by excluding women with 0.5% most common maiden names in the analyses(167). The implicit assumption used here is that the screening histories and the commonness of surnames are not associated.

Registered screening histories were organized into screening episodes. An episode starts with a primary test (a smear or a biopsy) followed by secondary tests in case this test was abnormal (at least borderline dyskaryosis) or of unsatisfactory quality. Follow-up or secondary tests were defined as the tests made within four years following the primary test, unless the follow-up of this primary smear had already been completed according to the guidelines (eg. with two consecutive negative smears after a borderline dyskaryotic

smear, or three consecutive negative smears after histologically confirmed CIN). All other tests were considered to be primary tests.

Analyses

Analyses were performed for different age groups, that is, 30-35 years, 30-40 years, and ≥30 years. The definition of age is the age of a woman at the end of the calendar year. This is because in the Netherlands women are invited to participate in screening during the year they become 30, 35, 40, 45, 50, 55 or 60 years. Therefore, some women are still 29, 34, and so forth, at the time they are being screened. Age groups were based on the fact that women below the age of 30 are not invited to participate into screening. In general, since women are screened with an interval of five years, the age group 30-35 years will show the interval cases that are missed at the first screening round. The age group 30-40 years will show the interval cases and cases picked up at the second screening round, that are missed at the first screening round.

We wanted to estimate how many cases were possibly the result of a failure of cytological screening [ie. a normal cytological result while a cancer (precursor) is present] performed before the age of 35, and could possibly have been prevented if a more sensitive test had been used, by analyzing the screening history of women with cervical cancer. To this end, we first identified cervical cancer cases diagnosed at ages 30-35 years, 30-40 years, and ≥30 years by selecting all PALGA records that included pathology codes for cervical cancer between 2004 and March 2009. For these women, we reviewed the free text of all histology reports in PALGA. Age of diagnosis was defined as the age at the date of the first registered pathology code for cervical cancer in PALGA.

Of all cervical cancer cases diagnosed in women aged 30-35 years, we first excluded those diagnosed at the first episode in a woman's lifetime (Figure 6-1). If the cancer was diagnosed at the first episode, it could not have been found earlier using a more sensitive test. Second, we analyzed whether the primary cytology test of the episode prior to the episode in

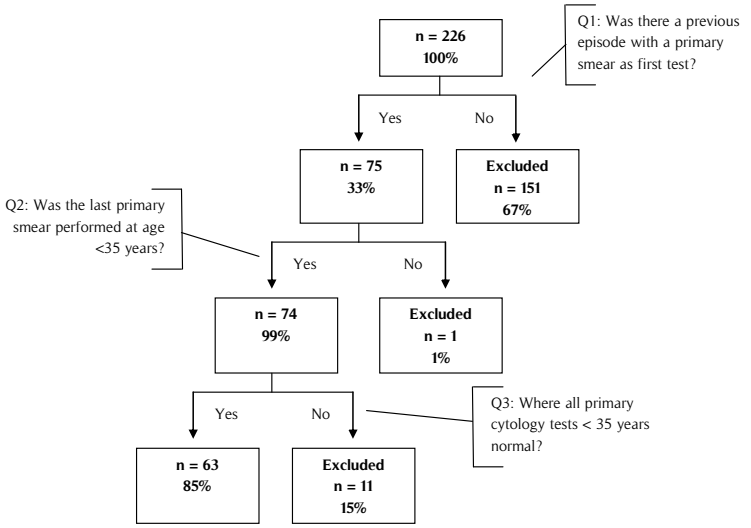

Figure 6-1. Analyses of screening history of women diagnosed with cervical cancer at age 30-35 years, in the period January 2004 to March 2009 (n = Number of cases; Q = Question). 'Excluded' means that these cases could not have been prevented using a more sensitive test at age 30.

which the cancer was diagnosed, was performed before the age of 35 years (Figure 6-1). If the previous episode started after the age of 35 years, we assumed that these cases were not the result of a false negative cytological test performed before the age of 35 years and therefore could not potentially be diagnosed earlier by HPV screening before the age of 35 years. Third, of the cases diagnosed in the second or later episode in a woman's lifetime and with a previous episode starting before the age of 35 years, we excluded those cases with an abnormal cytological result before the age of 35 (Figure 6-1). For these cases, the cancer (precursor lesion) was not missed by the primary cytological test, but by the follow up, like false-negative follow-up or no (complete) follow-up. If we assume a similar triage test in case of primary HPV testing and primary cytological testing, these cases would also not have been diagnosed by follow up after a positive HPV test. If there were

only negative cytological results before the age of 35 years, these cases (or their precursors) were possibly missed due to a false-negative test. These cases represent the cases that are potentially preventable by the more sensitive HPV test before age 35.

To explore the probability that the lesion was missed by the cytological test, we analyzed what the interval was between the date of first test of the episode in which the cancer was diagnosed and the date of the primary test of the previous episode. The shorter the interval, the more likely the lesion was already present at the time of screening. In other words, if the interval was short after a normal cytological result, the chance that the (pre)cancer had been missed by the preceding cytological test would be larger than if the interval was longer. Furthermore, we did the same subsequent analyses for women diagnosed with cervical cancer aged 30-40 years, and for women aged ≥30 years.

We calculated the percentage of cases that could potentially be prevented per age group, by dividing the number of cervical cancer cases that could possibly be prevented per age group and interval, by the number of cervical cancer cases diagnosed per age group. We calculated 95% binomial proportion confidence intervals using an exact method. We estimated the number of cervical cancer cases that could potentially be prevented per year and age group in the Netherlands using a more sensitive test before the age of 35 years, by multiplying the number of cancers diagnosed in 2011 in the Netherlands Cancer Registry(158) with the calculated percentage of cases that could be prevented. We calculated the crude rate of the number of cancer cases per 100,000 women using the Dutch female population in the specific age groups in the period 2011-2012(168).

Table 6-1. Analyses of screening history of women diagnosed with cervical cancer in the period January 2004 to March 2009, by age at cancer diagnosis (following the decision tree as presented in Figure 6-1 for age group 30-35 years)

	Age of cervical cancer diagnoses		
	30-35 years	30-40 years	≥30 years
Total cervical cancer cases, n	226	583	2,426
(a) Cases with a previous episode with a primary smear, n (% of total)	75 (33)	297 (51)	1,377 (57)
(b) Of (a), the cases of which the last primary smear was performed <35 years, n (% of total)	74 (33)	207 (36)	318 (13)
(c) Of (b), the cases with only normal smear results <35 years, n (% of total)	63 (28)	160 (27)	219 (9)

RESULTS

After the exclusion of the 0.5% most common surnames, 226 diagnosed cervical cancer cases were still registered in PALGA in the period January 2004 to March 2009 in women aged 30-35 years, 583 in women aged 30-40 years, and 2,426 in women aged ≥30 years (Figure 6-1, Table 6-1). Of the women aged 30-35 years, 67% had never been screened before the cancer was diagnosed. Of the remaining 33% (75 cases), 74 cases (99%) had the last cytological test before the episode in which the cancer was diagnosed, before the age of 35 years. In 63 of these 74 cases (85%), no cytological abnormality was found at that last screening test. In 83% (n = 52) of these cases, the interval between the first test of the episode in which cervical cancer was diagnosed and the previous primary smear was less than ten years, in 63% (n = 40) it was less than five years (Table 6-2).

In women aged 30-40 years, 583 cervical cancer cases were diagnosed (Table 6-1). Of these women, 49% were never screened before the cancer was diagnosed. Of the remaining 51% (297 cases), 207 cases (70%) had the

Table 6-2. Number of cervical cancer cases of which all previous cytology was under age 35 and negative in the period January 2004 to March 2009, by age at cancer diagnosis and the interval between the date of first test of the episode in which the cancer was diagnosed and the date of the primary test of the last previous episode.

Age of cervical cancer diagnoses	≤5 years, n (%)	≤10 years, n (%)	≤15 years, n (%)	Total, n (%)*
30-35 years	40 (63)	52 (83)	61 (97)	63 (100)
30-40 years	82 (51)	131 (82)	151 (94)	160 (100)
≥30 years	82 (37)	142 (65)	184 (84)	219 (100)

*Row (c) in Table 6-1.

last cytological test before the episode in which the cancer was diagnosed, before the age of 35 years. In 160 of these 207 cases (77%), no cytological abnormality was found at the last screening test performed before the cancer was diagnosed. In 82% (n = 131) of these cases, the interval between the first test of the episode in which cervical cancer was diagnosed and the previous primary smear was less than ten years, in 51% (n = 82) it was less than five years (Table 6-2).

In women aged ≥30 years, 2,426 cervical cancer cases were diagnosed (Table 6-1). Of these women, 43% were never screened before the cancer was diagnosed. Of the remaining 57% (1,377 cases), 318 cases (23%) had the last cytological test before the episode in which the cancer was diagnosed, before the age of 35 years. In 219 of these 318 cases (69%), no cytological abnormality was found at the last screening test performed before the cancer was diagnosed. In 65% (n = 142) of these cases, the interval between the first test of the episode in which cervical cancer was diagnosed and the previous primary smear was less than ten years, in 37% (n = 82) it was less than five years (Table 6-2).

In Table 6-3 and Figure 6-2, we show how many cases are potentially preventable per year and age group, by the use of a more sensitive screening test before the age of 35 years. If we assume that the lesion was already present at

Table 6-3. Nationwide annual number of cancer cases per 100,000 women (Source: Netherlands Cancer Registry) of which the last previous cytology was under age 35 and negative, and that therefore could potentially have been prevented with a more sensitive test at age 30, by age at cancer diagnosis and interval since that last negative smear.

	Age at cancer diagnosis		
	30-35 years	30-40 years	≥30 years
Total incidence in 2011			
Cases	73	164	711
Cases per 100,000 women	14.5	15.6	12.9
Preventable incidence in 2011			
Interval ≤5 years			
% cases (95% CI)	18% (12.7, 22.7)	14% (11.2, 16.9)	3% (2.7, 4.1)
Cases	13	23	21
Cases per 100,000 women	2.6	2.2	0.4
Interval ≤10 years			
% cases (95% CI)	23% (17.5, 28.5)	22% (19.1, 25.9)	6% (4.9, 6.8)
Cases	17	36	43
Cases per 100,000 women	3.3	3.4	0.8
Interval ≤15 years			
% cases (95% CI)	27% (21.2, 32.8)	26% (22.3, 29.5)	8% (6.5, 8.6)
Cases	20	43	57
Cases per 100,000 women	3.9	4.1	1.0
All intervals			
% cases (95% CI)	28% (22.0, 33.7)	27% (23.8, 31.1)	9% (7.9, 10.2)
Cases	20	44	64
Cases per 100,000 women	4.1	4.2	1.2

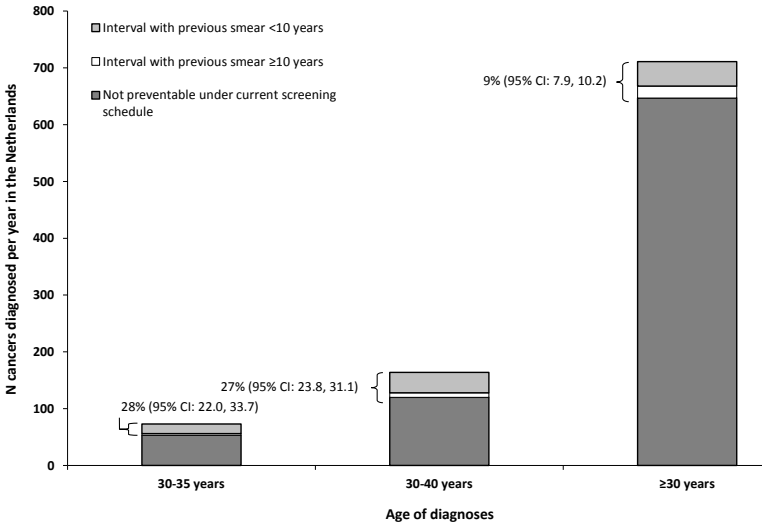

Figure 6-2. Maximum proportion of cervical cancer cases of which the last previous cytology was under age 35 and negative, and that therefore could potentially have been prevented with a more sensitive test at age 30, by age group and interval since that last negative smear (Table 6-3).

the screening moment less than five years before the cancer was diagnosed, 2.6 cancer cases per 100,000 women (18%, 95% CI: 12.7, 22.7) per year are missed and potentially preventable in age group 30-35 years. For the age groups 30-40 and ≥30 years, 2.2 (14%) and 0.4 (3%) cases per 100,000 women are potentially preventable per year. If we assume that the lesion was already present at the screening moment less than ten years before the cancer was diagnosed, 3.3 cases per 100,000 women (23%, 95% CI: 17.5, 28.5) are annually potentially preventable in age group 30-35 years, 3.4 (22%, 95% CI: 19.1, 25.9) in age group 30-40 years and 0.8 (6%, 95% CI: 4.9, 6.8) in age group ≥30 years. For a screening moment ≤15 years before the cancer was diagnosed, these figures are 3.5 (27%, 95% CI: 21.2, 32.8), 4.1 (26%, 95% CI: 22.3, 29.5), and 1.0 (8%, 95% CI: 6.5, 8.6) cases per 100,000 women per year, for age groups 30-35, 30-40, and ≥30 years, respectively. If all screening moments before the cancer was diagnosed are considered, regardless of the

interval of the previous test and the cancer diagnoses, these figures are 4.1 (28%, 95% CI: 22.0, 33.7), 4.2 (27%, 95% CI: 23.8, 31.1), and 1.2 (9%, 95% CI: 7.9, 10.2) cases per 100,00 women per year for age groups 30-35, 30-40, and ≥30 years, respectively.

DISCUSSION

We showed that only 9% of the women aged ≥30 years who were diagnosed with cervical cancer in recent years could potentially have benefited from a more sensitive test (than the Pap smear) also at age 30 instead of having this more sensitive test only from age 35 onward. The other women with cervical cancer either did not attend cervical screening before the cancer diagnoses (43%), attended screening only after age 35 (44%), or already had a positive test before age 35 using the less sensitive Pap smear (4%). Women in the latter categories would not have benefited from having had a more sensitive test offered instead of a Pap smear. Depending on the age at diagnosis and the interval since that last negative Pap smear considered, the percentage of cases that could potentially have been avoided by offering a more sensitive test varies from 3% in women aged ≥30 years with an interval of less than five years, to 28% in all women aged 30-35 years. This shows the maximum possible benefit when performing HPV testing at all screening ages (including age 30) compared to only at the ages over 35 years. Furthermore, the achieved benefit must be balanced with the relatively high positivity-rate of HPV testing at age 30 compared to the rate in older ages.

There are two explanations for a normal cytological test result in the episode before cervical cancer is diagnosed: (i) there is no neoplasia present yet, (ii) the lesion is present but missed by the test. The interval between the first manifestation of CIN I and the development of clinical cervical cancer is estimated to be on average 15 years(21, 169-171). Therefore, explanation two becomes less probable when the normal test result took place more than ten years before cervical cancer was diagnosed, and even less so in case of an

interval of more than 15 years. For the tests performed less than ten years before cancer diagnosis, and especially for the tests performed less than five years before diagnosis, however, explanation two seems most likely, meaning that a more sensitive test in principle could have detected the present lesions. We showed that, if we assume that the lesion was already present at the screening moment less than ten years before the cancer was diagnosed, 3.3 (23%) cases per 100,000 women have been annually missed and are potentially preventable in age group 30-35 years, 3.4 (22%) in age group 30-40 years and 0.8 (6%) in age group ≥30 years. Note that these figures are a maximum and one would have to assume 100% sensitivity to prevent them all by the alternative test. Although the HPV test is more sensitive than the Pap test, it is not 100% sensitive(159).

The presented proportions of cervical cancer cases that can be prevented by a more sensitive test at age 30 are representative for the Netherlands. The question is to what extent they are applicable for other countries. This depends on several factors, such as the coverage of the screening programme at young age, the frequency of screening, and the sensitivity of the cytological test in the specific situation. For example, in the United States, the percentage of young women that undergo screening is higher than in the Netherlands(172), therefore the probability that women had a previous primary smear before the cancer was diagnosed is higher. This would results in a higher number of potentially preventable cases. Conversely, as the frequency of screening is also higher in the United States compared to the Netherlands(172), the probability that a lesion was missed in a pre-invasive stage is lower. This would result in a lower number of potentially preventable cases.

Test-positive rates are up to three times higher with HPV DNA screening than with cytology(173). HPV DNA screening consequently also causes more false-positive tests, that is, positive screening tests without underlying (clinically relevant) CIN or cervical cancer. These women will have unnecessary follow-up, such as repeat testing or a referral for colposcopy, which have

important psychosocial consequences(174). Moreover, most women with high-grade CIN are of reproductive age, which will be picked up by HPV DNA screening applied <35 years of age. Treatment of CIN is associated with a small but real increase in risk of pregnancy-related morbidity, such as premature delivery (<37 weeks), low birth weight (<2500 g) and preterm prelabour rupture of membranes(26, 146, 175). These disadvantages must be weighed against the advantages of the extra cervical cancer cases that are potentially preventable by using a more sensitive screening test.

In the Netherlands, all smears and histologically diagnosed cervical cancer cases are in principle registered in PALGA. However, the total number of women aged ≥30 years with an incident cervical cancer in 2004-2008 in PALGA differs from the number published by the Cancer Registry (+14%). Nevertheless, PALGA is the only comprehensive registry in the Netherlands that links the cancer cases with their screening history.

In the future, studying interval cancers by stage would be useful, to get more insight into the cause of the occurrence of the cancer after a negative screening test. For example, low stage tumours are more likely newly developed tumours (not missed by the screening test). Unfortunately, tumour stage has not yet been included in the PALGA registry. To determine the total effect of applying a more sensitive screening test at younger ages, next to the effect on the cervical cancer incidence, the effect on cervical cancer mortality needs to be estimated. Currently, cervical cancer mortality is not included in the PALGA registry. In the future, we intend to link cervical cancer mortality to screening history to improve the evaluation of different cervical cancer screening methods. We showed that 43% of the cervical cancer patients did not attend screening before they were diagnosed. Some of these cases could possibly have been prevented by offering an HPV self-test to non-attending women. It has been shown that offering an HPV self-test to non-responding women increases the screening participation rate with approximately 6%(73). Furthermore, 25% of the women who never had been screened before, did attend the HPV self-test(176).

HPV vaccination was recently introduced in the Dutch National Immunisation Programme (NIP). In 2009, a catch-up programme was started for girls born between 1993 and 1996. Vaccination of the first NIP cohort (ie. girls born between January 1, 1997 and August 31, 1997) started in April 2010. As all women born after 1992 have been or will be invited for HPV vaccination, vaccinated women will reach the initial screening age (ie. 30 years) in 2023. As it is expected that the cervical cancer incidence will decrease as a result of the vaccination, in the future, we need to reconsider the cost-effectiveness of screening, and especially HPV screening. Determining the effectiveness of screening programmes for vaccinated women will require a separate analysis that can be performed when more is known about the long-term effectiveness of HPV vaccination.

In conclusion, we showed that the total incidence of cervical cancer at young screening age is not a good indicator for the potential gain when applying a more sensitive test at young age. For the Dutch situation, only 15-30% of all cases diagnosed in age group 30-40 years were preceded by a negative cytology based screening test and could therefore potentially have been prevented by a more sensitive screening test at age 30, for example, an HPV test. As long as the current screening pattern is not changed, the majority of the cervical cancer cases at young age would still occur even when applying a more sensitive screening test at the younger screening ages. The lower the screening coverage and the shorter the screening interval in those screened at young age, the lower the preventable proportion.

ACKNOWLEDGEMENTS

This study was funded by the Dutch National Institute for Public Health and the Environment (3022/07 DG MS/CvB/NvN).

CHAPTER 7.1

Offering self-sampling to non-attendees of organized primary HPV screening: When do harms outweigh the benefits?

Kirsten Rozemeijer[1], Inge MCM de Kok[1], Steffie K Naber, Folkert J van Kemenade[2], Corine Penning[1], Joost van Rosmalen[1,3], and Marjolein van Ballegooijen[1]

1. Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, the Netherlands
2. Erasmus MC, University Medical Center, Department of Pathology, Rotterdam, the Netherlands
3. Erasmus MC, University Medical Center, Department of Biostatistics, Rotterdam, the Netherlands

Cancer Epidemiol Biomarkers Prev. 2015 May;24(5):773-82.

ABSTRACT

Background. Human papillomavirus (HPV) self-sampling might be a promising method to increase effectiveness of primary HPV screening programmes when offered to non-attendees. However, effectiveness could decrease if regular attendees "switch" to self-sampling, because self-sampling test characteristics may be inferior. We examined under which conditions the harms would outweigh the benefits.

Methods. The MISCAN-cervix model was used to estimate quality-adjusted life-years (QALYs) gained and costs of offering HPV self-sampling to non-attendees. We varied the relative CIN II$^+$ sensitivity and specificity (self-sampling versus regular sampling), extra attendance, risk of extra attendees, and the switching percentage.

Results. Without switching, offering self-sampling is (cost-)effective under every studied condition. If the attendance due to self-sampling increases by ≥6 percentage points, higher primary underlying risk women (unscreened women who will never attend regular screening) attend and the relative CIN II$^+$ sensitivity and specificity are ≥0.95, it is (cost-)effective to offer self-sampling to non-attendees, even if all regular attendees switch. If the relative sensitivity decreases to 0.90 combined with either the absence of higher primary underlying risk women or a 3 percentage points extra attendance, QALYs are lost when more than 20 to 30% of the regular attendees switch.

Conclusions. Offering self-sampling will gain health effects if the relative CIN II$^+$ sensitivity is ≥0.95, unscreened attendees are recruited, and the total attendance increases by ≥6 percentage points. Otherwise, switching of regular attendees may decrease the total effectiveness of the programme.

Impact. Self-sampling needs to be implemented with great care and advantages of office-based sampling need to be emphasized to prevent switching.

Keywords. Self-sampling; Human papillomavirus; Cervical cancer screening; Microsimulation; Effectiveness; Non-attendees

BACKGROUND

In the Netherlands, cervical cancer incidence and mortality have decreased in the past decades to 6.5 and 1.3 per 100,000 women-years (age-adjusted to the World Population) in 2012(177). The introduction and improvements of the screening programme played a considerable role in this decrease(178). Since 1996, Dutch women of ages 30 to 60 years are invited to attend cervical cancer screening every five years. From 2016 onward, primary cytology will be replaced by primary high-risk human papillomavirus (HPV) testing(179), because the sensitivity for detecting CIN II^{+} lesions is higher when using HPV testing(65) and HPV testing can be performed on self-samples(69, 70). Although the current screening participation rate ranges from 65 to almost 70% [source: Dutch Network and National Database for Pathology (PALGA)(94)], it has been estimated that more than half of the invasive cervical cancers occur in women who did not participate in the previous six years. Moreover, some of these women had never been screened at all(74). This shows that addressing non-attendance can increase the effectiveness of the programme considerably.

Self-sampling devices, with which women can collect cervical cells themselves, have been developed recently. As self-sampling is more woman-friendly and less time consuming than letting a clinician, general practitioner, or midwife collect cervical cells, it probably increases participation in screening. Indeed, the Dutch PROHTECT study has shown that offering a self-sampling HPV test to non-attendees of the programme increased the overall screening participation rate by about 6 percentage points(73, 180). However, the gain in effectiveness of the programme [ie. gain in quality-adjusted life-years (QALYs)] probably not only depends on the increase in attendance, but also on the test characteristics of HPV self-sampling and on the ability to target higher risk non-attendees. It is likely that *unscreened women* (who were invited at least once but were never screened) have higher risks on developing cervical cancer than *one-time non-attendees* (who missed the last screening round, but have been screened in the past). Nevertheless,

including any non-attendee will probably increase the effectiveness of the programme. However, "switching" of *regular attendees* from office-based to self-sampling could, given a loss in detection (ie. more loss to follow-up, possible lower sensitivity), result in a decrease of the effectiveness of the programme (ie. losing QALYs). In other words, the QALYs gained by attracting non-attendees could be annulled by the QALYs lost by switching of regular attendees. It is unclear at which level of switching this will happen.

The aim of this study is to examine the effectiveness of offering HPV self-sampling to non-attendees of a primary HPV screening programme. We modelled effects of parameters such as the relative CIN II$^+$ sensitivity and specificity (self-sampling versus regular sampling), the extra attendance via self-sampling, and the risk of extra attendees. Given that the percentage of women who will switch from office-based to self-sampling is unknown, we determined the percentage of switching that would result in a decrease of the total effectiveness of the programme (ie. harms outweigh the benefits, QALYs are lost). We also examined the circumstances (ie. limits) under which it would not be cost-effective to offer HPV self-sampling to non-attendees.

METHODS

We used the MISCAN-Cervix model to estimate benefits, harms, and costs of offering a self-sampling HPV test to non-attendees(181). For detailed information on the model specifications, see the Supplementary Information.

Assumptions for screening and triage

The screening policy considered is primary HPV screening with cytology triage, as will be implemented in the Netherlands(182) (Figure 7.1-1). Women will be invited for screening at ages 30, 35, 40, 50, and 60 years. In addition, women will be invited at ages 45, 55 and 65 years if they not attend screening or have a positive HPV test in the previous screening round.

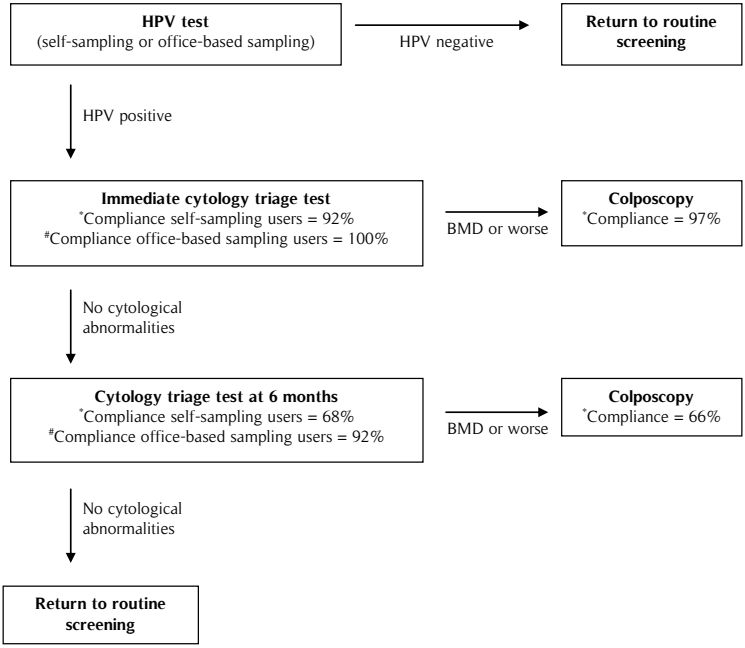

Figure 7.1-1. Triage strategy and compliance assumptions after a positive self-sampling and office-based sampling HPV test. We assumed that the compliance (ie. attendance for triage and colposcopy) behaviour does not differ between self-sampling, future, and current office-based sampling users.

*Compliance rates of the first triage test (ie. immediate cytology triage), second triage test (ie. cytology triage at 6 months), and of colposcopy are assumed equal to those observed within the current programme.
#The first triage test (ie. immediate cytology triage) after a positive office-based sampling test will be performed using co-collection, so the compliance will automatically be 100%. Compliance with the second triage test (ie. cytology triage at 6 months) is assumed to be equal to the first triage test in the current programme.
HPV = Human papillomavirus; BMD = Borderline and mildly dyskaryotic smears.

Assumptions for attendance

For HPV office-based sampling, we assumed an age-dependent overall 65% attendance rate as currently observed within the Dutch cytological screening programme (source: PALGA). On the basis of the findings of the PROHTECT trial (offering self-sampling to non-attendees after an opting-out letter), we assumed that a self-sampling kit was sent to 85% of the non-attendees(183), which resulted in an extra overall attendance of 6 percentage points(73). We assumed that 29% of these extra attendees are higher primary risk women (ie. unscreened women who will never attend via office-based sampling and who have a 1.7 times higher primary underlying risk for developing cervical cancer than women who are willing to attend office-based sampling) which is equal to the proportion in non-attendees

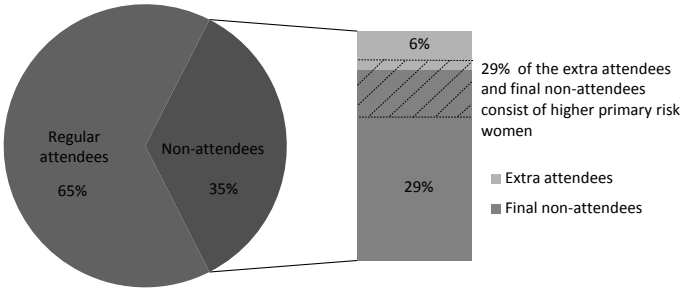

Figure 7.1-2. Distribution of regular attendees, non-attendees and extra attendees within the screening population. After receiving a screening invitation 65% of the invited women will attend via office-based sampling (ie. regular attendees) and 35% will not attend (ie. non-attendees). After a HPV self-sampling test has been offered to the non-attendees, 17% of them will attend (ie. extra attendees; = 6% of the screening population) and 83% will not (ie. final non-attendees; = 29% of the screening population). 29% of the non-attendees consist of higher primary risk women (= 10% of the screening population). We assumed that the proportion of higher primary risk women in the extra attendees and final non-attendees stayed equal to that in the non-attendees (= 1.7% of the screening population are higher primary risk women who attend via self-sampling, 8.3% are higher primary risk women who do not attend). Office-based sampling users consist of regular attendees, while self-sampling users consist of extra attendees and, in case of switching, of (part of the) regular attendees.

(ie. 10% / 35% = 29%) (Figure 7.1-2). In addition to their increased primary underlying risk, these women also have an increased cervical cancer risk due to never attending regular screening.

We assumed that the loss to follow-up after a positive self-sampling test was higher than after a positive office-based sampling test. On the basis of the observed data (source: PALGA), we assumed that 92% of the women comply with the first triage invitation and 68% with the second. With office-based sampling, the collected material can be used both for primary HPV and direct cytology triage testing (co-collection). Therefore, the first and only triage invitation is six months after the positive screening test. This results in a compliance of 100% for immediate cytology triage testing and 92% for triage testing six months after the positive office-based sampling test. In case of self-sampling, co-collection is not possible and women receive their first and second triage invitation directly and six months after the positive screening test. This results in a compliance of 92% for immediate cytology triage testing and 68% for triage testing six months after the positive self-sampling test (Figure 7.1-1).

As no data were available, we considered the two most extreme "switching" scenarios in the base-case analyses: no regular attendees and all regular attendees switch from office-based to self-sampling.

Base-case assumptions for test characteristics

The test characteristics of self-sampling were based on the assumption that a validated PCR test was used, as for instance the GP5$^+$/6$^+$(65, 184). According to the recent meta-analysis of Arbyn and colleagues(185), the point estimate for the relative sensitivity of CIN II$^+$ when comparing self-sampling with office-based sampling is approximately 0.95, whereas the point estimate for the relative specificity is probably higher than 1.00. Therefore, we assumed a 5 percentage points lower sensitivity for high-risk HPV infections when self-sampling (ie. 80 versus 85%), and an equal specificity of 100% (ie. the true but uncertain value of specificity is probably somewhat lower than

100% due to cross-reactivity with low-risk HPV types and contamination). By including fast clearing high-risk HPV infections, we were able to model a lack of specificity.

As women in our model can have multiple lesions at the same time, the CIN II$^+$ sensitivity not only depends on the sensitivity for a high-risk HPV infection, but also on the specificity. Therefore, a 5 percentage points lower sensitivity for high-risk HPV infections and an equal specificity corresponds with a 0.95 relative CIN II$^+$ sensitivity. On the other hand, the specificity for a CIN II$^+$ lesion depends on the specificity and sensitivity for a high-risk HPV infection. As the prevalence of high-risk HPV infections in women without CIN II$^+$ is higher in young women and relatively more young women use self-sampling, a 5 percentage points lower sensitivity for high-risk HPV infections and an equal specificity corresponds with a 0.99 relative CIN II$^+$ specificity.

Assumptions for costs and utilities

Table 7.1-1 presents the inputs for utilities and costs used in the analyses. Utilities were based on (inter)nationally published data(186). The unit costs were estimated from a societal perspective. As compared to office-based sampling, self-sampling was assumed to be less expensive, but the costs of immediate cytology triage were higher. Diagnostic costs of women referred for colposcopy, treatment costs, and costs of palliative care were equal between the two tests and were derived from previous cost studies performed in the Netherlands(187).

Cost-effectiveness analysis

We assumed that the evaluated alternative screening policies (ie. primary HPV screening with and without offering HPV self-sampling to non-attendees) started in 2013 and continued until all women reached the final screening age. The costs and effects of the simulated screening programmes were counted from 2013 onward until all simulated women (ie. born between 1953 and 1992) had died. We also simulated the last three screening rounds

Table 7.1-1. Model inputs: Costs and utilities under base-case assumptions.

		Utility loss	
Parameter	Costs in €	Fraction	Duration
Invitation	4.85		
Primary office-based sampling test			
Programme[1]	2.68 / 2.95		
Organisation	12.50		
Office-based sampling	12.09	0.006	2 weeks
Laboratory	29.00		
Time/travel	6.28		
Total	62.55 / 62.82		
Primary self-sampling test			
Self-sampling kit[2]	6.00		
Programme	2.68		
Organisation	12.50	0.006	2 weeks
Laboratory	29.00		
Time[3]	2.76		
Total	52.94		
Immediate cytology triage test after positive office-based sampling[4]			
Laboratory[5]	30.27	N.A.	N.A.
Total	30.27		
Immediate cytology triage test after positive self-sampling			
Organisation	10.00		
Office-based sampling	12.09		
Laboratory	32.27	0.006	2 weeks
Time/travel	6.28		
Total	60.64		
Cytology triage test at 6 months			
Organisation	10.00		
Office-based sampling	12.09		
Laboratory	32.27	0.006	0.5 year
Time/travel	6.28		
Total	60.64		

Table 7.1-1. (continued)

Parameter	Costs in €	Utility loss	
		Fraction	Duration
Diagnosis and treatment of pre-invasive stages			
False-positive referral	296	0.005	0.5 year
CIN grade I	924	0.03	0.5 year
CIN grade II	1,368	0.07	1 year
CIN grade III	1,602	0.07	1 year
Diagnosis and treatment of cancer			
FIGO 1A	5,246	0.062	5 years
FIGO 1B	12,440	0.062	5 years
FIGO 2^+ (screen-detected)	12,261	0.28	5 years
FIGO 2^+ (clinically detected)	11,451	0.28	5 years
Terminal care	27,859	0.712	1 month

Costs are in 2012 prices. N.A. = Not applicable.

[1] As the total programme costs were fixed, the costs per test were dependent on the number of women participating in the screening programme. As this number was higher with the inclusion of the self-sampling test, the costs per test were lower in the situation with versus without self-sampling.

[2] We assumed that 85% of the non-attendees received the self-sampling kit of €6.00 at home, irrespective of whether they used it or not. This price was estimated based on personal communication with multiple developers of brush and lavage HPV self-sampling kits. The remaining costs (eg. laboratory, organisation, etc.) were only taking into account among women who actually attended via self-sampling.

[3] Given that it was not required to go to the general practitioner's office, we assumed that women who attended via self-sampling spent half of the time to screening (€2.76 instead of €5.52) as compared with women who attended via office-based sampling, while travel costs (€0.76) were absent.

[4] Co-collection-based analysis was possible after positive office-based sampling and, therefore, women did not have to go to the general practitioner´s office for the immediate cytology triage test.

[5] We assumed that part of the material costs (€2.00) was already included in the price of the office-based sampling test. Therefore, laboratory costs of immediate cytology triage after a positive office-based sampling test were lower than after a positive self-sampling test.

before 2013 (ie. primary cytology screening with cytology triage), because they can influence the effectiveness of the screening programme after 2013.

We simulated ten million women for each strategy. Future costs and health effects [life-years (LYs) lived and utility losses] were discounted towards the year 2013 at an annual rate of 3%. We computed the net costs and number of QALYs gained by screening as the difference between the simulations with and without screening. The incremental cost-effectiveness ratio (ICER) was defined as the increase in costs per additional (QA)LY gained when self-sampling would be offered to non-attendees as compared to no such offer. The cost-effectiveness threshold was set to €20,000 per QALY gained, based on decisions of the Dutch government(188), and to €50,000, which is often used in an international perspective(189).

Multivariate sensitivity analyses

The relative CIN II$^+$ sensitivity and specificity can differ from the estimates we used in our base-case analysis, as there is uncertainty about the true value [eg. the 95% confidence interval (CI) for the pooled relative sensitivity and specificity when using the GP5$^+$/6$^+$ is 0.89 to 1.01 and 0.95 to 1.29, respectively(185)]. In addition, they depend on the type of HPV DNA test used(185), meaning that the values could be different when another validated HPV DNA test is used. Therefore, we choose to set the sensitivity for a high-risk HPV infection equally, 5, and 10 percentage points lower for self-sampling as compared to office-based sampling. The specificity was set equally, 5, and 15 percentage points lower. As the CIN II$^+$ sensitivity and specificity depend on both the sensitivity and specificity for high-risk HPV infections, the CIN II$^+$ sensitivity and specificity varied slightly between different combinations of self-sampling test characteristics for high-risk HPV infections. This resulted in a relative CIN II$^+$ sensitivity that varied between 0.89 and 1.02, and a relative CIN II$^+$ specificity that varied between 0.84 and 1.00.

The relative CIN II$^+$ sensitivity and specificity are expected to have a major influence on the effectiveness of the programme, especially when women switch. Therefore, we determined the percentage of women switching (0%, 10%, …., 90%, 100%) for which offering self-sampling is no longer

effective (ie. QALYs are lost) or cost-effective (ie. ICER is larger than the cost-effectiveness threshold). In addition, we varied the loss in quality of life associated with cytology triage, the costs of the self-sampling kit, the extra attendance via self-sampling, and the attendance of higher primary risk women women (ie. unscreened women who will never attend office-based sampling) and their underlying risk for cervical cancer.

Utility loss associated with cytology triage. True estimates of the utility loss due to having cytology triage are unavailable. Especially if self-sampling is associated with a lower specificity, this may influence the effectiveness of offering self-sampling. Therefore, we studied the effect of assuming no utility loss to 0.012 per week for being in triage (base-case: 0.006 per week).

Costs. The total price of a self-sampling kit depends on many factors (eg. type of self-sampling device, possibility to achieve economies of scale, and on-going innovations for the self-sampling test). Therefore, we varied unit self-sampling kit costs from €3.50 to €10.00 (base-case: €6.00).

Attendance via self-sampling. We varied the extra attendance rate due to self-sampling from 3 to 10 percentage points (base-case: 6 percentage points). Furthermore, we varied the proportion of higher primary risk women in extra attendees from 0 to 50% (base-case: 29%).

Underlying risk for cervical cancer of "higher primary risk" women. We assumed that all women have the same underlying risk for cervical cancer (base-case: "higher primary risk" women have a 1.7 times higher underlying risk as compared with regular attendees), although "higher primary risk" women still have an increased cervical cancer risk due to never attending regular screening.

Table 7.1-2. Undiscounted simulated effects and costs, compared with no screening, of primary HPV screening with and without offering self-sampling to non-attendees under base-case assumptions, per 100,000 simulated women.

	Without self-sampling	With self-sampling (difference versus situation without self-sampling, in %)			
		No switching		100% switching	
Effects, n					
Primary screens	219,953	234,171	(+6.5)	234,108	(+6.4)
Triage tests	12,983	13,952	(+7.5)	10,834	(-16.6)
False-positive referrals	163	173	(+5.7)	133	(-18.5)
CIN grade I diagnoses	786	842	(+7.2)	676	(-13.9)
CIN grade II diagnoses	523	565	(+8.1)	460	(-12.0)
CIN grade III diagnoses	844	924	(+9.5)	812	(-3.7)
CeCa cases	626	582	(-7.0)	631	(+0.8)
Screen-detected CeCa cases	77	86	(+12.2)	89	(+16.0)
Clinically-detected CeCa cases	549	496	(-9.6)	541	(-1.3)
CeCa deaths	250	227	(-9.2)	247	(-1.3)
LYs lost	5,929	5,388	(-9.1)	5,833	(-0.8)
QALYs lost	772	735	(-4.8)	760	(-1.6)
Costs, €					
Testing costs	16,022,798	17,134,705	(+6.9)	14,930,636	(-6.8)
Treatment costs	16,820,809	15,856,766	(-5.7)	16,485,808	(-2.0)
Total costs	32,843,608	32,991,471	(+0.5)	31,416,445	(-4.3)

CIN = Cervical intraepithelial neoplasia; CeCa = Cervical cancer; LYs = Life-years; QALYs = Quality-adjusted life-years.

Table 7.1-3. Discounted simulated effects and costs (both 3% per year) of providing non-attendees with a self-sampling test in a primary HPV screening programme under the base-case scenarios, per 100,000 simulated women.

	Base-case scenario, n (% versus no self-sampling)	
	No switching	100% switching
LYs gained	1,746 (+12.1)	1,573 (+1.0)
QALYs gained	1,880 (+12.1)	1,701 (+1.4)
Costs in €	8,184,676 (+5.5)	6,687,767 (-13.8)
ICER: Costs in € per LY gained	2,276	Cost-saving
ICER: Costs in € per QALY gained	2,115	Cost-saving

Cost-saving = Cervical cancer screening was both more effective and less costly with versus without offering HPV self-sampling test to non-attendees. LYs = Life-years; QALYs = Quality-adjusted life-years; ICER = Incremental cost-effectiveness ratio.

RESULTS

Base-case scenario

Table 7.1-2 presents the undiscounted effects and costs per 100,000 simulated women when offering HPV self-sampling to non-attendees of a primary HPV screening programme. Without switching, offering self-sampling increased the number of triage tests and false-positive referrals for colposcopy (+7.5% and +5.7%, respectively) and decreased the number of cervical cancer cases and deaths by 7.0% and 9.2%, respectively. Because the costs increased by only 5.5%, it was not only effective (+12.1% QALYs gained) but also cost-effective (ICER of €2,115 per QALY gained) to add self-sampling to the programme (Table 7.1-3).

As the sensitivity of self-sampling was lower than that of office-based sampling and because the probability of being lost to follow-up after a positive self-sampling test was higher than after a positive office-based sampling test, switching resulted in a decrease of the number of triage tests and subsequently false-positive referrals, an increase of the number of cervical cancers, and a decrease in the number of cervical cancer deaths prevented

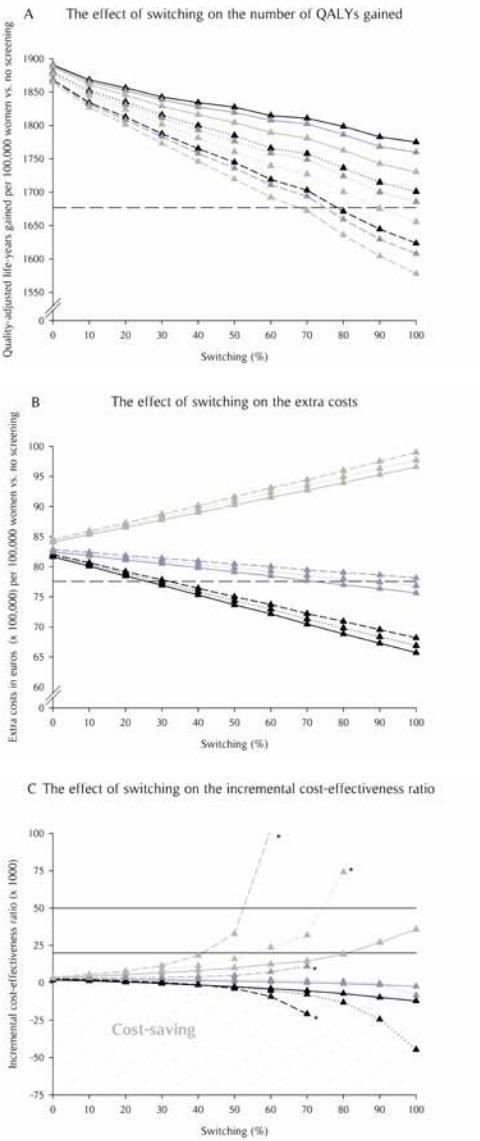

Figure 7.1-3. The effect of switching and the relative CIN II+ sensitivity and specificity on (A) the number of QALYs gained, (B) extra costs, and (C) ICER. Results are given per 100,000

simulated women (3% discounting for costs and effects). The relative CIN II$^+$ sensitivity and specificity (self-sampling versus office-based sampling) are indicated by the sensitivity and specificity in the legend.

(C) The combined effect of sensitivity, specificity and switching on the ICER is only shown when adding a self-sampling test resulted in a gain of QALYs as compared with primary HPV screening alone. Therefore, a negative ICER (ie. cost-saving) is also dominating (ie. primary HPV screening with offering a self-sampling test to non-attendees was both more effective and less costly than primary HPV screening alone).
*Beyond this level of switching, offering a self-sampling test resulted in a loss of QALYs as compared to primary HPV screening alone.
Black dashed line = Primary HPV-screening without offering self-sampling to non-attendees.

and QALYs gained. Still, when all women switched it was effective and cost-saving to offer self-sampling (Table 7.1-2 and 7.1-3).

Multivariate sensitivity analyses

Without switching, a decrease in the CIN II$^+$ sensitivity of self-sampling mainly resulted in fewer QALYs gained (Figure 7.1-3a), whereas a decrease in the CIN II$^+$ specificity mainly resulted in increased costs (Figure 7.1-3b). Both resulted in a higher ICER (Figure 7.1-3c). However, even when the relative sensitivity and specificity were inferior to that of office-based sampling (ie. 0.89-0.91 and 0.84-0.85, respectively), QALYs were gained and the ICER was below the threshold of €20,000 per QALY gained, if no women switched.

In all scenarios, switching resulted in fewer QALYs gained (Figure 7.1-3a). This effect was larger in case the relative sensitivity was lower than 1.00. However, even when the test characteristics of self-sampling were inferior to that of office-based sampling, QALYs were only lost when more than 60% of the women switched (Table 7.1-4). When they were slightly inferior (ie. 0.95-0.97 relative sensitivity and 0.94-0.95 relative specificity) or similar (ie. 1.01-1.02 relative sensitivity and 0.99-1.00 relative specificity), it was effective under every switching scenario. If self-sampling specificity was inferior, the costs of offering HPV self-sampling increased with increasing

Table 7.1-4. For what switching percentage does offering self-sampling to non-attendees lead to a loss in QALYs?

Relative CIN II* sensitivity and specificity (self versus office-based sampling)		Base-case	Utility loss due to cytology triage		Extra attendance (percentage points)		*Underlying risk of higher primary risk women	% Extra attendees consisting of higher primary risk women	
Sensitivity	Specificity		**0.000	0.012	10	3	***Average	50	0
1.01	0.99	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	>60
1.01	0.94	*Ind.*	*Ind.*	*Ind.*	*Ind.*	>90	*Ind.*	*Ind.*	>50
1.02	0.84	*Ind.*	*Ind.*	*Ind.*	*Ind.*	>70	>80	*Ind.*	>40
0.95	0.99	*Ind.*	*Ind.*	*Ind.*	*Ind.*	>50	>60	*Ind.*	>30
0.96	0.95	*Ind.*	*Ind.*	>90	*Ind.*	>50	>50	*Ind.*	>30
0.97	0.85	>80	*Ind.*	>70	*Ind.*	>40	>40	*Ind.*	>20
0.89	1.00	>70	>70	>70	*Ind.*	>40	>40	*Ind.*	>20
0.90	0.95	>70	>70	>70	*Ind.*	>30	>30	>90	>20
0.91	0.85	>60	>80	>50	>90	>30	>30	>80	>10

For every scenario the minimum switching percentage is given under which it is no longer effective (ie. QALYs are lost) to offer self-sampling to non-attendees. The switching percentage varies between 0 (ie. even when no women switch it is not effective to offer self-sampling), >90 (ie. when more than 90% of the women switch it is not effective to offer self-sampling) to independent (ind.; ie. independent of how many women switch, it is always effective to offer self-sampling).

Base-case assumptions: Utility loss due to cytology triage = 0.006 per week, extra attendance = 6 percentage points, underlying risk of higher primary risk women as compared with the rest of the screening population = 1.7 times higher, % of higher primary risk women in extra attendees = 29%. These variables, if not varied, were held constant at their base-case level.

*Underlying risk for developing cervical cancer.

**Because a higher utility loss for primary than triage testing does not seem realistic we also assume no utility loss for primary testing.

***Average = Equal to the rest of the screening population.

Table 7.1-5. For what switching percentage is offering self-sampling to non-attendees not (cost-)effective?

Relative CIN II* sensitivity and specificity (self versus office-based sampling)		Base-case	Utility loss due to cytology triage		Costs of self-sampling kit in €		Extra attendance (percentage points)		*Underlying risk of higher primary risk women	% Extra attendees consisting of higher primary risk women	
Sensitivity	Specificity		**0.000	0.012	3.50	10	10	3	***Average	50	0
1.01	0.99	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	>60
1.01	0.94	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	>90	*Ind.*	*Ind.*	>20
1.02	0.84	*>80*	*Ind.*	*>60*	*>90*	*>60*	*Ind.*	>40	>30	*Ind.*	>0
0.95	0.99	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	*Ind.*	>50	>60	*Ind.*	>20
0.96	0.95	*Ind.*	*Ind.*	>90	*Ind.*	*>80*	*Ind.*	>50	>40	*Ind.*	>10
0.97	0.85	>50	*>70*	>40	>60	>40	*>80*	>20	>20	*>80*	>0
0.89	1.00	>70	>70	>70	>70	>70	*Ind.*	>40	>30	*Ind.*	>10
0.90	0.95	>70	>70	>60	>70	>50	*Ind.*	>30	>20	>90	>0
0.91	0.85	>40	>50	>30	>40	>30	>60	>20	>10	>50	>0

For every scenario the minimum switching percentage is given under which it is no longer effective (ie. QALYs are lost) and/or cost-effective (ie. €20,000 per QALY gained) to offer self-sampling to non-attendees. The switching percentage varies between 0 (ie. even when no women switch it is not effective and/or cost-effective to offer self-sampling), >90 (ie. when more than 90% of the women switch it is not effective and/or cost-effective to offer self-sampling) to independent (ind.; ie. independent of how many women switch, it is always effective and cost-effective to offer self-sampling).

Base-case assumptions: Utility loss due to cytology triage = 0.006 per week, extra attendance = 6 percentage points, underlying risk of higher primary risk women as compared with the rest of the screening population = 1.7 times higher, % of higher primary risk women in extra attendees = 29%. These variables, if not varied, were held constant at their base-case level.

*Underlying risk for developing cervical cancer.

**Because a higher utility loss for primary than triage testing does not seem realistic we also assume no utility loss for primary testing.

***Average = Equal to the rest of the screening population.

percentages of switching (Figure 7.1-3b). Considering a cost-effectiveness threshold of 20,000 per QALY gained, the switching limit was up to 30 percentage points lower (Table 7.1-5). Therefore, offering self-sampling was not effective or cost-effective when more than 40% of the women switch and test characteristics of self-sampling were inferior to those of office-based sampling.

The effect of the level of utility loss associated with cytology triage was negligible (Table 7.1-4). Varying the extra attendance or underlying risk of higher primary risk women had more influence. When the extra attendance was halved (from 6 to 3 percentage points) or if higher primary risk women did not have an elevated underlying risk, QALYs were lost when more than 50% of the women switched and test characteristics were slightly inferior. When they were inferior, it was no longer effective if more than 30% of the women switched. The most influential parameter was the attendance of higher primary risk women. When they did not attend, it was not effective to offer self-sampling when more than 60% of the women switched and test characteristics were equal. In case they were inferior, this threshold decreased to 10%. For offering self-sampling to be cost-effective, these switching thresholds were even lower (Table 7.1-5).

DISCUSSION

The number of QALYs gained by offering HPV self-sampling to non-attendees was influenced by self-sampling test characteristics, the extra attendance via self-sampling, and the risk of extra attendees. When none of the regular attendees switched to self-sampling, it was always effective to offer HPV self-sampling. Switching resulted in fewer QALYs gained because the probability of being lost to follow-up after a positive self-sampling test was higher than after a positive office-based sampling test. If in addition the sensitivity of self-sampling was lower than that of office-based sampling, the number of QALYs gained decreased even more. However, even when test

characteristics were inferior, up to 60% of the regular attendees could switch before the QALYs gained by the 6 percentage points extra attendance were annulled by the QALYs lost by switching. This percentage dropped to 30% when the extra attendance halved from 6 to 3 percentage points or when higher primary risk women did not have an elevated underlying risk. It dropped to 10% if higher primary risk women did not attend self-sampling. When also considering a cost-effectiveness threshold of €20,000 per QALY gained, these switching thresholds were 10 to 20 percentage points lower.

Our base-case assumption of 6 percentage points extra attendance was based on the Dutch PROHTECT trial in which a self-sampling kit was sent to 85% of all non-attendees (ie. the remaining 15% opted-out via a letter) (73). Using another strategy will probably result in another extra attendance rate. If this rate will be lower than 3 percentage points (almost) no women can switch before more QALYs are lost than gained.

We assumed that a subset of the unscreened women have a 1.7 times higher underlying risk on cervical cancer (ie. higher primary risk women) than the rest of the screening population, which was based on model calibration. Dugué and colleagues' results have shown that non-attendees of cervical cancer screening (ie. no cervical smear taken in the past eight years) had a 3.8-fold increased risk of dying from non-cervical (ie. non-screened) HPV-associated cancers(190), which seems to confirm our assumption that at least part of the non-attendees have an increased underlying risk. Although the PROHTECT study showed that unscreened women (ie. invited for screening at least once but never attended) attended via self-sampling(176), it is uncertain whether this is the subset with an increased underlying risk. If these higher primary risk women do not attend via self-sampling, 10% to 60% of the women can switch before QALYs are lost by offering HPV self-sampling to non-attendees.

The relative sensitivity and specificity of self-sampling as compared with office-based sampling will depend on the type of HPV DNA test used(185, 191). However, even when a validated PCR is used (eg. $GP5^{+}/6^{+}$ or the real-time hrHPV test), it is possible that the sensitivity and specificity of self-

sampling are both inferior to that of office-based sampling. In fact, relative test characteristics of self-sampling might even be worse than we assumed in our sensitivity analyses(185). In that case, the maximum percentage of women that can switch before QALYs are lost is also lower.

Studies in Sweden(192), Finland(193), the United Kingdom(194), and Italy(195) have also shown that offering self-sampling to non-attendees increased screening participation rates. We expect that our conclusions to a large extent apply to other countries and regions with well-organized invitational screening programmes with a high compliance and an optimal age range and screening frequency. Even if this would mean that HPV self-sampling would be offered to non-attendees of a primary cytology instead of a primary HPV programme. For countries and regions with a lower underlying risk and/or a more intensive screening programme as compared to the Netherlands, benefits of increased participation due to self-sampling are probably lower. In countries without a highly organized invitational programme, it may not be feasible to offer a self-sampling test to unscreened women. Instead, it could be offered to the general population by selling it over the counter. However, when screening is not reimbursed by the government, it is questionable to what extent unscreened women will use self-sampling. Indeed, a discrete choice experiment in the USA showed that vulnerable adults valued costs higher than the kind of screening offered or the travel distance to obtain screening(196). When non-attendance is driven by other factors than feeling uncomfortable or having little time (ie. factors that can be overcome by using self-sampling instead of going to the clinician(197), the success of offering self-sampling may be limited.

To our knowledge, this is the first study on the harms and benefits of providing a self-sampling test to non-attendees of a cervical cancer screening programme. One of our key assumptions (ie. extra attendance via self-sampling) was based on observations from the PROHTECT trials(73, 176, 180). We extensively studied the effect of the level of switching in combination with the test characteristics of self-sampling and the underlying cervical cancer

risk of its users, which were important and uncertain parameters for the effectiveness of offering self-sampling.

A limitation of our study is that we only focused on unvaccinated women. Screening programmes will probably be adapted when vaccinated cohorts reach the start age of screening. A separate analysis for this future situation is beyond the scope of the present analysis. However, we expect that offering self-sampling to non-attendees will be less (cost-)effective, because we expect that fewer health effects can be gained by increasing attendance because of a lower underlying risk. Another drawback is the limited transposability to other health systems. We expect lower benefits of increased participation due to self-sampling in screening programmes that are more intensive than the Dutch future programme will be (ie. 5 lifetime screens at ages 30, 35, 40, 50, and 60 years). Moreover, we might have overestimated the colposcopy compliance after a positive self-sampling test, as this may be lower than after a positive office-based sampling test. This may have resulted in a slight overestimation of the effectiveness of self-sampling. In addition, the relative CIN II$^+$ specificity as described in our study will be somewhat higher when regular attendees switch, as the prevalence of high-risk HPV infections in women without a CIN II$^+$ is slightly lower in regular attendees as compared with non-attendees attending self-sampling. Furthermore, we did not account for other healthcare that women may get while attending clinic-based screening. This may have underestimated health losses in regular attendees switching to self-sampling, as well as health gains in the small group of extra attendees with a positive self-sampling test complying with their triage invitation.

Offering self-sampling to non-attendees clearly offers an opportunity to increase health benefits in cervical cancer screening if health providers make sure that (i) the relative CIN II$^+$ sensitivity is at least 0.95, (ii) unscreened attendees are recruited with self-sampling, and (iii) the total attendance increases by at least 6 percentage points. Otherwise, switching of regular attendees to self-sampling may annul the benefits of self-sampling and even decrease the effectiveness of a primary HPV screening programme.

ACKNOWLEDGEMENTS

This study was funded by the Dutch National Institute for Public Health and the Environment (007/12 V&Z NvdV/EM).

SUPPLEMENTARY INFORMATION

This Supplementary Material was provided by Steffie K. Naber, Kirsten Rozemeijer, Inge M.C.M. de Kok, Joost van Rosmalen and Marjolein van Ballegooijen. Here, we describe the model inputs of the MISCAN model for cervical cancer(181). This model can be used to assess the harms and benefits of different screening programmes for cervical cancer, as well as HPV vaccination. The model has been used previously for cost-effectiveness analyses of cervical cancer screening and HPV vaccination(84, 85, 188, 198).

The MISCAN model (Figure 7.1-S1) consists of the following four parts: demography, natural history, screening, and cost-effectiveness. The assumptions used in each of these parts are described below.

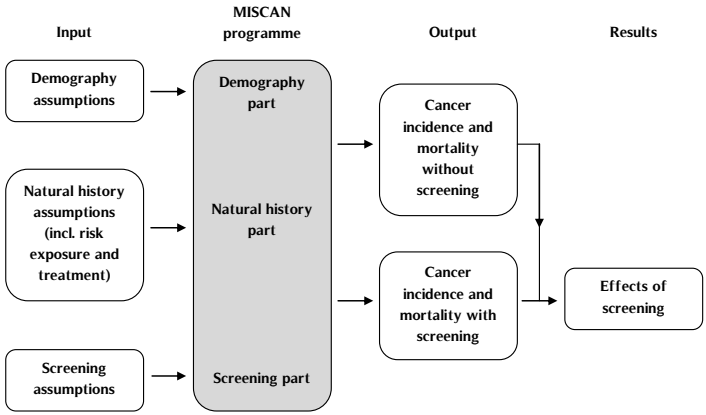

Figure 7.1-S1. Structure of the MISCAN model

Demography

The MISCAN model generates a simulated population of Dutch women born between 1953 and 1992. Women born after 1992 are eligible for HPV vaccination and are therefore not considered here. The relative sizes of the birth cohorts are based on the age distribution of women living in the Netherlands.

For each woman, a time of death from other causes (ie. causes other than cervical cancer) is generated; this time of death is independent of the cervical cancer disease model. In the model, a woman's lifetime cannot exceed 100 years. The time of death from other causes is generated using a life table for women from Statistics Netherlands(199). The assumed hysterectomy rates vary by age and by year of birth. These rates are based on data from Statistics

Table 7.1-S1. Model assumptions for the age-specific probability of having had a hysterectomy for reasons other than cervical cancer, for women with birth years 1953-1992.

Age	Birth years 1953-1958	Birth years 1959-1992
20	0.0000	0.0000
25	0.0003	0.0002
30	0.0030	0.0017
35	0.0134	0.0076
40	0.0372	0.0213
45	0.0755	0.0432
50	0.1138	0.0735
55	0.1367	0.0916
60	0.1484	0.1009
65	0.1601	0.1102
70	0.1732	0.1217
75	0.1862	0.1330
80	0.1963	0.1419
85	0.2023	0.1468

Linear interpolation is used to determine the probability of having had a hysterectomy at intermediate ages.

Netherlands and Information Centre for Health Care and are presented in Table 7.1-S1(200, 201).

Natural history

During her lifetime, each woman has an age-specific risk of acquiring high-risk HPV infections (ie. an infection caused by an HPV type that can cause cancer and that can be detected by the HPV test) and CIN lesion without a (detectable) high-risk HPV infection. Most HPV infections clear or regress naturally, some HPV infections can progress to CIN I, CIN II, CIN III, cervical cancer, and death from cervical cancer.

To account for the fact that HPV infections and CIN may clear or regress naturally, six disease pathways are distinguished in MISCAN. Each instance of these disease pathways represents an HPV infection or a 'lesion' (ie. CIN of a certain grade or a stage of cervical cancer). Each disease pathway starts as either an HPV infection or as an HPV negative CIN I lesion. The natural history (ie. in the situation without screening) of these six disease pathways is shown in Figure 7.1-S2 and can be described as follows.

A) HPV infections that clear naturally without ever leading to CIN
B) HPV infections that progress to CIN I and then regress
C) HPV infections that progress to CIN I and CIN II and then regress
D) HPV infections that progress to CIN I, CIN II, and CIN III and then regress
E) HPV negative CIN I lesions that regress naturally or become HPV negative CIN II and then regress naturally
F) HPV infections that progress to CIN I, CIN II, CIN III, preclinical FIGO 1A (micro-invasive) cervical cancer, and preclinical FIGO 1B cervical cancer. Preclinical FIGO 1B cervical cancer can either become clinically detected FIGO 1B cervical cancer or progress to preclinical FIGO 2^+ cervical cancer and then to clinical FIGO 2^+ cervical cancer. Clinically detected cervical cancer can progress to death from cervical cancer or remain in that state forever (if the woman is cured from cervical cancer).

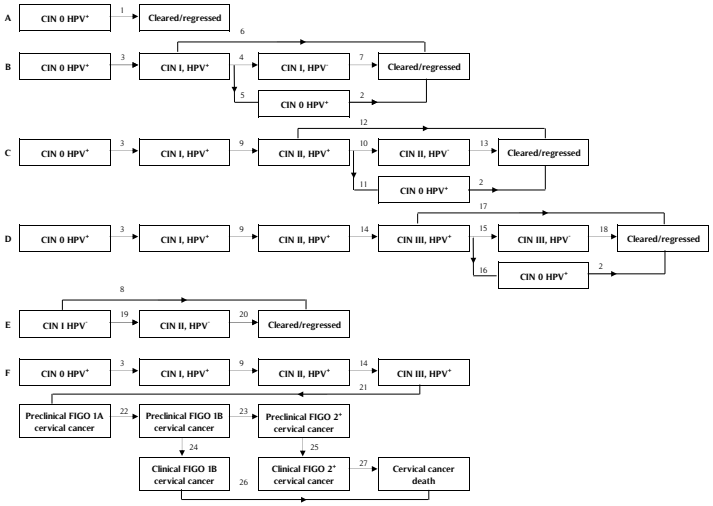

Figure 7.1-S2. Schematic representation of the MISCAN model, with disease pathways A through F.

Notes: There are six disease pathways (types A through F) in MISCAN. All lesions start as either an HPV infection without CIN (disease pathways A, B, C, D, and F) or as a CIN I lesion without HPV infection (disease pathway E). Cleared/regressed denotes the absence of CIN and HPV infection; CIN 0 denotes the absence of CIN and cervical cancer. All cervical cancer states are HPV positive. The arrows between the states show which types of transitions can occur; the numbers refer to the duration distributions shown in Table 7.1-S2. In every state before death, a transition to "other-cause death" can occur, and in every state before cancer, a transition to "hysterectomy" can occur; in these cases, the transition applies to all HPV infections and CIN lesions of that person simultaneously.

A woman can acquire multiple lesions and HPV infections during her lifetime, and multiple lesions and HPV infections may be present at the same time. In each simulated life history (ie. between ages 0 and 100), the number of lesions of each type follows a Poisson distribution. The annual probability of acquiring an HPV infection or CIN lesion is age-dependent and depicted in Figures 7.1-S3a (regressive disease pathways) and 7.1-S3b (progressive disease pathway). The transitions and sojourn times of the HPV infections or lesions are simulated based on a continuous-time semi-Markov process.

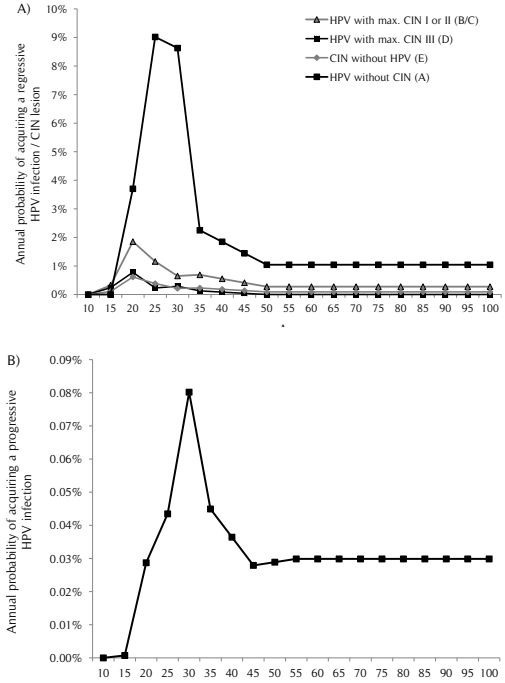

Figure 7.1-S3. Annual probability of acquiring A) a regressive HPV infection or CIN lesion, and B) a progressive HPV infection.

The sojourn times of most states in the model have either an exponential or a Weibull probability distribution (Table 7.1-S2).

The mean number of lesions of each type does not depend on the birth year, as we do not assume any cohort effects on the underlying risk of developing HPV infections, CIN, and cervical cancer for the cohorts included in this model. However, we do account for the possibility that the underlying risk may be higher for women who do not attend screening.

In the model, women who do not have cervical cancer have an age-specific probability of getting a hysterectomy for reasons other than cervical cancer.

Table 7.1-S2. Transitions and duration distributions used in MISCAN.

Transition number[1]	Disease pathway[1]	From state	To state	Probability of transition	Type of distribution	Mean duration (years)	Weibull shape parameter
1	A	CIN 0 HPV$^+$	Cleared/regressed	1	Exponential	1.0	1
2	B, C, D	CIN 0 HPV$^+$	Cleared/regressed	1	Exponential	1.0	1
3	B, C, D, F	CIN 0 HPV$^+$	CIN I HPV$^+$	1	Exponential	1.0	1
4	B	CIN I HPV$^+$	CIN I HPV$^-$	0.4	Exponential	1.5	1
5	B	CIN I HPV$^+$	CIN 0 HPV$^+$	0.3	Exponential	1.5	1
6	B	CIN I HPV$^+$	Cleared/regressed	0.3	Exponential	1.5	1
7	B	CIN I HPV$^-$	Cleared/regressed	1	Exponential	1.0	1
8	E	CIN I HPV$^-$	Cleared/regressed	1	Exponential	1.5	1
9	C, D, F	CIN I HPV$^+$	CIN II HPV$^+$	1	Exponential	1.5	1
10	C	CIN II HPV$^+$	CIN II HPV$^-$	0.4	Exponential	2.0	1
11	C	CIN II HPV$^+$	CIN 0 HPV$^+$	0.3	Exponential	2.0	1
12	C	CIN II HPV$^+$	Cleared/regressed	0.3	Exponential	2.0	1
13	C	CIN II HPV$^-$	Cleared/regressed	1	Exponential	1.0	1
14	D, F	CIN II HPV$^+$	CIN III HPV$^+$	1	Exponential	2.0	1
15	D	CIN III HPV$^+$	CIN III HPV$^-$	0.4	Weibull	3.1	1.67
16	D	CIN III HPV$^+$	CIN 0 HPV$^+$	0.3	Weibull	3.1	1.67
17	D	CIN III HPV$^+$	Cleared/regressed	0.3	Weibull	3.1	1.67
18	D	CIN III HPV$^-$	Cleared/regressed	1	Exponential	1.0	1
19	E	CIN I HPV$^-$	CIN II HPV$^-$	1	Exponential	1.5	1
20	E	CIN II HPV$^-$	Cleared/regressed	1	Exponential	2.0	1
21	F	CIN III HPV$^+$	Preclinical FIGO 1A	1	Weibull	11.8	1.67
22	F	Preclinical FIGO 1A	Preclinical FIGO 1B	1	Exponential	3.2	1
23	F	Preclinical FIGO 1B	Preclinical FIGO 2$^+$	Age-specific[2]	Exponential	0.5	1
24	F	Preclinical FIGO 1B	Clinical FIGO 1B	Age-specific[2]	Exponential	0.5	1
25	F	Preclinical FIGO 2$^+$	Clinical FIGO 2$^+$	1	Exponential	1.3	1
26	F	Clinical FIGO 1B	Cervical cancer death	Age-specific[3]	Piecewise uniform	Age-specific[4]	-
27	F	Clinical FIGO 2$^+$	Cervical cancer death	Age-specific[3]	Piecewise uniform	Age-specific[4]	-

[1] See Figure 7.1-S2

[2] Transition probability depends on age; see Table 7.1-S3.

[3] Transition probability depends on age; see Table 7.1-S4.

[4] See Table 7.1-S5 for the duration distribution.

Table 7.1-S3. Age-specific stage distribution of a clinically detected cervical cancer.

	Clinical detection in stage:	
Age	FIGO 1B	FIGO 2^+
0	25.4%	74.6%
25	25.4%	74.6%
40	35.0%	65.0%
55	61.4%	38.6%
70	75.4%	24.6%
100	75.4%	24.6%

Percentages in the table are estimated in the model calibration. Linear interpolation is used to determine the probabilities at intermediate ages.

A hysterectomy is assumed to remove all prevalent HPV infections and CIN lesions. Women with a hysterectomy will no longer acquire HPV infections or CIN lesions and are also no longer invited for screening tests.

The assumptions for the probability and the duration of survival after a clinically detected (ie. detected because of symptoms) cervical cancer are based on data from the Dutch Cancer Registry for the period 1989-2009. We assumed that all cervical cancer mortality occurs in the first ten years after diagnosis. The assumed probability of long-term survival depends on age and stage (FIGO 1B or FIGO 2^+); in the model, FIGO 1A cervical cancer cannot be clinically detected. Table 7.1-S3 shows what percentage of clinically detected cancers is detected in stages FIGO 1B and FIGO 2^+. The model assumptions for the long-term survival probabilities are shown in Table 7.1-S4 and the assumed duration distributions are shown in Table 7.1-S5.

Table 7.1-S4. Model assumptions for the age-specific probability that clinical FIGO 1B and FIGO 2+ cervical cancer will lead to death from cervical cancer (ie. 100% - probability of long-term survival), in the absence of other-cause mortality.

Age	Clinical FIGO 1B	Clinical FIGO 2+
0	9.7%	45.5%
30	9.7%	45.5%
45	10.8%	51.1%
60	22.9%	55.4%
80	34.5%	68.7%
100	34.5%	68.7%

Linear interpolation is used to determine the probabilities at intermediate ages. Source: observed age-specific and stage-specific survival for the periods 1989-2002 and 2003-2009, obtained from the Dutch Cancer Registry.

Table 7.1-S5. Model assumptions for the duration distribution of clinical FIGO 1B and FIGO 2+ cervical cancer, if the transition to death from cervical cancer occurs.

Years after detection	Clinical FIGO 1B	Clinical FIGO 2+
1	10.4%	37.6%
2	36.5%	64.6%
3	47.9%	78.1%
4	61.5%	84.5%
5	78.3%	88.5%
6	84.4%	90.5%
7	90.3%	93.3%
8	93.1%	96.4%
10	100.0%	100.0%

The values in this table represent the percentages of cervical cancer deaths that occur within a given number of years after the moment of clinical diagnosis. It is assumed that no cervical cancer mortality occurs more than ten years after clinical diagnosis. Source: observed age-specific and stage-specific survival for the periods 1989-2002 and 2003-2009, obtained from the Dutch Cancer Registry.

Screening

Screening can change the life histories of women. In the current analysis, women are invited for primary HPV screening at the age of 30, 35, 40, 50, and 60. In some strategies, a self-sampling HPV test is offered to non-attendees of office sampling (ie. collected in the general practitioner's office). Women aged 45, 55 and 65 years are only invited for screening if they had a positive test or did not respond in the last screening round.

Table 7.1-S4 shows the assumed age-specific attendance rates, which were based on the current Dutch screening programme. We assumed that 10% of the female population never attends screening via office sampling (ie. office-sampling refusers or higher primary risk women). Based on the proportion of cancers that are clinically detected in the Netherlands, we estimated that the women who never attend screening have a 1.71 times higher underlying risk than the other 90% of the female population (ie. women who are inclined to attend office sampling). The compliance in different stages of the follow-up is shown in Figure 7.1-1.

If an HPV test is applied, each HPV infection prevalent at the time of screening has a probability of producing a positive test (ie. the sensitivity). If the HPV test is positive, cytological inspection determines whether the woman is referred for colposcopy or invited for cytological triage after 6 months. The test characteristics of the office sampling HPV test and of the cytological inspection used in the triage are shown in Table 7.1-S5. The sensitivity of office-based sampling was based on the difference in CIN III$^+$ detection rates between cytology and HPV testing found in the POBASCAM study (9). The test characteristics of the HPV self-test are varied and the possible values are described in the methods of the main manuscript.

In case a woman is referred for colposcopy, she will visit the gynaecologist with a certain probability (see Figure 7.1-1). At the gynaecologist, all prevalent CIN lesions are assumed to be diagnosed and successfully removed (in the costs we accounted for the requirement of repeated CIN treatment in

Table 7.1-S6: Assumed age-specific attendance rates at primary HPV screening.

Age	Women who are inclined to attend office sampling (90% of the women)	Office-sampling refusers[1] (10% of the women)	Overall
30	61.4%	0%	55.2%
35	69.6%	0%	62.6%
40	72.9%	0%	65.6%
45	77.9%	0%	70.1%
50	79.4%	0%	71.5%
55	75.0%	0%	67.5%
60	76.2%	0%	68.6%
65[2]	76.2%	0%	68.6%

[1] Defined as higher primary risk women in the main manuscript.

[2] In the current screening programme, women aged 65 are not invited for screening. We assumed their attendance to be the same as for women aged 60.

Table 7.1-S7. Model assumptions for test characteristics of HPV office-based sampling and cytological inspection

Test	Parameter	Value
HPV office-based sampling	Probability of positive HPV test if HPV positive	85%
	Probability of positive HPV test if HPV negative	0%[1]
Cytology	Probability of at least ASCUS/LSIL for CIN 0	2.4%
	Probability of at least HSIL for CIN 0	0.03%
	Probability of at least ASCUS/LSIL for CIN I	40.0%
	Probability of at least HSIL for CIN I	3.6%
	Probability of at least ASCUS/LSIL for CIN II	50.0%
	Probability of at least HSIL for CIN II	18.0%
	Probability of at least ASCUS/LSIL for CIN III	75.0%
	Probability of at least HSIL for CIN III	55.9%
	Probability of at least ASCUS/LSIL for preclinical cervical cancer	75.0%
	Probability of at least HSIL for preclinical cervical cancer	59.7%

[1] The specificity of the HPV test is assumed to be 100%; a possible lack of specificity is modeled as fast-clearing HPV infections.

15% of the cases, but no cancer development will take place after CIN detection). HPV infections without CIN are not treated. For screen-detected cervical cancer, a stage-specific improvement (compared to the situation without screening) in the probability of cure is assumed.

The effects of early detection on survival

For screen-detected invasive cancers, survival was modelled as a reduction in the risk of dying compared with that risk in the situation without screening, when the cancer would have become clinical. This improvement of prognosis (89.4%, 50%, and 20% for FIGO IA, IB, and 2^+, respectively) was calibrated to reproduce recently observed stage specific survival given observed screening [Dutch Cancer Registry (NKR)].

Calculation of health effects and costs of screening

For each simulated woman who is alive, MISCAN can determine the state, which can be Normal, HPV infected, CIN I, CIN II, CIN III, FIGO 1A, FIGO 1B, and FIGO 2^+. A woman can have multiple HPV infections or CIN lesions at the same time. Her state is determined by the most severe disease stage present, using the order HPV infection, CIN I, CIN III, CIN III, FIGO 1A cervical cancer, FIGO 1B cervical, and FIGO 2^+ cervical cancer; if no HPV infections or CIN lesions are present, the woman's state is Normal.

The model produces the number of life-years spent in each state as well as the number of certain events (eg. screenings and cervical cancer diagnoses) in a lifetime. For each of these events, Table 7.1-1 presents the associated costs and disutility. To calculate the total costs of a screening strategy, a sum is taken over all the numbers of events multiplied by their associated costs. The same holds for the total disutility of a strategy.

In the current analysis, the number of life-years gained is calculated as the difference in total years lived by the population between the situation with and without offering self-sampling to non-attendees of office sampling. To determine the number of QALYs gained (or lost), we computed the dif-

ference in the total number of QALYs between both situations. Similarly, the net costs were determined. From these numbers, the incremental cost-effectiveness ratio (ICER) of offering HPV self-sampling to non-attendees could be calculated by dividing the additional costs by the additional QALYs gained.

CHAPTER 7.2

When is it effective to offer self-sampling to non-attendees — Response

Kirsten Rozemeijer[1], Inge MCM de Kok[1], Steffie K Naber, Folkert J van Kemenade[2], Corine Penning[1], Joost van Rosmalen[1,3], and Marjolein van Ballegooijen[1]

1. Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, the Netherlands
2. Erasmus MC, University Medical Center, Department of Pathology, Rotterdam, the Netherlands
3. Erasmus MC, University Medical Center, Department of Biostatistics, Rotterdam, the Netherlands

Cancer Epidemiol Biomarkers Prev. 2015 Aug;24(8):1296.

Our paper of chapter 7.1 evoked a response by dr. Castle that the editorial board of Cancer Epidemiology, Biomarkers & Prevention allowed us to provide with an answer. In his letter, dr. Castle expressed his concern on the performance of HPV self-sampling as compared to office-based sampling(202). He stated that the real-world performance of HPV self-sampling is likely to be less accurate than its performance in the artificial setting of a clinical study. Moreover, he could imagine that 20% to 30% of attendees switch from office-based to self-sampling, while the loss in sensitivity for cervical precancer and cancer is larger than 10%. Our response is reproduced in this chapter.

WHEN IS IT EFFECTIVE TO OFFER SELF-SAMPLING TO NON-ATTENDEES—RESPONSE

Kirsten Rozemeijer, Inge MCM de Kok, Steffie K Naber, Folkert J van Kemenade, Corine Penning, Joost van Rosmalen, and Marjolein van Ballegooijen

We appreciate Dr. Castle's concern that a loss in CIN II$^+$ sensitivity, when using self-sampling instead of office-based sampling, may still be an issue(202), even though data of a recently published meta-analysis reported otherwise(185).

Data of the Dutch PROHTECT study strongly suggested that the CIN II$^+$ sensitivity of HPV self-sampling was non-inferior to that of HPV office-based sampling(73). As self-sampling was offered to non-attendees at home, it is unlikely that the sensitivity profile in this study was biased because of "in office procedures", although the PROHTECT study estimates may have been biased due to higher risk profiles in the non-attendees. Therefore, studies are needed to validate self-sampling test characteristics in regular responders with lower risk profiles in case they switch from office-based sampling to self-sampling.

In our article, we showed that even if the loss in CIN II$^+$ sensitivity is 10%, and 20 to 30% of the regular attendees switch to self-sampling, it is still both effective and cost-effective to offer self-sampling to non-attendees as long as unscreened women attend, and the extra attendance rates is at least 6 percentage points(203). Results of the Dutch PROHTECT studies, where self-sampling was offered via an opt-out procedure (ie. a self-sampling kit was sent to all non-attendees except when they opted-out via a letter), showed that both assumptions were realistic(73, 176). However, as opt-in procedures (ie. involving a request for a self-sampler) may reduce response rates(71), the chosen strategy could be crucial in whether or not offering self-sampling is (cost-)effective. Therefore, we fully concur with Dr. Arbyn and Dr. Castle that the introduction of self-sampling strategies should be carefully prepared and evaluated in pilot studies integrated in well-organized settings before general rollout(71).

CHAPTER 8

The association between socioeconomic status and the underlying screen-independent cervical cancer risk

Kirsten Rozemeijer[1], Corine Penning[1], Folkert J van Kemenade[2], Steffie K Naber[1], Suzette M Matthijsse[1], Albert G Siebers[3, 4], Inge MCM de Kok[1], and Marjolein van Ballegooijen[1]

1. Erasmus MC, University Medical Center, Department of Public Health, Rotterdam, the Netherlands
2. Erasmus MC, University Medical Center, Department of Pathology, Rotterdam, the Netherlands
3. Radboud University Medical Center, Department of Pathology, Nijmegen, the Netherlands
4. PALGA, the nationwide network and registry of histo- and cytopathology in the Netherlands, Houten, the Netherlands

Submitted

ABSTRACT

Background. Lower socioeconomic status (SES) women are more often diagnosed with cervical cancer, which might solely be attributed to their lower screening participation rate, but could also be partly due to other factors. Therefore, we determined whether low SES women have an increased underlying risk to be diagnosed with cervical cancer, independent of differences in screening history and screening uptake.

Methods. Retrospective data were obtained from the Dutch nationwide network and registry of histo- and cytopathology (PALGA). First-time attendees without any history of cervical examinations were included in the study to eliminate influences of differences in screening history and uptake. We performed logistic regression analyses to compare the cervical cancer risk between women living in low, intermediate, and high SES neighbourhoods (ie. neighbourhood status is indicative of their SES) attending the organized cervical cancer screening programme between 2000 and 2007. We adjusted for differences in age distribution. For every postal code, SES was determined based on income, education and unemployment rate.

Results. We included 404,761 first-time attendees of whom 254 (0.06%) were diagnosed with cervical cancer. As compared to women living in high SES neighbourhoods, those in low and intermediate SES neighbourhoods had a similar cervical cancer risk [ie. OR of 0.98 (95% CI: 0.70, 1.39) and 1.07 (95% CI: 0.79, 1.45), respectively].

Conclusions. SES seems not to be associated with an increased underlying cervical cancer risk, although it cannot be ruled out. Any differences in cervical cancer risk between different SES groups, if present, are therefore probable caused by differences in screening participation.

Keywords. Socioeconomic status; Cervical cancer; Screening; Underlying risk

BACKGROUND

Having a lower socioeconomic status (SES) is considered to be a risk factor for developing cervical cancer(86, 87), which can be (partly) explained by the fact that women with lower SES are less likely to attend screening(90, 204, 205). One of the factors causing this difference in uptake is that lower SES women have a more negative attitude towards screening (ie. lower SES women anticipate a higher distress of screening)(206). In countries without organized screening, a lack of health insurance, and therefore less access to health care, also contributes to this difference in uptake between SES groups(87).

Benard et al. have found that lower education and higher poverty rates are associated with increased vaginal and penile cancer incidence rates(207). As both types of cancers are non-screened human papillomavirus (HPV) associated cancers, it suggests that lower SES is associated with an increased underlying (ie. background) risk of HPV related cancers. If true, and given that there is a causal relationship between the development of cervical cancer and a high-risk HPV infection(4, 5), one would expect that lower SES women also have an increased underlying risk for developing cervical cancer, independent of differences in former and current screening uptake. To test this hypothesis, we compared the cervical cancer rate between first-time attendees without any history of cervical examinations living in low, intermediate, and high SES neighbourhoods (ie. neighbourhood status is indicative of a woman's SES). We thereby eliminated the influence of differences in screening history and uptake between the SES groups (ie. ensuring that a possible risk difference was not caused by differences in former and current screening uptake).

METHODS

Since 1996, the Dutch cervical cancer screening programme consists of 5-yearly cytological screening in women aged 30 to 60 years. The participation rate ranges from 65% to almost 70% (programme only) and the coverage rate ranges from 77% to almost 80% [source: Dutch Network and National Database for Pathology (PALGA)](94). For a detailed screening protocol see Figure 8-1.

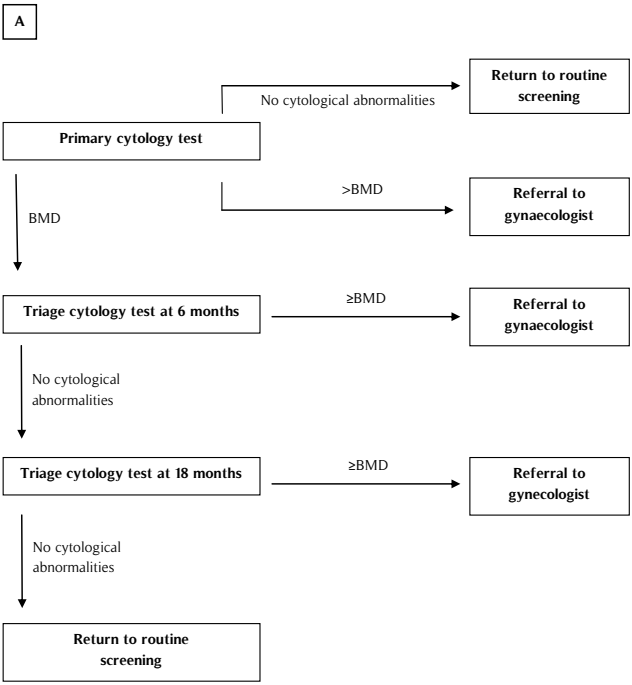

Figure 8-1. Triage protocol consisting of triage cytology (A) without HPV testing, and (B) with HPV testing. HPV = Human papillomavirus; BMD = Borderline and mildly dyskaryotic smears.

From 1990 onward, all cervical cytology and histology tests taken in the Netherlands, within and outside the screening programme, are registered in PALGA. Every woman in the database receives an identification code based on her birth date and the first four letters of her (maiden) family name, which enables linkage of the tests belonging to the same woman, and thus allowing us to follow individual women. However, the linkage is not unique and certainly with common last names, administrative fusions occur quite often. To correct for these false identity matches(208), we excluded test results of women with the 0.5% most common surnames, which corresponds with approximately 30% of the women registered in the PALGA database(167).

Retrospectively, we retrieved information from PALGA on all cervix uteri cytological and histological tests registered from January 2000 until March 2009. To ensure enough follow-up time we only included women with a first episode taken within the national screening programme between January 2000 and December 2007 (ie. almost all cervical cancers are detected within 15 months following an abnormal cervical smear). Histologically confirmed cervical cancer cases were identified by selecting all PALGA records that included corresponding pathology codes.

The status score of the four-digit postal code of the woman's place of residence at the time of the primary test was used to determine the SES of the woman's neighbourhood(95). Status scores per four-digit postal code were provided by the Netherlands Institute for Social Research(96). They were based on: (i) mean income, (ii) percentage of households with a low income, (iii) percentage of households with, on average, a low education, and (iiii) unemployment rate per postal code in 2006. These variables were merged into one score (ie. status score) using principal components analysis(209). Status scores were ranked and low SES was defined as having a status score in the first (ie. lowest) quartile; intermediate SES as having a status score in the second and third quartile; and high SES as having a status score in the fourth (ie. upper) quartile. For our study intake period 2000 to 2007, status

scores were also available based on income, education and employments rates in 2002. In a sensitivity analysis, we used this alternative set of scores.

Statistical analyses

As cervical cancer risk varies by age, it was considered a candidate for confounding(97, 98). Since women are invited in the year they turn 30, 35, .., 60, the woman's age was categorized as: 29-38, 39-48, and 49-63, at the time of the primary smear (ie. first test of the episode). We used a Pearson Chi-Square test to test whether the distribution of age differed between the three SES groups. Thus, we tested whether age was indeed a potential confounder.

We performed logistic regression analyses to compare the cervical cancer risk between the three SES groups, unadjusted and adjusted for confounding factors. The odds ratio (OR) was interpreted as relative risk since the cervical cancer prevalence was lower than 10%(99). The software programme SPSS (version 20) was used to perform the statistical analyses.

RESULTS

Status scores per four-digit postal code ranged from -6.09 to 2.65. The lowest quartile ranged from -6.09 to -0.59 and the upper quartile ranged from 0.60 to 2.65. Thus, low SES corresponded with a status score of ≤-0.59, intermediate SES with a status score of >-0.59 and <0.60, and high SES corresponded with a status score of ≥0.60.

A total of 404,761 first-time attendees without any history of cervical examinations were included in our study; 26.4% lived in a low (106,747 women), 48.2% in an intermediate (195,003 women) and 25.4% in a high SES neighbourhood (103,011 women) (Table 8-1). Cervical cancer rates were similar between the three SES groups (61.8 versus 64.6 versus 60.2 cervical cancer cases per 100,000 women living in low, intermediate and high SES neighbourhoods, respectively) (p = 0.891). First-time attendees living in low SES

Table 8-1. Population characteristics. The distribution of cervical cancer detection and age per socioeconomic status group.

	Low SES	Intermediate SES	High SES	*P* value
No. of women	106,747	195,003	103,011	
No. of cervical cancers (per 100,000 women)	66 (61.8)	126 (64.6)	62 (60.2)	0.891
Age				<0.001
29-38, n (%)	80,532 (75.4)	154,664 (79.3)	82,493 (80.1)	
39-48, n (%)	16,333 (15.3)	24,440 (12.5)	13,140 (12.8)	
49-63, n (%)	9,882 (9.3)	15,899 (8.2)	7,378 (7.2)	

SES = Socioeconomic status

neighbourhoods were more likely to attend at an older age than those in intermediate and high SES neighbourhoods (ie. 25 versus 21 and 20% were aged 39 years or older, respectively). Given these differences, and age being a risk factor for the development of cervical cancer, we considered age as a potential confounding factor.

As compared to first-time attendees living in high SES neighbourhoods, those in low and intermediate SES neighbourhoods had a similar cervical cancer risk [i.e. OR of 1.03 (95% CI: 0.73, 1.45) and 1.07 (95% CI: 0.79,

Table 8-2. Comparing cervical cancer risk between first-time attendees living in different socioeconomic status neighbourhoods. Odds ratios are given, unadjusted and adjusted for differences in the distribution of age.

Model	Low versus High	Intermediate versus High
SES	1.03 (0.73, 1.45)	1.07 (0.79, 1.46)
SES, adjusted for age	0.98 (0.70, 1.39)	1.07 (0.79, 1.45)

The odds ratio can be interpreted as relative risk since the cervical cancer prevalence is less than 10%. The 95% confidence intervals are given in brackets. A *p* value of <0.05 was considered to be statistically significant. No significant differences were detected.

1.46), respectively] (Table 8-2). When adjusted for age, cervical cancer risks were still equivocal between the SES groups [ie. OR of 0.98 (95% CI: 0.70, 1.39) and 1.07 (95% CI: 0.79, 1.45), respectively, when comparing first-time attendees in low and intermediate versus high SES neighbourhoods].

Sensitivity analysis

Status scores from 2002 ranged from -7.21 to 2.62. Low SES corresponded with a status score of ≤-0.79, intermediate SES with a status score of >-0.79 and <0.22, and high SES corresponded with a status score of ≥0.22. We found that the age-adjusted OR of being diagnosed with cervical cancer was 1.10 (95% CI: 0.79, 1.52) and 0.91 (95% CI: 0.67, 1.23) for first-time attendees living in low and intermediate SES neighbourhoods as compared to first-time attendees living in high SES neighbourhoods, respectively.

DISCUSSION

Our results showed no difference in cervical cancer risk between first-time attendees without any history of cervical examination living in low and intermediate versus high SES neighbourhoods. As we eliminated differences in former and current screening uptake, this indicates that the underlying cervical cancer risk is not associated with the SES of the neighbourhood the woman lives in, at least for the women considered: ever attending to screening. Therefore, any differences in cervical cancer risk between different SES groups, if present, are probably caused by differences in screening participation.

This finding contrasted to our expectation that the underlying cervical cancer risk was associated with a woman's SES. As a high-risk HPV infection is a necessary condition for the development of cervical cancer, an additional risk would result from a higher incidence of high-risk HPV infections, a higher probability for a high-risk HPV infection to progress into cancer, or a combination thereof. While a higher incidence can be caused by an

earlier sexual debut and more sexual partners(9-11), a higher probability of progression can be influenced by factors such as smoking, long-term contraceptive use and the incidence of sexual transmitted diseases (STD) (9). Thus, it seems that no significant differences in sexual behaviour and/or these factors were present between women living in different SES neighbourhoods. Another possibility is that differences were present, but that they levelled each other off (eg. the effect of an increased smoking incidence was compensated by the effect of a decreased STD incidence). In addition, the number of cervical cancer cases was such that a maximal 39% increased underlying cervical cancer risk (ie. upper bound 95% CI) for women living in neighbourhoods with a status score within the lowest quartile of all neighbourhoods compared to women living in a neighbourhood with a status score in the highest quartile cannot be ruled out.

Status score of a neighbourhood (ie. defined by postal code) was based on income, education, and employment rates of inhabitants living in that neighbourhood. Obviously, it is possible that high SES women live in low SES neighbourhoods and the other way around. However, if a woman's SES is not associated with her underlying cervical cancer risk difference, as indicated by our results, our effect estimates are unaffected when comparing SES on an individual instead of ecological level. If a woman's SES is inversely correlated with the cervical cancer risk (ie. lower SES is associated with an increased underlying cervical cancer risk) our effect estimates would be underestimated. If positively correlated (ie. lower SES is associated with a decreased underlying risk), our effect estimates would be overestimated when comparing SES on an individual instead of ecological level.

While it is possible that lower SES non-attendees are a different group in terms of sexual behaviour and exposure to other risk factors than lower SES attendees, this probable also applies for higher SES non-attendees versus attendees. Therefore, we expect that our findings (ie. lower SES is not associated with an increased underlying risk in first-time attendees) can be extrapolated to lower SES women in general. However, as in PALGA no data are available

on non-attendees without any cervical examinations taken, performing such a comparison was not possible without risking selection-bias.

SES was defined according to status scores in 2006. As first lifetime cervical examinations were included from 2000 to 2007 this may have caused bias if women were misclassified because of changes in status scores over time. Knol et al. have shown that the average status score per postal code has increased from -0.26 in 1998 to +0.17 in 2010(209). However, we found that differences in status scores over time did not bias our results as our conclusions were still valid when SES was based on 2002 instead of 2006 status scores.

While the association between SES and the underlying cervical cancer risk was not found in this Dutch study, this does not automatically mean it is absent in other countries. For instance, the study of Benard et al. suggested that lower education and higher poverty rates are associated with an increased underlying risk of HPV related cancers, and thus cervical cancer, in the United States.

Although multiple studies have shown that lower SES women have a higher cervical cancer risk and are more reluctant to attend screening, this is the first study that determined whether they still have an increased risk after eliminating differences in screening uptake.

Nonetheless, our study has some limitations. First, liquid-based cytology tests SurePath and ThinPrep have gradually been implemented in most laboratories processing primary screening tests. As the few studies which compared cervical cancer detection rates between liquid-based cytology and conventional cytology did not find significant differences(49, 58, 117), we believe it is unlikely that possible differences in the distribution of the type of cytology test used biased our results. Second, no data are present whether cytology triage testing at six months was combined with HPV testing, and whether this was correlated with SES. Even so, Siebers et al. have shown that additional HPV triage testing shortens follow-up without altering the

detection of cervical intraepithelial neoplasia grade III or cancer(39). Third, we could only include cervix uteri cytological and histological tests taken until March 2009 as it was impossible to correct for false identity matches thereafter. However, we believe it is highly unlikely that our conclusion is time-dependent.

Our results do not show that SES is associated with a woman's underlying cervical cancer risk, at least for the women considered: ever attending to screening. Therefore, any differences in cervical cancer risk between different SES groups in the Netherlands, if present, are probably caused by differences in screening participation.

ACKNOWLEDGEMENTS

This is an independent report commissioned and funded by Dutch National Institute for Public Health and the Environment (007/12 V&Z NvdV/EM). We wish to thank Jurjen Iedema from the Netherlands Institute for Social Research for his advice on status scores and socioeconomic status.

CHAPTER 9

General discussion

In this thesis, we studied the impact of recent and future implementation of new screening tests in the Dutch cervical cancer screening programme. First, we will focus on the effect of using liquid-based cytology (LBC) tests SurePath and ThinPrep as compared to conventional cytology. Second, we will focus on offering human papillomavirus (HPV) self-sampling to non-attendees in order to increase attendance rates. Thereafter, we will discuss monitoring of the future programme and future directions of Dutch cervical cancer prevention, followed by our conclusions and recommendations.

PART I: PRIMARY LIQUID-BASED CYTOLOGY TESTING

Question 1: What is the effect of using SurePath and ThinPrep as primary test method on CIN II$^+$ detection rates?

Our research showed that the use of SurePath as compared to conventional cytology led to an increased detection of CIN II$^+$ by 8%, although it simultaneously increased the detection of CIN I by 14%. The use of ThinPrep did not affect CIN II$^+$ or CIN I detection rates.

These findings correspond with literature. In chapter 2, we showed that our observed CIN II$^+$ detection rate ratio of 0.99 (95% CI: 0.96, 1.02) was compatible with CIN II$^+$ detection rate ratios of other studies comparing ThinPrep with conventional cytology. In addition, our observed CIN II$^+$ detection rate ratio of 1.08 (95% CI: 1.05, 1.12) when comparing SurePath with conventional cytology was compatible to the only other CIN II$^+$ detection rate ratio provided by literature, although their conclusion differed from ours [ie. they found similar CIN II$^+$ detection rates between these two tests: ratio of 1.01 (95% CI: 0.76, 1.33)].

As described in Table 1-2 there are multiple differences in preparation and protocol present between LBC tests SurePath and ThinPrep which may affect their performances'. First, the protocol concerned with handling of the collecting device is likely to be associated with the cell yield(100, 101)

and therefore, with the probability of transferring abnormal cells from the cervical specimen (if present) to the slide. Differences in this protocol could therefore be the cause of the differences in CIN II$^+$ sensitivity found between the tests. In addition, the difference in the method of cells transfer is possible associated with the risk on an unsatisfactory smear. Zhao et al. suggested that debris like blood and mucus can attach to the filter used for processing ThinPrep samples. As a consequence, this could reduce the number of cells transferred to the slide resulting in for instance more unsatisfactory smears(53). This was confirmed in other studies which simulated the preparation of smears of cervical samples containing excess blood and mucus(210, 211). They found that SurePath was able to handle larger amounts than ThinPrep. These results correspond with our own data. We found that the unsatisfactory rate was higher when using ThinPrep than using SurePath [odds ratio (OR) of 1.41 (95% CI: 1.38, 1.45)] (data not shown). Whether the method of cell transfer is associated with the sensitivity to detect CIN II$^+$ is currently unknown, as is the effect of using different types of preservative fluid and collecting devices. Thus, more research is needed to examine which differences in preparation of the smears cause the difference in CIN II$^+$ sensitivity between the different types of LBC tests. Then, it might be possible to adjust the ThinPrep protocol in order to improve its performance.

It is not only meaningful to know whether the use of SurePath or ThinPrep affected CIN II$^+$ detection rates, but also whether it affected the ability of cytologists to distinguish between women with and without a CIN II$^+$ lesion. This is important as the aim of screening is to prevent as many cervical cancers against as little inconvenience for the women as possible. When comparing SurePath with conventional cytology, we found that the finding of cell abnormalities (ie. Pap $\geq$ 2) more often resulted in the detection of CIN II$^+$ detection rates [ie. OR of 1.11 (95% CI: 1.07, 1.15)]. Thus, the positive predictive value (PPV) of a positive smear was higher and cytologists were better able to distinguish between women with and without CIN II$^+$ lesions when using SurePath. When comparing ThinPrep with conventional

cytology, the PPV was similar or even lower [ie. OR of 0.97 (95% CI: 0.93, 1.00)].

Question 2: Are increasing CIN detection rates caused by implementation of LBC tests?

We found that from 2003-2005 to 2009, trends of increased detection were present for all CIN grades. After adjusting for demographic factors, CIN I, CIN II, and CIN III detection rates were 2.11, 1.79, and 1.59 times higher in 2009 as compared to 2003, respectively. When also adjusted for the type of cytology test used, detection rates were still 1.90, 1.48, and 1.55 times larger in 2009 as compared to 2003. Thus, the gradual implementation of LBC over time contributed to the increased detection of CIN I, CIN II, and CIN III but it was not the only cause.

CIN I and CIN II detection rates decreased with 0.21 (ie. 2.11 to 1.90) and 0.31 (ie. 1.79 to 1.48) when correcting for the type of cytology test used, which was larger than expected. Based on results from chapter 2, we foresaw a maximal decrease of 0.14 for both CIN I and CIN II (which would have been the decrease if all women switched from using conventional cytology in 2003 to using SurePath in 2009). On the other hand, the 0.04 decrease in CIN III detection rates (ie. 1.59 to 1.55) was as expected, namely below 0.06 [which would have been the decrease if all women switched from using conventional cytology in 2003 to using SurePath in 2009 (Chapter 2)]. This discrepancy on the effect of CIN I and CIN II detection rates is caused by the inclusion of two two-way interaction terms between: (i) the type of primary cytology test and age, and (ii) the type of primary cytology test and screening region. Thus, the effect of implementing SurePath or ThinPrep differs per age group, which was confirmed by Matjka et al.(64). This can be explained by the fact that the sensitivity of conventional cytology is probably age-dependent(83, 212, 213). Moreover, it is likely that the effect of implementing SurePath or ThinPrep differs per laboratory, as we

used screening region as a proxy for the laboratories involved. This may be explained by the following three factors. First, differences in protocol could be present between laboratories. Second, implementation of automated reading in some of the laboratories could have affected their ability to detect CIN lesions when using LBC tests, although studies have shown heterogeneous results(61-64). Third, it is likely that the interobserver agreement of conventional cytology is lower than that of SurePath and ThinPrep as the quality of cell transfer, and therefore the quality of the conventional cytology smear, might differ between clinicians. Thus, how much there is to gain when replacing conventional cytology by another primary test method possibly differs between clinicians and therefore, between laboratories. Also, it is possible for the interobserver agreement of ThinPrep to be lower than that of SurePath as the thoroughness of rinsing the brush in the vial with preservative fluid might differ between clinicians and therefore, the cell yield might also differ(101). A possible explanation for the finding of a larger decrease in CIN I and CIN II detection rates than expected, while the decrease in CIN III detection rates was in line with our expectations, is that the detection of mild cell abnormalities has a lower interobserver agreement than severe cell abnormalities. Therefore, the effect of switching to LBC testing on mild cell abnormalities, and therefore CIN I and CIN II, can differ more per lab than the effect on severe cell abnormalities, and consequently CIN III.

If the effect of switching to either SurePath or ThinPrep differs per laboratory, this would probably mean that we under- or overestimated the contribution of LBC implementation on the trend in increased CIN detection rates because: (i) no data per laboratory were available, and (ii) laboratories switched to either SurePath or ThinPrep on different points in time. Thus, the gradual implementation of LBC over time could have contributed for a larger part to the CIN increase than estimated.

Question 3: Is there a difference in sensitivity to detect progressive CIN lesions between SurePath, ThinPrep, and conventional cytology?

To assess whether the ability to detect progressive CIN lesions differed between SurePath, ThinPrep, and conventional cytology, we compared interval cancer rates (ie. a cervical cancer diagnosed after a negative primary screening smear) between the three types of cytology tests. When adjusted for confounders, the overall risk of interval cancer was 17% lower after a negative Surepath smear and 20% higher after a negative ThinPrep smear than after a negative conventional cytology smear. This strongly suggests that the sensitivity to detect progressive CIN lesions is highest when using SurePath and lowest when using ThinPrep as primary test method.

The use of SurePath seemed more effective in finding progressive CIN lesions than ThinPrep and conventional cytology, both based on differences in CIN II$^+$ detection rates and interval cancer rates (Chapters 2 and 4). Based on differences in CIN II$^+$ detection rates, the use of ThinPrep seemed equally effective as the use of conventional cytology (Chapter 2). However, when using the same data but replacing the detection of CIN II$^+$ by the detection of interval cancers as outcome measure, we found strong indications that ThinPrep was less effective as conventional cytology in finding progressive CIN lesions. Thus, the use of surrogate measures led to the general consensus that ThinPrep was equally effective in preventing cervical cancer by finding equal numbers of CIN II$^+$, while the use of interval cancers showed that fewer cervical cancers were prevented by finding fewer progressive CIN lesions. Meanwhile, ThinPrep has been implemented in multiple countries with and without organized cervical cancer screening programmes, such as the Netherlands, Denmark, the UK, and the USA(92, 93), based on this consensus. Therefore, it is most important to carefully monitor the effect of implementing ThinPrep in these countries.

When implementing new screening tests or strategies one should not only focus on the benefits (ie. how many cervical cancer cases and associated deaths have been prevented) but also on the harms (ie. the loss in quality of life due to screening and associated treatments). For a screening programme to be effective the benefits need to outweigh the harms, which can be expressed by the number of quality-adjusted life-years (QALYs) gained and calculated using microsimulation. A good indicator for the negative effect of screening is the overdiagnosis rate (ie. the number of excess diagnoses when comparing the situation with versus without screening).

Question 4: What is the amount of overdiagnosis in the Dutch cervical cancer screening programme?

In general, more (pre-invasive) disease is found by screening. This excess in diagnosis by screening is defined as overdiagnosis (ie. number of extra diagnoses with screening divided by total number of diagnoses with screening). When assuming that conventional cytology was used as primary test method in the current primary cytological programme, the percentage of excess diagnoses was 70.6% (ie. including all CIN grades). When CIN II, CIN III, and cervical cancer were defined as excess diagnoses, this percentage was 63.2%. When CIN III and cervical cancer were defined as excess diagnoses, this percentage was 50.0%. As cervical cancers are prevented by the diagnosis and treatment of pre-invasive CIN lesions, the incidence of cervical cancer decreased by 55.2% with screening. Thus, when only cervical cancer was defined as excess diagnosis, no overdiagnosis was present. In contrast, breast cancer screening is aimed at finding breast cancer in an earlier stage. Therefore, the percentage of excess diagnoses is 1.5%.

Thus, the use of conventional cytology as primary test method results in an overdiagnosis rate of 70.6%, when including all CIN grades. The use of ThinPrep as primary test method will probably lead to a decreased sensitivity of the programme as the sensitivity to detect progressive CIN lesions is

decreased. Simultaneously, the number of detected CIN lesions is not significantly affected. Thus, the number of detected CIN per prevented cervical cancer will probably rise and therefore the overdiagnosis rate. Hence, the use of ThinPrep as primary test method will probably also lead to a decrease in the QALYs gained and consequently, in a decreased effectiveness of the programme. Therefore, its sensitivity should probable be improved in order to guarantee a similar harm-benefits ratio and effectiveness of the original programme. The use of Surepath as primary test method will probably lead to an increased sensitivity of the programme as more cervical cancers are prevented by detecting more progressive CIN lesions. An increased sensitivity often leads to a decreased specificity which is expressed by the finding of 3.1 extra CIN I (30.1 / 9.7), 3.2 extra CIN II (31.2 / 9.7) and 3.1 extra CIN III (30.3 / 9.7) diagnoses per extra prevented interval cancer (Table 9-1). Thus, the percentage of overdiagnosis will probably increase when using SurePath as primary test method and when including CIN diagnoses in the definition. However, it is possible that part of these extra detected CIN I and CIN II lesions are actually due to introduction of HPV triage testing instead of primary screening with SurePath(39). Especially since it is likely that the introduction of HPV triage testing is correlated to the use of LBC testing (ie. the possibility of co-testing is one of the advantages of LBC over conventional cytology(41, 44)). However, even then, the percentage of overdiagnosis will probably still increase when using SurePath as compared to

Table 9-1. The difference in CIN lesions and interval cancers per 100,000 SurePath and ThinPrep versus 100,000 conventional cytology smears.

	SurePath versus CC	ThinPrep versus CC
Extra CIN I lesions detected	**30.1**	-3.5
Extra CIN II lesions detected	**31.2**	9.4
Extra CIN III lesions detected	**30.3**	-12.2
Extra interval cancers prevented	**9.7**	**-11.1***

Bold = Significant. A *p* value of <0.05 was considered to be statistically significant. CC = Conventional cytology.

*Thus, the use of ThinPrep leads to 11.1 extra interval cancers per 100,000 smears.

conventional cytology. Whether this will result in a decrease in the number of QALYs gained and therefore in a decreased effectiveness of the Dutch cervical cancer screening programme is unknown.

Microsimulation can be used to estimate the impact of using SurePath as compared to conventional cytology as primary test method on the effectiveness of the Dutch cervical cancer screening programme using: (i) observed short-term indicators (eg. number of women referred to triage, number of women referred directly to the gynaecologist, and the number of detected CIN lesions), and (ii) long-term outcomes (eg. interval cancer rates). In addition, sensitivity analyses can be included as regards to the cause (of part) of the extra detected CIN I and CIN II lesions. Moreover, cost-effectiveness analyses can be performed to assess whether SurePath is cost-effective besides being effective. In an earlier microsimulation study it was shown that using LBC is cost-effective if the sensitivity is at least 3 to 5 percentage points higher than conventional cytology for CIN I, CIN II, CIN III, cervical cancer FIGO stage 1A, 1B, and 2^{+}(214). However, we found a similar sensitivity to detect cervical cancer, while the sensitivity to detect CIN II and CIN III lesions is 1.14 and 1.06 times higher, leading to 17% fewer interval cancers. Thus, we recommend performing new microsimulations, based on our recent findings, in order to assess whether the use of SurePath as primary test method is (cost-)effective or not.

Additional analyses

In light of all the new evidence gathered within this first part of this thesis, we performed two additional analyses. First, we were interested in whether the sensitivity of the programme was affected by implementing screening tests SurePath and ThinPrep. We compared the risk on an interval cancer after a negative primary screening test (which is an important indicator for the sensitivity of the programme) between the periods 2006-onward and 2000-2005 using Cox regression analyses. We did not find any changes in risk over time [the ratio was 0.92 (95% CI: 0.84, 1.02)]. Thus, we found no

indications that the interval cancer risk or the sensitivity of the programme was affected by gradual implementation of SurePath and ThinPrep over time. However, it cannot be ruled out that the sensitivity of the programme was altered while data to support this are not (yet) present. As the distribution of both LBC tests was similar until 2006 (Figure 9-1), it is possible that the negative effect of using ThinPrep on the sensitivity of the programme was levelled off by the use of SurePath. In addition, the effect of ThinPrep is time-dependent and the risk on an interval cancer is only increased after approximately four years. Thus, the first negative effects on the screening programme would appear somewhere in 2011. As data until March 2013 are present and interval cancers are rare, it is possible that not enough data are present yet to observe an increase in interval cancers. In addition, a decrease in the underlying risk of screened women could in theory have biased our results, although our data support the contrary (Chapter 3). Importantly,

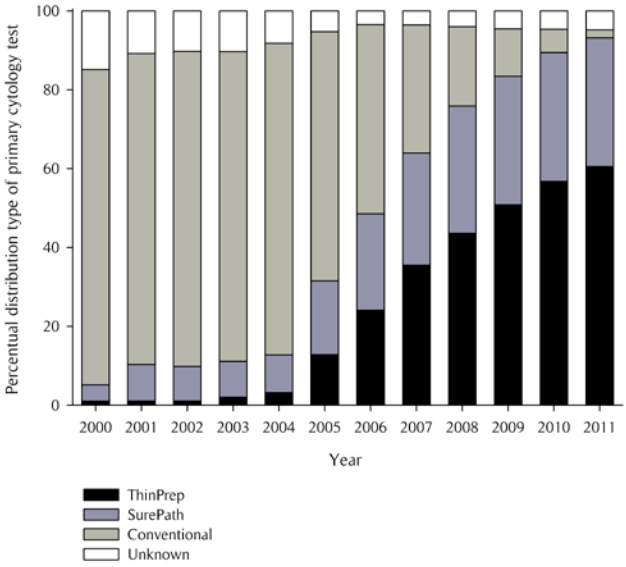

Figure 9-1. Percentual distribution of the type of primary cytology test used in the Dutch cervical cancer screening programme.

even when the sensitivity of the programme did not decrease over time, it is likely that the number of QALYs gained by screening did decrease because of: (i) the increase in women being referred to triage, (ii) the increase in women being referred directly to the gynaecologist, and (iii) the increase in CIN detection rates (Chapter 3). Thus, even when the sensitivity of the programme did not alter, the effectiveness of the programme probably did.

In our second additional analysis, we assessed whether the cervical cancer incidence rate in the Netherlands increased, as was suggested in Figure 1-1, using the Joinpoint Regression Programme(108). We found that the incidence rate non-significantly decreased by 1.2% (95% CI: -2.4, +0.0) annually from 1989-1998, followed by a non-significant decrease of 5.9% (95% CI: -17.8, +7.7) from 1998-2001 and a significant increase of 1.5% (95% CI; +0.7, +2.3) from 2001-2013. As it is not likely that this increase is caused by an increase in interval cancer rates, as shown earlier, it is probable caused by an increase in screen-detected and/or clinically detected cancer rates. Our finding that cervical cancer detection rates were nonsignificantly lower when using ThinPrep as compared to conventional cytology, and unaffected when using SurePath (Chapter 2), does not support the theory that screen-detected cancer rates increased over time. However, it is possible that the underlying cervical cancer risk of the general Dutch population increased, which was also pointed out in Chapter 3. This could have caused an increase in both screen-detected and clinically detected cancers. Comparison of the prevalence of high-risk HPV infections over time in the eligible Dutch screening population could indicate whether this is the case or not. Another factor which could have contributed to a possible increase in both screen-detected and clinically detected cancers is compensation of the rapid decrease in cervical cancer incidence from 1998 to 2001, which was caused by an increased screening intensity from 1996 to 1998(109). However, we we doubt it was the main cause for a 12-year period of increase. As cervical cancer screening programme coverage rates were stable from 2004 to 2013(79, 80), they can be excluded as possible explanation for a possible increased clinically detected cancer rate.

PART II: INCREASING ATTENDANCE BY OFFERING HPV SELF-SAMPLING TO NON-ATTENDEES

A study examining the screening history of Dutch women with cervical cancer revealed that more than half of the cervical cancers occurred in women who did not participate in screening in the previous six years while they were invited. In addition, less than 10% of the women had a previous negative screening smear within the previous six years(74). Thus, according to this study increasing the attendance has more potential in preventing cervical cancers than implementing a more sensitive screening test.

Question 5: How many cervical cancer cases in young women can potentially be prevented using a more sensitive screening test at age 30?

We analyzed the screening history of 30 to 35 year old women diagnosed with cervical cancer between 2004 and March 2009. Accordingly, we assessed the percentage of cervical cancer cases that were preceded by a negative cytology test under the age of 35, which is the age of the second screening round, and the percentage of cervical cancer cases without a history of cervical cancer screening. If we assume that the lesion was already present 10 years before the cancer was diagnosed, 23% of the cancers can maximally be prevented by using a more sensitive screening test under the age of 35, while 67% of the cancers occurred in women who did not have a history of cervical cancer screening. When assuming that the lesion was already present 5 years before the cancer was diagnosed, these percentages were 18% and 67%, respectively. Thus, the majority of cancers in women aged 30-35 years would still occur when applying a more sensitive screening test at first screening age 30, although increasing the attendance rate could make an important contribution in preventing them.

The study of Gok et al. showed that offering HPV self-sampling to non-attendees of Dutch cervical cancer screening is a promising method to increase participation rates, of invited women in general and of young women at their first screening age(73, 180). Thus, offering self-sampling to non-attendees could be an effective way in reducing the incidence and mortality of cervical cancer. As HPV testing can be performed on self-collected samples(69, 70, 215) and primary HPV testing is expected to replace primary cytology testing in the Dutch cervical cancer screening programme from 2016 onward, we estimated we estimated under which circumstances it would be effective (ie. QALYs are gained) to offer self-sampling to non-attendees of organized primary HPV screening.

Question 6: When is it effective to offer self-sampling to non-attendees of organized primary HPV screening?

Without "switching" of regular attendees from HPV office-based to HPV self-sampling, offering self-sampling is (cost-) effective under every studied condition. If all regular attendees switch, offering self-sampling is (cost-)effective if simultaneously the attendance increases by at least 6 percentage points, women with at least a 1.7 higher underlying risk are recruited by self-sampling and the relative CIN II$^+$ sensitivity and specificity of HPV self-sampling (as compared to office-based sampling) are at least 0.95. If the relative sensitivity decreases to 0.90 combined with either the absence of these higher underlying risk women or a 3 percentage points extra attendance, the effectiveness of the programme already decreases when more than 20% to 30% of the regular attendees switch.

The following variables are most influential in determining whether offering HPV self-sampling to non-attendees is effective or not: the relative CIN II$^+$ sensitivity of HPV self-sampling, the increase in attendance, the ability to target higher underlying risk non-attendees, and the percentage of regular attendees switching from office-based to HPV self-sampling.

The largest uncertainty about offering HPV self-sampling to non-attendees is how many women will switch from office-based to HPV self-sampling. As we believe it is likely that women will actually switch, it is important that the following conditions are met in order to ensure an equal or increased effectiveness of the future HPV programme when offering HPV self-sampling to its non-attendees: (i) the relative CIN II$^+$ sensitivity is at least 0.95, (ii) women with at least a 1.7 higher underlying risk are recruited by self-sampling, and (iii) the total attendance increases by at least 6 percentage points.

A recent meta-analysis has shown that a relative sensitivity of 1.00 can be met when using a validated polymerase chain reaction (PCR) test(185). However, as these results are uncertain (based on its confidence intervals) it is possible that the relative sensitivity of a validated PCR is lower than 0.95. Moreover, as many of these studies were done in clinical settings, these relative sensitivities may be overestimated and they may even be lower than 0.90 (ie. the real-world performances of HPV self-sampling might be lower than the performances of HPV self-sampling in the artificial setting of clinical studies)(202). Thus, whether the first criterion will be met is uncertain. However, even when the relative CIN II$^+$ sensitivity is 0.90, 70% of the regular attendees could switch before the effectiveness of the programme is decreased, as long as the second and third condition are met (Chapter 7). Therefore, women with at least a 1.7 higher underlying risk need to be identified and recruited by self-sampling and the total attendance needs to increase by at least 6 percentage points.

Question 7: Is lower socioeconomic status associated with an increased underlying cervical cancer risk?

We found no difference in cervical cancer risk between first-time attendees without any history of cervical examination living in low and intermediate versus high SES neighbourhoods. As we eliminated differences in former and current screening uptake, this indicates

> that the underlying cervical cancer risk is not associated with the SES of the neighbourhood the woman lives in, although it cannot be ruled out. As the SES of a neighbourhood is based on income levels, education levels and employment rates of its inhabitants, we expect most women to have similar SES statuses as the neighbourhood they live in. Therefore, we conclude that not only women living in lower SES neighbourhoods, but also lower SES women, probably do not have an increased underlying risk to be diagnosed with cervical cancer.

The number of cervical cancer cases was such that a maximal 39% increased underlying cervical cancer risk (ie. upper bound 95% CI) for low as compared with high SES women cannot be ruled out (Chapter 8). Compared to the general population this maximal increase will probably even be lower. Thus, the second criterion (ie. women with at least a 1.7 higher underlying risk as compared to the general population are recruited by self-sampling) will not be met by focussing on lower SES non-attendees.

Another important target group for offering HPV self-sampling can be underscreened and unscreened women, since it was indicated that at least a part of them have an increased underlying risk(190). Results of the Dutch PROHTECT trial showed that these women indeed attended via self-sampling when it was offered to them(176), although it is unknown whether this is the subset with an increased underlying risk. Moreover, self-sampling was offered via an opt-out procedure (ie. a self-sampling kit was sent to all non-attendees except when they opted-out via a letter), while the use of an opt-in procedure (ie. involving a request for a self-sampler) may reduce response rates (71). This would mean that fulfilment of the second criterion (ie. the recruitment of women with a 1.7 higher underlying risk) also depends on the chosen strategy. Thus, whether the second criterion will be met is uncertain. Then, the effectiveness of the programme decreases when more than 30% of the regular attendees switch. If in addition the first criterion is also not met, this percentage decreases to 20%. If the cost-effectiveness is also taken into account, any woman switching will result in a decreased effectiveness of the programme (Chapter 7).

Results of the Dutch PROHTECT trial have also shown that the extra attendance rate is at least 6 percentage points when self-sampling was offered to non-attendees(73). Therefore, fulfilment of the third criterion seems realistic, although this may also depend on the chosen strategy (ie. opt-in or opt-out). In case the criterion is not met, the effectiveness of the programme decreases when more than 50% of the regular attendees switch. If in addition, the first criterion is not met, this percentage decreases to 30%. When all three criteria are not met, this percentage decreases to less than 20%. Furthermore, it will be questionable whether it is cost-effective to offer HPV self-sampling to non-attendees even if no women switch at all (Chapter 7).

In conclusion, it is unknown whether all conditions for offering HPV self-sampling to be (cost-)effective are met in real-practice. Therefore, it is essential to carefully monitor the effects, both on the short term and on the long term, in order to guarantee equal or increased effectiveness when adding HPV self-sampling to the programme. Short term effects are defined as the extra attendance via HPV self-sampling, the percentage of regular attendees switching, the findings of self-sampling versus office-based sampling (eg. the number of women referred to triage, the number of women who comply to the given advice, CIN I, and CIN II^{+} detection rates), and the attendance of underscreened and unscreened women. Long term effects are defined as interval cancer incidence (which is an important indicator for the sensitivity of the programme) and mortality rates.

Future cervical cancer screening programme

It is expected that from 2016 onward, primary cytology screening will be replaced by primary HPV screening, combined with an extended screening interval from five to ten years at the ages of 40, 50, and 60. This decision was based on microsimulations which showed a mortality reduction of approximately 11%, while saving costs, when comparing the future with the current programme(216). However, assumptions of the HPV test characteristics could be overestimated as they were based on premature information, while at the same time an increased underlying risk of the screening population

cannot be ruled out. In combination with extending the screen interval to ten years, it is therefore essential to carefully monitor and evaluate the short and long term effects in order to guarantee equal or increased sensitivity and effectiveness of the future programme. Short term effects are defined as the number of women referred for triage testing, false-positive referrals, CIN I, and CIN II^{+} detection rates. Long term effects are defined as interval cancer incidence and mortality rates. In general, careful monitoring and evaluation of the effects of implementing new screening tests or new screening strategies is necessary to guarantee the effectiveness of the programme. Especially, as such decisions are often based on microsimulations and therefore, on assumptions which might differ from real-life settings. For instance, implementation of colon cancer screening in the Netherlands resulted in a referral rate for a colonoscopy of more than 12%, which was much higher than anticipated(217, 218). By carefull monitoring and evaluation of the programme there could be intervened in time (namely an increase of the threshold from 88 ng/mL to 275 ng/mL) in order to ensure that QALYs were still gained and screening for colon cancer was still effective.

Future directions of cervical cancer prevention

Since 2009, HPV vaccination has been implemented in the Dutch National Immunisation Programme. Twelve year old girls are vaccinated with the bivalent Cervarix vaccine against high-risk HPV types 16 and 18, while a catch-up campaign was organized for girls born between 1993 and 1996(75). Vink et al. estimated that the average HPV-16 reduction will be 44% for 20-30 aged women, 15% for 30-40 aged women, 4% for 40-50 aged women and 1% for 50-60 aged women by 2029, assuming current circumstances (ie. a coverage of 60% under 12-year old girls)(219). By 2059, these average HPV-16 reductions were estimated to be 42, 15, 26, and 39%, respectively(219). Estimated reductions for HPV-18 in the Netherlands are not (yet) available. Naturally, increasing the coverage in 12-year old girls would improve these numbers. Another solution could be the additional vaccination of 12-year old boys. Bogaards et al. have estimated that this would result in the prevention of one extra cancer case in males (ie. anal, penile, or oropharyngeal

cancer) per 795 boys vaccinated(220), while girls and women experience an increased protection against cervical cancer via herd-immunity(219). Furthermore, as men have a wider age preference for their sexual partners as compared to women(221, 222), older age cohorts will also be protected(219). However, as 35% of cervical cancers are estimated to be caused by oncogenic types other than HPV-16 and HPV-18(8), cervical cancer screening could still be effective. Whether this would be cost-effective and which screening strategy should be used needs to be determined. It probably depends on the vaccination coverage, herd-immunity levels, and distribution of unvaccinated women in the eligible screening population.

CONCLUSIONS

- The effect of using liquid-based cytology as primary test method for cervical cancer screening depends on the type of liquid-based cytology test used. Our results strongly suggest that SurePath has a higher sensitivity to detect progressive CIN lesions as compared to conventional cytology, accompanied by an increased detection of CIN I, while ThinPrep has a lower sensitivity to detect progressive CIN lesions.
- The gradual implementation of liquid-based cytology was responsible for a small but significant part of the increase in CIN I, CIN II, and CIN III detection rates as observed in the Dutch cervical cancer screening programme. Therefore, other factors must also have played an important role.
- More than 70% of the cervical cancers diagnosed within young women aged 30 to 35 years could not have been prevented by using a more sensitive screening test at the first screening at age 30. As most of these cervical cancers occurred in women who did not have a history of cervical cancer screening, increasing the attendance rate for screening at age 30, for instance by offering non-attendees HPV self-sampling, can make an important contribution in preventing cervical cancers within young women.

- Lower socioeconomic status women do not seem to have an increased underlying cervical cancer risk to be diagnosed with cervical cancer, although it cannot be ruled out.
- The following variables are most influential in determining whether offering HPV self-sampling to non-attendees is effective: the relative CIN II$^+$ sensitivity of HPV self-sampling, the extra attendance rate generated by HPV self-sampling, the ability to target higher underlying risk non-attendees, and the percentage of regular attendees switching from office-based to HPV self-sampling. In the absence of switching, offering self-sampling is effective under every studied condition. If all regular attendees would switch, offering self-sampling is effective if simultaneously (i) the attendance rate increases by at least 6 percentage points, (ii) women with at least a 1.7 higher underlying risk are recruited by self-sampling, and (iii) the relative CIN II$^+$ sensitivity of HPV self-sampling as compared to office-based sampling is at least 0.95.
- The effect of using new screening tests on preventing cervical cancers can be estimated incorrectly when using surrogate endpoints such as CIN detection rates instead of interval cancer rates.

RECOMMENDATIONS

- If the decision which primary cytology test to use depends on the ability to prevent cervical cancer cases, the use of SurePath in the Dutch cervical cancer screening programme is to be preferred.
- The sensitivity of ThinPrep to detect progressive CIN lesions should be improved in order to guarantee an equal harm-benefits ratio and effectiveness of the original Dutch cervical cancer screening programme.
- It is unknown whether all conditions for offering HPV self-sampling to be (cost-)effective are met in real-practice. Therefore, it is essential to carefully monitor the effects, both on the short term and on the long term, in order to guarantee the effectiveness of the Dutch cervical

cancer screening programme when offering HPV self-sampling to non-attendees.

- Careful monitoring of the effects of implementing new screening tests or new screening strategies using population-based data is necessary to guarantee the effectiveness of the programme.

CHAPTER 10

Summary

Samenvatting

SUMMARY

Cervical cancer is the fourth most common cancer in women all over the world, mainly affecting young women between 40 and 55 years. It was estimated that 528,000 women were diagnosed with cervical cancer worldwide in 2012 and 266,000 women died because of it. In the Netherlands, cervical cancer incidence and mortality have decreased towards 5.9 and 1.4 cases per 100,000 women in 2013, standardized to the world's age distribution.

Cervical intraepithelial neoplasia (CIN) are considered as asymptomatic prestadia of cervical cancer and they are ranked to the severity of the lesion (CIN I, II, or III). This preclinical detectable phase of cervical cancer is estimated to last on average more than 10 years which makes it ideal for screening. In the Netherlands cervical cancer screening already exists since the 1970s. Since 1996, women are invited every five years from the ages of 30 to 60 years to be tested by primary cytology and cytology triage, with or without addition of human papillomavirus (HPV) testing.

PRIMARY LIQUID-BASED CYTOLOGY TESTING

Within the last 10 to 15 years, most Dutch laboratories processing primary screening tests switched from using conventional cytology to liquid-based cytology (LBC) tests SurePath or ThinPrep. It was believed that the CIN II$^+$ sensitivity of LBC was similar to that of conventional cytology, although no studies have been published comparing CIN detection rates between different types of LBC tests and conventional cytology. In **chapter 2,** we demonstrated that the use of SurePath as primary test method led to an increased detection of CIN II$^+$ by 8%, although it simultaneously increased the detection of CIN I by 14%. The use of ThinPrep did not affect CIN detection rates. In **chapter 3**, we showed that from 2003-2005 to 2009 trends of increased detection were apparent for all CIN grades. Although implementation of LBC tests SurePath and ThinPrep contributed to this trend, it was

not the only cause. Therefore, an increased underlying risk cannot be ruled out. However, as it is likely that the effect of switching to either SurePath or ThinPrep differs per laboratory, and no data per laboratory were available, it is possible that we underestimated the contribution of LBC implementation on the trend in increased CIN detection rates. Based on results from chapter 2, one could argue that SurePath seems more effective in preventing cervical cancers than ThinPrep and conventional cytology (ie. because of their increased CIN II^+ detection), while ThinPrep and conventional cytology seemed equally effective (ie. because of their similar CIN II^+ detection). However, as in the absence of screening and associated treatment only a fraction of CIN lesions would progress to cervical cancer, detecting an equal or increased CIN rate is not necessarily equivalent to preventing equally or increased numbers of cervical cancers. Indeed, in **chapter 4** we showed that the effect of using new screening tests on preventing cervical cancers can be estimated incorrectly when using surrogate endpoints such as CIN detection rates instead of interval cancer rates (ie. cervical cancer diagnosed after a negative primary screening smear). The risk of an interval cancer was lowest when using SurePath as primary test method and highest when using ThinPrep as primary test method. This difference in interval cancer rates strongly suggests that the sensitivity to detect progressive CIN lesions is highest when using SurePath and lowest when using ThinPrep as primary test method. Thus, SurePath is likely to be more effective in preventing cervical cancers than conventional cytology, which in turn is likely to be more effective than ThinPrep. However, when implementing new screening tests or strategies one should not only focus on the benefits (how many cervical cancers have been prevented) but also on the harms (the loss in quality of life due to screening and associated treatments). A good indication for the negative effect of screening is the overdiagnosis rate, which is defined as: the number of excess diagnoses when comparing the situation with screening versus without screening divided by the number of diagnoses with screening. In **chapter 5**, we estimated the overdiagnosis rate to be 70.6% in the current programme, when including all CIN grades and assuming that conventional cytology is used as primary test method. The use of ThinPrep

as primary test method will probably lead to a higher overdiagnosis rate as fewer cervical cancers are prevented while the number of detected CIN lesions are not affected. Therefore, it is likely that fewer quality-adjusted life-years (QALYs) are gained and thus the effectiveness of the programme decreases by using ThinPrep as primary test method. The use of SurePath as primary test method will probably also lead to a higher overdiagnosis rate as we estimated that 3.1 extra CIN I, 3.2 extra CIN II, and 3.1 extra CIN III lesions are detected per extra prevented interval cancer. In what way this will affect the number of QALYs gained and thereby the effectiveness of the Dutch cervical cancer screening programme needs to be estimated using microsimulations.

INCREASING ATTENDANCE BY OFFERING HPV SELF-SAMPLING TO NON-ATTENDEES

A study examining the screening history of Dutch women with cervical cancer revealed that more than half of the cervical cancers occurred in women who did not participate in screening in the previous six years. In addition, less than ten percent of the women had a previous negative screening smear within the previous six years. This was confirmed in **chapter 6** where we found that the majority of cancers in women aged 30-35 would still occur when applying a more sensitive screening test at first screening age 30, while part of them could have been prevented by increasing the attendance rate. A promising method to enhance screening participation rates is by offering HPV self-sampling to non-attendees of primary HPV screening, which is expected to be implemented in the Netherlands from 2016 onward. However, this could result in a decrease of the effectiveness of the programme if too many regular attendees would "switch" from office-based sampling to self-sampling, given a loss in CIN II^{+} detection (ie. more loss to follow-up and a possible lower self-sampling sensitivity). In **chapter 7**, we showed that as long as women do not switch it is effective to offer self-sampling to non-attendees in every scenario studied. Otherwise, there has to be ensured

that (i) the relative CIN II$^+$ sensitivity is at least 0.95, (ii) women with at least a 1.7 higher underlying risk are recruited by self-sampling, and (iii) the total attendance increases by at least 6 percentage points. In **chapter 8**, we found no association between socioeconomic status (SES) and the underlying cervical cancer risk, although an 39% increased underlying cervical cancer risk for low as compared with high SES women cannot be ruled out. However, compared to the general population this maximal increase will probably be lower. Thus, the second criterion (ie. women with at least a 1.7 higher underlying risk as compared to the general population are recruited by self-sampling) will not be met by focussing on lower SES non-attendees. As we believe it is likely that women will switch, and because it is unknown whether all conditions will be met in real-practice, it is essential to carefully monitor the effects, both on the short term and on the long term, in order to guarantee equal or increased effectiveness when adding HPV self-sampling to the programme.

CONCLUSIONS

- The effect of using liquid-based cytology as primary test method for cervical cancer screening depends on the type of liquid-based cytology test used. Our results strongly suggest that SurePath has a higher sensitivity to detect progressive CIN lesions as compared to conventional cytology, accompanied by an increased detection of CIN I, while ThinPrep has a lower sensitivity to detect progressive CIN lesions.
- The gradual implementation of liquid-based cytology was responsible for a small but significant part of the increase in CIN I, CIN II, and CIN III detection rates as observed in the Dutch cervical cancer screening programme. Therefore, other factors must also have played an important role.
- More than 70% of the cervical cancers diagnosed within young women aged 30 to 35 years could not have been prevented by using a more sensitive screening test at the first screening at age 30. As most of these

cervical cancers occurred in women who did not have a history of cervical cancer screening, increasing the attendance rate for screening at age 30, for instance by offering non-attendees HPV self-sampling, can make an important contribution in preventing cervical cancers within young women.

- Lower socioeconomic status women do not seem to have an increased underlying cervical cancer risk to be diagnosed with cervical cancer, although it cannot be ruled out.
- The following variables are most influential in determining whether offering HPV self-sampling to non-attendees is effective: the relative CIN II$^+$ sensitivity of HPV self-sampling, the extra attendance rate generated by HPV self-sampling, the ability to target higher underlying risk non-attendees, and the percentage of regular attendees switching from office-based to HPV self-sampling. In the absence of switching, offering self-sampling is effective under every studied condition. If all regular attendees would switch, offering self-sampling is effective if simultaneously (i) the attendance rate increases by at least 6 percentage points, (ii) women with at least a 1.7 higher underlying risk are recruited by self-sampling, and (iii) the relative CIN II$^+$ sensitivity of HPV self-sampling as compared to office-based sampling is at least 0.95.
- The effect of using new screening tests on preventing cervical cancers can be estimated incorrectly when using surrogate endpoints such as CIN detection rates instead of interval cancer rates.

RECOMMENDATIONS

- If the decision which primary cytology test to use depends on the ability to prevent cervical cancer cases, the use of SurePath in the Dutch cervical cancer screening programme is to be preferred.
- The sensitivity of ThinPrep to detect progressive CIN lesions should be improved in order to guarantee an equal harm-benefits ratio and effectiveness of the original Dutch cervical cancer screening programme.

- It is unknown whether all conditions for offering HPV self-sampling to be (cost-)effective are met in real-practice. Therefore, it is essential to carefully monitor the effects, both on the short term and on the long term, in order to guarantee the effectiveness of the Dutch cervical cancer screening programme when offering HPV self-sampling to non-attendees.
- Careful monitoring of the effects of implementing new screening tests or new screening strategies using population-based data is necessary to guarantee the effectiveness of the programme.

SAMENVATTING

Baarmoederhalskanker is de vierde meest voorkomende kanker bij vrouwen wereldwijd en het treft met name jonge vrouwen tussen de leeftijd van 40 en 55 jaar. In 2012 werd wereldwijd bij circa 528.000 vrouwen de diagnose baarmoederhalskanker gesteld en circa 266.000 vrouwen overleden aan de gevolgen hiervan. In Nederland zijn de baarmoederhalskanker incidentie en mortaliteit gedaald naar 5,9 en 1,4 gevallen per 100.000 vrouwen in 2013, gestandaardiseerd naar de wereld leeftijdsdistributie.

CINnen, oftewel cervical intraepithelial neoplasia, worden beschouwd als asymptomatische prestadia van baarmoederhalskanker die gerangschikt zijn naar de ernst van de lesie (CIN I, II en III respectievelijk). Deze preklinische, maar detecteerbare fase van baarmoederhalskanker duurt volgens schattingen gemiddeld meer dan 10 jaar, wat het ideaal maakt voor screening. In Nederland is screening naar baarmoederhalskanker daarom alreeds in de jaren 70 geïntroduceerd. Sinds 1996 worden vrouwen tussen de 30 en 60 jaar elke 5 jaar uitgenodigd voor primaire cytologie screening met cytologie triage, eventueel gecombineerd met het testen op humaan papillomavirus (HPV).

DUNNELAAGCYTOLOGIE ALS PRIMAIRE TEST METHODE

In de afgelopen 10 tot 15 jaar zijn de meeste Nederlandse laboratoria die primaire screeningstesten verwerken geswitcht van conventionele cytologie naar dunnelaagcytologie (DLC) ("liquid-based cytology") testen SurePath en ThinPrep. Lange tijd werd aangenomen dat de CIN II$^+$ sensitiviteit van DLC gelijk was aan dat van conventionele cytologie, alhoewel tot nu toe geen studies zijn gepubliceerd die CIN detectie rates tussen verschillende DLC testen en conventionele cytologie hebben vergeleken. In **hoofdstuk 2** hebben we laten zien dat het gebruik van SurePath als primaire test methode leidde tot een 8% toename in CIN II$^+$ detectie rates, alhoewel het tegelijker-

tijd leidde tot een 14% toename in de detectie van CIN I. Het gebruik van ThinPrep had geen effect op de CIN I en CIN II^{+} detectie rates. In **hoofdstuk 3** toonden we aan dat een trend in toenemende CIN detecties plaatsvond van 2003-2005 tot 2009. Alhoewel de implementatie van DLC testen SurePath en ThinPrep heeft bijgedragen aan deze trend, was het niet de enige oorzaak. Een verhoogd achtergrondrisico kan hierdoor niet worden uitgesloten. Echter, het is mogelijk dat we de bijdrage van DLC implementatie hebben onderschat, aangezien het aannemelijk is dat het effect van switchen naar Surepath of ThinPrep verschilt per laboratorium en geen data per laboratorium beschikbaar zijn. Gebaseerd op resultaten van hoofdstuk 2 lijkt SurePath meer effectief in het voorkomen van baarmoederhalskanker dan ThinPrep en conventionele cytologie, terwijl ThinPrep en conventionele cytologie even effectief lijken. Echter, slechts een fractie van de CIN lesies zou zonder screening en de bijbehorende behandelingen doorgegroeid zijn tot baarmoederhalskanker. Het detecteren van een gelijk of toenemed aantal CINnen staat dus niet perse gelijk aan het voorkomen van evenveel of meer baarmoederhalskankers. Dit werd bevestigd in **hoofdstuk 4** waar we hebben aangetoond dat het effect van nieuwe screeningstesten verkeerd kan worden ingeschat indien gebruik wordt gemaakt van surrogaat eindpunten, zoals CIN detectie rates, in plaats van intervalkanker rates (een baarmoederhalskanker gediagnosticeerd na een negatief primair screeningsuitstrijkje). Het risico op een intervalkanker is het laagst bij gebruik van SurePath als primaire test methode en het hoogst bij gebruik van ThinPrep als primaire test methode. Dit verschil in intervalkanker rates suggereert ten zeerste dat de sensitiviteit om progressieve CIN lesies te detecteren het hoogst is wanneer gebruik wordt gemaakt van SurePath en het laagst wanneer gebruikt wordt gemaakt van ThinPrep als primaire test methode. Daarom is de kans groot dat SurePath meer effectief is in het voorkomen van baarmoederhalskankers dan conventionele cytologie, wat op zijn beurt weer meer effectief is dan ThinPrep. Echter, de focus moet niet alleen liggen op de voordelen van het implementeren van nieuwe screeningstesten of strategieën (het aantal voorkomen baarmoederhalskankers), maar ook op de nadelen (het verlies in kwaliteit van leven door screening en bijbeho-

rende behandelingen). Een goede indicator voor het negatieve effect van screening is de mate van overdiagnose, oftewel het aantal exces diagnoses in de situatie met screening versus de situatie zonder screening gedeeld door het aantal diagnoses met screening. In **hoofdstuk 5** hebben we de mate van overdiagnose in het huidige programma geschat op 70,6%, indien alle CIN grades worden meegeteld en conventionele cytologie als primaire test methode wordt gebruikt. Het gebruik van ThinPrep leidt waarschijnlijk tot een hogere mate van overdiagnose aangezien minder baarmoederhalskanker worden voorkomen en het aantal gedetecteerde CIN lesies gelijk blijft. Het is daarom aannemelijk dat minder levensjaren gecorrigeerd voor de kwaliteit van leven ("quality-adjusted life-years") worden verkregen en dat de effectiviteit van het programma afneemt bij gebruik van ThinPrep als primaire test methode. Het gebruik van Surepath leidt waarschijnlijk ook tot een hogere mate van overdiagnose aangezien we hebben berekend dat 3,1 extra CIN I, 3,2 extra CIN II en 3,1 extra CIN III lesies worden gedetecteerd per extra voorkomen intervalkanker. Echter, hoe het gebruik van SurePath het aantal verkregen levensjaren gecorrigeerd voor de kwaliteit van leven, en daarmee de effectiviteit van het Nederlandse bevolkingsonderzoek naar baarmoederhalskanker, zal beïnvloeden moet worden bepaald met gebruik van microsimulaties.

TOENAME IN OPKOMST DOOR HET AANBIEDEN VAN DE HPV ZELFAFNAMESET AAN NIET-OPKOMENDE VROUWEN

Een studie die de screengeschiedenis van Nederlandse vrouwen met baarmoederhalskanker onderzocht liet zien dat meer dan de helft van de kankers voorkomt in vrouwen die de afgelopen zes jaar niet deelnamen aan het Nederlandse bevolkingsonderzoek naar baarmoederhalskanker. En dat terwijl minder dan tien procent van de vrouwen met kanker een voorgaand negatief screeningsuitstrijkje had. Dit werd bevestigd in **hoofdstuk 6** waarin we hebben laten zien dat de meerderheid van de baarmoederhalskankers in 30 tot 35-jarige vrouwen nog steeds zou plaatsvinden indien een sensitievere

screeningstest op leeftijd 30 zou worden gebruikt. Echter, een gedeelte van deze kankers had voorkomen kunnen worden door het verbeteren van de screeningsopkomst. Een veelbelovende manier om de screeningsdeelname te verhogen is door het aanbieden van de HPV zelfafnameset (ZAS) aan niet-opkomende vrouwen van primaire HPV screening, wat naar verwachting vanaf 2016 zal worden ingevoerd. Dit zou echter ook kunnen resulteren in een afname van de effectiviteit van het programma indien teveel vrouwen zullen "switchen" van reguliere screening naar zelfafname, gegeven dat met behulp van de ZAS minder CIN II$^+$ wordt gedetecteerd (door meer loss to follow-up en een eventueel lagere sensitiviteit). In **hoofdstuk 7** hebben we laten zien dat het effectief is om de HPV ZAS aan te bieden aan niet-opkomende vrouwen in elk bestudeerd scenario, zolang er geen vrouwen switchen. Anders moet er worden gegarandeerd dat (i) de relatieve CIN II$^+$ sensitiviteit tenminste 0,95 is, (ii) vrouwen met tenminste een 1,7 keer verhoogd achtergrondrisico deelnemen en (iii) de deelname graad toeneemt met tenminste 6 percentage punten. In **hoofdstuk 8** hebben we geen associatie gevonden tussen sociaaleconomische status (SES) en het baarmoederhalskanker achtergrondrisico, alhoewel een 1,39 hoger achtergrondrisico voor lagere versus hogere SES vrouwen niet kan worden uitgesloten. In vergelijking met de algemene populatie zal dit cijfer waarschijnlijk nog wat lager uitvallen. Aan het tweede criterium (vrouwen met ten minste een 1,7 keer verhoogd achtergrondrisico moeten deelnemen via de HPV ZAS) zal dus niet worden voldaan door te focussen op lagere SES vrouwen. Het is daarom essentieel dat de kortetermijneffecten en langetermijneffecten van het aanbieden van de HPV ZAS zorgvuldig worden gemonitored zodat op die manier de effectiviteit van het programma kan worden gewaarborgd. Helemaal aangezien het aannemelijk is dat vrouwen zullen switchen en het onbekend is of aan alle criteria wordt voldaan in de praktijk.

CONCLUSIES

- Het effect van het gebruik van dunnelaagcytologie als primaire test methode hangt af van het type dunnelaagcytologie. Onze resultaten suggereren ten zeerste dat Surepath een hogere sensitiviteit heeft om progressieve CIN lesies te detecteren dan conventionele cytologie, terwijl het tegelijkertijd meer CIN I lesies detecteert. ThinPrep lijkt daarentegen een lagere sensitiviteit te hebben om progressieve CIN lesies te detecteren.
- De geleidelijke implementatie van dunnelaagcytologie is verantwoordelijk voor een klein, maar significant gedeelte van de toename in CIN I, CIN II en CIN III detectie binnen het Nederlandse bevolkingsonderzoek naar baarmoederhalskanker. Andere factoren moeten daarom ook een belangrijke rol hebben gespeeld.
- Meer dan 70% van de baarmoederhalskankers gediagnosticeert in 30 tot 35-jarige vrouwen had niet kunnen worden voorkomen door het gebruik van een sensitievere screeningstest op leeftijd 30. Een groot gedeelte van deze kankers komt echter voor in vrouwen zonder enige screeningsgeschiedenis. Het verhogen van de screeningsopkomst op screeningsleeftijd 30, bijvoorbeeld door het aanbieden van de HPV zelfafnameset, kan daarom een belangrijke bijdrage leveren aan het voorkomen van baarmoederhalskankers in jonge vrouwen.
- Lagere sociaaleconomische status vrouwen lijken geen verhoogd achtergrondrisico te hebben om baarmoederhalskanker te ontwikkelen, alhoewel dit niet kan worden uitgesloten.
- De volgende variabelen zijn het meest invloedrijk in het bepalen of het aanbieden van de HPV zelfafnameset aan niet-opkomende vrouwen effectief is of niet: De relatieve CIN II^{+} sensitiviteit van de HPV zelfafnameset, de extra opkomst gegenereerd door de HPV zelfafnameset, het vermogen om vrouwen met een hoger achtergrondrisico te bereiken en het percentage vrouwen wat switcht van reguliere screening naar zelfafname. In elk bestudeerd scenario is het effectief om de HPV zelfafnameset aan te bieden aan niet-opkomende vrouwen zolang er

geen vrouwen switchen. Indien alle reguliere opkomende vrouwen zouden switchen, is het aanbieden van de HPV zelfafnameset effectief indien tegelijkertijd (i) de deelname graad toeneemt met tenminste 6 percentage punten, (ii) vrouwen met ten minste een 1,7 keer verhoogd achtergrondrisico deelnemen en (iii) de relatieve CIN II^{+} sensitiviteit tenminste 0,95 is.

- Het effect van nieuwe screeningstesten op het voorkomen van baarmoederhalskanker kan verkeerd worden ingeschat indien gebruik wordt gemaakt van surrogaat eindpunten, zoals CIN detectie rates, in plaats van intervalkanker rates.

AANBEVELINGEN

- Als de beslissing welke primaire cytologietest te gebruiken afhangt van het vermogen om baarmoederhalskanker gevallen te voorkomen, dan is het gebruik van SurePath in het Nederlandse bevolkingsonderzoek naar baarmoederhalskanker aan te raden.
- De sensitiviteit van ThinPrep om progressieve CIN lesies te detecteren moet worden verbeterd om een gelijk nadeel-voordeel ratio en gelijke effectiviteit van het Nederlandse bevolkingsonderzoek naar baarmoederhalskanker te garanderen.
- Het is onbekend of aan alle condities voor de HPV zelfafnameset om (kosten)effectief te zijn wordt voldaan in de praktijk. Daarom is het essentieel dat de kortetermijneffecten en langetermijneffecten van het aanbieden van de HPV zelfafnameset zorgvuldig worden gemonitored zodat op die manier de effectiviteit van het Nederlandse bevolkingsonderzoek naar baarmoederhalskanker kan worden gewaarborgd.
- Monitoring van de effecten van het implementeren van nieuwe screeningstesten of nieuwe screeningsstrategieën, gebruik makende van populatie-data, is essentieel om de effectiviteit van het programma te waarborgen.

CHAPTER 11

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F.GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr, accessed on 27/02/2015.
2. Gustafsson L, Ponten J, Zack M, Adami HO. International incidence rates of invasive cervical cancer after introduction of cytological screening. Cancer Causes Control. 1997 Sep;8(5):755-63.
3. Dutch Comprehensive Cancer Center. Online Cancer Registry [Internet]. Available from http://www.cijfersoverkanker.nl, accessed on 27/02/2015. .
4. Franco EL, Rohan TE, Villa LL. Epidemiologic evidence and human papillomavirus infection as a necessary cause of cervical cancer. J Natl Cancer Inst. 1999 Mar 17;91(6):506-11.
5. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol. 1999 Sep;189(1):12-9.
6. de Villiers EM. Taxonomic classification of papillomaviruses. Papillomavirus Rep 2001;12:57– 63.
7. Schiffman M, Kjaer SK. Chapter 2: Natural history of anogenital human papillomavirus infection and neoplasia. J Natl Cancer Inst Monogr. 2003(31):14-9.
8. Munoz N, Bosch FX, Castellsague X, Diaz M, de Sanjose S, Hammouda D, et al. Against which human papillomavirus types shall we vaccinate and screen? The international perspective. Int J Cancer. 2004 Aug 20;111(2):278-85.
9. Franco EL, Duarte-Franco E, Ferenczy A. Cervical cancer: epidemiology, prevention and the role of human papillomavirus infection. CMAJ. 2001 Apr 3;164(7):1017-25.
10. Kjaer SK, van den Brule AJ, Bock JE, Poll PA, Engholm G, Sherman ME, et al. Determinants for genital human papillomavirus (HPV) infection in 1000 randomly chosen young Danish women with normal Pap smear: are there different risk profiles for oncogenic and nononcogenic HPV types? Cancer Epidemiol Biomarkers Prev. 1997 Oct;6(10):799-805.
11. Baseman JG, Koutsky LA. The epidemiology of human papillomavirus infections. J Clin Virol. 2005 Mar;32 Suppl 1:S16-24.
12. Jacobs MV, Walboomers JM, Snijders PJ, Voorhorst FJ, Verheijen RH, Fransen-Daalmeijer N, et al. Distribution of 37 mucosotropic HPV types in women with cytologically normal cervical smears: the age-related patterns for high-risk and low-risk types. Int J Cancer. 2000 Jul 15;87(2): 221-7.
13. Sellors JW, Mahony JB, Kaczorowski J, Lytwyn A, Bangura H, Chong S, et al. Prevalence and predictors of human papillomavirus infection in women in Ontario, Canada. Survey of HPV in Ontario Women (SHOW) Group. CMAJ. 2000 Sep 5;163(5):503-8.
14. Veldhuijzen NJ, Berkhof J, Gillio-Tos A, De Marco L, Carozzi F, Del Mistro A, et al. The age distribution of type-specific high-risk human papillomavirus incidence in two population-based screening trials. Cancer Epidemiol Biomarkers Prev. 2015 Jan;24(1):111-8.
15. Kulasingam SL, Hughes JP, Kiviat NB, Mao C, Weiss NS, Kuypers JM, et al. Evaluation of human papillomavirus testing in primary screening for cervical abnormalities: comparison of sensitivity, specificity, and frequency of referral. JAMA. 2002 Oct 9;288(14):1749-57.
16. Koutsky LA, Holmes KK, Critchlow CW, Stevens CE, Paavonen J, Beckmann AM, et al. A cohort study of the risk of cervical intraepithelial neoplasia grade 2 or 3 in relation to papillomavirus infection. N Engl J Med. 1992 Oct 29;327(18): 1272-8.
17. Rozendaal L, Walboomers JM, van der Linden JC, Voorhorst FJ, Kenemans P, Helmerhorst TJ, et al. PCR-based high-risk HPV test in cervical cancer screening gives objective risk assessment of women with cytomorphologically normal cervical smears. Int J Cancer. 1996 Dec 11;68(6): 766-9.
18. Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV. The causal relation between human papil-

lomavirus and cervical cancer. J Clin Pathol. 2002 Apr;55(4):244-65.

19. Ostor AG. Natural history of cervical intraepithelial neoplasia: a critical review. Int J Gynecol Pathol. 1993 Apr;12(2):186-92.

20. Arends MJ, Buckley CH, Wells M. Aetiology, pathogenesis, and pathology of cervical neoplasia. J Clin Pathol. 1998 Feb;51(2):96-103.

21. van Oortmarssen GJ, Habbema JD. Epidemiological evidence for age-dependent regression of pre-invasive cervical cancer. Br J Cancer. 1991 Sep;64(3):559-65.

22. Bos AB, van Ballegooijen M, van Oortmarssen GJ, van Marle ME, Habbema JD, Lynge E. Non-progression of cervical intraepithelial neoplasia estimated from population-screening data. Br J Cancer. 1997;75(1):124-30.

23. Plaxe SC, Saltzstein SL. Estimation of the duration of the preclinical phase of cervical adenocarcinoma suggests that there is ample opportunity for screening. Gynecol Oncol. 1999 Oct;75(1): 55-61.

24. Jordan J, Martin-Hirsch P, Arbyn M, Schenck U, Baldauf JJ, Da Silva D, et al. European guidelines for clinical management of abnormal cervical cytology, part 2. Cytopathology. 2009 Feb;20(1): 5-16.

25. Soutter WP, de Barros Lopes A, Fletcher A, Monaghan JM, Duncan ID, Paraskevaidis E, et al. Invasive cervical cancer after conservative therapy for cervical intraepithelial neoplasia. Lancet. 1997 Apr 5;349(9057):978-80.

26. Arbyn M, Kyrgiou M, Simoens C, Raifu AO, Koliopoulos G, Martin-Hirsch P, et al. Perinatal mortality and other severe adverse pregnancy outcomes associated with treatment of cervical intraepithelial neoplasia: meta-analysis. BMJ. 2008;337:a1284.

27. Dutch society for obstetrics and gynaecology [Internet]. Available from http://www.nvog.nl, accessed on 03/03/2015

28. Bulk S, Van Kemenade FJ, Rozendaal L, Meijer CJ. The Dutch CISOE-A framework for cytology reporting increases efficacy of screening upon standardisation since 1996. J Clin Pathol. 2004 Apr;57(4):388-93.

29. M R. Recent developments in the Dutch cervical cancer screening programme 2008.

30. Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. JAMA. 2002 Apr 24;287(16): 2114-9.

31. Giard RW, Hermans J, Doornewaard H. [National results of cervix cytology diagnosis in 1992; efficacy of screening could be improved] Landelijke resultaten van cervixcytologische diagnostiek in 1992; de screening kan doelmatiger. Ned Tijdschr Geneeskd. 1994 Jun 25;138(26):1325-30.

32. Habbema F, Lubbe JTN, van Agt HM, van Ballegooijen M, Koopmanschap MA, van Oortmarssen GJ. Costs and effects of the organised cervical cancer screening programme (In Dutch). Rotterdam, the Netherlands; Department of Public Health, Erasmus University, 1988.

33. van Elven EH, Griffioen HP, Wiegman HP. Survey into bottlenecks of the organised cervical cancer screening programme (in Dutch). KPMG, 1990.

34. Rebolj M, van Ballegooijen M, Berkers LM, Habbema D. Monitoring a national cancer prevention program: successful changes in cervical cancer screening in the Netherlands. Int J Cancer. 2007 Feb 15;120(4):806-12.

35. Hanselaar AG. Criteria for organized cervical screening programs. Special emphasis on The Netherlands program. Acta Cytol. 2002 Jul-Aug; 46(4):619-29.

36. Rebolj M, van Ballegooijen M, van Kemenade F, Looman C, Boer R, Habbema JD. No increased risk for cervical cancer after a broader definition of a negative Pap smear. Int J Cancer. 2008 Dec 1; 123(11):2632-5.

37. Laara E, Day NE, Hakama M. Trends in mortality from cervical cancer in the Nordic countries: association with organised screening programmes. Lancet. 1987 May 30;1(8544):1247-9.

38. Praktijkrichtlijn versie 3.2 (in Dutch). Voor kwaliteitsborging van cytopathologisch onder-

zoek van de baarmoederhals. Dutch Society of Pathologists. 12 Jan 2012. .

39. Siebers AG, Arbyn M, Melchers WJ, van Kemenade FJ, Vedder JE, van der Linden H, et al. Effectiveness of two strategies to follow-up ASC-US and LSIL screening results in The Netherlands using repeat cytology with or without additional hrHPV testing: a retrospective cohort study. Cancer Causes Control. 2014 Sep;25(9):1141-9.

40. Colgan TJ, McLachlin CM, Cotterchio M, Howlett R, Seidenfeld AM, Mai VM. Results of the implementation of liquid-based cytology-SurePath in the Ontario screening program. Cancer. 2004 Dec 25;102(6):362-7.

41. Albrow R, Kitchener H, Gupta N, Desai M. Cervical screening in England: the past, present, and future. Cancer Cytopathol. 2012 Apr 25;120(2): 87-96.

42. Beerman H, van Dorst EB, Kuenen-Boumeester V, Hogendoorn PC. Superior performance of liquid-based versus conventional cytology in a population-based cervical cancer screening program. Gynecol Oncol. 2009 Mar;112(3):572-6.

43. Cheung AN, Szeto EF, Leung BS, Khoo US, Ng AW. Liquid-based cytology and conventional cervical smears: a comparison study in an Asian screening population. Cancer. 2003 Dec 25;99(6): 331-5.

44. Cox JT. History of the use of HPV testing in cervical screening and in the management of abnormal cervical screening results. J Clin Virol. 2009 Jul;45 Suppl 1:S3-S12.

45. Ferenczy A, Robitaille J, Franco E, Arseneau J, Richart RM, Wright TC. Conventional cervical cytologic smears vs. ThinPrep smears. A paired comparison study on cervical cytology. Acta Cytol. 1996 Nov-Dec;40(6):1136-42.

46. Fremont-Smith M, Marino J, Griffin B, Spencer L, Bolick D. Comparison of the SurePath liquid-based Papanicolaou smear with the conventional Papanicolaou smear in a multisite direct-to-vial study. Cancer. 2004 Oct 25;102(5):269-79.

47. Sass MA. Use of a liquid-based, thin-layer Pap test in a community hospital. Impact on cytology performance and productivity. Acta Cytol. 2004 Jan-Feb;48(1):17-22.

48. Siebers AG, Klinkhamer PJ, Arbyn M, Raifu AO, Massuger LF, Bulten J. Cytologic detection of cervical abnormalities using liquid-based compared with conventional cytology: a randomized controlled trial. Obstet Gynecol. 2008 Dec;112(6): 1327-34.

49. Siebers AG, Klinkhamer PJ, Grefte JM, Massuger LF, Vedder JE, Beijers-Broos A, et al. Comparison of liquid-based cytology with conventional cytology for detection of cervical cancer precursors: a randomized controlled trial. JAMA. 2009 Oct 28; 302(16):1757-64.

50. Davey E, Barratt A, Irwig L, Chan SF, Macaskill P, Mannes P, et al. Effect of study design and quality on unsatisfactory rates, cytology classifications, and accuracy in liquid-based versus conventional cervical cytology: a systematic review. Lancet. 2006 Jan 14;367(9505):122-32.

51. Davey E, d'Assuncao J, Irwig L, Macaskill P, Chan SF, Richards A, et al. Accuracy of reading liquid based cytology slides using the ThinPrep Imager compared with conventional cytology: prospective study. BMJ. 2007 Jul 7;335(7609):31.

52. Taylor S, Kuhn L, Dupree W, Denny L, De Souza M, Wright TC, Jr. Direct comparison of liquid-based and conventional cytology in a South African screening trial. Int J Cancer. 2006 Feb 15; 118(4):957-62.

53. Zhao FH, Hu SY, Bian JJ, Liu B, Peck RB, Bao YP, et al. Comparison of ThinPrep and SurePath liquid-based cytology and subsequent human papillomavirus DNA testing in China. Cancer Cytopathol. 2011 Dec 25;119(6):387-94.

54. Thiryayi SA, Marshall J, Rana DN. An audit of liquid-based cervical cytology screening samples (ThinPrep and SurePath) reported as glandular neoplasia. Cytopathology. 2010 Aug;21(4):223-8.

55. Arbyn M, Bergeron C, Klinkhamer P, Martin-Hirsch P, Siebers AG, Bulten J. Liquid compared with conventional cervical cytology: a systematic review and meta-analysis. Obstet Gynecol. 2008 Jan;111(1):167-77.

56. Whitlock EP, Vesco KK, Eder M, Lin JS, Senger CA, Burda BU. Liquid-based cytology and human papillomavirus testing to screen for cervical cancer: a systematic review for the U.S. Preventive

Services Task Force. Ann Intern Med. 2011 Nov 15;155(10):687-97, W214-5.

57. Ronco G, Cuzick J, Pierotti P, Cariaggi MP, Dalla Palma P, Naldoni C, et al. Accuracy of liquid based versus conventional cytology: overall results of new technologies for cervical cancer screening: randomised controlled trial. BMJ. 2007 Jul 7;335(7609):28.

58. Sykes PH, Harker DY, Miller A, Whitehead M, Neal H, Wells JE, et al. A randomised comparison of SurePath liquid-based cytology and conventional smear cytology in a colposcopy clinic setting. BJOG. 2008 Oct;115(11):1375-81.

59. Hussein T, Desai M, Tomlinson A, Kitchener HC. The comparative diagnostic accuracy of conventional and liquid-based cytology in a colposcopic setting. BJOG. 2005 Nov;112(11):1542-6.

60. Coste J, Cochand-Priollet B, de Cremoux P, Le Gales C, Cartier I, Molinie V, et al. Cross sectional study of conventional cervical smear, monolayer cytology, and human papillomavirus DNA testing for cervical cancer screening. BMJ. 2003 Apr 5; 326(7392):733.

61. Klug SJ, Neis KJ, Harlfinger W, Malter A, Konig J, Spieth S, et al. A randomized trial comparing conventional cytology to liquid-based cytology and computer assistance. Int J Cancer. 2013 Jun 15;132(12):2849-57.

62. Palmer TJ, Nicoll SM, McKean ME, Park AJ, Bishop D, Baker L, et al. Prospective parallel randomized trial of the MultiCyte ThinPrep((R)) imaging system: the Scottish experience. Cytopathology. 2013 Aug;24(4):235-45.

63. Kitchener HC, Blanks R, Dunn G, Gunn L, Desai M, Albrow R, et al. Automation-assisted versus manual reading of cervical cytology (MAVARIC): a randomised controlled trial. Lancet Oncol. 2011 Jan;12(1):56-64.

64. Rebolj M, Rask J, van Ballegooijen M, Kirschner B, Rozemeijer K, Bonde J, et al. Cervical histology after routine ThinPrep or SurePath liquid-based cytology and computer-assisted reading in Denmark. Br J Cancer. 2015 Oct 8.

65. Arbyn M, Ronco G, Anttila A, Meijer CJ, Poljak M, Ogilvie G, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. Vaccine. 2012 Nov 20;30 Suppl 5: F88-99.

66. Ronco G, Dillner J, Elfstrom KM, Tunesi S, Snijders PJ, Arbyn M, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. Lancet. 2014 Feb 8;383(9916): 524-32.

67. Schippers EI. Improvement of the cervical cancer screening program (In Dutch) [Letter of the ministry of Welfare, Public Health and Sport, 17 October 2013 (156231-111126-PG)]. .

68. van der Veen N, Carpay MEM, van Delden JA, Brouwer E, Grievink L, Hoebee B, Lock AJJ, Salverda-Nijhof JGW. Feasiblity study for improvements to the population screening for cervical cancer 2013, 28 September 2014 (2014-0056).

69. Nobbenhuis MA, Helmerhorst TJ, van den Brule AJ, Rozendaal L, Jaspars LH, Voorhorst FJ, et al. Primary screening for high risk HPV by home obtained cervicovaginal lavage is an alternative screening tool for unscreened women. J Clin Pathol. 2002 Jun;55(6):435-9.

70. Brink AA, Meijer CJ, Wiegerinck MA, Nieboer TE, Kruitwagen RF, van Kemenade F, et al. High concordance of results of testing for human papillomavirus in cervicovaginal samples collected by two methods, with comparison of a novel self-sampling device to a conventional endocervical brush. J Clin Microbiol. 2006 Jul;44(7):2518-23.

71. Arbyn M, Castle PE. Offering self-sampling kits for HPV testing to reach women who do not attend in the regular cervical cancer screening program. Cancer Epidemiol Biomarkers Prev. 2015 Feb 24.

72. Racey CS, Withrow DR, Gesink D. Self-collected HPV testing improves participation in cervical cancer screening: a systematic review and meta-analysis. Can J Public Health. 2013 Mar-Apr; 104(2):e159-66.

73. Gok M, Heideman DA, van Kemenade FJ, Berkhof J, Rozendaal L, Spruyt JW, et al. HPV testing on self collected cervicovaginal lavage specimens as screening method for women who do not at-

tend cervical screening: cohort study. BMJ. 2010; 340:c1040.

74. Bos AB, Rebolj M, Habbema JD, van Ballegooijen M. Nonattendance is still the main limitation for the effectiveness of screening for cervical cancer in the Netherlands. Int J Cancer. 2006 Nov 15; 119(10):2372-5.

75. de Melker HE, Conyn-van Spaendonck MA, Boot HJ, Coutinho RA. [Introduction to vaccination against cervical cancer] Introductie van vaccinatie tegen baarmoederhalskanker. Ned Tijdschr Geneeskd. 2009 Apr 4;153(14):658-61.

76. van Lier EA, Oomen PJ, Mulder M, et al. Vaccinatiegraad Rijksvaccinatieprogramma Nederland. Rijksinstituut voor Volksgezondheid en Milieu (RIVM); 2013.

77. Coupe VM, Bogaards JA, Meijer CJ, Berkhof J. Impact of vaccine protection against multiple HPV types on the cost-effectiveness of cervical screening. Vaccine. 2012 Feb 27;30(10):1813-22.

78. Landelijke Evaluatie Bevolkingsonderzoek Baarmoederhalskanker, LEBA rapportage tot en met 2009 (in Dutch). Available from the National Institute for Public Health and the Environment; www.rivm.nl (accessed 2 April 2015).

79. Landelijke Evaluatie Bevolkingsonderzoek Baarmoederhalskanker, LEBA rapportage tot en met 2011 (in Dutch). Available from the National Institute for Public Health and the Environment; www.rivm.nl (accessed 2 April 2015).

80. Landelijke Evaluatie Bevolkingsonderzoek Baarmoederhalskanker, LEBA rapportage tot en met 2013 (in Dutch). Available from the National Institute for Public Health and the Environment; www.rivm.nl (accessed 2 April 2015).

81. van Ballegooijen M, Rebolj M, Essink-Bot ML, Meerding WJ, Berkers LM, Habbema JDF. The effects and costs of the Dutch cervical cancer screening programme after the reorganisation (in Dutch). Rotterdam: Department of Public Health, Erasmus MC 2006.

82. Insinga RP, Glass AG, Myers ER, Rush BB. Abnormal outcomes following cervical cancer screening: event duration and health utility loss. Med Decis Making. 2007 Jul-Aug;27(4):414-22.

83. Cuzick J, Clavel C, Petry KU, Meijer CJ, Hoyer H, Ratnam S, et al. Overview of the European and North American studies on HPV testing in primary cervical cancer screening. Int J Cancer. 2006 Sep 1;119(5):1095-101.

84. de Kok IM, van Rosmalen J, Dillner J, Arbyn M, Sasieni P, Iftner T, et al. Primary screening for human papillomavirus compared with cytology screening for cervical cancer in European settings: cost effectiveness analysis based on a Dutch microsimulation model. BMJ. 2012;344:e670.

85. van Rosmalen J, de Kok IM, van Ballegooijen M. Cost-effectiveness of cervical cancer screening: cytology versus human papillomavirus DNA testing. Bjog. 2012 May;119(6):699-709.

86. Jensen KE, Hannibal CG, Nielsen A, Jensen A, Nohr B, Munk C, et al. Social inequality and incidence of and survival from cancer of the female genital organs in a population-based study in Denmark, 1994-2003. Eur J Cancer. 2008 Sep; 44(14):2003-17.

87. Froment MA, Gomez SL, Roux A, DeRouen MC, Kidd EA. Impact of socioeconomic status and ethnic enclave on cervical cancer incidence among Hispanics and Asians in California. Gynecol Oncol. 2014 Jun;133(3):409-15.

88. Hansen BT, Hukkelberg SS, Haldorsen T, Eriksen T, Skare GB, Nygard M. Factors associated with non-attendance, opportunistic attendance and reminded attendance to cervical screening in an organized screening program: a cross-sectional study of 12,058 Norwegian women. BMC Public Health. 2011;11:264.

89. Ackerson K, Preston SD. A decision theory perspective on why women do or do not decide to have cancer screening: systematic review. J Adv Nurs. 2009 Jun;65(6):1130-40.

90. Kristensson JH, Sander BB, von Euler-Chelpin M, Lynge E. Predictors of non-participation in cervical screening in Denmark. Cancer Epidemiol. 2014 Apr;38(2):174-80.

91. Klinkhamer PJ, Meerding WJ, Rosier PF, Hanselaar AG. Liquid-based cervical cytology. Cancer. 2003 Oct 25;99(5):263-71.

92. Rask J, Lynge E, Franzmann M, Hansen B, Hjortebjerg A, Rygaard C, et al. Impact of technology on

cytology outcome in cervical cancer screening of young and older women. Int J Cancer. 2014 May 1;134(9):2168-79.

93. http://www.cancerscreening.nhs.uk/cervical/lbc.html [cited 2014-09-02].

94. Casparie M, Tiebosch AT, Burger G, Blauwgeers H, van de Pol A, van Krieken JH, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cell Oncol. 2007;29(1):19-24.

95. Knol FA. Van hoog naar laag; van laag naar hoog: de sociaal-ruimtelijke ontwikkeling van wijken tussen 1971-1995 [in Dutch]. Social Cultureel Planbureau; 1998.

96. Available from the Netherlands Institute for Social Research; www.scp.nl.

97. Rozemeijer K, van Kemenade FJ, Penning C, Matthijsse SM, Naber SK, van Rosmalen J, et al. Exploring the trend of increased cervical intraepithelial neoplasia detection rates in the Netherlands. J Med Screen. 2015 May 14.

98. Vesco KK, Whitlock EP, Eder M, Burda BU, Senger CA, Lutz K. Risk factors and other epidemiologic considerations for cervical cancer screening: a narrative review for the U.S. Preventive Services Task Force. Ann Intern Med. 2011 Nov 15;155(10):698-705, W216.

99. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA. 1998 Nov 18;280(19):1690-1.

100. Umana A, Dunsmore H, Herbert A, Jokhan A, Kubba A. Are significant numbers of abnormal cells lost on the discarded ThinPrep(R) broom when used for cervical cytology? Cytopathology. 2013 Aug;24(4):228-34.

101. Bigras G, Rieder MA, Lambercy JM, Kunz B, Chatelain JP, Reymond O, et al. Keeping collecting device in liquid medium is mandatory to ensure optimized liquid-based cervical cytologic sampling. J Low Genit Tract Dis. 2003 Jul;7(3):168-74.

102. Benoy IH, Vanden Broeck D, Ruymbeke MJ, Sahebali S, Arbyn M, Bogers JJ, et al. Prior knowledge of HPV status improves detection of CIN2+ by cytology screening. Am J Obstet Gynecol. 2011 Dec;205(6):569 e1-7.

103. Moriarty AT, Nayar R, Arnold T, Gearries L, Renshaw A, Thomas N, et al. The Tahoe Study: bias in the interpretation of Papanicolaou test results when human papillomavirus status is known. Arch Pathol Lab Med. 2014 Sep;138(9):1182-5.

104. Cuzick J, Ahmad A, Costa M, Lyons D, Wright C, Ho L, Terry G, Austin J, Ashdown-Barr L, Cadman L, Szarewski A. A comparison of the performance of different HPV tests in ThinPrep vs SurePath. Paper presented at Eurogin; November 2013; Florence, Italy. .

105. Richart RM. Colpomicroscopic Studies of Cervical Intraepithelial Neoplasia. Cancer 1966;19:395–405.

106. Richart RM. Causes and management of cervical intraepithelial neoplasia. Cancer. 1987 Oct 15;60(8 Suppl):1951-9.

107. Workgroup Gynecologic Oncology. Guideline Draft 1.1; Treatment of CIN [in Dutch]. 2004.

108. Statistical Research and Applications Branch of the US National Cancer Institute (http://surveillance.cancer.gov/joinpoint). Version 4.0.4.

109. de Kok IM, van der Aa MA, van Ballegooijen M, Siesling S, Karim-Kos HE, van Kemenade FJ, et al. Trends in cervical cancer in the Netherlands until 2007: has the bottom been reached? Int J Cancer. 2011 May 1;128(9):2174-81.

110. Schiffman M, Solomon D. Clinical practice. Cervical-cancer screening with human papillomavirus and cytologic cotesting. N Engl J Med. 2013 Dec 12;369(24):2324-31.

111. Carreon JD, Sherman ME, Guillen D, Solomon D, Herrero R, Jeronimo J, et al. CIN2 is a much less reproducible and less valid diagnosis than CIN3: results from a histological review of population-based cervical samples. Int J Gynecol Pathol. 2007 Oct;26(4):441-6.

112. Dalla Palma P, Giorgi Rossi P, Collina G, Buccoliero AM, Ghiringhello B, Gilioli E, et al. The reproducibility of CIN diagnoses among different pathologists: data from histology reviews from a multicenter randomized study. Am J Clin Pathol. 2009 Jul;132(1):125-32.

113. Barken SS, Rebolj M, Andersen ES, Lynge E. Frequency of cervical intraepithelial neoplasia treatment in a well-screened population. Int J Cancer. 2012 May 15;130(10):2438-44.

114. Finnish Cancer Registry. Statistics of cervical cancer screening in Finland, 2006 and 2009. In: https://cancer-fi.directo.fi/syoparekisteri/en/.

115. NHS Cervical screening programme. Annual review 2012. ISBN 978-1-84463-0943-3.

116. Sigurdsson K. Is a liquid-based cytology more sensitive than a conventional Pap smear? Cytopathology. 2013 Aug;24(4):254-63.

117. Rozemeijer K, Penning C, Siebers AG, Naber SK, Matthijsse SM, van Ballegooijen M, et al. Comparing SurePath, ThinPrep, and conventional cytology as primary test method: SurePath is associated with increased CIN II detection rates. Cancer Causes Control. 2015 Oct 12.

118. van de Velde-Blok M. The description of the PALEBA system and algorithms 2007 (in Dutch). Amsterdam: PSB, by order of the Department of Public Health (Erasmus MC) and Prismant 2007.

119. Harrel FE. Regression modelling strategies with applications to linear models, logistic regression, and survival analysis. New York, etc.: Springer, 2001.

120. Rozemeijer K PC, Siebers AG. Comparing Surepath and ThinPrep with conventional cytology: Surepath is associated with increased CIN II+ detection rates. Accepted by Cancer Causes and Control.

121. http: //www.nationaalkompas.nl/bevolking/sterfte/huidig/.

122. Wilson JMGJ, G. Principles and Practice of Screening For Disease. Geneva: WHO; 1968.

123. Boer R, Warmerdam P, de Koning H, van Oortmarssen G. Extra incidence caused by mammographic screening. Lancet. 1994 Apr 16; 343(8903):979.

124. Burstein HJ, Polyak K, Wong JS, Lester SC, Kaelin CM. Ductal carcinoma in situ of the breast. N Engl J Med. 2004 Apr 1;350(14):1430-41.

125. Bulk S, Visser O, Rozendaal L, Verheijen RH, Meijer CJ. Cervical cancer in the Netherlands 1989-1998: Decrease of squamous cell carcinoma in older women, increase of adenocarcinoma in younger women. Int J Cancer. 2005 Mar 1;113(6): 1005-9.

126. Esserman LJ, Thompson IM, Jr., Reid B. Overdiagnosis and overtreatment in cancer: an opportunity for improvement. JAMA. 2013 Aug 28;310(8):797-8.

127. de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. Epidemiol Rev. 2011 Jul;33(1):111-21.

128. McCredie MR, Sharples KJ, Paul C, Baranyai J, Medley G, Jones RW, et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. Lancet Oncol. 2008 May;9(5):425-34.

129. Bansal N, Wright JD, Cohen CJ, Herzog TJ. Natural history of established low grade cervical intraepithelial (CIN 1) lesions. Anticancer Res. 2008 May-Jun;28(3B):1763-6.

130. Moscicki AB, Ma Y, Wibbelsman C, Darragh TM, Powers A, Farhat S, et al. Rate of and risks for regression of cervical intraepithelial neoplasia 2 in adolescents and young women. Obstet Gynecol. 2010 Dec;116(6):1373-80.

131. Page DL, Dupont WD, Rogers LW, Jensen RA, Schuyler PA. Continued local recurrence of carcinoma 15-25 years after a diagnosis of low grade ductal carcinoma in situ of the breast treated only by biopsy. Cancer. 1995 Oct 1;76(7):1197-200.

132. Albrektsen G, Heuch I, Thoresen SO. Histological type and grade of breast cancer tumors by parity, age at birth, and time since birth: a register-based study in Norway. BMC Cancer. 2010;10:226.

133. van Ballegooijen M, Hermens R. Cervical cancer screening in the Netherlands. Eur J Cancer. 2000 Nov;36(17):2244-6.

134. Otten JD, van Dijck JA, Peer PG, Straatman H, Verbeek AL, Mravunac M, et al. Long term breast cancer screening in Nijmegen, The Netherlands: the nine rounds from 1975-92. J Epidemiol Community Health. 1996 Jun;50(3):353-8.

135. Morrell S, Barratt A, Irwig L, Howard K, Biesheuvel C, Armstrong B. Estimates of overdiagnosis of

invasive breast cancer associated with screening mammography. Cancer Causes Control. 2010 Feb;21(2):275-82.

136. Welch HG, Black WC. Overdiagnosis in cancer. J Natl Cancer Inst. 2010 May 5;102(9):605-13.

137. Kalager M, Adami HO, Bretthauer M, Tamimi RM. Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program. Ann Intern Med. 2012 Apr 3;156(7):491-9.

138. Duffy SW, Agbaje O, Tabar L, Vitak B, Bjurstam N, Bjorneld L, et al. Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. Breast Cancer Res. 2005;7(6):258-65.

139. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Br J Cancer. 2013 Jun 11;108(11):2205-40.

140. De Kok IM, van Rosmalen J, Van Ballegooijen M. Description of MISCAN-cervix, Web Appendix accompanying 'A comparison of primary HPV to cytology cervical cancer screening in different European settings: A costeffectiveness analysis based on a Dutch microsimulation model'. RePub EUR; [09/03/2013]; Available from: http://repub.eur.nl/res/pub/31582/deki874693.ww1.pdf.

141. Cijfers over Kanker. The Netherlands Cancer Registry; 2014; Available from: http://www.cijfersoverkanker.nl/.

142. R dG. Predicting the Benefits and Harms of Breast Cancer Screening: Current debates and future directions: Erasmus University, Rotterdam; 2012.

143. Tabar L, Yen MF, Vitak B, Chen HH, Smith RA, Duffy SW. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. Lancet. 2003 Apr 26;361(9367):1405-10.

144. de Gelder R, Fracheboud J, Heijnsdijk EA, den Heeten G, Verbeek AL, Broeders MJ, et al. Digital mammography screening: Weighing reduced mortality against increased overdiagnosis. Prev Med. 2011 Jun 21.

145. Drutskoy A, Abe K, Abe K, Adachi I, Aihara H, Anipko D, et al. Measurement of inclusive Ds, D0, and J/psi rates and determination of the Bs(*) Bs(*) production fraction in bb Events at the Upsilon(5S) resonance. Phys Rev Lett. 2007 Feb 2;98(5):052001.

146. Kyrgiou M, Koliopoulos G, Martin-Hirsch P, Arbyn M, Prendiville W, Paraskevaidis E. Obstetric outcomes after conservative treatment for intraepithelial or early invasive cervical lesions: systematic review and meta-analysis. Lancet. 2006 Feb 11;367(9509):489-98.

147. Verbeek AL, Broeders MJ, Otto SJ, Frachebou J, Otten JD, Holland R, et al. [Effects of the population screening into breast cancer] Effecten van het bevolkingsonderzoek naar borstkanker. Ned Tijdschr Geneeskd. 2013;157(10):A5218.

148. Bluekens AM, Holland R, Karssemeijer N, Broeders MJ, den Heeten GJ. Comparison of digital screening mammography and screen-film mammography in the early detection of clinically relevant cancers: a multicenter study. Radiology. 2012 Dec;265(3):707-14.

149. NABON. Breast cancer, Dutch guideline, version 2.0. oncoline; 2012 [09/03/2013]; Available from: http://www.oncoline.nl/uploaded/docs/mammacarcinoom/Dutch%20Breast%20Cancer%20Guideline%202012.pdf.

150. McLaughlin SA. Surgical management of the breast: breast conservation therapy and mastectomy. Surg Clin North Am. 2013 Apr;93(2):411-28.

151. Health Council of the Netherlands. Population screening for cervical cancer. The Hague: Health Council of the Netherlands, 2011; publication no. 2011/07E.ISBN 978-90-5549-866-6.

152. Mitka M. Physicians, patients not following advice from USPSTF on mammography screening. JAMA. 2013 May 22;309(20):2084.

153. ClinicalTrials.gov. ClinicalTrials.gov, A service of the U.S. National Institutes of Health. 2011 [cited 2014 28-01-2014]; Available from: http://clinicaltrials.gov/ct2/show/NCT01081288.

154. Dillner J. Cervical cancer screening in Sweden. Eur J Cancer. 2000 Nov;36(17):2255-9.

155. Denmark NboH. Cervical cancer screening. http://sundhedsstyrelsen.dk/publ/Publ2007/PLAN/

Kraeft/Anbef_screen_livmoderhals_en_samftn.pdf: Sundhedsstyrelsen; 2007.

156. NHS. NHS cervical screening programme. Public Health England; 2013 [cited 2014 28-01-2014]; Available from: http://www.cancerscreening.nhs.uk/cervical/about-cervical-screening.html.

157. de Gelder R, Fracheboud J, Heijnsdijk EA, den Heeten G, Verbeek AL, Broeders MJ, et al. Digital mammography screening: weighing reduced mortality against increased overdiagnosis. Prev Med. 2011 Sep;53(3):134-40.

158. www.cijfersoverkanker.nl, Incidence figures cervical cancer [database on the Internet]2011 [cited 14/03/2013].

159. Arbyn M, Sasieni P, Meijer CJ, Clavel C, Koliopoulos G, Dillner J. Chapter 9: Clinical applications of HPV testing: a summary of meta-analyses. Vaccine. 2006 Aug 31;24 Suppl 3: S3/78-89.

160. Berkhof J, Coupe VM, Bogaards JA, van Kemenade FJ, Helmerhorst TJ, Snijders PJ, et al. The health and economic effects of HPV DNA screening in The Netherlands. Int J Cancer. 2010 Nov 1;127(9):2147-58.

161. Goldie SJ, Kim JJ, Wright TC. Cost-effectiveness of human papillomavirus DNA testing for cervical cancer screening in women aged 30 years or more. Obstet Gynecol. 2004 Apr;103(4):619-31.

162. Kim JJ, Wright TC, Goldie SJ. Cost-effectiveness of human papillomavirus DNA testing in the United Kingdom, The Netherlands, France, and Italy. J Natl Cancer Inst. 2005 Jun 15;97(12):888-95.

163. Mittendorf T, Petry KU, Iftner T, Greiner W, von der Schulenburg JM. Economic evaluation of human papillomavirus screening in Germany. Eur J Health Econ. 2003 Sep;4(3):209-15.

164. Sherlaw-Johnson C, Philips Z. An evaluation of liquid-based cytology and human papillomavirus testing within the UK cervical cancer screening programme. Br J Cancer. 2004 Jul 5;91(1):84-91.

165. Goldhaber-Fiebert JD, Stout NK, Salomon JA, Kuntz KM, Goldie SJ. Cost-effectiveness of cervical cancer screening with human papillomavirus DNA testing and HPV-16,18 vaccination. J Natl Cancer Inst. 2008 Mar 5;100(5):308-20.

166. Wright TC, Jr., Schiffman M, Solomon D, Cox JT, Garcia F, Goldie S, et al. Interim guidance for the use of human papillomavirus DNA testing as an adjunct to cervical cytology for screening. Obstet Gynecol. 2004 Feb;103(2):304-9.

167. van den Akker-van Marle ME, van Ballegooijen M, Habbema JD. Low risk of cervical cancer during a long period after negative screening in the Netherlands. Br J Cancer. 2003 Apr 7;88(7): 1054-7.

168. CBS (Netherlands Central Bureau of Statistics). Population by sex and age. Voorburg.

169. Gustafsson L, Adami HO. Natural history of cervical neoplasia: consistent results obtained by an identification technique. Br J Cancer. 1989 Jul; 60(1):132-41.

170. van Oortmarssen GJ, Habbema JD. Duration of preclinical cervical cancer and reduction in incidence of invasive cancer following negative pap smears. Int J Epidemiol. 1995 Apr;24(2):300-7.

171. van Oortmarssen GJ, Habbema JD, van Ballegooijen M. Predicting mortality from cervical cancer after negative smear test results. BMJ. 1992 Aug 22;305(6851):449-51.

172. Habbema D, De Kok IM, Brown ML. Cervical cancer screening in the United States and the Netherlands: a tale of two countries. Milbank Q. 2012 Mar;90(1):5-37.

173. Lynge E, Rebolj M. Primary HPV screening for cervical cancer prevention: results from European trials. Nat Rev Clin Oncol. 2009 Dec;6(12):699-706.

174. Korfage IJ, van Ballegooijen M, Huveneers H, Essink-Bot ML. Anxiety and borderline PAP smear results. Eur J Cancer. 2010 Jan;46(1):134-41.

175. van de Vijver A, Poppe W, Verguts J, Arbyn M. Pregnancy outcome after cervical conisation: a retrospective cohort study in the Leuven University Hospital. BJOG. 2010 Feb;117(3):268-73.

176. Gok M, Heideman DA, van Kemenade FJ, de Vries AL, Berkhof J, Rozendaal L, et al. Offering self-sampling for human papillomavirus testing to non-attendees of the cervical screening programme: Characteristics of the responders. Eur J Cancer. 2012 Aug;48(12):1799-808.

177. Netherlands Cancer Registry. Incidence and mortality of cervical cancer; 2012. [cited 2014 Dec 11]. Available from: http://cijfersoverkanker.nl/selecties/dataset_1/img51e671de6dab7.

178. van der Aa MA, Pukkala E, Coebergh JW, Anttila A, Siesling S. Mass screening programmes and trends in cervical cancer in Finland and the Netherlands. Int J Cancer. 2008 Apr 15;122(8):1854-8.

179. Schippers EI. Improvement of the cervical cancer screening program (In Dutch) [Letter of the ministry of Welfare, Public Health and Sport, 17 October 2013 (156231-111126-PG)].

180. Gok M, van Kemenade FJ, Heideman DA, Berkhof J, Rozendaal L, Spruyt JW, et al. Experience with high-risk human papillomavirus testing on vaginal brush-based self-samples of non-attendees of the cervical screening program. Int J Cancer. 2012 Mar 1;130(5):1128-35.

181. Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. Comput Methods Programs Biomed. 1985 May;20(1):79-93.

182. Health Council of the Netherlands. Population screening for cervical cancer. The Hague: Health Council of the Netherlands, 2011; publication no. 2011/07E.ISBN 978-90-5549-866-6.

183. Verhoef VM, Bosgraaf RP, van Kemenade FJ, Rozendaal L, Heideman DA, Hesselink AT, et al. Triage by methylation-marker testing versus cytology in women who test HPV-positive on self-collected cervicovaginal specimens (PROHTECT-3): a randomised controlled non-inferiority trial. Lancet Oncol. 2014 Mar;15(3):315-22.

184. Meijer CJ, Berkhof J, Castle PE, Hesselink AT, Franco EL, Ronco G, et al. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. Int J Cancer. 2009 Feb 1;124(3):516-20.

185. Arbyn M, Verdoodt F, Snijders PJ, Verhoef VM, Suonio E, Dillner L, et al. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. Lancet Oncol. 2014 Jan 13.

186. Goldie SJ, Kohli M, Grima D, Weinstein MC, Wright TC, Bosch FX, et al. Projected clinical benefits and cost-effectiveness of a human papillomavirus 16/18 vaccine. J Natl Cancer Inst. 2004 Apr 21;96(8):604-15.

187. Van Ballegooijen M, Rebolj M, Essink-Bot ML, Meerding WJ, Berkers LM, Habbema JDF. De effecten en kosten van het bevolkingsonderzoek naar baarmoederhalskanker in Nederland na de herstructurering. Erasmus MC, afdeling Maatschappelijke Gezondheidszorg, 2006.

188. van den Akker-van Marle ME, van Ballegooijen M, van Oortmarssen GJ, Boer R, Habbema JD. Cost-effectiveness of cervical cancer screening: comparison of screening policies. J Natl Cancer Inst. 2002 Feb 6;94(3):193-204.

189. Eichler HG, Kong SX, Gerth WC, Mavros P, Jonsson B. Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? Value Health. 2004 Sep-Oct;7(5):518-28.

190. Dugue PA, Lynge E, Rebolj M. Mortality of non-participants in cervical screening: Register-based cohort study. Int J Cancer. 2013 Nov 8.

191. Snijders PJ, Verhoef VM, Arbyn M, Ogilvie G, Minozzi S, Banzi R, et al. High-risk HPV testing on self-sampled versus clinician-collected specimens: a review on the clinical accuracy and impact on population attendance in cervical cancer screening. Int J Cancer. 2013 May 15;132(10):2223-36.

192. Broberg G, Gyrd-Hansen D, Miao Jonasson J, Ryd ML, Holtenman M, Milsom I, et al. Increasing participation in cervical cancer screening: offering a HPV self-test to long-term non-attendees as part of RACOMIP, a Swedish randomized controlled trial. Int J Cancer. 2014 May 1;134(9):2223-30.

193. Virtanen A, Nieminen P, Luostarinen T, Anttila A. Self-sample HPV tests as an intervention for nonattendees of cervical cancer screening in Finland: a randomized trial. Cancer Epidemiol Biomarkers Prev. 2011 Sep;20(9):1960-9.

194. Szarewski A, Cadman L, Mesher D, Austin J, Ashdown-Barr L, Edwards R, et al. HPV self-sampling as an alternative strategy in non-attenders for cervical screening - a randomised controlled trial. Br J Cancer. 2011 Mar 15;104(6): 915-20.

195. Giorgi Rossi P, Marsili LM, Camilloni L, Iossa A, Lattanzi A, Sani C, et al. The effect of self-sampled HPV testing on participation to cervical cancer screening in Italy: a randomised controlled trial (ISRCTN96071600). Br J Cancer. 2011 Jan 18; 104(2):248-54.

196. Pignone M. Understanding preferences for CRC screening programs among vulnerable adults in rural North-Carolina: a discrete choice experiment. Paper presented at 35th annual North American meeting of the Society for Medical Decision Making; October 2013; Baltimore, MD.

197. Bosgraaf RP, Ketelaars PJ, Verhoef VM, Massuger LF, Meijer CJ, Melchers WJ, et al. Reasons for non-attendance to cervical screening and preferences for HPV self-sampling in Dutch women. Prev Med. 2014 Jul;64:108-13.

198. de Kok IM, van Ballegooijen M, Habbema JD. Cost-effectiveness analysis of human papillomavirus vaccination in the Netherlands. J Natl Cancer Inst. 2009 Aug 5;101(15):1083-92.

199. Statline Database [database on the Internet]2013.

200. SIG (Information Centre for Health Care). Hospital Diagnosis Statistics 1963-1985. Utrecht: SIG1985.

201. CBS (netherlands Central Bureau of Statistics). Death by cause of death, age and sex 1950-1992. Voorburg1994.

202. Castle PE. When Is It Effective to Offer Self-Sampling to Non-Attendees-Letter. Cancer Epidemiol Biomarkers Prev. 2015 Aug;24(8):1295.

203. Rozemeijer K, de Kok IM, Naber SK, van Kemenade FJ, Penning C, van Rosmalen J, et al. Offering Self-Sampling to Non-Attendees of Organized Primary HPV Screening: When Do Harms Outweigh the Benefits? Cancer Epidemiol Biomarkers Prev. 2015 May;24(5):773-82.

204. Segnan N. Socioeconomic status and cancer screening. IARC Sci Publ. 1997(138):369-76.

205. Link BG, Northridge ME, Phelan JC, Ganz ML. Social epidemiology and the fundamental cause concept: on the structuring of effective cancer screens by socioeconomic status. Milbank Q. 1998;76(3):375-402, 304-5.

206. Orbell S, Crombie I, Robertson A, Johnston G, Kenicer M. Assessing the effectiveness of a screening campaign: who is missed by 80% cervical screening coverage? J R Soc Med. 1995 Jul; 88(7):389-94.

207. Benard VB, Johnson CJ, Thompson TD, Roland KB, Lai SM, Cokkinides V, et al. Examining the association between socioeconomic status and potential human papillomavirus-associated cancers. Cancer. 2008 Nov 15;113(10 Suppl):2910-8.

208. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol. 1990 Sep;19(3):553-8.

209. Knol FA. Statusontwikkeling van wijken in Nederland 1998-2010 [in Dutch]. Sociaal en Cultureel Planbureau; 2012.

210. Kenyon S, Sweeney BJ, Happel J, Marchilli GE, Weinstein B, Schneider D. Comparison of BD Surepath and ThinPrep Pap systems in the processing of mucus-rich specimens. Cancer Cytopathol. 2010 Oct 25;118(5):244-9.

211. Sweeney BJ, Haq Z, Happel JF, Weinstein B, Schneider D. Comparison of the effectiveness of two liquid-based Papanicolaou systems in the handling of adverse limiting factors, such as excessive blood. Cancer. 2006 Feb 25;108(1):27-31.

212. Zappa M, Visioli CB, Ciatto S, Iossa A, Paci E, Sasieni P. Lower protection of cytological screening for adenocarcinomas and shorter protection for younger women: the results of a case-control study in Florence. Br J Cancer. 2004 May 4;90(9): 1784-6.

213. Sasieni P, Castanon A, Cuzick J. Effectiveness of cervical screening with age: population based case-control study of prospectively recorded data. BMJ. 2009;339:b2968.

214. de Bekker-Grob EW, de Kok IM, Bulten J, van Rosmalen J, Vedder JE, Arbyn M, et al. Liquid-based cervical cytology using ThinPrep

technology: weighing the pros and cons in a cost-effectiveness analysis. Cancer Causes Control. 2012 Aug;23(8):1323-31.

215. Garcia F, Barker B, Santos C, Brown EM, Nuno T, Giuliano A, et al. Cross-sectional study of patient- and physician-collected cervical cytology and human papillomavirus. Obstet Gynecol. 2003 Aug;102(2):266-72.

216. Health Council of the Netherlands. Population screening for cervical cancer. The Hague: Health Council of the Netherlands, 2011; publication no. 2011/07.

217. Rapport 2014 (in Dutch). Landelijke Monitoring & Evaluatie Bevolkingsonderzoek Darmkanker.

218. van Hees F, Zauber AG, van Veldhuizen H, Heijnen ML, Penning C, de Koning HJ, et al. The value of models in informing resource allocation in colorectal cancer screening: the case of the Netherlands. Gut. 2015 Jun 10.

219. Vink MA. Modeling the effects of cervical cancer prevention in the Netherlands [thesis]. Amsterdam, the Netherlands, Department of Epidemiology and Biostatistics, VU University Medical Center Amsterdam;2015.

220. Bogaards JA, Wallinga J, Brakenhoff RH, Meijer CJ, Berkhof J. Direct benefit of vaccinating boys along with girls against oncogenic human papillomavirus: bayesian evidence synthesis. BMJ. 2015;350:h2016.

221. de Graaf H, Meijer S, Poelman J, et al. Sex below the age f 25; Sexual Health of Youth in the Netherlands in 2005 [in Dutch]. Delft, the Netherlands, Eburon Academic Publishers 2005.

222. Bakker F and Vanwesenbeeck I. Sexual health in the Netherlands in 2006 [in Dutch]. Delft, the Netherlands, Eburon Academic Publishers 2006.

APPENDICES

Dankwoord

PhD Portfolio

Curriculum Vitae

Publications

DANKWOORD

Tot slot mijn persoonlijke favoriet, namelijk het dankwoord. Het is altijd leuk om anderen eens in het zonnetje te zetten, wat waarschijnlijk ook de reden is dat ik alreeds in mijn 2e jaar aan dit hoofdstuk begonnen ben!

Allereerst wil ik iedereen bedanken die mij door de loop van de tijd hebben begeleidt (Inge, Joost, Marjolein, Corine en natuurlijk mijn promotor Harry) en daarbij Inge en Corine in het bijzonder. Jullie zijn toch degenen die mij het langst hebben bijgestaan en waar ik dan ook ontzettend veel van heb geleerd. Inge, onze brainstormsessies waren van onschatbare waarde en hebben zeker bijgedragen aan de mooie lijn die in mijn papers en boekje te vinden zijn. Corine, je wist altijd precies de juiste toon aan te slaan en ik heb genoten van onze fijne samenwerking. Hopelijk gaan we binnenkort dat wijntje (of ons kennende meerdere wijntjes) nuttigen waar we het al zo lang over hebben.

Daarnaast wil ik al mijn co-auteurs (Bert, Caspar, Corine, Folkert, Inge, Joost, Lucy, Matejka, Marjolein, Steffie en Suzette) bedanken voor hun kritische noot. Af en toe erg frustrerend (de laptop is meerdere malen begonnen aan zijn weg richting open raam), maar het leidde altijd tot een sterk verbeterd paper! In het bijzonder dank aan Folkert voor je vele mee denken en behulpzame commentaar. Also many thanks to Matejka. Your nice words and stories from first hand really helped me from getting too frustrated and depressed when one of my papers was rejected again.

Roel, jij bedankt voor je vele geduld tijdens 1 van onze velen "hoe te programmeren met SQL" sessies. Ik denk dat je nog nooit zo vaak een opdracht hebt moeten "killen".

Ook dank aan de RCPen voor het verzamelen van de DLC data en natuurlijk Bert en Lucy voor het linken hiervan aan de PALGA/PALEBA database. Zonder jullie waren hoofdstuk twee en vier nooit tot stand gekomen!

Prof. van Kemenade (of moet ik gewoon Folkert zeggen zoals ik altijd doe?), Prof. Bindels en Prof. Massuger, bedankt voor het kritisch beoordelen van mijn boekje. Daarnaast wil ik jullie, en natuurlijk ook de andere leden van de commissie, bij voorbaat danken voor jullie kritische vragen tijdens mijn verdediging. Ik ben nu al peentjes te zweten!

Natuurlijk mag ook de gezelligheid niet ontbreken. Suzette, wat hebben we een hoop avonturen beleefd in Berlijn en Puerte Rico. Wijntjes, zwemmen, jacuzzi parties op ons dakterras... Je kan het zo gek niet bedenken. Je hebt me zelfs aan het salsa dansen gekregen! Natuurlijk ook dank richting de andere HPV gangers die er zo'n geweldige tijd van maakten. Ook Joost, Inge, Steffie en Corine bedankt voor de gezellige ontbijtjes, lunches en avondjes tijdens 1 van de vele congressen. Jesper, thank you for taking me under your wings (is that even an English saying?) at the cytology conference in Geneva and for all the nice talks we had.

My dear roomies: Thanks for all the nice moments we had together. Too bad we never confiscated that couch! Sonja en Alex, jullie bedank ik in het bijzonder. Ik heb genoten van onze vele wandelingen, lunches en niet te vergeten onze vele mopper momentjes (jaja, het leven van een promovendi bestaat uit frustraties, mopperen en weer doorgaan, zie ook stelling 11).

Ook speciale dank richting mijn twee lieve paranimfen, Sonja en Björn. Sonja, vanaf moment 1 roomies en friends! We hebben deze vier jaar samen doorlopen en wat was het fijn om samen te mopperen als één van de twee, of beide, er doorheen zat. Hopelijk zullen we elkaar blijven zien, ondanks dat we niet meer op dezelfde afdeling werken! Ik vind het een eer dat we elkaars paranimf mogen zijn! Mijn lieve "knur", ondanks onze verschillen (mega chaoot, of zoals jij altijd zegt verstrooide professor, regelkip en controlfreak versus netjes, opgeruimd maar ook erg lay-back) toch twee handen op één buik! Wat ben ik trots dat mijn allang niet meer zo kleine broertje naast me staat! Daar drinken we na de verdediging samen een paar overheerlijke biertjes op!

Mijn lieve Alphen-gang (Bart, Bart, Chang, Mark, Emilio, Björn, Cynthia, Marijn, Rock, Daniel, Anne, Sander, Stephie, Bobbo en Rianne). Wat zou ik toch zonder jullie zijn? Dank voor de vele mooie "we doen 1 drankje" avondjes, terasjes, weekendjes weg, afterparty's, koninginnedagavonturen, "diepzinnige" discussies met iets te veel drank op , pre-oud en nieuw party's en ga zo maar door. Alles is bij ons reden om gezellig bij elkander te komen en zonder jullie was mijn leven half niet zo leuk!

Mijn lieve Leiden vriendjes en vriendinnetjes, schaapjes en schaappinetjes. Ook jullie mogen niet in dit dankwoord ontbreken. Wat hebben we toch een mooie tijden beleefd en wat gaan we nog een mooie tijden beleven! Mijn lieve PB-matties (het worden er teveel om hier persoonlijk te bedanken), wat heb ik genoten van al onze gekke en maffe avonturen in Hepatho, de Hut, in een random Nederlandse stad of ergens in het verre buitenland! Mijn studentenleven had niet mooier kunnen zijn! Ik kan niet wachten tot we alweer op de 3e lustrumreis gaan! Ik stem wederom als locatie op die rokende vulkaan. Ooit zal ik toch eens gelijk krijgen? Ook nog een speciaal woord richting mijn geliefde Club-de-Groep. Jullie hebben me onder jullie hoede genomen toen ik nog een klein, schattig en vooral erg druk sjaarsje was. Door jullie raakte ik al snel ingeburgerd in het Leidse leven en ik dank jullie voor alle mooie momenten die we samen hebben gehad en die er zeker ook nog gaan komen!! Op een mooie vriendschap. Proost!

Ook dank aan mijn lieve schone ouders en zus! Ik ben erg blij dat ik jullie in mijn leven heb, ik had me geen lievere schoonfamily kunnen wensen!

En nu, last but not least, dank aan mijn geliefde family (Paps & Mams, en natuurlijk ook Björn & Cynthia, en niet te vergeten: de liefde van mijn leven, Marijn). De afgelopen vier jaar, en ook de jaren daarvoor, hebben jullie me op elk mogelijke manier gesteund. Jullie boden een luisterend oor als een van mijn papers weer eens was afgewezen, jullie stelden doordachte vragen over mijn onderzoek (waar ik soms het antwoord nog niet eens op wist)

en jullie proostten mee toen mijn 1e papers eindelijk werden geaccepteerd. Dank je wel dat jullie er altijd voor me zijn, deze is voor jullie!

Lieve Marijn, als allerlaatst (nu echt) richt ik me tot jou persoonlijk. Wat hebben we de afgelopen 10 jaar toch een hoop meegemaakt en wat vullen we elkaar toch goed aan. Door jou heb ik geleerd minder te plannen en het leven meer te nemen zoals het is. Ik kan niet wachten om de rest van mijn verdere leven met jou te delen. Ik hou van je. Op naar ons volgende avontuur!

PHD PORTFOLIO

Summary of PhD training and teaching

Name PhD student: Kirsten Rozemeijer
PhD Period: 2011-2015
Erasmus MC Department: Public Health
Promotor: Prof. dr. H.J. de Koning
Supervisor: dr. I.M.C.M. de Kok

1. PhD Training	Year	Workload (ECTS)
General courses		
Master of Public Health, Netherlands Institute for Health Sciences (NIHES), Rotterdam	2011 - 2014	70
Seminars and symposia		
Seminars at the Department of Public Health, Erasmus MC	2011 - 2015	3.6
QIAGEN symposium , Utrecht	2012	0.2
Symposium Baarmoederhalskanker Bevolkingsonderzoek Oost, Papendal	2012	0.2
QIAGEN symposium, Utrecht	2013	0.2
RIVM, expertmeeting, Utrecht.	2014	0.2
Wetenschappelijk symposium PALGA, Utrecht	2014	0.2
Symposium Baarmoederhalskanker Bevolkingsonderzoek Oost, Papendal	2014	0.2
Presentations		
Oral presentation. Eurogin, Prague, Czech Republic	2012	1.0
Oral presentation. Research meeting at the department of Public Health, Erasmus MC, Rotterdam	2012	0.6
Oral presentation. IPV conference, San Juan, Puorte Rico	2012	1.0
Oral presentation. Eurogin, Florence, Italy	2013	1.0
Oral presentation [in Dutch]. RIVM, expertmeeting, Utrecht.	2014	0.8

Oral presentation. European Congress of cytology, Geneve, Zwitserland	2014	1.0
Oral presentation [in Dutch]. Wetenschappelijk symposium PALGA, Utrecht	2014	0.8
Oral presentation [in Dutch]. Symposium Baarmoederhalskanker Bevolkingsonderzoek Oost, Papendal	2014	0.8
Oral presentation. Eurogin, Sevilla, Spain	2015	1.0

(Inter)national conferences

IPV conference, Berlin, Germany	2011	1.0
WEON, Rotterdam	2012	0.6
Eurogin, Prague, Czech Republic	2012	0.8
IPV conference, San Juan, Puerto Rico	2012	1.0
Eurogin, Florence, Italy	2013	0.8
IFCPC, London, United Kingdom	2014	0.6
European Congress of Cytology, Geneva, Zwitserland	2014	0.6
Eurogin, Sevilla, Spain	2015	0.8
ICSN, Rotterdam	2015	0.6

Other

Peer review for BMC Cancer	2015	0.5
Interview by Prof. dr. T.C. Wright for ReachMD	2015	5 hrs

2. Teaching	Year	Workload (ECTS)
Audit bachelor thesis: Curriculum medical students 3rd year	2012	36 hrs

CURRICULUM VITAE

Kirsten Rozemeijer was born on the 23th of March 1987, in Rijnsburg, the Netherlands. In 2005, she completed her secondary education at the Scala College in Alphen a/d Rijn and she started studying Biomedical Sciences in Leiden. During her bachelor internship at the department of Clinical Epidemiology in the LUMC, she compared the performance of two different test methods to measure kidney function. This resulted in a co-authorship in an international paper and she presented the results at a national and international conference for epidemiologists and students, respectively. During her research master in the Biomedical Sciences, she decided to focus more on the epidemiology and public health of scientific research. She wrote an essay about the advantages of including men in studies on female related risk factors which both resulted in a paper that has been submitted to an international scientific journal and a presentation at an international student conference. She performed laboratorial research comparing the antigen capacity of naïve versus Epstein-Barr virus transformed non-human primate B-cells at the Biomedical Primate Research Centre at Rijswijk. In addition, she compared the immune response after yellow fever vaccination between young and old vaccinees at the department of Infectious Diseases at the LUMC. She obtained her Master of Science degree in 2011.

From 2011 to 2015, she was employed as researcher at the department of Public Health at the Erasmus MC in Rotterdam where she evaluated the Dutch cervical cancer screening programme by using the nationwide network and registry of histo- and cytopathology (PALGA). During this period she also obtained her Master of Public Health degree at the Netherlands Institute for Health Sciences.

LIST OF PUBLICATIONS

Matthijsse SM, Hontelez JAC, Naber SK, **Rozemeijer K,** Penning C, Bakker R, van Ballegooijen M, van Rosmalen J, de Vlas SJ. Public health benefits of routine human papillomavirus vaccination for adults: a mathematical modelling study. *Soon to be submitted*

Rozemeijer K, Naber SK, Penning C, Overbeek LI, Looman CWN, de Kok IMCM, Matthijsse SM, Rebolj M, van Kemenade FJ, van Ballegooijen M. Cervical cancer incidence after a negative cytological smear in routine screening: Comparing SurePath, ThinPrep and conventional cytology. *Submitted*

Naber SK, Matthijsse SM, **Rozemeijer K**, Penning C, de Kok IMCM, van Ballegooijen M. Cervical cancer screening in a partly HPV vaccinated population - a cost-effectiveness analysis. *Submitted*

van Luijt P, **Rozemeijer K**, Naber SK, Heijnsdijk EAM, van Rosmalen J, van Ballegooijen M, de Koning HJ. The role of pre-invasive disease in overdiagnosis: a microsimulation study comparing mass screening for breast cancer and cervical cancer. *Submitted*

Rozemeijer K, le Cessie S, van Hylckama Vlieg A, Rosendaal FR, Vandenbroucke JP, Poole C, Cannegieter SC. Exposure opportunity: The advantages of including men in analyses on female-related risk factors. *Submitted*

Rozemeijer K, Penning C, van Kemenade FJ, Naber SK, Matthijsse SM, Siebers AG, de Kok IMCM, van Ballegooijen M. The association between socioeconomic status and the underlying screen-independent cervical cancer risk. *Submitted*

Rebolj M, Rask J, van Ballegooijen M, Kirschner B, **Rozemeijer K,** Bonde J, Rygaard C, Lynge E. Histological outcomes with liquid--based cytology

and computer--assisted reading: Data from routine cervical screening with ThinPrep and SurePath technologies in Denmark. Br J Cancer. 2015 Oct 8. *[Epub ahead of print]*

Rozemeijer K, Penning C, Siebers AG, Naber SK, Matthijsse SM, van Ballegooijen M, van Kemenade FJ, de Kok IMCM. Comparing SurePath, ThinPrep and conventional cytology as primary test method: SurePath is associated with increased CIN II^{+} detection rates. Cancer Causes Control. 2015 Oct 12. *[Epub ahead of print]*

Matthijsse SM, Hontelez JAC, Naber SK, van Rosmalen J, **Rozemeijer K,** Penning C, van Ballegooijen M, de Kok IMCM, de Vlas SJ. The estimated impact of natural immunity on the effectiveness of human papillomavirus vaccination. Vaccine. 2015 Oct 5;33(41):5357-64.

Rozemeijer K, de Kok IM, Naber SK, van Kemenade FJ, Penning C, van Rosmalen J, van Ballegooijen M. When is it effective to offer self-sampling to non-attendees—Response. Cancer Epidemiol Biomarkers Prev. 2015 Aug;24(8):1296.

Rozemeijer K, van Kemenade FJ, Penning C, Matthijsse SM, Naber SK, van Rosmalen J, van Ballegooijen M, de Kok IM. Exploring the trend of increased cervical intraepithelial neoplasia detection rates in the Netherlands. Journal of Medical Screening. J Med Screen. 2015 Sep;22(3):144-50.

Rozemeijer K, de Kok IM, Naber SK, van Kemenade FJ, Penning C, van Rosmalen J, van Ballegooijen M. Offering self-sampling to non-attendees of organized primary HPV screening: When do harms outweigh the benefits? Cancer Epidemiol Biomarkers Prev. 2015 May;24(5):773-82.

de Kok IM, van Rosmalen J, **Rozemeijer K,** Penning C, van Ballegooijen M. How many cervical cancer cases can potentially be prevented using a more sensitive screening test at young age? Int J Cancer. 2014 Jan 15;134(2):460-6.

Michels WM, Grootendorst DC, **Rozemeijer K,** Dekker FW, Krediet RT. Glomerular filtration rate measurements by 125I-iothalamate should be corrected for inaccurate urine collections with 131I-hippuran. Clin Nephrol. 2009 Nov;72(5):337-43.