

An ODE-based mixed modelling approach for B- and T-cell dynamics induced by Varicella-Zoster Virus vaccines in adults shows higher T-cell proliferation with Shingrix than with Varilrix

Peer-reviewed author version

Keersmaekers, Nina; OGUNJIMI, Benson; Van Damme, Pierre; Beutels, Philippe & HENS, Niel (2019) An ODE-based mixed modelling approach for B- and T-cell dynamics induced by Varicella-Zoster Virus vaccines in adults shows higher T-cell proliferation with Shingrix than with Varilrix. In: VACCINE, 37(19), p. 2537-2553.

DOI: 10.1016/j.vaccine.2019.03.075

Handle: <http://hdl.handle.net/1942/30031>

An ODE-based mixed modelling approach for B- and T-cell dynamics induced by Varicella-Zoster Virus vaccines in adults shows higher T-cell proliferation with Shingrix than with Varilrix

Nina Keersmaekers^{1,2,*}, Benson Ogunjimi^{1,2,3,4}, Pierre Van Damme^{2,5},
Philippe Beutels^{1,2}, Niel Hens^{1,2,6}

Abstract

Clinical trials covering the immunogenicity of a vaccine aim to study the longitudinal dynamics of certain immune cells after vaccination. The corresponding immunogenicity datasets are mainly analyzed by the use of statistical (mixed effects) models. This paper proposes the use of mathematical ordinary differential equation (ODE) models, combined with a mixed effects approach. ODE models are capable of translating underlying immunological post vaccination processes into mathematical formulas thereby enabling a

*Corresponding author

Email address: nina.keersmaekers@uantwerp.be (Niel Hens)

¹Centre for Health Economics Research & Modeling Infectious Diseases (CHERMID), Vaccine & Infectious Disease Institute (VAXINFECTIO), University of Antwerp, Antwerp, Belgium

²Antwerp Unit for Data Analysis and Computation in Immunology and Sequencing (AUDACIS), University of Antwerp, Antwerp, Belgium

³Antwerp Center for Translational Immunology and Virology (ACTIV), Vaccine & Infectious Disease Institute (VAXINFECTIO), University of Antwerp, Antwerp, Belgium

⁴Department of Paediatrics, Antwerp University Hospital, Edegem, Belgium

⁵Centre for the Evaluation of Vaccination (CEV), Vaccine & Infectious Disease Institute (VAXINFECTIO), University of Antwerp, Antwerp, Belgium

⁶Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, Diepenbeek, Belgium

January 23, 2019

testable data analysis. Mixed models include both population-averaged parameters (fixed effects) and individual-specific parameters (random effects) for dealing with inter- and intra-individual variability, respectively.

This paper models B-cell and T-cell datasets of a phase I/II, open-label, randomized, parallel-group study (NCT00492648) in which the immunogenicity of a new Herpes Zoster vaccine (Shingrix) is compared with the original Varicella Zoster Virus vaccine (Varilrix).

Since few significant correlations were found between the B-cell and T-cell datasets, each dataset was modeled separately. By following a general approach to both the formulation of several different models and the procedure of selecting the most suitable model, we were able to propose a mathematical ODE mixed-effects model for each dataset. As such, the use of ODE-based mixed effects models offers a suitable framework for handling longitudinal vaccine immunogenicity data. Moreover, this approach allows testing for differences in immunological processes between vaccines or schedules. We found that the Shingrix vaccination schedule led to a more pronounced proliferation of T-cells, without a difference in T-cell decay rate compared to the Varilrix vaccination schedule.

Keywords: Mathematical, models, ordinary differential equations, ODE, mixed effects, Varicella Zoster Virus, VZV, Herpes Zoster, vaccines, B-cells, T-cells, dynamics

1. Introduction

Vaccines are developed in order to activate (and subsequently cause proliferation) of B-cells and T-cells that are specifically directed against the

4 vaccine antigens. B-cells will (1) produce antigen-specific antibodies and (2)
5 differentiate into long-living plasma cells. Antibodies are the primary ef-
6 fectors of the so-called humoral immune response in combating circulating
7 pathogens. T-cells represent the cellular immune response and consist of
8 CD4+ T-cells and CD8+ T-cells (and some other classes not discussed in
9 this paper). CD4+ T-cells have an important role in helping other cell types
10 (such as B-cells and macrophages) combating pathogens. CD8+ T-cells have
11 a direct cytotoxic function and can target host cells that are infected by a
12 pathogen.

13 The vaccine-induced B-cells and T-cells are hypothesized to be capable of pre-
14 venting or minimizing the morbidity related to the infectious disease against
15 which the vaccine is targeted. Vaccine immunogenicity trials aim to study
16 the longitudinal dynamics of the specific immune response following vacci-
17 nation. These trials can range from several months to several decades. The
18 quantitative analysis of longitudinal immune response data has evolved from
19 between-group and time point comparisons to statistical regression analy-
20 ses [1, 2, 3]. Current state-of-the art statistical analyses of longitudinal
21 data consist of a mixed effects model approach in which a separation is
22 made between population-averaged parameters (so called fixed effects) and
23 individual-specific parameters (so called random effects). More recently, An-
24 draud et al. [4] and Le et al. [5] published the first papers in which the
25 mixed effects modeling approach was combined with the use of ordinary dif-
26 ferential equations (ODE), thereby more closely resembling immune response
27 dynamics post vaccination. Whereas [4] focused on the long term dynam-
28 ics following vaccination, [5] focused on the short term dynamics following

29 vaccination.

30 ODE-based mathematical models are capable of translating the underlying
31 immunological/biological theory into a testable data analysis. Moreover,
32 the combination with mixed effects modeling offers a methodology capable
33 of dealing with inter- and intra-individual variability. As such, the use of
34 ODE-based mixed effects models offers a suitable framework for handling
35 longitudinal vaccine immunogenicity data.

36 In this paper, we set out to use ODE-based mixed effects models to study
37 B-cell and T-cell dynamics following varicella-zoster virus (VZV) vaccina-
38 tions in VZV-immune adults. In particular, this framework will allow us to
39 disentangle the immunogenic differences between two different VZV-specific
40 vaccines.

41 We start by showing the immunogenicity data from two VZV vaccine
42 studies consisting of B-cells and CD4+ T-cells of participants at different time
43 points. We then present the differential equations, the ODE and the ODE-
44 based mixed effects models used to describe the immune response dynamics
45 within each individual as well as the associated model selection procedures.
46 By applying the above methods, we consequently select a suitable model for
47 each dataset. Next, we compare the results of the two VZV vaccines, using
48 a group-related effect on a chosen parameter. Correlations in and between
49 the datasets are also explored. We end by reviewing our findings, discussing
50 shortcomings and adding points for future research.

51 **2. Materials and methods**

52 *2.1. Data*

53 The phase I/II, open-label, randomized, parallel-group study EXPLOR-
54 CR004 (101501) investigated the safety and immunogenicity of an adjuvanted
55 recombinant glycoprotein E vaccine ("HZ/su", GSK) for VZV, by compar-
56 ing it with a live attenuated Oka strain VZV vaccine ("OKA", Varilrix[®],
57 GSK). To evaluate safety prior to administration in older adults, two groups
58 of young adults (18-30 years) were vaccinated with two vaccine doses two
59 months apart. The first group (GROUP 1; sample size: $n_1 = 10$) received
60 one dose of HZ/su and one dose of OKA concomitantly at month 0 and
61 month 2 (i.e. four doses in total), whereas the second group (GROUP 2;
62 $n_2 = 10$) received a dose of HZ/su both times (i.e. two doses in total).

63 After vaccine safety was confirmed, three groups of older adults (50-70 years)
64 were vaccinated two months apart, one group (GROUP 3; $n_3 = 45$) received
65 twice a single dose of HZ/su, the second (GROUP 4; $n_4 = 45$) twice a single
66 dose of OKA and the last (GROUP 5; $n_5 = 45$) twice two concomitant doses
67 of HZ/su and OKA. So, all in all, 155 participants were divided over these 5
68 groups. The properties of each group are summarized in Table 1.

69 Safety and immunogenicity were assessed for all groups up to 12 months
70 post-vaccination in the original study. In order to obtain long-term immuno-
71 genicity data on the newly proposed HZ/su vaccine, 23 individuals from the
72 groups solely receiving HZ/su (i.e. GROUPS 2 and 3) were assessed up to 42
73 months post-vaccination in the extension studies NCT00492648. Every indi-
74 vidual was considered a responder. Descriptive statistical tests were used to
75 assess the difference between the two vaccines. A significant higher immune

76 response was found in the groups receiving the HZ/su vaccine compared to
 77 the group solely receiving the OKA vaccine. We refer to [6] for a more in
 78 depth description of the design and results of these studies

GROUP	sample size	age	vaccine	schedule
1	10	18-30	HZ/su + OKA	1+1 dose mo. 0, 1+1 dose mo. 2
2	10	18-30	HZ/su	1 dose mo. 0, 1 dose mo. 2
3	45	50-70	HZ/su	1 dose mo. 0, 1 dose mo. 2
4	45	50-70	OKA	1 dose mo. 0, 1 dose mo. 2
5	45	50-70	HZ/su + OKA	1+1 dose mo. 0, 1+1 dose mo. 2

Table 1: **Properties of the different groups in the VZV vaccine trial.** Shown are sample size, age, vaccine and vaccination schedule (mo.=month). Group 4 is defined as reference group.

79 *2.1.1. B-cell data*

80 First, we used data on the number of antigen-specific memory B-cells,
 81 provided by a B-cell ELISPOT assay, at baseline, and at 1 month and 12
 82 months after receiving the first vaccine dose. Two tests were performed:
 83 the first used Varilrix[©] (1/20x) as stimulus in the B-cell ELISPOT assay,
 84 the second used 100 μ l of gE (10 μ g/ml) as stimulus. This resulted in two
 85 datasets comprising the frequencies of either "total" VZV-specific memory B-
 86 cells or gE-specific memory B-cells per million of total memory B-cells. The
 87 participants were split into 5 different groups, based on age and vaccine type
 88 (see Section 2.1), and their profiles relative to the total number of memory
 89 B-cells are plotted in Fig. 1 for the Varilrix-specific B-cells.

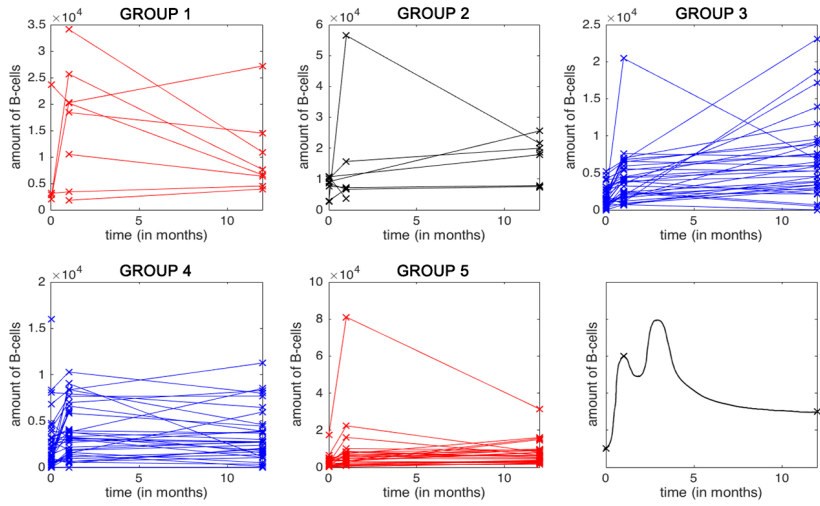


Figure 1: **Amount of VZV-specific memory B-cells. Measured by B-cell ELISPOT (Varilrix stimulus) per 10^6 total memory cells, up to 12 months.** Data are shown per study group. The last panel acts as an illustration of the vaccination dynamics and shows a hypothetical, smooth function of the expected change in number of memory B-cells over time (in months), based on the observed data points per individual and considering the second vaccination at month 2.

90 We observe an increase in memory B-cells (further denoted as B-cells),
91 after vaccination at time $t = 0$ months. At time $t = 2$ months, the subjects
92 were re-vaccinated, but no data were collected at that time point. Fig 1
93 shows only the time points for which data were available ($t = 0, 1$ and 12
94 months). Since it is reasonable to expect a (higher) peak in the data after the
95 second dose at $t = 2$ months, we will assume a time period $[0, h]$, $h > 0$ during
96 which the level of B-cells increases up to a point, h , after which it decreases.
97 The data plots of gE-specific memory B-cells show a similar pattern (see
98 Supplementary material, Fig S1).

99 2.1.2. *T-cell data*

100 Intracellular cytokine staining (ICS) in combination with a flow cyto-
101 metric readout was performed to measure the amount of CD4+ T-cells that
102 produced at least 2 cytokines (interferon-gamma, interleukin 2, CD40 Lig-
103 and, tumor necrosis factor alpha) using both Varilrix and gE as stimuli (in
104 separate experiments). The subsequent two datasets comprise the same 155
105 participants, but now with time points at baseline, and at 1, 2, 3 and 12
106 months after receiving the first vaccine dose. The total VZV-specific T-cell
107 profiles of the participants by study group are shown in Fig 2.

108 Given that T-cell data were collected at more time points than B-cell
109 data, we now observe a second peak in the group-specific data plots, as is
110 expected given the vaccine administrations at month 2. Therefore we will
111 use two time periods $[0, h_1]$ and $[2, h_2]$ (with $0 < h_1 < 2 < h_2$) during which
112 the level of T-cells first increases and then decreases. In case only one peak
113 is observed, we will assume $h_1 = 2 < h_2$.

114 As with the B-cell profiles, the gE specific T-cell profiles are shown in Fig

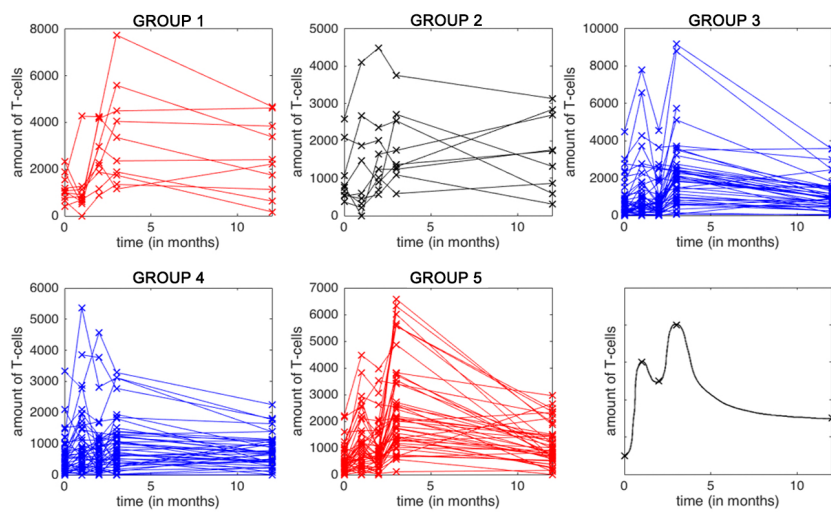


Figure 2: **Amount of VZV-specific CD4+ T-cells, producing at least 2 immune markers.** Measured by ICS per 10^6 CD4+ T-cells, shown per group and up to 12 months. The last panel acts as an illustration of the vaccination dynamics and shows a hypothetical, smooth function of the expected change in number of CD4+ T-cells over time (in months), based on the observed data points per individual.

115 S2.

116 2.2. Mathematical methods

117 We used systems of (nonlinear) ODEs to model the B-cell and T-cell
118 dynamics. We applied a systematic approach to fit and compare several
119 models in order to obtain the models that best describe the available data,
120 while providing sufficient biological interpretation. The detailed version of
121 all ODEs, along with their solution, can be found in Appendix A. In the
122 following subsections we provide the basic rationale of these ODEs for both
123 B-cell and T-cell dynamics, respectively.

124 2.2.1. B-cell dynamics models

125 We describe the dynamics of the memory B-cells using the following ODE:

126

$$\frac{dB}{dt} = f_1(B)I_{t \leq h} - f_2(B), \quad (1)$$

127 where $B_0 = B(0)$ denotes the initial number of memory B-cells at time=0
128 (months) and f_1 and f_2 are smooth functions of the number of memory B-
129 cells at time t (months), describing the change in the number of B-cells due
130 to activation and decay of B-cells. We assume that the activation of B-cells
131 happens during a certain time period $[0, h]$ and that after this time period,
132 no new B-cells are activated. The process of activation of B-cells is described
133 by the function f_1 . The decay in the number of B-cells occurs at all times
134 and is described by f_2 . In all models, the decay rate is assumed to remain
135 constant over time.

136 A first distinction between models can be made through the nature of the
137 rate at which B-cells will be activated. A schematic overview of the different

138 choices in activation/proliferation functions is given in Fig 3. In model B1,
 139 the proliferation rate is assumed to be constant over the time period $[0, h]$.
 140 Another option is to assume a rate which is proportional to the number of
 141 B-cells at time t , as in model B2.

142 A distinction can be made between memory B-cells with a short lifespan
 143 and B-cells with a longer lifespan. A similar line of reasoning is followed in
 144 [7], in which a distinction is made between memory B-cells and long living
 145 plasma cells. Models B3 and B4 incorporate this distinction by including
 146 different equations in the ODE system for short living B-cells (SB) and long
 147 living B-cells (LB). The dynamics of the SB in models B3 and B4 are similar
 148 to those in model B1, but at time 0 months, no SB are present in models B3
 149 and B4. LB however, are present in models B3 and B4 at time 0 months.

150 To distinguish between models B3 and B4, the dynamics of LB are consid-
 151 ered. First, in view of their long lifespan, no decay of LB is assumed and
 152 model B3 expresses no proliferation rate either which means that the total
 153 number of LB remains constant over time. Second, in model B4, a constant
 154 proliferation rate of LB is introduced during time period $[0, h]$ after which
 155 their number will remain constant over time. We refer to Appendix B for
 156 an overview of the parameters used in the dynamic B-cell models.

157 2.2.2. T-cell dynamic models

158 The design of the T-cell models follows a similar procedure as that of
 159 the B-cell models. The following ODE describes the basic dynamics of the
 160 stimulus-specific T-cell population:

$$\frac{dT}{dt} = f_1(T)I_{0 \leq t \leq h_1} + f_2(T)I_{2 \leq t \leq h_2} - f_3(T), \quad (2)$$

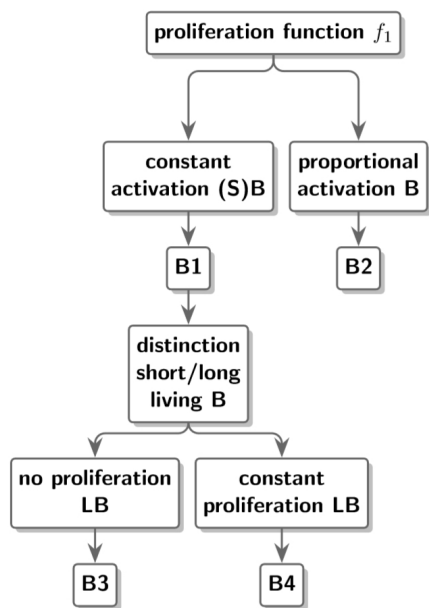


Figure 3: Schematic representation of the different proliferation functions used to describe the B-cells population.

161 with $T_0 = T(0)$ the number of T-cells at time 0 (months). In this equation,
 162 $f_1(T)$ describes the proliferation of T-cells after the first vaccination event
 163 at time 0, which will occur until a certain time point h_1 (with $0 < h_1 \leq 2$).
 164 Afterwards, no T-cells will be activated until the second vaccination event 2
 165 months after the first, which $f_2(T)$ describes as the proliferation of T-cells
 166 during the time period, $[2, h_2]$, with h_2 the time point at which the second
 167 peak in T-cells is reached. The decay of T-cells will happen during the whole
 168 time period, and is represented by the function $f_3(T)$. In all models, it will
 169 be assumed that the decay rate of T-cells remains constant over time (cf.
 170 B-cell models). Moreover, we assume that the activation of T-cells after
 171 each vaccination event happens according to a constant proliferation rate.
 172 It is noteworthy that a non-constant proliferation rate, proportional to the
 173 number of T-cells, was part of exploratory analyses, but these explorations
 174 did not result in a convergent model (see the Inference and model selection
 175 section).
 176 The rates after the first and second vaccination event are not necessarily
 177 equal and neither are the ranges of the time periods $[0, h_1]$ and $[2, h_2]$. Since
 178 T-cells are still present in the blood at the time of the second vaccination,
 179 we assume a different number of new T-cells will be activated.
 180 Fig 4 summarizes the difference between all T-cell models we will consider.
 181 We start by assuming that all T-cells can be regarded as one population.
 182 With the additional assumptions that f_1 and f_2 are identical functions (and
 183 that the proliferation rates of T-cells are equal after each vaccination), we
 184 arrive at model T1. Model T2 does not presume both proliferation rates are
 185 equal.

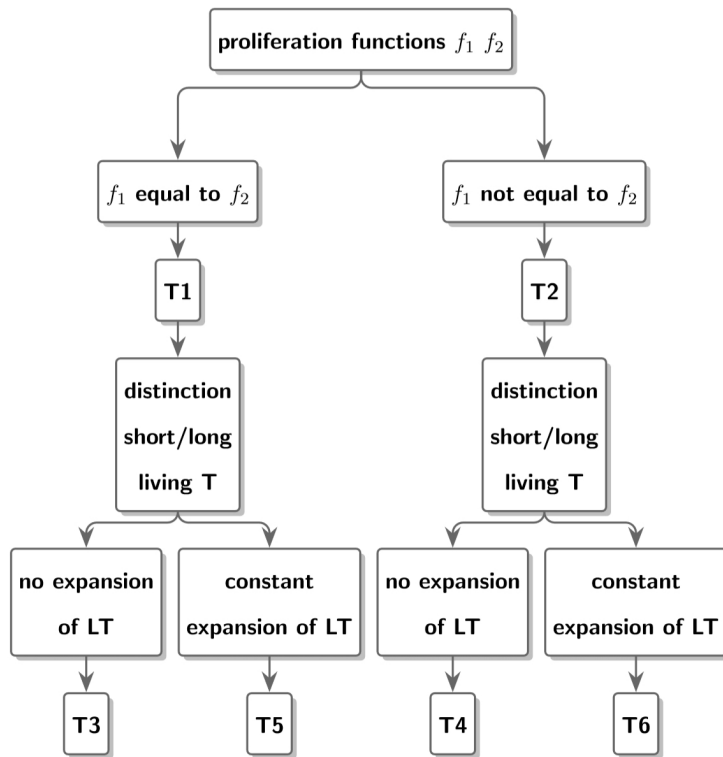


Figure 4: Schematic representation of the difference in proliferation functions used to describe the T-cell dynamics.

186 Next, models are considered in which the T-cell population is divided into
187 short living and long living T-cell populations, represented by ST and LT,
188 respectively. This can be considered as a distinction between effector T-cells
189 (short living) and memory T-cells (long living).
190 The dynamics of the short living T-cells are similarly described as in models
191 T1 and T2: a constant number of short living T-cells will be activated after
192 each vaccination (not necessarily the same number), while the decay of short
193 living T-cells occurs at all times at a constant decay rate. For models T3
194 and T4, the assumption is made that the total number of long living T-
195 cells remains constant over time. If we add the distinction between models
196 with equal and different functions f_1 and f_2 , we arrive at models T3 and
197 T4, respectively. Adding proliferation rates of long living T-cells after each
198 vaccination, models T3 and T4 are extended to yield models T5 and T6.
199 In order to restrict the total number of parameters, the long living T cell
200 proliferation rates are assumed to be equal. We refer to Appendix B for an
201 overview of the parameters used in the dynamic T-cell models.

202 *2.3. Nonlinear mixed models*

203 The dynamic models described in previous subsections can be formulated
204 as nonlinear mixed models in which the parameters are assigned distributions
205 through the specification of fixed and random effects. The fixed component
206 can be interpreted as a population parameter, i.e. an average for all individ-
207 uals, while the random component accounts for individual differences. More
208 specifically, each individual parameter P_i can thus be written as $P_i = u_i \times P_{pop}$
209 where P_{pop} is a population parameter and u_i is log-normally distributed with
210 $E(u_i) = 1$.

211 In case of the presence of a categorical variable (e.g. different group in vac-
212 cine trial study), the different groups can be compared against each other
213 by adding a component β_j to the distribution of a certain parameter. β_j de-
214 scribes how for group j , this parameter deviates from the (chosen) reference
215 group. This makes it possible to test whether one group has a significant
216 higher variable (e.g. rate of cell activation) compared to the reference group.
217 The Monolix software ©Lixoft was used for the estimation of the parameters.
218 A built in stochastic approximation of the standard expectation maximiza-
219 tion algorithm (SAEM) with simulated annealing, combined with a Markov
220 Chain Monte Carlo (MCMC) procedure which replaces the simulation step
221 of the SAEM algorithm, is used to obtain population parameters estimates.
222 Loglikelihood calculation was done by importance sampling, in which a fixed
223 t -distribution is assumed with 5 degrees of freedom. For more details on the
224 algorithms used we refer to [8]. Mostly, Monolix default parameter values
225 were used in the algorithms (see Appendix C) The two-step SAEM-MCMC
226 algorithm uses $10^6 + 10^5$ iterations in order to assess convergence for estimat-
227 ing the population parameters.

228 *2.4. Inference and model selection*

229 Although mathematical identifiability is guaranteed for the models pre-
230 sented in the Mathematical methods section, the complexity of these models
231 when combining them with many random effects in view of the data limita-
232 tions in terms of sampling times and sample sizes resulted in non-convergence.
233 Therefore, simplifying assumptions needed to be made. One such simplify-
234 ing assumption is presuming that the decay of B or T-cells is identical for
235 all individuals, implying that the random effect for that decay parameter is

236 omitted from the model.

237 For both the B-cell and T-cell data sets, the following procedure was used
238 for comparing and selecting the most suitable biologically plausible model
239 to describe the data. In a first step, a list of models was composed, con-
240 sisting of models B1 to B4 for the B-cell data and of models T1 to T6 for
241 T-cell data, together with assumptions on the parameters reflecting whether
242 or not individual variation on these parameters is present, i.e. whether or
243 not random effects were included for the different parameters. The model
244 parameters were then estimated with the Monolix software.

245 Models with poor SAEM convergence, likely because of abundant model
246 complexity, were discarded. Next, the candidate models were compared us-
247 ing Akaike's Information Criterion (AIC) and the model with lowest AIC
248 value was selected as first candidate model. Subsequently, a non-parametric
249 bootstrap, using 1000 bootstrap re-samples, was performed on the candidate
250 model. Since a sequential approach based on the candidate models with the
251 lowest AIC values was used, the need to perform bootstraps for all candidate
252 models was avoided, in order to decrease the number of computations. It was
253 found that for a bootstrap, either 65%-77% of the samples had proper SAEM
254 convergence, or the proportion of bootstrap samples with proper convergence
255 was less than 15%. For this reason the criterion for good bootstrap conver-
256 gence was defined as having at least 65% of bootstrap samples with proper
257 SAEM convergence. In case of poor bootstrap convergence, the candidate
258 model was rejected from the list of candidate models.

259 Next, a sensitivity analysis on the bootstrap results of the (converging) can-
260 didate model was performed by investigating whether the presence or absence

261 of certain profiles of individuals in the bootstrap samples, had influence on
262 its model convergence. For instance, if a single participant’s profile was more
263 frequently part of non-converging datasets, a new bootstrap was performed,
264 excluding the specific participant’s profile. Again, in case of poor bootstrap
265 convergence the candidate model was rejected. If convergence remained suf-
266 ficiently robust, it was investigated whether it was possible to improve the
267 biological plausibility of the candidate model. This was done by examining
268 the assumptions made on the parameters and adjusting those. An example
269 of this is shown in the Model selection of T-cell datasets results section. The
270 new model then became a candidate model.

271 *2.5. B-cell and T-cell dynamics associations*

272 As CD4+ T-cells may directly influence the activation of B-cells, we in-
273 vestigated the existence of associations between B-cell and T-cell dynamics.
274 We used a raw data complete cases analysis given that constructing a joint
275 model based on the same concepts was not successful, likely due to data lim-
276 itations (Inference and model selection section).
277 More specifics on this analysis, and its results, can be found in Appendix D.

278 **3. Results**

279 *3.1. Model selection of B-cell datasets*

280 We started by modeling the Varilrix-specific B-cell dataset, for which
281 the model selection procedure outlined in the Inference and model selection
282 section was followed. All details on the considered models and the following
283 of the procedure can be found in Appendix E. The resulting Varilrix-specific

284 B-cell model differentiates between SB and LB, LB are assumed to remain
 285 constant through time (model B3, see B-cell dynamics models section. In
 286 time period $[0, h]$, a constant number of SB is activated and this proliferation
 287 rate is considered a group-specific parameter. The decay rate is assumed to
 288 be equal for each individual.

289 The estimated population parameters of the selected B-cell model are shown
 290 in Table 2.

parameter	estimate	95% CI
LB_0	1852.84	(1425.79, 2407.80)
aSB	774.17	(367.67, 1630.12)
β_1	2.14	(0.79, 3.49)
β_2	1.44	(-0.02, 2.91)
β_3	0.67	(-0.12, 1.46)
β_5	0.90	(0.02, 1.79)
uSB	0.13	(0.11, 0.16)
h	3.30	(2.05, 5.30)

Table 2: **Varilrix-specific B-cell results.** Parameter estimates and corresponding 95% confidence intervals (CI) of final model B3b. $SB(0)$ is assumed to be zero, $LB_0 = LB(0)$ denotes the initial number of LB. The proliferation of SB is constant in time period $[0, h]$, at rate aSB and assumed to be group-specific. β_i ($i = 1, 2, 3, 5$) represents the increase of aSB , compared to the reference group 4 ($\beta_4 = 0$). Decay of SB happens at rate uSB . The number of LB remains constant through time.

291 The same models as with the Varilrix-specific B-cell data were used for the
 292 gE-specific B-cell data and a similar model selection procedure was followed.
 293 The outcome of the model selection was a model which does not differentiate

294 between SB and LB. In time period $[0, h]$, a constant number of B-cells are
 295 activated (model B1). All parameters are assumed to have random effects
 296 and the activation rate of B-cells is chosen as group-specific parameter. The
 297 parameter estimations, along with confidence interval, are shown in Table 3.

parameter	estimate	95% CI
B_0	564.39	(429.80, 741.14)
aB	566.08	(337.69, 948.95)
β_1	2.59	(1.52, 3.66)
β_2	2.13	(1.08, 3.17)
β_3	1.27	(0.69, 1.84)
β_5	1.25	(0.59, 1.91)
uB	0.035	(0.0060, 0.21)
h	2.51	(1.78, 3.52)

Table 3: **gE-specific B-cell results.** Parameter estimates and corresponding 95% confidence intervals (CI) of final model B1a. No distinction between SB and LB is presumed. $B_0 = B(0)$ denotes the initial number of B-cells. The proliferation of B-cells is constant in time period $[0, h]$, at rate aB and assumed to be group-specific. β_i ($i = 1, 2, 3, 5$) represents the increase of aB , compared to the reference group 4 ($\beta_4 = 0$). Decay of B-cells happens at rate uB .

298 3.2. Model selection of T-cell datasets

299 By following the T-cell model selection procedure (details in Appendix
 300 E), a model which differentiates between ST and LT, was selected as final
 301 Varilrix-specific T-cell model. This model furthermore assumes the number
 302 of LT remains constant through time and assumes activation of ST is constant
 303 in time periods $[0, h_1]$ and $[2, h_2]$ with $a_2ST = 0.15 \times a_1ST$ (model T3, see

304 T-cell dynamics models section). Moreover, all parameters are assumed to
 305 have random effects with the activation rate a group-specific parameter. The
 306 model parameter estimates are shown in Table 4.

parameter	estimate	95% CI
LT_0	517.53	(422.96, 633.25)
aST	1825.20	(893.64, 3727.84)
β_1	1.50	(0.29, 2.72)
β_2	0.76	(-0.81, 2.32)
β_3	1.57	(0.90, 2.25)
β_5	1.73	(1.07, 2.39)
uST	0.40	(0.18, 0.86)
h_1	0.026	(0.0088, 0.075)
h_2	3.85	(2.57, 5.78)

Table 4: **Varilrix-specific T-cell results.** Parameter estimates and corresponding 95% confidence intervals (CI) of final model T3a'. $ST(0)$ is assumed to be zero, $LT_0 = LT(0)$ denotes the initial number of LT. The proliferation of ST is constant in time period $[0, h_1]$ at rate aST and in time period $[2, h_2]$ at rate $0.15.aST$. aST is assumed to be group-specific with effects β_i ($i = 1, 2, 3, 5$), representing the increase of aST , compared to the reference group 4 ($\beta_4 = 0$). Decay of ST happens at rate uST . The number of LT remains constant through time.

307 The same T-cell models were used in the model selection procedure of the
 308 gE-specific T-cell data. The parameter estimations of the final gE-specific
 309 T-cell model are shown in Table 5. This model does not differentiate between
 310 ST and LT and assumes a constant activation of T-cells in time periods $[0, h_1]$
 311 and $[2, h_2]$ with $a_2ST = 0.66 \times a_1ST$ (model T1). All parameters were as-
 312 sumed to have random effects with the activation rate being a group-specific

313 parameter. Moreover, individuals 89 and 149 were viewed as statistical out-
 314 liers (with datapoints distant from other individuals) and their profiles were
 315 left out of this dataset.

parameter	estimate	95% CI
T_0	90.07	(45.31, 179.07)
aT	329.14	(162.70, 665.81)
β_1	2.41	(1.75, 3.32)
β_2	2.30	(1.65, 3.21)
β_3	2.53	(1.93, 3.31)
β_5	2.43	(1.83, 3.22)
uT	0.54	(0.26, 1.11)
h_1	0.21	(0.15, 0.31)
h_2	9.03	(5.61, 14.53)

Table 5: **gE-specific T-cell results.** Parameter estimates and corresponding 95% confidence intervals (CI) of final model T1a'. No distinction of ST and LT is presumed. $T_0 = T(0)$ denotes the initial number of T-cells. The proliferation of T-cells is constant in time period $[0, h_1]$ at rate aT and in time period $[2, h_2]$ at rate $0.66 \times aT$. aT is assumed to be group-specific with effects β_i ($i = 1, 2, 3, 5$), representing the increase of aT , compared to the reference group 4 ($\beta_4 = 0$). Decay of T-cells happens at rate uT .

316 3.3. Vaccine differences

317 Group-specific effects on chosen parameters make it possible to compare
 318 each group by examining the differences in these effects. This comparison
 319 focuses on different proliferations of B- and T-cells. Group-specific effects
 320 were also added on other parameters, more specifically the decay rate and
 321 time point h , which marks the end of the proliferation period after vaccina-

322 tion. However, models with these group-specific parameters did either not
323 have an increased AIC compared to a model without this effect, or did not
324 show SAEM convergence.

325 Since a group-specific component was added to the activation of B-/T-cells
326 for each final model, it was subsequently possible to examine whether the
327 HZ/su vaccine caused a higher increase in B- and or T-cells after vaccina-
328 tion, compared to the original OKA vaccine.

329 As a reminder, Table 1 summarizes the characteristics of the 5 different
330 groups in the vaccine trial. Group 4 received the original OKA vaccine and
331 was thereby defined as reference group.

332 Next, we calculated corresponding p-values of the group-specific parame-
333 ters that were estimated in the Model selection of B-cell datasets and T-cell
334 datasets sections, shown in Table 6. In view of the sample size of groups 1
335 and 2 (and age), we were mainly interested in β_3 and β_5 . A β_i higher than
336 zero indicates a higher activation rate of cells in the groups receiving the
337 HZ/su vaccine, compared to the activation rate in the reference group which
338 received the OKA vaccine. In the case of Varilrix- and gE-specific T-cells,
339 both groups 3 and 5 showed a significant higher activation rate ($p < 0.05$).
340 The activation rate of gE-specific B-cells also was significantly higher com-
341 pared to the reference group. Varilrix-specific B-cells also seemed to have
342 a higher proliferation rate, though in the case of groups receiving solely the
343 HZ/su vaccine, not significantly so ($p > 0.05$).

344 As age has an important influence on vaccine responses, it was considered
345 to compare the young groups (groups 1 and 2) to the older groups with the
346 same vaccination schedule (groups 3 and 5). However, we can observe from

347 Tables 2, 3, 4 and 5 that the confidence intervals of the group-specific param-
 348 eters 1-5 overlap. Therefore, no conclusions can be made in that respect. It
 349 is to be noted that the young cohorts are much smaller than the old cohorts
 350 (n=10 vs. n=45 respectively).

	Varilrix B-cells	gE B-cells	Varilrix T-cells	gE T-cells
β_1 (CI), p-value	2.14 (0.79, 3.49) 2.8×10^{-4}	2.59 (1.52, 3.66) $< 10^{-5}$	1.50 (0.29, 2.72) 3.8×10^{-4}	2.41 (1.75, 3.32) $< 10^{-5}$
β_2 (CI), p-value	1.44 (-0.02, 2.91) 0.013	2.13 (1.08, 3.17) $< 10^{-5}$	0.76 (-0.81, 2.32) 0.12	2.30 (1.65, 3.21) $< 10^{-5}$
β_3 (CI) p-value	0.67 (-0.12, 1.46) 0.10	1.27 (0.69, 1.84) 10^{-5}	1.57 (0.90, 2.25) $< 10^{-5}$	2.53 (1.93, 3.31) $< 10^{-5}$
β_5 (CI) p-value	0.90 (0.02, 1.79) 0.025	1.25 (0.59, 1.91) 3×10^{-5}	1.73 (1.07, 2.39) $< 10^{-5}$	2.43 (1.83, 3.22) $< 10^{-5}$

Table 6: **Differences in proliferation rates per group.** Group-specific parameter estimates, corresponding 95% confidence intervals (CI) and p-values, calculated for Varilrix B-cell, gE B-cell, Varilrix T-cell and gE T-cell data.

351 4. Discussion

352 In this study we used a nonlinear mixed modeling approach using ordi-
 353 nary differential equations (ODE) to describe B-cell and T-cell dynamics in
 354 adults following a 2-dose vaccination against VZV by means of the novel sub-
 355 unit VZV gE vaccine (Shingrix, GSK) and the live-attenuated VZV vaccine
 356 (Varilrix, GSK).

357 Whereas the latter vaccine was not intended to be used in a similar manner
 358 as Shingrix to augment protection against herpes zoster (HZ), we did use

359 the data available from the clinical trial performed by Leroux-Roels et al
360 (Leroux-Roels2012). In this trial, Shingrix and Varilrix were compared in
361 regard to safety and immunogenicity in adults. Using comparative group-
362 wise statistical tests, they found that gE-specific CD4+ T-cell levels were
363 much higher in the groups receiving Shingrix than those receiving Varilrix
364 alone from 3 until 42 months after vaccination. Additionally, they showed
365 that the addition of Varilrix to Shingrix did not significantly increase the
366 immunogenicity.

367 Our study was motivated by the difficulties in attributing differences between
368 vaccines and vaccination schedules to underlying immunological processes
369 when using the "classical" group-wise comparative statistical techniques that
370 do not take into account underlying immunological processes.

371 Recently, [4] and [5] showed that nonlinear ODE mixed modeling was able to
372 produce (plausible) estimates on several biological parameters in the setting
373 of vaccinations.

374 In our study, we assessed Varilrix and VZV gE-specific B-cell and T-cell re-
375 sponses elicited by the 2-dose vaccination schedule for three different sched-
376 ules (Shingrix only, Varilrix only and the combination of Shingrix and Varilrix
377 on both vaccination moments). We developed and used a suitable method-
378 ological framework to obtain for each setting (immune response and vaccina-
379 tion schedule) the most optimal ODE model, informed by immunological the-
380 ory while acknowledging data sparseness by adjusting the inferential method
381 using a nonparametric bootstrap approach. Using this robust approach we
382 were able to conclude that the best models did not have an overly complex
383 structure. Models with constant proliferation rates (after each vaccination,

384 in time periods $[0, h]$ (B-cells) and $[0, h_1]$ and $[2, h_2]$ (T-cells)) had lower AIC
385 compared to models with proportional proliferation rates. Restricting the
386 number of parameters, either by not making a distinction between short and
387 long living B-/T-cells, or by assuming the number of long living B-/T-cells
388 remains constant through time, resulted in models that were preferable com-
389 pared to the other considered model structures. Although some of these
390 other models may have had a more intuitively logical biological interpreta-
391 tion, they often did not have SAEM or bootstrap convergence.
392 Importantly, this way of modeling allowed us to directly compare specific
393 parameters between several vaccination schedules. We found that the Shin-
394 grix vaccination schedules led to a more pronounced proliferation of T-cells,
395 however without a difference in T-cell decay rate between Shingrix and Var-
396 ilrix vaccination schedules. This novel result underscores the benefit of using
397 mathematical mixed models that are based on the underlying immunological
398 processes instead of performing standard group-wise comparisons. Indeed, in
399 the latter case it is possible to prove significant differences between vaccines,
400 however, it is impossible to determine what drives these differences. That is,
401 the standard group-wise comparisons cannot show whether higher response
402 levels are the result of either a higher proliferation of cells, a lower decay
403 (mainly in the case of a restricted number of data points), or a longer time
404 period $[0, h]$ in which cells are activated.
405 We note that the adjuvant used for the Shingrix vaccine has been reported
406 to be a very potent adjuvant [10] and our modeling approach thus confirms
407 the increased proliferation of T-cells for the Shingrix vaccine.
408 We also assessed whether a correlation existed between the B-cell and T-cell

409 counts, but we did not find a significant association between the two immune
410 response types. This confirms previous findings concerning the glycoprotein-
411 E adjuvant, part of the AS01_B Adjuvant System family, in which it has been
412 shown that this family has been reported to show the lowest correlations
413 between B-cells and T-cells of all families [11].

414 During our modeling analyses we encountered several limitations. First, we
415 noted that given the limited sample size only models with moderate complex-
416 ity could be analysed. Second, the sparseness of time points for the B-cell
417 responses posed a significant limitation on the complexity of the B-cell mod-
418 els. Future work should focus on estimating an optimal sampling schedule
419 for subsequent modeling.

420 In this study, we wanted to focus on the advantages and possibilities of ODE
421 modeling combined with a mixed effect approach in the analysis of vaccine
422 trial immunogenicity data, rather than group-wise or time point-wise com-
423 parisons using standard comparative statistics between different vaccination
424 schedules. The techniques underlying this work can now be applied on novel
425 datasets to answer fundamental questions on the understanding of immune
426 responses.

427 We conclude that nonlinear mixed modeling by means of ODE shows that
428 Shingrix vaccination causes a significantly higher proliferation of T-cells com-
429 pared to Varilrix vaccination in VZV-immune adults.

430 **Appendix A. Detailed ODE models and solutions**

431 *Appendix A.1. Antibody secreting cell models*

432 The dynamics of the memory B-cell population is described using the
 433 following ODE:

$$\frac{dB}{dt} = f_1(B)I_{t \leq h} - f_2(B), \quad (\text{A.1})$$

434 where $B_0 = B(0)$ denotes the initial number of memory B-cells at time =
 435 0 (months) and f_1 and f_2 are smooth functions of the number of B-cells at
 436 time t (months) and describe the change in the number of B-cells due to
 437 activation and decay of B-celms. We assume that the activation of B-cells
 438 happens during a certain time period $[0, h]$ and that after this time period,
 439 no new B-cells are activated. The process of activation of B-cells is thereby
 440 described by the function f_1 . The decay of B-cells happens at all times and is
 441 described by f_2 . In all models, the decay rate is assumed to remain constant
 442 over time and the decay of B-cells is thereby proportional to the number of
 443 B-cells.

444 In model B1, the activation of B-cells is assumed to be constant in time, with
 445 aB the rate of activation. This means functions f_1 and f_2 can be expressed
 446 as

$$\begin{cases} f_1(B) = aB \\ f_2(B) = -uB \cdot B \end{cases} \quad (\text{A.2a})$$

447 The solution of this differential equation leads to the following equation

$$\begin{cases} B(t) = \frac{aB}{uB} + e^{-uB \cdot t} \left(B_0 - \frac{aB}{uB} \right) & t \leq h \\ B(t) = \left(\frac{aB}{uB} + e^{-uB \cdot h} \left(B_0 - \frac{aB}{uB} \right) \right) e^{-uB(t-h)} & t > h \end{cases} \quad (\text{A.2b})$$

448 In model B2, the proliferation of B-cells is assumed to be proportional to
 449 the number of B-cells, so the activation is now equal to $pB \times B$ with pB the
 450 activation rate. This yields equations

$$\begin{cases} f_1(B) = pB \cdot B \\ f_2(B) = -uB \cdot B \end{cases} \quad (\text{A.3a})$$

451 with solution

$$\begin{cases} B(t) = B_0 e^{(pB-uB)t} & t \leq h \\ B(t) = B_0 e^{pB \cdot h - uB \cdot t} & t > h \end{cases} \quad (\text{A.3b})$$

452 Previous two models can be combined in the following model, where the
 453 activation has both a constant and a proportional component:

$$\begin{cases} f_1(B) = aB + pB \cdot B \\ f_2(B) = -uB \cdot B \end{cases} \quad (\text{A.4a})$$

454 with solution

$$\begin{cases} B(t) = -\frac{aB}{paSC-uB} + \left(B_0 + \frac{aB}{pB-uB} \right) e^{(pB-uB)t} & t \leq h \\ B(t) = \left(-\frac{aB}{pB-uB} + \left(B_0 + \frac{aB}{pB-uB} \right) e^{(pB-uB)h} \right) \cdot e^{-uB(t-h)} & t > h \end{cases} \quad (\text{A.4b})$$

455 Since this model did not yield converging results, it was thereby omitted from
 456 the paper.

457 In a next step we distinguish the short-living B-cells $SB(t)$ from the B-cells
 458 with long lifespan $LB(t)$. The total number of B-cells is then equal to the
 459 sum of these two. We assume that, at baseline, only long-living and thus no

460 short-living B-cells are present. This model can be described as

$$\begin{cases} \frac{dSB}{dt} = f_1(SB)I_{t \leq h} + f_2(SB) \\ \frac{dLB}{dt} = g_1(LB)I_{t \leq h} + g_2(LB) \\ B(t) = SB(t) + LB(t) \end{cases} \quad (\text{A.5})$$

461 with $LB_0 = LB(0)$ denoting the initial number of LB and $SB(0) = 0$.

462 First, in model B3 we assume a constant proliferation rate, equal to aSB .

463 Furthermore, due to their long lifespan, we assume the decay rate of LB

464 to be equal to zero and that no new long living B-cells are activated. This

465 means the number of LB will remain constant over time. We can summarize

466 this as

$$\begin{cases} f_1(SB) = aSBS \\ f_2(SB) = -uSB \cdot SB \\ g_1(LB) = g_2(LB) = 0 \end{cases} \quad (\text{A.6a})$$

467 The solution of the ODE of model B3 is given by

$$\begin{cases} B(t) = -\frac{aSB}{uSB} (e^{-uSB \cdot t} - 1) + LB_0 & t \leq h \\ B(t) = -\frac{aSB}{uSB} (1 - e^{uSB \cdot h}) e^{-uSB \cdot t} + LB_0 & t > h \end{cases} \quad (\text{A.6b})$$

468 Instead of a constant number of LB , we now assume in model B4 that LB

469 will proliferate as well, according to a constant proliferation rate aLB during

470 time period h . Their decay will still be neglected due to their long lifespan.

$$\begin{cases} f_1(SB) = aSB \\ f_2(SB) = -uSB \cdot SB \\ g_1(LB) = aLB \\ g_2(LB) = 0 \end{cases} \quad (\text{A.7a})$$

471 with solution

$$\begin{cases} B(t) = -\frac{aSB}{uSB} (e^{-uSB \cdot t} - 1) + LB_0 + aLB \cdot t & t \leq h \\ B(t) = -\frac{aSB}{uSB} (1 - e^{uSB \cdot h}) e^{-uSB \cdot t} + \\ LB_0 + aLB \cdot h & t > h \end{cases} \quad (\text{A.7b})$$

472 *Appendix A.2. T-cell models*

473 The design of the T-cell models follows a similar procedure as with the B-
 474 cell models. The following differential equation describes the basic dynamics
 475 of the stimulus-specific T-cell population:

$$\frac{dT}{dt} = f_1(T)I_{0 \leq t \leq h_1} + f_2(T)I_{2 \leq t \leq h_2} + f_3(T) \quad (\text{A.8})$$

476 with $T_0 = T(0)$ the number of T-cells at time 0 (months). In this equation,
 477 $f_1(T)$ describes the proliferation of T-cells after the first vaccination event
 478 at time 0, which will occur until a certain time point h_1 (with $0 < h_1 \leq 2$).
 479 Afterwards, no T-cells will be activated until the second vaccination event 2
 480 months after the first vaccine, which $f_2(T)$ describes as the proliferation of
 481 T-cells during the time period $[2, h_2]$, with h_2 the time point at which the
 482 second peak in T-cells is reached. The decay of T-cells will happen during the
 483 whole time period, and is represented by the function $f_3(T)$. In all models,
 484 it will be assumed that the decay rate of T-cells remains constant over time
 485 (cfr. B-cell models). Moreover, we assume that the activation of T-cells after
 486 each vaccination event happens according to a constant proliferation rate.
 487 Assuming an equal activation rate of T cells after each vaccination leads to
 488 model T1. Functions f_i ($i = 1, 2, 3$) can now be written as

$$\begin{cases} f_1(T) = f_2(T) = aT \\ f_3(T) = -uT \cdot T \end{cases} \quad (\text{A.9a})$$

The solution of the ODE is now given by

$$\left\{ \begin{array}{ll} T(t) = \frac{aT}{uT} + e^{-uT \cdot t} \left(T_0 - \frac{aT}{uT} \right) & t \leq h_1 \\ T(t) = \left(\frac{aT}{uT} + e^{-uT \cdot h_1} \left(T_0 - \frac{aT}{uT} \right) \right) e^{-uT(t-h_1)} & h_1 < t < 2 \\ T(t) = \frac{aT}{uT} + \left(T_0 - \frac{aT}{uT} + \frac{aT}{uT} e^{uT \cdot h_1} - \frac{aT}{uT} e^{uT \cdot 2} \right) e^{-uT \cdot t} & 2 \leq t \leq h_2 \\ T(t) = \left(\frac{aT}{uT} + \left(T_0 - \frac{aT}{uT} + \frac{aT}{uT} e^{uT \cdot h_1} - \frac{aT}{uT} e^{uT \cdot 2} \right) e^{-uT \cdot h_2} \right) \cdot & \\ \quad e^{-uT(t-h_2)} & t > h_2 \end{array} \right. \quad (\text{A.9b})$$

489 Model T2 does not possess the assumption of an equal proliferation rate after
 490 each vaccination. Moreover, it is plausible to assume a changed rate after
 491 the second vaccination due to a memory response. In this case functions f_i
 492 are written as

$$\left\{ \begin{array}{l} f_1(T) = a_1 T \\ f_2(T) = a_2 T \\ f_3(T) = -uT \cdot T \end{array} \right. \quad (\text{A.10a})$$

493 with solution

$$\left\{ \begin{array}{ll} T(t) = \frac{a_1 T}{uT} + \left(T_0 + \frac{a_1 T}{-uT} \right) e^{-uT \cdot t} & t \leq h_1 \\ T(t) = \left(\frac{a_1 T}{uT} + \left(T_0 + \frac{a_1 T}{-uT} \right) e^{-uT \cdot h_1} \right) e^{-uT(t-h_1)} & h_1 < t < 2 \\ T(t) = \frac{a_2 T}{uT} + \left(\left(\frac{a_1 T}{uT} + \left(T_0 - \frac{a_1 T}{uT} \right) e^{-uT \cdot h_1} \right) e^{-uT(2-h_1)} - \frac{a_2 T}{uT} \right) \cdot & \\ \quad e^{-uT(t-2)} & 2 \leq t \leq h_2 \\ T(t) = \left(\frac{a_2 T}{uT} + \left(\left(\frac{a_1 T}{uT} + \left(T_0 - \frac{a_1 T}{uT} \right) e^{-uT \cdot h_1} \right) e^{-uT(2-h_1)} - \frac{a_2 T}{uT} \right) \cdot & \\ \quad e^{-uT(t-2)} \right) e^{-uT(t-h_2)} & t > h_2 \end{array} \right. \quad (\text{A.10b})$$

494 Again, in the next step we differentiate between the short living (ST) and
 495 long living (LT) T-cells. We express this with equations f_i, g_i ($i = 1 : 3$):

$$\begin{cases} \frac{dST}{dt} = f_1(ST)I_{0 < t \leq h_1} + f_2(ST)I_{2 < t \leq h_2} + f_3(ST) \\ \frac{LST}{dt} = g_1(LT)I_{0 < t \leq h_1} + g_2(LT)I_{2 < t \leq h_2} + g_3(LT) \\ T(t) = ST(t) + LT(t) \end{cases} \quad (\text{A.11})$$

496 with $LT_0 = LT(0)$ the initial number of long living T-cells and $ST(0) = 0$.

497 As in previous models, we will assume constant proliferation rates a_1ST
 498 and a_2ST of ST, no decay of LT and at first a constant number of LT. Model
 499 T3 assumes as well that activation rates of ST after both vaccinations are
 500 equal. Functions f_i and g_i are thereby written as

$$\begin{cases} f_1(ST) = f_2(ST) = aST \\ f_3(ST) = -uST \cdot ST \\ g_1(LT) = g_2(LT) = 0 \\ g_3(LT) = 0 \end{cases} \quad (\text{A.12a})$$

501 Its solution is given by

$$\begin{cases} T(t) = LT_0 - \frac{aST}{uST} (e^{-uST \cdot t} - 1) & t \leq h_1 \\ T(t) = LT_0 - \frac{aST}{uST} (e^{-uST \cdot h_1} - 1) e^{-uST(t-h_1)} & h_1 < t < 2 \\ T(t) = LT_0 - \frac{aST}{uST} (1 - e^{uST \cdot h_1}) e^{-uST \cdot t} + \\ \quad \frac{aST}{uST} (1 - e^{-uST(t-2)}) & 2 \leq t \leq h_2 \\ T(t) = LT_0 + \left(-\frac{aST}{uST} (1 - e^{uST \cdot h_1}) e^{-uST \cdot h_2} + \right. \\ \quad \left. \frac{aST}{uST} (1 - e^{-uST(h_2-2)}) \right) e^{-uST(h_2-t)} & t > h_2 \end{cases} \quad (\text{A.12b})$$

502 Model T4 does not make the assumption of equal ST activation rates, which
 503 leads to equations

$$\begin{cases} f_1(ST) = a_1 ST \\ f_2(ST) = a_2 ST \\ f_3(ST) = -uST \cdot ST \\ g_1(LT) = g_2(LT) = 0 \\ g_3(LT) = 0 \end{cases} \quad (\text{A.13a})$$

504 with solution

$$\begin{cases} T(t) = LT_0 - \frac{a_1 ST}{uST} (e^{-uST \cdot t} - 1) & t \leq h_1 \\ T(t) = LT_0 - \frac{a_1 ST}{uST} (e^{-uST \cdot h_1} - 1) e^{-uST(t-h_1)} & h_1 < t < 2 \\ T(t) = LT_0 - \frac{a_1 ST}{uST} (1 - e^{uST \cdot h_1}) e^{-uST \cdot t} + \\ \quad \frac{a_2 ST}{uST} (1 - e^{-uST(t-2)}) & 2 \leq t \leq h_2 \\ T(t) = LT_0 + \left(-\frac{a_1 ST}{uST} (1 - e^{uST \cdot h_1}) e^{-uST \cdot h_2} + \right. \\ \quad \left. \frac{a_2 ST}{uST} (1 - e^{-uST(h_2-2)}) \right) e^{-uST(h_2-t)} & t > h_2 \end{cases} \quad (\text{A.13b})$$

505 Next, we add a constant proliferation rate aLT of long-living T-cells after
 506 each vaccination (A.15a). Model T5 assumes equal activation rates of ST,
 507 with following equations:

$$\begin{cases} f_1(ST) = f_2(ST) = aST \\ f_3(ST) = -uST \cdot ST \\ g_1(LT) = g_2(LT) = aLT \\ g_3(LT) = 0 \end{cases} \quad (\text{A.14a})$$

508 and solution:

$$\left\{ \begin{array}{ll}
 T(t) = -\frac{aST}{uST}(e^{-uST \cdot t} - 1) + LT_0 + aLT \cdot t & t \leq h_1 \\
 T(t) = -\frac{aST}{uST}(e^{-uST \cdot h_1} - 1)e^{-uST(t-h_1)} + \\
 \quad LT_0 + aLT \cdot h_1 & h_1 < t < 2 \\
 T(t) = -\frac{aST}{uST}(1 - e^{uST \cdot h_1})e^{-uST \cdot t} + \\
 \quad \frac{aST}{uST}(1 - e^{-uST(t-2)}) + \\
 \quad LT_0 + aLT \cdot (h_1 + t - 2) & 2 \leq t \leq h_2 \\
 T(t) = \left(-\frac{aST}{uST}(1 - e^{uST \cdot h_1})e^{-uST \cdot h_2} + \right. \\
 \quad \left. \frac{aST}{uST}(1 - e^{-uST(h_2-2)}) \right) e^{-uST(h_2-t)} + \\
 \quad LT_0 + aLT(h_1 + h_2 - 2) & t > h_2
 \end{array} \right. \quad (\text{A.14b})$$

509 We conclude the T-cell models with model T6, in which different activation
 510 rates of ST after each vaccination are assumed. Equations f_i and g_i are
 511 written as

$$\left\{ \begin{array}{l}
 f_1(ST) = a_1 ST \\
 f_2(ST) = a_2 ST \\
 f_3(ST) = -uST \cdot ST \\
 g_1(LT) = g_2(LT) = aLT \\
 g_3(LT) = 0
 \end{array} \right. \quad (\text{A.15a})$$

512 The solution of the resulting ODE is given by

$$\left\{ \begin{array}{ll}
 T(t) = -\frac{a_1 ST}{uST} (e^{-uST \cdot t} - 1) + LT_0 + aLT \cdot t & t \leq h_1 \\
 T(t) = -\frac{a_1 ST}{uST} (e^{-uST \cdot h_1} - 1) e^{-uST(t-h_1)} + \\
 \quad LT_0 + aLT \cdot h_1 & h_1 < t < 2 \\
 T(t) = -\frac{a_1 ST}{uST} (1 - e^{uST \cdot h_1}) e^{-uST \cdot t} + \\
 \quad \frac{a_2 ST}{uST} (1 - e^{-uST(t-2)}) + \\
 \quad LT_0 + aLT \cdot (h_1 + t - 2) & 2 \leq t \leq h_2 \\
 T(t) = \left(-\frac{a_1 ST}{uST} (1 - e^{uST \cdot h_1}) e^{-uST \cdot h_2} + \right. \\
 \quad \left. \frac{a_2 ST}{uST} (1 - e^{-uST(h_2-2)}) \right) e^{-uST(h_2-t)} + \\
 \quad LT_0 + aLT(h_1 + h_2 - 2) & t > h_2
 \end{array} \right. \quad (\text{A.15b})$$

513 **Appendix B. Overview of the different parameters used in the B-**
 514 **cell and T-cell dynamic models**

515 Table B.7 summarizes the parameters used in the functions of the different
 516 B-cell models. Table B.8 summarizes the parameters used in the functions
 517 of the different T-cell models.

Parameter	Description
B_0	initial number of B-cells
aB	constant proliferation rate of B-cells
aSB	constant proliferation rate of SB (in case of a distinction)
aLB	constant proliferation rate of LB (in case of a distinction)
pB	proportional proliferation rate of B-cells
uB	decay of (short living, in case of a distinction) B-cells
h	time point after which no new B-cells are activated

Table B.7: **An overview of the different parameters in the B-cell dynamic models**

Parameter	Description
T_0	initial number of T-cells
a_T	constant proliferation rate of T-cells after the i th vaccination
a_{ST}	constant proliferation rate of short living T-cells (in case of a distinction) after i th vaccination
a_{LT}	constant proliferation rate of long living T-cells (in case of a distinction)
u_T	decay of (short living, in case of a distinction) T-cells
h_i	time point after which no new T-cells are activated after the i th vaccination

Table B.8: **An overview of the different parameters in the T-cell dynamic models**

518 **Appendix C. Algorithm parameter values used in Monolix**

519 The following list shows a summary of the values used in the SAEM-
 520 MCMC algorithm and loglikelihood estimation in Monolix:

Population parameters	SAEM	$K_0 = 100$ $K_1 = 10^6$ $K_2 = 10^5$ $a_1 = 0$ $a_2 = 1$
	MCMC	$m_1 = 2$ $m_2 = 0$ $m_3 = 2$ $m_4 = 2$ $\rho = 0.3$
	Simulated annealing	$\tau_1 = 0.95$ $\tau_2 = 0.95$
Individual parameters	MCMC	$m_1 = 2$ $m_2 = 0$ $m_3 = 2$ $m_4 = 2$ $\rho = 0.3$
	Stopping rule	$L_{mcmc} = 1000$ $r_{mcmc} = 0.01$
Loglikelihood	Importance sampling	Monte-Carlo size = 10^7

521 **Appendix D. B-cell and T-cell dynamics associations and dataset**
522 **correlations**

523 *Appendix D.1. B-cell and T-cell dynamics associations*

524 Since the process of B-cell activation is dependent on certain cytokine-
525 expressing T-cells (such as CD40L), the hypothesis is investigated whether
526 the increase in B-cells is proportional to the increase in T-cells. In order
527 to examine this, we define $T_{01} := T(1) - T(0)$ and $B_{01} := B(1) - B(0)$
528 and use the `minerva` package in R to calculate the maximal information
529 coefficient (MIC). This is a way to detect linear and non-linear relations
530 between variables, and can thereby be used to indicate whether a linear
531 relation is feasible by comparing it with the R-squared value. In addition we
532 compute the Spearman correlation.

533 We started by examining the Varilrix-specific B-cell and T-cell datasets. The
534 datasets were restricted to individuals without missing values of B-cell or
535 T-cell data at time points 0 and 1 month, 96 individuals in total. First,
536 the hypothesis was made that an increase in T-cells was proportional to an
537 increase in B-cells. We express the expected proportionality factor by m :

$$m = E \left(\frac{B(1) - B(0)}{T(1) - T(0)} \right). \quad (\text{D.1})$$

538 Fig 5 shows a scatterplot of $T_{01} := T(1) - T(0)$ plotted against $B_{01} :=$
539 $B(1) - B(0)$. At first sight, a linear relation between T_{01} and B_{01} might not
540 seem a reasonable assumption.

541 To further examine this, we calculated the Spearman correlation and
542 the maximal information coefficient (MIC) as a way to assess and measure
543 (non)linear relationships between datasets.

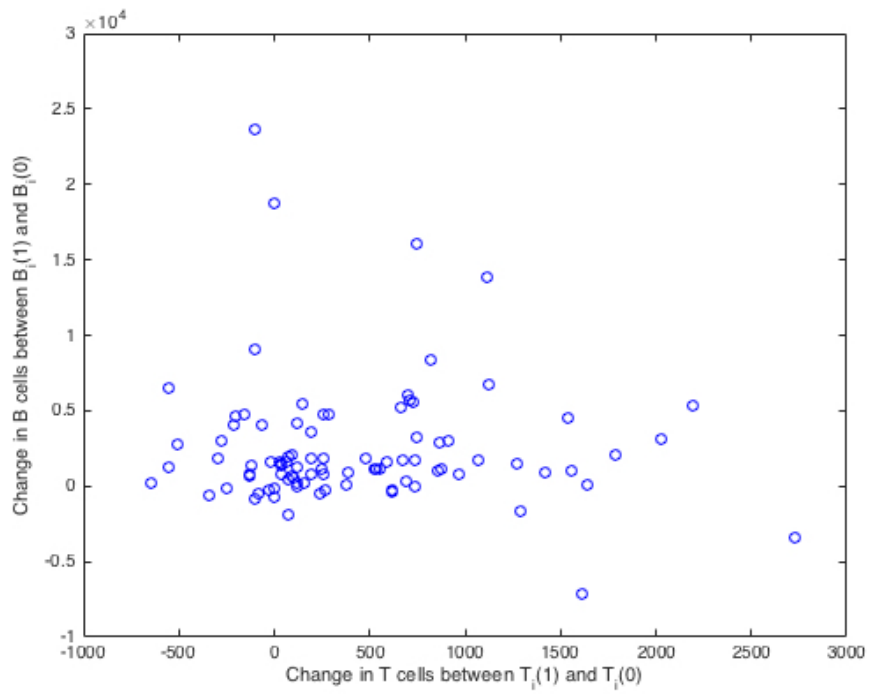


Figure D.5: Scatterplot of the increase in Varilrix-specific T-cells (T_{01}), plotted against the increase in Varilrix-specific B-cells (B_{01}).

544 The Spearman correlation between Varilrix-specific T_{01} and B_{01} was -
545 0.0274 ($p = 0.7914$), which rejected the hypothesis that increases in Varilrix-
546 specific T-cells were associated to increases in Varilrix-specific B-cells. A
547 non-significant MIC of 0.2161 (5% significance level) confirmed this result.
548 We also assessed MIC and Spearman correlations on the same datasets per
549 subgroup, reaching the same conclusion (Table D.9).

group	MIC	p-value	Spearman	p-value
3	0.2854	> 0.05	0.2096	0.2663
4	0.1673	> 0.05	-0.2192	0.2271
5	0.2611	> 0.05	0.1009	0.6380

Table D.9: **Correlations between B-cells and T-cells (Varilrix stimulus)**. MIC coefficients, Spearman correlation and corresponding p-values between increase in Varilrix-specific B-cells (B_{01}) and increase in Varilrix-specific T-cells (T_{01}), calculated for groups 3, 4 and 5. As the sample sizes of groups 1 and 2 were too small ($n_1 = 4$ and $n_2 = 6$), those groups were omitted from the analysis.

550 Potential associations between increases in gE-specific T-cells and B-cells
551 were investigated. The scatterplot of T_{01} plotted against B_{01} is shown in
552 Figure D.6. As before, we also studied the relations between T_{01} and B_{01}
553 per subgroup (see Table D.10). The Spearman p-values showed that only in
554 group 3 the Spearman correlation could be considered significant, together
555 with MIC, implying a nonlinear relationship. The Spearman correlation of
556 0.3833 ($p = 1.0647e^{-04}$) suggested there was indeed an association. The MIC
557 score was calculated as 0.4107, which was significant ($p < 0.001$). In case of
558 a linear relation, the R-square is expected to be close to this MIC score. As
559 the R-square was equal to 0.05593, we could exclude a linear relationship.

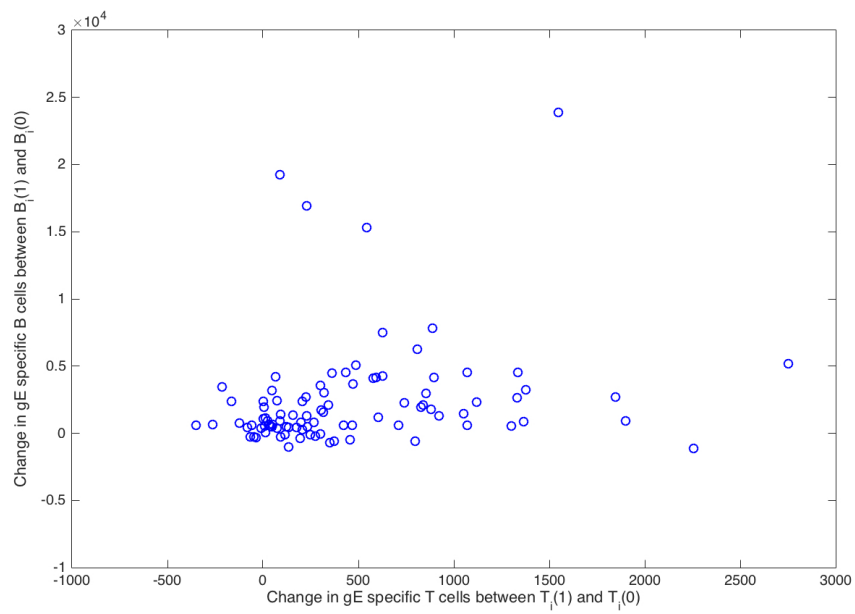


Figure D.6: Scatterplot of the increase in gE-specific T-cells (T_{01}), plotted against the increase in gE-specific B-cells (B_{01}).

group	MIC	p-value	MIC- R^2	Spearman	p-value
3	0.4734	0.01007	0.4665	0.4725	0.0084
4	0.3052	> 0.05		-0.2300	0.2053
5	0.3149	> 0.05		0.3138	0.1266

Table D.10: **Correlations between B-cells and T-cells (gE stimulus)**. MIC coefficients, Spearman correlation and corresponding p-values between increase in gE-specific B-cells (B_{01}) and increase in gE-specific T-cells (T_{01}), calculated for groups 3, 4 and 5. As the sample sizes of groups 1 and 2 were too small ($n_1 = 4$ and $n_2 = 6$), those groups were omitted from the analysis.

560 *Appendix D.2. Correlations between B-cell and T-cell datasets*

561 Instead of solely examining the increase in B-cells and T-cells, we also
562 looked at the correlation between the specific values of B- and T-cells at
563 time points 0 and 1; Spearman correlations for all individuals were calculated
564 between the data points B(0), B(1), T(0) and T(1).

565 Correlations between the initial number of B-cells B(0), the initial number
566 of T-cells T(0), the number of B-cells at month 1 B(1) and the number of
567 T-cells at month 1 T(1) have also been investigated.

568 We started with the Varilrix-specific B-cell and T-cell data. Again, we only
569 included individuals for whom we had data points B(0), B(1), T(0) and T(1).
570 Spearman correlations for all individuals were calculated between these data
571 points, and the results are shown in Fig D.7. This figure also shows the
572 Spearman correlations when we separated the individuals by group. When
573 looking at the correlations of all individuals, we noticed correlations between
574 B(0) and B(1), and T(0) and T(1), but no significant correlations between
575 B- and T-cells. Examining the correlations by group, it seemed those differ

576 greatly depending on the group. However, it has to be noted that group
577 1 and 2 contain a very small number of individuals (4 and 6 individuals
578 respectively). For groups 3 to 5, the results were similar to the results of the
579 correlation between all individuals, and we therefore found no convincing
580 evidence for an association between Varilrix specific B-cells and T-cells.

581 The correlation between data points at time 0 and 1 of gE-specific B-cells
582 and T-cells was examined next. Fig D.8 shows the Spearman correlations,
583 first for all individuals and then split by group. We observed that some corre-
584 lations seemed to be higher compared to the Varilrix-specific data, however,
585 when examining the Spearman matrices by group, the values between the
586 different groups seemed to vary widely. For this reason, we did not find de-
587 cisive evidence to include the number of gE-specific T-cells into the B-cell
588 models or vice versa.

589 *Appendix D.3. Influence of the initial number of B-cells to the remaining*
590 *B-cell data*

591 Apart from associations between the B-cell and T-cell datasets, it was
592 investigated as well whether the initial number of B-cells ($B(0)$) had an
593 influence on the short term number of B-cells ($B(1)$), the long term number
594 ($B(12)$) and the decay of B-cells ($B(12) - B(1)$). This was achieved by
595 plotting for each individual $B(0)$ first against $B(1)$, then against $B(12)$ and
596 lastly against $B(12) - B(1)$. If $B(0)$ had influence on either one of these
597 values, a linear relation should be clear from the plot. We started with the
598 Varilrix-specific B-cell data. As can be perceived from Fig D.9, no clear linear
599 relation was found. The same values were plotted afterwards, but now on a
600 logarithmic scale. An evident linear relation could not be detected either.

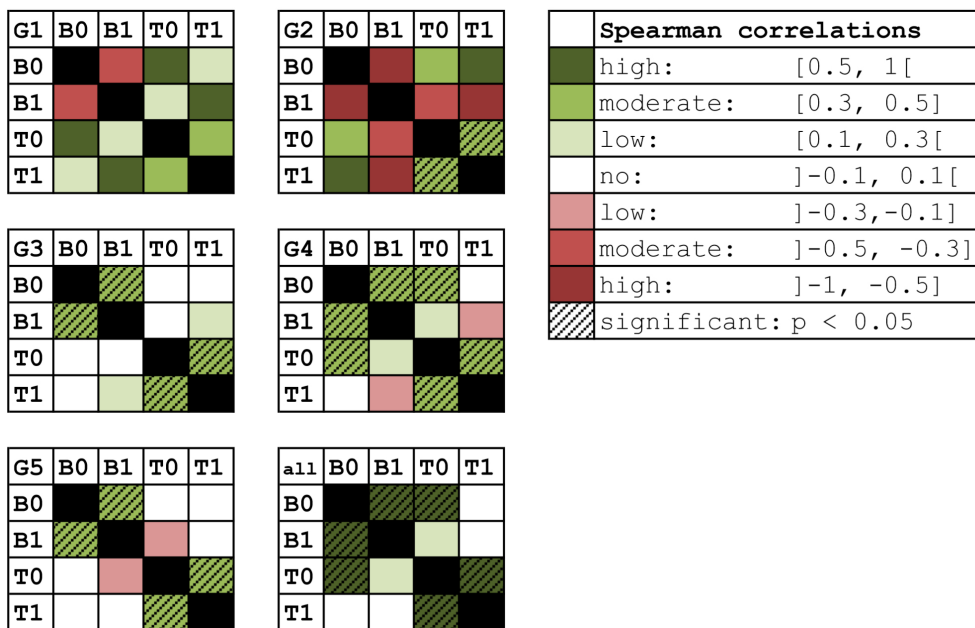


Figure D.7: Spearman correlation between Varilrix-specific data points B(0), B(1), T(0) and T(1), shown for all individuals and per group. Significant correlations are indicated.

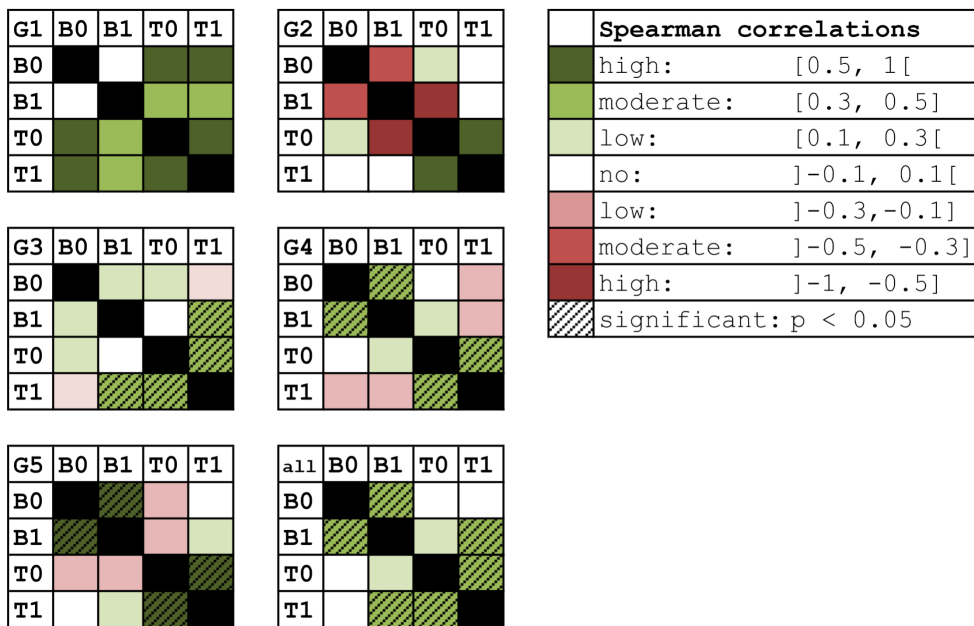


Figure D.8: Spearman correlation between gE-specific data points B(0), B(1), T(0) and T(1), shown for all individuals and per group. Significant correlations are indicated.

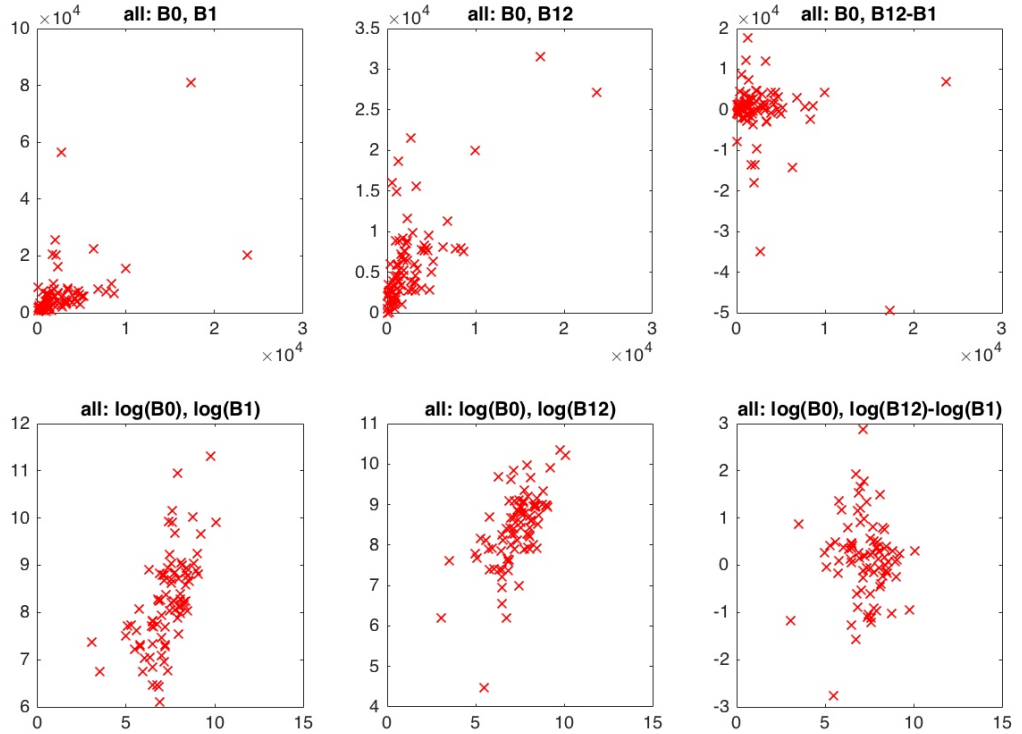


Figure D.9: **Associations between Varilrix-specific B-cell data points.** Above: Scatterplots of Varilrix-specific $(B(0), B(1))$, $(B(0), B(12))$, $(B(0), B(12) - B(1))$. Under: Scatterplots of $(B(0), B(1))$, $(B(0), B(12))$, $(B(0), B(12) - B(1))$, on logarithmic scale

601 Fig D.10 examines the influence of the initial gE-specific B-cell value on
 602 the number of short term B-cells, long term and decay of B-cells. There
 603 seemed to be even less evidence to assume a linear relation between B-cell
 604 values compared to the Varilrix-specific B-cells, on logarithmic scale neither.
 605 We can conclude the initial value of B-cells has no definite influence on fol-
 606 lowing B-cell values.

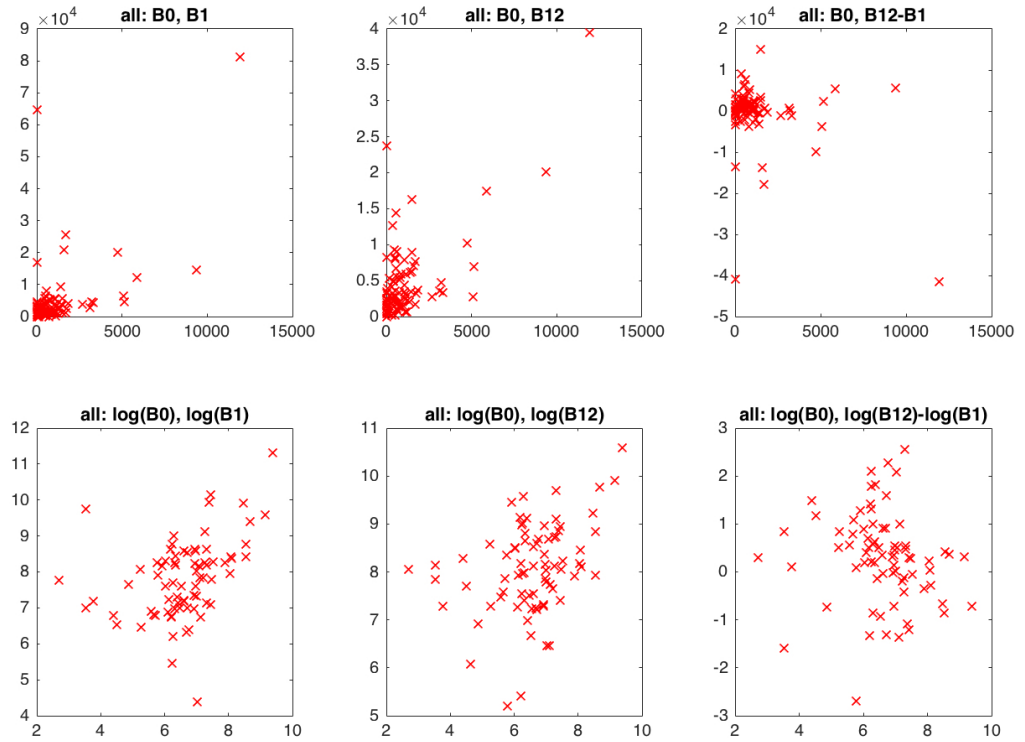


Figure D.10: **Associations between gE-specific B-cell data points.** Above: Scatterplots of gE-specific $(B(0), B(1))$, $(B(0), B(12))$, $(B(0), B(12) - B(1))$ Under: Scatterplots of $(B(0), B(1))$, $(B(0), B(12))$, $(B(0), B(12) - B(1))$, on logarithmic scale

607 **Appendix E. Detailed model selection procedures**

608 *Appendix E.1. Model selection of B-cell datasets*

609 We started by modeling the Varilrix-specific B-cell dataset. The upper
 610 part of Table E.11 summarizes the differences between all models that were
 611 tested.

model	fixed population parameter	group-specific parameter	convergence-AIC
B1a	-	aB	6559
B1b	uB	aB	6595
B1c	h	aB	no
B2	-	paB	6603
B3a	-	aSB	no
B3b	uB	aSB	6541
B3c	h	aSB	no
B4a(i/ii)	-	aSB / aLB	no / no
B4b(i/ii)	uB	aSB / aLB	no / no
B4c	actL	aSB	no
B4d(i/ii)	h	aSB / aLB	no / no
B4e(i/ii)	uB, h	aSB / aLB	no / no
B4f	aLB, h	aSB	no
B4g	aLB, uB	aSB	6524
candidate model	bootstrap convergence	deviating IDs	bootstrap w/o ID
B4g	no		

B3b	65%	ID54	73%
-----	-----	------	-----

Table E.11: **ODE Model formulations considered for Varilrix-specific B-cell data and model selection procedure.** Upper part: Overview of the different models used to model the Varilrix-specific B-cell data. First column: model identifier. Second column: parameters selected as fixed population parameter. Third column: parameter which is chosen to be group-specific. Fourth column: AIC value of each model, in case of convergence.

Lower part: Overview of the considered candidate models (first column) used in the Varilrix-specific B-cell data model selection procedure. Second column: convergence results of the performed bootstraps. Third column: the results of possible IDs with deviating presence in the converging bootstrap samples. Fourth column: results of a bootstrap performed on the Varilrix-specific B-cell dataset, in case IDs are found.

612 For model B1, we distinguished between the following scenarios: a sce-
613 nario with random effects for all parameters (model B1a), a scenario in which
614 the decay of B-cells (uB) was assumed having a fixed population parameter
615 (model B1b), and a scenario in which the time period where B-cells were
616 activated (h) was fixed (model B1c). SAEM convergence was obtained for

617 models B1a and B1b, with an AIC value of 6559 and 6595, respectively.
618 Model B1c did not converge within $10^6 + 10^5$ iterations and was therefore
619 not considered any further.

620 Next, a proportional proliferation rate was explored in model B2. Model
621 B2 assumed a group specific proliferation parameter paB and all parameters
622 having random effects. For this model, an AIC value of 6603 was obtained.
623 Consequently, we looked at models in which a distinction between SB and
624 LB was made. Model B3a assumed all parameters had random effects. In
625 models B3b and B3c, uB and h , respectively, were set as fixed population
626 parameter. Model B3b was the only model that showed convergence, with
627 an AIC value of 6541.

628 The last model examined was model B4, in which a proliferation rate for
629 LB was added. Many assumptions on the parameters were made; the decay
630 rate of B-cells (uB), proliferation rate of LB (aLB) and activation period
631 (h) were set as fixed parameters in models B4b, B4c and B4d, respectively.
632 Combinations of these fixed parameters were considered as well in models
633 B4f, B4g and B4h. Apart from this, we also looked at different group spe-
634 cific parameters, not only the proliferation rate of SB (aSB) was considered,
635 but the proliferation rate of long living B-cells (aLB) as well. Model B4g
636 was the only model that accounted for aSB and aLB and still converged. In
637 this model, both aLB and uB were set as fixed population parameters, and
638 aSB was considered to be a group specific parameter. This model had the
639 lowest AIC value of 6524 among all aforementioned models, and was selected
640 as first candidate model.

641 The bootstrap that subsequently was performed did not converge for model

642 B4g, and it was therefore in the end rejected.

643 Model B3b was selected as next-candidate model. The converging of its
644 bootstrap was successful, with 65% of bootstrap samples showing proper
645 SAEM convergence. An analysis was done to explore whether the presence
646 (or absence) of certain individuals was responsible for the convergence of
647 the datasets. When examining percentages of presence, ID54 showed de-
648 viant behavior: the profile was absent from a significant number of the non-
649 converging datasets. It was therefore removed from the B-cell dataset, after
650 which a new bootstrap with the candidate model was performed. Bootstrap
651 convergence remained well (73 % of bootstrap samples showed SAEM conver-
652 gence) and model B3b was therefore selected as final Varilrix-specific B-cell
653 model. Model B3b differentiates between SB and LB, LB are assumed to
654 remain constant through time. In time period $(0, h)$, a constant number of
655 SB is activated and this proliferation rate is considered a group-specific pa-
656 rameter. The decay rate is assumed to be equal for each individual.

657 This model selection procedure is summarized in the lower part of Table E.11.

658 In the upper part of Table E.12, models and convergence results are
659 shown. Model B3a proved to be the model with lowest AIC value that
660 still provided sufficient bootstrap convergence (68% of all datasets had good
661 SAEM convergence, see lower part of Table E.12). Next, it was investigated
662 whether the convergence for the bootstrap samples was influenced by the
663 presence or absence of certain individual profiles. Indeed, it was discovered
664 that 73% of the bootstrap samples resulting in non-converge, included ID48
665 in their dataset. For this reason, a new bootstrap was performed with model
666 B3a, but now on the gE-specific B-cell dataset, excluding ID48. Bootstrap

667 convergence was no longer reached and model B3a was discarded. Model
668 B1a, which also showed sufficient bootstrap convergence (68% of bootstrap
669 samples converged) was considered thereafter. As with previous model, we
670 investigated the relation between the bootstrap convergence and presence or
671 absence of individual profiles. No deviant occurrence was found and model
672 B1a was thereby selected as final gE-specific B-cell model.

model	fixed population parameter	group-specific parameter	convergence-AIC
B1a	-	aB	6233
B1b	uB	aB	6267
B1c	h	aB	6306
B2	-	paB	6304
B3a	-	aSB	6227
B3b	uB	aSB	6231
B3c	h	aSB	no
B4a(i/ii)	-	aSB / aLB	6222 / 6207
B4b(i/ii)	uB	aSB / aLB	no / no
B4c	actL	aSB	no
B4d(i/ii)	h	aSB / aLB	no / 6257
B4e(i/ii)	uB, h	aSB / aLB	6267 / 6246
B4f	aLB, h	aSB	no
B4g	aLB, uB	aSB	6222

candidate model	bootstrap convergence	deviating IDs	bootstrap w/o ID
B4a(ii)	no		
B4a(i)	no		
B4g	no		
B3a	68%	ID48	no
B3b	no		
B1a	68%	no	

Table E.12: **ODE Model formulations considered for gE-specific B-cell data and model selection procedure.** Upper part: Overview of the different models used to model the gE-specific B-cell data. First column: model identifier. Second column: parameters selected as fixed population parameter. Third column: parameter which is chosen to be group-specific. Fourth column: AIC value of each model, in case of convergence. Lower part: Overview of the considered candidate models (first column) used in the gE-specific B-cell data model selection procedure. Second column: convergence results of the performed bootstraps. Third column: the results of possible IDs with deviating presence in the converging bootstrap samples. Fourth column: results of a bootstrap performed on the gE-specific B-cell dataset, in case IDs are found.

673 *Appendix E.2. Model selection of T-cell datasets*

674 Just like with the B-cell models, the most simplistic T-cell model T1
675 was considered first, in which $a_1T = a_2T = aT$. Model T1a assumed all
676 parameters had random effects, and the activation of T-cells (aT) had a
677 group-specific effect. An AIC value of 11,664 was obtained.

678 When assuming $a_1T \neq a_2T$, we arrived at model T2. First, the assumption
679 was made that all parameters had random effects, and both a_1T and a_2T
680 had a group-specific effect in model T2a. This model converged, with an
681 AIC value of 11,658. When assuming only a_2T had a group-specific effect
682 (model T2b), a slightly lower AIC value was obtained at 11,655.

683 Next, a distinction between short and long living T-cells (ST and LT, respec-
684 tively) was considered. In model T3a, all parameters had random effects and
685 aST ($= a_1ST = a_2ST$) had a group-related effect resulting in an AIC value
686 of 11,637. Model T3b, in which uT was a fixed parameter, did not reach
687 convergence.

688 Model T4 assumed different activation rates of T-cells after each vaccina-
689 tion. When all parameters had random effects, and both a_1T and a_2T were
690 group specific, SAEM convergence was not reached (model T4a). When only
691 a group specific effect on a_2T was assumed (model T4b), convergence was
692 achieved resulting in an AIC value of 11,615. Subsequently, uT was set as
693 fixed parameter in models T4c and T4d, again with group specific effects
694 on both activation rates (T4c) and on a_2T only (T4d), respectively. Both
695 models showed SAEM convergence with an AIC value of 11,646 and a lower
696 AIC value of 11,626, respectively.

697 When assuming LT activation according to a constant proliferation rate

698 (equal after each vaccination in order to limit the number of parameters
699 to be estimated), models T5 and T6 were reached. In models T5, the activa-
700 tion rates of ST were presumed equal after each vaccination. Together with
701 the assumption that all parameters were random, and aST was a group spe-
702 cific parameter, this led to model T5a, where an AIC value of 11,631 was
703 found. We note that setting aLT as a group specific parameter was tested as
704 well, but none of these models (including the following) showed convergence
705 and thus were omitted from Table E.13.

model	fixed population parameter	group-specific parameter	convergence- AIC
T1a	-	aT	11664
T2a	-	a_1T, a_2T	11658
T2b	-	a_2T	11655
T3a	-	aST	11637
T3b	uT	aST	no
T4a	-	a_1ST, a_2ST	no
T4b	-	a_2ST	11615
T4c	uT	a_1ST, a_2ST	11646
T4d	uT	a_2ST	11626
T5a	-	aST	11631
T5b	uT	aST	11637
T5c	aLT	aST	11640
T5d	aLT, uT	aST	11667
T6a	-	a_1ST, a_2ST	no

T6b	-	a_2ST	no
T6c	uT	a_1ST, a_2ST	no
T6d	uT	a_2ST	11630
T6e	aLT	a_1ST, a_2ST	no
T6f	aLT	a_2ST	11624
T6g	aLT, uT	a_2ST	no
candidate model	bootstrap convergence	deviating IDs	bootstrap w/o ID
T4b	no		
T6f	no		
T4d	no		
T6d	no		
T5a	no		
T5b	no		
T3a	66 %	no	
T3a'	67 %	no	

Table E.13: **ODE Model formulations considered for Varilrix-specific T-cell data and model selection procedure.** Above: Overview of the different models used to model the Varilrix-specific T-cell data. First column: model identifier. Second column: parameters selected as fixed population parameter. Third column: parameter which is chosen to be group-specific. Fourth column: AIC value of each model, in case of convergence. Under: Overview of the considered candidate models (first column) used in the Varilrix-specific T-cell data model selection procedure. Second column: convergence results of the performed bootstraps. Third column: the results of possible IDs with deviating presence in the converging bootstrap samples. Fourth column: results of a bootstrap performed on the Varilrix-specific T-cell dataset, in case IDs are found.

706 In Models T5b and T5c, respectively, uT and aLT were assumed to be
 707 fixed population parameters. They showed SAEM convergence, with AIC
 708 values of 11,637 and 11,640, respectively. Setting both uT and aLT fixed in
 709 model T5d did not improve the model (AIC: 11,667).

710 Finally, model T6 was considered, with different activation rates after each
 711 vaccination event. Assuming all parameters were random and either both
 712 a_1ST and a_2ST (T6a), or only a_2ST (T6b), were group specific, did not lead

713 to convergence within $10^6 + 10^5$ iterations. For this reason, fixed parameters
714 uT and/or aLT were considered. When uT was fixed, and both activation
715 rates were group specific, convergence was not reached (model T6c). With
716 a_2ST being group specific (model T6d), convergence was obtained with an
717 AIC value of 11,630. In case of setting aLT as a fixed parameter, similar
718 results were found; assuming both activation rates to be group specific did
719 not lead to convergence, but assuming only a_2ST was group specific, did,
720 with a slightly lower AIC value equal to 11,624. The last scenario assumed
721 both aLT and uT were fixed population parameters, though no convergence
722 was obtained.

723

724 Since model T4b was the model with lowest AIC (11,615), it was selected
725 as first candidate model. However, model T4b was subsequently rejected as
726 a 1000 sample bootstrap failed to converge. Likewise models T6f, T4d, T6d,
727 T5a and T5b did not have proper bootstrap convergence. Next, model T3a
728 was selected as candidate model and showed bootstrap convergence; 66% of
729 the bootstrap samples reached SAEM convergence. As before, a search for
730 frequently deviant profiles in the converging and non-converging bootstrap
731 datasets was performed, but no such profile was identified.

732 Taking into account that the assumption that $a_1ST = a_2ST = aST$ might
733 not be a realistic assumption in a model that described a real life cellular
734 process, a difference in proliferation rates after each vaccination was inserted,
735 assuming a_2ST was proportional to a_1ST : $a_2ST = k \times a_1ST$. Adding this
736 parameter to the pool of parameters to be estimated in the SAEM algorithm,
737 did not yield a converging model. In order to limit the number of parameters

738 to be estimated by the SAEM-algorithm, different values of k were set fixed
739 and for each subsequent model, AIC values were compared. A model with
740 $k = 0.15$ showed the lowest AIC value (11,623), which we named model T3a',
741 and a bootstrap with 100 bootstrap samples was performed on this model.
742 From this bootstrap, 67% of samples resulted in SAEM convergence and
743 a search for disproportionate presence of deviant profile(s) did not identify
744 such profiles. Therefore, model T3a', a model which differentiates between
745 ST and LT, was selected as final Varilrix-specific T-cell model.

746 Table E.14 summarizes the models used in the selection procedure of
747 gE-specific T-cells. One individual (ID149), with deviating T-cell data, has
748 been left off from the model selection procedure since all models of the dataset
749 including this individual did not converge. Model T1a was the singular model
750 with a converging SAEM algorithm and was thereby selected as candidate
751 model. The bootstrap that was performed afterwards converged as well.

752 However, it was found that ID89 heavily influenced the converging value
753 of one of the parameters (h). For this reason, ID89 was omitted from the gE
754 T-cell dataset. Without this parameter, the bootstrap convergence of model
755 T1a remained.

756 In the next step, we included parameter a_2T in the model, with $a_2T = k \times a_T$.
757 k was estimated as 0.66 by comparing AIC-values for models with various
758 values of k . The resulting model, T1a', showed bootstrap convergence and
759 the bootstrap samples did not possess deviating presence of certain IDs. For
760 this reason, model T1a' was selected as final gE-specific T-cell model.

model	fixed population parameter	group-specific parameter	convergence- AIC
-------	-------------------------------	-----------------------------	---------------------

T1a	-	aT	11658
T2a	-	a_1T, a_2T	no
T2b	-	a_2T	no
T3a	-	aST	no
T3b	uT	aST	no
T4a	-	a_1ST, a_2ST	no
T4b	-	a_2ST	no
T4c	uT	a_1ST, a_2ST	no
T4d	uT	a_2ST	no
T5a	-	aST	no
T5b	uT	aST	no
T5c	aLT	aST	no
T5d	aLT, uT	aST	no
T6a	-	a_1ST, a_2ST	no
T6b	-	a_2ST	no
T6c	uT	a_1ST, a_2ST	no
T6d	uT	a_2ST	no
T6e	aLT	a_1ST, a_2ST	no
T6f	aLT	a_2ST	no
T6g	aLT, uT	a_2ST	no
candidate model	bootstrap convergence	deviating IDs	bootstrap w/o ID
T1a	yes	ID89	yes
T1a'	yes	no	

Table E.14: **ODE Model formulations considered for gE-specific T-cell data and model selection procedure.** Above: Overview of the different models used to model the gE-specific T-cell data. First column: model identifier. Second column: parameters selected as fixed population parameter. Third column: parameter which is chosen to be group-specific. Fourth column: AIC value of each model, in case of convergence.

Under: Overview of the considered candidate models (first column) used in the gE-specific T-cell data model selection procedure. Second column: convergence results of the performed bootstraps. Third column: the results of possible IDs with deviating presence in the converging bootstrap samples. Fourth column: results of a bootstrap performed on the gE-specific T-cell dataset, in case IDs are found.

761 **Acknowledgements**

762 This research was funded by the University of Antwerp [BOF Concerted
763 Research Action, Methusalem funding], NH acknowledges support from the
764 chair in evidence-based vaccinology sponsored by a gift from Pfizer (2009-
765 2018) and GSK (2016) and from the European Research Council (ERC) un-
766 der the European Union's Horizon 2020 research and innovation programme
767 (grant agreement 682540 - TransMID). This research was supported by the
768 Antwerp Study Centre for Infectious Diseases (BID). The funders had no role
769 in study design, data collection and analysis, decision to publish, or prepa-
770 ration of the manuscript.

771 All authors attest they meet the ICMJE criteria for authorship.

772 **References**

- 773 [1] Van Herck K, Beutels P, Van Damme P, Beutels M, Van den Dries
774 J, Briantais P, et al. Mathematical models for assessment of long-term
775 persistence of antibodies after vaccination with two inactivated hepatitis
776 A vaccines. *Journal of Medical Virology*. 2000;60(1):17.
- 777 [2] Hens N, Habteab Ghebretinsae A, Hardt K, Van Damme P, Van
778 Herck K. Model based estimates of long-term persistence of inac-
779 tivated hepatitis A vaccine-induced antibodies in adults. *Vaccine*.
780 2014;32(13):15071513. doi:10.1016/j.vaccine.2013.10.088.
- 781 [3] Theeten H, Van Herck K, Van Der Meeren O, Crasta P, Van Damme
782 P, Hens N. Long-term antibody persistence after vaccination with a 2-
783 dose Havrix (inactivated hepatitis A vaccine): 20 years of observed data,

- 784 and long-term model-based predictions. *Vaccine*. 2015;33(42):57235727.
785 doi:10.1016/j.vaccine.2015.07.008.
- 786 [4] Andraud M, Lejeune O, Musoro JZ, Ogunjimi B, Beutels P, Hens N.
787 Living on three time scales: The dynamics of plasma cell and antibody
788 populations illustrated for hepatitis A virus. *PLoS Computational Biol-*
789 *ogy*. 2012;8(3):18. doi:10.1371/journal.pcbi.1002418.
- 790 [5] Le D, Miller JD, Ganusov VV. Mathematical modeling provides ki-
791 netic details of the human immune response to vaccination. *Fron-*
792 *tiers in Cellular and Infection Microbiology*. 2015;4(January):113.
793 doi:10.3389/fcimb.2014.00177.
- 794 [6] Leroux-Roels I, Leroux-Roels G, Clement F, Vandepapelire P, Vassilev
795 V, Ledent E, et al. A Phase 1/2 Clinical Trial Evaluating Safety and
796 Immunogenicity of a Varicella Zoster Glycoprotein E Subunit Vaccine
797 Candidate in Young and Older Adults. *The Journal of Infectious Dis-*
798 *eases*. 2012;206(8):12801290. doi:10.1093/infdis/jis497.
- 799 [7] Amanna IJ, Slifka MK. Mechanisms that determine plasma cell lifes-
800 pan and the duration of humoral immunity. *Immunological Reviews*.
801 2010;236(1):125138. doi:10.1111/j.1600-065X.2010.00912.x.
- 802 [8] Lavielle M, Mentr F. Estimation of population pharmacokinetic param-
803 eters of saquinavir in HIV patients with the MONOLIX software. *Jour-*
804 *nal of Pharmacokinetics and Pharmacodynamics*. 2007;34(2):229249.
805 doi:10.1007/s10928-006-9043-z.

- 806 [9] Reshef D, Reshef Y, Finucane H, Grossman S, Mcvean G, Turnbaugh P,
807 et al. Detecting Novel Associations in Large Data Sets. *Science Transla-*
808 *tional Medicine*. 2011;334(6062):15181524. doi:10.1126/science.1205438.
- 809 [10] Didierlaurent AM, Laupze B, Di Pasquale A, Hergli N, Collignon
810 C, Garon N. Adjuvant system AS01: helping to overcome the chal-
811 lenges of modern vaccines. *Expert Review of Vaccines*. 2017;16(1):5563.
812 doi:10.1080/14760584.2016.121363
- 813 [11] Leroux-Roels G, Marchant A, Levy J, Van Damme P, Schwarz T, Hors-
814 mans Y, et al. Impact of adjuvants on CD4+ T cell and B cell re-
815 sponses to a protein antigen vaccine: Results from a phase II, random-
816 ized, multicenter trial. *Clinical Immunology*. 2016;169(2016):16-27. doi:
817 10.1016/j.clim.2016.05.007.

818 **Supplementary material**

819 **Fig S1. Amount of VZV IgG-secreting cells.** Measured by B-
820 cell ELISPOT (gE stimulus) per 10^6 IgG-secreting cells, up to 12 months.
821 Data are shown per study group. The last panel acts as an illustration
822 of the vaccination dynamics and shows a hypothetical, smooth function of
823 the expected change in number of memory B-cells over time (in months),
824 based on the observed data points per individual and considering the second
825 vaccination at month 2.

826 **Fig S2. Amount of gE-specific CD4+ T-cells, producing at least**
827 **2 immune markers.** Measured by ICS per 10^6 CD4+ T-cells, shown per
828 group and up to 12 months. The last panel acts as an illustration of the
829 vaccination dynamics and shows a hypothetical, smooth function of the ex-
830 pected change in number of CD4+ T-cells over time (in months), based on
831 the observed data points per individual.

832 **Dataset 1. Varilrix-specific memory B-cells.** Amount of VZV-
833 specific memory B-cells. Measured by B-cell ELISPOT (Varilrix stimulus)
834 per 10^6 total memory cells, up to 12 months.

835 **Dataset 2. gE-specific memory B-cells.** Amount of VZV-specific
836 memory B-cells. Measured by B-cell ELISPOT (gE stimulus) per 10^6 total
837 memory cells, up to 12 months.

838 **Dataset 3. Varilrix-specific CD4+ T-cells.** Amount of VZV-specific
839 CD4+ T-cells (Varilrix stimulus) producing at least 2 immune markers. Mea-
840 sured by ICS per 10^6 CD4+ T-cells up to 12 months.

841 **Dataset 4. gE-specific CD4+ T-cells.** Amount of VZV-specific
842 CD4+ T-cells (gE stimulus) producing at least 2 immune markers. Mea-

843 sured by ICS per 10^6 CD4+ T-cells up to 12 months.