RESEARCH CENTRE

Inria Centre
at the University of Bordeaux

IN PARTNERSHIP WITH:

Université de Bordeaux, INSERM

2023

ACTIVITY REPORT

Project-Team

SISTM

Statistics In System biology and Translational Medicine

DOMAIN

Digital Health, Biology and Earth

THEME

Modeling and Control for Life Sciences

Inria

Contents

Project-Team SISTM

Creation of the Project-Team: 2015 January 01

Keywords

Computer sciences and digital sciences

A3.1.1. – Modeling, representation

A3.1.10. – Heterogeneous data

A3.1.11. – Structured data

A3.3.2. – Data mining

A3.3.3. – Big data analysis

A3.4.1. – Supervised learning

A3.4.2. – Unsupervised learning

A3.4.3. – Reinforcement learning

A3.4.4. – Optimization and learning

A3.4.5. – Bayesian methods

A5.2. – Data visualization

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

A9.2. – Machine learning

A9.6. – Decision support

Other research topics and application domains

B1.1. – Biology

B1.1.5. – Immunology

B1.1.7. – Bioinformatics

B1.1.10. – Systems and synthetic biology

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

B9.1.1. – E-learning, MOOC

B9.5.6. – Data science

B9.8. – Reproducibility

1 Team members, visitors, external collaborators

Research Scientists

- Quentin Clairon [INRIA, ISFP]
- Boris Hejblum [INSERM, Researcher]
- Melanie Prague [INRIA, Researcher]

Faculty Members

- Rodolphe Thiebaut [Team leader, UNIV BORDEAUX, Professor, HDR]
- Marta Avalos Fernandez [UNIV BORDEAUX, Associate Professor, HDR]
- Robin Genuer [UNIV BORDEAUX, Associate Professor, HDR]
- Edouard Lhomme [UNIV BORDEAUX, Associate Professor]
- Laura Richert [UNIV BORDEAUX, Professor, HDR]
- Linda Wittkop [UNIV BORDEAUX, Associate Professor, HDR]

Post-Doctoral Fellows

- Marie Alexandre [INSERM]
- Corentin Segalas [UNIV BORDEAUX, from Nov 2023]

PhD Students

- Kalidou Ba [INSERM, from Oct 2023]
- Houreratou Barry [INSERM, until Nov 2023]
- Antonin Colajanni [UNIV BORDEAUX, from Oct 2023]
- Thomas Ferte [UNIV BORDEAUX]
- Auriane Gabaut [INRIA, from Oct 2023]
- Iris Ganser [UNIV BORDEAUX]
- Benjamin Hivert [UNIV BORDEAUX]
- Arthur Hugues [INSERM, from Oct 2023]
- Cyrille Kone [INRIA, co-supervision 50% with Inria Lille Scool team]
- Annesh Pal [INRIA, from Oct 2023]
- Justine Remiat [INSERM, from Oct 2023]
- Hélène Savel [INSERM, CIFRE, until Nov 2023]

Technical Staff

- Kalidou Ba [INSERM, Engineer, until Sep 2023]
- Antonin Colajanni [INSERM, Engineer, until Sep 2023]
- Sara Fallet [INSERM, Engineer, from Apr 2023]
- Mélanie Huchon [INSERM, Engineer]
- Zhe Li [INRIA, Engineer, from Oct 2023]
- Clément Nerestan [INRIA, Engineer]
- Anton Ottavi [INSERM]
- Panthea Tzourio [INSERM, Engineer]

Interns and Apprentices

- Marie Colin [INRIA, Intern, from May 2023 until Sep 2023]
- Auriane Gabaut [INSERM, Intern, from Mar 2023 until Sep 2023]
- Arthur Hugues [INSERM, Intern, from Feb 2023 until Jul 2023]
- Victoria Idier [INRIA, Intern, from Mar 2023 until Jun 2023]
- Junior Jumbong [INSERM, Intern, from May 2023 until Aug 2023]
- Camila Labarca Rosenbluth [INRIA, Intern, until Apr 2023]
- Iban Larrouture [UNIV BORDEAUX, Intern, from May 2023 until Oct 2023]
- Quentin Laval [INSERM, from Nov 2023]
- Fanny Moreau [INSERM, Intern, from Apr 2023 until Jul 2023]
- Maud Perpere [INSERM, Intern, from Mar 2023 until Jun 2023]
- Arthur Saracco [INRIA, Intern, from Feb 2023 until Jul 2023]
- Lydia Wilson [INSERM, Intern, from Mar 2023 until Aug 2023]

Administrative Assistants

- Sandrine Darmigny [INSERM]
- Anne-Lise Pernel [INRIA, from Mar 2023]

Visiting Scientist

- John Fricks [UNIV BORDEAUX AND ARIZONA STATE UNIV , from Sep 2023]

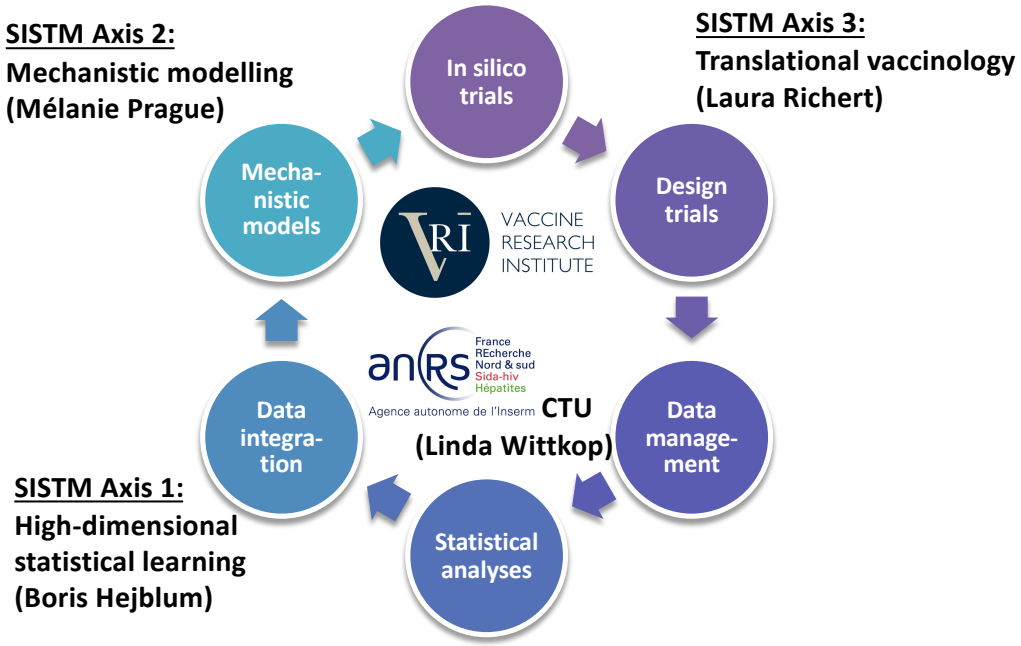

Figure 1: The SISTM wheel. Presentation of the three axes.

2 Overall objectives

The two main objectives of the SISTM team are:

i) to accelerate the development of vaccines by analyzing all the information available in early clinical trials and optimizing new trials

ii) to develop new data science approaches to analyze and model high dimensional data in small sample size studies.

The methods developed are relevant in many other applications than those encountered in the SISTM team. However, the focus devoted to vaccine development is justified by the importance of the objective from a public health point of view and a good knowledge of the application field that maximizes the relevance and the implementation of the methods developed. This equilibrium between the methodological and the applied work reached over the last years is a fundamental motivation for each member of the SISTM team even though the background could be very different from one researcher to the other (e.g. applied mathematics vs. public health). This equilibrium is maintained by the organization of the team as well as the collaborations established especially through the Vaccine Research Institute, Bordeaux University, Inserm and Inria. Hence, we are able to collaborate for a theoretical problem during the development of a new method (e.g. demonstration of the convergence of an estimator) as well as to translate the research outcomes (either new analytical methods or applied results) to clinicians and biologists first in our collaborative networks and then beyond. Figure (1) illustrates this synergism and materializes the three research axis of the team: high dimension statistical learning, mechanistic modelling and translational vaccinology.

Biological and clinical research has dramatically changed thanks to technological advances, leading to the possibility of measuring much more biological parameter in high-throughput methods than previously. Clinical research studies can include now traditional measurements such as clinical status, but also (tens of) thousands of cell populations, peptides, gene expressions for a given participant. This has facilitated the transfer of knowledge from basic to clinical science (from "bench to bedside") and vice versa, a process often called "Translational medicine". However, the analysis of these large amounts of data requires specific methods, especially when one wants to have a global understanding of the information inherent to complex systems through an "integrative analysis". Systems like the immune

system are complex because of the many interactions within and between several levels (inside cells, between cells, in different tissues, between individuals, between various species). This has led to a new field called "Systems biology" rapidly adapted to specific topics such as "Systems Immunology" [70], "Systems vaccinology" [68], "Systems medicine" [58]. From the statistical point of view, two main challenges arise: i) to adequately deal with the massive amount of data ii) to find relevant models capturing observed data.

First, with respect to the relatively moderate number of participants in vaccine studies and clinical trials, this profusion of high-throughput "omics" data often sets us in a ultra high-dimension context. This mandates updated statistical tools able to tackle this wealth of information. On top of the challenge signal extraction and dimension reduction, there is a redundancy of the information across data modalities, that in turn can be leveraged to boost statistical methods and harness artificial intelligence approaches to predict immunological surrogate endpoints from early indicators.

Second, once a small amount of markers has been selected, we use modeling approaches to understand the biological mechanism (specifically in vaccinology antibodies kinetics or viral dynamics [64]). In our work we are interested in the inverse problem: how can we infer the mechanism of a biological process from data. It can be modeled using differential equations (mainly ordinary but could extend to partial and stochastic). The challenge in our methods rely in the type of collected data which are sparse (as opposed to measured in continuous time), with measurement error and repeated across multiple individuals. Thus, we adopt nonlinear mixed-effects model population approach [62]. Construction of these models is a challenging process which requires confirmed expertise, advanced statistical methods and the development of software tools.

Finally, once a model has been defined and validated, it is possible to perform *in silico* trials to predict further strategies. In particular, a systems personalized vaccinology approach [65] using multidimensional immunogenicity data from clinical trials and statistical models (such as optimal control or reinforcement learning) can help improve the selection of optimized vaccine strategies that can then be tested again in subsequent clinical trials. Domains of application of our methods in vaccinology focuses on, but not limited to, Ebola virus, Human Immunodeficiency Virus (HIV) virus and SARS-CoV-2 virus. The choice of these applications is deliberate and important for the relevance of the results and their translation into practice, thanks to a longstanding collaboration with several immunology research teams and the implication of the team in VRI - the Labex Vaccine Research Institute.

The SISTM team benefits from a very rich ecosystem. Firstly, it is one of the rare (n~6 as of 01/2022) teams belonging to both Inserm and Inria national institutes, which helps establishing collaboration as testified by the co-supervision of PhD Students and co-publications with other researchers belonging to Inserm teams or Inria teams from the two research centres in Bordeaux. Secondly, the applications in clinical research are facilitated by the very close collaboration with Clinical Trial Units (CTUs): from the ANRS/VRI (UMS 54 MART directed by LW), from Bordeaux Hospital (USMR directed by LR and previously by RT), from F-CRIN (Euclid platform, directed by LR and EL), from the international consortia linked to the Vaccine Research Institute (for which SISTM is leading the data science division). Finally, the team is very much involved in teaching activities at Bordeaux University and ISPED Institute, especially through the Graduate's program Digital Public Health (directed by RT) and the Master of Public Health (first year in e-learning led by MA, Biostatistics led by RG and Public Health Data Science led by RT). A better description of all these interaction can be found in section Teaching (9.2) and section Fundings (8).

In term of positioning in regards of other teams at Inria and in France, the application domain (immunology and vaccine development) is nearly unique with the exception of DRACULA in Lyon. DRACULA like other teams at Inria (MONC, CARMEN, M3DISIM) or Inserm (IAME) or international groups (e.g. A. Perelson lab in Los Alamos) are also developing mathematical models but rarely with the integration of high dimensional data. In other hand, groups such as Raphael Gottardo lab in Lausanne (previously at the Fred Hutchinson in Seattle) are developing methods for high dimensional data in immunology but are not using dynamical models.

3 Research program

The team is organized in three research axes:

1. High Dimensional Statistical Learning (leader Boris Hejblum),

2. Mechanistic learning (leader Mélanie Prague),
3. Translational vaccinology (leader Laura Richert).

3.1 Axis 1 - High-Dimensional Statistical Learning

The specific objectives are:

- To unlock the analysis of high-dimensional longitudinal data by developing suitable statistical approaches, in particular for applications to longitudinal high-throughput data (e.g. microbiome, transcriptome, cytomics) generated in vaccine trials.
- To leverage prior biological knowledge and formally incorporate it into statistical models to tackle the small n large p setting, one of the characteristics of early phase vaccine trials.
- To advance adaptive clustering methods of high-dimensional data in both supervised and unsupervised settings, especially to infer the proportions of cellular population from gene expression measurements and also to identify gene whose expression is key in segmenting transcriptomic measurements across vaccine arms or disease severity for instance.
- To perform feature selection and dimension reduction of high-dimensional molecular and cellular data, as a first step to feed such information into mechanistic models.

Despite being high-dimensional, biomedical data from high-throughput technologies is rarely analyzed in its entirety due to its size or its complexity. For example, in cellular phenotyping data, only a limited number of markers are used to quantify a pre-defined set of cell types; this strategy precludes the discovery of new cell types defined by new combinations of markers. This issue is exacerbated by mass cytometry technologies, which enable the measurement of up to 100 markers on a single cell.

However, measuring specific cells across a large number of intracellular and surface markers requires substantial amounts of blood, ideally fresh, making it difficult to implement such measurements on large sample sizes with multiple repeated measurements. This motivates the exploration of replacing cell phenotyping with transcriptomics analysis in whole blood, as gene expression can be measured more easily and frequently with a much finer temporal resolution (using finger prick at-home self-sampling technology [73]). This ambitious endeavor goes beyond previous work done on this topic using standard deconvolution approaches [60]. By using more sophisticated statistical [53], machine learning [54], and artificial intelligence [75] models (in particular for adaptive clustering, robust to unobserved cell populations), by exploiting public databases of cytometry data coupled newly available single cell transcriptomics measurements, and by explicitly leveraging the repeated aspect of longitudinal observations from vaccine trial measurements, we set ourselves to successfully study and develop methods delivering accurate cell proportions estimates from gene expression data.

In addition, among high-throughput omics data, the microbiome is also becoming an increasingly important component in understanding the immune system[63]. The compositional nature of these data, along with their hierarchical phylogenic structure particularly suited to tree-based models, coupled with their high-dimension requires the use of adequate statistical tools [72].

Furthermore, while those high-throughput molecular and cellular data have an unquestionable value for diving into underlying mechanisms governing and deepening our understanding of the human immune system, we want to determine whether they could be used as early surrogate markers for correlates of protection in vaccine studies (such as antibody titers after vaccination). Due to their high-dimensional nature, answering this question requires the development of new mediation approaches [12] to develop this emerging field of vaccinomics epidemiology.

Outside biological data generated in clinical trials, electronic health records from hospital data warehouse systems are also representing an opportunity for studying infectious diseases and requires specific approaches. Several works have been done on this topic in the SISTM team [55, 57, 71, 78, 56].

Regarding this research axis, there are some common interest with other Inria teams such as HeKA and PreMeDICaL in regards of the use of machine learning approaches applied to medical data or Soda, Mind and Aramis that are more focus on brain applications. Applications in SISTM are focused on analyzing high-throughput omics data (nearly no imaging) in immunology and vaccine trials. Also,

modeling biological networks as done in Beagle or Dyliss Inria teams is not an objective of SISTM, the data recorded in human clinical trials being unsuited because of their sparsity. At the international level, the main competitors are groups engaged in biostatistical methods development for the analysis of omics data such as Jeff Leek (previously at John Hopkins, now at Fred Hutchinson, Seattle), Raphael Gottardo (previously at Fred Hutchinson, Seattle, now at Université de Lausanne, Switzerland) or Mark Robinson (Prof at the University of Zurich, Switzerland).

3.2 Axis 2 - Mechanistic learning

The specific objectives are:

- To develop methods for statistical inference of differential equations model parameters in population framework.
- Within-host modeling of immunological and virological dynamics in samples of individuals.
- Implementing control strategies toward personalized medicine.

When studying the dynamics of some given markers one can for instance use descriptive models summarizing the dynamics over time in term of slopes of the trajectories [74]. These slopes can be compared between treatment groups or according to patients' characteristics. Mechanistic modeling, that is dynamical models based on Ordinary Differential Equations (ODE), could be preferred as it integrates knowledge about the biological mechanism and it carries causal interpretation of the observed phenomenon [52, 67]. Thus, in this axis, we focus on inference of model parameters of mechanistic models in population of subjects (e.g. from a clinical trial). This modeling is constituted by three features: 1/ a dynamical model, which describes a phenomenon, often based on ODE (but also possibly partial and stochastic DE) 2/ a statistical model, which describes the variability that exists in data, and 3/ an observational model, which relates what is observable with error in the mathematical model.

The definition of the model needs to identify the parameter values that fit the data. Contrary to Inria team such as MAKUTU or BEAGLE, which are interested in simulation scheme for large differential equation systems, we focus on inverse problems for inference of parameters from data. In clinical research, this is challenging because data are sparse, and often unbalanced, coming from populations of individuals. A substantial inter-individual variability is always present and needs to be accounted as this is the main source of information. Many approaches have been developed to estimate the parameters of non-linear mixed models (NLME) including Bayesian approach [77], semi-parametric approaches [76] or penalized likelihood approach (in house NIMROD program [66]). The SAEM algorithm [61], as implemented in Monolix [69], is now also used for many of our projects. We however, continue to participate in the development of related methods in collaboration with Inria team XPOP. We also devote a large part of this axis methodological research to the development of alternative methods for estimation in NLME ODE models.

Having a good mechanistic model with a population approach in a biomedical context 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 biological mechanisms rather than simple regressions. Control theory (Inria team ASTRAL), game theory (Inria team SCOOL) and learning approaches (Inria team FLOWERS) may serve for defining optimal interventions or optimal designs to evaluate new interventions. In the last period of evaluation, we made a proof of concept of such open-loop control problem. We model the response to Interleukin-7 (IL-7) injections in HIV-infected patients, and that has allowed to design new trials finally implementing personalized medicine [59]. We still devote a large part of this axis methodological research to the development of methods around personalized medicine.

Finally, one core research direction of this axis, is the development of statistically proven (unbiased and efficient) methods for evaluation of intervention effects. Once adjusted to data, these models could be used to perform *in silico* trials to predict the effect of a various administration strategies in various

populations. Development of estimation/simulation tools allow the translational vaccinology axis to better design next trials.

Regarding this axis, the SISTM team compares to DRACULA, BIOCORE, MONC and COMPO Inria team. However, differences arise in two ways 1/ the application field is immunology and infectious diseases and 2/ we adopt a population approach. This last point results in using simpler models in which it is possible to infer parameters from sparse data by taking advantage of an underlying mechanism common to all patients.

3.3 Axis 3 - Translational vaccinology

The specific objectives are:

- To accelerate the vaccine development by in depth analysis of data generated in early clinical trials and
- designing the next trials with development of new adptative designs and in silico trials in collaboration with immunologists and clinicians.

Vaccines are one of the most efficient tools to prevent and control infectious diseases, and there is a need to increase the number of safe and efficacious vaccines against various pathogens. However, clinical development of vaccines - and of any other investigational product - is a lengthy and costly process. Considering the public health benefits of vaccines, their development needs to be supported and accelerated. During early phase clinical vaccine development (phase I, II, translational trials), the number of possible candidate vaccine strategies against a given pathogen that needs to be down-selected is potentially very large. Moreover, during early clinical development there are most often no validated surrogate endpoints to predict the clinical efficacy of a vaccine strategy based on immunogenicity results that could be used as a consensus immunogenicity endpoint and down-selection criterion. This implies considerable uncertainty about the interpretation of immunogenicity results and about the potential value of a vaccine strategy as it transits through early clinical development. Given the complexity of the immune system and the many unknowns in the generation of a protective immune response, early vaccine clinical development nowadays thus takes advantage of high throughput (or “omics”) methods allowing to simultaneously assess a large number of response markers at different levels (“multi-omics”) of the immune system. Outside of the context of emergency vaccine development during a pandemic, this has induced a paradigm shift towards early-stage and translational vaccine clinical trials including fewer participants but with thousands of data points collected on every single individual. This is expected to contribute to acceleration of vaccine development thanks to a broader search for immunogenicity signals and a better understanding of the mechanisms induced by each vaccine strategy. However, this remains a difficult research field, both from the immunological as well as from the statistical perspective. Extracting meaningful information from these multi-omics data and transferring it towards an acceleration of vaccine development requires adequate statistical methods (in close collaboration with axis 1), state-of-the art immunological technologies and expertise, and thoughtful interpretation of the results.

Our main current areas of application here are early phase trials of HIV and Ebola vaccine strategies, in which we participate from the initial trial design to the final data analyses. We are also involved in the development of next-generation pan-Coronavirus vaccines.

In regards of the number of trials we are dealing with, the complexity of the data (including clinical and biological high dimensional data), the need for a collaborative tool for data sharing that is respectful of GDPR and health data protection, we have set up a data warehouse system based on the Labkey solution (also used for the Immunespace funded by the NIH). We are currently plugging in our data analysis and data vizualization tools. This solution may constitute a very nice way to boost our collaborations but also to facilitate the access to the statistical tools we have developped.

To our knowledge, our specific application to vaccine trials is unique in France. Although some research teams have sometimes applications in this field (e.g. clinical epidemiology team at Inserm U1018 or Inria DRACULA team), there are less devoted to it. Internationally, the closest group to SISTM research axis 3 is the vaccine and infectious disease division of the Fred Hutchinson Institute (Seattle). There are also several groups working on systems immunology mainly in United States such as Mark Davis at Stanford University, Bali Pulendran at Emory University, Rafick Sekaly at Case Western Reserve

University, Galit Alter at the Ragon Institute. There are all immunologists integrating bioinformaticians in their groups therefore they are more applying than developping new methods. We have collaborated with several of these groups.

4 Application domains

The main application domain is the clinical immunology of infectious diseases and more specifically vaccine development.

The main infectious diseases concerned up to now are:

- Human Immunodeficiency Virus (HIV);
- Ebola virus (following the 2014 epidemics);
- SARS-Cov2 virus;
- Hepatitis B virus;
- NIPAH virus;

This is not a closed list and new studies are currently settled on other infectious agents (e.g. tuberculosis, Human Papilloma Virus...).

5 Social and environmental responsibility

5.1 Footprint of research activities

National and international programs

- **Coordination of the response to the Referral for primary care clinical research in France - Ministry of Health (September 2021 - April 2022)**: The objective was to make proposals to anticipate the implementation of future ambulatory trials in response to an emerging infectious disease and enable them to reach their recruitment targets quickly, and to structure research in primary care more broadly. The response includes a national and international review of COVID-19 ambulatory research and 20 proposals on research strategy, its structuring and the removal of budgetary and regulatory constraints.
- **Participation in Delphi consensus groups**: The objective was to extend the CONSORT and SPIRIT recommandations. Participated in the elaboration of SPIRIT/CONSORT Extension for Surrogate endpoints (2023)

5.2 Impact of research results

Drug licensure and patents

- Participant as "Inventor" (Décret n°96-858 du 2 octobre 1996) to the development and the authorization for commercialization (1/7/2020) of the Janssen Zabdeno® (Ad26.ZEBOV) and Mvabea® (MVA-BN-Filo) vaccines against Ebola virus infection.
- Patent 20 306 527.1 on "Use of CD177 as biomarker of worsening in patients suffering from COVID-19" (10/12/2020)

Public/Private partnership

- In the context of clinical trials: Johnson and Johnson (IMI-2 Anti-Ebola vaccine trial Ebovac and Prevac; Merck (Anti-Ebola vaccine trial Prevac/Prevac-up); Iliad Biotechnologies (Anti-pertussis vaccine trial BPZE-1); Gilead Sciences (IP-Cure-B)
- In the context of CIFRE PhD funding: Ipsen (LR HS, 2020-2023). Thesis defended in 2023.

Multicenter clinical trials on vaccine research

- Coordination clinical trials through the Euclid/F-CRIN, CIC1401 platform: Leading Phase II international clinical trials (steering and methodology) for projects BPZE-1, Ebovac2, IP-Cure-B, Prevac, Prevac-Up et PrimalVac (see fundings section).
- Methodology for clinical trials:
  - International phase II anti-Ebola vaccine trial PREVAC (NCT02876328) and EDCTP2 PREVAC-UP
  - International phase I anti-Malaria vaccine trial PRIMALVAC (NCT02658253)
  - French Phase I/II anti-HIV vaccine trial ANRS VRI01 (NCT02038842)
  - French Phase I anti-HIV vaccine trial ANRS VRI06 (NCT04842682)
  - Monocenter anti-pertussis phase I vaccine trial BPZE-1 (NCT02453048)
  - French phase II anti-pneumococal vaccine trial PNEUMOVAS (NCT03069703)
  - French phase II anti-pneumococal vaccine trial SPLENEVAC2 (NCT03873727)
  - French phase II anti-meningococcal vaccine trial SPLENMENGO (NCT04166656)
  - French phase II anti-HPV vaccine trial PRIMAVERA (NCT01687192)
  - French Phase IV anti-Dengue vaccine trial (LR, trial set-up ongoing)
  - Cohort study of anti-COVID-19 vaccination in specific populations (ANRS0001S COV-POPART)
  - Cohort study of HIV infected patients in Nouvelle-Aquitaine (ANRS CO3 Aquitaine)
  - Cohort study of HIV-2 infected patients in France (ANRS CO5 VIH-2)
  - Cohort study of co-infected patients with HIV and Hepatitis in France (ANRS CO13 HEPAVIH)
  - International phase II proof of concept trial IP-cure-B . Educating the liver immune environment through TLR8 stimulation followed by NUC discontinuation. (ANRS HB 07 IP-Cure-B Trial)
  - French phase I anti-SARS-COV2 nasal vaccine trial MUCO-BOOST.

6 Highlights of the year

Modelling in-host SARS-Cov2 neutralisation Because SARS-CoV-2 constantly mutates to escape from the immune response, there is a reduction of neutralizing capacity of antibodies initially targeting the historical strain against emerging Variants of Concern (VoC)s. That is why the measure of the protection conferred by vaccination cannot solely rely on the antibody levels, but also requires to measure their neutralization capacity. Here we used a mathematical model to follow the humoral response in 26 individuals that received up to three vaccination doses of Bnt162b2 vaccine, and for whom both anti-S IgG and neutralization capacity was measured longitudinally against all main VoCs. Our model could identify two independent mechanisms that led to a marked increase in measured humoral response over the successive vaccination doses. In addition to the already known increase in IgG levels after each dose, we identified that the neutralization capacity was significantly increased after the third vaccine administration against all VoCs, despite large inter-individual variability. Consequently, the model projects that the mean duration of detectable neutralizing capacity against non-Omicron VoC is between 348 days (Beta variant, 95% Prediction Intervals PI [307; 389]) and 587 days (Alpha variant, 95% PI [537; 636]). Despite the low neutralization levels after three doses, the mean duration of detectable neutralizing

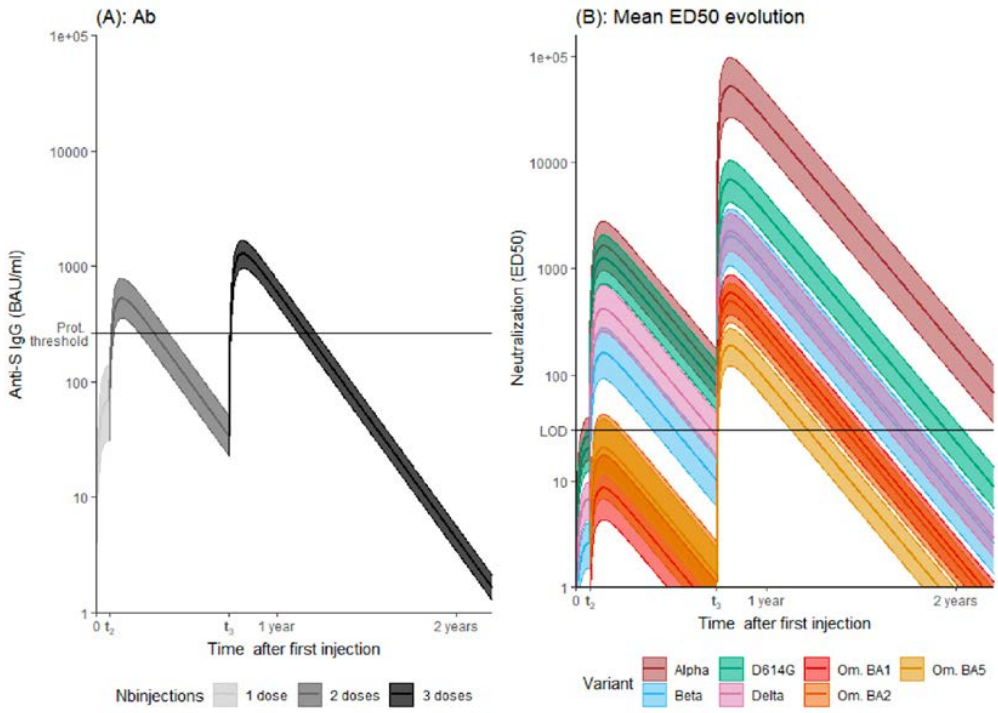

Figure 2: A: Predicted evolution of binding antibody concentration. B: Predicted kinetics of ED50. The shaded area is the 95% prediction interval. DOI

capacity against Omicron variants varies between 173 days (BA.5 variant, 95% PI [142; 200]) and 256 days (BA.1 variant, 95% PI [227; 286]). Our model shows the benefit of incorporating the neutralization capacity in the follow-up of patients to better inform on their level of protection against the different SARS-CoV-2 variants [22].

7 New results

7.1 High-dimensional statistical learning

7.1.1 Valid inference in high-dimension

Statistical inference in high dimension remains challenging. In particular, the sound analysis of transcriptomic data requires valid statistical testing procedures with well calibrated control of of the type I error. In addition, the analysis of high-dimensional longitudinal and time-to-event data bears its own layer of complexity. Another example of the complexity of the analysis of such data is the so-called “double dipping”, i.e. using the same data twice, first to identify clusters and second to identify statistically significant variables diffentiating those clusters. This leads to poor statistical performances overall without carefully adjusting for this double step – yet it is widely used in transcriptomic data analysis for instance. We are working on characterizing the limits for the application of new approaches (namely data-fission and data-thining) to practical post-clustering inference.

High dimensionnal data, such as transcriptomic data, comprise a tremendous amount of biological information. While this information can be hard to pinpoint, hidden by a small signal-to-noise ratio, it still carries the promise of clinical utility. Therefore, we have started to investigate their potential as surrogate markers in vaccine trials. A surrogate marker is a marker that can be measured earlier and/or more easily than the original clinical outcome, while retaining the ability to reliably assess the impact of a treatment. Those bear a particular interest in interventional studies (eg vaccine trials) where multiple omics data are measured a few hours or days after the intervention as it could significantly accelerate future studies. We have made new developments towards establishing a method for investigating and

validating high-dimensional markers, such as gene expression data or cellular marker data, as surrogate [12].

7.1.2 Optimal transport for cell-type proportions inference

Variations of cellular population abundance is of critical importance in following the immune state of vaccine trial participants. It is usually measured from flow-cytometry or mass-cytometry measurements, high-throughput single-cell technologies. Raw output data are manually processed through specific sequences of 2-dimensional projections form measured cellular markers, an expensive, subjective and time-consuming approach. We proposed new developments of optimal transport to infer mixture proportions robust to domain changes[25], a new methodology that can be applied to generalize a cytometry manual gating to newly measured, ungated samples.

7.1.3 Machine learning applied to EHR

Given the recent advances in machine learning, notably thanks to deep-learning and large language models, we are questioning the impact and interest of such approaches for biomedical research and training [18, 35].

7.1.4 Microbiota data

Recently, the team has been interested in microbiota data analysis. Several poster communications in international conferences showcased our effort in reviewing the literrature [47, 48, 49]

7.2 Mechanistic learning

7.2.1 Parameters estimation in mechanistic models

The stakes of estimating parameters in non-linear mixed effects models are high as it is a critical step in understanding the underlying processes that generate the data. Accurately estimating the parameters allows for making valid inferences and predictions about the system being studied. Inaccurate parameter estimates can lead to incorrect conclusions and biased results. A major pitfall of classic methods such as maximum likelihood is their inability to account for model error / misspecification during parameter inference. This could have a dramatic effect on estimation accuracy. To this end, Clairon et al. 2023 [21] proposed an estimation method based on optimal control theory to acknowledge model error presence at the subject level in non-linear mixed effect models.

7.2.2 Within host modeling of SARS-CoV-2 and Ebola

One result of the year was the better understanding of virus dynamics for SARS-CoV-2 variants of concern. Using non-human primates data, in Marc et al. 2023 [28], a mathematical model was developed to describe the observed dynamics, revealing that VoC, except for beta, showed an escape from the immune response. Notably, the delta variant exhibited increased viral production, while the omicron variant showed a significant reduction. The increased infectivity of omicron is likely to be mediated by an increase infectiousness of the virus.

Later we dedicated our work to improve the understanding of humoral response after vaccination. In Alexandre et al. 2023 [13] we demonstrated that the durability of antibodies response after the two-dose heterologous Ad26.ZEBOV,MVA-BN-Filo vaccination against Ebola may may last more than 15 years. In Clairon et al. 2023 [22], the mathematical model revealed two mechanisms contributing to increased humoral response with successive vaccine doses: elevated IgG levels and enhanced neutralization capacity after the third dose against all VoCs. The model predicts a considerable duration of detectable neutralizing capacity against non-Omicron VoCs, ranging from 348 to 587 days. However, for Omicron variants, detectable neutralization capacity lasts between 173 and 256 days.

7.2.3 Between host modeling of COVID-19 Epidemics

We investigated the effectiveness of non-pharmaceutical interventions (NPIs) implemented in France during the COVID-19 pandemic. We obtained similar, yet to be discussed results with two competing approaches. In Paireau et al. 2023 [29], we compute the reproductive number of the epidemics over time and utilize a log-linear model to predict it as a function of NPIs. In Collin et al. 2023 [23] we use an extended Susceptible-Infectious-Recovered (SIR) model with a dynamic transmission rate estimated with population Kalman filters appraoch. In both studies, we highlight the effectiveness of NPIs in France and the role of weather in reducing transmission in summer, and note increased transmissibility due to VoC. Works are currently ongoing to compare the approaches with a fully mechanistic one-step estimation using SAEM algorithm. On a side note, we investigated in Wang et al. 2023 [36] the use of network features to better understand the spread of an epidemic on a network. We found that features concerning infections and distance to infected individuals are most informative.

7.3 Translational vaccinology

7.3.1 Vaccine trial results and milestones

The ANRS VRI06 first-in-human phase I trial of a novel vaccine concept targeting dendritic cells (here as HIV vaccine) has completed all W48 follow-up visits and results are available. A manuscript is ready for submission to the Lancet HIV. We have prepared the designs of subsequent clinical trials of this vaccine candidate as well as of other candidates based on the same platform (HPV vaccine, SARS-COV2 and Pansarbeco vaccines ; Nipah vaccine).

A large randomized multi-arm Ebola vaccine trial (Prevac trial, evaluating three different Ebola vaccine strategies ; and its long-term follow-up in the EDCTP funded Prevac-Up project) has reached the 5-year follow-up visits for all participants. Results will become available in 2024.

Modeling of PREVAC antibody response and its associated factors has been presented at ASTMH 2023 (Chicago), and a manuscript is prepared for publication in Nature Communications.

We published systems vaccinology results of the Ebovac2 project (Blengio et al. 2023) [16].

Two new international projects, in which SISTM is a partner, have been funded : Horizon Europe Solve and CEPI Musicc Consortium. The MUSICC consortium has been funded by CEPI to develop and conduct Controlled Human Infection Models For Beta-coronaviruses in order to assess vaccine effects. The Solve project has been funded by the EU (Horizon Europe), to decipher the mechanisms of induction of long-lasting immunity through a comparison of vaccine platforms and to bring new vaccine concepts.

7.3.2 Covid treatment trials

Linda Wittkop is deputy head of the International coordinating centre ANRS MIE in STRIVE (Strategies and Treatments for Respiratory Infections and Viral Emergencies) which is a continuing global trials network consisting of a diverse mix of over 40 high-, middle- and low-income countries. STRIVE aims at improving the clinical outcomes of patients with infections while being prepared to respond to infectious disease emergencies, through the rapid implementation of clinical trials designed to inform practice guidelines, public health policy, and the delivery of health care.

7.3.3 Methodological results and milestones

Rodolphe Thiebaut co-coordinates one axis (SMATCH) within PEPR Santé Numérique. This includes PhD funding for the development of Bayesian adaptive phase I-II trials for early vaccine development. The PhD project of Cyrille Kone (co-directed with Emilie Kaufmann, Scool team, Inria Lille) on the development of novel early phase vaccine trial designs based on bandit algorithms, started in Nov 2022, has produced its first results on theoretical work for using reinforcement learning methods for adaptive trial designs, which were presented at Neurips 2023 [41].

7.3.4 Knowledge transfer

We have set-up a transfert unit (BVA, Bordeaux Vaccine Analytics) with Adera, University of Bordeaux, to facilitate the collaborations with private companies. We have continued to develop a data warehouse

system based on the Labkey solution where all raw data are organized and that includes meta-data on the design of the clinical trials and is used in international collaborations of facilitate data sharing and exploration (EHVA, EBOVAC, IP-Cure-B and CARE consortia).

Participants: Marta Avalos, Quentin Clairon, Robin Genuer, Boris Hejblum, Edouard Lhomme, Mélanie Prague, Laura Richert, Rodolphe Thiébaut, Linda Wittkop.

8 Partnerships and cooperations

8.1 International initiatives

8.1.1 Associate Teams in the framework of an Inria International Lab or in the framework of an Inria International Program

Title: DESTRIER DEfining Surrogacy of early Transcriptomics foR vaccInE Response

Duration: 2022 ->

Coordinator: Denis Agniel

Inria contact: Boris Hejblum

Partners: RAND Corporation (États-Unis)

Summary: This project seeks to develop statistical methods to evaluate to which extent can transcriptomics be used to capture vaccine effects. Gene expression is central to protein production and largely determines cellular function: it is thus a promising biomarker for quickly measuring effects of vaccines. Validated transcriptomic signatures could thus be developed to dramatically speed up vaccine trials for emerging infectious diseases like Ebola or COVID-19. Such a technology could also be used for identifying good vaccine responders in health care workers that can be deployed in case of an epidemic emergency, or identify poor responders among vaccine recipient that would benefit from an additional booster dose. In this project, we set to develop novel statistical methods for assessing the surrogacy potential of transcriptomic data in vaccine research. We will first develop methods to quantify how much of the vaccine effect is mediated by gene expression, establishing if gene expression is suitable for capturing the vaccine's effect. We will develop model-free approaches to estimating this quantity which will remove many of the modeling assumptions typically used in high-dimensional mediation analysis. Second, we will develop methods to construct an optimal gene expression signature for capturing the vaccine effect, and we will develop methods to operationalize its use in future studies, establishing how to build and use such a transcriptomic signature. These methods will similarly take advantage of modern machine learning approaches and doubly robust estimation to provide model-free estimators of key quantities. We will use these methods to study high-impact clinical trials from the Vaccine Research Institute in the context of Ebola and COVID-19.

8.1.2 STIC/MATH/CLIMAT AmSud projects

Participation to the Regional Program MATH-AmSud 2023, Chile, Uruguay, France. Project duration until December 2025.

8.1.3 Participation in other International Programs

In 2023, the MUSICC consortium has been selected for funding by CEPI (Coalition for Epidemic Preparedness Innovations). This project, in which SISTM is a partner, is in the last steps of its Grant and Consortium Agreement preparation. The objective of this project will be to develop and conduct Controlled Human Infection Models For Beta-coronaviruses in order to assess vaccine effects. We expect the project to be starting February 1st, 2024.

In addition, SISTM has continued working on the following projects: NIPAH (Chine) – Scientific co-operation program France/Chine. (2019-2022). M. Prague is workpackage co-PI - Sino-French Agreement Aviesan. Sep 2018 – Dec 2023, 150,000 euros. To raise the challenge caused by Nipah virus we propose to develop a program that shall led to a better understanding of the epidemiology of the virus as well as the associated physiopathology. To develop new tools in the field of diagnosis, treatment and prevention of the infection. This grant aims at funding a 2 years of postdoc, travel and equipment expenses.

8.2 International research visitors

8.2.1 Visits of international scientists

Julien Martinelli

Status: post-Doc

Institution of origin: Aalto University - Probabilistic Machine Learning

Country: Norway

Dates: Oct 2023

Context of the visit: Start a collaboration on Gaussian processes

Mobility program/type of mobility: research stay

Baptiste Elie

Status PhD

Institution of origin: MIVEGEC, CNRS IRD

Country: France

Dates: March 2023

Context of the visit: Collaboration on HPV modeling

Mobility program/type of mobility: PhD student exchange

John Fricks

Status Associate Professor

Institution of origin: School of Mathematical and Statistical Sciences, Arizona State University

Country: USA

Dates: Sept 23 - June 24

Context of the visit: Development of stochastic models for biological dynamics

Mobility program/type of mobility: sabbatical, IDEX UB

Camila Labarca

Status intern (eng)

Institution of origin: Universidad de Chile

Country: Chile

Dates: January 2023 - April 2023

Context of the visit: collaboration of Inria Chile, the International Relations Department (DRI) of Inria in France, Chilean universities on partnership and the Inria Teams in France

Mobility program/type of mobility: internship

Marie Weskamm

Status researcher

Institution of origin: IIRVD - Institute for Infection Research and Vaccine Development University Hospital Hamburg (UKE)

Country: Germany

Dates: April-May 2023

Context of the visit: Visit to the Sistm team

Mobility program/type of mobility: research stay

Rebecca Knowlton

Status PhD student

Institution of origin: University of Texas at Austin — Department of Statistics Texas

Country: USA

Dates: Nov 2023

Context of the visit: DESTRIER Inria Associate Team

Mobility program/type of mobility: research visit

8.2.2 Visits to international teams

Research stays abroad From 22nd to 26th May, R. Thiebaut has been visited the departement of Epidemiology, Biostatistics and Occupational Health at McGill University. He is now adjunct professor there. McGill university is a major partner of the Graduate's program Digital Public Health that RT is leading.

8.3 European initiatives

8.3.1 H2020 projects

SOLVE: In 2023, the SOLVE project, in which SISTM is workpackage leader (WP7 Data Science), has been selected by Horizon Europe for funding. It is currently finishing its negotiation with the EU (Horizon Europe). The project is expected to start on January 1st, 2024 to decipher the mechanisms of induction of long-lasting immunity through a comparison of vaccine platforms and to bring new vaccine concepts.

IP-CURE-B: Immune profiling to guide host-directed interventions to cure HBV infections. Co-ordinated by Inserm, the project includes a total of 13 Beneficiaries: Centre Hospitalier Universitaire Vaudois (Switzerland), Karolinska Institutet (Sweden), Institut Pasteur (France), Universita degli studi di Parma (Italy), Fondazione IRCCS CA' Granda – Ospedale maggiore policlinico (Italy), Universitaetsklinikum Freiburg (Germany), Ethniko Kai Kapodistriako Panepistimio Athi-non (Greece), Fundacio Hospital Universitari vall d'Hebron (Spain), Gilead Sciences Inc. (USA), Spring Bank Pharmaceuticals, Inc (USA), European Liver Patients Association (Belgium), Inserm Transfert SA (France). L Wittkop. Duration: 60 months 01/01/20-31/12/24. 409,632 Euros.

EHVA (European HIV Vaccine Alliance): European HIV Vaccine Alliance: a EU platform for the discovery and evaluation of novel prophylactic and therapeutic vaccine candidates. Coordinator: Inserm/University of Lausanne. Other partners: EHVA consortium gathers 41 partners. R. Thiébaut. Duration: 66 months. 01/01/2016 - 31/12/23 – 208 686 euros.

8.3.2 Other european programs/initiatives

EBOVAC1: Development of a Prophylactic Ebola Vaccine Using an Heterologous Prime-Boost Regimen. Coordinated by London School of Hygiene & Tropical Medicine (United Kingdom). Other beneficiaries: Janssen a Pharmaceutical Companies of Johnson and Johnson, The Chancellor, Masters and Scholars of the University of Oxford (United Kingdom), Inserm (France), University of Sierra Leone (Sierra Leone), R. Thiébaut. Duration: 84 months. 01/12/2014 - 31/05/2023. 552,050 Euros.

EBOVAC3: Bringing a prophylactic Ebola vaccine to licensure. Coordinated by the London School of Hygiene & Tropical Medicine (United Kingdom). Other beneficiaries: Janssen a Pharmaceutical Companies of Johnson and Johnson, Inserm (France), The University of Antwerpen (Belgium), University of Sierra Leone (Sierra Leone), R. Thiébaut. Duration: 60 months. 01 /06 /2018 - 30 /06/2023. 351,274 Euros.

PREVAC-UP: The Partnership for Research on Ebola VACcinations-extended follow-UP and clinical research capacity build-UP. SISTM is also involved in PREVAC-UP, an EDCTP2 project in direct link with the research carried out on the Ebola vaccines. Coordinated by Inserm (France). Other beneficia- ries: CNFRSR (Guinea), CERFIG (Guinea), LSHTM (UK), COMAHS (Sierra-Leone), NIAID (USA), NPHIL, (Liberia), USTTB (Mali), Centre pour le Développement des Vaccins (Mali), Inserm Transfert SA (France), R. Thiébaut. Duration: 60 months. 01 /01 /2019 - 31 /12 /2023. 328,000 Euros.

CARE: Corona Accelerated R&D in Europe is an IMI2 funded project coordinated by Inserm which gathers 36 globally renowned academic institutions, pharmaceutical companies and non-profit research organisations which have committed to rapidly and efficiently address the COVID-19 emergent heath threat. This major initiative aims at addressing two key objectives: the development of therapeutics to provide an emergency response towards the current COVID-19 pandemic and the development of therapeutics to address the current and/or future coronavirus outbreaks. To address both goals, the CARE consortium has carefully designed a comprehensive research and development (R&D) program around thoughtfully designed Target Product Profiles (TPP) of the urgently needed antiCOVID-19 drugs. This includes small and large molecule discovery and Phase 1 and 2 clinical trials centred around three main pillars: drug repositioning, small-molecule drug discovery, and virus neutralising antibody discovery. These pillars reflect a bifocal strategy where efforts are geared towards (a) a rapid response against current COVID-19 pandemic and (b) a longer-term preparedness strategy against future coronavirus outbreaks. This will maximize the screening landscape of relevant therapeutic avenues and ensure effective therapeutics can be rapidly identified, pre-clinically tested and optimised for clinical-grade manufacturing and clinical testing. In this project, SISTM and EUCLID are working closely together with the support of the CREDIM in the WP5, W7 and WP8 with the respective objectives of providing statistical analysis and data modelling of the immune assays carried out in the project, bring some expert support to the clinical work and develop a LabKey-based platform for the integration and management of the data. Duration: 60 months. 01/04/2020 - 30/03/2025. 1,256,003 Euros.

ASCENT: Acceleration of Novel Coronavirus Serological Test Development and Seroprevalence Study: An African-European Initiative. ASCENT is an EDCTP2 projects involving 7 partners (Inserm, CHUV, EuroVacc, Utrecht University, Centre Muraz, SAMRC and CERFIG) from 6 different countries in Africa and Europe which will aim at assessing the real prevalence of the infection, the projection of the immunity acquired by the populations, and the evaluation of measures aimed to break the transmission in Africa. To do so ASCENT will implement in Burkina Faso, South Africa and Guinea, a novel robust and reproducible luminex-based serological diagnostic test with high throughput, sensitivity, specificity and rapid turn- around time. In this project, SISTM will be involved in statistical analysis of the tests data and will lead the WP3 which aims at modelling the epidemics. Duration: 01/05/2020 - 31/01/2024. 37,500 Euros.

CoVICIS: The CoVICIS (EU-Africa Concerted Action on SAR-CoV-2 Virus Variant and Immunological Surveillance) program is proposing a global approach with a powerful state-of-the-art virologic and immunologic platforms coupled with large genomic surveillance studies and diverse cohorts

in EU and SSA CoVICIS aims to contribute to the early identification of emerging VOC and address key unanswered questions regarding: i) the susceptibility to infection with VOC after a prior infection in the setting of a long-COVID or after vaccination with different vaccines, ii) the risk posed by VOC in immunocompromised patients, and iii) the modalities of infection and immune responses in children. CoVICIS is coordinated by the CHUV (Switzerland). SISTM is involved in WP7 Data Science and Analysis which aims to utilize cutting edge computational and statistical analysis method to obtain comprehensive assessment of immunogenicity and immune correlates of protection. 11/2021-10/2024. Coordinated by the CHUV, CoVICIS counts 14 partners amongst which we can find Inserm, UNIMI, UNIGE, and 4 South-African partners. Total budget: 10M€, SISTM budget: 110k€.

8.4 National initiatives

Labex Vaccine Research Institute (VRI): Funded by the PIA under Laboratory of excellence initiative, VRI conducts research to accelerate the development of effective vaccines against HIV/AIDS and (re)-emerging infectious diseases. The SISTM team is leading the Data science division of the VRI. To this purpose, SISTM has established strong collaboration with immunologists. SISTM carries out biostatistical analysis of the data produced by the different other VRI teams together with a modelling approach of the immune response to the vaccines or other interventions. 2012-2024, Main partners: the VRI was established by the French National Agency for Research on AIDS and viral hepatitis (ANRS – France Recherche Nord & Sud Sida-HIV Hépatites) and the University of Paris-Est Créteil (UPEC). The other partners of the VRI are CEA, Inserm, Pasteur Institute, the University of Bordeaux, the Baylor Institute for immunology research and the University of Strasbourg. Total budget: 75M€, SISTM budget: 1.85M€ (about 170k€ a year since 2012).

RHU SHIVA: Since November 2019, the RHU SHIVA aims to better understand the determinants and consequences of cerebral small vessel disease through innovative imaging, molecular and analytical approaches, to develop new personalized diagnostic and preventive strategies and to accelerate the discovery of novel therapeutic targets. RT as workpackage leader is collaborating in the SHIVA RHU to work on the integration and systems biology of small vessel disease biomarkers. 12/2019-11/2024. Main partners: University of Bordeaux, Inserm, CNRS, CHU Bordeaux, CHNO AP-HP, Fealinx, Qynapse, ImaginEyes. Total budget: 8.2 M€, SISTM budget: 190k€.

Ecole Universitaire de Recherche "Digital Public Health" Funded under the PIA3 The Digital Public Health Graduate Program provides an interdisciplinary and international training from Master to Doctorate in epidemiology, biostatistics, computing and social sciences to explore the impact of digital public health on society. The whole program is directed by Rodolphe Thiébaut. The whole SISTM team is implicated in these activities. 2018-2028. Main partners: University of Bordeaux, Inserm, Inria, Sciences Po Bordeaux and University Bordeaux Montaigne. Total budget: 4.52 M€, SISTM budget: The budget is mostly dedicated to grants to students, running costs and indemnification of teachers.

Project Emergen: Consortium for Surveillance and Research on EMERgent Pathogens via Microbial GENomics. EMERGEN, coordinated by Sante publique France and ANRS-Emerging Infectious Diseases, aims to deploy a genomic surveillance system for SARS-CoV-2 infections throughout France. Its main objective is to to follow the genetic evolution of the SARS-CoV-2 virus in order to detect the emergence and the spatio-temporal distribution of variants, i.e. viruses with mutations likely to have functional consequences, such as infectivity, contagiousness, virulence or immune escape. In this context, the role of SISTM will be to contribute to modelling the impact on epidemic dynamics of SARS-CoV-2 variants based on the estimation of their characteristics. 01/2022-01/2024. Main partners: Santé Publique France, Inserm/ANRS, APHP, HCL, Pasteur Institute, Anses, IFB, CNRGH/CEA, Réseau Sentinelles. Total budget: 10M€, SISTM budget: 56k€.

PEPR Santé Numérique SMATCH: The PEPR SN SMATCH coordinated by Inria and co-coordinated by Sarah Zohar (HEKA) and Rodolphe Thiébaut (SISTM) is part of the France 2030 initiative to develop digital health in France. SMATCH objectives are to develop and apply statistical and

AI-based methods with the ultimate goal of accelerating the development of medical interventions (drugs and digital medical devices) during their evaluation in clinical trials based on the following assumptions: 1. The use of information generated in preclinical studies (animal studies, organoids, in silico studies) combined with adaptive designs should help the early phases of development; 2. The integration of multi-source data including real-world and in silico data should help to complete trials; 3. Specific adaptive designs should be defined for the evaluation of digital medical devices based on learning algorithms. The consortium counts 16 teams mainly from Inria and Inserm Centers recognized in this field, bringing a unique and complementary expertise in data sciences and AI applied to health problems and specifically to clinical trials. In addition, links with the regulatory bodies involved are already established within the consortium (e.g. HAS) and outside (e.g. EMA). Finally, many connections exist with the other axes of the PEPR Digital Health and more generally with the projects carried out within the framework of the digital health acceleration strategy. Thus, by providing innovative and adapted methodological tools that will have already been applied in a real context, we hope to participate in the acceleration of clinical research leading to major societal and economic impacts. 01/09/2023 – 31/08/2029. Total budget : 3M€, SISTM budget: 693 996 €

PIEEC MEDITWIN: MEDITWIN is a Projet Important d'Intérêt Européen Commun (PIEEC) part of the France 2030 strategy coordinated by Dassault Systems and Inria. The aim of the MEDITWIN project is to develop and validate digital twins to support personalised medical practices and strengthen the healthcare system in targeted therapeutic areas. These virtual twins will be multi-disciplinary and multi-physiological, and will be based on real clinical data, acquired prospectively and historically, at the molecular, genetic, cellular and tissue levels, right down to the organ, system, individual and population level. They will be based on structured, interoperable data hosted in sovereign infrastructures. In this frame, SISTM will develop innovative methods for adaptive clinical study designs for pilot (feasibility) and perpetual (after initial validation) clinical trial designs for the evaluation of patients' risk confronted to SaMD updates in collaboration with HEKA. 2024-2029, SISTM budget: 433 125 €

IHU VBHI : The Vascular Brain Health Institute (VBHI) is a joint-venture between the University of Bordeaux (UB), Bordeaux University Hospital (CHUB), the national institutes for medical and digital science research (Inserm, Inria), and the New Aquitaine region, aiming to create a Center of Excellence on Vascular Brain Health. It will establish an entirely novel paradigm to prevent stroke and dementia, two leading causes of death and disability worldwide, by taking a precision population health approach and leading an emerging global dynamic geared towards both innovation and inclusion. 11/2023-10/2032. The SISTM team will be involved mostly in WP1 to contribute to the analysis of high dimensional data and notably by conducting extensive bioinformatics analyses, including an original pipeline to identify miRNA-based candidate treatments for identified targets. In addition, the team will be involved in the design of omics- guided clinical trials design. Total budget: 40 M€ overall.

8.4.1 Various Partnership

Mélanie Prague: Chaire Digital Innovation and Health Data Science program of the Center for Applied Mathematics CMAP at the Ecole Polytechnique
The project team members are involved in:

- **F-CRIN** (French clinical research infrastructure network), initiated in 2012 by ANR under "Programme des Investissements d'avenir". (L Richert).
- **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).
- **TARPON** (Traitement Automatique des Résumés de Passages aux urgences pour un Observatoire National), laureate project from the 2nd Health Data Hub calls for projects, great challenge "Improving medical diagnostics through Artificial Intelligence" and Bpifrance 2020-2022, extended in

2023. (Principal PI E. Lagarde Inserm U1219 in collaboration with University Hospital of Bordeaux. Marta Avalos is listed as a collaborator).

- **CESIR IV** (Combination of Studies on Health and Road Safety - 4th project) funded by ONISR DSR. (Principal PI E. Lagarde Inserm U1219. Marta Avalos is listed as a collaborator).
- **EMERG** (Exposome microbien et Risque sanitaire : intérêt d'une Gestion One Health des enjeux liés aux grippes zoonotiques) funded by PSGAR (Programmes Scientifiques de Grande Ambition Régionale). (Principal PI L Delhaes and D Malvy. Marta Avalos is listed as a collaborator).
- Collaboration with **Inserm PRC** (pôle Recherche clinique).
- Collaboration with **Inserm REACTing** (REsearch and ACTion targeting emerging infectious diseases) network.
- Collaboration with **Inserm RECap** (Recherche en Epidémiologie Clinique et en Santé Publique) network.
- **STRIVE** (Strategies and Treatments for Respiratory and Viral Emergencies Study Payments). International Network for respiratory and viral emergency studies. (Collaborator: Linda Wittkop).

Participants: Marta Avalos, Quentin Clairon, Robin Genuer, Boris Hejblum, Edouard Lhomme, Mélanie Prague, Laura Richert, Rodolphe Thiébaut, Linda Wittkop.

9 Dissemination

9.1 Promoting scientific activities

9.1.1 Scientific events: organisation

General chair, scientific chair Robin Genuer is co-president of the organization committee of Journées de la Statistique 2024 (about 500 participants) in Bordeaux May 2024.

Member of the organizing committees Edouard Lhomme organized the first ANRS MIE Respiratory viruses Workshop – Paris, March 31 2023 (110 persons) on behalf of ANRS-MIE

Edouard Lhomme and Laura Richert organized the EUCLID annual scientific day, Bordeaux Nov 2023

Linda Wittkop was session Co-Chair in the 10th international HBV cure ANRS workshop: Session I: Clinical trial update, 22 Mai 2023, Paris, France and Co-Chair of the 'Hepatitis' session in IWHOD International Workshop on HIV and Hepatitis Observational Databases, 23-25 March 2023, Athens, Greece.

9.1.2 Scientific events: selection

Member of the conference program committees Rodolphe Thiébaut is a member of the scientific committee of the IWHOD International Workshop on HIV Observational Databases since 2013.

Boris Hejblum was member of the program commitee of CNC23 9th Channel Netwok Conference of the International Biometric Society 2023

Marta Avalos was member of the program commitee of

- DATAQUITAINE, Bordeaux, Feb 2023
- FLAIRS-36 (The 36th International Conference of the Florida Artificial Intelligence Research Society), AI in Healthcare Informatics track, Florida, May 2023
- Conférence d'apprentissage automatique, Cap2023, Strasbourg, Jul 2023
- Machine Learning for Health (ML4H) New Orleans, Dec 2023

Reviewer Marta Avalos was a reviewer for the conferences:

- Conference on Health, Inference, and Learning CHIL, Cambridge, USA, Jun 2023
- IASE 2023 Satellite Conference, Toronto, Canada, Jul 2023
- IA & santé, AFIA, Strasbourg, France, Jul 2023,

9.1.3 Journal

Member of the editorial boards Mélanie Prague is an Associate Editor of the International Journal of Biostatistics.

Reviewer - reviewing activities Quentin Clairon was a reviewer for the journal Mathematical Biosciences and Engineering (MDPI)

Mélanie Prague was a reviewer for the journals Biometrics, Statistical Methods in Medical research, Plos One, eLife and Plos Computational Biology.

Robin Genuer was a reviewer for the Journal of the American Statistical Association

Laura Richert was a reviewer for Trials

Boris Hejblum was a reviewer for the journals BMC Bioinformatics, Statistics in Medicine and Cell Reports Methods.

Marta Avalos was a reviewer for BMJ Open and the 'Public Health and Epidemiology Informatics' section of the IMIA Yearbook.

9.1.4 Invited talks

Robin Genuer presented a talk 'Towards a supervised clustering of variables', Marie Chavent, Robin Genuer and Jérôme Saracco, Journée de recherche du réseau Impulsion Public Health Data Sciences, Bordeaux, Jul 2023.

Mélanie Prague presented the following invited talks:

- Defining mechanistic correlates of protection. Prague M. World congress on basics and clinical pharmacology. , Glasgow UK, 2-7th July 2023.
- Joint modeling of viral and humoral response in Non-human primates to define mechanistic correlates of protection for SARS-CoV-2. Prague M. M. Alexandre, R. Marlin, Roger le Grand, Y. Levy and R. Thiébaut. Society of mathematical Biology conference, Columbus USA, 16-21th July 2023.
- Utilisation des données de transcriptomique pour informé les modèles mécanistes de réponse immunitaire. R. Thiébaut, B. Hejblum, K. Ba and Prague M. AC modélisation ANRS MIE, Paris, France. 24-25th Oct. 2023. (Invited oral presentation by collaborator)
- Defining Mechanistic correlates of protection MRC Cambridge, Biostatistics department, online, 24 May 2023.

Linda Wittkop presented the following invited talks:

- I-REIVAC webinar: ANRS0001S Cohort COV-POPART:Antibody levels and risk of breakthrough SARS-CoV-2 infection, 23 May 2023.
- I-REIVAC seminar : ANRS0001S Cohort COV-POPART: Défis de la cohorte COV-POPART, 6-7 April 2023.
- Swiss HIV Cohort – Young Investigator seminar: COVID-19 vaccination in specific populations: results from the ANRS0001S COV POPART cohort study, 6 September 2023.
- EUCLID seminar: Strategies and Treatments for Respiratory Infections and Viral Emergencies (STRIVE) (A global clinical trial infrastructure for current and emerging infectious disease threats), 8 December 2023.

Edouard Lhomme presented the following invited talks:

- ANRS MIE Respiratory viruses Workshop (Paris, March 31 2023) | Hightlight on “Over half of known human pathogenic diseases can be aggravated by climate change. Nat. Clim. Chang. (2022)”
- ASTHM 2023 (Chigaco): Symposium ’The use of mathematical modelling for vaccine development: the example of Ebola’
- BPH Seminar (8 Sept 2023, Bordeaux): ’Immunological aspects of Ebola vaccination’.
- VRI day 2023: “Determinants and durability of antibody response to rVSVΔG-ZEBOV-GP and Ad26.ZEBOV,MVA-BN-Filo Ebola virus disease vaccines: a modelling study from the PREVAC randomized trial”

Rodolphe Thiebaut presented the following talks:

- Journée de l’Action coordonnée « Modélisation » ANRS/MIE, November 24th, 2023, How gene expression can inform dynamical models? [communication orale invitée]
- 7eme édition du printemps du CIC, September 8, 2023, Limoges, L’avenir de la Recherche Clinique passe-t-il par l’intelligence artificielle ? [communication orale invitée]
- Modelling Workshop in Infectious Diseases, July 12-13, 2023, Lyon, Forecasting SARS-CoV-2 hospitalizations using EHR : application to Bordeaux hospital [communication orale invitée]
- 6th Workshop on Virus Dynamics, July 4-6, 2023, Nagoya, Japan, How gene expression can inform dynamical models? [communication orale]
- Conférence Internationale des Doyens des facultés de PHARMacie d’Expression Française Mai 30-Juin 2, 2023, Bordeaux, France, Conférence introductive sur l’utilisation de l’IA en santé [communication orale invitée]

9.1.5 Leadership within the scientific community

Mélanie Prague is a member of the Scientific Director board of the COVID immunity task force in Cananda (2021 - 2023). She co-leads with J. Guedj the working group Within-Host Modeling, Action Coordonnée Modélisation (ANRS-MIE) since 2021.

Boris Hejblum is a board member of the “MAchine Learning et Intelligence Artificielle” (MALIA) group of the French Society of Statistics (SFdS)

Hélène Savel is a board member of the “Biopharmacie et Santé” group of SFdS.

9.1.6 Scientific expertise

Rodolphe Thiébaut is

- a member of the Pasteur Institute evaluation committee, a member of the independent committee of international trials ODYSSEY, SMILE, BREATHER, 3D.
- the coordinator of the evaluation committee for the Italian « “Extended Partnerships with universities, research centres and companies to fund basic research projects” - within the framework of the National Recovery and Resilience Plan (160 M€ over 3 years).
- 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).
- 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 scientific council of Muraz’s Center (Bobo-Dioulasso, Burkina Faso).

Mélanie Prague is an expert for:

- Millenium Science Initiative (Chile) Application reviewer for the Natural/Exact Sciences application (3 projets).
- Swiss National Science Foundation (SNSF) (Swiss). Application reviewer (2 projets).
- ANR CES45 Member of evaluation commitee (since 2023), France.
- ANRS MIE CSS13 "Clinical research" Member of evaluation commitee (since 2019), France.
- She participates in the scientific committee for attribution of research grants for delegations, PhD and Postdoctoral fellowships (2017 - 2023). Commission Emploi-Recherche Inria.

Linda Wittkop is a member of the

- external ethics and scientific advisory board of the EU-funded project VACCELERATE,
- CESREES (Comité éthique et scientifique pour les Recherches, les Etudes et les Evaluation dans le domaine de la santé),
- a president/member of the data and safety monitoring boeard of international trials (B-FREE), and
- an expert for the PHRC (Programme hospitalier de recherche Clinique)

Laura Richert is member for the

- CNU 46.04 (Biostatistiques, informatique médicale et technologies de communication) and an expert for
- PHRC (Programme hospitalier de recherche Clinique).

Edouard Lhomme is member of the

- jury of the PHRC (Programme hospitalier de recherche Clinique) organized by French Ministry of Health since 2022
- jury of the FCRIN call for project for platform trials (2023)
- ANRS-MIE joints actions (JA): Modeling JA, Transmission JA, Vaccine Respiratory virosis JA.

Boris Hejblum is a national expert for the 2023 MESSIDORE project call from Inserm IReSP 'Méthodologie des ESSais cliniques Innovants, Dispositifs, Outils et Recherches Exploitant les données de santé et biobanques'.

9.1.7 Research administration

Rodolphe Thiébaut is a member of the Scientific Council of Inserm (since 2017).

Boris Hejblum is a member of the chairing committee of the Société Française de Biométrie, the French Chapter of the International Biometric Society.

Mélanie Prague is a member of the bureau 'Action coordonnée modélisation' (ANRS-MIE) since 2021.

Linda Wittkop is director of the mixed Inserm/Bordeaux university unit UMS 54 Methods and Applied research of Trials and coordinator of the axis 'Infectious diseases and Inflammation' of the CIC1401-E.

Laura Richert is coordinator of the Clinical epidemiology module of the Clinical Investigations Center (CIC1401 Bordeaux).

Edouard Lhomme is co-president of the joint action on respiratory viruses of ANRS-MIE since 2022.

9.2 Teaching - Supervision - Juries

9.2.1 Teaching

Each faculty member is involved in teaching with approximatively MA 220 h/year, RG 200 h/year, RT 130 h/year, and LW 110 h/year, LR 80 h/year, BH 70 h/year, EL 80 h/year, MP 30 h/year. These activities splits as follow.

- **In class teaching**
  - Rodolphe Thiébaut is head of the Digital Public Health graduate program, University of Bordeaux. Robin Genuer is head of the M2 Biostatistique, Master of Public Health, University of Bordeaux.
  - Master: All the permanent members and several PhD students teaches in the Master of Public Health (M1 Santé publique, M2 Biostatistique and/or M2 Epidemiology) and the Digital Public Health graduate program, University of Bordeaux.
  - Master: Marta Avalos teaches in the Master of Applied Mathematics and Statistics (1st and/or 2nd year), University of Bordeaux.
  - Bachelor: Laura Richert, Linda Wittkop and Edouard Lhomme teach in PASS and DFASM1-3 (Diplôme de Formation Approfondie en Sciences Médicales) for Medical degree at Univ. Bordeaux.
  - Master: Laura Richert and Edouard Lhomme teach in the Master of Vaccinology from basic immunology to social sciences of health (University Paris-Est Créteil, UPEC).
  - Teaching unit coordination: Laura Richert, Linda Wittkop, Rodolphe Thiébaut, Robin Genuer, Boris Hejblum and Marta Avalos coordinate several teaching units of Master in Public Health (Biostatistics, Epidemiology, Public Health).
  - Laura Richert coordinates the teaching unit "critical article reading" (across 4 years of medical school), University of Bordeaux; Edouard Lhomme coordinates the teaching “Evaluation of health innovation (M2 Health innovations); Linda Wittkop coordinates the teaching unit “Public Health and Statistics in Medicine” of the first year of Medical School, University of Bordeaux; Marta Avalos co-coordinates the teaching unit "Statistical analysis of high-dimensional data" (M2 Statistical and stochastic modeling MSS and M2 in Statistical and Computer Engineering CMI ISI of the UFMI of the University of Bordeaux).
  - Boris Hejblum teaches a 3-day graduate course “Bayesian analysis for biomedical research” and a 2-day graduate course ’EndeavR: a course on code efficiency and software development with R’ at the University of Copenhaguen.
  - Boris Hejblum teaches on Statistical learning in high-dimension in M2 Numerical sciences & bio-health, École Centrale Nantes
  - Mélanie Prague teaches "Missing Data" at ENSAI Master Level
  - ISPED Summer school: Introduction to R (Mélanie Prague), Advanced R (R Genuer & B Hejblum), Platform trial (Edouard Lhomme).
- **E-learning**
  - Marta Avalos is head of the e-learning program of the Master of Public Health, 1st year, University of Bordeaux.
  - Master: Marta Avalos teaches in the e-learning program of the Master of Public Health (1st and 2nd year).
  - ODL University Course: Robin Genuer is head of the Diplôme universitaire "Méthodes statistiques en santé". Mélanie Prague ans Robin Genuer teach in the Diplôme universitaire "Méthodes statistiques de régression en épidémiologie".
  - ODL University Course: Edouard Lhomme co-coordinates and teaches in the Diplôme universitaire "Recherche Clinique".
  - ODL University project: Robin Genuer participated to the IdEx Bordeaux University "Défi numérique" project BeginR.

9.2.2 Supervision

Supervision of intern students:

- Quentin Clairon: Victoria Idier (Master 1 Santé publique, ISPED, Université de Bordeaux) "Modélisation mécaniste de l'activité neutralisante post infection au Covid19 pour des sujets vaccinés". Apr-Jun 2023
- Mélanie Prague :
  - Marie COLIN (ENSAI 2 ieme année June – Aug. 2023) "Mechanistic modeling of antibodies response using transcriptomics data".
  - Junior JUMBONG (ENSAI 2 ieme année June – Aug. 2023) "Analysis of determinant of dynamical vaccination response using mechanistic models".
  - Fanny MOREAU (Licence 3 MIASCHG Bordeaux May – July 2023) "Analysis of determinant of vaccination response at 12 months using regression models".
  - Auriane GABAUT (Master 2 Modélisation science du vivant Polytechnique Paris Apr. – Sept. 2023) "Model building in nonlinear models in presence of a high number of covariates".
- Linda Wittkop : Lydia Wilson (M2)
- Boris Hejblum: Arthur Hughes (M2) and Maud Perpere (M1)
- Marta Avalos: Arthur Saracco (M2 ENSC - Bordeaux INP) co-supervision with Marie Chavent; Camila Labarca (eng Universidad de Chile) co-supervision with Laurence Delhaes; Dalia Cohen (M1 Isped) co-supervision with Emmanuel Lagarde and Cédric Gil-Jardiné.

Supervision of PhD students:

- Rodolphe Thiébaut:
  - PhD defense: Houreratou Barry "Response to Ebola vaccine and factors associated with the variablity of vaccine immune response in African countries" -directed by Rodolphe Thiébaut, Nov 2023
  - PhD in progress: Iris Ganser, Evaluation of event-based internet biosurveillance for multi-regional detection of seasonal influenza onset, co-directed by David Buckeridge (McGill University) and Rodolphe Thiébaut, from Oct 2020.
  - PhD in progress: Thomas Ferté "Contribution of health data warehouses for clinical research and epidemiological surveillance in the context of Covid-19", co-directed by Rodolphe Thiébaut and Vianney Jouhet, from Sept 2021.
  - PhD in progress: Antonin Colajanni, Université de Bordeaux, from Oct 2023, co-direction with Patricia Thebault.
  - PhD in progress: Annesh Pal, from Oct 2023, INRIA PEPR.
- Mélanie Prague: Auriane GABAUT from Oct 2023.
- Robin Genuer: Justine REMIAT (1st year PhD started in 10/2023) "Development of distance-based random forests methods for complex data analysis", co-supervision with Cécile Proust-Lima
- Laura Richert:
  - PhD Ipsen, CIFRE defense: Hélène Savel 'Statistical classification of treatment responses in mouse clinical trials for stratified medicine in oncology drug discovery' from Oct 2020 to Nov 2023.
  - PhD in progress: Cyrille Kone "Bandit algorithms for early phase clinical trials in vaccinology", codirected by Emilie Kaufmann (Inria Scool) and Laura Richert from November 2022.
- Boris Hejblum:

  - Benjamin Hivert (PhD co-direction 50%): "Hierarchical modeling for integrative analysis of high-dimensional, high-throughput, multi-modal cell and molecular data for immunology research", co-directed by Boris Hejblum and Rodolphe Thiébaut, from Sept 2020.
  - Arthur Hughes (PhD co-supervision)
  - Kalidou Ba (PhD co-supervision): “Reservoir computing for cellular composition prediction from longitudinal transcriptomics data in vaccine trials”, co-directed by Rodolphe Thiébaut and Xavier Hinaut, from Nov 2022.
  - Annesh Pal (PhD co-supervision)

Other supervision:

- Boris Hejblum: Sara Fallet (eng 100%)
- Robin Genuer: Corentin SEGALAS (Post-doc started in 02/2023), "Random forests for longitudinal data irregularly measured" co-supervision with Cécile Proust-Lima

9.2.3 Juries

Mélanie Prague:

- Jury hiring Inria: CRCN Inria NICE Cote d’Azur 2023
- National thesis award jury: SFdS Jury du prix Marie Jeanne Laurent Duhamel (since 2022).
- PhD defense: THOMAS BENETTEAU, thesis examiner, IRD Montpellier, « Modélisation mathématique des infections HPV: quel rôle du hasard dans la persistance et l’oncogénése ? » Dec. 2023.
- PhD Follow-up:
  - Ilona Suhanda (2023 - 2026 ; director Raphaëlle Metras - Sorbonne Université). Spatial Risk assessment of lyme borreliosis and thick-borne enceohalitis: a joint modeling approach.
  - Eve Rahbe (2023 - 2024; director Lulla Opatowski and Philippe Glaser - Institut Pasteur). Modélisation spatio-temporelle de l’antibiorésistance à l’échelle mondiale.
  - Erwan Gaymard (2023 - 2024; director Maxime Sermesant and Irene Balelli - Inria Nice, Exactcure CIFRE). Développement de méthodologies mathématiques en méta-modélisation PK à partir de sources hautement hétérogènes.
  - Maxime Beaulieu (2023 - 2026; director Jérémie Guedj; Inserm Iame Paris). Modélisation de l’efficacité des stratégies antivirales contre les variants du Sars-CoV-2 : de la population générale aux patients hospitalisés.
  - Benjamin Glemain (2022 - 2025; director Fabrice Carrat and Nathanaël Lapidus; Inserm iPLesp Paris). Corrélats de protection contre les différents variants du SARS-CoV-2.
  - Iris Granger (2020 - 2024; director David Buckeridge and Rodolphe Thiébaut - Mc Gills Canada U. Bordeaux). Modeling the epidemics of COVID-19.
  - Baptiste Elie (2021 - 2024 ; director Samuel Alizon and Nacho Bravo - U. Montpelier). Etude des infection génitales HPV.

Linda Wittkop

- Jury PhD defense: Laurent Lam, Sorbonne Université and Dathan Byonanebye, UNSW Sydney
- PhD Follow-up: Domecq Sandrine, Univ Bordeaux

Robin Genuer:

- Jury PhD defense: Ludovic Arnoult, Examinateur, Sorbonne Université, Paris ’When Random Forests Meet Neural Networks A Finite-Sample Analysis’ Oct. 2023.

- PhD Follow-up: Valentin Portmann, Université de Bordeaux, Sep. 2023

Laura Richert:

- Jury HDR defense: Vanessa Douet-Vannucci, De la biologie des systèmes intégrés à la médecine des systèmes : Une histoire d'interactions vers l'épidémiologie numérique ; Université de Nice 2023.
- Jury PhD defense:
  - François Camille Grolleau Raoux: Méthodes d'inférence causale pour la médecine personnalisée : une application au temps d'initiation de l'épuration extra-rénale, Université Paris Cité 2023
  - Drifa Belhadi: Bayesian decision analysis to design clinical trials in the context of viral hemorrhagic fevers, Université Paris Cité 2023
  - Li-Yun Lin: Characterization of the protective antibody response induced following vaccination or infection, Université de Strasbourg 2023
- PhD Follow-up:
  - Vincent Bouteloup: Apport des biomarqueurs sanguins dans le dépistage et le diagnostic précoce de la maladie d'Alzheimer ; Université de Bordeaux
  - Clement Massonnaud: Statistical challenges and pitfalls for the design and analysis of platform clinical trials in the context of emerging infectious diseases, Univ Paris Cité

Boris Hejblum:

- PhD Follow-up:
  - Manel Rakez (2021 - 2024 ; directors Virginie Rondeau and Brice Amadeo — Université de Bordeaux).
  - Lucie Fontaine (2023 - 2026; director Frédérique Alexandre — Université de Bordeaux). Modèles génératifs bio-inspirés pour l'IA
  - Catalina Gonzales-Gomez (2022-2025; director Manuel Rosa-calatrava — Université de Lyon 1). Développement d'outils in silico pour le repositionnement pharmaceutique de molécules basés sur des données d'expression omics et la notion de connectivité
  - Gustavo Magana-Lopez (2022-2025; director Loïc Paulevé — Université de Bordeaux). Synthèse et apprentissage d'ensembles de réseaux booléens prédictifs pour la reprogrammation cellulaire

Marta Avalos

- International PhD Follow-up: Jhon Erick Barrera Perez, Doctorado En Matematica, Jan 2023, 'New developments in the estimation of statistical models for complex longitudinal data', University of Valparaiso (Chile).
- PhD Follow-up: Dylan Russon, Université de Bordeaux, Sep. 2023
- Academic tutor of PhD student S. Zaïd, Ecole doctorale Sociétés, Politique, Santé Publique de l'Université de Bordeaux- EDSP2.

Rodolphe Thiebaut:

- Jury HDR defense: Vianney Jouhet (2023), Benjamin Bouyer (2023), Nolwenn L Meur (2023).
- Jury PhD defense: Paul Frelon (2023), Marc Labriffe (2023), Sophie Quenelle (2023)

9.3 Popularization

9.3.1 Articles and contents

Modéliser la COVID-19: de la population à l'individu. Vignals C., Hejblum B. and Prague M. Interstice Online, June 2023 [51].

9.3.2 Education

Developing Autonomy through an Entrepreneurial Project, presented in the ENLIGHT Teaching and Learning Conference 2023 - Empowering students for tomorrow, Oct 2023 [50]

9.3.3 Interventions

Interventions by Rodolphe Thiebaut

- « Les Grandes Tendances de la E-santé 2023 » (24/1/2023), Paris, France (4500 inscrits). Table ronde : Entrepôt de données de santé, évaluation des applications E-santé, souveraineté numérique... Les enjeux pour le patient ?
- IMI Impact on : Ebola, 2023 June 13th, Visioconference, The research side in EBOVAC projects www.ihi.europa.eu/news-events/events/imi-impact-ebola

10 Scientific production

10.1 Major publications

[1] D. Agniel and B. P. Hejblum. 'Variance component score test for time-course gene set analysis of longitudinal RNA-seq data'. In: *Biostatistics* 18.4 (2017), pp. 589–604. URL: https://hal.inria.fr/hal-01579077.

[2] H. Barry, G. Mutua, H. Kibuuka, Z. Anywaine, S. Sirima, N. Meda, O. Anzala, S. Eholie, C. Bétard, L. Richert, C. Lacabaratz, M. J. McElrath, S. De Rosa, K. Cohen, G. Shukarev, C. Robinson, A. Gaddah, D. Heerwegh, V. Bockstal, K. Luhn, M. Leyssen, M. Douoguih and R. Thiébaut. 'Safety and immunogenicity of 2-dose heterologous Ad26.ZEBOV, MVA-BN-Filo Ebola vaccination in healthy and HIV-infected adults: A randomised, placebo-controlled Phase II clinical trial in Africa'. In: *PLoS Medicine* 18.10 (29th Oct. 2021), e1003813. DOI: 10.1371/journal.pmed.1003813. URL: https://hal.inria.fr/hal-03481531.

[3] D. Commenges, C. Alkhassim, R. Gottardo, B. P. Hejblum and R. Thiébaut. 'cytometree: A binary tree algorithm for automatic gating in cytometry analysis'. In: *Cytometry Part A* 93.11 (2018), pp. 1132–1140. URL: https://hal.inria.fr/hal-01887966.

[4] D. Commenges and H. Jacqmin-Gadda. *Dynamical Biostatistical Models*. Chapman and Hall/CRC, 2015. URL: https://hal.inria.fr/hal-01580149.

[5] A. Jarne, D. Commenges, M. Prague, Y. Levy and R. Thiébaut. 'Modeling CD4 + T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous Interleukin 7'. In: *Annals of Applied Statistics* (2017). URL: https://hal.inria.fr/hal-01579008.

[6] P. Loubet, L. Wittkop, E. Tartour, B. Parfait, B. Barrou, J.-Y. Blay, M. Hourmant, M. Lachâtre, D.-A. Laplaud, M. Laville, B. Laviolle, J.-D. Lelievre, J. Morel, S. Nguyen, J.-P. Spano, B. Terrier, A. Thiebaut, J.-F. Viallard, F. Vrtovsnik, X. De Lamballerie and O. Launay. 'A French cohort for assessing COVID-19 vaccine responses in specific populations'. In: *Nature Medicine* 27.8 (12th July 2021), pp. 1319–1321. DOI: 10.1038/s41591-021-01435-1. URL: https://hal.sorbonne-universite.fr/hal-03290330.

[7] C. Pasin, F. Dufour, L. Villain, H. Zhang and R. Thiébaut. 'Controlling IL-7 injections in HIV-infected patients'. In: *Bulletin of Mathematical Biology* (2018).

[8] A. Pollard, O. Launay, J.-D. Lelievre, C. Lacabaratz, S. Grande, N. Goldstein, C. Robinson, A. Gaddah, V. Bockstal, M. Leyssen et al. 'Safety and immunogenicity of a two-dose heterologous Ad26.ZEBOV and MVA-BN-Filo Ebola vaccine regimen in adults in Europe (EBOVAC2): a randomised, observer-blind, participant-blind, placebo-controlled, phase 2 trial'. In: *The Lancet Infectious Diseases* (Nov. 2020). DOI: 10.1016/S1473-3099(20)30476-X. URL: https://hal.inria.fr/hal-03142752.

[9] M. Prague, D. Commenges, J. M. Gran, B. Ledergerber, J. Young, H. Furrer and 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* (2017). URL: https://hal.inria.fr/hal-01406614.

[10] 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 and 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* 20.9 (2017), pp. 2251–2261. URL: https://hal.inria.fr/hal-01579246.

[11] L. Villain, D. Commenges, C. Pasin, M. Prague and R. Thiébaut. 'Adaptive protocols based on predictions from a mechanistic model of the effect of IL7 on CD4 counts'. In: *Statistics in Medicine* 38.2 (2018), pp. 221–235.

10.2 Publications of the year

International journals

[12] D. Agniel, B. Hejblum, R. Thiébaut and L. Parast. 'Doubly-robust evaluation of high-dimensional surrogate markers'. In: *Biostatistics* 24.4 (2023), pp. 985–999. DOI: 10.1093/biostatistics/kxac020. URL: https://inria.hal.science/hal-03100499.

[13] M. Alexandre, M. Prague, C. Mclean, V. Bockstal, M. Douoguih and R. Thiébaut. 'Prediction of long-term humoral response induced by the two-dose heterologous Ad26.ZEBOV, MVA-BN-Filo vaccine against Ebola'. In: *NPJ vaccines* 8.174 (8th Nov. 2023). DOI: 10.1038/s41541-023-00767-y. URL: https://hal.science/hal-04226696.

[14] D. Barger, M. Hessamfar, D. Neau, S. Farbos, O. Leleux, C. Cazanave, N. Rouanes, P. Duffau, E. Lazaro, P. Rispal, F. Dabis, L. Wittkop and F. Bonnet. 'Factors associated with poorer quality of life in people living with HIV in southwestern France in 2018-2020 (ANRS CO3 AQUIVIH-NA cohort: QuAliV study)'. In: *Scientific Reports* 13.1 (2nd Oct. 2023), p. 16535. DOI: 10.1038/s41598-023-43434-x. URL: https://hal.science/hal-04342328.

[15] J. Ben Farhat, M. Hessamfar, S. Farbos, A. Desclaux, G. Dumondin, H. Ferrand, C. Greib, B. Castan, P. Rispal, P. Duffau, O. Leleux, A. Perrier, L. Wittkop, F. Bonnet and D. Barger. 'Time to treatment initiation and HIV viral suppression in people diagnosed with HIV-1 during COVID-19 pandemic in ex-Aquitaine, France (ANRS CO3 AQUIVIH-NA Cohort-QuAliCOV Study)'. In: *Journal of Acquired Immune Deficiency Syndromes - JAIDS* (28th Sept. 2023). DOI: 10.1097/QAI.0000000000003310. URL: https://hal.science/hal-04288500.

[16] F. Blengio, H. Hocini, L. Richert, C. Lefebvre, M. Durand, B. P. Hejblum, P. Tisserand, C. Mclean, K. Luhn, R. Thiébaut and Y. Levy. 'Identification of early gene expression profiles associated with long-lasting antibody responses to the Ebola vaccine Ad26.ZEBOV/MVA-BN-Filo'. In: *Cell Reports* 42.9 (26th Sept. 2023), p. 113101. DOI: 10.1016/j.celrep.2023.113101. URL: https://hal.science/hal-04273901.

[17] M. Chalouni, A. Trickey, S. M. Ingle, M. A. Sepuvelda, J. Gonzalez, A. Rauch, H. M. Crane, M. J. Gill, P. F. Rebeiro, J. K. Rockstroh, R. A. Franco, G. Touloumi, D. Neau, M. Laguno, M. Rappold, C. Smit, J. A. Sterne and L. Wittkop. 'Impact of Hepatitis C cure on risk of mortality and morbidity in people with HIV after ART initiation'. In: *AIDS. Official journal of the international AIDS Society* (10th May 2023). DOI: 10.1097/qad.0000000000003594. URL: https://hal.science/hal-04139288.

[18] G. Chenais, C. Gil-Jardiné, H. Touchais, M. Avalos Fernandez, B. Contrand, E. Tellier, X. Combes, L. Bourdois, P. Revel and E. Lagarde. 'Deep Learning Transformer Models for Building a Comprehensive and Real-time Trauma Observatory: Development and Validation Study'. In: *JMIR AI* 2 (2023), e40843. DOI: 10.2196/40843. URL: https://inria.hal.science/hal-03978442.

[19] E. M.-L. Choi, B. Lacarra, M. O. Afolabi, B. M. Ale, F. Baiden, C. Betard, J. Foster, B. Hamze, C. Schwimmer, D. Manno, E. d'Ortenzio, D. Ishola, C. M. Keita, B. Keshinro, Y. Njie, W. van Dijck, A. Gaddah, D. Anumendem, B. Lowe, R. Vatrinet, B. J. Lawal, G. T. Otieno, M. Samai, G. F. Deen, I. B. Swaray, A. B. Kamara, M. M. Kamara, M. A. Diagne, D. Kowuor, C. Mclean, B. Leigh, A. H. Beavogui, M. Leyssen, K. Luhn, C. Robinson, M. Douoguih, B. Greenwood, R. Thiebaut and D. Watson-Jones. 'Safety and immunogenicity of the two-dose heterologous Ad26.ZEBOV and MVA-BN-Filo Ebola vaccine regimen in infants: a phase 2, randomised, double-blind, active-controlled trial in Guinea and Sierra Leone'. In: *The Lancet Global Health* 11.11 (Nov. 2023), e1743–e1752. DOI: 10.1016/S2214-109X(23)00410-2. URL: https://hal.science/hal-04384862.

[20] O. Ciani, A. M. Manyara, P. Davies, D. Stewart, C. J. Weir, A. E. Young, J. Blazeby, N. J. Butcher, S. Bujkiewicz, A.-W. Chan, D. Dawoud, M. Offringa, M. Ouwens, A. Hrobjartssson, A. Amstutz, L. Bertolaccini, V. D. Bruno, D. Devane, C. D. C. M. Faria, P. B. Gilbert, R. Harris, M. Lassere, L. Marinelli, S. Markham, J. H. Powers, Y. Rezaei, L. Richert, F. Schwendicke, L. G. Tereshchenko, A. Thoma, A. Turan, A. Worrall, R. Christensen, G. S. Collins, J. S. Ross and R. S. Taylor. 'A framework for the definition and interpretation of the use of surrogate endpoints in interventional trials'. In: *EClinicalMedicine* 65 (Nov. 2023), p. 102283. DOI: 10.1016/j.eclinm.2023.102283. URL: https://hal.science/hal-04385165.

[21] Q. Clairon, C. Pasin, I. Balelli, R. Thiébaut and M. Prague. 'Parameter estimation in nonlinear mixed effect models based on ordinary differential equations: An optimal control approach'. In: *Computational Statistics* (18th Sept. 2023). URL: https://hal.science/hal-03335826.

[22] Q. Clairon, M. Prague, D. Planas, T. Bruel, L. Hocqueloux, T. Prazuck, O. Schwartz, R. Thiébaut and J. Guedj. 'Modeling the evolution of the neutralizing antibody response against SARS-CoV-2 variants after several administrations of Bnt162b2'. In: *PLoS Computational Biology* (31st July 2023). URL: https://inria.hal.science/hal-03946556.

[23] A. Collin, B. P. Hejblum, C. Vignals, L. Lehot, R. Thiébaut, P. Moireau and M. Prague. 'Using Population Based Kalman Estimator to Model COVID-19 Epidemic in France: Estimating the Effects of Non-Pharmaceutical Interventions on the Dynamics of Epidemic'. In: *The international journal of biostatistics* (2023). DOI: 10.1515/ijb-2022-0087. URL: https://hal.science/hal-04151651.

[24] L. Dong, E. E. M. Moodie, L. Villain and R. Thiébaut. 'Evaluating the Use of Generalized Dynamic Weighted Ordinary Least Squares for Individualized Hiv Treatment Strategies'. In: *Annals of Applied Statistics* 17.3 (Sept. 2023), pp. 2432–2451. DOI: 10.1214/22-aoas1726. URL: https://hal.science/hal-04344321.

[25] P. Freulon, J. Bigot and B. Hejblum. 'CytOpT: Optimal Transport with Domain Adaptation for Interpreting Flow Cytometry data'. In: *Annals of Applied Statistics* 17.2 (2023), pp. 1086–1104. DOI: 10.1214/22-AOAS1660. URL: https://inria.hal.science/hal-03100405.

[26] H. Hocini, A. Wiedemann, F. Blengio, C. Lefebvre, M. Cervantes-Gonzalez, E. Foucat, P. Tisserand, M. Surenaud, S. Coléon, M. Prague et al. 'Neutrophil Activation and Immune Thrombosis Profiles Persist in Convalescent COVID-19'. In: *Journal of Clinical Immunology* 43.5 (21st Mar. 2023), pp. 882–893. DOI: 10.1007/s10875-023-01459-x. URL: https://hal.science/hal-04373844.

[27] P. Loubet, L. Wittkop, L. Ninove, M. Chalouni, B. Barrou, J.-Y. Blay, M. Hourmant, E. Thouvenot, M. Laville, B. Laviolle, J.-D. Lelievre, J. Morel, S. N. Quoc, J.-P. Spano, B. Terrier, A. Thiebaut, J.-F. Viallard, F. Vrtovsnik, S. Circosta, L. Esterle, A. Levier, P. Vanhems, E. Tartour, B. Parfait, X. de Lamballerie and O. Launay. 'One-month humoral response following two or three doses of mRNA Covid-19 vaccines as primary vaccination in specific populations in France: first results from the ANRS0001S COV-POPART cohort'. In: *Clinical Microbiology and Infection* 29.3 (2023), 388.e1–388.e8. DOI: 10.1016/j.cmi.2022.10.009. URL: https://hal.science/hal-03822592.

[28] A. Marc, R. Marlin, F. Donati, M. Prague, M. Kerioui, C. Hérate, M. Alexandre, N. Dereuddre-Bosquet, J. Bertrand, V. Contreras, S. Behillil, P. Maisonnasse, S. van der Werf, R. Le Grand and J. Guedj. 'Impact of variants of concern on SARS-CoV-2 viral dynamics in non-human primates'. In: *PLoS Computational Biology* 19.8 (Aug. 2023), e1010721. DOI: 10.1371/journal.pcbi.1010721. URL: https://hal.science/hal-04225046.

[29] J. Paireau, M.-L. Charpignon, S. Larrieu, C. Calba, N. Hozé, P.-Y. Boëlle, R. Thiébaut, M. Prague and S. Cauchemez. 'Impact of non-pharmaceutical interventions, weather, vaccination, and variants on COVID-19 transmission across departments in France'. In: *BMC Infectious Diseases* 23.190 (30th Mar. 2023). DOI: 10.1186/s12879-023-08106-1. URL: https://hal.science/hal-03791761.

[30] S. M. Rein, F. C. Lampe, S. M. Ingle, J. a. C. Sterne, A. Trickey, M. J. Gill, V. Papastamopoulos, L. Wittkop, M. van der Valk, M. Kitchen, J. L. Guest, D. D. Satre, G. Wandeler, P. Galindo, J. Castilho, H. M. Crane and C. J. Smith. 'All-cause hospitalisation among people living with HIV according to gender, mode of HIV acquisition, ethnicity, and geographical origin in Europe and North America: findings from the ART-CC cohort collaboration'. In: *Lancet Public Health* 8.10 (Oct. 2023), pp. 776–787. DOI: 10.1016/s2468-2667(23)00178-0. URL: https://hal.science/hal-04327414.

[31] S. M. Rein, S. Lodi, R. W. Logan, G. Touloumi, A. C. Antoniadou, L. Wittkop, F. Bonnet, A. I. van Sighem, M. van der Valk, P. Reiss, M. B. Klein, J. Young, I. Jarrin, A. D. d'Arminio Monforte, A. Tavelli, L. Meyer, L. Tran, M. J. Gill, R. Lang, B. Surial, A. D. Haas, A. C. Justice, C. T. Rentsch, A. N. Phillips, C. A. Sabin, J. M. Miro, A. Trickey, S. M. Ingle, J. A. Sterne and M. A. Á. Hernán. 'Integrase strand-transfer inhibitor use and cardiovascular events in adults with HIV: an emulation of target trials in the HIV-CAUSAL Collaboration and the Antiretroviral Therapy Cohort Collaboration'. In: *The Lancet HIV* 10.11 (2023), e723–e732. DOI: 10.1016/S2352-3018(23)00233-3. URL: https://hal.science/hal-04294680.

[32] E. Ruiz-Burga, D. S. Tariq, G. Touloumi, J. Gill, E. J. Nicholls, C. A. Sabin, C. Mussini, L. Meyer, A. Volny Anne, C. Carlander, S. G. Grabar, I. Jarrin, M. van der Valk, L. Wittkop, B. Spire, N. Pantazis, F. M. Burns and K. Porter. 'CASCADE protocol: exploring current viral and host characteristics, measuring clinical and patient-reported outcomes, and understanding the lived experiences and needs of individuals with recently acquired HIV infection through a multicentre mixed-methods observational study in Europe and Canada'. In: *BMJ Open* 13.5 (2023). DOI: 10.1136/bmjopen-2022-070837. URL: https://hal.science/hal-04193750.

[33] H. Savel, S. Barbier, C. Proust-Lima, V. Rondeau, R. Thiébaut, F. Meyer-Losic and L. Richert. 'On the Choice of Longitudinal Models for the Analysis of Antitumor Efficacy in Mouse Clinical Trials of Patient-derived Xenograft Models'. In: *Cancer Research Communications* 3.1 (26th Jan. 2023), pp. 140–147. DOI: 10.1158/2767-9764.CRC-22-0238. URL: https://inria.hal.science/hal-04403048.

[34] H. Savel, F. Meyer-Losic, C. Proust-Lima and L. Richert. 'Statistical classification of treatment responses in mouse clinical trials for stratified medicine in oncology drug discovery'. In: *Scientific Reports* 14.1 (9th Jan. 2024), p. 934. DOI: 10.1038/s41598-023-51055-7. URL: https://inria.hal.science/hal-04403059.

[35] R. Thiebaut, B. P. Hejblum, F. Mougin, C. Tzourio and L. Richert. 'ChatGPT and beyond with artificial intelligence (AI) in health: Lessons to be learned.' In: *Joint Bone Spine* 90.5 (1st Sept. 2023), p. 105607. DOI: 10.1016/j.jbspin.2023.105607. URL: https://hal.science/hal-04236822.

[36] M. Wang, P. Staples, M. Prague, R. Goyal, V. Degruttola and J.-P. Onnela. 'Leveraging Contact Network Information in Clustered Randomized Studies of Contagion Processes'. In: *Observational Studies* 9.2 (2023), pp. 157–175. DOI: 10.1353/obs.2023.0021. URL: https://hal.science/hal-04225014.

[37] L. Wittkop and T. Boettler. 'More than shots in the dark: driving vaccine efficacy in cirrhosis.' In: *Gut* (7th June 2023). DOI: 10.1136/gutjnl-2023-329867. URL: https://hal.science/hal-04228574.

[38] M. Zysman, J. Asselineau, O. Saut, E. Frison, M. Oranger, A. Maurac, J. Charriot, R. Achkir, S. Regueme, E. Klein, S. Bommart, A. Bourdin, G. Dournes, J. Casteigt, A. Blum, G. Ferretti, B. Degano, R. Thiébaut, F. Chabot, P. Berger, F. Laurent and I. Benlala. 'Development and external validation of a prediction model for the transition from mild to moderate or severe form of COVID-19'. In: *European Radiology* (2023). DOI: 10.1007/s00330-023-09759-x. URL: https://hal.science/hal-04153209.

International peer-reviewed conferences

[39] M. Avalos and G. Coureau. 'Interdisciplinary and differentiated learning of Statistics in an online MPH'. In: *Proceedings of the IASE 2023 Satellite Conference: "Fostering Learning of Statistics and Data Science"*. IASE 2023 - The 13th IASE Satellite Conference. Toronto, Canada, 2023. URL: https://inria.hal.science/hal-04173057.

[40] C. Blot, K. Weickert, A. Payros, M. Avalos and C. Gil-Jardiné. 'Serious Games for Training in Patient Flow Management in Emergency Departments: State of the Art and Perspectives'. In: *Florida Online Journals*. FLAIRS 36 - 36th International Florida Artificial Intelligence Research Society Conference. Vol. 36. Clearwater Beach, FL., United States, 8th May 2023, p. 3. DOI: 10.32473/flairs.36.133276. URL: https://inria.hal.science/hal-04098447.

[41] C. Kone, E. Kaufmann and L. Richert. 'Adaptive Algorithms for Relaxed Pareto Set Identification'. In: NeurIPS 2023 - 37th Conference on Neural Information Processing Systems. La Nouvelle Orléans, LA, United States, 10th Dec. 2023. URL: https://hal.science/hal-04306210.

Conferences without proceedings

[42] M. Prague. 'Joint modeling of viral and humoral response in Non-human primates to define mechanistic correlates of protection for SARS-CoV-2'. In: Society for Mathematical Biology Annual Meeting - 2023. Colombus, United States, 16th July 2023. URL: https://hal.science/hal-04416297.

[43] M. Prague. 'Pharmacometric modelling to inform vaccine development'. In: 19th World Congress of Basic & Clinical Pharmacology 2023. Glasgow, United Kingdom, 2nd July 2023. URL: https://hal.science/hal-04416259.

Doctoral dissertations and habilitation theses

[44] H. Dufaure-Savel. 'Modélisation statistique de la réponse au traitement de modèles souris de tumeurs dérivées de patient. Application à la recherche translationnelle en oncologie'. PhD thesis. Université de Bordeaux, Nov. 2023. URL: http://www.theses.fr/2023BORD0296.

Reports & preprints

[45] A. Devaux, C. Helmer, R. Genuer and C. Proust-Lima. *Random survival forests with multivariate longitudinal endogenous covariates*. Jan. 2023. URL: https://hal.science/hal-03747106.

[46] A. Marc, R. Marlin, F. Donati, M. Prague, M. Kerioui, C. Hérate, M. Alexandre, N. Dereuddre-Bosquet, J. Bertrand, V. Contreras, S. Behillil, P. Maisonnasse, S. van der Werf, R. L. Grand and J. Guedj. *Impact of variants of concern on SARS-CoV-2 viral dynamics in non-human primates*. 9th Nov. 2022. DOI: 10.1101/2022.11.09.515748. URL: https://pasteur.hal.science/pasteur-04072755.

Other scientific publications

[47] A.-P. Ake, M. Diomande, D. Valoura-Esteves, R. Enaud, L. Delhaes and M. Avalos. 'The respiratory microbiome alpha-diversity and chronic respiratory diseases in children: A systematic review'. In: WoM 2023 - 4th International World of Microbiome Conference. Sofia, Bulgaria, 26th Oct. 2023. URL: https://inria.hal.science/hal-04260986.

[48] C. Salaun-Ferron, T. Oufkir, N. P. Ouedraogo and M. Avalos. 'Association between indoor environmental microbiota of living spaces and chronic asthma and respiratory allergies in Europe: A systematic review'. In: WoM 2023 - 4th International World of Microbiome Conference. Sofia, Bulgaria, 26th Oct. 2023. URL: https://inria.hal.science/hal-04260931.

[49] A. Saracco, M. Chavent and M. Avalos. 'Utility of multivariate data analysis and penalized meta-regression to explore sources of heterogeneity in microbiome meta-analyses'. In: WoM 2023 - 4th International World of Microbiome Conference. Sofia, Bulgaria, 26th Oct. 2023. URL: https://inria.hal.science/hal-04260888.

[50] L. Sebastien and M. Avalos. 'Developing Autonomy through an Entrepreneurial Project'. In: ENLIGHT Teaching and Learning Conference 2023 -Empowering students for tomorrow. Bordeaux, France, 11th Oct. 2023. URL: https://inria.hal.science/hal-04243910.

10.3 Other

Scientific popularization

[51] C. Vignals, B. P. Hejblum and M. Prague. 'Modéliser la COVID-19 : de la population à l'individu'. In: *Interstices* (15th June 2023). URL: https://inria.hal.science/hal-04230705.

10.4 Cited publications

[52] O. O. Aalen, K. Røysland, J. M. Gran and B. Ledergerber. 'Causality, mediation and time: a dynamic viewpoint'. In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 175.4 (2012), pp. 831–861. DOI: https://doi.org/10.1111/j.1467-985X.2011.01030.x.

[53] C. Bouveyron. 'Adaptive Mixture Discriminant Analysis for Supervised Learning with Unobserved Classes'. In: *Journal of Classification* 31.1 (2014), pp. 49–84. DOI: 10.1007/s00357-014-9147-x.

[54] L. Capitaine, J. Bigot, R. Thiébaut and R. Genuer. 'Fréchet random forests for metric space valued regression with non Euclidean predictors'. working paper or preprint. Dec. 2020. URL: https://hal.archives-ouvertes.fr/hal-03066146.

[55] T. Ferté, S. Cossin, T. Schaeverbeke, T. Barnetche, V. Jouhet and B. Hejblum. 'Automatic phenotyping of electronical health record: PheVis algorithm'. In: *Journal of Biomedical Informatics* 117 (2021), p. 103746. DOI: 10.1016/j.jbi.2021.103746. URL: https://hal.inria.fr/hal-03100435.

[56] T. Ferté, V. Jouhet, R. Griffier, B. Hejblum and R. Thiébaut. 'The benefit of augmenting open data with clinical data-warehouse EHR for forecasting SARS-CoV-2 hospitalizations in Bordeaux area, France'. In: *JAMIA open* 5.4 (Dec. 2022). DOI: 10.1093/jamiaopen/ooac086. URL: https://hal.science/hal-03908260.

[57] B. P. Hejblum, G. M. Weber, K. P. Liao, N. P. Palmer, S. Churchill, N. Shadick, P. Szolovits, S. N. Murphy, I. Kohane and T. Cai. 'Probabilistic record linkage of de-identified research datasets with discrepancies using diagnosis codes'. In: *Scientific Data* 6 (2019), p. 180298. DOI: 10.1038/sdata.2018.298. URL: https://hal.inria.fr/hal-01942279.

[58] L. Hood and Q. Tian. 'Systems approaches to biology and disease enable translational systems medicine'. In: *Genomics Proteomics Bioinformatics* 10.4 (2012), pp. 181–5.

[59] A. Jarne, D. Commenges, L. Villain, M. Prague, Y. Lévy and R. Thiébaut. 'Modeling CD4+ T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous Interleukin 7'. In: *The Annals of Applied Statistics* 11.3 (2017), pp. 1593–1616.

[60] H. Jin and Z. Liu. 'A Benchmark for RNA-seq Deconvolution Analysis under Dynamic Testing Environments'. In: *Genome Biology* 22.1 (2021), p. 102. DOI: 10.1186/s13059-021-02290-6.

[61] E. Kuhn and M. Lavielle. 'Coupling a stochastic approximation version of EM with an MCMC procedure'. In: *ESAIM: Probability and Statistics* 8 (2004), pp. 115–131.

[62] M. Lavielle. *Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools*. Chapman and Hall/CRC, 2014. URL: https://hal.science/hal-01122873.

[63] K. Nganou-Makamdop, A. Talla, A. A. Sharma, S. Darko, A. Ransier, F. Laboune, J. G. Chipman, G. J. Beilman, T. Hoskuldsson, S. Fourati, T. E. Schmidt, S. Arumugam, N. S. Lima, D. Moon, S. Callisto, J. Schoephoerster, J. Tomalka, P. Mugyenyi, F. Ssali, P. Muloma, P. Ssengendo, A. R. Leda, R. K. Cheu, J. K. Flynn, A. Morou, E. Brunet-Ratnasingham, B. Rodriguez, M. M. Lederman, D. E. Kaufmann, N. R. Klatt, C. Kityo, J. M. Brenchley, T. W. Schacker, R. P. Sekaly and D. C. Douek. 'Translocated Microbiome Composition Determines Immunological Outcome in Treated HIV Infection'. In: *Cell* 184.15 (2021), 3899–3914.e16. DOI: 10.1016/j.cell.2021.05.023.

[64] A. S. Perelson and R. M. Ribeiro. 'Introduction to modeling viral infections and immunity'. In: *Immunological Reviews* 285.1 (2018), pp. 5–8. DOI: https://doi.org/10.1111/imr.12700.

[65] G. A. Poland, I. G. Ovsyannikova and R. B. Kennedy. 'Personalized vaccinology: a review'. In: *Vaccine* 36.36 (2018), pp. 5350–5357.

[66] M. Prague, D. Commenges, J. Guedj, J. Drylewicz and R. Thiébaut. 'NIMROD: A program for inference via a normal approximation of the posterior in models with random effects based on ordinary differential equations'. In: *Computer Methods and Programs in Biomedicine* 111.2 (2013), pp. 447–458.

[67] M. Prague, D. Commenges, J. M. Gran, B. Ledergerber, J. Young, H. Furrer and 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* 73.1 (Mar. 2017), pp. 294–304. DOI: 10.1111/biom.12564. URL: https://hal.inria.fr/hal-01406614.

[68] B. Pulendran. 'Learning immunology from the yellow fever vaccine: innate immunity to systems vaccinology'. In: *Nature Reviews Immunology* 9.10 (2009), pp. 741–7.

[69] L. SAS. 'Monolix version 2021R1'. In: *Antony, France* (2022), http://lixoft.com/products/monolix/.

[70] C. Schubert. 'Systems immunology: complexity captured'. In: *Nature* 473.7345 (2011), pp. 113–4.

[71] J. A. Sinnott, F. Cai, S. Yu, B. P. Hejblum, C. Hong, I. Kohane and K. P. Liao. 'PheProb: probabilistic phenotyping using diagnosis codes to improve power for genetic association studies'. In: *Journal of the American Medical Informatics Association* (May 2018). DOI: 10.1093/jamia/ocy056. URL: https://hal.inria.fr/hal-01887930.

[72] P. Soret. 'Régression pénalisée de type Lasso pour l'analyse de données biologiques de grande dimension : application à la charge virale du VIH censurée par une limite de quantification et aux données compositionnelles du microbiote'. Theses. Université de bordeaux, Nov. 2019. URL: https://hal.inria.fr/tel-02425113.

[73] D. Stein, D. O'Connor, C. Blohmke, M. Sadarangani and A. Pollard. 'Gene Expression Profiles Are Different in Venous and Capillary Blood: Implications for Vaccine Studies'. In: *Vaccine* 34.44 (2016), pp. 5306–5313. DOI: 10.1016/j.vaccine.2016.09.007.

[74] R. Thiébaut, H. Jacqmin-Gadda, A. Babiker and 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* 24.1 (2005), pp. 65–82.

[75] N. Trouvain, L. Pedrelli, T. T. Dinh and X. Hinaut. 'ReservoirPy: An Efficient and User-Friendly Library to Design Echo State Networks'. In: *Artificial Neural Networks and Machine Learning – ICANN 2020*. Ed. by I. Farkaš, P. Masulli and S. Wermter. Vol. 12397. Cham: Springer International Publishing, 2020, pp. 494–505. DOI: 10.1007/978-3-030-61616-8_40.

[76] L. Wang, J. Cao, J. O. Ramsay, D. Burger, C. Laporte and J. K. Rockstroh. 'Estimating mixed-effects differential equation models'. In: *Statistics and Computing* 24.1 (2014), pp. 111–121.

[77] H. Wu. 'Statistical methods for HIV dynamic studies in AIDS clinical trials'. In: *Statistical Methods in Medical Research* 14.2 (2005), pp. 171–192.

[78] H. Zhang, B. P. Hejblum, G. Weber, N. Palmer, S. Churchill, P. Szolovits, S. Murphy, K. Liao, I. Kohane and T. Cai. 'ATLAS: an automated association test using probabilistically linked health records with application to genetic studies'. In: *Journal of the American Medical Informatics Association* 28.12 (Dec. 2021), pp. 2582–2592. DOI: 10.1093/jamia/ocab187. URL: https://hal.inria.fr/hal-03478039.