

Mixed-effects models for growth curves: an application to the study of reconstitution kinetics of lymphocyte subpopulations

Laura Azzimonti, Claudia De Lalla, Daniela Montagna, Laura M. Sangalli, Anna Maria Paganoni

Abstract In this work we describe an application of non linear mixed-effects models to biological growth curves. The growth curves represent the reconstitution of lymphocyte subpopulations in the peripheral blood of pediatric patients subjected to Hematopoietic Stem Cell Transplantation; in particular, we focus on iNKT frequency among T cells and on iNKT CD161+ cells among iNKT CD4+ and iNKT CD4- cells. The aim of the study is to describe reconstitution kinetics of these subpopulations and to highlight potential dependencies of these curves on relapse insurgence. We use a logistic mixed-effects model to describe iNKT frequency reconstitution and an asymptotic exponential mixed-effects model to describe iNKT CD161+ frequency growth.

Key words: Longitudinal data analysis, Multilevel models, Biostatistics and bioinformatics, Analysis of censored and incomplete data

Laura Azzimonti
MOX - Dipartimento di Matematica, Politecnico di Milano, e-mail: laura.azzimonti@gmail.com

Claudia De Lalla
Hospital San Raffaele Scientific Institute, Milano, e-mail: delalla.claudia@hsr.it

Daniela Montagna
Fondazione IRCCS Policlinico San Matteo, Pavia

Laura M. Sangalli
MOX - Dipartimento di Matematica, Politecnico di Milano, e-mail: laura.sangalli@polimi.it

Anna Maria Paganoni
MOX - Dipartimento di Matematica, Politecnico di Milano, e-mail: anna.paganoni@polimi.it

1 Introduction

We analyze data consisting of serial measurements of lymphocyte subpopulations in peripheral blood of myeloablated pediatric patients subjected to Hematopoietic Stem Cell Transplantation (HSCT) from haplo-identical donor. These data have been collected by the Experimental Immunology Unit DIBIT of Hospital San Raffaele Scientific Institute (Milano) and Fondazione IRCCS Policlinico San Matteo (Pavia). The quantities that we analyze are invariant Natural Killer T cells (iNKT) and their CD4 and CD161 subpopulations [1, 2, 3]. Measurements have been collected for 17 patients in the period between the first and the 18th month after transplantation, approximately every month, through blood samples. Between the 17 patients considered in the study, 4 of them had a relapse before the 18th month whereas 13 had not. The aim of this study is to analyze and describe reconstitution kinetics of different lymphocyte subsets and to highlight potential dependencies of these curves on relapse insurgence [4, 5].

These data present many peculiarities that make the analysis difficult: there are only few subjects in the study; the times of observation for each patient are not evenly spaced and differ from patient to patient. Other difficulties are caused by the presence of censoring: the censoring may be on the left side, since for some patients the first measurement is sampled later than the first month after transplantation; the censoring may also be on the right side, since for many patients measurements have not been collected until the 18th month and some patients left the study due to a relapse.

2 Applications of mixed-effects models to cell reconstitution

We analyze the data presented in Section 1 using non-linear mixed-effects models [6]. These models are particularly suited to describe longitudinal data. Moreover, they are also able to overcome the difficulties discussed in the previous section. Mixed-effects models include parameters associated with the entire population, called fixed effects, and subject-specific parameters, called random effects. For this reason mixed-effects models are able to describe the common dynamic of the phenomenon under investigation, even in presence of high between-subjects variability, yielding estimates both for the entire population and subject-specific.

2.1 Analysis of iNKT frequency reconstitution

The first quantity under investigation is the frequency of iNKT cells among T cells. We use a logistic growth curve to describe the kinetic of reconstitution of this quantity. A logistic model is justified by the fact that in healthy children the frequency of iNKT cells stabilizes at the second year from birth. The aim of this analysis is

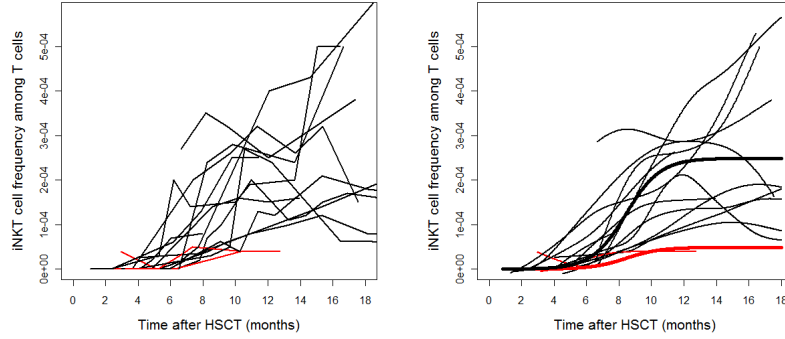


Fig. 1 iNKT cell frequency among T cells vs time after HSCT, in red for patients who had a relapse and in black for patients who had not a relapse. Left panel: observed data. Right panel: fitted fixed-effects logistic curves (bold lines) superimposed to smoothed growth curves.

to describe the kinetic of reconstitution of iNKT cell frequency and to focus on differences between relapsed and non-relapsed patients. Figure 1, left panel, displays the iNKT frequency among T cells vs time after HSCT for all patients considered in the study: the observations of each patient are linked by lines, red for patients who incurred in relapse and black for patients who did not incur in relapse. A better visualization of growth curves is provided in the right panel of the same figure, where raw data have been smoothed via local linear regression.

A visual inspection of Figure 1 shows that iNKT cell frequency for patients having a relapse, in red, does not increase as much as that of non-relapsed patients, in black, and that the variability between relapsed patients is lower. We shall describe iNKT frequency growth by the following model: for $i = 1, \dots, M$, $j = 1, \dots, n_i$

$$y_{ij} = \frac{\alpha^R \cdot I_{\{i \in R\}} + (\alpha^{NR} + a_i^{NR}) \cdot I_{\{i \in NR\}}}{1 + e^{-(t_{ij} - (\mu + m_i^{NR} I_{\{i \in NR\}})) / \gamma}} + \varepsilon_{ij} \quad (1)$$

$$\begin{aligned} a_i^{NR} &\sim N(0, \sigma_a^2) && i.i.d. \\ m_i^{NR} &\sim N(0, \sigma_m^2) && i.i.d. \\ \varepsilon_{ij} &\sim N(0, \sigma^2) && i.i.d. \end{aligned}$$

where $M = 17$ is the number of patients, n_i is the number of measurements for patient i , y_{ij} is the frequency of iNKT cells among T cells, t_{ij} is time (in months) elapsed from transplantation to the j th measurement for patient i , R and NR are respectively the index subsets of relapsed and non-relapsed patients. The errors ε_{ij} are normally distributed and independent errors associated with the j th measurement of patient i .

The parameters used in the model (1) for the fixed effects and estimated from the data are: α^R , the asymptote for relapsed patients; α^{NR} , the asymptote for non-relapsed patients; μ , the inflection point, which correspond to the time at which

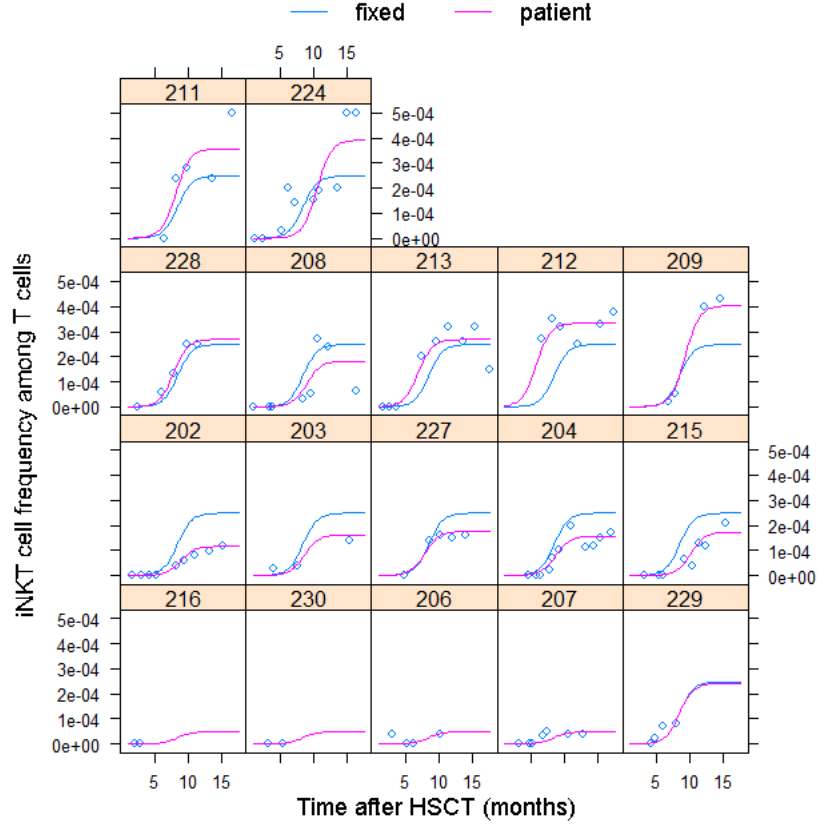


Fig. 2 Predicted curves for iNKT cell frequency among T cells for the various patients through logistic mixed-effects model.

the growth curve reaches the half of the asymptotic level; γ , the time elapsed between μ and the time at which the growth curve reaches $3/4$ of the asymptotic level. The model also includes random effects for the asymptote, a_i^{NR} , and for the inflection point, m_i^{NR} . Random effects are used only for non-relapsed patients and not for relapsed ones. If included in the model, random effects for relapsed patients would have approximately null variance (8 magnitude orders less than variance of random effects for non-relapsed patients); this is consistent with what we have commented above about the little variability between relapsed patients. Random effects are assumed normally distributed, independent among themselves and independent of errors.

All parameters are estimated in R through the maximum likelihood method based on Lindstrom and Bates algorithm [6] implemented in the `nlme` package [8]. The bold lines displayed in the left panel of Figure 1, besides curves for all the patients

	Value	std error	p-value		Value	95% IC lower bound	95% IC upper bound
α^R	0.000049	0.0000352	0.1629	σ_a	$1.026079 \cdot 10^{-4}$	$6.543785 \cdot 10^{-5}$	$1.608914 \cdot 10^{-4}$
α^{NR}	0.000249	0.0000327	0.0000	σ_m	1.473184	0.852571	2.545560
μ	8.437210	0.5197482	0.0000	σ	$5.507862 \cdot 10^{-5}$	$4.708399 \cdot 10^{-5}$	$6.443070 \cdot 10^{-5}$
γ	1.055258	0.2647638	0.0001				

Table 1 Parameter estimates for logistic mixed-effects model for iNKT cell frequency among T cells.

considered in the study, are fitted fixed-effects curves for patients who had a relapse, in red, and for patients who had not, in black. We can see that the fitted fixed-effects curve of patients who had a relapse reaches only a very low frequency of iNKT cells at 18 months. The t-test on the parameter α^R yields p-value 0.1629 (see Table 1), thus suggesting that iNKT cell frequency of relapsed patients does not increase over time and that the kinetic of iNKT cell frequency for these patients degenerate into a constantly null growth curve.

Figure 2 shows the predicted growth curves for the various patients: the blue curve is the fitted fixed-effects curve and the pink one is the predicted subject-specific curve. Notice that for patients 206, 207, 216 and 230, who incurred in relapse, the fixed-effects and subject-specific curves overlap, since random effects have not been included for these patients. Predicted curves provide very good descriptions of the reconstitution of iNKT cell frequency for almost every patient.

2.2 Analysis of iNKT CD161⁺ frequency reconstitution

We now want to investigate the frequency reconstitution of two subpopulations: iNKT cells that express CD161⁺ marker among CD4⁺ iNKT cells and iNKT cells that express CD161⁺ marker among CD4⁻ iNKT cells. CD161⁺ is a typical marker of Natural Killer lymphocytes (NK), which represents the phenotypic maturation of iNKT cells. The goal of this analysis is to describe the kinetic of reconstitution of these two subpopulations and to highlight differences between them. Only 13 patients are considered in this analysis because for 4 patients iNKT CD161⁺ cells were not detectable.

Figure 3 shows iNKT CD161⁺ frequency among iNKT CD4⁺ (top panel) and iNKT CD161⁺ frequency among iNKT CD4⁻ (bottom panel), raw data and curves smoothed via local linear regression left and right respectively. We shall use the following asymptotic exponential mixed-effects model to describe the growth of and to verify possible differences between these two subpopulations: for $i = 1, \dots, M$, $j = 1, \dots, n_i$ and $k = 1, 2$

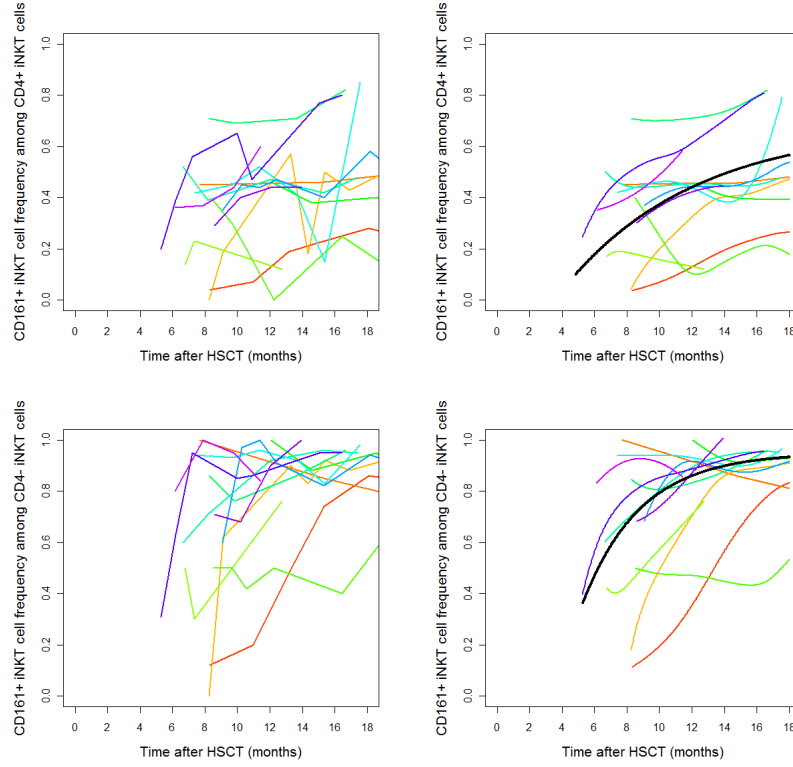


Fig. 3 Top panels: iNKT CD161⁺ frequency among iNKT CD4⁺ vs time after HSTC: observed data (left) and fitted fixed-effects exponential curve superimposed to smoothed growth curves (right). Bottom panels: iNKT CD161⁺ frequency among iNKT CD4⁻ vs time after HSTC: observed data (left) and fitted fixed-effects exponential curve superimposed to smoothed growth curves (right).

$$\begin{aligned}
 y_{ijk} &= (\alpha + \alpha^{CD4} \cdot I_{\{k=1\}}) \left(1 - e^{-e^{(\lambda + \lambda^{CD4} \cdot I_{\{k=1\}} + l_i)(t_{ijk} - \delta)}} \right) I_{\{t_{ijk} > \delta\}} + \varepsilon_{ijk} \\
 l_i &\sim N(0, \sigma_l^2) && i.i.d. \\
 \varepsilon_{ijk} &\sim N(0, \sigma^2) && i.i.d.
 \end{aligned} \tag{2}$$

where $M = 13$, y_{ij1} and y_{ij2} represents respectively iNKT CD161⁺ cell frequency among iNKT CD4⁺ and among CD4⁻ cells and t_{ijk} is time (in months) elapsed from transplantation to the j th measurement for patient i . The errors ε_{ijk} are independent normally distributed errors associated with the j th measurement of the subpopulation k for patient i .

The parameters used in the model (2) for the fixed effects and estimated from the data are: α and α^{CD4} , which are respectively the baseline asymptote and the incremental difference in the asymptote for the CD4⁺ subpopulation with respect to the

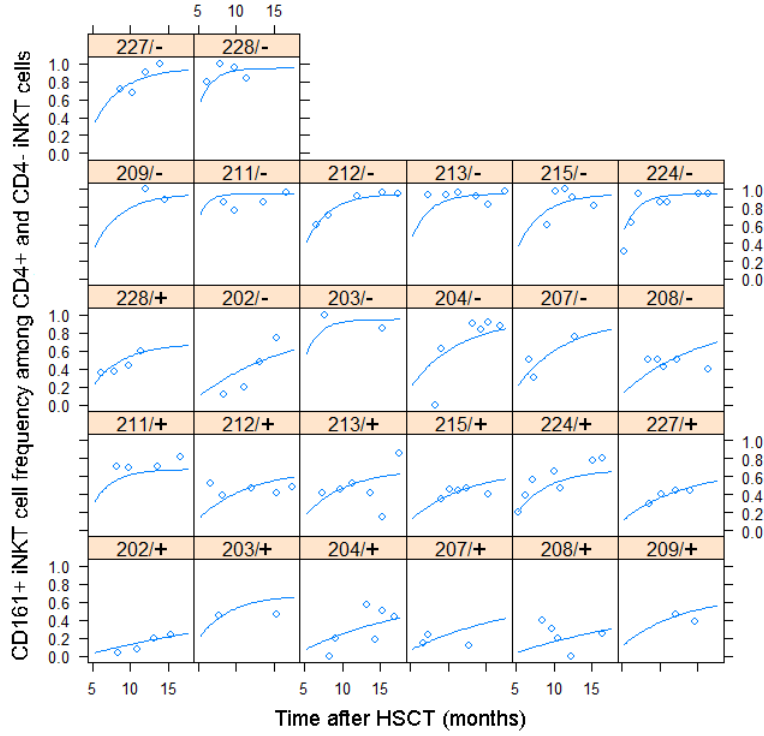


Fig. 4 Predicted curves for NKT CD161⁺ frequency among iNKT CD4⁺ and among iNKT CD4⁻ for the various patients through exponential mixed-effects model.

	Value	std error	p-value			
α	0.949636	0.0382750	0.0000		95% IC	95% IC
α^{CD4}	-0.269804	0.0791130	0.0010	Value	lower bound	upper bound
λ	-1.275433	0.3417345	0.0003	σ_i	0.765874	0.475989 1.232304
λ^{CD4}	-0.814391	0.2423680	0.0012	σ	0.134895	0.117371 0.155035
δ	3.544281	0.8880844	0.0001			

Table 2 Parameter estimates for exponential mixed-effects model for iNKT CD161⁺ cell frequency among iNKT CD4⁺ or among CD4⁻ cells.

CD4⁻ subpopulation; δ , which is a translation in time; λ and λ^{CD4} , which are respectively a baseline parameter related to growth speed at time δ and the incremental difference for the CD4⁺ subpopulation with respect to the CD4⁻ subpopulation. The model also includes random effects on growth speed, l_i , for all the patients. As in the previous section, the parameters are estimated via maximum likelihood.

The bold black lines in Figure 3 are the fitted fixed-effects growth curves for iNKT CD161⁺ frequency among iNKT CD4⁺ (top right panel) and iNKT CD161⁺ frequency among iNKT CD4⁻ (bottom right panel). The t-test on parameters α^{CD4} and λ^{CD4} yield p-values 0.0010 and 0.0012 respectively (see Table 2.2); these results show that the reconstitution dynamic of the two subpopulations are statistically different. In particular, this analysis suggests that not only the asymptote for the two subpopulations is different, as discussed in the medical literature, but also that the initial growth speed is different. Figure 4 shows the predicted subject-specific curves for the two subpopulations.

3 Discussion

We have described an application of mixed-effects models to growth curves concerning lymphocyte subpopulations. In particular, using these models we have been able to explore and describe kinetics of reconstitution of different lymphocyte subsets in pediatric patients after HSCT. The analyses yielded results amenable of clear biological interpretation.

References

1. Bendelac A., Savage P.B. and Teyton L.: The Biology of NKT Cells. *Annu. Rev. Immunol.* **25**, 297–336 (2007)
2. Berzins S.P., Cochrane A.D., Pellicci D.G., Smyth M.J. and Godfrey D.I.: Limited correlation between human thymus and blood NKT cell content revealed by an ontogeny study of paired tissue samples. *Eur. J. Immunol.* **35**, 1399–1407 (2005)
3. De Lalla C., Festuccia N., Albrecht I., Chang H.D., Andol G., Benninghoff U., Bombelli F., Borsellino G., Aiuti A., Radbruch A., Dellabona P. and Casorati G.: Innate-Like Effector Differentiation of Human Invariant NKT Cells Driven by IL-7. *J. Immunol.* **180**, 4415–4424 (2008)
4. De Lalla C., Rinaldi A., Montagna D., Azzimonti L., Sangalli L.M., Paganoni A.M., Maccario R., Bernardo M.E., Locatelli F., Dellabona P. and Casorati G.: iNKT cell reconstitution in paediatric leukaemia patients following haploidentical stem cell transplantation suggests contribution to leukaemia control and reveals independent CD4⁺ and CD4⁻ subset maturation programmes. Conference proceeding of European Group for Blood and Marrow Transplantation 2010.
5. De Lalla C., Rinaldi A., Montagna D., Azzimonti L., Sangalli L.M., Paganoni A.M., Maccario R., Bernardo M.E., Locatelli F., Dellabona P. and Casorati G.: Contribution to leukemia control and independent CD4⁺ and CD4⁻ subset maturation of iNKT cells reconstituting children undergoing haploidentical stem cell transplantation. Work in progress.
6. Lindstrom M. J. and Bates D. M.: Nonlinear mixed-effects models for repeated measures data. *Biometrics* **46**, 673–687 (1990)
7. Pinheiro J.C. and Bates D.M.: *Mixed-Effects Models in S and S-plus*. Springer-Verlag, New York (2000)
8. Pinheiro J.C., Bates D.M., DebRoy S., Sarkar D. and the R Core team (2009). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-93.