**Quality of life assumptions determine which cervical cancer screening strategies are cost-effective**

Inge M.C.M. de Kok, PhD [a]*, i.dekok@erasmusmc.nl
Ida J. Korfage, PhD [a]*, i.korfage@erasmusmc.nl
Wilbert B. van den Hout, PhD [b], w.b.van_den_hout@lumc.nl
Theo J.M. Helmerhorst, MD PhD [c], t.helmerhorst@erasmusmc.nl
J. Dik F. Habbema, PhD [a], j.d.f.habbema@erasmusmc.nl
Marie-Louise Essink-Bot, MD PhD [d†]
Marjolein van Ballegooijen, MD PhD [a], m.vanballegooijen@erasmusmc.nl

[a] Erasmus University Medical Center, Department of Public Health, P.O. Box 2040, 3000 CA Rotterdam, the Netherlands
[b] Leiden University Medical Center, Department of Medical Decision Making, P.O. Box 9600, 2300 RC Leiden, the Netherlands
[c] Erasmus University Medical Center, Department of Gynaecology and Obstetrics, P.O. Box 2040, 3000 CA Rotterdam, the Netherlands
[d] University of Amsterdam, Academic Medical Center, Department of Public Health, PO Box 22660, 1100 DD Amsterdam, the Netherlands.
* This concerns a shared first authorship, because both authors contributed equally

**Novelty and Impact**
This is the first study that empirically obtained utility scores for quality of life associated with cervical cancer screening and treatment in women in all relevant health states, ranging from being invited for screening to being disease-free after primary treatment of cervical cancer or having advanced cancer. Furthermore, we show that quality of life assumptions determine the cost-effectiveness of cervical cancer screening and that the measure used to empirically assess utilities is crucial for cost-effectiveness conclusions.

**Funding**: The Dutch Cancer Society (EMCR 2003–2775)
**Corresponding author**:
Inge M.C.M. de Kok
Erasmus MC, Dept. of Public Health
P.O. Box 2040, 3000 CA Rotterdam, the Netherlands
Tel. no.: +31 10 704 3721
Fax no.: +31 10 703 8475
E-mail: i.dekok@erasmusmc.nl

**ABSTRACT**

Quality adjusted life years are used in cost-effectiveness analyses (CEAs). To calculate QALYs, a 'utility' (0-1) is used for each health state induced or prevented by the intervention. We aimed to estimate the impact of quality-of-life (QoL) assumptions (utilities and durations of health states) on CEAs of cervical cancer screening. To do so, twelve alternative sets of utility assumptions were retrieved from published cervical cancer screening CEAs. Two additional sets were based on empirical QoL data that were integrally obtained through two different measures (SF-6D and EQ-5D) from eight groups of women (total n= 3,087), from invitation for screening to diagnosis with cervical cancer. Per utility set we calculated the number of quality-adjusted days lost (QADL) for each relevant health state in cervical cancer screening, by multiplying the study-specific assumed disutilities (i.e. 1-utility) with study-specific durations of the loss in QoL, resulting in 14 'QADL-sets'. With microsimulation model MISCAN we calculated cost-effectiveness of 342 alternative screening programs (varying in primary screening test [Human Papillomavirus (HPV) versus cytology], starting ages, and screening interval) for each of the 14 QADL-sets. Utilities used in CEAs appeared to differ largely. We found that ten QADL-sets from the literature resulted in HPV and two in cytology as preferred primary test. The SF-6D empirical QADL-set resulted in cytology and the EQ-5D one in HPV as preferred primary test. In conclusion, assumed utilities and health state durations determine cost-effectiveness of cervical cancer screening. Also, the measure used to empirically assess utilities can be crucial for CEA conclusions.

## INTRODUCTION

Cervical cancer screening, still widely based on cytological Pap test, decreases mortality and incidence from this cancer.[1] However, changes are ongoing in cervical cancer prevention, such as the introduction of human papilloma virus (HPV) screening and HPV vaccination programs. An important criterion for appraising the validity of a new screening programme is that its benefit (i.e. preventing clinical cancer and cancer death) should outweigh its physical, psychological, and societal harm (caused by the test and screening-related diagnostic procedures and treatment).[2] This is assessed in a cost-effectiveness analysis (CEA). CEAs can use 'quality adjusted life years' (QALYs) as the summary effect measure. The QALY is a generic effect variable of disease burden, including both the quality and the length of life lived.[3] To calculate QALYs, a quality-adjustment weight or 'utility' is applied to each relevant health state. A utility is a number anchored at 0 (dead) and 1 (perfect health), indicating the preferability of that health state in terms of quality of life. Because of their crucial impact on CEAs, and thus on health care decision making about the introduction of e.g. screening programs or the choice about specific screening tests, reliable estimates of the utility scores of all health states induced and prevented by screening are important.
For cervical cancer screening, after early detection via cytological abnormalities, women with cervical intraepithelial neoplasia (CIN) can be treated so that potential development of these abnormalities into cancer is prevented. The downsides of cervical cancer screening are lack of specificity, overdiagnosis and overtreatment, since most of the abnormalities found in the screening and follow up will never develop into clinical cervical cancer. Notably, it was shown that in the Netherlands one prevented cervical cancer death due to screening entails an estimated 2,097 women being screened, 64 women being sent to triage, and 42 women being referred for colposcopy, with punch biopsy and eventually treatment of mostly preinvasive neoplastic conditions.[4] The numbers of women needed to be screened or sent to triage or colposcopy to prevent one cervical cancer death increase by implementation of a more sensitive (but less specific) screening test. Recently it was shown, that the implementation of primary HPV screening in the Netherlands may lead to a threefold increase in the number of (false-positive) referrals to the gynaecologist.[5] Women

screened and referred to a gynaecologist may experience distress and anxiety.[6-8] Per woman this may have a limited effect, but considering the large numbers of women screened and referred, small quality of life losses for these harms might still result in relevant effects at population level. We currently lack empirically derived utility estimates for all relevant health states in cervical cancer screening. Currently, available utility estimates in the literature are conflicting, which results in parameter uncertainty.[9]

The present study first aims to estimate the impact of the variation in currently used sets of utility assumptions on the estimated cost-effectiveness of cervical cancer screening programs. To do so, we will review utility assumptions in the CEA literature. Next, we will use microsimulation modelling to estimate the impact of the differences in utility assumptions on the cost-effectiveness of cervical cancer screening (in a population unvaccinated against HPV) and on the preferred screening strategies.
Second, we derived two sets of utilities based on a state of the art empirical study. In this study, we have performed an integrated questionnaire study addressing all states related to the cervical cancer-screening programme in a generic and standardized way, resulting in two types of utilities for all relevant health states. Selected parts of this empirical study have been published earlier [6, 10-12] The study results will be presented here as a whole. Next, we will evaluate how use of our empirical utility measurements affects the microsimulation cost-effectiveness results, and how this compares to using utility estimates applied in the cervical cancer screening CEA literature.

## METHODS

### Part 1: Literature review

To collect the utilities that have been used so far, we reviewed the literature for CEAs of cervical cancer screening published between 2003 and 2015. We used the following search terms in Pubmed: cost-benefit analysis; cost-effectiveness analysis; uterine cervical neoplasms; cervical cancer; screening; early detection of cancer; quality-adjusted life years; QALY. We selected papers that examined cost-effectiveness of screening from Europe, North-America and Australia that are published in English. We identified 19 studies that used QALYs.[4, 13-30] Two pairs of studies used the same set of utility assumptions (i.e. same model and research group), and each of these pairs were combined ([4]and [29]; [17] and [18]). One study was excluded since it also used a common set of utility assumptions that was already included.[23] One study was excluded since the utilities used were not reported.[25] The resulting 15 studies defined different health states in cervical cancer screening. We first summarized the published health states into 13 health states (primary screening; positive primary screentest; false positive test; referral for triage test; referral for colposcopy; false positive referral; CIN1; CIN2; CIN3; FIGO 1; FIGO2+; Survivors; Palliative phase). We then for each study assessed the number of days lost due to diminished QoL ('quality-adjusted days lost' (QADL)) for these different health states, by multiplying the study-specific assumed disutilities (i.e. 1-utility) with the study-specific mean durations of the loss (or gain) in QoL. Three studies[20, 28, 30] were excluded because of absence of data on the durations. For five studies[4, 16, 24, 26, 27] we had to make assumptions (based on limited available information) on durations (see footnotes Table 2). So, the literature review resulted in 12 different 'QADL-sets'. Appendix Table 3 shows the assumed utilities in the twelve included published CEAs per health stage, as well as the source of each of the assumed utilities.

### Part 2: Questionnaire study

We empirically obtained utility scores, indicating quality of life, for 7 study groups of women: invited for cervical cancer screening (n=1,023 respondents, of whom 905 participated in screening, response 60%); having borderline or mildly dyskaryotic (BMD) pap

test results (n=270; response 49%); referred for colposcopy (n=132); treated for a precursor of cervical cancer (n=81; response 49%); diagnosed with cervical cancer (n=77); disease-free after primary treatment of cervical cancer (n=285; response 69%), and a reference group (n=835, response 46%, Appendix Table 2). The methods and results have been published in detail for the EQ-5D scores of the screening participants,[10] the group with borderline or mildly dyskaryotic pap test results,[6] women referred for colposcopy,[12] and women who are disease-free after treatment.[11] Data were collected by self-administered questionnaires, containing both the EQ-5D and SF-6D questions. At the time of data-collection HPV screening had not been introduced yet. A description of the data-collection procedures and the background characteristics of the participants is included in the Appendix (Appendix Tables 1 and 2). No women with advanced cervical cancer (i.e. palliative phase) could be included due to logistic reasons; for this group we used utility estimates of women with breast or lung cancer.[31, 32] The ethics review committee of Erasmus University Medical Center Rotterdam approved the research protocol.

*Utility measures*

A quality-adjustment weight or 'utility' is a number anchored at 0 and 1, with "perfect health" carrying a weight of 1 and dead carrying a weight of 0. We used both the SF-6D and the EQ-5D. The SF-6D, based on a subset of SF-12 responses,[33] is composed of six dimensions: physical functioning, role limitations, social functioning, bodily pain, mental health, and vitality. It was derived from a valuation study among a representative sample of the general public in the UK, using the standard gamble valuation technique.[33] The EQ-5D classification consists of five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.[34] Response categories indicate 1) no problems; 2) some problems; or 3) serious problems. Responses were linked to utility scores as obtained in the UK using time trade-off methodology.[34]Information on age, marital status, education, and profession of respondents was obtained through the questionnaires.

*Statistical analyses*

Paired t-tests were used to assess the statistical significance of differences between SF-6D and EQ-5D scores. Pearson product-moment correlations were used to calculate within-person correlations between SF-6D and EQ-5D assessments. All analyses were performed in SPSS, version 20.

*Calculation of quality adjusted days lost (QADL)*

Based on the questionnaire study we calculated the number of QADL for those health states for which we empirically found a significant different utility (measured by the EQ-5D and the SF-6D, separately) compared to the reference population (screening ages in case of screening health states and all ages in case of cancer health states). For these health states we calculated the disutility by subtracting the measured utility (the point estimate) in that specific health state from the utility measured in the general population (screening ages in case of screening stages). In case of 'terminal care' we subtracted the utility in breast cancer patients in the last year of life from the utility in the general population, both measured by the EQ-5D. We then calculated the number of QADL for the different health states, by multiplying the measured disutilities with the durations of the loss in QoL per state. The durations where based on 1) the Dutch screening guideline in case of a BMD result (i.e. 15 months), 2) measured duration of significant QoL loss in case of longitudinal measurements for referral to the gynaecologist and cancer health states, 3) 10 years in case of cancer survivor based on the definition used in the questionnaire study, 4) last year of life in case of terminal care (i.e. 1 year).

**Part 3: Microsimulation modelling**

We used the microsimulation screening analysis (MISCAN) model[29] to estimate the costs and QALYs gained for alternative cervical cancer screening programs, in a Dutch (unvaccinated against HPV) population. The model and the inputs used are described in the Appendix, as well as in an earlier publication.[29] Also, the calibration and validation of the model are described in an online technical appendix [http://hdl.handle.net/1765/31582].

We simulated 1,000,000 unvaccinated women born between 1939 and 1992. The alternative strategies included primary cytology and primary HPV (followed by cytology triage) screening, with respective triage strategies that according to the MISCAN model used were cost-effective.[29] We considered all screening policies with starting ages of 25, 27, 30 or 32 years that comprise at least three and at most ten screenings in a woman's lifetime. Policies had an interval of at least three years and at most ten years, policies that include screenings over the age of 70 years were excluded. This resulted in 342 simulated screening policies. Based on monitoring (program), trial and administrative (program and hospital) data we made assumptions on screening attendance, test characteristics and costs (see Appendix Table 4). The sensitivity of the HPV test (the probability of a positive test result if an HPV infection is present) was estimated at 94%, and the sensitivity of cytology was assumed to be 40% for CIN 1, 50% for CIN 2, and 75% for CIN 3 and invasive cervical cancer (see Appendix Table 4). The specificity of the HPV test (probability of a negative test for women without high-risk HPV infections) is assumed to be 100%, and the specificity of cytology (probability of a negative test for women without CIN or cancer) is estimated to be 98.5%. Lack of specificity in case of HPV screening was accounted for by the inclusion of fast-clearing HPV infections.

The lifelong costs and effects of each simulated screening program were counted for the period from 2011 onwards (until all women have died), and discounted at an annual rate of 3% towards the year 2011. For each woman, we calculated the number of QALYs as the weighted sum of the number of years spent in each of the health states, using the state specific utility weights. The total effectiveness of screening (QALYs gained) is determined as the difference in the number of QALYs between the situation with screening and the situation without screening. We used a similar approach to determine the net costs of screening.

Programs that were more costly and less effective than other programs were ruled out as non-efficient (by both simple and extended dominance). The remaining programs constitute the frontier of efficient screening programs, see the lines in Figure 2. For each 'QADL-set' we determined which screening programs were located on the cost-effectiveness frontier.

Based on the incremental cost-effectiveness ratios (ICERs) between these programs we determined which screening program was preferred considering cost-effectiveness thresholds of €20,000 and €50,000 per QALY gained.

## RESULTS

### Part 1: Literature review

Both for screening and treatment phases we found a large variation in number of QADL between the published studies (Table 1). For example, the number of QADL for a cervical intraepithelial neoplasia (CIN) 2 or CIN 3 lesions varied from 0 to 47.5. For a FIGO (International Federation of Gynecology and Obstetrics) stage 2+ cancer, the number of QADL varied from 51 to 1460, which is an 80-fold difference. This variation is the result of differences in utility values (see Appendix Table 3), as well as differences in durations. We found that all CEAs ultimately refer to 12 different original utility studies, published between 1991 and 2012 (see Appendix Table 4).

### Part 2: Questionnaire study

*SF-6D utility scores*

The mean utility score was 0.83 (confidence interval (CI): 0.82-0.84) in the general adult female population and 0.84 (CI: 0.83-0.85) in the on average somewhat younger women in screening ages (30-60 years), see Table 2 and Figure 1A. Screen participants' mean scores were slightly higher: 0.85 before and after screening and 0.86 after receiving the test results. The mean utility of women who were recommended to have a repeat Pap test because of borderline or mild dyskaryotic test results were lower than those of the reference population and the screening invitees. Before onset of treatment, both FIGO groups reported similar utility scores. The FIGO 1 group (n=47) improved steadily from 0.73 at baseline to 0.81 one year later, while the FIGO 2+ group (n=16) reported 0.71 at baseline, 0.63 at three months follow-up, and 0.71 and 0.75 at six and twelve months follow-up. The mean utility of tumorfree cancer survivors was 0.80, and thus comparable to that of women with cervical cancer FIGO 1 at 1-year post –diagnosis.

Overall, SF-6D utility scores were best (0.86, CI: 0.85-0.87) in screen participants after receipt of Pap test results and worst in women with cervical cancer (FIGO 2+) at three months after diagnosis (0.63, CI:0.57-0.70).

*Comparison of SF-6D and EQ-5D utility scores*

SF-6D and EQ-5D utility scores were always significantly correlated and the patterns of SF-6D and EQ-5D utility scores were roughly similar (Figures 1.A and 1.B). On average, SF-6D utility scores were lower than EQ-5D utility scores, but in more serious health states, such as diagnosis with cervical cancer, SF-6D and EQ-5D scores no longer significantly differed.

*Number of QADL*

The number of QADL based on our questionnaire study differ substantially from those used in the CEAs obtained from the literature. The QADL, calculated with the SF-6D, due to having gone through a triage episode is 14 days. With the EQ-5D, no significant different utility was measured for women in triage compared to the references population (screening ages) (Table 2), so no QADL were measured in the screening stages. The QADL (SF-6D) due to a FIGO2+ cancer diagnosis is 161 (51+110) days for survivors and 234 (51+183) days for those who die from it (Table 2). In case of the EQ-5D, these figures are 166 (21+146) and 204 (21+183), respectively.

**Part 3: Microsimulation modelling**

Per 'QADL-set' different screening programs are presented on the cost-efficient frontier (Figure 2). Table 3 shows per QADL-set the preferred screening strategy at cost-effectiveness thresholds of €20,000 and €50,000 per QALY gained, according to the MISCAN model. At a willingness to pay of €20,000 per QALY, eleven QADL-sets (ten from the literature and the EQ-5D empirical data) preferred HPV screening, all starting at age 30 or 32, performing 3 or 4 tests with a 5 or 6 year interval. Our empirical SF-6D data and two other sets from the literature preferred cytology screening, starting at age 30 or 32, performing 4 or 5 tests at a 4, 5 or 6 year interval.

At €50,000 per QALY, the optimal strategies tended to have lower starting ages (age 27 to 30), with more tests (7 to 10) and shorter intervals (4 to 6 years). However, the preferred screening tests for each QADL-set did not change. The QADL-set 'Kitchener (Simonella)'[22]

resulted in less frequent screening, since this set showed a considerable loss in QoL related to attending screening.

## DISCUSSION

We showed that differences in quality of life assumptions as they are found in published cervical cancer screening CEAs and as measured in an integral questionnaire study using utility measures, lead to different conclusions about cost-effectiveness of strategies. Differences in quality of life assumptions result in different number of QALYs gained for the same screening strategy and, as a result, different ICERs between screening strategies. In general, it largely depends on utility losses assumed for positive screening results including the associated follow up, which screening modality is preferred. Primary HPV screening, with relatively more positive screen test results than cytology screening, is preferred in case of lower utility loss assumptions for the screening phase (used in ten studies), and cytology in case of higher utility losses (used in two studies). For example, the assumed number of QADL due to referral in the study of Karnon et al. [24] is significantly higher than in the other scenarios, and therefore having a (false-positive/clinically not relevant) referral for colposcopy has more impact in this scenario than in the other scenarios. In this example, one could argue that the fact that a large majority of studies used lower utility losses in the screening phases, provides some strong certainty that HPV screening is to be preferred. In a situation with high uncertainty, however, such counting of studies not necessarily leads to the correct conclusions. Clearly, focussing further screening related QoL research on follow up after positive screen results is important.

The use of the presented empirical utility data measured by the SF-6D resulted in the preference for primary cytology screening, whereas in most utility sets used in the CEA literature, primary HPV screening was preferred. This was caused by the fact that relatively larger quality of life losses were found with the SF-6D for triage and referral for colposcopy, which is a health state that will occur more frequently with primary HPV screening than with cytology.[35] As a result, more quality of life will be lost due to the screening itself with primary HPV screening compared to cytology and therefore primary cytology becomes the preferred strategy. Our finding that quality of life assumptions influence the results of CEA of cervical cancer prevention is in line with previous studies.[22, 36, 37]

We found large differences in utility estimates and durations used in CEAs. It appeared that none of these CEAs was based on utility estimates derived from empirical data of cervical cancer screenees or actual patients who were faced with the studied diagnosis in real life, but on data collected in other (patient) groups (i.e. patients with other conditions than (precursors of) cervical cancer). However, patients (that experience the studied disease) often report better quality of life than healthy people who are asked to imagine the patient's circumstances; the so-called disability paradox.[38] Causes of this paradox include loss aversion, focusing illusion, and underestimation of their own adaptation in healthy people and adaptation processes in patients.[39-41] These phenomena indicate the importance of empirically derived utility estimates.

In this first study to empirically assess the QoL effects of all relevant stages in population-based organized cervical cancer screening, we observed, per woman involved, only limited QoL loss due to screening as long as there was no diagnosis of cervical cancer. This held for women invited for screening, referred for colposcopy, and treated for a CIN lesion. A study on the impact of abnormal cervical smear results found that at baseline, shortly after the news, women who had received abnormal results reported significantly worse overall quality of life (EQ-5D and SF-6D) than women with normal results. At 12 weeks follow-up, only SF-6D results still significantly differed between groups. The QALYs lost during the 16 weeks after being informed of an abnormal smear result were estimated to be 0.007–0.009, which is equivalent to 2.4 to 3.2 days of healthy life lost.[42] Studies that focused on condition-specific effects like anxiety and worry did report negative effects of screen-detected non-cancer abnormalities.[6, 12, 42-44] Our findings may thus be related to the use of generic measures, which are less sensitive to small, condition-specific changes in health or quality of life. The patterns of the two generic measures used (EQ-5D and SF-6D) were roughly similar, although EQ-5D utility scores were often significantly higher than SF-6D scores. This is in line with the literature.[45] With different items and scoring mechanisms, the EQ-5D and the SF-6D were not expected to generate completely similar utilities. However, the EQ-5D and the SF-6D each resulted in a different preferred screening strategy, which clearly shows the dependence of CEA results on the choice of generic measure.

*Limitations and strengths*

With regard to the questionnaire study, we acknowledge that some of the (sub) groups were of limited size, that response rates are unavailable for the data collection among women referred for colposcopy and women diagnosed with cervical cancer, and that data considering the final stages of life were collected among patients with primary breast or lung cancer, and not among patients with cervical cancer. Also, we reported SF-6D as based on responses to the SF-12 items although SF-6D is better able to discriminate between conditions if based on responses to the SF-36 items.[46] Furthermore, screening has some rare undesired effects, like pregnancy complications after treatment for CIN, which were not captured in this study, but are important to be mentioned.[47] Also, we acknowledge that we assessed utilities in the Netherlands, and that we cannot determine if and to which extent this has influenced our results. However, we used widely accepted measures, that are each available in over 170 languages and that have been applied in numerous studies. We conclude that we have not solved all issues of how to determine the impact of cervical cancer on quality of life. However, at least we now have an idea about extent of the implications of screening and treatment for quality of life. We provided a set of empirical data and we do think that such a set is vital for the validity of CEAs. This data was collected in one country and we recommend more international data-collection to study external validity. Since we do not know what the results will be in another setting, we recommend to use the utilities as collected in the current study in modelling exercises, at least as an extra dataset, to enable sensitivity analyses. Still, this is the first study that made the extensive effort of comprising all health states induced or prevented by screening on cervical cancer. Data were collected in 7 cohorts, of which three were followed longitudinally. An additional strength is that the available clinical information enabled subdividing respondent groups according to FIGO stage. The design of our study has proven its feasibility and value, and can serve as a basis for measurements in other populations and situations.

This is also the first study focusing on the impact of utility assumptions on the recommended cervical cancer screening strategy. We, however, only used one model (for

the Dutch situation) to estimate the impact. Also, in absence of utility measurements for being HPV positive, we had to make the assumption that the effect on quality of life is similar for having a positive cytology test compared to having a positive HPV test. If higher quality of life losses will be found for being HPV positive than cytology positive, HPV screening might be less cost-effective than indicated by our calculations. Furthermore, in case of the HPV test, we assumed similar sensitivity for all disease stages with an HPV infection. Recent studies indicated lower sensitivity of HPV screening in women without any neoplasia.[48] We, however, calibrated our model to the observed HPV positivity rate (measured with a clinical validated HPV test) in the Dutch population, which determines the effect of HPV screening in the population. If we assume a lower sensitivity, we would have to increase the number of infections in the model, to reproduce that HPV positivity rate. Finally, the literature included in our analyses is not an exhaustive summary. We, for example, did not include HPV vaccination CEAs, even if vaccination was applied in a screening setting. Since the information on utility assumptions and the durations of the health states was sometimes lacking in the publications, we had to do some approximations to calculate the QADL. But even if our assumptions differ from the assumptions in the specific publications, our study clearly shows the importance of the utility assumptions for preferred cervical cancer screening strategies. We used the example of cervical cancer screening, but think the central message of our findings is generalizable to other settings, such as other (cancer) screening programs and vaccination programs.

*Implications*

The ongoing changes in cervical cancer prevention, such as the introduction of HPV screening and HPV vaccination programs require new policy, which is nowadays unacceptable without CEAs. The empirical utility scores that now have become available will enable more valid estimates of cost-effectiveness of screening programs and will contribute to a better evidence base for policy recommendations, which is especially useful given limited resources. Our results also showed that measuring utility values for different health states in screening is feasible but challenging, because these often involve small, condition-

specific changes in quality of life. This indicates that in economic evaluations, besides the QALY, other measurements that include the harms and benefits of screening (such as, number needed to screen or number needed to treat) need to be regarded as well.

*Conclusion*

This is the first study that empirically estimated utilities for the majority of the cervical cancer health states. We found that QoL assumptions in decision analyses vary in a range that is crucial for preferences regarding screening strategies. Empirical data show that primary HPV screening might not be the most cost-effective screening strategy, if QALYs are used in this decision. However, the empirical data also show that the measurement of utilities is challenging. This indicates that other measurements that include the harms and benefits of screening needs to be regarded as well.

**Acknowledgements**

We are grateful to the women who completed questionnaires for participating in the study. This work was funded by the Dutch Cancer Society (EMCR 2003–2775).

**Role of the funding source**

The authors were completely independent from funders in conducting and writing this study.

**Conflict of interest statement**

The authors have no conflicts of interest to disclose.

**REFERENCES**


1. IARC. Cervic cancer screening *IARC Handbook of cancer prevention*ed.: IARC Press, 2004.

2. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease] Principios y metodos del examen colectivo para identificar enfermedades. *Bol Oficina Sanit Panam* 1968;**65**: 281-393.

3. Gold M, Siegel JE, Russell L, Weinstein MC. *Cost-Effectiveness in Health and Medicine* ed. Oxford: Oxford University Press, 1996.

4. de Kok IM, van Rosmalen J, Dillner J, Arbyn M, Sasieni P, Iftner T, van Ballegooijen M. 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.

5. Naber SK, Matthijsse SM, Jansen EEL, de Kok IMCM, de Koning HJ, van Ballegooijen M. Effecten en kosten van het vernieuwde bevolkingsonderzoek naar baarmoederhalskanker in Nederland naar aanleiding van recente ontwikkelingen Rotterdam, 2016.

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

7. Tahseen S, Reid PC. Psychological distress associated with colposcopy: patients' perception. *Eur J Obstet Gynecol Reprod Biol* 2008;**139**: 90-4.

8. Valdini A, Esielionis P. Measurement of colposcopy-associated distress using the psychosocial effects of having an abnormal pap smear-questionnaire in a Latina population. *J Low Genit Tract Dis* 2004;**8**: 25-32.

9. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD, Force I-SMGRPT. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--6. *Value Health* 2012;**15**: 835-42.

10. Korfage I, van Ballegooijen M, Wauben B, Looman C, Habbema J, Essink-Bot ML. Having a Pap smear, quality of life before and after cervical screening: a questionnaire study. *BJOG* 2012;**119**: 936-44.

11. Korfage IJ, Essink-Bot ML, Mols F, van de Poll-Franse L, Kruitwagen R, van Ballegooijen M. Health-related quality of life in cervical cancer survivors: a population-based survey. *Int J Radiat Oncol Biol Phys* 2009;**73**: 1501-9.

12. Korfage IJ, Essink-Bot ML, Westenberg SM, Helmerhorst T, Habbema JD, van Ballegooijen M. How distressing is referral to colposcopy in cervical cancer screening?: a prospective quality of life study. *Gynecol Oncol* 2014;**132**: 142-8.

13. Accetta G, Biggeri A, Carreras G, Lippi G, Carozzi FM, Confortini M, Zappa M, Paci E. Is human papillomavirus screening preferable to current policies in vaccinated and unvaccinated women? A cost-effectiveness analysis. *J Med Screen* 2010;**17**: 181-9.

14. Balasubramanian A, Kulasingam SL, Baer A, Hughes JP, Myers ER, Mao C, Kiviat NB, Koutsky LA. Accuracy and cost-effectiveness of cervical cancer screening by high-risk human papillomavirus DNA testing of self-collected vaginal samples. *J Low Genit Tract Dis* 2010;**14**: 185-95.

15. Berkhof J, Coupe VM, Bogaards JA, van Kemenade FJ, Helmerhorst TJ, Snijders PJ, Meijer CJ. The health and economic effects of HPV DNA screening in The Netherlands. *Int J Cancer* 2010;**127**: 2147-58.

16. Chuck A. Cost-effectiveness of 21 alternative cervical cancer screening strategies. *Value Health* 2010;**13**: 169-79.

17. 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;**30**: 1813-22.

18. Coupe VM, de Melker HE, Snijders PJ, Meijer CJ, Berkhof J. How to screen for cervical cancer after HPV16/18 vaccination in The Netherlands. *Vaccine* 2009;**27**: 5111-9.

19. de Bekker-Grob EW, de Kok IM, Bulten J, van Rosmalen J, Vedder JE, Arbyn M, Klinkhamer PJ, Siebers AG, van Ballegooijen M. Liquid-based cervical cytology using ThinPrep technology: weighing the pros and cons in a cost-effectiveness analysis. *Cancer Causes Control* 2012;**23**: 1323-31.

20. Elbasha EH, Dasbach EJ, Insinga RP. A multi-type HPV transmission model. *Bull Math Biol* 2008;**70**: 2126-76.

21. 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;**100**: 308-20.

22. Kitchener HC, Canfell K, Gilham C, Sargent A, Roberts C, Desai M, Peto J. The clinical effectiveness and cost-effectiveness of primary human papillomavirus cervical screening in England: extended follow-up of the ARTISTIC randomised trial cohort through three screening rounds. *Health Technol Assess* 2014;**18**: 1-196.

23. Huh WK, Williams E, Huang J, Bramley T, Poulios N. Cost effectiveness of human papillomavirus-16/18 genotyping in cervical cancer screening. *Applied Health Economics & Health Policy* 2015;**13**: 95-107.

24. Karnon J, Peters J, Platt J, Chilcott J, McGoogan E, Brewer N. Liquid-based cytology in cervical screening: an updated rapid and systematic review and economic analysis. *Health Technol Assess* 2004;**8**: iii, 1-78.

25. Kim JJ, Campos NG, Sy S, Burger EA, Cuzick J, Castle PE, Hunt WC, Waxman A, Wheeler CM, New Mexico HPVPRSC. Inefficiencies and High-Value Improvements in U.S. Cervical Cancer Screening Practice: A Cost-Effectiveness Analysis. *Ann Intern Med* 2015;**163**: 589-97.

26. Kitchener HC, Blanks R, Cubie H, Desai M, Dunn G, Legood R, Gray A, Sadique Z, Moss S, Group MTS. MAVARIC - a comparison of automation-assisted and manual cervical screening: a randomised controlled trial. *Health Technol Assess* 2011;**15**: iii-iv, ix-xi, 1-170.

27. Kulasingam SL, Myers ER, Lawson HW, McConnell KJ, Kerlikowske K, Melnikow J, Washington AE, Sawaya GF. Cost-effectiveness of extending cervical cancer screening intervals among women with prior normal pap tests. *Obstet Gynecol* 2006;**107**: 321-8.

28. Scoggins JF, Ramsey SD, Jackson JC, Taylor VM. Cost effectiveness of a program to promote screening for cervical cancer in the Vietnamese-American population. *Asian Pac J Cancer Prev* 2010;**11**: 717-22.

29. van Rosmalen J, de Kok IM, van Ballegooijen M. Cost-effectiveness of cervical cancer screening: cytology versus human papillomavirus DNA testing. *BJOG* 2012;**119**: 699-709.

30. Voko Z, Nagyjanosi L, Margitai B, Kovi R, Toth Z, Laszlo D, Kalo Z. Modeling cost-effectiveness of cervical cancer screening in Hungary. *Value Health* 2012;**15**: 39-45.

31. van den Hout WB, Kramer GW, Noordijk EM, Leer JW. Cost-utility analysis of short- versus long-course palliative radiotherapy in patients with non-small-cell lung cancer. *J Natl Cancer Inst* 2006;**98**: 1786-94.

32. van den Hout WB, van der Linden YM, Steenland E, Wiggenraad RG, Kievit J, de Haes H, Leer JW. Single- versus multiple-fraction radiotherapy in patients with painful bone metastases: cost-utility analysis based on a randomized trial. *J Natl Cancer Inst* 2003;**95**: 222-9.

33. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004;**42**: 851-59.

34. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**: 1095-108.

35. Ronco G, Dillner J, Elfstrom KM, Tunesi S, Snijders PJ, Arbyn M, Kitchener H, Segnan N, Gilham C, Giorgi-Rossi P, Berkhof J, Peto J, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet* 2014;**383**: 524-32.

36. Simms K, Laprise J, Smith M, Lew J, Caruana M, Brisson M, Canfell K. Cost-effectiveness of the next generation nonavalent human papillomavirus vaccine in the context of primary human papillomavirus screening in Australia: a comparative modelling analysis. *The Lancet Public Health* 2016;**1**: e66 - e75

37. Lew JB, Simms K, Smith M, Lewis H, Neal H, Canfell K. Effectiveness Modelling and Economic Evaluation of Primary HPV Screening for Cervical Cancer Prevention in New Zealand. *PLoS One* 2016;**11**: e0151619.

38. Albrecht GL, Devlieger PJ. The disability paradox: high quality of life against all odds. *Social science & medicine* 1999;**48**: 977-88.

39. Tversky A, Kahneman D. Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics* 1991: 1039-61.

40. Ubel PA, Loewenstein G, Jepson C. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. *Quality of life Research* 2003;**12**: 599-607.

41. Schkade DA, Kahneman D. Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science* 1998;**9**: 340-6.

42. Drolet M, Brisson M, Maunsell E, Franco EL, Coutlee F, Ferenczy A, Fisher W, Mansi JA. The psychosocial impact of an abnormal cervical smear result. *Psychooncology* 2012;**21**: 1071-81.

43. Frederiksen ME, Njor S, Lynge E, Rebolj M. Psychological effects of diagnosis and treatment of cervical intraepithelial neoplasia: a systematic review. *Sexually transmitted infections* 2014: sextrans-2014-051754.

44. O'Connor M, Gallagher P, Waller J, Martin CM, O'Leary JJ, Sharp L. Adverse psychological outcomes following colposcopy and related procedures: a systematic review. *BJOG: An International Journal of Obstetrics & Gynaecology* 2016;**123**: 24-38.

45. Whitehurst DG, Bryan S, Lewis M. Systematic review and empirical comparison of contemporaneous EQ-5D and SF-6D group mean scores. *Med Decis Making* 2011;**31**: E34-44.

46. Luo N, Wang P, Fu AZ, Johnson JA, Coons SJ. Preference-based SF-6D Scores Derived From the SF-36 and SF-12 Have Different Discriminative Power in a Population Health Survey. *Med Care* 2012;**50**: 627-32.

47. Habbema D, Weinmann S, Arbyn M, Kamineni A, Williams AE, I MCMdK, van Kemenade F, Field TS, van Rosmalen J, Brown ML. Harms of cervical cancer screening in the United States and the Netherlands. *Int J Cancer* 2017;**140**: 1215-22.

48. Rebolj M, Bonde J, Preisler S, Ejegod D, Rygaard C, Lynge E. Human Papillomavirus Assays and Cytology in Primary Cervical Screening of Women Aged 30 Years and Above. *PLoS One* 2016;**11**: e0147326.

**Table 1.** Number of quality adjusted days lost (QADL) (utility loss x duration in days) found in published CEAs and empirical data per health state. Publications with unknown durations of health states are excluded. [20, 28, 30]

| Source | Induced by screening: screening and pre-invasive health states | | | | | | | | | Prevented by screening: invasive cancer health states | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Prim. screens | Pos. prim. screens | False-pos. prim. screens | Triage tests | Referrals for colposcopy | False-pos. Referral | CIN1 | CIN2 | CIN3 | FIGO 1 | FIGO 2+ | Palliative phase | Survivors |
| Empirical data EQ-5D | - | - | - | - | - | - | - | - | - | 11.0 | 21.0 | 182.5 | 146.0 |
| Empirical data SF-6D | - | - | - | 13.7 | 0.9 | - | - | - | - | 29.2 | 51.1 | 182.5 | 109.5 |
| Accetta[13] | - | - | - | - | - | - | - | - | - | 116.8 | 167.3 | - | - |
| Balasubramanian[14] | - | - | 1.1 | - | - | - | - | - | - | 438.0 | 602.3 | - | - |
| Berkhof[15] | - | 0.9 | - | - | - | - | 5.5 | 25.6 | 25.6 | 113.2 | 328.5 | - | - |
| Chuck[16] | - | - | - | - | - | - | 32.9 | 47.5 | 47.5 | 178.9 | 232.2 | - | - |
| Coupé[17, 18] | - | 0.9 | - | - | - | - | 5.5 | 25.6 | 25.6 | 113.2 | 328.5 | - | - |
| de Bekker-Grob[19] | 0.1 | - | - | 0.1 | - | 0.9 | 5.5 | 25.6 | 25.6 | 113.2 | 511.0 | 21.9 | - |
| de Kok[4]. van Rosmalen[29] | 0.1 | - | - | 1.1 | - | 0.9 | 5.5 | 25.6 | 25.6 | 113.2 | 511.0 | 21.9 | - |
| Goldhaber-Fiebert[21] | - | - | - | - | - | - | - | - | - | 116.8 | 167.3 | - | - |
| Karnon[24] | - | - | - | 9.0 | 18.0 | - | - | - | - | 1460.0 | 1460.0 | - | - |
| Kitchener[26] | - | - | 14.6 | - | - | - | 40.2 | 43.8 | 40.2 | 438.0 | 767.5 | - | - |
| Kitchener (Simonella)[22] | 1.2 | 9.7 | - | - | - | 10.1 | 10.1 | 10.8 | 10.8 | 438.0 | 767.5 | - | - |
| Kulasingam[27] | - | - | 0.9 | - | - | - | - | - | - | 273.8 | 821.3 | - | - |

Accetta, Chuck and Goldhaber-Fiebert: Duration health states unknown, assumed 1 year; Chuck: Assumed per invasive stage: 1 year without treatment and 1 year with treatment, total duration 2 years; de Kok: Assumed 'duration since last test' in case of triage testis 6 months; Karnon: Based on 'This mortality is based on an average life expectancy with invasive cancer present in an unscreened population of approximately 10 years', we assumed a duration of 10 years for 'remainder of lifetime'; Kitchener: Similar durations assumed for cancer as in 'Kitchener_Simonella'; Kulasingam: Only QoL loss due to cancer treatment (duration 5 years) included.

**Table 2.** Quality of life based utility scores, obtained with SF-6D and EQ-5D, Dutch questionnaire study.

| Stage | Assessments | SF-6D | | EQ-5D | | Paired t-test |
|---|---|---|---|---|---|---|
| | | (0-1, SD) | CI (mean ±1.96SE) | (0-1, SD) | CI (mean ±1.96SE) | *p*-value |
| **Reference** | | | | | | |
| 1. General population | 1. Between screening rounds | 0.83 (0.13) | 0.82-0.84 | 0.86 (0.20) | 0.85-0.88 | <0.001 |
| *Subgroup: screening ages (30-60 years)* | 1. Between screening rounds | 0.84 (0.12) | 0.83-0.85 | 0.88 (0.19) | 0.86-0.89 | <0.001 |
| **Screening** | | | | | | |
| 2. Invited for screening | 1. At invitation | 0.84 (0.10) | 0.84-0.85 | 0.89 (0.19) | 0.88-0.90 | <0.001 |
| *Subgroup:* Screen participants | 1. At invitation | 0.84 (0.10) | 0.84-0.85 | 0.89 (0.19) | 0.88-0.91 | <0.001 |
| | 2. After Pap test | 0.85 (0.10) | 0.85-0.86 | 0.90 (0.18) | 0.89-0.91 | <0.001 |
| | 3. After Pap test result | 0.86 (0.11) | 0.85-0.87 | 0.91 (0.17) | 0.90-0.92 | <0.001 |
| **Triage** | | | | | | |
| 3. Receiving a repeat Pap test after BMD[1.] result | 1. 6-24 months after BMD[1.] Pap test | 0.80 (0.13) | 0.79-0.82 | 0.87 (0.21) | 0.84-0.89 | <0.001 |
| *Subgroup: NOT (yet) referred for colposcopy* | 1. 6-24 months after BMD Pap test | 0.81 (0.12) | 0.79-0.83 | 0.88 (0.21) | 0.84-0.91 | <0.001 |
| **Referral for colposcopy** | | | | | | |
| 4. Initial six months following referral for colposcopy | 1. Shortly after suspicious Pap test | 0.81 (0.10) | 0.79-0.83 | 0.91 (0.13) | 0.89-0.94 | <0.001 |
| | 2. At 1 month f-up | 0.82 (0.12) | 0.79-0.84 | 0.89 (0.18) | 0.86-0.93 | <0.001 |
| | 3. At 3 months f-up | 0.82 (0.10) | 0.80-0.85 | 0.92 (0.17) | 0.89-0.95 | <0.001 |
| | 4. At 6 months f-up | 0.83 (0.12) | 0.81-0.86 | 0.91 (0.20) | 0.87-0.95 | <0.001 |
| 5. 6-35 months following referral for colposcopy[.] | 1. 6-35 months after treatment of CIN | 0.83 (0.11) | 0.80-0.85 | 0.91 (0.15) | 0.88-0.94 | <0.001 |

**(Follow up) Table 2.** Quality of life based utility scores, obtained with SF-6D and EQ-5D, Dutch questionnaire study.

| | | | | | | |
|---|---|---|---|---|---|---|
| **Cervical cancer** | | | | | | |
| 6. Diagnosed with cervical cancer: FIGO[2.] 1A&1B | 1. After diagnosis | 0.73 (0.13) | 0.69-0.77 | 0.79 (0.21) | 0.73-0.85 | 0.004 |
| | 2. At 3 months f-up | 0.75 (0.14) | 0.70-0.79 | 0.81 (0.14) | 0.77-0.85 | 0.001 |
| | 3. At 6 months f-up | 0.76 (0.15) | 0.71-0.81 | 0.82 (0.21) | 0.76-0.89 | 0.031 |
| | 4. At 12 months f-up | 0.81 (0.13) | 0.77-0.85 | 0.87 (0.13) | 0.83-0.91 | 0.016 |
| Diagnosed with cervical cancer: FIGO[2.] 2+ | 1. After diagnosis | 0.71 (0.18) | 0.62-0.80 | 0.72 (0.34) | 0.55-0.89 | 0.939 |
| | 2. At 3 months f-up | 0.63 (0.11) | 0.57-0.70 | 0.63 (0.31) | 0.46-0.80 | 0.628 |
| | 3. At 6 months f-up | 0.71 (0.13) | 0.64-0.79 | 0.74 (0.24) | 0.61-0.87 | 0.621 |
| | 4. At 12 months f-up | 0.75 (0.19) | 0.63-0.87 | 0.73 (0.38) | 0.51-0.94 | 0.603 |
| 7. Tumorfree, 2-10 years after diagnosis | 1. Disease free | 0.80 (0.14) | 0.79-0.82 | 0.82 (0.25) | 0.79-0.85 | 0.007 |
| 8. Having advanced cancer | 1. Last year of life<br>Breast cancer<br>Lung cancer<br>2. Last three months of life<br>Breast cancer<br>Lung cancer | - | - | <br>0.36 (0.37)<br>0.11 (0.38)<br><br>0.13 (0.39)<br>0.10 (0.36) | <br>0.33-0.40<br>0.02-0.21<br><br>0.08-0.18<br>-0.01-0.21 | - |

[1.]BMD=borderline or mild dyskaryosis;[2.]FIGO= International Federation of Gynecology and Obstetrics; f-up relates to follow-up in terms of timing of questionnaire assessments

**Table 3**. Preferred screening strategy at a threshold of €20,000 and €50,000 per QALY gained (3% discounting) by different sources (i.e. literature and empirical data) of QoL assumptions. Costs and QALYs gained presented are absolute, compared to no screening.

| Threshold | QoL assumptions based on | Primary test | Start age | Number of tests | Interval (years) | Costs (€, x1000) | QALYs gained | ICER (€) |
|---|---|---|---|---|---|---|---|---|
| €20,000 per QALY gained | Empirical data EQ-5D | HPV | 30 | 4 | 6 | 369 | 63 | 17,857 |
| | Empirical data SF-6D | Cytology | 32 | 5 | 5 | 409 | 63 | 19,963 |
| | Accetta[13] | HPV | 30 | 4 | 6 | 369 | 62 | 18,007 |
| | Balasubramanian[14] | HPV | 30 | 4 | 6 | 369 | 70 | 16,350 |
| | Berkhof[15] | HPV | 30 | 4 | 6 | 369 | 60 | 18,901 |
| | Chuck[16] | HPV | 30 | 4 | 6 | 369 | 57 | 20,420 |
| | Coupe[17, 18] | HPV | 30 | 4 | 6 | 369 | 60 | 18,901 |
| | de Bekker-Grob[19] | HPV | 30 | 4 | 6 | 369 | 62 | 18,844 |
| | de Kok[4], van Rosmalen[29] | HPV | 30 | 4 | 6 | 369 | 62 | 19,220 |
| | Goldhaber-Fiebert[21] | HPV | 30 | 4 | 6 | 369 | 62 | 18,007 |
| | Karnon[24] | Cytology | 30 | 6 | 4 | 472 | 89 | 17,151 |
| | Kitchener[26] | Cytology | 32 | 4 | 5 | 288 | 56 | 16,099 |
| | Kitchener (SImonella)[22] | HPV | 32 | 3 | 6 | 252 | 42 | 16,100 |
| | Kulasingam[27] | HPV | 30 | 5 | 6 | 520 | 79 | 19,706 |
| €50,000 per QALY gained | Empirical data EQ-5D | HPV | 27 | 9 | 5 | 1,074 | 84 | 47,378 |
| | Empirical data SF-6D | Cytology | 27 | 10 | 4 | 931 | 79 | 47,739 |
| | Accetta[13] | HPV | 27 | 9 | 5 | 1,074 | 83 | 47,977 |
| | Balasubramanian[14] | HPV | 27 | 9 | 5 | 1,074 | 92 | 41,426 |
| | Berkhof[15] | HPV | 30 | 7 | 6 | 833 | 75 | 43,015 |
| | Chuck[16] | HPV | 30 | 7 | 6 | 833 | 70 | 48,966 |
| | Coupe[17, 18] | HPV | 30 | 7 | 6 | 833 | 75 | 43,015 |
| | de Bekker-Grob[19] | HPV | 27 | 8 | 5 | 921 | 78 | 49,879 |
| | de Kok[4], van Rosmalen[29] | HPV | 30 | 7 | 6 | 833 | 75 | 46,202 |
| | Goldhaber-Fiebert[21] | HPV | 27 | 9 | 5 | 1,074 | 83 | 47,977 |
| | Karnon[24] | Cytology | 30 | 10 | 4 | 992 | 107 | 41,144 |
| | Kitchener[26] | Cytology | 27 | 7 | 5 | 586 | 66 | 39,120 |
| | Kitchener (SImonella)[22] | HPV | 32 | 3 | 6 | 252 | 42 | 16,100 |
| | Kulasingam[27] | HPV | 27 | 9 | 5 | 1,074 | 95 | 40,037 |

## Figure legends

**Figure 1.** Utilities measured by the SF-6D and EQ-5D questionnaire per stage in the screening and follow up process. Dutch questionnaire study.

**Figure 2.** Net costs and health effects (compared to no screening) of the efficient screening programmes for the fourteen different QADL-sets, for a cohort of 100,000 unvaccinated women, for the period from 2011 onwards. Costs and effects discounted with 3% towards year 2011. Each point corresponds with a screening policy on the efficient frontier. Black marker points represent the policy at the €20,000 (left-side on the frontier) and €50,000 (right-side at the frontier) per quality-adjusted life year (QALY) gained threshold (see table 3).

**APPENDIX 'Quality of life assumptions determine which cervical cancer screening strategies are cost-effective'**

This appendix gives a description of the data-collection procedures and the background characteristics of the participants of the questionnaire study (pages 1-3), summarized in Appendix Tables 1 and 2. Furthermore, Appendix Table 3 shows the assumed utilities in included published cost-effectiveness analyses (CEAs) per health stage. Table 4 presents the sources of the utility values used in the different CEAs. Finally, a description of the model and an overview of the input used for the cost-effectiveness calculations is given (Appendix Table 5).

## The questionnaire study

In the Netherlands, at time of the questionnaire study, women aged 30–60 years, were invited every 5 years to attend cytological screening free of charge. Nationally in 2014, the 5-years coverage was 77%.[1] Of the women screened for cervical cancer, 94.3% had normal test results and 1.7% of Pap tests were of inadequate quality requiring repeat tests. In 3.2% test results were borderline or mildly dyskaryotic (BMD, Pap 2/3a1), and in 0.8% high-grade cytological abnormalities, including moderately dyskaryotic (Pap 3a2) or worse, were found.[1]

Following the screening protocol, women with borderline or mildly dyskaryotic results (Pap2/3a1) were advised to have follow-up Pap tests with their general practitioner.[2] If follow-up tests were again borderline or mildly dyskaryotic, these women were referred for colposcopy. Women with high-grade cytological abnormalities (Pap 3a2 or worse) were immediately referred for colposcopy and punch biopsy.

## Study groups in the questionnaire study

1. **Reference population**
   To include a reference group from the general population 1,800 randomly selected women (aged 30–70 years, stratified in 10-year age groups) were sent a questionnaire through the regional screening organization in Maastricht (The Netherlands) between April and June 2006. 835 women completed the questionnaire (response rate 46%).

2. **Screening**
   A random sample of 2,300 women, aged 30-60 years, was obtained at the moment of delivery of the screening invitation from the regional screening organization in Maastricht (The Netherlands) between April and August 2006 [18]. Sampled women received a letter, a questionnaire, and a consent form. Of the 2,300 addressed women

1,551 had the screening test. 1,023 women completed the baseline questionnaire (response rate 44%). At a group level we know that 924 women who completed the first questionnaire had the screening test, at an individual level this is known to us about 905 women. The response rate among screen participants was 60% (924/1,551 screen participants).
*Subgroup of screen participants:* Women who gave consent to participate in our questionnaire study and who had a screening test within four months following their invitation were sent a second questionnaire following their Pap test and a third one attached to the letter informing them about their test results.

3. **Triage**
Between April and August 2006 we conducted a cross-sectional survey in cooperation with the regional screening organization in Maastricht (The Netherlands) [2]. Five hundred and fifty women who participated in the Dutch cervical cancer-screening programme with a borderline (Pap 2) or mildly (Pap 3a1) dyskaryotic (BMD) Pap smear result in the previous 6–24 months were sent a questionnaire. 270 women completed the questionnaire (response rate 49%).

REFERRAL FOR COLOSCOPY

4. **Undergoing gynecological evaluation**
Between February 2006 and April 2008 a prospective longitudinal cohort study was conducted in two Dutch hospitals [19], aiming to include all consecutive women who were referred for gynecological evaluation because of abnormal Pap test results in the national screening program. These women were sent a letter, asking for informed consent to participate in the study. This involved completion of four questionnaires in the course of six months. Women were also asked for permission to consult their patient files and/or the gynecologist about their colposcopy follow-up. If the patient file or the information of the gynecologist showed that women were ineligible, e.g. since they had not been referred in the context of the screening program, women were excluded.

5. **Treated for CIN**
From February 2006 onwards in two Dutch hospitals women who had been referred for gynaecological evaluation 6 to 35 months earlier and had been treated for a precursor of cervical cancer were identified in the medical files and approached with a letter and a questionnaire. 81 women completed the questionnaire (response rate 49 %).

CERVICAL CANCER

**6. Diagnosed with cervical cancer**

Between March 2006 and April 2008 a prospective longitudinal cohort study was conducted in four Dutch hospitals, aiming to include all consecutive women who were diagnosed with cervical cancer. These women were sent a letter, asking for informed consent to participate in the study. This involved completion of four questionnaires in the course of 12 months. Women were also asked for permission to consult their patient files and/or the gynaecologist about their treatment.

7. **Tumorfree 2-10 years after diagnosis**

In cooperation with the Eindhoven regional cancer registry and all gynaecologic oncology departments in the region, we conducted a population-based cross-sectional survey[17]. We identified all women who had been diagnosed with cervical cancer between January 1995 and December 2003 (n = 691). Women who had died could be excluded through linkage with the Central Bureau for Genealogy that collects data for all deceased Dutch citizens. We identified 444 women (64%) who were still alive on January 31, 2006. Of this group, 8 women were not contacted as advised by their physician, e.g., because of serious (mental) illness, and 15 addresses could not be verified. The remaining 421 survivors were sent a questionnaire by their (former) gynaecologists. In the accompanying letter, the women were asked to complete the questionnaire and were informed that by returning the completed questionnaire, they consented to linkage of the questionnaire data with their disease history as registered by the cancer registry. 291 women completed the questionnaire (response rate 69%). For this paper we excluded women who were known to have recurrence at time of questionnaire completion (n=6).

8. **Having advanced cancer**

For logistic reasons we were not able to include women with advanced cancer in our study. For this patient group we opted for using data that were obtained in another study context[15,16]. Data were available for women with advanced stages of breast (n=340) and lung cancer (n=44) in their last year of life.

**Appendix Table 1** Overview of study groups and timing of assessments

| STAGE | Description of groups (n) | Inclusion site(s) | Utility assessments |
|---|---|---|---|
| **Reference** | | | |
| 1. General population | Women who had not been invited for cervical cancer screening in the previous 2 years, n= 835 | Screening organisation | 1. In between screening rounds |
| *Subgroup: screening ages (30-60 years)* | Women eligible for cervical cancer screening, n=612 | | 1. In between screening rounds |
| **Screening** | | | |
| 2. Invited for cervical cancer screening | Women invited for the next screening round, n=1,023 | Screening organisation | 1. At moment of invitation |
| *Subgroup: screen participants* | Women who had a Pap test following the invitation | | 1. Before screening (n= 905).<br>2. After screening (n=802)<br>3. After test result (n=843) |
| **Triage** | | | |
| 3. Receiving a repeat Pap test after BMD[1.] result | Women who were advised to have a repeat Pap test because of a BMD Pap test, n=270 | Screening organisation | 1. 6-24 months after BMD Pap test result |
| *Subgroup: not (yet) referred for colposcopy* | n=159 | | 1. 6-24 months after BMD Pap test result |
| **Referral for colposcopy** | | | |
| 4. Initial six months following referral for colposcopy | Women who underwent a gynaecological evaluation after abnormal Pap results, n=132 | Two gynaecology departments of general hospitals | 1. Shortly after suspicious test result (n=132);<br>2. At 1 month follow-up* (n=114);<br>3. At 3 months follow-up* (n=110);<br>4. At 6 months follow-up* (n=108) |
| 5. 6-35 months following referral for colposcopy | Women treated for CIN[2.], n=81 | Two gynaecology departments of general hospitals | 1. 6-35 months after treatment of CIN |
| **Cervical cancer** | | | |
| 6. Diagnosed with cervical cancer<br>*Subdivided into 'FIGO[3.] 1a & 1b' and 'FIGO[3.] 2+'* | Women who were recently diagnosed with invasive cervical cancer, n=77 | Four gynaecology departments of general hospitals | 1. Shortly after diagnosis (n=77)<br>2. At 3 months follow-up* (n=66)<br>3. At 6 months follow-up* (n=64)<br>4. At 12 months follow-up* (n=62) |
| 7. Tumorfree 2-10 years after diagnosis | Women who were treated for cervical cancer, n=285 | One cancer registry region (17 hospitals) | 1. 2-10 years after cervical cancer diagnosis and treatment |
| *8. Having advanced cancer* | *Women with advanced cancer*<br>*Breast cancer, n=340*<br>*Lung cancer, n=44* | *Secondary data analysis* | *1. Last 12 months of life*<br>*2. Last 3 months of life* |

[1.]BMD=borderline of mild dyskaryosis; [2.]CIN = cervical intraepithelial neoplasia; [3.]FIGO= International Federation of Gynecology and Obstetrics; follow-up refers to timing of questionnaire assessment

**Appendix Table 2** Background characteristics of the participants

| Stage | N | **Education** [1.] (N (%)) | | | **Employment status** (N (%)) | | | | **Marital status** (N (%)) | | **Age** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Low | Medium | High | Paid job | Housewife/unpaid job/student | No job | Retired | Married/ cohabiting | Living without partner | Mean (SD) | range |
| **Reference** | | | | | | | | | | | | |
| 1. General population | 835 | 273 (35%) | 358 (46%) | 150 (19%) | 372 (51%) | 238 (32%) | 52 (7%) | 72 (10%) | 654 (78%) | 181 (22%) | 51.8 (11.4) | 27-72 |
| *Subgroup: screening ages (30-60 years)* | 612 | 154 (27%) | 296 (51%) | 129 (22%) | 358 (67%) | 125 (23%) | 44 (8%) | 9 (2%) | 499 (82%) | 113 (19%) | 47.2 (8.9) | 27-62 |
| **Screening** | | | | | | | | | | | | |
| 2. Invited for cervical cancer screening | 1,023 | 209 (23%) | 457 (49%) | 263 (28%) | 608 (68%) | 201 (22%) | 74 (8%) | 16 (2%) | 807 (79%) | 211 (21%) | 45.1 (9.4) | 29-60 |
| *Subgroup: screen participants* | 905 | 181 (22%) | 410 (50%) | 227 (28%) | 541 (68%) | 178 (22%) | 64 (8%) | 14 (2%) | 715 (79%) | 185 (21%) | 45.3 (9.4) | 29-60 |
| **Triage** | | | | | | | | | | | | |
| 3. Receiving a repeat Pap test after BMD[2.] result | 270 | 39 (16%) | 132 (56%) | 67 (28%) | 168 (74%) | 34 (15%) | 26 (11%) | - | 188 (74%) | 65 (26%) | 43.0 (7.9) | 30-62 |
| *Subgroup: NOT yet referred for colposcopy* | 159 | 19 (14%) | 87 (62%) | 35 (25%) | 100 (74%) | 25 (18%) | 11 (8%) | - | 113 (75%) | 38 (25%) | 42.6 (7.7) | 30-62 |
| **Referral for colposcopy** | | | | | | | | | | | | |
| 4. Initial six months following referral for colposcopy | 114 | 17 (16%) | 67 (63%) | 22 (21%) | 82 (83%) | 11 (11%) | 6 (6%) | - | 81 (74%) | 29 (26%) | 40.8 (8.4) | 29-60 |
| 5. 6-35 months following referral for colposcopy. | 81 | 14 (18%) | 46 (58%) | 20 (25%) | 61 (82%) | 9 (12%) | 4 (5%) | - | 63 (78%) | 18 (22%) | 41.3 (8.0) | 30-62 |
| **Cervical cancer** | | | | | | | | | | | | |
| 6. Diagnosed with cervical cancer: FIGO[4.] 1A&1B | 47 | 7 (18%) | 27 (68%) | 6 (15%) | 28 (72%) | 7 (18%) | 3 (8%) | 1 (3%) | 27 (66%) | 14 (34%) | 42.8 (9.0) | 29-72 |
| Diagnosed with cervical cancer: FIGO[4.] 2+ | 16 | 2 (18%) | 8 (73%) | 1 (9%) | 7 (70%) | 1 (10%) | 1 (10%) | 1 (10%) | 7 (58%) | 5 (42%) | 46.7 (15.4) | 30-89 |
| 7. Tumorfree, 2-10 years after diagnosis | 285 | 114 (44%) | 115 (45%) | 29 (11%) | 94 (41%) | 83 (37%) | 28 (12%) | 22 (10%) | 184 (67%) | 91 (33%) | 52.7 (13.6) | 31-88 |
| 8. Having advanced cancer[5] | 340 (breast cancer)<br>44 (lung cancer) | | | | | | | | | | 59.6 (12.5)<br>63.3 (10.1) | 32-89<br>45-87 |

[1.] Educational level was classified as low (primary school or lower technical education), intermediate, or high (college/university degree); [2.] BMD=borderline of mild dyskaryosis; [3.] CIN = cervical intraepithelial neoplasia; [4.] FIGO= International Federation of Gynecology and Obstetrics; [5.] these data were collected by others, with information about age being available for the whole dataset, and data on employment status being available for a subset of 61 participants

**The literature review**

To collect the utilities that have been used so far, we reviewed the literature for CEAs of cervical cancer screening published between 2003 and 2015. We used the following search query in Pubmed:

*((("cost-benefit analysis"[MeSH Terms] OR ("cost-benefit"[All Fields] AND "analysis"[All Fields]) OR "cost-benefit analysis"[All Fields] OR ("cost"[All Fields] AND "effectiveness"[All Fields] AND "analysis"[All Fields]) OR "cost effectiveness analysis"[All Fields]) AND (("uterine cervical neoplasms"[MeSH Terms] OR ("uterine"[All Fields] AND "cervical"[All Fields] AND "neoplasms"[All Fields]) OR "uterine cervical neoplasms"[All Fields] OR ("cervical"[All Fields] AND "cancer"[All Fields]) OR "cervical cancer"[All Fields]) AND ("diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields])))) AND ("quality-adjusted life years"[MeSH Terms] OR ("quality-adjusted"[All Fields] AND "life"[All Fields] AND "years"[All Fields]) OR "quality-adjusted life years"[All Fields] OR "qaly"[All Fields])) AND ("2003/01/01"[PDAT] : "2015/12/31"[PDAT])*

Appendix Table 3 shows the assumed utilities in included published CEAs per health stage, as well as the source of the utilities reported in the specific CEAs. Appendix Table 4 presents the original sources of assumed utilities. We found that all CEAs finally refer to 12 different sources, published between 1991 and 2012:

- de Haes et al. Int J Cancer 1991
- Fryback DG, et al. Med Decis Making 1993
- Wolfson MC. Health Rep 1996
- Gold MR, et al. Med Care.1998
- Stratton et al. Vaccines for the 21th century: A tool for decision making. Washington: National Academy Press, 2000
- Conference abstract: Myers ER. 21st International Papillomavirus Conference Mexico City, 2004.
- Maissi E, et al Br J Cancer 2005
- Hanmer J, et al. Med Decis Making 2006
- Insinga et al. Med Decis Making 2007
- Conference abstract: Insinga RP, et al. 22nd International Papillomavirus Conference Vancouver, Canada, 2007.
- Thesis: Simonella LM. NSW: University of Sydney, 2012.
- Website: Center for the Evaluation of value and Risk in Health. Catalog of Preference Scores: the CEA Registry. Boston, MA: Tufts New England Medical Center.

**Appendix Table 3.** Assumed utilities in included published CEAs per health stage (in case of two utility values for one stage, the first utility value represents the short term effect and the second value represents the long term effect)

| | Induced by screening: screening and pre-invasive health states | | | | | | | | | Prevented by screening: invasive cancer health states | | | | Sources reported for utility values |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Prim. screens | Pos. prim. screens | False-pos. prim. screens | Triage tests | Referrals for colpos copy | False-pos. Referral | CIN1 | CIN2 | CIN3 | FIGO 1 | FIGO 2+ | Palliative phase | Survivors | |
| Accetta[3] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.68 | 0.56/0.48 | 0 | 0 | [4] |
| Balasubramanian[5] | 0 | 0 | 0.98 | 0 | 0 | 0 | 0 | 0 | 0 | 0.76 | 0.67 | 0 | 0 | [6] |
| Berkhof[7] | 0 | 0.97 | 0 | 0 | 0 | 0 | 0.97 | 0.93 | 0.93 | 0.65/0.97 | 0.55/0.85 | 0 | 0 | [8-10] |
| Chuck[11] | 0 | 0 | 0 | 0 | 0 | 0 | 0.91 | 0.87 | 0.87 | 0.65/0.86 | 0.67/0.83[1] | 0 | 0 | [12] |
| Coupe[13, 14] | 0 | 0.995 | 0 | 0 | 0 | 0 | 0.97 | 0.93 | 0.93 | 0.65/0.97 | 0.55/0.85 | 0 | 0 | [8, 10] |
| de Bekker-Grob[15] | 0 | 0 | 0 | 0.97 | 0.994 | 0.994 | 0.97 | 0.93 | 0.93 | 0.938 | 0.72 | 0.288 | 0 | [8, 16] |
| de Kok[17]. van Rosmalen[18] | 0 | 0 | 0 | 0.995 | 0.994 | 0.994 | 0.97 | 0.93 | 0.93 | 0.938 | 0.72 | 0.288 | 0 | [8, 10, 16] |
| Goldhaber-Fiebert[4] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.68 | 0.56[2] | 0 | 0 | [19-21] |
| Karnon[22] | 0 | 0 | 0 | 0.975[3] | 0.95[3] | 0 | 0 | 0 | 0 | 0.6 | 0.6 | 0 | 0 | |
| Kitchener[23] | 0 | 0 | 0.96 | 0 | 0 | 0 | 0.89 | 0.88 | 0.89 | 0.76 | 0.67[4] | 0 | 0 | [6, 8, 10, 21, 24-26] |
| Kitchener (Simonella)[27] | 0.997 | 0.974 | 0 | 0 | 0 | 0.972 | 0.972 | 0.970 | 0.970 | 0.76 | 0.67[4] | 0 | 0 | [10, 28, 29] |
| Kulasingam[30] | 0 | 0 | 0.97 | 0 | 0 | 0 | 0 | 0 | 0 | 0.85 | 0.55 | 0 | 0 | [8, 21] |

[1] Represents the utility values for FIGO 2. assumptions for other stages are: FIGO3 - 0.56/0.83. FIGO4 - 0.48/0.63

[2] Represents utility value for FIGO2/3. for FIGO 4 a value of 0.48 is assumed.

[3] Average base case utility value (the study uses two other base case utility values as well: 0.95 and 0.98 for triage. 0.9 and 0.97 for colposcopy

[4] Represents the utility values for FIGO 2. assumptions for other stages are: FIGO3 - 0.56. FIGO 4 - 0.48

**Appendix Table 4**. Literature sources of assumed utilities in included published CEAs. Gray shaded sources are the original utility studies.

| | | Sources reported for utility values | Sources of sources |
|---|---|---|---|
| 1 | Accetta[3] | Goldhaber-Fiebert JD, et al. J Natl Cancer Inst 2008[4] | See 8 |
| 2 | Balasubramanian[5] | Myers ER. 21st IPV Conference Mexico City, 2004. [6] | |
| 3 | Berkhof[7] | Mandelblatt JS,et al JAMA 2002[8] | Gold MR, et al. Med Care.1998[31] |
| | | Maissi E, et al Br J Cancer 2005[9] | |
| | | Goldie SJ, et al. J Natl Cancer Inst 2004[10] | Stratton et al. Vaccines for the 21th century: A tool for decision making. Washington: National Academy Press, 2000[24]<br>Wolfson MC. Health Rep 1996[32]<br>Gold MR, et al. Med Care.1998[31] |
| 4 | Chuck[11] | Krahn M, et al. Canadian Agency for Drugs and Technologies in Health, 2008. [12] | Myers ER. 21st IPV Conference Mexico City, 2004. [6] |
| | | | Goldie SJ et al. Int J Cancer 2003*[33] |
| | | | Stratton et al. Vaccines for the 21th century: A tool for decision making. Washington: National Academy Press, 2000[24]<br>Wolfson MC. Health Rep 1996[32]<br>Gold MR, et al. Med Care.1998[31] |
| | | | Mandelblatt et al. JAMA 1988*[34] |
| 5 | Coupe[13, 14] | Mandelblatt JS, et al JAMA 2002[8] | Gold MR, et al. Med Care.1998[31] |
| | | Goldie SJ, et al. J Natl Cancer Inst 2004[10] | Stratton et al. Vaccines for the 21th century: A tool for decision making. Washington: National Academy Press, 2000[24]<br>Wolfson MC. Health Rep 1996[32]<br>Gold MR, et al. Med Care.1998[31] |
| 6 | de Bekker-Grob[15] | Mandelblatt JS,et al. JAMA 2002[8] | Gold MR, et al. Med Care.1998[31] |
| | | **Thesis:** van Ballegooijen M. Erasmus University Rotterdam, 1998. [16] | de Haes et al. Int J Cancer 1991[35] |
| 7 | de Kok[17]van Rosmalen[18] | Mandelblatt JS, et al. JAMA 2002[8] | Gold MR, et al. Med Care.1998[31] |
| | | Goldie SJ, et al. J Natl Cancer Inst 2004[10] | Stratton et al. Vaccines for the 21th century: A tool for decision making. Washington: National Academy Press, 2000[24]<br>Wolfson MC. Health Rep 1996[32]<br>Gold MR, et al. Med Care.1998[31] |
| | | **Thesis:** van Ballegooijen M. Erasmus University Rotterdam, 1998. [16] | de Haes et al. Int J Cancer 1991[35] |

**Appendix Table 4 (continued)**. Literature sources of assumed utilities in included published CEAs. Gray shaded sources are the original utility studies

| | | | |
|---|---|---|---|
| 8 | Goldhaber-Fiebert[4] | Hanmer J, et al. Med Decis Making 2006[19] | |
| | | **Website:** Center for the Evaluation of value and Risk in Health. Boston, MA: Tufts New England Medical Center.[20] | |
| | | Kim JJ, et al. JAMA 2002[21] | Gold MR, et al. Med Care.1998[31]<br>Fryback DG, et al. Med Decis Making 1993 |
| 9 | Karnon[22] | *Not reported* | |
| 10 | Kitchener[23] | Myers ER. 21st IPV Conference Mexico City, 2004. [6] | |
| | | Mandelblatt JS, et al. JAMA 2002[8] | Gold MR, et al. Med Care.1998[31] |
| | | Goldie SJ, et al. J Natl Cancer Inst 2004[10] | Stratton et al. Vaccines for the 21th century: A tool for decision making. Washington: National Academy Press, 2000[24]<br>Wolfson MC. Health Rep 1996[32]<br>Gold MR, et al. Med Care.1998[31] |
| | | Kim JJ, et al. JAMA 2002[21] | Gold MR, et al. Med Care.1998[31]<br>Fryback DG, et al. Med Decis Making 1993[36] |
| | | Stratton et al. Vaccines for the 21th century: A tool for decision making. Washington: National Academy Press, 2000[24] | |
| | | Sanders GD, Taira AV. Emerg Infect Dis 2003[25] | Stratton et al. Vaccines for the 21th century: A tool for decision making. Washington: National Academy Press, 2000[24]<br>Fryback DG, et al. Med Decis Making 1993[36] |
| | | Insinga RP, et al. 22nd International Papillomavirus Conference Vancouver, Canada, 2007. [26] | |
| 11 | Kitchener (Simonella)[27] | Goldie SJ, et al. J Natl Cancer Inst 2004[10] | Stratton et al. Vaccines for the 21th century: A tool for decision making. Washington: National Academy Press, 2000[24]<br>Wolfson MC. Health Rep 1996[32]<br>Gold MR, et al. Med Care.1998[31] |
| | | Elbasha EH, et al. Emerg Infect Dis 2007[28] | Gold MR, et al. Med Care.1998[31]<br>Myers ER. 21st IPV Conference Mexico City, 2004. [6]<br>Insinga et al. Med Decis Making 2007[37] |
| | | **Thesis:** Simonella LM. NSW: University of Sydney, 2012. [29] | |
| 12 | Kulasingam[30] | Mandelblatt JS, et al. JAMA 2002[8] | Gold MR, et al. Med Care.1998[31] |
| | | Kim JJ, et al. JAMA 2002[21] | Gold MR, et al. Med Care.1998[31]<br>Fryback DG, et al. Med Decis Making 1993[36] |

*Sources of sources of sources could not be found

**Model description for cost-effectiveness calculations**

costs and the effects of different programmes were estimated using the microsimulation screening analysis (MISCAN) model. In MISCAN a large population is simulated that consists of hypothetical (but realistic) individual life histories. If applied to cervical cancer some women will develop high-risk HPV infection(s), cervical neoplasia or cancer. This simulation yields an age-specific output of the prevalence of HPV infections and cervical neoplasia, and the incidence and the mortality of cervical cancer. This simulated population then undergoes simulated (but realistic) screening, which changes some of the life histories. The effects of screening can be determined from the changes in the life histories using the numbers of events and stages induced or prevented. This approach yields the changes in the number of life years, the quality of life and the costs.

*Model specifications: demography, epidemiology and natural history*
For this manuscript we simulated the Dutch population at risk for cervical cancer based on demographic and hysterectomy data;[38 39] mortality from other causes was estimated using the observed age-specific mortality in the Netherlands in 2008.[38] The age-specific incidence of HPV infections that progress to cervical cancer was calibrated to the age distribution of the prescreening mortality in the Netherlands; the latter distribution was corrected for cohort effects based on an age–period–cohort analysis.[16] The estimated cumulative incidence of HPV infections that progress to cervical cancer is 1.06% for women born in the period 1939–1948, and 1.48% for women born after 1948; these assumptions have been used previously.[40] The age-specific incidence of pre-invasive lesions that do not progress to cancer was calibrated so that the simulated detection rates of cervical intraepithelial neoplasia (CIN) fit the observed CIN detection rates in the Netherlands; the observed detection rates were obtained from the Dutch Network and National Database for Pathology (PALGA) for the period 1997–2001. Finally, the age-specific incidence of high-risk HPV infections that do not progress to CIN was calibrated so that the simulated prevalence of all high-risk HPV infections fits the observed high-risk HPV prevalence.[41]
In the model, the disease is subdivided into seven sequential stages: high-risk HPV infection (low-risk HPV infections are not simulated in the model); three pre-invasive stages (CIN 1, 2 and 3), and three invasive stages (International Federation of Obstetrics and Gynecology staging; FIGO 1A, 1B and 2+). Pre-invasive stages and FIGO 1A cases can only be diagnosed by screening, as they are assumed to be asymptomatic, whereas FIGO 1B and 2+ cases can also be clinically diagnosed. HPV infections are usually not progressive; in the model, more than 90% of HPV infections will clear without resulting in CIN, and most lesions in the pre-invasive stages regress naturally. For example, in the absence of cervical screening. approximately 72% of the CIN 3 lesions in the model would not become cancer, which corresponds well with estimates from a retrospective cohort study of women with untreated CIN 3.[42] In a CIN lesion, a high-risk HPV infection may or may not be present. In the model it is assumed that, without an HPV infection, a CIN lesion will not progress to cervical cancer. In the invasive stages, all women are assumed to be HPV infected. A woman can develop multiple HPV infections and neoplasias in her lifetime, and multiple infections and neoplasias can exist at the same time. Weibull probability distributions are used to assume variation among women in the durations of the different stages. The inputs on stage-specific survival in clinical cases are age specific, and are based on observed survival

data and on Dutch mortality-to-incidence ratios from the prescreening period in the Netherlands.[16]
We used a population model that simulates the life histories of 8 million unvaccinated women born between 1939 and 1992; women born before 1939 are too old to attend screening after 2011, and women born after 1992 are eligible for HPV vaccination. The simulated screening programmes start in 2011 and continue until all women have completed their screening programmes. Results of screening practices before 2011, can influence the effectiveness of the screening programme after 2011. Therefore, we also simulated the last three screening rounds before 2011, based on the assumption that screening rounds before 1996 will not affect the screening results after 2011. Information on the screening activities before 2011 was obtained from PALGA.

*Assumptions for screening and treatment*
In our analyses, we varied the ages at which screening takes place. We considered all screening policies with starting ages of 25, 27, 30 or 32 years that comprise at least three and at most ten screenings in a woman's lifetime, and that have an interval of at least 3 years and at most 10 years; policies that include screenings over the age of 70 years were not simulated.
We assumed that 10% of the population never attends screening and has a three times higher background risk (i.e. using the cervical cancer incidence in a situation without screening) than the 90% potential attenders. This assumption is based on an analysis of data from the start of organised screening.[43] We assumed that the potential attenders attend 80% of all primary screenings. so that the overall attendance rate is 72%; under the base-case assumptions, follow-up screenings and referrals for colposcopy are always attended. The sensitivity of the HPV test (the probability of a positive test result if an HPV infection is present) was estimated at 94%, and the sensitivity of cytology was assumed to be 40% for CIN 1, 50% for CIN 2, and 75% for CIN 3 and invasive cervical cancer (see Appendix Table 5).[7] The specificity of the HPV test (probability of a negative test for women without high risk HPV infections) is assumed to be 100%, and the specificity of cytology (probability of a negative test for women without CIN or cancer) is estimated to be 98.5%, based on the observed false positive rate of Pap smears in the Dutch screening programme. Several screening strategies distinguish between smears read as ASC-US/LSIL (atypical squamous cells and low-grade cervical squamous intraepithelial lesions, equivalent to borderline/mild dyskaryosis) and smears read as at least HSIL (high-grade cervical squamous intraepithelial lesions, equivalent to moderate dyskaryosis). Therefore, the probability of at least HSIL is also specified for each disease stage in Appendix Table 5. The detection and the associated management (including retreatment, if necessary) of pre-invasive lesions were assumed to lead to a 100% cure rate. For screen-detected invasive cancers, the survival was modelled as a reduction in the risk of dying from cervical cancer compared with that of dying from clinically diagnosed cancer: in the model, detection by screening of an invasive cancer results in a reduction of the risk of dying of cervical cancer of 80% (FIGO 1A), 60% (FIGO 1B) or 20% (FIGO 2+).

*Assumptions for costs*
Appendix Table 5 presents the costs used in the analysis. The estimated costs are based on a societal perspective, and are reported in 2010 euros (€). The screening costs include the costs for the invitational system and quality assurance, the time and travel costs of the

woman being screened, the costs of smear taking, the costs of cytological evaluation, the costs of repeat tests after an inadequate test result and the costs of registration in PALGA. The diagnosis costs for women referred for colposcopy, the treatment costs for detected pre-invasive lesions, the costs of primary treatment for invasive cervical cancer, and the costs of treatment and palliative care for advanced cervical cancer were derived from previous cost studies performed in the Netherlands.[44]

*Cost-effectiveness analysis*

The costs and the effects of each simulated screening programme are counted for the period from 2011 onwards. Future costs and health effects (life years and utility losses) are discounted in the base-case analysis towards the year 2011 at a rate of 3%. Programmes that are more costly and less effective than other programmes are ruled out as nonefficient (i.e. by simple dominance). Programmes that are more costly and less effective than a combination of other programmes are also ruled out as non-efficient (i.e. by extended dominance). The remaining programmes constitute the frontier of efficient screening programmes. The total costs consist of the costs of the invitations (including the costs of the invitational system and the quality assurance), the primary and follow-up screenings, the treatment of pre-invasive and invasive lesions, and terminal care. We compute the net costs of screening as the difference in the total costs between the simulation in which the screening programme is implemented and a simulation without cervical screenings after 2011. The total number of QALYs is the number of years lived by the population minus the utility losses associated with attending screening, receiving treatment and having a terminal stage of cervical cancer. The number of QALYs gained by screening is the total number of QALYs in the simulation with the screening programme, minus the total number of QALYs in a simulation without screening after 2011.

**Appendix Table 5**. Model assumptions which are identical for all QoL scenario's: test characteristics and costs

| Parameter | Value |
|---|---|
| Attendance of potential attenders* | 80% |
| **Sensitivity of cytology** | |
| Probability of at least ASCUS (at least triage) for: | |
| CIN grade I | 40% |
| CIN grade II | 50% |
| CIN grade III and cervical cancer | 75% |
| Probability of at least HSIL (referral for colposcopy) for: | |
| CIN grade I | 4% |
| CIN grade II | 19% |
| CIN grade III and cervical cancer | 47% |
| Specificity of cytology (CIN grade I or worse) | 98.5% |
| Sensitivity of HPV test‡ | 94% |
| Specificity of HPV test§ | 100% |
| **Costs of screening (€)** | |
| Invitation | 4.65 |
| Cytology (first) | 60.39 |
| Cytology (repeat after at least 6 months) | 30.27 |
| HPV-test (first) | 62.80 |
| HPV-test (repeat) | 29.00 |
| **Costs of treatment pre-invasive disease (€)** | |
| False positive | 279 |
| CIN1 | 869 |
| CIN2 | 1.287 |
| CIN3 | 1.507 |
| **Costs of treatment of invasive cancer (€)** | |
| FIGO 1A | 4.935 |
| FIGO 1B | 11.703 |
| FIGO 2+ | 10.773 |
| Terminal care | 26.209 |

*The potential attenders consist of 90% of the female population; the remaining women are assumed to never attend screening.
‡Probability to detect an HPV infection. regardless of whether a CIN lesion or cancer is present.
§A possible lack of specificity was modelled by including fast-clearing HPV infections.

**References**


1. Erasmus MC, Screening program cervical cancer, report on 2014 [bevolkingsonderzoek baarmoederhalskanker, LEBA rapportage 2015] 2015.

2. 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;**120**: 806-12.

3. Accetta G, Biggeri A, Carreras G, Lippi G, Carozzi FM, Confortini M, Zappa M, Paci E. Is human papillomavirus screening preferable to current policies in vaccinated and unvaccinated women? A cost-effectiveness analysis. *J Med Screen* 2010;**17**: 181-9.

4. 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;**100**: 308-20.

5. Balasubramanian A, Kulasingam SL, Baer A, Hughes JP, Myers ER, Mao C, Kiviat NB, Koutsky LA. Accuracy and cost-effectiveness of cervical cancer screening by high-risk human papillomavirus DNA testing of self-collected vaginal samples. *J Low Genit Tract Dis* 2010;**14**: 185-95.

6. Myers ER. Patient preferences for health states related to HPV infection. Visual analog scalesversus time trade-off elicitation 21st International Papillomavirus Conference Mexico City, 2004.

7. Berkhof J, Coupe VM, Bogaards JA, van Kemenade FJ, Helmerhorst TJ, Snijders PJ, Meijer CJ. The health and economic effects of HPV DNA screening in The Netherlands. *Int J Cancer* 2010;**127**: 2147-58.

8. Mandelblatt JS, Lawrence WF, Womack SM, Jacobson D, Yi B, Hwang YT, Gold K, Barter J, Shah K. Benefits and costs of using HPV testing to screen for cervical cancer. *JAMA* 2002;**287**: 2372-81.

9. Maissi E, Marteau TM, Hankins M, Moss S, Legood R, Gray A. The psychological impact of human papillomavirus testing in women with borderline or mildly dyskaryotic cervical smear test results: 6-month follow-up. *Br J Cancer* 2005;**92**: 990-4.

10. Goldie SJ, Kohli M, Grima D, Weinstein MC, Wright TC, Bosch FX, Franco E. Projected clinical benefits and cost-effectiveness of a human papillomavirus 16/18 vaccine. *J Natl Cancer Inst* 2004;**96**: 604-15.

11. Chuck A. Cost-effectiveness of 21 alternative cervical cancer screening strategies. *Value Health* 2010;**13**: 169-79.

12. Krahn M, McLachlin M, Pham B, Rosen B, Sander B, Grootendorst P, Tomlinson G, John-Baptiste A, Frikemerid M, Hong Chen M, Woo G, Anonychuk A, et al., Liquid-based techniques for cervical cancer screening: systematic review and cost-effectiveness analysis. Canadian Agency for Drugs and Technologies in Health, 2008.

13. 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;**30**: 1813-22.

14. Coupe VM, de Melker HE, Snijders PJ, Meijer CJ, Berkhof J. How to screen for cervical cancer after HPV16/18 vaccination in The Netherlands. *Vaccine* 2009;**27**: 5111-9.

15. de Bekker-Grob EW, de Kok IM, Bulten J, van Rosmalen J, Vedder JE, Arbyn M, Klinkhamer PJ, Siebers AG, van Ballegooijen M. Liquid-based cervical cytology using

ThinPrep technology: weighing the pros and cons in a cost-effectiveness analysis. *Cancer Causes Control* 2012;**23**: 1323-31.

16. van Ballegooijen M. Effects and Costs of Cervical Cancer Screening Rotterdam: Erasmus University Rotterdam, 1998.

17. de Kok IM, van Rosmalen J, Dillner J, Arbyn M, Sasieni P, Iftner T, van Ballegooijen M. 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.

18. van Rosmalen J, de Kok IM, van Ballegooijen M. Cost-effectiveness of cervical cancer screening: cytology versus human papillomavirus DNA testing. *BJOG* 2012;**119**: 699-709.

19. Hanmer J, Lawrence WF, Anderson JP, Kaplan RM, Fryback DG. Report of nationally representative values for the noninstitutionalized US adult population for 7 health-related quality-of-life scores. *Med Decis Making* 2006;**26**: 391-400.

20. Center for the Evaluation of value and Risk in Health. Catalog of Preference Scores: the CEA Registry. Boston, MA: Tufts New England Medical Center.

21. Kim JJ, Wright TC, Goldie SJ. Cost-effectiveness of alternative triage strategies for atypical squamous cells of undetermined significance. *JAMA* 2002;**287**: 2382-90.

22. Karnon J, Peters J, Platt J, Chilcott J, McGoogan E, Brewer N. Liquid-based cytology in cervical screening: an updated rapid and systematic review and economic analysis. *Health Technol Assess* 2004;**8**: iii, 1-78.

23. Kitchener HC, Blanks R, Cubie H, Desai M, Dunn G, Legood R, Gray A, Sadique Z, Moss S, Group MTS. MAVARIC - a comparison of automation-assisted and manual cervical screening: a randomised controlled trial. *Health Technol Assess* 2011;**15**: iii-iv, ix-xi, 1-170.

24. Institute of Medicine. *Vaccines for the 21st Century: A Tool for Decisionmaking*ed. Washington, DC: The National Academies Press, 2000. 476.

25. Sanders GD, Taira AV. Cost-effectiveness of a potential vaccine for human papillomavirus. *Emerg Infect Dis* 2003;**9**: 37-48.

26. Insinga RP, Glass AG, Rush BB. Health state transitions following an abnormal Pap smear: implications for health utility assessment in cost-effectiveness analyses 22nd International Papillomavirus Conference Vancouver, Canada, 2007.

27. H CK, Canfell K, Gilham C, Sargent A, Roberts C, Desai M, Peto J. The clinical effectiveness and cost-effectiveness of primary human papillomavirus cervical screening in England: extended follow-up of the ARTISTIC randomised trial cohort through three screening rounds. *Health Technol Assess* 2014;**18**: 1-196.

28. Elbasha EH, Dasbach EJ, Insinga RP. Model for assessing human papillomavirus vaccination strategies. *Emerg Infect Dis* 2007;**13**: 28-41.

29. Simonella LM. Characterisation and Assessment of Organised Screening, Model Validation, and Health State Preference Scores in Decision Analytic Models of Human Papillomavirus (HPV) Vaccination in Australia and New Zealand Sydney, NSW: University of Sydney, 2012.

30. Kulasingam SL, Myers ER, Lawson HW, McConnell KJ, Kerlikowske K, Melnikow J, Washington AE, Sawaya GF. Cost-effectiveness of extending cervical cancer screening intervals among women with prior normal pap tests. *Obstet Gynecol* 2006;**107**: 321-8.

31. Gold MR, Franks P, McCoy KI, Fryback DG. Toward consistency in cost-utility analyses: using national measures to create condition-specific values. *Med Care* 1998;**36**: 778-92.

32. Wolfson MC. Health-adjusted life expectancy. *Health Rep* 1996;**8**: 41-6 (Eng); 3-9 (Fre).

33. Goldie SJ, Grima D, Kohli M, Wright TC, Weinstein M, Franco E. A comprehensive natural history model of HPV infection and cervical cancer to estimate the clinical impact of a prophylactic HPV-16/18 vaccine. *Int J Cancer* 2003;**106**: 896-904.

34. Mandelblatt JS, Fahs MC. The cost-effectiveness of cervical cancer screening for low-income elderly women. *JAMA* 1988;**259**: 2409-13.

35. de Haes JC, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van Der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. *Int J Cancer* 1991;**49**: 538-44.

36. Fryback DG, Dasbach EJ, Klein R, Klein BE, Dorn N, Peterson K, Martin PA. The Beaver Dam Health Outcomes Study: initial catalog of health-state quality factors. *Med Decis Making* 1993;**13**: 89-102.

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

38. Statistics Netherlands (CBS), StatLine database. Statistic Netherlands (CBS), 2011.

39. SIG (Information Centre for Health Care), Hospital Diagnosis Statistics 1963-1985. SIG, 1985.

40. de Kok IM, van Ballegooijen M, Habbema JD. Cost-effectiveness analysis of human papillomavirus vaccination in the Netherlands. *J Natl Cancer Inst* 2009;**101**: 1083-92.

41. Jacobs MV, Walboomers JM, Snijders PJ, Voorhorst FJ, Verheijen RH, Fransen-Daalmeijer N, Meijer CJ. 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;**87**: 221-7.

42. 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**: 425-34.

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

44. 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.