Cost-effectiveness of primarily human papillomavirus–based cervical cancer screening in settings with currently established Pap screening: A systematic review commissioned by the German Federal Ministry of Health

Nikolai Mühlberger, Gaby Sroczynski, Eva Esteban
UMIT—University for Health Sciences, Medical Informatics and Technology

Thomas Mittendorf
Leibniz University Hannover

Rebecca A. Miksad
Beth Israel Deaconess Medical Center and Massachusetts General Hospital

Uwe Siebert
UMIT—University of Health, Sciences, Medical Informatics and Technology and Harvard Medical School

Objectives: The aim of this study was to summarize the current evidence for the cost-effectiveness of primarily human papillomavirus (HPV) -based cervical cancer screening in settings with already established Papanicolaou test (Pap) programs. Emphasis was placed on the German situation with annual Pap screening.

Methods: Medical, economic, and health technology assessment (HTA) databases were systematically searched for cost-effectiveness studies comparing HPV to Pap screening. Study data were extracted, standardized, and summarized in cost-effectiveness plots contrasting HPV strategies to Pap screening with 1-, 2-, 3-, and 5-years interval. For each Pap setting, the likelihood of cost-effective HPV screening was assessed depending on willingness-to-pay.
Results: We reviewed twelve decision-analytic cost-effectiveness models. Study results showed wide variation due to methodical heterogeneity. Data synthesis revealed that the cost-effectiveness of HPV screening depends on the interval of the established Pap screening strategy. In comparison with Pap screening every 2 years, only 25 percent of the HPV-based screening strategies were cost-effective. However, in comparison with Pap screening every 1, 3, or 5 years, 83 percent, 55 percent, and 92 percent of HPV screening strategies were cost-effective, respectively. Results for settings with annual Pap screening are based on models assuming 100 percent screening coverage.

Conclusions: The introduction of HPV-based screening programs is cost-effective if the screening interval of the established Pap program exceeds 2 years. In settings with biennial Pap screening, introduction of HPV-based screening is unlikely to be cost-effective. Results also suggest cost-effectiveness of HPV-based screening in settings with annual Pap screening; however, this finding should be confirmed under realistic screening adherence assumptions.

Keywords: Cervical cancer, Human papillomavirus, Screening, Cost-effectiveness, Health technology assessment

Worldwide, cervical cancer is the second most common cancer in women (2;33). In industrialized countries, the incidence of cervical cancer dropped substantially after the 1960s due to the introduction of mass screening with the Papanicolaou test (Pap) (9;19).

Pap screening can detect premalignant cytological lesions, allowing early treatment to prevent the development of invasive cervical cancer. However, cytological screening is not optimal, because a considerable number of women are still diagnosed with invasive cervical cancer despite regular participation in screening (23;30). The primary weakness of Pap screening is its poor sensitivity, caused by the test's susceptibility to sampling and reading error. Systematic reviews report that the Pap test has a 30 to 87 percent sensitivity to detect advanced precancerous lesions and cancer (15;18). German studies conducted in routine screening settings found an even lower Pap sensitivity (20–43.5 percent) (21;26). In contrast to sensitivity, the specificity of the Pap test is high (86–100 percent) (15;18). Because the sensitivity of a single Pap test is poor, a successful Pap screening program requires frequent screening rounds. However, frequent screening is expensive, burdens participants, and strongly relies on good adherence. Presently, most industrialized countries screen in intervals of 3–6 years, whereas Germany, Austria, and Luxembourg recommend annual Pap screening (10).

Nowadays, there is strong evidence that persistent infections with sexually transmitted oncogenic human papillomaviruses (HPV) are causal for the development of high-grade cervical intraepithelial neoplasia (CIN) and invasive cancer (1;17;32). Most HPV infections, especially in younger women, are transient and do not play a role in cervical cancerogenesis. However, a fraction of infections persist and may cause stepwise progression to intraepithelial neoplasia and, ultimately, to cervical cancer over a period of 10 to 30 years. The success of cervical cancer treatment is directly related to the stage of disease. Whereas survival is virtually 100 percent when treated in the preinvasive stage, the prognosis for invasive disease is poor.

Presently, approximately 30 HPV types, which are classified as no- or low-risk, intermediate-risk, and high-risk types, are known to be prevalent in the cervix. The thirteen HPV types that are most commonly screened for account for more than 90 percent of all cervical cancer cases (12). Because HPV infection marks the potential starting point of cancerogenesis, HPV testing literally allows risk detection at the earliest stage. Therefore, HPV-based screening programs are less sensitive to prolonged screening intervals and incomplete adherence patterns than Pap-based screening programs.

Clinical studies evaluating the accuracy of HPV testing have consistently reported higher sensitivity and lower specificity than for conventional cytology (3;4;6;21;22;25;26). The high sensitivity of HPV testing improves cancer prevention by increasing the detection of treatment-relevant precancerous lesions. However, the lower specificity of HPV testing increases the number of false positives, potentially reducing quality of life through unnecessary anxiety and diagnostic work-up. Higher sensitivity and lower specificity of HPV screening also have economic implications. Whereas expenditures for cancer treatment may decrease because of fewer cases, the work-up of false-positive results requires additional recourses. Apart from that finding, resources might be saved by increasing the screening interval for HPV-negative women, which would be feasible in view of the slow cancerogenesis and the highly predictive value of a negative HPV result.

The clinical and economic trade-off associated with the higher sensitivity and lower specificity of HPV testing raises questions about the cost-effectiveness of introducing HPV-based screening programs. Whether HPV screening is cost-effective or not, depends on the cost per unit of health gained by replacing the current screening standard, and the price society is willing to pay for an additional unit of health.
Therefore, countries with established Pap-based cervical cancer screening programs must consider the cost-effectiveness of HPV-based screening relative to their current Pap-based screening programs.

HPV testing can be integrated into cervical cancer screening programs in several ways. Choosing a conservative approach, HPV testing may be used in secondary screening to triage women with equivocal cytology results for further diagnostic work-up. Another approach would be using HPV testing as a primary screening test, performed as an alternative to, in parallel with, or before cytology. Several modeling studies have evaluated the cost-effectiveness of primarily HPV-based cervical cancer screening in the healthcare context of developed countries with established Pap-based screening (5;8;13;14;16;28;31). Although recent studies (8;13;14;16) suggest that HPV screening may be a cost-effective alternative to Pap screening, overall results are conflicting. Previous reviews did not systematically analyze the implication of already established cytological screening programs and reach uncertain conclusions (5;10;11;20).

Therefore, the objective of this review, which was performed as part of a health technology assessment (HTA) commissioned by the German Agency for Health Technology Assessment at DIMDI (a subsidiary of the German Federal Ministry of Health), was to specifically assess current evidence for the cost-effectiveness of primarily HPV-based cervical cancer screening in settings with already established Pap screening programs. As Pap-based screening programs primarily differ by length of the screening interval, the cost-effectiveness of HPV screening was analyzed compared with commonly used cytology intervals. The focus of this review is on the implications of the current evidence for the German healthcare context with an annual Pap screening program.

METHODS

We performed a systematic literature search of medical, economic, and health technology assessment (HTA) databases up to March 2006 to identify published cost-effectiveness assessments of primarily HPV-based cervical cancer screening in the health care context of developed countries with established Pap-based screening programs. Original cost-effectiveness studies published in English or German were retrieved for the review. We excluded studies that (i) did not present a full economic evaluation; (ii) did not evaluate HPV testing as a primary screening test; (iii) evaluated HPV testing in resource-poor settings; and (iv) assessed HPV-testing in selected populations.

The methodological quality, study characteristics, and results of the included cost-effectiveness studies were assessed using a standardized format (29). Reported average lifetime cost and health-effect data for primarily HPV-based and conventional Pap-based screening strategies were extracted. Data for strategies involving liquid-based cytology (LBC) or HPV triage of atypical cells of undetermined significance (ASCUS) were not evaluated in this review.

Extracted data were processed to reach a consistent database for the review. First, we converted all cost and effect estimates to increments compared with no screening, because some studies reported only lifetime costs and effects incremental to no screening. Second, we converted cost data to year 2005 Euros. Non-Euro currencies were first converted to Euros using gross domestic product purchasing power parities (GDP-PPPs; Source: Eurostat http://epp.eurostat.ec.europa.eu/), and then inflated to year 2005 Euros using the German consumer price index (CPI; Source: German Federal Statistical Office http://www.destatis.de/).

Using the processed data, we recalculated the incremental cost-effectiveness ratios (ICER) of the original studies, which was necessary due to the exclusion of HPV-triage and LBC strategies in our review. The ICER is the cost per unit of health gained by a strategy that is in excess to the units of health produced by the next less expensive and economically rational strategy. Economically irrational strategies, that are either more costly and less effective than others (dominance), or yield additional health at higher costs than more effective strategies (extended dominance) were identified and excluded from the ICER calculations.

To visually explore the extent of variation in study results, we summarized all results into a single cost-effectiveness plot displaying the incremental costs (y-axis) and incremental effectiveness (x-axis) of each single screening strategy in comparison with no screening (reference strategy). The ICER of each strategy relative to the reference strategy is equal to the slope of the line connecting a strategy with the origin. Strategies positioned on a cost-effectiveness plane can be compared with a line with a slope representing the societal willingness-to-pay threshold. Strategies with an ICER below a chosen threshold demonstrate acceptable trade-offs between costs and gains in health units. For descriptive reasons, our plots display a willingness-to-pay (WTP) threshold of 50,000 Euro per life-year or quality-adjusted life-year (QALY) gained, which in view of already established medical technologies can be considered a well-accepted benchmark for cost-effectiveness in developed countries (7).

To specifically analyze the cost-effectiveness of primarily HPV-based cervical cancer screening in settings with already established Pap screening programs, we performed additional cost-effectiveness analyses comparing HPV-based screening strategies directly versus Pap-based screening programs with 1-, 2-, 3-, and 5-year intervals. Pap-interval specific analyses were solely based on studies specifically evaluating the respective Pap interval. Results were summarized in four cost-effectiveness plots displaying cost and effectiveness differences to the compared Pap reference. Each plot is divided into four quadrants by the crossing cost and effectiveness axes. HPV strategies located in the lower right quadrant are more effective and less expensive than the Pap
comparator and, therefore, cost-effective, while strategies located in the upper left quadrant are less effective and more expensive and, therefore, not cost-effective. Strategies located in the remaining two quadrants are cost-effective if they fall below the willingness-to-pay threshold.

In the final step of information synthesis, we contrasted the proportions of cost-effective HPV strategies yielded by the Pap interval-specific cost-effectiveness plots, which reflect the certainty that the introduction of HPV screening is cost-effective in the respective setting with established Pap screening. Proportions of cost-effective HPV screening strategies were calculated for willingness-to-pay thresholds varying from zero to infinity. Results of this analysis were summarized in a stratified line-plot, displaying the willingness-to-pay dependent proportion of cost-effective HPV screening strategies derived from comparison with 1-, 2-, 3-, and 5-yearly Pap screening.

RESULTS

Identified Studies

Our literature search identified 322 publications mentioning HPV testing in conjunction with economic keywords. A total of seven publications comprising six journal articles (8;13;14;16;28;31) and one HTA report (5) met the inclusion criteria for the review. All seven publications described decision-analytic studies modeling the long-term costs and effects of primary HPV and Pap screening in developed countries. Six of these studies (5;8;13;16;28;31) were cost-effectiveness analyses that expressed cost-effectiveness in terms of cost per life-year gained. One study (14) was a cost-utility analysis, expressing cost-effectiveness in terms of cost per QALY gained.

Four studies presented models for the healthcare context of individual countries, whereas the study by Kim et al. (13) included four models for different European countries. Cuzick et al. (5) and Van Ballegooijen et al. (31) each presented two models: one favorable (Model A) and one unfavorable (Model B) for HPV testing, to account for the present uncertainty about the true HPV attack and remission rates. Altogether, the included studies described twelve different decision-analytic cost-effectiveness models. Each of these was regarded as a separate entity in our analysis. Only one model by Mittendorf et al. (16) evaluated the cost-effectiveness of HPV screening in the German healthcare context.

Study Characteristics

Supplementary Table 1 (which can be viewed online at http://www.journals.cambridge.org/jid_thc) summarizes important characteristics of the included decision-analytic studies. Studies differed in many respects. First of all, the investigated screening strategies differ in respect of the applied tests or test combinations, screening intervals, and target ages. Eleven of the twelve models evaluated parallel combination of Pap and HPV testing (Pap+HPV), six models evaluated HPV testing alone (HPV), and one model evaluated sequential combination of HPV and Pap testing (HPV/Pap). HPV screening intervals ranged from 1 to 10 years. However, only three of the models evaluated intervals shorter than 3 years. Additional heterogeneity arises from different management of positive screening results, precancerous lesions, and cancer in the various settings. In addition to strategy-related variations, the models used different methodological approaches. Methodological variations occurred in general topics (type of model, analytic time horizon, and the perspective of the analysis) as well as in specific issues such as cost calculations, valuation of lost or gained health (i.e., consideration of quality of life, included costs components, correction for inflation, and the use of discounting), and with regard to parameter estimates (HPV incidence, regression rates, sensitivity and specificity of HPV and Pap testing, and screening compliance). In particular, the methodology applied in the German model differed from other models. Constructed as a decision tree model, the German model may not be completely adequate in modeling repetitive events (HPV infection) or time-dependent events (regression of HPV infections, progressive cancerogenesis and age-specific mortality). Furthermore, this model did not specifically account for competing background mortality and assumes that women will not experience another HPV infection after remission of a precancerous lesion. Finally, the German analysis did not include costs for medication and chemotherapy, and only costs (not health effects) were discounted.

Heterogeneity of Results

Supplementary Table 2 (which can be viewed online at http://www.journals.cambridge.org/jid_thc) presents the results of the cost-effectiveness analyses performed with homogeneously processed data derived from the twelve models. Figure 1 summarizes the findings. Results of the twelve models are widely scattered over the cost-effectiveness plot, even for models assessing similar screening strategies. In particular, the German model deviates from others. Because it yields approximately ten times higher gains in life expectancy due to screening than other models, its results are located far outside the cost-effectiveness plane. In contrast to the German analysis, the results of the only cost-utility analysis are located within the range of the other studies. In view of the large intermodel variability, we refrained from a strictly quantitative synthesis of the modeling results.

Cost-effectiveness Versus Pap Screening

As shown in Figure 1, all screening Pap strategies have an ICER below 50,000 Euro per life-year or QALY gained if compared with no screening. Figure 2 describes the cost-effectiveness of HPV screening in comparison to Pap screening with 1-, 2-,...
Figure 1. Cost-Effectiveness plot summarizing the results of decision analytic studies modeling primarily human papillomavirus (HPV) -based screening in developed countries. Shorthand notation indicates the applied test or test combination (Test + Test = parallel testing, Test/Test = sequential testing) and the screening interval (e.g., 3y = every 3 years). Symbol shape specifies screening procedure, color specifies interval length (see key). QALY, quality-adjusted life-year; LY, life-year; EUR, euro.

3-, and 5-year intervals. None of the twelve reviewed models evaluated all four Pap screening intervals.

Annual Pap screening (Figure 2A) was evaluated by only two models: the German model (16) and a U.S. model (8). Of the six HPV strategies compared with annual Pap screening by the same models, five (83 percent) are located in the cost-effective area below the willingness-to-pay threshold. As shown by the plot, five strategies with screening intervals ranging from 2 to 10 years were cost-effective compared with annual Pap screening primarily due to lower costs. Only one strategy is located outside the cost-effective area of the plot, which is annual testing with both Pap and HPV.

Biennial Pap screening (Figure 2B) was evaluated by two U.S. models (8;14), including the cost-utility model. HPV screening strategies evaluated by these models enabled eight comparisons. Two HPV strategies were evaluated by both models. Six (75 percent) of the eight HPV strategies fall above the willingness-to-pay threshold, mostly due to higher costs than biennial Pap screening. Only one data point, representing triennial combined Pap and HPV testing, is located in a cost-effective area of the cost-effectiveness plane. However, as one data point of the same strategy can also be found in the least favorable upper left quadrant, the underlying models yielded contradictory results.

Triennial Pap screening (Figure 2C) was evaluated by all, but the German (16) and Dutch (13) models. Evaluated HPV screening strategies permitted twenty-nine comparisons. Five of the HPV strategies were evaluated by more than one model. Sixteen (55 percent) of the twenty-nine data points are located in a cost-effective area. In contrast to the previous plots, a shift of data points to the right of the cost-effectiveness plane can be observed, which indicates increasing incremental effectiveness of HPV screening.

Pap screening every 5 years (Figure 2D) was evaluated by four models (5;13;27). Twelve comparisons with various HPV screening strategies were computable. Eleven (92 percent) data points are located in a cost-effective area. Compared with the preceding plot, the previously seen shift of data points to the right of the cost-effectiveness plane is even more pronounced.

Figure 3 displays the proportion of HPV-based screening algorithms that are cost-effective when compared with Pap screening every 1, 2, 3, and 5 years under various willingness-to-pay assumptions. As shown in the graph, HPV strategies analyzed by decision-analytic models were least frequently cost-effective in comparison with 2-yearly Pap screening, while comparisons with 1-, 3-, and 5-yearly Pap screening more frequently yielded cost-effective results. For example, at a willingness-to-pay of 50,000 EUR per life-year or QALY gained, 83 percent, 25 percent, 55 percent, and 92 percent of the analyzed HPV strategies were found to be cost-effective in comparison with 1-, 2-, 3-, and 5-yearly Pap screening, respectively. This “U”-shaped distribution suggests that, in a setting with established Pap screening, the cost-effectiveness of HPV screening strongly depends on the interval of the established Pap program.
DISCUSSION

The objective of this review was to summarize current evidence for the cost-effectiveness of introducing primarily HPV-based cervical cancer screening programs into settings with existing Pap-based screening programs. In particular, we focused on the situation with annual Pap screening, which is relevant in the German healthcare context.

We reviewed and further evaluated twelve decision-analytic cost-effectiveness models. The results illustrate that the incremental cost-effectiveness of HPV-based screening strongly depends on the screening interval of the existing Pap-based screening program. In comparisons with annual Pap screening, HPV-based strategies with less frequent screening rounds (≥ 2 years) were equally effective but less costly. This finding indicates that the introduction of HPV screening in settings with annual Pap screening might be cost-effective primarily due to cost savings. However, the proportion of cost-effective HPV strategies detected in comparison with different Pap screening intervals showed a “U”-shaped distribution. In comparison with biennial Pap screening, HPV-based screening no longer yields a clear cost advantage, and health gains are still marginal, which results in an unfavorable ICER. Only when Pap screening is carried out less frequently than biennial, its effectiveness is reduced to a degree that HPV screening becomes cost-effective again—in this case, primarily due to superior effectiveness. That most HPV strategies compared with triennial Pap screening had a similar or even longer screening interval indicates that HPV screening is less sensitive to extensions of the screening interval than Pap screening.

Our review suggests that HPV screening with a screening interval greater than 1 year is a cost-effective alternative to annual Pap screening. However, this finding must be interpreted in the context of this study. Our results for settings with annual Pap screening are based on only two models, of which one deviates from common methodological standards. A second limitation is that both models assume a 100 percent screening coverage and compliance, which means that every woman in the target age group is screened every year. In
Figure 3. Percentage of human papillomavirus (HPV)-based screening programs that are cost-effective when compared to Pap screening every 1, 2, 3, and 5 years under various willingness-to-pay assumptions. QALY, quality-adjusted life-year; EUR, euro.

reality, screening adherence is incomplete and varies among women. Hence, the average screening interval is greater than one year and the variation is likely to be skewed toward longer intervals. Adherence to annual screening in Germany has been reported to be approximately 50 percent (24). Because the distribution of individual screening intervals is unknown, the impact on the cost-effectiveness of HPV-based screening is unclear. In an extreme scenario, a 50 percent adherence rate could mean that all women are uniformly screened every other year. Because this case is equivalent to 100 percent compliance with a biennial Pap-screening program HPV screening would have to be evaluated versus biennial Pap screening, which, as shown before, increases its ICER. A more realistic scenario of 50 percent adherence to annual Pap smears allows for heterogeneous screening behavior, where the majority of women are screened every year, while a minority are screened at longer intervals that may regularly exceed 2 years. This scenario argues in favor of HPV-based screening because, in addition to the acceptable ICER of HPV-based screening compared with annual Pap smears, women attending screening less frequently than every 2 years have been shown to gain from HPV testing at acceptable costs.

Because the cost-effectiveness of HPV screening in settings with annual Pap screening strongly depends on adherence patterns, cost-effectiveness evaluations of HPV screening versus annual Pap screening require a more thorough consideration of coverage aspects than evaluations versus less frequent Pap programs. In settings with less frequent than annual screening, incomplete coverage also causes a prolongation of the screening interval. However, as the incremental effectiveness of HPV screening improves with an extension of the screening interval beyond 2 years, this would only further increase the likelihood that HPV screening is cost-effective in contrast to Pap-based screening.

Our systematic review differs from prior publications in several ways. Previous reviews did not systematically assess the cost-effectiveness of primarily HPV-based screening in comparison with currently established Pap screening programs (5;10;11;20). In addition, the implications of the currently available evidence for countries like Germany with annual Pap-based screening programs have not been investigated previously. Whereas most prior reviews were descriptive, Holmes et al. (11) reviewed recent health economic modeling studies and summarized the cost-effectiveness of HPV testing in comparison with no screening. However, most developed countries already have a Pap-based screening program in place and, therefore, need to be informed about the incremental cost-effectiveness of introducing HPV-based screening programs in settings with existing Pap-based screening programs. Our study specifically addresses this issue by evaluating HPV-based screening programs in the context of various existing Pap-based programs.

Our study has several limitations. A strictly quantitative synthesis of the results from the different modeling studies was not feasible due to considerable methodological heterogeneity. In particular, the model most relevant for the German healthcare context (16) differed methodologically from other models, which weakens conclusions drawn from the study and hampers comparison with other models. Unfortunately
not all models assessed the same HPV-based and Pap-based screening strategies. Therefore, the number of models available for comparisons of individual Pap-based and HPV-based strategies varied considerably. Whereas only two models contributed data for the comparison of HPV screening with annual or biennial Pap screening, ten contributed data for the comparison with Pap-screening every 3 years, and four models contributed data for the comparison with Pap screening every 5 years. Conclusions drawn for the context of triennial Pap screening may, therefore, be more reliable than those drawn for the context with annual Pap screening. Finally, only the German model evaluated HPV testing as a replacement of annual Pap testing. Information relevant for the introduction of HPV testing as a sole screening test in those settings is, therefore, even less reliable.

POLICY IMPLICATIONS

Based on our review, the introduction of primarily HPV-based screening is cost-effective in settings with Pap-based screening programs with a screening interval longer than 2 years. In settings with biennial Pap screening, HPV screening is unlikely to be cost-effective. For settings with annual Pap screening, our analysis indicates cost-effectiveness of HPV screening. However, this result is affected by the unrealistic assumption of complete screening coverage. Further modeling incorporating more realistic screening adherence scenarios is recommended to confirm the finding.

CONTACT INFORMATION

Nikolai Mühlberger, DVM, MPH (nikolai.muehberger@umit.at), Senior Scientist; Deputy Director of EURIDICE—European Infectious Disease Research Center, Department of Public Health, Medical Decision Making and Health Technology Assessment, Gaby Sroczynski, MPH (gaby.sroczynski@umit.at), Senior Scientist, Department of Public Health, Medical Decision Making and Health Technology Assessment; Director of the Program on Decision Analysis, Department of Public Health, Medical Decision Making and Health Technology Assessment, Eva Esteban, MPH (evaesteban@web.de), Research Scientist, Department of Public Health, Medical Decision Making and Health Technology Assessment, Research Scientist, Department of Public Health, Medical Decision Making and Health Technology Assessment, UMIT—University for Health Sciences, Medical Informatics and Technology, EWZ I, A-6060 Hall i.T., Austria; Associate Professor of Radiology (Harvard University), Institute for Technology Assessment and Dept. of Radiology, Harvard Medical School, 101 Merrimac Street, Boston, Massachusetts, 02114

Uwe Siebert, MD, MPH, MSc, ScD (public-health.at), Chair, Professor of Public Health (UMIT), Department of Public Health, Medical Decision Making and Health Technology Assessment, UMIT—University of Health, Sciences, Medical Informatics and Technology, Eduard Wallnoefer Center I, A-6060 Hall i.T., Austria; Associate Professor of Radiology (Harvard University), Institute for Technology Assessment and Dept. of Radiology, Harvard Medical School, 101 Merrimac Street, Boston, Massachusetts, 02114

REFERENCES