Model based estimates of long-term persistence of induced HPV antibodies; a flexible subject-specific approach

Peer-reviewed author version

Model Based Estimates of Long-Term Persistence of Induced HPV Antibodies: A Flexible Subject-Specific Approach

Mehreteab Aregay[1]   Ziv Shkedy[2]   Geert Molenberghs[2,1]

Marie-Pierre David[3]   Fabián Tibaldi[3]

[1] *I-BioStat, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium*

[2] *I-BioStat, Universiteit Hasselt, B-3590 Diepenbeek, Belgium*

[3] *GlaxoSmithKline Biologicals, 89 Rue de l'Institut, B-1330, Rixensart, Belgium*

Abstract

In infectious diseases, it is important to predict the long-term persistence of vaccine-induced antibodies and to estimate the time points where the individual titers are below the threshold value for protection. This article focuses on HPV-16/18, and uses a so-called fractional-polynomial model to this effect, derived in a data-driven fashion. Initially, model selection was done from among the second- and first-order fractional polynomials on the one hand, and the linear mixed model on the other. According to a functional selection procedure, the first-order fractional polynomial was selected. Apart from the fractional polynomial model, we also fitted a power law model, which is a special case of the fractional polynomial model. Both models were compared using Akaike's Information Criterion. Over the observation period, the fractional polynomials fitted the data better than the power-law model; this, of course, does not imply that it fits best over the long run and hence caution ought to be used when prediction is of interest. Therefore, we point out that the persistence of the anti-HPV responses induced by these vaccines can only be ascertained empirically by long-term follow-up analysis.

Some Keywords: Akaike's information criterion; Fractional polynomial model; Functional selection procedure; Power-law model.

1   Introduction

It has been well-established that HPV is a sexually transmitted agent that causes most invasive cervical cancers and their associated pre-cancerous lesions. HPV infection is also a cause of penile, vaginal, and anal cancers (Rubin *et al.*, 2001; Daling *et al.*, 2002; Clark *et al.*, 2004). HPV-16/18 is estimated to account for 70% of all cervical cancers worldwide (Muñoz *et al.*, 2003).

The growth of serum antibodies in HPV-16 infected women appears to be a slow process and antibodies are not necessarily found in all infected women (Baseman and Koutsky, 2005). A median time of IgG

1

sero-conversion to HPV-16 is about 8.3 month (Ho *et al.*, 2004; Carter *et al.*, 1996).

To prevent cervical pre-cancerous lesions and cancer, many studies have focused on the prospects of vaccination against HPV. Using only two HPV-16 expressed genes (L1 and L2) in eukaryotic cells, Zhou *et al.* (1991) produced virus-like particles (VLPs). A multi-center phase IIb study on prophylactic HPV L1 VLP vaccines reported 100% protection against persistent HPV-16 infection (Koutsky *et al.*, 2002).

Several studies have mathematically modeled antibody decay following vaccination and/or natural infection for the purpose of predicting long-term immunity (Bovier *et al.*, 2002; Gesemann and Scheiermann, 1995; Glass and Grenfell, 2004; Honorati *et al.*, 1999; Mossong *et al.*, 1999). When developing a model to predict long-term immunity, several dynamic factors must be considered, including rates of B-cell decay and proliferation, B-cell immune memory, cell mediated immunity, and individual variability. David *et al.* (2009) and Fraser *et al.* (2007) studied the HPV-16/18 data set (GlaxoSmithKline Biologicals, HPV Protocols 001/007), which will be used in this manuscript, and the HPV-16 L1 VLP vaccine trial (Merck Research Laboratories, HPV Protocol 005), respectively, using a conventional power law as well as a modified power law model. However, the first model has been limited to the assumption of a progressive decay of antibody and antibody-producing B-cells, while the second model assumes, in addition, that the proportion of memory B-cells remains stable and identical for all women, which is biologically unlikely.

The objective of this paper is to predict the long-term persistence of vaccine-induced anti-HPV-16 and anti-HPV-18 antibodies, to obtain the time at which the antibody level of an individual crosses the threshold value and to predict the proportion of subjects above the threshold value using a flexible subject-specific model based on fractional polynomials; FPs.

This paper is structured as follows. In Section 2, we present the vaccination trials of HPV-16 and HPV-18. In Section 3, we formulate the mean structure for antibody decay using fractional polynomial models. In Section 4, the long-term prediction with subject-specific models based on FPs will be discussed.

2 The HPV-16 and HPV-18 Vaccination Trials

The data analyzed here consist of healthy women aged 15–25 years, who were cytologically negative, sero-negative for HPV-16 and HPV-18 antibodies by enzyme-linked immunosorbent assay (ELISA), and HPV-DNA-negative by PCR for 14 high-risk HPV types no more than 90 days before study entry. 393 women who received three doses of vaccine and participated in the initial efficacy study were enrolled into an extended follow-up study (Bovier *et al.*, 2002).

In the initial phase, blood samples from the 393 women in the HPV group of the extension study, and who came from North America (USA and Canada) and Brazil, were evaluated at months 7, 12, and 18, and annually thereafter up to 6.4 years after first vaccination, for the presence of HPV-16/18 antibodies using ELISA, developed in-house by GSK. The assay cut-off value was 8 EU/mL for HPV-16 and 7 EU/mL for HPV-18. For the current evaluation, we included women who had received three doses of AS04-adjuvanted HPV-16/18 vaccine and had at least one time point after the third dose with serology results available for at least one vaccine antigen component. For more information about this trial we refer to Harper *et al.* (2006) and David *et al.* (2009).

2.1 Exploratory Data Analysis

Sample sizes for each categorized month (visit) are shown in Figure 1. A drop in sample size reveals itself at the months in the bracket 25–32. The highest sample sizes for HPV-16 are encountered at month 7; thereafter there is a small increase in number of blood samples from 364 to around 366 in month 12 and then a small decrease to 365 in month 18, but a rapid decrease to 89 shows in the month in the range 25–32. From months around 25–32 to months in the range 69–74, the sample size varies between 130 to 234, followed by a drop to 66 in the interval [M75–...]. A similar pattern is observed for HPV-18 in the right panel of Figure 1.

Individual profile curves for both studies are shown in Figure 2 and reveal substantial variability between subjects. We note that the decline in antibody level is higher in the first few months followed by a moderate decrease until the end of the follow-up period. This pattern can be seen in Figure 3 at the top, displaying the evolution of mean antibody over time and a loess estimate for the mean antibody. The plots reveal a pattern
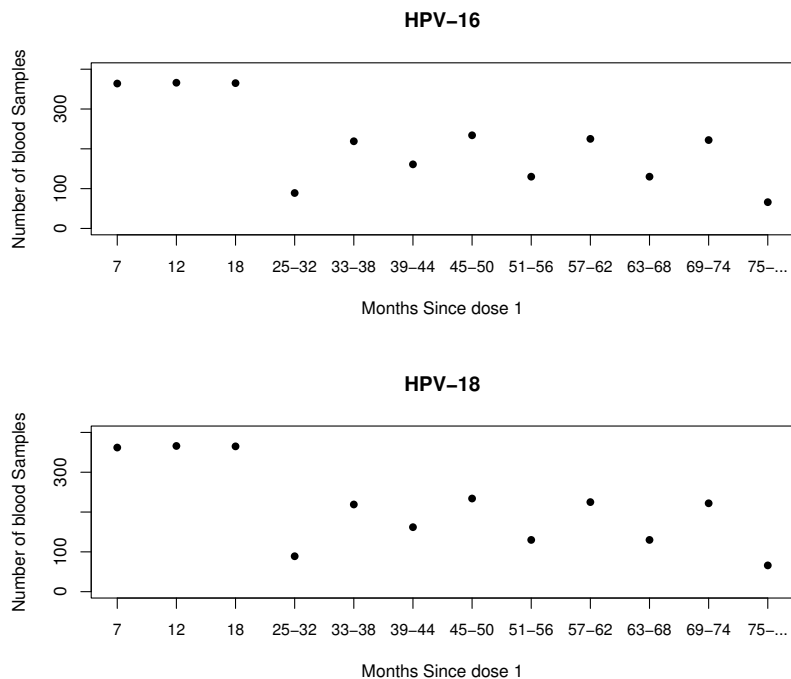
3

Figure 1: *Number of blood samples for HPV-16 and HPV-18 by visit (time intervals: continuous time was categorized into 12 time points).*

of a quick drop in the mean of antibody titers of HPV-16 from 3.58 to 2.86 until month 18, followed by a stabilized level between 2.59 and 2.82 for the rest of the month. A similar pattern is observed for HPV-18.

The observed variances at each visit are shown in the bottom two graphs of Figure 3. We note that the variance is not constant over time, implying that a random-intercept model might not be an appropriate model for these studies; hence, random slopes are incorporated as well.

3 Modeling Mean Antibody Using Subject-specific Models

3.1 The Power-law(PL) and Piece-wise Models

Fraser *et al.* (2007) estimated the persistence of anti-HPV levels over time using two mixed-effects models (Verbeke and Molenberghs 2000) that took into account the rates of B-cell decay. The first mixed-effect model that they used is a conventional power law model, given by:

$$Y_{ij} = f(t_{ij}) + \varepsilon_{ij}, \;\; i = 1, \ldots, n; j = 1, \ldots, t_i. \tag{1}$$
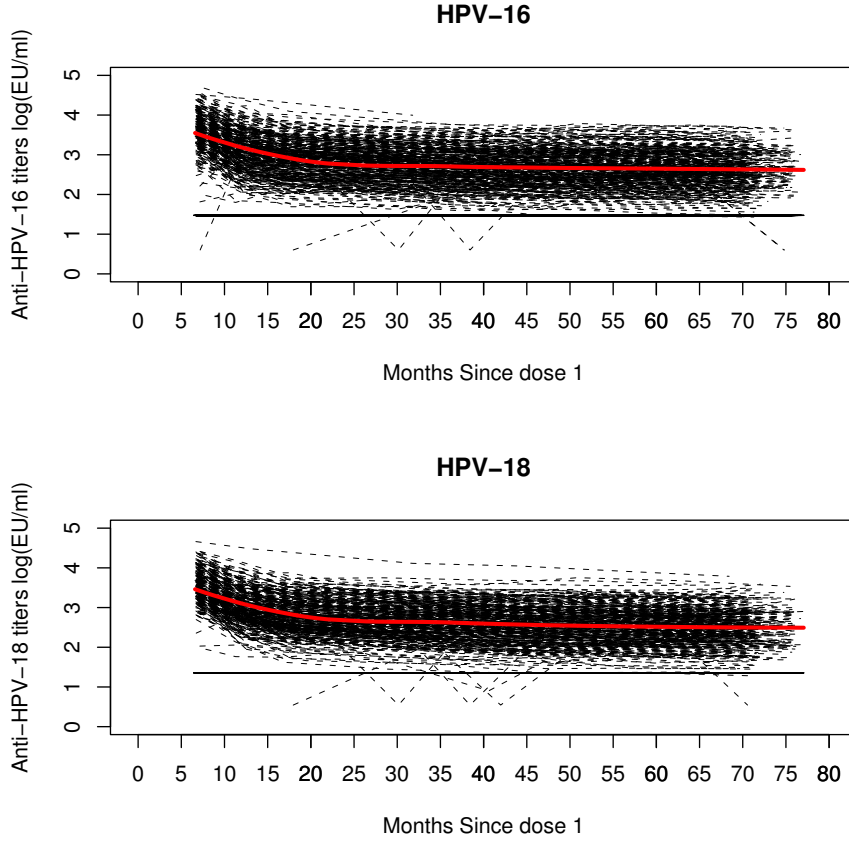
4

Figure 2: *Individual profiles of HPV-16 (top panel) and HPV-18 (bottom panel). The horizontal solid line indicates the natural infection level and the thick line shows the loess estimate for the mean antibody.*

Here, Y_{ij} is the log antibody level of subject i at time j, $f(t_{ij}) = k_i - a_i \log(c + t_{ij})$ is the mean log antibody titer at time t_{ij} post vaccination, ε_{ij} is the random error term for subject i at time j, assumed to be normally distributed, $\varepsilon_i \sim N(\mathbf{0}, W_i)$. The parameter k_i is the peak log level, a_i is the decay rate, c is an arbitrary small constant (often set to zero), n is the number of subjects, and t_i is the number of time points for subject i. Fraser *et al.* (2007) fitted a mixed model and they assumed the random effects k_i and a_i are drawn from a bivariate normal distribution. The linear mixed model representation of the power law model is:

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \boldsymbol{b}_i + \boldsymbol{\varepsilon}_i, \tag{2}$$

for subject i, where $\mathbf{Y}_i$ is the vector of observed random variables, $\mathbf{X}_i$ and $\mathbf{Z}_i$ are known design matrices with dimensions of $t_i \times p$ and $t_i \times q$ respectively, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters, which are called
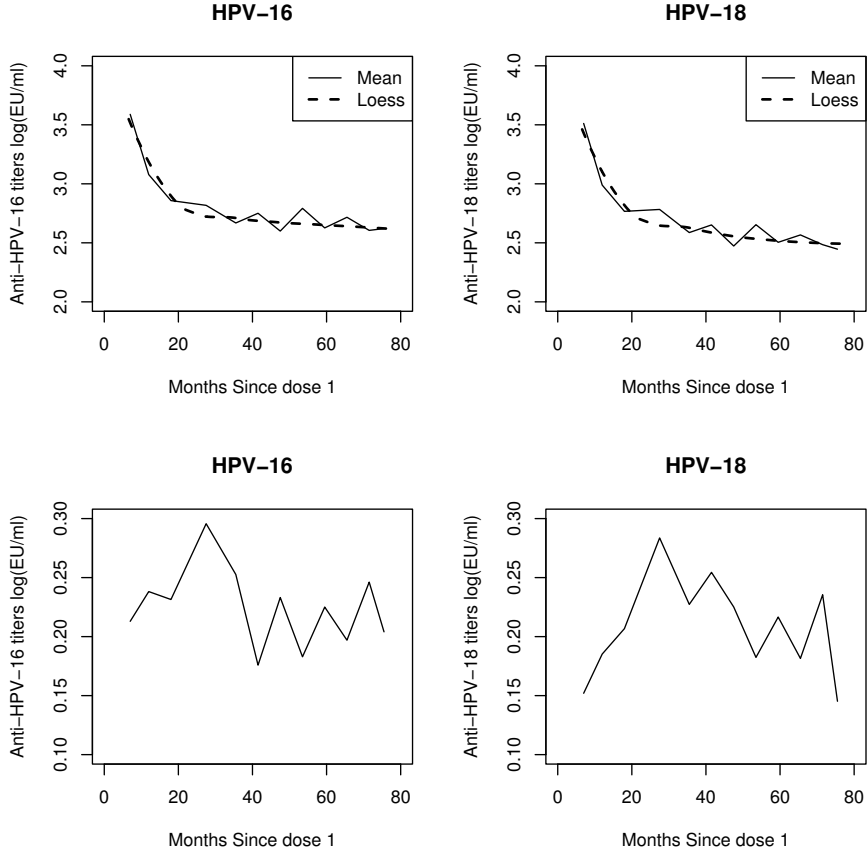
5

Figure 3: *Mean (top panel) and variance (bottom panel) structure of HPV-16 and HPV-18.*

the fixed effects, $\boldsymbol{b}_i$ is a $q \times 1$ vector of random effects, and ε_i is a $n \times 1$ vector of unobserved measurement errors. For the random effects and the random error, we assume

$$\left[\begin{array}{c} \boldsymbol{b}_i \\ \varepsilon_i \end{array} \right] \sim \left(\left[\begin{array}{c} 0 \\ 0 \end{array} \right], \left[\begin{array}{cc} D & 0 \\ 0 & W_i \end{array} \right] \right).$$

Hence, the mean structure for the power-law model with random intercept and slope is given by:

$$f(t_{ij}) = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i}) \log(t_{ij}).$$

Here, b_{0i} and b_{1i} are subject-specific intercepts and slopes, respectively. Fraser *et al.* (2007) extended this model to account for two populations of B-cells, including activated and memory B-cells, which impose a long-term antibody plateau. Their model takes the form:

$$f(t_{ij}) = \{k_i + \log[(1 - \pi)(c + t_{ij})^{-a_i} + \pi]\},$$

6

where π is the relative level of antibody produced in the long term memory plateau (between 0 and 1). A value of $\pi > 0$ indicates long-term antibody persistence. Note that for $\pi = 0$, the modified power law model is reduced to standard power law model (1). The modified power law model proposed by Fraser *et al.* (2007) imposes an asymptote for the antibody levels at $t_{ij} \to \infty$ and, as a result, the expected value of individual antibody level reaches a constant value in the long run.

In addition to the power law and modified power law models, David *et al.* (2009) proposed a subject-specific piece-wise constant model with mean structure:

$$
f(t_{ij}) = \begin{cases} \beta_0 + \beta_1 t_{ij} + b_{0i} & \text{month } 7 \leq t_{ij} < \text{ month } 12, \\ \beta_0 + \beta_1 t_{ij} + \beta_2(t_{ij} - 12) + b_{0i} & \text{month } 12 \leq t_{ij} < \text{ month } 21, \\ \beta_0 + \beta_1 t_{ij} + \beta_2(t_{ij} - 12) + \beta_3(t_{ij} - 21) + b_{0i} & \text{if } t_{ij} \geq \text{month } 21. \end{cases}
$$

They select the three break points, months 7, 12, and 21, using Akaike's Information Criterion (AIC, Pan *et al.*, 2001). The main difficulty of the modified model proposed by Fraser *et al.* (2007) and the piece-wise model proposed by David *et al.* (2009) is that the first assumes *a priori* that the antibody levels will reach an asymptote in the long run while the later assumes that the rate of decline is constant within a specified period. In what follows, we generalized the power law model and use the fractional polynomial family of models as a flexible, and data driven, modeling approach to estimate subject-specific models for the antibody decline. Moreover, we included a serial correlation function to account for the possible dependency of the correlation between the different time points.

3.2   Modeling Mean Antibody Using Subject-specific Fractional Polynomials

Fractional polynomials (FP; Royston and Altman, 1994; Royston and Sauerbrei, 2008) are proposed as a flexible parametric approach to describe the dependency between a response of primary interest and a covariate. In our example, the response of primary interest is the log-transformed antibodies and the covariate is time. The mean structure of a fractional polynomial model can be formulated in the following way:

$$
\sum_{l=0}^{m} \beta_l H_l(t_{ij}) + \sum_{l=0}^{m} b_{li} H_l(t_{ij}), \tag{3}
$$

7

where m is an integer, $p_1 \leq p_2 \leq \cdots \leq p_m$ is a sequence of powers and $H_l(a)$ is a transformation function given by

$$H_l(t_{ij}) = \begin{cases} t_{ij}^{p_l} & \text{if } p_l \neq p_{l-1}, \\ H_{l-1}(t_{ij}) \cdot \log(t_{ij}) & \text{if } p_l = p_{l-1}, \end{cases} \tag{4}$$

with $p_0 = 0$ and $H_0 = 1$. To take subject heterogeneity into account, we assume two components in the mean structure of (3). The first consists of the fixed parameters β_l and the latter the subject-specific parameters b_{li}. For the analysis presented in this paper, first-order FPs were used. The mean structure for the first-order mixed fractional polynomial can be written as

$$f(t_{ij}) = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_{ij}^p. \tag{5}$$

Note that, for $p = 0$ FP model (5) reduces to power law model (1) (Box and Cox, 1964).

In addition, higher-order FP models for $m > 1$ can be considered as well. Royston and Sauerbrei (2008) argued that the second order FP ($m = 2$) provides enough flexibility for modeling many of the types of continuous functions that we encounter in the health science and elsewhere. Therefore, in this study, we used $m \leq 2$. The closed testing procedure for function selection (FSP), which was discussed by Royston and Sauerbrei (2008), was implemented for model selection from among FP1, FP2, and the linear mixed model.

Mixed model (2) assumes that the covariance matrix W_i is equal to $\sigma^2 I_{n_i}$, where I_{n_i} denotes the identity matrix of dimension n_i. The model can be extended to include a serial correlation (Diggle *et al.*, 2002; Verbeke and Molenberghs, 2000). The resulting linear mixed model can be written as

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \boldsymbol{b}_i + \boldsymbol{\varepsilon}_{(1)i} + \boldsymbol{\varepsilon}_{(2)i}.$$

The error term $\boldsymbol{\varepsilon}_{(2)i}$ is a component of serial correlation and $\boldsymbol{\varepsilon}_{(1)i}$ is an extra component of measurement error, reflecting variation added by the measurement process itself (Verbeke and Molenberghs, 2000) and assumed to be independent of $\boldsymbol{\varepsilon}_{(2)i}$. Hence, $b_i \sim N(0, D)$, $\boldsymbol{\varepsilon}_{(1)i} \sim N(0, \sigma^2 I_{n_i})$, and $\boldsymbol{\varepsilon}_{(2)i} \sim N(0, \tau^2 H_i)$. The serial correlation matrix H_i depends on i through the number n_i of observations and through the time points t_{ij} at which measurements were taken. For an elaborate discussion regarding serial-correlation functions, we refer to Verbeke and Molenberghs (2000).

4 Long-term Prediction Using Subject-specific Fractional Polynomials

4.1 Model Building Over the Data Period

The FP models discussed above were used to estimate subject-specific evolutions for the log antibody within the follow-up period of the study (the estimation period). In the first stage, we selected an appropriate model for the serial correlation process, as well as the FP model for the mean structure. To this effect, four models for the serial correlation process were considered: (1) a model without a serial correlation process, (2) a local exponential model, (3) a Gaussian serial correlation, and (4) an exponential serial correlation. Next, to select the power of the FP1 model, powers in the range $\{-3, -2.75, -2.5, \ldots, 2.5, 2.75, 3\}$ were considered. Figure 4 shows the AIC values *versus* the power for the first-order fractional polynomials. For both HPV-16 and HPV-18, the model with the smallest AIC was obtained for $p = -1.25$.

For HPV-16, the best serial correlation model is the local exponential model (AIC: 264.9) while for HPV-18 the Gaussian serial correlation model fits best (AIC: -197.9). Note that we reached almost the same AIC value (approximately 265) for $p = -1.5$ and $p = -1.25$; that was for HPV-16 using the local exponential model. However, the estimate of d_{22} (the variance of the random slope) and its standard error obtained from p=-1.5 was equal to 13.9 (s.e. 5.6) which is larger than the result obtained with $p = -1.25$; 5.4 (s.e. 2.1). Furthermore, using the simple covariance structure, the AIC of $p = -1.25$ (433.6) was smaller than the AIC of $p = -1.5$ (438.9). Hence, we decided to use $p = -1.25$. Similarly, second-order FP were fitted and for both HPV-16 and HPV-18, the model with smallest AIC obtained for p; (-1.25;-1.25). Note that the optimal serial correlation for FP2 was the same as for FP1 above. Next, the aforementioned FSP was applied to select the best model from among the linear mixed model, FP1, and FP2. The result implies that the first-order FP, reported above, is to preferred (Appendix A; Table A.1).

The AIC values for the power-law model, piece-wise, FP1, and modified power-law model are shown in Table 1. It is clearly seen from the table that the AIC of the FP is smallest for both HPV-16 and HPV-18. Hence, the fractional polynomial model fits the data best among these four models. Table 2 shows the parameter estimates and their associated standard errors, based on the FP1 model. Observed and predicted
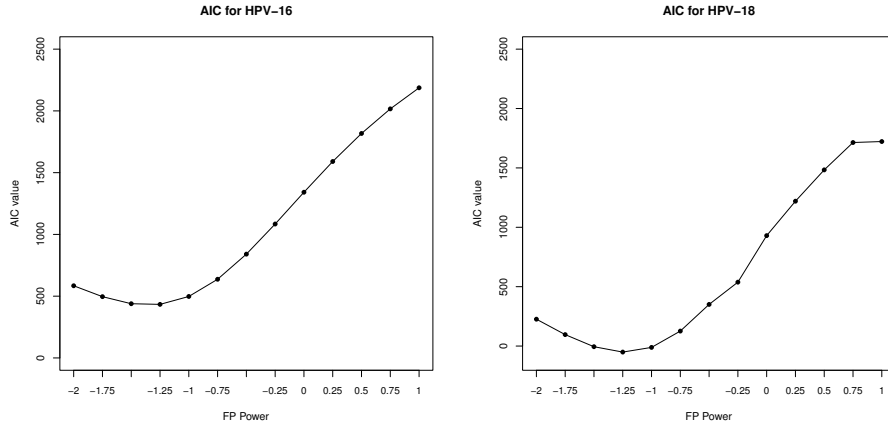
9

Figure 4: *Selection of the powers for the first order fractional polynomial using AIC. The circular points in the plots indicate the AIC values for the pre-specified powers.*

Table 1: *AIC value for power Law (PL), piece-wise linear (PW), fractional polynomial (FP), and modified power-law (MPL) models, for HPV-16 and HPV-18.*

Model	HPV-16	HPV-18
FP	264.9	−197.9
MPL	517.1	61.8
PL	1341.5	929.8
PW	662.5	264.5

means for the FPs and the power law models are shown in Figure 5. The fractional polynomial returned a predicted mean more similar to the observed mean than the power law model. The 95% pointwise and bootstrap confidence bands for the FP1 are displayed in Figure 5. Comparing these two methods, the second method results in estimates more similar to the observed values. Moreover, as expected, it leads to wider confidence intervals than the first method.

To evaluate the prediction performance of the FP model, individual predicted profiles were obtained for the time points within the estimation period: [M57–M62], [M63–M68], [M69–M74], and [M75–. . .] months. For each one of these categorized months in the estimation period, the density of the predicted values and the observed value for HVP-16 and HVP-18 are shown in Figures 6 and 7, respectively. From the figures, we

Table 2: *Model parameter estimates and their associated standard error using FP1. The p-values are based on the Wald test.*

Parameter	Estimate (s.e.)	p-value
HPV-16		
β_0	2.582 (0.024)	<0.001
β_1	11.539 (0.242)	<0.001
HPV-18		
β_0	2.468 (0.024)	<0.001
β_1	11.985 (0.208)	<0.001

can clearly see that the observed values and the model-based predicted values in these months are similar. This indicates that the model is performing well in terms of prediction at time points within the estimation period.

However, in this approach, both the estimation and the evaluation are based on the same data. Hence, a cross-validation approach has been applied to avoid overly optimistic conclusions. First, we considered the data up to [M69–M74] month for model estimation and then the performance of the FP model was evaluated by plotting the density of the observed values against the density of the predicted values at [M75–...] months (Figure A.1). From this figure, it can be seen that both densities are similar, indicating that the model is performing well in terms of prediction. Second, the data up to [M45–M50] months were used for estimation, and then the model-based and observed densities were compared at [M51–M56], [M57–M62], [M63–M68], [M69–M74], and [M75–...] months. The results showed that both densities are similar (details not shown). Hence, it is reasonable to state that the FP model gives good prediction.

4.2 Model-based Long-term Prediction for 30 and 50 Years

In practice, it is impracticable and costly to follow up subjects over the long term, such as 30 or even 50 years. Hence, to save resources, it is unavoidable to predict the persistence of the anti-HPV responses induced by the vaccine outside the range of the data. In this section, we discuss the subject-specific long-term
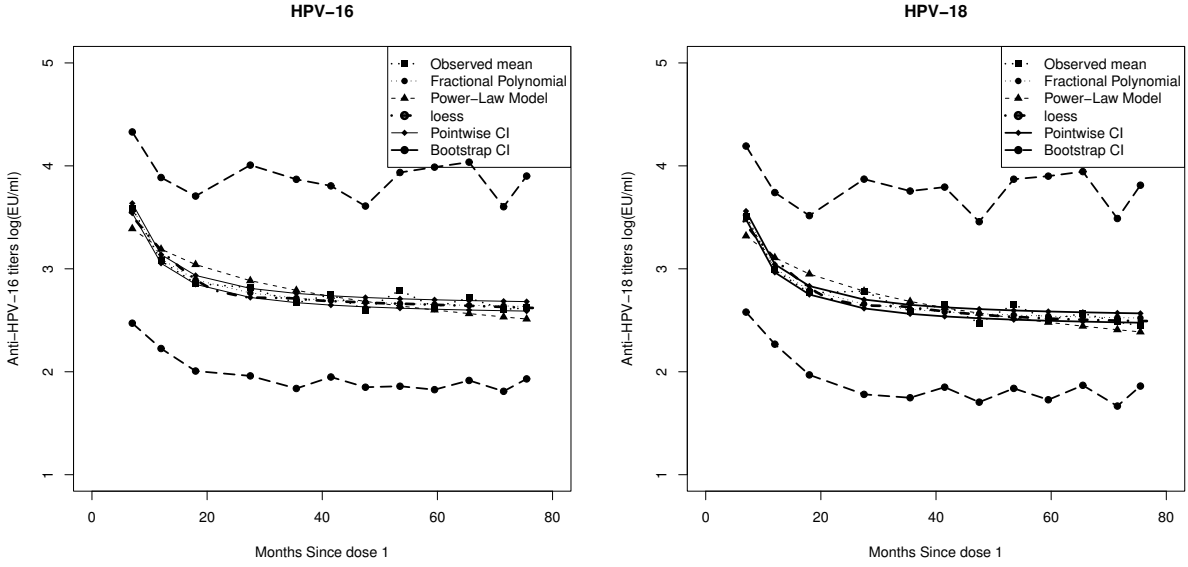
Figure 5: *Predicted means* versus *observed means with 95% pointwise and bootstrap confidence bands using FP1 for HPV-16 and HPV-18.*

prediction for the antibody level. Using the maximum likelihood parameter estimates for the fixed effects and the empirical Bayes estimates for the random effects, we can use the model estimated within the estimation period to predict the antibody level for 30 and 50 years. Figure 8 shows the long-term predicted means for 50 years for the population mean with 95% pointwise and bootstrap confidence bands, while Figures 9 and 10 exhibit an example of individual predictions. From Figure 8 we can clearly see that on average, the antibody level of the population is significantly above the natural infection level for 50 years. In addition, over the long term prediction, the 95% bootstrap confidence bands result in a pattern similar to the observed mean, whereas the 95% pointwise confidence bands remain constant.

Let τ be a threshold value such that

$$
\begin{cases}
\hat{y}_{it} > \tau & \text{above threshold,} \\
\hat{y}_{it} \leq \tau & \text{below threshold.}
\end{cases}
$$

Our aim is to estimate the proportion of individuals in the study who are above the threshold value for a given period. Two different threshold values (τ) were used. For HPV-16, $\tau = 1.474$ log (EU/ml) and 2.621 log (EU/ml) and for HPV-18, $\tau = 1.355$ log (EU/ml), and 2.446 log (EU/ml). The threshold values 1.474 and 1.355 are the natural infection level (David *et al.*, 2009) while the threshold values of $\tau = 2.621$ and
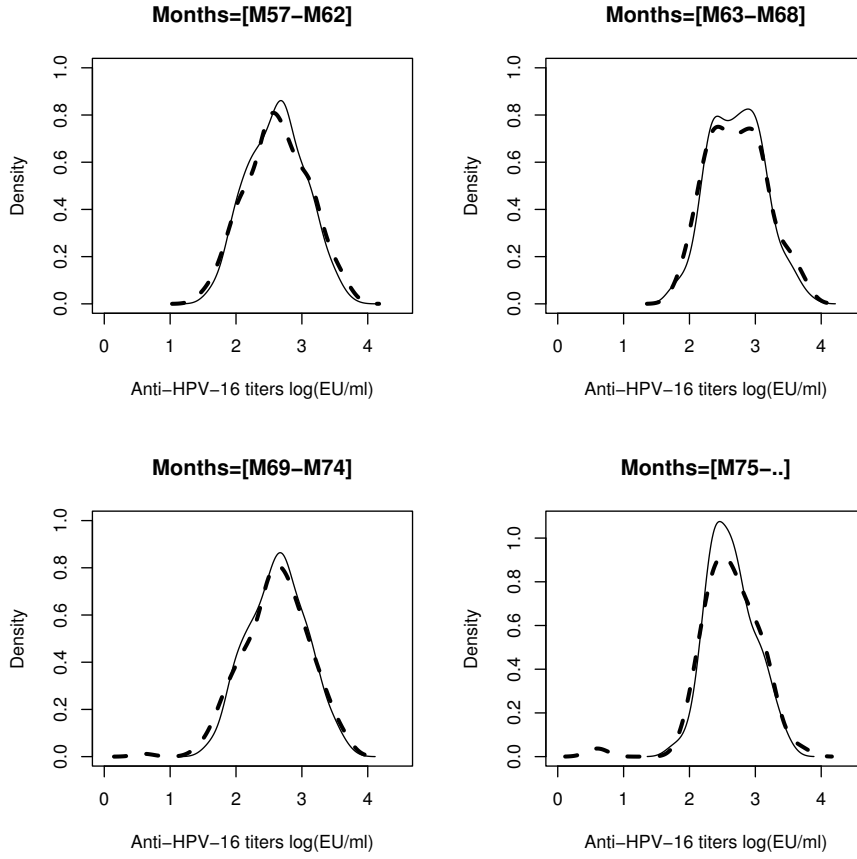
Figure 6: *The densities of observed values (dashed line) and model-based predictions (solid line) at categories [M57–M62], [M63–M68], [M69–M74], and [M75–. . .] for HPV-16. using FP1*

$\tau = 2.446$ are the mean of log anti-HPV-16 titers and log anti-HPV-18 titers at the end of the study period (Fraser *et al.*, 2007).

Figure 9 shows an example of one subject that crosses the threshold $\tau = 1.474$ for HPV-16 and two subjects that cross the threshold $\tau = 1.355$ for HPV-18 before 50 years, while Figure 10 shows an example of two subjects who remain above the threshold for a period of 50 years.

For each subject in the study, the time to cross a given threshold value, t_τ, can be calculated from the predicted serological result. Figure 11 shows the density estimate for the model-based antibody level and the two thresholds mentioned above. We note that for lower thresholds (1.474 and 1.355 for HPV-16 and HPV-18, respectively), the proportion of subjects below the threshold is 0.3% (only one subject) for HPV-16
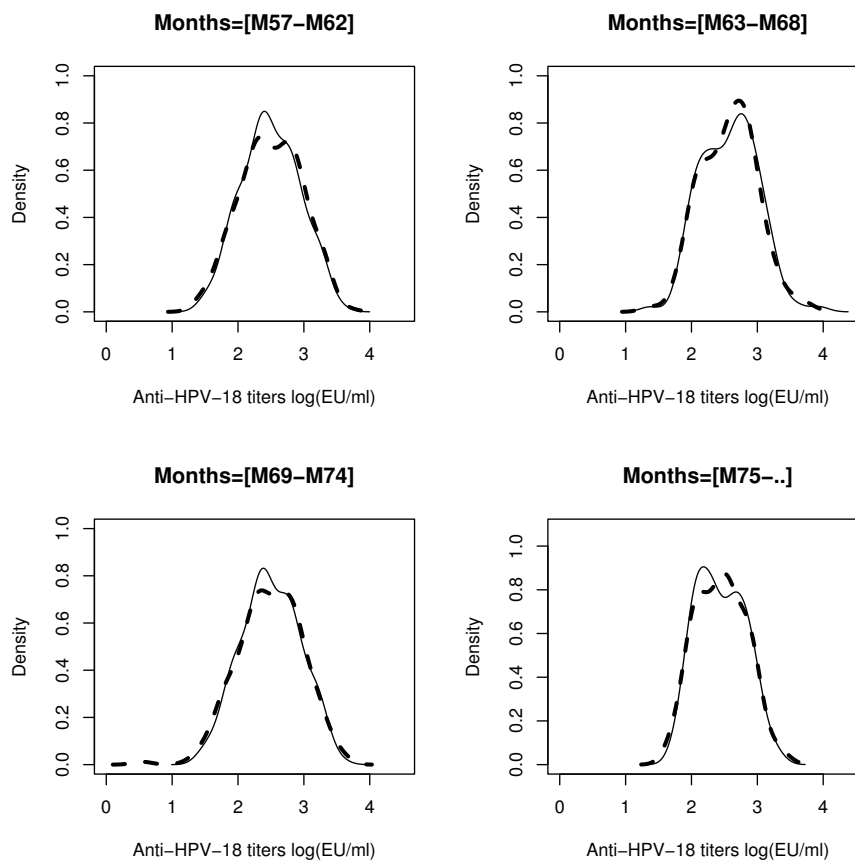
Figure 7: *The densities of observed values (dashed line) and model-based predictions (solid line) at categories [M57–M62], [M63–M68], [M69–M74], and [M75–...] for HPV-18 using FP1*

and 0.5% (only two subjects) for HPV-18. Table 3 shows the observed distribution of subjects above the threshold over the estimation period, while Tables 4 and 5 show the model-based distribution of subjects for different thresholds within the follow-up period and after the vaccination period, respectively. When we use $\tau = 1.474$ for HPV-16 and $\tau = 1.355$ for HPV-18, it can be seen from the tables that the model-based proportion is similar to the observed proportion for all months. However, if we use $\tau = 2.621$ for HPV-16 and $\tau = 2.446$ for HPV-18, there is a small discrepancy at months 12 and [M75–...] because the model-based proportion is not as close as the observed proportion. Thus, considering the bias and the variability of the estimated FP model coefficients, a bootstrap method was applied to estimate the proportion and the percentile confidence interval. Because the HPV-16/18 data set is longitudinal, we used a cluster bootstrap method in which clusters (or individuals) are selected by simple random sampling with replacement and there
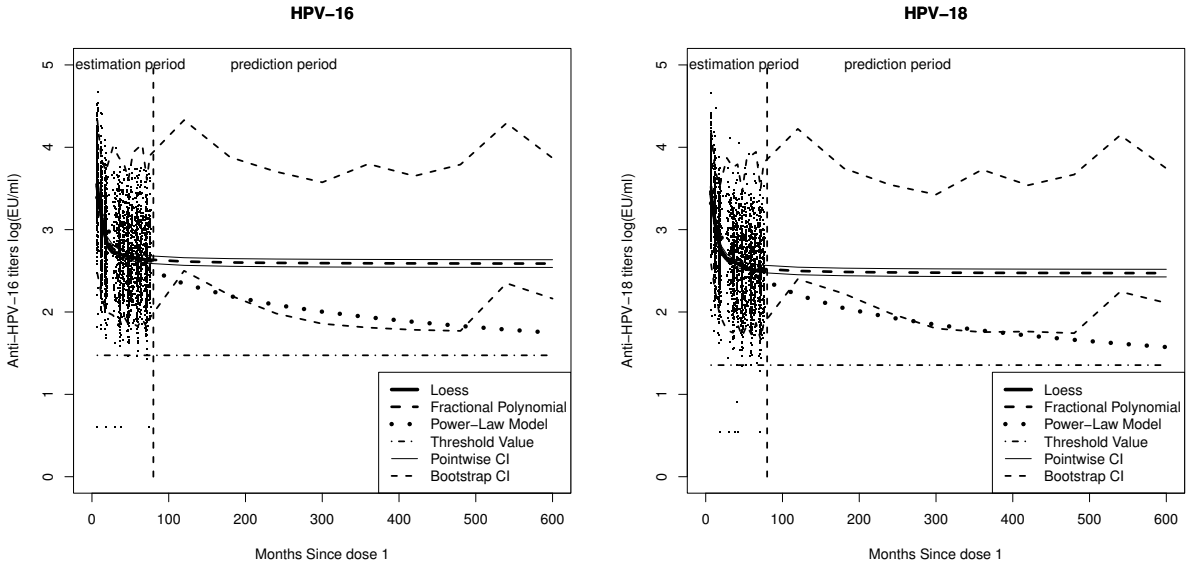
Figure 8: *Long-term prediction (50 years) with 95% pointwise and bootstrap confidence band using FP1 for HPV-16 and HPV-18.*

is no subsequent permutation (Field and Welsh, 2007). The results are shown in Appendix B (Table A.2). It can be clearly seen that the bootstrap method returned a slightly closer estimate to the observed proportion than the classical approach.

Figure 12 shows the distribution of the time to cross a given threshold in the two studies, respectively. Two subjects crossed the natural infection level for HPV-18, while one subject did so for HPV-16; for subject 8650, the time to cross the natural infection level was 260 months for HPV-16 while it was 131 months for HPV-18. For subject 8579, the time to cross the natural infection level was 76 months for HPV-18 while it was more than 600 months for HPV-16. Note that the dots below the threshold in Figure 8 are the antibody level for different subjects. The model is not flexible enough to predict accurately these antibody values. To assess the impact of these subjects, they were removed from analysis and parameter estimates and predictions re-calculated. Changes were, fortunately, very minor. Based on these observations, it seems best to keep these values in the data analysis, so as neither to lose information nor to make arbitrary decisions.

Figure 13 shows the proportion of subjects above the threshold over the estimation period and for a period of 50 years; proportions are calculated based on model-based prediction for antibody levels. It is clearly
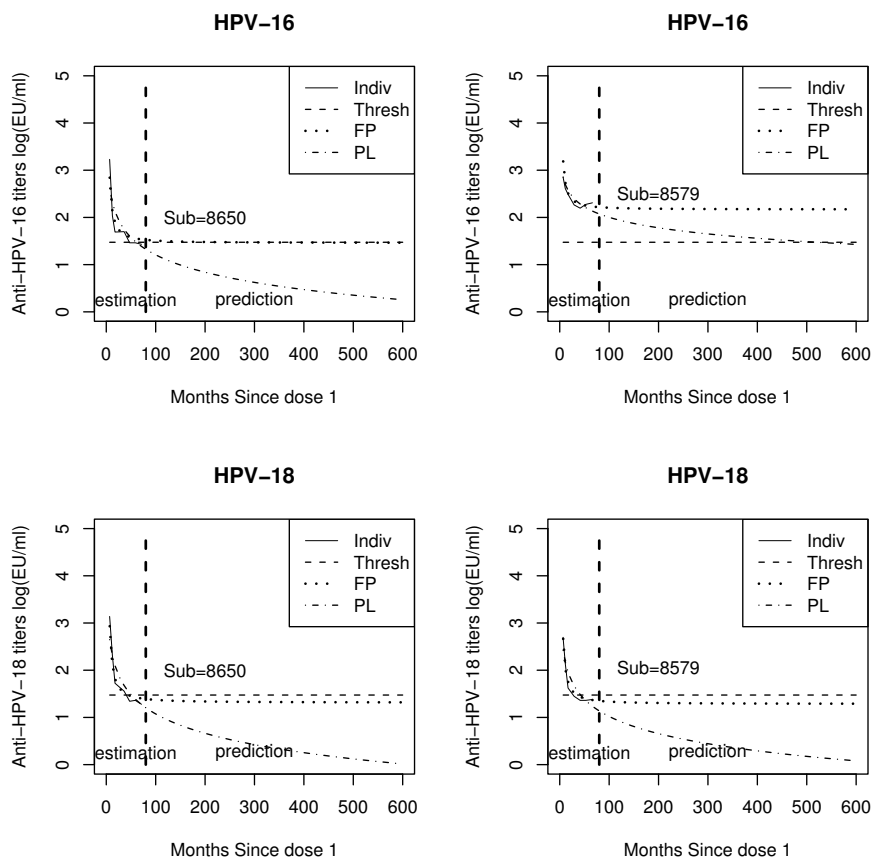
Figure 9: *Long-term prediction (50 years) for selected subjects 8650 and 8579 for HPV-16 (top panel) and HPV-18 (bottom panel). The dots (.) in the figure over the estimation period indicate individual profiles*

seen from the figure that the FP estimated the proportion more similar to the observed proportion than the power-law models. This implies that the FP is more preferable than the power-law model in this study. For $\tau = 2.621$, 47.4% [C.I.: (42.6,52.4)] and for $\tau = 2.446$, 52.5% [C.I.: (47.6,57.5)] are above the threshold during 50 years for HPV-16 and HPV-18, respectively.

The SAS software was used to fit the models. Code is presented in Appendix C.

5   Concluding Remarks

In this article, a fractional polynomial mixed model was applied to predict the proportion of subjects above a given threshold value, to obtain the estimated time points where the individual titers are below the
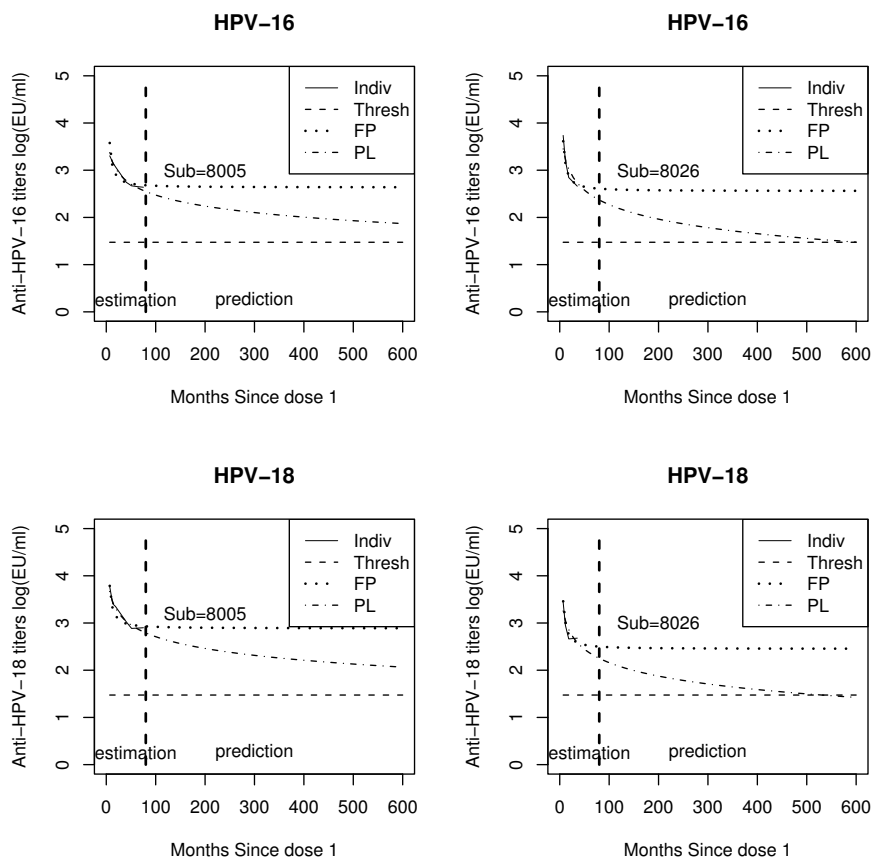
Figure 10: *Long-term prediction (50 years) for selected subjects 8005 and 8026 for HPV-16 (top panel) and HVP-18 (bottom panel).*

threshold value and to predict the long-term protection of vaccine which is given against HPV-16 and HPV-18 antigens. Generally, we found that the predicted mean anti-HPV-16 and anti-HPV-18 titers were above those associated with natural infection for at least 50 years when given to women aged 15–25 years.

It should be pointed out that, according to the fractional polynomial model, out of the 393 subjects, only one subject for HPV-16 and two subjects for HPV-18 have predicted values below the threshold value associated with natural infection. 99.7% and 99.5% of the subjects have a predicted serological result above the threshold value for life-time for HPV-16 and HPV-18, respectively. Nevertheless, when we use another threshold value other than the natural infection level, $\tau = 2.621$ for HPV-16 and $\tau = 2.446$ for HPV-18, around 47.4% and 52.6% of subjects are above the threshold for at least 50 years for HPV-16 and HPV-18, respectively.
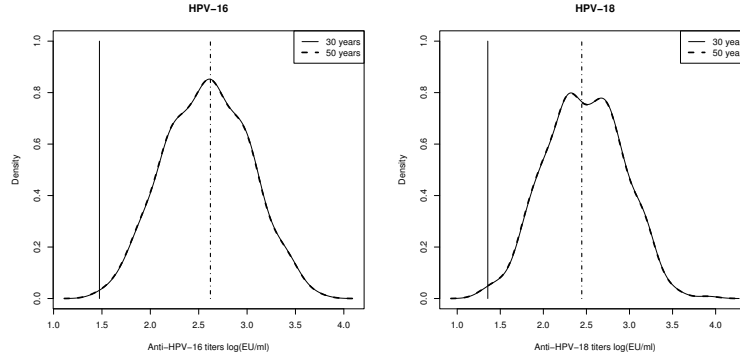
Figure 11: *The densities of model-based predictions of the antibody level for 30 and 50 years and for HPV-16 and HPV-18 using FP1. The solid and bold dashed lines indicate the thresholds 1.474 and 2.621 (left panel) and 1.355 and 2.446 (right), respectively.*

The results obtained for the estimated time points where the individual titers are below those associated with natural infection level is greater than 50 years for all subjects. However, for one subject, the estimated time point associated with natural infection is 260 month for HPV-16. There were two subjects for HPV-18 who have the estimated time points 76 and 131 months. Furthermore, if we take $\tau = 2.621$ and $\tau = 2.446$, the estimated time point is less than 120 month (10 years) for around 50.8% and 45.8% subjects for HPV-16 and HPV-18, respectively, while less than 360 month (30 years) for 52.5% and 47.2% of the subjects for HPV-16 and HPV-18, respectively.

To evaluate the predictive power of the FP model, a density of the observed value and the model based prediction was considered for different months. It could be claimed that the FP model gives good prediction within the estimation period. This result seems to imply that the model may be applied for long-term prediction, while at the same time the clinical relevance of long-term persistence of HPV-16 and HPV-18 antibodies is being investigated in long-term follow-up studies.

Variability was studied by calculating the proportion of vaccines with anti-HPV responses above the natural infection, and vaccine induced level at the last month in this study, even though an anti-HPV level associated with protection has not been defined. Both FP and power law models described the proportions of antibody responses above these thresholds. However, the FP model provided proportion estimates more similar to those observed proportions than the power-law model (Figure 13). This indicates that the fractional model

Table 3: *Observed proportion of subjects with anti HPV-16 and anti HPV-18 responses greater than the thresholds. P indicates for the proportion*

| | response > 1.474 | | | response > 1.355 | | | response > 2.621 | | | response > 2.446 | | |
| | HPV-16 | | | HPV-18 | | | HPV-16 | | | HPV-18 | | |
Month	n	P	95% CI	n	P	95% CI	n	P	95% CI	n	P	95% CI
7	363	0.997	(0.983,1.00)	362	1.00	(0.987,1.00)	357	0.981	(0.96,0.992)	359	0.992	(0.975,0.998)
12	366	1.00	(0.988,1.00)	366	1.00	(0.988,1.00)	292	0.798	(0.754,0.834)	330	0.902	(0.867,0.928)
18	364	0.997	(0.983,1.00)	364	0.997	(0.983,1.00)	249	0.682	(0.633,0.728)	276	0.756	(0.709,0.798)
[M25–M32]	88	0.989	(0.933,1.00)	88	0.989	(0.933,1.00)	58	0.652	(0.548,0.743)	67	0.753	(0.653,0.831)
[M33–M38]	218	0.995	(0.972,1.00)	218	0.995	(0.972,1.00)	125	0.571	(0.505,0.635)	133	0.607	(0.541,0.669)
[M39–M44]	161	1.00	(0.972,1.00)	162	1.00	(0.972,1.00)	98	0.609	(0.532,0.681)	109	0.673	(0.597,0.741)
[M45–M50]	233	0.996	(0.974,1.00)	233	0.996	(0.974,1.00)	113	0.483	(0.42,0.547)	113	0.483	(0.42,0.547)
[M51–M56]	130	1.00	(0.966,1.00)	130	1.00	(0.966,1.00)	88	0.677	(0.592,0.751)	91	0.70	(0.616,0.772)
[M57–M62]	224	0.996	(0.973,1.00)	225	1.00	(0.979,1.00)	110	0.489	(0.446,0.576)	119	0.529	(0.464,0.593)
[M63–M68]	130	1.00	(0.966,1.00)	130	1.00	(0.966,1.00)	71	0.546	(0.461,0.629)	80	0.615	(0.529,0.695)
[M69–M74]	220	0.991	(0.966,0.999)	220	0.991	(0.966,0.999)	112	0.505	(0.439,0.569)	118	0.532	(0.466,0.596)
[M75–…]	66	0.985	(0.911,1.00)	66	1.00	(0.934,1.00)	34	0.515	(0.397,0.632)	34	0.515	(0.397,0.632)

is a better model in predicting anti-HPV-16 and anti-HPV-18 persistence in this study.

The results of long term prediction using an FP model corroborate the findings of previous work done on the same data set by David *et al.* (2009), up to 6.4 years with modified power law model. Both models contain long-term plateaus. The modified power law model assumes a decline of antibodies and it assumes also the memory B-cells will remain constant over long term, which may not be correct biologically. Hence, this model allows bias towards asymptotic constant in predicting antibodies levels over the long term. However, the FP model is very flexible (Royston and Altman, 1994) because of its data-driven nature and also its appropriate fitting property in the study at hand.

It should be noted that, even though the AIC values shows that the FP model best fitted the data within the follow-up period when compared to the power law model, piece-wise and modified power law model, this does not mean it best fits for the prediction as well. Therefore, we would like to point out that the persistence of the anti-HPV responses induced by these vaccines can only be ascertained empirically by long-term follow-up

Table 4: *Model-based proportion of subjects with anti HPV-16 and anti HPV-18 responses greater than the thresholds within the estimation period using FP1. P indicates the proportion*

| | response > 1.474 | | | response > 1.355 | | | response > 2.621 | | | response > 2.446 | | |
| | HPV-16 | | | HPV-18 | | | HPV-16 | | | HPV-18 | | |
Month	n	P	95% CI	n	P	95% CI	n	P	95% CI	n	P	95% CI
7	364	1.00	(0.988,1.00)	362	1.00	(0.987,1.00)	364	1.00	(0.988,1.00)	361	0.997	(0.983,1.00)
12	366	1.00	(0.987,1.00)	366	1.00	(0.988,1.00)	321	0.878	(0.839,0.907)	345	0.943	(0.914,0.963)
18	365	1.00	(0.987,1.00)	365	1.00	(0.987,1.00)	258	0.707	(0.658,0.751)	286	0.784	(0.738,0.823)
[M25–M32]	89	1.00	(0.950,1.00)	89	1.00	(0.950,1.00)	61	0.685	(0.583,0.773)	69	0.775	(0.678,0.850)
[M33–M38]	219	1.00	(0.979,1.00)	219	1.00	(0.979,1.00)	124	0.566	(0.500,0.630)	131	0.598	(0.532,0.661)
[M39–M44]	161	1.00	(0.972,1.00)	160	0.988	(0.953,0.999)	98	0.609	(0.532,0.681)	105	0.648	(0.572,0.718)
[M45–M50]	234	1.00	(0.981,1.00)	234	1.00	(0.981,1.00)	116	0.496	(0.441,0.568)	121	0.517	(0.453,0.580)
[M51–M56]	130	1.00	(0.966,1.00)	130	1.00	(0.966,1.00)	87	0.669	(0.584,0.744)	88	0.677	(0.592,0.751)
[M57–M62]	225	1.00	(0.979,1.00)	225	1.00	(0.979,1.00)	114	0.507	(0.442,0.571)	121	0.538	(0.473,0.602)
[M63–M68]	130	1.00	(0.966,1.00)	130	1.00	(0.966,1.00)	73	0.562	(0.476,0.644)	77	0.592	(0.506,0.673)
[M69–M74]	222	1.00	(0.979,0.999)	222	1.00	(0.979,0.999)	113	0.509	(0.444,0.574)	114	0.514	(0.448,0.579)
[M75–...]	66	1.00	(0.934,1.00)	66	1.00	(0.934,1.00)	29	0.439	(0.326,0.559)	31	0.469	(0.35,0.588)

analysis. Moreover, modeling using FP functions has some limitations. Perhaps the most important one is sufficient power to detect a non-linear function and possible sensitivity to extreme values at either end of the distribution of a covariate (Royston and Sauerbrei, 2008). Note that the FP1 model considered in this study has asymptotic behavior, which may result in overestimation of the log antibody level when t_{ij} approaches infinity. Hence, caution should be used when using such model for large values of t_{ij}.

In conclusion, according to the FP mixed model, vaccination of healthy women with AS04-adjuvanted HPV-16/18 vaccine results in antibody level above the natural infection level from 50 years to near-life long persistence for more than 99% of the subjects. The model we considered in this study is a fractional polynomial model with only time as a covariate. Therefore, it is natural to extend this study to an FP model, including other covariates deemed relevant. Finally, further work is needed to define antibody levels associated with protection.

Table 5: *Model-based prediction for the proportion of individuals with anti HPV-16 and anti HPV-18 greater than the threshold using FP1. P indicates the proportion*

| | response > 1.474 | | | response > 1.355 | | | response > 2.621 | | | response > 2.446 | | |
| | HPV-16 | | | HPV-18 | | | HPV-16 | | | HPV-18 | | |
Month	n	P	95% CI	n	P	95% CI	n	P	95%CI	n	P	95% CI
120	390	1.00	(0.988,1.00)	389	0.997	(0.984,1.00)	192	0.492	(0.458,0.557)	211	0.541	(0.491,0.589)
180	390	1.00	(0.988,1.00)	388	0.995	(0.980,0.999)	186	0.477	(0.43,0.53)	208	0.533	(0.484,0.582)
240	390	1.00	(0.988,1.00)	388	0.995	(0.980,0.999)	185	0.474	(0.426,0.524)	207	0.531	(0.481,0.579)
300	389	0.997	(0.984,1.00)	388	0.995	(0.980,0.999)	185	0.474	(0.426,0.524)	206	0.528	(0.479,0.578)
360	389	0.997	(0.984,1.00)	388	0.995	(0.980,0.999)	185	0.474	(0.426,0.524)	206	0.528	(0.479,0.578)
420	389	0.997	(0.984,1.00)	388	0.995	(0.980,0.999)	185	0.474	(0.426,0.524)	206	0.528	(0.479,0.578)
480	389	0.997	(0.984,1.00)	388	0.995	(0.980,0.999)	185	0.474	(0.426,0.524)	205	0.526	(0.476,0.575)
540	389	0.997	(0.984,1.00)	388	0.995	(0.980,0.999)	185	0.474	(0.426,0.524)	205	0.526	(0.476,0.575)

Acknowledgments

References

Baseman, J.G., and Koutsky, L.A. (2005). The epidemiology of human papillomavirus infections. *Journal of Clinical Virology*, **32**, 16-24.

Bovier, P.A., Bock, J., Farinelli, T., Glueck, R., and Herzog, C. (2002). Long-term immunogenicity of an inactivated virosome hepatitis A vaccine. *Journal of Medical Virology*, **68**, 489–493.

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, **26**, 211-252.
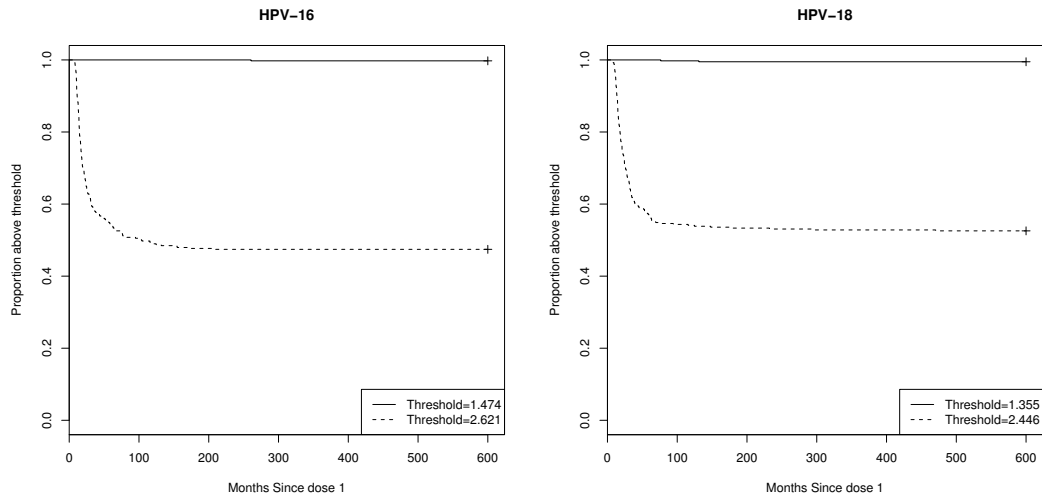
Figure 12: *Kaplan-Meier estimate for the distribution of the time to cross the threshold value for HPV-16 (left panel) and HPV-18 (right panel) using FP1 model.*

Carter, J.J., Koutsky, L.A., Wipf, G.C., Christensen, N.D., Lee, S.K., and Kuypers, J. (1996). The natural history of human papillomavirus type 16 capsid antibodies among a cohort of university women. *Journal of Infectious Diseases*, **174**, 927-936.

Daling, J.R., Madeleine. M.M., Schwartz, S.M., Shera, K.A., Carter, J.J., McKnight, B., *et al.*. (2002). A population-based study of squamous cell vaginal cancer: HPV and cofactors. *ogical Oncology*, **84**, 263-270.

David, M., Van Herck, K., Hardt, K., Tibaldi, F., Dubin, G., Descamps, D. and Van Damme, P. (2009). Long-term persistence of anti-HPV-16 and -18 antibodies induced by vaccination with the AS04-adjuvanted cervical cancer vaccine: Modeling of sustained antibody responses. *Gynecologic Oncology*, **115**, S1-S6.

Diggle, P.J., Heagerty, P.J., Liang, K.Y., and Zeger, S.L. (2002). *Analysis of Longitudina Data*. Oxford: Clarendon Press.

Field, C.A. and Welsh, A.H. (2007). Bootstrapping clustered data. *J.R.Statist.Soc.B /*, **69**, 369-390.

Fraser, C., Tomassini, J.E., Xi, L., Golm, G., Watson, M., and Giuliano, A.R. (2007). Modeling the long-term antibody response of a human papillomavirus (HPV) virus-like particle (VLP) type 16 prophylactic
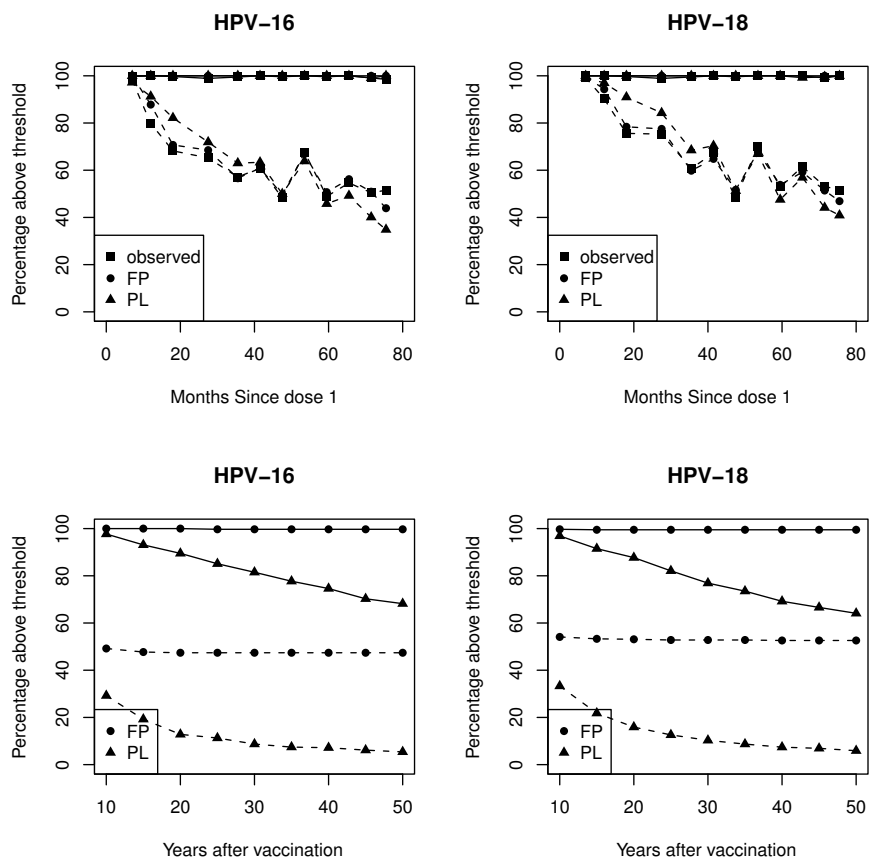
Figure 13: *Observed proportion and model-based proportion above different threshold values [τ = 1.474 and τ = 1.355 (solid line) and τ = 2.621 and τ = 2.446 (dashed line)] for HPV-16 (left) and HPV-18 (right), using the power-law model and FP models.*

vaccine. *Vaccine*, **25**, 4324-4333.

Gesemann, M. and Scheiermann, N. (1995). Quantification of hepatitis B vaccine-induced antibodies as a predictor of anti-HBs persistence. *Vaccine*, **13**, 443-447.

Glass, K. and Grenfell, B.T. (2004). Waning immunity and sub-clinical measles infections in England. *Vaccine*, **22**, 4110-4116.

Harper, D.M., Franco, E.L., Wheeler, C.M., Moscicki, A.B., Romanowski, B., Roteli-Martins, C.M., *et al.*. (2006). Sustained efficacy up to 4.5 years of a bivalent L1 virus-like particle vaccine against human papillomavirus types 16 and 18: follow-up from a randomised control trial. *Lancet*, **367**, 1247-1255.

Ho, G.Y., Studentsov, Y.Y., Bierman, R., and Burk, R.D. (2004). Natural history of human papillomavirus type 16 virus-like particle antibodies in young women. *Cancer Epidemiology and Biomarkers Prevention*, **13**, 110-116.

Honorati, M.C., Palareti, A., Dolzani, P., Busachi, C.A., Rizzoli, R., and Facchini A. (1999). A mathematical model predicting anti-hepatitis B virus surface antigen (HBs) decay after vaccination against hepatitis B. *Clinical Experimental Immunology*, **116**, 121-126.

Koutsky, L.A., Ault, K.A., Wheeler, C.M., Brown, D.R., Barr, E., and Alvarez, F.B. (2002). A controlled trial of a human papillomavirus type 16 vaccine. *New England Journal of Medicine*, **347**, 1645-1651.

Mossong, J., Nokes, D.J., Edmunds, W.J., Cox, M.J., Ratnam, S., and Muller, C.P. (1999). Modeling the impact of sub-clinical measles transmission in vaccinated populations with waning immunity. *American Journal of Epidemiology*, **150**, 1238-1249.

Muñoz, N., Bosch F.X., de Sanjose, S., Herrero, R., Castellsague, X., and Shah, K.V. (2003). Epidemiologic classification of human papillomavirus types associated with cervical cancer. *New England Journal of Medicine*, **348**, 518-527.

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. *Biometrics*, **57**, 120-125.

Royston, P. and Altman, D.G. (1994). Regression using fractional polynomials of continuous covariates: parsimonious parametric modeling, *Applied Statistics*, **43**, 429–467.

Royston, P. and Sauerbrei,W. (2008). *Multivariate Model Building; A pragmatic Approach to Regression Analysis Based on Fractional Polynomials for Modeling Continuous Variables.* New York: John Wiley & Sons.

Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data.* New York: Springer.

Zhou, J., Sun, X.Y., Stenzel, D.J., and Frazer, I.H. (1991). Expression of vaccinia recombinant HPV 16 L1 and L2 ORF proteins in epithelial cells is sufficient for assembly of HPV virion-like particles. *Virology*, **185**, 251-257.

Table A.1: *Application of the Functional Selection Procedure to HPV-16 and HPV-18. The selected model is FP1.*

Model	Deviance	Power	Comparison	Dev. diff.	p-value
			HPV-16		
FP2	243.3	-1.25;-1.25	FP2 vs null	3217.0	<0.001
FP1	246.9	-1.25	FP2 vs linear	1931.7	<0.001
Linear	2175.1	1	FP2 vs FP1	3.6	0.1653
Null	3460.3	-			
			HPV-18		
FP2	-214.0	-1.25;-1.25	FP2 vs null	3685.7	<0.001
FP1	-213.9	-1.25	FP2 vs linear	1920.0	<0.001
Linear	1706.0	1	FP2 vs FP1	0.1	0.9512
Null	3471.7	-			

APPENDIX

A   Functional Selection Procedure

The Functional Selection Procedure (FSP), which was discussed by Royston and Sauerbrei (2008), was applied to select the most suitable predictor function for the data. Results are shown in Table A.1. First, the FP2 was tested against the null model, and found to be significant. Then it was compared to the linear mixed model, leading to significance as well. Finally, FP2 was tested against FP1 and the former was non-significant for both HPV-16 and HPV-18. Hence, the selected model was FP1.

B   Cross-validation and Bootstrapping

We applied cross-validation to evaluate the performance of the prediction. The result for [M75–...] months is shown in Figure A.1. Besides, to obtain a confidence interval of the proportion estimate, a cluster bootstrap method was implemented and the results are shown in Table A.2.
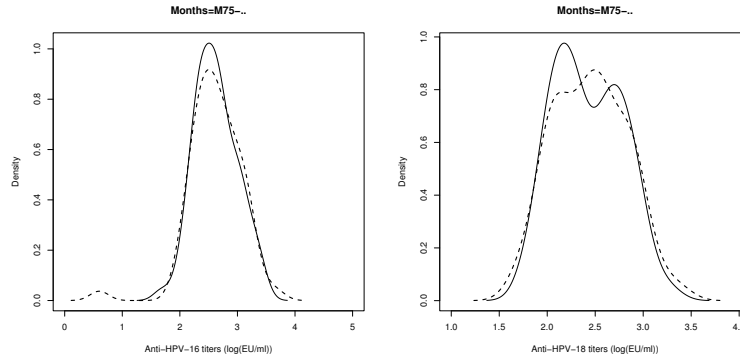
Figure A.1: *The densities of observed values (dashed line) and model-based predictions (solid line) using FP1 at [M75–...] month for HPV-16 (left panel) and HPV-18 (right panel), using cross-validation.*

C   SAS Implementation

A SAS program, using the procedure MIXED with local exponential serial correlation for the fractional polynomial (FP1) of HPV-16 is as follows:

```
* Local, exp(time) covariance structure;
%macro fp1(p1);
data fracpol; set hpv16; hx1=time**&p1; if &p1=0 then hx1=log(time);
varfp1=hx1; pbmthcls=time; run;
proc mixed data=fracpol method=ml covtest update info scoring=5 noclprint maxiter=200;* empirical;
  id pid log_val timing time;
 class pid pbmthcls;
 model log_val = varfp1 / s outp=p112a outpm=pm112a;
 random int varfp1/s sub=pid type=un;
 ods output SolutionR=Ebh16;
 repeated pbmthcls/  sub=pid type=sp(exp)(time) local=exp(time);
run;
```

The Gaussian serial correlation was the best covariance structure for HPV-18 and it was fitted as follows:

```
* Gaussian covariance structure;
%macro fp1(p1);
data fracpol; set hpv18; hx1=time**&p1; if &p1=0 then hx1=log(time);
varfp1=hx1;
```

26

Table A.2: *Proportion of subjects for FP1 model with anti HPV-16 and anti HPV-18 responses greater than the thresholds within the estimation period using bootstrap estimation; P indicates the proportion.*

| | response > 2.621 | | response > 2.446 | |
| | HPV-16 | | HPV-18 | |
Month	P	95% CI	P	95% CI
7	0.984	(0.967,0.997)	0.995	(0.986,1.00)
12	0.854	(0.816,0.893)	0.934	(0.908,0.957)
18	0.713	(0.664,0.759)	0.789	(0.744,0.829)
[M25–M32]	0.663	(0.560,0.759)	0.759	(0.654,0.849)
[M33–M38]	0.567	(0.505,0.624)	0.593	(0.528,0.652)
[M39–M44]	0.604	(0.529,0.679)	0.651	(0.576,0.723)
[M45–M50]	0.496	(0.436,0.558)	0.514	(0.450,0.573)
[M51–M56]	0.661	(0.571,0.744)	0.687	(0.605,0.765)
[M57–M62]	0.493	(0.431,0.557)	0.534	(0.468,0.595)
[M63–M68]	0.564	(0.479,0.648)	0.598	(0.504,0.680)
[M69–M74]	0.499	(0.437,0.563)	0.521	(0.458,0.586)
[M75–…]	0.443	(0.317,0.563)	0.473	(0.352,0.589)

```
pbmthcls=time; run;
 proc mixed data=fracpol method=ml covtest update info scoring=5 noclprint maxiter=200;
  id pid log_val timing time;
 class pid pbmthcls;
 model log_val = varfp1 / s outp=p112a outpm=pm112a;
 random int varfp1/s sub=pid type=un;
 ods output SolutionR=Ebh18;
 repeated pbmthcls/  sub=pid  type=sp(gau)(time) local  ;
 PARMS (0.1081) (-0.00232) (0.006719) (0.03579) (8.4780) (0.01957)/HOLD=0.006719,8.4780 NOBOUND ;
run;
```