IN PARTNERSHIP WITH:
INSERM

Université de Bordeaux

Activity Report 2016

Project-Team SISTM

Statistics In System biology and Translational Medicine

RESEARCH CENTER
Bordeaux - Sud-Ouest

THEME
Computational Neuroscience and Medecine

Table of contents

Project-Team SISTM

Creation of the Team: 2013 April 02, updated into Project-Team: 2015 January 01

Keywords:

Computer Science and Digital Science:

- 3.3.2. - Data mining
- 3.3.3. - Big data analysis
- 3.4.1. - Supervised learning
- 3.4.2. - Unsupervised learning
- 3.4.4. - Optimization and learning
- 3.4.5. - Bayesian methods
- 6.1.1. - Continuous Modeling (PDE, ODE)
- 6.2.4. - Statistical methods
- 6.3.1. - Inverse problems
- 6.3.4. - Model reduction
- 6.4.2. - Stochastic control

Other Research Topics and Application Domains:

- 1.1. - Biology
- 1.1.6. - Genomics
- 1.1.7. - Immunology
- 1.1.9. - Bioinformatics
- 1.1.11. - Systems biology
- 1.4. - Pathologies
- 2.2.4. - Infectious diseases, Virology
- 2.2.5. - Immune system diseases
- 2.3. - Epidemiology
- 2.4.1. - Pharmaco kinetics and dynamics
- 2.4.2. - Drug resistance
- 2.8. - Sports, performance, motor skills

1. Members

Research Scientists

Daniel Commenges [INSERM, Senior Researcher, HDR]
Melanie Prague [Inria, Researcher]

Faculty Members

Rodolphe Thiébaut [Team leader, PUPH, Univ. Bordeaux, Professor, HDR]
Marta Avalos Fernandez [Univ. Bordeaux, Associate Professor]
Robin Genuer [Univ. Bordeaux, Associate Professor]
Laura Richert [MCUPH, Univ. Bordeaux, Associate Professor]
Boris Hejblum [Univ. Bordeaux, Associate Professor]

Engineers

Henri Bonnabau [Univ. Bordeaux]

Chariff Alkhassim [INSERM, from Mar 2015]
Aaron Ayllon Benitez [INSERM, from until Oct 2016]

PhD Students

Perrine Soret [INSERM]
Laura Villain [Univ. Bordeaux]
Chloe Pasin [Univ. Bordeaux]
Edouard Lhomme [Univ. Bordeaux]
Hadrien Lorenzo [Univ. Bordeaux]

Post-Doctoral Fellow

Marius Kwemou [INSERM, from May 2016]

Administrative Assistants

Sandrine Darmigny [INSERM]
Catherine Cattaert Megrat [Inria]
Sylvie Embolla [Inria]
Anton Ottavi [INSERM, European Projects coordinator]
Destandau Eugénie [INSERM, Communication]

Others

David Myers [Master Internship, from Apr 2016 until Sep 2016]
Chloé Belguermi [Master Internship, from Apr 2016 until Sep 2016]
Thomas Esnaud [Master Internship, from Mar 2016 until Aug 2016]
Jean-Louis Palgen [Master Internship, from Feb 2016 until Jun 2016]
Hao Ren [Master Internship, from Mar 2016 until Aug 2016]
Madelyn Yiseth Rojas Castro [Master Internship, from Jul 2016 until Sep 2016]
Hugo Thiessard [Assistant ingenior Internship, Jul 2016]

2. Overall Objectives

2.1. Overall Theoretical Objectives

The overall objective of SISTM is to develop statistical methods for the integrative analysis of health data, especially those related to clinical immunology to answer specific questions risen in the application field. To reach this objective we are developing statistical methods belonging to two main research areas:

- Statistical and mechanistic modeling, especially based on ordinary differential equation systems, fitted to population and sparse data
- Statistical learning methods in the context of high-dimensional data

These two approaches are used for addressing different types of questions. Statistical learning methods are developed and applied to deal with the high dimensional characteristics of the data. The outcome of this research leads to hypotheses linked to a restricted number of markers. Mechanistic models are then developed and used for modeling the dynamics of a few markers. For example, regularized methods can be used to select relevant genes among 20000 measured with microarray technology, whereas differential equations can be used to capture the dynamics and relationship between several genes followed over time by a q-PCR assay or RNA-seq.

2.2. Overall Applied Objectives

Data are generated in clinical trials or biological experimentations. Our main application of interest is the immune response to vaccine or other immune interventions (such as exogenous cytokines), mainly in the context of HIV infection. The methods developed in this context can be applied in other circumstances but the focus of the team on immunology is important for the relevance of the results and their translation into practice, thanks to a longstanding collaboration with several immunologists and the implication of the team in the Labex Vaccine Research Institute (http://vaccine-research-institute.fr/fr/). Exemples of objectives related to this application field are:

- To understand how immune response is generated with immune interventions (vaccines or interleukines)
- To predict what would be the immune response to a given immune intervention for designing next studies and adapting interventions to individual patients

3. Research Program

3.1. Mecanistic modelling

When studying the dynamics of a given marker, say the HIV concentration in the blood (HIV viral load), one can for instance use descriptive models summarising the dynamics over time in term of slopes of the trajectories [51]. These slopes can be compared between treatment groups or according to patients' characteristics. Another way for analysing these data is to define a mathematical model based on the biological knowledge of what drives HIV dynamics. In this case, it is mainly the availability of target cells (the CD4+ T lymphocytes), the production and death rates of infected cells and the clearance of the viral particles that impact the dynamics. Then, a mathematical model most often based on ordinary differential equations (ODE) can be written [41]. Estimating the parameters of this model to fit observed HIV viral load gave a crucial insight in HIV pathogenesis as it revealed the very short half-life of the virions and infected cells and therefore a very high turnover of the virus, making mutations a very frequent event [40].

Having a good mechanistic model in a biomedical context such as HIV infection opens doors to various applications beyond a good understanding of the data. Global and individual predictions can be excellent because of the external validity of a model based on main biological mechanisms. Control theory may serve for defining optimal interventions or optimal designs to evaluate new interventions [30]. Finally, these models can capture explicitly the complex relationship between several processes that change over time and may therefore challenge other proposed approaches such as marginal structural models to deal with causal associations in epidemiology [28].

Therefore, we postulate that this type of model could be very useful in the context of our research that is in complex biological systems. The definition of the model needs to identify the parameter values that fit the data. In clinical research this is challenging because data are sparse, and often unbalanced, coming from populations of subjects. A substantial inter-individual variability is always present and needs to be accounted as this is the main source of information. Although many approaches have been developed to estimate the parameters of non-linear mixed models [44], [54], [33], [42], [36], [53], the difficulty associated with the complexity of ODE models and the sparsity of the data leading to identifiability issues need further research.

3.2. High dimensional data

With the availability of omics data such as genomics (DNA), transcriptomics (RNA) or proteomics (proteins), but also other types of data, such as those arising from the combination of large observational databases (e.g. in pharmacoepidemiology or environmental epidemiology), high-dimensional data have became increasingly common. Use of molecular biological technics such as Polymerase Chain Reaction (PCR) allows for amplification of DNA or RNA sequences. Nowadays, microarray and Next Generation Sequencing (NGS) techniques

give the possibility to explore very large portions of the genome. Furthermore, other assays have also evolved, and traditional measures such as cytometry or imaging have became new sources of big data. Therefore, in the context of HIV research, the dimension of the datasets has much grown in term of number of variables per individual than in term of number of included patients although this latter is also growing thanks to the multi-cohort collaborations such as CASCADE or COHERE organized in the EuroCoord network [1]. As an exemple, in a recent phase 1/2 clinical trial evaluating the safety and the immunological response to a dendritic cell-based HIV vaccine, 19 infected patients were included. Bringing together data on cell count, cytokine production, gene expression and viral genome change led to a 20 Go database [50]. This is far from big databases faced in other areas but constitutes a revolution in clinical research where clinical trials of hundred of patients sized few hundred of Ko at most. Therefore, more than the storage and calculation capacities, the challenge is the comprehensive analysis of these datasets.

The objective is either to select the relevant information or to summarize it for understanding or prediction purposes. When dealing with high dimensional data, the methodological challenge arises from the fact that datasets typically contain many variables, much more than observations. Hence, multiple testing is an obvious issue that needs to be taken into account [45]. Furthermore, conventional methods, such as linear models, are inefficient and most of the time even inapplicable. Specific methods have been developed, often derived from the machine learning field, such as regularization methods [52]. The integrative analysis of large datasets is challenging. For instance, one may want to look at the correlation between two large scale matrices composed by the transcriptome in the one hand and the proteome on the other hand [37]. The comprehensive analysis of these large datasets concerning several levels from molecular pathways to clinical response of a population of patients needs specific approaches and a very close collaboration with the providers of data that is the immunologists, the virologists, the clinicians...

4. Application Domains

4.1. Systems Biology and Translational medicine

Biological and clinical researches have dramatically changed because of the technological advances, leading to the possibility of measuring much more biological quantities than previously. Clinical research studies can include now traditional measurements such as clinical status, but also thousands of cell populations, peptides, gene expressions for a given patient. This has facilitated the transfer of knowledge from basic to clinical science (from "bench side to bedside") and vice versa, a process often called "Translational medicine". However, the analysis of these large amounts of data needs specific methods, especially when one wants to have a global understanding of the information inherent to complex systems through an "integrative analysis". These systems like the immune system are complex because of many interactions within and between many levels (inside cells, between cells, in different tissues, in various species). This has led to a new field called "Systems biology" rapidly adapted to specific topics such as "Systems Immunology" [47], "Systems vaccinology" [43], "Systems medicine" [32]. From the statistician point of view, two main challenges appear: i) to deal with the massive amount of data ii) to find relevant models capturing observed behaviors.

4.2. The case of HIV immunology

The management of HIV infected patients and the control of the epidemics have been revolutionized by the availability of highly active antiretroviral therapies. Patients treated by these combinations of antiretrovirals have most often undetectable viral loads with an immune reconstitution leading to a survival which is nearly the same to uninfected individuals [39]. Hence, it has been demonstrated that early start of antiretroviral treatments may be good for individual patients as well as for the control of the HIV epidemics (by reducing the transmission from infected people) [31]. However, the implementation of such strategy is difficult especially in developing countries. Some HIV infected individuals do not tolerate antiretroviral regimen or did not

[1] see online at http://www.eurocoord.net

reconstitute their immune system. Therefore, vaccine and other immune interventions are required. Many vaccine candidates as well as other immune interventions (IL7, IL15) are currently evaluated. The challenges here are multiple because the effects of these interventions on the immune system are not fully understood, there are no good surrogate markers although the number of measured markers has exponentially increased. Hence, HIV clinical epidemiology has also entered in the era of Big Data because of the very deep evaluation at individual level leading to a huge amount of complex data, repeated over time, even in clinical trials that includes a small number of subjects.

4.3. The case of Ebola vaccine development

In response to the recent outbreak of Ebola virus disease in West Africa, the clinical development of some candidate to Ebola vaccine has been accelerated. Several vectors, mostly encoding glycoprotein of the virus, were tested in Phase I-II studies in order to assess their safety and immunogenicity. One of the main question of interest there is the antibody response induced by vaccination, as some non-human primates studies have shown protection against the virus when antibody levels were high enough. Although bridging studies still have to be developed, antibodies are thus considered as a criterium of interest. The challenge is then to evaluate the durability of the antibody response, whether it be at an individual or population level, in order to evaluate the impact of a vaccine strategy in case of an epidemic. Moreover, we are interested in the factors associated to this antibody response, and even more the other immune markers (from both innate and adaptative immune response) able to predict antibody levels. As those relationship are non-linear, sophisticated statistical and mathematical methods are developed in order to address these questions. A systems medicine approach using multidimensional immunogenicity data from clinical trials and statistical models can help to understand vaccine mechanisms and improve the selection of optimised vaccine strategies for clinical trials.

5. Highlights of the Year

5.1. Highlights of the Year

Modeling clinical trials of IL7

We have published the results of two clinical trials [17] that are showing the feasability of repeating IL-7 cycles and confirmed the predictions performed with our dynamical model published in [49]. This mecanistic modeling allow to propose protocol which decrease the number of injection within each IL-7 cycle while keeping the same efficacy [35].

Awards Mélanie Prague published an invited paper on her PhD works (which was supervized by Daniel Commenges and co-suppervized by Rodolphe Thiébaut) as a perks for the attribution of the "Marie-Jeanne Laurent-Duhamel PhD award (2015) by the SFdS (Société Francaise de statistiques). [15]

6. New Software and Platforms

6.1. New software

6.1.1. clogitLasso

Lasso Estimation of Conditional Logistic Regression Models for Case-Crossover and Matched Case-Control Studies

- KEYWORDS: Classification , Statistics , Cluster , Machine learning , Regression
- FUNCTIONAL DESCRIPTION Fit a sequence of conditional logistic regression models with lasso, for small to large sized samples.
- Contact: Marius Kwémou
- URL: https://github.com/robingenuer/

6.1.2. COVVSURF

Combination of Clustering Of Variables and Variable Selection Using Random Forests

- KEYWORDS: Classification , Statistics , Cluster , Machine learning , Regression
- FUNCTIONAL DESCRIPTION This package implements a two stage strategy, where first we use ClustOfVar package to perform a clustering of variables and second we use VSURF package to select features (i.e. synthetic variables built in the first step).
- Contact: Robin Genuer
- URL: https://github.com/robingenuer/CoVVSURF

6.1.3. NPflow

Bayesian Nonparametrics for Automatic Gating of Flow-Cytometry Data

- KEYWORDS: Bayesian estimation , Bioinformatics , Biostatistics
- FUNCTIONAL DESCRIPTION Dirichlet process mixture of multivariate normal, skew normal or skew t-distributions modeling oriented towards flow-cytometry data pre-processing applications.
- Contact: Boris Hejblum
- URL: https://cran.r-project.org/web/packages/NPflow/

6.1.4. tcgsaseq

Time-Course Gene Set Analysis for RNA-Seq Data

- KEYWORDS: Genomics , Biostatistics , Statistical modeling , RNA-seq , Gene Set Analysis
- FUNCTIONAL DESCRIPTION Gene set analysis of longitudinal RNA-seq data with variance component score test accounting for data heteroscedasticity through precision weights.
- Contact: Boris Hejblum
- URL: https://cran.r-project.org/web/packages/tcgsaseq/

6.1.5. CD4 Shiny

Reference curves for CD4 response to antiretroviral therapy in HiV infected patients

- KEYWORDS: HIV infection, antiretroviral therapy, cd4 response, reference curves, quantile regression
- FUNCTIONAL DESCRIPTION References curves for CD4 response to antiretroviral therapy in HIV infected patients derived from large cohorts and estimated according to known factors associated with the response to antiretroviral therapy.
- Contact: Rodolphe Thiébaut
- URL: http://shiny.isped.u-bordeaux.fr/CD4refcurves/

6.2. Older software still maintained by SISTM

6.2.1. NIMROD

Normal approximation Inference in Models with Random effects based on Ordinary Differential equations

- KEYWORDS: Biostatistics - Optimization
- FUNCTIONAL DESCRIPTION We have written a specific program called NIMROD for estimating parameter of ODE based population models.
- Contact: Rodolphe Thiebaut
- URL: http://etudes.isped.u-bordeaux2.fr/BIOSTATISTIQUE/NIMROD/documentation/html/index.html

6.2.2. R2GUESS

Graphical processing Unit Evolutionary Stochastic Search

- KEYWORDS: Bioinformatics - Biostatistics
- FUNCTIONAL DESCRIPTION R2GUESS package is a wrapper of the GUESS (Graphical processing Unit Evolutionary Stochastic Search) program. GUESS is a computationally optimised C++ implementation of a fully Bayesian variable selection approach that can analyse, in a genome-wide context, single and multiple responses in an integrated way. The program uses packages from the GNU Scientific Library (GSL) and offers the possibility to re-route computationally intensive linear algebra operations towards the Graphical Processing Unit (GPU) through the use of proprietary CULA-dense library.
- Contact: Rodolphe Thiebaut
- URL: https://cran.r-project.org/web/packages/R2GUESS/index.html

6.2.3. TcGSA

Time-course Gene Set Analysis

- KEYWORDS: Bioinformatics - Genomics
- FUNCTIONAL DESCRIPTION An R package for the gene set analysis of longitudinal gene expression data sets. Under development, and soon to be available on the CRAN website, this package implements a Time-course Gene Set Analysis method and provides useful plotting functions facilitating the interpretation of the results.
- Contact: Boris Hejblum
- URL: https://cran.r-project.org/web/packages/TcGSA/index.html

6.2.4. VSURF

Variable Selection Using Random Forests

- KEYWORD: Bioinformatics
- FUNCTIONAL DESCRIPTION An R package for Variable Selection Using Random Forests. Available on CRAN, this package performs an automatic (meaning completely data-driven) variable selection procedure. Originally designed to deal with high dimensional data, it can also be applied to standard datasets.
- Contact: Robin Genuer
- URL: https://cran.r-project.org/web/packages/VSURF/index.html

6.2.5. marqLevAlg

Function optimization (minimization or maximization)

- KEYWORDS: Optimization - Biostatistics
- FUNCTIONAL DESCRIPTION An R package for function optimization. Available on CRAN, this package performs a minimization of function based on the Marquardt-Levenberg algorithm. This package is really useful when the surface to optimize is non-strictly convex or far from a quadratic function. A new convergence criterion, the relative distance to maximum (RDM), allows the user to have a better confidence in the stopping points, other than basic algorithm stabilization.
- Contact: Melanie Prague
- URL: https://cran.r-project.org/web/packages/marqLevAlg/index.html

7. New Results

7.1. High dimensional data

Approaches Applied in Genomics Context [13]
Motivation: The association between two blocks of ?omics? data brings challenging issues in computational biology due to their size and complexity. Here, we focus on a class of multivariate statistical methods called partial least square (PLS). Sparse version of PLS (sPLS) operates integration of two datasets while simultaneously selecting the contributing variables. However, these methods do not take into account the important structural or group effects due to the relationship between markers among biological pathways. Hence, considering the predefined groups of markers (e.g. genesets), this could improve the relevance and the efficacy of the PLS approach. Results: We propose two PLS extensions called group PLS (gPLS) and sparse gPLS (sgPLS). Our algorithm enables to study the relationship between two different types of omics data (e.g. SNP and gene expression) or between an omics dataset and multivariate phenotypes (e.g. cytokine secretion). We demonstrate the good performance of gPLS and sgPLS compared with the sPLS in the context of grouped data. Then, these methods are compared through an HIV therapeutic vaccine trial. Our approaches provide parsimonious models to reveal the relationship between gene abundance and the immunological response to the vaccine.

Combining clustering of variables and feature selection using random forests: the CoV/VSURF procedure [26]
High-dimensional data classification is a challenging problem. A standard approach to tackle this problem is to perform variables selection, e.g. using step-wise or LASSO procedures. Another standard way is to perform dimension reduction, e.g. by Principal Component Analysis or Partial Least Square procedures. The approach proposed in this paper combines both dimension reduction and variables selection. First, a procedure of clustering of variables is used to built groups of correlated variables in order to reduce the redundancy of information. This dimension reduction step relies on the R package ClustOfVar which can deal with both numerical and categorical variables. Secondly, the most relevant synthetic variables (which are numerical variables summarizing the groups obtained in the first step) are selected with a procedure of variable selection using random forests, implemented in the R package VSURF. Numerical performances of the proposed methodology called CoV/VSURF are compared with direct applications of VSURF or random forests on the original p variables. Improvements obtained with the CoV/VSURF procedure are illustrated on two simulated mixed datasets (cases n>p and n<p)

Arbres CART et Forêts aléatoires, Importance et sélection de variables [27]
Two algorithms proposed by Leo Breiman : CART trees (Classification And Regression Trees for) introduced in the first half of the 80s and random forests emerged, meanwhile, in the early 2000s, are the subject of this article. The goal is to provide each of the topics, a presentation, a theoretical guarantee, an example and some variants and extensions. After a preamble, introduction recalls objectives of classification and regression problems before retracing some predecessors of the Random Forests. Then, a section is devoted to CART trees then random forests are presented. Then, a variable selection procedure based on permutation variable importance is proposed. Finally the adaptation of random forests to the Big Data context is sketched.

Comments on: " A Random Forest Guided Tour " [8]
This paper is a comment on the survey paper by Biau and Scornet (2016) about random forests. We focus on the problem of quantifying the impact of each ingredient of random forests on their performance. We show that such a quantification is possible for a simple pure forest , leading to conclusions that could apply more generally. Then, we consider " hold-out " random forests, which are a good middle point between " toy " pure forests and Breiman's original random forests.

Targeting HIV-1 Env gp140 to LOX-1 Elicits Immune Responses in Rhesus Macaques. [18]
Improved antigenicity against HIV-1 envelope (Env) protein is needed to elicit vaccine-induced protective immunity in humans. Here we describe the first tests in non-human primates (NHPs) of Env gp140 protein fused to a humanized anti-LOX-1 recombinant antibody for delivering Env directly to LOX-1-bearing antigen presenting cells, especially dendritic cells (DC). These data, as well as the safety of this protein vaccine, justify further exploration of this DC-targeting vaccine approach for protective immunity against HIV-1.

Significant changes in HIV-1 Capsid stability induced by common CTL-driven viral sequence mutations. [46]
HIV-1-infected individuals with protective HLA class I alleles exhibit better control of viremia and slower disease progression. Virus control in these individuals has been associated with strong and potent HIV-1-specific cytotoxic-T-lymphocyte (CTL) responses restricted by protective HLA alleles, but control of viremia also occurs in the presence of selected CTL escape mutations. Taken together, these data demonstrate that CTL-driven escape mutations within p24 Gag restricted by protective HLA class I alleles have a significant impact on capsid stability that might contribute to the persistent control of viral replication observed despite viral escape from CTL responses.

Optimization and evaluation of luminex performance with supernatants of Peripheral Blood Mononuclear Cell culture. [48]
The Luminex bead-based multiplex assay is useful for quantifying immune mediators such as cytokines and chemokines. Cross-comparisons of reagents for this technique from different suppliers have already been performed using serum or plasma but rarely with supernatants collected from antigen-stimulated peripheral blood mononuclear cells (PBMC). Here, we first describe an optimization protocol for cell culture including quantity of cells and culture duration to obtain reproducible cytokine and chemokine quantifications. Then, we compared three different Luminex kit suppliers.

7.2. Modeling biomarkers and Mecanistic modeling

- **Dynamic models for estimating the effect of HAART on CD4 in observational studies: Application to the Aquitaine Cohort and the Swiss HIV Cohort Study. [15]**
  Highly active antiretroviral therapy (HAART) has proved efficient in increasing CD4 counts in many randomized clinical trials. Because randomized trials have some limitations (e.g., short duration, highly selected subjects), it is interesting to assess the effect of treatments using observational studies. This is challenging because treatment is started preferentially in subjects with severe conditions. This general problem had been treated using Marginal Structural Models (MSM) relying on the counterfactual formulation. Another approach to causality is based on dynamical models. We present three discrete-time dynamic models based on linear increments models (LIM): the first one based on one difference equation for CD4 counts, the second with an equilibrium point, and the third based on a system of two difference equations, which allows jointly modeling CD4 counts and viral load. We also consider continuous-time models based on ordinary differential equations with non-linear mixed effects (ODE-NLME). These mechanistic models allow incorporating biological knowledge when available, which leads to increased statistical evidence for detecting treatment effect. Because inference in ODE-NLME is numerically challenging and requires specific methods and softwares, LIM are a valuable intermediary option in terms of consistency, precision, and complexity. We compare the different approaches in simulation and in illustration on the ANRS CO3 Aquitaine Cohort and the Swiss HIV Cohort Study.
- **Use of dynamical models for treatment optimization in HIV infected patients : a sequential Bayesian analysis approach. [15]**
  The use of dynamic mechanistic models based on ordinary differential equations (ODE) has greatly improved the knowledge of the dynamics of HIV and of the immune system. Their flexibility for fitting data and prediction abilities make them a good tool for optimization of the design delivery and efficacy of new intervention in the HIV field. We present the problem of inference in ODE

models with mixed effects on parameters. We introduce a Bayesian estimation procedure based on the maximization of the penalized likelihood and a normal approximation of posteriors, which is implemented in the NIMROD software. We investigate the impact of pooling different data by using a sequential Bayesian analysis (SBA), which uses posteriors of a previous study as new priors. We show that the normal approximation of the posteriors, which constrains the shape of new priors, leads to gains in accuracy of estimation while reducing computation times. The illustration is from two clinical trials of combination of antiretroviral therapies (cART): ALBI ANRS 070 and PUZZLE ANRS 104. This paper reproduces some unpublished work from my PhD thesis. It is an extension of my oral presentation on the same topic at the 47th Journées de Statistique organized by the French Statistical Society (SFdS) in Lille, France, May 2015, when being awarded the Marie-Jeanne Laurent-Duhamel prize.

- **Surveillance of $\gamma\delta$T Cells Predicts Cytomegalovirus Infection Resolution in Kidney Transplants. [11]**
  Cytomegalovirus (CMV) infection in solid-organ transplantation is associated with increased morbidity and mortality, particularly if a CMV mutant strain with antiviral resistance emerges. Monitoring CMV specific T cell response could provide relevant information for patient care. We assessed if V delta 2 neg gamma delta T cell kinetics in peripheral blood predict CMV infection resolution and emergence of a mutant strain in high risk recipients of kidney transplants, including 168 seronegative recipients receiving organs from seropositive donors and 104 seropositive recipients receiving antithymocyte globulins (R+/ATG). In conclusion, longitudinal surveillance of V delta 2 neg gamma delta T cells in recipients of kidney transplants may predict CMV infection resolution and antiviral drug resistance.
- **Early CD4+ T Cell Responses Are Associated with Subsequent CD8+ T Cell Responses to an rAd5-Based Prophylactic Prime-Boost HIV Vaccine Strategy. [12]**
  Initial evaluation of a candidate vaccine against HIV includes an assessment of the vaccine's ability to generate immune responses. However, the dynamics of vaccine-induced immune responses are unclear. We hypothesized that the IFN-gamma producing cytotoxic CD8+ T cell responses could be predicted by early IL-2 producing CD4+ helper T cell responses, and we evaluated this hypothesis using data from a phase I/II prophylactic HIV vaccine trial. The objective was to assess the dynamics after vaccination with a recombinant adenoviral serotype 5 (rAd5) HIV vaccine. Regression models confirmed this relationship with a significant association between the two markers. These results suggest an early and leading role of CD4+ T cells in the cellular response to the rAd5-rAd5 vaccine and in particular the stimulation of cytotoxic CD8+ T cell responses. These results could inform better timing of CD4+ T cell measurements in future clinical trials.
- **Reference curves for CD4 T-cell count response to combination antiretroviral therapy in HIV-1-infected treatment-naïve patients. [29]**
  The aim of this work was to provide a reference for the CD4 T-cell count response in the early months after the initiation of combination antiretroviral therapy (cART) in HIV-1-infected patients. All patients in the Collaboration of Observational HIV Epidemiological Research Europe (COHERE) cohort who were aged > 18 years and started cART for the first time between 1 January 2005 and 1 January 2010 and who had at least one available measurement of CD4 count and a viral load < 50 HIV-1 RNA copies/mL at 6 months (+- 3 months) after cART initiation were included in the study. Unadjusted and adjusted references curves and predictions were obtained using quantile regressions. Reference curves aid the evaluation of the immune response early after antiretroviral therapy initiation that leads to viral control.
- **Repeated Cycles of Recombinant Human Interleukin 7 in HIV-Infected Patients With Low CD4 T-Cell Reconstitution on Antiretroviral Therapy: Results of 2 Phase II Multicenter Studies. [17].**
  Phase I/II studies in human immunodeficiency virus (HIV) infected patients receiving antiretroviral therapy have shown that a single cycle of 3 weekly subcutaneous injections of recombinant human interleukin 7 (r-hIL-7) is safe and improves immune CD4 T-cell restoration. Herein, we report data

from 2 phase II trials evaluating the effect of repeated cycles of r-hIL-7 (20 microg/kg) with the objective of restoring a sustained CD4 T-cell count >500 cells/microL. INSPIRE 2 was a single-arm trial conducted in the United States and Canada. INSPIRE 3 was a 2 arm trial with 3:1 randomization to r-hIL-7 versus control conducted in Europe and South Africa. Participants with plasma HIV RNA levels <50 copies/mL during antiretroviral therapy and with CD4 T-cell counts between 101 and 400 cells/microL were eligible. A repeat cycle was administered when CD4 T-cell counts fell to <550 cells/microL. A total of 107 patients were treated and received 1 (n = 107), 2 (n = 74), 3 (n = 14), or 4 (n = 1) r-hIL-7 cycles during a median follow-up of 23 months. r-hIL-7 was well tolerated. Four grade 4 events were observed, including 1 case of asymptomatic alanine aminotransferase elevation. After the second cycle, anti-r-hIL-7 binding antibodies developed in 82% and 77% of patients in INSPIRE 2 and 3, respectively (neutralizing antibodies in 38% and 37%), without impact on the CD4 T-cell response. Half of the patients spent >63% of their follow-up time with a CD4 T-cell count >500 cells/microL. CONCLUSIONS: Repeated cycles of r-hIL-7 were well tolerated and achieved sustained CD4 T-cell restoration to >500 cells/microL in the majority of study participants.

7.3. Implication in analysis of results from Clinical trials and cohorts

- **Superior efficacy of an HIV vaccine combined with ARV prevention in SHIV challenged non-human primates. [38]**
  Although vaccines and antiretroviral (ARV) prevention have demonstrated partial success against human immunodeficiency virus (HIV) infection in clinical trials, their combined introduction could provide more potent protection. Furthermore, combination approaches could ameliorate the potential increased risk of infection following vaccination in the absence of protective immunity. We used a nonhuman primate model to determine potential interactions of combining a partially effective ARV microbicide with an envelope-based vaccine. These important findings suggest that combined implementation of new biomedical prevention strategies may provide significant gains in HIV prevention.
- **A Method to Estimate the Size and Characteristics of HIV-positive Populations Using an Individual-based Stochastic Simulation Model. [14]**
  It is important not only to collect epidemiologic data on HIV but to also fully utilize such information to understand the epidemic over time and to help inform and monitor the impact of policies and interventions. We describe and apply a novel method to estimate the size and characteristics of HIV-positive populations. In the pseudo-epidemic example, HIV estimates have narrower plausibility ranges and are closer to the true number, the greater the data availability to calibrate the model. We demonstrate that our method can be applied to settings with less data, however plausibility ranges for estimates will be wider to reflect greater uncertainty of the data used to fit the model.
- **Immunologic response in treatment-naïve HIV-2-infected patients: the IeDEA West Africa cohort. [9]**
  Response to antiretroviral therapy (ART) among individuals infected with HIV-2 is poorly described. We compared the immunological response among patients treated with three nucleoside reverse-transcriptase inhibitors (NRTIs) to boosted protease inhibitor (PI) and unboosted PI-based regimens in West Africa. In this observational study using African data, boosted PI-containing regimens had better immunological response compared to triple NRTI combinations and unboosted PI-based regimens at 12 months. A randomized clinical trial is still required to determine the best initial regimen for treating HIV-2 infected patients.
- **Intrinsic defect in keratinocyte function leads to inflammation in Hidradenitis suppurativa. [10]**
  Hidradenitis suppurativa (HS) is a chronic, inflammatory, debilitating, follicular disease of the skin. Despite a high prevalence in the general population, the physiopathology of HS remains poorly understood. The use of antibiotics and immunosuppressive agents for therapy suggests a deregulated immune response to microflora. These findings point out a functional defect of keratinocytes in

HS leading to a balance prone to inflammatory responses. This is likely to favor a permissive environment for bacterial infections and chronic inflammation characterizing clinical outcomes in patients with HS.

- **Uptake of Home-Based HIV Testing, Linkage to Care, and Community Attitudes about ART in Rural KwaZulu-Natal, South Africa: Descriptive Results from the First Phase of the ANRS 12249 TasP Cluster-Randomised Trial. [34]**
  The 2015 WHO recommendation of antiretroviral therapy (ART) for all immediately following HIV diagnosis is partially based on the anticipated impact on HIV incidence in the surrounding population. We investigated this approach in a cluster-randomised trial in a high HIV prevalence setting in rural KwaZulu-Natal. We present findings from the first phase of the trial and report on uptake of home-based HIV testing, linkage to care, uptake of ART, and community attitudes about ART. Home-based HIV testing was well received in this rural population, although men were less easily contactable at home; immediate ART was acceptable, with good viral suppression and retention. However, only about half of HIV-positive people accessed care within 6 mo of being identified, with nearly two-thirds accessing care by 12 mo. The observed delay in linkage to care would limit the individual and public health ART benefits of universal testing and treatment in this population.

7.4. Conferences

Members of the team were involved in more than 20 talks during conferences and colloquium. In particular, [20], [21], [23], [19], [22], [24] and [25] have proceedings.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Implication in research for the development of vaccine has lead to a direct contracts with industry such as withs Iliad Biotechnologies. This contract had been signed for the BPZE-1 pertussis vaccine trial. This study evaluates the safety and immunogenicity of a higher dose formulation of a new live attenuated vaccine, BPZE1, intended to prevent Bordetella pertussis nasopharyngeal colonization and pertussis disease, and investigates whether higher doses of BPZE1 induce the live vaccine to colonize subjects' nasopharynx. The study is a Phase Ib (high dose), single centre, dose-escalating, placebo-controlled study of the live attenuated B. pertussis strain BPZE1 given as a single intranasal dose to healthy adult volunteer. This contrat is part of the EUCLID platform (via the CIC 1401) in which Laura Richert and Rodolphe Thiébaut are involved.

8.2. Bilateral Grants with Industry

Implication in research for the development of Ebola vaccine has lead to several indirect contracts with industry:

- The EBOVAC2 project, which is presented in Section 'FP7 & H2020 Projects', leads to collaboration with Janssen from Johnson et Johnson.
- The BPZE-1 pertussis vaccine trial , which is presented in Section 'Bilateral Contracts with Industry', leads to collaboration with Iliad Biotechnologies. (Via the EUCLID platform and CIC 1401)
- The Prevac trial vaccine trial leads to collaboration with Merck. The purpose of this study is to evaluate the safety and immunogenicity of three vaccine strategies that may prevent Ebola virus disease (EVD) events in children and adults. Participants will receive either the Ad26.ZEBOV (rHAd26) vaccine with a MVA-BN-Filo (MVA) boost, or the rVSVΔG-ZEBOV-GP (rVSV) vaccine with or without boosting, or placebo. (Via the EUCLID platform and CIC 1401)

9. Partnerships and Cooperations

9.1. Regional Initiatives

The team have strong links with :

Université de Bordeaux

ISPED (Institut de Santé Publique et du Développement)

Bordeaux CHU ("Centre Hospitalier Universitaire").

Limoges CHU ("Centre Hospitalier Universitaire").

Research teams of the research center INSERM U1219 : "Injury Epidemiology, Transport, Occupation" (IETO), Biostatistics, "Pharmacoepidemiology and population impact of drugs", "Multimorbidity and public health in patients with HIV or Hepatitis" (MORPH3Eus) and "Maladies infectieuses dans les pays à ressources limitées" (IDLIC).

Institut Bergonié, Univ Bordeaux through the EUCLID platform

Inria Project-team MONC and CQFD

9.2. National Initiatives

9.2.1. Labex Vaccine Research Institute (VRI)

There are strong collaborations with immunologists involved in the Labex Vaccine Research Institute (VRI) as Rodolphe Thiébaut is leading the Biostatistics/Bioinformatics division.

9.2.2. Expert Appraisals

Rodolphe Thiébaut is an expert for INCA (Institut National du Cancer) for the PHRC (Programme hospitalier de recherche Clinique en cancérologie) and for the PRME (Programme de recherche médico-économique en cancérologie).

Mélanie Prague is an expert for ANRS (France Recherche Nord&Sud Sida-HIV Hépatites) in the CSS 3 (Recherches cliniques et physiopathologiques dans l'infection à VIH).

Rodolphe Thiébaut is a member of the Membre du CNU 46.04 (Biostatistiques, informatique médicale et technologies de communication).

Laura Richert is an expert for the PHRC (Programme hospitalier de recherche Clinique).

Laura Richert is a member of F-CRIN Steering Committee.

Marta Avalos is an expert for L'ASNM (Agence nationale de sécurité du médicament et des produits de santé)

9.2.3. Various Partnership

The project team members are involved in:

Convention between the "Fédération francaise de natation" and Inria (18950 euros) for the R&D project "Quels schémas de périodisation pour la préparation des Jeux Olympiques à Rio ?" (Marta Avalos).

DRUGS-SAFE platform funded by ANSM (Marta Avalos).

F-CRIN (French clinical research infrastructure network) was initiated in 2012 by ANR under a PIA founding (Programme des Investissements d'avenir) named "INBS/Infrastructures nationales en biologie et en santé". (Laura Richert)

The project team members also collaborate with:

I-REIVAC is the French vaccine research network. This network is part of the Consortium de Recherche en Vaccinologie (CoReVac) created by the Institut de Microbiologie et des Maladies Infectieuses (IMMI). (Laura Richert)

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

The member of SISTM Team are involved in EHVA (European HIV Vaccine Alliance):

Program: Most information about this program can be found at http://www.ehv-a.eu/.

Coordinator: Rodolphe Thiébaut is Work Package leader of the WP10 "Data Integration".

Other partners: The EHVA encompasses 39 partners, each with the expertise to promote a comprehensive approach to the development of an effective HIV vaccine. The international alliance, which includes academic and industrial research partners from all over Europe, as well as sub-Saharan Africa and North America, will work to discover and progress novel vaccine candidates through the clinic.

Abstract: With 37 million people living with HIV worldwide, and over 2 million new infections diagnosed each year, an effective vaccine is regarded as the most potent public health strategy for addressing the pandemic. Despite the many advances in the understanding, treatment and prevention of HIV made over the past 30 years, the development of broadly-effective HIV vaccine has remained unachievable. EHVA plans to develop and implement:

- Discovery Platform with the goal of generating novel vaccine candidates inducing potent neutralizing and non-neutralizing antibody responses and T-cell responses
- Immune Profiling Platform with the goal of ranking novel and existing (benchmark) vaccine candidates on the basis of the immune profile
- Data Management/Integration/Down-Selection Platform, with the goal of providing statistical tools for the analysis and interpretation of complex data and algorithms for the efficient selection of vaccines
- Clinical Trials Platform with the goal of accelerating the clinical development of novel vaccines and the early prediction of vaccine failure.

The member of SISTM Team and particularly Laura Richert are also involved in other H2020 projects such as SenseCog, Medit'aging and Orthunion.

9.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: The EBOVAC2 project is one of 8 projects funded under IMI Ebola+ programme that was launched in response to the Ebola virus disease outbreak. The project aims to assess the safety and efficacy of a novel prime boost preventive vaccine regimen against Ebola Virus Disease (EVD).

Project acronym: EBOVAC2

Project title: EBOVAC2

Coordinator: Rodolphe Thiébaut

Other partners: Inserm (France), Labex VRI (France), Janssen Pharmaceutical Companies of Johnson & Johnson, London School of Hygiene & Tropical Medicine (United Kingdom), The Chancellor, Masters and Scholars of the University of Oxford (United Kingdom), Le Centre Muraz (Burknia Faso), Inserm Transfert (France)

Abstract: Given the urgent need for an preventive Ebola vaccine strategy in the context of the current epidemic, the clinical development plan follows an expedited scheme, aiming at starting a Phase 2B large scale safety and immunogenicity study as soon as possible while assuring the safety of the trial participants.

- Phase 1 trials to assess the safety and immunogenicity data of the candidate prime-boost regimen in healthy volunteers are ongoing in the UK, the US and Kenya and Uganda. A further study site has been approved to start in Tanzania. Both prime-boost combinations (Ad26.ZEBOV prime + MVA-BN-Filo boost; and MVA-BN-Filo prime + Ad26.ZEBOV boost) administered at different intervals are being tested in these trials.

- Phase 2 trials (this project) are planned to start as soon as the post-prime safety and immunogenicity data from the UK Phase I are available. Phase 2 trials will be conducted in healthy volunteers in Europe (France and UK) and non-epidemic African countries (to be determined). HIV positive adults will also be vaccinated in African countries. The rationale for inclusion of European volunteers in Phase 2, in addition to the trials in Africa, is to allow for higher sensitivity in safety signal detection in populations with low incidence of febrile illnesses, to generate negative control specimens for assay development, to allow for inclusion of health care workers or military personnel that may be deployed to Ebola-endemic regions.

9.3.3. Collaborations with Major European Organizations

University of Oxford;

London School of Hygiene and Tropical Medicine;

University Hospital Hamburg;

Heinrich Pette Institute for Experimental Virology, Hambourg;

MRC, University College London

9.4. International Initiatives

Scharp, Seattle;

Fred Hutchinson Cancer center, Seattle;

Baylor Institute;

NIH for the Prevac trial;

NGO Alima for the Prevac trial;

Several African clinical sites for Ebovac2 and Prevac trials.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Cristian Meza, Associate Professor of the Universidad de Valparaiso (Chili), member of the research center CIMFAV : http://cmeza.cimfav.cl/ collaborates on the project entitled "Longitudinal high-dimensional data" (septembre)

David Conesa, Associate Professor of the Universidad de Valencia (Espagne), member of the research group GEEITEMA : http://www.geeitema.org/conesa/ collaborates on the project entitled "Bayesian predictive methods with application to the home and leisure injuries in France study MAVIE" (septembre)

Sam Doerken, PhD student of the University of Freiburg (Allemagne), member of the Institute for Medical Biometry and Statistics : http://portal.uni-freiburg.de/imbi/employees/persons/doerken collaborates on the project entitled "Penalization regression methods for sparse exposures with application to pharmacoepidemiology" (septembre - octobre)

Jessica Gronsbell, PhD student of the Harvard T.H. Chan School of Public Health, came as a visiting scholar on a subject of "analysis of high dimensional genetic data" (May).

9.5.2. Visits to International Teams

Marta Avalos will be a research visitor at CSIRO's Data61 in Canberra, Australia from Dec. 2016 until June 2017. Collaboration with Cheng Soon Ong http://www.ong-home.my/

Marta Avalos (in April and october) visited David Conesa through the Erasmus+ program Universidad de Valencia (Espagne).

Perrine Soret (from 26/12/15 to 28/01/16) visited Cristian Meza and Karine Bertin (Inria Chili) at CIMFAV (Centre for Research and Modeling of Random Phenomena, Valparaíso), Univ Valparaiso, Chili, concerning the project "New challenges in mixed-effects models".

Laura Richtert spent 6 months as visiting researcher at Heinrich Pette Institut for experimental virology, department virus immunology (Pr M. Altfeld), Hamburg Germany in 2016

Boris Hejblum is a Visiting Scientist appointment at Harvard University (not paid), Department of Biostatistics

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

Daniel Commenges organised a SFB (Société Française de Biometrie) in Montpellier (3 Juin 2016),

Daniel Commenges Co-organised the "Journées GDR-SFB" in Lyon (27-28 Juin)

Robin Genuer Co-organised a reading group called Smiling in Bordeaux (http://www.math.u-bordeaux.fr/~machaven/smiling)

Rodolphe Thiébaut organized the scientific program of the Bordeaux Modelling Workshop (1 et 2 Juin 2016)

Rodolphe Thiébaut organised a 10 hours seminar on "Bayesian filters and particule methods" at Inria Bordeaux (Nov. and Dec. 2016)

10.1.2. General Chair, Scientific Chair

Rodolphe Thiébaut is a member of the scientific committee of the Muraz Center (Bobo-Dioulasso, Burkina Faso), since 2016

Rodolphe Thiébaut is a member of the Scientific Advisory Board de l'Institut Pierre Louis d'Epidémiologie et de Santé Publique (UPMC, Dir : Dominique Costagliola), since 2015

Daniel Commenges is a member of the scientific committee of de "Journée de la SFdS (Société Francaise de Statistiques)" (Montpellier, 30 Mai-3 Juin)

10.1.3. Member of the Organizing Committees

All the team members helped in the ground organisation of the Bordeaux Modelling Workshop

10.1.4. Member of the Journal Editorial Boards

Lifetime Data Analysis (Daniel Commenges)

Statistics Surveys (Daniel Commenges)

Journal de la Société Francaise de Statistique (Daniel Commenges)

Daniel Commenges DC Principales revues de Statistique (Biometrics, JASA, JRSS, Stat Med, LIDA,...)

10.1.5. Reviewer - Reviewing Activities

AIDS (Rodolphe Thiébaut)

Annals of Applied Statistics (Boris Hejblum)

BioData Mining (Boris Hejblum)

Biometrics (Daniel Commenges, Mélanie Prague)

International Journal of Biostatistics (Robin Genuer)

International Journal of Epidemiology (Daniel Commenges)

Journal of Applied Statistics (Marta Avalos)

Journal of the Royal Statistical Society: Interaction (Mélanie Prague)

Machine Learning (Robin Genuer)

Neural Information Processing Systems (Robin Genuer)

Pattern Recognition Letters (Robin Genuer)

Statistics and Computing (Robin Genuer)

Statistical Methods and Applications (Marta Avalos)

Statistics in Medicine (Daniel Commenges, Rodolphe Thiébaut, Mélanie Prague)

Statistics Surveys (Daniel Commenges)

10.1.6. Invited Talks

Daniel Commenges gave 3 invited talks in Vienne (19 Mai), Vigo (26 Ocotbre) and Berlin (25 Novembre).

Laura Richert gave 2 invited talks in Webinar about "big data in epidemiology" (20 juin) and in Paris for the Colloquium "One Health" about "Signature transcriptomique post-vaccinale chez l'Homme" (3 Novembre).

Rodolphe Thiébaut gave 3 invited talks

Mélanie Prague gave 2 invited talks one in Nancy (20 fev.) and one un Summer Sim (26 july).

10.1.7. Leadership within the Scientific Community

Daniel Commenges is the president of the SFB (Société Française de Biométrie) which is the French satellite for the IBS (International Biometrics society).

10.1.8. Research Administration

Daniel Commenges is the director of the Biostat-Info axis in the Inserm BPH (Bordeaux Public Health) institute.

Rodolphe Thiébaut is a member of the department of life science in University of Bordeaux

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

In class teaching

Master : Marta Avalos teaches in the two years of the Master of Public Health at ISPED, Univ. Bordeaux, France.

Master : Robin Genuer, teaches in the two years of the Master of Public Health (M1 Santé publique, M2 Biostatistique, M2 Informatique médicale, M2 Santé internationale, M2 épidémiologie).

Master : Boris Hejblum, teaches in the two years of the Master of Public Health (M1 Santé publique, M2 Biostatistique, M2 Informatique médicale, M2 Santé internationale, M2 épidémiologie).

Master : Robin Genuer, MSS du collège ST, intervention dans le cours de Statistique en grande dimension.

Master : Rodolphe Thiébaut, teaches in the two years of the Master of Public Health, and he is head of the Epidemiology specialty of the second year of the Master of Public Health.

Master : Laura Richert teaches in the two years of the Master of Public Health at ISPED, Univ. Bordeaux, France (M2 Biostatistiques).

Master : Laura Richert teaches in the Master of Vaccinology at UPEC (University Paris-Est-Créteil), France.

Master : Chloe Pasin is a teaching assistant for the two years of the Master of Public Health at ISPED, Univ. Bordeaux, France.

Master : Laura Villain is a teaching assistant for the two years of the Master of Public Health at ISPED, Univ. Bordeaux, France

Bachelor : Laura Richert teaches in PACES and DFASM1-3 for Medical degree at Univ. Bordeaux, France

Summer School: All the SISTM team member teach in the ISPED Summer school.

E-learning

Marta Avalos is head of the first year of the e-learning program of the Master of Public Health, and teaches in it.

Mélanie Prague teaches in the Diplôme universitaire "Méthodes statistiques de régression en épidémiologie".

Laura Richert teaches in the Diplôme universitaire "Recherche Clinique".

Rodolphe Thiébaut is head of the Epidemiology specialty of the second year of the e-learning program of the Master of Public Health, and teaches in it.

Robin Genuer and Perrine Soret participate to the IdEx Bordeaux University "Défi numérique" project "BeginR" (http://beginr.moutault.net/).

10.2.2. Supervision

PhD in progress : Wenjia Wang "Modèle de Rasch" (CIFRE, co-direction avec Mickael Guedj Pharnext), from Oct 2015, directed by Daniel Commenges.

PhD in progress : Laura Villain "Modélisation de l'effet du traitement par injection IL7" (CIFRE, co-direction avec Rodolphe Thiébaut), from Oct 2015, directed by Daniel Commenges.

PhD in progress : Perrine Soret, *Modélisation de données longitudinales en grande dimension*, from Oct 2014, directed by Marta Avalos.

PhD in progress : Mélanie Née *Recherche et caractérisation de profils attentionnels : mieux comprendre la place de l'attention dans la survenue des accidents de la vie courante*, from Oct 2015, co-directed by Emmanuel Lagarde (60%), Cédric Galera (20%), Marta Avalos (20%)

PhD in progress : Chloé Pasin, *Modelling the immune response to HIV vaccine*, from Sep 2015, co-directed by Rodolphe Thiébaut and Francois Dufour

PhD in progress : Edouard Lhomme, *Analyse des déterminants de la réponse immunitaire post-vaccination dans des stratégies vaccinales expérimentales*, from Oct 2016, directed by Laura Richert.

PhD in progress : Hadrien Lorenzo, *Analyses de données longitudinales de grandes dimensions appliquées aux essais vaccinaux contre le VIH et Ebola*, from Oct 2016, co-directed by Rodolphe Thiébaut and Jérôme Saracco.

Master internship : Hao Ren "Contribution au développement d'un outil statistique d'aide à la décision en sport de haut niveau", directed by Marta Avalos and Perrine Soret (01/03/2016-12/08/2016)

Master internship : Madelyn Rojas "Practices for the provision of prior information in Bayesian Logistic Regression: Application in MAVIE project", directed by Marta Avalos and David Conesa (11/07/2016-09/09/2016)

Master internship : Thomas Blondel "Application of Bayesian linear models to sports science data", directed by David Conesa and Marta Avalos (05/04/2016 - 04/06/2016)

Master internship : Julie Havas "Application of Bayesian Logistic Regression to the mavie study of home and leaisure injury", directed by David Conesa and Marta Avalos (05/04/2016 - 04/06/2016)

Master internship : Thomas Esnaud "Etude de la méthode de clustering par forêts aléatoires, applications à la reconnaissance automatique de populations cellulaires.", directed by Robin Genuer (14/03/2016 - 31/08/2016)

Master internship : Lise Mandigny "Revue systématique et Méta-analyse des essais cliniques publiés de développement de vaccins contre le virus Ebola", directed by Rodolphe Thiébaut (1/04/2016 - 31/09/2016)

Master internship : Stella Huang "Modélisation de l'infection à pseudomonas aeruginosa dans les services de réanimation ? étude DYNAPYO", directed by Rodolphe Thiébaut (11/02/2016 - 15/08/2016)

Master project : B Dufoyer, A Chevalier, H Aassif, A Labchri, projet de programmation du Master 1 Informatique, Univ Bordeaux. Titre : " Développement d'un outil de prévention des accidents de la vie courante à partir de méthodes de machine learning : site web et bases de données ", directed by Marta Avalos and L Orriols, M Travanca, L Divert, INSERM U1219. (11/01/2016-12/04/2016)

Master project : N Craeye, C Elassaoui, F Elouazi, B Faltrept, projet de programmation du Master 1 Informatique, Univ Bordeaux. Titre : " Développement d'un outil de mesure de l'attention via internet ", directed by Marta Avalos and M Née, L Divert, E Lagarde, INSERM U1219. (11/01/2016-12/04/2016)

10.2.3. Juries

Daniel Commenges was involved in two PhD defences as president of the jury: Leila Azarang (Vigo), Anais Rouanet (Bordeaux).

Robin Genuer was in charge of the reports of the PhD of Havelund Welling, entitled "Characterization of absorption enhancers for orally administered therapeutic peptides in tablet formulations", defended on 30/09/2016 in Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby

Mélanie Prague is a member of the follow-up dissertation comity of Sébastien Benzkcry's PhD student (Inria Bor- deaux Sud-ouest, MONC team). Nicolo Chiara is working on "Mathematical modeling of systemic aspects of cancer and cancer therapy".

Rodolphe Thiébaut took part in the HDR committee of Vivian Viallon (2016) and Francesco Salvo (2016)

Robin Genuer took part in the recruitment commission MCF CNU 26 (Toulouse 2016)

Rodolphe Thiébaut took part in the recruitment commissions PU CNU 26 (Paris Descartes 2016), MCF CNU 26 (Bordeaux 2016), MCF CNU 85 (Bordeaux 2016).

10.3. Popularization

Marta Avalos, Marius Kwémou and Perrine Soret animated "Mais qui est le coupable ? (Ou comment les maths contribuent à conduire une enquête épidémiologique)" for high school students through the "Fête de la Science" organized at Inria, Oct 2016.

Laura Richert participated to "Nuit Européenne des Chercheurs" with speed dating and a radio interview, Cap Sciences, Bordeaux, September 2016.

11. Bibliography

Major publications by the team in recent years

[1] R. Bouteloup, R. Sabin, R. Mocroft, M. Gras, M. Pantazis, M. Le Moing, M. d'Arminio Monforte, R. Mary-Krause, R. Roca, C. Miro, R. Battegay, R. Brockmeyer, M. Berenguer, R. Morlat, C. Obel, L. De Wit, R. Fätkenheuer, R. Zangerle, C. Ghosn, R. Pérez-Hoyos, R. Campbell, C. Prins, R. Chêne, R. Meyer, R. Dorrucci, C. Torti, R. Thiébaut. *Reference curves for CD4 T-cell count response to combination antiretroviral therapy in HIV-1-infected treatment-naïve patients*, in "HIV Medicine", September 2016, https://hal.inria.fr/hal-01406622

[2] C. C. Iwuji, J. Orne-Gliemann, J. Larmarange, N. Okesola, F. Tanser, R. Thiebaut, C. Rekacewicz, M.-L. Newell, F. Dabis, N. Low. *Uptake of Home-Based HIV Testing, Linkage to Care, and Community Attitudes about ART in Rural KwaZulu-Natal, South Africa: Descriptive Results from the First Phase of the ANRS 12249 TasP Cluster-Randomised Trial*, in "PLoS Medicine", August 2016, vol. 13, n° 8, https://hal.inria.fr/hal-01406616

[3] A. Jarne, D. Commenges, L. Villain, M. Prague, Y. Lévy, R. Thiébaut. *Modeling CD4+ T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous Interleukin 7*, in "Submitted in Annals of Applied statistics", Sept 2016, vol. X, n° X

[4] R. Le Grand, N. Dereuddre-Bosquet, S. Dispinseri, L. Gosse, D. Desjardins, X. Shen, M. Tolazzi, C. Ochsenbauer, H. Saidi, G. Tomaras, M. Prague, S. W. Barnett, R. Thiebaut, A. Cope, G. Scarlatti, R. J. Shattock, G. Silvestri. *Superior Efficacy of a Human Immunodeficiency Virus Vaccine Combined with Antiretroviral Prevention in Simian-Human Immunodeficiency Virus-Challenged Nonhuman Primates*, in "Journal of Virology", May 2016, vol. 90, n° 11, pp. 5315 - 5328, https://hal.inria.fr/hal-01406364

[5] P. Schommers, G. Martrus, U. Matschl, M. Sirignano, M. Lütgehetmann, L. Richert, T. J. Hope, G. Fätkenheuer, M. Altfeld, F. Kirchhoff. *Changes in HIV-1 Capsid Stability Induced by Common Cytotoxic-T-Lymphocyte-Driven Viral Sequence Mutations*, in "Journal of Virology", July 2016, vol. 90, n° 16, pp. 7579 - 7586 [*DOI :* 10.1128/JVI.00867-16], https://hal.archives-ouvertes.fr/hal-01404506

[6] M. Surenaud, C. Manier, L. Richert, R. Thiébaut, Y. Levy, S. Hue, C. Lacabaratz. *Optimization and evaluation of Luminex performance with supernatants of antigen-stimulated peripheral blood mononuclear cells*, in "BMC Immunology", December 2015, vol. 17, n° 1, 44 p. [*DOI :* 10.1186/s12865-016-0182-8], http://www.hal.inserm.fr/inserm-01395760

[7] R. Thiébaut, J. Drylewicz, M. Prague, C. Lacabaratz, S. Beq, A. Jarne, T. Croughs, R.-P. Sekaly, M. M. Lederman, I. Sereti. *Quantifying and Predicting the Effect of Exogenous Interleukin-7 on CD4+ T Cells in HIV-1 Infection*, in "PLoS computational biology", 2014, vol. 10, n° 5, e1003630

Publications of the year

Articles in International Peer-Reviewed Journals

[8] S. ARLOT, R. GENUER. *Comments on: "A Random Forest Guided Tour" by G. Biau and E. Scornet*, in "Test", 2016, vol. 25, n^{o} 2, pp. 228–238, The final publication is available at Springer: http://dx.doi.org/10.1007/s11749-016-0484-4 [*DOI :* 10.1007/s11749-016-0484-4], https://hal.archives-ouvertes.fr/hal-01297557

[9] E. BALESTRE, D. K. EKOUEVI, B. TCHOUNGA, S. P. EHOLIE, E. MESSOU, A. SAWADOGO, R. THIÉBAUT, M. T. MAY, J. AC STERNE, F. DABIS. *Immunologic response in treatment-naïve HIV-2-infected patients: the IeDEA West Africa cohort*, in "Journal of the International AIDS Society", February 2016, vol. 19, n^{o} 1 [*DOI :* 10.7448/IAS.19.1.20044], https://hal.inria.fr/hal-01288896

[10] C. HOTZ, M. BONIOTTO, A. GUGUIN, M. SURENAUD, F. JEAN-LOUIS, P. TISSERAND, N. ORTONNE, B. HERSANT, R. BOSC, F. POLI, H. BONNABAU, R. THIÉBAUT, V. GODOT, P. WOLKENSTEIN, H. HOCINI, Y. LÉVY, S. HÜE. *Intrinsic Defect in Keratinocyte Function Leads to Inflammation in Hidradenitis Suppurativa*, in "Journal of Investigative Dermatology", September 2016, vol. 136, n^{o} 9, pp. 1768 - 1780, https://hal.inria.fr/hal-01406612

[11] H. KAMINSKI, I. GARRIGUE, L. COUZI, B. TATON, T. BACHELET, J.-F. MOREAU, J. DECHANET-MERVILLE, R. THIEBAUT, P. MERVILLE. *Surveillance of T Cells Predicts Cytomegalovirus Infection Resolution in Kidney Transplants*, in "Journal of the American Society of Nephrology", 2016, vol. 27, n^{o} 2, pp. 637-645 [*DOI :* 10.1681/ASN.2014100985], https://hal.inria.fr/hal-01288893

[12] E. LHOMME, L. RICHERT, Z. MOODIE, C. PASIN, S. A. KALAMS, C. MORGAN, S. SELF, S. C. DE ROSA, R. THIÉBAUT. *Early CD4+ T Cell Responses Are Associated with Subsequent CD8+ T Cell Responses to an rAd5-Based Prophylactic Prime-Boost HIV Vaccine Strategy*, in "PLoS ONE", April 2016, vol. 11, n^{o} 4, https://hal.inria.fr/hal-01406611

[13] B. LIQUET, P. LAFAYE DE MICHEAUX, B. P. HEJBLUM, R. THIÉBAUT. *Group and sparse group partial least square approaches applied in genomics context*, in "Bioinformatics", 2016, vol. 32, n^{o} 1, pp. 35-42 [*DOI :* 10.1093/BIOINFORMATICS/BTV535], https://hal.inria.fr/hal-01288891

[14] F. NAKAGAWA, A. VAN SIGHEM, R. THIEBAUT, C. SMITH, O. RATMANN, V. CAMBIANO, J. ALBERT, A. AMATO-GAUCI, D. BEZEMER, C. CAMPBELL, D. COMMENGES, M. DONOGHOE, D. FORD, R. KOUYOS, R. LODWICK, J. LUNDGREN, N. PANTAZIS, A. PHARRIS, C. QUINTEN, C. THORNE, G. TOULOUMI, V. DELPECH, A. PHILLIPS. *A method to estimate the size and characteristics of HIV-positive populations using an individual-based stochastic simulation model*, in "Epidemiology", 2016, vol. 27, n^{o} 2, pp. 247-256 [*DOI :* 10.1097/EDE.0000000000000423], https://hal.inria.fr/hal-01288889

[15] M. PRAGUE, D. COMMENGES, J. M. GRAN, B. LEDERGERBER, J. YOUNG, H. FURRER, R. THIÉBAUT. *Dynamic models for estimating the effect of HAART on CD4 in observational studies: Application to the Aquitaine Cohort and the Swiss HIV Cohort Study*, in "Biometrics", July 2016, https://hal.inria.fr/hal-01406614

[16] M. PRAGUE. *Use of dynamical models for treatment optimization in HIV infected patients : a sequential Bayesian analysis approach.* , in "Journal de la Société Française de Statistiques", September 2016, vol. 157, n^{o} 2, 20 p. , https://hal.inria.fr/hal-01416102

[17] R. THIÉBAUT, A. JARNE, J.-P. ROUTY, I. SERETI, M. FISCHL, P. IVE, R. F. SPECK, G. D'OFFIZI, S. CASARI, D. COMMENGES, S. FOULKES, V. NATARAJAN, T. CROUGHS, J.-F. DELFRAISSY, G. TAMBUSSI, Y. LEVY, M. M. LEDERMAN. *Repeated Cycles of Recombinant Human Interleukin 7 in HIV-Infected Patients*

With Low CD4 T-Cell Reconstitution on Antiretroviral Therapy: Results of 2 Phase II Multicenter Studies, in "Clinical Infectious Diseases", May 2016, vol. 62, n° 9, pp. 1178 - 1185, https://hal.inria.fr/hal-01406362

[18] G. ZURAWSKI, S. ZURAWSKI, A.-L. FLAMAR, L. RICHERT, R. WAGNER, G. D. TOMARAS, D. C. MONTEFIORI, M. ROEDERER, G. FERRARI, C. LACABARATZ, H. BONNABAU, P. KLUCAR, Z. WANG, K. E. FOULDS, S.-F. KAO, N. L. YATES, C. LABRANCHE, B. L. JACOBS, K. KIBLER, B. ASBACH, A. KLICHE, A. SALAZAR, S. REED, S. SELF, R. GOTTARDO, L. GALMIN, D. WEISS, A. CRISTILLO, R. THIEBAUT, G. PANTALEO, Y. LEVY. *Targeting HIV-1 Env gp140 to LOX-1 Elicits Immune Responses in Rhesus Macaques*, in "PLoS ONE", April 2016, vol. 11, n° 4 [*DOI :* 10.1371/JOURNAL.PONE.0153484], https://hal.inria.fr/hal-01389087

International Conferences with Proceedings

[19] M. F. AVALOS, M. LE GOFF, P. JOLY, M.-A. JUTAND, A. ALIOUM. *Evolution of Teaching Strategies in a French ODL University Course*, in "9th Australian Conference on Teaching Statistics (OZCOTS)", Canberra, Australia, H. MACGILLIVRAY, M. MARTIN, B. PHILLIPS (editors), Proceedings of the Ninth Australian Conference on Teaching Statistics, The International Association for Statistical Education (IASE), December 2016, 183 p. , ISBN: 978-0-9805950-2-4, https://hal.inria.fr/hal-01396772

[20] M. F. AVALOS, P. SORET, C. MEZA, K. BERTIN, H. REN, P. HELLARD. *Exploring variable selection in additive mixed effects models using group lasso*, in "23rd Australian Statistical Conference", Canberra, Australia, Statistical Society of Australia (SSA), December 2016, https://hal.inria.fr/hal-01396767

[21] G. LEFORT, M. F. AVALOS, P. SORET, P. DAVID, J. F. TOUSSAINT, P. HELLARD. *A comparison of unsupervised curve classification methods for sport training data*, in "3rd conference of the International Society for Non-Parametric Statistics (ISNPS)", Avignon, France, The International Society for NonParametric Statistics (ISNPS), June 2016, https://hal.inria.fr/hal-01396366

[22] G. LEFORT, M. F. AVALOS, P. SORET, P. DAVID, J. F. TOUSSAINT, P. N. HELLARD. *Clustering of temporal sport training curves: a comparison of FDA and LDA approaches*, in "23rd Australian Statistical Conference", Canberra, Australia, December 2016, https://hal.inria.fr/hal-01396372

[23] B. SILENOU CHAWO, M. F. AVALOS, A. PARIENTE, H. JACQMIN-GADDA. *Adjustment for Unobserved Confounders in Health Administrative Databases*, in "32nd International Conference on Pharmacoepidemiology & Therapeutic Risk Management", Dublin, Ireland, INTERNATIONAL SOCIETY FOR PHARMACOEPIDEMIOLOGY (ISPE), August 2016, https://hal.inria.fr/hal-01396349

[24] P. SORET, C. MEZA, M. F. AVALOS, K. BERTIN, R. THIÉBAUT. *Lasso-type estimators for non-parametric mixed-effects models: application to high-dimensional data from a vaccine clinical trial for HIV*, in "3rd conference of the International Society for Non-Parametric Statistics (ISNPS)", Avignon, France, June 2016, https://hal.inria.fr/hal-01396378

National Conferences with Proceedings

[25] P. SORET, C. MEZA, M. AVALOS, K. BERTIN. *Estimateur de type Lasso pour modèle mixte non-paramétrique*, in "48èmes Journées de Statistique", Montpellier, France, Société Française de Statistique (SFdS), May 2016, https://hal.inria.fr/hal-01396802

Other Publications

[26] M. CHAVENT, R. GENUER, J. SARACCO. *Combining clustering of variables and feature selection using random forests: the CoV/VSURF procedure*, July 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01345840

[27] R. GENUER, J.-M. POGGI. *Arbres CART et Forêts aléatoires,Importance et sélection de variables*, October 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01387654

References in notes

[28] O. AALEN, K. KJETIL ROYSLAND, J. GRAN, B. LEDERGERBER. *Causality, mediation and time: a dynamic viewpoint*, in "Journal of the Royal Statistical Society: Series A (Statistics in Society)", 2007, vol. 175, n^{o} 4, pp. 831–861

[29] R. BOUTELOUP, R. SABIN, R. MOCROFT, M. GRAS, M. PANTAZIS, M. LE MOING, M. D'ARMINIO MONFORTE, R. MARY-KRAUSE, R. ROCA, C. MIRO, R. BATTEGAY, R. BROCKMEYER, M. BERENGUER, R. MORLAT, C. OBEL, L. DE WIT, R. FÄTKENHEUER, R. ZANGERLE, C. GHOSN, R. PÉREZ-HOYOS, R. CAMPBELL, C. PRINS, R. CHÊNE, R. MEYER, R. DORRUCCI, C. TORTI, R. THIÉBAUT. *Reference curves for CD4 T-cell count response to combination antiretroviral therapy in HIV-1-infected treatment-naïve patients*, in "HIV Medicine", September 2016, https://hal.inria.fr/hal-01406622

[30] F. CASTIGLIONE, B. PICCOLI. *Cancer immunotherapy, mathematical modeling and optimal control*, in "Biometrical Journal", 2007, vol. 247, n^{o} 4, pp. 723-32

[31] R. M. GRANICH, C. F. GILKS, C. DYE, K. M. DE COCK, B. G. WILLIAMS. *Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model*, in "Lancet", 2009, vol. 373, n^{o} 9657, pp. 48-57, 0140 6736 English

[32] L. HOOD, Q. TIAN. *Systems approaches to biology and disease enable translational systems medicine*, in "Genomics Proteomics Bioinformatics", 2012, vol. 10, n^{o} 4, pp. 181–5

[33] Y. HUANG, D. LIU, H. WU. *Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system*, in "Biometrics", 2006, vol. 62, n^{o} 2, pp. 413–423

[34] C. C. IWUJI, J. ORNE-GLIEMANN, J. LARMARANGE, N. OKESOLA, F. TANSER, R. THIEBAUT, C. REKACEWICZ, M.-L. NEWELL, F. DABIS, N. LOW. *Uptake of Home-Based HIV Testing, Linkage to Care, and Community Attitudes about ART in Rural KwaZulu-Natal, South Africa: Descriptive Results from the First Phase of the ANRS 12249 TasP Cluster-Randomised Trial*, in "PLoS Medicine", August 2016, vol. 13, n^{o} 8, https://hal.inria.fr/hal-01406616

[35] A. JARNE, D. COMMENGES, L. VILLAIN, M. PRAGUE, Y. LÉVY, R. THIÉBAUT. *Modeling CD4+ T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous Interleukin 7*, in "Submitted in Annals of Applied statistics", Sept 2016, vol. X, n^{o} X

[36] E. KUHN, M. LAVIELLE. *Maximum likelihood estimation in nonlinear mixed effects models*, in "Computational Statistics & Data Analysis", 2005, vol. 49, n^{o} 4, pp. 1020–1038

[37] K.-A. LE CAO, P. MARTIN, C. ROBERT-GRANIÉ, P. BESSE. *Sparse canonical methods for biological data integration: application to a cross-platform study*, in "BMC bioinformatics", 2009, vol. 10, 34 p.

[38] R. LE GRAND, N. DEREUDDRE-BOSQUET, S. DISPINSERI, L. GOSSE, D. DESJARDINS, X. SHEN, M. TOLAZZI, C. OCHSENBAUER, H. SAIDI, G. TOMARAS, M. PRAGUE, S. W. BARNETT, R. THIEBAUT, A. COPE, G. SCARLATTI, R. J. SHATTOCK, G. SILVESTRI. *Superior Efficacy of a Human Immunodeficiency Virus Vaccine Combined with Antiretroviral Prevention in Simian-Human Immunodeficiency Virus-Challenged Nonhuman Primates*, in "Journal of Virology", May 2016, vol. 90, n° 11, pp. 5315 - 5328, https://hal.inria.fr/hal-01406364

[39] C. LEWDEN, D. SALMON, P. MORLAT, S. BEVILACQUA, E. JOUGLA, F. BONNET, L. HERIPRET, D. COSTAGLIOLA, T. MAY, G. CHÊNE. *Causes of death among human immunodeficiency virus (HIV)-infected adults in the era of potent antiretroviral therapy: emerging role of hepatitis and cancers, persistent role of AIDS*, in "International Journal of Epidemiology", 2005, vol. 34, n° 1, pp. 121-130, 0300 5771 English

[40] A. S. PERELSON, A. U. NEUMANN, M. MARKOWITZ, J. M. LEONARD, D. D. HO. *HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time*, in "Science", 1996, vol. 271, n° 5255, pp. 1582-6

[41] A. S. PERELSON. *Modelling viral and immune system dynamics*, in "Nature Reviews Immunology", 2002, vol. 2, n° 1, pp. 28-36

[42] J. PINHEIRO, D. BATES. *Approximations to the log-likelihood function in the nonlinear mixed-effects model*, in "Journal of Computational and Graphical Statistics", 1995, vol. 4, n° 1, pp. 12–35

[43] B. PULENDRAN. *Learning immunology from the yellow fever vaccine: innate immunity to systems vaccinology*, in "Nature Reviews Immunology", 2009, vol. 9, n° 10, pp. 741-7

[44] H. PUTTER, S. HEISTERKAMP, J. LANGE, F. DE WOLF. *A Bayesian approach to parameter estimation in HIV dynamical models*, in "Statistics in Medicine", 2002, vol. 21, n° 15, pp. 2199–2214

[45] A. REINER, D. YEKUTIELI, Y. BENJAMINI. *Identifying differentially expressed genes using false discovery rate controlling procedures*, in "Bioinformatics", 2003, vol. 19, n° 3, pp. 368–375

[46] P. SCHOMMERS, G. MARTRUS, U. MATSCHL, M. SIRIGNANO, M. LÜTGEHETMANN, L. RICHERT, T. J. HOPE, G. FÄTKENHEUER, M. ALTFELD, F. KIRCHHOFF. *Changes in HIV-1 Capsid Stability Induced by Common Cytotoxic-T-Lymphocyte-Driven Viral Sequence Mutations*, in "Journal of Virology", July 2016, vol. 90, n° 16, pp. 7579 - 7586 [*DOI :* 10.1128/JVI.00867-16], https://hal.archives-ouvertes.fr/hal-01404506

[47] C. SCHUBERT. *Systems immunology: complexity captured*, in "Nature", 2011, vol. 473, n° 7345, pp. 113-4

[48] M. SURENAUD, C. MANIER, L. RICHERT, R. THIÉBAUT, Y. LEVY, S. HUE, C. LACABARATZ. *Optimization and evaluation of Luminex performance with supernatants of antigen-stimulated peripheral blood mononuclear cells*, in "BMC Immunology", December 2015, vol. 17, n° 1, 44 p. [*DOI :* 10.1186/s12865-016-0182-8], http://www.hal.inserm.fr/inserm-01395760

[49] R. THIÉBAUT, J. DRYLEWICZ, M. PRAGUE, C. LACABARATZ, S. BEQ, A. JARNE, T. CROUGHS, R.-P. SEKALY, M. M. LEDERMAN, I. SERETI. *Quantifying and Predicting the Effect of Exogenous Interleukin-7 on CD4+ T Cells in HIV-1 Infection*, in "PLoS computational biology", 2014, vol. 10, n° 5, e1003630

[50] R. THIÉBAUT, B. HEJBLUM, L. RICHERT. *[The analysis of "Big Data" in clinical research.]*, in "Epidemiology and Public Health / Revue d'Epidémiologie et de Santé Publique", January 2014, vol. 62, n^{o} 1, pp. 1–4 [*DOI :* 10.1016/J.RESPE.2013.12.021], http://www.hal.inserm.fr/inserm-00933691

[51] R. THIÉBAUT, H. JACQMIN-GADDA, A. BABIKER, D. COMMENGES. *Joint modelling of bivariate longitudinal data with informative dropout and left-censoring, with application to the evolution of CD4+cell count and HIV RNA viral load in response to treatment of HIV infection*, in "Statistics in Medicine", 2005, vol. 24, n^{o} 1, pp. 65-82

[52] R. TIBSHIRANI. *Regression shrinkage and selection via the lasso*, in "Journal of the Royal Statistical Society: Series B (Statistical Methodology)", 1996, vol. 58, pp. 267–288

[53] Y. WANG. *Derivation of various NONMEM estimation methods*, in "Journal of Pharmacokinetics and pharmacodynamics", 2007, vol. 34, n^{o} 5, pp. 575–593

[54] H. WU. *Statistical methods for HIV dynamic studies in AIDS clinical trials*, in "Statistical Methods in Medical Research", 2005, vol. 14, n^{o} 2, pp. 171–192