MODELING CD4$^+$ T CELLS DYNAMICS IN HIV-INFECTED PATIENTS RECEIVING REPEATED CYCLES OF EXOGENOUS INTERLEUKIN 7

By Ana Jarne[*], Daniel Commenges[*], Laura Villain[*], Mélanie Prague[*], Yves Lévy[†,‡] and Rodolphe Thiébaut[*,‡] for inspire 2&3 study group

INSERM U1219-INRIA SISTM team[*], INSERM U955[†], VRI Paris[‡]

Combination antiretroviral therapy succeeds to control viral replication in most HIV infected patients. This is normally followed by a reconstitution of the CD4$^+$ T cells pool; however, this does not happen for all patients. For these patients, an immunotherapy based on injections of Interleukin 7 (IL-7) has been recently proposed in the hope of obtaining long-term reconstitution of the T cells pool. Several questions arise as to the long-term efficiency of this treatment and the best protocol to apply. Mathematical and statistical models can help answering these questions.

We develop a model based on a system of ordinary differential equations and a statistical model of variability and measurement. We can estimate key parameters of this model using the data from the main studies for this treatment, the INSPIRE, INSPIRE 2 & INSPIRE 3 trials. In all three studies, cycles of three injections have been administered; in the last two studies, for the first time, repeated cycles of IL-7 have been administered. Repeated measures of total CD4$^+$ T cells count in 128 patients as well as CD4$^+$Ki67$^+$ T cells count (the number of cells expressing the proliferation marker Ki67) in some of them were available. Our aim was to estimate the possibly different effects of successive injections in a cycle, to estimate the effect of repeated cycles and to assess different protocols.

The use of dynamical models together with our complex statistical approach allow us to analyze major biological questions. We found a strong effect of IL-7 injections on the proliferation rate; however, the effect of the third injection of the cycle appears to be much weaker than the first ones. Also, despite a slightly weaker effect of repeated cycles with respect to the initial one, our simulations show the ability of this treatment of maintaining adequate CD4$^+$ T cells count for years. We were also able to compare different protocols, showing that cycles of two injections should be sufficient in most cases.

1. Introduction. Infection by the Human Immunodeficiency Virus (HIV) leads to the decrease of the number of CD4$^+$ T-lymphocytes which induces a general immune dysfunction. Combination antiretroviral therapy (cART)

1

allows controlling viral load in most patients and often leads to an adequate immune restoration. However, not all patients get a satisfactory immune reconstitution despite undetectable viral load. A non-negligible proportion of patients experience an insufficient increase of $CD4^+$ T-lymphocytes, and can be called "low immunological responders".

A treatment based on injections of Interleukin-7 (IL-7) has been recently proposed to increase the population of $CD4^+$ T-lymphocytes, and is for the moment the only promising approach in this context (Sereti et al., 2009; Levy et al., 2009, 2012). Endogenous IL-7 is a cytokine discovered in 1988 (Namen et al., 1988); it has been found to play an important role in the maintenance of the T cells population. (Fry and Mackall, 2002; Mackall, Fry and Gress, 2011). Different effects of IL-7 have been uncovered and include enhancing production (thymopoiesis) (Mackall et al., 2001; Okamoto et al., 2002), proliferation (Vieira et al., 1998; Sportès et al., 2008) and survival (Seddon, Tomlinson and Zamoyska, 2003; Kondrack et al., 2003) of $CD4^+$ T cells.

Mathematical representations of the behavior of the immune system in the context of HIV infection have been useful to describe and quantify biological processes that are not directly observed; the interaction between HIV virions and $CD4^+$ T cells was firstly modeled by Ho et al. (1995) and Perelson et al. (1996). For modeling the effect of IL-7 administration, it is not useful to model virus concentration (because viral load is undetectable under cART), but it is necessary to distinguish between quiescent and proliferating cells, because we already know that the main effect of IL7 is through stimulating the proliferation of T cells (a proliferating cell divides and yields two cells). In this context, Thiébaut et al. (2014) have quantified the contribution of several biological mechanisms in $CD4^+$ T cells homeostasis. They have studied the effect of a single cycle of IL-7 administration, where a cycle included three consecutive injections. Here, we extend this approach with a modified statistical model for analyzing repeated cycles, based on data from 3 clinical studies, INSPIRE, INSPIRE 2 and INSPIRE 3. We focus on several major clinical questions. What is the effect of the different injections in a cycle? What is the effect of repeated cycles? What is the long-term efficacy of this therapy in maintaining $CD4^+$ T cells count at a satisfactory level (over 500 cells/μL)? What is the best protocol of injections?

This paper is divided into 6 sections. Section 2 gives an overview of the INSPIRE studies and the available data. Section 3 describes the main structure of the mathematical and statistical models. Section 4 presents and compares different statistical models: the "basic" model studying the effect of IL-7 over a cycle as a whole, the "3 β's" model allowing the successive

injections of a cycle to have different effects, and the "cycle effect model" investigating the long term effect when administering repeated cycles. Section 5 compares four possible protocols (varying the number of injections of a cycle) on their ability of maintaining CD4 counts over 500 and the average number of injection required. Section 6 concludes.

2. Data and materials. The data have been compiled from three studies: INSPIRE (Levy et al., 2012), INSPIRE 2 and INSPIRE 3 (Thiébaut et al., 2016). These studies investigated the effect of IL-7 treatment on immune restoration. All participants were aged ≥ 18 years, were under stable cART for at least one year, presenting CD4 counts between 100 and 400 cells/μL, and undetectable viral load for at least 6 months prior to screening.

In the first study, INSPIRE, 21 patients received three weekly injections IL-7 at different weight-dependent doses: 10, 20 and 30 μg/kg. INSPIRE 2 and INSPIRE 3 (with 23 and 84 treated patients, respectively) further studied the biological activity of *repeated cycles* of IL-7 at 20 μg/kg.

The two measurements that interest us are the total CD4$^+$ T cells count and the number of CD4$^+$ T cells expressing the Ki67 proliferation marker, hereafter called "CD4 count" and "Ki67 count", respectively. The Ki67 count measure the number of proliferating cells. The patients had clinic visits at weeks 1, 2 and 3 (at the moment of the injections), weeks 4, 5, 6, 9 and 12, and after, one visit every 3 months. Measurements of CD4 counts were made at each visit, while Ki67 counts were only measured at weeks 1, 2, 3, 5 and 12.

For INSPIRE 2 and 3, repeated cycles were available. After the first cycle, if CD4 counts were found to be below 550 cells/μL in one of the quarterly visits, a new IL-7 cycle was administered. Within these repeated cycles, clinic visits were scheduled at weeks 1, 2 and 3 (at the moment of the injections), weeks 5 and 12, and once again quarterly visits were made to check the CD4 count. A maximum of 4 cycles within 21 months and a maximum of 3 cycles within 12 months were established, and all patients have been followed up at least 3 months after the last cycle. CD4 counts were measured at all visits for all patients, while Ki67 counts were measured only for the first cycles of the first 12 patients of INSPIRE 2 at weeks 1, 2, 3, 5 and 12. The total duration of the studies was 12, 24 and 21 months for INSPIRE, INSPIRE 2 and INSPIRE 3, respectively.

In this paper, data for all patients from the three studies (N=128) have been included from the time of the first injection. Overall, 197 IL-7 cycles were administered (41 of them were incomplete cycles consisting of 1 or 2

injections). More details are provided in Appendix A and in Thiébaut et al. (2016).

3. Mathematical and statistical structure.

3.1. *Mathematical and statistical models.* A Markov jump process could be written for the dynamics of the quiescent and proliferating CD4 cells populations. When the number of cells is moderately large, a linear noise approximation leads to a stochastic differential equation Finkenstädt et al. (2013). In our case, the number of cells is very large: one can estimate the order of magnitude of the numbers of both types of cells to be larger than 100 millions. Thus the stochastic term is negligible which allows us to work with ordinary differential equations (ODE) for the concentration per volume unit (here, μL). We use the same system of ODE as proposed by Thiébaut et al. (2014). For patient i this model can be written as:

$$
\begin{cases}
\frac{dQ^i}{dt} & = \quad \lambda^i + 2\rho^i P^i - \pi^i Q^i - \mu_Q^i Q^i \\
\frac{dP^i}{dt} & = \quad \pi^i Q^i - \rho^i P^i - \mu_P^i P^i
\end{cases}
$$

The initial condition is assumed to be the equilibrium point (specified by $\frac{dQ^i}{dt}(0) = 0$, $\frac{dP^i}{dt}(0) = 0$).

A graphical representation of the system can be found in Figure 1. This model includes two state variables: P, the concentration of proliferating cells expressing the Ki67 proliferation marker and Q, the concentration of quiescent cells. We have also investigated a model with a feedback term, obtained by multiplying the basic proliferation rate by $\frac{1}{(P^i+Q^i)^\nu}$, where ν is a parameter to be estimated. We did not retain this feedback term because it did not lead to major improvement of the fit while requiring much more computation time (see Appendix B).

The vector of parameters of the ODE system is $\boldsymbol{\xi^i} = [\lambda^i, \rho^i, \pi^i, \mu_Q^i, \mu_P^i]$. These parameters have a biological interpretation: λ is the production rate, ρ is the reversion rate, π is the proliferation rate and μ_Q and μ_P are the mortality rates of Q and P cells, respectively. The logarithmic transformation ensures positivity of these biological parameters: $\boldsymbol{\tilde{\xi}^i} = \log(\boldsymbol{\xi^i})$.

Modeling the variability of the parameters is a crucial ingredient in our model because it allows to have a joint estimation of parameters across the population instead of fitting the model patient-by-patient. A mixed-effect model can be assumed for each transformed parameter l, $l = 1,...,p$ (here $p = 5$):

$$
\tilde{\xi}_l^i(t) = \tilde{\xi}_l^0 + \beta_l^\top z_l^i(t) + u_l^i,
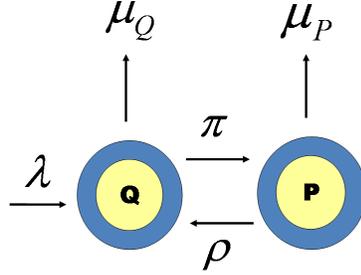$$

Fig 1: Graphical representation of the mathematical model

where $\tilde{\xi}_l^0$ is the baseline value, β_l is a vector of regression coefficients, $\boldsymbol{z_l^i}$ is a vector of explanatory variables, and u_l^i are random effects assumed to be independently and identically normally distributed. Thus, the parameters can vary between subjects, but also with time through the time-dependent explanatory variables. In practice, for parsimony, random effects and explanatory variables are included for a subset of the parameters.

In Section 4 of this paper, we present and discuss two models for the effect of injections in one cycle and a model for repeated cycles. The random effects have been applied on λ and ρ: $u_\lambda^i \sim \mathcal{N}(0, \sigma_\lambda^2)$, $u_\rho^i \sim \mathcal{N}(0, \sigma_\rho^2)$ for all the models. The explanatory variables used are functions of the dose and of the timing of the IL-7 injections and are used to model the proliferation rate (π) and the mortality rate of quiescent cells (μ_Q). These choices are based on many trials and on previous results of the literature (as in Thiébaut et al. (2014)).

We also need a model for the observations. The state variables $(P^i(t), Q^i(t))$ are not directly observable; we only have discrete-time observations of some functions of the components of this vector. Let Y_{1j}^i and Y_{2k}^i be the CD4 count and the Ki67 count for patient i at time t_{ij} and t_{ik}, respectively. The following observation scheme is assumed:

$$\begin{cases} (Y_{1j}^i)^{0.25} &= (P^i(t_{ij}) + Q^i(t_{ij}))^{0.25} &+ \epsilon_{1ij} \\ (Y_{2k}^i)^{0.25} &= P^i(t_{ik})^{0.25} &+ \epsilon_{2ik} \end{cases}$$

with independently normally distributed measurement errors: $\epsilon_{i1j} \sim \mathcal{N}(0, \sigma_{CD4}^2)$, $\epsilon_{i2k} \sim \mathcal{N}(0, \sigma_P^2)$. The fourth-root transformation for achieving approximate normality and homoscedasticity for cell counts has been studied by Thiébaut et al. (2003). Note that the times of observations may be different for the two observed components; indeed there were less observations of Ki67 counts

than of CD4 counts.

3.2. *Inference.* The vector θ to be estimated includes the intercepts of the biological parameters $(\tilde{\lambda}^0, \tilde{\rho}^0, \tilde{\pi}^0, \tilde{\mu}_Q^0, \tilde{\mu}_P^0)$, the regression coefficients $(\beta_\pi, \beta_{\mu_Q})$, the standard deviations of the random effects $(\sigma_\lambda, \sigma_\rho)$ and the standard deviations of the measurement errors (σ_{CD4}, σ_P). The two main approaches are the maximum likelihood and the Bayesian approaches; in both cases we have to compute the likelihood. As in Guedj, Thiébaut and Commenges (2007a), first the individual likelihoods given the random effects can be computed; then, the individual likelihoods are computed by integrating over the random effects via the adaptive Gaussian quadrature (Genz and Keister, 1996; Pinheiro and Bates, 2000); the global log-likelihood is the sum of the individual log-likelihoods. The parameters can then in principle be estimated by maximum likelihood. However, due to identifiability problems, it is useful to adopt a Bayesian approach. The prior distribution $\pi(\theta)$ allows incorporating prior knowledge taken from the literature. In such very complex models MCMC algorithm may fail, so we use an approximate Bayesian inference as in Drylewicz, Commenges and Thiebaut (2012), simpler than the INLA approach of Rue, Martino and Chopin (2009) which is also difficult to apply here. Bayes theorem gives

$$\log[P(\theta \mid Y)]) = L(\theta) + \log[\pi(\theta)] + C,$$

where $P(\theta \mid Y)$ is the posterior distribution, $L(\theta)$ is the log-likelihood and C is the normalization constant. The Bernstein-Von Mises theorem (Van der Vaart, 2000) justifies a normal approximation of the posterior (NAP). The NAP can be computed by maximizing the penalized log-likelihood $L^P(\theta) = L(\theta) + log[\pi(\theta)]$ and computing the inverse of the Hessian of $-L^P(\theta)$, $H_{L^P}^{-1}$. Thus, the NAP is $\mathcal{N}(\tilde{\theta}, H_{L^P}^{-1}(\tilde{\theta}))$, where $\tilde{\theta}$ is maximizes $L^P(\theta)$.

This computation can be achieved with the NIMROD program (Prague et al., 2013) which uses the so-called RVS algorithm (Commenges et al., 2006); parallel computing is implemented to achieve acceptable computation times. Other approaches have been proposed for fitting ODE-based models: Ramsay et al. (2007) proposed a penalized likelihood approach for the trajectories of the state variables circumventing the need of solving the ODE system; the original proposal did not treat models with random effects but it was extended by Wang et al. (2014) who used an approximation of the integrals; Kuhn and Lavielle (2005) have proposed the stochastic approximation expectation maximisation (SAEM) algorithm which can be used for maximising a log-likelihood or a penalized log-likelihood. Also, full Bayesian approach using the MCMC algorithm have been proposed in the context

of HIV modeling (Huang, Liu and Wu, 2006) and in that of chemical reaction networks (Finkenstädt et al., 2013). In the context of HIV modeling, Drylewicz, Commenges and Thiebaut (2012) made a comparison of the full Bayesian approach implemented in Winbugs and the penalized likelihood implemented in NIMROD and found results in favour of the latter. One advantage of the RVS algorithm is the possibility of computing a stopping criterion which can be interpreted as the ratio of numerical error over statistical error. See Appendix C for details.

3.3. *Comparison of different models.* Here, we present more than one possible statistical model to describe the effect of IL-7 on biological parameters. To compare them, apart direct likelihood comparison and individual fits, we use an approximate cross-validation criterion, LCVa, proposed by Commenges et al. (2007). LCVa is an extension of Akaike criterion (AIC), similar to the General Information Criterion (GIC) (Konishi and Kitagawa, 2008) that corrects not only for the number of parameters but also for the penalization; LCVa is normalized on the number of observations (see Commenges et al. (2008) and Commenges et al. (2015) for further developments). This criterion is:

$$\mathrm{LCVa} = -\,\mathrm{n}^{-1}[\,\mathrm{L}(\tilde{\theta}) - \mathrm{Trace}(\mathrm{H}_{L^P}^{-1}(\tilde{\theta})\mathrm{H}_L(\tilde{\theta}))],$$

where H_L is the Hessian of minus the log-likelihood. Since LCVa estimates a "risk" (cross-entropy or Kullback-Leibler risk equivalently), the smaller the better. Differences in criteria values between two models can be considered as "large" beyond 0.1 when the response is univariate. However, when the response is multivariate, the threshold for considering a difference as "large" should be higher, because LCVa, as defined here, is normalized on the number of subjects and does not take into account the number of observations per subject.

4. Main results.

4.1. *Basic model: A cycle as a whole entity.* Firstly, we are interested in estimating the global effect of the first cycle of IL-7. To begin with, only first received cycles for each patient have been considered. As in Thiébaut et al. (2014) the effect of IL-7 is considered to be dose-dependent. In our case, we have chosen to consider a power of the dose (as is common in pharmacology), that was fixed as 0.25 by profile likelihood.

The effect on proliferation π is taken into account during 7 days (this time was also fixed by profile likelihood) after each injection. Besides, the effect

on the mortality rate μ_Q is considered to be constant from two days after the first injection during twelve months, followed by a linear decrease during another twelve months. As already mentioned, random effects are added on the production rate λ and the reversion rate ρ. Let d_i the dose received for patient i, and let N_t^i the number of injections that patient i has received until time t. The statistical description for this first model is as follows:

$$\begin{cases} \tilde{\pi}^i(t) & = & \tilde{\pi}^0 & & + & \beta_\pi & d_i^{0.25} & \mathbb{1}_{\{N_t^i - N_{t-7}^i = 1\}} \\ \tilde{\lambda}^i(t) & = & \tilde{\lambda}^0 & + & u_\lambda^i & & & \\ \tilde{\mu}_Q^i(t) & = & \tilde{\mu}_Q^0 & & + & \beta_{\mu_Q} & d_i^{0.25} & f(t) \\ \tilde{\rho}^i(t) & = & \tilde{\rho}^0 & + & u_\rho^i & & & \\ \tilde{\mu}_P^i(t) & = & \tilde{\mu}_P^0 & & & & & \end{cases}$$

where $\mathbb{1}_{\{N_t^i - N_{t-7}^i = 1\}}$ is an indicator function taking value 1 if an injection has been administrated in the last 7 days, and

$$(1) \qquad f(t) = \begin{cases} 1 & if & 2 < t \leq 360 \\ 1 - (t - 360)/360 & if & 360 < t \leq 720 \\ 0 & if & 720 < t \end{cases}$$

Taking the same priors as Thiébaut et al. (2014), we ran the analysis with the NIMROD program. The results are displayed in Table 1: IL-7 injections increase the proliferation rate (π) from 0.041 per day at baseline to 0.135 per day during 7 days after each injection (for the dose equal to 20 $\mu g/kg$). Also the estimated mortality rate of Q cells decreases from 0.104 per day at baseline to 0.072 during the first year after the treatment.

4.2. *"3 β's" model: A cycle as three different injections.* Here we focus on a major question: have all the three injections the same quantitative effect on proliferation of CD4$^+$ T cells? Or, more accurately, what is the role of every single injection in the whole effect of a cycle? For this model too, we only consider the first received cycle for each patient. The statistical model for π was:

$$\tilde{\pi}^i(t) = \tilde{\pi}^0 + \sum_{k=1}^{3} \mathbb{1}_{\{N_t^i = k\}} \beta_{\pi_k} d_i^{0.25} \mathbb{1}_{\{N_t^i - N_{t-7}^i = 1\}}$$

The results are displayed in Table 2. This model is largely better than the "basic" model in terms of LCVa (2.136 vs 2.558). The quantitative effects of the successive injections are not equal: the first and second one are similar but the effect of the third one is considerably weaker.

Priors and estimated mean and standard deviation (sd) of all parameters (in logarithmic and natural scales) for the "basic" model when considering only the first cycle for all patients from INSPIRE 1, 2 & 3; Penalized (P) and Non Penalized (NP) likelihoods, and LCVa criteria

	PRIOR (log-scale)		POSTERIOR (log-scale)		POSTERIOR (natural-scale)	
	mean	sd	mean	sd	mean	sd
λ	1.000	1.000	2.967	0.062	**19.440**	*1.196*
ρ	0.000	0.250	0.680	0.095	**1.973**	*0.187*
π	-4.000	1.000	-3.185	0.115	**0.041**	*0.005*
μ_Q	-3.600	0.500	-2.264	0.073	**0.104**	*0.008*
μ_P	-2.500	0.500	-1.550	0.202	**0.212**	*0.043*
β_π					**0.997**	*0.058*
β_{μ_Q}					**-0.305**	*0.020*
σ_λ					**0.254**	*0.025*
σ_ρ					**0.534**	*0.096*
σ_{CD4}					**0.254**	*0.003*
σ_P					**0.299**	*0.023*
P likelihood			**-338.7**			
NP likelihood			**-327.4**			
LCVa			**2.558**			

4.3. *Cycle effect model: Effect of successive cycles.* Among the 128 treated patients from all the three studies, 74 have received more than one cycle. A key question is: have these repeated cycles the same quantitative effect with respect to initial ones? CD4 counts are higher before starting repeated cycles. Also, antibodies anti-IL-7 could appear after an initial cycle, modifying the effect of IL-7 when cycles are repeated. The second goal of this paper is to estimate possible quantitative differences in repeated versus initial cycles. To make this possible, we included data from all received cycles and we estimated a new fixed effect: the "cycle effect" β_C. We keep the notation t_{i1} for the time when patient i receives the first injection of a cycle. If C(t) counts the number of cycles received at time t, let $\mathbb{1}_{C(t)>1}$ be 1 if a cycle has been received before time t, 0 otherwise. The cycle effect is incorporated into the statistical model of proliferation rate as follows:

$$\tilde{\pi}^i(t) = \tilde{\pi}^0 + \left[\beta_C \mathbb{1}_{\{C(t)>1\}} + \sum_{k=1}^{3} \mathbb{1}_{\{N_t^i=k\}} \, \beta_{\pi_k} \, d_i^{0.25}\right] \mathbb{1}_{\{N_t^i - N_{t-7}^i = 1\}}$$

The results are displayed in Table 3. Checks of fit of this model appear satisfactory: see Appendix D. Appendix E shows some fits of real data from INSPIRE 2 and 3 obtained with this model. Individual predicted trajecto-

TABLE 2

Priors and estimated mean and standard deviation (sd) of all parameters (in logarithmic and natural scales) for the "3 β's" model when considering only the first cycle for all patients from INSPIRE 1, 2 & 3; Penalized (P) and Non Penalized (NP) likelihoods, and LCVa criteria

	PRIOR (log-scale)		POSTERIOR (log-scale)		POSTERIOR (natural-scale)	
	mean	sd	mean	sd	mean	sd
λ	1.000	1.000	2.355	0.087	**10.541**	*0.920*
ρ	0.000	0.250	0.635	0.102	**1.887**	*0.192*
π	-4.000	1.000	-3.306	0.125	**0.037**	*0.005*
μ_Q	-3.600	0.500	-2.617	0.080	**0.073**	*0.006*
μ_P	-2.500	0.500	-2.187	0.258	**0.112**	*0.029*
$\beta_{\pi 1}$					**1.155**	*0.079*
$\beta_{\pi 2}$					**1.120**	*0.081*
$\beta_{\pi 3}$					**0.622**	*0.073*
β_{μ_Q}					**-0.239**	*0.022*
σ_λ					**0.267**	*0.025*
σ_ρ					**0.575**	*0.108*
σ_{CD4}					**0.241**	*0.003*
σ_P					**0.305**	*0.025*
P likelihood			**-279.8**			
NP likelihood			**-273.3**			
LCVa			**2.136**			

ries were computed using the Parametric Empirical Bayes (PEB) for the parameters having a random effect (λ and ρ).

The posterior distribution of the cycle effect β_C has mean equal to -0.163 and standard deviation equal to 0.015. In other words, the cycle effect is found to be significantly negative. In natural scale, the effect on proliferation rate for successive cycles is found to be $e^{-0.163} = 0.85$ times the effect of the first cycle. The biological interpretation of the cycle effect is not yet clearly explained. One explanation may be that the first cycle has modified the reaction of the immune system to further injections; one possibility is that antibodies against IL-7 decrease the efficient concentration of IL-7 obtained at the target. However, we must take into consideration differences in mean CD4 count before the initial and repeated cycles. The mean CD4 count at baseline was 266 cells/μL whereas it was 456 cells/μL before repeated cycles. Considering the homeostatic regulation of the population of $CD4^+$ cells, that prevents CD4 counts from exceeding 1200-1300 cells/μL, a feedback mechanism may explain an apparent cycle effect. With the aim to deeper study this phenomenon, we have incorporated a feedback term (see Appendix B). We found that a feedback effect could indeed be detected, but this had no major influence on the estimate of the cycle effect.

Table 3

Priors and estimated mean and standard deviation (sd) of all parameters (in logarithmic and natural scales) for the "cycle effect" model when considering all cycles for all patients from INSPIRE 1, 2 & 3; Penalized (P) and Non Penalized (NP) likelihoods, and LCVa criteria

	PRIOR (log-scale)		POSTERIOR (log-scale)		POSTERIOR (natural-scale)	
	mean	sd	mean	sd	mean	sd
λ	1.000	1.000	1.672	0.061	**5.323**	*0.326*
ρ	0.000	0.250	0.892	0.093	**2.440**	*0.226*
π	-4.000	1.000	-2.853	0.074	**0.058**	*0.004*
μ_Q	-3.600	0.500	-2.610	0.068	**0.074**	*0.005*
μ_P	-2.500	0.500	-2.567	0.200	**0.077**	*0.015*
$\beta_{\pi 1}$					**0.931**	*0.042*
$\beta_{\pi 2}$					**0.707**	*0.043*
$\beta_{\pi 3}$					**0.229**	*0.042*
β_{μ_Q}					**-0.082**	*0.006*
β_C					**-0.163**	*0.015*
σ_λ					**0.243**	*0.026*
σ_ρ					**0.515**	*0.084*
σ_{CD4}					**0.289**	*0.003*
σ_P					**0.281**	*0.019*
P likelihood			**-618.6**			
NP likelihood			**-609.4**			
LCVa			**4.762**			

5. Comparing different protocols. We have used the "cycle effect" model to compare four administration protocols of IL-7, with the hope that we can find protocols with equivalent efficiency as the protocol actually applied, but necessitating less injections. Protocol A was the protocol actually applied in the INSPIRE studies. For the three other protocols considered, CD4 counts are measured every three months, and a new cycle is administered when CD4 count < 550 cells/μL, as in Protocol A. While in Protocol A, all the cycles include three injections, we examined the possibility of reducing the number of injections by cycle. In Protocol B, the patient receives a first three-injection cycle, followed by repeated two-injection cycles; in Protocol C the patient receives a first three-injection cycle followed by repeated one-injection cycles; in Protocol D the patient always receives two-injection cycles (including the initial one).

The protocols were compared according to three criteria. The three quantities of interest were: number of injections received, mean CD4 count, and time spent below 500 cells/μL over a four-year period. The criteria to be compared over the four protocols were the expectations of these quantities, over both the random effects and the posterior distribution of the parame-

ters. These expectations were computed by simulation. Based on the results
of the cycle-effect model (Section 4.3) we drew at random 200 values of the
parameters (including the variances of the random effects) from their poste-
rior distribution (approximated to be multivariate normal); then 100 random
effects values were drawn for each value of the fixed parameters. We applied
the inclusion criterion, that is the initial value of the CD4 counts between
100 and 400 (draws leading to values outside of this range were eliminated)
in order to generate the target population of "low immunological respon-
ders". The trajectories were then computed which allowed to compute the
quantities of interest. The medians of these quantities were then computed.
We also computed the variance of these medians for different values of the
parameters to evaluate the uncertainty of the result due to the uncertainty
about the value of the parameters. The results for the whole target popula-
tion are displayed in the first part of Table 4, which gives the three criteria
for the four protocols as well as the standard deviation due to the uncertainty
about the fixed parameters.

Protocol B leads to similar results as Protocol A in terms of mean CD4
count, with the non negligible advantage that it requires significantly less
injections. Protocol C uses less injections but leads to an increased time
spent under 500 cells/μL and a lower mean CD4 count than Protocol A.
Protocol D is very similar to Protocol B, showing that one can spare the
third injection even in the first cycle.

As an illustration, Figure 2 shows the trajectories of an average patient
(with null random effects) for the four protocols, and we see that that the
trajectories for Protocols B and D, although using less injections, and nearly
the same as for Protocol A.

While the target population for IL7 treatment are "low immunological
responders" (to cART) with initial CD4 count < 400, we may distinguish
the group having initial CD4 counts lower than 200 that we call "very low
responders" and the group with initial CD4 counts higher (or equal) to 300
called "not too low responders". We did the same analysis as above on these
two groups. The results are given in Table 4. For "very low responders",
Protocol C is clearly inadequate, while Protocols B and D achieve similar
result as Protocol A with less injections. For "not too low responders", Pro-
tocol C is acceptable, using less injections than the other protocols, although
leading to a slightly lower mean CD4 count. Protocols B and D however,
achieve nearly the same reduction of the number of injections as Protocol C
with less cycles: protocol D uses five cycles while Protocol C uses 8 cycles.

TABLE 4

Comparison of the medians over four years of number of injections and cycles received, time under 500 CD4 count and mean CD4 count for the four protocols: A, three-injection cycles; B, three-injection cycle followed by repeated two-injection cycles; C, three-injection cycle followed by repeated one-injection cycles ; D, two-injection cycles (including the initial one). The standard errors of the medians due to uncertainty of the parameters are given. Analyses done for the whole target population ("low responders"), and the subpopulations "Very low responders" (initial CD4 count < 200) and "Not too low responders" (initial CD4 count > 300).

	A	**B**	**C**	**D**
"Low responders"				
Number of injections received	27 (1.8)	19 (0.9)	13 (1.2)	18 (0.9)
Time under 500 CD4/μL (days)	81 (43.8)	78 (41.1)	114 (63.7)	78 (41.1)
Mean CD4 count	641 (8.2)	633 (8.4)	599 (4.5)	633 (8.4)
"Very low responders"				
Number of injections received	39 (5.7)	29 (3.8)	19 (0)	28 (3.8)
Time under 500 CD4/μL (days)	396 (96.6)	409 (102.9)	998 (306.3)	409 (102.9)
Mean CD4 count	599 (27.8)	591 (28.7)	475 (53.5)	592 (28.8)
"Not too low responders"				
Number of injections received	15 (2.2)	11 (1.5)	8 (1.1)	10 (1.5)
Time under 500 CD4/μL (days)	5 (0.2)	5 (0.2)	5 (0.2)	5 (0.2)
Mean CD4 count	650 (8.7)	641 (8.4)	613 (4.9)	642 (8.4)

6. Discussion. INSPIRE 2 and INSPIRE 3 are the first studies where repeated cycles of IL-7 were administrated to test the long-term restoration of the immune system in low immunological responders. Here we have used a simple mathematical model with complex statistical approaches to model the effect of these repeated cycles on CD4$^+$ T cells concentration. We worked with two CD4$^+$ T cells populations: quiescent and proliferating (presenting the Ki67$^+$ marker). The checks of fit were reasonably good although the model had some difficulties to catch the peak leading to some more extreme values of the residuals than expected.
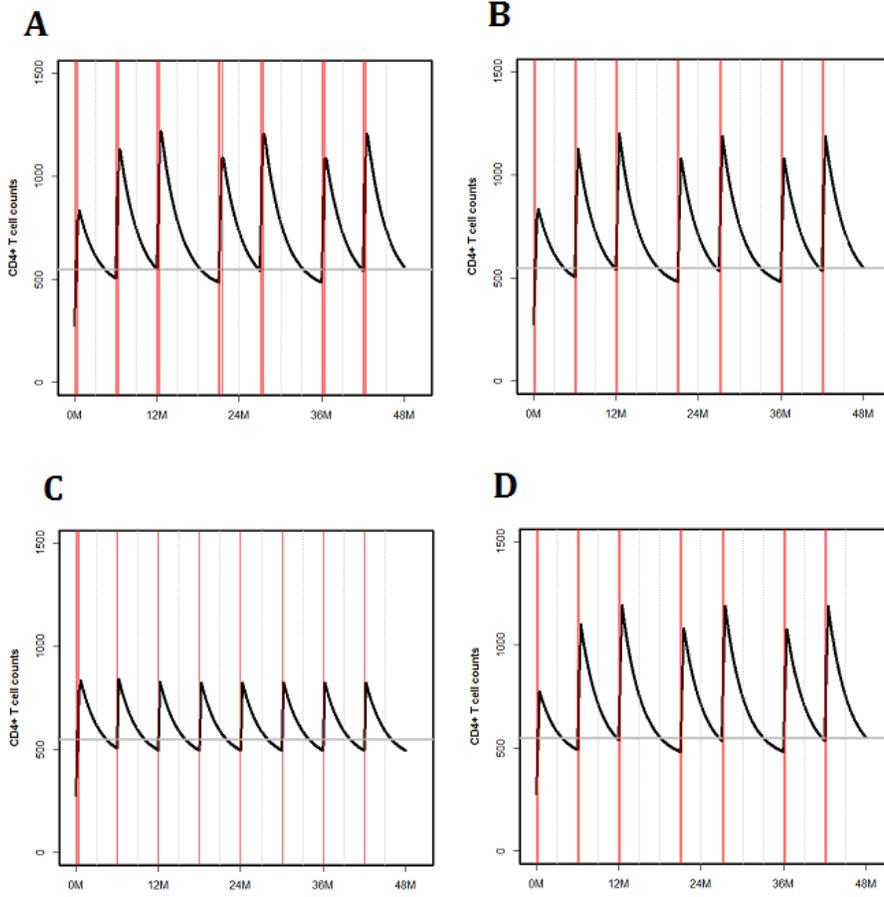
Fig 2: CD4 count (cells/μL) predictions for 4 years for a patient having $u^i_\rho = u^i_\lambda = 0$, for the four protocols A, B, C, D. Vertical dotted lines are CD4 count assessments (every three months) and vertical solid lines are injections. Horizontal line marks the CD4 threshold of 550 cells/μL.

When considering every injection separately, the first important result of this paper is that there is a decreasing effect of successive injections on proliferation rate; the third injection appears to have a weaker effect. We also found that the effect of repeated cycles on proliferation rate was slightly weaker than the effect of the initial one. This can be due to the natural homeostatic regulation of CD4$^+$ T cells, since repeated cycles start at a higher CD4 count. In order to investigate this question, we have introduced a feedback term; in this case the feedback term slightly improved the fit but the

estimate of the cycle effect did not change much. Thus although a feedback mechanism is plausible, there may be other reasons of this phenomenon; one possibility is the presence of antibodies against IL7 after the first cycle. In spite of this phenomenon, simulations show how these repeated cycles are able to maintain adequate CD4 counts for a long time.

We have compared four protocols and shown that cycles of two injections (Protocols B or D) should be sufficient, sparing a certain number of injections without detrimental effect on CD4 count. The conclusion holds even when analysing two different groups, "very low responders" and "not too low responders". Our results agree with a survival analysis presented in Thiébaut et al. (2016) who compared the time spent over 500 cells/μL after a three-injection cycle and a two-injection cycle.

Also, the inclusion of random effects is a key ingredient when considering dynamic models as assistance for treatment personalized decisions. Inter-individual differences in parameters imply inter-individual differences in expected trajectories that can be used for devising adaptive treatment strategies (Prague et al., 2012). We could use this mechanistic model for guiding the treatment, with the aim of minimizing the number of administered injections within repeated cycles ensuring the expected response. Predictions could also be made for different time lapses between cycles or thresholds for receiving a new cycle.

References.

COMMENGES, D., JACQMIN-GADDA, H., PROUST, C. and GUEDJ, J. (2006). A newton-like algorithm for likelihood maximization: The robust-variance scoring algorithm. *arXiv preprint math/0610402*.

COMMENGES, D., JOLY, P., ANNE, G.-P. and LIQUET, B. (2007). Choice between Semiparametric Estimators of Markov and Non-Markov Multi-state Models from Coarsened Observations. *Scandinavian Journal of Statistics* **34** 33–52.

COMMENGES, D., SAYYAREH, A., LETENNEUR, L., GUEDJ, J., BAR-HEN, A. et al. (2008). Estimating a difference of Kullback–Leibler risks using a normalized difference of AIC. *The Annals of Applied Statistics* **2** 1123–1142.

COMMENGES, D., PROUST-LIMA, C., SAMIERI, C. and LIQUET, B. (2015). A universal approximate cross-validation criterion for regular risk functions. *The International Journal of Biostatistics* **11** 51–67.

DRYLEWICZ, J., COMMENGES, D. and THIEBAUT, R. (2012). Maximum a Posteriori estimation in dynamical models of primary HIV infection. *Statistical Communications in Infectious Diseases* **4**.

FINKENSTÄDT, B., WOODCOCK, D. J., KOMOROWSKI, M., HARPER, C. V., DAVIS, J. R., WHITE, M. R., RAND, D. A. et al. (2013). Quantifying intrinsic and extrinsic noise in gene transcription using the linear noise approximation: An application to single cell data. *The Annals of Applied Statistics* **7** 1960–1982.

FRY, T. J. and MACKALL, C. L. (2002). Interleukin-7: from bench to clinic. *Blood* **99** 3892–3904.

GENZ, A. and KEISTER, B. (1996). Fully symmetric interpolatory rules for multiple integrals over infinite regions with Gaussian weight. *Journal of Computational and Applied Mathematics* **71** 299–309.

GUEDJ, J., THIÉBAUT, R. and COMMENGES, D. (2007a). Maximum likelihood estimation in dynamical models of HIV. *Biometrics* **63** 1198–1206.

GUEDJ, J., THIÉBAUT, R. and COMMENGES, D. (2007b). Practical identifiability of HIV dynamics models. *Bulletin of mathematical biology* **69** 2493–2513.

HO, D. D., NEUMANN, A. U., PERELSON, A. S., CHEN, W., LEONARD, J. M., MARKOWITZ, M. et al. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373** 123–126.

HUANG, Y., LIU, D. and WU, H. (2006). Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system. *Biometrics* **62** 413–423.

KONDRACK, R. M., HARBERTSON, J., TAN, J. T., MCBREEN, M. E., SURH, C. D. and BRADLEY, L. M. (2003). Interleukin 7 regulates the survival and generation of memory CD4 cells. *The Journal of experimental medicine* **198** 1797–1806.

KONISHI, S. and KITAGAWA, G. (2008). *Information criteria and statistical modeling.* Springer Science & Business Media.

KUHN, E. and LAVIELLE, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis* **49** 1020–1038.

LEVY, Y., LACABARATZ, C., WEISS, L., VIARD, J.-P., GOUJARD, C., LELIÈVRE, J.-D., BOUÉ, F., MOLINA, J.-M., ROUZIOUX, C., AVETTAND-FÉNOÊL, V. et al. (2009). Enhanced T cell recovery in HIV-1–infected adults through IL-7 treatment. *The Journal of clinical investigation* **119** 997.

LEVY, Y., SERETI, I., TAMBUSSI, G., ROUTY, J., LELIEVRE, J., DELFRAISSY, J., MOLINA, J., FISCHL, M., GOUJARD, C., RODRIGUEZ, B. et al. (2012). Effects of recombinant human interleukin 7 on T-cell recovery and thymic output in HIV-infected patients receiving antiretroviral therapy: results of a phase I/IIa randomized, placebo-controlled, multicenter study. *Clinical infectious diseases* **55** 291–300.

MACKALL, C. L., FRY, T. J. and GRESS, R. E. (2011). Harnessing the biology of IL-7 for therapeutic application. *Nature Reviews Immunology* **11** 330–342.

MACKALL, C. L., FRY, T. J., BARE, C., MORGAN, P., GALBRAITH, A. and GRESS, R. E. (2001). IL-7 increases both thymic-dependent and thymic-independent T-cell regeneration after bone marrow transplantation. *Blood* **97** 1491–1497.

NAMEN, A., SCHMIERER, A., MARCH, C., OVERELL, R., PARK, L., URDAL, D. and MOCHIZUKI, D. (1988). B cell precursor growth-promoting activity. Purification and characterization of a growth factor active on lymphocyte precursors. *The Journal of experimental medicine* **167** 988–1002.

OKAMOTO, Y., DOUEK, D. C., MCFARLAND, R. D. and KOUP, R. A. (2002). Effects of

exogenous interleukin-7 on human thymus function. *Blood* **99** 2851–2858.

Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M. and Ho, D. D. (1996). HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271** 1582–1586.

Pinheiro, J. C. and Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS.* Springer.

Post, T. M., Freijer, J. I., Ploeger, B. A. and Danhof, M. (2008). Extensions to the visual predictive check to facilitate model performance evaluation. *Journal of pharmacokinetics and pharmacodynamics* **35** 185–202.

Prague, M., Commenges, D., Drylewicz, J. and Thiébaut, R. (2012). Treatment Monitoring of HIV-Infected Patients based on Mechanistic Models. *Biometrics* **68** 902–911.

Prague, M., Commenges, D., Guedj, J., Drylewicz, J. and Thiébaut, R. (2013). NIMROD: A program for inference via a normal approximation of the posterior in models with random effects based on ordinary differential equations. *Computer methods and programs in biomedicine* **111** 447-458.

Ramsay, J. O., Hooker, G., Campbell, D. and Cao, J. (2007). Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69** 741–796.

Rue, H., Martino, S. and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)* **71** 319–392.

Seddon, B., Tomlinson, P. and Zamoyska, R. (2003). Interleukin 7 and T cell receptor signals regulate homeostasis of CD4 memory cells. *Nature immunology* **4** 680–686.

Sereti, I., Dunham, R. M., Spritzler, J., Aga, E., Proschan, M. A., Medvik, K., Battaglia, C. A., Landay, A. L., Pahwa, S., Fischl, M. A. et al. (2009). IL-7 administration drives T cell–cycle entry and expansion in HIV-1 infection. *Blood* **113** 6304-6314.

Sportès, C., Hakim, F. T., Memon, S. A., Zhang, H., Chua, K. S., Brown, M. R., Fleisher, T. A., Krumlauf, M. C., Babb, R. R., Chow, C. K. et al. (2008). Administration of rhIL-7 in humans increases in vivo TCR repertoire diversity by preferential expansion of naive T cell subsets. *The Journal of experimental medicine* **205** 1701–1714.

Thiébaut, R., Jacqmin-Gadda, H., Leport, C., Katlama, C., Costagliola, D., Le Moing, V., Morlat, P., Chêne, G., Group, A. S. et al. (2003). Bivariate longitudinal model for the analysis of the evolution of HIV RNA and CD4 cell count in HIV infection taking into account left censoring of HIV RNA measures. *Journal of biopharmaceutical statistics* **13** 271–282.

Thiébaut, R., Drylewicz, J., Prague, M., Lacabaratz, C., Beq, S., Jarne, A., Croughs, T., Sekaly, R.-P., Lederman, M. M., Sereti, I. et al. (2014). Quantifying and Predicting the Effect of Exogenous Interleukin-7 on CD4+ T Cells in HIV-1 Infection. *PLoS computational biology* **10** e1003630.

Thiébaut, R., Jarne, A., Routy, J.-P., Sereti, I., Fischl, M., Ive, P., Speck, R. et al. (2016). Repeated Cycles of Recombinant Human Interleukin 7 in HIV-Infected Patients with low CD4 T cell reconstitution on Antiretroviral Therapy: Results of two Phase II Multicenter Studies. *Clinical Infectious Diseases* **62** 1178–1185.

Van der Vaart, A. W. (2000). *Asymptotic statistics* **3**. Cambridge university press.

Vieira, M., Soares, D., Borthwick, N. J., Maini, M. K., Janossy, G., Salmon, M. and Akbar, A. N. (1998). IL-7-dependent extrathymic expansion of CD45RA+ T cells enables preservation of a naive repertoire. *The Journal of Immunology* **161** 5909–5917.

Wang, L., Cao, J., Ramsay, J., Burger, D., Laporte, C. and Rockstroh, J. (2014).

Estimating mixed-effects differential equation models. *Statistics and Computing* **24** 111–121.

APPENDIX A: DATA AND MATERIALS

A.1. Data source and subjects. The data have been compiled from three phase I/II multicenter studies: INSPIRE (Levy et al., 2012), INSPIRE 2 and INSPIRE 3 (Thiébaut et al., 2016). These studies investigated the effect of a *purified glycosylated recombinant human Interleukin 7* (IL-7) treatment on immune restoration in immunological low responder patients. All participants were aged $\geq$ 18 years, were under stable cART for at least one year, presenting $CD4^+$ T cells count between 100-350 cells/μL (100-400 cells/μL for INSPIRE 2), and undetectable viral load for at least 6 months prior to screening.

In the first study, INSPIRE, 21 patients received three weekly injections (a "complete cycle") of IL-7 at different weight-dependent doses: 10, 20 and 30 μg/kg and the main objective was to evaluate the safety of this treatment. INSPIRE 2 and INSPIRE 3 (with 23 and 84 treated patients, respectively) further studied the biological activity (as well as the safety) of repeated cycles of IL-7 at 20 μg/kg. In this paper, data for all treated patients from the three studies (N=128) have been included from the time of the first injection. Overall, 197 IL-7 cycles were administered (41 of them were incomplete cycles consisting of 1 or 2 injections). More details are provided in a previous publication (Thiébaut et al., 2016).

A.2. Study design and observations. Within the first INSPIRE study, all patients received complete cycles. They had clinic visits at weeks 1, 2 and 3 (at the moment of the injections), weeks 4, 5, 6, 9 and 12, and after, one visit every 3 months; see Levy et al. (2012) for more information. Among many measured biomarkers, our model uses total $CD4^+$ T cells count and the number of $CD4^+$ T cells expressing the Ki67 proliferation marker, the "CD4 count" and "Ki67 count", respectively. Measurements of CD4 counts were made at each visit, while Ki67 counts were only measured at weeks 1, 2, 3, 5 and 12.

For the first twelve patients of INSPIRE 2, clinic visits within the initial cycles were scheduled as for the INSPIRE study (for the rest of them, visits at week 9 were not performed). After, if CD4 counts were found to be below 550 cells/μL in one of the quarterly visits, a new IL-7 cycle was administered (with the exception of the first 12 patients, who wait a year before receiving a new cycle). Within these repeated cycles, clinic visits were scheduled at weeks 1, 2 and 3 (at the moment of the injections), weeks 5 and 12, and once

again quarterly visits are made to check the CD4 count. A maximum of 4 cycles within 21 months and a maximum of 3 cycles within 12 months were established, and all patients have been followed up at least 3 months after the last cycle. CD4 counts were measured at all visits for all patients, while Ki67 counts were measured only for the first cycles of the first 12 patients at weeks 1, 2, 3, 5 and 12.

For INSPIRE 3, patients were randomized into two arms: "IL-7 arm" and "Control arm" with a ratio 3:1 (3 IL-7 : 1 Control). Patients of the "IL-7 arm" received the same treatment scheme as patients from INSPIRE 2. Patients of the "Control arm" were first followed up without receiving the IL-7 for one year, and if CD4 count was still below 500 cells/μL, IL-7 treatment was started as for the other group (Thiébaut et al., 2016). CD4 counts were measured at all visits. No Ki67 counts measurements were available.

The total duration of the studies was 12, 24 and 21 months for INSPIRE, INSPIRE 2 and INSPIRE 3, respectively.

APPENDIX B: MODEL WITH A FEEDBACK TERM

Trajectories satisfying an ODE system have an intrinsic tendency to return to the equilibrium point, when it exists, which is the case for the systems proposed in this paper. In this sense, a feedback term is not necessary to ensure homeostasis, a key concept in physiology. We have, however, considered adding a feedback term in the mathematical model in order to examine the cycle effect β_C in depth. This term will explicitly avoid CD4$^+$ T cells to proliferate without control and possibly ensure a faster return to an equilibrium point. The simplest feedback term is $[\frac{1}{P+Q}]^\nu$, and can be added in both equations to the proliferation term. The system with feedback is as follows:

$$\begin{cases} \frac{dQ^i}{dt} = \lambda^i + 2\rho^i P^i - \mu_Q^i Q^i - \pi^i Q^i \frac{1}{(P^i+Q^i)^\nu} \\ \\ \frac{dP^i}{dt} = \pi^i Q^i \frac{1}{(P^i+Q^i)^\nu} - \rho^i P^i - \mu_P^i P^i \end{cases}$$

Models with feedback were fitted using the 39 patients of INSPIRE who had Ki67 count measurements. The feedback coefficient was estimated at $\nu = 0.119$. In Table 5 we compare some models with and without feedback term.

The feedback term does not lead to a great improvement of the LCVa criterion, especially for the "3 β's" model.

The detection of a cycle effect raises anew the issue of a possible feedback. It may be that the feedback could not be detected when starting with very low CD4 count, but could be more visible when starting at higher CD4

TABLE 5
Comparison of loglikelihoods and LCVa criteria of models with and without feedback for all INSPIRE patients with CD4 and Ki67 count measurements (N=39)

	WITHOUT FEEDBACK		WITH FEEDBACK	
	Basic model	"3 β's" model	Basic model	"3 β's" model
NP loglike	-44.643	-36.549	-41.735	-36.419
P loglike	-49.393	-41.306	-46.965	-41.015
LCVa	1.146	0.940	1.073	0.963

count; this feedback might explain the apparent cycle effect. To answer this question we ran the model for repeated cycles with feedback. With this more complicated model and larger data set, we could not directly estimate the parameter ν, so we resort to profile likelihood. Computing the likelihood for $\nu = 0.05, 0.1, 0.15, 0.20, 0.25, 0.30$ we found that the best likelihood was obtained for $\nu = 0.1$, a value close to what was estimated in the small data set ($\nu = 0.119$). The results are shown in Table 6.

TABLE 6
Priors and estimated mean and standard deviation (sd) of all parameters (in logarithmic and natural scales) for the "cycle effect" model when considering all cycles for each patient including a feedback term with $\nu = 0.1$; Penalized (P) and Non Penalized (NP) likelihood and LCVa criteria

	PRIOR (log-scale)		POSTERIOR (log-scale)		POSTERIOR (natural-scale)	
	mean	sd	mean	sd	mean	sd
λ	1.000	1.000	0.275	0.157	1.316	0.207
ρ	0.000	0.250	1.052	0.083	2.863	0.238
π	-4.000	1.000	-1.975	0.068	0.139	0.009
μ_Q	-3.600	0.500	-2.538	0.067	0.079	0.005
μ_P	-2.500	0.500	-2.212	0.138	0.109	0.015
$\beta_{\pi 1}$					0.806	0.038
$\beta_{\pi 2}$					0.626	0.037
$\beta_{\pi 3}$					0.212	0.035
β_{μ_Q}					-0.063	0.005
β_C					-0.153	0.015
σ_λ					-0.608	0.097
σ_ρ					-0.440	0.071
σ_{CD4}					0.286	0.004
σ_P					0.301	0.021
P likelihood			**-598.0**			
NP likelihood			**-584.5**			
LCVa			**4.567**			

For the repeated cycles data set, the feedback term leads to an improvement of the LCVa criterion. This may reflect a biological feedback mechanism. However, this does not modify the cycle effect β_C.

APPENDIX C: IDENTIFIABILITY AND CONVERGENCE

As can be easily verified, both models with and without the feedback term present no problems regarding the "theoretical" identifiability (that depends on the model structure) but even so, they could present "practical" identifiability problems as explained in Guedj, Thiébaut and Commenges (2007b). In fact, practical identifiability problems are a mix of statistical and numerical problems which are difficult to disentangle; with scarce information, the variances of the estimators are large, but it comes also with a flat shape of the log-likelihood, making it difficult to maximize. The difficulty is enhanced by the fact that there are several layers of numerical computation needed to compute the likelihood, leading to an accumulation of numerical errors.

A crucial point in an iterative algorithm is the stopping criteria. Besides the displacement in the parameter space and the variation of the likelihood function, another convergence criterion proposed by Commenges et al. (2006) has been implemented in NIMROD. It is the Relative Distance to Maximum (RDM) defined as

$$\mathrm{RDM}(\theta^{(k)}) = \frac{U^P(\theta^{(k)})^T G^{-1}(\theta^{(k)}) U^P(\theta^{(k)})}{p}$$

where $U^P(\cdot)$ is the penalized score and $G(\cdot)$ is an approximation of the Hessian of minus the penalized likelihood. This criterion can be interpreted as the ratio of the numerical error over the statistical error, and is asymptotically invariant near the maximum to any one-to-one transformation of the parameters. Prague et al. (2013) propose 0.1 as a good default value.

APPENDIX D: CHECK OF FIT

We performed three graphical procedures to check the fit of of the "3 β's" model with cycle effect ("Cycle effect" model) presented in Section 4.3. A Q-Q plot in Figure 3 looking at the normality of the errors showed good agreement except at the extremes. The fit of residuals for CD4 counts displayed in Figure 4 exhibited no trend in mean nor in dispersion. The Visual Predictive Check (VPC) is a popular tool for checking non-linear mixed-effect models (Post et al., 2008). The VPC compares the percentiles of the real data and the percentiles of the data simulated from the statistical model. We simulated 2000 replicates of the original dataset design. For every replicate, we took the parameters in their a-posteriori laws estimated in Table 3. For each patient we computed the trajectory. Then we added a measurement error (with the estimated variance) and kept only the observation times that were on the original dataset design. We chose to keep only
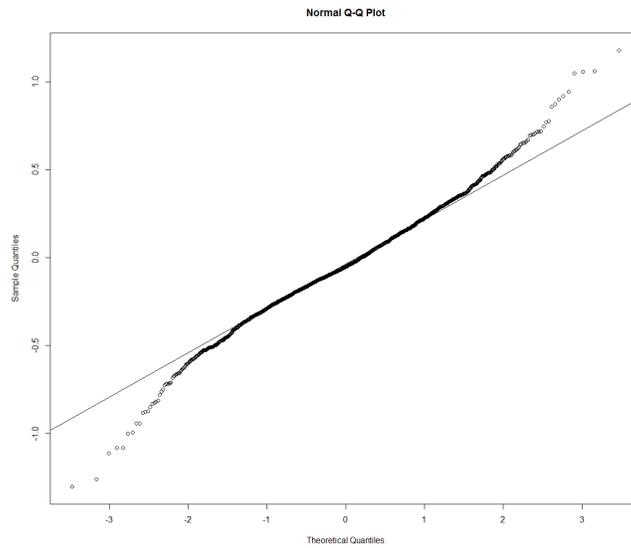
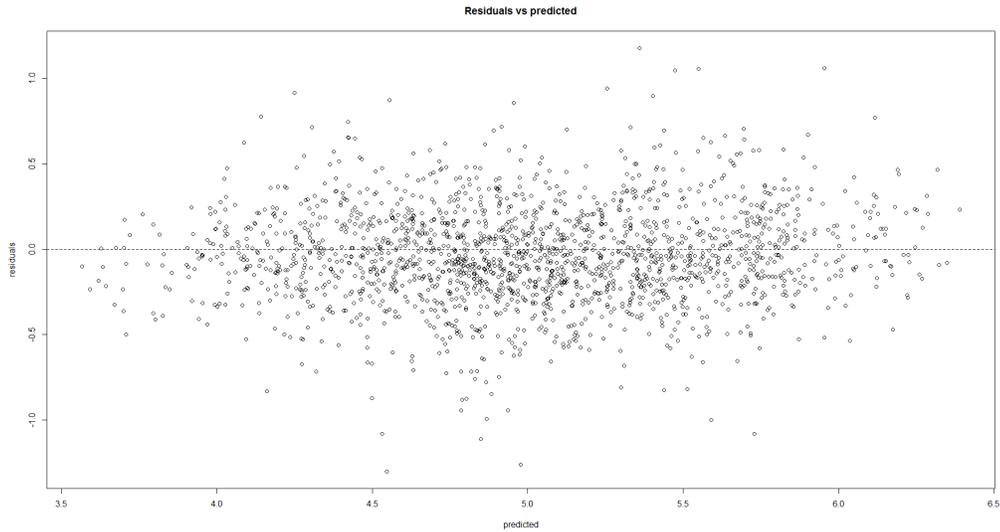Fig 3: *"Cycle effect" model*: Q-Q plot for CD4 counts.

Fig 4: *"Cycle effect" model*: All residuals for CD4 counts plotted against predicted values.

the data from the first cycle of injection, as the time of the second cycle is different for most patient it would be difficult to have a visual interpretation.

The second cycle is administrated when the CD4 levels are too low, so to avoid a selection effect, we chose to keep only the observations that were before 180 days, as it is a time where only a few patients have already had their second cycle. The black lines represent the median and 95% and 5% percentiles of the observed data, and the blue area represent the confidence interval of the same percentiles but with the 2000 simulated datasets. We used the vpc package on R (http://ronkeizer.github.io/vpc/). The graph is shown in Figure 5 and appears rather satisfactory in that the percentiles of the observed data are most of the time within the confidence intervals predicted by the model.
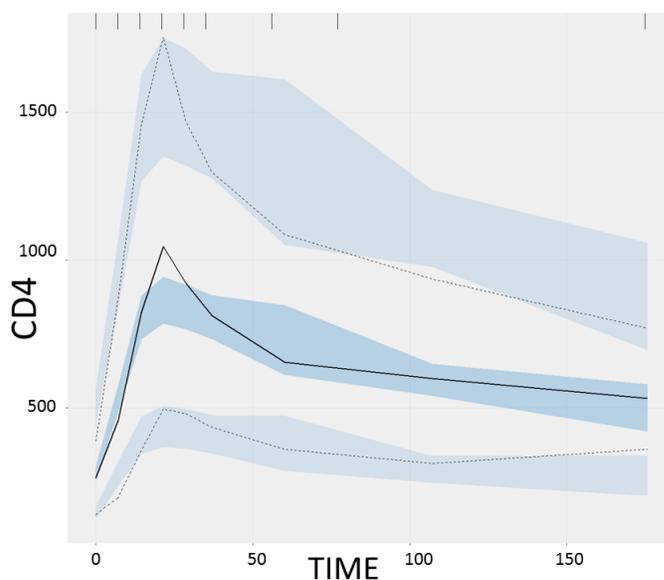

Fig 5: *"Cycle effect" model*: Visual check for CD4 counts.

The same analyses were done for Ki67 counts and exhibited similar features.

APPENDIX E: SOME FITS OF TOTAL CD4 AND KI67 COUNTS

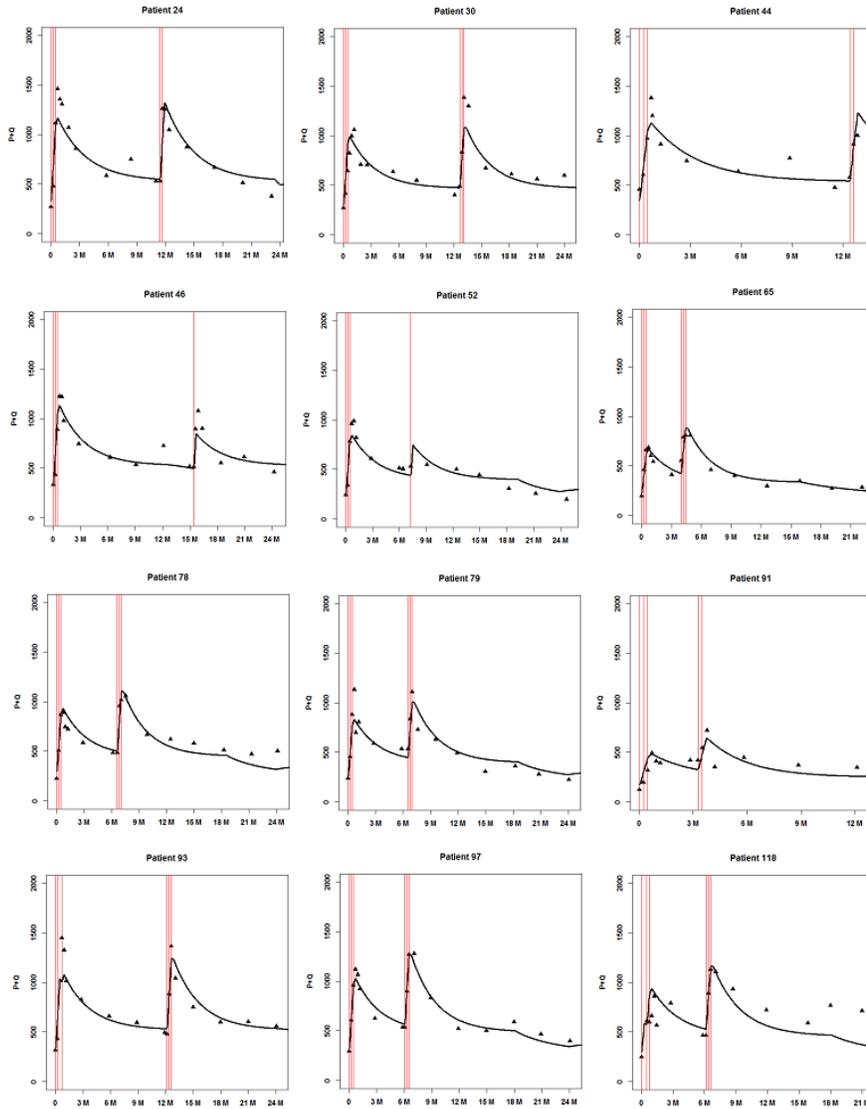E-MAIL: rodolphe.thiebaut@isped.u-bordeaux2.fr

Fig 6: *"Cycle effect" model*: Fits of total CD4 count for 12 patients from INSPIRE 2 and 3 chosen randomly among those who received more than a cycle.
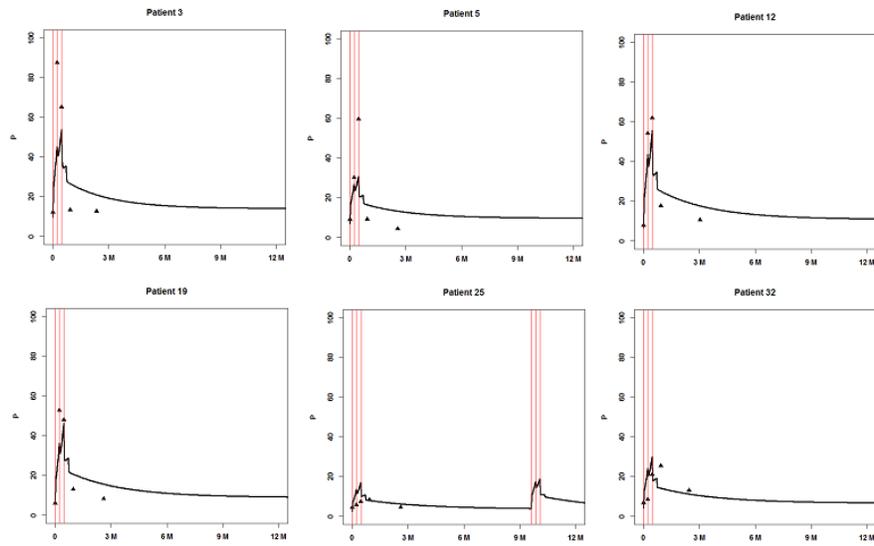
Fig 7: *"Cycle effect" model*: Fits of Ki67 count for 6 patients from INSPIRE and INSPIRE 2 chosen randomly among those who had measurements for this biomarker (only during the first cycle).