Modeling Human Papillomavirus Vaccine Effectiveness: Quantifying the Impact of Parameter Uncertainty

Nicolas Van de Velde¹, Marc Brisson¹,², and Marie-Claude Boily¹,³

¹ Department of Social and Preventive Medicine, Laval University, Quebec, Canada.
² Health Economics and Outcomes Research, Merck Frosst Canada Ltd., Montreal, Canada.
³ Department of Infectious Disease Epidemiology, Imperial College, London, United Kingdom.

Received for publication June 7, 2006; accepted for publication September 15, 2006.

The development of models is based on assumptions, which inevitably embed a level of uncertainty. Quantifying such uncertainty is particularly important when modeling human papillomavirus (HPV) vaccine effectiveness; the natural history of infection and disease is complex, and age- and type-specific data remain scarce and incomplete. The aim of this study was to predict the impact of HPV-6/11/16/18 vaccination, using a cohort model and measuring parameter uncertainty. An extensive fitting procedure was conducted, which identified 164 posterior parameter combinations (out of 200,000 prior parameter sets) that fit simultaneously HPV type-specific incidence and prevalence data for infection, cervical intraepithelial neoplasia (CIN), and squamous cell carcinoma (SCC). Results based on these posterior parameter sets suggest that vaccinating girls aged 12 years (vaccine efficacy = 95%, no waning) would reduce their lifetime risk of HPV infection, CIN1, CIN2/3, and SCC by 21% (80% credibility interval: 17, 29), 24% (80% credibility interval: 17, 31), 49% (80% credibility interval: 36, 60), and 61% (80% credibility interval: 47, 73), respectively. If vaccine efficacy is reduced or vaccine protection is assumed to wane, uncertainty surrounding predictions widens considerably. Important priorities for future research are to understand the role of natural immunity and to measure the duration of vaccine protection because results were most sensitive to these parameters.

cancer vaccines; computer simulation; models, theoretical; papillomavirus infections; papillomavirus vaccines; uncertainty; uterine cervical neoplasms; vaccination

Abbreviations: CIN, cervical intraepithelial neoplasia; HPV, human papillomavirus; HR, other high oncogenic risk types; SCC, squamous cell carcinoma.

Infection with human papillomavirus (HPV) is a necessary cause of cervical cancer (1–5), which is the second most common cancer in women worldwide (6). HPV also contributes to the development of a number of other cancers, including anogenital carcinomas (vulva, vaginal, anus, penile) (7–16), and head and neck cancers (17–23). Genital warts are also associated with HPV (14, 24, 25), with an estimated lifetime risk close to 10 percent (26). More than 40 HPV genotypes are known to infect the anogenital tract (27). On the basis of their epidemiologic associations with invasive cervical cancer, the International Agency for Research on Cancer classified these types as high oncogenic risk types or low oncogenic risk (3). Oncogenic types HPV-16 and -18 are the most common cause of anogenital carcinomas and account for approximately 70 percent of all cervical cancers (3, 28, 29). Low-oncogenic-risk HPV-6/11 are responsible for approximately 90 percent of genital warts (24, 25).

Two HPV prophylactic vaccines are currently in clinical trials, which target HPV-6/11/16/18 (Gardasil; Merck & Co., Inc., Whitehouse Station, New Jersey) and HPV-16/18...
FIGURE 1. Model flow diagram. Left: natural history of squamous cell carcinoma (SCC). The different compartments represent individuals in each mutually exclusive state of human papillomavirus (HPV) infection and disease. The arrows represent the flow between these states (solid lines represent progression, dashed lines clearance and regression, dotted lines flow due to screening and treatment). Right: infection and natural immunity. The different compartments represent the underlying pattern of infection within each disease state. Solid lines represent acquisition of another HPV type, dashed lines clearance of an HPV type, and dotted lines clearance with a probability of developing natural immunity. Women have a certain probability of developing lifelong immunity following natural infection with HPV-16 and HPV-18 (but not other high oncogenic risk types (HR) or low oncogenic risk (LR)). The model allows for dual coinfections with HPV-16 and HPV-18 (but not other high oncogenic risk types (HR) or low oncogenic risk (LR)). The model allows for dual coinfections with HPV-16 and HPV-18. CIN, cervical intraepithelial neoplasia.

Despite considerable uncertainty surrounding type-specific natural history parameters, previous modeling studies have derived predictions regarding the impact of HPV vaccination by using only a base-case parameter set (40–47). In these studies, one-way sensitivity analyses were performed, which are limited in scope because they do not take into account the conjoint uncertainty of all natural history parameters.

On the basis of these considerations, the aim of our study was to 1) develop and parameterize an HPV natural history model, 2) estimate the potential effectiveness of HPV vaccination, and 3) assess the sensitivity of model predictions to parameter uncertainty.

MATERIALS AND METHODS

Model structure

We developed a model of the natural history of HPV infection and squamous cell carcinoma (SCC) in a cohort of women followed over their lifetimes. The natural history of HPV is represented by six mutually exclusive disease states: susceptible, immune, infected, CIN1, CIN2/3, and SCC. Women can transition between these states, as described in figure 1, left. Transition probabilities between disease states are allowed to vary according to HPV genotype and age. We do not model adenocarcinoma because its natural history is different from that of SCC (48).

The model includes four classes of HPV genotypes: HPV-16, HPV-18, other high-oncogenic risk types (HR), and low-risk oncogenic types (49–51). The model incorporates the following dual coinfections—16–18, 16–HR, 16–low oncogenic risk, 18–HR, and 18–low oncogenic risk (figure 1, right)—because they represent the large majority of multiple concurrent coinfections (3, 49–56). Coinfected women will progress and regress through the HPV disease states according to the transition rates of the most aggressive type.
Following clearance from HPV-16 or -18, women have a probability of developing type-specific lifelong natural immunity (figure 1, right). However, we assume that infection from HR or low oncogenic risk types does not result in lifelong natural immunity since these groups comprise many genotypes and have a small probability of reexposure to the same genotype.

The model also accounts for screening and treatment outcomes. Women have an age-specific probability of being screened and a lesion-specific test sensitivity of being detected. Following treatment, women can remain infected with the genotype, clear the infection, or stop being susceptible to HPV infection and disease because of hysterectomy or ablative treatment (57–66). Women also face age-specific mortality and hysterectomy for conditions unrelated to HPV.

Parameter values and model fit

Parameter values and references are summarized in tables 1 and 2. All natural history parameters are type and age specific. Parameterization and model validation were performed by using the following five-step process:

- **Step 1.** Setting prior distributions from prospective studies: A comprehensive search of prospective HPV studies published between 1995 and 2005 was conducted. Each type-specific natural history parameter was assigned a prior uniform distribution between the minimum and maximum estimates found in the literature review. With the exception of the type-specific force of HPV infection and progression rate from CIN2/3 to SCC (appendix 1), the prior distributions are not age specific because the important heterogeneity or the lack of information in the literature did not enable us to make any assumptions regarding age-specific patterns. Refer to table 2 for prior parameter ranges.
  - **Step 2.** Sampling parameter sets: A total of 200,000 parameter sets were drawn from the uniform prior parameter distributions by using Latin Hypercube Sampling (67).
  - **Step 3.** Identifying epidemiologic data for parameter fit and validation: North American–specific epidemiologic data were identified, from a comprehensive review of articles and reports, for model fit and cross-validation. The literature review included studies published between 1995 and 2005. Studies among populations at “high risk” of HPV infection were excluded. Whenever possible, data specific to Canada were used. Appendix 2 lists epidemiologic outcomes and references.
  - **Step 4.** Fitting the model and identifying posterior parameter sets: Parameter sets from step 2 were judged to produce “acceptable” fit if the associated model predictions were simultaneously within prespecified targets (ranges) defined by using the epidemiologic data from step 3. Refer to appendix 3 for a description of how targets were chosen.
  - **Step 5.** Cross-validating model fit: To assess the predictive value of the model by using the posterior (“acceptable”) parameter sets, we compared model predictions for the cross-validation outcomes (those not used in step 4) with observed epidemiologic data from step 3 (appendix 2).

Model simulations

Model predictions of the natural history of HPV and the impact of HPV vaccination are based on the posterior parameter sets identified during model fitting. The uncertainty

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Reference number(s)</th>
<th>All</th>
<th>15–19</th>
<th>20–29</th>
<th>30–39</th>
<th>40–49</th>
<th>50–59</th>
<th>60–69</th>
<th>70–100</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of women in the cohort</td>
<td>100,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at entry (years)</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural mortality rate (per 1,000 woman-years)</td>
<td>0.003</td>
<td>0.004</td>
<td>0.006</td>
<td>0.016</td>
<td>0.040</td>
<td>0.101</td>
<td>1.071</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Non-HPV* hysterectomy rate (per 1,000 woman-years)</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>4.85</td>
<td>10.6</td>
<td>1.7</td>
<td>1.3</td>
<td>70</td>
<td></td>
</tr>
</tbody>
</table>

* HPV, human papillomavirus; CIN, cervical intraepithelial neoplasia.
† Assumption because no prior distribution was found in the literature.
### TABLE 2. Natural history parameters

| Natural history parameter | Prior parameter ranges | Posterior parameter ranges | posterior/prior
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>16†</td>
<td>18†</td>
<td>HR†,§</td>
</tr>
<tr>
<td>Infection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Force of infection (per woman-year)</td>
<td>0.000–0.139</td>
<td>0.000–0.042</td>
<td>0.000–0.278</td>
</tr>
<tr>
<td>Clearance infection (per woman-year)</td>
<td>0.08–1.54</td>
<td>0.08–1.71</td>
<td>0.33–2.06</td>
</tr>
<tr>
<td>Natural immunity following infection (%)</td>
<td>0–100</td>
<td>0–100</td>
<td>NA§</td>
</tr>
<tr>
<td>CIN§ (per woman-year)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Progression of HPV§ to CIN1</td>
<td>0.01–0.18</td>
<td>0.01–0.18</td>
<td>0.02–0.18</td>
</tr>
<tr>
<td>Progression of HPV to CIN2/3</td>
<td>0.001–0.205</td>
<td>0.001–0.205</td>
<td>0.001–0.102</td>
</tr>
<tr>
<td>Progression of CIN1 to CIN2/3</td>
<td>0.00–0.39</td>
<td>0.00–0.39</td>
<td>0.00–0.39</td>
</tr>
<tr>
<td>Clearance of CIN1</td>
<td>0.06–1.75</td>
<td>0.06–1.75</td>
<td>0.22–2.62</td>
</tr>
<tr>
<td>Clearance of CIN2/3</td>
<td>0.04–1.15</td>
<td>0.04–1.15</td>
<td>0.15–1.73</td>
</tr>
<tr>
<td>Regression of CIN1 to HPV</td>
<td>0.06–1.00</td>
<td>0.06–1.00</td>
<td>0.06–1.00</td>
</tr>
<tr>
<td>Regression of CIN2/3 to HPV</td>
<td>0.00–1.00</td>
<td>0.00–1.00</td>
<td>0.00–1.00</td>
</tr>
<tr>
<td>Regression of CIN2/3 to CIN1</td>
<td>0.00–0.30</td>
<td>0.00–0.30</td>
<td>0.00–0.30</td>
</tr>
</tbody>
</table>

* The natural history parameters are annual transition rates.
† Fraction of the prior parameter range occupied by the 80% credibility interval of the posterior parameter range. A nonzero ratio is a measure of residual nonuniqueness.
‡ Minimum and maximum of the age-specific prior parameter ranges. Age groups are 13–14, 15–19, 20–24, 25–34, 35–49, 50–79, and ≥80 years.
§ HR, other high oncogenic risk types; LR, low oncogenic risk; NA, not available; CIN, cervical intraepithelial neoplasia; HPV, human papillomavirus.
¶ Minimum 10th and maximum 90th percentiles of the age-specific posterior parameter ranges.
# Parameter values were not available from the literature review.
of model predictions is presented as the median, 10th and 90th percentiles of results from the posterior parameter sets, which we call the 80 percent credibility intervals. Impact of vaccination was assessed under current screening algorithms in Canada (68).

**Vaccination strategies**

The base-case vaccine strategy assumes vaccination of girls aged 12 years. Sensitivity analysis is performed to evaluate the impact of age at vaccination and booster scenarios.

**Vaccine characteristics**

Base-case vaccine characteristics are assumed to be as follows: 1) The proportion of females protected following vaccination (take) is 100 percent, 2) vaccine duration is lifelong, and 3) reduction in susceptibility to HPV-6/11/16/18 (vaccine efficacy) is 95 percent. A sensitivity analysis is performed to explore the impact of vaccine parameters on model predictions. Vaccine duration is varied by assuming a constant waning of vaccine protection, resulting in an exponential decay of efficacy for the protected population.

**RESULTS**

**Model fit and validation**

Of 200,000 different combinations of parameters sampled from the prior parameter distributions, 164 parameter sets produced model results within the prespecified targets. The model predictions based on the 164 parameter sets produced good visual fit to the age- and type-specific epidemiologic data listed in appendix 2, which include overall HR HPV prevalence, overall SCC incidence, and the proportion of SCC attributable to HPV-16 and -18 (figure 2, upper left, upper right, lower left). In addition, model predictions for...
outcomes not used in the fitting procedure (cross-validation outcomes) were also consistent with their corresponding epidemiologic data (figure 2, lower right).

Parameter uncertainty

The posterior parameter sets that best fit the observed epidemiologic data cover a wide range of values. Figure 3 illustrates the age-specific prior and posterior ranges for HPV-16 force of infection. Here, and for the other age-specific parameters, the range of values of the posterior parameter sets is only 40–50 percent narrower than the wide ranges given to the prior distributions (table 2), suggesting that the natural history of HPV and SCC can be fitted by using a broad range of parameter combinations that seem equally probable. The parameter with the least amount of information gained from the fitting procedure and with the greatest uncertainty is natural immunity (percentage of individuals who acquire natural lifelong immunity) because its posterior range covers most of the prior distribution.

Base-case vaccine scenario

Under our base-case vaccine assumptions, the model predicts that vaccinating girls aged 12 years against HPV-6/11/16/18 would reduce their overall lifetime risk of HPV infection, CIN1, CIN2/3, and SCC by 21 percent (80 percent credibility interval: 17, 29), 24 percent (80 percent credibility interval: 17, 31), 49 percent (80 percent credibility interval: 36, 60), and 61 percent (80 percent credibility interval: 47, 73), respectively (figure 4, left). The proportions of these lifetime risk reductions achieved before 30 years of age for HPV infection, CIN1, CIN2/3, and SCC are, respectively, 40 percent, 44 percent, 55 percent, and 20 percent. As expected, the HPV-16 component of the vaccine contributes most to the overall reduction in the burden of HPV disease (figure 4, right). Although the HPV-6/11 component of the vaccine is primarily thought of for its potential impact on genital warts, these genotypes also contribute significantly to the reduction in HPV infection and CIN1 (figure 4, right). Of note, the model predicts an increase in the number of CINs and SCCs attributable to the HR genotypes not included in the vaccine. This increase occurs mainly because, in the model, we assume that women with coinfection progress and regress according to the transition rates of the most aggressive HPV genotype, and disease is attributed to this type. Hence, by preventing disease due to HPV-6/11/16/18, vaccination allows greater opportunity for HR genotypes to cause CIN and SCC even if a woman’s risk of infection remains unchanged (figure 4, right). Since multiple coinfections are allowed and cross-protection is not modeled, the model predicts that vaccination will have very little impact on the incidence of HPV infection from HR (figure 4, right).
Impact of vaccine characteristics and vaccination scenarios

The model shows that small increases in vaccine efficacy can result in a disproportionately higher reduction in disease (figure 5, upper left). The source of this amplification occurs at the level of genotype-specific infection. For example, a 5 percent increase (from 95 percent to 100 percent) in HPV-16 vaccine efficacy increases the lifetime reduction in HPV-16–related infection by 14 percent. This effect occurs only if clearance of infection confers natural immunity, resulting in depletion of susceptible women who would otherwise become infected following successive exposure to the same genotype. In a model that assumes no natural immunity, vaccine efficacy will be equal to effectiveness.

Model predictions are particularly sensitive to the duration of vaccine protection (figure 5, upper right). A vaccine with an average duration of 30 years and 95 percent efficacy given to girls aged 12 years would reduce their lifetime risk of SCC by 6 percent compared with 61 percent, if lifelong protection is assumed. This effect is a consequence of shifts in age at infection, and the magnitude of the difference is mainly due to the level of natural immunity and the age-specific progression/regression of disease. A vaccine with waning immunity prevents natural infection at early stages of sexual activity and, if natural immunity occurs, moves the pool of susceptibles toward older ages. If the force of HPV infection remains high among older women and/or progression rates toward SCC are higher in older women, then waning of vaccine protection can have an important impact on vaccine effectiveness. On the basis of our best-fit parameter sets, the force of infection remains high, the progression from CIN2/3 to SCC is higher in age groups older than 30 years, and lifelong natural immunity occurs. As a
consequence, with an average vaccine duration of 30 years, some of the SCC that would have occurred in younger age groups is temporarily prevented and delayed to when women reach age 45 years or older (figure 5, upper right). If the HPV vaccine has significant waning, booster shots are likely to be needed to maintain and lengthen vaccine-induced immunity. The impact of additional booster doses to vaccination of girls aged 12 years is shown in figure 5, upper right.

Under base-case vaccine characteristics, results suggest that vaccination of older age groups can still be highly effective in reducing HPV infection, CIN, and SCC (figure 5, lower left). In fact, vaccinating women aged 30 years can reduce SCC by 49 percent (80 percent credibility interval: 31, 65) because of the large fraction of women still susceptible at this age (e.g., 54 percent are estimated to be susceptible to HPV-16).

Figure 5 shows that the variability in the predicted lifetime reduction in HPV infection and disease depends on vaccine characteristics and the vaccination scenarios investigated. The uncertainty is smaller when vaccine characteristics such as vaccine efficacy (figure 5, lower right) and duration of protection are high; when vaccine efficacy is high and lifelong protection is assumed, age-specific associations between natural history parameter values have less opportunity to affect results. In such cases, the variability around predictions mostly corresponds to the variability of the epidemiologic data used for model fit.

Impact of natural immunity

The most uncertain natural history model parameter is natural immunity (table 2). Figure 6 shows that vaccine effectiveness increases as the probability of developing natural immunity decreases. Most of the difference is explained by women over 50 years of age who benefit from natural protection acquired at younger ages, when the risk of infection is high but clearance of infection is more likely.

DISCUSSION

Model results suggest that, for a wide range of ages at vaccination, the overall incidence of HPV infection, CIN, and SCC, and, by association, repeat Pap test, smears, treatment, and mortality, are likely to be significantly reduced by vaccination against HPV. The overall effectiveness of HPV immunization will depend on the characteristics of the vaccine, the type of vaccination strategy, and the age-specific natural history of SCC.

Results from large, randomized controlled trials have shown that the efficacy of HPV vaccines is close to 100 percent at preventing type-specific CIN (30–34). When 100 percent vaccine efficacy against HPV-6/11/16/18 and lifelong vaccine protection are assumed, the model estimates that vaccinating girls aged 12 years would reduce their overall lifetime risk of SCC by 72 percent (80 percent credibility interval: 58, 82). This reduction is in line with previously published estimates from Goldie et al. (42) and Taira et al. (46), who reported that cervical cancer would be reduced by 62 percent and 66 percent, respectively. Models that assume HPV vaccine to have close to perfect characteristics (or close to 100 percent vaccine type-specific effectiveness) are expected to produce similar results because they are insensitive to natural history assumptions. In such cases, model results are mostly dependent on the incidence and type distribution to which the model has been calibrated. However, when vaccine efficacy is reduced or vaccine protection is assumed to wane over time, then predictions regarding the impact of HPV vaccination become highly sensitive to natural history assumptions, and the uncertainty surrounding vaccine effectiveness predictions widens considerably. It is therefore important when examining the potential impact of vaccine characteristics on effectiveness to do so by varying the natural history of HPV to illustrate the level of confidence one can have in model predictions.

Of the different vaccine characteristics, waning immunity has the greatest impact on effectiveness and generates high levels of uncertainty surrounding model predictions, mostly as a result of three factors. First, vaccine effectiveness is reduced when natural immunity is longer than vaccine-induced immunity. Currently, little is known about the level of natural immunity following clearance of infection. Only a few of the posterior parameter combinations included low probabilities of natural immunity following HPV infection, which suggests that some degree of type-specific immunity occurs. Second, vaccine effectiveness is reduced if progression from CIN2/3 to cancer increases with age. Again, no epidemiologic evidence exists regarding this factor, mostly because of ethical reasons. Third, the effectiveness is lower if the force of infection remains high in women over 30 years of age. Previous studies have shown...
that the mean rate of new partner acquisition (refer to appendix 1), which should correlate with the force of infection, in women aged 35–39 years is only slightly lower than for those aged 20–24 years (69).

The main strength of the model is that it involves an extensive fitting procedure that identifies multiple parameter sets, which fit simultaneously different epidemiologic data. The parameter sets identified can loosely be considered as different models that enable thorough investigation of the impact of natural history assumptions and parameter uncertainty on predictions of vaccine effectiveness. This strength contrasts with previous modeling approaches, which were centered mainly on a base-case scenario and simple univariate sensitivity analyses. Quantifying parameter uncertainty is particularly important in the context of HPV vaccination because empirical data are scarce and incomplete (especially in older age groups), and the course of disease from infection to development of SCC may never entirely be measured.

There are several methodological advantages to the fitting procedure used. First, we fit HPV type-specific incidence and prevalence data for infection, CIN, and cervical cancer. It is important to fit the model to epidemiologic data from the different stages of disease in order to accurately represent the entire natural history of HPV and SCC. When we fit the model to only observed SCC incidence and genotype distribution, we identify 4,798 posterior parameter sets compared with 164 when HPV infection and CIN data are included in the fitting procedure. These 4,798 parameter sets produce effectiveness results that are much wider than those for the 164 parameter sets we used. Hence, models that use only a base case and fit few epidemiologic outputs have a significant chance of using a parameter set that lacks validity. Second, we do not assume, a priori, age dependencies in our progression and regression rates because no conclusive evidence exists as to these associations. This advantage enables us to explore the impact of age-specific natural history assumptions. Other strengths of our model are that it is, to our knowledge, the first to incorporate genotype-specific susceptibility, natural immunity, and coinfection.

Tangible implications of the modeling approach presented here are that it can help identify the most important data to collect, and it provides decision makers with a measure of uncertainty surrounding predictions. In terms of value of information, model results suggest that efforts should be made to better characterize natural immunity following HPV infection and the epidemiology of HPV in older age groups, because they are important causes of uncertainty in vaccine effectiveness. Similarly, more studies should be focused on quantifying the rate of waning protection following vaccination by measuring antibody decay in clinical trials and/or duration of vaccine protection from surveillance data.

The main limitation of our modeling approach is that it does not take into account the change in the transmission dynamics of infection following vaccination, and not all HPV-related diseases have been factored in. Because model predictions are based on a static model, the effectiveness of vaccination is underestimated because it does not account for herd immunity effects. The rationale behind this choice is that we first need to understand thoroughly the role of parameter uncertainty and the impact of natural history assumptions within a static structure before superimposing the greater uncertainty inherent to dynamic modeling. This model focuses on the role of HPV in the development of SCC because it is currently the most important HPV-associated disease in terms of mortality and health resource use. However, to capture the full potential of the HPV-6/11/16/18 vaccine, one should ideally account for the potential to prevent cases of adenocarcinoma, other anogenital cancers, head and neck cancers, and genital warts.

In summary, our modeling approach, incorporating the extensive parameterization procedure, provides a flexible framework to evaluate vaccination strategies that takes into account uncertainty in parameter assumptions. Model predictions suggest that an HPV-6/11/16/18 prophylactic vaccine would significantly reduce the incidence of HPV infection, CIN, and SCC. Important priorities for future research include understanding the role of natural immunity and modeling the duration of vaccine-induced protection.

ACKNOWLEDGMENTS

Unconditional funding for this research was provided by Merck Frosst, Canada.
Conflict of interest: Marc Brisson was an employee of Merck Frosst, Canada, when the manuscript was submitted.

REFERENCES


Am J Epidemiol 2007;165:762–775
27. de Villiers EM, Fauquet C, Broker TR, et al. Classification of
26. Insinga RP, Dasbach EJ, Myers ER. The health and eco-
22. Paz IB, Cook N, Odom-Maryon T, et al. Human papilloma-
21. McKaig RG, Baric RS, Olshan AF. Human papillomavirus
16. Salazar EL, Mercado E, Calzada L. Human papillomavirus
6. Salazar EL, Mercado E, Calzada L. Human papillomavirus
2. Paz IB, Cook N, Odom-Maryon T, et al. Human papilloma-
1. McKaig RG, Baric RS, Olshan AF. Human papillomavirus

47. Barnabas RV, Laukkanen P, Koskela P, et al. Epidemiology of HPV 16 and cervical cancer in Finland and the potential...


APPENDIX 1

Prior Distributions for the Force of Infection and Progression to SCC Parameters

Force of infection. The age-specific force of infection $\lambda_{i,j}$ is the rate at which susceptible individuals in age group $i$ acquire infection with HPV genotype $j$. The age-specific incidence of infection $I_{i,j}$ is the overall genotype-specific rate of HPV infection within an age group $i$, including in the denominator individuals who are not susceptible. The relation between these rates can be defined as

$$I_{i,j}/S_{i,j} = \lambda_{i,j},$$

with $S_{i,j} \leq 1$.

Here, $S_{i,j}$ is the proportion of individuals in age group $i$ susceptible to HPV genotype $j$.

North American empirical estimates of the incidence rate of HPV exist. However, estimates for the age-specific force of infection of HPV are not available because it is difficult to identify individuals who are immune; type-specific seroconversion does not always occur after natural HPV infection (96). Hence, we defined wide prior distributions for the force of infection by using age-specific incidence rate...
estimates and rates of new partner acquisition. The upper bound of the prior distribution is based on the assumption that the fraction of all individuals susceptible to genotype-specific HPV is close to 100% ($S_{ij} \approx 1$); thus, $I_{ij} \approx \lambda_{ij}$ (equation 1). Based on this assumption and the fact that, in Canada, the number of new sexual partners is largest among women aged 20–24 years (69), the upper bound of the force of infection for women aged $\geq 15$ years was fixed to the maximum HPV incidence in women in the age category 20–24 years (51, 56, 74–80). For girls less than 15 years of age, the upper bound of the force of infection was divided by four because only a fourth of girls between ages 13 and 15 years report being sexually active (97). The lower bound of this prior distribution is fixed to the age-specific incidence of infection, since $I_{ij} < \lambda_{ij}$ (equation 1) for all age groups.

**Progression from CIN2/3 to cancer.** Because progression rates from CIN2/3 to SCC cannot be measured empirically for ethical reasons, we have to fit these rates to observed epidemiologic data. Under our model structure, progression to SCC is solely defined by CIN2/3 prevalence and SCC incidence as

$$X_{ij}^{\text{CIN2/3}} \times \tau_{ij}^{\text{CIN2/3}\rightarrow\text{SCC}} = I_{ij}^{\text{SCC}},$$

where $X_{ij}^{\text{CIN2/3}}$ is the prevalence of women with CIN2/3 caused by genotype $j$ in age group $i$, $\tau_{ij}^{\text{CIN2/3}\rightarrow\text{SCC}}$ is the progression rate of HPV genotype $j$ from CIN2/3 to SCC in age group $i$, and $I_{ij}^{\text{SCC}}$ is the incidence of SCC in age group $i$ due to HPV genotype $j$.

Because reported epidemiologic data on CIN2/3 prevalence and SCC incidence are not genotype specific, we first fit the overall CIN2/3 prevalence to the overall SCC incidence with non-genotype-specific progression rates:

$$X_{i}^{\text{CIN2/3}} \times \tau_{i,\text{opt}}^{\text{CIN2/3}\rightarrow\text{SCC}} = I_{i}^{\text{SCC}}.$$  

Here $\tau_{i,\text{opt}}^{\text{CIN2/3}\rightarrow\text{SCC}}$ is thus the best fitting overall progression rate from CIN2/3 to SCC in age group $i$.

Second, the uncertainty of observed data is taken into account by defining uniform age-specific distributions around these best fitting values because it allows for some flexibility in this age pattern:

$$\tau_{i}^{\text{CIN2/3}\rightarrow\text{SCC}} = \text{Uniform}[\tau_{i,\text{opt}}^{\text{CIN2/3}\rightarrow\text{SCC}} - 0.01; \tau_{i,\text{opt}}^{\text{CIN2/3}\rightarrow\text{SCC}} + 0.01].$$

The value of 0.01 woman-year is arbitrary and corresponds to a degree of variation in the age pattern that is still acceptable in terms of the ratio of parameter sets that fell within $I_{i}^{\text{SCC}}$ targets.

Finally, the genotype-specific rates are obtained by moving the sampled age pattern over a wide range of progression values and fitting the outcomes to the CIN2/3 and SCC genotype distributions:

$$\tau_{i,j}^{\text{CIN2/3}\rightarrow\text{SCC}} = \tau_{i,\text{opt}}^{\text{CIN2/3}\rightarrow\text{SCC}} \times \text{Uniform}_{j}[0.2;3].$$

Here, $\text{Uniform}_{j}[0.2;3]$ is a constant over age for HPV genotype $j$.

### APPENDIX 2. Epidemiologic data for parameter fit and validation*

<table>
<thead>
<tr>
<th>Outcomes used for model fitting (step 4)</th>
<th>Age range†</th>
<th>HPV† types§</th>
<th>Reference number(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV prevalence (%)</td>
<td>15–70</td>
<td>All, 16, 18, all HR, LR</td>
<td>73, 76, 78, 82, 86, 87, 98–101</td>
</tr>
<tr>
<td>Type distribution among CIN1 (%)</td>
<td>Not age specific</td>
<td>16, 18, all HR, LR</td>
<td>102–106</td>
</tr>
<tr>
<td>Type distribution among CIN2/3 (%)</td>
<td>Not age specific</td>
<td>16, 18, all HR, LR</td>
<td>102–107</td>
</tr>
<tr>
<td>CC† incidence (per 100,000 woman-years)</td>
<td>20–70</td>
<td>All</td>
<td>68, 88–92, 108</td>
</tr>
<tr>
<td>Type distribution among CC (%)</td>
<td>Not age specific</td>
<td>All, 16, 18, HR, LR</td>
<td>29, 54‡</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcomes used for cross-validation (step 5)</th>
<th>Age range†</th>
<th>HPV† types§</th>
<th>Reference number(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV incidence (per 100,000 women-years)</td>
<td>15–24</td>
<td>All, 16, 18, all HR</td>
<td>51, 76–78, 80, 109</td>
</tr>
<tr>
<td>CIN1 prevalence (%)</td>
<td>15–70</td>
<td>All</td>
<td>82, 86, 89, 93, 94</td>
</tr>
<tr>
<td>CIN2/3 prevalence (%)</td>
<td>15–70</td>
<td>All</td>
<td>82, 86, 89, 93, 94</td>
</tr>
<tr>
<td>CIN1/2/3 prevalence (%)</td>
<td>15–70</td>
<td>All</td>
<td>86, 94, 100, 107</td>
</tr>
<tr>
<td>CC prevalence (per 100,000 women)</td>
<td>20–70</td>
<td>All</td>
<td>90</td>
</tr>
<tr>
<td>CC mortality (per 100,000 woman-years)</td>
<td>20–70</td>
<td>All</td>
<td>48, 90, 108</td>
</tr>
</tbody>
</table>

* To summarize cytologic and histologic studies, we assumed cervical intraepithelial neoplasia (CIN)1 = low-grade squamous intraepithelial lesion and CIN2/3 = carcinoma in situ = high-grade squamous intraepithelial lesion.

† Age categories: 5-year groups until age 70 years.

‡ HPV, human papillomavirus; HR, high oncogenic risk types; LR, low oncogenic risk; CC, cervical cancer.

§ All, all HPV types combined; 16, HPV-16 only; 18, HPV-18 only; all HR, all HR HPV types including 16 and 18; LR, all LR HPV types.

¶ Only North American data from these papers were used.
APPENDIX 3

Target Definition

Given the variation between data sources, populations, and collection methods, parameter sets were considered acceptable if their predictions fell within prespecified ranges around the observations (110). These ranges are built based on target values $\xi$, which are defined as

$$\xi_l = f \times \text{Max}_i(O_{i,l}), \tag{1}$$

where $l$ represents each genotype-specific outcome used in model fitting (appendix 2), $\text{Max}_i()$ takes the maximum over all age groups $i$, and $O_{i,l}$ is the set of observations for outcome $l$ in age group $i$. Then, the target ranges $r_{i,l}$ are built as

$$r_{i,l} = \{\text{Max}(O_{i,l}) + \xi_l; \text{Min}(O_{i,l}) - \xi_l\}. \tag{2}$$

A parameter set is considered acceptable if the associated model prediction for each fitting outcome falls within its respective target range, $r_{i,l}$, that is, for all type- and age-specific outcomes. Here, we fix $f = 0.5$, which defines a wide target range. Thus, in the fitting procedure, the best parameter sets will be identified first by their ability to fit multiple outcomes rather than fitting a few outcomes precisely. This was preferable given the great heterogeneity in empirical estimates of a given outcome even if derived from prospective studies. $\xi_l$ was defined independently of age to avoid defining a more restrictive range for given age groups because we preferred to include more uncertainty than not. An extra constraint, $\xi_l > 0.1$, was also specified for the same reason. The only exception was for HPV-16 distribution among SCC. Because good estimates of this outcome have been reported by several high-quality studies, the $f = 0.5$ parameter overestimates the uncertainty around this parameter. Thus, $f$ was reduced to 0.1.