HHS Public Access

Author manuscript

Int J Cancer. Author manuscript; available in PMC 2018 March 01.

Published in final edited form as:

Int J Cancer. 2017 March 01; 140(5): 1215–1222. doi:10.1002/ijc.30524.

HARMS OF CERVICAL CANCER SCREENING IN THE UNITED STATES AND THE NETHERLANDS

Dik Habbema, **Sheila Weinmann**, **Marc Arbyn**, **Aruna Kamineni**, **Andrew E Williams**, **Inge MCM de Kok**, **Folkert van Kemenade**, **Terry S Field**, **Joost van Rosmalen**, and **Martin Brown**

Abstract

We studied harms related to cervical cancer screening and management of screen-positive women in the United States (US) and the Netherlands. We utilised data from four US integrated health care systems (SEARCH), the US National Health Interview Survey, New Mexico state, the Netherlands national histopathology registry, and included studies on adverse health effects of cervical screening. We compared the number of Papanicolaou (Pap) smear tests, abnormal test results, punch biopsies, treatments, health problems (anxiety, pain, bleeding, and discharge), and preterm births associated with excisional treatments. Results were age-standardized to the 2007 US population. Based on SEARCH, an estimated 36 million Pap tests were performed in 2007 for 91 million US women aged 21–65 years, leading to 2.3 million abnormal Pap tests, 1.5 million punch biopsies, 0.3 million treatments for precancerous lesions, 5 thousand preterm births, and over 8 million health problems. Under Netherlands screening practice, fewer Pap tests (58%), abnormal test results (64%), punch biopsies (75%), treatment procedures (40%), preterm births (60%), and health problems (63%) would have occurred. The SEARCH data did not differ much from other US data for 2007 or from more recent data up to 2013. Thus compared with the less intensive screening practice in the Netherlands, US practice of cervical cancer screening may have resulted in two- to three-fold higher harms, while the effects on cervical cancer incidence and mortality are similar. The results are also of high relevance in making recommendations for HPV screening. Systematic collection of harms data is needed for monitoring and for better incorporation of harms in making screening recommendations.

Keywords

Cervical cancer; Screening; Pap test; Harms; United States; Netherlands; Health problems

INTRODUCTION

Reduced cervical cancer incidence and mortality are well-known benefits of cervical cancer screening (1,2). However, screening can also produce harms. Screening-driven diagnostic and therapeutic events - including the cytological Papanicolaou (Pap) test, abnormal cytology test results, punch biopsies, and treatment procedures - can cause physical and psychological problems (3–5). Treatment of pre-cancerous cervical lesions can lead to

Adress of corresponding author: Dik Habbema: Dep. of Public Health, Erasmus MC University Medical Center Rotterdam, PO box 2040, 3000 CA Rotterdam, the Netherlands, j.d.f.habbema@erasmusmc.nl.

adverse pregnancy outcomes, such as preterm birth (6,7). A variety of reviews have concluded that harms have been understudied in screening trials (8) and in other studies (9,10), that estimates of studied harms can be biased (11), and that potential harms are underestimated by patients (12). Moreover, harms resulting from possible overdiagnosis and overtreatment in screening are relevant to the discussion of "too much medicine." (13, www.bmj.com/too-much-medicine).

Because approaches to delivering preventive health care differ across countries, cross-national comparisons of harms of preventive practices may be informative. We previously compared the benefits of cervical cancer screening practices in the United States (US) and the Netherlands (NL) by using the number of Pap tests as a measure of screening intensity and cervical cancer mortality and incidence rates and trends as a measure of screening benefit. We observed that mortality and incidence trends were similar in the two countries, despite a much greater screening intensity in the US which is no surprise in view of the difference in screening recommendations between the US and the NL, with many more life-time screenings advised in the US (14).

The current study intends to enrich the medical harms literature with an empirical study of the harms related to screening and management of screen-positive women in the US and NL. Although we expected beforehand that the US, which screens more intensively than the NL does, would accrue more harms, we were very uncertain about the possible magnitude of the difference.

METHODS

screening data

United States—To estimate harms of cervical cancer screening, the Screening Effectiveness and Research in Community-based Healthcare (SEARCH) database is used. SEARCH was created by a National Cancer Institute-funded research project conducted in four integrated health care systems in the states of Hawaii, Massachusetts, Oregon and Washington affiliated with the Health Care Systems Research Network. The SEARCH database links cytology, biopsy, histology, treatment, and health plan membership data (15,16). Since it is the most comprehensive US population-based database on cervical cancer screening and follow-up, we used it as a baseline for our analysis. Data quality was assessed with tests for missing data, temporal stability, logical consistency, and value range plausibility. Patient-level inclusion criteria were female gender, age 21–65 years during any study year, and no known history of a gynaecologic cancer or total hysterectomy at cohort entry. As one of the SEARCH sites had incomplete cervical histology information, we used histology data from the three other sites only.

The Netherlands—We used data from the NL Pathological Anatomical National Automated Archive (PALGA), a nationwide histopathology and cytopathology network-based archive that encompasses all pathology laboratories in the country (17). PALGA includes a cervical cancer screening database complete for all cytology and histology testing, regardless of the technique, setting, and reason. For every test registered in PALGA, the woman's (coded) identification, the date, and the topographic and morphologic codes

(18) are also recorded, making screening and diagnostic histories available for all women. PALGA has instructions and structured data entry for pathology and cytology reports. Labs are notified on missing or paradoxical data. See (17) and www.palga.nl. PALGA does not include treatment, and there are no other treatment data for NL.

screening-related events

Pap Test—We included all Pap tests (conventional or liquid based) in the SEARCH and PALGA registries, including primary screening tests, follow-up tests, and Pap tests done for other reasons. SEARCH results were considered abnormal unless they were defined as "normal" or "unsatisfactory/inadequate/rejected." PALGA results were considered abnormal unless they were classified as either "normal" or "inadequate."

Punch Biopsy—For SEARCH, histological records that could not be matched to CPT codes for cervical treatment were considered punch biopsies. For NL we estimated the number of punch biopsies by subtracting the number of treatments from the total number of histological exams, under the plausible assumption of an average of one histological exam per treatment. Both SEARCH and PALGA register all histological procedures and their outcomes. We included CIN1 and CIN2+ as histological outcome categories, with CIN2+ consisting of CIN2, CIN3, and all invasive cancers of the cervix.

Treatment—We based age-specific US treatment rates on data from the three SEARCH health systems for which 2007 CPT and histology information was complete. PALGA does not track NL treatment data. As follow-up compliance is very high in the NL(19), we assumed that all CIN2+ cases have been treated, realizing that this is an upper bound. Making an assumption on the percentage CIN1 treated is more problematic. We used as a proxy 2008 data from an Italian study (20) showing that 25% of women with CIN1 had undergone treatment. Like in the Netherlands, follow-up is well organized in Italy, and virtually all women with a histological diagnosis of CIN2+ had undergone treatment (20).

screening-related health problems

To estimate screening harms, we identified three population-based studies of psychological and physical problems associated with cervical cancer screening. Korfage et al. (5) studied symptoms of the Pap test. Drolet et al. (3) reported on women who experienced anxiety after being informed about an abnormal cytology result. The TOMBOLA research group (4) studied adverse health effects in women following cervical punch biopsy and loop electrosurgical excision treatment procedures (LEEP; also known as LLETZ, large loop excisions of the transformation zone). Table 1 summarizes the results of these studies.

We confined the assessment of reproductive harms to increased risk of preterm delivery caused by excisional treatment of CIN. Kyrgiou et al. (7) found that only excisions exceeding 10mm in depth lead to measurable harms. An update of (7) found relative risk (RR) of preterm delivery of 1.75 for women with excisional treatment exceeding 10 mm in depth compared to women with smaller excisions (see ANNEX FIGURE). This approach might be conservative. We cannot exclude the possibility that shallow excisions also lead to a somewhat increased risk of preterm delivery. On the other hand, women with large

excisions might be statistically different from women with shallow excisions on unknown factors related to preterm birth.

The SEARCH study had no information on depth of excision. However, a study in a cohort of women of reproductive age at one of the four SEARCH health systems found that 23% of excisions exceeded 10 mm (21).

Apart from the baseline calculations with 23% deep excisions and a RR of 1.75, we also considered a higher- and a lower-risk scenario for studying the impact of the uncertainty in these assumptions. Numbers of about 50% deep excisions were found in studies conducted in Belgium and England (22,23). We used these European values for the higher-risk scenario for the US, as deep excisions are probably less frequent in the US than in Europe (24). For the lower-risk scenario, we arbitrarily chose a value of 18.4% (80% of the baseline value). We used RR values of 2.19 and 1.31 for the higher and lower risk scenarios (125% and 75%, respectively, of the baseline value).

analysis of 2007 data

We chose 2007 as the year of analysis for the US and NL data, as it was the most recent year for which SEARCH has complete data. The SEARCH database is restricted to women age 21–65 years. We clustered age in 5-year intervals within this range. Age-specific rates of events (Pap tests, abnormal Pap test results, punch biopsies, and treatments) in SEARCH and the NL database were calculated by dividing the number of events by the total number of women in each age category. For calculating overall population figures, age-specific results were standardized to the female US population in 2007 (25). All rates are presented per 1000 women. The number of US women experiencing a specific event was calculated by multiplying the standardized rate with the total number of US women aged 21–65 years in 2007. We compared US and NL screening practices using ratios of the standardized rates and numbers.

The number of health problems associated with events (Pap test, abnormal test result, punch biopsy, and treatment) was estimated by multiplying the age-standardized rate of the event by the number of women estimated to experience those health problems. We calculated rates, rate ratios, and number of US women experiencing health problems under US and NL screening practices.

To estimate the number of preterm births in the US in 2007 that could be associated with screening, we multiplied the age-specific rates of excisional treatment in the SEARCH population by the estimated fraction of deep excisions. In turn, that result was multiplied by the expected number of childbirths after the age of treatment, according to the 2007 US birth statistics (26). The number of preterm births caused by treatment was estimated using 2007 data on US preterm births (27) and the RR of preterm birth after deep excisional treatment, assuming no increased risk after shallow excisions. We also estimated the number of preterm births when applying NL treatment rates.

RESULTS

Among women aged 31–60 years, the rate of Pap testing was roughly twice as high in the SEARCH based US data as in the NL (Table 2). Because NL rarely tests women younger than 31 or older than 60, the overall Pap testing rates in theses age classes is much higher in the US than in the NL. For abnormal Pap test results and punch biopsies, the difference in age-specific rates between US and NL was greater than that for the Pap test, whereas the difference was smaller for treatment (Table 2). See the Annex Table for additional age-specific results.

Table 3 shows that compared with NL screening and follow-up practice, age-standardized rates for the US female population in 2007 were higher for Pap testing (2.4), abnormal test result (2.8), punch biopsy (3.9), and CIN treatment (1.7). The rightmost columns of Table 3 show the estimated number of events in the total female US population 21–65.

Table 4 gives estimates of the number of women in 2007 who experienced health problems stemming from Pap tests and their downstream events. Under US screening practice, an estimated 4.6 million women experienced health problems due to the Pap test itself, 0.8 million women felt anxious because of an abnormal test result, and almost 3.3 million had health problems owing to punch biopsy or treatment (41% moderate or severe) (Table 4). The number of estimated health problems due to cervical cancer screening-related events totalled 8.7 million. If NL screening practice had been used in the US population, an estimated 63% fewer health problems would have occurred. (Table 4).

With the baseline assumptions of a 1.75 relative risk of preterm delivery associated with a history of excisional treatment exceeding 10 mm in depth and an average of 23% of excisions exceeding 10 mm, US screening practice in 2007 led to an estimated 5,300 preterm births (Table 5). For the lower risk scenario, the number of preterm births was three times lower compared to the baseline assumptions, while for the higher risk scenario it was three times higher. Under NL screening practice, all these numbers of preterm births would have been about 60% lower.

DISCUSSION

We compared the harms of cervical cancer screening between the United States (US) and the Netherlands (NL) using data from four US integrated health plans and a national database for NL. Our main finding is that harms occur much more frequently in US than in NL, while the levels of incidence and mortality have been quite comparable between the two countries from the time that screening could have influenced these levels, suggesting that baseline risk is similar in the US and the NL (14). This is confirmed by similar HPV prevalence, also in women with normal cytology. For example, for women aged 25–34 HPV prevalence is estimated as 9.3% in the US and 8.5% in the NL (Figures 28 in the USA and NL country reports of the HPV information centre, accessed at www.hpvcentre.net on September 11 2016).

We found a higher rate of Pap tests (RR=2.4) in the US. The difference in rates was about the same for abnormal test results (RR=2.8), higher for punch biopsy (RR=3.9), and smaller

for CIN treatment (RR=1.7). From the rates in Table 3 it can be calculated that there were 6.3% abnormal test in the US compared to 5.5% in the NL. The number of punch biopsies per abnormal test was 0.67 in the US and 0.48 in the NL. And the number of treatments per punch biopsy was 0.18 in the US and 0.42 in the NL. These findings suggest that the threshold for taking a biopsy in case of an abnormal test result was lower in the US than in NL. In view of the earlier finding of similar cervical cancer mortality rates in the two countries (14), the 70% higher treatment rates with US screening practice might predominantly involve regressive lesions (28,29).

The SEARCH integrated health care systems data were used for estimating overall US rates. Although the SEARCH population had socio-economic and racial/ethnic profiles similar to the local communities (16), it only consists of participants in integrated health care systems, and may not be representative of the entire US population, which includes uninsured women and those in fee-for-service arrangements. Also, the SEARCH preventive services practice might not be representative for the US. We therefore compared SEARCH with other US population based data. The 2007 US National Health Interview Survey (NHIS) indicates a Pap test intensity of about 400 tests per 1000 women per year, including women below 21 and over 65 (14). The rate for the SEARCH sites is 400 for ages 21–65 and would be about 260–290 for all ages, depending on the amount of screening outside the 21–65 age range (30). These differences may partly be explained by the fact that the NHIS estimates are based on women's self-report, which may lead to an upward bias of 10% – 30% (31). Also, guidelines in SEARCH type of integrated health care systems recommended about 17 lifetime Pap tests (34), compared to about 25 tests for national guidelines (14,33–35).

In 2010, the NHIS asked US women about having an abnormal test results in the past 3 years (36,37). When linked with the NHIS data on having a Pap test in the last 3 years, about 10% of Pap tests gave an abnormal result for the age group 21–65, which is higher than the 6.5% in SEARCH. Again, differences might (partially) have been caused by over-reporting of abnormal tests (31).

State-wide cervical cancer screening data have been published for New Mexico for the years 2008–2011 (38). We used 2008 data to compare with the 2007 SEARCH data. New Mexico has an atypical population with 10% (US 1%) American Indian population and 47% Hispanic population (US 17%), and a 20% poverty rate (US 15%) (http://quickfactscensus.gov). Nevertheless, the age-specific screening rates were quite similar, about 2% lower than in SEARCH (38, supplemental Table 1). The rate of abnormal smears was almost identical for all age groups except for ages 40–50, with 24 per 1000 in SEARCH and 18 per 1000 in New Mexico (38, Supplemental Table 2).

In summary, the rates of Pap smear intensity, abnormal rates and treatment rates in SEARCH do not differ much from data from available national and state-wide US sources.

We have explored whether the 2007 results are applicable to more recent years. For NL, a small decease in screening (1% relative decrease per year) was documented between 2007 and 2013, along with an increasing trend in abnormal test results (3% relative increase per year) (39). A possible cause of the increase in abnormal tests might be the increase in liquid

based cytology in the NL from 65% in 2007 to nearly 100% by 2011(40). In the US, cytology was already 100% liquid-based in 2007 in SEARCH (16). The US National Health Interview Survey data show a small 1% relative decrease per year between 2008 and 2013 in the percent of women 25–64 having a Pap test within the past 3 years (39,40). In New Mexico there was a downward trend in Pap test intensity between 2008 and 2011 (5% relative decrease per year), and an equal upward trend in the rate of abnormal test results (38, supplementary Tables 1 and 2). Together, these data tell us that changes in screening rates and abnormal test results between 2007 and 2013 were modest and balanced, suggesting that a comparison of the US and the NL for these years would have led to quite similar results. SEARCH-based screening recommendations are not yet outdated for Pap test screening, as the 2007 guidelines in these integrated health plans (32) are comparable to current national US guidelines for cytology alone screening, both recommending a Pap test every 3 years for women aged 21–65 years (43,44).

One should not expect, however, that the impact of the new 2012 US guidelines will already be visible in 2013, the most recent year of comparison, since it takes a few years for new recommendations to be adopted in the decentralized screening practice in the US. The NL cytology screening guideline, which promotes Pap screening every 5 years to women aged 30–60 years, has not changed since 2007 (45).

In a comparison of harms-associated events, CIN treatment causes more serious health problems than a punch biopsy (Table 1). However, due to their greater volume punch biopsies cause more health problems on a population level (Table 4). Estimates of the frequency of health problems caused by CIN treatment and punch biopsies are uncertain given the paucity of studies. Nevertheless, the number of health problems associated with screening affects millions of US women each year.

Preterm birth can cause substantial human suffering as well as costs for neonatal intensive care and continued care after hospital discharge. A trend toward less aggressive treatment and a shift from extensive cold knife conisation to more conservative loop excisions have been observed over the last decades (46), but concerns have been raised regarding a concurrent decrease in therapeutic effectiveness (47). It is possible that our baseline assumption of 23% of excisions exceeding 10mm is not representative of overall practice in the US in 2007. Therefore we also explored the values of 18% and 50% in lower- and higher-risk scenarios. Assumptions on treatment rates for NL are not based on national data. It is nevertheless unlikely that the assumptions of 100% CIN2+ treatment and 25% CIN1 treatment are far off the mark in view of the high follow-up rates in the NL (19). When NL treatment rates would have been lower than assumed, the difference in preterm births between US and NL practice would increase. The harms to the female population (Table 4) would be reduced, but not very much, because less treatments implies more punch biopsies with their associated problems. Importantly, the lack of data concerning the treatment of CIN in both US and NL highlights the need for better registration and more research to define the oncologic and obstetric safety of available treatment modalities.

Our finding that harms are more frequent under US than NL cervical screening practice while the level and downward trends in cervical cancer mortality and incidence are

comparable between the two countries (14) suggests that changes in US screening and follow-up practice would reduce the harms of cervical cancer screening without substantially reducing the benefits for US women. Apart from re-examining the frequency of and age range for screening, these results may encourage revision of protocols for follow-up and treatment (48). It is noteworthy that the American Cancer Society, the American Society for Colposcopy and Cervical Pathology, and the American Society for Clinical Pathology considered harms when revising cervical cancer screening guidelines in 2012, using colposcopies as a proxy for all harms (44). Likewise, the US Preventive Services Task Force 2012 guidelines concluded that the potential harms outweigh the benefits for women under the age of 21 years and over the age of 65 years (43).

We did not include screening data for women younger than 21 years or older than 65 years because these ages are outside the age range in the SEARCH database (16). Including those women would increase the differences between the US and the NL because almost no screening is conducted at younger ages in the NL, and for women age 65 and older the 5-year screening rate was only 10% in the NL, compared to 60% in the US (14).

Hysterectomy rates have been roughly a factor 2 higher in the US than in the NL (49,50). This implies that if we had considered the female population at risk (with uterus), the difference in results between the US and the NL would have been larger, especially at older ages

A limitation of our study is our focus only on Pap screening. Cervical cancer screening is becoming increasingly complex, with recommendations involving combinations of cytological and human papillomavirus (HPV) testing (44). HPV-based screening is likely to result in higher screen-test positivity rates (51,52). Without strict adherence to evidence-based policies, such as not starting HPV testing before the age of 30 years, harms could increase (51). In the future, HPV vaccination will reduce the incidence of cervical cancer and therefore lead to a less favourable balance between the benefits and harms of screening in HPV-vaccinated women. Guidelines for HPV-based screening recommend more screening visits in the USA, 5-year intervals for co-testing with cytology and HPV tests (44) or 3-year intervals for HPV testing alone (53) than in the NL, where starting 2017, HPV testing alone will be offered every 5 years between ages 30 and 40, followed by two screens at 50 and 60 if HPV negative at the age of 40 (54). In 2007 and the following years, HPV testing still played a minor role. In the US, HPV co-testing increased in New Mexico from 5% in 2007 to 19% in 2013 (55), and in SEARCH from no HPV testing in 2002 to about 7% in 2007(16). In the NL, HPV testing was only done in trials and in some cases of triage, counting for a few percent of all tests. This may change completely when the new HPV testing program is implemented in the NL, foreseen for 2017 (54).

US cervical cancer screening guidelines have been updated to recommend less intensive screening, particularly for younger women. Our evidence suggests that such revisions are warranted. Further revisions should be based on quantitative evidence on the balance of harms and benefits that, to date, has not been available on a population basis for the US. While the NL has more extensive population-based data for cervical cancer screening and follow-up, the lack of data on treatment for precancerous lesions is a serious gap. The

development of such data resources would greatly facilitate evidenced-based evaluation and could hasten the improvement of screening practice in both countries (14).

In conclusion, although the benefits of cervical cancer screening are similar between the US and the NL, many more women in the US experience harms than would be the case under the less intensive NL screening practice. There is a lack of systematically collected harms data in both countries. Filling this gap is critically needed for better informed balancing of benefits against harms in making screening recommendations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Elizabeth Hess (the KPNW Center for Health Research) has repeatedly and carefully reviewed earlier versions of this manuscript.

REFERENCES

1. Laara E, Day NE, Hakama M. Trends in mortality from cervical cancer in the Nordic countries: association with organised screening programmes. Lancet. 1987; 1:1247–1249. [PubMed: 2884378]
2. IARC working group on evalution of cervical cancer screening programs. Screening for squamous cervical cancer: duration of low risk after negative results of cervical cytology and its implications for screening programs. BMJ. 1986; 293:659–664. [PubMed: 3092971]
3. Drolet M, Brisson M, Maunsell E, Franco EL, Coutlee F, Ferenczy A, Fishr W, Mansi JA. The psychosocial impact of an abnormal cervical smear result. Psycho-oncology. 2012; 21:1071–1081. [PubMed: 21695747]
4. The TOMBOLA Group. After-effects reported by women following colposcopy, cervical biopsies and LLETZ: results from the TOMBOLA trial. BJOG. 2009; 116:1506–1514. [PubMed: 19583712]
5. Korfage IJ, van Ballegooijen M, Wauben B, Looman CW, Habbema JD, Essink-Bot ML. Having a Pap smear, quality of life before and after cervical screening: a questionnaire study. BJOG. 2012; 119:936–944. [PubMed: 22568482]
6. Arbyn M, Kyrgiou M, Simoens C, Raifu AO, Koliopoulos G, Martin-Hirsch P, Prendiville W, Paraskevaidis E. Perinatal mortality and other severe adverse pregnancy outcomes associated with treatment of cervical intraepithelial neoplasia: meta-analysis. BMJ. 2008; 337:a1284. [PubMed: 18801868]
7. 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; 367:489–498. [PubMed: 16473126]
8. Heleno KJ, Thomsen MF, Rodrigues DS, Jorgerson B, Brodersen J. Quantification of harms in cancer clinical trials: literature review. BMJ. 2013; 347:f5334. [PubMed: 24041703]
9. Saini P, Loke YK, Gamble C, Altman DG, Williamson PR, Kirkham JJ. Selective reporting bias of harms outcomes within studies: findings from a cohort of systematic reviews. BMJ. 2014; 349:g6501. [PubMed: 25416499]
10. Zorzela L, Golder S, Liu Y, Pilkington K, Hartling L, Joffe A, Loke Y, Vohra S. Quality of reporting in systematic reviews of adverse events: systematic review. BMJ. 2014; 348:f7668. [PubMed: 24401468]
11. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. BMJ. 2015; 350:g7773. [PubMed: 25569206]
12. Hoffmann TC, Del Mar C. Patients expectations of the benefits and harms of treatments, screening, and tests. A systematic review. JAMA Intern Med. 2015; 175:274–286. [PubMed: 25531451]

13. Caverly TJ, Combs BP, Moriates C, Shab N, Grady D. Too much medicine happens too often. JAMA Intern Med. 2014; 174:8–9. [PubMed: 24080955]
14. 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; 90:5–37. [PubMed: 22428690]
15. Buist, DS., Tuzzio, L., Greenlee, RT., Field, T., Williams, A., Weinmann, S., Stout, NK., Kamineni, A., Doubeni, CA. Clinical guideline development, implementation and dissemination: connection points for researchers and health care delivery systems for comparative effectiveness research. 2012. 18th Annual HMO Research Network Conference; 2012 Apr 30; Seattle, WA. 2012.
16. Weinmann S, Williams AE, Kamineni A, Buist DSM, Masterson EE, Stout NK, Stark A, Ross TR, Owens CL, Field TS, Doubeni CA. Cervical Cancer Screening and Follow Up in4 Geographically Diverse US Health Care Systems, 1988 Through 2007. Cancer. 2015; 121:2976–2983. [PubMed: 25989253]
17. Casparie M, Tiebosch AT, Burger G, Blauwgeers H, van de Pol A, van Krieken JH, Meijer GA. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cell Oncol. 2007; 29:19–24. [PubMed: 17429138]
18. 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; 57:388–393. [PubMed: 15047743]
19. 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. 2006; 120:806–812.
20. Volante R, Giubilato P, Ronco G. Quality of colposcopy and treatment: data from the national survey of Italian organised cervical screening programmes. 2008 activity. Epidemiol Prev. 2010; 34:73–80.
21. Weinmann S, Naleway A, Swamy G, Krishnarajah G, Arondekar B, Fernandez J, Myers E. Pregnancy outcomes after treatment for cervical cancer precursor lesions. Submitted for publication. 2016
22. Simoens C, Goffin F, Simon P, Barlow P, Antoine J, Foidart JM, Arbyn M. Adverse obstetrical outcomes after treatment of precancerous cervical lesions: a Belgian multicentre study. BJOG. 2012; 119:1247–1255. [PubMed: 22804838]
23. Ang C, Mukhopadhyay A, Burnley C, Faulkner K, Cross P, Martin-Hirsch P, Naik R. Histological recurrence and depth of loop treatment of the cervix in women of reproductive age: incomplete excision versus adverse pregnancy outcome. BJOG. 2012; 21:776–782.
24. Werner CL, Lo JY, Heffernan T, Griffith WF, McIntire DD, Leveno KJ. Loop electrosurgical excision procedure and risk of preterm birth. Obstet Gynecol. 2010; 115:605–608. [PubMed: 20177293]
25. United States Census Bureau. 2012. www.census.gov
26. Martin JA, Hamilton BE, Sutton PD, Ventura SJ, Mathews TJ, Kirmeyer S, Osterman MJK. Births: final data for 2007. Nat Vital Statist Rep. 2010; 58:nr24.
27. Martin JA, Osterman MJK, Sutton PD. Are Preterm Births on the Decline in the United States? Recent Data From the National Vital Statistics System. NCHS. 2010 Data Brief nr. 39.
28. Castle PE, Schiffman M, Wheeler CM, Solomon D. Evidence for frequent regression of cervical intraepithelial neoplasia-grade 2. Obstet Gynecol. 2009; 113:18–25. [PubMed: 19104355]
29. McCredie MR, Sharples KJ, Paul C, Baranyai J, Medley G, Jones RW, Skegg DC. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. Lancet Oncol. 2008; 9:425–434. [PubMed: 18407790]
30. Insinga RP, Glass AG, Rush BB. Pap screening in a U.S. health plan. Cancer Epidemiol Biomarkers Prev. 2004; 13:355–360. [PubMed: 15006908]
31. Howard M, Agarwal G, Lytwyn A. Accuracy of self-reports of Pap and mammography screening compared to medical record. Cancer Causes&Control. 2009; 20:1–13. [PubMed: 18802779]
32. Kaiser Permanente National Cervical Cancer Screening Guideline Development Team. Cervical cancer screening guideline. Oakland: Kaiser Permanente Care Management Institute; 2006 Oct.

33. Saslow D, Runowicz CD, Solomon D, Moscicki A, Smith RA, Eyre HJ, Cohen C. American Cancer Society Guideline for the Early Detection of Cervical Neoplasia and Cancer. CA Cancer J Clin. 2002; 52:342–362. [PubMed: 12469763]
34. ACOG Committee on Practice Bulletins. ACOG Practice Bulletin: clinical management guidelines for obstetrician-gynecologists. Number 45, August 2003. Cervical cytology screening (replaces committee opinion 152, March 1995). Obstet Gynecol. 2003; 102:417–427. [PubMed: 12907124]
35. U.S. Preventive Services Task Force. [accessed February 10th, 2011] Screening for Cervical Cancer. 2003. Available at http://www.uspreventiveservicestaskforce.org/uspstf/uspscerv.htm
36. National Center for Health Statistics. National Health Interview Survey Methods. 2011a. 2011. http://www.cdc.gov/nchs/nhis/methods.htm
37. National Center for Health Statistics. National Health Interview Survey, NHIS Questionnaires 1964–1996, 1997-Present. 2011. 2011b; 118:685–692. http://www.cdc.gov/nchs/nhis/quest_doc.htm.
38. Cuzick J, Myers O, Hunt WC, Robertson M, Joste NE, Castle PE, Benard VB, Wheeler CM. A Population-based Evaluation of Cervical Screening in the United States: 2008–2011. Cancer Epidemiol Biomarkers Prev. 2014; 35:765–773.
39. LEBA reports 2013 and 2011, dept. Public Health, ErasmusMC University Medical Center Rotterdam. accessible at http://www.rivm.nl/Documenten_en_publicaties/Algemeen_Actueel/Uitgaven/Preventie_Ziekte_Zorg/baarmoederhalskankerscreening/Landelijke_Evaluatie_van_het_Bevolkingsonderzoek_Baarmoederhalskanker_LEBA
40. Rozemeijer K, Penning C, Siebers AG, Naber SK, Matthijsse M, 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. 2016; 27:15–25. [PubMed: 26458884]
41. table on use of Pap smears in Health, United States. National Center for Health Statistics; 2014. www.cdc.gov/nchs/hus/contents2014.htm#077
42. Sabatino SA, White MC, Thompson TD, Klabunde CN. Cancer screening test use – United States, 2013. MMWR. 2015; 64:464–468. [PubMed: 25950253]
43. Moyer VA. on behalf of the U.S. Preventive Services Task Force. Screening for Cervical Cancer: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med. 2012; 156:880–891. [PubMed: 22711081]
44. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, Garcia FAR, Moriarty AT, Waxman AG, Wilbur DC, Wentzensen N, Downs LS, Spitzer M, Moscicki A, Franco EL, Stoler MH, Schiffman M, Castle PE, Myers ER. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. CA Cancer J Clin. 2012; 62:147–172. [PubMed: 22422631]
45. National Health Insurance Council. Principles for the restructuring of the organised cervical cancer screening programme-Final decision (in Dutch). Amstelveen, The Netherlands: Ziekenfondsraad; 1993.
46. Albrechtsen S, Rasmussen S, Thoresen S, Irgens LM, Iversen OE. Pregnancy outcome in women before and after cervical conisation: population based cohort study. BMJ. 2008; 337:a13439.
47. Strander B, Hallgren J, Sparen P. Effect of ageing on cervical or vaginal cancer after Cervical Intraepithelial Neoplasia grade 3 (CIN3) – a population-based cohort study of long term incidence and mortality. BMJ. 2014:16. in-press.
48. Briët MC, Berger TH, van Ballegooijen M, Boon ME, Rebolj M. Effects of streamlining cervical cancer screening the Dutch way: consequences of changes in the Dutch KOPAC-based follow-up protocol and consensus-based limitation of equivocal cytology. Acta Cytol. 2010 Nov-Dec; 54:1095–1100. [PubMed: 21428155]
49. Farquhar CM, Steiner CA. Hysterectomy rates in the United States 1990–1997. Obstet Gynecol. 2002; 99:229–234. [PubMed: 11814502]
50. Hanstede MM, Burger MJ, Timmermans A, Burger MP. Regional and temporal variation in hysterectomy rates and surgical routes for benign diseases in the Netherlands. Acta Obstet Gynecol Scand. 2012; 91:220–225. [PubMed: 22043840]

51. Goodman A. HPV testing as a screen for cervical cancer. BMJ. 2015; 350:h2372. [PubMed: 26126623]

52. Malila N, Leinonen M, Kotaniemi-Talonen L, Laurila P, Tarkkanen J, Hakama M. The HPV test has similar sensitivity but more overdiagnosis than the Pap test- A randomised health services study on cervical cancer screening in Finland. Int J Cancer. 2013; 132:2141–2147. [PubMed: 22987601]

53. Huh WK, Ault KA, Chelmow D, Davey DD, Goulart RA, Garcia FA, Kinney WK, Massad LS, Mayeaux EJ, Saslow D, Schiffman M, Wentzensen N, Lawson HW, Einstein MH. Use of primary high-risk human papillomavirus testing for cervical cancer screening: interim clinical guidance. Obstet Gynecol. 2015; 125:300–307.

54. Health Council of the Netherlands. Screening for cervical cancer. Report 2011/07. The Hague, The Netherlands: 2011.

55. Cuzick J, Myers O, Hunt WC, Saslow D, Castle PE, Kinney W, Waxman A, Robertson M, Wheeler CM. Human Papillomavirus testing 2007–2012: co-testing and triage utilization and impact on subsequent clinical management. Int J Cancer. 2015; 136:2854–2863. [PubMed: 25447979]

NOVELTY AND IMPACT

This is the first study on population-wide harmful health effects of cervical cancer screening, based on empirical data. We show that two countries with different recommendations and practice of screening, and with comparable beneficial effects of screening, experience markedly different harmful effects, emphasizing the importance of carefully considering harms alongside benefits. We recommend that adverse health effects of screening should be measured routinely and be included in monitoring and decision making.

Table 1

Frequency of selected health problems in cervical cancer screening that are associated with a Pap test, an abnormal test result, a punch biopsy, and treatment

Event	Health Problem	Proportion of women reporting a health problem		
		Total	Light*	Moderate+**
Pap test (5)	Symptoms*** at least 2–7 days	13%		
Abnormal test (3)	Anxiety at least 12 weeks	35%		
Punch biopsy (4)	Pain	53%	25%	28%
	Bleeding	79%	58%	21%
	Discharge	46%	32%	14%
Treatment**** (4)	Pain	67%	34%	33%
	Bleeding	77%	24%	53%
	Discharge	63%	21%	42%

* Very light or light (4).

** Moderate, severe, or very severe (4).

*** Lower abdominal pain, urinary discomfort, feeling sick, feeling dizzy, and/or painful sexual activity.

**** Loop electrosurgical excision (LEEP or LLETZ).

Table 2

Age-specific rates of Pap tests, abnormal test results, punch biopsies, and treatments in 2007 for the United States (US) and the Netherlands (NL)

Age group (years)	Per 1,000 women							
	Pap test		Abnormal* test result		Punch biopsy		Treatment**	
	US	NL	US	NL	US	NL	US	NL
21–25	455	30	58	4	31.6	1.1	6.2	0.8
26–30	533	145	43	11	25.5	2.2	6.8	3.1
31–35	479	225	29	16	19.2	4.8	4.5	3.7
36–40	421	205	22	12	17.1	4.0	3.2	3.4
41–45	371	208	26	12	14.6	4.9	2.2	2.6
46–50	349	207	23	10	16.7	5.8	1.9	1.8
51–55	321	162	14	6	13.5	5.1	1.4	1.0
56–60	299	179	9	4	9.8	4.1	1.1	0.7
61–65	301	51	7	2	8.3	2.9	0.6	0.4

Note: see Methods section for background and references.

* Cytology result of ASCUS or higher

** Including excisional and ablative procedures

Table 3

Cervical cancer-screening-related events in the United States (US) and the Netherlands (NL) screening practice in 2007.

	Per 1000 women			In US population*	
Event	**US**	**NL**	**US:NL ratio**	**US practice**	**NL practice**
Pap test	394	164	2.4	35,800,000	14,900,000
Abnormal test	25	9	2.8	2,270,000	820,000
Punch biopsy	16.9	4.3	3.9	1,540,000	390,000
Treatment	3.0	1.8	1.7	273,000	164,000

Note: US and NL rates were obtained by standardizing the age-specific data from Table 2 to the 2007 US female population 21–65 years (27).

*Total of 90,905,000 women age 21–65 years (27).

Table 4

Health problems due to cervical cancer screening-related events

Harm	Per 1000 women			In US population***	
	US	NL	US:NL ratio	US practice	NL practice
Pap test:					
2–7 days symptoms *	51	21	2.4	4,600,000	1,900,000
Abnormal test result:	9	3	2.8	800,000	290,000
Anxiety 12 weeks					
Punch biopsy:					
Light ** symptoms	19	5	3.9	1,700,000	400,000
Moderate or strong ** symptoms	11	3	3.9	1,000,000	280,000
Treatment:					
Light ** symptoms	2.4	1.4	1.7	218,000	127,000
Moderate or strong symptoms	3.8	2.3	1.7	345,000	205,000

Note: Number of events were obtained by applying the data of Table 1 to the United States (US) and the Netherlands (NL) screening practice in 2007. Data are standardized to the 2007 US female population 21–65 years (27).

* Lower abdominal pain, urinary discomfort, feeling sick, feeling dizzy, and/or painful sexual activity.

** "light" refers to (very) light pain, bleeding or discharge, and "moderate+" to moderate or (very) severe pain, bleeding or discharge, see (4).

*** 90,905,000 women aged 21–65 years in 2007 (27).

Table 5

Number of preterm deliveries attributable to deep excisional treatment of cervical pre-cancer (CIN) in screening practice in the United States (US) and the Netherlands (NL).

	Relative risk of preterm delivery after deep excisional treatment	Deep excisional treatment* (%)	Number of preterm deliveries		US:NL ratio
			US	NL	
Baseline assumptions	1.75	23%	5,300	2,100	2.6
Lower risk scenario	1.31	18.4%	1,800	700	2.6
Higher risk scenario	2.19	50%	17,600	7,000	2.5

Note: The lower/higher risk scenario assumes a lower/higher relative risk of preterm delivery after deep excisional treatment and a lower/higher percentage of deep excisional treatments, compared to baseline assumptions. See the Methods section. Numbers are standardized to the 2007 US female population (27).

*Exceeding 10 mm in depth