Made available by Hasselt University Library in https://documentserver.uhasselt.be

Model structure analysis to estimate basic immunological processes and maternal risk for parvovirus B19

Supplementary material

GOEYVAERTS, Nele; HENS, Niel; AERTS, Marc & Beutels, Philippe (2011) Model structure analysis to estimate basic immunological processes and maternal risk for parvovirus B19. In: BIOSTATISTICS, 12(2). p. 283-302.

DOI: 10.1093/biostatistics/kxq059
Handle: http://hdl.handle.net/1942/11227

Biostatistics (), **1**, 1, *pp.* 1–24

doi:10.1093/biostatistics/

Supplementary Material to 'Model structure analysis to estimate basic immunological processes and maternal risk for parvovirus B19'

NELE GOEYVAERTS*

Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, Agoralaan 1 Gebouw D, B3590 Diepenbeek, Belgium

nele.goeyvaerts@uhasselt.be

NIEL HENS

Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, and Centre for Health Economics Research and Modeling Infectious Diseases & Centre for the Evaluation of Vaccination, Vaccine & Infectious Disease Institute, University of Antwerp, Belgium

MARC AERTS

Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, Belgium

PHILIPPE BEUTELS

Centre for Health Economics Research and Modeling Infectious Diseases & Centre for the Evaluation of Vaccination, Vaccine & Infectious Disease Institute, University of Antwerp, Belgium

*To whom correspondence should be addressed.

© The Author . Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org.

APPENDIX

A. PARVOVIRUS B19 EPIDEMIOLOGY

We observe that after an initial monotone increase with age, the seroprevalence profiles for parvovirus B19 (PVB19) from five European countries show a decrease or plateau between the ages of 20 and 40, after which the prevalence continues to monotonically increase with age. It seems very unlikely that this would reflect a cohort effect due to an epidemic or a demographical shift, since Nascimento et al. (1990) noted a similar decrease in adults for serological surveys conducted in the 1980's in Rio De Janeiro (Brazil), England and Wales, Japan, and Germany (Nascimento et al., 1990; Cohen and Buckley, 1988; Nunoue et al., 1985; Schwarz et al., 1987). Additionally, we find a decrease or plateau in the age class 20-40 years for PVB19 seroprevalence studies conducted in the 1990's in Japan, Australia and The Netherlands (Matsunaga et al., 1995; Kelly et al., 2000; Zaaijer et al., 2004). Furthermore, Schoub et al. (1993) used an avidity test to establish that most PVB19 infections in pregnancy are not primary infections but re-infections, and in 2007 a case report was published of a secondary symptomatic PVB19 infection in a healthy, immunocompetent adult two years after a positive PVB19 IgG antibody test during prenatal care (Kaufmann et al., 2007). This may imply that reinfection with PVB19 remains possible after an adequate level of antibodies is produced upon primary infection.

B. DATA

B.1 *Serological and demographic data*

Table S1 presents a short summary of the univocal serological data for PVB19 collected in Belgium (BE), England and Wales (EW), Finland (FI), Italy (IT) and Poland (PL). Some demographic figures for each country from the time of data collection will be used when modeling the serological data. First, to make the serological data representative of the different populations, post-stratification weights w_i are calculated from demographic data on population sizes per age class, obtained from EUROSTAT (http://epp.eurostat.ec.europa.eu) and the Office for National Statistics, United Kingdom (http://www.statistics.gov.uk/popest). The reference years chosen for BE, EW, FI, IT and PL, are 2003, 1996, 1998, 2004 and 1999, respectively. The weights w_i are truncated, applying a cut-off c, $\tilde{w}_i = \min(w_i, c)$, to

Table S1. Summary of the serological data sets, life expectancy L, total population size N, and total number of live births B.

country	serological data			demographic data		
	year of collection	age range	sample size	L	N	B
BE	2001-2003	0-65	3075	79	10355197	114001
EW	1996	1-79	2822	77	51125400	649034
FI	1997-1998	1-79	2499	78	5146965	57108
IT	2003-2004	1-79	2514	81	57880478	562603
PL	1995-2004	1-79	2495	73	38651893	382002

reduce the influence of individuals with extreme weights and to avoid excessive variability. Based on the distributions of the post-stratification weights for all countries, we have chosen c equal to 7.

Further, under the assumption of demographic equilibrium, the following relation holds between the age-specific population size $N(a)$ and mortality rate $\mu(a)$:

$$N(a) = N(0)\exp\big(-\Omega(a)\big) \text{ where } \Omega(a) = \int_0^a \mu(u)du.$$

Since the total population size N equals $\int_0^\infty N(a)da$ (Table S1), the number of newborns equals $N(0) = N/L$ where L is the life expectancy given by $L = \int_0^\infty \exp\big(-\Omega(a)\big)da$. The mortality rates $\mu(a)$ are estimated from the population sizes and additional data on age stratified numbers of deaths in the reference year, obtained from EUROSTAT. A Poisson generalized additive model with log link is used to model the number of deaths as a function of age with population size as an offset factor (Hens et al., 2010). Thin plate regression splines are chosen via the `gam` function (`mgcv 1.3-30` package, `R` software). Then, the life expectancy L is estimated from $\hat{\mu}(a)$ using the above formula (Table S1).

Finally, to estimate the frequency and burden of PVB19 infection during pregnancy, data on the number of live births in the reference year stratified by age of the mother at her last birthday, are retrieved from EUROSTAT. The maternal age distribution for live births is denoted by $B(a)$, thus the total number of live births equals $B = \int_0^\infty B(a)da$ (Table S1).

Table S2. Method of recruitment and various sample sizes for the contact surveys.

country	recruitment (max # contact entries)	# participants (missing age)	# reported contacts (missing age)	# close contacts $>$ 15 min
BE	random digit dialling (90)	749 (0)	12775 (3)	5666 (44%)
FI	population registers (34)	1006 (0)	11128 (0)	4215 (38%)
GB	face-to-face interview (29)	1012 (0)	11876 (3)	4961 (42%)
IT	random digit dialling (45)	849 (7)	16623 (3)	7740 (47%)
PL	face-to-face interview (45)	1012 (0)	16501 (2)	8036 (49%)

B.2 *Social contact surveys*

Close contacts, i.e. with physical skin-to-skin touching, are likely to play an important role in the transmission of PVB19, considering the reports of school outbreaks (Woolf et al., 1989; Rice and Cohen, 1996; Gonçalves et al., 2005), high attack rates in households (Chorba et al., 1986) and outbreaks in hospital wards (Bell et al., 1989; Pillay et al., 1992). Also in different studies, high occupational risk estimates are reported for day-care and after-school clubs personnel, nursery and elementary school teachers (Valeur-Jensen et al., 1999; Gillespie et al., 1990; Cartter et al., 1991), indicating that young children are the main spreaders of PVB19. Furthermore, exposure to children, particularly in the household, has been identified as the main risk factor for PVB19 infection in pregnant women (Valeur-Jensen et al., 1999).

The diary-based questionnaires from the POLYMOD contact survey consist of participant-related information such as age and gender, and details about each contact made during one randomly assigned day: age and gender of the person made contact with, location, duration and frequency. Moreover, a distinction between two types of contacts was made: non-close contacts, defined as two-way conversations of at least three words in each others proximity, and close contacts that involve any sort of physical skin-to-skin touching. Using EUROSTAT census data on population sizes of different age by household size combinations for the year 2000, post-stratification weights are given to the participants in order to make the data representative of the different populations. The weights are truncated to a maximum of 5; a value chosen based on the weight distributions. A short summary of the data collection and sample sizes for each country are provided in Table S2.

C. DISCRETIZED FORMULAS

The integral equation (3.1) has no closed form solution and therefore, we solve the system numerically by turning to a discrete age framework, assuming a constant force of infection in each age-class. For this purpose, denote the first age interval $(a_{[1]}, a_{[2]})$ and the jth age interval $[a_{[j]}, a_{[j+1]})$, where $a_{[1]} = A$. For the MSIRW models, the proportion of susceptibles of age a (3.2), with $a \in [a_{[j]}, a_{[j+1]})$, reduces to

$$s(a) = \exp\left(-\sum_{k=1}^{j-1} \lambda_k (a_{[k+1]} - a_{[k]}) - \lambda_j (a - a_{[j]})\right).$$

Making use of the latter formula, the force of infection for age class i is approximated by:

$$\begin{aligned}\lambda_i =& \frac{ND}{L} \exp(-\mu_1 A) \sum_j \beta_{ij} \frac{\lambda_j}{\lambda_j + \mu_j} \left[\exp\left(-\sum_{k=1}^{j-1} (\lambda_k + \mu_k)(a_{[k+1]} - a_{[k]})\right)\right. \\ & \left. - \exp\left(-\sum_{k=1}^{j} (\lambda_k + \mu_k)(a_{[k+1]} - a_{[k]})\right)\right],\end{aligned}$$

β_{ij} denoting the per capita rate at which an individual of age class j makes effective contacts with a person of age class i, per year. The fraction of seropositives for the MSIRWb-ext model is approximated by

$$\begin{aligned}r(a) =& (1-\varphi) \sum_{\ell=1}^{j-1} \frac{\lambda_\ell}{(1-\varphi)\lambda_\ell - \varepsilon_\ell} \exp\left(-\sum_{k=1}^{\ell-1} \lambda_k (a_{[k+1]} - a_{[k]}) - \sum_{m=\ell+1}^{j-1} (\varphi\lambda_m + \varepsilon_m)(a_{[m+1]} - a_{[m]})\right. \\ & \left. - (\varphi\lambda_j + \varepsilon_j)(a - a_{[j]})\right) \cdot \left[\exp\left(-(\varphi\lambda_\ell + \varepsilon_\ell)(a_{[\ell+1]} - a_{[\ell]})\right) - \exp\left(-\lambda_\ell (a_{[\ell+1]} - a_{[\ell]})\right)\right] \\ & + \frac{(1-\varphi)\lambda_j}{(1-\varphi)\lambda_j - \varepsilon_j} \cdot \exp\left(-\sum_{k=1}^{j-1} \lambda_k (a_{[k+1]} - a_{[k]})\right) \left[\exp\left(-(\varphi\lambda_j + \varepsilon_j)(a - a_{[j]})\right)\right. \\ & \left. - \exp\left(-\lambda_j (a - a_{[j]})\right)\right] + \varphi \sum_{\ell=1}^{j-1} \frac{\lambda_\ell}{\varphi\lambda_\ell + \varepsilon_\ell} \cdot \exp\left(-\sum_{m=\ell+1}^{j-1} (\varphi\lambda_m + \varepsilon_m)(a_{[m+1]} - a_{[m]})\right. \\ & \left. - (\varphi\lambda_j + \varepsilon_j)(a - a_{[j]})\right) \left[1 - \exp\left(-(\varphi\lambda_\ell + \varepsilon_\ell)(a_{[\ell+1]} - a_{[\ell]})\right)\right] + \frac{\varphi\lambda_j}{\varphi\lambda_j + \varepsilon_j} \\ & \cdot \left[1 - \exp\left(-(\varphi\lambda_j + \varepsilon_j)(a - a_{[j]})\right)\right],\end{aligned} \tag{C.1}$$

if a belongs to the jth age interval. The variants for the more parsimonious, nested MSIRW models are obtained by substituting φ with 0 and 1, respectively.

From the integral equation (3.4) it follows that in case of an MSIRS-ext model, the force of infection for age class i can be approximated by:

$$\begin{aligned}\lambda_i =& \frac{ND}{L}\exp(-\mu_1 A)\sum_j\Bigg[\beta_{1ij}\frac{\lambda_j}{\lambda_j+\mu_j}\Bigg\{\exp\Bigg(-\sum_{\ell=1}^{j-1}(\lambda_\ell+\mu_\ell)(a_{[\ell+1]}-a_{[\ell]})\Bigg)\\ &-\exp\Bigg(-\sum_{\ell=1}^{j}(\lambda_\ell+\mu_\ell)(a_{[\ell+1]}-a_{[\ell]})\Bigg)\Bigg\}+\beta_{2ij}\frac{\lambda_j}{\lambda_j+\sigma_j+\mu_j}\exp\Big(-\sum_{\ell=1}^{j-1}\mu_\ell(a_{[\ell+1]}-a_{[\ell]})\Big)\\ &\cdot\Big\{1-\exp\big(-(\lambda_j+\sigma_j+\mu_j)(a_{[j+1]}-a_{[j]})\big)\Big\}\Bigg\{\sum_{\ell=1}^{j-1}\exp\Big(-\sum_{m=\ell+1}^{j-1}(\lambda_m+\sigma_m)(a_{[m+1]}-a_{[m]})\Big)\\ &\cdot\Bigg[\frac{\sigma_\ell}{\lambda_\ell+\sigma_\ell}\Big\{1-\exp\big(-(\lambda_\ell+\sigma_\ell)(a_{[\ell+1]}-a_{[\ell]})\big)\Big\}-\exp\Big(-\sum_{k=1}^{\ell}\lambda_k(a_{[k+1]}-a_{[k]})\Big)\\ &\cdot\Big\{1-\exp\big(-\sigma_\ell(a_{[\ell+1]}-a_{[\ell]})\big)\Big\}\Bigg]+\frac{\sigma_j}{\lambda_j+\sigma_j}\Bigg[\frac{\lambda_j+\sigma_j+\mu_j}{\mu_j}\\ &\cdot\frac{1-\exp\big(-\mu_j(a_{[j+1]}-a_{[j]})\big)}{1-\exp\big(-(\lambda_j+\sigma_j+\mu_j)(a_{[j+1]}-a_{[j]})\big)}-1\Bigg]-\exp\Big(-\sum_{\ell=1}^{j-1}\lambda_\ell(a_{[\ell+1]}-a_{[\ell]})\Big)\\ &\cdot\Bigg[\frac{\lambda_j+\sigma_j+\mu_j}{\lambda_j+\mu_j}\cdot\frac{1-\exp\big(-(\lambda_j+\mu_j)(a_{[j+1]}-a_{[j]})\big)}{1-\exp\big(-(\lambda_j+\sigma_j+\mu_j)(a_{[j+1]}-a_{[j]})\big)}-1\Bigg]\Bigg\}\Bigg].\end{aligned} \tag{C.2}$$

In case $\beta_{1ij}=\beta_{2ij},\forall i,j$, it can be shown analytically that this expression reduces to the corresponding formula for the MSIRS model. The formula for the fraction of seropositives for both the MSIRS and MSIRS-ext model is identical to the one derived for the MSIRW model (C.1) with $\varphi=1$, where the waning rate ε is replaced by the rate of re-entering the susceptible state, σ.

For the MSIRWS model, we partition the age classes $[a_{[j]},a_{[j+1]})$ into smaller intervals of length δ and approximate the system of differential equations by a set of difference equations. The proportion of susceptibles and the proportion of individuals in the low immunity state, are then calculated as follows:

$$\begin{cases} s_{i+1} &= s_i+\sigma_i\delta w_i-\lambda_i\delta s_i,\\ w_{i+1} &= w_i+\varepsilon_i\delta\{1-s_i\}-\{\varphi\lambda_i+\sigma_i+\varepsilon_i\}\delta w_i,\end{cases}$$

where $s_1 = 1$ and $w_1 = 0$. The force of infection for age class i is approximated by:

$$\lambda_i = \frac{NDe^{-\mu_1 A}}{L} \sum_j \frac{\beta_{ij}\lambda_j s_j}{\mu_j} \left[\exp\left(-\sum_{k=1}^{j-1} \mu_k (a_{[k+1]} - a_{[k]}) \right) - \exp\left(-\sum_{k=1}^{j} \mu_k (a_{[k+1]} - a_{[k]}) \right) \right]$$

and the fraction of seropositives $r(a)$ is approximated by $1 - s_i - w_i$, where the index i is chosen such that age a is located in the i^{th} age interval of length δ.

D. Contact and transmission rates

It has been shown that the method of estimating contact rates from social contact surveys and using them to inform transmission rates for infections transmitted predominantly through non-sexual social contacts, is more efficient than the traditional Anderson and May approach (Anderson and May, 1991) of imposing parametric mixing patterns on the so-called Who-Acquires-Infection-From-Whom matrix (Wallinga et al., 2006; Ogunjimi et al., 2009; Goeyvaerts et al., 2010). The social contact hypothesis states that the age-specific transmission rates are directly proportional to the age-specific rates of making social contact (Wallinga et al., 2006): $\beta(a, a') = q \cdot c(a, a')$. Here, $c(a, a')$ denotes the per capita rate at which an individual of age a' makes contact with a person of age a, per year. Note that we have two q parameters in the mass action principle (C.2) for the MSIRS-ext scenario, q_1 and q_2, to differentiate between infectivity of individuals with primary infection and reinfection.

The contact rates $c(a, a')$ are estimated from the POLYMOD contact survey (Appendix A) by applying a smooth-then-constrain-approach as described in Goeyvaerts et al. (2010). In short, the mean contact surface is estimated using a bivariate smoothing approach with a thin plate regression spline basis (Wood, 2006), assuming a negative binomial distribution for the number of reported contacts over one year age intervals and taking into account post-stratification weights (`gam` function, `mgcv 1.3-30` package, `R` software). Subsequently, the estimated contact surface is constrained using age-specific population sizes (Appendix A) such that the reciprocal nature of contacts is taken into account (Wallinga et al., 2006). In this paper, we assume that PVB19 transmission rates are proportional to rates of making close contact, i.e. involving physical skin-to-skin touching, and particularly those for which the total contact time per day exceeds fifteen minutes (Goeyvaerts et al., 2010).

D.1 *Age-dependent proportionality of the transmission rates*

We assess the sensitivity of the results from our model structure analysis for PVB19 (Section 4) with respect to the constant proportionality (CP) assumption by allowing for an age-dependent q: $\beta(a, a') = q(a, a') \cdot c(a, a')$. This age-specific proportionality factor $q(a, a')$ may reflect, for instance, discrepancies between the social contact proxies measured in the contact survey and the 'true' contact rates underlying infectious disease transmission, or differences in characteristics related to susceptibility or infectiousness, though the latter is not estimable from serological surveys. As in Goeyvaerts et al. (2010), we consider the following three matrix structures for $q(a, a')$:

$$\boldsymbol{D_1} = \begin{pmatrix} \theta_1 & \theta_2 \\ \theta_2 & \theta_2 \end{pmatrix}, \ \boldsymbol{D_2} = \begin{pmatrix} \theta_1 & \theta_1 \\ \theta_2 & \theta_2 \end{pmatrix}, \ \boldsymbol{D_3} = \begin{pmatrix} \theta_1 & \theta_2 \\ \theta_2 & \theta_1 \end{pmatrix}, \tag{D.1}$$

involving two transmission parameters θ_1 and θ_2 for the population dichotomized by a cut-off at a predetermined age G. In the sensitivity analysis, we restrict attention to MSIR, MSIRW(b) and MSIRS to ensure estimability and choose the same dichotomy of the population, namely with a cut-off point at age $G = 12$ years, which performed well in our application of the MSIR model to varicella zoster virus (VZV) serology described in Goeyvaerts et al. (2010).

By parameterizing $q(a, a')$ according to the matrix structures $\boldsymbol{D_1}, \boldsymbol{D_2}, \boldsymbol{D_3}$ (D.1), the evidence of waning immunity arising from the CW models for BE, EW and IT, is almost completely absorbed, which is expressed by the very small estimates for the waning rates. For these countries, under the assumption of lifelong immunity, the age-dependent proportionality (AP) model is always selected according to AIC/BIC. For Belgium, the AP constant waning models fit the seroprofile much better than the CP models and the estimates for R_0 vary around the estimates obtained previously (Table 1), with a pronounced dependence on the configuration type $\boldsymbol{D}$ which is similar to what we observed for VZV (Goeyvaerts et al., 2010). When making pairwise comparisons of the CP versus AP constant waning models for EW, the AIC values are always in favor of the AP scenarios, while the selection based on BIC depends on the waning scenario and the parametric model considered for q. For Italy, however, the BIC values always select the CP constant waning models over their AP counterpart. For BE, EW and IT, the force of infection is now estimated to be smaller in adults which reduces the estimated maternal frequency of PVB19 infection

(Table 3). Finally for FI and PL, allowing q to be age-dependent, does not substantially affect the fit of the CW scenarios to the serological data and nearly preserves the estimates obtained previously (Tables 1 and 3). The CP-models are better in terms of AIC and BIC, the latter with the exception of $\boldsymbol{D_2}$ for Poland.

For the AW models, however, the evidence in favor of waning immunity is sustained for BE, EW and PL, under the age-dependent proportionality assumption for the transmission rates. Furthermore, the ranking of the different waning scenarios according to AIC/BIC remains approximately the same for each country compared to the results in Table 2. Under AP, the estimates for the waning rates $\boldsymbol{\varepsilon}, \boldsymbol{\sigma}$ slightly decrease for BE and slightly increase for EW and FI, while for IT and PL these fluctuate around the estimates obtained before depending on the parametric structure for q. Further, the estimates for R_0 are generally close to the estimates obtained previously (Table 2), though we observe somewhat larger deviations in case of $\boldsymbol{D_1}$ and $\boldsymbol{D_2}$ for the MSIRWb scenario for BE, EW and PL. For these three countries, information criteria based pairwise selection of the CP versus AP counterparts differs depending on the waning scenario and the configuration type $\boldsymbol{D}$ considered, but overall the smallest AIC/BIC values are obtained for the MSIRS AP scenario based on $\boldsymbol{D_3}$ for BE and $\boldsymbol{D_1}$ for EW and PL. A visual inspection of the fit to the serological data shows that this model more pronouncedly captures the shoulder effect in teenagers and 20 year olds for BE and EW, and that the fit to the initial prevalence rise in children is improved for PL, compared to the scenarios depicted in Figure 2. For Finland, there is still no evidence against lifelong immunity and the MSIR CP model (depicted in Figure 2) is ultimately the best one according to information criteria based selection. For Italy, similar to the constant waning case, the evidence of waning immunity is absorbed by the age-dependent q and the CP age-dependent waning models are selected over their AP counterpart. The frequency of PVB19 infection in pregnancy is now estimated to be lower in BE, slightly higher for FI, and for EW, IT and PL it fluctuates around the estimates displayed in Table 3 depending on the AP matrix $\boldsymbol{D}$. For all countries, the annual number of PVB19-induced fetal deaths estimated from the MSIRS scenario seems to be the most sensitive with respect to the proportionality assumption.

E. Small simulation study

We conduct a small simulation study to assess the performance of the different mathematical scenarios and the model selection criteria AIC and BIC. Without loss of generality, we simulate serological responses

Table S3. Results of a small simulation study ($n = 198$) considering MSIRWb-ext AW as the 'true' model with parameter values: $q = 0.085$, $\varepsilon_1 = 0.013$, $\varepsilon_2 = 0.000$, and $\varphi = 0.35$ ($R_0 = 3.75$, $\bar{s}_p = 0.12$, $\bar{\lambda}_p = 0.054$).

model	waning	$\bar{\hat{q}}$	$\widehat{\text{s.d.}}\,(\hat{q})$	$\text{MSE}(\bar{\hat{q}})$ $(\cdot 10^{-3})$	$\bar{\hat{R}}_0$	$\widehat{\text{s.d.}}\left(\hat{R}_0\right)$	$\text{MSE}\left(\bar{\hat{R}}_0\right)$	$\pi_{\text{sel,AIC}}$
MSIR		0.056	0.001	0.818	2.49	0.06	1.59	0.0%
MSIRW	CW	0.072	0.003	0.172	3.18	0.13	0.34	0.0%
	AW	0.079	0.003	0.042	3.49	0.14	0.08	6.1%
MSIRWb	CW	0.076	0.003	0.093	3.35	0.15	0.18	0.0%
	AW	0.086	0.003	**0.012**	3.79	0.14	**0.02**	21.2%
MSIRWb-ext	CW	0.076	0.004	0.088	3.36	0.16	0.17	0.5%
	AW	0.086	0.003	**0.013**	3.79	0.15	**0.03**	**43.9%**
MSIRS	CW	0.064	0.002	0.434	2.83	0.08	0.84	0.0%
	AW	0.065	0.002	0.384	2.88	0.08	0.75	28.3%
MSIRS-ext	CW	0.076	0.006	0.128	3.33	0.28	0.25	0.0%

model	waning	$\bar{\hat{s}}_p$	$\widehat{\text{s.d.}}\,(\hat{s}_p)$	$\text{MSE}\left(\bar{\hat{s}}_p\right)$ $(\cdot 10^{-2})$	$\bar{\hat{\lambda}}_p$	$\widehat{\text{s.d.}}\left(\hat{\lambda}_p\right)$	$\text{MSE}\left(\bar{\hat{\lambda}}_p\right)$ $(\cdot 10^{-3})$	$\pi_{\text{sel,BIC}}$
MSIR		0.27	0.01	2.07	0.034	0.001	0.412	0.0%
MSIRW	CW	0.17	0.01	0.24	0.046	0.002	0.079	0.0%
	AW	0.14	0.01	0.05	0.051	0.002	0.019	16.7%
MSIRWb	CW	0.15	0.01	0.12	0.048	0.002	0.042	3.0%
	AW	0.12	0.01	**0.01**	0.055	0.002	**0.005**	**33.8%**
MSIRWb-ext	CW	0.15	0.01	0.11	0.049	0.002	0.040	0.5%
	AW	0.12	0.01	**0.01**	0.055	0.002	**0.006**	3.5%
MSIRS	CW	0.24	0.01	1.41	0.058	0.006	0.049	0.0%
	AW	0.30	0.02	3.03	0.081	0.007	0.743	**42.4%**
MSIRS-ext	CW	0.25	0.01	1.59	0.050	0.007	0.068	0.0%

for the Belgian data set considering MSIRWb-ext AW as the 'true' model with the ML-estimates from Belgium (Table 2) as parameter values: $q = 0.085$, $\varepsilon_1 = 0.013$, $\varepsilon_2 = 0.000$, and $\varphi = 0.35$. The bias of the estimator $\bar{\hat{q}} = \frac{1}{n}\sum_{i=1}^{n}\hat{q}_i$, for the proportionality factor q is defined as $\text{bias}(\bar{\hat{q}}) = \bar{\hat{q}} - q$, and the mean squared error (MSE) is computed as:

$$\text{MSE}\left(\bar{\hat{q}}\right) = \text{bias}^2(\bar{\hat{q}}) + \widehat{\text{Var}}(\hat{q}), \text{ with } \widehat{\text{Var}}(\hat{q}) = \frac{1}{n-1}\sum_{i=1}^{n}(\hat{q}_i - \bar{\hat{q}})^2,$$

the estimated sample variance of $\hat{q}$. Further, we calculate the same figures for a few other 'global' parameters: the basic reproduction number R_0, the average maternal proportion of susceptibles $\bar{s}_p$, and the average maternal force of infection $\bar{\lambda}_p$.

Table S4. The average number of transitions per person during their lifetime (*) and the average age at which these transitions occur.

variable	formula	interpretation
$\bar{n}_{SI}$	$\int_0^\infty \lambda(a)s(a)N(a)/N(0)da$	average number of infections *
$\bar{A}_{SI}$	$\{\int_0^\infty a\lambda(a)s(a)N(a)/N(0)da\}/\bar{n}_{SI}$	average age at infection
$\bar{n}_{RW}$	$\int_0^\infty \varepsilon(a)r(a)N(a)/N(0)da$	average number of transitions from high to low immunity *
$\bar{A}_{RW}$	$\{\int_0^\infty a\varepsilon(a)r(a)N(a)/N(0)da\}/\bar{n}_{RW}$	average age at transition from high to low immunity
$\bar{n}_{WR}$	$\int_0^\infty \varphi\lambda(a)w(a)N(a)/N(0)da$	average number of boosts of immunity from low to high *
$\bar{A}_{WR}$	$\{\int_0^\infty a\varphi\lambda(a)w(a)N(a)/N(0)da\}/\bar{n}_{WR}$	average age at boosting of immunity from low to high
$\bar{n}_{RS}$	$\int_0^\infty \sigma(a)r(a)N(a)/N(0)da$	average number of losses of disease-acquired immunity *
$\bar{A}_{RS}$	$\{\int_0^\infty a\sigma(a)r(a)N(a)/N(0)da\}/\bar{n}_{RS}$	average age at loss of disease-acquired immunity

All mathematical scenarios considered for PVB19 are fitted to the simulated serological data sets and the resulting average estimator, estimated sample standard deviation ($\widehat{\text{s.d.}}$), and MSE for the global parameters are presented in Table S3, together with AIC and BIC model selection percentages: $\pi_{\text{sel,AIC}}$ and $\pi_{\text{sel,BIC}}$, respectively. As expected, the MSE's for all parameters are small for MSIRWb-ext AW, but interestingly, the MSE's for MSIRWb AW are even smaller. The latter model keeps the boosting rate fixed at the force of infection, therefore reducing the variability of the other parameter estimates. AIC is selecting the correct underlying process of age-specific waning and boosting of low immunity in roughly 65% of the simulation runs. In 28% of the cases, however, AIC selects an MSIRS AW scenario allowing for reinfections. This percentage increases to 42% if selection is performed using BIC instead of AIC. BIC hardly ever selects the MSIRWb-ext AW model due to the more severe penalization of the φ parameter. The AIC selection probability of 0.44 for MSIRWb-ext AW almost gets equally distributed over the BIC selection probabilities for the three more parsimonious AW models: MSIRW, MSIRWb and MSIRS.

F. IMMUNITY TRANSITIONS

Following Rouderfer et al. (1994), we estimate the number of certain transitions per person during their lifetime and the average age at which these transitions occur (Table S4), hereby using the ML-estimates for the scenario-specific parameters. Note that for the MSIRS-ext model, the total fraction of susceptibles equals $s(a) = s_1(a) + s_2(a)$. For each country and each transmission scenario considered, the resulting estimates are presented in Table S5.

Table S5. ML-estimates for the average number of transitions per person during their lifetime and the average age at which these transitions occur (see Table S4 for notations), together with 95% bootstrap-based percentile confidence intervals in square brackets. First entry: constant waning (CW); second entry (if available): age-specific waning (AW) with $H = 35$ years.

country	model	waning	$\hat{\bar{n}}_{SI}$		$\hat{\bar{A}}_{SI}$		$\hat{\bar{n}}_{RW*}$ / $\hat{\bar{n}}_{RS+}$		$\hat{\bar{A}}_{RW*}$ / $\hat{\bar{A}}_{RS+}$		$\hat{\bar{n}}_{WR}$		$\hat{\bar{A}}_{WR}$	
BE	MSIR		0.90	[0.88, 0.92]	16.8	[15.6, 17.9]								
	MSIRW	CW	0.95	[0.93, 0.97]	13.3	[12.2, 14.2]	0.21 $*$	[0.13, 0.28]	41.8 $*$	[41.0, 42.5]				
		AW	0.96	[0.95, 0.98]	12.1	[11.2, 13.0]	0.15 $*$	[0.11, 0.19]	20.4 $*$	[20.2, 20.8]				
	MSIRWb	CW	0.96	[0.93, 0.98]	12.7	[11.6, 13.8]	0.50 $*$	[0.28, 0.72]	41.7 $*$	[41.0, 42.4]	0.27	[0.13, 0.43]	43.1	[41.6, 45.0]
		AW	0.97	[0.95, 0.98]	11.5	[10.6, 12.5]	0.55 $*$	[0.34, 0.76]	31.3 $*$	[22.3, 37.2]	0.39	[0.24, 0.56]	36.7	[33.6, 40.6]
	MSIRWb-ext	CW	0.96	[0.93, 0.98]	12.7	[11.6, 13.7]	0.47 $*$	[0.24, 1.06]	41.7 $*$	[40.9, 42.4]	0.24	[0.06, 0.80]	43.2	[41.4, 45.3]
		AW	0.97	[0.95, 0.98]	11.4	[10.4, 12.4]	0.26 $*$	[0.16, 0.48]	20.1 $*$	[19.9, 31.3]	0.14	[0.02, 0.35]	37.3	[34.3, 40.6]
	MSIRS	CW	1.38	[1.12, 1.78]	21.5	[18.2, 25.0]	0.68 +	[0.34, 1.16]	40.3 +	[39.0, 41.5]				
		AW	1.70	[1.38, 2.22]	21.2	[18.7, 23.9]	0.95 +	[0.57, 1.55]	30.6 +	[24.7, 35.2]				
	MSIRS-ext	CW	1.23	[1.07, 2.86]	19.3	[17.1, 30.7]	0.51 +	[0.29, 2.44]	40.6 +	[37.2, 41.6]				
EW	MSIR		0.78	[0.75, 0.80]	16.9	[15.9, 17.8]								
	MSIRW	CW	0.83	[0.78, 0.86]	15.6	[14.8, 16.6]	0.12 $*$	[0.02, 0.20]	42.8 $*$	[42.0, 43.8]				
		AW	0.87	[0.82, 0.90]	14.4	[13.6, 15.3]	0.14 $*$	[0.07, 0.20]	20.8 $*$	[20.6, 35.7]				
	MSIRWb	CW	0.84	[0.79, 0.87]	15.5	[14.6, 16.5]	0.20 $*$	[0.03, 0.34]	42.7 $*$	[41.9, 43.8]	0.06	[0.01, 0.12]	43.8	[41.9, 45.6]
		AW	0.89	[0.85, 0.92]	13.5	[12.7, 14.7]	0.37 $*$	[0.18, 0.52]	27.2 $*$	[20.2, 35.4]	0.21	[0.09, 0.32]	36.4	[34.1, 39.8]
	MSIRWb-ext	CW	0.84	[0.79, 0.87]	15.4	[14.5, 16.4]	0.36 $*$	[0.08, 1.09]	42.6 $*$	[41.9, 43.7]	0.22	[0.03, 0.91]	42.8	[40.8, 44.5]
		AW	0.90	[0.85, 0.92]	13.5	[12.7, 14.7]	0.62 $*$	[0.19, 1.22]	31.1 $*$	[20.3, 36.5]	0.44	[0.05, 1.04]	35.4	[32.2, 38.9]
	MSIRS	CW	0.91	[0.80, 1.01]	17.8	[16.2, 19.2]	0.21 +	[0.03, 0.37]	42.1 +	[41.2, 43.3]				
		AW	1.20	[0.98, 1.38]	18.8	[17.0, 20.0]	0.47 +	[0.23, 0.68]	26.9 +	[19.6, 34.2]				
	MSIRS-ext	CW	0.90	[0.81, 1.14]	17.5	[16.3, 22.3]	0.20 +	[0.04, 0.48]	42.1 +	[41.1, 43.3]				
FI	MSIR		0.72	[0.69, 0.75]	16.5	[15.6, 17.5]								
	MSIRW(b)	CW	0.72	[0.69, 0.75]	16.5	[15.6, 17.3]								
	MSIRS	CW	0.72	[0.69, 0.76]	16.5	[15.6, 17.6]								
IT	MSIR		0.75	[0.71, 0.77]	16.6	[15.2, 17.8]								
	MSIRW	CW	0.81	[0.75, 0.84]	15.2	[14.3, 16.4]	0.13 $*$	[0.02, 0.20]	43.0 $*$	[42.3, 44.0]				
		AW	0.83	[0.76, 0.86]	14.6	[13.9, 15.9]	0.13 $*$	[0.04, 0.20]	31.0 $*$	[20.7, 56.4]				
	MSIRWb	CW	0.82	[0.75, 0.85]	15.0	[14.2, 16.2]	0.20 $*$	[0.03, 0.32]	42.9 $*$	[42.2, 44.0]	0.06	[0.01, 0.11]	46.4	[45.1, 48.1]
		AW	0.84	[0.77, 0.87]	14.4	[13.5, 15.7]	0.24 $*$	[0.07, 0.36]	35.4 $*$	[20.8, 56.3]	0.10	[0.01, 0.16]	42.8	[38.4, 58.7]
	MSIRS	CW	0.89	[0.76, 0.98]	17.5	[15.5, 19.2]	0.21 +	[0.03, 0.34]	42.4 +	[41.7, 43.6]				
		AW	0.97	[0.79, 1.10]	17.9	[15.7, 19.6]	0.28 +	[0.07, 0.44]	34.6 +	[20.5, 55.9]				
	MSIRS-ext	CW	0.88	[0.77, 0.99]	17.2	[15.7, 19.2]	0.20 +	[0.05, 0.35]	42.4 +	[41.7, 43.6]				
PL	MSIR(W)(b)	CW	0.86	[0.83, 0.87]	16.0	[15.1, 17.1]								
	MSIRW	AW	0.87	[0.83, 0.89]	15.5	[14.8, 16.5]	0.02 $*$	[0.00, 0.07]	21.5 $*$	[21.3, 55.4]				
	MSIRWb	AW	0.92	[0.85, 0.94]	13.4	[12.1, 15.3]	0.23 $*$	[0.02, 0.41]	20.8 $*$	[20.4, 21.6]	0.16	[0.01, 0.32]	35.2	[33.8, 37.1]
	MSIRWb-ext	AW	0.92	[0.88, 0.94]	13.2	[12.5, 14.5]	0.52 $*$	[0.32, 0.81]	20.6 $*$	[20.4, 21.8]	0.48	[0.28, 0.79]	30.0	[26.5, 32.8]
	MSIRS	CW	0.86	[0.83, 0.88]	16.0	[15.1, 17.1]								
		AW	1.65	[1.11, 2.33]	19.1	[16.9, 20.2]	0.81 +	[0.26, 1.57]	18.9 +	[18.0, 20.5]				

G. MATLAB CODE

The data sets and `R` code that are used in the paper are available from the authors on request. The MSIRW and MSIRS scenarios considered for PVB19 are implemented in `Matlab` and ML-estimates are obtained using `fminsearch`. We provide the `Matlab` functions below for the two most extensive models `MSIRWboostext` and `MSIRSext`, since all other scenarios are special cases. Both functions make use of the function `read` (not displayed here) to import the country-specific data: the estimated daily contact rates matrix `rij(:,:)`, the vectors containing the serological data i.e. the individuals' age `age(:,)`, serological status `resp(:,)` and post-stratification weight `weight(:,)`, the life expectancy `L`, the total population size `N`, the age-specific mortality rates `mu(:,)`, and the maternal age distribution `bi(:,)` for live births. Further, both functions make use of the function `Rfrac` displayed below to calculate the age-specific fraction of seropositives according to formula (C.1). Finally, both functions require the following input parameters: the `country{''}` specification, the cut-off point `H` for the age-specific waning scenario, the starting values `init(,:)` for the optimization procedure, and the `model{''}` specification for the waning rates.

```
function r = Rfrac(age,epsilon,phi,C,B1,Cb,foi,k)
alow = floor(age);
if length(epsilon)>1
    theta = exp(-(phi*foi(alow+1)+epsilon(alow+1)).*(age-max(0.5,alow)));
    r2 = sum(Cb(:,alow+1).*(theta*ones(1,k))')'+(foi(alow+1)./(phi*foi(alow+1)
    +epsilon(alow+1))).*(1-theta);
else
    theta = exp(-(phi*foi(alow+1)+epsilon).*(age-max(0.5,alow)));
    r2 = sum(Cb(:,alow+1).*(theta*ones(1,k))')'+(foi(alow+1)./(phi*foi(alow+1)
    +epsilon)).*(1-theta);
end
r1 = sum(C(:,alow+1).*(theta*ones(1,k))')'+(B1(alow+1).*(theta-exp(-foi(alow+1)
    .*(age-max(0.5,alow)))));
r1 = max(0,r1);
r2 = max(0,r2);
r = (1-phi)*r1 + phi*r2;
end
```

G.1 *MSIRWb-ext model*

```
function [parhat,R0,risk,trans,aic,bic,exitflag,output] = MSIRWboostext(country,H,init,
                                                          model)

[rij,age,resp,weight,L,N,mu,bi] = readd(country);

% age of maternal antibody waning (0<=A<1)
A = 0.5;

% mean duration of infectiousness
D = 6/365;

% k right-open age-intervals are considered: (A,1), [1,2),..., [k-1,k)
k = 80;
step = [1-A, ones(1,k-1)]';
ageint = [A+[0;cumsum(step(1:end-1))] A+cumsum(step)];
rij = rij(1:k,1:k);
mu = mu(1:k);
bi = bi(1:k);

% ages <= A and >= k are removed from the serological data
resp = resp(age>A & age<k);
age = age(age>A & age<k);

% Function "qestim" to calculate the FOI and likelihood
% conditional on the parameter values
%*****************************************************
function dev = qestim(par)
    q = exp(-par(1));
    if strcmp(model,'constant')
        epsilon = exp(-par(2));
    end
    if strcmp(model,'discrete')
        % piecewise constant function
        epsilon = [exp(-par(2))*ones(H,1) ; exp(-par(3))*ones(k-H,1)];
```

```
end
phi = exp(-par(end));
bij = 365*q.*rij;
foi = 0.1*ones(k,1);
tol = 1;
it = 0;
while (tol>1D-15) && (it<2000)
    S = (N/L)*exp(-mu(1)*A)*exp(-cumsum([0;(foi+mu).*step]));
    I = foi./(foi+mu).*(S(1:end-1)-S(2:end));
    foinext = D*bij*I;
    tol = sum((foinext-foi).^2);
    it = it+1;
    foi = foinext;
end
if it==2000
    error('Maximum number of iterations exceeded')
end
% input from MSIRW framework
s = exp(-cumsum([0;foi.*step]));
if length(epsilon)>1
    f = @(i,j) exp(-sum((phi*foi(i+1:j-1)+epsilon(i+1:j-1)).*step(i+1:j-1)));
else
    f = @(i,j) exp(-sum((phi*foi(i+1:j-1)+epsilon).*step(i+1:j-1)));
end
F = zeros(k);
for j = 1:k
    for i = 1:j-1
        F(i,j) = f(i,j);
    end
end
E = [f(0,2) diag(F,2)' f(k-1,k+1)];
B1 = foi./((1-phi)*foi-epsilon).*s(1:end-1);
B2 = B1.*(E'-(s(2:end)./s(1:end-1)));
C = (B2*ones(1,k)).*F;
% input from MSIRWboost framework
B = (foi./(phi*foi+epsilon)).*(1-E');
```

```
    Cb = (B*ones(1,k)).*F;
    % fraction of seropositives
    r = Rfrac(age,epsilon,phi,C,B1,Cb,foi,k);
    ll = resp.*log(r)+(1-resp).*log(1-r);
    dev = -2*sum(weight.*ll);
end

% Non-linear optimization of the function "qestim"
%*************************************************
[parhat,dev,exitflag,output] = fminsearch(@qestim,init,optimset('FunValCheck',
                                'on','Display','final','MaxFunEvals',1500));
parhat = exp(-parhat);

% Next generation matrix and R0
%******************************
Na = (N/L)*exp(-mu(1)*A)*exp(-cumsum([0;mu.*step]));
M = (Na(1:end-1)-Na(2:end))./mu;
G = D*diag(M)*bij;
R0 = max(real(eig(G)));

% Risk in pregnancy
%******************
Iy = sum(bi.*(s(1:end-1)-s(2:end)));
slb = sum(bi./foi.*(s(1:end-1)-s(2:end)));
sp = slb/sum(bi);
foip = Iy/slb;
Ip = 0.77*Iy;
freqp = sum(bi)/Ip;
fetaldeath = Ip*0.077*(20/40);
risk = [sp foip Ip freqp fetaldeath];

% Transitions
%************
U1 = (1-exp(-(phi*foi+epsilon+mu).*step))./(phi*foi+epsilon+mu);
U2 = (1-exp(-(foi+mu).*step))./(foi+mu);
U3 = (1-exp(-mu.*step))./mu;
```

```
T1 = (ageint(:,1)-ageint(:,2).*exp(-(phi*foi+epsilon+mu).*step))./(phi*foi+epsilon+mu)
    +U1./(phi*foi+epsilon+mu);
T2 = T1-((ageint(:,1)-ageint(:,2).*exp(-(foi+mu).*step))./(foi+mu)+U2./(foi+mu));
T3 = ((ageint(:,1)-ageint(:,2).*exp(-mu.*step))./mu+U3./mu)-T1;
nSI = sum((L/N)*I);
ASI = sum((L/N)*foi./(foi+mu).*(S(1:end-1).*(ageint(:,1)+U2)-S(2:end)
    .*(ageint(:,2))))/nSI;

r1 = sum(C)'+(B1.*(1-(U2./U1)));
r2 = sum(Cb)'+(foi./(phi*foi+epsilon).*((U3./U1)-1));
radapt = (1-phi)*r1+phi*r2;
r1A = T1.*sum(C)'+T2.*B1;
r2A = T1.*sum(Cb)'+T3.*foi./(phi*foi+epsilon);
radaptA = (1-phi)*r1A+phi*r2A;
nRW = sum((L/N)*epsilon.*Na(1:end-1).*U1.*(radapt));
ARW = sum((L/N)*epsilon.*Na(1:end-1).*(radaptA))/nRW;

B1 = (epsilon./(phi*foi+epsilon)).*(1-E');
C1 = (B1*ones(1,k)).*F;
B2 = (epsilon./((1-phi)*foi-epsilon)).*s(1:end-1).*(E'-s(2:end)./s(1:end-1));
C2 = (B2*ones(1,k)).*F;
wadapt = sum(C1)'-sum(C2)'+epsilon./(phi*foi+epsilon).*((U3./U1)-1)-epsilon
        ./((1-phi)*foi-epsilon).*s(1:end-1).*(1-(U2./U1));
wadaptA = T1.*(sum(C1)'-sum(C2)')+epsilon./(phi*foi+epsilon).*T3-epsilon
        ./((1-phi)*foi-epsilon).*s(1:end-1).*T2;
nWR = sum((L/N)*phi*foi.*Na(1:end-1).*U1.*(wadapt));
AWR = sum((L/N)*phi*foi.*Na(1:end-1).*(wadaptA))/nWR;

trans = [nSI ASI nRW ARW nWR AWR];

% Information criteria
%********************
aic = dev+2*length(parhat);
bic = dev+log(length(resp))*length(parhat);

end
```

G.2 *MSIRS-ext model*

```
function [parhat,R0,risk,trans,aic,bic,exitflag,output] = MSIRSext(country,H,init,
                                                                   model)

[rij,age,resp,weight,L,N,mu,bi] = readd(country);

% age of maternal antibody waning (0<=A<1)
A = 0.5;

% mean duration of infectiousness
D = 6/365;

% k right-open age-intervals are considered: (A,1), [1,2),..., [k-1,k)
k = 80;
step = [1-A, ones(1,k-1)]';
ageint = [A+[0;cumsum(step(1:end-1))] A+cumsum(step)];
rij = rij(1:k,1:k);
mu = mu(1:k);
bi = bi(1:k);

% ages <= A and >= k are removed from the serological data
resp = resp(age>A & age<k);
age = age(age>A & age<k);

% Function "qestim" to calculate the FOI and likelihood
% conditional on the parameter values
%*****************************************************
function dev = qestim(par)
    q1 = exp(-par(1));
    q2 = exp(-par(2));
    if strcmp(model,'constant')
        sig = exp(-par(end));
    end
    if strcmp(model,'discrete')
        % piecewise constant function
```

```
    sig = [exp(-par(end-1))*ones(H,1) ; exp(-par(end))*ones(k-H,1)];
end
b1ij = 365*q1*rij;
b2ij = 365*q2*rij;
V2 = exp(-sig.*step);
V3 = exp(-mu.*step);
CV3 = cumprod([1;V3]);
foi = 0.1*ones(k,1);
tol = 1;
it = 0;
while (tol>1D-15) && (it<2000)
    % foi = term*(I1+I2)
    V1 = exp(-foi.*step);
    CV1 = cumprod([1;V1]);
    V12 = V1.*V2;
    CV13 = CV1.*CV3;
    % constructing the number of primary infectious individuals I1
    I1 = (N/L)*exp(-mu(1)*A)*foi./(foi+mu).*(CV13(1:end-1)-CV13(2:end));
    % constructing the number of secondary infectious individuals
    % I2 = term*(Q1+Q2-Q3)
    % constructing Q1
    f = @(l,j) prod(V12(l+1:j-1));
    g = @(l,j) prod(V2(l+1:j-1));
    F = zeros(k);
    G = zeros(k);
    for j = 1:k
        for l = 1:j-1
            F(l,j) = f(l,j);
            G(l,j) = g(l,j);
        end
    end
    B1 = (sig./(foi+sig)).*(1-V12);
    T1 = (B1*ones(1,k)).*F;
    B2 = CV1(1:end-1)*ones(1,k);
    T2 = ((1-V2)*ones(1,k)).*G.*(B2');
    Q1 = sum(T1)'-sum(T2)';
```

```
        % constructing Q2 en Q3
        Q2 = sig./(sig+foi).*((foi+sig+mu)./mu.*((1-V3)./(1-V1.*V2.*V3))-1);
        Q3 = CV1(1:end-1).*((foi+sig+mu)./(foi+mu).*((1-V1.*V3)./(1-V1.*V2.*V3))-1);
        I2 = (N/L)*exp(-mu(1)*A)*foi./(foi+sig+mu).*CV3(1:end-1).*(1-V1.*V2.*V3)
            .*(Q1+Q2-Q3);
        foinext = D*(b1ij*I1+b2ij*I2);
        tol = sum((foinext-foi).^2);
        it = it+1;
        foi = foinext;
    end
    if it==2000
        error('Maximum number of iterations exceeded')
    end
    V1 = exp(-foi.*step);
    V12 = V1.*V2;
    f = @(l,j) prod(V12(l+1:j-1));
    F = zeros(k);
    for j = 1:k
        for l = 1:j-1
            F(l,j) = f(l,j);
        end
    end
    E = [f(0,2) diag(F,2)' f(k-1,k+1)];
    B = (foi./(foi+sig)).*(1-E');
    C = (B*ones(1,k)).*F;
    r = Rfrac(age,sig,1,zeros(k),zeros(k,1),C,foi,k);
    ll = resp.*log(r)+(1-resp).*log(1-r);
    dev = -2*sum(weight.*ll);
end

% Non-linear optimization of the function "qestim"
%*************************************************
[parhat,dev,exitflag,output] = fminsearch(@qestim,init,optimset('FunValCheck',
                                'on','Display','final','MaxFunEvals',1500));
parhat = exp(-parhat);
```

```
% Next generation matrix and R0
%*****************************
Na = (N/L)*exp(-mu(1)*A)*exp(-cumsum([0;mu.*step]));
M = (Na(1:end-1)-Na(2:end))./mu;
G = D*diag(M)*b1ij;
R0 = max(real(eig(G)));

% Risk in pregnancy
%******************
Tp = exp(-cumsum([0;(foi+sig).*step]));
% constructing Q1p
Q1p = (1-Tp(2:end)./Tp(1:end-1)).*sum(C)';
% constructing Q3p
Q3p = foi./(foi+sig).*(((foi+sig).*step)-(1-Tp(2:end)./Tp(1:end-1)));
Iy = sum(foi.*bi.*step)-sum(foi.*bi./(foi+sig).*(Q1p+Q3p));
slb = sum(bi.*step)-sum(bi./(foi+sig).*(Q1p+Q3p));
sp = slb/sum(bi);
foip = Iy/slb;
Ip = 0.77*Iy;
freqp = sum(bi)/Ip;
fetaldeath = Ip*0.077*(20/40);
risk = [sp foip Ip freqp fetaldeath];

% Transitions
%************
% constructing P
T = exp(-cumsum([0;(foi+sig+mu).*step]));
P = T(1:end-1)-T(2:end);
% constructing Q1
B1 = (sig./(foi+sig)).*(1-E');
C1 = (B1*ones(1,k)).*F;
Q1 = (1-T(2:end)./T(1:end-1)).*sum(C1)';
% constructing Q2
Q2 = sig./(foi+sig).*(((foi+sig+mu)./mu.*(1-exp(-mu.*step)))-(1-T(2:end)./T(1:end-1)));
% constructing Q
Q = exp(-cumsum([0;mu(1:end-1).*step(1:end-1)])).*(Q1+Q2);
```

```
nSI = sum(exp(-mu(1)*A)*foi./(foi+sig+mu).*(P+Q));
U1 = (1-exp(-mu.*step))./mu;
U2 = (ageint(:,1)-ageint(:,2).*exp(-mu.*step))./mu;
U3 = (1-T(2:end)./T(1:end-1))./(foi+sig+mu);
U4 = (ageint(:,1)-ageint(:,2).*T(2:end)./T(1:end-1))./(foi+sig+mu);
V1 = (L/N)*(foi./mu).*(Na(1:end-1).*(ageint(:,1)+U1)-Na(2:end).*(ageint(:,2)));
V2 = (L/N)*Na(1:end-1).*foi.*((U3./(foi+sig+mu)+U4).*sum(C)'+foi./(foi+sig)
    .*((U1./mu)+U2-(U3./(foi+sig+mu))-U4));
ASI = sum(V1-V2)/nSI;

T1 = U4+U3./(foi+sig+mu);
T3 = U2+U1./mu-T1;
radapt = sum(C)'+(foi./(foi+sig).*((U1./U3)-1));
radaptA = T1.*sum(C)'+T3.*foi./(foi+sig);
nRS = sum((L/N)*sig.*Na(1:end-1).*U1.*(radapt));
ARS = sum((L/N)*sig.*Na(1:end-1).*(radaptA))/nRS;

trans = [nSI ASI nRS ARS];

% Information criteria
%********************
aic = dev+2*length(parhat);
bic = dev+log(length(resp))*length(parhat);

end
```

REFERENCES

Anderson, R. and R. May (1991). *Infectious Diseases of Humans: Dynamics and Control*. Oxford: Oxford University Press.

Bell, L., S. Naides, P. Stoffman, R. Hodinka, and S. Plotkin (1989). Human parvovirus B19 infection among hospital staff members after contact with infected patients. *The New England Journal of Medicine 321*, 485–491.

Cartter, M., T. Farley, S. Rosengren, D. Quinn, S. Gillespie, G. Gary, and J. Hadler (1991). Occupational risk factors for infection with parvovirus B19 among pregnant women. *The Journal of Infectious Diseases 163*, 282–285.

Chorba, T., P. Coccia, R. Holman, P. Tattersall, L. Anderson, J. Sudman, N. Young, E. Kurczynski, U. Saarinen, R. Moir, et al. (1986). The role of parvovirus B19 in aplastic crisis and erythema infectiosum (fifth disease). *The Journal of Infectious Diseases 154*, 383–393.

Cohen, B. and M. Buckley (1988). The prevalence of antibody to human parvovirus B19 in England and Wales. *Journal of Medical Microbiology 25*, 151–153.

Gillespie, S., M. Cartter, S. Asch, J. Rokos, G. Gary, C. Tsou, D. Hall, L. Anderson, and E. Hurwitz (1990). Occupational risk of human parvovirus B19 infection for school and day-care personnel during an outbreak of erythema infectiosum. *Journal of the American Medical Association 263*, 2061–2065.

Goeyvaerts, N., N. Hens, B. Ogunjimi, M. Aerts, Z. Shkedy, P. Van Damme, and P. Beutels (2010). Estimating infectious disease parameters from data on social contacts and serological status. *Applied Statistics 59*, 255–277.

Gonçalves, G., A. Correia, P. Palminha, H. Rebelo de Andrade, and A. Alves (2005). Outbreaks caused by parvovirus B19 in three Portuguese schools. *Eurosurveillance 10*, pii=549.

Hens, N., Z. Shkedy, M. Aerts, C. Faes, P. Van Damme, and P. Beutels (2010). *Modeling Infectious Disease Parameters Based on Serological and Social Contact Data: a modern statistical perspective*. Springer-Verlag New York Inc. Forthcoming.

Kaufmann, J., J. Buccola, W. Stead, C. Rowley, M. Wong, and C. Bates (2007). Secondary symptomatic parvovirus B19 infection in a healthy adult. *Journal of General Internal Medicine 22*, 877–878.

Kelly, H., R. Siebert, R. Hammond, J. Leydon, and W. Maskill (2000). The age-specific prevalence of human parvovirus immunity in Victoria, Australia compared with other parts of the world. *Epidemiology and Infection 124*, 449–457.

Matsunaga, Y., N. Takeda, S. Yamazaki, K. Kamata, and D. Kurosawa (1995). Seroepidemiology of human parvovirus

B19 using recombinant VP1+VP2 particle antigen. *Kansenshogaku Zasshi 69*, 1371–1375.

Nascimento, J., M. Buckley, K. Brown, and B. Cohen (1990). The prevalence of antibody to human parvovirus B19 in Rio De Janeiro, Brazil. *Revista do Instituto de Medicina Tropical de Sao Paulo 32*, 41–45.

Nunoue, T., K. Okochi, P. Mortimer, and B. Cohen (1985). Human parvovirus (B19) and erythema infectiosum. *The Journal of Pediatrics 107*, 38–40.

Ogunjimi, B., N. Hens, N. Goeyvaerts, M. Aerts, P. Van Damme, and P. Beutels (2009). Using empirical social contact data to model person to person infectious disease transmission: an illustration for varicella. *Mathematical Biosciences 218*, 80–87.

Pillay, D., G. Patou, S. Hurt, C. Kibbler, and P. Griffiths (1992). Parvovirus B19 outbreak in a children's ward. *Lancet 339*, 107–109.

Rice, P. and B. Cohen (1996). A school outbreak of parvovirus B19 infection investigated using salivary antibody assays. *Epidemiology and Infection 116*, 331–338.

Rouderfer, V., N. Becker, and H. Hethcote (1994). Waning immunity and its effects on vaccination schedules. *Mathematical Biosciences 124*, 59–82.

Schoub, B., N. Blackburn, S. Johnson, and J. McAnerney (1993). Primary and secondary infection with human parvovirus B19 in pregnant women in South Africa. *South African Medical Journal 83*, 505–506.

Schwarz, T. F., M. Roggendorf, and F. Deinhardt (1987). Hufigkeit der Parvoviras-B19-Infektionen: Seroepidemiologische Untersuchungen. *Deutsche Medicinische Wochenschrift 112*, 1526–1531.

Valeur-Jensen, A., C. Pedersen, I. Westergaar, I. Jensen, M. Lebech, P. Andersen, et al. (1999). Risk factors for parvovirus B19 infection in pregnancy. *Journal of the American Medical Association 281*, 1099–1105.

Wallinga, J., P. Teunis, and M. Kretzschmar (2006). Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *American Journal of Epidemiology 164*, 936–944.

Wood, S. (2006). *Generalized Additive Models: an Introduction with R*. Chapman and Hall/CRC Press.

Woolf, A., G. Campion, A. Chishick, S. Wise, B. Cohen, P. Klouda, O. Caul, and P. Dieppe (1989). Clinical manifestations of human parvovirus B19 in adults. *Archives of Internal Medicine 149(5)*, 1153–1156.

Zaaijer, H., M. Koppelman, and C. Farrington (2004). Parvovirus B19 viraemia in Dutch blood donors. *Epidemiology and Infection 132*, 1161–1166.