

RESEARCH CENTRE

**Inria Center
at Université Côte d'Azur**

IN PARTNERSHIP WITH:

**INSERM, Université Paul-Valéry
Montpellier 3, Université de Montpellier**

2022

ACTIVITY REPORT

Project-Team

PREMEDICAL

Precision Medicine by Data Integration and Causal Learning

IN COLLABORATION WITH: Institut Desbrest d'Épidémiologie et de
Santé Publique (IDESP)

DOMAIN

Digital Health, Biology and Earth

THEME

**Computational Neuroscience and
Medicine**

Inria

Contents

Project-Team PREMEDICAL	1
1 Team members, visitors, external collaborators	2
2 Overall objectives	2
3 Research program	3
3.1 Research Axis 1: Personalized medicine by optimal prescription of treatment	3
3.2 Research axis 2: Personalized medicine by integration of different data sources	4
4 Application domains	4
5 Social and environmental responsibility	4
5.1 Impact of research results	4
6 Highlights of the year	5
6.1 Awards	5
6.2 Other highlights	5
7 New software and platforms	5
7.1 New software	5
7.1.1 factominer	5
7.1.2 missMDA	6
7.2 New platforms	6
8 New results	6
8.1 Generalization of clinical trials	6
8.2 Handling missing values	7
8.3 Uncertainty quantification with conformal prediction	7
9 Bilateral contracts and grants with industry	8
9.1 Bilateral contracts with industry	8
9.2 Bilateral Grants with Industry	9
10 Partnerships and cooperations	9
10.1 International initiatives	9
10.1.1 Participation in other International Programs	9
10.2 International research visitors	10
10.2.1 Visits of international scientists	10
10.2.2 Visits to international teams	10
10.3 European initiatives	10
10.3.1 Horizon Europe	10
10.3.2 Other european programs/initiatives	10
11 Dissemination	11
11.1 Promoting scientific activities	11
11.1.1 Scientific events: organisation	11
11.1.2 Scientific events: selection	11
11.1.3 Journal	11
11.1.4 Invited talks	11
11.1.5 Leadership within the scientific community	12
11.2 Teaching - Supervision - Juries	12
11.2.1 Teaching	12
11.2.2 Supervision	12
11.2.3 Juries	13
11.3 Popularization	13

11.3.1 Internal or external Inria responsibilities	13
11.3.2 Articles and contents	13
11.3.3 Interventions	13
12 Scientific production	13
12.1 Major publications	13
12.2 Publications of the year	13
12.3 Cited publications	17

Project-Team PREMEDICAL

Creation of the Project-Team: 2022 June 01

Keywords

Computer sciences and digital sciences

A6.1. – Methods in mathematical modeling

A9.2. – Machine learning

A9.6. – Decision support

Other research topics and application domains

B2. – Health

B2.2. – Physiology and diseases

B2.3. – Epidemiology

1 Team members, visitors, external collaborators

Research Scientist

- Julie Josse [Team leader, INRIA, Senior Researcher, HDR]

Faculty Members

- Pascal Demoly [UNIV MONTPELLIER, Professor, from Jun 2022]
- Pierre Lafaye De Micheaux [UNIV MONTPELLIER III, Associate Professor, from Jun 2022]
- Nicolas Molinari [UNIV MONTPELLIER, Professor, from Jun 2022]

PhD Students

- Marie Felicia Beclin [UNIV MONTPELLIER, from Jun 2022]
- Benedicte Colnet [inria]
- Maxime Fosset [UNIV MONTPELLIER, from Jun 2022]
- Margaux Zaffran [EDE, CIFRE]
- Pan Zhao [UNIV MONTPELLIER]

Administrative Assistant

- Annie Aliaga [INRIA]

2 Overall objectives

The objective of the PreMeDICAL team (**P**recision **M**edicine by **D**ata **I**ntegration and **C**ausal **L**earning) is to develop the next generation of methods/algorithms to extract knowledge from health data and improve the care of patients. More specifically, the aim is to develop learning tools for **personalized treatment effect prediction** and for **predicting outcome**, while **integrating different data sources to guide decisions** made by clinicians and authorities. PreMeDICAL has two research axes:

1. Personalized medicine by optimal prescription of treatment. We will develop causal inference techniques for **(dynamic) policy learning** (allocating the best treatment for each person at the right time), that handle missing values and leverage both RCTs and observational data. Using both data sources allow to **better design future RCTs** or to **launch a drug without running RCTs** and **in the longer term to rethink the evidence needed to bring treatments to the market** and to do so more quickly.
2. Personalized medicine by integration of different data sources. We will build predictive models for heterogeneous data: for instance given monitoring data in continuous time, images and clinical data what is the risk for an event to occur? Is it useful to have all the sources or do they provide the same information? We will additionally develop **solutions to handle missing values** in a supervised learning setting and **to improve the confidence of the outputs of the predictive models**.

The aim is to **push methodological innovation up to the stakeholders** (patients, clinicians, regulators, etc.). Consequently, beyond these methodological developments, innovative responses to the public health challenge posed by respiratory allergies are targeted. In addition to **leveraging machine learning algorithms and leveraging appropriate data, combining them with clinical expertise and existing recommendations** is necessary. Long- term aims are to have both a strong scientific and societal impact with a substantial impact on the quality of care for patients and major consequences for the medical profession by **providing a much earlier access to innovative solutions and more efficient treatment**

and care. With a successful proof of concept in the domain of allergies, by having clear **reproducible pipelines, methodologies, software** (by providing clinical decision making system tools) we could thereafter consider other pathologies (such as traumatology and oncology studied at IDESP). Hence, a joint team between Inria and Inserm provides a unique opportunity for trans-disciplinary research and collaboration bringing together mathematical, methodological, technological and medical expertise. The PreMeDiCaL team contributes to precision medicine (where the treatment/device is adapted on a patient basis) and to translational medicine which aims at bridging the gap between fundamental research and its practical use.

3 Research program

3.1 Research Axis 1: Personalized medicine by optimal prescription of treatment

Randomized controlled trials are considered the gold standard approach for assessing the causal effect (i.e., the treatment effect) of an intervention or a treatment on an outcome of interest. Indeed, the allocation of the treatment is under control, which implies that there is no confounding factors (the distribution of covariates for treated and control patients is asymptotically balanced) that could interfere with the treatment and simple estimators (such as the difference in mean effect between the treated and controls) can be used to consistently estimate the average treatment effect (ATE). However, RCTs can come with drawbacks. They can be expensive, take a long time to set up, and be compromised by insufficient sample size due to either recruitment difficulties or restrictive inclusion/exclusion criteria. These criteria can lead to a narrowly defined trial sample that differs markedly from the population potentially eligible for the treatment (distributional shift). Therefore, the findings from RCTs can lack generalizability (or *external validity*). This has been largely published in the field of respiratory and allergic diseases, see for instance (Pahus et al, 2019) which highlights that the population from RCTs represents less than 10% of the population that will receive treatments.

In contrast, there is an abundance of observational data, collected without systematically designed interventions. Such data can come from different sources: they can be collected from research sources (such as disease registries, cohorts, biobanks, epidemiological studies), or they can be routinely collected (through electronic health records, insurance claims, administrative databases, patients' App, etc). In that sense, observational data can be readily available, can include large samples representative of the target populations, and can be less costly than RCTs. To leverage observational data for treatment effect estimation in health domains, several laws built on studies by the USA Food and Drug Administration (FDA) encourage the use of "real world data" (RWD), defined as data "derived from sources other than randomized clinical trials", for regulatory decision making. Clinical evidence regarding the usage and potential benefits or risks of a medical product derived from the analysis of RWD is named Real World Evidence (RWE). The European Medicines Agency (EMA) is also a very active regulatory authority working with RWD to facilitate development and access to medicines. However, despite the large number of methods available to estimate the causal treatment effect from observational data such as matching, inverse probability weighting (IPW) or more recent doubly robust methods based on machine learning there are often concerns about the quality of these "big data" and causal claims. Indeed, building on observational data is still not consensual due to the lack of controlled experimental interventions, which opens the door to confounding biases (lack of *internal validity*).

Observational data and clinical trial data can provide different perspectives when evaluating an intervention or a medical treatment. Combining the information gathered from experimental and observational data is a promising avenue for medical research, because the knowledge acquired from integrative analyses could not be gathered from a single-source analysis alone. Three potential high impact applications of observational and clinical data are:

1. Predicting the effect of a treatment estimated on a RCT, on a new target population (generalization);
2. Comparing RCTs and RWE to validate observational methods;
3. Better estimation of heterogeneous treatment effects.

There is an abundant literature on bridging the findings from an RCT to a target population and combining both sources of information. Similar problems have been termed as *transportability*, and

data fusion and have connections to the *covariate shift/domain generalization* problem in ML. Colnet et al, 2020. reviewed the methods to generalize the treatment effect while integrating the distributional shift (IPSW, g-formula, AIPSW, calibration weighting, etc.) (a), or improve the estimate of the conditional average treatment effect (CATE, i.e. heterogeneous effect) while correcting for confounding factors not measured in the observational study (c). However, these methods have many shortcomings and there are still many challenges to address. We provide below examples of methodological locks we will overcome.

- Handling missing values and unmeasured covariates with multi-source data;
- Transfert Learning of optimal individualized treatment regimes with right-censored survival data;
- Policy learning and dynamic treatment policy with missing values.

3.2 Research axis 2: Personalized medicine by integration of different data sources

Integrating heterogeneous data (time series, images, text, numerical or categorical data) potentially from different centers to establish predictive models involves many obstacles. In the case where a patient is described by several sources, problems of high dimensionality are exacerbated. We will start by the question of the links between the sources before tackling some challenges posed by missing values:

- Relationship between the different sources;
- (Informative) missing values in time series and structured by blocks;
- Conformal prediction with missing values;
- Federative learning with missing values.

4 Application domains

The first application domain of PreMeDICaL is **respiratory diseases** and in particular **Asthma**. For more than 30 years, there has been an increase in a number of chronic non-communicable diseases (NCD), such as asthma and allergies, respiratory diseases. Allergies are the fourth most common chronic disease in the world. The World Health Organization (WHO) predicts that by 2050, one in two people in the world will suffer from allergies. In France, the number of people suffering from allergies has doubled in 20 years, particularly among children and young people. Although the expression of these diseases results from the interaction between the genetic background and the environment, especially through epigenetic mechanisms, their sudden increase is solely due to the environmental changes that occurred in the last decades because of the Western lifestyle, the genetic heritage requiring centuries to change. A full understanding of the complexity of chronic NCD prompts researchers to analyse large data utilizing proper markers and tools (e.g., biological, clinical, behavioural, economic, social, demographic, environmental data, patient experience, patient social networks) in an etiological and evaluative way to determine phenotypical patients' pathways, explain their impacts, their causes, their influences, prevent them and improve their prognosis. Integrating these different sources of information, collected by several actors (healthcare professionals, public authorities or patients themselves), thus offer new opportunities to design personalized solutions by adapting treatment to the patient and the organizational context, leading to improved patient care and prevention policies.

With a successful proof of concept in the domain of allergies, by having clear **reproducible pipelines, methodologies, software** we will thereafter consider other pathologies (such as traumatology and oncology studied at IDESP).

5 Social and environmental responsibility

5.1 Impact of research results

From a methodological point of view, the aim is to improve and develop new statistical and ML methods for establishing evidence on the efficiency of treatment by data enrichment (data fusion), by taking the

example of AIT in respiratory diseases. An important output of this research is that these methodological works have a concrete impact on designing future clinical trials and that the new methodology will be supported by regulatory authorities. Indeed, exploiting both RCTs and observational data serve different purposes such as prediction of the treatment effect on new populations, increasing the generalization of clinical trials (so that they are more representative of the patient population who may benefit from the treatment) and also defining new inclusion criteria (because we identify subgroups who can benefit from treatment). This research is part of the PEPR project "Next methodological challenges in clinical trials in the era of digital health".

From a technological point of view, the aim is to provide software (starting with open access) for these methods to be applied in practice by studies stakeholders, clinicians and the clinical trial community.

From the clinical and patients point of view, the different projects aim at quantifying the clinical benefit of treatment (over time), taking into account all patient characteristics, and provide useful clinical prognosis tools allowing clinicians to optimally treat every patient. The aim is to give patients better care and early access to innovation. In addition, these works can lead to a better adoption by the medical community of certain (advanced) techniques used to estimate the effects of treatment on patients (by comparing the results obtained in an RCT with the RWE).

From a public-health point of view, the aim is to guide decisions made by investigators, sponsors and authorities. Better trials' designs may also have an important impact in terms of cost reduction. Finally, we aim at having a significant impact in the field of allergy treatments providing new knowledge that may change guidelines and practice.

6 Highlights of the year

6.1 Awards

Margaux Zaffran received the **Séphora Berrebi Women in Advanced Mathematics & Computer Science Scholarship** which aims at encouraging active involvement of young women in scientific research, especially in the mathematics and computer science areas. They recognized her PhD work on conformal prediction and the grant allowed her to spend a research stay at the Technion Israel.

6.2 Other highlights

- Pascal Demoly renew the direction (with Luciana Kase Tanno) of the WHO Collaborating Center for "Scientific Support for Classifications" (2018-2022 and 2022-2026)
- Pascal Demoly was elected President of the French Society of Allergology (2022-2024)
- Pascal Demoly was appointed to the strategic and scientific orientation council of the recent ExposUM Institute (created in 2022 as part of the PIA4 Excellencies project)

7 New software and platforms

7.1 New software

7.1.1 factominer

Keywords: Dimensionality reduction, PCA, Text mining, Clustering

Functional Description: The FactoMineR package is dedicated to performing principal components methods to explore, sum-up and visualize data. Dimensionality reduction methods include PCA, correspondence analysis (CA) for count data such as documents-words data, multiple correspondence analysis (MCA) for categorical data such as survey data, factorial analysis of mixed data (FAMD) for both types of variables as well as methods for groups of variables, of individuals (multiple factorial analysis, MFA), for hierarchy . . .

References: https://husson.github.io/MOOC_AnaDo/index.html <https://husson.github.io/MOOC.html#PCAcourse>

URL: http://factominer.free.fr/index_fr.html

Contact: Julie Josse

Partner: AGROCAMPUS

7.1.2 missMDA

Keyword: Missing data

Functional Description: The missMDA package is dedicated to missing values in and with Multivariate Data Analysis. It allows one to apply PCA, MCA, FAMD and MFA on incomplete data. It performs single and multiple imputation for continuous, categorical and mixed data based on principal components methods

URL: <http://factominer.free.fr/missMDA/index.html>

Contact: Julie Josse

Partner: AGROCAMPUS

7.2 New platforms

Causal inference taskview: to list and organize all the R packages on causal inference

Participants: Julie Josse, Pan Zhao.

8 New results

8.1 Generalization of clinical trials

Participants: Benedicte Colnet, Julie Josse.

Context

RCTs are the current gold-standard to empirically measure a causal effect of a given intervention on an outcome. But more recently, concerns have been raised on the limited scope of RCTs: stringent eligibility criteria, unrealistic real-world compliance, short timeframe, limited sample size, etc. Such limitations threaten the external validity of RCT studies to other situations or populations [51]. The usage of complementary non-randomized data, referred to as *observational* or from *real world*, brings promises as additional sources of evidence, in particular combined to trials. **Transportability** (also known as **generalization**, *recoverability from sampling bias*, or *data-fusion* [52, 49]) allows to generalize or transport the trial findings toward a target population of interest, potentially subject to a covariate distributional shift.

Results: Reweighting the trial - Publications [11, 42]

RCT and observational data are seldom acquired as part of a homogeneous effort. As a result, they come with different covariates. Restricting the analysis to the shared covariates raises the risk of omitting an important one leading to identifiabilities issues. This problem is reminiscent of unobserved confounding in causal inference with one observational data. In [1], we suggest a **sensitivity analysis** to handle cases where such covariates (namely treatment effect modifiers that are shifted between the two sets when

studying risk difference) are missing in one or both sets. We also completed proofs on the consistency of generalization estimators that use either **weighting (Inverse Propensity of Sampling Weighting, IPSW)**, **outcome modeling**, or combine the two in **doubly robust approaches with Augmented IPSW (AIPSW)**. We further analysed the IPSW estimator, which consists of re-weighting the trial so that it resembles the observational sample, in [42]. In particular, we established **finite sample bias and variance** (the literature mostly focuses on asymptotic results) and upper bound on the risk of different versions of the estimator: oracle, semi-oracle, etc. This work can lead to practical recommendations in terms of data collection (e.g., doubling the size of the observational data leads to a smaller asymptotic variance than doubling the size of the trial).

Results: Optimal policy for survival data

Participants: Pan Zhao, Julie Josse.

The optimal individualized treatment regime (ITR) learned from a source population, due to covariate shift, may not generalize well to the target population that we aim to apply the ITR on. We propose a transfer learning framework, where covariate information from the target population is available, for ITR estimation with heterogeneous populations and right-censored survival data, which is common in clinical studies and motivated by our medical application.

We characterize the efficient influence function (EIF) and propose a doubly robust estimator of the targeted value function, which accommodates a broad class of functionals of survival distributions. For a pre-specified class of ITRs, we establish the $N^{-1/3}$ rate of convergence for the estimated parameter indexing the optimal ITR. Based on the Neyman orthogonality of the EIF, we also propose a cross-fitting procedure and show that the proposed optimal value estimator is consistent and asymptotically normal with flexible machine learning methods for nuisance parameter estimation.

8.2 Handling missing values

Results: Handling MNAR data in clustering

Participants: Julie Josse.

Model-based unsupervised learning, as any learning task, stalls as soon as missing data occurs. This is even more true when the missing data are informative, or said missing not at random (MNAR). In this paper, we propose model-based clustering algorithms designed to handle very 5 general types of missing data, including MNAR data. To do so, we introduce a mixture model for different types of data (continuous, count, categorical and mixed) to jointly model the data distribution and the MNAR mechanism, remaining vigilant to the degrees of freedom of each. Eight different MNAR models which depend on the class membership and/or on the values of the missing variables themselves are proposed. For a particular type of MNAR models, for which 10 the missingness depends on the class membership, we show that the statistical inference can be carried out on the data matrix concatenated with the missing mask considering a MAR mechanism instead; this specifically underlines the versatility of the studied MNAR models. Then, we establish sufficient conditions for identifiability of parameters of both the data distribution and the mechanism. Regardless of the type of data and the mechanism, we propose to perform 15 clustering using EM or stochastic EM algorithms specially developed for the purpose. Finally, we assess the numerical performances of the proposed methods on synthetic data and on the real medical registry TraumaBase R as well.

8.3 Uncertainty quantification with conformal prediction

Participants: Julie Josse, Margaux Zaffran.

Context

Most statistical learning and artificial intelligence methodologies provide point predictions, without any indication of the degree of confidence that can be given to these predictions (i.e. without predictive intervals). This lack of uncertainty quantification of predictive models is a major barrier to the adoption of powerful machine learning methods by society. Probabilistic forecasts, i.e. predicting the entire distribution probability and not only the conditional expectation, could partially tackle this issue but they are only valid asymptotically, require strong assumptions on the data (e.g. normality) or/and are model-dependent. The emergent field of conformal prediction (CP) [53, 48, 47] is a promising framework for distribution-free uncertainty quantification. It is a general procedure to build predictive intervals for any predictive model (including black-box methods such as deep learning), which are valid (i.e. achieve nominal marginal coverage), in finite sample, and without assumption on the data generation process except the exchangeability. This is extremely promising for decision support tools in critical applications: healthcare, autonomous driving, etc. An extension of CP (Conformalized Quantile Regression, [50]) was used to predict the U.S. presidential elections (2020) by the Washington Post.

Results: Conformal prediction for time series - publication [3]

Given the non-exchangeability of time series data, CP can not be applied as such to this framework. To achieve this task, we study and extend Adaptive Conformal Inference (ACI) [46] in the context of time series with general dependency. ACI is a method designed to handle an online setting, with distributional shift. It relies on using an adaptive miscoverage rate, that is updated according to previous performances and to a hyper-parameter, playing the role of a learning rate. First, we study theoretically, using Markov Chain theory, the impact of the learning rate on the length of the predictive intervals, in the exchangeable and auto-regressive case, in order to describe not only the validity but also the efficiency of ACI. This is hardly useful in practice: the optimal learning rate depends on the unknown data distribution. This is why we introduce AgACI, a parameter-free method using online expert aggregation [45]. Finally, we compare ACI, AgACI and other methods slightly adapted to time series, on extensive synthetic experiments. These experiments highlight that AgACI achieves good performances in terms of validity and efficiency. To allow for better benchmarking of existing and new methods, we provide implementations in Python of (all) the described methods and a complete pipeline of analysis on GitHub.

Results: Conformal prediction with missing values

Uncertainty quantification has not been addressed with missing values. In the finite-sample regime, we show that, for almost all imputations and missing values mechanisms, the imputed data set is exchangeable. Thus, CP properties still hold and marginal guarantees are met. Nevertheless, we emphasize that the average coverage varies depending on the pattern of missing values: it tends to construct prediction intervals that often under-cover the response conditional on a given missing pattern. After theoretically studying the case of a linear model, we propose a methodology, *missing data augmentation*, to achieve approximate conditional guarantees conditional on the 2^d patterns of missing values, where d is the data dimension.

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

Participants: Nicolas Molinari, Pascal Demoly.

- Participation to the Fondation TEZOS (Vigicard digital health card project) with the startup CodInsight
- Co-creation of the startup **AdviceMedica** (collective intelligence for solving complex cases in medicine)
- Title: Apprentissage et modélisation statistique en médecine du sommeil : du diagnostic au traitement
- Company: **groupe Adène**
- Duration: May 2020 - May 2023

9.2 Bilateral Grants with Industry

Participants: Julie Josse, Pascal Demoly.

- Title: Combining RCT and observational data. Educational Grant
- Company: Allergologisk Laboratorium Kobenhavn (ALK)
- Duration: Sept 2022 - Sept 2023

Participants: Nicolas Molinari.

- Title: Etude IRIS - Efficacité en vie Réelle du dispositif de télésurveillance médicale Chronic Care Connect™ chez les patients atteints d'Insuffisance cardiaque chronique
- Company: IQVIA
- Duration: Oct 2022 - Oct 2023
- Title: étude de vie réelle analysant la mortalité des patients BPCO sévères
- Company: Sanofi
- Duration: Nov 2022 - Dec 2023
- Title: étude « Home-Care SIMEOX »,
- Company: Agir à Dom
- Duration: Nov 2019 - Nov 2023

10 Partnerships and cooperations

10.1 International initiatives

10.1.1 Participation in other International Programs

International medical trials with AB science: member of the Data Safety Monitoring Board (DSMB) which will oversee the safety of study AB20001 entitled: "A Randomized, Open-label Phase 2 Clinical Trial to Evaluate the Safety and Efficacy of Masitinib combined with Isoquercetin, and Best Supportive Care in Hospitalized Patients with Moderate and Severe COVID-19."

Participants: Nicolas Molinari.

10.2 International research visitors

10.2.1 Visits of international scientists

Other international visits to the team

Mats Julius Stensrud

Status researcher (Ass. Professor)

Institution of origin: EPFL

Country: Switzerland

Dates: June 1

Context of the visit: Work on dynamic treatment regime

Mobility program/type of mobility: research stay

Michael Elliott

Status researcher (Professor)

Institution of origin: University of Michigan

Country: USA

Dates: June 3

Context of the visit: Talk on combining experimental and non experimental data

Mobility program/type of mobility: lecture

10.2.2 Visits to international teams

Research stays abroad

- 3 months visit at Departments of Electrical Engineering and of Computer Science at the Technion—Israel Institute of Technology to work with **Yanniv Romano** on conformal prediction - Margaux Zaffran
- 3 months visit at the Statistic Department of Stanford University hosted by **Trevor Hastie** - Bénédicte Colnet

10.3 European initiatives

10.3.1 Horizon Europe

HORIZONHLTH-2021-ENVHLTH-02 SynAir-G (500k€)

Participants: Pascal Demoly.

10.3.2 Other european programs/initiatives

- Participation to the consortium ECRHS (cohort European Community Respiratory Health Survey) with 2 ANR
- Participation to the ENDA (European Network for Drug Allergy)

Participants: Pascal Demoly.

11 Dissemination

Dissemination from the creation of the PreMeDiCaL team in June 2022 to December 2022

11.1 Promoting scientific activities

11.1.1 Scientific events: organisation

General chair, scientific chair

- [Journée des jeunes statisticiens et probabilistes](#) - Margaux Zaffran

11.1.2 Scientific events: selection

Chair of conference program committees

- [Journée des jeunes statisticiens et probabilistes](#) - Margaux Zaffran

Member of the conference program committees

- Journée de la société Française de Statistique, Bruxelles, Belgium 2023 - Julie Josse
- Statlearn, Montpellier, France, April 2023 - Julie Josse
- IMS International Conference on Data Science, Florence, Italy, December 2022 - Julie Josse
- Mathematical Methods of Modern Statistics 3. Luminy, France, June 2022 - Julie Josse

11.1.3 Journal

Member of the editorial boards

- Journal of Statistical Software - Pierre Lafaye de Micheaux
- Journal allergy clin immunology in practice, Allergy Asthma and Clin Immunol, Rev fr d'Allergologie - Pascal Demoly
- European Respiratory Journal - Nicolas Molinari

Reviewer - reviewing activities Annals of Statistics, Journal of Statistical Computations and Simulations, Journal of Statistical Software, WAO Journal, Allergy, Clin Exp Allergy

11.1.4 Invited talks

- Online Causal Inference Seminar, Nov., 2022 - Julie Josse
- Celebrating causal inference in medicine and public health, Oct. 2022, Ghent University, Belgium (before the Rousseeuw prize) - Julie Josse
- Exposome Symposium, Oct. 2022, Montpellier, France - Julie Josse
- AutoML conference (with ICML conference), Online, July 2022, Baltimore, USA - Julie Josse
- IMS 2022, Talk in session on Modern Approaches to Missing Data. Online, June 2022, London, UK - Julie Josse
- International Meeting on Statistical Methods in Biopharmacy, Sep. 2022 - Benedicte Colnet
- Traumabase scientific comitee and Capgemini, Dec. 2022 - Benedicte Colnet
- Causal Tau team, Oct. 2022 - Benedicte Colnet

- Yaniv Romano's Group Meeting, Haifa, Israel, July 2022 - Margaux Zaffran
- Mathematical Methods of Modern Statistics 3, Marseille, France, July 2022 - Margaux Zaffran
- Journées de statistiques, Lyon, France, June 2022 - Margaux Zaffran
- Journées de statistiques, Lyon, France, June 2022 - Pan Zhao

11.1.5 Leadership within the scientific community

- **Traumatrix** project: Since 2016, Julie Josse coordinate the Traumatrix project in partnership with the École Polytechnique (IPP), the CNRS, EHESS (École des Hautes Études en Sciences Sociales), Inria, the traumabase group (30 hospitals) and APHP (les hopitaux de Paris) to improve the management of polytrauma patients by leveraging observational data. The project is funded by ARS (Agence Régionale de Santé) and by the skill sponsorship of Capgemini Invent (10 data scientists, project managers, data managers). The project is mature and we have launched prospective studies with a mobile application in the ambulances that collects data and predicts the risk of the patient going into hemorrhagic shock.
- Vice-president of the **Young Statisticians Group** of the French Statistical Society - Margaux Zaffran

11.2 Teaching - Supervision - Juries

11.2.1 Teaching

- Master: ESEEC, 19.5eqTD, Les outils statistiques du diagnostic (modélisation et statistiques économiques et sociales), Université Paul Valéry, Montpellier - Pierre Lafaye De Micheaux
- Master: ESEEC, 17.5eqTD, Analyse de Données Multidimensionnelles SHS, Université Paul Valéry, Montpellier - Pierre Lafaye De Micheaux
- Master: MIASHS, 22.5eqTD, Statistique et probabilités bivariées, Université Paul Valéry, Montpellier - Pierre Lafaye De Micheaux
- Bachelor: MIASHS, 26eqTD, Science des données 2, Université Paul Valéry, Montpellier - Pierre Lafaye De Micheaux
- Master: MIASHS, head of projects of " Travaux d'Études et de Recherche (TER)", Université Paul Valéry, Montpellier - Pierre Lafaye De Micheaux
- Master: MIASHS, head of "Marathon du Web", Université Paul Valéry, Montpellier - Pierre Lafaye De Micheaux
- Master: Institut de formation en masso-kinésithérapie, 9eqTD, statistics, Montpellier - Nicolas Molinari
- Master: Institut de formation en masso-kinésithérapie, head of the program, statistics, Montpellier - Nicolas Molinari
- Master: EDSB « Epidémiologie, Données de Santé, Biostatistique », head of « Grands enjeux en santé », Université de Montpellier - Pascal Demoly

11.2.2 Supervision

- PhD in progress: Marie Felicia Beclin - Pierre Lafaye De Micheaux and Nicolas Molinari
- PhD in progress: Maxime Fosset - Julie Josse and Nicolas Molinari
- PhD in progress: Margaux Zaffran - Julie Josse
- PhD in progress: Pan Zhao - Julie Josse
- Master in progress: Mame Fatou Gueye - Pierre Lafaye De Micheaux

11.2.3 Juries

- PhD [with report] Elise Dumas, France - Julie Josse
- Comité de Sélection to hire two permanent researchers "Chargé de Recherche" (Inrae), Apprentissage statistique pour les sciences du vivant et de l'environnement - Julie Josse

11.3 Popularization

11.3.1 Internal or external Inria responsibilities

- Member of "Comité du suivi doctoral" Inria Sophia, to allocate PhD grants extensions, etc - Julie Josse
- Member of the board for Region Occitanie to allocate PhD/postdoc grants in AI for health - Julie Josse

11.3.2 Articles and contents

- [Interview for Montpellier University](#) - Julie Josse

11.3.3 Interventions

- Women in Machine Learning profile highlight invitation - Julie Josse

12 Scientific production

12.1 Major publications

- [1] B. Colnet, J. Josse, G. Varoquaux and E. Scornet. 'Causal effect on a target population: a sensitivity analysis to handle missing covariates'. In: *Journal of Causal Inference* 10.1 (15th Sept. 2022), pp. 372–414. DOI: [10.1515/jci-2021-0059](https://doi.org/10.1515/jci-2021-0059). URL: <https://hal.science/hal-03473691>.
- [2] I. Mayer, A. Sportisse, N. Tierney, N. Vialaneix and J. Josse. 'R-miss-tastic: a unified platform for missing values methods and workflows'. In: *The R Journal* (30th July 2022). URL: <https://hal.archives-ouvertes.fr/hal-02879337>.
- [3] M. Zaffran, O. Féron, Y. Goude, A. Dieuleveut and J. Josse. 'Adaptive Conformal Predictions for Time Series'. In: ICML 2022 - International Conference on Machine Learning. Baltimore, United States, 17th July 2022. URL: <https://hal.archives-ouvertes.fr/hal-03573934>.

12.2 Publications of the year

International journals

- [4] A. Agustí, E. Rapsomaniki, R. Beasley, R. Hughes, H. Müllerová, A. Papi, I. Pavord, M. van den Berge, R. Faner, G. Benhabib et al. 'Treatable traits in the NOVELTY study'. In: *Respirology* 27.11 (Nov. 2022), pp. 929–940. DOI: [10.1111/resp.14325](https://doi.org/10.1111/resp.14325). URL: <https://hal.science/hal-03936394>.
- [5] D. Bessis, S. Trouillet-Assant, L.-P. Secco, N. Bardin, B. Blanc, V. Blatière, C. Chable-Bessia, C. Delfour, C. Girard, J.-C. Richard, N. Gros, V. Le Moing, N. Molinari, V. Pallure, A. Pisoni, N. Raison-Peyron, E. Reynaud, É. Schwob, R. Pescarmona, Q. Samaran, M. Willems, T. Vincent, M. T. Sofonea, A. Belot and É. Tuailon. 'COVID-19 pandemic-associated chilblains: more links for SARS-CoV-2 and less evidence for high interferon type I systemic response'. In: *British Journal of Dermatology* 187.6 (1st Dec. 2022), pp. 1032–1035. DOI: [10.1111/bjd.21820](https://doi.org/10.1111/bjd.21820). URL: <https://hal.science/hal-03930679>.

- [6] A. Blet, J. Mcneil, J. Josse, B. Cholley, R. Cinotti, G. Cotter, A. Dauvergne, B. Davison, K. Duarte, J. Duranteau, M.-C. Fournier, E. Gayat, S. Jaber, S. Lasocki, T. Merklng, K. Peoc'h, I. Mayer, M. Sadoune, P.-F. Laterre, R. Sonnevile, L. Ware, A. Mebazaa and A. Kimmoun. 'Association between in-ICU red blood cells transfusion and 1-year mortality in ICU survivors'. In: *Critical Care* 26.1 (Dec. 2022), p. 307. DOI: [10.1186/s13054-022-04171-1](https://doi.org/10.1186/s13054-022-04171-1). URL: <https://hal.archives-ouvertes.fr/hal-03806977>.
- [7] A. Bourdin, S. Bommart, G. Marin, I. Vachier, A. S. Gamez, E. Ahmed, C. Suehs and N. Molinari. 'Obesity in women with asthma: baseline disadvantage plus greater small-airway responsiveness'. In: *Allergy* (2022). DOI: [10.1111/all.15509](https://doi.org/10.1111/all.15509). URL: <https://hal.science/hal-03772214>.
- [8] D. Caimmi and P. Demoly. 'A review of allergen immunotherapy in asthma'. In: *Allergy and Asthma Proceedings* 43.4 (1st July 2022), pp. 310–313. DOI: [10.2500/aap.2022.43.210113](https://doi.org/10.2500/aap.2022.43.210113). URL: <https://hal.science/hal-03936077>.
- [9] B. Cardoso, T. Moscoso, M. Morais-Almeida, P. Demoly and L. Tanno. 'Economic burden of drug-induced anaphylaxis: what can we do better?'. In: *Current Opinion in Allergy and Clinical Immunology* 22.4 (2022), pp. 234–241. DOI: [10.1097/ACI.0000000000000836](https://doi.org/10.1097/ACI.0000000000000836). URL: <https://hal.umontpellier.fr/hal-03779292>.
- [10] A. M. Chiriach, R. Zhu, L. Izquierdo, N. Molinari and P. Demoly. 'The minimal important difference of the allergic rhinitis control test'. In: *The Journal of Allergy and Clinical Immunology: In Practice* 10.7 (July 2022), 1924–1926.e2. DOI: [10.1016/j.jaip.2022.02.050](https://doi.org/10.1016/j.jaip.2022.02.050). URL: <https://hal.science/hal-03778482>.
- [11] B. Colnet, J. Josse, G. Varoquaux and E. Scornet. 'Causal effect on a target population: a sensitivity analysis to handle missing covariates'. In: *Journal of Causal Inference* 10.1 (15th Sept. 2022), pp. 372–414. DOI: [10.1515/jci-2021-0059](https://doi.org/10.1515/jci-2021-0059). URL: <https://hal.science/hal-03473691>.
- [12] E. Debourdeau, G. Planells, C. Chamard, D. Touboul, M. Villain, P. Demoly, F. Babeau, P. Fournie and V. Daien. 'New Keratoconus Risk Factors: A Cross-Sectional Case—Control Study'. In: *Journal of Ophthalmology* 2022 (28th Sept. 2022), p. 6605771. DOI: [10.1155/2022/6605771](https://doi.org/10.1155/2022/6605771). URL: <https://hal.science/hal-03930828>.
- [13] P. Demoly, C. Leroyer, E. Serrano, A. Le Maux, G. Magnier and A. Chartier. 'The SQ HDM SLIT-Tablet is safe and well tolerated in patients with House Dust Mite allergic rhinitis with or without asthma: A "real-life" French study'. In: *Clinical and Translational Allergy* 12.3 (2022), e12129. DOI: [10.1002/ctt2.12129](https://doi.org/10.1002/ctt2.12129). URL: <https://hal.umontpellier.fr/hal-03641065>.
- [14] P. Descloux, C. Boyer, J. Josse, A. Sportisse and S. Sardy. 'Robust Lasso-Zero for sparse corruption and model selection with missing covariates'. In: *Scandinavian Journal of Statistics* (2022). URL: <https://hal.archives-ouvertes.fr/hal-02569696>.
- [15] A. Desgagné, P. Lafaye de Micheaux and F. Ouimet. 'Goodness-of-fit tests for Laplace, Gaussian and exponential power distributions based on λ -th power skewness and kurtosis'. In: *Statistics* (24th Nov. 2022), pp. 1–29. DOI: [10.1080/02331888.2022.2144859](https://doi.org/10.1080/02331888.2022.2144859). URL: <https://hal.science/hal-03929998>.
- [16] A. Fillard, A. Licari, N. Molinari, G. Marseglia, P. Demoly and D. Caimmi. 'Sensitivity of FEV1 and Clinical Parameters in Children With a Suspected Asthma Diagnosis'. In: *The Journal of Allergy and Clinical Immunology: In Practice* 11.1 (Jan. 2023), pp. 238–247. DOI: [10.1016/j.jaip.2022.10.011](https://doi.org/10.1016/j.jaip.2022.10.011). URL: <https://hal.umontpellier.fr/hal-03930657>.
- [17] T. Garaix, S. Gaubert, J. Josse, N. Vayatis and A. Véber. 'Decision-making tools for healthcare structures in times of pandemic'. In: *Anaesthesia Critical Care & Pain Medicine* 41.2 (Mar. 2022), p. 101052. DOI: [10.1016/j.accpm.2022.101052](https://doi.org/10.1016/j.accpm.2022.101052). URL: <https://hal.archives-ouvertes.fr/hal-03606559>.
- [18] T. Gauss, J. Richards, C. Tortù, F.-X. Ageron, S. Hamada, J. Josse, F. Husson, A. Harrois, T. Scalea, V. Vivant et al. 'Association of Early Norepinephrine Administration With 24-Hour Mortality Among Patients With Blunt Trauma and Hemorrhagic Shock'. In: *JAMA Network Open* 5.10 (3rd Oct. 2022), e2234258. DOI: [10.1001/jamanetworkopen.2022.34258](https://doi.org/10.1001/jamanetworkopen.2022.34258). URL: <https://hal.archives-ouvertes.fr/hal-03807657>.

- [19] T. Godet, A. de Jong, C. Garin, R. Guérin, B. Rieu, L. Borao, B. Pereira, N. Molinari, J.-E. Bazin, M. Jabaudon, G. Chanques, E. Futier and S. Jaber. 'Impact of Macintosh blade size on endotracheal intubation success in intensive care units: a retrospective multicenter observational MacSize-ICU study'. In: *Intensive Care Medicine* (2022). DOI: [10.1007/s00134-022-06832-9](https://doi.org/10.1007/s00134-022-06832-9). URL: <https://hal.science/hal-03752646>.
- [20] C. Hédon, F. Schnell, P. Sosner, F. Chagué, I. Schuster, M. Julia, A. Duparc, J.-M. Guy, N. Molinari, L. Michaux, F. Cransac and S. Cade. 'Cardiac screening before returning to elite sport after SARS-CoV-2 infection'. In: *Archives of cardiovascular diseases* 115.11 (Nov. 2022), pp. 562–570. DOI: [10.1016/j.acvd.2022.06.005](https://doi.org/10.1016/j.acvd.2022.06.005). URL: <https://hal.science/hal-03930065>.
- [21] L. Klimek, R. Brehler, R. Mösges, P. Demoly, J. Mullol, D. Wang, R. O'Hehir, A. Didier, M. Kopp, C. Bos and E. Karagiannis. 'Update about Oralair® as a treatment for grass pollen allergic rhinitis'. In: *Human Vaccines & Immunotherapeutics* 18.5 (2022), p. 2066424. DOI: [10.1080/21645515.2022.2066424](https://doi.org/10.1080/21645515.2022.2066424). URL: <https://hal.science/hal-03935968>.
- [22] C. Krantz, S. Accordini, K. Alving, A. G. Corsico, P. M. Demoly, D. S. Ferreira, B. Forsberg, J. Garcia-Aymerich, T. Gislason, J. Heinrich, R. Jögi, A. Johannessen, B. Leynaert, A. Marcon, J. Martínez-Moratalla Rovira, E. Nerpin, D. A. Nowak, A. C. Olin, M. Olivieri, A. R. Pereira-Vega, C. Raheison-Semjen, F. Real, T. Sigsgaard, G. Squillacioti, C. Janson and A. Malinovsky. 'Cross-sectional study on exhaled nitric oxide in relation to upper airway inflammatory disorders with regard to asthma and perennial sensitization'. In: *Clinical and Experimental Allergy* 52.2 (2022), pp. 297–311. DOI: [10.1111/cea.14019](https://doi.org/10.1111/cea.14019). URL: <https://hal.science/hal-03604113>.
- [23] E. Macy, A. Trautmann, A. M. Chiriach, P. Demoly and E. Phillips. 'Advances in the Understanding of Drug Hypersensitivity: 2012 Through 2022'. In: *The Journal of Allergy and Clinical Immunology: In Practice* 11.1 (Jan. 2023), pp. 80–91. DOI: [10.1016/j.jaip.2022.10.025](https://doi.org/10.1016/j.jaip.2022.10.025). URL: <https://hal.umontpellier.fr/hal-03930648>.
- [24] I. Mayer, A. Sportisse, N. Tierney, N. Vialaneix and J. Josse. 'R-miss-tastic: a unified platform for missing values methods and workflows'. In: *The R Journal* (30th July 2022). URL: <https://hal.archives-ouvertes.fr/hal-02879337>.
- [25] L. Merliot-Gailhoustet, C. Raimbert, O. Garnier, J. Carr, A. de Jong, N. Molinari, S. Jaber and G. Chanques. 'Discomfort improvement for critically ill patients using electronic relaxation devices: results of the cross-over randomized controlled trial E-CHOISIR (Electronic-CHOIce of a System for Intensive care Relaxation)'. In: *Critical Care* 26.1 (Dec. 2022), p. 263. DOI: [10.1186/s13054-022-04136-4](https://doi.org/10.1186/s13054-022-04136-4). URL: <https://hal.science/hal-03768880>.
- [26] A. Perez-Lebel, G. Varoquaux, M. Le Morvan, J. Josse and J.-B. Poline. 'Benchmarking missing-values approaches for predictive models on health databases'. In: *GigaScience* (2022). DOI: [10.1093/gigascience/giac013](https://doi.org/10.1093/gigascience/giac013). URL: <https://hal.archives-ouvertes.fr/hal-03526292>.
- [27] C. Raimbert, O. Garnier, J. Carr, A. de Jong, N. Molinari, S. Jaber and G. Chanques. 'Publisher Correction to: Discomfort improvement for critically ill patients using electronic relaxation devices: results of the cross-over randomized controlled trial E-CHOISIR (Electronic-CHOIce of a System for Intensive care Relaxation)'. In: *Critical Care* 26.1 (Dec. 2022), p. 291. DOI: [10.1186/s13054-022-04169-9](https://doi.org/10.1186/s13054-022-04169-9). URL: <https://hal.science/hal-03930075>.
- [28] N. Roche, G. Garcia, A. de Larrard, C. Cancalon, S. Bénard, V. Perez, A. Mahieu, L. Vieu and P. Demoly. 'Real-life impact of uncontrolled severe asthma on mortality and healthcare use in adolescents and adults: findings from the retrospective, observational RESONANCE study in France'. In: *BMJ Open* 12.8 (Aug. 2022), e060160. DOI: [10.1136/bmjopen-2021-060160](https://doi.org/10.1136/bmjopen-2021-060160). URL: <https://hal.umontpellier.fr/hal-03930939>.
- [29] A. Rolle, A. de Jong, E. Vidal, N. Molinari, E. Azoulay and S. Jaber. 'Cardiac arrest and complications during non-invasive ventilation: a systematic review and meta-analysis with meta-regression'. In: *Intensive Care Medicine* (2022). DOI: [10.1007/s00134-022-06821-y](https://doi.org/10.1007/s00134-022-06821-y). URL: <https://hal.science/hal-03780151>.

- [30] F. Roubille, E. Matzner-Lober, S. Aguilhon, M. Rene, L. Lecourt, M. Galinier, J.-e. Ricci and N. Molinari. 'Impact of global warming on weight in patients with heart failure during the 2019 heatwave in France'. In: *ESC Heart Failure* (2022). DOI: [10.1002/ehf2.14129](https://doi.org/10.1002/ehf2.14129). URL: <https://hal.science/hal-03783874>.
- [31] D. Ryan, B. Flokstra-de Blok, E. Clark, C. Gaudin, M. Mamodaly, J. Kocks, J. L. van Der Velde, L. Angier, K. Romberg, R. Gawlik, P. Demoly and L. K. Tanno. 'Allergic and hypersensitivity conditions in non-specialist care: Flow diagrams to support clinical practice'. In: *Allergy* (2022). DOI: [10.1111/all.15273](https://doi.org/10.1111/all.15273). URL: <https://hal.umontpellier.fr/hal-03641107>.
- [32] Q. Samaran, R. Samaran, E. Ferreira, N. Haddad, A. Fottorino, H. Maillard, B. Dreno, N. Meyer, D. Azria, E. Maubec, C. Gaudy-Marqueste, N. Molinari, P.-E. Stoebner and O. Dereure. 'Anti-PD-1 for the treatment of advanced cutaneous squamous cell carcinoma in elderly patients: a French multicenter retrospective survey'. In: *Journal of Cancer Research and Clinical Oncology* (12th Aug. 2022), Online ahead of print. DOI: [10.1007/s00432-022-04246-0](https://doi.org/10.1007/s00432-022-04246-0). URL: <https://www.hal.inserm.fr/inserm-03778969>.
- [33] E. S. Sigurðardóttir, T. Gislason, B. Benediktsdóttir, S. Hustad, P. Dadvand, P. Demoly, K. A. Franklin, J. Heinrich, M. Holm, D. A. van der Plaats, R. Jögi, B. Leynaert, E. B. Lindberg, J. Martinez-Moratalla, L. S. de Aja, G. Pesce, I. Pin, C. Raheison, A. R. Pereira-Vega, F. Real and K. Triebner. 'Female sex hormones and symptoms of obstructive sleep apnea in European women of a population-based cohort'. In: *PLoS ONE* 17.6 June (2022). DOI: [10.1371/journal.pone.0269569](https://doi.org/10.1371/journal.pone.0269569). URL: <https://hal.science/hal-03765544>.
- [34] W. Srisuwatchari, C. Tacquard, A. Borushko, S. Viville, R. Stenger, Y. Ehrhard, J.-M. Malinovsky, P.-M. Mertes, P. Demoly and A. M. Chiriac. 'Diagnostic performance of serial serum total tryptase measurement to differentiate positive from negative allergy testing among patients with suspected perioperative hypersensitivity'. In: *Clinical and Experimental Allergy* 52.2 (Feb. 2022), pp. 334–344. DOI: [10.1111/cea.14040](https://doi.org/10.1111/cea.14040). URL: <https://hal.umontpellier.fr/hal-03641426>.
- [35] W. Srisuwatchari, T. Vo, A. Gauthier, N. Molinari, R. Schrijvers, P. Demoly and A. M. Chiriac. 'Hypersensitivity reactions to iodinated radiocontrast media: Cluster analysis reveals distinct clinical phenotypes'. In: *World Allergy Organization Journal* 15.9 (2022), p. 100680. DOI: [10.1016/j.waojou.2022.100680](https://doi.org/10.1016/j.waojou.2022.100680). URL: <https://hal.science/hal-03778478>.
- [36] L. Tanno, A. Gauthier, S. Allichon and P. Demoly. 'Global patterns of drug allergy-induced fatalities: a wake-up call to prevent avoidable deaths'. In: *Current Opinion in Allergy and Clinical Immunology* 22.4 (2022), pp. 215–220. DOI: [10.1097/ACI.0000000000000835](https://doi.org/10.1097/ACI.0000000000000835). URL: <https://hal.umontpellier.fr/hal-03779327>.
- [37] L. K. Tanno and P. Demoly. 'Food allergy in the World Health Organization's International Classification of Diseases (ICD)-11'. In: *Pediatric Allergy and Immunology* 33.11 (Nov. 2022), e13882. DOI: [10.1111/pai.13882](https://doi.org/10.1111/pai.13882). URL: <https://hal.umontpellier.fr/hal-03930620>.
- [38] L. K. Tanno, N. Molinari, I. Annesi-maesano, P. Demoly and A. L. Bierrenbach. 'Anaphylaxis in Brazil between 2011 and 2019'. In: *Clinical and Experimental Allergy* 52 (20th July 2022), pp. 1071–1078. DOI: [10.1111/cea.14193](https://doi.org/10.1111/cea.14193). URL: <https://hal.umontpellier.fr/hal-03779229>.
- [39] J. Vitte, M. Michel, A. Malinovschi, M. Caminati, A. Odebode, I. Annesi-maesano, D. P. Caimmi, C. Cassagne, P. Demoly, E. Heffler, E. Menu, B. I. Nwaru, Y. Sereme, S. Ranque, M. Raulf, W. Feleszko, C. Janson, C. Galán and E. T. F. O. Allergic Bronchopulmonary Aspergillosis. 'Fungal exposome, human health, and unmet needs: A 2022 update with special focus on allergy'. In: *Allergy* 77.11 (Nov. 2022), pp. 3199–3216. DOI: [10.1111/all.15483](https://doi.org/10.1111/all.15483). URL: <https://hal.umontpellier.fr/hal-03931288>.
- [40] M. Volpato, J. Vialaret, C. Hirtz, A. Petit, C. Suehs, J. Patarin, E. Matzner-Lober, I. Vachier, N. Molinari, A. Bourdin and J. Charriot. 'Rheology predicts sputum eosinophilia in patients with muco-obstructive lung diseases'. In: *Biochemical and Biophysical Research Communications* 622 (2022), pp. 64–71. DOI: [10.1016/j.bbrc.2022.07.025](https://doi.org/10.1016/j.bbrc.2022.07.025). URL: <https://hal.science/hal-03725444>.

Edition (books, proceedings, special issue of a journal)

- [41] M. Zaffran, O. Féron, Y. Goude, J. Josse and A. Dieuleveut, eds. *Adaptive Conformal Predictions for Time Series*. 2022. URL: <https://hal.archives-ouvertes.fr/hal-03573934>.

Reports & preprints

- [42] B. Colnet, J. Josse, G. Varoquaux and E. Scornet. *Reweighting the RCT for generalization: finite sample error and variable selection*. 4th Nov. 2022. URL: <https://hal.archives-ouvertes.fr/hal-03822662>.
- [43] I. Mayer, J. Josse and T. Group. *Generalizing treatment effects with incomplete covariates*. 7th Jan. 2022. URL: <https://hal.archives-ouvertes.fr/hal-03517373>.

Other scientific publications

- [44] M. Calderon and P. Demoly. ‘Integrating the evidence in allergen immunotherapy: Why real-world data should be important for the practicing clinician?’ In: *Allergy and Asthma Proceedings* 43.4 (1st July 2022), pp. 305–309. DOI: [10.2500/aap.2022.43.220006](https://doi.org/10.2500/aap.2022.43.220006). URL: <https://hal.science/hal-03936101>.

12.3 Cited publications

- [45] N. Cesa-Bianchi and G. Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.
- [46] I. Gibbs and E. Candès. ‘Adaptive Conformal Inference Under Distribution Shift’. en. In: *arXiv:2106.00170 [stat]* (June 2021). arXiv: 2106.00170. URL: <http://arxiv.org/abs/2106.00170> (visited on 09/07/2021).
- [47] J. Lei, M. G’Sell, A. Rinaldo, R. J. Tibshirani and L. Wasserman. ‘Distribution-Free Predictive Inference for Regression’. In: *Journal of the American Statistical Association* 113.523 (2018), pp. 1094–1111.
- [48] H. Papadopoulos, K. Proedrou, V. Vovk and A. Gammerman. ‘Inductive Confidence Machines for Regression’. In: *Machine Learning: ECML 2002*. Springer, 2002, pp. 345–356.
- [49] J. Pearl and E. Bareinboim. ‘Transportability of Causal and Statistical Relations: A Formal Approach’. In: *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*. AAAI’11. San Francisco, California: AAAI Press, 2011, pp. 247–254.
- [50] Y. Romano, E. Patterson and E. Candès. ‘Conformalized Quantile Regression’. en. In: *Advances in Neural Information Processing Systems*. Vol. 32. 2019. URL: <https://papers.nips.cc/paper/2019/hash/5103c3584b063c431bd1268e9b5e76fb-Abstract.html>.
- [51] P. Rothwell. ‘External validity of randomised controlled trials: “To whom do the results of this trial apply?”’ In: *Lancet* 365 (Jan. 2007), pp. 82–93. DOI: [10.1016/S0140-6736\(04\)17670-8](https://doi.org/10.1016/S0140-6736(04)17670-8).
- [52] E. A. Stuart, S. R. Cole, C. P. Bradshaw and P. J. Leaf. ‘The use of propensity scores to assess the generalizability of results from randomized trials’. In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 174 (2011), pp. 369–386.
- [53] V. Vovk, A. Gammerman and G. Shafer. *Algorithmic Learning in a Random World*. en. Springer US, 2005.