Clinical Progression of High-Grade Cervical Intraepithelial Neoplasia: Estimating the Time to Preclinical Cervical Cancer From Doubly Censored National Registry Data

Margaretha A. Vink*, Johannes A. Bogaards, Folkert J. van Kemenade, Hester E. de Melker, Chris J. L. M. Meijer, and Johannes Berkhof

* Correspondence to Margaretha A. Vink, National Institute for Public Health and the Environment, P.O. Box 1, 3720 BA Bilthoven, the Netherlands (e-mail: Annelie.Vink@rivm.nl).

Initially submitted September 21, 2012; accepted for publication March 27, 2013.

Little is known about the time span of progression from high-grade cervical intraepithelial neoplasia (CIN2/3) to invasive cervical cancer. Estimation of this duration from longitudinal studies is not permitted, as CIN2/3 should be treated when detected. Cross-sectional data on the age-specific incidence of detected CIN2/3 and cervical cancer cases are readily available in national registries, but these data are difficult to interpret because neither the moment of lesion development nor the onset of invasive cancer is observed. We developed a statistical model for estimating the duration of time between CIN2/3 and preclinical cancer using Dutch national registries for the years 2000–2005. Human papillomavirus (HPV) genotype data were used to separate CIN2/3 and cancer incidences to obtain estimates for HPV-16-positive and HPV-16-negative lesions. The median time from CIN2/3 to cancer was estimated to be 23.5 years (95% confidence interval: 20.8, 26.6), and 1.6% of the lesions progressed to cancer within 10 years. The median duration for HPV-16-positive lesions was similar, but 2.4% of the HPV-16-positive lesions progressed to cancer within 10 years, as compared with 0.6% for HPV-16-negative lesions. Estimated durations of time to cancer are essential for reassessment of the optimal screening interval in light of vaccination and novel screening tests.

Infection with human papillomavirus (HPV), which is transmitted sexually, is a necessary cause for the development of cervical cancer, the second most common cancer among women worldwide (1). The majority of sexually active people will experience HPV infection at some point in life, with estimates of lifetime risk of approximately 80% for any oncogenic type (2). Fewer than 10% of these infections are persistent, and only a few persistent infections progress to cervical intraepithelial neoplasia (CIN) grade 2 or 3 (CIN2/3) (3). CIN2/3 is considered a precursor of cervical cancer and is treated when detected, even though the possibility of regression to a normal state exists. Whereas CIN2/3 typically develops within a few years of infection with HPV (4–6), progression to invasive carcinoma is generally thought to require much more time. However, the usual duration of time from CIN2/3 to cancer is largely unknown. This time span strongly determines the eventual impact of screening programs on cervical cancer incidence and the time scale on which cervical cancer reductions will be realized in vaccinated cohorts.

Longitudinal studies on the progression of precancerous cervical lesions to preclinical cervical cancer are not available, with the exception of a study conducted in New Zealand in 1965, in which treatment was withheld from women with large CIN3 lesions (7). In that study, 31.3% of the CIN3 lesions progressed to cancer within 30 years. However, those CIN3 lesions were more advanced than the high-grade lesions typically
found in screening programs (8), and the biopsy taken could have influenced the course of disease. The corresponding estimates of clinical progression may have been positively biased despite the longitudinal nature of that study, and under those assumptions, cervical cancer rates among young women should be much higher than those observed in regularly screened populations (9).

A different approach is to infer the time span between onset of CIN2/3 and cancer from cross-sectional data on the occurrence of CIN2/3 and cervical cancer. These data are readily available in national screening databases and cancer registries, but they are difficult to interpret because neither the moment of development of CIN2/3 nor the onset of cancer is observed. In the statistical literature, these data are referred to as doubly censored current status data (10). Moreover, cross-sectional data can be used to reliably estimate longitudinal parameters only if the underlying dynamical system is in a state of equilibrium. Regarding the epidemiology of HPV in the Netherlands, this assumption probably holds for the decade between 2000 and 2009, when the age-adjusted incidence of high-grade cervical lesions and cervical cancer had stabilized after a 30-year period of decreasing cervical cancer rates (11). Previous studies have used cross-sectional data to estimate the time from CIN2/3 to cancer, but the variation in reported mean time spans is considerable; estimates range from 11.8 years to 24.3 years (12–15). To our knowledge, no statistical models have been developed for estimating the time span from CIN2/3 to cervical cancer that have explicitly incorporated the natural history from HPV infection to CIN2/3 to cancer. Van Oortmarsen and Habbema (12) and Bos et al. (14) developed a statistical model based on cytological and histological screening data, but they did not incorporate the age-specific incidence of HPV infection. This means that onset of CIN2/3 was determined without the use of viral data, which are now abundantly available via population-based screening trials (16), affording a more precise estimate of the onset of CIN2/3. Other investigators have used compartmental models to describe the natural history of HPV infection up to development of cervical cancer, but these approaches require simultaneous estimation of numerous parameters, and fitting procedures are not well defined for such complex models (15, 17, 18).

In this paper, we present a statistical model for estimating the time span between CIN2/3 and cervical cancer, using data from national registries in the Netherlands. This model accounts for the age-specific incidence of HPV infection, the age-specific receipt of cervical screening, and the sensitivity of screening by conventional Papanicolaou (Pap) smear. We describe the time from CIN2/3 to cancer by means of a gamma density, the parameters of which are estimated using maximum likelihood. We also estimate separate gamma distributions for HPV-16-positive and HPV-16-negative lesions.

MATERIALS AND METHODS

Data

We used cross-sectional registry data on women diagnosed with either CIN2/3 or cervical cancer in the Netherlands during the years 2000–2005. For each woman, we observed the disease status \( Y \) (0 if CIN2/3, 1 if cancer) and the categorized age of diagnosis \( A_d \) (1 if 14–18 years, 2 if 19–23 years, . . . , 17 if 94–98 years). The CIN2/3 data were obtained from PALGA (Pathologisch-Anatomisch Landelijk Geautomatiseerd Archief), the nationwide network and registry of histo- and cytopathology in the Netherlands (19). PALGA registers all cytological and histological excerpts that are made inside and outside of the national screening program (NSP). Women aged 30–60 years are invited to participate in cervical screening once every 5 years as part of the NSP. Cervical cancer diagnoses were derived from the Netherlands Cancer Registry (www.iknl.nl), a nationwide registry that retrieves and confirms histological information from PALGA. The data are given in Web Table 1, available at http://aje.oxfordjournals.org/.

Model

Figure 1 provides a schematic representation of the data that enter the model. A woman is infected with HPV at age \( A_{HPV} \) and develops a CIN2/3 lesion at age \( A_{CIN} = A_{HPV} + T_1 \) (see Web Appendix). The inclusion of HPV incidence, together with an assumed duration of time from HPV to onset of CIN2/3, \( T_1 \), allowed us to estimate the CIN2/3 incidence. This incidence cannot be retrieved from screening data only, since one observes prevalent lesions instead of incident lesions. After CIN2/3 onset, the lesion either may be detected by screening, after which it is treated, or it may progress to cervical cancer (CA) in \( T_2 \) years at age \( A_{CA} = A_{CIN} + T_2 \).

We are interested in the unknown probability distribution \( P(T_2 = t_2) \), the time from CIN2/3 to cancer. This problem addresses a form of doubly censored current status data, since we observe for each woman only \( (Y, A_d) \); thus, both

\[
\begin{align*}
A_{HPV}, A_{CIN}, A_{CA} & \geq 0 \\
T_1, T_2 & \geq 0
\end{align*}
\]

Figure 1. Schematic representation of the Dutch registry data from the years 2000–2005 that enter the model. Black diamonds (●) denote \( A_{HPV} \), the age of human papillomavirus (HPV) infection based on the age-specific HPV incidence in the Netherlands. White triangles (△) denote \( A_{CIN} \), the age of development of a cervical intra-epithelial neoplasia (CIN) grade 2 or 3 (CIN2/3) lesion (unobserved). Black triangles (◆) denote the age of CIN2/3 lesion detection (from the nationwide histo- and cytopathology registry). White squares (□) denote \( A_{CA} \), the age of cervical cancer (CA) development (unobserved). Black squares (■) denote the age of cervical cancer detection (Dutch cancer registry). Between the dotted lines (at ages 30–60 years), women are invited every 5 years to participate in a national screening program. A woman progresses from \( A_{HPV} \) to \( A_{CIN} \) in \( T_1 \) years and from \( A_{CIN} \) to \( A_{CA} \) in \( T_2 \) years.
$A_{\text{CIN}}$ and $A_{\text{CA}}$ are interval-censored (if $Y = 0$: $A_{\text{HPV}} < A_{\text{CIN}} < A_j < A_{\text{CA}}$; and if $Y = 1$: $A_{\text{HPV}} < A_{\text{CIN}} < A_{\text{CA}} < A_j$). We modeled the probability of detecting $X_j$ CIN2/3 cases and $Z_j$ cervical cancer cases in age group $A_d = j$ ($j = 1, \ldots, N$) given the total number of detected cases in this age group: $X_j + Z_j$. This probability depends on the unknown probability distribution of the time between CIN2/3 and cervical cancer and on the probability of detection by cervical screening. By conditioning on the age of detection, model adjustments for competing risks such as population mortality and hysterectomy will only influence parameter estimates via their effect on the ratio of detected CIN2/3 to detected cancer cases. This effect is likely to be negligible, and therefore we did not adjust for competing risks in the model. Our model assumes that a woman with CIN2/3 either persists in having CIN2/3 or progresses to cancer. However, lesions may also regress, particularly CIN2 (20–22). Ignoring the competing risk of regression of CIN2/3 will lead to a downward bias in the duration of time from CIN2/3 to cancer. This bias depends on the duration of regression of the lesion and will be small if the mean regression time is substantially smaller than the mean progression time. This seems to be an acceptable assumption, since longitudinal studies have shown that regression is likely to occur quickly, with approximately 40% regression of undiagnosed CIN2 over 2 years (21). The time from CIN2/3 to cancer, $T_2$, is modeled by a mixture of a proportion $\omega$ with increasing hazard and a proportion $1 - \omega$ with zero hazard. The function with increasing hazard is assumed to follow a gamma density, $f(t|\kappa, \theta) = \Gamma^{-1}(\kappa)\theta^{-\kappa}t^{\kappa-1}e^{-t/\theta}$, with shape parameter $\kappa$ and scale parameter $\theta$.

### Estimation method

We estimated the parameters $\varphi = \{\omega, \kappa, \theta\}$ by maximizing the log-likelihood function

$$
\ell(\varphi|X, Z) = \sum_{j=1}^{N} X_j \log[P_\varphi(Y = 0|A_d = j)] + Z_j \log[P_\varphi(Y = 1|A_d = j)],
$$

with $N$ representing the total number of age groups. We optimized the log-likelihood function in the statistical package R, version 2.13.1 (R Foundation for Statistical Computing, Vienna, Austria). The full derivations of the log-likelihood function can be found in the Web Appendix. Ninety-five percent confidence intervals were calculated using the nonparametric bootstrap method. The confidence bounds were corrected for bias using the accelerated bias-corrected (BCa) method (23).

### Parameter assumptions

**CIN2/3 incidence.** We calculated the probability distribution for the age of CIN2/3 development $A_{\text{CIN}}$ as the sum of the incidence of HPV infection and the time from HPV infection to CIN2/3 (for complete derivations, see the Web Appendix). The HPV incidence for the Netherlands was based on data from a large randomized screening trial, POBASCAM, in which 44,102 women were tested for HPV (16) and for which transmission model estimates from prevalence data have been previously published (24); see Table 1 for incidence figures. The time from HPV infection to CIN2/3 development was assumed to follow an exponential distribution and to be independent of age at infection. Its mean value was also determined using the POBASCAM data and was estimated at 3.0 years (see Web Appendix).

**Screening.** We assumed that 90% of the women participated in the NSP (25, 26) and that the remaining 10% had only a small probability of receiving an opportunistic Pap smear. The participants in the NSP have an age-specific rate of attendance, which we calculated from PALGA data. These rates are large at screen-eligible ages and small at other ages. All age-specific screening rates are given in Table 2.

**Lesion detection.** The sensitivity of the Pap test for detection of CIN2/3 was set at 0.7 (27), and treatment was assumed to be successful upon detection. For cervical cancer, we assumed Pap test sensitivity of 1 at the first screen after onset of invasive disease and detection within 2 screening rounds, leading to a preclinical phase of cervical cancer with a mean of 5 years and a maximum of 10 years.

**Type specificity.** We used 2 Dutch studies in which CIN2/3 and cancer cases were evaluated with regard to HPV-16 status (28, 29). We pooled data from the 2 studies and obtained a smooth estimator of the age-specific HPV-16 positivity in CIN2/3 and in cervical cancer by performing logistic regression analysis with a (fractional) polynomial function of age as the explanatory variable (see Web Tables 2 and 3 and Web Figures 1 and 2). We used the predicted age-specific proportions of HPV-16-positive and HPV-16-negative CIN2/3 and cancer cases to separate the registry data and to obtain progression parameters for HPV-16-positive and -negative lesions.

### Sensitivity analysis

We performed a sensitivity analysis to investigate the influence of parameter assumptions on the cumulative incidence of cancer development after onset of CIN2/3. The sensitivity

<table>
<thead>
<tr>
<th>Age Group, years</th>
<th>HPV Incidence Rate* (per 5 Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>14–18</td>
<td>0.57</td>
</tr>
<tr>
<td>19–23</td>
<td>0.76</td>
</tr>
<tr>
<td>24–28</td>
<td>0.45</td>
</tr>
<tr>
<td>29–33</td>
<td>0.28</td>
</tr>
<tr>
<td>34–38</td>
<td>0.20</td>
</tr>
<tr>
<td>39–43</td>
<td>0.14</td>
</tr>
<tr>
<td>44–48</td>
<td>0.10</td>
</tr>
<tr>
<td>49–53</td>
<td>0.08</td>
</tr>
<tr>
<td>54–58</td>
<td>0.06</td>
</tr>
<tr>
<td>59–63</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Abbreviation: HPV, human papillomavirus.

*Rates were obtained from the model created by Bogaards et al. (24).*
of the Pap test was varied from 0.5 to 0.8. The proportion of women participating in the NSP was set at 0.80, 0.85, 0.90 (base case), and 0.95. Recent research suggests the duration from HPV infection to CIN2/3 to be longer for HPV-16-negative infections than for HPV-16-positive infections (30). We set this duration at 2, 4, 6, 8, and 10 years for both HPV-16-positive and HPV-16-negative infections.

**RESULTS**

The model’s fit to the data is presented in Figure 2. The peak of observed CIN2/3 cases appears at age 30 years, the starting age of organized screening in the Netherlands. The age distribution of cervical cancer detection has 2 peaks, one around age 35 years and a second one around age 70 years. The shape of this graph shows the effect of screening; at young ages, the fast-progressing CIN2/3 lesions develop to cancer and are detected by screening at age 30–35 years. After age 35 years, screening starts paying off, leading to a reduction in the number of detected cancers. The preventive effect of screening will decline from age 60 years (the upper end of the eligible screening age range) onward, leading to an increase in symptomatic cancers. Without screening, the peak of cancer incidence would be at age 50 years, approximately 25 years after the peak of HPV infection. For HPV-16-positive cervical cancers, the peak at older age is less pronounced in comparison with HPV-16-negative cancers. For both CIN2/3 and cervical cancer, the shape and absolute height of the observed figures are fitted well by the model (see Web Figure 3).

The parameter estimates and their 95% bootstrapped confidence intervals are given in Table 3. The time to cancer in women with progressive CIN2/3 (i.e., the CIN2/3 cases that have a positive hazard of developing cancer) follows a gamma distribution with shape \( \kappa = 5.1 \) (95% confidence interval (CI): 1.7, 6.9) and scale \( \theta = 4.9 \) (95% CI: 3.3, 52.0). The median time from onset of these CIN2/3 lesions to onset of cervical cancer is estimated at 23.5 years (95% CI: 20.8, 26.6). The cumulative cancer incidence is shown in Figure 3. Within 10 years, 1.6% of the CIN2/3 cases will

---

**Table 2.** Rates of Cervical Screening Among Participants and Nonparticipants in the National Screening Program, by Age Group, the Netherlands, 2000–2005

<table>
<thead>
<tr>
<th>Age Group, years</th>
<th>Proportion Screened</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Participants(^a)</td>
<td>Nonparticipants</td>
<td></td>
</tr>
<tr>
<td>14–18</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>19–23</td>
<td>0.06</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>24–28</td>
<td>0.10</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>29–33</td>
<td>0.76</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>34–38</td>
<td>0.74</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>39–43</td>
<td>0.80</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>44–48</td>
<td>0.82</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>49–53</td>
<td>0.80</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>54–58</td>
<td>0.70</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>59–63</td>
<td>0.71</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>64–68</td>
<td>0.04</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>69–73</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>74–78</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>79–83</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>84–88</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>89–93</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>94–98</td>
<td>0.004</td>
<td>0.004</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) We assumed that 90% of screen-eligible women (ages 30–60 years) participated in the National Screening Program (25).

---

**Figure 2.** The black lines represent the expected number of high-grade cervical lesions (cervical intraepithelial neoplasia (CIN) grade 2 or 3 (CIN2/3)) (left) and cervical cancer cases (right) in the Netherlands, by age group, estimated by the model. The large dots denote the observed annual numbers of CIN2/3 lesions and cervical cancer cases observed in the Netherlands for the years 2000–2005. The vertical lines represent 95% confidence intervals.
develop into cervical cancer. Within 20 years, 12% will develop into cancer. The bootstrapped curves show that there is considerable uncertainty about the cumulative cancer incidence after 20 years from onset of CIN2/3. We estimate that 60% (95% CI: 56, 65) of the cancers are found in NSP participants.

### Type specificity

The proportion of HPV-16-positive lesions that have a positive hazard for progression to cancer is estimated at 0.44 (95% CI: 0.30, 0.99), whereas this proportion is estimated at 0.31 (95% CI: 0.26, 0.36) for HPV-16-negative lesions. The small shape parameter for HPV-16-positive lesions compared with HPV-16-negative lesions (Table 3) indicates that the duration of time to cancer shows a much higher variability for HPV-16-positive lesions than for HPV-16-negative lesions. This is in line with the data: Most cancers diagnosed at a young age are HPV-16-positive, but cancers detected in older women still have a moderate probability of being HPV-16-positive as well (see Web Figure 2). This variance is also reflected in the median duration for progressive HPV-16-positive CIN2/3 versus progressive HPV-16-negative CIN2/3 (progressive HPV-16-positive CIN2/3 vs. progressive HPV-16-negative CIN2/3: 29.0 years (95% CI: 22.0, 49.7) vs. 21.9 years (95% CI: 19.8, 24.2)). The cumulative incidence of cervical cancer is higher for HPV-16-positive lesions than for HPV-16-negative lesions in the first 20 years (see Figure 3). Although this seems to be a small difference, these fast-progressing lesions are important when the length of the screening interval must be re-optimized (for instance, when implementing HPV screening or when screening vaccinated women). This figure also shows the larger heterogeneity in progression of HPV-16-positive CIN2/3, which may progress to cancer relatively quickly, as well as after 40 years. In comparison, the hazard becomes almost zero for HPV-16-negative CIN2/3 after 40 years. Of the HPV-16-positive cancers, 58% (95% CI: 49, 65) are found in NSP participants, which is comparable to the proportion of HPV-16-negative cancers, 59% (95% CI: 54, 64).

### Sensitivity analysis

We performed a sensitivity analysis on the estimated cumulative incidence of cervical cancer with respect to the proportion of women participating in the NSP, the sensitivity of the Pap test (Web Figure 4), and the duration from HPV infection to CIN2/3 (Web Figure 5). For the first 15 years after onset of CIN2/3, the cumulative incidence of cancer was robust against a change in the proportion of women participating in the NSP. Thereafter, the cumulative incidence was positively associated with the proportion of women participating in the NSP. The cancer incidence was weakly related to the sensitivity of the Pap test, with an increased sensitivity corresponding to a higher cumulative incidence of cervical cancer. The duration of time from HPV infection to CIN2/3 had a positive association with cancer incidence, for both HPV-16-positive and HPV-16-negative infections. However, the median time from HPV infection to onset of cervical cancer remained fairly stable. For HPV-16-positive infections, median times were 32.2, 32.2, 33.0, 35.0, and 37.7 years for durations from HPV infection to CIN2/3 of 2, 4, 6, 8, and 10 years, respectively. For HPV-16-negative infections, the corresponding estimates were 24.7, 25.1, 25.9, 27.1, and 28.5 years, respectively. Irrespective of the duration of time from HPV infection to CIN2/3, the cumulative cancer incidence for HPV-16-positive lesions remained higher than the cumulative incidence for HPV-16-negative lesions for the first 15 years after onset of CIN2/3.

### DISCUSSION

We developed a statistical model for the detected CIN2/3 and cervical cancer cases that accounted for age-specific rates of cervical screening and for the sensitivity of screening via conventional Pap smear. With this model, we estimated a median time to cancer development of 23.5 years for progressive CIN2/3. We further estimated that of all CIN2/3 lesions (progressive and nonprogressive combined), 1.6% will progress to cancer within 10 years after CIN2/3 onset. The 10-year risk of progression was 2.4% for HPV-16-positive CIN2/3 and 0.6% for HPV-16-negative CIN2/3.

In our model, we were able to estimate the shape of the gamma probability distribution for time to cancer. This parameter was estimated at 5.13, indicating a strongly increasing cancer hazard since onset of CIN2/3. This is in line with the idea that the clinical significance of small, early CIN2/3 is still uncertain but large CIN2/3 cases are likely to eventually progress to cancer (7, 8). However, most compartmental models used for assessing the effectiveness and cost-effectiveness of prevention programs assume a direct link between the disease state CIN3 and cancer (17, 31). Such models assume an exponentially distributed waiting time between CIN3 and...
The use of an exponential (i.e., a gamma distribution with shape parameter 1) instead of a gamma distribution may lead to overestimation of the contribution of fast-progressing lesions. Consequently, aggressive prevention programs with frequent screening may be recommended to prevent interval carcinomas, whereas a less aggressive program would also provide effective protection if the shape parameter had been markedly larger than 1, as indicated by our model.

Our model can be interpreted as a mover-stayer model with the movers (a proportion $\omega$ of the CIN2/3 lesions) progressing to cancer following a gamma distribution. Similar work has been done using longitudinal data on colorectal cancer (32). Our model facilitated description of the heterogeneity in the duration of time to cancer. However, for HPV-16-positive CIN2/3, the cumulative incidence function still showed a large variability in duration compared with HPV-16-negative CIN2/3. The large heterogeneity in duration resulted in a small shape parameter and a flat hazard function, leading to a larger median duration in comparison with HPV-16-negative lesions.

Our estimate of the median time from CIN2/3 to cancer (23.5 years) is similar to the estimate of Insinga et al. (15). However, it is clearly larger than earlier estimates (12–14), a difference which may be explained by the absence of HPV in the early models. We now know that most HPV infections happen early in life, and hence the incidence of CIN2/3 may precede the detection of (prevalent) CIN2/3 by several years. In earlier studies, the onset of CIN2/3 was informed by screening data only and could therefore be strongly associated with screening characteristics. Estimation of the duration from CIN2/3 to cervical cancer depended on the assumed duration from HPV infection to onset of CIN2/3. The sensitivity analysis showed that estimates of the durations from HPV infection to CIN2/3 and from CIN2/3 to cancer were negatively correlated, but the total duration from HPV to cancer remained fairly stable.

We assumed successful treatment for the detected CIN2/3 cases, implying that women have a lesion detected only once. However, lesions may recur, which in our case would have led to overrepresentation of the detected CIN2/3. We accounted for this by choosing the lesion with the highest histological grade per woman per calendar year from the PALGA database. Because CIN2/3 lesions most often recur within 6 months (33), we would expect there to have been a negligible number of double counts.

The cancer data did not contain information on which cancer cases were detected in regular screening and which cancers were found because of symptoms. It is likely that symptomatic cancers are further along in terms of progression and hence older than screen-detected cancers. We did adjust for this heterogeneity, since cancers could be missed in case of nonparticipation in screening. These cancers were assumed to be detected in the subsequent round of screening, which means that the preclinical cancers in our model will become clinical within 10 years, with an average of 5 years, which corresponds to figures reported in the Netherlands (34, 35). In addition, our model estimate of the proportion of screen-detected cancers agrees well with a recent meta-analytical estimate made by Spence et al. (36).

In addition to our estimate of the median time to cancer, we provided estimates for the cumulative incidence of cancer after onset of CIN2/3. Whereas the median time may be sensitive to the tail of the distribution function of the duration to cancer, the cumulative cancer incidence appeared to be a
robust measure for clinical progression for the first 25 years after onset of CIN2/3. It is useful for clinicians and other health professionals, since accurate estimates of cumulative cancer risk have not been widely available. However, these cumulative incidence rates should be interpreted with caution and probably overestimate the absolute risk of a CIN2/3 lesion’s progressing, because our model does not account for regression of CIN2/3. Just as for progression, relatively little is known about regression of CIN2/3, but it is thought to be fast (21). For HPV-16-positive CIN2/3, the cumulative incidence curve displayed in Figure 3 lies above the HPV-16-negative CIN2/3 for the first 15 years, indicating that there is a difference in cumulative incidence rates. This difference will probably be larger for the absolute cancer risk, since HPV-16-positive lesions are thought to be less regressive than lesions caused by another HPV type.

Information on the proportion of women who participate in the NSP for several screening rounds was not available from our data but was taken from another study (25). In a sensitivity analysis, we showed that our estimates for the time to cancer were fairly robust against changes in this proportion. It may also be useful to compare the proportion of cancers detected among women participating in the NSP (60% in our study) with that in other retrospective studies. Bos et al. (37) investigated the screening history of Dutch women with cancer detected between 1994 and 1997, a period in which the NSP was restructured. In that study, 70% of all cancers were detected among women in the screening-eligible age cohort, and 45% of these women had a screening history. Gök et al. (38) investigated the screening history of cervical cancers detected between 2005 and 2007 and found that 74.5% of the cancer cases occurred in women with a screening history. Our estimate lies between those of these 2 Dutch studies and is similar to the proportion of cancers detected in screened women in a meta-analysis (36). Note that our definition includes cancers detected in women who have been screened or who intend to participate in future screening rounds. The nonparticipants have an increased risk of cancer development, which is explained by the poor screening attendance. We did not assume an increased background risk for these women, since screening is widely accepted and accessible in the Netherlands. This assumption is supported by data from a self-sampling study among nonparticipants (26) showing that at 30 years of age, HPV prevalence was comparable to that in persons undergoing regular screening.

Including type specificity improved the model fit and led to type-specific differences in estimates of the progression of CIN2/3, with a higher incidence of cervical cancer for HPV-16-positive lesions than for HPV-16-negative lesions. This finding is in line with previous research in the Netherlands showing that 33% of HPV-positive women eligible for screening were infected with HPV-16, whereas in women with a CIN2+ lesion, this proportion was approximately 55% (39). This increase suggests a more progressive nature of HPV-16-positive lesions in comparison with other types of HPV (27, 40). The more aggressive nature of HPV-16 was also studied by Wentzensen et al. (30), who showed that HPV-16-positive lesions grew faster to CIN2 and CIN3, and by Khan et al. (41) and Kjær et al. (42), who showed that HPV-16-positive women had the highest risk of CIN3 or worse.

Cost-effectiveness studies have suggested that the screening interval may be extended if cytological analysis is replaced by the more sensitive HPV DNA test (43, 44). Our results are helpful in determining the optimal screening interval, since we have calculated the proportion of high-grade lesions progressing to cancer after a certain time since onset of CIN2/3. With the introduction of vaccination against the oncogenic types of HPV, types 16 and 18, the current screening program must be reconsidered as well (45); it is probably safe to offer vaccinated women fewer Pap smears, which lowers screening-related costs and harm. In addition, the results of this study indicate that HPV-16-positive lesions progress to cancer more quickly than HPV-16-negative lesions. This may have an additional impact on the expected decrease in the lifetime number of screens per vaccinated woman, since our estimates indicate that interval cancers are more likely for HPV-16-related lesions than for other lesions.

In summary, we have provided parameter estimates for the distribution function of the waiting time from the development of CIN2/3 lesions to preclinical cervical cancer. In our study, we included current insights into the natural history of cervical disease and used national registry data to obtain reliable estimates. To the best of our knowledge, this is the first study that has modeled the natural history of CIN2/3 according to HPV type, and we found significant differences between HPV-16-positive and HPV-16-negative lesions. These results are essential in mathematical modeling studies that aim to predict the effect of HPV-based screening algorithms and vaccination strategies, and they are also particularly interesting when studying type-specific referral strategies in organized screening.

ACKNOWLEDGMENTS

Author affiliations: Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, the Netherlands (Margaretha A. Vink, Johannes A. Bogaards, Hester E. de Melker); Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, the Netherlands (Margaretha A. Vink, Johannes A. Bogaards, Johannes Berkhof); and Department of Pathology, VU University Medical Center, Amsterdam, the Netherlands (Folkert J. van Kemenade, Chris J. L. M. Meijer).

Dr. Johannes Berkhof and Prof. Chris Meijer have received funding from the European Community’s Seventh Framework Programme under grant agreement 242061 (PREHDICT).

We thank Dr. L. Overbeek of PALGA (Utrecht, the Netherlands) for her contributions to data collection. Conflict of interest: none declared.

REFERENCES


