IN PARTNERSHIP WITH:
INSERM

Université de Bordeaux

Activity Report 2017

Project-Team SISTM

Statistics In System biology and Translational Medicine

RESEARCH CENTER
Bordeaux - Sud-Ouest

THEME
Modeling and Control for Life Sciences

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:

A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.2.4. - Statistical methods
A6.3.1. - Inverse problems
A6.3.4. - Model reduction
A6.4.2. - Stochastic control

Other Research Topics and Application Domains:

B1.1. - Biology
B1.1.6. - Genomics
B1.1.7. - Immunology
B1.1.9. - Bioinformatics
B1.1.11. - Systems biology
B1.1.14. - Microbiology
B2.2.4. - Infectious diseases, Virology
B2.2.5. - Immune system diseases
B2.3. - Epidemiology
B2.4.1. - Pharmaco kinetics and dynamics
B2.4.2. - Drug resistance

1. Personnel

Research Scientists

Daniel Commenges [Inserm, Senior Researcher, HDR]
Mélanie Prague [Inria, Researcher]

Faculty Members

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

Post-Doctoral Fellows

Marius Kwémou [Univ. Bordeaux, until Apr 2017]
Irene Balelli [Inserm]

PhD Students
Perrine Soret [Univ Bordeaux, PhD Student]
Laura Villain [Univ Bordeaux, PhD Student]
Chloé Pasin [Univ Bordeaux, PhD Student]
Edouard Lhomme [Univ Bordeaux, PhD Student]
Hadrien Lorenzo [Inserm, PhD Student]
Louis Capitaine [Univ Bordeaux, PhD Student, from Oct 2017]

Technical staff
Henri Bonnabau [Univ Bordeaux]
Chariff Alkhassim [Inserm]
Fabien Barbier [half-time, Bordeaux University Hospital]
Myrian Lopez [Inserm]
Solenne Delahaye [Univ Bordeaux]

Interns
Nicolas Lafosse [INSERM, from Feb 2017 until Jun 2017]
Paul Tauzia [Inria, from Mar 2017 until Sep 2017]
Louis Capitaine [Inria, from Feb 2017 until Aug 2017]
Augusta Alphonse [INSERM, from Feb 2017 until Aug 2017]
Marie Alexandre [INSERM, from Jun 2017 until Aug 2017]

Administrative Assistants
Sandrine Darmigny [Inserm, Assistant]
Sylvie Embolla [Inria, Assistant]
Nathalie Robin [Inria, Assistant]
Anton Ottavi [Inserm, European Projects coordinator]
Destandau Eugénie [Inserm, Communication]

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). 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 [57]. 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 [50]. 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 [49].

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 [42]. 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 [41].

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 [53], [60], [45], [51], [46], [59], 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 [56]. 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 [54]. 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 [58]. 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 [47]. 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" [55], "Systems vaccinology" [52], "Systems medicine" [44]. 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 [48]. 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) [43]. 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

Funding by PIA3 of the Bordeaux Graduate's School in Digital Public Health, headed by Rodolphe Thiébaut. This Master/PhD program is built with the expertise coming from the Inria Sistm project team and in collaboration with several other teams (MONC, CARMEN, PHOENIX).

Successful application of integrative analyses tools on high dimensional immunogenicity data from an Ebola vaccine trial with identification of early correlates of later antibody responses [30]

We published a milestone paper in Biometrics comparing descriptive models (Marginal structural models) and mechanistic models (Ordinary differential equations with mixed effect models on parameters). This is impactful as it shows that mechanistic models can adequately estimate a treatment effect in time-varying confounders settings as it is in observational studies. This opens the perspective of in silico trials based on predictions based on the analysis of available cohorts. [26]

We published a robust and powerful statistical method to analyzed longitudinal RNAseq data, largely outperforming state-of-the-art methods. With the surge in RNAseq data production, e.g. in system vaccinology, this principled methodology has a broad impact in deepening our understanding of underlying molecular mechanisms in various contexts, paving the way for further biological innovation. [16]

5.1.1. Awards

The University of Bordeaux Initiative of Excellence (IdEx) and Zellidja travel grants for a research PhD student visit of 3 months to the CSIRO's machine learning Data61 team, Canberra, Australia (Perrine Soret).

6. New Software and Platforms

6.1. marqLevAlg

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

6.2. VSURF

Variable Selection Using Random Forests
KEYWORDS: Classification - Statistics - Machine learning - Regression
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://github.com/robingenuer/VSURF

6.3. 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/

6.4. COVVSURF

Combination of Clustering Of Variables and Variable Selection Using Random Forests
KEYWORDS: Classification - Statistics - Cluster - Machine learning - Regression

- Contact: Robin Genuer
- URL: https://github.com/robingenuer/CoVVSURF

6.5. clogitLasso

KEYWORDS: Biostatistics - Bioinformatics - Machine learning - Regression
FUNCTIONAL DESCRIPTION: R package to fit a sequence of conditional logistic regression models with lasso, for small to large sized samples.

- Partner: DRUGS-SAFE
- Contact: Marta Avalos Fernandez
- URL: https://cran.r-project.org/web/packages/clogitLasso/index.html

6.6. 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. 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

6.7. NIMROD

Normal approximation Inference in Models with Random effects based on Ordinary Differential equations
KEYWORDS: Ordinary differential equations - Statistical modeling
FUNCTIONAL DESCRIPTION: We have written a specific program called NIMROD for estimating parameter of ODE based population models.

- Contact: Melanie Prague
- URL: http://etudes.isped.u-bordeaux2.fr/BIOSTATISTIQUE/NIMROD/documentation/html/index.html

6.8. 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/index.html

6.9. cytometree

KEYWORDS: Clustering - Biostatistics - Bioinformatics
FUNCTIONAL DESCRIPTION: Given the hypothesis of a bimodal distribution of cells for each marker, the algorithm constructs a binary tree, the nodes of which are subpopulations of cells. At each node, observed cells and markers are modeled by both a family of normal distributions and a family of bimodal normal mixture distributions. Splitting is done according to a normalized difference of AIC between the two families.

- Contact: Boris Hejblum
- URL: https://cran.r-project.org/web/packages/cytometree/index.html

6.10. CRTgeeDR

KEYWORDS: Missing data - Statistics - Regression
FUNCTIONAL DESCRIPTION: The CRTgeeDR package allows you to estimates parameters in a regression model (with possibly a link function). It allows treatment augmentation and IPW for missing outcome. It is particularly of use when the goal is to estimate the intervention effect of a prevention strategy agains epidemics in cluster randomised trials.

- Contact: Melanie Prague
- URL: https://cran.r-project.org/web/packages/CRTgeeDR/index.html

6.11. ludic

KEYWORDS: Probability - Biostatistics

FUNCTIONAL DESCRIPTION: An R package to perform probabilistic record Linkage Using only DIagnosis Codes without direct identifiers, using C++ code to speed up computations. Available on CRAN, development version on github.

- Contact: Boris Hejblum
- URL: https://cran.r-project.org/web/packages/ludic/index.html

7. New Results

7.1. Statistical and mechanistic modeling

- Prague M, Commenges D, Gran JM, Ledergerber B, Young J, Furrer H, Thiébaut R. Dynamic models for estimating the effect of HAART on CD4 in observational studies: Application to the Aquitaine Cohort and the Swiss HIV Cohort Study. Biometrics. 2017;73:294-304. [26]
  Comparison of descriptive models (Marginal structural models) and mechanistic models (Ordinary differential equations with mixed effect models on parameters) performances for estimating treatment effect from observational studies.
- Turner EL, Li F, Gallis JA, Prague M, Murray DM. Review of recent methodological developments in group-randomized trials: part 1 design. American Journal of Public Health 107(6), 907-915. [27]
- Mélanie Prague, Elisabeth Turner, Gallis John, Li Fan, Murray David. Review of recent Methodological Developments in group-randomized trials: Part 2 - Analysis. American Journal of Public Health, American Public Health Association, 2017. [28]
  Review of the literature on how to analyse data from a cluster randomised trial

  Review of literature regarding Design of cluster randomised trials.
- Jarne Ana, Daniel Commenges, Mélanie Prague, Yves Levy, Rodolphe Thiébaut and inspire 2/3 study group. Modelling CD4+ T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous interleukin 7. Annals of applied statistics 11(3) 1593-1616. [21]
  Modelling the disability of CD4 restauration by repeated cycles of Intereukine-7 injections using mechanistic models.
- Prague M., Wang R., Stephens A., Tchetgen Tchetgen E, DeGruttola V. Accounting for interference variables using semi-parametric augmentation for improving efficiency in clustered randomized trials with missing at random outcomes. Biometrics 72(4) 1066-1077. [29]
  Doubly robust approach to estimate the treatment effect in Cluster randomised trials.

7.2. Statistical learning methods for high-dimensional data

- Genuer R, Poggi J-M, Tuleau-Malot C, Villa-Vialaneix N. Random Forests for Big Data. Big Data Research, 9 (2017). [18]
  Addresses the analysis of Big Data with Random Forests, review of existing algorithms, simulation study and recommandations.
- Agniel D and Hejblum BP, Variance component score test for time-course gene set analysis of longitudinal RNA-seq data, Biostatistics, 18(4):589–604, 2017.[16]
  We propose tcgsaseq, a principled, model-free, and efficient method for detecting longitudinal changes in RNA-seq gene sets defined a priori. Applied to both simulated data and two real datasets, tcgsaseq is shown to exhibit very good statistical properties, with an increase in stability and power when compared to state-of-the-art methods

- Hejblum BP, Alkhassim C, Gottardo R, Caron F, Thiébaut R. Sequential Dirichlet Process Mixtures of Multivariate Skew t-distributions for Model-based Clustering of Flow Cytometry Data, preprint on ArXiv. [39]

  We propose to use a Bayesian nonparametric approach with Dirichlet process mixture of multivariate skew t-distributions to perform model based clustering of flow-cytometry data, robustly estimating the number of cell populations from the data.

7.3. Software tools

- Mougin F, Auber D, Bourqui R, Diallo G, Dutour I, Jouhet V, Thiessard F, Thiébaut R, Thébault P. Visualizing omics and clinical data: Which challenges for dealing with their variety? Methods. 2017. [23]
  This is a review on the methods to visualize the big data in the context of clinical research.
- Prague M., Wang R. and de Grutolla V. CRTgeeDR: An R package for generalized estimating equations with missing data in cluster randomized trials. R journal (in press). [29]
  Diffusion of a package to estimate estimate the intervention effect of a prevention strategy agains epidemics in cluster randomised trials. Estimation is based on GEE

7.4. Analysis of results from Clinical trials and cohorts in HIV

- Bouteloup V, Sabin C, Mocroft A, Gras L, Pantazis N, Le Moing V, d'Arminio Monforte A, Mary-Krause M, Roca B, Miro JM, Battegay M, Brockmeyer N, Berenguer J, Morlat P, Obel N, De Wit S, Fätkenheuer G, Zangerle R, Ghosn J, Pérez-Hoyos S, Campbell M, Prins M, Chêne G, Meyer L, Dorrucci M, Torti C, Thiébaut R; Standard Reference Distribution of CD4 Response to HAART Project Team for the Collaboration of Observational HIV Epidemiological Research Europe (COHERE) in EuroCoord. Reference curves for CD4 T-cell count response to combination antiretroviral therapy in HIV-1-infected treatment-naïve patients. HIV Medicine. 2017;18:33-44.
  This is a tool that should help clinicians to evaluate the immunological response to antiretroviral therapy in HIV infected patients. Thanks to the analyse of one of the largest observational database in the world, we provide with an online tool references on the CD4 count during the first year of antiretroviral therapy.
- Picat MQ, Pellegrin I, Bitard J, Wittkop L, Proust-Lima C, Liquet B, Moreau JF, Bonnet F, Blanco P, Thiébaut R; ANRS CO3 Aquitaine Cohort. Integrative Analysis of Immunological Data to Explore Chronic Immune T-Cell Activation in Successfully Treated HIV Patients. PLoS One. 2017;12:e0169164. [25]
  This is an analysis of data on gene expression and factors associated to the immune activation in HIV-infected patients. Using structural models, we disentangle the effect of factors such as CMV and the mediation through type 1 interferon pathway.
- Thiébaut R, Hue S, Le Marec F, Lelièvre JD, Dupon M, Foucat E, Lazaro E, Dabis F, Duffau P, Wittkop L, Surenaud M, Pellegrin I, Lacabaratz C, Bonnet F, Lévy Y; ANRS CO3 Aquitaine Cohort. Serum ST2 level is an independent predictor of all-cause mortality in HIV-infected patients. Aquitaine Cohort, France. AIDS. 2017 [32]
  In this work, we have demonstrated the independent effect of the biomarker ST2 on the overall mortality in a large cohort of HIV infected patients.
- Vladimir Novitsky, Mélanie Prague, Sikhulile Moyo, Tendani Gaolathe, Mompati Mmalane, et al.. High HIV-1 RNA among Newly Diagnosed People in Botswana. AIDS Research and Human Retroviruses, Mary Ann Liebert, To appear.

7.5. Analysis of results from Clinical trials and cohorts in Ebola

- Rechtien A, Richert L, Lorenzo H, Martrus G, Hejblum B, Dahlke C, Kasonta R, Zinser M, Stubbe H, Matschl U, Lohse A, Krähling V, Eickmann M, Becker S; VEBCON Consortium, Thiébaut R, Altfeld M, Addo MM. Systems Vaccinology Identifies an Early Innate Immune Signature as a Correlate of Antibody Responses to the Ebola Vaccine rVSV-ZEBOV. Cell Report. 2017;20:2251-2261. [30]
  In this work, we have analyzed high-dimensional gene expression and cell characterization data. We showed the predictive capacity of the innate immune response to the Ebola vaccine to define the antibody response established beyond one month. This is a successful application of integrative analyses tools on high dimensional immunogenicity data from an Ebola vaccine trial with identification of early correlates of later antibody responses.

7.6. Analysis of results from clinical trials and cohorts in other fields (Epidemiology, Medical Sciences, Neuroimaging, Sport Sciences)

- Zago L, Hervé PY, Genuer R, Laurent A, Mazoyer B, Tzourio-Mazoyer N, Joliot M. Predicting hemispheric dominance for language production in healthy individuals using support vector machine. Hum Brain Mapp. 2017 Dec;38(12):5871-5889. doi: 10.1002/hbm.23770. [34]
  Joint work with the GIN-IMN team, application of a variable selection procedure based on SVM method to analyze functional MRI data.
- Tabue-Teguo M, Grasset L, Avila-Funes JA, Genuer R, Proust-Lima C, Péres K, Féart C, Amieva H, Harmand MG, Helmer C, Salles N, Rainfray M, Dartigues JF. Prevalence and Co-Occurrence of Geriatric Syndromes in People Aged 75 Years and Older in France: Results From the Bordeaux Three-city Study. J Gerontol A Biol Sci Med Sci. (2017) [31]
  Application of Multiple Correspondence Analysis which enlights frailty and dependent profile of people from the Three-city study.
- Née M, Avalos M, Luxcey A, Contrand B, Salmi LR, Fourrier-Réglat A, Gadegbeku B, Lagarde E, Orriols L. Prescription medicine use by pedestrians and the risk of injurious road traffic crashes: A case-crossover study. PLoS Medicine. Jul 18;14(7):e1002347 (2017) [24]
  Exploration of the association between the use of medicinal drugs and the risk of being involved in a road traffic crash as a pedestrian. We applied the Lasso methodology that we previously developed for the case-crossover design in a high-dimensional setting. This design controls for time-invariant factors by using each case as its own control. This study highlights the necessity of improving awareness of the effect of medicines on pedestrians.
- Hellard P, Scordia C, Avalos M, Mujika I, Pyne DB. Modelling of optimal training load patterns during the 11 weeks preceding major competition in elite swimmers. Applied Physiology, Nutrition, and Metabolism. Jun 26 (2017) [20]
  Quantification of the relationships between the effects of periodization variables and competitive performance in elite swimmers using semiparametric mixed effects models. In the framework of the 2014-2016 R&D project "Quels schémas de périodisation pour la préparation des Jeux Olympiques à Rio ?" with the French Swimming Federation.
- Hejblum BP, Cui J, Lahey LJ, Cagan A, Sparks JA, Sokolove J, Cai T, Liao KP, Association between anti-citrullinated fibrinogen antibodies and coronary artery disease in rheumatoid arthritis, Arthritis Care & Research, in press, 2017. [19].
  We show that anti-cit-fibrinogen antibodies as a group were associated with CAD outcomes in our RA cohort, with the strongest signal for association arising from a subset of the autoantibodies.
- Liao KP, Sparks JA, Hejblum BP, Kuo IH, Cui J, Lahey LJ, Cagan A, Gainer VS, Liu W, Cai TT, Sokolove J, Cai T, Phenome-wide association study of autoantibodies to citrullinated and non-citrullinated epitopes in rheumatoid arthritis, Arthritis & Rheumatology, 69: 742–749, 2017. [22]
  We demonstrated application of a bioinformatics method, the PheWAS, to screen for the clinical significance of RA-related autoantibodies. Using the PheWAS approach, we identified potentially significant links between variations in the levels of autoantibodies and comorbidities of interest in RA.

7.7. Conferences

Members of the team were involved in 12 talks during conferences and colloquium.

Mélanie Prague has her work presented in 2017 in 2 peer-reviewed international conferences (Society of clinical trials Liverpool UK and Keystone symposium of mathematical modeling of virus infection, Este Park, May 2017).
Robin Genuer presented his work in the peer-reviewed International Conference of the European Research Consortium for Informatics and Mathematics Working Group (ERCIM WG) on Computational and Methodological Statistics, University of London, UK.
Boris Hejblum presented his work in the peer-reviewed 38th Annual Conference of the International Society for Clinical Biostatistics.
Chloé Pasin presented her work in the peer-reviewed Systems Immunology & Vaccine Design symposium, Heidelberg, Germany and the French Applied and Industrial Mathematics Society (SMAI) conference (Ronce-les-Bains).

Members of the team participated in French conferences: GDR Stat santé Bordeaux, GDR mathematical modelling of life Lyon and Journées de la statistique Francaise, Avignon (Perrine Soret, Mélanie Prague, Boris Hejblum). Mélanie Prague and Boris Hejblum also presented 4 posters in workshops.

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 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.

8.2. Bilateral Grants with Industry

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

- The EBOVAC1, EBOVAC2 and EBOVAC3 project, 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.
- The Prevac trial vaccine trial leads to collaboration with Merck and Janssen. 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.

9. Partnerships and Cooperations

9.1. Regional Initiatives

The team have strong links with :

- 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), "Computer research applied to health" (ERIAS) emerging research team.
- Bordeaux and Limoges CHU ("Centre Hospitalier Universitaire").
- Institut Bergonié, Univ Bordeaux through the Euclid platform
- Inria Project-team MONC and CQFD

The project team members are involved in:

- EUCLID/F-CRIN clinical trials platform (Laura Richert)
- The research project “Self-management of injury risk and decision support systems based on predictive computer modelling. Development, implementation and evaluation in the MAVIE cohort study” funded by the Nouvelle-Aquitaine regional council (Marta Avalos).
- Phenotyping from Electronic Health Records pilot project in cooperation with with the ERIAS Inserm emerging team in Bordeaux and the Rheumatology service from the Bordeaux Hospital (Boris Hejblum)

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 http://vaccine-research-institute.fr.

Collaboration with Inserm PRC (pôle Recherche clinique).

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).
- Rodolphe Thiébaut is a member of the CNU 46.04 (Biostatistiques, informatique médicale et technologies de communication).
- Rodolphe Thiébaut is a member of the Scientific Council of INSERM.
- 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).
- Laura Richert is an expert for the PHRC (Programme hospitalier de recherche Clinique).
- Marta Avalos is an expert for L’ANSM (Agence nationale de sécurité du médicament et des produits de santé)

9.2.3. Various Partnership

The project team members are involved in:

- DRUGS-SAFE platform funded by ANSM (Marta Avalos).
- F-CRIN (French clinical research infrastructure network) was initiated in 2012 by ANR under two sources of founding "INBS/Infrastructures nationales en biologie et en santé" and "Programme des Investissements d’avenir". (Laura Richert)

- 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)
- INCA (Institut National du Cancer) funded the project « Evaluation de l’efficacité d’un traitement sur l’évolution de la taille tumorale et autres critères de survie : développement de modèles conjoints. » (Principal PI Virginie Rondeau Inserm U1219, Mélanie Prague is responsible of Work package 4 mechanistic modeling of cancer: 5800 euros).
- Contrat Initiation ANRS MoDeL-CI: Modeling the HIV epidemic in Ivory Coast (Principal PI Eric Ouattara Inserm U1219 in collaboration with University College London, Mélanie Prague is listed as a collaborator).

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

9.4.1. Inria International Labs

Fred Hutchinson Cancer center, Seattle;

Baylor Institute for Immunology (Dallas);

Duke University;

Collaborations through clinical trials: NIH for the Prevac trial, NGO Alima for the Prevac trial, Several African clinical sites for Ebovac2 and Prevac trials;

NIH program project grant "Revealing Reservoirs During Rebound", Harvard School of Public Health (HSPH) and the University of California, San Diego (P01AI131385, total budget $1.5M/yr for 5 years starting Oct 2017, both university manage the funding. Mélanie Prague is part of modelling unit of the "Quantitative Methods" research project (budget $220,000/yr). The principal investigator for this core is Victor de Grutolla (HSPH) The overall goal of this grant is to characterize viral rebound following antiretroviral therapy cessation in cohorts of patients who have started therapy early in infection, as well as in a cohort of terminally-ill patients who will interrupt therapy before death and subsequently donate their bodies to research.

Project submitted by the Inria DYNMO-HIVE team with the laboratory “Program for evolutionary Dynamics” at Harvard (head Martin Nowak).

Denis Agniel from the RAND Corporation on developing statistical methods for the analysis of RNA-seq data (Boris Hejblum).

Tianxi Cai from Harvard University on developing methods for the linkage and analysis of Electronic Health Records data (Boris Hejblum).

Katherine Liao from Harvard University on the analysis of Electronic Health Records data in the context of Rheumatoid Arthritis (Boris Hejblum).

Machine learning team Data61 at CSIRO, Australia

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Alison Hill from "Program for evolutionary Dynamics" at Harvard visited the SISTM team twice (each time for 5 days) in May 2017 and July 2017. Main topic discussed was mechanistic modelling of new agents in HIV cure.

Linda Valeri from "Harvard medical school" visited the SISTM team 3 days. Main topic discussed was mediation analysis in high dimension.

Denis Agniel (RAND Corporation) visited B. Hejblum in Bordeaux for a week in May for a research collaboration

Visiting PhD student from Marcus Altfeld's team: Annika Niehrs (2 week stay with SISTM).

9.5.2. Visits to International Teams

Marta Avalos visited David Conesa 1 week in October through the Erasmus+ program Universidad de Valencia (Espagne).

Mélanie Prague got invited in University of Pennsylvania (Philadelphia) for a 2-days research trip in the Biostatistics department on April 2-3 2017.

Mélanie Prague spend 10 days in Boston as an invited researcher in Harvard School of Public Health, Biostatistics department on April 10-15 2017.

Boris Hejblum visited Harvard University for a week in November 2017 for a research collaboration with Katherine Liao & Tianxi Cai.

9.5.2.1. Research Stays Abroad

Marta Avalos was 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/

Perrine Soret was a research student visitor at CSIRO's Data61 in Canberra (Australia) from Feb. 2017 to April 2017. Collaboration with Cheng Soon Ong. Funding: The University of Bordeaux Initiative of Excellence and Zellidja travel grants for a research visit of 3 months.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

Daniel Commenges co-organised a SFB (French Region of the International Biometric Society) and GdR "Statistique et Santé" conference in Bordeaux (http://gdr-stat-sante.math.cnrs.fr/spip/spip.php?rubrique20),

Daniel Commenges organised a session of the French Region at the "Journées de Statistiques" (Avignon),

Robin Genuer Co-organises a reading group called Smiling in Bordeaux

Boris Hejblum organizes the Biostatistics Seminar Series at the Bordeaux Public Health Inserm Research Center

Mélanie Prague organized the "Déjeuners scientifiques" at the "Journées de la statistique française" 2017 and 2018.

Rodolphe Thiébaut organised a summer school course « Statistical analysis of big data in immunology », 14 participants from The Netherlands, Germany, UK and France. All the team members helped in the ground organisation and were involved in teaching. (http://bss-publichealth.u-bordeaux.fr/en/Teaching-team/Course-n-1-Statistical-analysis-of-big-data-linked-to-immunology-systems/r745.html

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

Daniel Commenges was a member of the scientific committee of the SFB (French Region of the International Biometric Society) and GdR "Statistique et Santé" conference in Bordeaux (http://gdr-stat-sante.math.cnrs.fr/spip/spip.php?rubrique20),

10.1.2.2. Member of the Conference Program Committees

Daniel Commenges is a member of the scientific committee of the International Biometric Conference Barcelona, July 2018 (http://2018.biometricconference.org),

Mélanie Prague is a member of the scientific committee of CIMI conference "Statistics in Health - personalised medicine" (http://www.cimi.univ-toulouse.fr/mib/en/conference-statistics-and-health), Toulouse 2018, 10-12 January

Mélanie Prague is a member of the scientific committee of the "Déjeuners scientifiques" at the "Journées de la statistique française" 2017 and 2018.

Rodolphe Thiébaut was a member of the scientific committee of the national conference on clinical research (EPICLIN)

Rodolphe Thiébaut was a member of the scientific committee of the IWHOD International Workshop on HIV Observational Databases since 2013 (http://newsite.iwhod.org/Committee)

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Lifetime Data Analysis (Daniel Commenges)

Statistics Surveys (Daniel Commenges)

2017 IMIA Yearb Med Inform, section editor (Rodolpe Thiébaut)

10.1.3.2. Reviewer - Reviewing Activities

AIDS (Rodolphe Thiébaut)

Annals of Applied Statistics (Boris Hejblum)

Annals of Statistics (Robin Genuer)

Am J Epidemiol (Marta Avalos)

Am J Public Health (Mélanie Prague)

BioData Mining (Boris Hejblum)

Biostatistics (Laura Richert)

Biometrics (Mélanie Prague)

IMIA Yearb Med Inform (Marta Avalos)

International Journal of Epidemiology (Daniel Commenges)

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

Machine Learning (Robin Genuer)

Neural Information Processing Systems (Robin Genuer)

Pattern Recognition Letters (Robin Genuer)

JRSS-B (Mélanie Prague)

Scientific Reports (Laura Richert)

Society of clinical trial (Mélanie Prague)

Statistical Methods in Medical Research (Robin Genuer, Mélanie Prague)

Statistical science (Mélanie Prague)

Statistics in Medicine (Marta Avalos)

10.1.4. Invited Talks

Rodolpe Thiébaut gave 3 invited talks. He gave an invited talk intitled “Objets connectés et Big Data” at the conference “10 ANS DE L’IRESP : Journées de la recherche en santé publique”

Daniel Commenges gave 2 invited talks: Evidence based Medicine (Canterbury) and Biophamaceutic group of the French Statistical Society (Paris)

Mélanie Prague gave 3 invited talks (Philadelphia Upenn, Boston Harvard school of public health, Inria Bordeaux) and had an invited session in Society for clinical trials, Liverpool, UK (11-14 Mai 2017) on "Integrate approaches for analysis of cluster randomised trials. New development in analysis “.

Robin Genuer was invited to the "Recent advances in tree-based methods (EO380)" session of ERCIM2017, by Ruoqing Zhu (University of Illinois Urbana-Champaign).

Boris Hejblum gave 2 invited talks on Dirichlet process mixtures of multivariate skew t-distributions for unsupervised clustering of cell populations from flow-cytometry data at The Biostatistics Unit at the Cambridge University (UK)

Marta Avalos gave 2 invited talks at the National Centre for Epidemiology & Population Health at the Australian National University, Canberra (Australia), April 2017 and at CSIRO’s Data61, Canberra (Australia), February 2017

Chloé Pasin gave an invited talk at the Probability-Statistics seminair of the Institute Montpelliérain Alexander Grothendieck, Montpellier

10.1.5. Leadership within the Scientific Community

Rodolpe Thiébaut and Chloé Pasin are elected members of the “collège des écoles doctorales’, University of Bordeaux

Daniel Commenges is President of the French Region of the International Biometric Society

Mélanie Prague is an elected member of the “Young statistician group” of SFdS (French Society of Statistics)

Mélanie Prague and Boris Hejblum are part of the group responsible for the communication of the SFdS - in charge of organising the sponsoring of the society by public and private companies.

Laura Richert is a member of F-CRIN Steering Committee

10.1.6. Scientific Expertise

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

- Rodolphe Thiébaut is a member of the commitee “Biologie des Systèmes et Cancer (Plan Cancer)”, a member of the Scientific Advisory Board of the “Institut Pierre Louis d’Epidémiologie et de Santé Publique” (UPMC, Dir : Dominique Costagliola), a member of the independent committee of international trials ODYSSEY and SMILE, a member of the scientific council of Muraz’s Center (Bobo-Dioulasso, Burkina Faso)
- 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).
- Laura Richert is an expert for the PHRC (Programme hospitalier de recherche Clinique).
- Marta Avalos is an expert for L’ASNM (Agence nationale de sécurité du médicament et des produits de santé)

10.1.7. Research Administration

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

Rodolphe Thiébaut is an elected member of the research committee (health sector) in University of Bordeaux and a member of the INSERM Scientific Council

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

In class teaching

Master : Robin Genuer, teaches in the two years of the Master of Public Health (M1 Santé publique, M2 Biostatistique) and 2nd year of the "Modélisation Stochastique et Statistique" Master, University of Bordeaux.

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

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 Master of Public Health at ISPED, Univ. Bordeaux, France (M2 Biostatistiques, M2 Eidémiologie).

Master : Mélanie Prague teaches in the Master of Public Health at ISPED, Univ. Bordeaux, France (M2 Biostatistiques).

Master : Marta Avalos teaches in the two years of the Master of Public Health (M1 Santé publique, M2 Biostatistique) at ISPED, Univ. Bordeaux and the 2nd year of the Biostatistics Master, University of Valencia (Spain).

Master : Chloe Pasin, Laura Villain, Hadrien Lorenzo and Louis Capitaine are teaching assistants for the two years.

Edouard Lhomme teaches in the Master of Public Health at ISPED, Univ. Bordeaux (M2 Epidémiologie) and in the Master of Vaccinology from basic immunology to social sciences of health (University Paris-Est Créteil, UPEC)

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

Edouard Lhomme teaches in PACES and DFASM1-3 for Medical degree at Univ. Bordeaux

Bachelor : Mélanie Prague and Boris Hejblum teach in the third year ingenious school ENSAI, Rennes.

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".

Boris Hejblum teaches in the Diplôme universitaire "Méthodes statistiques en santé.

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

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

10.2.2. Supervision

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

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 : Wenjia Wang "Modèle de Rasch", co-directed by Daniel Commenges with Mickael Guedj CIFRE Pharnext, from Oct 2015.

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

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, Cédric Galera and Marta Avalos.

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

PhD in progress : Louis Capitaine, *Random forests for high dimensional longitudinal data*, from Oct 2017, co-directed by Robin Genuer and Rodolphe Thiébaut.

PhD in progress : Madelyn Rojas *Self-management of injury risk and decision support systems based on predictive computer modelling. Development, implementation and evaluation in the MAVIE cohort study*, from Oct 2017, co-directed by Emmanuel Lagarde, David Conesa and Marta Avalos.

Master internship : Nicolas Lafosse "Déterminants de la réponse cellulaire T mesurée par Elispot après vaccination avec différentes stratégies vaccinales "prime-boost" préventives contre le VIH", directed by Laura Richert (1/03/2017 - 1/08/2017)

Master internship : Louis Capitaine, *Random forests for high dimensional longitudinal data*, directed by Robin Genuer (27/02/2017 - 31/08/2017)

Master internship : Paul Tauzia, *Utilisation de la déconvolution cellulaire pour détecter des différences d'expression génique*, directed by Mélanie Prague and Boris Hejblum (01/03/2017 - 31/08/2017)

Master internship : Augusta Alphonse, *Proof of concept for an automated gating tool applied to flow cytometry data from a HIV therapeutic vaccine trial*, directed by Mélanie Prague and Boris Hejblum (01/02/2017 - 31/07/2017)

Master internship : Marie Alexandre, *Correlation between cellular and antibodies response in Ebola vaccine*, directed by Mélanie Prague (01/06/2017 - 31/08/2017)

10.2.3. Juries

Daniel Commenges was involved in the PhD defences jury of Mr Adjakossa (Paris).

Robin Genuer was a PhD thesis examinator of Wei FENG thesis, defended the 07/19, "Investigation of training data issues in ensemble classification based on margin concept. Application to land cover mapping".

Mélanie Prague is a member of the follow-up dissertation comity of Nicolo Chiara (working on "Mathematical modeling of systemic aspects of cancer and cancer therapy"), Sébastien Benzkcry's PhD student (Inria Bordeaux Sud-ouest, MONC team) and Thiebaut Larivière (working on "Population Kalman estimation in Partial differential equations"), Annabelle Colin's PhD student (Inria Bordeaux Sud-ouest, MONC team).

Marta Avalos was a member of the follow-up dissertation comity of her PhD students Perrine Soret and Mélanie Née.

Rodolphe Thiébaut took part in the HDR committee of Patricia Thebault, Benoit Lepage, Simon Cauchemez, Sébastien Benzekry and Romulus Breban.

Rodolphe Thiébaut was involved in the PhD defences jury of Nicky De La Mata and Edouard Ollier.

Mélanie Prague took part in the recruitment commission MCF CNU 26 (CNAM).

Laura Richert, Rodolphe Thiébaut, Robin Genuer, Boris Hejblum and Marta Avalos participated to the juries of Master in Public Health (Biostatistics, Epidemiology)

Edouard Lhomme participated to the juries of two medical thesis defense, Medical School of Bordeaux University

10.3. Popularization

Participation to the Inria magazines "Plug-in" and "So news" (Mélanie Prague).

Participation in "The ou café" in Inria Bordeaux (Mélanie Prague).

Lightning talk to present the NIMROD software in Dev Days at Inria (Mélanie Prague).

Edouard Lhomme, as President of AquitHealth, a non profit organisation for the development of e-Health in south west France organized the When Doctors Meet Hackers (WDMH) Congress 2017. The WDMH Congress is a 3 day event with one day of conference on the futur of health and 50 hours of hackathon. The Hackathon is a human adventure where healthcare professionals, patients, developers, designers and entrepreneurs collaborate over a weekend to develop prototypes of eHealth solutions.Six WDMH meetups were also organised each month from January to June 2017 in Bordeaux to discuss about several e-heath topics (telemedecine, data protection, simulation, ...)

11. Bibliography

Major publications by the team in recent years

[1] M. AVALOS, H. POUYES, Y. GRANDVALET, L. ORRIOLS, E. LAGARDE. *Sparse conditional logistic regression for analyzing large-scale matched data from epidemiological studies: a simple algorithm*, in "BMC Bioinformatics", 2015, vol. 16, n^{o} Suppl 6, S1 p. [*DOI :* 10.1186/1471-2105-16-S6-S1], https://hal.inria.fr/hal-01217312

[2] D. COMMENGES, A. GÉGOUT-PETIT. *The stochastic system approach for estimating dynamic treatments effect*, in "Lifetime Data Analysis", 2015, 18 p. [*DOI :* 10.1007/s10985-015-9322-3], https://hal.inria.fr/hal-01205328

[3] D. COMMENGES, H. JACQMIN-GADDA. *Dynamical Biostatistical Models*, Chapman and Hall/CRC, October 2015, https://hal.inria.fr/hal-01580149

[4] B. P. HEJBLUM, J. SKINNER, R. THIÉBAUT. *Time-Course Gene Set Analysis for Longitudinal Gene Expression Data*, in "PLoS Computational Biology", 2015, vol. 11, n^{o} 6, e1004310 p. [*DOI :* 10.1371/JOURNAL.PCBI.1004310], https://hal.inria.fr/hal-01203446

[5] 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

[6] 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^{o} 11, pp. 5315 - 5328, https://hal.inria.fr/hal-01406364

[7] 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

[8] 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

[9] S. M. MOYO, T. M. MOHAMMED, K. M. WIRTH, M. M. PRAGUE, K. M. BENNETT, M. P. HOLME, L. M. MUPFUMI, P. M. SEBOGODI, N. M. MORAKA, C. M. BOLEO, C. M. MAPHORISA, B. M. SERAISE, S. M. GASEITSIWE, R. M. MUSONDA, E. M. VAN WIDENFELT, K. M. POWIS, T. M. GAOLATHE, E. M. TCHETGEN TCHETGEN, J. M. MAKHEMA, M. M. ESSEX, S. M. LOCKMAN, V. M. NOVITSKY. *Point-of-Care Cepheid Xpert HIV-1 Viral Load Test in Rural African Communities Is Feasible and Reliable*, in "Journal of Clinical Microbiology", 2016, vol. 54, n^{o} 12, pp. 3050 - 3055 [*DOI :* 10.1128/JCM.01594-16], https://hal.inria.fr/hal-01579540

[10] 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

[11] L. RICHERT, E. LHOMME, C. FAGARD, Y. LEVY, G. CHÊNE, R. THIÉBAUT. *Recent developments in clinical trial designs for HIV vaccine research*, in "Human vaccines & immunotherapeutics", 2015, vol. 11, n^{o} 4, pp. 1022–9 [*DOI :* 10.1080/21645515.2015.1011974], https://hal.archives-ouvertes.fr/hal-01164644

[12] 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

[13] 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

[14] R. THIÉBAUT, J. DRYLEWICZ, M. PRAGUE, C. LACABARATZ, S. BEQ, A. JARNE, T. CROUGHS, R.-P. SEKALY, M. M. LEDERMAN, I. SERETI, D. COMMENGES, Y. LÉVY. *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

[15] 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

Publications of the year

Articles in International Peer-Reviewed Journals

[16] D. AGNIEL, B. P. HEJBLUM. *Variance component score test for time-course gene set analysis of longitudinal RNA-seq data*, in "Biostatistics", 2017, vol. 18, n° 4, pp. 589-604, https://arxiv.org/abs/1605.02351 [*DOI :* 10.1093/BIOSTATISTICS/KXX005], https://hal.inria.fr/hal-01579077

[17] V. BOUTELOUP, C. SABIN, A. MOCROFT, L. GRAS, N. PANTAZIS, V. LE MOING, A. D'ARMINIO MONFORTE, M. MARY-KRAUSE, R. ROCA, J. M. MIRO, M. BATTEGAY, N. H. BROCKMEYER, M. BERENGUER, P. MORLAT, C. OBEL, S. DE WIT, G. FÄTKENHEUER, R. ZANGERLE, J. GHOSN, R. PÉREZ-HOYOS, M. CAMPBELL, C. PRINS, G. CHÊNE, L. MEYER, M. 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", 2017, vol. 18, n° 1, pp. 33 - 44 [*DOI :* 10.1111/HIV.12389], https://hal.archives-ouvertes.fr/hal-01579158

[18] R. GENUER, J.-M. POGGI, C. TULEAU-MALOT, N. VILLA-VIALANEIX. *Random Forests for Big Data*, in "Big Data Research", September 2017, vol. 9, pp. 28-46, https://arxiv.org/abs/1511.08327 [*DOI :* 10.1016/J.BDR.2017.07.003], https://hal.archives-ouvertes.fr/hal-01233923

[19] B. P. HEJBLUM, J. CUI, L. J. LAHEY, A. CAGAN, J. A. SPARKS, J. SOKOLOVE, T. CAI, K. P. LIAO. *Association between anti-citrullinated fibrinogen antibodies and coronary artery disease in rheumatoid arthritis*, in "Arthritis Care & Research", 2017, forthcoming [*DOI :* 10.1002/ACR.23444], https://hal.inria.fr/hal-01654403

[20] P. HELLARD, C. SCORDIA, M. F. AVALOS, I. MUJIKA, P. DAVID. *Modelling of optimal training load patterns during the 11 weeks preceding major competition in elite swimmers*, in "Applied Physiology, Nutrition, and Metabolism", 2017, https://hal.inria.fr/hal-01577936

[21] A. JARNE, D. COMMENGES, M. PRAGUE, Y. LEVY, R. THIÉBAUT. *Modeling CD4 + T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous Interleukin 7*, in "Annals Of Applied Statistics", June 2017, https://hal.inria.fr/hal-01579008

[22] K. LIAO, J. A. SPARKS, B. P. HEJBLUM, I.-H. KUO, J. CUI, L. J. LAHEY, A. CAGAN, V. S. GAINER, W. LIU, T. T. CAI, J. SOKOLOVE, T. CAI. *Phenome-Wide Association Study of Autoantibodies to Citrullinated and Noncitrullinated Epitopes in Rheumatoid Arthritis*, in "Arthritis & rheumatology", April 2017, vol. 69, n^{o} 4, pp. 742 - 749 [*DOI :* 10.1002/ART.39974], https://hal.inria.fr/hal-01654405

[23] F. MOUGIN, D. AUBER, R. BOURQUI, G. DIALLO, I. DUTOUR, V. JOUHET, F. THIESSARD, R. THIEBAUT, P. THEBAULT. *Visualizing omics and clinical data: which challenges for dealing with their variety?*, in "Methods", 2017 [*DOI :* 10.1016/J.YMETH.2017.08.012], https://hal.archives-ouvertes.fr/hal-01671067

[24] N. MÉLANIE, M. F. AVALOS, A. LUXCEY, B. CONTRAND, L. R. SALMI, A. FOURRIER-REGLAT, B. GADEGBEKU, E. LAGARDE, L. ORRIOLS. *Prescription medicine use by pedestrians and the risk of injurious road traffic crashes: A case-crossover study*, in "PLoS Medicine", 2017, https://hal.inria.fr/hal-01577941

[25] M.-Q. PICAT, I. PELLEGRIN, J. BITARD, L. WITTKOP, C. PROUST-LIMA, B. LIQUET, J.-F. MOREAU, F. BONNET, P. BLANCO, R. THIÉBAUT. *Integrative Analysis of Immunological Data to Explore Chronic Immune T-Cell Activation in Successfully Treated HIV Patients*, in "PLoS ONE", January 2017, vol. 12, n^{o} 1 [*DOI :* 10.1371/JOURNAL.PONE.0169164], https://hal.archives-ouvertes.fr/hal-01579159

[26] 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", March 2017, vol. 73, n^{o} 1, pp. 294 - 304 [*DOI :* 10.1111/BIOM.12564], https://hal.inria.fr/hal-01670680

[27] M. PRAGUE, E. TURNER, L. FAN, G. JOHN, M. DAVID. *Review of recent Methodological Developments in group-randomized trials: Part 1 - Design*, in "American Journal of Public Health", June 2017, https://hal.inria.fr/hal-01579073

[28] M. PRAGUE, E. TURNER, G. JOHN, L. FAN, M. DAVID. *Review of recent Methodological Developments in group-randomized trials: Part 2 - Analysis*, in "American Journal of Public Health", June 2017, https://hal.inria.fr/hal-01579075

[29] M. PRAGUE, R. WANG, V. DE GRUTTOLA. *CRTgeeDR: An R Package for Doubly Robust Generalized Estimating Equations Estimations in Cluster Randomized Trials with Missing Data*, in "R journal", September 2017, https://hal.inria.fr/hal-01579080

[30] A. RECHTIEN, L. RICHERT, H. LORENZO, G. MARTRUS, B. P. HEJBLUM, C. DAHLKE, R. KASONTA, M. ZINSER, H. STUBBE, U. MATSCHL, A. LOHSE, V. KRÄHLING, M. EICKMANN, S. BECKER, R. THIÉBAUT, M. ALTFELD, M. ADDO. *Systems Vaccinology Identifies an Early Innate Immune Signature as a Correlate of Antibody Responses to the Ebola Vaccine rVSV-ZEBOV*, in "Cell Reports ", 2017, vol. 20, n^{o} 9, pp. 2251 - 2261 [*DOI :* 10.1016/J.CELREP.2017.08.023], https://hal.inria.fr/hal-01579246

[31] M. TABUE-TEGUO, L. GRASSET, J. A. AVILA-FUNES, R. GENUER, C. PROUST-LIMA, K. PÉRÈS, C. FEART, H. AMIEVA, M. G.-C. HARMAND, C. HELMER, N. SALLES, M. RAINFRAY, J. F. DARTIGUES. *Prevalence and Co-Occurrence of Geriatric Syndromes in People Aged 75 Years and Older in*

France: Results From the Bordeaux Three-city Study, in "Journals of Gerontology, Series A", May 2017 [*DOI :* 10.1093/GERONA/GLX068], https://hal.inria.fr/hal-01652204

[32] R. THIEBAUT, S. HUE, F. LE MAREC, J.-D. LELIEVRE, M. DUPON, E. FOUCAT, E. LAZARO, F. DABIS, P. DUFFAU, L. WITTKOP, M. SURENAUD, I. PELLEGRIN, C. LACABARATZ, F. BONNET, Y. LÉVY. *Serum ST2 level is an independent predictor of all-cause mortality in HIV-infected patients. Aquitaine Cohort, France*, in "AIDS", 2017, https://hal.archives-ouvertes.fr/hal-01671081

[33] R. THIEBAUT, F. THIESSARD. *Public Health and Epidemiology Informatics*, in "IMIA Yearbook of Medical Informatics", 2017 [*DOI :* 10.15265/IY-2017-035], https://hal.archives-ouvertes.fr/hal-01671095

[34] L. ZAGO, P.-Y. HERVÉ, R. GENUER, A. LAURENT, B. MAZOYER, N. TZOURIO-MAZOYER, M. JOLIOT. *Predicting hemispheric dominance for language production in healthy individuals using support vector machine*, in "Human Brain Mapping", September 2017 [*DOI :* 10.1002/HBM.23770], https://hal.inria.fr/hal-01590497

National Conferences with Proceedings

[35] M. PRAGUE, R. THIÉBAUT, P. MOIREAU, A. COLLIN. *Joint-state and parameters estimation using nudging and SEIK filters for HIV mechanistic models*, in "Journée de la statistique francaise", Avignon, France, June 2017, https://hal.inria.fr/hal-01579068

Conferences without Proceedings

[36] H. LORENZO, R. THIEBAUT, J. SARACCO. *Multi-block high-dimensional lasso-penalized analysis with imputation of missing data applied to postgenomic data in an Ebola vaccine trial*, in "Annual workshop on Statistical Methods for Post Genomic Data - SMPGD 2018", Montpellier, France, January 2018, https://hal.archives-ouvertes.fr/hal-01664610

[37] P. SORET, M. F. AVALOS, S. CHENG, R. THIEBAUT. *High-dimensional compositional microbiota data: state-of-the-art of methods and software implementations*, in "2017 - GDR « Statistiques et santé »", Bordeaux, France, October 2017, https://hal.inria.fr/hal-01667295

Other Publications

[38] R. GENUER, J.-M. POGGI. *Arbres CART et Forêts aléatoires,Importance et sélection de variables*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01387654

[39] B. P. HEJBLUM, C. ALKHASSIM, R. GOTTARDO, F. CARON, R. THIÉBAUT. *Sequential Dirichlet Process Mixtures of Multivariate Skew t-distributions for Model-based Clustering of Flow Cytometry Data*, August 2017, https://arxiv.org/abs/1702.04407 - 39 pages, 11 figures, https://hal.inria.fr/hal-01579063

[40] M. ´. PRAGUE, D. COMMENGES, R. THIÉBAUT. *In Silico Evaluation of HIV Short-cycle Therapies with Dynamical Models*, August 2017, Keystone Symposium , Poster, https://hal.inria.fr/hal-01579070

References in notes

[41] 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

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

[43] 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

[44] 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

[45] 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

[46] 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

[47] 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.

[48] 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^{o} 1, pp. 121-130, 0300 5771 English

[49] 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^{o} 5255, pp. 1582-6

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

[51] 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^{o} 1, pp. 12–35

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

[53] 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^{o} 15, pp. 2199–2214

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

[55] C. SCHUBERT. *Systems immunology: complexity captured*, in "Nature", 2011, vol. 473, n^{o} 7345, pp. 113-4

[56] 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

[57] 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

[58] 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

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

[60] 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