Estimation of age-specific rates of reactivation and immune boosting of the varicella zoster virus

Isabella Marinelli[a,b], Alies van Lier[c], Hester de Melker[c], Andrea Pugliese[a], Michiel van Boven[c,*]

[a]*Dipartimento di Matematica, Università di Trento, Trento, Italy*
[b]*Basque Center for Applied Mathematics, Bilbao, Spain*
[c]*Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, the Netherlands*

Abstract

Studies into the impact of vaccination against the varicella zoster virus (VZV) have increasingly focused on herpes zoster (HZ), which is believed to be increasing temporally in vaccinated populations with decreasing infection pressure. This idea can be traced back to Hope-Simpson's hypothesis, in which a person's immune status determines the likelihood that he/she will develop HZ. Immunity decreases over time, and can be boosted by contact with a person experiencing varicella (exogenous boosting) or by a reactivation attempt of the virus (endogenous boosting). Here we use transmission models to estimate age-specific rates of reactivation and immune boosting, exogenous as well as endogenous, using zoster incidence data from the Netherlands (2002-2011, $n = 7,026$). The boosting and reactivation rates are estimated with splines, enabling these quantities to be optimally informed by the data. The analyses show that models with high levels of exogenous boosting and estimated or zero endogenous boosting, constant rate of loss of immunity, and reactivation rate increasing with age (to more than 5% per year in the elderly) give the best fit to the data. Estimates of the rates of immune boosting and reactivation are strongly correlated. This has important implications as these parameters determine the fraction of the population with waned immunity. We conclude that independent evidence on rates of immune boosting and reactivation in persons with waned immunity are needed to robustly predict the impact of varicella vaccination on the incidence of HZ.

Highlights

- We estimate age-specific rates of VZV reactivation and immune boosting using HZ incidence data

- We find high rates of exogenous and lower rates of endogenous immune boosting, while reactivation rates increase strongly with age

- A substantial part of the adult population is expected to be unavailable for reactivation leading to HZ

- High estimated rates of exogenous immune boosting lend credibility to the idea that HZ incidence depends on VZV circulation

---

*Corresponding author

Introduction

Varicella zoster virus (VZV) is a herpes virus causing a disease known as chickenpox or varicella. After resolving of the systemic primary infection, the virus remains latently present in the host. The virus can reactivate later in life, resulting in a disease known as herpes zoster (HZ) or shingles. Dating back to the work of Hope-Simpson [12], it has been hypothesized that boosting of the immune system of a latently infected person by contact with a person experiencing chickenpox could lower the probability of HZ developing at some point later in life [3, 10, 20, 19, 29]. If true, such exogenous immune boosting could have profound implications for vaccination programs aimed at reducing chickenpox, as these are expected to reduce virus transmission and exogenous immune boosting. Over the years this has become a popular hypothesis to explain and predict the dynamics of HZ in unvaccinated and vaccinated populations, and an important reserve for countries to introduce VZV vaccination [2, 3, 10, 25].

While earlier studies focused exclusively on the implications of exogenous boosting [1, 3, 10, 11, 21, 25, 31], Hope-Simpson's original hypothesis states that both exogenous and endogenous boosting, by reactivation attempts of the virus that are successfully countered by the immune system of the host, contribute to the maintenance of immunity. Although data are scarce, some studies indeed show subclinical reactivation in specific populations, most likely in response to stress [5, 14, 16, 24, 33]. It is therefore of particular interest to evaluate the joint ability of endogenous and exogenous boosting to explain the available HZ incidence data, and to obtain insight in the relative strength of endogenous versus exogenous immune boosting. Here we build on earlier studies [1, 3, 10, 11, 25, 21, 31] to obtain estimates of key parameters determining the dynamics of VZV in unvaccinated and vaccinated populations. Specifically, we use long-term HZ incidence data from the Netherlands (2002-2011, $n = 7,026$ cases [31]) to estimate age-specific rates of immune boosting, loss of immunity, and reactivation in a population without VZV vaccination. Our approach extends earlier analyses by i) including the possibility of endogenous next to exogenous immune boosting, and ii) allowing the age-specific rates of reactivation and immune boosting to have flexible shapes, using natural cubic splines. In this manner, the shapes and values of the rate functions are optimally informed by the data, in contrast to previous studies that have made restrictive assumptions from the onset (e.g., [10, 31]).

The analyses reveal that rates of loss of immunity are high ($0.048 - 0.10$ per year), that reactivation rates are low up to 40 years, increase strongly to $0.03 - 0.10$ per year at old age, and that the fraction of the population with waned immunity that is prone to HZ decreases from $45\% - 55\%$ at age 25 to $10\% - 40\%$ at age 80, and depends sensitively on whether or not endogenous boosting is taken into account. Our analyses show that exogenous boosting is stronger than endogenous immune boosting in children and adults with young children, and that it is conceivable that endogenous boosting exceeds exogenous boosting in other age strata. We discuss the implications for vaccination programs aiming to reduce chickenpox, and possible future avenues to obtain further quantitative information on the magnitude of the Hope-Simpson effect.
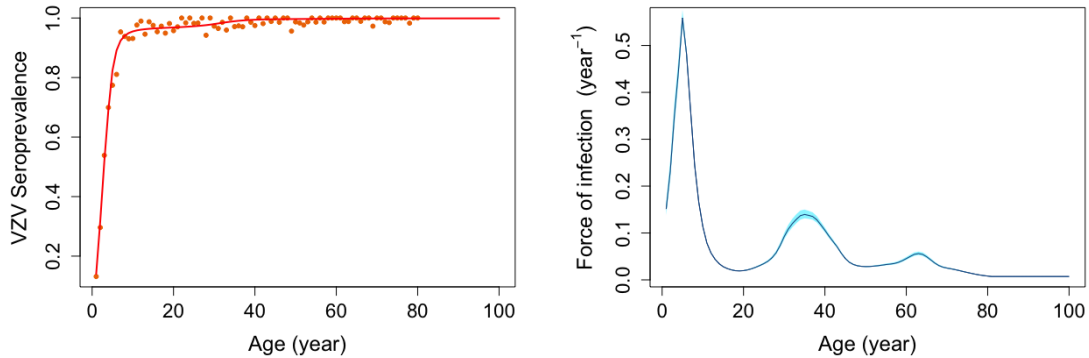
Figure 1: Overview of varicella zoster virus (VZV) seroprevalence (i.e. fraction of the population infected) (left-hand panel) and estimated force of infection (right-hand panel)[31]. The shaded area represents the 95% confidence band of the force of infection.

Methods

Varicella prevalence

The prevalence of VZV by age is based on Dutch serological data, derived from the second cross-sectional serosurveillance study (PIENTER2), conducted among people aged 0-79 years in the Netherlands in 2006/2007 [30]. The seroprevalence of VZV in the PIENTER2 study is quantitatively very similar to the seroprevalence in the earlier PIENTER1 study (conducted in 1995/1996), indicating little change of the age at infection with VZV over the period 1996-2006 [7, 32]. To avoid the interference of maternal antibodies we only include the 6,251 samples of participants aged 6 months or older. The VZV seroprevalence data provide the basis for estimation of the basic reproduction number of VZV using the so-called social contact hypothesis [31]. Here we use the earlier estimates of the force of infection as input for estimation of the rates of reactivation and immune boosting. The data and earlier model fit are shown in Figure 1, and can be downloaded in the data supplement of an earlier publication [31].

Herpes zoster incidence

Sentinel data on the incidence of general practitioner (GP) consultations due to HZ by age in the period 2002-2011 (based on 7,026 cases) are provided by NIVEL, the Netherlands Institute for Health Services Research [31]. The majority of HZ patients consult their GP because it is a painful condition. Because HZ complaints are highly specific and accompanied by typical lesions, with a positive predictive value of clinical judgment of 90.8% (95%CI: $87.3\% - 94.3\%$), we expect that misclassification of the diagnosis by the GP is rare [22]. Hence, we use the sentinel GP data as a proxy for the total incidence of HZ in the catchment population. The HZ data can be downloaded in the data supplement of an earlier publication [31], and shows no increase or decrease of incidence over time, and no relative increase or decrease in any age group in the study period. The HZ data are the main source of information on which estimates of the rates of reactivation and boosting are based.
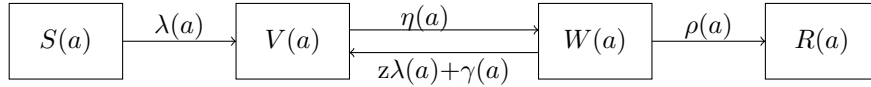
Figure 2: Schematic of the model under the short-disease approximation. $S(a)$ denotes the age-specific prevalence of uninfected persons, $V(a)$ the age-specific prevalence of latently infected persons with sufficient immunity to prevent reactivation, $W(a)$ the prevalence of latently infected persons with waned immunity, and $R(a)$ the prevalence of persons who have experienced HZ. The force of infection and rate of exogenous immune boosting are given by $\lambda(a)$ and $z\lambda(a)$, and the rates of loss of immunity, endogenous immune boosting, and reactivation are given by $\eta(a)$, $\gamma(a)$, and $\rho(a)$.

Model structure and analysis

The mathematical model to describe progression to HZ is based on earlier studies [10, 25, 31]. We extend these models by i) including endogenous next to exogenous immune boosting, as in Hope-Simpson's original hypothesis [12], and ii) allowing for the rates of reactivation and immune boosting to be relatively unconstrained, using natural cubic splines.

The population is assumed to be in demographic equilibrium, and infectious contacts (i.e. contacts sufficient for transmission) between persons of different ages are made according to the social contact hypothesis. This implies that the probability of an infectious contact is proportional to observed contact patterns in the Netherlands [17, 31]. We also assume, as in earlier studies and backed by empirical evidence, that primary varicella is infectious and that infectiousness of HZ is negligible [10, 28, 31]. Next, we assume in our main analyses that reactivation can occur at most once during a lifetime [12, 23]. We have also considered an extension of the main model in which multiple reactivations are possible, and obtained similar results (Appendix E). For this reason, we focus here on the model with at most one reactivation in a person's lifetime (Figure 2).

Further, the duration of varicella and HZ episodes (weeks to months) are short compared with the human lifespan (decades), and we may therefore use the short-disease approximation in which the varicella and HZ states are not explicitly modeled [9]. Finally, as in earlier studies [10, 25] but contrasting with others [15], we model the distribution of persons over classes by age with a fixed contact structure, with demographic composition as in the Netherlands during collection of the contact data.

Figure 2 shows a schematic of the model, and Equations (1)-(2) give a mathematical description. Individuals are born susceptible to VZV infection (in the S class). From the S compartment individuals are infected at an age-specific rate $\lambda(a)$ (the force of infection), after which they enter the V compartment (latently infected with immunity to reactivation). Immunity to reactivation is lost at a rate $\eta(a)$, and gained at rates $z\lambda(a)$ and $\gamma(a)$ by exogenous and endogenous boosting respectively. Reactivation from the latently infected compartment with waned immunity W occurs at a rate $\rho(a)$, and individuals who have experienced HZ after reactivation enter the R compartment.

Mathematically, the model dynamics are specified by the following set of ordinary differential equations

$$\frac{dS(a)}{da} = -\lambda(a)S(a),$$
$$\frac{dV(a)}{da} = \lambda(a)S(a) - \eta(a)V(a) + (z\lambda(a) + \gamma(a))W(a),$$
$$\frac{dW(a)}{da} = \eta(a)V(a) - (z\lambda(a) + \gamma(a) + \rho(a))W(a),$$
$$\frac{dR(a)}{da} = \rho(a)W(a) ,$$

(1)

with force of infection $\lambda(a)$ given by

$$\lambda(a) = q \int_0^M c(a,s)\lambda(s)S(s)ds ,$$

(2)

and initial conditions $S(0) = 1$, $V(0) = 0$, $W(0) = 0$, and $R(0) = 0$. In the above, $M = 100$ (yr) is the maximum age, $c(a,s)$ is the contact rate between individuals of age s and those of age a, and q is the proportionality parameter that has been estimated from seroprevalence data. Specifically, with proper scaling of the contact function the proportionality parameter q can be equated with the basic reproduction number $\mathcal{R}_0$. Throughout we take $\widehat{\mathcal{R}_0} = 8.7$ (95%CI: $8.1 - 9.3$), as estimated using seroprevalence data [31]. Notice that this implies that in the Netherlands infection of infants and children occurs at a relatively young age [18, 32]. Notice furthermore that the above implicitly implies that we do not use the explicit Equation (2) in the estimation procedure, but rather insert the estimate of force of infection $\widehat{\lambda(a)}$ in Equation (1) (Figure 1) as obtained earlier [31]. In this manner, all other estimates of transition rates are fully compatible with the serological data (Figure 1), at the (slight) cost that the force of infection may not be optimally informed by the HZ incidence data [31].

Discretization

Statistical inference of transition rates is based on a discretized approximation of the expanded state space [9, 10, 31]. First, we extend the state space by keeping track not only of whether individuals are latently infected with or without immunity to reactivation (i.e. in compartments V and W), but also by counting the number of times that individuals have passed through the V and W compartments. Hence, V_i and W_i refer to individuals who are in the V and W compartments for the i-th time. We can then solve Equation (1) recursively for all V_i and W_i ($i = 1, 2, \ldots$) in terms of the force of infection $\lambda(a)$. Next we approximate the system by assuming that individuals can only cycle through the V_i and W_i states for a certain number of times. Backed by preliminary analyses we take this number to be 5, as the error so introduced remains small when compared to results with higher maximal number of cycles (not shown). Finally, we discretize the solutions for V_i and W_i to be able to tie the modeled to the observed incidence of HZ. Since the HZ incidence data are available in age groups of 1 year, we also take age groups of 1 year in the discretization. Details are given in Appendix A.

Natural cubic splines

Flexible estimation of the transition rates $\eta(a)$, $\gamma(a)$, and $\rho(a)$ is ensured by using natural cubic splines [6]. A cubic spline is a smooth function that is constructed from cubic polynomial pieces that are joined

at certain knots, and that have continuous first and second derivatives [6]. For cubic splines the number of polynomial basis functions (pieces) is $m = n + 3$, where n is the number of age intervals. For simplicity, we assume throughout that the intervals are of the same length.

Next, each rate $f(a)$ can be written as a linear combination of elements of a B-spline basis $\{B_j\}_{j=1}^m$

$$f(a) = \sum_{j=1}^m \boldsymbol{\beta}_j B_j(a) \ ,$$

where $\boldsymbol{\beta}^T = (\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_m)$ are the coefficients of the basis functions that need to be estimated. Based on the above, the discretized age-specific rates $\mathbf{f}^T = (f(a_1), ...f(a_{100}))$ are given by

$$\mathbf{f} = \mathbf{B}\boldsymbol{\beta},$$

where $\mathbf{B} = (b_{ij}) = B_j(a_i)$. We evaluate the performance of models for a range of values for n using Akaike information criterion (AIC) and Bayesian information criterion (BIC) to select the optimal number of age intervals [4]. Appendix B shows results for $n = 4, ..., 7$. The analyses suggest that $n = 6$ is optimal (i.e. knots at $x_i = \frac{100i}{6}$ years, with $i = 0, 1, ..., 6$), and we will use this value throughout. Hence, $m = 9$ and $\mathbf{B}$ is a 100×9 matrix.

Estimation and model selection

Estimates of the parameters are obtained by maximization of the likelihood of the HZ incidence. As the number of cases is small compared to the number of person-years, both with regard to the total number of cases and total number of person-years ($7,026$ and $2,049,362$) and the numbers per age group, we use a Poisson likelihood with parameter $\rho_i W_i N_i$, where $\rho_i W_i$ is the expected (modeled) incidence rate of HZ in age group i, and N_i is the number of person-years. Confidence intervals of parameters (η, ρ, γ, and z) are calculated by direct resampling of the data. Here, 200 resampled data sets are generated assuming that the number of HZ cases in each of the age-classes follows a Binomial distribution with binomial totals N_i and parameter I_i/N_i), where N_i and I_i are the reported number of individuals and reported number of cases in age-class i, respectively. Throughout we use age classes of one year, so that with the exception of the youngest age classes (< 4 years) each age class contains more than 50 cases. For each of the resampled data sets the parameters are re-estimated, and percentile bootstrap confidence intervals are calculated. Model selection is based on AIC and BIC differences [4]. All analyses are performed using R 3.2 [26].

Scenarios

We consider a suite of scenarios to estimate the rate of loss of immunity to reactivation ($\eta(a)$), the rate of endogenous immune boosting ($\gamma(a)$), the relative susceptibility of latently infected individuals to exogenous immune boosting (z), and the reactivation rate ($\rho(a)$). The scenarios differ by whether the rates are assumed constant or estimated with age-dependent splines. In the baseline scenario (Scenario 1) all rates are estimated with splines. In Scenarios 2-4 one rate parameter is assumed constant at a time, in Scenarios 5-7 two rates are assumed constant, and in Scenario 8 all three rates are constant. Finally, based on preliminary analyses we also include a Scenario (Scenario 6*) in which the exogenous reactivation rate is

Scenario	Constant rates	Label	$log(L_{max})$	ΔAIC	ΔBIC	df
1	-		-347.5	16.0	60.2	28
2	γ		-348.0	0.8	24.2	20
3	ρ		-352.7	10.4	33.7	20
4	η		-351.9	8.7	32.1	20
5	γ, ρ		-360.2	9.2	11.8	12
6	γ, η	Model with endogenous boosting	-354.8	-1.6	0.98	12
6*	$\gamma \equiv 0, \eta$	Model without endogenous boosting	-356.6	**0**	**0**	11
7	ρ, η		-383.7	56.3	58.9	12
8	all rates		-1363.7	2000.3	1982.1	4

Table 1: Overview of model fits. A suite of simplifications of the baseline model (Scenario 1) are considered in which one or more of the parameters are constant instead of being modeled with splines. Shown are the maximized log-likelihood ($log(L_{max})$), the associated AIC and BIC differences with the best-fitting model Scenario 6* in the main text, and the degrees of freedom (df). The best-fitting model does not include endogenous immune boosting ($\gamma \equiv 0$), and has a constant rate of loss of immunity ($\widehat{\eta} = 0.076 \ (yr^{-1})$). See Table 2 for an overview of parameter estimates.

assumed to be zero ($\gamma \equiv 0$). All these scenarios include exogenous boosting. To validate this choice, we also consider a similar suite of scenarios that do not include exogenous boosting (see Appendix C). As expected, these models give a worse fit to the data.

Results

Table 1 gives an overview of fits of the different Scenarios based on AIC and BIC differences. Scenarios with all rates assumed constant (Scenario 8) or with constant rates of reactivation and loss of immunity (Scenario 7) do not perform well (ΔAIC and ΔBIC with the best fitting model are well over 50 in both cases). Better fits are obtained with models in which all parameters are estimated with splines (Scenario 1), in which one of the rate parameters is assumed constant (Scenarios 2-4), or in which both the rate of endogenous boosting (γ) and the reactivation rate (ρ) are constant (Scenario 5). Overall, the best-fitting models have a constant or zero rate of endogenous immune boosting, constant rate of loss of immunity, and age-specific rate of reactivation. We henceforth focus on these two models, which we will label the model with (Scenarios 6) and without (Scenarios 6*) endogenous boosting.

Parameter estimates of the best-fitting models with and without endogenous boosting are given in Table 2 and Figure 3. Both scenarios produce good and similar fits to the HZ incidence data (Figure 3). Waning of immunity occurs on a time scale that varies from 20 years ($\widehat{\eta} = 0.048$ per year ($95\%CI : 0.033 - 0.068$)) in the model with endogenous boosting to 13 years ($\widehat{\eta} = 0.076$ per year ($95\%CI : 0.063 - 0.087$)) in the model without endogenous boosting. Further, while the endogenous reactivation rate is 0 in the model without endogenous boosting, it is estimated at $\widehat{\gamma} = 0.093$ per year ($95\%CI : 0.069 - 0.11$) in the model with endogenous boosting. Susceptibility of latently infected individuals relative to the susceptibility of uninfected persons per infectious contact is estimated to be approximately 1 ($\widehat{z} = 1$ ($95\%CI : 0.0.998 - 1$) in the model with endogenous boosting and $\widehat{z} = 0.987$ ($95\%CI : 0.986 - 0.990$) in the model without endogenous boosting), this indicates that an infectious contact that would lead to infection of an uninfected person is

Scenario	Parameter		
	η (year^{-1}) (95%CI)	γ (year^{-1}) (95%CI)	z (95%CI)
6	0.048 $(0.033 - 0.068)$	0.093 $(0.069 - 0.114)$	1 $(0.998 - 1)$
6*	0.076 $(0.063 - 0.087)$	-	0.987 $(0.986 - 0.990)$

Table 2: Parameter estimates of the scenarios with and without endogenous immune boosting (cf. Table 1). Shown are maximum likelihood estimates with 95% confidence bounds for parameters that are assumed constant. See Figure 2 for parameters estimated with splines.

also sufficient to lead to immune boosting. As a consequence, exogenous immune boosting almost equals the force of infection (Figure 1), and in children and in the age group of adults with young children ($\approx 30 - 40$ years) exogenous immune boosting exceeds endogenous boosting (if present). In older adults (> 40 years) and older children and young adults (< 30 years) who have a low estimated force of infection, exogenous immune boosting is expected to be less prominent, and endogenous boosting exceeds exogenous boosting in the model with immune boosting, perhaps with exception of grandparents with young grandchildren (≈ 65 years).

Estimated reactivation rates are age-dependent in both scenarios with high statistical support, increasing with age from, say, 40 years onwards. Specifically, estimates are approximately 0.10 per year in elderly in the model with endogenous immune boosting and over 0.02 per year in the model without endogenous immune boosting (Figure 3). Interestingly, the reactivation rates are substantially higher in the model with endogenous boosting than in the model without endogenous boosting from the age of 60 years onwards, while the estimated fraction that is available for reactivation $(\widehat{W(a)})$ is substantially higher in the model without endogenous immune boosting (Figure 3). Hence, both scenarios make different interpretations on the fraction of the population that is available for reactivation in older adults and elderly. Hence, while the models provide similar fits to the HZ incidence data, there are significant underlying differences with regard to the rates of endogenous immune boosting, reactivation, and fraction of the population that is available for reactivation.

Although it is expected that the majority of cases of HZ is reported to a GP [22], some degree of under-ascertainment is conceivable. Additionally, even though HZ symptoms are highly specific, it is possible that some reported cases of HZ may have been misdiagnosed. To investigate the robustness of the above results to under- and over-ascertainment, we have have reanalyzed the data assuming 10% under-ascertainment and 10% over-ascertainment. The results of the analyses are laid down in Appendix D (see Table 5 and Figure 5 for details) and show that the results with either under- or over-ascertainment are quantitatively close to the results obtained in Tables 1-2 and Figure 3.

Discussion

Our analyses provide support for the hypothesis of Hope-Simpson [12] that immune boosting affects the incidence of HZ. Our results add to earlier studies [1, 10, 11, 25, 31] by showing that immune boosting is a potent force in a model that includes both endogenous and exogenous boosting, and that allows
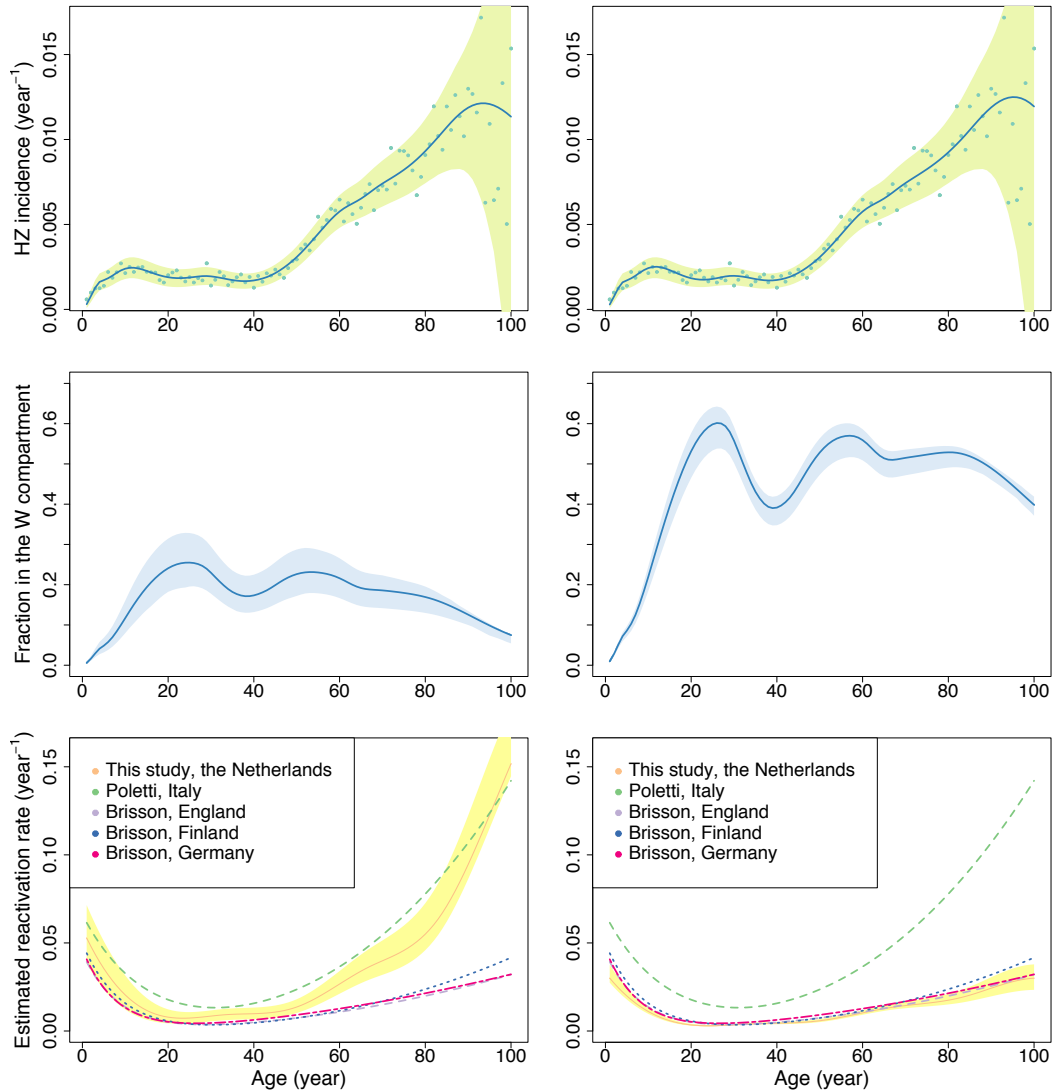
Figure 3: HZ incidence data and overview of the best-fitting models (left-hand panels: with endogenous immune boosting, right-hand panels: without endogenous immune boosting). Shown are the observed and predicted incidence of HZ in both scenarios (top panels; dots: data; lines: prediction), the fraction of the population with waned immunity ($W(a)$) (middle panels), and the estimated reactivation rates ($\rho(a)$)(bottom panels). Lines correspond to maximum likelihood estimates and shaded areas to bootstrapped 95% confidence ranges. Also shown are the reactivation rates as estimated for Italy, Finland, the UK, and Germany in more restrictive analyses [3, 25].

the reactivation rate to take flexible shapes. Our quantitative estimations further show that in most age groups the rate of exogenous immune boosting is substantially higher than the rate of endogenous immune boosting, but also underline that even low estimated rates of endogenous immune boosting significantly affect the distribution of latently infected individuals with and without waned immunity. More generally, the distribution of latently infected individuals with and without immunity to reactivation is shaped by an intricate interplay between the age-specific rates at which immunity is lost and gained (see also [1, 11]), and even small rates of endogenous boosting could affect this distribution significantly. The main reason is that there is no direct information underpinning which part of the adult population is latently infected with and without waned immunity. The implication is that not all individuals who are latently infected may be available for developing HZ, but that the exact fraction cannot be estimated with precision. Hence, for public health considerable uncertainty remains about the potential effectiveness of HZ vaccination of elderly.

Waning of immunity and, to a lesser extent, immune boosting are well documented, and have been included in an agent-based transmission model to anticipate the impact of varicella vaccination of HZ incidence [21]. In this model, immunity is a continuous quantity, and each individual is equipped with his/her own level of (cellular) immunity. This contrasts with our model in which immunity to reactivation is either complete (compartment V) or completely absent (compartment W). From a biological perspective, the earlier model [21] is biologically more plausible, while our current model is simpler and more amenable to formal statistical inference. In both models, much of the remaining variation and uncertainty derives from the fact that there is very little data to pinpoint exactly which fraction of the adult population is available for reactivation.

A number of assumptions need scrutiny. First, we have assumed, as in earlier studies (e.g., [10, 25]), that the force of infection is known. Here we have taken the force of infection as estimated from Dutch seroprevalence data [31]. Ideally, all parameters of interest would be estimated in one go, taking into account both the seroprevalence and HZ incidence data. This is, however, computationally demanding. Further, it has been argued that the basic reproduction number and force of infection are strongly identified by serological data, thereby diminishing the need to include all data in one synthesizing analysis [31]. Earlier studies have taken a similar approach in which the force of infection is estimated first, and then in a second step the transition rates are estimated from HZ incidence [10, 25, 31].

Second, our analyses are based on the arbitrary assumption that the splines have five interior knots. This seems to produce reasonably flexible shapes of the reactivation rate without grossly overparameterizing the model. It should be noted, however, that whether or not models with constant or flexible rates of boosting, loss of immunity, and reactivation are preferred by information criteria such as AIC or BIC depends not only on the data but also on the number of additional parameters (and hence the number of knots). A computationally more involved alternative could for instance make use of penalized splines in which a penalty is used to avoid excessive fluctuations of the rate functions [8]. This, however, would also necessitate estimation of the effective number of parameters, and cannot address the question whether it is likely or not that certain parameters are constant, as is invariably assumed in modeling studies. In a sensitivity analysis we have decreased and increased the number of knots and obtained quantitatively similar

results, showing that our results do not depend sensitively on the degree of freedom of the rate functions (Appendix B).

Figure 3 shows the reactivation rates as estimated using our model from Dutch HZ incidence data together with estimates of the reactivation rates as estimated earlier using HZ incidence data from Germany, Finland, Italy, and the UK [3, 25]. Interestingly, our estimates are close to earlier estimates by Brisson and colleagues up to the age of 60 [3]. In older age groups our estimates of the reactivation rates are substantially higher in the scenario with endogenous immune boosting. This can be attributed to the fact that in this scenario more older persons have functional immunity due to endogenous boosting, and higher reactivation rates are required to explain the HZ incidence data. Also interesting is the finding in our study and the earlier studies by Brisson and colleagues and Poletti and colleagues that the estimated reactivation rate decreases in childhood until the age of 23 years. In the earlier studies one could have argued that this is a consequence of the relative inflexibility of the function describing the reactivation rate. This explanation, however, does not hold in our analyses, and suggests that the high estimated reactivation rates in children may reflect a true biological phenomenon. A reasonable explanation might be that for some reason the immune system of children has not (yet) acquired full control of the virus.

Our results have potential implications for large-scale varicella and combined varicella and zoster vaccination programs. Our two main scenarios (Figure 3) make different explanations on how the observed zoster cases are produced. In the model with endogenous boosting the fraction of persons with waned immunity is small while the reactivation rate is high, while in the model without endogenous boosting the fraction with waned immunity is higher but the reactivation rate is lower. Indiscriminate zoster vaccination is expected to prevent similar number of cases of HZ in both cases. However, this expectation may not hold anymore if large-scale varicella vaccination is added to the mix. In this case, the force of infection is expected to be reduced to a similar extent in both scenarios as estimates of the exogenous boosting parameter z are very close in both scenarios (Table 2). As a result, the impact of exogenous boosting is expected to be reduced to a similar extent after the introduction of varicella vaccination. However, immune boosting is still a factor in the model with endogenous boosting (cf. Table 2). As a consequence, one would expect that persons that are exogenously boosted in the absence of varicella vaccination could still be boosted endogenously in the presence of varicella vaccination. This could potentially lead to a less dramatic transient increase in the incidence of HZ after the introduction of varicella vaccination than predicted earlier [3, 10, 11, 13, 27, 31]. The downside would be that the added value of zoster vaccination would also be diminished. Of course, in an ideal scenario it would be possible to target zoster vaccination specifically to those persons with waned immunity who are at risk of developing zoster.

Acknowledgements

References

[1] M. Betta, M. Laurino, A. Pugliese, G. Guzzetta, A. Landi, and P. Manfredi. Perspectives on optimal control of varicella and herpes zoster by mass routine varicella vaccination. *Proceedings. Biological sciences / The Royal Society*, 283(1826):20160054, mar 2016.

[2] P. Bonanni, J. Breuer, A. Gershon, M. Gershon, W. Hryniewicz, V. Papaevangelou, B. Rentier, H. Rümke, C. Sadzot-Delvaux, J. Senterre, C. Weil-Olivier, and P. Wutzler. Varicella vaccination in Europe - taking the practical approach. *BMC medicine*, 7(1):26, jan 2009.

[3] M. Brisson, G. Melkonyan, M. Drolet, G. De Serres, R. Thibeault, and P. De Wals. Modeling the impact of one- and two-dose varicella vaccination on the epidemiology of varicella and zoster. *Vaccine*, 28(19):3385–3397, apr 2010.

[4] K. P. Burnham and D. R. Anderson. *Model Selection and Multimodel Inference*, volume 45. Springer, 2003.

[5] R. J. Cohrs, S. K. Mehta, D. S. Schmid, D. H. Gilden, and D. L. Pierson. Asymptomatic reactivation and shed of infectious varicella zoster virus in astronauts. *Journal of Medical Virology*, 80(6):1116–1122, jun 2008.

[6] C. de Boor. A Practical Guide to Splines - Revised Edition. *Springer-Verlag, New York*, page 366, 2001.

[7] H. de Melker, G. Berbers, Susan Hahé, H. Rümke, S. van den Hof, A. de Wit, and H. Boot. The epidemiology of varicella and herpes zoster in The Netherlands: Implications for varicella zoster virus vaccination. *Vaccine*, 24(18):3946–3952, may 2006.

[8] P. H. C. Eilers and B. D. Marx. Flexible smoothing with B -splines and penalties. *Statistical Science*, 11(2):89–121, 1996.

[9] N. Goeyvaerts, N. Hens, B. Ogunjimi, M. Aerts, Z. Shkedy, P. Van Damme, and P. Beutels. Estimating infectious disease parameters from data on social contacts and serological status. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, 59(2):255–277, mar 2010.

[10] G. Guzzetta, P. Poletti, E. Del Fava, M. Ajelli, G. P. Scalia Tomba, S. Merler, and P. Manfredi. Hope-Simpson's progressive immunity hypothesis as a possible explanation for Herpes Zoster incidence data. *American Journal of Epidemiology*, 177(10):1134–1142, may 2013.

[11] G. Guzzetta, P. Poletti, S. Merler, and P. Manfredi. The Epidemiology of Herpes Zoster After Varicella Immunization Under Different Biological Hypotheses: Perspectives From Mathematical Modeling. *American Journal of Epidemiology*, 183(8):765—-773, 2016.

[12] R. E. Hope-Simpson. The Nature of Herpes Zoster: A Long-term Study and a New Hypothesis. *Proc.R.Soc.Med.*, 58(0035-9157 (Print)):9–20, jan 1965.

[13] M. Karhunen, T. Leino, H. Salo, I. Davidkin, T. Kilpi, and K. Auranen. Modelling the impact of varicella vaccination on varicella and zoster. *Epidemiology and infection*, 138(4):469–81, 2010.

[14] P. Ljungman, B. Lönnqvist, G. Gahrton, O. Ringdén, V. A. Sundqvist, and B. Wahren. Clinical and subclinical reactivations of varicella-zoster virus in immunocompromised patients. *Journal of Infectious Diseases*, 153(5):840–847, 1986.

[15] V. Marziano, P. Poletti, G. Guzzetta, M. Ajelli, P. Manfredi, and S. Merler. The impact of demographic changes on the epidemiology of herpes zoster: Spain as a case study. *Proceedings of the Royal Society of London B: Biological Sciences*, 282(1804), 2015.

[16] S. K. Mehta, R. J. Cohrs, B. Forghani, G. Zerbe, D. H. Gilden, and D. L. Pierson. Stress-Induced Subclinical Reactivation of Varicella Zoster Virus in Astronauts. *Journal of Medical Virology*, 72(1):174–179, jan 2004.

[17] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J. Edmunds. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*, 5(3):0381–0391, mar 2008.

[18] A. Nardone, F. de Ory, M. Carton, D. Cohen, P. van Damme, I. Davidkin, M. C. Rota, H. de Melker, J. Mossong, M. Slacikova, A. Tischer, N. Andrews, G. Berbers, G. Gabutti, N. Gay, L. Jones, S. Jokinen, G. Kafatos, M. V. de Aragón, F. Schneider, Z. Smetana, B. Vargova, R. Vranckx, and E. Miller. The comparative sero-epidemiology of varicella zoster virus in 11 countries in the European region. *Vaccine*, 25(45):7866–7872, nov 2007.

[19] B. Ogunjimi, E. Smits, S. Heynderickx, J. Van Den Bergh, J. Bilcke, H. Jansens, R. Malfait, J. Ramet, H. T. Maecker, N. Cools, P. Beutels, and P Van Damme. Influence of frequent infectious exposures on general and varicella-zoster virus-specific immune responses in pediatricians. *Clinical and Vaccine Immunology*, 21(3):417–426, 2014.

[20] B. Ogunjimi, P. Van Damme, and P. Beutels. Herpes Zoster Risk Reduction through Exposure to Chickenpox Patients: A Systematic Multidisciplinary Review. *PLoS ONE*, 8(6), jun 2013.

[21] B. Ogunjimi, L. Willem, P. Beutels, and N. Hens. Integrating between-host transmission and within-host immunity to analyze the impact of varicella vaccination on zoster. *eLife*, 4(AUGUST2015), jul 2015.

[22] W. Opstelten, A. M. van Loon, M. Schuller, A. J. van Wijck, G. A. van Essen, K. G. Moons, and T. J. Verheij. Clinical diagnosis of herpes zoster in family practice. *Ann Fam Med*, 5(4):305–309, 2007.

[23] M. N. Oxman. Herpes zoster pathogenesis and cell-mediated immunity and immunosenescence. *The Journal of the American Osteopathic Association*, 109(6 Suppl 2):S13–7, jun 2009.

[24] V. Papaevangelou, M. Quinlivan, J. Lockwood, O. Papaloukas, G. Sideri, E. Critselis, I. Papassotiriou, J. Papadatos, and J. Breuer. Subclinical VZV reactivation in immunocompetent children hospitalized in the ICU associated with prolonged fever duration. *Clinical Microbiology and Infection*, 19(5), may 2013.

[25] P. Poletti, A. Melegaro, M. Ajelli, E. del Fava, G. Guzzetta, L. Faustini, G. Scalia Tomba, P. Lopalco, C. Rizzo, S. Merler, and P. Manfredi. Perspectives on the Impact of Varicella Immunization on Herpes Zoster. A Model-Based Evaluation from Three European Countries. *PLoS ONE*, 8(4), apr 2013.

[26] R Foundation For Statistical Computing R Development Core Team. R: A Language and Environment for Statistical Computing. *Vienna Austria R Foundation for Statistical Computing*, 1(10):ISBN 3–900051–07–0, 2008.

[27] M. C. Schuette and H. W. Hethcote. Modeling the effects of varicella vaccination programs on the incidence of chickenpox and shingles. *Bulletin of mathematical biology*, 61(6):1031–64, 1999.

[28] H. E. Seiler. A study of herpes zoster particularly in its relationship to chickenpox. *The Journal of Hygiene*, 47(3):253–262, 1949.

[29] S. L. Thomas, J. G. Wheeler, and A. J. Hall. Contacts with varicella or with children and protection against herpes zoster in adults: A case-control study. *Lancet*, 360(9334):678–682, aug 2002.

[30] F. R. M. Van Der Klis, L. Mollema, G. A. M. Berbers, H. E. De Melker, and R. A. Coutinho. Second national serum bank for population-based seroprevalence studies in the Netherlands. *Netherlands Journal of Medicine*, 67(7):301–308, jul 2009.

[31] A. van Lier, A. Lugnér, W. Opstelten, P. Jochemsen, J. Wallinga, F. Schellevis, E. Sanders, H. de Melker, and M. van Boven. Distribution of Health Effects and Cost-effectiveness of Varicella Vaccination are Shaped by the Impact on Herpes Zoster. *EBioMedicine*, 2(10):1494–1499, oct 2015.

[32] A. Van Lier, G. Smits, L. Mollema, S. Waaijenborg, G. Berbers, F. Van Der Klis, H. Boot, J. Wallinga, and H. De Melker. Varicella zoster virus infection occurs at a relatively young age in the Netherlands. *Vaccine*, 31(44):5127–5133, oct 2013.

[33] A. Wilson, M. Sharp, C. M. Koropchak, S. F. Ting, and A. M. Arvin. Subclinical varicella-zoster virus viremia, herpes zoster, and t lymphocyte immunity to varicella-zoster viral antigens after bone marrow transplantation. *Journal of Infectious Diseases*, 165(1):119–126, 1992.

Appendix A. Model Solution and Discretization

We base our estimation procedures on the ODEs specified by Equations (1)-(2). In the analyses we use earlier estimates of the force of infection $\lambda(a)$ [31], and obtain maximum likelihood estimates of other parameters using the discretized solution of Equation (1). Specifically, the population is split into $J = 100$ age classes of 1 year.

As a first step we expand the state space by keeping track of the number of times that an individual has been in the V and W compartments [10]. Hence, $V_i(a)$ represents the fraction of the population of age a who are in the V compartment for the i-th time ($i = 1, 2, \ldots$). See Figure 4 for a schematic of the model with expanded state space (see also [10]). The ODEs of Equation 1 for $V(a)$ and $W(a)$ are replaced in the expanded model by

$$
\begin{aligned}
\frac{dV_1(a)}{da} &= \lambda(a)S(a) - \eta(a)V_1(a) \\
\frac{dW_1(a)}{da} &= \eta(a)V_1(a) - (z\lambda(a) + \gamma(a) + \rho(a))W_1(a) \\
&\quad \ldots \\
\frac{dV_i(a)}{da} &= (z\lambda(a) + \gamma(a))W_{i-1}(a) - \eta(a)V_i(a) \\
\frac{dW_i(a)}{da} &= \eta(a)V_i(a) - (z\lambda(a) + \gamma(a) + \rho(a))W_i(a) \ ,
\end{aligned}
\tag{3}
$$

with $i = 2, 3, \ldots$ and appropriate initial conditions. In the above, the (total) fractions $V(a)$ and $W(a)$ in the latently infected compartments are given by infinite sums $V(a) = \sum_{i=1}^{\infty} V_i(a)$ and $W(a) = \sum_{i=1}^{\infty} W_i(a)$. Reasonable approximations for these sums can be obtained by assuming that there is a finite but sufficiently large number of times n that an individual can pass through the V and W compartments during its lifetime. Hence, $V_i(a)$ and $W_i(a)$ are approximated by

$$
V(a) \approx \sum_{i=1}^{n} V_i(a) \qquad \text{and} \qquad W(a) \approx \sum_{i=1}^{n} W_i(a) \ ,
$$

with $n \in \mathbb{Z}^+$. In practice, these approximations work well already for small n [10]. Here we take $n = 5$, which gives results that quantitatively close to those obtained when taking $n = 10$ (results not shown).

The ODEs with expanded state space can be solved recursively, starting with $S(a)$, $V_1(a)$, and $W_1(a)$. The solution is given by

$$
\begin{aligned}
S(a) &= e^{-\int_0^a \lambda(s)ds} \\
V_1(a) &= \int_0^a \lambda(s)S(s)e^{-\int_s^a \eta(r)dr}ds \\
W_1(a) &= \int_0^a \eta(s)V_1(s)e^{-\int_s^a [z\lambda(r)+\gamma(r)+\rho(r)]dr}ds \\
V_i(a) &= \int_0^a (z\lambda(s) + \gamma(s))W_{i-1}(s)e^{-\int_s^a \eta(r)dr}ds \\
W_i(a) &= \int_0^a \eta(s)V_i(s)e^{-\int_s^a [z\lambda(r)+\gamma(r)+\rho(r)]dr}ds
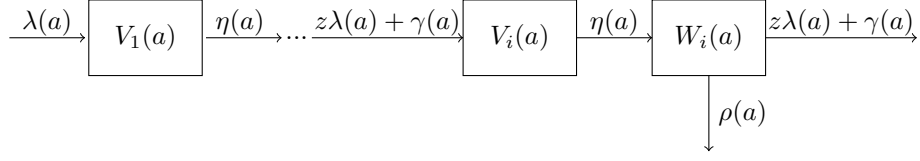\end{aligned}
\tag{4}
$$

14

Figure 4: Schematic of the model with expanded state space (cf. Figure 2).

with $i \geq 2$ and $S(0) = 1$ the only non-zero initial condition. It follows that $V(a)$ and $W(a)$ can be approximated by

$$V(a) \approx V_1(a) + \sum_{i=2}^{n} \phi_1(W_{i-1})(a)$$

$$W(a) \approx W_1(a) + \sum_{i=2}^{n} (\phi_1 \circ \phi_2)(W_{i-1})(a) \ ,$$

where

$$\phi_1(f) = \int_0^a (z\lambda(s) + \gamma(s))f(s)e^{-\int_s^a \eta(\tau)d\tau}ds$$

$$\phi_2(f) = \int_0^a \eta(s)f(s)e^{-\int_s^a [z\lambda(\tau)+\gamma(\tau)+\rho(\tau)]d\tau}ds \ .$$

Next, we discretize Equation (4) as follows. If the limits of the $J = 100$ age classes are defined by the vector $\mathbf{a} = (0, 1, ..., J)$, then the jth class contains all persons with age in the interval $[a_{[j]}, a_{[j+1]})$, where $a_{[j]}$ denotes the j-th element of $\mathbf{a}$. Using this notation, the force of infection $\lambda(a)$ and the rates $\eta(a)$, $\rho(a)$, and $\gamma(a)$ are replaced by their discretized counterpart λ_j, η_j, ρ_j, and γ_j, and V_i and W_i $(i, = 1, ..., n, j = 1, ..., J)$ in the j-th age-class are given by

$$(V_1)_j = \sum_{k=1}^{j} \frac{\lambda_k}{\lambda_k - \eta_k} \exp\left(-\sum_{i=1}^{k-1} \lambda_i(a_{i+1} - a_i)\right) \exp\left(-\sum_{i=k+1}^{j} \eta_i(a_{i+1} - a_i)\right)$$
$$\times \left[\exp\left(-\eta_k(a_{k+1} - a_k)\right) - \exp\left(-\lambda_k(a_{k+1} - a_k)\right)\right]$$

$$(W_1)_j = \sum_{k=1}^{j} \frac{\eta_k}{z\lambda_k + \gamma_k + \rho_k} (V_1)_k \exp\left(-\sum_{i=k+1}^{j} (z\lambda_i + \gamma_i + \rho_i)(a_{i+1} - a_i)\right)$$
$$\times \left[1 - \exp\left(-(z\lambda_k + \gamma_k + \rho_k)(a_{k+1} - a_k)\right)\right]$$

$$(V_i)_j = \sum_{k=1}^{j} \frac{z\lambda_k + \gamma_k}{\eta_k} (W_{i-1})_k \exp\left(-\sum_{i=k+1}^{j} \eta_i(a_{i+1} - a_i)\right)$$
$$\times \left[1 - \exp\left(-\eta_k(a_{k+1} - a_k)\right)\right]$$

$$(W_i)_j = \sum_{k=1}^{j} \frac{\eta_k}{z\lambda_k + \gamma_k + \rho_k} (V_i)_k \exp\left(-\sum_{i=k+1}^{j} (z\lambda_i + \gamma_i + \rho_i)(a_{i+1} - a_i)\right)$$
$$\times \left[1 - \exp\left(-(z\lambda_k + \gamma_k + \rho_k)(a_{k+1} - a_k)\right)\right].$$

15

Appendix B. Optimal Number of Age Intervals

To determine the optimal number of age intervals we have re-analyzed the data using the scenarios of Table 1, while varying the number of age intervals from $n = 4$ to $n = 7$ (5 to 8 knots). The results are shown in Table 3 below.

Number of age intervals	Scenario	Rates assumed constant	$log(L_{max})$	ΔAIC	ΔBIC	df
	1	-	-355.4	19.7	48.2	22
	2	γ	-357.2	11.3	24.3	16
	3	ρ	-359.5	15.9	28.8	16
4	4	η	-356.8	10.6	23.5	16
	5	γ, ρ	-370.3	25.5	22.9	10
	6	γ, η	-362.2	13.3	15.9	10
	6*	$\gamma \equiv 0, \eta$	-382.3	51.5	51.5	9
	7	ρ, η	-413.9	112.7	110.0	10
	1	-	-354.9	24.7	61.1	25
	2	γ	-353.7	8.4	26.6	18
	3	ρ	-357.9	16.7	35.0	18
5	4	η	-355.5	12.0	30.2	18
	5	γ, ρ	-370.9	28.8	28.7	11
	6	γ, η	-369.0	26.9	29.4	11
	6*	$\gamma \equiv 0, \eta$	-380.0	46.9	46.9	10
	7	ρ, η	-395.4	77.8	77.7	11
	1	-	-347.5	16.0	60.2	28
	2	γ	-348.0	0.8	24.2	20
	3	ρ	-352.7	10.4	33.7	20
6	4	η	-351.9	8.7	32.1	20
	5	γ, ρ	-360.2	9.2	11.8	12
	6	γ, η	-354.8	-1.6	0.98	12
	6*	$\gamma \equiv 0, \eta$	-356.6	**0**	**0**	11
	7	ρ, η	-383.7	56.3	58.9	12
	1	-	-352.9	32.8	84.9	31
	2	γ	-346.9	2.7	31.3	22
	3	ρ	-353.8	16.5	45.1	22
7	4	η	-353.3	15.6	44.2	22
	5	γ, ρ	-356.0	2.8	8.0	13
	6	γ, η	-358.7	8.3	13.5	13
	6*	$\gamma \equiv 0, \eta$	-403.5	95.8	98.4	12
	7	ρ, η	-373.4	37.7	42.8	13

Table 3: Overview of analyses with varying number of age intervals. See main text for details.

Appendix C. A Model Without Exogenous Boosting

In the main text we focus on model scenarios with exogenous immune boosting, as put forward by Hope-Simpson [12]. Here we provide a sensitivity analysis showing that model scenarios without exogenous boosting have low empirical support compared to scenarios with exogenous boosting. In Table 4 we present $log(L_{max})$, and AIC and BIC differences of these models with the best-fitting model Scenario 6* in the main text.

Scenario	Rates assumed constant	$log(L_{max})$	ΔAIC	ΔBIC	df
1	-	-360.5	40.0	81.7	27
2	γ	-351.9	6.76	27.6	19
3	ρ	-358.9	20.6	41.5	19
4	η	-358.8	20.4	41.2	19
5	γ, ρ	-373.8	34.5	34.5	11
6	γ, η	-366.7	20.2	20.2	11
6*	$\gamma \equiv 0, \eta$	-379.5	44.0	41.4	10
7	ρ, η	-387.5	61.8	61.8	11
8	all rates	-2996.8	5266.5	5248.3	3

Table 4: Overview of fits of model scenarios without exogenous boosting (cf. Table 1). A suite of simplifications of the baseline model (Scenario 1) are considered in which one or more of the parameters are constant instead of being modeled with splines. Shown are the maximized log-likelihood ($log(L_{max})$), the associated AIC and BIC differences with the best-fitting model Scenario 6* in the main text, and the degrees of freedom (df).

Appendix D. Under- and Over-Ascertainment

To further investigate the robustness of our estimates, we analyze two additional scenarios assuming 10% under-ascertainment and 10% over-ascertainment. Specifically, we increase or decrease the observed number of HZ infections by 10%, and re-estimate the parameters. Table 5 and Figure 5 show results of the best-fitting Scenarios 6 and 6*.

Scenario	Assumption	Parameter		
		η (year^{-1})	γ (year^{-1})	z
6	10% over-ascertainment	0.048	0.11	1
6	10% under-ascertainment	0.053	0.080	1
6*	10% over-ascertainment	0.076	-	0.981
6*	10% under-ascertainment	0.078	-	0.987

Table 5: Parameter estimates of model scenarios with under- or over-ascertainment and with or without endogenous immune boosting (cf. Table 1). Shown are maximum likelihood estimates with 95% confidence bounds for parameters that are assumed constant. See also Figure 5.
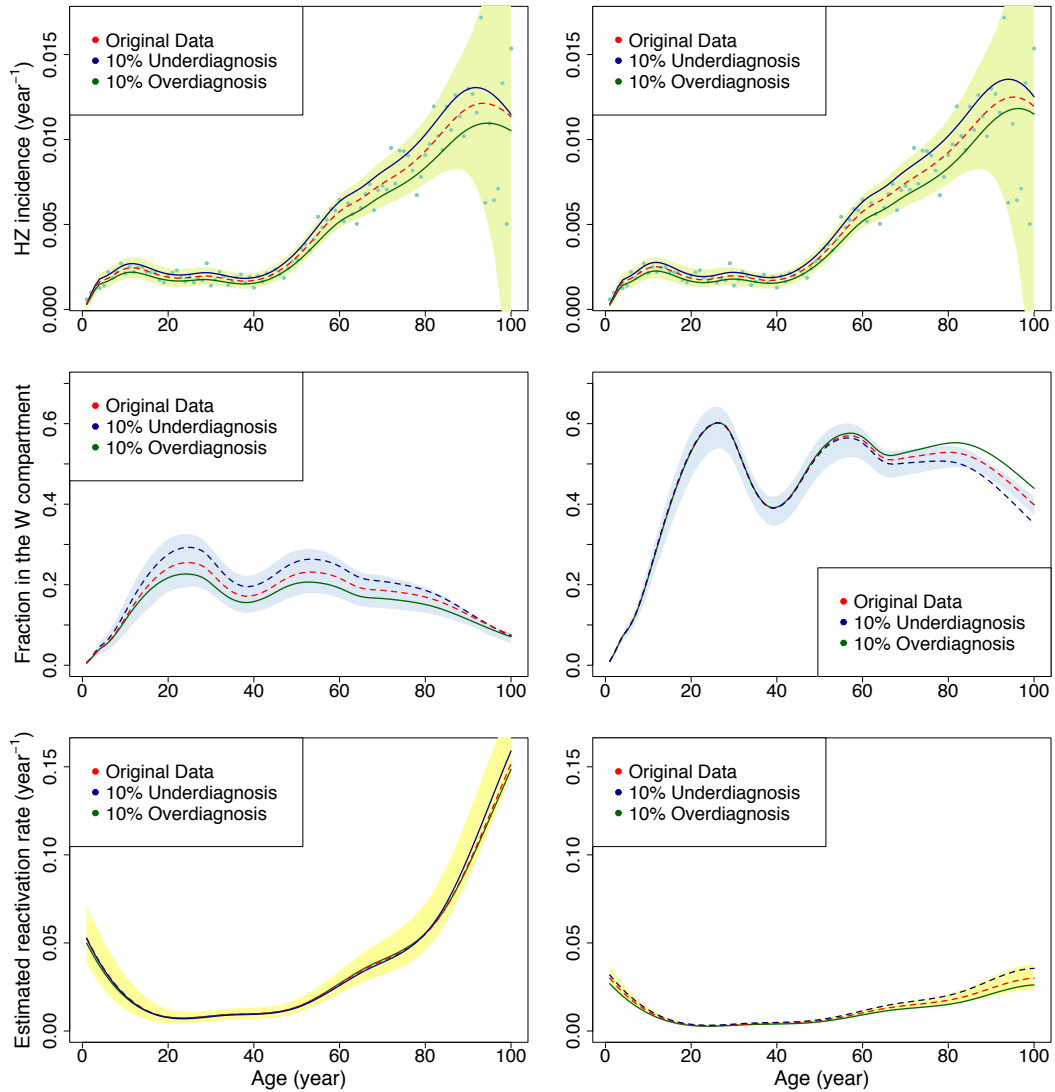
Figure 5: Model fits in the presence of under- and over-ascertainment(left-hand panels: with endogenous immune boosting; right-hand panels: without endogenous immune boosting). Shown are the observed and predicted incidence of HZ in both scenarios (top panels), the fraction of the population with waned immunity ($W(a)$) (middle panels), and the estimated reactivation rates ($\rho(a)$)(bottom panels). Dashed lines correspond to maximum likelihood estimates presented in the main text (Figure 3) for the scenarios. Continuous lines refer to maximum likelihood estimates assuming 10% under-ascertainment (blue) and 10% over-ascertainment (green).

Appendix E. A Model with Multiple Reactivations

The model in the main text assumes that only one reactivation can take place in a person's life. Here we investigate whether the results so obtained still hold in a model with multiple reactivations. The model structure is shown in Figure 6. In order not to overcomplicate the model we assume that reactivation leads to immune boosting, and that immunity wanes at an age-specific rate $\eta(a)$ that does not depend on the number of reactivation events that a person has already gone through. The results of the analyses are given in Tables 6-7 and Figure 7 (cf. Tables 1-2 and Figure 3). Overall, these analyses give similar fits to the data as analyses with models that do not include multiple reactivation (Table 7), and mainly differ with respect to the estimated (unobserved) fraction of the population in the W compartment (Figure 7).
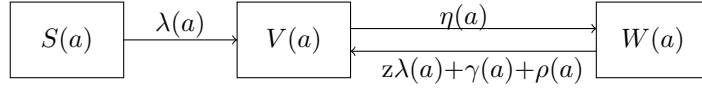

Figure 6: Schematic of the model with multiple reactivations. $S(a)$ denotes the age-specific prevalence of uninfected persons, $V(a)$ the age-specific prevalence of latently infected persons with sufficient immunity to prevent reactivation, and $W(a)$ the prevalence of latently infected persons with waned immunity. The force of infection and rate of exogenous immune boosting are given by $\lambda(a)$ and $z\lambda(a)$, and the rates of loss of immunity, endogenous immune boosting, and reactivation are given by $\eta(a)$, $\gamma(a)$, and $\rho(a)$. The age-specific incidence of reactivation is given by $\rho(a)W(a)$

The set of ordinary differential equations of this dynamical system is given by

$$\frac{dS(a)}{da} = -\lambda(a)S(a),$$

$$\frac{dV(a)}{da} = \lambda(a)S(a) - \eta(a)V(a) + (z\lambda(a) + \gamma(a) + \rho(a))W(a),$$

$$\frac{dW(a)}{da} = \eta(a)V(a) - (z\lambda(a) + \gamma(a) + \rho(a))W(a) .$$

Following the same reasoning as in Appendix A, we arrive at the following set of discretized equations

$$(V_1)_j = \sum_{k=1}^{j} \frac{\lambda_k}{\lambda_k - \eta_k} \exp\left(-\sum_{i=1}^{k-1} \lambda_i(a_{i+1} - a_i)\right) \exp\left(-\sum_{i=k+1}^{j} \eta_i(a_{i+1} - a_i)\right)$$
$$\times \left[\exp\left(-\eta_k(a_{k+1} - a_k)\right) - \exp\left(-\lambda_k(a_{k+1} - a_k)\right)\right]$$

$$(W_1)_j = \sum_{k=1}^{j} \frac{\eta_k}{z\lambda_k + \gamma_k + \rho_k} (V_1)_k \exp\left(-\sum_{i=k+1}^{j} (z\lambda_i + \gamma_i + \rho_i)(a_{i+1} - a_i)\right)$$
$$\times \left[1 - \exp\left(-(z\lambda_k + \gamma_k + \rho_k)(a_{k+1} - a_k)\right)\right]$$

$$(V_i)_j = \sum_{k=1}^{j} \frac{z\lambda_k + \gamma_k + \rho_k}{\eta_k} (W_{i-1})_k \exp\left(-\sum_{i=k+1}^{j} \eta_i(a_{i+1} - a_i)\right)$$
$$\times \left[1 - \exp\left(-\eta_k(a_{k+1} - a_k)\right)\right]$$

$$(W_i)_j = \sum_{k=1}^{j} \frac{\eta_k}{z\lambda_k + \gamma_k + \rho_k} (V_i)_k \exp\left(-\sum_{i=k+1}^{j} (z\lambda_i + \gamma_i + \rho_i)(a_{i+1} - a_i)\right)$$
$$\times \left[1 - \exp\left(-(z\lambda_k + \gamma_k + \rho_k)(a_{k+1} - a_k)\right)\right].$$

Scenario	Parameter		
	η (year^{-1}) (95%CI)	γ (year^{-1}) (95%CI)	z (95%CI)
6	$0.084\ (0.073 - 0.100)$	$0.10\ (0.093 - 0.13)$	$0.978\ (0.977 - 0.984)$
6*	$0.042\ (0.025 - 0.044)$	-	$0.990\ (0.989 - 0.994)$

Table 6: Parameter estimates of the scenarios with and without endogenous immune boosting and multiple reactions (cf. Table 2). Shown are maximum likelihood estimates with 95% confidence bounds for parameters that are assumed constant.

Scenario	Rates assumed constant	$log(L_{max})$	ΔAIC	ΔBIC	df
1	-	-347.7	16.3	60.6	28
2	γ	-353.1	11.1	34.5	20
3	ρ	-352.7	10.4	33.8	20
4	η	-351.2	7.4	30.8	20
5	γ, ρ	-360.6	10.1	12.7	12
6	γ, η	-354.8	-1.4	1.2	12
6*	$\gamma \equiv 0, \eta$	-356.6	0.2	0.2	11
7	ρ, η	-383.0	54.9	57.5	12
8	all rates	-1176.0	1626.0	1607.7	4

Table 7: Overview of fits of model scenarios with multiple reactivations (cf. Table 1). A suite of simplifications of the baseline model (Scenario 1) are considered in which one or more of the parameters are constant instead of being modeled with splines. Shown are the maximized log-likelihood ($log(L_{max})$), the associated AIC and BIC differences with the best-fitting model Scenario 6* in the main text, and the degrees of freedom (df).
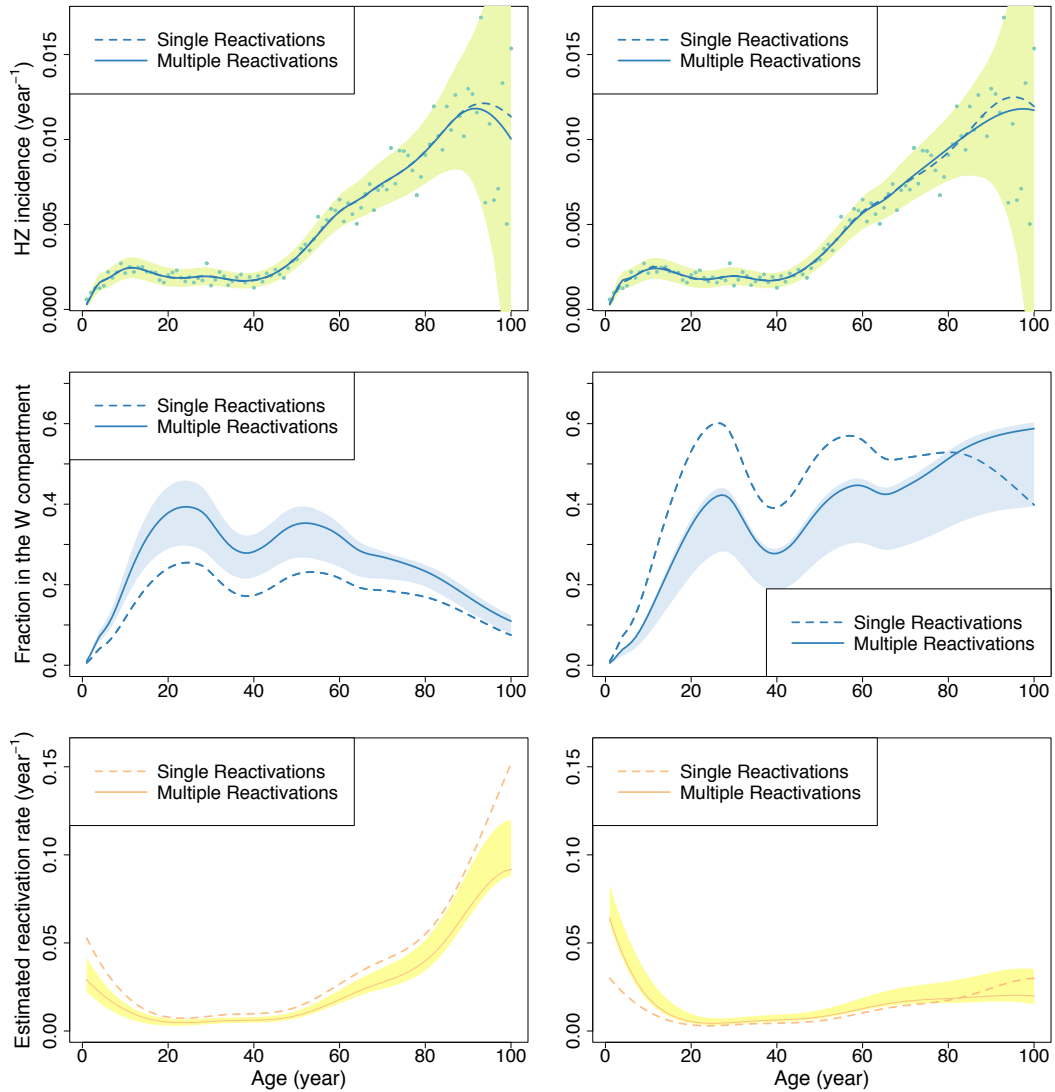
Figure 7: HZ incidence data and overview of the best-fitting models (left-hand panels: with endogenous immune boosting, right-hand panels: without endogenous immune boosting). Shown are the observed and predicted incidence of HZ in both scenarios (top panels; dots: data; lines: prediction), the fraction of the population with waned immunity ($W(a)$) (middle panels), and the estimated reactivation rates ($\rho(a)$)(bottom panels). Lines correspond to maximum likelihood estimates and shaded areas to bootstrapped 95% confidence ranges. Also shown are the corresponding estimates in the model without multiple reactivation (Figure 3).