RESEARCH CENTRE

Inria Paris Center

IN PARTNERSHIP WITH:

INSERM, Université Paris Cité

2022

ACTIVITY REPORT

Project-Team

HEKA

Health data- and model- driven approaches for Knowledge Acquisition

IN COLLABORATION WITH: CENTRE DE RECHERCHE DES CORDELIERS

DOMAIN

Digital Health, Biology and Earth

THEME

Computational Neuroscience and Medicine

Inria

Contents

Project-Team HEKA

Creation of the Project-Team: 2021 November 01

Keywords

Computer sciences and digital sciences

A3.3. – Data and knowledge analysis

A3.4. – Machine learning and statistics

A6.1. – Methods in mathematical modeling

A6.2. – Scientific computing, Numerical Analysis & Optimization

A9.1. – Knowledge

A9.2. – Machine learning

A9.4. – Natural language processing

A9.6. – Decision support

Other research topics and application domains

B2.2. – Physiology and diseases

B2.3. – Epidemiology

B2.6. – Biological and medical imaging

1 Team members, visitors, external collaborators

Research Scientists

- Sarah Zohar [Team leader, INSERM, Senior Researcher, HDR]

- Adrien Coulet [INRIA, Researcher, HDR]

- Jean Feydy [Inria, Researcher]

- Nicolas Garcelon [Fondation Imagine]

- Agathe Guilloux [Inria, Senior Researcher, from Sep 2022, HDR]

- Moreno Ursino [INSERM, Researcher]

Faculty Members

- Stéphanie Allassonniere [UNIV PARIS, Professor, HDR]

- François Angoulvant [UNIV PARIS, Professor, HDR]

- David Drummond [UNIV PARIS, Associate Professor]

- Anne-Sophie Jannot [UNIV PARIS, Associate Professor, HDR]

- Sandrine Katsahian [UNIV PARIS, Professor, HDR]

- Antoine Neuraz [UNIV PARIS, Associate Professor]

- Bastien Rance [UNIV PARIS, Associate Professor, HDR]

- Brigitte Sabatier [AP/HP, Hospital Staff, HDR]

Post-Doctoral Fellows

- Nadim Ballout [Inria, from Dec 2022]

- Sarah Berdot [AP/HP]

- Sandrine Boulet [INSERM]

- Germain Perrin [AP/HP, from Mar 2022]

- Rosy Tsopra [UNIV PARIS]

PhD Students

- Safa Alsaidi [Inria, from Sep 2022]

- Nesrine Bannour [UNIV PARIS SACLAY]

- Linus Bleistein [ENS PARIS]

- Tom Boeken [UNIV PARIS]

- Clément Chadebec [UNIV PARIS]

- Pierre Clavier [LIX]

- Aziliz Cottin [DASSAULT SYSTEMES]

- Benjamin Duputel [EXYSTAT, CIFRE]

- Thibaut Fabacher [UNIV STRASBOURG]

- Fleur Gaudfernau [UNIV PARIS]

- Enora Laas-Faron [INSTITUT CURIE, from Mar 2022]

- Ivan Lerner [INSERM]

- Juliette Murris [PIERRE FABRE, CIFRE]

- Lillian Muyama [Inria]

- Sophie Quennelle [UNIV PARIS]

- Alice Rogier [INSERM]

- Pierre Sabatier [APHP]

- Agathe Senellart [UNIV PARIS - CITE, from Oct 2022]

- Agathe Senellart [ECOLE POLY PALAISEAU, from Apr 2022 until Aug 2022]

- Stylianos Tzedakis [UNIV PARIS]

- Alexis Van Straaten [AP/HP, from Apr 2022]

- Louis Vincent [IMPLICITY, CIFRE]

Technical Staff

- Deycy Camila Arias Villamil [Inria, Engineer]

- Armelle Arnoux [AP/HP, Engineer, from Jun 2022]

- Olivier Birot [Inria, Engineer]

- Xiaoyi Wang [INSERM, Engineer]

Administrative Assistant

- Christelle Guiziou [INRIA]

2   Overall objectives

2.1   Context

Clinicians routinely have to take decisions upon the diagnosis and treatment of complex patients for which clinical guidelines do not provide clear recommendations or do not exist. This is particularly the case for very heterogeneous diseases (e.g., rare diseases or cancer in which clinical manifestations or response to treatment differ frequently from one individual to another) or when patients are seen for a new emerging disease for which no recommendation has been established yet, as it was the case in March, 2020 for the management of COVID-19 infections. Similar situations, leading to delicate decisions, happened in the past but, unfortunately, this experience is hardly taken into account or rationalized for clinicians. Indeed, data related to these past experiences are captured, but until now, these data are not accessible to clinicians and not transformed into high level evidences. This level of evidence is currently only reached by highly controlled analyses (such as controlled clinical trials), for which patients might differ strongly from those treated in routine care, or might have never been seen before in the case of new conditions. Besides, hospital information systems are used at every step of patient care, collecting continuously longitudinal data, both unstructured and structured, including clinical reports, drug prescriptions, physiology, laboratory results, imaging and omics data. These data may even be

enriched with medical wearable devices or large-scale claims data such as those of the French Health Data Hub that informs on the global clinical course of patients. Likewise, progresses on artificial intelligence approaches, such as supervised machine learning, have found applications in healthcare enabling for instance clinical decision support systems, patient prioritization, drug repurposing or monitoring drug safety. However, many particularities of health data (e.g., their sensitivity, noisiness, incompleteness, heterogeneity and small volume when one very specific feature or several time points are required), and particularities of health data analysis (e.g., the risk of confoundedness, the need for explanations and fairness) make it challenging to develop tools that are both reliable and usable within hospital work- and patient- flows. For instance, precision medicine requires stratifying smaller and smaller groups of patients, which may seem contradictory in regards with the general strategy of deep learning that requires large amounts of data to be efficient. Another challenging task is the development of tools that enable gaining knowledge from data agilely, i.e., to update knowledge gain continuously (without compromising on reliability). In summary, methodological developments are required at each step of the health data chain, including: (1) data access, (2) data transformation e.g., via representation learning, (3) data analysis, predictive modelling and knowledge discovery with data- and model- driven approaches, and (4) agile, fast and reliable access to data, implementation of these approaches through applications such as decision support systems, medical devices, next generation clinical trials for the assessment of medical knowledge.

2.2 General aim

The main objective of HeKA is to develop methodologies, tools and their applications in clinics towards a learning health system, i.e., a health system that leverages clinical data collected to extract agilely and reliably novel medical knowledge that, in turn, continuously improves healthcare. Indeed, the availability of EHRs (Electronic Health Records), cohorts and other linked data such as the national Health Data Hub, offers the opportunity to develop models for stratification and prediction with the potential of improving the precision and the personalization of treatments, and thus the quality of healthcare.

3 Research program

The HeKA project-team is following 3 interdependent axes, that are (1) knowledge extraction from clinical data, (2) stochastic and data-driven predictive modelling of health data, and (3) data-driven and designs for next generation clinical trials. Theses axes can be view and interpreted form a patient care point of view as (1) from patient data to patient representations, (2) from patient data to prediction and decision, and (3) from models to improve patient-related knowledge, respectively. All these axes participate in the development of a learning health system. Axes 1 and 2 can be related to observational studies (either retrospective or prospective) and Axis 3 is related to interventional studies. As a remind, in observational studies the investigator is not acting upon study participants, but only observing relationships between factors and outcomes, while in interventional studies (i.e., clinical trials) the investigator intercedes as part of the study design.

3.1 Axis 1 - Knowledge extraction from clinical data

The development of clinical decision support and statistical predictive models has been historically made by manually selecting and tuning sets of predictive variables. This is a task-dependent and time-consuming operation that neglects most of the available data. Real-word data, such as EHRs or cohorts, offer an access to many variables, even those not initially thought of as predictive. For instance, EHRs consist of structured data, such as demographics, diagnosis, procedures, biological laboratory results, and medication exposures, which can be associated with unstructured data, such as clinical notes, discharge summaries, pathology and imaging reports. In addition, this core EHR data may be complemented with others, including images, omics data, patient-reported outcomes, or conversation transcriptions. However, the use of EHR data for any precision medicine application represents an initial and significant information extraction challenge because of their heterogeneity, incompleteness, and dynamic nature. The aim of this axis is to develop methods and tools for leveraging patients' data in their wide variety and

complexity. This encompasses the extraction and transformation of raw data into engineered, featured and learned representations of good quality that will enable or facilitate the development of further clinical decision support and knowledge discovery approaches, as those presented in Axes 2 and 3.

3.1.1 Methods

Methods developed in Axis 1 can be associated with three type of tasks: (i) deep phenotyping, (ii) patient representation and (iii) reasoning with clinical knowledge. (i) Deep phenotyping consists in defining algorithms that enable to identify patients with a particular and potentially complex profile within large healthcare databases. It encompasses the development of natural language processing tools capable of extracting complex features and their context out of clinical texts; it also includes the ability to consider simultaneously structured and unstructured data of these databases to identify relevant patients. To this aim our methods rely on expert rules, distant-supervision and deep learning language representations. (ii) Regarding patient representation we focus on two distinct kind of representations. The first one is an explicit representation of patients in the form of knowledge graphs, using Semantic Web standards and tools. The second one is a representation of patients within a latent space, using representation learning methods largely inspired from results obtained by deep learning models to learn language representations. (iii) Tasks concerned with reasoning on clinical knowledge are mutliple. It encompasses methods to measure patients' similarity between elaborated patients representations (either in the form of knowledge graph or embeddings), hybrid approaches for analogical reasoning and logical and statistical inference.

3.2 Axis 2 - Stochastic and data-driven predictive modelling of health trajectories

The recent availability of high dimensional health data enables the emergence of data-driven models for description, analyses with further possibility to guide clinical decisions. In this high dimensional setting, machine learning-based prediction tasks have been demonstrated to be efficient, although they may not be the best option in every setting. We are interested in these borders between settings, where deep learning approaches fail, but alternatives succeed and reciprocally. Particularly, we are considering borders found in temporal modelling and small-sample settings. Health data provided by EHRs have several specificities among which: (i) patient care trajectories are high dimensional, and (ii) are censored, i.e., data are observed until a certain timepoint. Current models do not succeed in simultaneously tackling these two previous concerns. For example, patient trajectories include comorbidities pathways where one disease may impact the others. Embedding patient trajectories in prognostic models remains a challenge, especially, when low-sample-size high-dimension setting. Machine learning in such setting should be adapted and compared with efficient statistical learning models.

3.2.1 Methods

During this year, we focused on adapting machine learning based methods on the following challenges: signal detection using spatio-temporal data to be able to build health geographical vigilance system, medical image classification in low sample setting, pharmacovigilance signal detection using large database. We proposed a method to model and monitor population distributions over space and time, in order to build an alert system for spatio-temporal data changes based on a new version of the Expectation-Maximization (EM) algorithm to better estimate the number of clusters and their parameters at the same time. We validate this approach on a real data set of positive diagnosed patients to coronavirus disease 2019. We show that our pipeline correctly models evolving real data and detects epidemic changes. We proposed a new method to perform data augmentation in a reliable way in the High Dimensional Low Sample Size setting using a geometry-based variational autoencoder. We validated this approach on a medical imaging classification task where a small number of 3D brain magnetic resonance images are considered and augmented using the proposed framework. We adapted a very classical pharmacoepidemiological method to high dimensional setting, the Weigthed Cumulative Exposure statistical model, which makes it possible to model the temporal relationship between the prescription of a drug and a side effect, by implementing it using Graphics Processing Unit (GPU) programming. We analysed several real life datasets using such implementation and show that it was possible using this adaptation to apply this method to the National Health Insurance Database (SNDS).

3.3   Axis 3: Data-driven and designs for next generation clinical trials

New model-based fundamental researches could help in screening patients and predicting response prior to clinical trials, thanks to biomarkers or EHRs based predictive modelling of patients' response.Digital Medical Devices, are health technologies falling into the definition of medical devices intending to "Prevention, diagnosis, monitoring, treatment or alleviation of disease" (Regulation EU 2017/745) and are mostly based on machine learning algorithms. Among them, learning software as medical devices technologies (SaMD) now have the potential to be translated to patients' care. As such, they should be evaluated in clinical trials. However, by contrast to drugs for which the chemical formula does not change over time, the performances of a predictive model are constantly enriched by new observational data. The challenge we face now is to be able at any time to ensure the safety and efficacy of SaMD and other advances, for patient care. The objective of this axis is twofold : how clinical trials can help machine learning? And inversely, how can machine learning and other information sources help clinical trials? Accordingly, the first objective is to propose clinical trial methods and designs adapted to continuous learning and adapting tools (DMDs). The second objective, is to produce innovative clinical trial designs acquiring all possible patient-related knowledge: disease and translational models, EHRs, clinical trial data and synthetic patients.

3.3.1   Methods

This year we focused more on the second objective. We developed (i) dose-finding methods that include pharmacokinetics and pharmacodynamics modelling to better choose the appropriate dose-regimen in oncology ; (II) methods based on logistic regression and random effects meta-analysis to syntetise information from multiple early phase trials with several doses ; (iii)methods which allow for analsing disruped clinical trial by COVID pandemic ; (iv) deep learning models that allow to analyse multistate EHR ...

4   Application domains

4.1   Cancer

4.1.1   EraPerMed PeCAN (Personalized medicine in CANcer) (EU funding 2020-2023)

Breast cancer is the most common cancer in women worldwide. The most aggressive type is triple negative breast cancer (TNBC), characterized by a very poor prognosis. New targeted and personalized therapies are urgently required. In this context, the European project PeCAN aims at developing and assessing new "AI algorithms" able to predict treatment response and personalize therapy of TNBC patients. They will propose new approaches for personalization of drug therapy based on a deep molecular characterization of individual tumors and patients and the establishment and use of digital medicine approaches to model effects and side effects of all therapy options. In these projects, our team is in charge of developing multi-scale models that integrate more clinical data in the models (as biomarkers, patients' characteristics, chemo and targeted molecule treatments, biology, clinical outcomes, etc.). We are designing a framework for assessing the clinical performance and safety of the AI algorithms itself, by using real-world data collected in clinical data warehouses and biobanks. These projects are associated with Axes 1-3.

4.1.2   Optimizing patients' treatment trajectory in metastatic colon cancer (Ongoing)

In metastatic colorectal cancer, biomarkers that explain certain toxicities have recently been identified. For instance, the UGT1A1 gene is strongly associated with bilirubin levels, for which a variant is linked to an increased risk of irinotecan toxicity. This finding suggests that patient subgroups defined by these variables should be treated differently. While standard protocols for dose adjustment in chemotherapy exist, these recommendations do not account for possible associations between other covariates and comorbidities. Thus, in clinical practice, physicians adapt doses and schedules of successive cycles of treatment by accounting for patient's characteristics and treatment history, including doses and outcomes. As a result, each patient's dose regimen over multiple cycles often differs from what is recommended in standard protocols. We used Bayesian inference accounting for elicited clinical relevance weights from

physicians to reflect their medical experience. This allowed us to model and optimize the treatment trajectory over cycles based on observed toxicities. Further, we will optimize patient treatment trajectory to avoid cancer progression and increase overall survival (in Axis 2). For this project we got an INCA funding and we will collaborate with the HEGP oncology department.

4.1.3 Optimizing patients' treatment trajectory in metastatic liver cancer (Ongoing)

Metastatic liver cancer can be treated through mini invasive surgery, typically via interventional radiology. The aims of this project are (1) to cluster the possible responders and non-responders to these particular therapies (using imaging), (2) to evaluate the quantity of required treatment (either burn or freeze an area including all the tumoral matter) and follow-up. We will optimize patient treatment trajectory in interventional radiology by better selecting, guiding and following up patients. We will adapt and develop models for patient care trajectories, as proposed in Axis 2, and using features extracted from Axis 1. This project will be done in collaboration with Pr Marc Sapoval (HEGP) and its team in the interventional radiology department. This project will also be part of a bilateral collaboration with Singapore where we will co-develop methods to use both imaging and ctDNA data for the selection and prediction of patients' outcomes.

4.1.4 Cancer follow-up and modelling through liquid biopsies in HPV-induced cancers (Ongoing)

Human Papillomavirus (HPV) is a small virus responsible for benign genital warts that are susceptible to transform into in-situ to invasive cancers. The cancerization process often involves integration of HPV DNA in the host DNA, which is identified by circulating HPV-DNA found in the blood stream (Veyer et al. 2019). The objective of this project is to provide data structures enabling longitudinal omics follow-up of HPV-induced cancers. We will propose, on one hand, methods providing longitudinal omics representation in clinical data warehouses using graph structures and content-addressable storage and, on the other hand, modeling patients' trajectories using longitudinal omics data for clustering and predicting patients' sub-group outcomes. This project is in relation with Axis 1 for the development of temporal representation of omics data in graphs and in relation with Axis 2 for the modelling of patient care trajectories. This project will be in collaboration with HEGP virology and oncology departments and Curie Institute.

4.1.5 PersoProCaRisk (Ongoing)

Risk stratification for prostate cancer (PCa) included prostate specific antigen (PSA), (if indicated) needle biopsy, histopathological examination and consistent tumor grading. In this project, we aim at (1) stratify patient's treatment (active surveillance vs. surgical intervention vs. radiation therapy), (2) identify patients with high-risk tumors and (3) improve disease progression monitoring, employing a) deep proteomics of tissue specimens, b) Next Generation Sequencing of seminal vesicle plasma and matched tissue; and c) imaging techniques (PSMA PET/CT, and MRI) as a radiomics approach. These high dimensional data will be integrated into statistical/learning models for onset of events and time-dependent covariates. For this axis-2 project, we collaborate with the University Medical Center Freiburg's Department of Radiation Oncology, the University Medical Center Freiburg's Institute of Molecular Medicine and the University of Toronto's Clinical Biochemistry, Laboratory Medicine and Pathobiology department.

4.1.6 Twinonco (Ongoing)

Almost the totality studies in oncology, use medical imaging to assess (i-)RECIST criteria leading to a standard measure of disease progression or overall survival. However, images reading relies on radiologist expertise and manual and repetitive processes. In this project, we will in position to test new services for automatic images analyses on dedicated platform. More specifically, our team will have to provide the methodology and the results of tests of several sofware prototypes helping lesions (both target and non-target) readings. All these development aims at providing a virtual twin of the patient further accompanying medical or therapeutic decisions, inside the axis-2 framework. This project is done in collaboration with Prof. Laure Fournier and the Radiology-HEGP team, and Dassault Systèmes.

4.1.7 Dysplastic and invasive squamous lesions of the upper aerodigestive pathways-PREINV (Ongoing)

This project is dedicated to the development of an algorithm for automatic imaging analysis, with Deep Learning, helping the automatic classification of lesions and the prediction of dysplasia evolution. This algorithm should have the same level of performance thatn the pathologist. Dysplastic images will be studied in order to search for elements of evolution into invasive cancer and pronostic prediction. Taking into account the lack of inter-observers agreement and lack of precision with actual criteria and human analyses, will be key in this process. In this project, not so close but still related to axis-2, we are data controller for clinical data, and will closely with Prof. C. Badoual and Y Bellahsen-Harrar (Pathological Anatomy and Cytology Department-HEGP).

4.1.8 RHU – Optimisation and imProved Efficacy of targeted RAdioNuclide therapy in Digestive cancers by Imagomics - Operandi (Ongoing)

The OPERANDI project (Optimisation and imProved Efficacy of targeted RAdioNuclide therapy in Digestive cancers by Imagomics) aims to address unmet clinical needs in the current management of advanced stage digestive tumours treated with targeted radionuclide therapy (TRT) by exploring new opportunities provided by imaging-based artificial intelligence (AI) and data augmentation, simultaneous PET-MRI imaging, and novel approaches to increase TRT efficacy (genomic profiling, radiopotentiators, and new radionuclides).

4.2 Rare diseases and pediatrics

4.2.1 Diagnostic dead-ends and rare diseases (Ongoing)

The identification of rare disease variations or of their oligogenic forms is hampered by the lack of a precise phenotypic description of patients. The ability of describing precisely these diseases, from a phenotypic point of view, would enable to search for relevant associated genes. The aim of this project is to avoid diagnostic dead-ends by proposing methods that relate the diagnostic to a relevant group of genes. This motivates the development of algorithms to predict the most likely diagnoses. We will use the information that is available in the "rare diseases fact sheets" set up by the National Rare Diseases Database (BNDMR). By cross-referencing the information extracted from EHR texts and from the rare disease sheets and modeling patient's care trajectories, we will propose to build diagnostic algorithms that will allow identifying most probable diagnoses for each patient in a diagnostic dead-end by prioritizing the variations to be searched for. This project is related to Axis 1 by developing extraction and phenotyping tools and to Axis 2 by modelling patient care trajectories. This project will be carried out in collaboration with the BNDMR, the AnDDI-Rares rare diseases Healthcare Network, Necker hospital and Imagine Institute.

4.2.2 CIL LICO- Ciliopathies: group of disorders associated with genetic mutation leading to rare and severe genetic diseases - Projet RHU - Recherche Hospitalo-Universitaire- 20118-22

Ciliopathies are a large group of rare and severe genetic diseases caused by ciliary dysfunction associated with clinical and genetic heterogeneity, as well as a lack of knowledge on patients' natural history (i.e., the evolution of the disease). The aim of this RHU-3 project, funded through the Investissements d'avenir program, is to develop innovative, diagnostic, prognostic and tailored therapeutic approaches for patients suffering from ciliopathies to prevent them from developing renal failure. Following this aim, we will develop a mechanistic stratification of ciliopathies, in order to regroup suspected and already diagnosed ciliopathies in a treatment-orientated classification. One goal is to engineer a ready-to-use bio-kit for assessing both diagnosis and prognosis of developing renal alteration, in ciliopathy patients. Finally, personalized therapeutic approaches will be proposed for patients using predictive approaches. We will use methods developed in Axis 1 as well as methods proposed in Axis 2 to model patient trajectory. This project is in collaboration with Imagine Institute, AP-HP, the Medetia company and Ecole Polytechnique.

4.2.3 Optimizing and personalizing the treatment of children's asthma via machine learning and at-home data entry (under evaluation)

The objective of the CHIASMA X project is to integrate personalized approaches into the optimization and personalization of childhood asthma. First, it will implement a remote, continuous and extensive assessment of children's asthma, through the integration of various data sources including home-based monitoring (mobile app, smart inhalers, home spirometers) and environmental data. Second, it will allow timely delivery of effective personalized treatments through the development of two data-driven models (in relation with Axis 2): one indicating when a medical review is required, the other suggesting the most appropriate treatment to each child. It will also be part of the next-generation clinical trials (in relation with Axis 3) proposing data adherence algorithms. This project will be in collaboration with the university hospital Necker.

4.2.4 BNDMR- Banque Nationale des Maladies Rares

The French National Registry for Rare Diseases (BNDMR) is a national tool for epidemiology and public health purposes in the field of rare diseases. In line with the objectives defined by the 2nd and 3rd French National Plan for Rare Diseases, the BNDMR team develops a secure national information system which gathers anonymized clinical data of patients affected by rare diseases in its BNDMR data warehouse. As medical head of the BNDMR, AS Jannot has several research projects strongly connected with HeKA team including CDE.AI and Dromos project. CDE.ai aims to create a set of natural language processing algorithms that will allow the semi-automatic completion of the rare disease minimal data set that is currently completed manually for all patients followed up in the rare disease expert centres. In this project, we will use methods developped in Axe 1 (collaboration with N Garcelon). The DROMOS project is a project that uses the National Data Bank for Rare Diseases by linking it to health insurance data. This matching will allow the description of the care of rare disease patients at the national level for rare diseases, including the characteristic care of the most frequent rare diseases. We will use methods developped in the from of Axis 2 to model these longitudinal data.

4.2.5 Genomic variability transversal program (2018-2023)

The objective of the INSERM Genomic Variability cross-cutting programme is to understand the role played by genes and their variants on the development of pathologies. This programme, which is based on the longitudinal follow-up of cohorts of individuals and their phenotyping, aims to promote the development of new methods for the analysis of longitudinal data. In this program, we develop in the frame of axis 2 longitudinal approaches to define patients trajectories from truncated data such as those in France National Health Database (SNDS).

4.3 Other

4.3.1 Antibiotic resistance – FAIR project EU Horizon (on going)

The aim of the FAIR project is to evaluate Flagelin aerosol therapy for stimulation of immunity as an alternative treatment against pneumonia with multidrug resistant bacteria. In this project, we are developing a full model using pharmacometrics expertise as well as statistical designs for extrapolation purpose and the design of dose-finding study in healthy volunteers. As written above, in this project, S Zoharvco-lead along with C Kloft (Freie Universitaet Berlin) the WP entitled "Development of a translational modelling and simulation platform for flagellin PK/PD". The aim of the WP is to propose an optimal design for the first-in-man clinical trials, maximizing knowledge gained from in vitro experimentation, expert knowledge and pre-clinical experiments along the way. By incorporating mechanistic approaches earlier in the development process along with a continuous learning modelling under Bayesian inference, we hope to increase the probability of success of the translation process to the clinical setting and thus, optimizing the statistical design and sample size. This project is in relation with axis 3.

4.3.2 Endocrine/Primary Hypertension - HT-ADVANCE (funded)

Conducted by several hypertension Centers of Excellence, and following the successful ENSAT-HT project, the HT-ADVANCE project targets a deep change in personalized management of arterial hypertension (HT) by using multi-omics (MOMICS) stratification biomarkers for existing drugs prescription. The objective of HT-ADVANCE is to validate two multicomponent stratification biomarkers in patients with HT in order to (1) identify patients with endocrine hypertension (EHT), and (2) predict response to treatment in patients with primary hypertension (PHT). Three clinical trials are planned and apply machine learning techniques to integrate the genetic, genomic and metabolomic features in order to generate accurate diagnostic and therapeutic response predictions for clinicians. In this project, related to axis-2, we will provide (i) methological support for all the 3 trials, (ii) validity- and (iii) usefulness-assessements of the omics approach to deliver improved health outcomes in HT through MOMICS (WP5). The whole project is funded by the HORIZON-HLTH-2022-TOOL-11 call and relies on an international consortium, including 14 research teams representing 7 EU/non-EU states, coordinated by the Paris Cardiovascular Research Center (INSERM).

4.3.3 Heart Failure with preserved Ejection Fraction – PACIFIC (Ongoing)

The prevalence of Heart Failure with preserved Ejection Fraction (HFpEF) is increasing especially on the grounds of the aging of the population. The pathology presents a wide variety of cardiac and extra-cardiac abnormalities causing problems in terms of diagnosis and therapeutic management. As no global and deep analysis of all anomalies has ever been provided, it becomes essential to understand underlying phenotypes heterogeneities, improve patients clustering. In addition, this project tries also to ensure the simultaneous and continuous measures of key follow-up parameters via a connected tool (a T-shirt) and link the abnormalities observed via the connected tool with clinical and paraclinical anormalities observed in standard of care. In this project, not so close but still related to axis-2, we are data controller for clinical data, and will closely with Coordinating Investigator (Prof. JS Hulot at the CIC1418-PT) and Scientific Responsible (Dr H. Tachouaft, Resp. of Clinical Operation Dpt at Sanofi).

4.3.4 Virtual reality (Ongoing)

Several projects led at the HEGP are currently ongoing to evaluate the analgesia provided by the use of Virtual Reality in different care settings (extracorporeal lithotripsy, after colorectal cancer surgery, and fiberoptic bronchoscopy in critical care). In these projects, not so close but still related to axis-2, we will provide methodological approach and use statistical methods to conclude on the clinical questions, by working closely with all Coordinating Investigators (Prof. D Clausse, G Manceau and A Rastello)

5 Social and environmental responsibility

5.1 Impact of research results

Our methods and designs are applied in collaboration with medical research team members at HEGP and Necker hospitals (among others). During the last two years, and in consequence of the COVID-19 sanitary pandemic, we developed and deployed tools for knowledge extraction from clinical text to the central clinical data warehouse of the AP-HP (Entrepôt de Données de Santé, AP-HP) . Through others, these tools enable the recognition of named entities, their context (e.g., negation, personal or family history, hypothesis). They have been widely reused for the development of clinical studies, which led to 13 international publication of clinical epidemiology related to COVID-19. It also serve as a basis for our implication in the 4CE international consortium (https://www.covidclinical.net/), which led to 15 publications. These tools have been refactored and their use is facilitate within the medkit library.

6 Highlights of the year

- CRCN Inria – recruitment of Jean Feydy (Dec 2021)

- DR Inria (Detachement) - Agathe Guilloux (Sept 2022)

- CRCN Inserm – recruitment of Moreno Ursino (Oct 2022)

- Medical Head of BNDMR (Banque Nationale des Maladies Rares) – nomination of Anne-Sophie Jannot (Sept 2022)

- AT2TA ANR grant project (2022-25), AT2TA project aims at developping the use of representation learning and deep machine learning for analogical reasoning. We are leading a workpackage that will investigate clinical applications of such novel approach for analogical reasoning.

- ADHERENCE BOAS grant project (2022-24) ADHERENCE aims at developping, evaluating and sharing algorithms for the evaluation of therapeutical adherence in the context of three chronic diseases. This project is funded by the Health Data Hub.

7 New software and platforms

7.1 New software

7.1.1 medkit

Name: a toolkit for a learning health system

Keywords: Learning health system, Biomedical data, Decision support, Python, Information extraction, Natural language processing, Audio signal processing, Machine learning

Functional Description: This library aims at (1) facilitating the manipulation of healthcare data of various modalities (e.g., structured, text, audio data) for the extraction of relevant features and (2) developing supervized models from these various modalities for decision support in healthcare.

Release Contributions: This release 0.3.1 fixes a pip install issue found for release 0.3.0.

URL: https://github.com/TeamHeka/medkit

Contact: Adrien Coulet

Participants: Deycy Camila Arias Villamil, Olivier Birot, Kim-Tam Huynh, Antoine Neuraz, Ivan Lerner, Bastien Rance, Adrien Coulet

7.1.2 Pythae

Keywords: Generative Models, Benchmarking, Reproducibility

Functional Description: This library implements some of the most common (Variational) Autoencoder models under a unified implementation. In particular, it provides the possibility to perform benchmark experiments and comparisons by training the models with the same autoencoding neural network architecture. The feature make your own autoencoder allows you to train any of these models with your own data and own Encoder and Decoder neural networks. It integrates experiment monitoring tools such wandb, mlflow or comet-ml and allows model sharing and loading from the HuggingFace Hub in a few lines of code.

URL: https://github.com/clementchadebec/benchmark_VAE

Contact: Clément Chadebec

7.1.3 Pyraug

Keywords: Generative Models, Data augmentation

Functional Description: This library provides a way to perform Data Augmentation using Variational Autoencoders in a reliable way even in challenging contexts such as high dimensional and low sample size data.

URL: https://github.com/clementchadebec/pyraug

Contact: Clément Chadebec

8 New results

The team have generated many results in the last year, here are five illustrations for each axis.

8.1 Axis 1

Rogier, A., Coulet, A., and Rance, B. (2022). Using an Ontological Representation of Chemotherapy Toxicities for Guiding Information Extraction and Integration from EHRs. Studies in health technology and informatics, 290, 91-95.

The detection of toxicities and their severity from EHRs is of importance for many downstream applications. However toxicity information is dispersed in various sources in the EHRs, making its extraction challenging. We developped OntoTox, an ontology designed to represent chemotherapy toxicities, its attributes and provenance. OntoTox enable the integration of toxicities and grading information extracted from three heterogeneous sources: EHR questionnaires, semistructured tables, and free-text. We instantiated 53,510, 2,366 and 54,420 toxicities from questionnaires, tables and free-text respectively, and compared the complementarity and redundancy of the three sources.

> **Participants:** Rogier A ., Coulet A ., Rance B ..

Lerner, I., Serret-Larmande, A., Rance, B., Garcelon, N., Burgun, A., Chouchana, L. and Neuraz, A., 2022. Mining Electronic Health Records for Drugs Associated With 28-day Mortality in COVID-19: Pharmacopoeia-wide Association Study (PharmWAS). JMIR medical informatics, 10(3), p.e35190.

We developed a pharmacopeia-wide association study (PharmWAS) pipeline inspired from the PheWAS methodology, which systematically screens for associations between the whole pharmacopeia and a clinical phenotype. First, a fully data-driven procedure based on adaptive least absolute shrinkage and selection operator (LASSO) determined drug-specific adjustment sets. Second, we computed several measures of association, including robust methods based on propensity scores (PSs) to control indication bias. We applied this method in a multicenter retrospective cohort study using electronic medical records from 16 university hospitals of the Greater Paris area. We investigated the association between drug prescription within 48 hours from admission and 28-day mortality for COVID-19. 4 drugs were associated with increased in-hospital mortality. Among these, diazepam and tramadol could worsen COVID-19, what needs to be further assessed.

> **Participants:** Lerner, I ., Garcelon, N ., Rance B ., Burgun, A ., Neuraz, A ..

Chen, X., Faviez, C., Vincent, M., Briseño-Roa, L., Faour, H., Annereau, J.P., Lyonnet, S., Zaidan, M., Saunier, S., Garcelon, N. and Burgun, A., 2022. Patient-Patient Similarity-Based Screening of a Clinical Data Warehouse to Support Ciliopathy Diagnosis. Frontiers in Pharmacology, 13.

A timely diagnosis is a key challenge for many rare diseases. As an expanding group of rare and severe monogenic disorders with a broad spectrum of clinical manifestations, ciliopathies, notably renal ciliopathies, suffer from important underdiagnosis issues. We developed an approach for screening

large-scale clinical data warehouses and detecting patients with similar clinical manifestations to those from diagnosed ciliopathy patients. A ranking model based on the best-subtype-average similarity was proposed to address the phenotypic overlapping and heterogeneity of ciliopathies. Our results showed that using less than one-tenth of learning sources, our language and center specific embedding provided comparable or better performances than other existing medical concept embeddings. Our approach offer the opportunity to identify candidate patients who could go through genetic testing for ciliopathy.

Participants: Faviez, C ., Garcelon, N ., Burgun, A ..

Faviez, C., Vincent, M., Garcelon, N., Michot, C., Baujat, G., Cormier-Daire, V., Saunier, S., Chen, X. and Burgun, A., 2022. Enriching UMLS-Based Phenotyping of Rare Diseases Using Deep-Learning: Evaluation on Jeune Syndrome. In Challenges of Trustable AI and Added-Value on Health (pp. 844-848). IOS Press.
The phenotype extraction from EHR narrative reports can be performed by using either dictionary-based or data-driven methods. We developed a hybrid pipeline using deep learning to enrich the UMLS Metathesaurus for automatic detection of phenotypes from EHRs. The pipeline was evaluated on a French database of patients with a rare disease characterized by skeletal abnormalities, Jeune syndrome. The results showed a 2.5-fold improvement regarding the number of detected skeletal abnormalities compared to the baseline extraction using the standard release of UMLS. Our method help enrich the coverage of the UMLS and improve phenotyping, especially for languages other than English.

Participants: Faviez, C ., Garcelon, N ., Chen, X ., Burgun, A ..

Monnin, P., Raïssi, C., Napoli, A., and Coulet, A. (2022). Discovering alignment relations with Graph Convolutional Networks: A biomedical case study. Semantic Web, 13(1), 379-398.
We focused here on the identification, within an aggregated knowledge graph, of nodes of various sources, that are equivalent, more specific, or weakly related. We proposed to match nodes within a knowledge graph by (i) learning node embeddings with Graph Convolutional Networks such that similar nodes have low distances in the embedding space, and (ii) clustering nodes based on their embeddings, in order to suggest alignment relations. We conducted experiments on the real world application of aligning knowledge in the field of pharmacogenomics, which motivated our study. We observed that distances in the embedding space are coherent with the "strength" of these different relations (e.g., smaller distances for equivalences), letting us considering clustering and distances in the embedding space as a means to suggest alignment relations.

Participants: Coulet, A ..

8.2 Axis 2

We adapted and developed innovative methods based both on machine and deep learning for longitudinal health data in high dimensional setting in different contexts including cluster detection using geographic data, image classification, cut-point detection of prognostic factor, drug side effect detection.

Pruilh, S., Jannot, A. S., and Allassonnière, S. (2022). Spatio-temporal mixture process estimation to detect dynamical changes in population. Artificial intelligence in medicine, 126, 102258.
We proposed a method to model and monitor population distributions over space and time, in order to build an alert system for spatio-temporal data changes based on a new version of the Expectation-Maximization (EM) algorithm to better estimate the number of clusters and their parameters at the same time. We validate this approach on a real data set of positive diagnosed patients to coronavirus disease 2019. We show that our pipeline correctly models evolving real data and detects epidemic changes.

Participants: Pruilh, S ., Jannot, A. S ., Allassonnière, S ..

Chadebec, C., Thibeau-Sutre, E., Burgos, N., and Allassonnière, S. (2022). Data augmentation in high dimensional low sample size setting using a geometry-based variational autoencoder. IEEE Transactions on Pattern Analysis and Machine Intelligence. We proposed a new method to perform data augmentation in a reliable way in the High Dimensional Low Sample Size setting using a geometry-based variational autoencoder. We validated this approach on a medical imaging classification task where a small number of 3D brain magnetic resonance images are considered and augmented using the proposed framework.

Participants: Chadebec, C ., Allassonnière, S ..

Sabatier, P., Feydy, J., and Jannot, A. S. (2022). Accelerating High-Dimensional Temporal Modelling Using Graphics Processing Units for Pharmacovigilance Signal Detection on Real-Life Data. In Challenges of Trustable AI and Added-Value on Health(pp. 83-87). IOS Press. We adapted a pharmacoepidemiological method to high dimensional setting, the Weigthed Cumulative Exposure statistical model, which makes it possible to model the temporal relationship between the prescription of a drug and a side effect, by implementing it using Graphics Processing Unit (GPU) programming.

Participants: Sabatier, P ., Feydy, J ., Jannot, A. S ..

Rives-Lange C, Poghosyan T, Phan A, et ... Jannot AS. Risk-Benefit Balance Associated With Obstetric, Neonatal, and Child Outcomes After Metabolic and Bariatric Surgery . JAMA Surg. 2022;e225450. doi:10.1001/jamasurg.2022.5450 We applied a self-control approach combined with a falsification analysis to analyse safety of bariatric surgery regarding pregnancies and their resulting children.To take into account potential confounding factors including age, we considered both pregnancies before and after bariatric surgeries but also consecutive pregnancies before bariatric surgeries. Mutiple comparisons without any a priori hypothesis allowed us to pinpoint potential increase of respiratory illnesses in the newborn due to bariatric surgery.

Participants: Jannot A. S .
                                  .

Lartigue, T., Durrleman, S., and Allassonnière, S. (2022). Deterministic approximate EM algorithm; Application to the Riemann approximation EM and the tempered EM. Algorithms, 15(3), 78. In this paper, we introduce a theoretical framework, with state-of-the-art convergence guarantees, for any deterministic approximation of the E step of the expectation-maximisation algorithm. We analyse theoretically and empirically several approximations that fit into this framework. First, for intractable E-steps, we introduce a deterministic version of MC-EM using Riemann sums. A straightforward method, not requiring any hyper-parameter fine-tuning, useful when the low dimensionality does not warrant a MC-EM. Then, we consider the tempered approximation, borrowed from the Simulated Annealing literature and used to escape local extrema. We prove that the tempered EM verifies the convergence guarantees for a wider range of temperature profiles than previously considered. We showcase empirically how new non-trivial profiles can more successfully escape adversarial initialisations. Finally, we combine the Riemann and tempered approximations into a method that accomplishes both their purposes.

Participants: Allassonnière, S ..

8.3   Axis 3

Cottin A, Pecuchet N, Zulian M, Guilloux A, Katsahian S. IDNetwork: A deep illness-death network based on multi-state event history process for disease prognostication. Statistics in Medicine. 2022; 41(9): 1573– 1598. doi:10.1002/sim.9310 Multi-state models can capture the different patterns of disease evolution. In particular, the illness-death model is used to follow disease progression from a healthy state to an intermediate state of the disease and to a death-related final state. We aim to use those models in order to adapt treatment decisions according to the evolution of the disease. In state-of-the-art methods, the risks of transition between the states are modeled via (semi-) Markov processes and transition-specific Cox proportional hazard (P.H.) models. The Cox P.H. model assumes that each variable makes a linear contribution to the model, but the relationship between covariates and risks can be more complex in clinical situations. To address this challenge, we propose a neural network architecture called illness-death network (IDNetwork) that relaxes the linear Cox P.H. assumption within an illness-death process. IDNetwork employs a multi-task architecture and uses a set of fully connected subnetworks in order to learn the probabilities of transition. Through simulations, we explore different configurations of the architecture and demonstrate the added value of our model. IDNetwork significantly improves the predictive performance compared to state-of-the-art methods on a simulated data set, on two clinical trials for patients with colon cancer and on a real-world data set in breast cancer.

> **Participants:**   Cottin A ., Guilloux A ., Katsahian S ..

Deep learning segmentation algorithm for sarcopenia Roblot V, Giret Y, Mezghani S, Auclin E, Arnoux A, Oudard S, Duron L, and Fournier L. Validation of a deep learning segmentation algorithm to quantify the skeletal muscle index and sarcopenia in metastatic renal carcinoma. Eur Radiol 32, 4728–4737 (2022) To validate a deep learning (DL) algorithm for measurement of skeletal muscular index (SMI) and prediction of overall survival in oncology populations. A retrospective single-center observational study included patients with metastatic renal cell carcinoma between 2007 and 2019. A set of 37 patients was used for technical validation of the algorithm, comparing manual vs DL-based evaluations. Segmentations were compared using mean Dice similarity coefficient (DSC), SMI using concordance correlation coefficient (CCC) and Bland-Altman plots. Overall survivals (OS) were compared using log-rank (Kaplan-Meier) and Mann-Whitney tests. Generalizability of the prognostic value was tested in an independent validation population (N = 87). Differences between two manual segmentations (DSC = 0.91, CCC = 0.98 for areas) or manual vs. automated segmentation (DSC = 0.90, CCC = 0.98 for areas, CCC = 0.97 for SMI) had the same order of magnitude. Bland-Altman plots showed a mean difference of -3.33 cm(2) [95%CI: -15.98, 9.1] between two manual segmentations, and -3.28 cm(2) [95% CI: -14.77, 8.21] for manual vs. automated segmentations. With each method, 20/37 (56%) patients were classified as sarcopenic. Sarcopenic vs. non-sarcopenic groups had statistically different survival curves with median OS of 6.0 vs. 12.5 (p = 0.008) and 6.0 vs. 13.9 (p = 0.014) months respectively for manual and DL methods. In the independent validation population, sarcopenic patients according to DL had a lower OS (10.7 vs. 17.3 months, p = 0.033). A DL algorithm allowed accurate estimation of SMI compared to manual reference standard. The DL-calculated SMI demonstrated a prognostic value in terms of OS.

> **Participants:**   Arnoux A ..

Hua, W., Mei, H., Zohar, S., Giral, M., and Xu, Y. (2022). Personalized dynamic treatment regimes in continuous time: a Bayesian approach for optimizing clinical decisions with timing. Bayesian Analysis, 17(3), 849-878. We developed a two-step Bayesian approach to optimize clinical decisions, estimating personalized optimal dynamic treatment regimes (DTRs), taking into account intervention timing. In the first step, we build a generative model for a sequence of medical interventions with a marked temporal point process (MTPP) where the mark is the assigned treatment or dosage. Then this clinical action model is embedded into a Bayesian joint framework where the other components model clinical observations including longitudinal medical measurements and time-to-event data conditional on

treatment histories. In the second step, we proposed a policy gradient method to learn the personalized optimal clinical decision that maximizes the patient survival by interacting the MTPP with the model on clinical observations while accounting for uncertainties in clinical observations learned from the posterior inference of the Bayesian joint model in the first step.

Participants: Zohar, S ..

Gerard, E., Zohar, S., Thai, H. T., Lorenzato, C., Riviere, M. K., and Ursino, M. (2022). Bayesian dose regimen assessment in early phase oncology incorporating pharmacokinetics and pharmaco-dynamics. Biometrics, 78(1), 300-312.   We proposed a Bayesian dose regimen assessment method (DRtox) using pharmacokinetics/pharmacodynamics (PK/PD) to estimate the maximum tolerated dose regimen (MTD-regimen) at the end of the dose-escalation stage of a immunoteraphy trial. We modeled the binary toxicity via a PD endpoint and estimated the dose regimen toxicity relationship through the integration of a dose regimen PD model and a PD toxicity model. For the first model, we considered non-linear mixed-effects models, and for the second one, we proposed the following two Bayesian approaches: a logistic model and a hierarchical model.

Participants: Gerard, E ., Zohar, S ., Ursino, M ..

Röver C, Ursino M, Friede T, Zohar S. A straightforward meta-analysis approach for oncology phase I dose-finding studies. Stat Med. 2022 Sep 10;41(20):3915-3940.
Phase I early-phase clinical studies aim at investigating the safety and the underlying dose-toxicity relationship of a drug or combination. While little may still be known about the compound's properties, it is crucial to consider quantitative information available from any studies that may have been conducted previously on the same drug. A meta-analytic approach has the advantages of being able to properly account for between-study heterogeneity, and it may be readily extended to prediction or shrinkage applications. Here we propose a simple and robust two-stage approach for the estimation of maximum tolerated dose(s) utilizing penalized logistic regression and Bayesian random-effects meta-analysis methodology. Implementation is facilitated using standard R packages. The properties of the proposed methods are investigated in Monte Carlo simulations. The investigations are motivated and illustrated by two examples from oncology.

Participants: Ursino M ., Zohar S ..

9   Bilateral contracts and grants with industry

9.1   Bilateral Grants with Industry

- Sandrine Katsahian signed a new cifre contract with Sanofi for a PhD under a co-supervision with Agathe Guilloux.

- Bourse cifre Exystat- ANRT pour l'encadrement de B Duputel (2020-2023) encadré par Sarah Zohar et Moreno Ursino

- Bourse cifre Dassault Système - ANRT pour l'encadrement d'Aziliz Cottin encadrée par Sandrine Katsahian et Agathe Guilloux are still ongoing (2020-2023).

- Bourse cifre Pierre Fabre - ANRT pour l'encadrement de Abir Tadmouri Sellier encadré par Sandrine Katsahian et Agathe Guilloux (2021-2024)

- Bourse cifre Implicty - ANRT pour l'encadrement de Louis Vincent encadré par Stéphanie Allasson-nière (2021-2024)

10 Partnerships and cooperations

10.1 International initiatives

10.1.1 Participation in other International Programs

4CE is an international consortium for electronic health record (EHR) data-driven studies of the COVID-19 pandemic. The goal of this effort—led by the i2b2 international academics users group—is to inform doctors, epidemiologists and the public about COVID-19 patients with data acquired through the health care process.

10.2 International research visitors

10.2.1 Visits to international teams

David Drummont was a visiting Fuculty at the Centre For Health Informatics, University of Manchester (Sept2021-Sept2022).

10.3 European initiatives

10.3.1 Other european programs - initiatives

The European taskforce lead by EIT Health and French Ministry of Solidarity and Health, for the "harmonization of clinical studies criteria and methodologies in Europe for the evaluation of digital medical devices". In this taskforce, Sarah Zohar co-lead the WP2 on "Evidence in clinical evaluation" with Corinne Collignon (Head of the Digital Mission at HAS, France) and Barbara Hofgen (Head of the Unit DiGA-Fast-Track at Bfarm, Germany. Guidelines will generated to be used by the European commission to establish regulatory policy.

10.4 National initiatives

ReCAP - Moreno Ursino is the referent of the « Early trials » group.The national network RECaP of Research in Clinical Epidemiology and in Public Health has for purpose to mutualize original research projects and to produce innovations in clinical epidemiology and public health.

11 Dissemination

11.1 Promoting scientific activities

Member of the organizing committees

- Members of HeKA are part of program committees of several events, including: Medical Informatics Europe 2022, 2022 OHDSI Symposium, the international SWAT4HCLS Conference, La journée IA and Santé de la plateforme PFIA.

11.1.1 Journal

Member of the editorial boards

- Sarah Zohar is Associated Editor of two scientific journals; "Biometrics" and "Statistics in Biopharmaceutical Research"

- Bastien Rance is member of the editorial board of the "International Journal of Medical Informatics".

Reviewer - reviewing activities

- Members of HeKA are regular reviewers for the following journals: Scientific Data, Scientific Reports, Semantic Web Journal, Journal of the Semantic Web, Artificial Intelligence in Medicine, statistics in medicine, biometrics, statistics in medical research, etc.

11.1.2 Leadership within the scientific community

- Antia Burgun serves as the Representative of the French medical informatics community at the IMIA (International Medical Informatics Association). She also serves as a member of the Executive Board of the Imagine Institute.

11.1.3 Scientific expertise

- Sarah Zohar is a voting member at the Cnedimts ("Commission nationale d'évaluation des dispositifs médicaux et des technologies de santé") at HAS ("Haute Autorité de Santé").

- Sarah Zohar was in the review committee of BpiFrance grant evaluation "évaluation DM à base d'IA et de numérique".

- Members of the HeKA team are involved in several ethic committees, such as the *Comité d'Ethique de la Recherche APHP.Centre*, the EDS AP-HP *Comité Scientifique et Ethique*. Members of HeKA the scientific board of the ANR generic call *Axe H.14 : Interfaces : mathématiques, sciences du numérique – biologie, santé* and of the FC3R (French Center for the 3R).

- Stéphanie Allassonière and Anita Burgun hold the Helath Chair at PRAIRIE institute.

11.2 Teaching - Supervision - Juries

- Anne Sophie Jannot and Sandrine Katsahian lead the speciality "Big Data in Health" in the Master of Science of Public Health at the Paris Cité University.

- Anne Sophie Jannot co-leads the quantitative biomedicine course, which is part from the medical degree course.

- Anne Sophie Jannot leads a professional degree of "health data reuse" at the Paris Cité University in collaboration with Marseille and Bordeaux Universities.

- Sandrine Katsahian is responsible for the PRIME department dedicated to the Research, Innovation, Digital Medicine located in the HEGP in APHP.Centre.

- Moreno URSINO co-leads the course "Science des données" in the L2SIAS at the Paris Cité University.

- Members of the HeKA team are in charge of two majors of the Master in Public Health of the Université Paris Cité: one on Biomedical Informatics (IBM) and one on Data Science in Healthcare (DMS).

11.3 Popularization

11.3.1 Internal or external Inria responsibilities

- Stéphanie Allassonière was the chair of the CRCN-ISFP INRIA Lyon committee.

11.3.2 Interventions

- Sarah Zohar has participated to "Leem - Les Entreprises du médicament - Atelier presse « Essais Cliniques 2030 »"

- Sarah Zohar has presented HeKA's project to the 21e eHealth Network Meeting de la commission européenne dans le cadre de la présidence Française de l'UE

- Sarah Zohar has presented "Conférence au grand public à la Cité de l'Economie (Paris) - "Quand le numérique révolutionne la santé, quel avenir pour la HealthTech ?""

12 Scientific production

12.1 Major publications

[1] C. Chadebec, E. Thibeau-Sutre, N. Burgos and S. Allassonnière. 'Data Augmentation in High Dimensional Low Sample Size Setting Using a Geometry-Based Variational Autoencoder'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (24th June 2022). URL: `https://hal.archives-ouvertes.fr/hal-03214093`.

[2] A. Cottin, N. Pecuchet, M. Zulian, A. Guilloux and S. Katsahian. 'IDNetwork: A deep illness-death network based on multi-state event history process for disease prognostication'. In: *Statistics in Medicine* 41.9 (30th Apr. 2022), pp. 1573–1598. DOI: `10.1002/sim.9310`. URL: `https://hal.archives-ouvertes.fr/hal-03822456`.

[3] E. Gerard, S. Zohar, H.-t. Thai, C. Lorenzato, M.-k. Riviere and M. Ursino. 'Bayesian dose-regimen assessment in early phase oncology incorporating pharmacokinetics and pharmacodynamics'. In: *Biometrics* (2nd Feb. 2022). DOI: `10.1111/biom.13433`. URL: `https://hal.archives-ouvertes.fr/hal-03141298`.

[4] I. Lerner, A. Serret-Larmande, B. Rance, N. Garcelon, A. Burgun, L. Chouchana and A. Neuraz. 'Mining Electronic Health Records for Drugs Associated With 28-day Mortality in COVID-19: Pharmacopoeia-wide Association Study (PharmWAS)'. In: *JMIR Medical Informatics* 10.3 (2022), e35190. DOI: `10.2196/35190`. URL: `https://hal.archives-ouvertes.fr/hal-03792405`.

[5] S. Pruilh, A.-S. Jannot and S. Allassonnière. 'Spatio-temporal mixture process estimation to detect population dynamical changes'. In: *Artificial Intelligence in Medicine* 126 (2022), p. 102258. DOI: `10.1016/j.artmed.2022.102258`. URL: `https://hal.archives-ouvertes.fr/hal-02933217`.

[6] A. Rogier, A. Coulet and B. Rance. 'Using an ontological representation of chemotherapy toxicities for guiding information extraction and integration from EHRs'. In: Medinfo 2021 - 18th World Congress on Medical and Health Informatics. Virtual conference, Australia, 2nd Oct. 2021. URL: `https://hal.inria.fr/hal-03364585`.

12.2 Publications of the year

International journals

[7] D. Agher, K. Sedki, S. Despres, J.-P. Albinet, M.-C. Jaulent and R. Tsopra. 'Encouraging Behavior Changes and Preventing Cardiovascular Diseases Using the Prevent Connect Mobile Health App: Conception and Evaluation of App Quality'. In: *Journal of Medical Internet Research* 24.1 (20th Jan. 2022), e25384. DOI: `10.2196/25384`. URL: `https://hal.sorbonne-universite.fr/hal-03539411`.

[8] N. Bannour, P. Wajsbürt, B. Rance, X. Tannier and A. Névéol. 'Privacy-preserving mimic models for clinical named entity recognition in French'. In: *Journal of Biomedical Informatics* 130 (June 2022), p. 104073. DOI: `10.1016/j.jbi.2022.104073`. URL: `https://hal.archives-ouvertes.fr/hal-03655039`.

[9]   M. Barat, A.-S. Jannot, A. Dohan and P. Soyer. 'How to report and compare quantitative variables in a radiology article'. In: *Diagnostic and Interventional Imaging* 103.12 (Dec. 2022), pp. 571–573. DOI: 10.1016/j.diii.2022.09.007. URL: https://hal.science/hal-03911632.

[10]   C. Baudrier, Y. Tran, N. Delanoy, S. Katsahian, B. Sabatier and G. Perrin. 'Identifying homogeneous healthcare use profiles and treatment sequences by combining sequence pattern mining with care trajectory clustering in kidney cancer patients on oral anticancer drugs: A case study'. In: *Health Informatics Journal* 28.2 (Jan. 2022), p. 146045822211015. DOI: 10.1177/14604582221101526. URL: https://hal.archives-ouvertes.fr/hal-03794007.

[11]   T. Bertrand, C. Rives-Lange, A.-S. Jannot, C. Baratte, F. de Castelbajac, E. Lu, S. Krivan, M. Le Gall, C. Carette, S. Czernichow, J.-M. Chevallier and T. Poghosyan. '150-cm Versus 200-cm Biliopancreatic Limb One-Anastomosis Gastric Bypass: Propensity Score–Matched Analysis'. In: *Obesity Surgery* 32.9 (Sept. 2022), pp. 2839–2845. DOI: 10.1007/s11695-022-06203-9. URL: https://hal.archives-ouvertes.fr/hal-03911638.

[12]   A. Boukobza, A. Burgun, B. Roudier and R. Tsopra. 'Deep Neural Networks for Simultaneously Capturing Public Topics and Sentiments During a Pandemic: Application on a COVID-19 Tweet Data Set'. In: *JMIR Medical Informatics* 10.5 (2022), e34306. DOI: 10.2196/34306. URL: https://hal.archives-ouvertes.fr/hal-03738943.

[13]   D. Bouzid, J. Mullaert, A. Ghazali, V. M. Ferré, F. Mentré, C. Lemogne, P. Ruszniewski, A. Faye, A. T. Dinh, T. Mirault, N. P. Smadja, L. Muller, L. F. Pierrotin, M. Thy, M. Assadi, S. Yung, C. de Tymowski, Q. Le Hingrat, X. Eyer, P. H. Wicky, M. Oualha, V. Houdouin, P. Jabre, D. Vodovar, M. D. Burgio, N. Zucman, R. Tsopra, A. Tazi, Q. Ressaire, Y. Nguyen, M. Girard, A. Frachon, F. Depret, A. Pellat, A. de Masson, H. Azais, N. de Castro, C. Jeantrelle, N. Javaud, A. Malmartel, C. J. de Margerie, B. Chousterman, L. Fournel, M. Holleville and S. Blanche. 'eOSCE stations live versus remote evaluation and scores variability'. In: *BMC Medical Education* 22.1 (Dec. 2022), p. 861. DOI: 10.1186/s12909-022-03919-1. URL: https://hal.archives-ouvertes.fr/hal-03910361.

[14]   T. Caruba, A. Charles-Nelson, J. Alexander, R. Mehta, M. Heringlake, W. Toller, A. Yavchitz, B. Sabatier, G. Chatellier and B. Cholley. 'Prophylactic levosimendan in patients with low ejection fraction undergoing coronary artery bypass grafting: A pooled analysis of two multicentre randomised controlled trials'. In: *Anaesthesia Critical Care & Pain Medicine* 41.4 (Aug. 2022), p. 101107. DOI: 10.1016/j.accpm.2022.101107. URL: https://hal.archives-ouvertes.fr/hal-03794008.

[15]   C. Chadebec, E. Thibeau-Sutre, N. Burgos and S. Allassonnière. 'Data Augmentation in High Dimensional Low Sample Size Setting Using a Geometry-Based Variational Autoencoder'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (24th June 2022). DOI: 10.1109/TPAMI.2022.3185773. URL: https://hal.science/hal-03214093.

[16]   A. Challine, L. Maggiori, S. Katsahian, H. Corté, D. Goere, A. Lazzati, P. Cattan and M. Chirica. 'Outcomes Associated With Caustic Ingestion Among Adults in a National Prospective Database in France'. In: *JAMA Surgery* 157.2 (1st Feb. 2022), p. 112. DOI: 10.1001/jamasurg.2021.6368. URL: https://hal.archives-ouvertes.fr/hal-03825949.

[17]   A. Challine, O. Soubrane and J. H. Lefevre. 'The first year of Junior Doctors in visceral and digestive surgery in France: For what position and with which function?' In: *Journal of Visceral Surgery* 159.2 (Apr. 2022), pp. 179–180. DOI: 10.1016/j.jviscsurg.2022.01.012. URL: https://hal.inria.fr/hal-03864966.

[18]   A. Challine, T. Voron, L. O'connell, N. Chafai, C. Debove, M. Collard, Y. Parc and J. Lefèvre. 'Does An Ileo-anal Anastomosis Decrease the Rate of Successful Pregnancy Compared to an Ileorectal Anastomosis? A National Study of 1,491 Patients'. In: *Annals of Surgery* Publish Ahead of Print (2022). DOI: 10.1097/SLA.0000000000005569. URL: https://hal.inria.fr/hal-03864955.

[19] F. Charbit-Henrion, M. Haas, S. Chaussade, C. Cellier, N. Cerf-Bensussan, G. Malamut, S. Khater, A. Khiat, S. Cording, M. Parlato, M.-A. Dragon-Durey, F. Beuvon, N. Brousse, B. Terris, C. Picard, M. Fusaro, F. Rieux-Laucat, M.-C. Stolzenberg, A.-S. Jannot, A. Mathian, M. Allez, M. Malphettes, C. Fieschi, A. Aubourg, C. Zallot, X. Roblin, V. Abitbol, A. Belle, P. Wils, M. Cheminant, T. Matysiak-Budnik, L. Vuitton, P. Pouderoux, L. Abramowitz, M. Castelle, F. Suarez, O. Hermine, F. Ruemmele and L. Mouthon. 'Genetic Diagnosis Guides Treatment of Autoimmune Enteropathy'. In: *Clinical Gastroenterology and Hepatology* (Aug. 2022). DOI: 10.1016/j.cgh.2022.07.030. URL: https://hal.science/hal-03911634.

[20] X. Chen, C. Faviez, M. Vincent, L. Briseño-Roa, H. Faour, J.-P. Annereau, S. Lyonnet, M. Zaidan, S. Saunier, N. Garcelon and A. Burgun. 'Patient-Patient Similarity-Based Screening of a Clinical Data Warehouse to Support Ciliopathy Diagnosis'. In: *Frontiers in Pharmacology* 13 (25th Mar. 2022). DOI: 10.3389/fphar.2022.786710. URL: https://hal.science/hal-03629207.

[21] A. Cottin, N. Pecuchet, M. Zulian, A. Guilloux and S. Katsahian. 'IDNetwork: A deep illness-death network based on multi-state event history process for disease prognostication'. In: *Statistics in Medicine* 41.9 (30th Apr. 2022), pp. 1573–1598. DOI: 10.1002/sim.9310. URL: https://hal.archives-ouvertes.fr/hal-03822456.

[22] B. Denis, I. Gendre, N. Tuzin, J. Murris, A. Guignard, P. Perrin and G. Rahmi. 'Adenoma detection rate is enough to assess endoscopist performance: a population-based observational study of FIT-positive colonoscopies'. In: *Endoscopy International Open* 10.09 (14th Sept. 2022), E1208–E1217. DOI: 10.1055/a-1859-8277. URL: https://hal.archives-ouvertes.fr/hal-03822394.

[23] D. Drummond. 'Methionine supplementation for multi-organ dysfunction in MetRS-related pulmonary alveolar proteinosis'. In: *Revue des Maladies Respiratoires* 39.3 (Mar. 2022), pp. 241–257. DOI: 10.1016/j.rmr.2022.01.018. URL: https://hal.archives-ouvertes.fr/hal-03866739.

[24] D. Drummond and A. Coulet. 'Technical, Ethical, Legal, and Societal Challenges With Digital Twin Systems for the Management of Chronic Diseases in Children and Young People'. In: *Journal of Medical Internet Research* 24.10 (2022), e39698. DOI: 10.2196/39698. URL: https://hal.inria.fr/hal-03835708.

[25] D. Drummond, J. Dana, L. Berteloot, E. Schneider-Futschik, F. Chedevergne, C. Bailly-Botuha, T. Nguyen-Khoa, M. Cornet, M. Le Bourgeois, D. Debray, M. Girard and I. Sermet-Gaudelus. 'Lumacaftor-ivacaftor effects on cystic fibrosis-related liver involvement in adolescents with homozygous F508 del-CFTR'. In: *Journal of Cystic Fibrosis* 21.2 (Mar. 2022), pp. 212–219. DOI: 10.1016/j.jcf.2021.07.018. URL: https://hal.archives-ouvertes.fr/hal-03866719.

[26] D. Drummond, A. Hadchouel, A. Petit, N. Khen-Dunlop, C. Lozach, C. Delacourt and L. Berteloot. 'Strategies for recognizing pneumonia look-alikes'. In: *European Journal of Pediatrics* 181.10 (Oct. 2022), pp. 3565–3575. DOI: 10.1007/s00431-022-04575-9. URL: https://hal.archives-ouvertes.fr/hal-03866696.

[27] D. Drummond, C. Thumerelle, A. Roux, C. Mordacq, P. Frange, M. Leruez-Ville, L. Gibault, L. Berteloot, C. Roy, M. Pontailler, V. Lopez, M. Oualha, M. Grimaud, L. Saint Blanquat, F. Parquin and I. Sermet-Gaudelus. 'Severe COVID-19 evolving towards organizing pneumonia in a pediatric lung transplant recipient'. In: *Pediatric Pulmonology* 57.2 (Feb. 2022), pp. 583–585. DOI: 10.1002/ppul.25744. URL: https://hal.archives-ouvertes.fr/hal-03866717.

[28] C. Durand, S. Alfandari, G. Béraud, R. Tsopra, F.-X. Lescure and N. Peiffer-Smadja. 'Clinical Decision Support Systems for Antibiotic Prescribing: An Inventory of Current French Language Tools'. In: *Antibiotics* 11.3 (Mar. 2022), p. 384. DOI: 10.3390/antibiotics11030384. URL: https://hal.archives-ouvertes.fr/hal-03738939.

[29] L.-A. Eveillard, P. Quartier, N. Ouldali, B. Bader-Meunier, F. Aeschlimann, C. Abasq, C. Ballot, P. Bouric, A. Desdoits, C. Dumaine, C. Galeotti, V. Hentgen, A. Lefevre-Utile, A. Chausset, T. Hubiche, I. Kupfer-Bessaguet, S. Leclerq-Mercier, S. Mallet, I. Melki, E. Merlin, J. Miquel, M. Piram, D. Talmud, N. Garcelon, C. Vinit, A. Welfringer, E. Bourrat and U. Meinzer. 'Association of atypical skin manifestations at the onset of systemic juvenile idiopathic arthritis with difficult-to-treat disease: A retrospective multicenter study'. In: *Journal of The American Academy of Dermatology*

(Aug. 2022). DOI: 10.1016/j.jaad.2022.07.055. URL: https://hal.archives-ouvertes.f
r/hal-03794615.

[30] A. Fayol, M. Wack, M. Livrozet, J.-B. Carves, O. Domengé, E. Vermersch, M. Mirabel, A. Karras, J. Le Guen, A. Blanchard, M. Azizi, L. Amar, M.-C. Bories, E. Mousseaux, C. Carette, E. Puymirat, A. Hagège, A.-S. Jannot and J.-S. Hulot. 'Aetiological classification and prognosis in patients with heart failure with preserved ejection fraction'. In: *ESC Heart Failure* 9.1 (Feb. 2022), pp. 519–530. DOI: 10.1002/ehf2.13717. URL: https://hal.science/hal-03560925.

[31] F. Gaillard, M. Ould Rabah, N. Garcelon, M. Touam, A. Neuraz, C. Legendre, D. Anglicheau, D. Prié and F. Bienaimé. 'Allograft function and muscle mass evolution after kidney transplantation'. In: *Journal of Cachexia, Sarcopenia and Muscle* (15th Sept. 2022). DOI: 10.1002/jcsm.13066. URL: https://hal.archives-ouvertes.fr/hal-03794605.

[32] S. Gaujoux, G. Martin, E. Mirallié, N. Regenet, M. Le Bras, F. Pattou, B. Carnaille, C. Cardot-Bauters, L. Groussin, M. Faron, P. Chanson, H. Najah, A. Tabarin, A. Sauvanet, P. Ruszniewski, J. C. Lifante, T. Walter, N. Carrère, P. Caron, S. Deguelte, B. Delemer, C. Binquet, A. S. Jannot and P. Goudet. 'Life expectancy and likelihood of surgery in multiple endocrine neoplasia type 1: AFCE and GTE cohort study'. In: *British Journal of Surgery* 109.9 (1st Sept. 2022), pp. 872–879. DOI: 10.1093/bjs/znac006. URL: https://hal.archives-ouvertes.fr/hal-03883938.

[33] N. Gendron, L. Khider, C. Le Beller, B. Espinasse, C. Auditeau, W. Amara, G. Perrin, D. Lebeaux, A. Gaiffe, S. Combret, B. Bertin, A. Lillo-Le Louet, T. Mirault, D. Smadja, O. Sanchez, C. Tromeur, B. Planquette and F. Couturaud. 'Bleeding risk of intramuscular injection of COVID-19 vaccines in adult patients with therapeutic anticoagulation'. In: *Journal of Thrombosis and Haemostasis* 20.6 (June 2022), pp. 1507–1510. DOI: 10.1111/jth.15709. URL: https://hal.archives-ouverte
s.fr/hal-03866352.

[34] E. Gerard, S. Zohar, H.-t. Thai, C. Lorenzato, M.-k. Riviere and M. Ursino. 'Bayesian dose-regimen assessment in early phase oncology incorporating pharmacokinetics and pharmacodynamics'. In: *Biometrics* (2nd Feb. 2022). DOI: 10.1111/biom.13433. URL: https://hal.archives-ouvert
es.fr/hal-03141298.

[35] A. Guédon, J.-B. D. Freminville, T. Mirault, N. Mohamedi, B. Rance, N. Fournier, J.-L. Paul, E. Messas and G. Goudot. 'Elevated Lipoprotein(a) levels increase Major Adverse Limb Event'. In: *European Heart Journal* 43.Supplement_2 (1st Oct. 2022), e20773. DOI: 10.1093/eurheartj/eh
ac544.1980. URL: https://hal.archives-ouvertes.fr/hal-03865725.

[36] C. Hauw-Berlemont, L. Lamhaut, J.-L. Diehl, C. Andreotti, O. Varenne, P. Leroux, J.-B. Lascarrou, P. Guerin, T. Loeb, E. Roupie et al. 'Emergency vs Delayed Coronary Angiogram in Survivors of Out-of-Hospital Cardiac Arrest'. In: *JAMA Cardiology* 7.7 (1st July 2022), p. 700. DOI: 10.1001/jam
acardio.2022.1416. URL: https://hal.archives-ouvertes.fr/hal-03825951.

[37] N. Hoertel, D. Boulware, M. Sánchez-Rico, A. Burgun and F. Limosin. 'Prevalence of Contraindications to Nirmatrelvir-Ritonavir Among Hospitalized Patients With COVID-19 at Risk for Progression to Severe Disease'. In: *JAMA Network Open* 5.11 (1st Nov. 2022), e2242140. DOI: 10.1001/jam
anetworkopen.2022.42140. URL: https://hal.inria.fr/hal-03889233.

[38] C. Hong, H. Zhang, S. L'Yi, G. Weber, P. Avillach, B. Tan, A. Gutiérrez-Sacristán, C.-L. Bonzel, N. Palmer, A. Malovini et al. 'Changes in laboratory value improvement and mortality rates over the course of the pandemic: an international retrospective cohort study of hospitalised patients infected with SARS-CoV-2'. In: *BMJ Open* 12.6 (23rd June 2022), e057725. DOI: 10.1136/bmjopen-
2021-057725. URL: https://hal.inria.fr/hal-03886983.

[39] A.-S. Jannot, C. Messiaen, A. Khatim, T. Pichon and A. Sandrin. 'The ongoing French BaMaRa-BNDMR cohort: implementation and deployment of a nationwide information system on rare disease'. In: *Journal of the American Medical Informatics Association* 29.3 (1st Mar. 2022), pp. 553–558. DOI: 10.1093/jamia/ocab237. URL: https://hal.archives-ouvertes.fr/hal-03560
936.

[40] L. Joseph, A. Corbasson, S. Manceau, D. Khimoud, B. Meunier, G. Cheminet, F. Lefrere, A.-s. Jannot, E. Lu and J.-B. Arlet. 'Safety of coronavirus disease 2019 vaccines in 213 adult patients with sickle cell disease'. In: *British Journal of Haematology* (2022). DOI: 10.1111/bjh.18547. URL: https://hal.science/hal-03911627.

[41] A. Kalimouttou, I. Lerner, C. Cheurfa, A.-S. Jannot and R. Pirracchio. 'Machine-learning-derived sepsis bundle of care'. In: *Intensive Care Medicine* (2022). DOI: 10.1007/s00134-022-06928-2. URL: https://hal.science/hal-03911622.

[42] L. Lahmi, M.-F. Mamzer, A. Burgun, C. Durdux and J.-E. Bibault. 'Ethical Aspects of Artificial Intelligence in Radiation Oncology'. In: *Seminars in Radiation Oncology* 32.4 (Oct. 2022), pp. 442–448. DOI: 10.1016/j.semradonc.2022.06.013. URL: https://hal.inria.fr/hal-0388923 1.

[43] J. Lambert, E. Lengliné, R. Porcher, R. Thiébaut, S. Zohar and S. Chevret. 'Enriching single-arm clinical trials with external controls: possibilities and pitfalls'. In: *Blood Advances* (2022). DOI: 10.1182/bloodadvances.2022009167. URL: https://hal.inria.fr/hal-03914596.

[44] T. Lartigue, S. Durrleman and S. Allassonnière. 'Deterministic Approximate EM Algorithm; Application to the Riemann Approximation EM and the Tempered EM'. In: *Algorithms*. Stochastic Algorithms and Their Applications 15.3 (25th Feb. 2022), p. 78. DOI: 10.3390/a15030078. URL: https://hal.inria.fr/hal-02513593.

[45] A. Lazzati, S. Epaud, M. Ortala, S. Katsahian and E. Lanoy. 'Effect of bariatric surgery on cancer risk: results from an emulated target trial using population-based data'. In: *British Journal of Surgery* 109.5 (13th Apr. 2022), pp. 433–438. DOI: 10.1093/bjs/znac003. URL: https://hal.ar chives-ouvertes.fr/hal-03825952.

[46] I. Lerner, A. Serret-Larmande, B. Rance, N. Garcelon, A. Burgun, L. Chouchana and A. Neuraz. 'Mining Electronic Health Records for Drugs Associated With 28-day Mortality in COVID-19: Pharmacopoeia-wide Association Study (PharmWAS)'. In: *JMIR Medical Informatics* 10.3 (2022), e35190. DOI: 10.2196/35190. URL: https://hal.archives-ouvertes.fr/hal-03792405.

[47] M. Levy, M. Recher, H. Hubert, E. Javouhey, O. Fléchelles, S. Leteurtre and F. Angoulvant. 'Multisystem Inflammatory Syndrome in Children by COVID-19 Vaccination Status of Adolescents in France'. In: *JAMA Cardiology* 327.3 (18th Jan. 2022), p. 281. DOI: 10.1001/jama.2021.23262. URL: https://hal.science/hal-03941204.

[48] C. Mantoux, S. Durrleman and S. Allassonnière. 'Asymptotic Analysis of a Matrix Latent Decomposition Model'. In: *ESAIM: Probability and Statistics* 26 (2022), pp. 208–242. DOI: 10.1051/ps/2 022004. URL: https://hal.archives-ouvertes.fr/hal-03674722.

[49] F. Mifsud, D. Messager, A.-S. Jannot, B. Védie, N. A. Balanant, T. Poghosyan, E. Flamarion, C. Carette, L. Lucas-Martini, S. Czernichow and C. Rives-Lange. 'Clinical diagnosis, outcomes and treatment of thiamine deficiency in a tertiary hospital'. In: *Clinical Nutrition* 41.1 (Jan. 2022), pp. 33–39. DOI: 10.1016/j.clnu.2021.10.021. URL: https://hal.archives-ouvertes.fr /hal-03560923.

[50] L. Mimram, S. Magréault, Y. Kerroumi, D. Salmon, B. Kably, S. Marmor, A.-S. Jannot, V. Jullien and V. Zeller. 'Population Pharmacokinetics of Orally Administered Clindamycin to Treat Prosthetic Joint Infections: A Prospective Study'. In: *Antibiotics* 11.11 (Nov. 2022), p. 1462. DOI: 10.3390/ant ibiotics11111462. URL: https://hal.science/hal-03911624.

[51] P. Monnin and A. Coulet. 'Étudier les mécanismes des effets indésirables des médicaments avec l'IA explicable : expériences avec la fouille de graphes de connaissances'. In: *Bulletin de l'Association Française pour l'Intelligence Artificielle* (Apr. 2022). URL: https://hal.inria.fr/hal-0366906 5.

[52] P. Monnin and A. Coulet. 'Projet ANR (2016-2021) « PractiKPharma » : extraction, comparaison et découverte de connaissances en pharmacogénomique'. In: *1024 : Bulletin de la Société Informatique de France* 19 (Apr. 2022), pp. 109–120. DOI: 10.48556/sif.1024.19.109. URL: https://hal.inria.fr/hal-03669026.

[53] P. Monnin, C. Raïssi, A. Napoli and A. Coulet. 'Discovering alignment relations with Graph Convolutional Networks: A biomedical case study'. In: *Semantic Web – Interoperability, Usability, Applicability* (1st Jan. 2022), pp. 1–20. DOI: 10.3233/SW-210452. URL: https://hal.inria.fr/hal-03452182.

[54] A. Mouazer, R. Leguillon, B. Leroy, K. Sedki, C. Simon, H. Falcoff, M. Schuers, S. Dubois, C. Bousquet, L. Soualmia, C. Letord, S. Darmoni, J. Bouaud, B. Séroussi, R. Tsopra and J.-B. Lamy. 'ABiMed: Towards an Innovative Clinical Decision Support System for Medication Reviews and Polypharmacy Management'. In: *Studies in Health Technology and Informatics*. Studies in Health Technology and Informatics 289 (14th Jan. 2022), pp. 61–64. DOI: 10.3233/SHTI210859. URL: https://hal.archives-ouvertes.fr/hal-03738935.

[55] A. Mouazer, R. Tsopra, K. Sedki, C. Letord and J.-B. Lamy. 'Decision-support systems for managing polypharmacy in the elderly: A scoping review'. In: *Journal of Biomedical Informatics* (Apr. 2022), p. 104074. DOI: 10.1016/j.jbi.2022.104074. URL: https://hal.archives-ouvertes.fr/hal-03651749.

[56] A. Nassar, A. Challine, L. O'connell, T. Voron, N. Chafaï, C. Debove, Y. Parc and J. Lefèvre. 'Effective initial management of anastomotic leak in the maintenance of functional colorectal or coloanal anastomosis'. In: *Surgery Today* (16th Nov. 2022). DOI: 10.1007/s00595-022-02603-7. URL: https://hal.inria.fr/hal-03864962.

[57] T. Nguyen-Khoa, A. Hatton, D. Drummond, L. Aoust, J. Schlatter, C. Martin, S. Ramel, S. Kiefer, E. Gachelin, N. Stremler, L. Cosson, A. Gabsi, N. Remus, M. Benhamida, A. Hadchouel, I. Fajac, A. Munck, E. Girodon and I. Sermet-Gaudelus. 'Reclassifying inconclusive diagnosis for cystic fibrosis with new generation sweat test'. In: *European Respiratory Journal* 60.2 (4th Aug. 2022), p. 2200209. DOI: 10.1183/13993003.00209-2022. URL: https://hal.archives-ouvertes.fr/hal-03866713.

[58] E. Payen, A. Neuraz, L. Zenzeri, C. Talbotec, E. Abi Nader, L. Chatenoud, S. Chhun, O. Goulet, F. Ruemmele and B. Pigneur. 'Adalimumab therapy in pediatric Crohn's Disease: a two-year follow-up comparing 'top-down' and 'step-up' strategies.' In: *Journal of Pediatric Gastroenterology and Nutrition* Publish Ahead of Print (2022). DOI: 10.1097/MPG.0000000000003643. URL: https://hal.inria.fr/hal-03886928.

[59] L. Payet, A. Challine and J. H. Lefevre. 'Strangulated hedrocele after a fall'. In: *Journal of Visceral Surgery* 159.5 (Oct. 2022), pp. 439–440. DOI: 10.1016/j.jviscsurg.2022.01.005. URL: https://hal.inria.fr/hal-03864871.

[60] G. Perrin, A. Arnoux, S. Berdot, S. Katsahian, N. Danchin and B. Sabatier. 'Association Between Exposure to Effervescent Paracetamol and Hospitalization for Acute Heart Failure: A Case-Crossover Study'. In: *Journal of Clinical Pharmacology* 62.7 (2022), pp. 883–890. DOI: 10.1002/jcph.2027. URL: https://hal.sorbonne-universite.fr/hal-03644630.

[61] E. Pinet, P. Sabatier, M.-P. Fernandez-Gerlinger, A.-S. Jannot, J.-L. Mainardi, B. Sabatier and T. Caruba. 'Impact of medical and pharmaceutical interventions on anti-infective prescriptions: an observational study'. In: *European Journal of Clinical Microbiology and Infectious Diseases* 41.7 (July 2022), pp. 1077–1086. DOI: 10.1007/s10096-022-04465-w. URL: https://hal.archives-ouvertes.fr/hal-03794006.

[62] B. Poitier, R. Chocron, C. Peronino, A. Philippe, Y. Pya, N. Rivet, U. Richez, M. Bekbossynova, N. Gendron, M. Grimmé, M. C. Bories, J. Brichet, A. Capel, J. Rancic, B. Vedie, J. C. Roussel, A.-S. Jannot, P. Jansen, A. Carpentier, P. Ivak, C. Latremouille, I. Netuka and D. Smadja. 'Bioprosthetic Total Artificial Heart in Autoregulated Mode Is Biologically Hemocompatible: Insights for Multimers of von Willebrand Factor'. In: *Arteriosclerosis, Thrombosis, and Vascular Biology* 42.4 (Apr. 2022), pp. 470–480. DOI: 10.1161/ATVBAHA.121.316833. URL: https://hal.archives-ouvertes.fr/hal-03911974.

[63] S. Pruilh, A.-S. Jannot and S. Allassonnière. 'Spatio-temporal mixture process estimation to detect population dynamical changes'. In: *Artificial Intelligence in Medicine* 126 (2022), p. 102258. DOI: 10.1016/j.artmed.2022.102258. URL: https://hal.archives-ouvertes.fr/hal-02933217.

[64] C. Rives-Lange, T. Poghosyan, A. Phan, A. van Straaten, Y. Girardeau, J. Nizard, D. Mitanchez, C. Ciangura, M. Coupaye, C. Carette, S. Czernichow and A.-S. Jannot. 'Risk-Benefit Balance Associated With Obstetric, Neonatal, and Child Outcomes After Metabolic and Bariatric Surgery'. In: *JAMA Surgery* (9th Nov. 2022). DOI: 10.1001/jamasurg.2022.5450. URL: https://hal.science/hal-03911629.

[65] C. Röver, M. Ursino, T. Friede and S. Zohar. 'A straightforward meta-analysis approach for oncology phase I dose-finding studies'. In: *Statistics in Medicine* 41.20 (10th Sept. 2022), pp. 3915–3940. DOI: 10.1002/sim.9484. URL: https://hal.inria.fr/hal-03864767.

[66] P. Sabatier, M. Wack, J. Pouchot, N. Danchin and A.-S. Jannot. 'A data-driven pipeline to extract potential adverse drug reactions through prescription, procedures and medical diagnoses analysis: application to a cohort study of 2,010 patients taking hydroxychloroquine with an 11-year follow-up'. In: *BMC Medical Research Methodology* 22.1 (8th June 2022), p. 166. DOI: 10.1186/s12874-022-01628-3. URL: https://www.hal.inserm.fr/inserm-03694380.

[67] T. Schmitz, M. Doret-Dion, L. Sentilhes, O. Parant, O. Claris, L. Renesme, J. Abbal, A. Girault, H. Torchin, M. Houllier et al. 'Neonatal outcomes for women at risk of preterm delivery given half dose versus full dose of antenatal betamethasone: a randomised, multicentre, double-blind, placebo-controlled, non-inferiority trial'. In: *The Lancet* 400.10352 (Aug. 2022), pp. 592–604. DOI: 10.1016/S0140-6736(22)01535-5. URL: https://hal.archives-ouvertes.fr/hal-03872891.

[68] D. M. Smadja, O. Bory, J.-L. Diehl, A. Mareau, N. Gendron, A.-S. Jannot and R. Chocron. 'Daily Monitoring of D-Dimer Allows Outcomes Prediction in COVID-19'. In: *TH Open : Companion Journal To Thrombosis and Haemostasis* 06.01 (17th Jan. 2022), e21–e25. DOI: 10.1055/a-1709-5441. URL: https://hal.archives-ouvertes.fr/hal-03560901.

[69] L. Soussand, M. Kuchenbuch, C. Messiaen, A. Sandrin, A.-S. Jannot and R. Nabbout. 'Impact of the COVID-19 pandemic on the care of rare and undiagnosed diseases patients in France: a longitudinal population-based study'. In: *Orphanet Journal of Rare Diseases* 17.1 (Dec. 2022), p. 430. DOI: 10.1186/s13023-022-02580-7. URL: https://hal.archives-ouvertes.fr/hal-03911619.

[70] A. Tarhini, C. Rives-Lange, A.-S. Jannot, C. Baratte, N. Beaupel, V. Guillet, S. Krivan, M. Le Gall, C. Carette, S. Czernichow, J.-M. Chevallier and T. Poghosyan. 'One-Anastomosis Gastric Bypass Revision for Gastroesophageal Reflux Disease: Long Versus Short Biliopancreatic Limb Roux-en-Y Gastric Bypass'. In: *Obesity Surgery* (17th Jan. 2022). DOI: 10.1007/s11695-022-05892-6. URL: https://hal.archives-ouvertes.fr/hal-03560922.

[71] M. von Tottleben, K. Grinyer, A. Arfa, L. Traore, D. Verdoy, S. Lim Choi Keung, I. Larranaga, M.-C. Jaulent, E. de Manuel Keenoy, M. Lilja, M. Beach, C. Marguerie, M. Yuksel, G. B. Laleci Erturkmen, G. Klein, P. Lindman, J. Mar, D. Kalra, T. Arvanitis and C.-C. R.-T. Collaborators. 'An Integrated Care Platform System (C3-Cloud) for Care Planning, Decision Support, and Empowerment of Patients With Multimorbidity: Protocol for a Technology Trial'. In: *JMIR Research Protocols* 11.7 (2022), e21994. DOI: 10.2196/21994. URL: https://hal.archives-ouvertes.fr/hal-03738964.

[72] M. Ursino, L. Biard and S. Chevret. 'DICE: A Bayesian model for early dose finding in phase I trials with multiple treatment courses'. In: *Biometrical Journal* (2022). DOI: 10.1002/bimj.202000369. URL: https://hal.archives-ouvertes.fr/hal-03873768.

[73] G. Weber, C. Hong, Z. Xia, N. Palmer, P. Avillach, S. L'yi, M. Keller, S. Murphy, A. Gutiérrez-Sacristán, C.-L. Bonzel et al. 'International comparisons of laboratory values from the 4CE collaborative to predict COVID-19 mortality'. In: *npj Digital Medicine* 5.1 (Dec. 2022), p. 74. DOI: 10.1038/s41746-022-00601-0. URL: https://hal.inria.fr/hal-03886991.

International peer-reviewed conferences

[74] S. Alsaidi, M. Couceiro, E. Marquer, S. Quennelle, A. Burgun, N. Garcelon and A. Coulet. 'An analogy based framework for patient-stay identification in healthcare'. In: ATA@ICCBR 2022 - Workshop Analogies: from Theory to Applications. Nancy, France, 12th Sept. 2022. URL: https://hal.inria.fr/hal-03763772.

[75] C. Chadebec, L. J. Vincent and S. Allassonnière. 'Pythae: Unifying Generative Autoencoders in Python - A Benchmarking Use Case'. In: Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmark. New Orleans, United States, 28th Nov. 2022. URL: https://hal.inria.fr/hal-03697439.

[76] P. Monnin and A. Coulet. 'Matching Pharmacogenomic Knowledge: Particularities, Results, and Perspectives'. In: 17th International Workshop on Ontology Matching : OM-2022. Virtual, China, 23rd Oct. 2022. URL: https://hal.inria.fr/hal-03835712.

Conferences without proceedings

[77] E. Marquer, S. Alsaidi, A. Decker, P.-A. Murena and M. Couceiro. 'A Deep Learning Approach to Solving Morphological Analogies'. In: *Lecture Notes in Computer Science*. 30th International Conference on Case-Based Reasoning (ICCBR2022). Nancy, France: Springer, 12th Sept. 2022, pp. 159–174. DOI: 10.1007/978-3-031-14923-8_11. URL: https://hal.inria.fr/hal-03660625.

[78] P. Monnin, C. Raïssi, A. Napoli and A. Coulet. 'Combiner plongements de graphes et clustering pour l'alignement de connaissances pharmacogénomiques'. In: Santé & IA 2022 - Journée Santé et Intelligence Artificielle (Evènement affilié à PFIA 2022). Saint-Etienne, France, 2022. URL: https://hal.inria.fr/hal-03754776.

Scientific book chapters

[79] A. Boukobza, M. Wack, A. Neuraz, D. Geromin, C. Badoual, A.-S. Bats, A. Burgun, M. Koual and R. Tsopra. 'Determining the Set of Items to Include in Breast Operative Reports, Using Clustering Algorithms on Retrospective Data Extracted from Clinical DataWarehouse'. In: *Advances in Informatics, Management and Technology in Healthcare*. Vol. 295. Studies in Health Technology and Informatics. IOS Press, 29th June 2022, pp. 45–48. DOI: 10.3233/SHTI220656. URL: https://hal.archives-ouvertes.fr/hal-03738958.

[80] C. Durchon, S. Vanderlan, A. Jegard, H. Saram, M. Falchi, F. Campeotto, L. Dupic, A. Burgun, B. Vivien and R. Tsopra. 'An Interactive Interface for Displaying Recommendations on Emergency Phone Triage in Pediatrics'. In: *Challenges of Trustable AI and Added-Value on Health*. Vol. 294. Studies in Health Technology and Informatics. IOS Press, 25th May 2022, pp. 430–434. DOI: 10.3233/SHTI220495. URL: https://hal.archives-ouvertes.fr/hal-03738948.

[81] C. Faviez, M. Vincent, N. Garcelon, C. Michot, G. Baujat, V. Cormier-Daire, S. Saunier, X. Chen and A. Burgun. 'Enriching UMLS-Based Phenotyping of Rare Diseases Using Deep-Learning: Evaluation on Jeune Syndrome'. In: *Challenges of Trustable AI and Added-Value on Health*. Vol. 294. Studies in Health Technology and Informatics. IOS Press; IOS Press, 25th May 2022, pp. 844–848. DOI: 10.3233/SHTI220604. URL: https://hal.science/hal-03790710.

[82] J. Lambert, A.-L. Leutenegger, A.-S. Jannot and A. Baudot. 'Tracking Temporal Clusters from Patient Networks'. In: *Challenges of Trustable AI and Added-Value on Health*. Studies in Health Technology and Informatics. IOS Press, 25th May 2022. DOI: 10.3233/SHTI220427. URL: https://hal.science/hal-03911967.

[83] J.-B. Lamy, A. Mouazer, K. Sedki and R. Tsopra. 'Translating the Observational Medical Outcomes Partnership – Common Data Model (OMOP-CDM) Electronic Health Records to an OWL Ontology'. In: *MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation*. Vol. 290. Studies in Health Technology and Informatics. IOS Press, 6th June 2022, pp. 76–80. DOI: 10.3233/SHTI220035. URL: https://hal.archives-ouvertes.fr/hal-03738953.

[84] E. Michot, J. Woo, L. Mouline, C. Sinnappan, A. Boukobza, F. Campeotto, L. Dupic, A. Burgun, B. Vivien and R. Tsopra. 'Towards a Clinical Decision Support System for Helping Medical Students in Emergency Call Centers'. In: *Challenges of Trustable AI and Added-Value on Health*. Vol. 294. Studies in Health Technology and Informatics. IOS Press, 25th May 2022, pp. 425–429. DOI: 10.3233/SHTI220494. URL: https://hal.archives-ouvertes.fr/hal-03738946.

[85] A. Mouazer, K. Sedki, R. Tsopra and J.-B. Lamy. 'Speak-PIM, Towards a Framework for the Automatic Detection of Potentially Inappropriate Prescriptions'. In: *Challenges of Trustable AI and Added-Value on Health*. Vol. 294. Studies in Health Technology and Informatics. IOS Press, 25th May 2022, pp. 460–464. DOI: 10.3233/SHTI220501. URL: https://hal.archives-ouvertes.fr/hal-03738951.

[86] V. Pohyer, D. Baudoin, L. Fournier and B. Rance. 'Extraction of Tumor Response Criteria in Semi-Structured Imaging Report'. In: *Challenges of Trustable AI and Added-Value on Health*. Studies in Health Technology and Informatics. IOS Press, 25th May 2022. DOI: 10.3233/SHTI220424. URL: https://hal.archives-ouvertes.fr/hal-03865993.

[87] A. Saadi, A. Rogier, A. Burgun and R. Tsopra. 'Design of an ontology-based triage system for painful patients'. In: *MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation*. Vol. 290. Studies in Health Technology and Informatics. IOS Press; IOS Press, 6th June 2022, pp. 81–85. DOI: 10.3233/SHTI220036. URL: https://hal.sorbonne-universite.fr/hal-03694355.

[88] P. Sabatier, J. Feydy and A.-S. Jannot. 'Accelerating High-Dimensional Temporal Modelling Using Graphics Processing Units for Pharmacovigilance Signal Detection on Real-Life Data'. In: *Challenges of Trustable AI and Added-Value on Health*. Studies in Health Technology and Informatics. IOS Press, 25th May 2022. DOI: 10.3233/SHTI220401. URL: https://hal.science/hal-03911970.

[89] K. Sedki, J.-B. Lamy and R. Tsopra. 'A Qualitative Method for Learning Medical Expert Reasoning'. In: *MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation*. Vol. 290. Studies in Health Technology and Informatics. IOS Press, 6th June 2022, pp. 645–649. DOI: 10.3233/SHTI220157. URL: https://hal.archives-ouvertes.fr/hal-03738956.

Doctoral dissertations and habilitation theses

[90] A. Challine. 'Use of the PMSI in the evaluation of morbi-mortality in oeso-gastric surgery'. Université Paris Cité, 12th Dec. 2022. URL: https://hal.archives-ouvertes.fr/tel-03903166.

Reports & preprints

[91] F. Gaudfernau, E. Blondiaux and S. Allassonnière. *Analysis of the Anatomical Variability of Fetal Brains with Corpus Callosum Agenesis*. 28th Jan. 2022. URL: https://hal.archives-ouvertes.fr/hal-03546165.

[92] F. Gaudfernau, E. Blondiaux, S. Allassonnière and E. Le Pennec. *Wavelet-Based Multiscale Initial Flow For Improved Atlas Estimation in the Large Diffeomorphic Deformation Model Framework*. 25th Mar. 2022. URL: https://hal.archives-ouvertes.fr/hal-03620367.

[93] F. Gaudfernau, A. Lefebvre, D.-A. Engemann, A. Pedoux, A. Bánki, F. Baillin, B. Landman, A. Maruani, F. Amsellem, T. Bourgeron, R. Delorme and G. Dumas. *Cortico-Cerebellar Neurodynamics during Social Interaction in Autism Spectrum Disorder*. 17th Oct. 2022. URL: https://hal.archives-ouvertes.fr/hal-03818265.

[94] J. Lambert, A.-L. Leutenegger, A.-S. Jannot and A. Baudot. *Tracking patient clusters over time enables to extract all the information available in the medico-administrative databases*. 4th Oct. 2022. DOI: 10.1101/2022.08.05.22278468. URL: https://hal.archives-ouvertes.fr/hal-03796086.

Other scientific publications

[95] M. Aubart, C.-J. Roux, C. Durrleman, C. Gins, M. Hully, M. Kossorotoff, C. Gitiaux, R. Levy, F. Moulin, A. Debray, Z. Belhadjer, E. Georget, T. Kom, P. Blanc, S. Wehbi, M. Mazeghrane, J. Tencer, V. Gajdos, S. Rouget, L. de Pontual, R. Basmaci, K. Yacouben, F. Angoulvant, M. Leruez-Ville, D. Sterlin, F. Rozenberg, M. Robert, S.-Y. Zhang, N. Boddaert and I. Desguerre. 'Neuroinflammatory Disease following Severe Acute Respiratory Syndrome Coronavirus 2 Infection in Children'. In:

The Journal of Pediatrics 247 (Aug. 2022), 22–28.e2. DOI: 10.1016/j.jpeds.2022.05.018. URL: https://hal.science/hal-03941125.

[96]   C. Bousquet and A. Coulet. *Supporting Diagnosis With Next-Generation Artificial Intelligence.* 12th Apr. 2022. DOI: 10.1001/jama.2022.2303. URL: https://hal.inria.fr/hal-03700431.

[97]   E. Bresso, P. Monnin, C. Bousquet, F.-É. Calvier, N. C. Ndiaye, N. Petitpain, M. Smaïl-Tabbone and A. Coulet. *Investigating ADR mechanisms with Explainable AI: a feasibility study with knowledge graph mining.* Rennes, France, 5th July 2022. URL: https://hal.inria.fr/hal-03754898.

[98]   R. Cohen, C. Levy, A. Rybak, F. Angoulvant, N. Ouldali and E. Grimprel. 'Immune debt: Recrudescence of disease and confirmation of a contested concept'. In: *Infectious Diseases Now* (Dec. 2022). DOI: 10.1016/j.idnow.2022.12.003. URL: https://hal.science/hal-03940278.

[99]   V. Goetz, D. D. Yang, F. Lacaille, M. Pelosi, F. Angoulvant, A. Brassier, J.-B. Arnoux, M. Schiff, C. Heilbronner, E. Salvador, D. Debray, M. Oualha, S. Renolleau, M. Girard and P. de Lonlay. 'What are the clues for an inherited metabolic disorder in Reye syndrome? A single Centre study of 58 children'. In: *Molecular Genetics and Metabolism* 135.4 (Apr. 2022), pp. 320–326. DOI: 10.1016/j.ymgme.2022.02.001. URL: https://hal.science/hal-03941178.

[100] A. Lansier, C. Bourillon, C.-A. Cuénod, E. Ragot, A. Follin, S. Hamada, O. Clément, P. Soyer and A.-S. Jannot. 'CT-based diagnostic algorithm to identify bowel and/or mesenteric injury in patients with blunt abdominal trauma'. In: *European Radiology* (4th Dec. 2022). DOI: 10.1007/s00330-022-09200-9. URL: https://hal.archives-ouvertes.fr/hal-03883935.

[101] P. Monnin, J. Legrand and A. Coulet. 'PGxCorpus and PGxLOD: two shared resources for knowledge management in pharmacogenomics'. In: JOBIM 2022 - Journées Ouvertes en Biologie, Informatique et Mathématiques. JOBIM 2022 Proceedings - Posters, Démos. Rennes, France, 5th July 2022. URL: https://hal.inria.fr/hal-03754888.

[102] N. Ouldali, H. Bagheri, F. Salvo, D. Antona, A. Pariente, C. Leblanc, M. Tebacher, J. Micallef, C. Levy, R. Cohen et al. 'Hyper inflammatory syndrome following COVID-19 mRNA vaccine in children: A national post-authorization pharmacovigilance study'. In: *The Lancet Regional Health - Europe* 17 (June 2022), p. 100393. DOI: 10.1016/j.lanepe.2022.100393. URL: https://hal.science/hal-03941143.

[103] A. Rybak, C. Levy, F. Angoulvant, A. Auvrignon, P. Gembara, K. Danis, S. Vaux, D. Levy-Bruhl, S. van der Werf, S. Béchet, S. Bonacorsi, Z. Assad, A. Lazzati, M. Michel, F. Kaguelidou, A. Faye, R. Cohen, E. Varon and N. Ouldali. 'Association of Nonpharmaceutical Interventions During the COVID-19 Pandemic With Invasive Pneumococcal Disease, Pneumococcal Carriage, and Respiratory Viral Infections Among Children in France'. In: *JAMA Network Open* 5.6 (1st June 2022), e2218959. DOI: 10.1001/jamanetworkopen.2022.18959. URL: https://hal.science/hal-03941118.

[104] D. D. Yang, M. Rio, C. Michot, N. Boddaert, W. Yacoub, N. Garcelon, B. Thierry, D. Bonnet, S. Rondeau, D. Herve, S. Guey, F. Angoulvant and V. Cormier-Daire. 'Natural history of Myhre syndrome'. In: *Orphanet Journal of Rare Diseases* 17.1 (Dec. 2022), p. 304. DOI: 10.1186/s13023-022-02447-x. URL: https://hal.science/hal-03941105.