

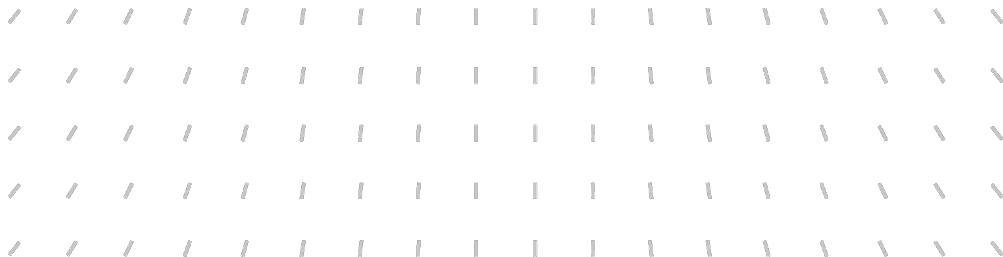
2025 Activity Report

RESEARCH CENTRE: Inria Centre at the University of Bordeaux
IN PARTNERSHIP WITH: Université de Bordeaux, INSERM

Project-Team

SISTM

Statistics In System biology and Translational
Medicine



Project-Team SISTM

Creation of the Project-Team: 2015 January 01

Each year, Inria research teams publish an Activity Report presenting their work and results over the reporting period. These reports follow a common structure, with some optional sections depending on the specific team. They typically begin by outlining the overall objectives and research programme, including the main research themes, goals, and methodological approaches. They also describe the application domains targeted by the team, highlighting the scientific or societal contexts in which their work is situated. The reports then present the highlights of the year, covering major scientific achievements, software developments, or teaching contributions. When relevant, they include sections on software, platforms, and open data, detailing the tools developed and how they are shared. A substantial part is dedicated to new results, where scientific contributions are described in detail, often with subsections specifying participants and associated keywords. Finally, the Activity Report addresses funding, contracts, partnerships, and collaborations at various levels, from industrial agreements to international cooperations. It also covers dissemination and teaching activities, such as participation in scientific events, outreach, and supervision. The document concludes with a presentation of scientific production, including major publications and those produced during the year.

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
- 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
- A9.2.1. – Supervised learning
- A9.2.2. – Unsupervised learning
- A9.2.3. – Reinforcement learning
- A9.2.4. – Optimization and learning
- A9.2.5. – Bayesian methods
- A9.2.6. – Neural networks
- 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. – Pharmacokinetics and dynamics
- B2.4.2. – Drug resistance
- B9.5.6. – Data science
- B9.8. – Reproducibility

Contents

Project-Team SISTM	1
1 Team members, visitors, external collaborators	5
2 Overall objectives	7
3 Research program	9
3.1 Axis High-Dimensional Statistical Learning	9
3.2 Axis Mechanistic learning	10
3.3 Axis Translational vaccinology and design	11
4 Application domains	12
5 Social and environmental responsibility	13
5.1 Footprint of research activities	13
5.2 Impact of research results	13
6 Highlights of the year	14
6.1 Major scientific Publication - RISE: Identification des marqueurs de substitution en haute dimension, appliquée à la vaccinologie	14
6.2 Major scientific Publication - Comparative evaluation of methodologies for estimating the effectiveness of non-pharmaceutical interventions in the context of COVID-19	14
6.3 Organisation of highly recognized national and international conferences	15
6.4 Awards-fundings	15
6.4.1 Renewal vaccine Research institute	15
6.4.2 Launch of ANRS PEPR MIE Previx (WP leader within-host modelling 2025-2028)	15
6.4.3 Launch of ANRS PEPR MIE Previx (WP leader within-host modelling 2025-2028)	15
7 Latest software developments, platforms, open data	16
7.1 Latest software developments	16
7.1.1 SurrogateRank	16
7.1.2 RastaRocket	16
7.1.3 reservoirnet	16
7.1.4 REMixed	17
7.1.5 LSAMBA	17
7.2 New platforms	17
8 New results	18
8.1 High-dimensional statistical learning	18
8.1.1 Identification of surrogates of protection	18
8.1.2 Reservoir computing for epidemic prediction	18
8.1.3 Machine learning/AI and bias	18
8.1.4 Compositional microbiome data	19
8.2 Mechanistic learning	19
8.2.1 Between host modeling of COVID-19 Epidemics	19
8.2.2 Modeling viruses and immune dynamics	19
8.2.3 PhD Defense	20
8.3 Translational vaccinology and design	20
8.3.1 Bandit Approach for trial design and PhD Defense	20
8.3.2 Response to COVID-19 vaccine in immunosuppressed populations	20
8.3.3 High-resolution transcriptomics	20
8.3.4 Knowledge transfer	21

9 Partnerships and cooperations	21
9.1 International initiatives	21
9.1.1 Associate Teams in the framework of an Inria International Lab or in the framework of an Inria International Program	21
9.1.2 STIC/MATH/CLIMAT AmSud projects	22
9.1.3 Participation in other International Programs	22
9.2 International research visitors	22
9.2.1 Visits of international scientists	22
9.2.2 Visits to international teams	23
9.3 European initiatives	24
9.3.1 H2020 projects	24
9.3.2 Other european programs/initiatives	25
9.4 National initiatives	25
9.4.1 Various Partnership	28
9.5 Regional initiatives	29
10 Dissemination	29
10.1 Promoting scientific activities	29
10.1.1 Scientific events: organisation	29
10.1.2 Scientific events: selection	30
10.1.3 Journal	30
10.1.4 Invited talks	31
10.1.5 Leadership within the scientific community	31
10.1.6 Scientific expertise	31
10.1.7 Research administration	32
10.2 Teaching - Supervision - Juries - Educational and pedagogical outreach	32
10.2.1 Teaching	32
10.2.2 Supervision	33
10.2.3 Juries	34
10.2.4 Educational and pedagogical outreach	34
10.3 Popularization	35
10.3.1 Specific official responsibilities in science outreach structures	35
10.3.2 Productions (articles, videos, podcasts, serious games, ...)	35
10.3.3 Participation in live events	35
10.3.4 Other relevant science outreach activities	35
11 Scientific production	35
11.1 Major publications	35
11.2 Publications of the year	37
11.3 Cited publications	42

1 Team members, visitors, external collaborators

Research Scientists

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

Faculty Members

- Marta Avalos Fernandez [UNIV BORDEAUX, Associate Professor, HDR]
- Robin Genuer [UNIV BORDEAUX, Associate Professor, until Jun 2025, HDR]
- Edouard Lhomme [UNIV BORDEAUX, Associate Professor]
- Laura Richert [UNIV BORDEAUX, Professor, HDR]
- Rodolphe Thiebaut [UNIV BORDEAUX, Professor, HDR]
- Linda Wittkop [UNIV BORDEAUX, Professor, HDR]

Post-Doctoral Fellows

- Iris Ganser [UNIV BORDEAUX, Post-Doctoral Fellow, until Mar 2025]
- Bastien Reyné [INSERM, Post-Doctoral Fellow, until Aug 2025]
- Corentin Segalas [UNIV BORDEAUX, Post-Doctoral Fellow, from Feb 2025 until Jun 2025]

PhD Students

- Kalidou Ba [INSERM]
- Antonin Colajanni [UNIV BORDEAUX]
- Lisa Crépin [UNIV BORDEAUX, from Oct 2025]
- Emie Delrieu [UNIV BORDEAUX, from Oct 2025]
- Sara Fallet [UNIV BORDEAUX]
- Thomas Ferte [UNIV BORDEAUX, until Oct 2025]
- Auriane Gabaut [INSERM, until Sep 2025]
- Ange-Marie Gouna [INSERM]
- Ariel Oscar Guerra Adames [UNIV BORDEAUX]
- Céline Hosteins [UNIV BORDEAUX, from Sep 2025]
- Arthur Hughes [UNIV BORDEAUX]
- Zhe Li [INSERM]
- Perrine Lunel [INRIA, from Oct 2025]
- Adrien Mitard De Girardier [INSERM]
- Takashi Noda [INSERM, from Oct 2025]

- Ansh Pal [INRIA]
- Justine Remiat [INSERM]
- Theo Rene [INRIA, from Oct 2025]
- Anne-Andrée Ruiz [INSERM, from Oct 2025]
- Nam-Anh Tran [INSERM, from Sep 2025]

Technical Staff

- Nicolas Boespflug [INSERM]
- Zeinab El Hajj [INRIA, Engineer]
- Jad El Karchi [UNIV BORDEAUX, from Feb 2025]
- Benjamin Hivert [INSERM]
- Mélanie Huchon [INSERM, until Aug 2025]
- Quentin Laval [INSERM]
- Myriam Maherzi [UNIV BORDEAUX]
- Yuriy Melnyk [INRIA, Engineer, from Sep 2025]
- Clément Nerestan [INRIA, Engineer, until Feb 2025]
- Anton Ottavi [INSERM]
- François Plessier [UNIV BORDEAUX]
- Alice Simon [INSERM]
- Panthea Tzourio [INSERM, Engineer, until Mar 2025]

Interns and Apprentices

- Lisa Crépin [UNIV BORDEAUX, Intern, from Apr 2025 until Sep 2025]
- Emie Delrieu [UNIV BORDEAUX, Intern, from Feb 2025 until Jul 2025]
- Marla Donnio [INRIA, Intern, from May 2025 until Jul 2025]
- Chloe Dumas De La Roque [UNIV BORDEAUX, Intern, until Aug 2025]
- Theodora Georgakopoulou [INSERM, Intern, until Jul 2025]
- Tatiana Gopinauth [UNIV BORDEAUX, Intern, until Mar 2025]
- Céline Hosteins [UNIV BORDEAUX, Intern, until Jul 2025]
- Diego Kauer [INRIA - Univ Chile, Intern, from Feb 2025 until Apr 2025]
- Lore Lafuente [UNIV BORDEAUX, Intern, from Feb 2025 until Jun 2025]
- Débora Marques De Oliveira [UNIV BORDEAUX, Intern, from Apr 2025 until Jun 2025]
- Marie Pelletier [UNIV BORDEAUX, Intern, from May 2025 until Aug 2025]
- Ahlam Pketoni [Lycée, Intern, from Feb 2025 until Feb 2025]

Administrative Assistants

- Ellie Correa Da Costa De Castro Pinto [INRIA]
- Sandrine Darmigny [INSERM]

Visiting Scientists

- John Barrera [UNIV VALPARAISO, from Jul 2025 until Jul 2025]
- Morgan Craig [Université de Montreal, from Jun 2025 until Sep 2025]
- Susana Eyheramendy [Universidad Adolfo Ibáñez, Chile, from Jul 2025 until Jul 2025]
- Cristian Meza [UNIV VAPARAISO, from Jul 2025 until Jul 2025]

External Collaborators

- Lucie Bourguignon [UNIV BORDEAUX]
- Hélène Savel [UNIV BORDEAUX]

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 beyond those encountered in the SISTM team. However, the focus devoted to vaccine development is justified by its importance from a public health perspective, and a long-standing expertise in this application field that maximizes the relevance and implementation of the methods developed. This equilibrium between the methodological and applied work reached over the last years is a fundamental motivation for each member of the SISTM team, regardless of complementary backgrounds across researchers (from applied mathematics to 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. Thus, we are able to collaborate for the development of new methods, and also to translate our innovations (either new analytical methods or applied results) to clinicians and immunologists – first in our collaborative networks, and then beyond. Figure 1 illustrates this synergy and materializes the three research axis of the team: high dimension statistical learning, mechanistic modelling, and translational vaccinology and design.

Biological and clinical research has dramatically changed thanks to technological advances, leading to the possibility of measuring many more biological parameters than previously thanks to high-throughput methods. Clinical research studies can now include traditional measurements such as clinical status, but also (tens of) thousands of cell populations, peptides, gene expressions, etc. for a given participant at a single time point. This has facilitated knowledge transfer 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 to obtain 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 scales (within 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" [104], "Systems vaccinology" [101], "Systems medicine" [87]. From the statistical point of view, two main challenges arise: i) to adequately deal with the massive amount of data, and ii) to find relevant models capturing observed data.

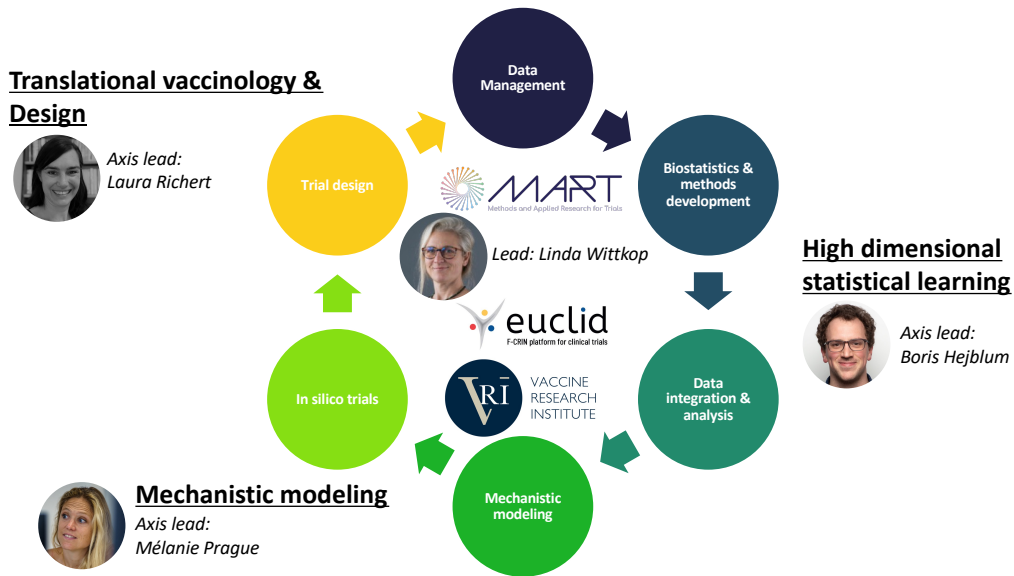


Figure 1: The SISTM wheel. Presentation of the three axes.

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 [97]). 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 [92]. 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 [98] 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 (also represented in part in the figure 1). Firstly, it is one of the rare 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 either Inserm teams or Inria teams from the two distinct 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 (10.2) and section Fundings (9).

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, Schiffer lab in Fred Hutchinson Cancer center) 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 and design (leader Laura Richert).

3.1 Axis 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 [107]). This ambitious endeavor goes beyond previous work done on this topic using standard deconvolution approaches [89]. By using more sophisticated statistical [78], machine learning [79], and artificial intelligence [109] 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[94]. The compositional nature of these data, along with their hierarchical phylogenetic structure particularly suited to tree-based models, coupled with their high-dimension requires the use of adequate statistical tools [106].

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 [75] 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 [83, 86, 105, 112, 84].

Regarding this research axis, there are some common interest with other Inria teams such as **HeKA**, **Soda**, and **PreMeDICaL** in regards of the use of machine learning approaches applied to medical data or **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 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.
- Between-host modeling of dynamics of epidemics in populations.
- Use mechanistic model as in silico platform for exploration of counterfactual scenarios with application in 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 [108]. 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 [74, 100]. 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 the heterogeneity between individuals, 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 [111], semi-parametric approaches [110] or penalized likelihood approach (in house NIMROD program [99]). The SAEM algorithm [91], as implemented in Monolix [103], is now also used for many of our projects. We however, continue to participate in the development of related methods in collaboration with ex - Inria team XPOP. We also devote a large part of this axis methodological research to the development of alternative methods for estimation in NLME-ODEs models.

From a computational perspective, the stochastic approximation of the EM algorithm (SAEM) provides accurate estimations for medium-sized parametric NLME-ODEs. For high-dimensional settings, alternative approaches to SAEM, such as those based on variational inference [90], have been proposed for generalized linear mixed models [95]. However, these methods have not yet been extended to NLME-ODEs. In the context of semi-parametric inference of ODEs, the universal approximation property of neural networks (NNs) has justified their use as proxies for missing model structures [113]. Nevertheless, this is usually limited to single-subject settings. While some studies have begun to consider population contexts [102, 93], these approaches remain inadequate for sparse data scenarios. A great amount of this axis work now focuses on estimation methods using concepts/devices coming from NNs, variational inference and inverse problem regularization, to construct high-dimensional, semi-parametric and properly regularized inference methods for mechanistic models, in the vein of hybrid modeling.

The integration of ordinary differential equation (ODE) models in our work enables a detailed examination of within-host and between-host dynamics of infectious diseases. At the within-host level, ODE models describe the interactions between pathogens and host immune responses, such as viral replication and immune clearance. These models provide insights into mechanisms like virus propagation and immune cell dynamics, as demonstrated for example in studies on HIV, Ebola [96, 77] and SARS-CoV-2 [76, 80]. Regarding the between-host dynamics, we extensively described the COVID-19 pandemics inferring the effect of vaccination and non-pharmaceutical interventions [82, 85].

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 SCOOOL) and learning approaches (Inria team FLOWERS) may serve for defining optimal interventions or optimal designs to evaluate new interventions. We made a proof of concept of such open-loop control problem in the within-host setting. 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 [88]. We also made a proof of concept of such open-loop control problem in the between-host setting. [81]. We introduced *EpidemiOptim*, a Python toolbox designed to optimize epidemic control policies through the integration of epidemiological models and machine learning algorithms, including reinforcement learning and evolutionary algorithms. The toolbox's utility is demonstrated through a case study optimizing COVID-19 lockdown policies, balancing health outcomes and economic impacts using a Susceptible-Exposed-Infectious-Removed (SEIR) model fitted to French data. We still devote a large part of this axis methodological research to the development of methods around personalized medicine and targeted numerical public health.

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, vaccinology 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. Regarding the modeling, our international competitors and collaborators are Perelson's lab in Los Alamos USA and Schiffer's lab in Fred Hutchinson USA. Finally, our work on in silico simulation is closely related to a digital twin of clinical trials. In this sense, it can be compared to the work conducted by the SIMBIOTX team at Inria.

3.3 Axis Translational vaccinology and design

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 adaptative 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.

Research on novel trial designs for early phase vaccine trials is carried out by the team within PEPR Santé Numérique SMATCH, and with PhD thesis (such as the Inserm-Inria funded thesis on multi-armed bandit algorithms for vaccine trials by Cyrille Kone; co-supervisor E Kaufmann Inria Lille).

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 visualization 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 developed.

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 developing 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 recommendations. Participated in the elaboration of SPIRIT/CONSORT Extension for Surrogate endpoints (2023)
- Laura Richert is the coordinator of the working group "Greener Clinical Research" (Décarbonation de la recherche clinique) within the Recap/Inserm network. She is also a member of the "Greener Trials" network (MRC, UK) and a member of the "Sustainable Development" working group of the CNCR. Thomas Ferté has developed an R package called CarbPack R, designed to facilitate the estimation of the carbon footprint of statistical analyses in R on a local computer. The package serves as a wrapper for the [Green Algorithm calculator](#).

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)
- Participant as "Inventor" (1/7th) for patent WO2021058914A1/FR1910515 on "Prediction of the content of omega-3 polyunsaturated fatty acids in the retina by measuring 7 cholesterol ester molecules"

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-pneumococcal vaccine trial PNEUMOVAS (NCT03069703)
 - French phase II anti-pneumococcal 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

6.1 Major scientific Publication - RISE: Identification des marqueurs de substitution en haute dimension, appliquée à la vaccinologie

This work [32], published in statistics in Medicine, proposes a new method – called RISE – for identifying surrogate markers from very high-dimensional biological data in vaccine trials. The method combines a non-parametric rank-based test to select candidate variables, followed by validation of the selected markers on independent data. Simulations show that RISE correctly controls the type I error rate while maintaining good power, even in low-sample, high-dimensional contexts. The application of RISE to an influenza vaccination trial, using gene expression data to identify candidate genes as surrogates for immune responses, highlights a set of genes, particularly in pathways related to interferon and innate antiviral activation, that may predict early vaccine-induced immune responses. These surrogate markers could enable faster and less costly evaluation of vaccine immunogenicity.

6.2 Major scientific Publication - Comparative evaluation of methodologies for estimating the effectiveness of non-pharmaceutical interventions in the context of COVID-19

The article [62], published in american journal of epidemiology, analyses the reliability of methods used to estimate the effectiveness of non-pharmaceutical interventions implemented during the COVID-19 pandemic, such as lockdowns and mask wearing. The authors carry out a series of simulations that reproduce different transmission dynamics and varying levels of contact heterogeneity. They show that certain approaches can produce biased estimates when models oversimplify the temporal evolution of the epidemic or neglect

the diversity of social behaviours. The study highlights that confidence intervals are highly dependent on structural assumptions. It thus emphasises that rigorous methodological analyses, particularly those based on mechanistic models, are essential before using these estimates to inform public health decisions.

6.3 Organisation of highly recognized national and international conferences

The EPICLIN/JSCLCC 2025 conference, organised by Laura Richert in Bordeaux from 14 to 16 May 2025 at the University of Bordeaux, brought together the French clinical epidemiology community (around 300 French-speaking participants). It covered methodological advances in clinical trials, causality, artificial intelligence and biomedical data. Discussions focused on the integration of new quantitative approaches to improve clinical evaluation and health research.

The Workshop on Virus Dynamics 2025, organised by Mélanie Prague at the University of Bordeaux, was held from 14 to 16 October 2025 (approximately 150 international participants). This workshop brought together virologists, immunologists and mathematical modellers to discuss viral dynamics and immune responses. This edition placed particular emphasis on vaccine development, adaptive and innate immunity, modelling of chronic, emerging or viral infections, and host-pathogen interaction.

6.4 Awards-fundings

6.4.1 Renewal vaccine Research institute

In 2025, the Vaccine Research Institute (VRI), a LabEx laboratory of excellence in Créteil closely linked to the BPH-SISTM team, was favourably evaluated by the French scientific authorities and had its funding renewed as part of major public programmes (ANR, PIA, France 2030). Within the VRI, the SISTM team leads the data science division, developing statistical and mechanistic modelling methods for early-stage (phase I/II) vaccine trials, with a particular focus on HIV, Ebola and pan-Coronavirus. SISTM is involved in every stage of the process, from the design of vaccine strategies to their *in silico* optimisation. The team collaborates closely with the Bordeaux teams at UMS 54 MART and methodological platforms (USMR Bordeaux and F-CRIN Euclid) and receives funding for vaccine trials coordinated by the VRI, as well as for Inserm-Inria co-funded theses and CIFREs for methodological development applied to vaccines. The VRI-SISTM collaboration is part of structural funding: LabEx VRI (PIA), major ANRS/VRI projects and national programmes, including the PEPR Santé Numérique (SMATCH project on innovative vaccine trial designs).

6.4.2 Launch of ANRS PEPR MIE Previs (WP leader within-host modelling 2025-2028)

The Previs project, funded by the ANRS MIE PEPR, aims to strengthen national preparedness for a future unknown virus ('virus X') by developing quantitative methods for early risk assessment, modelling and decision support. It is led by Mircea Sofonea, University of Montpellier. It focuses on six areas of work relating to pathogen characterisation, intra-host dynamics, phylodynamics, epidemiological inference, behavioural responses and hospital preparedness, with a view to producing operational tools for surveillance and control. It will provide methodological advances, interoperable tools and trained specialists to support public health agencies and optimise real-time responses to future health crises. Mélanie Prague (SISTM) is leading the intra-host modelling component.

6.4.3 Launch of ANRS PEPR MIE Previs (WP leader within-host modelling 2025-2028)

The CAIR project aims to improve the efficiency of randomised clinical trials by integrating real-world data to increase control groups. It is led by Yohann Foucher, University of Poitiers. It develops methods based on G-computation, mechanistic models and reinforcement learning to reduce the sample sizes required. Linda Wittkop (SISTM) is leading the section on the use of mechanistic models.

7 Latest software developments, platforms, open data

7.1 Latest software developments

7.1.1 SurrogateRank

Name: SurrogateRank R Package

Keywords: Biostatistics, Surrogate markers, Transcriptomics

Scientific Description: Implementation of a rank-based, nonparametric approach to evaluate low or high-dimensional surrogate markers in the small sample size setting.

Functional Description: Uses a novel rank-based nonparametric approach to evaluate a surrogate marker in a small sample size setting. Details are described in Parast et al (2024) <doi:10.1093/biomtc/ujad035> and Hughes A et al (2025) <doi:10.1002/sim.70241>. A tutorial for this package can be found at <<https://www.laylaparast.com/surrogaterank>> and a Shiny App implementing the package can be found at <<https://parastlab.shinyapps.io/SurrogateRankApp/>>.

Release Contributions: Extension to the high-dimensional setting using the methods described in Hughes A et al (2025) <doi:10.1002/sim.70241>.

URL: <https://cran.r-project.org/web/packages/SurrogateRank/index.html>

Publication: hal-04933460

Contact: Arthur Hughes

Partners: INSERM, INSERM U1219 Bordeaux Population Health, Equipe SISTM

7.1.2 RastaRocket

Name: RastaRocket: Rocket-Fast Clinical Research Reporting

Keyword: Biostatistics

Functional Description: Description of the tables, both grouped and not grouped, with some associated data management actions, such as sorting the terms of the variables and deleting terms with zero numbers.

URL: <https://github.com/biostatusmr/RastaRocket>

Contact: Thomas Ferté

Partner: CHU de Bordeaux

7.1.3 reservoirnet

Name: reservoirnet: Reservoir Computing and Echo State Networks

Keywords: Reservoir Computing, Recurrent network

Functional Description: Une bibliothèque simple et conviviale basée sur le module Python `reservoirpy`. Elle offre une interface flexible pour implémenter des architectures efficaces de `Reservoir Computing (RC)`, avec un accent particulier sur les `Echo State Networks (ESN)`. Parmi ses fonctionnalités, on trouve : l'entraînement en mode hors ligne et en ligne, l'implémentation parallèle, le calcul avec des matrices creuses, une initialisation spectrale rapide, des règles d'apprentissage avancées (par ex. `Plasticité intrinsèque`), etc.

Elle permet également de créer facilement des architectures complexes avec plusieurs réservoirs (par ex. `réservoirs profonds`), des couches de sortie (`readouts`) et des boucles de rétroaction complexes. De plus, des outils graphiques sont inclus pour explorer facilement les hyperparamètres. Enfin,

plusieurs tutoriels sont disponibles, abordant la **prédiction de séries temporelles**, la classification et l'optimisation des hyperparamètres.

Pour plus d'informations sur **reservoirpy**, veuillez consulter **Trouvain et al. (2020)** <doi:10.1007/978-3-030-61616-8_40>. Ce package a été développé dans le cadre du programme **IdEx "Investissements d'Avenir"** de l'Université de Bordeaux / **RRI PHDS**.

Contact: Thomas Ferté

7.1.4 REMixed

Keywords: Mechanistic modeling, Nonlinear mixed effects models, Non-Linear Mixed Effects modelling

Functional Description: REMixed is an R package that enables parameter estimation in nonlinear mixed-effects models, based on the integration of longitudinal measurements assumed to be linked to a latent compartment of a mechanistic model.

URL: <https://cran.r-project.org/web/packages/REMixed/index.html>

Contact: Melanie Prague

7.1.5 LSAMBA

Keywords: Nonlinear mixed effects models, Non-Linear Mixed Effects modelling, Mechanistic modeling

Functional Description: LSAMBA is an R package that extends the SAMBA (Stochastic Approximation for Model Building Algorithm) by incorporating a LASSO-type penalization to facilitate the automatic construction of mechanistic nonlinear mixed-effects models in the presence of high-dimensional covariates.

URL: <https://cran.r-project.org/web/packages/LSAMBA/index.html>

Contact: Melanie Prague

7.2 New platforms

The developed platform, named DTR (Data To Research), is a web-based solution designed as a direct alternative to LabKey (used for 10 years in the SISTM team for data management, access and sharing), with a more flexible architecture and closer alignment with current research needs. It is based on a fully containerized architecture and structures research activities around projects, clinical trials, and datasets that are integrated in a seamless way. Compared with LabKey, it provides a wider choice of customizable working environments, including RStudio with multiple R versions, Python, and JupyterLab, improved usability, and stronger control over data security, access rights, and data export traceability. It also does not link SISTM with an exterior provided who may raise the cost - thus ensuring sovereignty. In 2025, a functional proof of concept was designed and implemented, demonstrating the technical feasibility of replacing LabKey for both data management and execution of research environments, while addressing several key limitations of the existing solution. In the longer term, the objective is to evolve this platform into a scalable, multi-tenant solution that can be deployed across multiple organizations, with advanced security mechanisms, fine-grained role management, and the progressive integration of data analysis and AI-based tools, in order to provide a sustainable platform aligned with future needs in SISTM team and broader health research.

Participants: Selected data management system adopted by the team to handle health data in compliance with legal requirements.

8 New results

8.1 High-dimensional statistical learning

8.1.1 Identification of surrogates of protection

Participants: Boris Hejblum, Arthur Hughes, Rodolphe Thiébaud.

High dimensional 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 developed RISE, a two-stage, rank-based framework to screen and evaluate surrogate markers in high-dimensional settings with limited sample sizes. RISE focuses on trial-level surrogacy, comparing treatment effects on candidate markers to the treatment effect on the primary outcome using nonparametric inference and sample splitting. Applied to influenza vaccination, RISE identifies and validates an early transcriptomic signature that reliably captures the vaccine effect on antibody responses [9]. The RISE method is implemented in the R package `surrogateRank` (7.1.1).

8.1.2 Reservoir computing for epidemic prediction

Participants: Thomas Ferté, Boris Hejblum, Rodolphe Thiébaud.

We developed and evaluated reservoir computing approaches for epidemic forecasting using real-world COVID-19 data. First, we introduced `reservoirnet` (7.1.3), an R package allowing to perform reservoir computing in R through an API to the `reservoirpy` python library, and illustrated its application to the 14-day forecasting of SARS-CoV-2 hospitalizations in France using public data (high-dimensional epidemiological, clinical, and environmental time series) showing that reservoir models outperformed elastic-net regression despite strong non-stationarity in the data [29]. We also discuss the pro and cons of such a prediction framework in an emergency context such as the emergence of a new pathogen like the early COVID-19 epidemics [61], compared to a simpler, knowledge-driven, epidemiological approach.

Thomas Ferté defended their PhD on this topic in 2025 [58].

8.1.3 Machine learning/AI and bias

Clinical decision-making may be affected by cognitive biases that introduce non-medical, socially constructed factors into care pathways. When machine learning methods are used for decision support, they may be trained on biased data, thereby reproducing or amplifying healthcare inequities. Ariel Guerra's PhD (co-supervised by Marta Avalos) focuses on identifying biases in clinical data, with key challenges including disentangling bias sources and developing experimental protocols and evaluation metrics, as addressed in recent work [30, 63]. These activities were also disseminated to the general public through outreach events (University of Bordeaux Inclusivity Month, March; Fête de la Science at Cap Sciences, October) and a France Culture interview on medical biases and women's health (December).

In addition, an ENLIGHT ETN project on gender impact in health (IMPULSE), involving seven universities from the ENLIGHT European alliance, has been funded. Our contribution focuses on expertise in public health, biostatistics, and medical informatics, with particular emphasis on biases that may be reproduced by algorithms. On the other hand, current research explores counterfactual approaches in collaboration with the REGALIA team.

Participants: Marta Avalos, Ariel Guerra-Adames.

8.1.4 Compositional microbiome data

The VALPO Associated Team was created in 2025 as an extension of the AMSud SMILE collaboration between the University of Valparaíso and the SISTM team, and expanded to include the PLEIADE team and Inria Chile. This initiative strengthened collaborative research activities. A cotutelle PhD (Céline Hosteins, University of Valparaíso / University of Bordeaux) began in September 2025, and two PhD students completed 3-4 week research stays. An invited session on microbiome-based biomarkers was organized at the 65th ISI World Statistics Congress with participation from VALPO partners [47]. First joint publications have been produced [46, 49, 50, 67, 70].

Participants: Marta Avalos, Antonin Colajanni, Céline Hosteins, Diego Kauer, Rodolphe Thiébaud.

8.2 Mechanistic learning

8.2.1 Between host modeling of COVID-19 Epidemics

Participants: Mélanie Prague, Rodolphe Thiébaud.

We investigated the performance of mechanistic models, compared with traditional regression approaches, to estimate the basic reproduction number R_0 and the effectiveness of non-pharmaceutical interventions (NPIs). We showed that mechanistic models perform better in situations where traditional methods can be biased, for example when the susceptible population size decreases rapidly [62]. This is one of the highlight of the year.

Two more theoretical work on epidemics have been released. [42] compares polarised and leaky immunity assumptions in epidemiological models using a non-Markovian framework that accounts for immune age and waning. It shows that leaky immunity leads to more frequent reinfections and more infections overall, but fewer hospitalisations, highlighting the strong impact of immunity modelling choices on long-term epidemic predictions and public health decisions. [41] extends evolutionary invasion analysis by allowing immunity to wane and epidemiological conditions to vary over time, using a two-strain non-Markovian model with cross-immunity inspired by SARS-CoV-2. It shows that higher transmissibility, immune escape, and the timing of mutant introduction strongly determine invasion success, highlighting the role of immune waning and non-equilibrium dynamics in pathogen evolution.

8.2.2 Modeling viruses and immune dynamics

Participants: Mélanie Prague.

[38] shows that exposure history strongly determines SARS-CoV-2 Omicron viral dynamics in non-human primates. Bivalent vaccination reduces viral replication more than monovalent vaccination, while hybrid immunity (prior infection plus vaccination) provides near-complete protection. A mechanistic model separates the effects of antibody binding, neutralisation, and a strong non-antibody effect of prior infection that accelerates clearance of infected cells. The novelty lies in quantifying how exposure history, beyond antibody levels alone, shapes viral control and enables prediction of correlates of protection against infection and transmission. This is an important work mechanistically linking a virus and immune response model.

8.2.3 PhD Defense

Participants: Mélanie Prague.

Auriane Gabaut defended her PhD [59] under the supervision of mélanie Prague and Cécile Proust-lima (Inserm BPH, Biostat). This thesis is about developing new statistical and mechanistic methods to better understand variability in vaccine-induced immune responses. It focuses on integrating high-dimensional transcriptomic data into mechanistic ODE models with mixed effects, using penalised approaches combined with SAEM. The work addresses both fixed-time and longitudinal transcriptomic data, aiming to link biomarkers to unobserved immune processes. The methods are implemented in R packages and applied to Ebola, varicella, and SARS-CoV-2 vaccine studies.

8.3 Translational vaccinology and design

8.3.1 Bandit Approach for trial design and PhD Defense

Participants: Laura Richert.

These works [51] and [52] are motivated by applications in multi-criteria decision-making, such as recommendation systems, resource allocation, and experimental design, where several conflicting objectives must be optimized simultaneously under uncertainty. This is occurring in vaccine trial designs. The first studies Pareto Set identification with feasibility constraints and proposes a near-optimal fixed-confidence algorithm with matching lower bounds. The second addresses Pareto Set identification in structured linear bandits and introduces optimal design-based algorithms with near-optimal guarantees. Together, they advance theory and methods for efficient multi-objective learning.

Cyril Kone defended his PhD under the supervision of Laura Richert and Emilie Kauffman (Inria Lille, SCOOL).

8.3.2 Response to COVID-19 vaccine in immunosuppressed populations

Participants: Linda Wittkop.

Linda Wittkop is involved as a principal investigator in the COV-POPART cohort. The cohort was established to study the immune response to Covid-19 vaccination and its persistence in individuals with immune disorders, with a particular focus on characterizing vaccine failures (immunological and virological). A total of 6,112 adults affected by 10 different pathologies are participating in the study. The statistical analysis of the data generated has started leading to first publications [36].

8.3.3 High-resolution transcriptomics

Participants: Boris Hejblum, Edouard Lhomme, Rodolphe Thiébaud.

Whole-blood gene expression analysis is essential for understanding molecular vaccine responses, yet its use typically relies on venous sampling, limiting feasibility for frequent and remote monitoring. In an ancillary study of the COVERAGE France platform trial (NCT04356495), which enrolled at-risk outpatients with mild coronavirus disease 2019 (COVID-19) monitored at home, we compared transcriptomic profiles obtained from paired venous blood and ultralow-volume, self-collected, finger-prick capillary samples. We observed moderate to good concordance at the individual gene level and excellent agreement at the gene-set

level between the two sampling approaches. High-frequency finger-prick sampling enabled daily resolution of immune dynamics, revealing early interferon responses, sustained neutrophil activation, and evolving erythroid and inflammatory signatures during the initial phase of mild COVID-19 [44]. These results pave the way for deploying finger-prick sampling as a reliable approach for at-home transcriptomic profiling, promising a powerful tool for longitudinal immune monitoring and advanced clinical research in future trials.

8.3.4 Knowledge transfer

Participants: Rodolphe Thiébaud, Laura Richert, Melanie Prague, H el ene Savel.

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).

9 Partnerships and cooperations

9.1 International initiatives

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

VALPO

Title: Valid statistical Analysis of Longitudinal compositional and high-dimensional microbiome data to Predict health Outcomes

Duration: 2025 → 2028

Coordinator: Marta Avalos

Partners:

- Universidad de Valpara so, Chile
- Universidad Adolfo Ib a nez, Chile
- Inria Chile
- PLEIADE project-team
- SISTM project-team
- CHU Bordeaux

Inria contact: VALPO Bordeaux, VALPO Chile

Summary: The VALPO project (Valid statistical Analysis of Longitudinal compositional and high-dimensional microbiome data to Predict health Outcomes) aims to advance statistical methods for analyzing complex microbiome data. This collaboration focuses on longitudinal, compositional, and high-dimensional datasets, which are challenging due to their sparsity, zero-inflation, and intricate dependencies over time. The project builds on previous efforts involving the initial core teams of Inria SISTM, Universidad de Valpara so (CIMFAV), CHU Bordeaux, and Inserm. It is further strengthened by the addition of new teams, including Pleiade, Universidad Adolfo Ib a nez, and Inria Chile, bringing complementary expertise. The research will extend previous work through new methodologies for visualizing longitudinal microbiome data, SAEM-based approaches for identifying disease-associated microbial features, and machine learning techniques to predict health outcomes. The collaboration combines statistical expertise from Chile and France, with applications in chronic diseases such as

asthma, cystic fibrosis, and the analysis of microbial translocation in blood samples to predict vaccine responses. These examples are not exhaustive, as other applications using open-access data from published studies will also be explored to illustrate the utility of the algorithmic developments.

9.1.2 STIC/MATH/CLIMAT AmSud projects

SMILE

Participants: Marta Avalos.

Title: Program MATH AmSud 2023, Chile, Uruguay, France (SMILE, code 23-MATH-12)

Partner Institution(s): Universidad de Valparaiso, and Universidad Adolfo Ibáñez, Chile

Date/Duration: 2 years (until Dec 2025)

Additional info/keywords: Statistical Modeling, nonparametric Inference, and model selection for complex data

9.1.3 Participation in other International Programs

MUSICC

Participants: Rodolphe Thiébaud, Boris Hejblum .

Title: Controlled Human Infection Models For Beta-coronaviruses in order to assess vaccine effects

Partner Institution(s):

Date/Duration: The project has started on February 1st, 2024. Duration 60 months, 01/02/2024 - 31/01/29

Additional info/keywords: Selected for funding by CEPI (Coalition for Epidemic Preparedness Innovations). This project is rather unique in Europe by both the quality of the participants and its approach. In this context, SISTM will contribute to the data analysis and the modeling of the immune response. 355,000 USD.

9.2 International research visitors

9.2.1 Visits of international scientists

Cristian Meza

Status Professor

Institution of origin: Universidad de Valparaiso

Country: Chile

Dates: 2-9 Jul 2025

Context of the visit: Research meeting of the Associate Team VALPO

Mobility program/type of mobility: MATH AmSud

Susana Eyheramendy

Status Professor

Institution of origin: Universidad Adolfo Ibáñez

Country: Chile

Dates: 2-9 Jul 2025

Context of the visit: Research meeting of the Associate Team VALPO

Mobility program/type of mobility: MATH AmSud

John Barrera

Status PhD student

Institution of origin: Universidad de Valparaiso

Country: Chile

Dates: 4-25 Jul 2025

Context of the visit: Research stay - Associate Team VALPO

Mobility program/type of mobility: Associate Team VALPO

Morgan Craig

Status Professor

Institution of origin: University of McGill

Country: Canada

Dates: May-July 2025

Context of the visit: Research stay

Mobility program/type of mobility: Sabbatical

John Fricks

Status Professor

Institution of origin: Arizona state University

Country: USA

Dates: July 2025

Context of the visit: Research stay

Mobility program/type of mobility: Follow-up from last year sabbatical

9.2.2 Visits to international teams

Research stays abroad

Céline Hosteins**Visited institution:** Universidad de Valparaiso**Country:** Chile**Dates:** 5 Nov - 5 Dec 2025**Context of the visit:** Associate Team VALPO - Cotutelle agreement between University of Bordeaux and Universidad de Valparaiso**Mobility program/type of mobility:** research stay**Ariel Guerra****Visited institution:** Department of Preventive Medicine, Kyoto School of Public Health, Kyoto University**Country:** Japan**Dates:** 11 Jul - 30 Jul 2025**Context of the visit:** SP+ Fund ECR Program of Kyoto University**Mobility program/type of mobility:** research stay**Sara Fallet****Visited institution:** UC Berkeley**Country:** USA**Dates:** 18 Apr - 16 May 2025**Context of the visit:** Collaboration with Elizabeth Purdom on scRNA-seq data**Mobility program/type of mobility:** France Berkeley Fund**9.3 European initiatives****9.3.1 H2020 projects**

SOLVE: The project funded by Horizon Europe has started on January 1st, 2024 to decipher the mechanisms of induction of long-lasting immunity through a comparison of vaccine platforms and to advance new vaccine concepts. In the project, SISTM is workpackage leader (WP7 Data Science) and will analyze the consortium's data to model the immune response of the 4 main different types of COVID19 vaccine platforms and some variants. SISTM will thus contribute to the comparison of these platforms and the discussion to present recommendations to stakeholders to support future epidemic preparation decision-making. Duration: 60 months 01/01/24 - 31/12/28. 563 330 Euros.

Participants: Mélanie Prague, Rodolphe Thiébaud, Boris Hejblum, Linda Wittkop.

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. In this project, SISTM will work on the analysis of data from the clinical intervention and the modelisation of the response to the treatment. L Wittkop. Duration: 60 months 01/01/20-31/12/25. 409,632 Euros.

Participants: Mélanie Prague, Boris Hejblum, Linda Wittkop.

9.3.2 Other european programs/initiatives

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 health 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.

Participants: Edouard Lhomme, Rodolphe Thiébaud, Laura Richert, Boris Hejblum, Mélanie Prague.

ETN IMPULSE: IMPULSE “Interdisciplinary Network on the Impact of Sex and Gender in Health” has been successfully selected as part of the ENLIGHT THEMATIC NETWORKS 2025 initiative. This project aims to establish a sustainable and interdisciplinary collaborative platform to address a critical but often overlooked dimension of health and wellbeing: the role of sex and gender differences in health outcomes, research, education, and clinical practice. This network brings together a diverse alliance of ENLIGHT institutions to promote structural change, foster innovation, and enhance equity in the health sciences by integrating sex and gender perspectives. In our journey, we intend to put sex & gender factors on the agenda of health education, research and practice, looking to achieve stronger consolidation of equity, diversity and multiculturalism.

Main coordinator: University of the Basque Country, Funding Granted: 98,500 € from Nov 2025 to Oct 2027. Partners: University of Bordeaux (Lead contact: Marta Avalos, budget granted 21,000 €), University of Galway, Comenius University Bratislava, University of Göttingen, Uppsala University, University of Bern.

Bordeaux is contributing based on our work on detecting clinical judgment biases using AI-based tools. Bordeaux will host a summer school on this topic on June 17–18–19, 2026: **IMPULSE**

Participants: Marta Avalos, Ariel Guerra.

9.4 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. 2012-2025, 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 ParisEst 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).

Participants: Mélanie Prague, Laura Richert, Boris Hejblum, Rodolphe Thiébaud, Edouard Lhomme, Quentin Clairon, Linda Wittkop.

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 Thiébaud. 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.

Participants: Rodolphe Thiébaud.

PEPR Santé Numérique SMATCH: The PEPR SN SMATCH coordinated by Inria and co-coordinated by Sarah Zohar (HEKA) and Rodolphe Thiébaud (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 €

Participants: Mélanie Prague, Laura Richert, Boris Hejblum, Rodolphe Thiébaud.

PEPR Santé Numérique Programme 1 Axe AI4scMed: Cell-based precision medicine holds revolutionary potential for healthcare, but realizing its full potential demands a deep understanding of disease variability and multiscale aspects. Single-cell (sc) multi-omics offers a unique way to obtain molecular profiles of individual cells and predict disease trajectories. To harness this complexity, new AI breakthroughs are needed. Our consortium will tackle methodological challenges to bridge the gap between

sc data and personalized treatments, resolving cell type differences and integrating sc-multi-omics with imaging for spatial insights. Addressing the complexity of the human body and combining genomics with other assays, the goal is to develop AI-based methods to handle, integrate, analyze, and visualize multiscale complexity in diseases, and to leverage cutting-edge AI for sc-genomic data analysis. To infer causal mechanisms at different levels, causal/logical/stochastic modeling can be used to integrate heterogeneous data and account for temporal scales and biophysical priors. Boris Hejblum is task leader within this consortium. SISTM budget: 136 388 €

Participants: Boris Hejblum.

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 €

Participants: Mélanie Prague, Laura Richert, Rodolphe Thiébaud, Linda Wittkop.

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.

Participants: Rodolphe Thiébaud.

PEPR PREVIX: The Pandemic preparedness to Respiratory Virus X, integrative modelling from first cases to early public health countermeasures (PREViX) is a The last two decades have been marked by a series of outbreaks and pandemics of respiratory viruses. These were due either to new strains of influenza A virus (IAV) – avian influenza H5N1 in 2005 and 2024, swine flu H1N1 in 2009 – or to new species of betacoronaviruses – SARS-CoV in 2002-2004, MERS-CoV in 2014-2015, SARS-CoV-2 since 2020. Although different in terms of biology, time, space and impact, all these episodes highlight the challenges faced by governments to anticipate, assess, manage, and control emerging and re-emerging respiratory viruses that pose a threat to health security (and even beyond, as illustrated by worldwide lockdowns in spring 2020). For the next (re-)emergent respiratory virus, addressing these challenges will be much more feasible if we can

1. map the public health threat it poses to France before importation, using solely key indicators from abroad,

2. extrapolate accurately the within-human viral-immune dynamics using early non-human primate data,
3. estimate the effective reproduction number and infection duration based on the viral sequences isolated from the first patients
4. better characterise antibody dynamics and forecast the epidemic trajectory using viral antigenic distance,
5. anticipate the effectiveness of behavioural leverages and non-pharmaceutical countermeasures
6. plan, before the outbreak, the optimal hospital activity management at the peak.

The PREViX project aims to develop the tools for each of these six open problems and to solve related methodological questions in corresponding six work packages, focusing on the specific case of respiratory viruses. 01/09/2025 - 31/08/2028. Total budget : 1,416 ,037 €, SISTM budget: 144 000 €

Participants: Mélanie Prague.

AAP Messidore CAIR: The project Clinical trials Augmented with Real-word data (CAIR) has for overall objective to investigate original methods for a hybrid RWD-RCT and more specifically:

1. To augment control arms of the RCT with RWD-based super learner
2. To augment control arms with mechanistic models
3. To augment control arms with reinforcement learning and/or generative AI
4. To illustrate the usefulness and disseminate the previous developments

The project started on 01/05/2025 and will last until 30/06/2028. Total budget: 623 023 € SISTM budget: 145 114 €.

Participants: Linda Wittkop, Rodolphe Thiébaud, Laura Richert, Mélanie Prague

9.4.1 Various Partnership

Mélanie Prague: Chaire Digital Innovation and Health Data Science program of the Center for Applied Mathematics CMAP at the Ecole Polytechnique

Rodolphe Thiébaud is Adjunct professor, Department of Epidemiology, Biostatistics and Occupational Health, McGill University since 2023

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).
- 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).

9.5 Regional initiatives

EMERG (Exposome microbien et Risque sanitaire: intérêt d’une Gestion One Health des enjeux liés aux gripes zoonotiques) funded by PSGAR (Programmes Scientifiques de Grande Ambition Régionale). (Principal PI L Delhaes and D Malvy).

Participants: Marta Avalos.

10 Dissemination

Participants: Marta Avalos, Quentin Clairon, Boris Hejblum, Ariel Guerra, Edouard Lhomme, Mélanie Prague, Laura Richert, Rodolphe Thiébaud, Linda Wittkop.

10.1 Promoting scientific activities

10.1.1 Scientific events: organisation

General chair, scientific chair

- 7th Workshop on Viral Dynamics, Bordeaux, Oct. 14–16, 2025 (General Chair: Mélanie Prague).
- 3rd Workshop on Respiratory Viruses of the ANRS–MIE Joint Action, Paris, Dec. 11–12, 2025 (Co–General Chair: Edouard Lhomme).
- International Conference “Social Inequalities and the Exposome”, Bordeaux, Sept. 18, 2025 (General Chair: Quentin Clairon).
- Workshop “Causal Inference in High–Dimensional Settings”, Bordeaux, Sept. 19, 2025 (General Chair: Quentin Clairon).
- Workshop “In silico Case Studies: From Research to Pedagogical Innovation”, Bordeaux, July 7, 2025 (Organizer: Marta Avalos), within the CAP IA and CAP Santé Numérique projects, University of Bordeaux.
- Marta Avalos organized an invited session, ISP 927 Statistical Tools for Microbiome–Based Biomarker Identification and Disease Prediction, within the 65th ISI World Statistics Congress of the International Statistical Institute, on October 9, 2025, in The Hague, Netherlands. Three members of EA Valpo took part, representing the University of Valparaíso (Cristian Meza), the PLEIADE team (Simon Labarthe), and the SISTM team (Antonin Colajanni).

Member of the organizing committees

- 7th Workshop on Viral Dynamics, Bordeaux, Oct. 14–16, 2025 (all SISTM team members).
- EUCLID Annual Scientific Day on challenges of clinical trials conducted in the Global South, Bordeaux, Dec. 2025 (Edouard Lhomme).
- Workshop on Platform Clinical Trials, training session within the EPICLIN Conference, Bordeaux, May 13, 2025 (Edouard Lhomme).
- International Conference “Social Inequalities and the Exposome”, Bordeaux, Sept. 18, 2025 (Quentin Clairon).
- Workshop “Causal Inference in High–Dimensional Settings”, Bordeaux, Sept. 19, 2025 (Quentin Clairon).

10.1.2 Scientific events: selection

Chair of conference program committees

- 7th Workshop on Viral Dynamics, Bordeaux, Oct. 14–16, 2025 (Mélanie Prague).

Member of the conference program committees

- ANRS–MIE AC Modelling Scientific Day, Rennes, Nov. 12–14, 2025 (Mélanie Prague).
- EPICLIN 2025 and JSCLCC 2025 Conferences – Member of the Scientific Committee since 2024 (Linda Wittkop).
- International Workshop on HIV and Hepatitis Observational Databases (IWHOD) – Member of the Scientific Committee (Linda Wittkop).
- International Workshop on HIV and Hepatitis Observational Databases (IWHOD) – Member of the Scientific Committee since 2013 (Rodolphe Thiébaud).
- DATAQUITAINE Conference, Bordeaux, March 2025 (Marta Avalos).
- ML4H – Machine Learning for Health, San Diego, USA, Dec. 2025 (Marta Avalos).
- 10th CNC – Channel Network Conference, Liège, Belgium, 2025. (Boris Hejblum).
- “Mathematics of Single-Cell Data-Analysis” research school at C.I.R.M., Marseille, France, 2025. (Boris Hejblum).

Reviewer

- CHIL 2025 – Conference on Health, Inference, and Learning, New York, USA, June 2025 (Marta Avalos).

10.1.3 Journal

Member of the editorial boards

- Editor of a topical collection on Virus Dynamics and Immunity, *Bulletin of Mathematical Biology*, 2025–2026 (Mélanie Prague).
- Reproducible Research Editor, *Biometrical Journal* (Boris Hejblum).
- Associate Editor, *International Journal of Biostatistics* (Mélanie Prague).
- Associate Editor, *Biometrics* (Mélanie Prague and Boris Hejblum).

Reviewer – reviewing activities

- Biometrics; CPT: Pharmacometrics & Systems Pharmacology; Biometrical Journal; Frontiers in Immunology; eLife; PLOS Computational Biology (Mélanie Prague).
- Biometrics; PCI Mathematical & Computational Biology (Boris Hejblum).
- New England Journal of Medicine (Laura Richert).
- Mathematical Modelling of Natural Phenomena (2025); Biometrics (2025); PLOS Computational Biology (2025, Guest Academic Editor) (Quentin Clairon).
- PLOS Computational Biology (Linda Wittkop).
- IMIA Yearbook, Public Health and Epidemiology Informatics Section (Marta Avalos).

10.1.4 Invited talks

- “Ebola Vaccine Development: How Modeling Helped?”, Invited Symposium, Society for Mathematical Biology, Edmonton, Canada, July 2025 (Mélanie Prague).
- “High-Dimensional Marker Integration in Mechanistic Models”, Journée Biostat/Martha Bio Santé, Nov. 2025 (Mélanie Prague).
- “Integrating large to high markers in mechanistic models”, Pharmacometrics in France, Sept. 2025 (Mélanie Prague).
- “Platform Trials: Clinical Opportunities and Methodological Challenges”, SESTIM Webinar, Oct. 2025, and GIRCI Île-de-France Webinar, Dec. 2025 (Edouard Lhomme).
- “Optimizing Early-Phase Clinical Development of Vaccines”, ANRS Workshop on Modelling Tools for Vaccination, March 2025 (Laura Richert).
- “Carbon Footprint of Clinical Research”, I-Reivac Workshop, April 2025 (Laura Richert).
- “Decarbonization of Clinical Research Activities”, CIC Rennes Scientific Day, Dec. 2025 (Laura Richert).
- “STRIVE: An International Clinical Trials Network”, invited talks at ANRS-MIE Scientific Days, EUCLID Seminars, OpenReMIE Kick-Off Meeting, and EU-Response General Assembly, 2025 (Linda Wittkop).
- Invited Paper Session “Novel Statistical Approaches in Biomarker Discovery, Analysis and Disease Screening”, World Statistics Congress, The Hague, Oct. 2025 (Marta Avalos).
- “Detection and Quantification of Cognitive Biases in Healthcare Using AI”, Public Health Seminar, CERPOP Toulouse, Oct. 2025 (Ariel Guerra).
- Keynote “How AI Can Reveal Gender Biases in Healthcare”, AI and Nephrology Conference, Paris, Nov. 2025 (Ariel Guerra).
- “Testing and Perturbation”, *Mathematics of Single-Cell Data-Analysis* research school at C.I.R.M., Marseille, Jul. 2025. (Boris Hejblum)

10.1.5 Leadership within the scientific community

- Vice-President of the ANR CES45 Evaluation Committee, 2025 (Mélanie Prague).
- Bureau Member, ANRS-MIE Coordinated Action on Modelling (Mélanie Prague).
- Co-leader, ANRS-MIE Coordinated Action on Respiratory Viruses since 2022 (Edouard Lhomme).
- Coordinator, Working Group “Greener Clinical Research”, RECAPP/Inserm Network (Laura Richert).
- Member, CNU Subsection 46.04 (Laura Richert).
- Correspondant for the French chapter to the Channel Network region of the International Biometrics Society (Boris Hejblum).

10.1.6 Scientific expertise

- Grant reviewer for the Millennium Science Initiative (Chile) and the Swiss National Science Foundation, 2025 (Mélanie Prague).
- Jury member for PHRC-N and PRFI calls; expert for Horizon Europe and EDCTP3 since 2025 (Edouard Lhomme).

- Member of scientific steering committees and data and safety monitoring boards of international clinical trials (Linda Wittkop).
- Grant reviewer for UKRI (UK) and PHRC–N (France) (Laura Richert).
- Member of the ANRS MIE CSS13 ("Clinical research") evaluation committee (Boris Hejblum)
- Member of the ANR "Thématique Spécifique en IA (TSIA) – Biologie et Santé" evaluation committee (Boris Hejblum)
- Expert reviewer for the "DIM1HEALTH 2.0" call (Boris Hejblum)

10.1.7 Research administration

- Member, Inria CR Hiring Committee, Bordeaux, May 2025 (Mélanie Prague).
- Coordinator, EUCLID Clinical Trial Unit (CIC1401), F–CRIN labelled platform, since 2025 (Edouard Lhomme).
- Coordinator, CIC1401 Public Health Module, since 2025 (Edouard Lhomme).
- Deputy Director, UMR 1219 Bordeaux Population Health (Laura Richert).
- Director, UMS 54 Methods and Applied Research of Trials; Coordinator of the "Infectious Diseases and Inflammation" axis, CIC1401 Public Health Module (Linda Wittkop).
- Member of the chairing committee of the Société Française de Biométrie, the French Chapter of the International Biometric Society (Boris Hejblum)

10.2 Teaching - Supervision - Juries - Educational and pedagogical outreach

10.2.1 Teaching

- **International teaching.**
 - ESPIDAM European Summer Program in Infectious Disease Analysis and Modelling: 2.5–day workshop on within–host modelling of infectious diseases using population approaches (Mélanie Prague).
 - 3-day graduate course on "Bayesian analysis for biomedical research" at the University of Copenhagen (Boris Hejblum).
- **Engineering and Master levels.**
 - Missing Data (ENSAI, M2 level), 14h lectures (Mélanie Prague).
 - Dynamical Models (ISPED Biostatistics, M2 level), 6h lectures (Mélanie Prague).
 - Statistical learning in high-dimension in M2 Numerical sciences & bio-health, École Centrale Nantes
- **Bachelor level.** Advanced Linear Regression and Analysis of Variance (L1 Public Health), lectures and tutorials, distance learning and forum moderation, 2017–ongoing (Mélanie Prague).
- **Programme and course responsibilities** (Edouard Lhomme):
 - Program Director, University Diploma (DU) "Methods in Clinical Research", ISPED.
 - Head, Modelling Teaching Unit (UER), College of Health, University of Bordeaux.
 - Course Director, "Principles of Clinical Trials" (RCL201, eRCL201, RCL203), Master 2 in Epidemiology and Biostatistics, ISPED.
 - Course Director, "Evaluation of Health Innovations", Master 2 in Health Innovations, University of Bordeaux.

- Course Director, Optional Training Unit in Clinical Research, MD–PhD Dual Degree Program, University of Bordeaux.
- Co–Director, Inter–University Diploma “Clinical Research Adapted to the African Context” (University of Conakry & University of Bordeaux), since 2025.
- All permanent members and several PhD students contribute to teaching in the Master of Public Health (M1 Public Health, M2 Biostatistics and/or Epidemiology) and the Digital Public Health graduate program, University of Bordeaux.
- Linda Wittkop coordinates the teaching unit “Public Health and Statistics in Medicine”, first year of Medical School, University of Bordeaux.
- Laura Richert, Linda Wittkop and Edouard Lhomme teach in the medical curriculum (PASS and DFASM1–3), University of Bordeaux.
- Linda Wittkop co–coordinates the Master of Epidemiology, ISPED, University of Bordeaux, since 2024, and coordinates several teaching units within the program.
- Linda Wittkop coordinates the seminar series “Support for Medical Thesis”, Medical School, University of Bordeaux.

10.2.2 Supervision

- **PhD students** (Mélanie Prague):
 - Anne André Ruiz, “Building a Digital Twin for Vaccination”, since 2025 (co–supervision: Véronique Godot, VRI).
 - Adrien Mitard, “Modélisation de la réponse viro-immunologique dans les modèles du SARS-CoV2 : Implication pour l’optimisation thérapeutique et vaccinale”, since 2025 (co–supervision: Jérémie Guedj Inserm IAME Paris).
 - Lisa Crépin, “Methods for Latent Variable Models in Mechanistic Models”, since 2025 (co–supervision: Morgan Craig, University of Montreal).
 - Théo René, “Evaluation of Digital Twins”, since 2025 (co–supervision: Moreno Ursino, Inria Heka).
 - Auriane Gabaut, “Méthodes de régularisation pour l’intégration de données de grande dimension dans les modèles mécanistes : application pour le développement de vaccins”, defended Nov 2025 (co–supervision: Cécile Proust Lima Inserm Bordeaux Population Health Biostat).
- **Master interns** (Mélanie Prague): Lisa Crépin (AgroParisTech, April–September 2025).
- **PhD students** (Edouard Lhomme): Co–supervision (50%) with Linda Wittkop of Daniela Gouna, “Modelling and Analysis of Determinants of Post–Vaccination Immunogenicity and Tolerance”.
- **PhD students** (Laura Richert):
 - Cyrille Koné (co–supervision with Emilie Kaufmann; defended December 2025).
 - Nam–Anh Tran (co–supervision with Shirin Golchi, McGill University; ongoing).
 - Perrine Lunel (co–supervision with Sarah Zohar, Inria Heka; ongoing).
- **Master interns** (Laura Richert): Yanis Barteau (M1 Public Health), Perrine Lunel (M2 Biostatistics), Emilie Mesa (MD Public Health).
- **PhD students** (Linda Wittkop):
 - Emie Delrieu, “Methodological Approaches and Feasibility Assessment of External Data Integration in HIV Clinical Trials”, since October 2025.
 - Daniela Gouna, co–supervision with Edouard Lhomme.

- **Master interns** (Linda Wittkop): Tatiana Gropinauth (M2 PHDS, ISPED/McGill), Emie Delrieu (M2 Epidemiology, ISPED), Lucas Balihaut (M2 SITIS, ISPED).
- **PhD students** (Quentin Clairon): Aurore Li (since October 2024).
- **Master interns** (Quentin Clairon): Lore Lafuente (M2), Marie Pelletier (M1).
- **PhD students** (Marta Avalos): Céline Hosteins (co-supervision with Cristian Meza, since September 2025); Ariel Guerra (co-supervision with Emmanuel Lagarde, since September 2024).
- **Master interns** (Marta Avalos): Céline Hosteins (University of Bordeaux), Diego Kauer (University of Chile), 2025.
- **PhD students** (Boris Hejblum):
 - Arthur Hughes (PhD co-direction 50%): “Approches par groupes de gènes pour le développement de vaccins : association, prédiction et marqueurs de substitution” co-directed with R Thiébaud, from Oct 2023.
 - Kalidou Ba (PhD co-direction 50%): “Reservoir computing for cellular composition prediction from longitudinal transcriptomics data in vaccine trials”, co-directed with Xavier Hinaut (Inria Bordeaux), from Nov 2022.
 - Anness Pal (PhD co-direction 50%): “Modèles de mélange bayésien pour la déconvolution de proportions cellulaires à partir de données transcriptomiques en masse” co-directed with R Thiébaud, from Dec 2023.
 - Sara Fallet Pal (PhD co-direction 50%): “Analyse différentielle par groupes de gènes de données scRNA-seq issues d’échantillons multiples”, co-directed with Pierre Neuvial (Institut Mathématiques de Toulouse, CNRS), from Oct 2024.
 - Alice Simon (PhD co-direction 50%): “MA random forest-based clustering method for the unsupervised analysis of high-dimensional data applied to immunology” co-directed with R Genier, from Oct 2025.
- **Master interns** (Boris Hejblum):
 - Theodora Georgakopoulou (M2)

10.2.3 Juries

- **PhD juries and reports** (Mélanie Prague): reviewer (3), examiner (1), and member of follow-up committees (3) for several PhD theses.
- **Juries** (Laura Richert): several MD and Master juries; one PhD jury (Sorbonne University).
- **PhD juries** (Linda Wittkop): reporter, jury president, and follow-up committee member for PhD theses in epidemiology and biostatistics.
- **PhD juries** (Marta Avalos): member of PhD jury and follow-up committees.
- **PhD juries** (Boris Hejblum): reviewer for 2 PhD thesis, and member of several follow-up committees, all in statistics and in biostatistics.

10.2.4 Educational and pedagogical outreach

Serious games to address cognitive bias and patient flow in Emergency Departments. The 5th ENLIGHT Teaching & Learning Conference - Playfulness for the Future of Higher Education, Oct 2025, Uppsala, Sweden [72]

10.3 Popularization

10.3.1 Specific official responsibilities in science outreach structures

- Marta Avalos is a member of the Administrative Council (Research Division) of the competitiveness cluster ENTER (Digital Excellence in Service of Environmental and Responsible Transitions) and a member of the Labeling Committee.

10.3.2 Productions (articles, videos, podcasts, serious games, ...)

- Mélanie Prague: Interview in the Inria series “Elles font le numérique” (#1).
- Linda Wittkop: Interview/article in *La Gazette du Laboratoire*, “Actualité en Afrique – STRIVE: a global research and clinical trials network on infectious diseases”, December 2025, No. 325.
- Ariel Guerra: Interview for France Culture podcast series “Santé des femmes: comment la médecine répare ses biais” [France Culture](#).
- Diego Kauer: Focus on Diego Kauer: intern at SISTM. Article in Inria NUMIN by M Kazolea [Numin](#)
- Marta Avalos: Discover VALPO, the new team associated with Chile. Article in Inria NUMIN by M Kazolea [Numin](#)

10.3.3 Participation in live events

- Mélanie Prague: Geekfest, May 24, 2025, “Zombies, pandemics and mathematical models: what *The Last of Us* and *The Walking Dead* teach us about disease spread”.

10.3.4 Other relevant science outreach activities

- Mélanie Prague: Inria opening seminar and middle school outreach seminar, “Zombies, pandemics and mathematical models: what popular culture teaches us about disease spread”.
- Marta Avalos, Ariel Guerra, Cédric Gil-Jardiné, Céline Hosteins, Diego Kauer: University of Bordeaux Inclusivity Month, March 2025, “AI and cognitive biases: a double-edged sword”.
- Marta Avalos, Ariel Guerra-Adames, Nadia Elorga-Castagnet: Fête de la Science 2025 (theme “Intelligence(s)”), Cap Sciences, October 11–12, 2025, “AI and misconceptions”.
- Ariel Guerra and Océane Dorémus: Dataquitaine seminar, “Synthetic medical data: state of the art and challenges of automatic generation”.
- Ariel Guerra: Co-organizer of workshops “AI and health: rethinking medicine at the human–machine frontier”, Paris (December 9, 2025) and Bordeaux (December 13, 2025).

11 Scientific production

11.1 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] M. Alexandre, R. Marlin, M. Prague, S. Coleon, N. Kahlaoui, S. Cardinaud, T. Naninck, B. Delache, M. Surenaud, M. Galhaut, N. Dereuddre-Bosquet, M. Cavarelli, P. Maisonnasse, M. Centlivre, C. Lacabaratz, A. Wiedemann, S. Zurawski, G. Zurawski, O. Schwartz, R. W. Sanders, R. Le Grand, Y. Levy and R. Thiébaud. ‘Modelling the response to vaccine in non-human primates to define SARS-CoV-2 mechanistic correlates of protection’. In: *eLife* 11 (8th July 2022). DOI: [10.7554/elife.75427](https://doi.org/10.7554/elife.75427). URL: <https://hal.science/hal-04510052>.

- [3] M. Alexandre, M. Prague, C. Mclean, V. Bockstal, M. Douoguih and R. Thiébaud. ‘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](https://doi.org/10.1038/s41541-023-00767-y). URL: <https://hal.science/hal-04226696>.
- [4] 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. Leysen, M. Douoguih and R. Thiébaud. ‘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](https://doi.org/10.1371/journal.pmed.1003813). URL: <https://hal.inria.fr/hal-03481531>.
- [5] Q. Clairon, C. Pasin, I. Balelli, R. Thiébaud 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>.
- [6] Q. Clairon, M. Prague, D. Planas, T. Bruel, L. Hocqueloux, T. Prazuck, O. Schwartz, R. Thiébaud 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* 19.8 (7th Aug. 2023), e1011282. doi: [10.1371/journal.pcbi.1011282](https://doi.org/10.1371/journal.pcbi.1011282). URL: <https://inria.hal.science/hal-03946556>.
- [7] D. Commenges and H. Jacqmin-Gadda. *Dynamical Biostatistical Models*. Chapman and Hall/CRC, 2015. URL: <https://hal.inria.fr/hal-01580149>.
- [8] I. Ganser, D. L. Buckeridge, J. Heffernan, M. Prague and R. Thiébaud. ‘Estimating the population effectiveness of interventions against COVID-19 in France: A modelling study’. In: *Epidemics* 46 (1st Mar. 2024), p. 100744. doi: [10.1016/j.epidem.2024.100744](https://doi.org/10.1016/j.epidem.2024.100744). URL: <https://hal.science/hal-04505758>.
- [9] A. Hughes, L. Parast, R. Thiébaud and B. P. Hejblum. ‘RISE: Two-Stage Rank-Based Identification of High-Dimensional Surrogate Markers Applied to Vaccinology’. In: *Statistics in Medicine* 44.20-22 (2025), e70241. doi: [10.1002/sim.70241](https://doi.org/10.1002/sim.70241). URL: <https://hal.science/hal-04933460> (cit. on p. 18).
- [10] C. Kone, E. Kaufmann and L. Richert. ‘Constrained Pareto Set Identification with Bandit Feedback’. In: *Proceedings of Machine Learning Research*. ICML 2025 - 42nd International Conference on Machine Learning, Vancouver, Canada, 17th July 2025. URL: <https://hal.science/hal-05125339>.
- [11] Y. Lévy, Y. Lévy, C. Moog, C. Moog, A. Wiedemann, A. Wiedemann, O. Launay, O. Launay, F. Candotti, F. Candotti et al. ‘Safety and immunogenicity of CD40.HIVRI.Env, a dendritic cell-based HIV vaccine, in healthy HIV-uninfected adults: a first-in-human randomized, placebo-controlled, dose-escalation study (ANRS VRI06)’. In: *EClinicalMedicine* 77.8 (Nov. 2024), p. 102845. doi: [10.1016/j.eclim.2024.102845](https://doi.org/10.1016/j.eclim.2024.102845). URL: <https://hal.science/hal-04832417>.
- [12] 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. Vrtovsniak, 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](https://doi.org/10.1038/s41591-021-01435-1). URL: <https://hal.sorbonne-universite.fr/hal-03290330>.
- [13] C. Pasin, F. Dufour, L. Villain, H. Zhang and R. Thiébaud. ‘Controlling IL-7 injections in HIV-infected patients’. In: *Bulletin of Mathematical Biology* (2018).
- [14] A. Pollard, O. Launay, J.-D. Lelievre, C. Lacabaratz, S. Grande, N. Goldstein, C. Robinson, A. Gaddah, V. Bockstal, M. Leysen 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](https://doi.org/10.1016/S1473-3099(20)30476-X). URL: <https://hal.inria.fr/hal-03142752>.

- [15] M. Prague, D. Commenges, J. M. Gran, B. Ledergerber, J. Young, H. Furrer and R. Thiébaud. ‘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>.
- [16] M. Prague and M. Lavielle. ‘SAMBA: a Novel Method for Fast Automatic Model Building in Nonlinear Mixed-Effects Models’. In: *CPT: Pharmacometrics and Systems Pharmacology* 11.2 (2022). DOI: [10.1002/psp4.12742](https://doi.org/10.1002/psp4.12742). URL: <https://hal.inria.fr/hal-03410025>.
- [17] A. Rechten, 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. Thiébaud, 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>.
- [18] S. Valayer, M. Alexandre, M. Prague, A. H. Beavogui, S. Doumbia, M. Kieh, B. Greenwood, B. Leigh, M. Poupelin, C. Schwimmer, S. Sow, I. M. Berry, J. Kuhn, D. Fusco, N. D. Cauwelaert, D. Watson-Jones, R. Thiébaud, Y. Lévy, Y. Yazdanpanah, L. Richert and E. Lhomme. ‘Evaluation of waning of IgG antibody responses after rVSVΔG-ZEBOV-GP and Ad26.ZEBOV, MVA-BN-Filo Ebola virus disease vaccines: a modelling study from the PREVAC randomized trial’. In: *Emerging microbes & infections* (19th Nov. 2024), Online ahead of print. DOI: [10.1080/22221751.2024.2432353](https://doi.org/10.1080/22221751.2024.2432353). URL: <https://inria.hal.science/hal-04817450>.
- [19] L. Villain, D. Commenges, C. Pasin, M. Prague and R. Thiébaud. ‘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.

11.2 Publications of the year

International journals

- [20] F. Ader, É. Aubret, N. Bergaud, M. Bouscambert-Duchamp, S. Circosta, S. Couffin-Cadiergues, F. Danion, C. Delmas, A. Diallo, E. Faure, C. Fougere-Leurent, A. Gaymard, K. Lacombe, S. Le Mestre, C. Locher, A. Merelle, N. Peiffer-Smadja, I. Pellegrin, V. Petrov-sanchez and L. Wittkop. ‘Rapid-sequence clinical research before and during a pandemic: Lessons learned and the way forward’. In: *Infectious Diseases Now* 55.7 (2025), p. 105135. DOI: [10.1016/j.idnow.2025.105135](https://doi.org/10.1016/j.idnow.2025.105135). URL: <https://hal.science/hal-05390594>.
- [21] T. Barre, C. Ramier, L. Wittkop, P. Sogni, D. Zucman, R. Tardieu, P. Carrieri and F. Marcellin. ‘Low CD4 cell count is associated with post-hepatitis C virus cure mortality in people living with HIV (ANRS CO13 HEPAVIH cohort)’. In: *Clinical Infectious Diseases* 81.2 (15th Aug. 2025), e28–e30. DOI: [10.1093/cid/ciaf016](https://doi.org/10.1093/cid/ciaf016). URL: <https://hal.science/hal-04956589>.
- [22] T. Barré, C. Ramier, K. Ory, T. Saidi, P. Sogni, F. Zoulim, M. Bureau, C. Protopopescu, F. Marcellin, P. M. Carrieri, A. C. H. S. Group, L. Wittkop and P. Morlat. ‘Liver Enzyme Elevation After Hepatitis C Virus Cure: Is There a Sex Effect? (ANRS CO13 HEPAVIH Cohort)’. In: *Journal of Viral Hepatitis* 32.3 (1st Mar. 2025), e70007. DOI: [10.1111/jvh.70007](https://doi.org/10.1111/jvh.70007). URL: <https://hal.science/hal-05045695>.
- [23] C. Bellera and L. Richert. ‘19th Francophone Conference on Clinical Epidemiology and 32nd Meeting of the Statisticians from Comprehensive Cancer Centers’. In: *Journal of Epidemiology and Population Health* 73.Suppl 2 (5th Apr. 2025), p. 202991. DOI: [10.1016/j.jeph.2025.202991](https://doi.org/10.1016/j.jeph.2025.202991). URL: <https://hal.science/hal-05158335>.
- [24] J. Ben Farhat, M. Hessamfar, D. Neau, S. Farbos, E. Lazaro, P. Duffau, N. Rouanes, C. Cazanave, T. Pistone, P. Rispal, M.-A. Vandenhende, C. Krzyzanowsky, O. Leleux, L. Wittkop, F. Bonnet and D. Barger. ‘Exposure to COVID-19 Pandemic-Related Stressors and Their Association With Distress, Psychological Growth and Drug Use in People With HIV in Nouvelle Aquitaine, France (ANRS CO3 AQUIVIH-NA Cohort-QuAliV-QuAliCOV Study)’. In: *AIDS and Behavior* (6th Jan. 2025). DOI: [10.1007/s10461-024-04588-5](https://doi.org/10.1007/s10461-024-04588-5). URL: <https://hal.science/hal-04871453>.

- [25] J. Bigot, P. Freulon, B. P. Hejblum and A. Leclaire. ‘On the potential benefits of entropic regularization for smoothing Wasserstein estimators’. In: *Electronic Journal of Statistics* 19.2 (2025), pp. 3867–3894. DOI: [10.1214/25-EJS2430](https://doi.org/10.1214/25-EJS2430). URL: <https://hal.science/hal-03908229>.
- [26] M. Chalouni, D. K. van Santen, J. Berenguer, I. Jarrin, J. M. Miro, M. B. Klein, J. Young, J. Torgersen, C. T. Rentsch, M. J. Gill, R. L. Epstein, B. Linas, R. Zangerle, B. Surial, A. Rauch, G. Touloumi, A. Papadopoulos, L. Wittkop, M. van der Valk, A. Boyd, A. d. Monforte, M. Puoti, R. W. Logan, S. M. Rein, M. A. Hernan and S. Lodi. ‘Time to direct-acting antivirals initiation and liver-related events in people with HIV and Hepatitis C virus’. In: *AIDS. Official journal of the international AIDS Society* (19th Feb. 2025). DOI: [10.1097/qad.0000000000004161](https://doi.org/10.1097/qad.0000000000004161). URL: <https://hal.science/hal-05029812>.
- [27] T. Chouleur, C. Etchegaray, L. Villain, A. Lesur, T. Ferté, M. Rossi, L. Andrique, C. Simoncini, A.-S. Giacobbi, M. Gambaretti, E. Lopci, B. Fernandes, G. Dittmar, R. Bjerkvig, B. Hejblum, R. Thiebaut, O. Saut, L. Bello and A. Bikfalvi. ‘A strategy for multimodal integration of transcriptomics, proteomics, and radiomics data for the prediction of recurrence in patients with IDH-mutant gliomas’. In: *International Journal of Cancer* 157.3 (1st Aug. 2025), pp. 573–587. DOI: [10.1002/ijc.35441](https://doi.org/10.1002/ijc.35441). URL: <https://hal.science/hal-05069366>.
- [28] O. Dorémus, D. Russon, B. Contrand, A. Guerra-Adames, M. Avalos-Fernandez, C. Gil-Jardiné and E. Lagarde. ‘Harnessing Moderate-Sized Language Models for Reliable Patient Data De-identification in Emergency Department Records: An Evaluation of Strategies and Performance’. In: *JMIR AI* 4 (2025), e57828. DOI: [10.2196/57828](https://doi.org/10.2196/57828). URL: <https://inria.hal.science/hal-04489355>.
- [29] T. Ferté, K. Ba, D. Dutartre, P. Legrand, V. Jouhet, R. Thiébaud, X. Hinaut and B. Hejblum. ‘Reservoir Computing in R: a Tutorial for Using reservoirnet to Predict Complex Time-Series’. In: *Computo* (27th June 2025). DOI: [10.57750/arxn-6z34](https://doi.org/10.57750/arxn-6z34). URL: <https://hal.science/hal-05208084> (cit. on p. 18).
- [30] A. Guerra Adames, M. Avalos, O. Dorémus, C. Gil-Jardiné and E. Lagarde. ‘Uncovering Judgment Biases in Emergency Triage: A Public Health Approach Based on Large Language Models’. In: *Proceedings of Machine Learning Research* 259 (2025), pp. 420–439. URL: <https://hal.science/hal-04846735> (cit. on p. 18).
- [31] B. L. Harney, R. Sacks-Davis, D. K. van Santen, A. C. Stewart, G. V. Matthews, J. M. Carson, M. B. Klein, K. Lacombe, L. Wittkop, D. Salmon, O. Leleux, L. Merchadou, M. van der Valk, C. Smit, M. Prins, A. Boyd, J. Berenguer, I. Jarrin, A. Rauch, M. E. Hellard and J. S. Doyle. ‘Unsuccessful Direct Acting Antiviral Hepatitis C Treatment Among People With HIV: Findings From an International Cohort’. In: *Liver International* 45.1 (1st Jan. 2025), pp. 1–13. DOI: [10.1111/liv.16203](https://doi.org/10.1111/liv.16203). URL: <https://hal.science/hal-04956541>.
- [32] A. Hughes, L. Parast, R. Thiébaud and B. P. Hejblum. ‘RISE: Two-Stage Rank-Based Identification of High-Dimensional Surrogate Markers Applied to Vaccinology’. In: *Statistics in Medicine* 44.20-22 (2025), e70241. DOI: [10.1002/sim.70241](https://doi.org/10.1002/sim.70241). URL: <https://hal.science/hal-04933460> (cit. on p. 14).
- [33] S. M. Ingle, A. Trickey, A. Lankina, K. A. McGinnis, A. Justice, M. Cavassini, A. d’Arminio Monforte, A. van Sighem, M. J. Gill, H. M. Crane, N. Obel, I. Jarrin, E. Wallner, J. Guest, M. J. Silverberg, G. Vourli, L. Wittkop, T. R. Sterling, D. D. Satre, G. A. Burkholder, D. Costagliola and J. a. C. Sterne. ‘Harmonization of alcohol use data and mortality across a multi-national HIV cohort collaboration’. In: *Alcohol, Clinical and Experimental Research* 49.2 (1st Feb. 2025), pp. 407–417. DOI: [10.1111/acer.15522](https://doi.org/10.1111/acer.15522). URL: <https://hal.science/hal-04951925>.
- [34] L.-Y. Lin, T. Ferté, M. Chachage, C. Casteano, G. Laumond, S. Schmidt, O. Tahar, R. Carapito, L.-G. Bekker, G. Churchyard, M. Keefer, Z. Moodie, E. Viegas, C. Geldmacher, E. Lhomme and C. Moog. ‘Deciphering HIV vaccine-induced Antibody response according to ethnicity’. In: *AIDS. Official journal of the international AIDS Society* (1st July 2025). DOI: [10.1097/qad.0000000000004196](https://doi.org/10.1097/qad.0000000000004196). URL: <https://hal.science/hal-05039069>.

- [35] T. Loosli, N. Han, A. Hauser, J. Josi, S. M. Ingle, A. van Sighem, L. Wittkop, J. Vehreschild, F. Ceccherini-Silberstein, G. Maartens, M. J. Gill, C. A. Sabin, L. F. Johnson, R. Lessells, H. F. Gunthard, M. Egger and R. D. Kouyos. ‘Predicted dolutegravir resistance in people living with HIV in South Africa during 2020-35: a modelling study’. In: *The Lancet global health* 13.4 (1st Apr. 2025), e698–e706. DOI: [10.1016/s2214-109x\(24\)00553-9](https://doi.org/10.1016/s2214-109x(24)00553-9). URL: <https://hal.science/hal-05086220>.
- [36] L. B. Luong Nguyen, L. Magloire, A. François, D. Billard, F. Priou, J. Arrondeau, C. Linassier, I. Ben Ghezala, M. Gross-Goupil, J. Charles, N. Dohollou, P. Vanhems, C. Cracowski, A. M. Leroi, F. Lainé, F. Galtier, K. Barthelemy, S. Priet, M. Gharib, M. Chalouni, A. Barquin, P. Loubet, X. de Lamballerie, O. Launay, L. Wittkop, J.-Y. Blay and J.-P. Spano. ‘Humoral immune response to Covid-19 vaccination in patients with cancer – Results from the ANRS0001S COV-POPART study’. In: *Vaccine* 63 (Sept. 2025), p. 127633. DOI: [10.1016/j.vaccine.2025.127633](https://doi.org/10.1016/j.vaccine.2025.127633). URL: <https://normandie-univ.hal.science/hal-05288940> (cit. on p. 20).
- [37] G. Margue, J.-C. Bernhard, J. Giai, A. Bouzit, S. Ricard, M. Jaffredo, B. Guillaume, E. Jambon, G. Fiard, P. Bigot, T. Waeckel, L. Surlemont, S. D. Vergie, N. Branger, N. Doumerc, R. Boissier, H. Lang, F. Audenet, J.-B. Beauval, K. Bensalah, A. Descazeaud, S. David-Tchouda, L. Richert, J.-A. Long and J.-L. Descotes. ‘Clinical Trial Protocol for ACCURATE: A CCafU-UroCCR Randomized Trial: Three-dimensional Image-guided Robot-assisted Partial Nephrectomy for Renal Complex Tumor (UroCCR 99)’. In: *European Urology Oncology* (Apr. 2025). DOI: [10.1016/j.euo.2025.03.012](https://doi.org/10.1016/j.euo.2025.03.012). URL: <https://hal.science/hal-05032990>.
- [38] A. Mitard de Girardier, C. Herate, R. Marlin, F. Donati, Y. Rahou, L. Bossevoit, Q. Sconosciuti, M. Cavarelli, N. Dereuddre-Bosquet, F. Relouzat, S. van der Werf, E. Simon-Loriere, M. Prague, J. Guedj and R. Le Grand. ‘Exposure history shapes SARS-CoV-2 viral dynamics in Non-Human Primates and provides insights into correlates of protection against infection and transmission’. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* (2025). DOI: [10.1098/rstb.2024.0343](https://doi.org/10.1098/rstb.2024.0343). URL: <https://inserm.hal.science/inserm-05291797>. In press (cit. on p. 19).
- [39] M. Pattou, S. Masanet, M.-A. Jutand, H. Hoarau, J. Sarrazin, A. Pitout, L. Richert, H. Larribere, F. Rubat Baleuri, M. Jaffredo, S. Ricard, M. Faessel, J. Sabatier, J.-C. Bernhard and G. Margue. ‘Effects of a personalized or generic three-dimensional tumoral kidney model on patient experience and caregiver-patient interactions, before and after partial nephrectomy, a randomized trial (Rein 3D Print Personalize—UroCCR 114)’. In: *PLoS ONE* 20.8 (18th Aug. 2025), e0323515. DOI: [10.1371/journal.pone.0323515](https://doi.org/10.1371/journal.pone.0323515). URL: <https://hal.science/hal-05229557>.
- [40] A. Pitout, G. Margue, J. Sarrazin, L. Richert, H. Larribère, S. Masanet, T. Waeckel, P. Bigot, R. Boissier, B. Parier, S. de Vergie, M. Jaffredo, S. Ricard, H. Hoarau, M.-A. Jutand, M. Faessel, J. Sabatier and J.-C. Bernhard. ‘Effect on preoperative anxiety of a personalized three-dimensional kidney model prior to nephron-sparing surgery for renal tumor: study protocol for a randomized controlled trial (Rein 3D Print-Anxiety – UroCCR 113)’. In: *PLoS ONE* 20.4 (24th Apr. 2025), e0321747. DOI: [10.1371/journal.pone.0321747](https://doi.org/10.1371/journal.pone.0321747). URL: <https://hal.science/hal-05046676>.
- [41] B. Reyné, R. Djidjou-Demassee, M. T. Sofonea and S. Alizon. ‘Mutant emergence timing and population immunisation status impact epidemiological dynamics’. In: *Journal of Theoretical Biology* 608 (2025), p. 112140. DOI: [10.1016/j.jtbi.2025.112140](https://doi.org/10.1016/j.jtbi.2025.112140). URL: <https://hal.science/hal-05069967> (cit. on p. 19).
- [42] B. Reyné, T. Kamiya, R. Djidjou-Demassee, S. Alizon and M. T. Sofonea. ‘Leaky or polarised immunity: Non-Markovian modelling highlights the impact of immune memory assumptions’. In: *PLoS Computational Biology* 21.8 (2025), e1013399. DOI: [10.1371/journal.pcbi.1013399](https://doi.org/10.1371/journal.pcbi.1013399). URL: <https://hal.science/hal-05283311> (cit. on p. 19).
- [43] R. Tatoud, Y. Levy, R. Le Grand, J. Alcami, G. Barbareschi, C. Brander, A. Cara, B. Combadiere, F. Dabis, S. Fidler, T. Hanke, C. Herrera, G. B. Karlsson Hedestam, H. Kuipers, S. McCormack, C. Moog, G. Pantaleo, L. Richert, R. W. Sanders, R. Shattock, H. Streeck, R. Thiébaud, A. Trkola, K. Ueberla, M. J. van Gills, R. Wagner, W. Weissenhorn, Y. Yazdanpanah, G. Scarlatti and J. D. Lelievre. ‘In danger: HIV vaccine research and development in Europe’. In: *PLoS Global Public Health* 5.4 (8th Apr. 2025), e0004364. DOI: [10.1371/journal.pgph.0004364](https://doi.org/10.1371/journal.pgph.0004364). URL: <https://hal.science/hal-05084236>.

- [44] R. Thiébaud, E. Lhomme, H. Hocini, I. Pellegrin, A. Boizard-Moracchini, A. Duvignaud, M. Perpère, M. Huchon, M. Prague, C. Lacabaratz, M. Surenaud, X. Anglaret, D. Malvy, B. P. Hejblum and Y. Levy. ‘Self-collected finger-prick blood for gene expression profiling: Unveiling early immune responses in mild COVID-19’. In: *iScience* 29.2 (31st Dec. 2025), p. 114593. DOI: [10.1016/j.isci.2025.114593](https://doi.org/10.1016/j.isci.2025.114593). URL: <https://inria.hal.science/hal-05490229> (cit. on p. 21).
- [45] J. Young, S. Wang, R. Sacks-Davis, A. Stewart, D. K. van Santen, M. van der Valk, J. S. Doyle, G. Matthews, J. Berenguer, L. Wittkop, K. Lacombe, A. Rauch, M. Stooove, M. Hellard and M. B. Klein. ‘Liver fibrosis regression in people living with HIV after successful treatment for hepatitis C’. In: *Journal of Acquired Immune Deficiency Syndromes - JAIDS* (20th Feb. 2025). DOI: [10.1097/qai.0000000000003646](https://doi.org/10.1097/qai.0000000000003646). URL: <https://hal.science/hal-05030578>.

International peer-reviewed conferences

- [46] M. Avalos, J. Barrera, C. Meza, S. Eyheramendy, R. Enaud, L.-E. Vandenborgh and L. Delhaes. ‘Indoor Microbiome in Asthma: A Differential Abundance Framework Accounting for Zero-Inflation, Compositionality, Spatial Clustering, and Covariate Adjustment’. In: <https://www.bioinformatics.polimi.it/CIBB2025/index.html>. CIBB 2025 - 20th conference on Computational Intelligence methods for Bioinformatics and Biostatistics. Milan, Italy, 2025. URL: <https://hal.science/hal-05377271> (cit. on p. 19).
- [47] J. Barrera, C. Meza, M. Avalos, C. Hosteins, S. Eyheramendy, R. Enaud and L. Delhaes. ‘Mixed effects models for longitudinal compositional data using the SAEM algorithm: Application to identifying respiratory microbiome-based predictive’. In: <https://www.isi-next.org/conferences/wsc2025/>. ISI 2025 - 65th ISI World Statistics Congress. Hague, Netherlands, 2025. URL: <https://hal.science/hal-05402381> (cit. on p. 19).
- [48] O. Doremus, A. Guerra-Adames, M. Avalos-Fernandez, V. Jouhet, C. Gil-Jardiné and E. Lagarde. ‘Synthetic medical data generation: state of the art and application to trauma mechanism classification’. In: *Collections of Short Papers CIBB 2025*. CIBB 2025 - International Conference on Computational Intelligence Methods for Bioinformatics and Biostatistics. Milan, Italy, 2025. URL: <https://inria.hal.science/hal-05334858>.
- [49] C. Hosteins, M. Avalos, D. Kauer, L. Delhaes and R. Enaud. ‘Predicting Response to Cystic Fibrosis Modulators from Compositional Microbiome Data: A Comparison of Adapted Machine Learning Approaches’. In: <https://www.bioinformatics.polimi.it/CIBB2025/index.html>. CIBB 2025 - 20th conference on Computational Intelligence methods for Bioinformatics and Biostatistics. Milan, Italy, 2025. URL: <https://hal.science/hal-05401017> (cit. on p. 19).
- [50] C. Hosteins, M. Avalos, D. Kauer, C. Meza, L. Delhaes and R. Enaud. ‘Machine Learning for Compositional Microbiome Data to Predict a Clinical Outcome: Are Interpretable Methods Up to the Task?’ In: *Proceedings of the 15th International Conference on Pattern Recognition Systems (ICPRS), 2025*. ICPRS 2025 - IEEE 15th International Conference on Pattern Recognition Systems. Viña del Mar, Chile, Dec. 2025. DOI: [10.1109/ICPRS66293.2025.11302842](https://doi.org/10.1109/ICPRS66293.2025.11302842). URL: <https://hal.science/hal-05377246> (cit. on p. 19).
- [51] C. Kone, E. Kaufmann and L. Richert. ‘Bandit Pareto Set Identification in a Multi-Output Linear Model’. In: *Proceedings of Machine Learning Research*. AISTATS 2025 - 28th International Conference on Artificial Intelligence and Statistics. Phuket, Thailand, 5th May 2025. URL: <https://hal.science/hal-05125311> (cit. on p. 20).
- [52] C. Kone, E. Kaufmann and L. Richert. ‘Constrained Pareto Set Identification with Bandit Feedback’. In: *Proceedings of Machine Learning Research*. ICML 2025 - 42nd International Conference on Machine Learning. Vancouver, Canada, 17th July 2025. URL: <https://hal.science/hal-05125339> (cit. on p. 20).
- [53] L. Rodriguez-Idiazabal, M. Avalos, J. M. Quintana-Lopez and I. Barrio. ‘Can mortality prediction be enhanced integrating patient clinical and medical imaging data?’ In: *Libro de actas de la XX Conferencia Española de Bioestadística (CEB 2025)*. CEB 2025 - XX Conferencia Española de Bioestadística. Elche, Spain, 16th Sept. 2025. URL: <https://inria.hal.science/hal-05372878>.

- [54] L. Rodriguez-Idiazabal, M. Avalos, D. Russon, I. Barrio and J. M. Quintana-Lopez. ‘Balancing Predictive Performance and Explainability for Integrating Clinical, Biological, and Radiological Data in the Basque Country COVID-19 Cohort’. In: ISI 2025 - 65th ISI World Statistics Congress. Hague, Netherlands, 2025. URL: <https://inria.hal.science/hal-05385401>.
- [55] M. Ursino, S. Boulet, C. Collignon, F. Francis-Oliviero, E. Lhomme, R. Porcher, F. Saillour, G. Varoquaux, V. Vercamer, R. Thiébaud and S. Zohar. ‘Innovative clinical trial approach for evaluating digital medical devices under European HTA fast-Track frameworks’. In: ISCB 2025 - 46th Annual Conference of the International Society for Clinical Biostatistics. Basel, Switzerland, 24th Aug. 2025. URL: <https://hal.science/hal-05364910>.

National peer-reviewed Conferences

- [56] A. Guerra-Adames, O. Dorémus, M. Avalos, V. Jouhet, E. Lagarde and C. Gil-Jardiné. ‘Génération de données de santé synthétiques : état de l’art en vue de la classification des mécanismes des traumatismes’. In: RJCIA 2025 - Rencontres des Jeunes Chercheurs en Intelligence Artificielle - Plate-Forme Intelligence Artificielle (PFIA). Dijon, France, 30th June 2025. URL: <https://inria.hal.science/hal-05334934>.

Conferences without proceedings

- [57] S. Boulet, V. Thévenet, J. Abécassis, C. Collignon, B. Giraudeau, E. Lhomme, F. Petit, R. Porcher, L. Richert, F. Saillour, M. Ursino, G. Varoquaux, V. Vercamer, S. Zohar and R. Thiébaud. ‘Narrative review on the clinical evaluation of AI-based digital medical devices from a methodological perspective’. In: PEPR-SanteNum 2025 - Journées Annuelles du PEPR Santé Numérique. Lille, France, 14th Oct. 2025. URL: <https://hal.science/hal-05364958>.

Doctoral dissertations and habilitation theses

- [58] T. Ferté. ‘Short-Term COVID-19 forecasting using Bordeaux University Hospital data warehouse and Reservoir Computing’. Université de Bordeaux, 24th Nov. 2025. URL: <https://theses.hal.science/tel-05407547> (cit. on p. 18).
- [59] A. Gabaut. ‘Regularization methods for high-dimensional data integration into mechanistic models : application for vaccine development’. Université de Bordeaux, 25th Nov. 2025. URL: <https://theses.hal.science/tel-05404943> (cit. on p. 20).

Reports & preprints

- [60] Q. Clairon, J. Fricks and M. Prague. *Curve registration for non-linear mixed effect ordinary differential equation models*. 11th Dec. 2025. URL: <https://inria.hal.science/hal-05412063>.
- [61] T. Ferté, V. Thevenet, X. Hinaut, P. Legrand, D. Dutartre, R. Griffier, V. Jouhet, B. P. Hejblum and R. Thiébaud. *Epidemic Forecasting: Lessons Learned from the SARS-CoV-2 Pandemic to Balance Accuracy, Feasibility, and Impact*. 6th Nov. 2025. DOI: [10.1101/2025.11.03.25339385](https://doi.org/10.1101/2025.11.03.25339385). URL: <https://hal.science/hal-05392781> (cit. on p. 18).
- [62] I. Ganser, J. Paireau, D. L. Buckeridge, S. Cauchemez, R. Thiébaud and M. Prague. *Comparative evaluation of methodologies for estimating the effectiveness of non-pharmaceutical interventions in the context of COVID-19: a simulation study*. 28th Apr. 2025. DOI: [10.1101/2024.10.14.24314896](https://doi.org/10.1101/2024.10.14.24314896). URL: <https://hal.science/hal-05073613> (cit. on pp. 14, 19).
- [63] A. Guerra-Adames, M. Avalos-Fernandez, O. Dorémus, L. A. Celi, C. Gil-Jardiné and E. Lagarde. *A Counterfactual LLM Framework for Detecting Human Biases: A Case Study of Sex/Gender in Emergency Triage*. 21st Nov. 2025. URL: <https://hal.science/hal-05385521> (cit. on p. 18).
- [64] Z. Li, M. Prague, R. Thiébaud and Q. Clairon. *Variational autoencoder for inference of nonlinear mixed effect models based on ordinary differential equations*. 22nd Jan. 2026. DOI: [10.13140/RG.2.2.23271.71841](https://doi.org/10.13140/RG.2.2.23271.71841). URL: <https://hal.science/hal-05472212>.

- [65] A. Pal, A. Mimoun, R. Thiébaud and B. P. Hejblum. *Adaptive Dirichlet Process mixture model with unknown concentration parameter and variance: Scaling high dimensional clustering via collapsed variational inference*. 28th Jan. 2026. URL: <https://inria.hal.science/hal-05490235>.
- [66] D. Zhou, H. Tong, L. Wang, S. Liu, X. Xiong, Z. Gan, R. Griffier, B. Hejblum, Y.-C. Liu, C. Hong, C.-L. Bonzel, T. Cai, K. Pan, Y.-L. Ho, L. Costa, V. Panickan, J. M. Gaziano, K. Mandl, V. Jouhet, R. Thiebaut, Z. Xia, K. Cho, K. Liao and T. Cai. *Representation Learning to Advance Multi-institutional Studies with Electronic Health Record Data*. 2025. URL: <https://hal.science/hal-05222545>.

Other scientific publications

- [67] M. Avalos, C. Hosteins, D. Kauer, C. Renault, R. Enaud and L. Delhaes. ‘Data Transformations in Machine Learning Approaches for Studying Microbiota as a Biomarker of Non-Response Risk to CFTR Modulators’. In: 46th Annual Conference of the International Society for Clinical Biostatistics. Basel, Switzerland, 24th Aug. 2025. URL: <https://inria.hal.science/hal-05377184> (cit. on p. 19).
- [68] S. Boulet, V. Thévenet, J. Abécassis, C. Collignon, B. Giraudeau, E. Lhomme, F. Petit, R. Porcher, L. Richert, F. Saillour, M. Ursino, G. Varoquaux, V. Vercamer, S. Zohar and R. Thiébaud. ‘Narrative review on the clinical evaluation of AI-based digital medical devices from a methodological perspective’. In: PEPR-SanteNum 2025 - Journées Annuelles du PEPR Santé Numérique. Lille, France, 14th Oct. 2025. URL: <https://hal.science/hal-05347091>.
- [69] S. Boulet, V. Thévenet, J. Abécassis, C. Collignon, B. Giraudeau, E. Lhomme, F. Petit, R. Porcher, L. Richert, F. Saillour, M. Ursino, G. Varoquaux, V. Vercamer, S. Zohar and R. Thiébaud. ‘Narrative review on the clinical evaluation of AI-based digital medical devices from a methodological perspective’. In: SbN 2025 - 10th Statistics & Biopharmacy Conference. Paris, France, 8th Oct. 2025. URL: <https://hal.science/hal-05364943>.
- [70] D. Kauer, C. Hosteins, M. Avalos, L. Delhaes and R. Enaud. ‘Predicting Non-Response to Cystic Fibrosis Modulators from Microbiota Using Standard and Custom Machine Learning’. In: JOBIM 2025 - Journées Ouvertes Biologie Informatique Mathématiques. Bordeaux, France, 8th July 2025. URL: <https://inria.hal.science/hal-05337115> (cit. on p. 19).
- [71] M. Ursino, S. Boulet, C. Collignon, F. Francis-Oliviero, E. Lhomme, R. Porcher, F. Saillour, G. Varoquaux, R. Thiébaud and S. Zohar. ‘Innovative Clinical Trial Approach for Evaluating Digital Medical Devices under European Fast-Track Regulatory Frameworks’. In: EPICLIN 2025 - Conférence francophone d’EPIdémologie CLINique. Bordeaux, France, 14th May 2025. URL: <https://hal.science/hal-05085050>.

Scientific popularization

- [72] M. Avalos, N. Elorga-Castagnet, D. Russon, F. Larrue and C. Gil-Jardiné. ‘Serious Games to Address Cognitive Bias and Patient Flow in Emergency Departments: A Multidisciplinary Perspective on Reducing Health Inequities’. In: The 5th ENLIGHT Teaching & Learning Conference 2025 - Playfulness for the Future of Higher Education. Uppsala, Sweden, 8th Oct. 2025. URL: <https://inria.hal.science/hal-05334672> (cit. on p. 34).
- [73] A. Guerra Adames, O. Dorémus, M. Avalos, E. Lagarde, V. Jouhet and C. Gil-Jardiné. ‘Données médicales synthétiques : État de l’art et enjeux de la génération automatique de données tabulaires et textuelles’. In: Dataquitaine 2025. Talence, France, 20th Mar. 2025. URL: <https://inria.hal.science/hal-05334979>.

11.3 Cited publications

- [74] 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> (cit. on p. 10).

- [75] D. Agniel, B. Hejblum, R. Thiébaud and L. Parast. ‘Doubly-robust evaluation of high-dimensional surrogate markers’. In: *Biostatistics* 24.4 (2023), pp. 985–999. doi: [10.1093/biostatistics/kxac020](https://doi.org/10.1093/biostatistics/kxac020). URL: <https://inria.hal.science/hal-03100499> (cit. on p. 10).
- [76] M. Alexandre, R. Marlin, M. Prague, S. Coleon, N. Kahlaoui, S. Cardinaud, T. Naninck, B. Delache, M. Surenaud, M. Galhaut et al. ‘Modelling the response to vaccine in non-human primates to define SARS-CoV-2 mechanistic correlates of protection’. In: *elife* 11 (2022), e75427 (cit. on p. 11).
- [77] M. Alexandre, M. Prague, C. McLean, V. Bockstal, M. Douoguih, R. Thiébaud, E. 1 and E. 2. Consortia. ‘Prediction of long-term humoral response induced by the two-dose heterologous Ad26.ZEBOV, MVA-BN-Filo vaccine against Ebola’. In: *NPJ vaccines* 8.1 (2023), p. 174 (cit. on p. 11).
- [78] 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](https://doi.org/10.1007/s00357-014-9147-x) (cit. on p. 9).
- [79] L. Capitaine, J. Bigot, R. Thiébaud and R. Genuer. ‘Fréchet random forests for metric space valued regression with non Euclidean predictors’. In: *Journal of Machine Learning Research* 25.355 (2024), pp. 1–41 (cit. on p. 9).
- [80] Q. Clairon, M. Prague, D. Planas, T. Bruel, L. Hocqueloux, T. Prazuck, O. Schwartz, R. Thiébaud and J. Guedj. ‘Modeling the kinetics of the neutralizing antibody response against SARS-CoV-2 variants after several administrations of Bnt162b2’. In: *PLoS Computational Biology* 19.8 (2023), e1011282 (cit. on p. 11).
- [81] C. Colas, B. Hejblum, S. Rouillon, R. Thiébaud, P.-Y. Oudeyer, C. Moulin-Frier and M. Prague. ‘Epidemioptim: A toolbox for the optimization of control policies in epidemiological models’. In: *Journal of Artificial Intelligence Research* 71 (2021), pp. 479–519 (cit. on p. 11).
- [82] A. Collin, B. P. Hejblum, C. Vignals, L. Lehot, R. Thiébaud, 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) (cit. on p. 11).
- [83] T. Ferté, S. Cossin, T. Schaefferbeke, T. Barnetche, V. Jouhet and B. Hejblum. ‘Automatic phenotyping of electronic health record: PheVis algorithm’. In: *Journal of Biomedical Informatics* 117 (2021), p. 103746. doi: [10.1016/j.jbi.2021.103746](https://doi.org/10.1016/j.jbi.2021.103746). URL: <https://hal.inria.fr/hal-03100435> (cit. on p. 10).
- [84] T. Ferté, V. Jouhet, R. Griffier, B. Hejblum and R. Thiébaud. ‘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](https://doi.org/10.1093/jamiaopen/ooac086). URL: <https://hal.science/hal-03908260> (cit. on p. 10).
- [85] I. Ganser, D. L. Buckeridge, J. Heffernan, M. Prague and R. Thiébaud. ‘Estimating the population effectiveness of interventions against COVID-19 in France: A modelling study’. In: *Epidemics* 46 (Mar. 2024), p. 100744. doi: [10.1016/j.epidem.2024.100744](https://doi.org/10.1016/j.epidem.2024.100744). URL: <https://hal.science/hal-04505758> (cit. on p. 11).
- [86] 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](https://doi.org/10.1038/sdata.2018.298). URL: <https://hal.inria.fr/hal-01942279> (cit. on p. 10).
- [87] 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 (cit. on p. 7).
- [88] A. Jarne, D. Commenges, L. Villain, M. Prague, Y. Lévy and R. Thiébaud. ‘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 (cit. on p. 11).
- [89] 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](https://doi.org/10.1186/s13059-021-02290-6) (cit. on p. 9).

- [90] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola and L. K. Saul. ‘An introduction to variational methods for graphical models’. In: *Machine learning* 37 (1999), pp. 183–233 (cit. on p. 11).
- [91] 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 (cit. on p. 11).
- [92] 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> (cit. on p. 8).
- [93] C. J. Martensen, N. Korsbo, V. Ivaturi and S. Sager. ‘Data-Driven Discovery of Feedback Mechanisms in Acute Myeloid Leukaemia: Alternatives to classical models using Deep Nonlinear Mixed Effect modeling and Symbolic Regression’. In: *bioRxiv* (2024), pp. 2024–06 (cit. on p. 11).
- [94] 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. Mugenyi, 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](https://doi.org/10.1016/j.cell.2021.05.023) (cit. on p. 10).
- [95] J. T. Ormerod and M. P. Wand. ‘Gaussian variational approximate inference for generalized linear mixed models’. In: *Journal of Computational and Graphical Statistics* 21.1 (2012), pp. 2–17 (cit. on p. 11).
- [96] C. Pasin, I. Balelli, T. Van Effelterre, V. Bockstal, L. Solforosi, M. Prague, M. Douoguih and R. Thiébaud. ‘Dynamics of the humoral immune response to a prime-boost Ebola vaccine: quantification and sources of variation’. In: *Journal of virology* 93.18 (2019), pp. 10–1128 (cit. on p. 11).
- [97] 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> (cit. on p. 8).
- [98] G. A. Poland, I. G. Ovsyannikova and R. B. Kennedy. ‘Personalized vaccinology: a review’. In: *Vaccine* 36.36 (2018), pp. 5350–5357 (cit. on p. 8).
- [99] M. Prague, D. Commenges, J. Guedj, J. Drylewicz and R. Thiébaud. ‘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 (cit. on p. 11).
- [100] M. Prague, D. Commenges, J. M. Gran, B. Ledergerber, J. Young, H. Furrer and R. Thiébaud. ‘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](https://doi.org/10.1111/biom.12564). URL: <https://hal.inria.fr/hal-01406614> (cit. on p. 10).
- [101] B. Pulendran. ‘Learning immunology from the yellow fever vaccine: innate immunity to systems vaccinology’. In: *Nature Reviews Immunology* 9.10 (2009), pp. 741–7 (cit. on p. 7).
- [102] Z. Qian, W. Zame, L. Fleuren, P. Elbers and M. van der Schaar. ‘Integrating expert ODEs into neural ODEs: pharmacology and disease progression’. In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 11364–11383 (cit. on p. 11).
- [103] L. SAS. ‘Monolix version 2021R1’. In: *Antony, France* (2022), <http://lixoft.com/products/monolix/> (cit. on p. 11).
- [104] C. Schubert. ‘Systems immunology: complexity captured’. In: *Nature* 473.7345 (2011), pp. 113–4 (cit. on p. 7).
- [105] 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](https://doi.org/10.1093/jamia/ocy056). URL: <https://hal.inria.fr/hal-01887930> (cit. on p. 10).

- [106] 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> (cit. on p. 10).
- [107] 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](https://doi.org/10.1016/j.vaccine.2016.09.007) (cit. on p. 9).
- [108] R. Thiébaud, 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 (cit. on p. 10).
- [109] 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](https://doi.org/10.1007/978-3-030-61616-8_40) (cit. on p. 9).
- [110] 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 (cit. on p. 11).
- [111] H. Wu. ‘Statistical methods for HIV dynamic studies in AIDS clinical trials’. In: *Statistical Methods in Medical Research* 14.2 (2005), pp. 171–192 (cit. on p. 11).
- [112] 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](https://doi.org/10.1093/jamia/ocab187). URL: <https://hal.inria.fr/hal-03478039> (cit. on p. 10).
- [113] B. J. Zou, M. E. Levine, D. P. Zaharieva, R. Johari and E. B. Fox. ‘Hybrid Square Neural ODE Causal Modeling’. In: *arXiv preprint arXiv:2402.17233* (2024) (cit. on p. 11).