SUPPLEMENTARY APPENDIX

Human papillomavirus genotype replacement: still too early to tell?

Irene Man\textsuperscript{1,2*}, Simopekka Vänskä\textsuperscript{3,4}, Matti Lehtinen\textsuperscript{5,6}, Johannes A Bogaards\textsuperscript{1,7}

\textsuperscript{1} Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, The Netherlands
\textsuperscript{2} Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands
\textsuperscript{3} Infectious Disease Control and Vaccinations, National Institute for Health and Welfare, Helsinki, Finland
\textsuperscript{4} School of Health Sciences, University of Tampere, Finland
\textsuperscript{5} Department of Laboratory Medicine, Karolinska Institute, Stockholm, Sweden
\textsuperscript{6} Division of Infections and Cancer Epidemiology, Deutsches Krebsforschungszentrum, Heidelberg, Germany
\textsuperscript{7} Department of Epidemiology and Biostatistics, Amsterdam University Medical Center, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

\textsuperscript{*}Author for correspondence
Email: irene.man@rivm.nl
Address: RIVM, P.O. Box 1, 3720 BA Bilthoven, The Netherlands
## Contents

### Appendices

| Section | Page |
|---------|------|
| A. Transmission model | 3 |
| A.1 Demographics | 3 |
| A.2 Infection states | 3 |
| A.3 Sexual contact patterns | 4 |
| A.3.1 Per capita contact rate | 4 |
| A.3.2 Sexual contact mixing distribution | 5 |
| A.4 Transition rates | 7 |
| A.4.1 Acquisition rates | 7 |
| A.4.2 Clearance rates | 7 |
| A.4.3 Rates of waning immunity | 8 |
| A.4.4 Transition matrix | 8 |
| A.5 The system of partial differential equations | 8 |
| A.6 Approximation method for simulation | 8 |
| B. Parametrization and calibration | 9 |
| B.1 Overview of parameters | 9 |
| B.2 Target prevalence | 9 |
| B.3 Calibration under independence | 10 |
| B.4 Calibration under competition | 10 |
| B.5 Vaccination schemes of different analyses | 10 |
| B.5.1 High-uptake vaccination scheme | 10 |
| B.5.2 Low-uptake vaccination scheme | 12 |
| B.5.3 Excessively-high-uptake vaccination scheme | 12 |
| B.6 Overview of the analyses and results | 15 |
| C. Simulation results | 16 |
| C.1 Age-specific prevalence over time | 16 |
| C.2 Type replacement across age groups | 29 |
| C.2.1 Base-case analysis | 29 |
| C.2.2 Sensitivity analyses | 29 |
| C.3 Evaluating final type replacement with short-term measures | 30 |
| C.3.1 Final type replacement | 30 |
| C.3.2 Evaluation using short-term type replacement | 30 |
| C.3.3 Evaluation using vaccine effectiveness for the non-vaccine type | 30 |
Appendices

A Transmission model

A.1 Demographics

We describe a transmission model of two competing HPV types in an open, heterosexual population of age between $a_{\text{enter}} = 10$ and $a_{\text{exit}} = 70$ years. We assume a constant rate of individuals entering the population at age $a_{\text{enter}}$. The number of individuals newly entering the population is $m$ per year, which is set to unity to ease computation, without loss of generality. Throughout all birth cohorts, the proportions of male and female are $\{p_m, p_f\} = \{\frac{1}{2}, \frac{1}{2}\}$, respectively. All individuals leave the population at age 70 years exactly, and there is no exit due to death before this age.

The population is further stratified into a low-risk and a high-risk group with proportions $\{p_l, p_h\} = \{\frac{4}{5}, \frac{1}{5}\}$, respectively. We assume each individual to stay in the same risk group throughout his/her lifetime. Individuals of the same age but from different risk groups differ in their contact rates, with the one of the high-risk group being higher than one of the low-risk group.

In addition, the population is stratified into vaccinated and unvaccinated individuals. Within each sex, vaccinated individuals from the same birth cohorts are all vaccinated at the same age, which can be an age later than $a_{\text{enter}}$. Coverage of vaccination depends on time and sex but not on the risk-group stratification.

To denote the sets of strata in the population, we introduce the following notations:

- $\mathcal{A} = [a_{\text{enter}}, a_{\text{exit}}]$; age range for the sexually active population;
- $\mathcal{G} = \{f, m\}$: female and male;
- $\mathcal{R} = \{l, h\}$: low-risk and high-risk;
- $\mathcal{W} = \{u, v\}$: unvaccinated and vaccinated.

See Figure 1 for the division of a given birth cohort into smaller strata.

![Figure 1: Stratification of individuals aged $a$ into strata according to sex, risk group, and vaccination status.](image)

A.2 Infection states

The model prescribes the transmission dynamics of two competing HPV types, one vaccine type ($vt$) and one non-vaccine type ($nvt$). Both types follow susceptible-infected-recovered-susceptible ($SIRS$) dynamics. Hence, there are nine infection states in total, each characterized by a pairwise combination of $S$, $I$ and $R$ (Figure 2). $\mathcal{J} = \{SS, IS, SI, II, RS, SR, RI, IR, RR\}$. The $i$-th letter of a pairwise combination indicates the status with respect to type $i$. Acquisition and clearance of infection are assumed to occur sequentially, so that only transitions between infection states differing with respect to the status of only one type are possible (Figure 2).
**eFigure 2:** The nine infection states and the respective transitions in the two-type susceptible-infected-recovered-susceptible model. The arrows correspond to transitions of acquisition, clearance, and waning of immunity. In red are the transitions of clearance that are accelerated by a factor of $h$ due to competition through naturally acquired, i.e. infection-induced, cross-immunity.

Each individual enters the sexual mixing population being susceptible to both types (SS). Subsequently, he/she transits between infection states due to acquisition, clearance, and waning of natural immunity with respect to either of the two types until age 70 years.

To denote the number of individuals in infection state $X$ within a given stratum $(a, g, r, w)$, we use $X(a, g, r, w)$, where $X \in \mathcal{X}$, $a \in \mathcal{A}$, $g \in \mathcal{G}$, $r \in \mathcal{R}$ and $w \in \mathcal{W}$.

### A.3 Sexual contact patterns

Patterns of sexual contact depends on sex, age, and risk stratification but not on vaccination status. Firstly, we allow only heterosexual contact. Secondly, contacts between individuals of different ages and risk groups are assortative such that contacts between similar age and the same risk group are preferred. In the following, we first specify the per capita contact rates of individuals in different strata in Section A.3.1. In Section A.3.2, we then specify the mixing distribution, i.e. how the contact of an individual of any stratum is distributed among other strata. The contact pattern we use here is a simplified version of the contact pattern suggested by Vänskä et al.

#### A.3.1 Per capita contact rate

The mean per capita contact rate of different ages $\bar{c}(a)$ is given by the following equation (the black line in eFigure 3a):

$$
\bar{c}(a) = \begin{cases} 
\bar{c}_{\text{max}}(a - a_{\text{enter}})^\kappa \exp\left(-\frac{a - a_{\text{enter}}}{\omega}\right), & \text{for } a \in [a_{\text{enter}}, a_{\text{exit}}], \\
0, & \text{otherwise.}
\end{cases}
$$

The parameters in the above expression are

- $a_{\text{enter}}$: age of entering sexual mixing population as introduced earlier;
- $a_{\text{max}}$: age at which the maximum mean contact rate is attained;
- $\bar{c}_{\text{max}}$: maximum mean contact rate among all ages;
- $\omega$: width of decrease in mean contact rate when deviating from age $a_{\text{max}}$;
- $\kappa = (a_{\text{max}} - a_{\text{enter}})/\omega$;
- $\tau = (\kappa \omega)^\kappa \exp(-\kappa)$. 

4
The terms in the above equations are $q(a, a')$ of sexual contact by individuals of different age $a = 10, 20, \ldots, 70$ to another age $a'$. Note that, by construction, it follows that $c_{\text{max}} = \max c(a)$ and $a_{\text{max}} = \arg \max_a c(a)$. From the mean contact rate $c(a)$, we then derive the risk-group-specific contact rates, $c(a, l)$ and $c(a, h)$ by solving the following system of equations:

$$
\begin{align*}
\begin{cases}
\hat{c}(a) = p_l c(a, l) + p_h c(a, h) \\
(c_r \cdot \hat{c}(a))^2 = p_l (c(a, l) - \hat{c}(a))^2 + p_h (c(a, h) - \hat{c}(a))^2.
\end{cases}
\end{align*}
$$

(2)

The terms in the above equations are

- $c_r$: coefficient of variance of contact rate among risk groups;
- $p_l, p_h$: proportions of individuals in the low- and high-risk group, respectively, as introduced earlier.

Solving this system of equations yields the following expressions for $c(a, l)$ and $c(a, h)$ (Figure 3a):

$$
\begin{align*}
\hat{c}(a) &= (1 - c_r \sqrt{p_h/p_l}), \\
\hat{c}(a) &= c(a)(1 + c_r \sqrt{p_l/p_h}),
\end{align*}
$$

(3)

A.3.2 Sexual contact mixing distribution

The mixing distribution of sexual contact between individuals of different strata is constructed in the following three steps:

- constructing the mixing distribution of contact between risk groups;
- constructing the mixing distribution of contact between age groups;
- combining the two mixing distributions.

In the first step, the mixing distribution between risk groups $Q_{\alpha} = \begin{pmatrix} q_{\alpha}(l, h) & q_{\alpha}(l, l) & q_{\alpha}(h, h) \end{pmatrix}$ is computed according to the following equations:

$$
\begin{align*}
q_{\alpha}(l, h) &= \frac{p_l (1 - c_r \sqrt{p_h/p_l}) (1 - r_l) + p_h (1 - c_0) (1 - r_h)}{p_l (1 - c_r \sqrt{p_h/p_l}) (1 - r_l) + p_h (1 - c_0) (1 - r_h)}; \\
q_{\alpha}(h, l) &= \frac{p_l (1 - c_r \sqrt{p_h/p_l}) (1 - r_l) + p_h (1 - c_0) (1 - r_h)}{p_l (1 - c_r \sqrt{p_h/p_l}) (1 - r_l) + p_h (1 - c_0) (1 - r_h)}; \\
q_{\alpha}(l, l) &= 1 - q_{\alpha}(l, h); \\
q_{\alpha}(h, h) &= 1 - q_{\alpha}(h, l).
\end{align*}
$$

(4)

The terms in the above equations are
Mean age of the contact by individuals of different age \( a, \mu(a) \). Here, \( \delta_{\text{max}} \) is set as 3 years. (b) Variance of the age of the contact by individuals of different age \( a, \sigma^2(a) \). Here, \( \sigma^2_{\text{min}}, \sigma^2_{\text{max}} \) are set as 4 and 16 years\(^2\), respectively.

- \( r_l, r_h \): proportions of contacts of individuals in the low- and high-risk group that are reserved among themselves (assortativity parameters);
- \( p_l, p_h \): proportions of individuals in the low- and high-risk group, respectively, as introduced earlier;
- \( c_v \): coefficient of variance as introduced earlier.

In the second step, the mixing distribution between age groups \( q_{af}(a, a') \) is computed according to the following beta distribution (eFigure 3b):

\[
q_{af}(a, a') = \frac{\gamma(\alpha(a) + \beta(a))\phi(a')^{\alpha(a)-1}(1 - \phi(a'))^{\beta(a)-1}}{\gamma(\alpha(a))\gamma(\beta(a))(\alpha(a) - \alpha(a') + 1)},
\]

where

- \( \alpha(a) = \mu(a)/\frac{1}{\alpha(a)} - 1 \);
- \( \beta(a) = (1 - \mu(a))/\frac{1}{\beta(a)} - 1 \).

The mean age of the contact by an individual of age \( a, \mu(a) \), is assumed to be larger than \( a \) when \( a \) is small and smaller than \( a \) when \( a \) is large (see Figure 4b). This reflects the preference of young individuals to have older partners and vice versa. We assumed \( \mu(a) \) to increase linearly starting from \( a_{\text{enter}} + \delta_{\text{max}} \) at \( a_{\text{enter}} \) to \( a_{\text{exit}} - \delta_{\text{max}} \) at \( a_{\text{exit}} \). The corresponding variance of the age of the contact by an individual of age \( a, \sigma^2(a) \), is also assumed to be a linearly increasing function, which equals to \( \sigma^2_{\text{min}} \) at \( a_{\text{enter}} \) to \( \sigma^2_{\text{max}} \) at \( a_{\text{exit}} \) (see Figure 4b).

From the mixing distributions between risk groups and age groups, we then derive the whole mixing distribution, given by \( c(a, g, r, a', g', r') \). This function describes how many contacts an individual from a given stratum \((a, g, r)\) has with individuals of another stratum \((a', g', r')\), where \( a, a' \in \mathcal{A}, g \neq g' \in \mathcal{G} \) and \( r, r' \in \mathcal{R} \):

\[
c(a, g, r, a', g', r') = \frac{p_r c(a, r) q_{af}(a, a') q_{gr}(r, r') + p_{r'} c(a', r) q_{af}(a', a') q_{gr}(r', r')}{p_r}.
\]

The multiplication with \( p_r \) and \( p_{r'} \) in the above equation ensures the total number of contacts between any two strata is balanced, meaning that the total number of contacts from individuals of any given stratum to individuals of another stratum is exactly equal to the total number of contact vice versa. Note that due to this balance correction, the resulting value of \( c(a, g, r, a', g', r') \) may slightly differ from the product \( c(a, r) q_{af}(a, a') q_{gr}(r, r') \).
A.4 Transition rates

Now the rates of sexual contact are defined, we continue with introducing the related rates of acquiring infection and the rates of clearance and waning of natural immunity. The per capita rate of any given transition is determined by one or more of the following three components:

- baseline rate;
- multiplicative factor due to between-type competition;
- multiplicative factor due to the vaccination status.

In the following subsections, we define the rates of state transitions due to acquisition, clearance and waning of natural immunity, consecutively. We use $P_{XY}(a,g,r,w)$ to denote the rate of an individual from stratum $(a,g,r,w)$ to move from state $X$ to state $Y$. For completeness, transition rates corresponding to impossible transitions are set to zero, e.g. $P_{SS\to RS}(a,g,r,w) = 0$ (Figure 2).

A.4.1 Acquisition rates

The type-specific baseline acquisition rate $\lambda_i(a,g,r)$ of stratum $(a,g,r)$ has the following expression:

$$\lambda_i(a,g,r) = \beta_i \sum_{r' \in \mathcal{R}} \int_{\mathcal{I}} c(a,g,r,a',g',r') \frac{I(a',g',r')}{mp_g p'} da', \quad \text{where} \quad (7)$$

- $g'$: opposite sex;
- $m$: birth rate, which is set to unity, i.e. $m = 1$;
- $\beta_i$: probability of successfully acquiring type-$i$ infection given established contact with an individual infected with type $i$, where $i \in \{vt, nvt\}$;
- $I_{\mathcal{R}}(a',g',r') = \sum_{w' \in \mathcal{W}} IS(a',g',r',w') + II(a',g',r',w') + IR(a',g',r',w')$: all individuals in stratum $(a',g',r')$ infected with the vaccine type;
- $I_{\mathcal{R}}(a',g',r') = \sum_{w' \in \mathcal{W}} SI(a',g',r',w') + II(a',g',r',w') + RI(a',g',r',w')$: all individuals in stratum $(a',g',r')$ infected with the non-vaccine type.

Together with the multiplicative factors, the acquisition rates of an individual belonging to stratum $(a,g,r,w)$ from different infection states are

$$P_{SS\to IS}(a,g,r,w) = P_{SI\to II}(a,g,r,w) = P_{SR\to IR}(a,g,r,w) = \lambda_i(a,g,r) \cdot (1 - \theta_i) \mathbb{1}[w = vt],$$
$$P_{SS\to SS}(a,g,r,w) = P_{SI\to SI}(a,g,r,w) = P_{RS\to RI}(a,g,r,w) = \lambda_{vt}(a,g,r) \cdot (1 - \theta_i) \mathbb{1}[w = vt]. \quad \text{(8)}$$

Here, $\theta_i$ denotes the vaccine efficacy for type $i$. The indicator function $\mathbb{1}[w = vt]$ ensures that multiplication with $(1 - \theta_i)$ only applies to those that are vaccinated.

A.4.2 Clearance rates

The type-specific baseline clearance rates, $\mu_v$ and $\mu_{nvt}$, are constant and identical for all strata. Together with the multiplicative factors, the clearance rates from different infection states of an individual from stratum $(a,g,r,w)$ are

$$P_{IS\to RS}(a,g,r,w) = P_{IR\to RI}(a,g,r,w) = \mu_v \cdot h,$$
$$P_{SI\to RS}(a,g,r,w) = P_{IR\to RI}(a,g,r,w) = \mu_{nvt}, \quad \text{(9)}$$
$$P_{RS\to RR}(a,g,r,w) = P_{IR\to RI}(a,g,r,w) = \mu_{nvt} \cdot h.$$
A.4.3 Rates of waning immunity

The rate of waning immunity depends only on the sex. It is denoted by $\gamma_g$, where $g \in G$:

$$P_{RS \rightarrow SS}(a,g,r,w) = P_{RI \rightarrow SI}(a,g,r,w) = P_{SR \rightarrow SS}(a,g,r,w) = P_{IR \rightarrow IS}(a,g,r,w) = P_{RR \rightarrow RS}(a,g,r,w) = \gamma_g. \quad (10)$$

A.4.4 Transition matrix

For brevity, we abbreviate the notation for strata using $\Theta = (a,g,r,w)$. Then, we summarize all possible transition rates corresponding to an individual of stratum $\Theta$ compactly in a nine-by-nine transition matrix $\Psi(\Theta)$. In this matrix, the entries corresponding to different incoming and outgoing infection states are indexed in the order of $\{SS, IS, SI, II, RS, SR, RI, IR, RR\}$. For any pair of states $X, Y \in \mathcal{S}$, the entry of $\Psi(\Theta)$ in the row of $X$ and the column of $Y$ is defined by $P_{X \rightarrow Y}(\Theta)$, the rate of an individual of stratum $\Theta$ moving from state $X$ to $Y$.

Furthermore, the diagonal elements of the transition matrix $\Psi(\Theta)$ are defined by $(-1)$ times the sum of all other elements in the same row, i.e., the diagonal element of the row of state $X$ is equal to $-\sum_{Y \neq X} P_{X \rightarrow Y}(\Theta)$.

A.5 The system of partial differential equations

Until now, we have suppressed the dependence of the defined quantities on time $t$ and have mention how their evolve in time. Here, we write down $X(\Theta,t)$, $\lambda_t(\Theta,t)$ and $\Psi(\Theta,t)$ formally as quantities that change in time, as described above due to events of acquisition, clearance and waning of immunity throughout the population. With these new notation including time, we can now write down the system of partial differential equations describing how the model evolves in time. We do that in the following vector form. The equations in this section describe for each stratum how the corresponding individuals transit between state and becomes older as time passes by applying the transition matrix to the current situation:

$$\frac{\partial X(\Theta,t)}{\partial t} + \frac{\partial X(\Theta,t)}{\partial a} = \Psi(\Theta,t)^{\top}X(\Theta,t), \quad (11)$$

where $X(\Theta,t) = [SS(\Theta,t), IS(\Theta,t), SI(\Theta,t), II(\Theta,t), RS(\Theta,t), SR(\Theta,t), RI(\Theta,t), IR(\Theta,t), RR(\Theta,t)]^{\top}$ is the state vector of a given stratum $\Theta = (a,g,r,w)$. Note that the entire system of equations consists of the above equation for all strata.

In addition, the system of equations also has a boundary condition, which governs that the influx of individuals into the population. In the model, we assume all individuals to enter unvaccinated. Hence the boundary condition for stratum $(g,r,u)$ at age $a_{enter}$, where $g \in G$ and $r \in \mathcal{R}$, is:

$$SS(a_{enter},g,r,u,t) = m \cdot p_g \cdot p_r. \quad (12)$$

To model vaccination, we just transfer a proportion of a cohort from an unvaccinated stratum into the corresponding vaccinated stratum.

A.6 Approximation method for simulation

Simulation of the system of partial differential equations, described in the previous section, is done using the a numerical approximation called escalator boxcar train. This method discretizes the continuous influx of individuals by lumping together individuals that are born within some pre-fixed interval. Here, this interval is set to be a year. Furthermore, we discretize $c(a,g,r,a',g',r')$ into a stepwise constant function with step size of one year as well. A detailed description of this approximation method is provided by de Roos et al.\[3\]
B Parametrization and calibration

B.1 Overview of parameters

Parametrization of the model was partly based on data from literature and partly derived from calibration. Parametrization of the contact patterns was based on surveys of sexual behaviours in the Netherlands (Figure 3 and Figure 4). The derived parameter values are given in Table 1. This table also presents the chosen values of the demographic parameters. The transmission probabilities and clearance rates were obtained by calibrating the model to data of pre-vaccination age-specific prevalence of HPV in females in the Netherlands. The target prevalence is described in more details in Section B.2.

The calibration procedure briefly goes as follows. First, the parameters of waning of natural, infection-induced immunity were fixed: γ_f = 0.1, γ_m = 1 in the base-case analysis and γ_f = 0.2, γ_m = 4 in sensitivity analysis A, which explores the impact of assuming shorter duration of natural immunity. Then, the transmission probabilities (β_vt, β_{nvt}) and clearance rates (µ_vt, µ_{nvt}) were obtained assuming no competition between the two types, i.e. with the competition parameter h = 1. Results of this step are shown in Section B.3. The obtained transmission probabilities were fixed subsequently while we re-calibrated the clearance rates to match other pre-fixed values for h(= 1.33, 1.67, · · ·, 3). This step is further described in Section B.4. The parameter values obtained from this entire calibration procedure are shown in Table 2.

In addition, there were parameters that only affect the post-vaccination transmission dynamics but not the pre-vaccination one, such as the parameters concerning the vaccination scheme and vaccine efficacy. The applied values of these parameters are shown in B.5.

In Section B.6, we summarize the differences in applied parameter values between the performed analyses, which are the base-case analysis and the three sensitivity analyses.

Fixed parameters

| Demographics | Values |
|--------------|--------|
| Birth rate m | 1 per year * |
| Age of entering sexual mixing population a_{enter} | 10 years |
| Age of exiting sexual mixing population a_{exit} | 70 years |
| Fraction of female and male p_f, p_m | 1/2, 1/2 |
| Fraction of low- and high-risk groups p_l, p_h | 4/5, 1/5 |
| Contact patterns | |
| Age of maximum mean contact rate a_{c_{max}} | 22 years |
| Maximum mean contact rate c_{max} | 2 persons per year |
| Width of decrease in mean contact rate \(\omega\) | 5 persons per year |
| Assortative mixing parameters r_l, r_h | 1/2, 1/2 |
| Coefficient of variance c_v | 0.8 |
| Maximum mean age difference \(\delta_{max}\) | 3 years |
| Minimum variance in age of contacts \(\sigma_{min}^2\) | 4 years^2 |
| Maximum variance in age of contacts \(\sigma_{max}^2\) | 16 years^2 |

eTable 1: Values of the parameters that were fixed throughout all analyses. *To ease computation, we normalized the birth rate to unity without loss of generality.

B.2 Target prevalence

The prevalence of the model vaccine type was calibrated to broadly match the pre-vaccination age-specific prevalence of HPV-16 among females in the Netherlands. This age-specific prevalence is increasing until around age 22 years old and then decreasing, with a peak value of approximately 8%. Accordingly, the maximum value of the model vaccine type prevalence was required to be equal to 8% in the range 22 – 24 years. The target of the model non-vaccine type followed the same curve but was four times smaller, i.e. with a maximum value of 2%.

B.3 Calibration under independence

We obtained parameter values of the transmission probabilities (β_vt, β_{nvt}) and clearance rates (µ_vt, µ_{nvt}) that were able to reproduce the target prevalence assuming no competition between types. By assuming independence between the two types, parameters corresponding to either types could be calibrated separately. For both types, we obtained pairs of type-specific transmission probability and clearance rate that were able to reproduce the
target prevalence. These pairs were characterized by positive correlation, i.e., high values of $\beta_i$ needed to be compensated by high values of $\mu_i$ (Figure 5a for the base-case analysis, Figure 5b for sensitivity analysis A). In addition, the higher the values of a pair were, the smaller the age at the maximum contact rate became (Figure 6a and Figure 6b for the base-case analysis, Figure 7a and Figure 7b for the base-case analysis). Although more than one parameter set was able to reproduce the age-specific prevalence with the peak falling in the required range (22–24 years), we only chose one of them to continue with in the next calibration steps. For the base-case analysis, the chosen parameter value was $\beta_{vt} = \beta_{nvt} = 0.45$. For sensitivity analysis A, it was $\beta_{vt} = \beta_{nvt} = 0.75$.

### B.4 Calibration under competition

Using values of transmission probabilities ($\beta_{vt}, \beta_{nvt}$) obtained from the calibration step under independence, we then searched for combinations of the clearance rates ($\mu_{vt}, \mu_{nvt}$) that were able to reproduce the target prevalence of the vaccine and non-vaccine type simultaneously for assuming various pre-fixed values of the competition parameter $h > 1$. As we previously found, the clearance rates corresponding to $h = 1$ were $\mu_{vt} = 0.674$ and $\mu_{nvt} = 1.178$. The clearance rates we found for other values of the competition parameter ($h = 1.33, 1.67, \cdots, 3$) are shown in Table 2. Notably, the resulting clearance rates had an inverse relation with respect to the competition parameter; to retain the same prevalence, a stronger competition needed to be compensated by slower clearance. In Figure 8, this inverse relation can be observed by the shifting of the location where the two red lines cross towards the left lower corner. The same inverse relation was observed in sensitivity analysis A, where we assumed higher rates of waning of natural immunity (Figure 9).

### B.5 Vaccination schemes of different analyses

In the base-case analysis and sensitivity analysis A, we implemented a girl-only vaccination scheme with high uptake, mimicking the setting in Scotland. In addition, we also implemented two other vaccination schemes for two other sensitivity analyses:

- a girl-only vaccination scheme with medium uptake, mimicking the setting in Netherlands (sensitivity analysis B);
- a sex-neutral vaccination scheme with higher uptake, mimicking the setting in Australia (sensitivity analysis C).

Each of the following subsections describes one of the three vaccination schemes.

#### B.5.1 High-uptake vaccination scheme

The high-uptake vaccination scheme of the base-case analysis and sensitivity analysis A mimicking the Scottish setting consists of a regular girls-only vaccination program at age 12 years with 95% uptake and a catch-up
**eFigure 5:** Contour plots of the maximum value of the age-specific prevalence belonging to type \( i \) assuming no interactions to other types for different combinations of values of \( \beta_i \) and \( \mu_i \). The red lines indicate the combinations that were able to reproduce the same maximum value as the target prevalence. The combinations that also reproduced a peak value within the target range are shown in eFigure 6. (a) In the base-case analysis, \( \gamma_f = 0.1 \) and \( \gamma_m = 1 \). (b) In sensitivity analysis A, \( \gamma_f = 0.2 \) and \( \gamma_m = 4 \).

**eFigure 6:** (Base-case analysis) Shifting of age-specific prevalence of (a) the vaccine type and (b) the non-vaccine type produced by different combinations of values of \( \beta_{vt}, \mu_{vt} \) and \( \beta_{nvt}, \mu_{nvt} \), respectively. Increasing values of \( \beta_i \) and \( \mu_i \) lead to a shift to earlier age.
**eFigure 7: (Sensitivity analysis A)** Shifting of age-specific prevalence of (a) the vaccine type and (b) the non-vaccine type produced by different combinations of values of $\beta_{vt}, \mu_{vt}$ and $\beta_{nvt}, \mu_{nvt}$, respectively. Increasing values of $\beta_i$ and $\mu_i$ lead to a shift to earlier age.

program at the introduction of vaccination for girls up to 18 years old (eFigure 10). The vaccination coverages in different catch-up cohorts were:

- age 13-14 years: 80%;
- age 14-16 years: 60%;
- age 17-18 years: 40%.

**B.5.2 Low-uptake vaccination scheme**

The low-uptake vaccination scheme of sensitivity analysis B mimicking the Dutch setting consists of a regular girls-only vaccination program at age 12 years with 60% uptake and a catch-up program at the introduction of vaccination for girls up to 18 years old (eFigure 10). The vaccination coverages in different catch-up cohorts were:

- age 13-14 years: 50%;
- age 14-16 years: 40%;
- age 17-18 years: 30%.

**B.5.3 Excessively-high-uptake vaccination scheme**

The excessively-high-uptake vaccination scheme of sensitivity analysis C mimicking the Australian setting consists of a girls-only program at the introduction of vaccination. After six years of vaccination, the program was extended to boys with stable uptake of 80% (eFigure 10). The regular program for girls is at age 12 years with 80% uptake, complemented by a catch-up program at the introduction of vaccination for girls up to 18 years old. The vaccination coverages in different catch-up cohorts were:

- age 13-14 years: 70%;
- age 14-16 years: 60%;
- age 17-18 years: 50%.
**Figure 8:** (Base-case analysis) Contour plots of the maximum values of the age-specific prevalence belonging to the vaccine (dashed lines) and non-vaccine (solid lines) type for different combinations of $\mu_{vt}$ and $\mu_{nvt}$. Each sub-figure corresponds to a different value of the competition parameter $h (= 1, 2, 3)$. The red contour lines correspond to the maximum values of the target prevalence.
eFigure 9: (Sensitivity analysis A) Contour plots of the maximum values of the age-specific prevalence belonging the vaccine (dashed lines) and non-vaccine (solid lines) type for different combinations of $\mu_{vt}$ and $\mu_{nvt}$. Each subfigure corresponds to a different value of the competition parameter $h (= 1, 2, 3)$. The red contour lines correspond to the maximum values of the target prevalence.
**B.6 Overview of the analyses and results**

The section summarizes the key differences in the parameter values between the performed analyses. In the base-case analysis, the waning rates of naturally acquired, i.e., infection-induced, immunity in female and male were set to $\gamma_f = 0.1$ and $\gamma_m = 1$, respectively. The vaccination scheme was implemented with high uptake as described in B.5.1. In sensitivity analysis A, the rates of waning of natural immunity in female and male were higher: $\gamma_f = 0.2$ and $\gamma_m = 4$. The vaccination scheme was identical to the one in the base-case analysis. In sensitivity analysis B, the rates of waning natural immunity were changed back to the values of the base-case analysis, while the vaccination scheme was implemented with low uptake as described in B.5.2. Lastly, in sensitivity analysis C, the vaccination scheme was implemented excessively-high uptake as described in B.5.3. See eFigure 11 for a schematic diagram summarizing the parameter values used in the different analyses.

The values of the parameter obtained from the calibration process concerning the transmission probabilities ($\beta_{vt}$, $\beta_{nt}$), the clearance rates ($\mu_{vt}$, $\mu_{nt}$), and the competition parameter ($h$) for the different analyses can be found in Table 2.

---

**Sensitivity analysis B**

(low uptake)

$\gamma_f = 0.1, \gamma_m = 1$

$ff = 60\%, fm = 0\%$

**Base-case analysis**

(high uptake)

$\gamma_f = 0.1, \gamma_m = 1$

$ff = 95\%, fm = 0\%$

**Sensitivity analysis C**

(excessively high uptake)

$\gamma_f = 0.1, \gamma_m = 1$

$ff = 80\%, fm = 80\%$

---

**Sensitivity analysis A**

$\gamma_f = 0.2, \gamma_m = 4$

$ff = 95\%, fm = 0\%$

---

**eFigure 10:** Lexis diagrams depicting the vaccinated cohorts and vaccination moments in (a) the base-case analysis, sensitivity analyses A, B and (b) sensitivity analysis C. Girl-only and sex-neutral vaccinated cohorts are indicated in red and black, respectively. Time points of vaccination are indicated by circles.

---

**eFigure 11:** A schematic diagram summarizing the parameter values that characterize the base-case analysis and sensitivity analyses A, B, and C. Here, $ff$ and $fm$ denote the coverage of the regular vaccination scheme, whereas $\gamma_f$ and $\gamma_m$ denote the rate of natural immunity in female and male, respectively.
This chapter presents the simulation results across different analyses (defined in [3.6]) and is organized as follows. Section C.1 shows the age-specific prevalence over time since the introduction of vaccination. Section C.2 presents the results regarding short-term type replacement for different age groups. In Section C.3 we present the correspondence between, on one hand, final type replacement and, on the other hand, short-term type replacement or vaccine effectiveness for the non-vaccine type.

C.1 Age-specific prevalence over time

In this section, we present the age-specific prevalence of the vaccine and non-vaccine types over time. The prevalence of the vaccine and non-vaccine type are presented in separate figures:

- vaccine-type prevalence at \( t = 0, 5, 10, 15, 25, 50 \) years since vaccination (eFigure 12, eFigure 15, eFigure 18, eFigure 21);
- non-vaccine-type prevalence at \( t = 0, 5, 10, 15, 25, 50 \) years since vaccination (eFigure 13, eFigure 16, eFigure 19, eFigure 22).

In addition, there are separate figure showing non-vaccine-type prevalence at only \( t = 0 \) and \( t = 50 \) since vaccination to better visualize the difference of prevalence in the pre- and post-vaccination equilibria (eFigure 14, eFigure 17, eFigure 20, eFigure 23). To ease the comparison to Figure 3 and 4 in the main text, the subfigures of all figures in this section are arranged in the same grid format; the x-axis corresponds to the level of competition (\( h \)) and the y-axis the level of cross-protection (\( \theta_{nvt} \)). Furthermore, in the figures with the non-vaccine-type prevalence at \( t = 0, 50 \) years, we indicated in red the scenarios where 1) the non-vaccine-type prevalence had decreased among young ages and increased among old ages at \( t = 50 \) as compared to \( t = 0 \), and 2) final type replacement occurred at the population level.

Throughout all analyses, the post-vaccination equilibrium was reached within 50 years. In the base-case analysis, and sensitivity analyses A and C, the uptake was high enough to eliminate the vaccine type in the post-vaccination equilibrium but it was not the case in sensitivity analysis B (eFigure 12, eFigure 15, eFigure 18, eFigure 21).
**eFigure 12: (Base-case analysis)** Vaccine-type age-specific prevalence at $t = 0, 5, 10, 15, 25, 50$ years since vaccination.
**eFigure 13: (Base-case analysis)** Non-vaccine-type age-specific prevalence at $t = 0, 5, 10, 15, 25, 50$ years since vaccination.
eFigure 14: (Base-case analysis) Non-vaccine-type age-specific prevalence at the pre-vaccination ($t = 0$ years) and post-vaccination ($t = 50$ years) equilibria. Indicated in red are the scenarios where the non-vaccine-types prevalence had decreased among young ages but increased among old ages while final type replacement occurred.
eFigure 15: (Sensitivity analysis A) Vaccine-type age-specific prevalence at $t = 0, 5, 10, 15, 25, 50$ years since vaccination.
eFigure 16: (Sensitivity analysis A) Non-vaccine-type age-specific prevalence at $t = 0, 5, 10, 15, 25, 50$ years since vaccination.
**eFigure 17: (Sensitivity analysis A)** Non-vaccine-type age-specific prevalence at the pre-vaccination ($t = 0$ years) and post-vaccination ($t = 50$ years) equilibria. Indicated in red are the scenarios where the non-vaccine-types prevalence had decreased among young ages but increased among old ages while final type replacement occurred.
eFigure 18: (Sensitivity analysis B) Vaccine-type age-specific prevalence at $t = 0, 5, 10, 15, 25, 50$ years since vaccination.
eFigure 19: (Sensitivity analysis B) Non-vaccine-type age-specific prevalence at $t = 0, 5, 10, 15, 25, 50$ years since vaccination.
eFigure 20: (Sensitivity analysis B) Non-vaccine-type age-specific prevalence at the pre-vaccination ($t = 0$ years) and post-vaccination ($t = 50$ years) equilibria. Indicated in red are the scenarios where the non-vaccine-types prevalence had decreased among young ages but increased among old ages while final type replacement occurred.
eFigure 21: (Sensitivity analysis C) Vaccine-type age-specific prevalence at $t = 0, 5, 10, 15, 25, 50$ years since vaccination.
eFigure 22: (Sensitivity analysis C) Non-vaccine-type age-specific prevalence at $t = 0, 5, 10, 15, 25, 50$ years since vaccination.
eFigure 23: (Sensitivity analysis C) Non-vaccine-type age-specific prevalence at the pre-vaccination ($t = 0$ years) and post-vaccination ($t = 50$ years) equilibria. Indicated in red are the scenarios where the non-vaccine-types prevalence had decreased among young ages but increased among old ages while final type replacement occurred.
C.2 Type replacement across age groups

In the main text, we defined short-term type replacement to be an increase in non-vaccine-type prevalence among $a$-year-old women at $t$ years since the introduction of vaccination as compared to the corresponding prevalence in the pre-vaccination equilibrium. The corresponding prevalence ratio is denoted by $PR(a,t)$. In this section, we present this prevalence ratio evaluated at different age over time.

C.2.1 Base-case analysis

We first discuss the results obtained in the base-case analysis. In the absence of cross-protection and competition, the non-vaccine-type prevalence remained unchanged for all age groups (flat blue lines in Figure 24). When there was competition (but still in absence of cross-protection), the prevalence ratio corresponding to short-term type replacement ($PR(a,t)$) increased for approximately the first 18 years of vaccination. After a slight overshoot, the prevalence ratio stabilized at the new post-vaccination equilibria. Moreover, the qualitative behaviour is uniform across all age groups in the sense that, within each scenario (each colour), if the prevalence ratio increased in one age group, it also increased in other age groups.

When the level of cross-protection was set to $\theta_{nvt} = 40\%$, the prevalence ratio first decreased in young age groups (top row of Figure 25). After 8 to 12 years, the prevalence ratio either continued to decrease or rebounded, depending on the strength of competition. In the discussion of the main text, we called the initial period of decrease a honeymoon period before type replacement. Such a honeymoon period was not observed in older age groups (bottom row of Figure 25). Hence, in the presence of cross-protection, the qualitative behaviour of the prevalence ratio differs across age groups; within each scenario (each colour), the prevalence ratio may increase in some age groups but decrease in other age groups.

C.2.2 Sensitivity analyses

In the sensitivity analyses, where natural immunity was shorter or coverage was different, we encountered the same result of differences in qualitative behaviour among different age groups whenever the was cross-protection (Figure 25, Figure 25, and Figure 28).
C.3 Evaluating final type replacement with short-term measures

In this section, we present the results regarding final type replacement in the sensitivity analyses (the results of the base-case analysis are presented in the main text). We then check how well final type replacement corresponds with short-term type replacement and vaccine effectiveness for the non-vaccine type evaluated after 10 years of vaccination.

C.3.1 Final type replacement

In the main text, we defined final type replacement as an increase in non-vaccine-type prevalence in all women throughout the population in the post-vaccination equilibrium as compared to the corresponding prevalence in the pre-vaccination equilibrium ($t = 0$ years), with we expressed using prevalence ratio $PR_{final}$. As shown in Section C.1, the post-vaccination equilibrium was reached within 50 years throughout all simulations. Hence, we evaluated final type replacement at $t = 50$ years.

Firstly, when comparing the base-case analysis and sensitivity analysis A, where the duration of natural immunity is shorter, we found a similar scope of final type replacement.

Regarding the other sensitivity analyses on the vaccination coverage, we note that increasing uptake increased the scope of final type replacement; the red area shrinks when comparing eFigure 30 and eFigure 31 to Figure 3 of the main text. Moreover, increasing vaccination uptake above the level of vaccine-type elimination expands the amount of cross-protection in the population, while the extent of type replacement is already saturated. Hence, further increasing uptake then decreases the scope and effect size of final type replacement.

C.3.2 Evaluation using short-term type replacement

We now contrast the findings of final type replacement to the corresponding short-term type replacement evaluated after 10 years among 20-year-old and 25-year-old women. In sensitivity analysis A (shorter natural immunity), due to the faster occurrence of type replacement, evaluation of final type replacement using short-term type replacement after 10 years of vaccination performed better than in the base-case analysis (compare eFigure 29 to Figure 4 in the main text).

In sensitivity analysis B (lower uptake), the scope of final type replacement was larger than in the base-case analysis, while the evaluation of short-term type replacement was similar. Hence, there were more scenarios with final type replacement that were missed by the evaluation using short-term type replacement than in the base-case analysis (compare eFigure 30 to Figure 4 in the main text).

Conversely, in sensitivity analysis C (higher uptake), the scope of final type replacement was smaller than in the base-case analysis, while the evaluation of short-term type replacement was similar. Hence, short-term type replacement evaluated after 10 years of vaccination was better (compare eFigure 31 to Figure 4 in the main text).

Peculiarly, note furthermore that in sensitivity analysis C, there were a number of scenarios (e.g. $h = 3$ and $\theta_{nvt} = 40\%$) in which short-term type replacement occurred after 10 years among 25 year-old-women, while final type replacement did not occurred in the long run. The reason for this discrepancy is the later inclusion of boys vaccination. In these scenarios, final type replacement would have occurred if vaccination remained for girls only (check Figure 3 in the main text). The additional benefits of also vaccinating boys only became apparent after approximately 15 years of vaccination (eFigure 28).

C.3.3 Evaluation using vaccine effectiveness for the non-vaccine type

In the main text, we defined vaccine effectiveness (VE) as the reduction in the risk of non-vaccine-type infection in vaccinated relative to unvaccinated individuals. It was also defined among women and could be evaluated among different ages and time points. Here, it was evaluated after 10 years among 20-year-old or 25-year-old women.

Throughout all sensitivity analyses, VE for the non-vaccine type among 20-year-old women corresponded poorly with the occurrence of final type replacement but better among 25-year-old women (eFigure 29, eFigure 30, and eFigure 31). However, the higher sensitivity among older ages did last long in all scenarios. eFigure 32 shows the evaluation of VE among different ages and at different time points. For each age group, the sensitivity decreased as the time since vaccination increased. Eventually, VE only expresses the level of cross-protection. As a result, VE always converges to a non-negative value so that it becomes immaterial for indicating type replacement. For example, among 25-year-old women, the sensitivity for detecting final type replacement was almost negligible after 20 years of vaccination.
**eFigure 24:** (Base-case analysis) Prevalence ratio $PR(a,t)$ over time since the introduction of vaccination evaluating short-term type replacement. Each panel shows $PR(a,t)$ for a different age $a \in \{15, 20, 25, 30, 35, 40\}$. Different levels of competition $h$ are indicated by different colours. The level of cross-protection was $\theta_{nt} = 0\%$. 
eFigure 25: (Base-case analysis) Prevalence ratio $PR(a,t)$ over time since the introduction of vaccination evaluating short-term type replacement. Each panel shows $PR(a,t)$ for a different age $a \in \{15, 20, 25, 30, 35, 40\}$. Different levels of competition $h$ are indicated by different colours. The level of cross-protection was $\theta_{cvt} = 40\%$. 

32
eFigure 26: (Sensitivity analysis A) Prevalence ratio $PR(a, t)$ over time since the introduction of vaccination evaluating short-term type replacement. Each panel shows $PR(a, t)$ for a different age $a \in \{15, 20, 25, 30, 35, 40\}$. Different levels of competition $h$ are indicated by different colours. The level of cross-protection was $\theta_{int} = 40\%$. 
Figure 27: (Sensitivity analysis B) Prevalence ratio $PR(a,t)$ over time since the introduction of vaccination evaluating short-term type replacement. Each panel shows $PR(a,t)$ for a different age $a \in \{15, 20, 25, 30, 35, 40\}$. Different levels of competition $h$ are indicated by different colours. The level of cross-protection was $\theta_{\text{prv}} = 40\%$. 

\[ \theta_{\text{prv}} = 40\% \]
eFigure 28: (Sensitivity analysis C) Prevalence ratio $PR(a,t)$ over time since the introduction of vaccination evaluating short-term type replacement. Each panel shows $PR(a,t)$ for a different age $a \in \{15, 20, 25, 30, 35, 40\}$. Different levels of competition $h$ are indicated by different colours. The level of cross-protection was $\theta_{int} = 40\%$. 
**eFigure 29: (Sensitivity analysis A)** Top panel: Prevalence ratio $PR_{\text{final}}(a = 20, t = 10)$ evaluating final type replacement for different levels of competition $h$ and cross-protection $\theta_{nvt}$. Red (blue) corresponds to the (non-)occurrence of final type replacement. Middle row: Prevalence ratio $PR(a, t)$ evaluating short-term type-replacement among (left) 20-year-old and (right) 25-year-old women at 10 years after vaccination. Bottom row: Vaccine effectiveness for the non-vaccine type $VE_{\text{neg}}(a, t)$ evaluated among (left) 20-year-old and (right) 25-year-old women at 10 years after vaccination. In the middle and bottom rows, red (blue) corresponds to (no) indication of type replacement based on the applied measure.
eFigure 30: (Sensitivity analysis B) Top panel: Prevalence ratio $PR_{final}$ evaluating final type replacement for different levels of competition $h$ and cross-protection $\theta_{nvt}$. Red (blue) corresponds to the (non-)occurrence of final type replacement. Middle row: Prevalence ratio $PR(a,t)$ evaluating short-term type-replacement among (left) 20-year-old and (right) 25-year-old women at 10 years after vaccination. Bottom row: Vaccine effectiveness for the non-vaccine type $VE_{nvt}(a,t)$ evaluated among (left) 20-year-old and (right) 25-year-old women at 10 years after vaccination. In the middle and bottom rows, red (blue) corresponds to (no) indication of type replacement based on the applied measure.
**eFigure 31: (Sensitivity analysis C)** Top panel: Prevalence ratio $PR^{\text{final}}$ evaluating final type replacement for different levels of competition $h$ and cross-protection $\theta_{nvt}$. Red (blue) corresponds to the (non-)occurrence of final type replacement. Middle row: Prevalence ratio $PR(a, t)$ evaluating short-term type-replacement among (left) 20-year-old and (right) 25-year-old women at 10 years after vaccination. Bottom row: Vaccine effectiveness for the non-vaccine type $VE_{\text{neg}}(a, t)$ evaluated among (left) 20-year-old and (right) 25-year-old women at 10 years after vaccination. In the middle and bottom rows, red (blue) corresponds to (no) indication of type replacement based on the applied measure.
eFigure 32: Vaccine effectiveness for the non-vaccine type evaluated among different ages \( a \) and different time points \( t \) since introduction of vaccination as indicated by the black squares in the Lexis diagram at the top.
References

1 Vänskä S, Auranen K, Leino T, et al. Impact of vaccination on 14 high-risk HPV type infections: a mathematical modelling approach. *PLoS One*, 2013;8:e72088.

2 De Roos A, Diekmann O, Metz J. Studying the dynamics of structured population models: a versatile technique and its application to Daphnia. *Am Nat*, 1992;139:123–147.

3 De Graaf H, Wijsen C. Seksuele gezondheid in Nederland 2017. [https://www.rutgers.nl/feiten-en-cijfers/seksuele-gezondheid-en-gedrag/seksuele-gezondheid-nederland-2017](https://www.rutgers.nl/feiten-en-cijfers/seksuele-gezondheid-en-gedrag/seksuele-gezondheid-nederland-2017), 2017.

4 Coupé V, Berkhof J, Bulkmaens N, Snijders P, Meijer C. Age-dependent prevalence of 14 high-risk HPV types in the Netherlands: implications for prophylactic vaccination and screening. *Br J Cancer*, 2008;98:646–651.

5 Kavanagh K, Pollock KG, Cuschieri K, et al. Changes in the prevalence of human papillomavirus following a national bivalent human papillomavirus vaccination programme in Scotland: a 7-year cross-sectional study. *Lancet Infect Dis*, 2017;17:1293–1302.

6 van Lier E, Geraedts J, Oomen P, et al. Vaccinatiegraad en jaarverslag Rijksvaccinatieprogramma Nederland 2017. [https://www.rivm.nl/publicaties/vaccinatiegraad-en-jaarverslag-rijksvaccinatieprogramma-nederland-2017](https://www.rivm.nl/publicaties/vaccinatiegraad-en-jaarverslag-rijksvaccinatieprogramma-nederland-2017), 2018.

7 National HPV Vaccination Program Register of australia. [http://www.hpvregister.org.au/research/coverage-data](http://www.hpvregister.org.au/research/coverage-data), 2018.