

# Application BASTRI

## Fiches Equipes

### SISTM (SR0678GR)

Statistics In System biology and Translational Medicine  
SISTM (SR0575YR) □ SISTM

**Statut:** Décision signée

**Responsable :** Melanie Prague

**Mots-clés de "A - Thèmes de recherche en Sciences du numérique - 2024" :** A3.1.1. Modélisation, représentation, A3.1.10. Données hétérogènes, A3.1.11. Données structurées, A3.3.2. Fouille de données, A3.3.3. Analyse de données massives, A3.4.1. Apprentissage supervisé, A3.4.2. Apprentissage non supervisé, A3.4.3. Apprentissage par renforcement, A3.4.4. Optimisation pour l'apprentissage, A3.4.5. Méthodes bayésiennes, A5.2. Visualisation de données, A6.1.1. Modélisation continue (EDP, EDO), A6.2.4. Méthodes statistiques, A6.3.1. Problèmes inverses, A6.3.4. Réduction de modèles, A6.4.2. Contrôle stochastique, A9.2. Apprentissage, A9.6. Aide à la décision

**Mots-clés de "B - Autres sciences et domaines d'application - 2024" :** B1.1. Biologie, B1.1.5. Immunologie, B1.1.7. Biologie computationnelle, B1.1.10. Biologie des systèmes et biologie synthétique, B2.2.4. Maladies infectieuses, Virologie, B2.2.5. Maladies du système immunitaire, B2.3. Epidémiologie, B2.4.1. Pharmacologie et toxicologie, B2.4.2. Résistance aux médicaments, B9.5.6. Science des données, B9.8. Recherche reproductible

**Domaine :** Santé, biologie et planète numériques  
**Thème :** Modélisation et commande pour le vivant

**Période :** 01/01/2015 -> 31/12/2027

**Dates d'évaluation :** 12/10/2017, 15/05/2022

**Etablissement(s) de rattachement :** U. DE BORDEAUX, INSERM  
**Laboratoire(s) partenaire(s) :** <sans UMR>

**CRI :** Centre Inria de l'université de Bordeaux  
**Localisation :** Institut de Santé Publique d'Epidémiologie et de Développement  
**Code structure Inria :** 091058-1

**Numéro RNSR :** 201321095C  
**N° de structure Inria:** SR0678GR

### Présentation

## Objectifs

Le défi est d'analyser ces BIG DATA pour répondre à des questions cliniques et biologiques à l'aide de méthodes statistiques appropriées. Pour comprendre les données d'une cellule à l'état clinique des individus dans toutes les circonstances, y compris dans les essais cliniques, de nouveaux outils sont nécessaires pour traduire les informations obtenues à partir des systèmes complexes. Par extension, cela a conduit au domaine de la « médecine des systèmes » « biologie des systèmes », qui a lieu naturellement dans le cadre de la médecine translationnelle reliant la recherche clinique et biologique.

L'analyse statistique de ces données est confronté à plusieurs problèmes :

- Il y a plus de paramètres ( $p$ ) à estimer que d'individus ( $n$ )
- Les types / nature des données sont différentes
- La relation entre les variables est souvent complexe (par exemple non linéaire) et peut changer au fil du temps. Pour faire face à ces problèmes, nous développons des approches spécifiques souvent liées à l'immunologie.

Les méthodes sont principalement basées sur la modélisation mécanistique utilisant des systèmes d'équations différentielles ou sur les méthodes d'apprentissage statistique. Le paradigme de notre approche est d'inclure autant d'informations que disponible pour répondre à une question donnée. Ces informations proviennent des données disponibles, mais aussi de l'information biologique initiale définissant la structure du modèle ou de restreindre l'espace des valeurs de paramètres. Nous développons et appliquons nos méthodes principalement pour des applications appartenant à la recherche clinique en particulier l'immunologie du VIH. Par exemple, plusieurs projets sont consacrés

### Contact

- **Responsable :** Melanie Prague
- **Tél :** 05.57.57.45.21
- **Secrétariat Tél :**

### En savoir plus

- Site de l'équipe
- Site sur [inria.fr](http://inria.fr)
- Site du responsable
- Derniers Rapports d'Activité : 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023

### Documents sur la structure

- [Intranet](#)
- [Privés](#)

### Décisions

- **10607** (05/01/2015) : création
- **12209** (15/05/2017) : modification
- **13241** (10/12/2018) : prolongation
- **15196** (14/09/2022) : prolongation
- **16553** (09/11/2023) : prolongation
- **17197** (09/07/2024) : cessation du responsable
- **17198** (09/07/2024) : nomination responsable

### Localisation

- **Adresse postale :** ISPED - Université de Bordeaux 146 rue Léo Saignat 33076 Bordeaux cedex France
- **Coordonnées GPS :** 44.4929, 0.3624

à la modélisation de la réponse aux traitements antirétroviraux, les interventions immunitaires ou vaccin chez les patients infectés par le VIH.

Les applications sont réalisées en collaboration avec the Vaccine Research Institute (VRI), d'autres équipes du centre ainsi que l'Unité de Soutien Méthodologique à la Recherche Clinique et Epidémiologique (USMR) du CHU de Bordeaux.

## Axes de recherche

### Axe 1 : Modélisation mécanistique

Lorsqu'on étudie la dynamique d'un marqueur donné, comme la concentration du VIH dans le sang (charge virale du VIH), on peut par exemple utiliser des modèles descriptifs résumant la dynamique au fil du temps en terme (termes) de pentes des trajectoires. Ces pentes peuvent être comparées entre les groupes de traitement ou selon les caractéristiques des patients. Une autre façon d'analyser ces données est de définir un modèle mathématique basé sur la connaissance biologique qui explique la dynamique de l'infection. Avoir un bon modèle mécanistique dans un contexte biomédical ouvre des portes à diverses applications, au-delà d'une bonne compréhension des données.

### Axe 2 : Données de grande dimension

Lorsqu'on analyse des données de grande dimension comme des données omiques telles que les données génomiques (ADN), transcriptomiques (ARN) ou protéomiques (protéines), mais aussi d'autres types de données, comme celles qui résultent de la combinaison de grandes bases de données d'observation (par exemple en pharmacoépidémiologie ou épidémiologie environnementale), le défi méthodologique vient du fait que les ensembles de données contiennent généralement de nombreuses variables, beaucoup plus que d'observations. En outre, les méthodes classiques, tels que les modèles linéaires sont inefficaces et la plupart du temps, même inapplicable. Par conséquent, plus que les capacités de stockage et de calcul, le défi est l'analyse complète de ces ensembles de données partant de plusieurs niveaux de voies moléculaires jusqu'à la réponse clinique d'une population de patients, qui se fait à l'aide de méthodes spécifiques et d'une collaboration très étroite avec les fournisseurs de données (immunologistes, virologues, cliniciens...) L'objectif est de sélectionner les informations pertinentes ou de les résumer en vue d'une meilleure compréhension ou dans un but de prédiction.

## Relations industrielles et internationales

### Collaborations internationales

- **Vaccine Research Institute** (Hôpital Henri Mondor, Creteil), Labex laboratoire d'excellence, est une extension de l'ANRS (L'Agence nationale de recherches sur le sida et les hépatites virales) qui a pour but d'accélérer le développement de vaccin contre le VIH et hépatites C.
- Le department d'immunologie à Institute of Child Health, University College London
- Centre de statistiques pour VIH/Sida Recherche & Prévention (SHARP),
- Le département vaccins et maladies infectieuses de Fred Hutchinson, Centre de Recherche du Cancer
- Le département des systemes and bio informatique à Albert Einstein College of Medicine, New York
- Ecole des Mathématiques et Physiques à University of Queensland

### Des chercheurs et étudiants internationaux visitent régulièrement l'équipe SISTM, avec qui ils travaillent en étroite collaboration dont :

- Robin CALLARD (UCL, London, UK),
- **Raphael GOTTARDO** (Vaccine and Infectious Disease Division of the Fred Hutchinson Cancer Research Center, Seattle ,USA),
- Jessica GRONSBELL (Harvard T H Chan, School of Public Health, Boston, USA),
- **Samson KOELLE** (University of Washington, Seattle, USA)
- **Cristian Meza**, Associate Professor of the Universidad de Valparaíso (Chili), member of the research center CIMFAV
- **David Conesa**, Associate Professor of the Universidad de Valencia (Espagne), member of the research group GEEITEMA
- **Sam Doerken**, PhD student of the University of Freiburg (Allemagne), member of the Institute for Medical Biometry and Statistics

### Projets internationaux

- Le projet **EBOVAC2** a été fondé à l'initiative de IMI2 Ebola+ programme Ebovac 2, en réponse à la forte épidémie du virus Ebola afin d'étudier l'efficacité de la réponse immunitaire déclenchée par une stratégie vaccinale préventive et prometteuse « prime-boost » contre le virus

Ebola. Cet essai clinique de phase 2 est sous la responsabilité scientifique de Rodolphe Thiébaud de l'unité Inserm 1219.

- European HIV Vaccine Alliance (EHVA) a été fondé par EU Horizon 2020 afin de favoriser le développement de vaccin efficace via une plateforme européenne dédiée à la découverte et l'évaluation de nouveaux candidats vaccins prophylactiques et thérapeutiques. L'équipe SISTM est impliquée dans le projet en tant que leader de groupes de travail (WP10) sur l'intégration des datas.