

2025 Activity Report

RESEARCH CENTRE: Inria Paris Centre

IN PARTNERSHIP WITH: INSERM, Université Paris Cité

Project-Team

HEKA

Health data- and model- driven approaches for
Knowledge Acquisition

In collaboration with CENTRE DE RECHERCHE DES CORDELIERS



Project-Team HEKA

Creation of the Project-Team: 2021 November 01

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

Keywords

Computer sciences and digital sciences

- A3.1.1. – Modeling, representation
- A3.1.9. – Database
- A3.1.10. – Heterogeneous data
- A3.1.11. – Structured data
- A3.2.1. – Knowledge bases
- A3.2.2. – Knowledge extraction, cleaning
- A3.2.3. – Inference
- A3.2.5. – Ontologies
- A3.3.2. – Data mining
- A3.3.3. – Big data analysis
- A6.1.2. – Stochastic Modeling
- A6.2.3. – Probabilistic methods
- A6.2.4. – Statistical methods
- A6.3.3. – Data processing
- A6.3.5. – Uncertainty Quantification
- A9.1. – Knowledge
- A9.2. – Machine learning
- A9.4. – Natural language processing
- A9.6. – Decision support
- A9.10. – Hybrid approaches for AI
- A9.11. – Generative AI
- A9.12.6. – Object localization
- A9.14. – Evaluation of AI models
- A9.16. – Societal impact of AI

Other research topics and application domains

- B2.2.1. – Cardiovascular and respiratory diseases
- B2.2.3. – Cancer
- B2.2.5. – Immune system diseases
- B2.2.7. – Virtual human twin
- B2.3. – Epidemiology
- B2.4. – Therapies
- B2.4.1. – Pharmaco kinetics and dynamics
- B2.4.2. – Drug resistance
- B2.4.3. – Surgery
- B2.6.1. – Brain imaging
- B2.6.3. – Biological Imaging
- B2.7.2. – Health monitoring systems

Contents

Project-Team HEKA	1
1 Team members, visitors, external collaborators	5
2 Overall objectives	9
3 Research program	9
3.1 Axis 1 - Knowledge extraction from clinical data	9
3.1.1 Methods	10
3.2 Axis 2 - Models & Learning for Complex Health Data	11
3.2.1 Methods	11
3.3 Axis 3: Data-driven and designs for next generation clinical trials	12
3.3.1 Methods:	12
4 Application domains	13
4.1 Multimodal approaches generalizable for several diseases	13
4.1.1 PEPR Digital Health	13
4.1.2 SurvivalGPU – “Using Graphics Processing Units (GPUs) to scale up survival analysis to nation-wide cohorts”	13
4.1.3 Messidore-Inserm BEEP - “Bayesian methods for Early Enriched Platform trials”	13
4.1.4 ANR AT2TA - “Analogies: from Theory to Tools and Applications”	14
4.1.5 iDEMO Meditwin - Dassault Systems - "Virtual twin for personalised medicine "(started in 2024)	14
4.1.6 RHU ReBone - “Surgery planning for multiple fractures”	14
4.2 Cancer	14
4.2.1 SIRIC InsiTu - “Insights into cancer: from inflammation to Tumor”	14
4.2.2 Combo - Sanofi - "Evaluating drug combinations in oncology with Real-World Data and state-of-the-art knowledge”	15
4.2.3 ARC “Accélération de la Recherche Clinique”	15
4.2.4 RHU OPERANDI - “Optimisation and imProved Efficacy of targeted RAdioNuclide therapy in Digestive cancers by Imagomics”	15
4.3 Rare diseases and pediatrics	16
4.3.1 EU INVENTS Horizon project - “Innovative designs, extrapolation, simulation methods and evidence-tools for rare diseases addressing regulatory needs” (started in 2024)	16
4.3.2 EU MSCA Doctoral Networks Orgestra project (started 2024)	16
4.3.3 BNDMR- Banque Nationale des Maladies Rares	16
4.4 Other diseases	17
4.4.1 Antibiotic resistance – FAIR project EU Horizon (on going)	17
4.4.2 Virtual reality (Ongoing)	17
5 Social and environmental responsibility	17
5.1 Societal impact	17
6 Highlights of the year	17
7 Latest software developments, platforms, open data	18
7.1 Latest software developments	18
7.1.1 medkit	18
7.1.2 Pythae	18
7.1.3 Pyraug	18
7.1.4 MultiVae	19
7.1.5 FrailtyCompRisk	19
7.1.6 survinsights	19

7.1.7	survivalGPU	19
7.1.8	recforest	20
7.1.9	KeOps	20
7.1.10	scikit-shapes	20
8	New results	20
8.1	Axis 1	21
8.2	Axis 2	21
8.3	Axis 3	23
9	Bilateral contracts and grants with industry	24
9.1	Bilateral contracts with industry	24
9.2	Bilateral Grants with Industry	25
9.3	Technology transfer and socio-economic impact	25
10	Partnerships and cooperations	26
10.1	International research visitors	26
10.1.1	Visits of international scientists	26
10.2	European initiatives	26
10.3	National initiatives	27
10.4	Public policy support	31
11	Dissemination	32
11.1	Promoting scientific activities	32
11.1.1	Scientific events: organisation	32
11.1.2	Scientific events: selection	32
11.1.3	Journal	32
11.1.4	Leadership within the scientific community	33
11.1.5	Scientific expertise	33
11.1.6	Research administration	33
11.2	Teaching - Supervision - Juries - Educational and pedagogical outreach	33
11.3	Publications 2025	33
12	Scientific production	33
12.1	Major publications	33
12.2	Publications of the year	34
12.3	Cited publications	43

1 Team members, visitors, external collaborators

Research Scientists

- Sarah Zohar [Team leader, INSERM, Senior Researcher, HDR]
- Yannick Binois [INRIA, Starting Research Position, from Sep 2025]
- Adrien Coulet [INRIA, Researcher, from Sep 2025, HDR]
- Adrien Coulet [INRIA, Associate Professor Detachement, until Aug 2025, HDR]
- Jean Feydy [INRIA, Researcher]
- Agathe Guilloux [INRIA, Professor Detachement, HDR]
- Moreno Ursino [INSERM, Researcher, HDR]

Faculty Members

- Stephanie Allasonniere [UNIV PARIS - CITE, Professor, HDR]
- François Angoulvant [UNIV PARIS, Professor, until Oct 2025, HDR]
- Sarah Berdot [AP/HP, Hospital Staff]
- Tom Boeken [AP/HP, Associate Professor]
- Guillaume Chassagnon [AP/HP, Professor, HDR]
- David Drummond [UNIV PARIS, Associate Professor]
- Anne-Sophie Jannot [AP/HP, Professor, from Nov 2025, HDR]
- Anne-Sophie Jannot [AP/HP, Associate Professor, until Oct 2025, HDR]
- Sandrine Katsahian [UNIV PARIS, Professor, HDR]
- Andrea Lazzati [APHP USPN, Professor, HDR]
- Claire Leconte Rives-Lange [AP/HP, Associate Professor]
- Ivan Lerner [UNIV PARIS - CITE, Hospital Staff]
- Germain Perrin [AP/HP, Hospital Staff, until Oct 2025]
- Marie Pierre Revel [APHP, Professor, HDR]
- Brigitte Sabatier [APHP, Hospital Staff]
- Stylianos Tzedakis [APHP, Hospital Staff]

Post-Doctoral Fellows

- Charbella Abou Khalil [INRIA, Post-Doctoral Fellow, until Sep 2025]
- Nadim Ballout [INRIA, Post-Doctoral Fellow, until May 2025]
- Sandrine Boulet [INSERM, Post-Doctoral Fellow, until Apr 2025]
- Perrine Chassat [INRIA, Post-Doctoral Fellow]
- Lucas Ducrot [INRIA, Post-Doctoral Fellow, until Aug 2025]
- Romain Jaquet [GHNE, until Oct 2025]

- Jong Ho Jhee [INRIA, Post-Doctoral Fellow]
- Minon'Tsikpo-Kossi Kodji [INRIA, Post-Doctoral Fellow, from Apr 2025]
- Letao Li [INSERM, Post-Doctoral Fellow]
- Robin Magnet [UNIV PARIS - CITE, Post-Doctoral Fellow, from Nov 2025]
- Robin Magnet [INRIA, Post-Doctoral Fellow, until Oct 2025]
- Romain Michelucci [INRIA, Post-Doctoral Fellow, from Jun 2025]
- Remi Trimbour [UNIV PARIS - CITE, Post-Doctoral Fellow, from Oct 2025]

PhD Students

- Eya Abid [SORBONNE UNIVERSITE]
- Safa Alsaidi [IMAGINE OPTIC, from Dec 2025]
- Safa Alsaidi [INRIA, until Nov 2025]
- Jean-Baptiste Baitairian [SANOFI, CIFRE]
- Ariane Bercu [INSERM]
- Hadrien Bigo-Balland [MyFit Solutions, CIFRE]
- Ilona Blanchard [INRIA]
- Théau Blanchard [GE HELTHCARE, CIFRE]
- Linus Bleistein [UNIV EVRY, until Mar 2025]
- Lea Comin [CEA]
- Charles De Ponthaud [UNIV PARIS - CITE]
- Louie-David Desachy [INRIA, from Nov 2025]
- Louise Durand-Janin [DOCTOLIB, CIFRE, from Sep 2025]
- Pierre Epron [INRIA]
- Foucauld Estignard [DOCTOLIB, CIFRE, from Nov 2025]
- Thibaut Fabacher [UNIV STRASBOURG, until Aug 2025]
- Corentin Faujour [Cemka, CIFRE]
- Alexandre Pierre Edouard Faure [Univ Paris Cite]
- Fabrice Gambaraza [CHU AVICIENNE AP-HP, until Aug 2025]
- Mohamed Ghebriout [CNRS]
- Louis Goldenberg [DASSAULT SYSTEMES, CIFRE]
- Guillaume Houry [INRIA]
- Emilien Jemelen [EPICONCEPT, CIFRE]
- Antoine Laforgue [UNIV PARIS - CITE, from Oct 2025]
- Renee Le Clech [INRIA]

- Svetlana Le Ralle [ROCHE, from Nov 2025]
- Philomène Letzelter [WITHINGS, CIFRE, from Dec 2025]
- Nicolas Loche [Unvi Paris Cite, from Nov 2025]
- Hugo Malafosse [INSERM, from Oct 2025]
- Tristan Margate [DASSAULT SYSTEMES, CIFRE]
- Benjamin Maurel [INSERM]
- Fabien Maury [INSERM, from Oct 2025]
- Yifan Mei [UNIV PARIS - CITE]
- Giulia Monchietto [INSERM]
- Lillian Muyama [INRIA, until Feb 2025]
- Antoine Poirot-Bourdain [UNIV PARIS -Dauphine, from Nov 2025]
- Asok Rajkumar [CHU AVICIENNE AP-HP, until Nov 2025]
- Theo Rene [INRIA, from Nov 2025]
- Louis Romengas [AP/HP]
- Ababacar Sembede [INRIA, from Nov 2025]
- Agathe Senellart [UNIV PARIS - CITE]
- Guillaume Serieys [UNIV PARIS - CITE, until Sep 2025]
- Woosub Shin [UNIV PARIS - CITE]
- Tarini Singh [INSERM]
- Marie Pauline Talabard [AP/HP, from Nov 2025]
- Dylan Vellas [UNIV PARIS - CITE]
- Louis Vincent [IMPLICITY, CIFRE, until Mar 2025]
- Axel Vuorinen [INSERM]

Technical Staff

- Baptiste Archambaud [INRIA, Engineer, from Mar 2025]
- Armelle Arnoux [AP/HP, Engineer]
- Sandrine Boulet [INRIA, Engineer, from May 2025]
- Loubna Cadi [INSERM, Engineer, from Jul 2025]
- Francisco De Lima Andrade [INRIA, Engineer, from Feb 2025]
- Claire Dechaux [INRIA, Engineer, from May 2025]
- Nicolas Garcelon [Fondation Imagine, Engineer, until Oct 2025]
- Caroline Lawless [INRIA, Engineer, from Aug 2025]
- Diana Mandache [UNIV PARIS CITE, Engineer]

- Martial Marzloff [INSERM, Engineer, from Mar 2025]
- Sebastian Mendez Pineda [INRIA, Engineer, from Nov 2025]
- Van Tuan Nguyen [INRIA, Engineer]
- Antoine Poirot-Bourdain [CNRS, Engineer, until Sep 2025]
- Charlotte Ronde-Roupie [INRIA, Engineer, from Feb 2025 until Oct 2025]
- Erick Tavares Penate [INRIA, Engineer, from Mar 2025]
- Caglayan Tuna [INRIA, Engineer, until Oct 2025]
- Ghislain Vaillant [INRIA, Engineer]

Interns and Apprentices

- Gabriel Agossou [INRIA, Intern, from Feb 2025 until Aug 2025]
- Iman Bensalami [INRIA, Apprentice, from Nov 2025]
- Lina Benyamina [INRIA, Intern, from Jul 2025]
- Felix Berthou [INRIA, Apprentice]
- Arnaud Cournil [INRIA, Intern, from Jul 2025]
- Benjamin Delmas [INRIA, Intern, from Jun 2025 until Aug 2025]
- Louie-David Desachy [INRIA, Intern, from Apr 2025 until Oct 2025]
- Bastien Franja [INRIA, Intern, from Nov 2025]
- Antoine Laforgue [ENSMP, Intern, from Apr 2025 until Aug 2025]
- Clara Leducq [INRIA, Intern, from May 2025 until Jun 2025]
- Nicolas Loche [AP/HP, Intern, from Feb 2025 until Aug 2025]
- Hugo Malafosse [INRIA, from Jul 2025 until Sep 2025]
- Fabien Maury [INSERM, until Sep 2025]
- Leo Megret [Agence du Numérique en Santé, Intern, until Jun 2025]
- Theo Rene [INRIA, Intern, from May 2025 until Sep 2025]
- Thomas Trang [INRIA, Intern, from Jul 2025 until Sep 2025]
- Theo Vandenhole [Doshas Consulting, until Apr 2025]
- Imen Wafra [INRIA, Intern, from Apr 2025 until Oct 2025]
- Abdell-Aziz Youssouf [INSERM, Intern, from Jul 2025 until Aug 2025]

Administrative Assistants

- Vincent Damotte [INSERM]
- Ibtissam Fadiz [INSERM, from Feb 2025]
- Meriem Guemair [INRIA]
- Dimitri Varis Bado [Univ Paris Cite]

Visiting Scientist

- Alberto Marfaglia [UNIV BOLOGNE, from Nov 2025]

External Collaborator

- Lillian Muyama [UNIV MAKERERE, from Mar 2025]

2 Overall objectives

The primary objective of HeKA is to develop methods, models, and tools aimed at creating, evaluating, and validating learning health systems—that is, health systems capable of leveraging data collected throughout the care pathway to produce validated, data-driven medical insights that iteratively refine clinical decision-making and outcomes. The increasing availability of clinical trial data, real-world data (e.g., EHRs, cohorts, registries), and digital biomarker data from Digital Medical Devices (DMDs), together with linked sources such as the SNDS (the French National Health Data System), provides a unique opportunity to design multimodal and multidimensional modeling strategies for improved patient stratification, prediction, and future care. Importantly, many of these data sources now include imaging data, including high-resolution and 3D images, further enriching the clinical and biological information available for analysis. However, despite this wealth of heterogeneous data, we frequently face the HDLSS (high-dimension, low-sample-size) problem associated with data structures that are themselves usually complex. All in all, this requires developing robust statistical, mathematical, machine-learning, and possibly causal approaches to fully exploit the potential of such complex health datasets. By addressing these challenges, HeKA aims to advance precision and personalized approaches to monitoring, diagnosis, therapy, and prognosis, thereby contributing to higher-quality healthcare.

Methodological developments are required at every stage of the health-data pipeline, from securing and accessing data, to transforming them through approaches such as representation learning, to analyzing them using predictive modeling and data- or model-driven knowledge discovery. These advances must ultimately support their implementation in decision-support systems, medical devices, and next-generation clinical evaluation in real life setting for the assessment of medical knowledge.

3 Research program

HeKA addresses these objectives through three interdependent axes (Figure 1): (1) learning and reasoning over patient representations, (2) models and learning for complex health data, and (3) data- and model-driven designs for next-generation clinical trials. Together, these axes contribute to the development of a comprehensive learning health system. Axes 1 and 2 primarily involve observational studies, retrospective or prospective, where investigators observe associations between factors and outcomes without intervening. Axis 3, in contrast, relates mostly to prospective interventional studies (i.e., clinical trials or studies), in which the investigator actively intervenes as part of the study design, without disregarding hybrid design (incorporating external information).

Each axis addresses a set of highly competitive and timely challenges in contemporary medical research, and our work is positioned within a strong network of national and international collaborations, detailed in the research focus below, alongside a clear identification of the main competing teams in the field.

3.1 Axis 1 - Knowledge extraction from clinical data

The Axis 1 of the HeKA team focuses on the development of **methods to extract, represent, learn and reason from individual healthcare data**, with the general aim of **supporting personalized clinical decision-making**. One of our particularities here is to consider real-world healthcare data, along with biomedical domain knowledge.

The heterogeneity, incompleteness, and dynamic nature of EHR data, along with their large disconnection from background knowledge, make their use for personalized medicine applications a significant extraction and representation challenge. This challenge can be summarized by the following question, “*how to optimally*

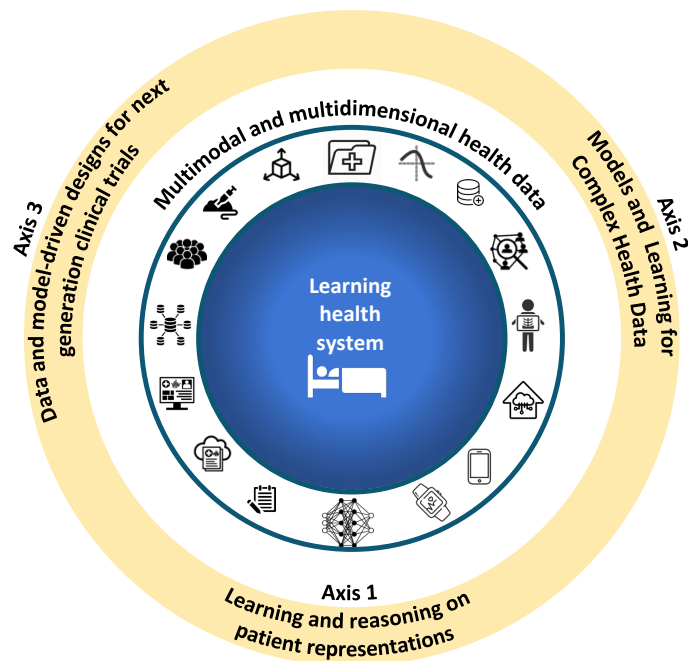


Figure 1: HeKA's three research axes.

represent patient data for knowledge discovery and decision support?” and motivates most of the works of the research Axis 1. Besides, a substantial amount of high quality biomedical knowledge is available online, in various levels of formalization, structuration and interoperability, going from formal ontologies and knowledge graphs to the scientific literature. This motivates the second scientific question investigated by the Axis 1, which is *“how biomedical knowledge in its variety can be leverage to improve knowledge discovery and decision support?”*

Figure 2 illustrates the three tasks the Axis 1 focus on: (i) information extraction, (ii) patient representation learning, and (iii) clinical decision support, and highlight the consideration of biomedical knowledge, in its various form, for each of these tasks.

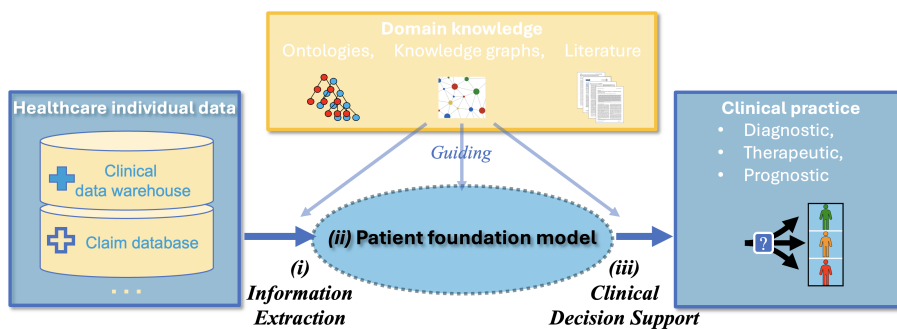


Figure 2: Main focuses of the HeKA Axis 1 named *Learning and reasoning on patient representations*.

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 multiple. 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 - Models & Learning for Complex Health Data

The increasing availability of complex, multimodal health data — ranging from electronic health records (EHR) and administrative claims to clinical trials, longitudinal cohorts, and medical imaging — offers unprecedented opportunities for developing predictive and causal models to support clinical decision-making. Yet, these datasets are intrinsically challenging: observations are often irregular and heterogeneous, trajectories may span multiple disease states, imaging and biological measurements introduce high-dimensional feature spaces, and clinical trials provide well-controlled but limited samples, while real-world data (RWD) remains rich but biased.

Developing robust and credible algorithms in this ecosystem requires a scientific approach that integrates statistical learning, machine learning, geometry and causal inference. Our research focuses precisely on these boundaries where methods succeed or fail: the interface between deep learning and more classical statistical models; between high-dimensional representation learning and low-sample-size inference; and between predictive accuracy and causal validity. Rather than assuming that modern deep architectures are universally optimal, we aim to understand when and why certain approaches outperform others in temporal modeling, multimodal integration, and rare-disease or small-sample settings. This perspective naturally aligns with the broader mission of HeKA to develop learning health systems that are trustworthy, explainable, and clinically actionable.

Embedded within HeKA's interdisciplinary environment, our team positions itself at the intersection of modeling, learning, and clinical applicability. We contribute methodological innovations for high-dimensional and longitudinal medical data, while ensuring that algorithms remain interpretable, reliable, and suited for real-world deployment within learning health systems. Our ambition is to advance models that not only perform well, but also explain, predict, and inform clinical decisions in a robust and transparent manner.

3.2.1 Methods

During this year, our team's scientific achievements have centred on four main themes: predictive modelling for longitudinal health data, causal inference, algorithms for data augmentation, and topology-aware anatomical analysis. We have led major projects such as REWIND (PEPR Digital Health), MEDITWIN, and multiple CIFRE and industrial collaborations, enabling strong academic, clinical, and industrial partnerships.

A first cornerstone is the development of predictive tools for longitudinal and time-series data, including algorithms based on latent evolving states, and high-dimensional joint models [88]. We advanced interpretability for survival and longitudinal prediction [106], and released the open-source `survinsights` library. Finally, large-scale SNDS access (HDH platform) allowed us to pioneer GPU-accelerated survival modelling, with the `SurvivalGPU` package now applied to millions of patients, and to tackle multimodal care pathways.

Causal inference is a second theme, with a focus on target trial emulation and sensitivity analysis for unmeasured confounding [104], [103]. The third topic, data augmentation and synthetic data, includes geometry-aware VAEs capable of generating realistic data from small cohorts [57] [105]. These approaches supports data completion, anonymisation, and the construction of artificial control arms for clinical trials, complementing our broader work on synthetic control arms. We were also awarded funding through the Boas call for proposals to develop an anonymized SNDS dataset that will facilitate methodological development using these data.

Finally, our fourth topic focuses on anatomical modelling methods that remain robust under topology changes. Classical shape analysis assumes smooth deformations between anatomies, an assumption that breaks down in applications such as orthopedic surgery or interventional radiology. We therefore develop

approaches suited to discontinuities, including bone fractures (RHU Rebone) and large-scale vascular network analyses (MEDITWIN). This work balances theory and transfer: our garment-tailoring methods (e.g., compression gloves) are already used by hundreds of patients through MyFit Solutions, and our vascular-visualisation tools, published at MICCAI 2025 [72], are generating interest from GE Healthcare.

Taken together, these contributions form a coherent and impactful scientific agenda that combines methodological innovation, large-scale data science and clinically driven applications across oncology, pharmacovigilance, public health surveillance and medical imaging.

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

New model-based fundamental research, spanning from preclinical to clinical stages, has the potential to play a pivotal role in patient screening and in predicting individual responses prior to and during clinical trials. By leveraging modelling approaches based on, for example, biomarkers, clinical outcomes, pharmacokinetics/pharmacodynamics (PK/PD), electronic health records (EHRs), and real-world data, these methods enable a more precise and data-driven selection of patients and treatment strategies. This paradigm is particularly critical in rare diseases, paediatrics, and other small-population settings, where patient numbers are limited and the optimal use of all available knowledge is essential. At the same time, the landscape of clinical evidence is rapidly evolving through the emergence of new digital sources of patient data. Digital Medical Devices (DMDs), as defined under Regulation EU 2017/745 for the prevention, diagnosis, monitoring, treatment, or alleviation of disease, increasingly rely on machine learning algorithms and Software as a Medical Device (SaMD) technologies. These innovations hold major promise for personalized and adaptive patient care. However, unlike conventional drugs, whose chemical composition usually remains stable over time, the performance of DMDs may vary since the underlying models are often continuously enriched by new observational data. This creates a major methodological and regulatory challenge: ensuring, at any point in time, the safety, robustness, and clinical effectiveness of adaptive and learning-based medical technologies. Within this context, our main aims address two challenges: (1) the development of model-based and simulation-driven methodologies tailored to small populations, rare diseases, and surgical or metabolic cohorts, and (2) the design, evaluation, and life-cycle monitoring of clinical studies for digital and medical devices.

3.3.1 Methods:

This year, we conducted and completed a comprehensive investigation into the added value of incorporating pharmacokinetic information into dose-finding studies. We adapted existing methodologies and compared their performance across a wide range of realistic clinical scenarios. In parallel, we advanced our work on the translational use of PET microdosing by developing modeling strategies that bridge preclinical and clinical development. On the clinical trial methodology side, we developed an innovative hybrid framework in which Bayesian and frequentist approaches are combined to enable adaptive interim analyses, with the objective of identifying the patient population most likely to benefit from treatment. We also concluded our contributions to the CONSORT and SPIRIT guidelines for early-phase clinical trials, helping to shape international standards for trial design and reporting. In applied clinical research, our team produced significant analyses in the areas of surgical outcomes, artificial nutrition, and pregnancy following bariatric surgery. Regarding DMDs, we contributed to the development of one of the first structured evaluation frameworks specifically designed for these technologies, addressing critical unmet needs in regulation, clinical validation, and real-world evidence generation. We are also actively developing new methods to improve the interpretability and explainability of AI-based medical technologies and are disseminating this work through international conferences.

4 Application domains

4.1 Multimodal approaches generalizable for several diseases

4.1.1 PEPR Digital Health

The PEPR ("Programmes et équipements prioritaires de recherche") Digital Health aims at gathering national multidisciplinary community active in digital health for the development and exploitation of the concept of digital twin in health (started in September 2023).

HeKA's involvement in this PEPR are the following; (i) within project **ShareFAIR**, to learn protocols from clinical data collected along healthcare activity in Electronic Health Records (EHRs) to explicitize the medical decision processes, (steps to reach a particular diagnosis or therapeutic choice) and the management of particular conditions (steps in the management of a particular condition). Protocols extracted from EHRs provide a view on the real-world clinical practice and may then be compared together or with CPG (clinical practice guidelines) which can be seen as more theoretical protocols in that they provide recommendations, or clinical pathways (CP) to standardize clinical practice. It will be applied within **NEUROVASC** in the impact reduction of intracranial aneurysm and stroke, in which we will extract and the proposed clinical pathways, (ii) within **REWIND**, to develop of new mathematical and statistical approaches for the analysis of multimodal multiscale longitudinal data to predict patient's response. These models will be designed, implemented as prototypes and then transferred to an easy-used-well-documented platform where people from diverse communities, in particular physicians, will be able to use them on their own data set, (iii) within **DIGPHAT**, to develop Bayesian modelling of meta-models pathways for the development of digital pharmacological twin; it consists in the analysis of data from omics experiments and selection of relevant covariables and to combine meta-models in pathways to select the most reliable twin model, (iv) within project **M4DI**, to develop a generic method for identifying subgroups of patients with the same phenotype from health databases, using jointly variable correlations and expert data, and to implement it within a computer package, (v) all these previous projects will purpose models or Clinical Decision Support Systems (CDSS) to be translated to clinical practice, however, proof based on data only is not sufficient and it should be evaluated in real life through prospective and interventional clinical trials or studies, this will be done within **SMATCH**. In this project we will propose new methodological paradigms for the clinical evaluation of Digital Medical Devices (DMD) including CDSS and AI based models and algorithms.

4.1.2 SurvivalGPU – “Using Graphics Processing Units (GPUs) to scale up survival analysis to nation-wide cohorts”

The recent availability of health insurance databases such as the SNDS opens the door to the detection of adverse drug reactions in the general population. The aim is to generalize the survival analyses usually carried out during clinical trials on cohorts of N=1k to 10k patients to the full French population. This line of research is appealing but poses real methodological challenges. Notably, it requires the development of statistical analysis models that meet the robustness and interpretability requirements of public health physicians while taking full advantage of recent hardware accelerators to scale up to millions of patients per cohort. In this context, our team has been working since 2022 on an efficient re-implementation, on Graphics Processing Units (GPUs), of the standard software tool in the field: the R package "survival". The new "**survivalGPU**" library leverages recent software tools (PyTorch, PyG, KeOps, reticulate) to bridge the gap between high-performance computing and traditional survival analysis. It now provides a complete re-implementation of the Cox proportional hazards model that is around 100 times faster on GPU than the survival package on CPU. Going further, it supports time-varying drug exposures via the Weighted Cumulative Exposures model and is accessible via an R interface which is fully retro-compatible with that of the survival package. We now intend to perform extensive validation and comparison with other models, prior to pharmaco-epidemiological studies on the SNDS data via the Health Data Hub platform.

4.1.3 Messidore-Inserm BEEP - “Bayesian methods for Early Enriched Platform trials”

The recent pandemic has shown the need of speeding up the clinical trial development of novel or repurposed therapies. Indeed, following the usual drug development paradigm, where clinical trial phases are performed sequentially and separately, the time required to the full process easily exceeds a decade. Our objective is to

propose innovative Bayesian enriched “platform” designs for early phase trials, which are adapted to the clinical context and go towards precision medicine. Since we are focusing on early phases of clinical trial, in this setting “platform” cannot be linked to classical RCT. Thus, we aim at defining how “platform” trial should be translated into these early phases. As in the original definition, early platform phases will allow for flexibility, such as adding new arm or stopping treatments for futility (and/or safety in our case). The word enriched refers to the use of new information, or at least not usually used in such early trials, such as positron emission tomography (PET) scan, pharmacokinetics/pharmacodynamics (PK/PD) modelling, mathematical modelling of immune responses, and to the enrichment of the enrolled patient based on their biomarkers. The project is built around workpackages (WPs). In WP1, we develop platform trials in phase 0/I, based on PET-scan; microdosing on several (preclinical) animal species and humans will be adaptively compared, added or deleted to better characterize the extrapolation to human. In WP2, we develop phase I/II dose-finding trial using PK/PD or mechanistic PD models. In WP3, enrichment designs for phase I/II, in survival settings, are proposed when selected biomarkers are available, and the design will be extended in case of combination therapies.

4.1.4 ANR AT2TA - “Analogies: from Theory to Tools and Applications”

Analogical reasoning is a kind of reasoning that is based on finding a common relational system between two situations, exemplars, or domains. In computer science, analogical reasoning can be supported by two main axes of artificial intelligence: knowledge representation and reasoning, and machine learning. The AT2TA projet particularly aims at studying the role that machine learning can play in analogical reasoning; and the HeKA team is in charge of exploring the application of their interplay in the healthcare domain. A PhD student, co-supervised with Inria Paris, IHU Imagine and Université de Lorraine, is learning representations of patients, relying on clinical texts, and study how these representations can first compose analogical propositions, and second serve as bricks to a machine learning architecture for analogical reasoning.

4.1.5 iDEMO Meditwin - Dassault Systems - “Virtual twin for personalised medicine ”(started in 2024)

Meditwin is a collaborative project funded by BPI (“Banque Publique d’Investissement”) with Dassault Systems (leader of the projects), Inria, IHU institutes across France and Medtech startups. The aim of the project is to provide a digital platform relying on virtual twins of individuals who faithfully reproduce their state of health and which make it possible to test different therapeutic options. It will promote interdisciplinarity by facilitating interoperability of multimodal medical data. Our team will use AI approaches to propose Clinical Decision Support Systems (CDSS) in cardiovascular diseases and cancer. We will also develop the clinical trial methodology evaluation these CDSS. In particular, HeKA will develop stratification and classification algorithms, synthetics patient’s generators, statistical and mathematical models for multi-modal and multidimensional health data and clinical evaluation methods for the resulted CDSS as Digital Medical Devices.

4.1.6 RHU ReBone - “Surgery planning for multiple fractures”

The RHU ReBone is a French consortium led by the orthopedic surgery unit of the Nice hospital. It is funded by the ANR from 2024 to 2029, and aims at producing robust anatomical software to automate the planning of complex fracture reductions. Jean Feydy and Stéphanie Allasonnière work on the image pre-processing and analysis, in close collaboration with Hervé Delingette from the Epione team at Inria Sophia.

4.2 Cancer

4.2.1 SIRIC InsiTu - “Insights into cancer: from inflammation to Tumor”

To turn scientific knowledge into sustainable healthcare, cancer research must identify who is at risk of cancer, when and in whom a new cancer arises, and how best to treat it and gauge response. Aligned with Europe’s Beating Cancer Plan, InsiTu takes on the three challenges of cancer prevention, interception, and treatment in digestive, lung, skin cancer, and heme malignancies. Chronic inflammation is a key cancer niche fostering tumor initiation. Leveraging a transformative Tissue-Hub interfacing diagnostics and research,

our program 'From inflammation to clonal emergence and cancer' will unite experts in chronic diseases damage to monitor patients with chronic tissue inflammation and cancer predisposition, mirrored by animal modelling, to understand the critical transition from chronic tissue damage to cancer progression, opening opportunities for prevention and interception. Such longitudinal (and sometimes invasive) interactions between patients and healthcare practitioners can be improved by empowering patients, taking psychic, social and ethical dimensions in consideration. Our program 'Imaging cancer and its environment' will take a different approach to this challenge. Through synergetic interactions with mathematicians and physicists, it will provide novel frameworks for multiscale integration of molecular alterations, cellular processes, and tissue complexity. This effort will result in image-based, non-invasive 'virtual biopsies' as proxies of key biological processes underlying tumor heterogeneity and drug resistance. Along with novel biomarkers such as circulating extracellular vesicles, these virtual biopsies will gauge responses to new therapeutic approaches developed in our third program 'From new targets to new trials'. There, experts in leukemias and skin cancers will use cutting-edge in vivo functional screens and multi-omic interrogation of Tissue-Hub samples to identify new targetable vulnerabilities and develop next-generation cell-based immunotherapies. To fasten the transfer of these innovations into care, new adaptive clinical trial designs will be engineered.

4.2.2 Combo - Sanofi - "Evaluating drug combinations in oncology with Real-World Data and state-of-the-art knowledge"

Combo is a collaborative project with Industry, national health data platforms and cancer institute: Sanofi Pharma, The Health Data Hub, Centre Léon Berard and Inria-Inserm-HeKA. The objective of the project is to identify promising families of drug combinations in oncology using multisource and multi-modal data modelling and prediction, including RWD (cancer patients' care data from CLCC cancer centre), genomic public databases, literature, clinical trials depository and expert's opinion. Once these combinations will be identified mechanistic models will be used to determine dose-regimen and build dose-finding trial designs for the combinations to be evaluated through formal clinical trials. In this project we lead the following WPs (1) AI based analysis of the multimodal RWD and subgroup discovery for the identification of relevant combinations, (2) multi-modal analysis accounting RWD modelling as well literature and public clinical platforms, and (3) proposing candidates.

4.2.3 ARC "Accélération de la Recherche Clinique"

Real-world data (RWD) from electronic records, wearables, and other digital sources provide valuable insights into routine clinical practice and help define the target population of a trial. Artificial patient data and virtual control arms can simulate comparators, reducing recruitment needs. These approaches are especially useful in diseases with difficult enrolment, such as oncology. By shortening trial timelines and lowering costs, synthetic trials enhance industrial competitiveness. Ensuring their mathematical and clinical validity remains essential for producing reliable evidence.

4.2.4 RHU OPERANDI - "Optimisation and imProved Efficacy of targeted RADioNuclide therapy in Digestive cancers by Imagomics"

Advanced stage hepatocellular carcinoma (HCC) and gastroenteropancreatic neuroendocrine tumours (GEP-NET) are currently treated with targeted radionuclide therapy (TRT), a highly advanced method that consists of either intra-arterial injections of radioactive microspheres (transarterial radioembolisation - TARE) or targeted peptides radioactively labelled and administered systemically (Peptide Receptor Radionuclide Therapy - PRRT). While highly effective, patient stratification and early identification of responders are currently managed insufficiently due to the lack of pertinent imaging biomarkers, either non-invasive or invasive. Furthermore, therapy-induced DNA damage leads to tumour resistance, reducing TRT efficacy. We aim to overcome those current limits through the OPERANDI project via innovative approaches in engineering, novel imaging biomarkers, and new concepts for DNA repair mechanisms, combined with a fundamental understanding of causal links. Our ambitions go beyond the current state-of-art, embracing even new combinations of drugs and α -emitters to enhance dose localization and efficiency. Methodology will try to understand fundamentally whether current patient management using CT/PET/MRI allows to predict response and survival using cutting edge imaging-based artificial intelligence (AI) approaches in combination with data augmentation techniques to reach statistical significance.

4.3 Rare diseases and pediatrics

4.3.1 EU INVENTS Horizon project - “Innovative designs, extrapolation, simulation methods and evidence-tools for rare diseases addressing regulatory needs” (started in 2024)

The evaluation of new medicines for rare diseases (RD) including rare paediatric RDs is challenging for several reasons, among which are the small patient sample sizes, heterogeneity of patients and diseases and heterogeneity in disease knowledge. Due to these difficulties, access to effective treatments and the number of treatment options are often limited in RDs. INVENTS aims to provide clinical trial trialists, researchers and regulators with a global framework encompassing methods, workflows and evidence assessment tools to be implemented in orphan and paediatric drug development. Our ambition is to significantly improve the evaluation of evidence and regulatory decision-making through the development and validation of: refined longitudinal model-based diseases trajectories and treatment effect, improved extrapolation models, in silico trials (e.g., virtual patient cohorts), optimised model-based clinical trial designs and evidence synthesis methods. These will be evaluated through simulation studies and tested on extensive data from a range of use cases provided by our industrial partners Roche and Novartis and Real World data (RWD) from RD registry. The INVENTS framework will improve consistency and efficiency of the drug evaluation process for RD by augmenting clinical evidence without compromising its scientific integrity and providing regulators assessment credibility criteria. At the end of this 5 years project, the European industry will be able to exploit novel and improved clinical trial designs, in silico trials and RWD analysis approaches supporting drug development in RD. The European Medicine Agency and European national regulators (including Health Technology Assessment bodies) will be supplied with a general framework allowing better informed decision making. Most importantly, RD patients will benefit from an increased and faster access to efficacious and safe treatments.

4.3.2 EU MSCA Doctoral Networks Orgestra project (started 2024)

Organoids experimental models are in vitro 3D cell cultures which can be generated from embryonic stem cells, induced pluripotent stem cells or adult stem cells, and can replicate organs functionally and structurally. Their physiological resemblance to target organs and ability to cryopreserve make organoids a powerful tool for biomedical research and advancing understanding of the mechanisms underlying certain disorders, including rare diseases. The ORGESTRA Joint Doctoral Network will propose innovative organoid technologies for two genetic disorders, i.e., cystic fibrosis and cystinosis. In this project we will supervise 2 PhDs project which will propose statistical development for; (1) linking in silico trials to organoids data and innovative trial design. These designs will incorporate biomarkers-based findings, as organoids, i.e., that reduce unnecessary exposure of patients (screening) or allow drugs to be screened more effectively for non-effectiveness before embarking on human trials. This will be done via a joint doctoral degree with University of Utrecht. (2) Estimand framework involving Bayesian principles on organoids data for clinical trial outcomes and models. The estimands framework will be based on expert’s elicitation to understand which questions are more relevant in term of clinical efficacy/toxicity, to select the proper outcomes, to identify the possible intercurrent events and to provide a robust statistical model whose parameters will be estimated under a Bayesian setting. This will be done via a joint doctoral degree with Katholieke Universiteit Leuven.

4.3.3 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 developed 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 developed in the form of Axis 2 to model these longitudinal data.

4.4 Other diseases

4.4.1 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.

4.4.2 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 Clause, G Manceau and A Rastello)

5 Social and environmental responsibility

5.1 Societal impact

Communicating and disseminating our research to the general public is becoming part of our communication activities. The questions raising from the general public highlights how digital health is still misunderstood and fears regarding AI have been largely spread. This is why, as researchers in the field it is important to explain and answer these important concerns. S. Allasonnière is a member of the scientific advisory board for "AI for Health" summit since 2020 and she regularly moderate round tables on the topic for industrial and general public. Likewise, she's a member of the scientific advisory board of "MedInTechs" since 2023 and moderate round tables in this context.

6 Highlights of the year

- HeKA became an UMR (INRIA - INSERM - Univ. Paris Cité) on January 1st 2025.
- S. Zohar received the Innovation award of Inserm in November 2025.
- S. Allasonnière won the ‘Visionary in Health’ Award at the MedInTechs exhibition in March 2025.
- The HAS, in collaboration with INRIA, has launched a self-referral initiative to develop a methodological guide for assessing the organizational impact of Digital medical devices (DMD), primarily digital therapeutics and remote patient monitoring solutions. The project will focus on DMDs currently covered by HAS's evaluation scope, while also providing value beyond this scope, particularly in supporting the development of new evaluation frameworks for DMDs that are not currently assessed by HAS.

7 Latest software developments, platforms, open data

7.1 Latest software developments

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

Scientific Description: medkit aims to facilitate information and knowledge extraction from data of various modalities by providing software modules for data preparation or analysis and facilitating their chaining. The development of new modules is motivated either by needs within the core of the library or by application projects. The initial projects of the team that motivated developments were related to knowledge extraction from healthcare data warehouses, particularly their textual content.

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 supervised models from these various modalities for decision support in healthcare.

Release Contributions: ## Fixed

- Use ISO 8601 timestamp for model checkpoint paths - Fix test of iamsystem matcher on Python 3.12

URL: <https://github.com/medkit-lib/medkit>

Contact: Adrien Coulet

Participant: 8 anonymous participants

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/clementchadbec/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/clementchadbec/pyraug>

Contact: Clément Chadebec

7.1.4 MultiVae

Keywords: Multimodality, Variational Autoencoder

Functional Description: This library gathers some of the most common multi-modal Variational AutoEncoder (VAE) implementation in PyTorch as well as benchmarking tools (datasets, metrics...).

Contact: Agathe Senellart

7.1.5 FrailtyCompRisk

Name: competing risks survival analysis in multicenter studies

Keyword: Survival analysis

Functional Description: This library provides tools for competing risks survival analysis in multicenter studies, which is of particular interest for statisticians and epidemiologists analyzing time-to-event data. The code is available on GitHub, released under a reciprocal license (GPL-3.0).

URL: <https://github.com/TeamHeKA/FrailtyCompRisk>

Contact: Sandrine Katsahian

7.1.6 survinsights

Name: analysing and interpreting machine-learning-based survival models

Keywords: Survival analysis, Machine learning

Functional Description: This library provides tools for analysing and interpreting machine-learning-based survival models. It features local and global state-of-the-art methods of explainability for ML models, as well as a common prediction analysis and evaluation framework. The code is available on GitHub, released under a permissive license (MIT).

URL: <https://github.com/TeamHeKA/survinsights>

Contact: Agathe Guilloux

7.1.7 survivalGPU

Name: survival analysis methods with support for GPU acceleration

Keywords: Survival analysis, GPU

Functional Description: This library implements survival analysis methods with support for GPU acceleration for faster computation. It is made available for R and Python, and has been tested and validated on a challenging use case on the SNDS (Système national de données de santé) dataset, where the goal was to associate side effects for a treatment without prior hypothesis. This task can become very expensive to compute on such a large dataset as the SNDS', and results showed run times short enough to make routine use possible. The code is available on GitHub, released under a permissive license (LGPL-2.1).

URL: <https://github.com/jeanfeudy/survivalGPU>

Contact: Jean Feydy

7.1.8 recforest

Name: Random forests for recurrent events

Keyword: Machine learning

Functional Description: This library implements the methodology called RecForest, a new ensemble approach for the analysis of recurrent events in a survival framework, with or without a terminal event. It outperforms traditional methods like the Cox model, in use cases with repeated events (hospital readmission for instance) and terminal events like death. It yields more accurate predictions, even with right-censored data, ultimately contributing to better decision-making and patient care. The code is available on CRAN, released under a permissive license (Apache-2.0), and features reasonable test coverage and documentation.

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

Contact: Sandrine Katsahian

7.1.9 KeOps

Keyword: High-Performance Computing

Functional Description: KeOps is a high-performance mathematical library that accelerates core operations far beyond standard numerical packages. It has been validated across diverse domains—including machine learning, computational biology, and shape analysis—and offers broad adoption thanks to support for MATLAB, R, Python, and C++. The library has 1.1k GitHub stars, over 800k PyPI downloads and 136k on CRAN, and received the 2023 French Open Science Award. Maintenance is jointly ensured by Benjamin Charlier (Inrae), Jean Feydy (Inria), and Joan Alexis Glaunès (UPC).

Contact: Jean Feydy

7.1.10 scikit-shapes

Name: tools for analysing shapes in 2D and 3D

Keyword: Shape recognition

Functional Description: This library provides tools for analysing shapes in 2D and 3D that may be encoded as point clouds, Gaussian splats, curves, surfaces or segmented images, and is tailored for registration, atlas construction and dimensionality reduction applications. `scikit-shapes` solves an important dilemma in geometry analysis: leverage the readability and ease-of-use of the Python language, whilst not sacrificing on performance compared to traditional C++ implementations. The code is available on GitHub, released under a permissive license (MIT), and has extensive documentation. It is still under very active development at HeKA, with the intent to open it to community contributions in 2026.

URL: <https://github.com/jeanfeudy/survivalGPU>

Contact: Jean Feydy

8 New results

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

8.1 Axis 1

Mohamed Imed Eddine Ghebriout, Gaël Guibon, Ivan Lerner, Emmanuel Vincent. QUARTZ: QA-based Unsupervised Abstractive Refinement for Task-oriented Dialogue Summarization. EMNLP 2025 : The 2025 Conference on Empirical Methods in Natural Language Processing, Nov 2025, Suzhou, China. (hal-05300943)

Dialogue summarization condenses conversations into concise text, reducing dialogue complexity in dialogue-heavy applications. Existing approaches heavily rely on costly human-written data, and the resulting summaries often lack task-specific focus, leading to suboptimal performance for downstream tasks, such as medical ones. In this paper, we introduce QUARTZ, a framework for task-oriented unsupervised dialogue summarization. QUARTZ starts by generating multiple summaries and task-specific question-answer pairs using large language models (LLMs). Summaries are evaluated by having the LLMs respond to task-related questions before (i) selecting the best candidate responses and (ii) identifying the most informative summary. Finally, we finetune the best LLM on the selected summaries. When validated on multiple datasets, QUARTZ achieves competitive zero-shot performance, rivaling fully-supervised State-of-the-Art (SoTA) approaches.

Participants: Ivan Lerner.

Jong Ho Jhee, Alberto Megina, Pacôme Constant Dit Beaufile, Matilde Karakachoff, Richard Redon, et al.. Predicting clinical outcomes from patient care pathways represented with temporal knowledge graphs. ESWC 2025 - 22nd European Semantic Web Conference, Jun 2025, Portorož, Slovenia. pp.282-300

With the increasing availability of healthcare data, predictive modeling finds many applications in the biomedical domain, such as the evaluation of the level of risk for various conditions, which in turn can guide clinical decision making. However, it is unclear how knowledge graph data representations and their embedding, which are competitive in some settings, could be of interest in biomedical predictive modeling. Method: We simulated synthetic but realistic data of patients with intracranial aneurysm and experimented on the task of predicting their clinical outcome. We compared the performance of various classification approaches on tabular data versus a graph-based representation of the same data. Next, we investigated how the adopted schema for representing first individual data and second temporal data impacts predictive performances. Results: Our study illustrates that in our case, a graph representation and Graph Convolutional Network (GCN) embeddings reach the best performance for a predictive task from observational data. We emphasize the importance of the adopted schema and of the consideration of literal values in the representation of individual data. Our study also moderates the relative impact of various time encoding on GCN performance.

Participants: Jong Ho Jhee, Adrien Coulet.

8.2 Axis 2

Do MH, Feydy J, Mula O, "Sparse Wasserstein barycenters and application to reduced order modeling", Journal of Scientific Computing 102, 64 (2025). online: [10.1007/s10915-024-02766-0](https://doi.org/10.1007/s10915-024-02766-0).

We develop a general theoretical and algorithmic framework for sparse approximation and structured prediction in spaces of probability measures with Wasserstein barycenters. The barycenters are sparse in the sense that they are computed from an available dictionary of measures, but the approximations only involve a reduced number of atoms. We show that the best reconstruction from the class of sparse barycenters is characterized by a notion of best n-term barycenter which we introduce, and which can be understood as a natural extension of the classical concept of best n-term approximation in Banach spaces. We show that the best n-term barycenter is the minimizer of a highly non-convex, bi-level optimization problem, and we develop algorithmic strategies for practical numerical computation. We next leverage this approximation tool to build interpolation strategies that involve a reduced computational cost, and that can be used for structured prediction, and metamodeling of parametrized families of measures. We illustrate the potential of the method

through the specific problem of Model Order Reduction (MOR) of parametrized PDEs. Since our approach is sparse, adaptive and preserves mass by construction, it has potential to overcome known bottlenecks of classical linear methods in hyperbolic conservation laws transporting discontinuities. It also paves the way towards MOR for measure-valued PDE problems such as gradient flows.

Participants: Jean Feydy.

Senellart, A., Chadebec, C., & Allasonnière, S. (2025). MultiVae: A Python package for Multimodal Variational Autoencoders on Partial Datasets. *Journal of Open Source Software*, 10(110), 7996.

In recent years, multimodal machine learning has seen significant growth, especially in representation learning and data generation. Recently, Multimodal Variational Autoencoders (VAEs) have been attracting growing interest for both tasks, thanks to their versatility, scalability, and interpretability as latent variable models. They are particularly useful in partially observed settings, such as medical applications, where available datasets are often incomplete (Antelmi et al., 2019; Lawry Aguila et al., 2023). We introduce MultiVae, an open-source Python library offering unified implementations of multimodal VAEs. It is designed for easy and customizable use of these models on fully or partially observed data. It facilitates the development and benchmarking of new algorithms by including standard benchmark datasets, evaluation metrics and tools for monitoring and sharing models.

Participants: Agathe Senellart, Stephanie Allasonniere.

Jemelen, E., Orchard, F., Madie, W., Valentin, B., Belin, J., Laas, E., ... & Guilloux, A. (2025). Evaluating breast cancer screening performance without registries using medico-administrative data. *Scientific Reports*, 15(1), 25096.

The French Breast Cancer Screening Program (DOCS) was created to detect early Breast Cancer (BC). Key performance indicators for digital mammography include sensitivity (SE), positive predictive value (PPV), interval cancer rate (ICR) and cancer detection rate (CDR). Calculating these metrics requires a linkage between screening data and BC registries; however, registries are scarce in France and often inaccessible for research. We therefore used medico-administrative data as an alternative. We linked regional screening data to the French National Health Data System (SNDS) between 2011 and 2020. Women were followed for 24 months post-screening. Screen-detected cancers and those identified with the SNDS were included. Performance metrics were calculated based on these linked datasets. A total of 252,786 screening exams were analyzed, covering 29,661-33,447 screenings annually, with a mean age of 61 years. SE was 77.9% (95% CI 76.3-79.3), indicating that approximately four in five cancers were detected through mammography. PPV was 19.8% (95% CI 19-20.5), meaning that one in five women with a positive screening test were confirmed with cancer within 24 months. CDR was 10.9 per 1000 exams (95% CI 10.5-11.3), equating to one detected case per 100 screenings. ICR was 2.4 per 1000 exams (95% CI 2.2-2.6), meaning that more than two interval cancers were detected per 1000 screenings. This identification approach using medico-administrative data offers a reproducible alternative for regions where cancer registries are unavailable. A future study applying this methodology in a registry-covered region could further validate the effectiveness of linking screenings to SNDS data for systematic cancer identification.

Participants: Emilien Jemelen, Sandrine Katsahian, Agathe Guilloux.

Houry, G., Boeken, T., Allasonnière, S., & Feydy, J. (2025, September). Untangling Vascular Trees for Surgery and Interventional Radiology. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 669-679). Cham: Springer Nature Switzerland.

The diffusion of minimally invasive, endovascular interventions motivates the development of visualization methods for complex vascular networks. We propose a planar representation of blood vessel trees which preserves the properties that are most relevant to catheter navigation: topology, length and curvature. Taking

as input a three-dimensional digital angiography, our algorithm produces a faithful two-dimensional map of the patient's vessels within a few seconds. To this end, we propose optimized implementations of standard morphological filters and a new recursive embedding algorithm that preserves the global orientation of the vascular network. We showcase our method on peroperative images of the brain, pelvic and knee artery networks. On the clinical side, our method simplifies the choice of devices prior to and during the intervention. This lowers the risk of failure during navigation or device deployment and may help to reduce the gap between expert and common intervention centers. From a research perspective, our method simulates the cadaveric display of artery trees from anatomical dissections. This opens the door to large population studies on the branching patterns and tortuosity of fine human blood vessels. Our code is released under the permissive MIT license as part of the scikit-shapes Python.

Participants: Guillaume Houry, Tom Boeken, Stephanie Allasonniere, Jean Feydy.

Perrine Chassat, Van Tuan Nguyen, Lucas Ducrot, Emilie Lanoy, Agathe Guilloux. Toward Valid Generative Clinical Trial Data with Survival Endpoints. ML4H - Machine Learning for Health Symposium, Dec 2025, San Diego (CA), United States.

Clinical trials face mounting challenges: fragmented patient populations, slow enrollment, and unsustainable costs, particularly for late-phase trials in oncology and rare diseases. While external control arms built from real-world data have been explored, a promising alternative is the generation of synthetic control arms using generative AI. A central challenge is the generation of time-to-event outcomes, which constitute primary endpoints in oncology and rare disease trials, but are difficult to model under censoring and small sample sizes. Existing generative approaches, largely GAN-based, are data-hungry, unstable, and rely on strong assumptions such as independent censoring. We introduce a variational autoencoder (VAE) that jointly generates mixed-type covariates and survival outcomes within a unified latent variable framework, without assuming independent censoring. Across synthetic and real trial datasets, we evaluate our model in two realistic scenarios: (i) data sharing under privacy constraints, where synthetic controls substitute for original data, and (ii) control-arm augmentation, where synthetic patients mitigate imbalances between treated and control groups. Our method outperforms GAN baselines on fidelity, utility, and privacy metrics, while revealing systematic miscalibration of type I error and power. We propose a post-generation selection procedure that improves calibration, highlighting both progress and open challenges for generative survival modeling.

Participants: Perrine Chassat, Van Tuan Nguyen, Sandrine Katsahian, Agathe Guilloux.

8.3 Axis 3

Comin L, Marie S, Ursino M, Zohar S, Tournier N, Comets E. Modeling Whole-Body Dynamic PET Microdosing Data to Predict the Whole-Body Pharmacokinetics of Glyburide in Humans. Clin Pharmacokinet. 2025 Nov;64(11):1709-1722

This work presents a novel pharmacokinetic modeling framework based on whole-body dynamic PET imaging with microdosed [^{14}C]glyburide in humans. By applying nonlinear mixed-effects modeling to multi-organ PET and arterial blood data, the study characterizes organ-level drug biodistribution and quantifies the effect of rifampicin on hepatic uptake, highlighting the potential of PET-based PK modeling as a translational tool complementary to conventional blood-based PK studies. This approach supports the integration of PET microdosing into model-informed drug development.

Participants: Léa Comin, Moreno Ursino, Sarah Zohar.

Bel Lassen P, Tropeano AI, Arnoux A, Lu E, Romengas L, Katsahian S, Ségrestin B, Lelièvre B, Mitanchez D, Gascoïn G, Poghosyan T, Lazzati A, Heude B, Nizard J, Czernichow S, Ciangura C,

Rives-Lange C. Maternal and neonatal outcomes of pregnancies after metabolic bariatric surgery: a retrospective population-based study. *Lancet Reg Health Eur.* 2025 Mar 22;51:101263.

This nationwide study evaluates maternal and neonatal outcomes of pregnancies following metabolic bariatric surgery in France. While post-surgery pregnancies were associated with reduced risks of gestational hypertension, preeclampsia, and gestational diabetes, they were also linked to increased risks of small-for-gestational-age infants, prematurity, stillbirth, and perinatal death. These neonatal risks were higher with shorter intervals between surgery and conception, gastric bypass procedures, and malnutrition.

Participants: Sandrine Katsahian, Andrea Lazzati, Claire Rives-Lange.

Boers M, Rochereau A, Stuwe L, Miguel LS, Klucken J, Mezei F, Fabiano J, Boulet S, Perchant A, Tarricone R, Petracca F, Hoefgen B, Collignon C, Zohar S; European Taskforce for Harmonised Evaluation of Digital Medical Devices (DMDs) (EvalEUDMD). Classification grid and evidence matrix for evaluating digital medical devices under the European union landscape. *NPJ Digit Med.* 2025 May 24;8(1):304

This work presents the development of a harmonised European taxonomy and evidence-based evaluation framework for Digital Medical Devices (DMDs) to support their integration into healthcare systems across the EU. Based on a comprehensive mapping of existing frameworks, a survey of Health Technology Assessment (HTA) bodies, and expert consensus, the Common European Classification Grid for DMDs (CEUGrid-DMD) and its associated Evidence Matrix were created. These tools aim to standardise scientific assessment practices and facilitate convergence of DMD evaluation within the context of the EU HTA Regulation.

Participants: Sandrine Boulet, Sarah Zohar.

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

CIFRE contact and similar PhD fundings (range 45-75k€)

Dassault Systèmes - S. Katsahian and A. Guilloux are co-supervising Tristan Margaté on the project: Prédiction dynamique de survie pour le suivi individualisé de patients atteints de cancer, prise en compte de variables longitudinales.

GE Healthcare - S. Allasonnière is co-supervising Theau Blanchard on the project: Virtual liver tumor pathology using self-supervised learning and multimodal data integration

Dassault Systèmes - J. Feydy supervises Louis Goldenberg on the population-wide statistical analysis of blood vessel networks. Formally, this is a collaboration contract instead of a genuine CIFRE contract funded by the ANRT.

Doctolib, - A. Coulet and I. Lerner co-supervise Louise Durand--Janin and Foucauld Estignard CIFRE PhDs on the diagnosis decision support from generative model and medical knowledge, and the mitigation of LLM uncertainty, respectively.

Epiconcept - S. Katsahian and A. Guilloux are co-supervising Emilien Jemelen on the project: Evaluation of the contribution of the use of artificial intelligence in the French breast cancer screening program.

Sanofi - S. Katsahian and A. Guilloux are co-supervising Jean-Baptiste Baitairian on the project: Quantitative Bias Assessment for causal inference.

Withings - A. Guilloux and S. Katsahian are co-supervising Philomène Letzelter on the project: Caractérisation et modélisation de la progression de la ménopause par analyse de données longitudinales issues de dispositifs connectés de santé Withings.

CEMKA-eval - A.S. Jannot supervises Corentin Faujour on the project : Development of Classifiers from the French National Health Data System.

MyFit Solutions - J. Feydy supervises Hadrien Bigo-Balland on the automated extraction of anatomical measurements from 3D scans reconstructed by a smartphone, with applications to the manufacturing of therapeutic garments (compression gloves, etc.).

Participants: Stephanie Allasonniere, Adrien Coulet, Jean Feydy,, Agathe Guilloux, Anne-Sophie Jannot, Sandrine Katsahian, Jean-Baptiste Baitairian, Hadrien Bigo-Balland, Théau Blanchard, Louise Durand-Janin, Foucauld Estignard, Corentin Faujour, Louis Goldenberg, Emilien Jemelen, Philomène Letzelter, Tristan Margate. .

9.2 Bilateral Grants with Industry

iDEMO FRANCE 2030 MEDITWIN - Dassault Systems **Virtual twin for personalised medicine (stating in 2024, Team members involved: S. Zohar (Partner leader), S. Allasonnière, A. Guilloux, S. Katashian and M. Ursino. 2.5M€ / >100M€)**. MEDITWIN is a collaborative project funded by BPI ("Banque Publique d'Investissement") with Dassault Systems (leader of the projects), Inria, IHU institutes across France and Medtech startups. The aim of the project is to provide a digital platform relying on virtual twins of individuals who faithfully reproduce their state of health and which make it possible to test different therapeutic options.

Combo - Sanofi **Evaluating drug combinations in oncology with Real-World Data and state-of-the-art knowledge (starting spring 2023, Team members involved: A. Coulet (Partner leader), S. Zohar (co-lead) and M. Ursino. 298k€ / 970k€)** Combo is a collaborative project with Sanofi, the HDH and the cancer institute of Lyon and HeKA. The objective of the project is to identify promising families of drug combinations in oncology using multisource and multi-modal data prediction, including RWD (data from Lyon cancer centre) and public databases.

9.3 Technology transfer and socio-economic impact

We highlight here our patents and startup initiatives.

HeKA members have been collaborating with SME and Industry. Some of these collaborating work have resulted in patents, as follows:

- EP4526469A1 "Methods for assessing the exhaustion of hematopoietic stems cells induced by chronic inflammation". Inventors: M. Cavazzana, E. Six, A. Guilloux, A. Denis, S. Sobrino, A. Rausell, L. Martignett, A.Cortal
- US20250087352A2: "Characteristics of patient influencing disease progression". Inventors: A. Cottin, N. Pecuchet, M. Zullian, S. Katsahian, A. Guilloux

Regarding the start-up founding: S. Allasonnière is co-founder of Sonio which is a spin-off of her researches in collaboration with E. Le Pennec (Ecole Polytechnique). Birth defects affect 1 in 33 babies born in Europe and in the US. In 50% of cases, the defects had not been detected during the ultrasound, showing how complex fetal medicine is. Physicians in charge of prenatal diagnostics face hundreds of signs visible on an ultrasound. Sonio Diagnostics accompanies them daily, and guides them through 250 syndromes and 700 hundred anomalies. Sonio guides practitioners in charge of prenatal screening and diagnosis before, during and after the examination.

10 Partnerships and cooperations

10.1 International research visitors

10.1.1 Visits of international scientists

Tina Hernandez-Boussard

Status: Professor

Institution of origin: Stanford

Country: US

Dates: summer 2025

Context of the visit: This visit led to the writing of a 5-year joint research program about the consideration of clinical guidelines in decision support system and in particular the evaluation of their fairness and the mitigation of potential unfairness. We submitted this joint research project to the Inria International Chair Program for 2026-2031.

Participants: Adrien Coulet, Sarah Zohar.

10.2 European initiatives

EU INVENTS Horizon project - "Innovative designs, extrapolation, simulation methods and evidence-tools for rare diseases addressing regulatory needs" (2024-2029, PI: S. Zohar, other team members involved: A.S. Jannot (WP leader) and M. Ursino 1.3M€ / 8.8M€)

In this project (made of 15 partners) our objectives include improving the robustness of model-based treatment-effect estimation and extrapolation methods, and developing in silico trial workflows that combine modelling and simulation, clinical trial data, and RWD to address gaps in disease knowledge. We aim to strengthen the reliability of confirmatory trials in small populations by using validated and credible models, while proposing advanced evidence-synthesis approaches that integrate computational models, clinical studies, RWD, and virtual patient cohorts. In addition, we are developing evidence-assessment tools tailored to regulatory decision-making in rare diseases and ensuring that patient engagement is fully integrated throughout the process.

Participants: Anne-Sophie Jannot, Moreno Ursino, Sarah Zohar.

EU MSCA Doctoral Networks Orgestra project - "Organoid technologies for disease modeling, drug discovery and development for rare diseases" (2024-2027, Team members involved: S. Zohar (WP leader) and M. Ursino 565K€ / 3.5M€) In this project (involving 13 partners) we supervise 2 PhDs project which will propose statistical development for linking in silico trials to organoids data and innovative trial design (within joint doctoral degree with University of Utrecht) and estimand framework involving Bayesian principles on organoids data for clinical trial outcomes and models (within a joint doctoral degree with Katholieke Universiteit Leuven).

Participants: Moreno Ursino, Sarah Zohar.

EU FAIR Horizon 2020 project - "Flagelin aerosol therapy for stimulation of immunity as an alternative treatment against pneumonia with multidrug resistant bacteria" (2020 - 2026, PI: J.C. Sirard (Inserm), Team members involved: S. Zohar (WP leader) and M. Ursino. 306k€ / 10M€)

In this project (regrouping 14 EU participants), 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. By incorporating mechanistic approaches earlier in the development process along with a continuous learning modeling 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.

Participants: Moreno Ursino, Sarah Zohar.

EU Joint action ERDERA - "the European Rare Diseases Research Alliance" (2024-2030, Team members involved: S. Zohar and A.S. Jannot 150K€) This is a European partnership uniting over 170 public and private organizations across 37 countries around a single goal: turning cutting-edge science into tangible benefits for the thirty million Europeans living with a rare disease. We are involved in the WP19 on innovative methodologies for small sample population like in rare diseases.

Participants: Anne-Sophie Jannot, Sarah Zohar.

EU Joint action eCAN+ - "Enhancing the digital capabilities of cancer centers across Europe"(2025-2030, Team members involved: S. Zohar, L. Cadi 250K€) In this project the aim is to enhance the digital capabilities of cancer centers across Europe, with a particular attention paid to opportunities in Eastern Europe. Coordinated by Sciensano, this 4-year initiative brings together 81 partners from 23 European countries, including public health institutes, universities, hospitals, cancer centers, and patient associations. We are task leader in WP9 for defining a classification framework for DMD used in cancer. We are co-supervising L. Cadi with the "Digital Health Delegation" (DNS) at tyhe French Minister of Health.

Participants: Sarah Zohar.

EU IHI Realised - "compRehensive mEthodological and operational Approach to clinical trials in ultra rarE Diseases" (2025-2029, Team members involved: S. Zohar and S. Katsahian 150K€) This project unites nearly 40 partners from academia, regulatory bodies, clinical research institutes and hospitals, patient organizations, pharmaceutical companies or European Research Infrastructures to establish new gold standards for clinical trials in rare and ultra-rare diseases. We are involved in the WP6 on innovative methodologies under regulatory perspective and acceptability.

Participants: Sandrine Katsahian, Sarah Zohar.

10.3 National initiatives

SIRIC InsiTu - "Insights into cancer: from inflammation to Tumor" (SIRIC, WP leader S.Allassonniere, Team members involved: J Feydy. 250k€/6.558M€) In this project, our aim is to turn scientific knowledge into sustainable healthcare by advancing cancer prevention, interception, and treatment. We will provide novel frameworks for multiscale integration of molecular alterations, cellular processes, and tissue complexity to create non-invasive "virtual biopsies" to map tumor ecosystems and monitor therapeutic response. We will also identify novel biomarkers. The project is composed by seven institutions: INCa, Inserm, APHP.Nord, UPC, Sorbonne Université, Institut du Cancer and CNRS.

Participants: Stephanie Allasonniere, Jean Feydy.

SurvivalGPU – "Using Graphics Processing Units (GPUs) to scale up survival analysis to nation-wide cohorts" (PI: A.S. Jannot, Team member involved: J. Feydy, 150k€ EPI-PHARE) This project aims at providing a complete re-implementation of the Cox proportional hazards model that is around 100 times faster on GPU than the survival package on CPU. Going further, it supports time-varying drug exposures via the Weighted Cumulative Exposures model and is accessible via an R interface which is fully retro-compatible with that of the survival package.

Participants: Anne-Sophie Jannot, Jean Feydy.

MESSIDORE, BEEP - "Bayesian methods for Early Enriched Platform trials" (PI: M.Ursino, Team member involved: S. Zohar. 274k€/700k€) It aims at proposing innovative Bayesian enriched "platform" designs for early phase trials (namely, phase I/II), which are adapted to the clinical context (healthy volunteers, patients and different indications) and go towards precision medicine. Teams involved: ECSTRRA, DRIVE, BioMaps, APHP Service Hématologie Adultes, CHRU de Tours Service de Médecine Intensive-Réanimation.

Participants: Moreno Ursino, Sarah Zohar.

RHU OPERANDI - "Optimisation and imProved Efficacy of targeted RADioNuclide therapy in Digestive cancers by Imagomics" (Team member involved: S.Allasonniere (WP leader). 431k€/8.5M€) This project (which involves 10 partners) aims to improve targeted radionuclide therapy (TARE and PRRT) for advanced hepatocellular carcinoma and astroenteropancreatic neuroendocrine tumours by developing imaging biomarkers and modelling approaches that enable better patient stratification and early identification of responders. Methodology will try to understand fundamentally whether current patient management using CT/PET/MRI allows to predict response and survival using cutting edge imaging-based AI approaches in combination with data augmentation techniques to reach statistical significance.

Participants: Stephanie Allasonniere.

ANR AT2TA - "Analogies: from Theory to Tools and Applications" (ANR 2022, Team members involved: A. Coulet (WP leader). 166k€/670k€) The AT2TA projet aims at studying the role that machine learning can play in analogical reasoning; and the HeKA team is in charge of exploring this interplay in the healthcare domain. A PhD student, co-supervised with Inria Paris, IHU Imagine and Université de Lorraine, is investigating this topic.

Participants: Adrien Coulet.

ANR LLM4All - "Up-to-date LLM for all" (ANR 2023, Team members involved: I. Lerner (WP leader). 70k€/715€) LLM4All aims to develop up-to-date and open-source LLMs. It focuses in particular on models that achieve performance comparable to proprietary models, as well as on creating methods for automatic updating them and reducing their computational requirements. A PhD student, co-supervised with AP-HP and Loria, Nancy (CNRS, Université de Lorraine), is investigating the refinement of models for task in the medical emergency settings, such as dialogues summarization and triage decision modeling.

Participants: Ivan Lerner.

CASCADE - "Lung Cancer Screening in French women using low-dose CT and Artificial intelligence for Detection" (INCA14771 and by the French Ministry of Health financement dérogatoire SERI 2020, Team member involved: M.P. Revel (PI) 2.2M€) This project addresses the under-representation of women in lung cancer screening trials and the need to validate AI-assisted CT reading. This prospective cohort study recruits 2,400 at-risk women to compare AI-supported reading by trained radiologists with double reading by thoracic experts, and to evaluate AI as an autonomous reader, including for coronary artery calcification assessment.

Participants: Brigitte Sabatier.

PRISONCO - "Cancer care in prison" (INCA SHS-E-SP-RISP 2022, Team member involved: A. Lazzati. 66k€/407k€) This project examines access to cancer care among incarcerated individuals, a population with poorer health outcomes and limited data available in France. Using SNDS data, we analyse screening, diagnosis, treatment pathways, and supportive care, comparing inmates' trajectories with those of matched individuals from the general population.

Participants: Andrea Lazzati.

ARC "Accélération de la Recherche Clinique" (PI: A. Guilloux (HeKA), Team members involved: S. Katsahian, S. Allasonnière, 200k€, INCA funding). Real-world data (RWD) from electronic records, wearables, and other digital sources provide valuable insights into routine clinical practice and help define the target population of a trial. Artificial patient data and virtual control arms can simulate comparators, reducing recruitment needs. These approaches are especially useful in diseases with difficult enrolment, such as oncology. By shortening trial timelines and lowering costs, synthetic trials enhance industrial competitiveness. Ensuring their mathematical and clinical validity remains essential for producing reliable evidence.

Participants: Agathe Guilloux, Sandrine Katsahian, Stephanie Allasonniere.

RHU Rebone - "Preoperative 3D reconstruction in real time for a better reflexion in bone repair" (PI: Pr. Marc-Olivier Gauci, orthopedic surgeon in the Nice hospital, Team member involved: J. Feydy, 200k€/8.3M€). ReBone aims to minimize complications and recovery time in complex bone trauma by developing and validating personalized, automated, and collaborative pre-operative planning, and its execution by the surgical team. In close collaboration with Hervé Delingette (Inria Université Côte d'Azur, Epione team), J. Feydy is responsible for segmenting bone fragments in the original, 3D Computed Tomography image. Downstream work-packages then combine this information with finite element simulations to propose fracture reduction strategies and personalized surgical implants.

Participants: Jean Feydy.

EDyLES - "Dynamic Models and Estimands for Longitudinal Epidemiological Studies" (PI : Cécile Proust, DR INSERM, Bordeaux Public Health, Team member involved: Agathe Guilloux) EDyLES is a

collaborative multi-disciplinary project gathering researchers in biostatistics, computer science, and epidemiology, along with clinicians in neurology and nephrology, coming from 12 research teams. The project's added value is manifold. From a biostatistical perspective, EDyLES will deliver novel analytical tools along with open-source software solutions for modeling the complex longitudinal data collected in cohort studies. These methodologies will leverage complementary assets of biostatistical models and ML techniques. The project will also address pivotal epidemiological questions developed directly by epidemiologists and clinicians from each domain. Beyond a better understanding of the disease progression and the mechanisms at play, they will help determine optimal therapeutic approaches in MSA and CKD, and preventive targets in cerebral aging. Although initially motivated by 3 pathologies, we anticipate that the techniques developed within EDyLES will benefit other areas of Public Health.

Participants: Agathe Guilloux.

ANR FRANCE 2030 PEPR Digital Health

The PEPR ("Programmes et équipements prioritaires de recherche") Digital Health2 aims at gathering national multidisciplinary community active in digital health for the development and exploitation of the concept of digital twin in health (started in September 2023). HeKA's involvement in this PEPR is through 5 projects, that are, A. Guilloux and S. Allassonnière are co-leading the project REWIND, S. Zohar is co-leading with R. Thiebaut (SISTM Inria Bordeaux) the project SMATCH, A. Coulet is WP leader in ShareFAIR, M. Ursino is WP leader in DIGIPATH and A.S. Jannot is WP leader in M4DI.

ShareFAIR - "Modelling and Extracting EHR data to create clinical pathways and to standardize clinical practice" (Team members involved: A. Coulet (WP leader). 135k€/1.8M€) This PEPR SN project, involving 9 partners, aims to automatically learn diagnostic and therapeutic protocols from EHRs by analysing the traces left by real-world clinical decisions. We combine reinforcement-learning with LLMs to extract, compare, and optimise clinical pathways.

Participants: Adrien Coulet.

NEUROVASC - "A 5P medicine program to reduce the global impact of intracranial aneurysm and stroke" (Team members involved: A. Coulet (partner). 10k€/1.5M€) The ambition of NEUROVASC is to set up an optimal digital ecosystem to develop predictive tools for 5P medicine (personalized, predictive, preventive, participatory, populational) against ICA (intracranial aneurysm) and stroke outcome, relying mostly on the distinctive resources developed by the French clinical network in neuroradiology. .

Participants: Adrien Coulet.

REWIND – "Modelling multimodal longitudinal health data to predict patient response" (Team members: A. Allassonnière (PI), A. Guilloux (co-PI), 650k€/1.8M€) This PEPR SN project develops methodological and AI tools to analyse multimodal longitudinal data for early diagnosis, prognosis, and prediction of treatment response. It spans four axes: new time-to-event models integrating repeated measures; spatio-temporal feature extraction for disease and treatment dynamics; model-selection criteria for longitudinal settings; and interpretable deep-learning approaches combining mechanistic and data-driven components.

Participants: Agathe Guilloux, Stephanie Allassonniere.

DIGIPHAT - "Modelling multi-scale pharmacological data to predict patient's response to treatments" (Team members involved: M. Ursino (WP leader), S Zohar, 130k€/1.8M€) This PEPR SN project develops approaches to integrate heterogeneous data, spanning multi-omic, pharmacokinetic, pharmacodynamic, clinical, and environmental sources, through a combination of advanced mechanistic modeling and machine-learning approaches. We primarily work on meta-model building, that is, linking models developed by other project partners (nine in total) across different levels (from microscopic to macroscopic) to enable the creation of "digital pharmacological twins."

Participants: Moreno Ursino, Sarah Zohar.

M4DI – "Modelling health data to identify phenotype-based patient subgroups" (Team members: A.-S. Jannot (WP leader), 140k€/1.8M€) This project develops methods to characterise clinical phenotypes and cluster patients using heterogeneous health-database variables across multiple modalities. We compare expert-driven and data-driven strategies by leveraging metadata (e.g., ontologies) and observed correlations between variables. The work evaluates weighted clustering approaches and latent-variable models (EM and variants), with a focus on interpretable methods that clearly expose how each feature contributes to phenotype definition.

Participants: Anne-Sophie Jannot.

SMATCH - "Clinical study and trial designs for the evaluation of models and DMDs for their translation to patient s care" (PI: S. Zohar (HeKA), coPI: R. Thiebaut (SISTM, U Bordeaux), Team members involved: M. Ursino, 660k€/3M€) In this PEPR SN project we develop adaptive randomized study designs to evaluate drugs and DMDs in a way that aligns with HTA expectations. We include secondary endpoints based on digital biomarkers collected manually or automatically through the DMDs. We explore multiple adaptive designs, including designs with or without direct comparisons, interim analyses for feasibility and efficacy, and platform trials sharing a common control arm.

Participants: Moreno Ursino, Sarah Zohar.

10.4 Public policy support

HeKA members are also actively involved in national health initiatives that play a critical role in shaping public policy. Notably, S. Zohar was a voting member for 6 years at Medical Device reimbursement at the French HTA, e.g., Cnedimts ("Commission nationale d'évaluation des dispositifs médicaux et des technologies de santé") at HAS ("Haute Autorité de Santé"). She reviewed the methodological aspect of medical devices applications asking to be reimbursed by the Health Insurance Fund as well as the "Forfait innovation" providing funding to promising technologies. S. Zohar was part of 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, she co-lead the WP2 on "Evidence in clinical evaluation" with Corinne Collignon (Head of the Digital Mission at HAS, France) and Barbara Höfgen (Head of the Unit DiGA-Fast-Track at Bfarm, Germany). This work and guidelines were published this year in Nature Digital Medicine. She has also collaborated with the governmental AIS ("Agence innovation en Santé") as part of the working group on "« Evolutions méthodologiques en recherche cliniques : valeur ; conditions de recours »".

Following these collaborations with public policy makers and regulators, S.Zohar has initiated a collaboration between HeKA and the "DNS" (Direction au Numérique en Santé) of the French Ministry of Health in which under the EU project eCAN+ was recruited L. Cadi that is spending half of her time at the DNS and half at HeKA. This resulting in collaborative work on methodological validation methods for

DMDs acceptable by the regulatory authorities. Likewise, S. Zohar has initiated a collaboration between HeKA and the “Mission Numérique en Santé” at the French HTA, HAS, under the PEPR Santé Numérique SMATCH project was recruited Y. Binnois that is spending half of his time at the HAS and half at HeKA. This collaborative work resulted in an “auto saisie” validated by the “college” of the HAS to work on a guideline for industry and health care professionals on methods and models to evaluate health care reorganization due to DMD embedding AI.

A.S. Jannot and A. Guilloux contribute as members of the Research Ethics Committee (CER) and the Scientific and Ethics Committee (CES) of AP-HP, Europe’s largest university hospital system.

11 Dissemination

11.1 Promoting scientific activities

11.1.1 Scientific events: organisation

S. Allasonnière served on the Scientific Advisory Board of the AI for Health Summit. A. Coulet co-organises the annual PFIA “IA & Santé” day (>70 participants), bringing together the French community working at the intersection of knowledge representation and machine learning in healthcare. J. Feydy organises a monthly seminar on 3D shape analysis at Inria Paris and is a member of the organising committee of the 11th International Conference on Curves and Surfaces (2026). A. Guilloux co-organises the 2026 workshop on recent advances in machine learning for healthcare.

Participants: Stephanie Allasonniere, Adrien Coulet, Jean Feydy, Agathe Guilloux.

11.1.2 Scientific events: selection

The team is also active in scientific communication and public engagement. Members are regularly invited to present their research or participate in roundtables targeting diverse audiences. For example, A. Coulet has been invited to a roundtable at AdoptAI, Paris an international symposium on societal applications of artificial intelligence. S. Allasonnière contributed to high-level discussions at the Convention on Health Analysis and Management (CHAM) and at the AFCDP annual colloquium on the European Health Data Space. A. Guilloux has been invited to roundtables at Futurapolis Santé and the HealthTech Days, and has participated in several public events focused on the role of AI in future medical innovation.

Participants: Stephanie Allasonniere, Agathe Guilloux.

11.1.3 Journal

Member of the editorial boards

A. Guilloux is Associate Editor for Biometrics and the International Journal of Biostatistics

Moreno Ursino is Associate Editor for Statistics in Medicine

S. Zohar is Associate Editor for Biometrics and Statistics in Biopharmaceutical Research

Participants: Moreno Ursino, Sarah Zohar, Agathe Guilloux.

Reviewer - reviewing activities All team members serve as reviewers for journals and conferences in their respective fields.

11.1.4 Leadership within the scientific community

S. Zohar was a voting member of the CNEDiMTS at the Haute Autorité de Santé.

Participants: Sarah Zohar.

11.1.5 Scientific expertise

Several members serve on high-level national committees: during the evaluation period, Team members also sit on ethics committees such as the Comité d'Éthique de la Recherche AP-HP, Centre and the Comité Scientifique et Éthique de l'EDS AP-HP, and serve on scientific boards including the ANR Generic Call (Axe H.14: Interfaces – mathematics and digital sciences with biology and health), ANR TSIA Call 2025, the FC3R (French Centre for the Replacement, Reduction and Refinement of Animal Testing) and MESSIDORE program (Méthodologie des Essais cliniques Innovants, Dispositifs, Outils et Recherches Exploitant les données de santé et biobanques), Bpifrance grant evaluations on AI-based medical devices

11.1.6 Research administration

S. Allassonnière is the vice president of innovation and valorization at UPC.

11.2 Teaching - Supervision - Juries - Educational and pedagogical outreach

A.S. Jannot and S. Katsahian lead the speciality “Méthodes et Outils pour les Données des Entrepôts en Santé” (ex - Big Data in Healthcare”) in the Master of Science of Public Health at UPC.

A.S. Jannot co-leads the quantitative biomedicine course, which is part from the medical degree course. She also leads a professional degree of “health data reuse” at UPC in collaboration with Marseille and Bordeaux Universities.

S. Allassonniere coordinates the SMPS Bioentrepreneur (UPC). S. Allassonniere runs a course focusing on the analysis of real-life health data within the M2 MVA (Mathematics, Vision, Learning).

J. Feydy runs a class of geometric data analysis in the same program, while also teaching linear statistical models to 2nd year medical students at UPC.

Since 2023 Moreno Ursino teaches the course “Clinical trials” at the third year of ENSAI (École Nationale de la Statistique et de l'Analyse de l'Information).

11.3 Publications 2025

Due to a bug in HAL, 88 publications of the team in 2025 could not be up-loaded to this report.

12 Scientific production

12.1 Major publications

- [1] P. Bel Lassen, A.-I. Tropeano, A. Arnoux, E. Lu, L. Romengas, S. Katsahian, B. Ségrestin, B. Lelièvre, D. Mitanchez, G. Gascoin, T. Poghosyan, A. Lazzati, B. Heude, J. Nizard, S. Czernichow, C. Ciangura and C. Rives-Lange. ‘Maternal and neonatal outcomes of pregnancies after metabolic bariatric surgery: a retrospective population-based study’. In: *The Lancet Regional Health - Europe* 51 (Apr. 2025), p. 101263. DOI: [10.1016/j.lanepe.2025.101263](https://doi.org/10.1016/j.lanepe.2025.101263). URL: <https://hal.science/hal-05383106>.
- [2] M. Boers, A. Rochereau, L. Stuwe, L. S. Miguel, J. Klucken, F. Mezei, J. Fabiano, S. Boulet, A. Perchant, R. Tarricone, F. Petracca, B. Hoefgen, C. Collignon and S. Zohar. ‘Classification grid and evidence matrix for evaluating digital medical devices under the European union landscape’. In: *npj Digital Medicine* 8.1 (24th May 2025), p. 304. DOI: [10.1038/s41746-025-01697-w](https://doi.org/10.1038/s41746-025-01697-w). URL: <https://hal.science/hal-05084641>.

- [3] P. Chassat, V. T. Nguyen, L. Ducrot, E. Lanoy and A. Guilloux. ‘Toward Valid Generative Clinical Trial Data with Survival Endpoints’. In: *ML4H - Machine Learning for Health Symposium*. San Diego (CA), United States, 2025. URL: <https://hal.science/hal-05375245>.
- [4] L. Comin, S. Marie, M. Ursino, S. Zohar, N. Tournier and E. Comets. ‘Modeling Whole-Body Dynamic PET Microdosing Data to Predict the Whole-Body Pharmacokinetics of Glyburide in Humans’. In: *Clinical Pharmacokinetics* (17th Sept. 2025). DOI: [10.1007/s40262-025-01562-9](https://doi.org/10.1007/s40262-025-01562-9). URL: <https://cea.hal.science/cea-05273848>.
- [5] M. I. E. Ghebriout, G. Guibon, I. Lerner and E. Vincent. ‘QUARTZ: QA-based Unsupervised Abstractive Refinement for Task-oriented Dialogue Summarization’. In: *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*. EMNLP 2025 : The 2025 Conference on Empirical Methods in Natural Language Processing. Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing. Suzhou, China, 4th Nov. 2025. URL: <https://hal.science/hal-05300943>.
- [6] G. Houry, T. Boeken, S. Allasonnière and J. Feydy. ‘Untangling Vascular Trees for Surgery and Interventional Radiology’. In: *MICCAI 2025 - 28th International Conference on Medical Image Computing and Computer Assisted Intervention*. Vol. 15968. Lecture Notes in Computer Science. Daejeon, South Korea: Springer Nature Switzerland, 21st Sept. 2025, pp. 669–679. DOI: [10.1007/978-3-032-05114-1_64](https://doi.org/10.1007/978-3-032-05114-1_64). URL: <https://hal.science/hal-05271418>.
- [7] E. Jemelen, F. Orchard, W. Madie, B. Valentin, J. Belin, E. Laas, G. Jeannerod, P. Mares, S. Katsahian and A. Guilloux. ‘Evaluating breast cancer screening performance without registries using medico-administrative data’. In: *Scientific Reports* 15.1 (11th July 2025), p. 25096. DOI: [10.1038/s41598-025-10115-w](https://doi.org/10.1038/s41598-025-10115-w). URL: <https://hal.science/hal-05369094>.
- [8] J. H. Jhee, A. Megina, P. Constant Dit Beaufils, M. Karakachoff, R. Redon, A. Gaignard and A. Coulet. ‘Predicting clinical outcomes from patient care pathways represented with temporal knowledge graphs’. In: *ESWC 2025 - 22nd European Semantic Web Conference*. Vol. LNCS-15718. Lecture Notes in Computer Science. Portorož, Slovenia: Springer Nature Switzerland, 30th Apr. 2025, pp. 282–300. DOI: [10.1007/978-3-031-94575-5_16](https://doi.org/10.1007/978-3-031-94575-5_16). URL: <https://hal.science/hal-05136324>.
- [9] A. Senellart, C. Chadebec and S. Allasonnière. ‘MultiVae: A Python package for Multimodal Variational Autoencoders on Partial Datasets.’ In: *Journal of Open Source Software* 10.110 (5th June 2025), p. 7996. DOI: [10.21105/joss.07996](https://doi.org/10.21105/joss.07996). URL: <https://hal.science/hal-05372674>.

12.2 Publications of the year

International journals

- [10] J. Abisoror, R. Asgari, Q. de Baynast, C. Le Beller, N. Karam, B. Sabatier and T. Caruba. ‘Hémorragie digestive due à une probable interaction médicamenteuse warfarine/oseltamivir : à propos d’un cas clinique’. In: *La Revue de Médecine Interne* (Oct. 2025). DOI: [10.1016/j.revmed.2025.09.008](https://doi.org/10.1016/j.revmed.2025.09.008). URL: <https://hal.science/hal-05368918>.
- [11] P. Addeo, M. Muzzolini, C. Laurent, B. Heyd, A. Sauvanet, J. Garnier, M. S. Alfano, S. Gaujoux, C. de Ponthaud, U. Marchese, D. da Silva, E. Buc, R. Souche, J. M. Fabre, P.-E. Colombo, L. Ferre, M. Foguene, C. Hubert, M. E. Amrani, S. Truant, L. Schwartz, N. Regenet, A. Dupre, R. Brustia, R. Cherif, J. Navez, B. Darnis, O. Facy, R. Grellet, G. Piessen, J. Veziant, R. Rhaïem, R. Kianmanesh, E. Fernandez-De-Sevilla, M. Gelli, A. Thaïbi, P. Georges, J. Y. Mabrut, M. Lesurtel, A. Doussot and P. Bachellier. ‘ASO Visual Abstract: Prognosis Associated with Complete Pathological Response Following Neoadjuvant Treatment for Pancreatic Adenocarcinoma in the FOFLIRINOX Era: The Multicenter TONO Study’. In: *Annals of Surgical Oncology* (29th Jan. 2025). DOI: [10.1245/s10434-025-16873-1](https://doi.org/10.1245/s10434-025-16873-1). URL: <https://hal.science/hal-04927134>.

- [12] P. Addeo, M. Muzzolini, C. Laurent, B. Heyd, A. Sauvanet, J. Garnier, M. S. Alfano, S. Gaujoux, C. de Ponthaud, U. Marchese, D. da Silva, E. Buc, R. Souche, J. M. Fabre, P.-E. Colombo, L. Ferre, M. Foguette, C. Hubert, M. El Amrani, S. Truant, L. Schwartz, N. Regenet, A. Dupre, R. Brustia, R. Cherif, J. Navez, B. Darnis, O. Facy, R. Grellet, G. Piessen, J. Veziat, R. Rhaïem, R. Kianmanesh, E. Fernandez-De-Sevilla, M. Gelli, A. Taïbi, P. Georges, J. Y. Mabrut, M. Lesurtel, A. Doussot and P. Bachellier. 'ASO Author Reflections: Complete Pathological Response after Neoadjuvant Treatment for Pancreatic Ductal Adenocarcinomas—Curative Surgery in Cured Patients?' In: *Annals of Surgical Oncology* (1st Feb. 2025). DOI: [10.1245/s10434-025-16925-6](https://doi.org/10.1245/s10434-025-16925-6). URL: <https://hal.science/hal-04927128>.
- [13] P. Addeo, M. Muzzolini, C. Laurent, B. Heyd, A. Sauvanet, J. Garnier, M. S. Alfano, S. Gaujoux, C. de Ponthaud, U. Marchese, D. da Silva, E. Buc, R. Souche, J. M. Fabre, P.-E. Colombo, L. Ferre, M. Foguette, C. Hubert, M. El Amrani, S. Truant, L. Schwartz, N. Regenet, A. Dupre, R. Brustia, R. Cherif, J. Navez, B. Darnis, O. Facy, R. Grellet, G. Piessen, J. Veziat, R. Rhaïem, R. Kianmanesh, E. Fernandez-De-Sevilla, M. Gelli, A. Taïbi, P. Georges, J. Y. Mabrut, M. Lesurtel, A. Doussot and P. Bachellier. 'Prognosis Associated with Complete Pathological Response Following Neoadjuvant Treatment for Pancreatic Adenocarcinoma in the FOFLIRINOX Era: the Multicenter TONO Study'. In: *Annals of Surgical Oncology* (8th Jan. 2025). DOI: [10.1245/s10434-024-16735-2](https://doi.org/10.1245/s10434-024-16735-2). URL: <https://hal.science/hal-04927152>.
- [14] J. Anibal, H. Huth, T. Boeken, D. Daye, J. Gichoya, F. G. Muñoz, J. Chapiro, B. Wood, D. Sze and K. Hausegger. 'Interventional Radiology Reporting Standards and Checklist for Artificial Intelligence Research Evaluation (iCARE)'. In: *CardioVascular and Interventional Radiology* (25th June 2025). DOI: [10.1007/s00270-024-03956-x](https://doi.org/10.1007/s00270-024-03956-x). URL: <https://hal.science/hal-05158206>.
- [15] T. Baron, C. de Ponthaud and S. Gaujoux. 'Intraoperative pancreatoscopy during left pancreatectomy for intraductal papillary mucinous neoplasia (with video)'. In: *Journal of Visceral Surgery* (Jan. 2025). DOI: [10.1016/j.jviscsurg.2025.01.004](https://doi.org/10.1016/j.jviscsurg.2025.01.004). URL: <https://hal.science/hal-04927130>.
- [16] P. Bel Lassen, A. Bellicha, C. Carette, E. Roze, J.-M. Oppert, S. Czernichow, K. Clément, C. Poitou and C. Rives-Lange. 'Addressing Explicit Weight Bias in Medical Students: Contribution of Demographics and Educational Factors'. In: *Obesity Facts* (25th Sept. 2025), pp. 1–9. DOI: [10.1159/000547818](https://doi.org/10.1159/000547818). URL: <https://hal.sorbonne-universite.fr/hal-05365947>.
- [17] P. Bel Lassen, A.-I. Tropeano, A. Arnoux, E. Lu, L. Romengas, S. Katsahian, B. Ségrestin, B. Lelièvre, D. Mitanché, G. Gascoïn, T. Poghosyan, A. Lazzati, B. Heude, J. Nizard, S. Czernichow, C. Ciangura and C. Rives-Lange. 'Maternal and neonatal outcomes of pregnancies after metabolic bariatric surgery: a retrospective population-based study'. In: *The Lancet Regional Health - Europe* 51 (Apr. 2025), p. 101263. DOI: [10.1016/j.lanep.2025.101263](https://doi.org/10.1016/j.lanep.2025.101263). URL: <https://hal.science/hal-05383106>.
- [18] S. Berdot, G. Perrin, E. Disse, A. Arnoux, E. Bertin, A.-L. Borel, M.-C. Brindisi, A. de Luca, E.-j. Laforgue, J.-D. Lalau, Y. Matussiere, E. Montastier, A. Sallé, B. Segrestin, V. Suteau, S. Czernichow and B. Sabatier. 'Delphi Consensus Among French Obesity Experts on Clinical Recommendations for Drug Prescription in Patients With Severe Obesity'. In: *Obesity Reviews* 26.12 (8th July 2025), e13977. DOI: [10.1111/obr.13977](https://doi.org/10.1111/obr.13977). URL: <https://hal.science/hal-05228358>.
- [19] M. Boers, A. Rochereau, L. Stuwe, L. S. Miguel, J. Klucken, F. Mezei, J. Fabiano, S. Boulet, A. Perchant, R. Tarricone, F. Petracca, B. Hoefgen, C. Collignon and S. Zohar. 'Classification grid and evidence matrix for evaluating digital medical devices under the European union landscape'. In: *npj Digital Medicine* 8.1 (24th May 2025), p. 304. DOI: [10.1038/s41746-025-01697-w](https://doi.org/10.1038/s41746-025-01697-w). URL: <https://hal.science/hal-05084641>.
- [20] A. Bonan, S. Huynh, J. Pineau, B. Sabatier and N. Martelli. 'Impact du Règlement Européen sur les dispositifs médicaux sur leur disponibilité : scoping review et analyse sur deux ans du risque pour les patients dans le domaine cardiovasculaire'. In: *Annales Pharmaceutiques Françaises* 83.3 (May 2025), pp. 428–437. DOI: [10.1016/j.pharma.2025.01.008](https://doi.org/10.1016/j.pharma.2025.01.008). URL: <https://hal.science/hal-05368921>.

- [21] N. Bougas, T. S. Dabakuyo Yonli, E. Assogba, N. Haddy, F. de Vathaire, A. Janah, M. Michel, A. Bourmaud, H. Jeanningros, M. Ursino, R. Debailly and A. Dumas. ‘Effects of the right to be forgotten for childhood or breast cancer survivors: results of a quasi-experimental study in France’. In: *Journal of Cancer Survivorship* (4th June 2025). DOI: [10.1007/s11764-025-01830-4](https://doi.org/10.1007/s11764-025-01830-4). URL: <https://hal.science/hal-05109463>.
- [22] S. Boulet and A. Chambaz. ‘Draw Me a Simulator’. In: *Computo* (12th Sept. 2025). DOI: [10.57750/w1hj-dw22](https://doi.org/10.57750/w1hj-dw22). URL: <https://hal.science/hal-05252426>.
- [23] A. Cantais, S. Pillet, J. Rigaille, F. Angoulvant, C. Gras-Le-Guen, P. Cros, C. Thuiller, C. Molly, L. Tripodi, A. Desbree, N. Annino, P. Verhoeven, A. Carricajo, T. Bourlet, C. Chapelle, I. Claudet, A. Garcin, J. Izopet, O. Mory and B. Pozzetto. ‘Impact of respiratory pathogens detection by a rapid multiplex polymerase chain reaction assay on the management of community-acquired pneumonia for children at the paediatric emergency department. A randomized controlled trial, the Optimization of Pneumonia Acute Care (OPTIPAC) study’. In: *Clinical Microbiology and Infection* 31.1 (Jan. 2025), pp. 64–70. DOI: [10.1016/j.cmi.2024.08.001](https://doi.org/10.1016/j.cmi.2024.08.001). URL: <https://hal.science/hal-04895967>.
- [24] G. Chassagnon. ‘Are we really ready to use radiomics and deep learning for clinical decision support in radiology?’ In: *Diagnostic and Interventional Imaging* 106.11 (Nov. 2025), pp. 371–372. DOI: [10.1016/j.diii.2025.07.004](https://doi.org/10.1016/j.diii.2025.07.004). URL: <https://hal.science/hal-05364848>.
- [25] G. Chassagnon, R. Marini, V. Ong, J. da Silva, D. Habip Gatenyo, I. Honore, R. Kanaan, N. Carlier, J. Fesenbeckh, E. Burnet, M.-P. Revel, C. Martin and P.-R. Burgel. ‘Identification of structural predictors of lung function improvement in adults with cystic fibrosis treated with elexacaftor-tezacaftor-ivacaftor using deep-learning’. In: *Diagnostic and Interventional Imaging* (Sept. 2025). DOI: [10.1016/j.diii.2025.09.003](https://doi.org/10.1016/j.diii.2025.09.003). URL: <https://hal.science/hal-05364842>.
- [26] P. Chassat, J. Park and N. J.-B. Brunel. ‘Intrinsic Estimation of Frenet Curvatures for 3D Trajectory Analysis’. In: *Journal of Mathematical Imaging and Vision* 68.1 (5th Jan. 2026), p. 2. DOI: [10.1007/s10851-025-01277-3](https://doi.org/10.1007/s10851-025-01277-3). URL: <https://hal.science/hal-05466937>.
- [27] B. Chen, L. Chantal Schneider, C. Röver, E. Comets, M. Christian Elze, A. Hooker, J. Inthout, A.-S. Jannot, D. Julkowska, Y. Mimouni, M. Savelieva, N. Stallard, M. Ursino, M. Vandemeulebroecke, S. Weber, M. Posch, S. Zohar and T. Friede. ‘In Silico Clinical Trials in Drug Development: A Systematic Review’. In: *Therapeutic Innovation & Regulatory Science* (2025), Online ahead of print. DOI: [10.1007/s43441-025-00893-w](https://doi.org/10.1007/s43441-025-00893-w). URL: <https://hal.science/hal-05405192>.
- [28] C. Ciangura, C. Rives-Lange, B. Lelièvre, M. Laville, M.-A. Charles, B. Ségrestin and B. Heude. ‘Why and How to Improve Nutritional Care for Pregnant Women after Bariatric Surgery: The NUMASURG Study Protocol.’ In: *Obesity Facts* (2025), pp. 1–11. DOI: [10.1159/000545123](https://doi.org/10.1159/000545123). URL: <https://inserm.hal.science/inserm-05302431>.
- [29] L. Comin, S. Marie, M. Ursino, S. Zohar, N. Tournier and E. Comets. ‘Modeling Whole-Body Dynamic PET Microdosing Data to Predict the Whole-Body Pharmacokinetics of Glyburide in Humans’. In: *Clinical Pharmacokinetics* (17th Sept. 2025). DOI: [10.1007/s40262-025-01562-9](https://doi.org/10.1007/s40262-025-01562-9). URL: <https://cea.hal.science/cea-05273848>.
- [30] J. Didier, J. Murriss, H. Schopper, E. Guérin, N. Flajoliet, M. Amiot, B. Hermann, S. Germain, A. Cariou, J.-L. Diehl, S. Katsahian and N. Bréchet. ‘Fluid balance after cardiac arrest: Any impact on outcome? Insights from the MIMIC IV database’. In: *Resuscitation Plus* 25 (Sept. 2025), p. 101037. DOI: [10.1016/j.resplu.2025.101037](https://doi.org/10.1016/j.resplu.2025.101037). URL: <https://college-de-france.hal.science/hal-05285460>.
- [31] K. Dotsikas, E. Lanoy, E. Lequy, E. Wiernik and M. Mary-Krause. ‘Neighbourhood deprivation and smoking cessation: a survival analysis using the French constances cohort’. In: *BMC Public Health* 25.1 (3rd July 2025), p. 2317. DOI: [10.1186/s12889-025-23261-7](https://doi.org/10.1186/s12889-025-23261-7). URL: <https://hal.science/hal-05148767>.
- [32] B. Duputel, N. Stallard, F. Montestruc, S. Zohar and M. Ursino. ‘A Seamless hybrid Phase II/III design with Bayesian interim subgroup selection’. In: *Statistics in Medicine* 44.13-14 (3rd June 2025). DOI: [10.1002/sim.70144](https://doi.org/10.1002/sim.70144). URL: <https://hal.science/hal-04695747>.

- [33] T. Fabacher, E.-A. Sauleau, E. Arcay, B. Faye, M. Alter, A. Chahard, N. Miraillet, A. Coulet and A. Névéal. 'Efficient extraction of medication information from clinical notes: an evaluation in 2 languages'. In: *Journal of the American Medical Informatics Association* (10th Oct. 2025), ocaf113. DOI: [10.1093/jamia/ocaf113](https://doi.org/10.1093/jamia/ocaf113). URL: <https://inria.hal.science/hal-05375038>.
- [34] F. Gaudfernau, E. Blondiaux, S. Allasonnière and E. Le Pennec. 'Wavelet-Based Multiscale Flow For Realistic Image Deformation in the Large Diffeomorphic Deformation Model Framework'. In: *Journal of Mathematical Imaging and Vision* 67.2 (9th Jan. 2025), p. 10. doi: [10.1007/s10851-024-01219-5](https://doi.org/10.1007/s10851-024-01219-5). URL: <https://hal.science/hal-04919330>.
- [35] H. Ghannouchi, J. Wang, T. Boeken, E. Rius, X. Guerra, K. Premat, J. Allard, P.-M. Chiaroni, S. Lenck, M. Elhorany, A. Talbi, B. Mathon, A. Nouet, A. Jacquens, M. Drir, N.-A. Sourour, F. Clarençon and E. Shotar. 'Correlating brain arteriovenous malformations location with clinical and angio-architectural characteristics: A voxel-based lesion symptom mapping approach'. In: *Journal of Neuroscience Methods* 422 (Oct. 2025), p. 110519. DOI: [10.1016/j.jneumeth.2025.110519](https://doi.org/10.1016/j.jneumeth.2025.110519). URL: <https://hal.science/hal-05158204>.
- [36] I. Haffaf, J. Cortese, L. Garzelli, B. Mathon, G. Guillaumet, A. Dimancea, T. Boeken, S. Lenck, A. Jacquens, C. Amouyal, A. Talbi, K. Premat, N.-A. Sourour, V. Degos, F. Clarençon and E. Shotar. 'Predictive factors of external ventricular drain placement in ruptured brain arteriovenous malformations with intraventricular hemorrhage'. In: *Neurosurgical Review* 48.1 (10th July 2025), p. 552. DOI: [10.1007/s10143-025-03700-0](https://doi.org/10.1007/s10143-025-03700-0). URL: <https://hal.science/hal-05158202>.
- [37] S. Huynh, A. Bonan, J. Pineau, E. Marijon, B. Sabatier and N. Martelli. 'Retraitement des dispositifs médicaux à usage unique : points de vue et perceptions des cardiologues et pharmaciens hospitaliers'. In: *Annales Pharmaceutiques Françaises* 83.4 (June 2025), pp. 720–728. DOI: [10.1016/j.pharma.2025.01.011](https://doi.org/10.1016/j.pharma.2025.01.011). URL: <https://hal.science/hal-05368915>.
- [38] E. Jemelen, F. Orchard, W. Madie, B. Valentin, J. Belin, E. Laas, G. Jeannerod, P. Mares, S. Katsahian and A. Guilloux. 'Evaluating breast cancer screening performance without registries using medico-administrative data'. In: *Scientific Reports* 15.1 (11th July 2025), p. 25096. DOI: [10.1038/s41598-025-10115-w](https://doi.org/10.1038/s41598-025-10115-w). URL: <https://hal.science/hal-05369094>.
- [39] E. Karam, M. Hollenbach, C. Heise, E. Abou Ali, A. Gulla, B. Terris, F. Auriemma, F. Souche, C. de Ponthaud, T. Hank, F. Caillol, R. Rhaïem, A. Sauvanet, B. Napoléon, S. Regner and S. Gaujoux. 'Ampullary composite gangliocytoma/neuroma and neuroendocrine tumor management'. In: *Endocrine-Related Cancer* 32.3 (2nd Jan. 2025). DOI: [10.1530/ERC-24-0238](https://doi.org/10.1530/ERC-24-0238). URL: <https://hal.science/hal-04927153>.
- [40] H. Leroy, L. Z. Wang, A. Jimenez, N. Mohamedi, C. Papadacci, P. Julia, S. E. Batti, J. M. Alsac, J. Sitruk, A. Arnoux-Guenegou, P. Bruneval, E. Messas, T. Mirault, G. Goudot and M. Pernot. 'Assessment of microvascular flow in human atherosclerotic carotid plaques using ultrasound localization microscopy'. In: *EBioMedicine* (Jan. 2025). URL: <https://hal.science/hal-05374974>.
- [41] A. Lesaunier, J. Khlaut, C. Dancette, L. Tselikas, B. Bonnet and T. Boeken. 'Artificial intelligence in interventional radiology: Current concepts and future trends'. In: *Diagnostic and Interventional Imaging* 106.1 (Jan. 2025), pp. 5–10. DOI: [10.1016/j.diii.2024.08.004](https://doi.org/10.1016/j.diii.2024.08.004). URL: <https://hal.science/hal-05158239>.
- [42] L. Machado, L. Alberge, H. Philippe, E. Ferreres, J. Khlaut, J. Dupuis, K. Le Floch, D. Habip Gatenyo, P. Roux, J. Grégory, M. Ronot, C. Dancette, T. Boeken, D. Tordjman, P. Manceron and P. Hérent. 'A promptable CT foundation model for solid tumor evaluation'. In: *npj Precision Oncology* 9.1 (25th Apr. 2025), p. 121. DOI: [10.1038/s41698-025-00903-y](https://doi.org/10.1038/s41698-025-00903-y). URL: <https://hal.science/hal-05158219>.
- [43] U. Marchese, X. Lenne, G. Naveendran, S. Tzedakis, M. Gaillard, Y. Richa, L. Boyer, D. Theis, A. Bruandet, S. Truant, D. Fuks and M. El Amrani. 'Nationwide analysis of one-year mortality following pancreatectomy in 17,183 patients with pancreatic cancer'. In: *HPB* 27.1 (Jan. 2025), pp. 123–129. DOI: [10.1016/j.hpb.2024.10.011](https://doi.org/10.1016/j.hpb.2024.10.011). URL: <https://hal.science/hal-04935509>.

- [44] S. Mouyal, X. Guerra, T. Boeken, A. Di Gaeta, C. Bordeianu, M. Gargiulo, O. Pellerin, M. Sapoval and M. Al Ahmar. 'Endovascular treatment of a subclavian artery pseudoaneurysm in a patient with a pancoast tumor'. In: *Cvir Endovascular* 8.1 (18th June 2025), p. 52. DOI: [10.1186/s42155-025-00572-9](https://doi.org/10.1186/s42155-025-00572-9). URL: <https://hal.science/hal-05158209>.
- [45] L. Noé, Z. Valtuille, E. Lanoy, S. Katsahian and F. Kaguelidou. 'Evaluation of Methods for Joinpoint Analysis of Time Series Using Simulated and Real-World Data'. In: *Journal of Clinical Epidemiology* 187 (1st Sept. 2025). DOI: [10.1016/j.jclinepi.2025.111966](https://doi.org/10.1016/j.jclinepi.2025.111966). URL: <https://hal.science/hal-05374790>.
- [46] A. Nun, O. Birot, G. Guibon, F. Lapostolle and I. Lerner. 'SIMSAMU - A French medical dispatch dialog open dataset'. In: *Computer Methods and Programs in Biomedicine* 268 (Aug. 2025), p. 108857. DOI: [10.1016/j.cmpb.2025.108857](https://doi.org/10.1016/j.cmpb.2025.108857). URL: <https://inria.hal.science/hal-05325611>.
- [47] L. Parlati, C. Mouliade, E. Nguyen Khac, M. Collier, S. Tzedakis, S. Bouam, A. Courtois, M. Corouge, A. Louvet, S. Pol, P. Sogni, A. Benyamina, J. Rehm, P. Mathurin, V. Mallet, D. Karinthi, J. Nabarro, V. D. Venier and A. V. Pichard. 'Alcohol Rehabilitation Within 3 Months after Alcohol Hepatitis and Survival - A National Analysis'. In: *Clinical Gastroenterology and Hepatology* 23.5 (2025), 808–815.e5. DOI: [10.1016/j.cgh.2024.07.037](https://doi.org/10.1016/j.cgh.2024.07.037). URL: <https://u-picardie.hal.science/hal-04687148>.
- [48] C. Pastier, C. de Ponthaud, A. Nassar, O. Soubrane, A. Mazzotta, F.-R. Souche, L. Brunaud, R. Kianmanesh, L. Sulpice, L. Schwarz, E. Karam, E. Lermite, S. Dokmak, D. Fuks and S. Gaujoux. 'Laparoscopic pancreatic enucleation: how far should we go for parenchyma preservation? A study by the French National Association of Surgery'. In: *Surgical Endoscopy* 39.3 (13th Jan. 2025), pp. 1696–1708. DOI: [10.1007/s00464-024-11453-y](https://doi.org/10.1007/s00464-024-11453-y). URL: <https://hal.science/hal-04927145>.
- [49] G. Perrin, A. Thomas-Schoemann, F. Goldwasser, P. Boudou-Rouquette, N. Delanoy, C. Gervais, J. Touchard, R. Batista, B. Blanchet, J.-S. Giraud, B. Sabatier and V. Korb-Savoldelli. 'Factors associated with medication adherence trajectory among patients treated with oral anticancer drugs: a multicentre cohort study'. In: *Supportive Care in Cancer* 33.11 (15th Oct. 2025), p. 951. DOI: [10.1007/s00520-025-09996-4](https://doi.org/10.1007/s00520-025-09996-4). URL: <https://hal.science/hal-05368907>.
- [50] C. de Ponthaud, F. Menegaux and S. Gaujoux. 'Approaching Immediate Total Thyroidectomy in First-Side Loss of Neuromonitoring Signal With Caution'. In: *Archives of Otorhinolaryngology - Head & Neck Surgery* (9th Jan. 2025). DOI: [10.1001/jamaoto.2024.4674](https://doi.org/10.1001/jamaoto.2024.4674). URL: <https://hal.science/hal-04927151>.
- [51] C. de Ponthaud, A. Nassar, S. Dokmak, T. Bertrand, J. de Martino, C. Pastier, A. Castel, R. de Rosa, M. Boubaddi, A. Iben-Khayat et al. 'Conversion during Minimally Invasive Left Pancreatectomy: a nationwide study of causes and consequences'. In: *Annals of Surgery* (28th Feb. 2025). DOI: [10.1097/SLA.0000000000006685](https://doi.org/10.1097/SLA.0000000000006685). URL: <https://u-picardie.hal.science/hal-04981841>.
- [52] E. Rangelova, T. Stoop, T. van Ramshorst, M. Ali, E. van Bodegraven, A. Javed, D. Hashimoto, E. Steyerberg, A. Banerjee, A. Jain et al. 'The impact of neoadjuvant therapy in patients with left-sided resectable pancreatic cancer: an international multicenter study'. In: *Annals of Oncology* (Jan. 2025). DOI: [10.1016/j.annonc.2024.12.015](https://doi.org/10.1016/j.annonc.2024.12.015). URL: <https://hal.science/hal-04927124>.
- [53] J. Rekowski, C. Guo, O. Solovyeva, M. Dimairo, M. Rouhifard, D. Patel, E. Alger, D. Ashby, J. Berlin, O. Boix, M. Calvert, A.-W. Chan, C. Coschi, J. de Bono, T. Evans, R. Golub, K. Hayward, S. Hopewell, J. Isaacs, S. P. Ivy, T. Jaki, O. Kholmanskikh, A. Kightley, S. Lee, R. Liu, I. Maia, A. Mander, L. Marshall, J. Matcham, R. Peck, K. R. Rantell, D. Richards, L. Seymour, Y. Tanaka, M. Ursino, C. Weir and C. Yap. 'CONSORT-DEFINE explanation and elaboration: recommendations for enhancing reporting quality and impact of early phase dose-finding clinical trials'. In: *EClinicalMedicine* 79 (Jan. 2025), p. 102987. DOI: [10.1016/j.eclinm.2024.102987](https://doi.org/10.1016/j.eclinm.2024.102987). URL: <https://inria.hal.science/hal-04954290>.
- [54] J.-C. Rozé, G. Cambonie, C. Flamant, J. Patkai, T. Mühlbacher, G. Gascoin, A. Rideau Batista Novais, M. Tauzin, K. Le Duc, A. Beuchée et al. 'Prophylactic Treatment of Patent Ductus Arteriosus With Acetaminophen'. In: *Medicine Archives of Pediatrics & Adolescent - JAMA Pediatrics* (16th Feb. 2026), e256150. DOI: [10.1001/jamapediatrics.2025.6150](https://doi.org/10.1001/jamapediatrics.2025.6150). URL: <https://hal.inrae.fr/hal-05539868>.

- [55] M. Sapoval, O. Pellerin, A. Boyer, C. Déan and T. Boeken. ‘Does genicular artery embolization compromise future knee surgery in patients with knee osteoarthritis? A strategic call to the community’. In: *Diagnostic and Interventional Imaging* 106.3 (Mar. 2025), pp. 105–106. DOI: [10.1016/j.diii.2024.12.006](https://doi.org/10.1016/j.diii.2024.12.006). URL: <https://hal.science/hal-05158225>.
- [56] S. Selvam, O. Peyrony, A. Elezi, A. Braganca, A.-M. Zagdanski, L. Biard, J. Assouline, G. Chassagnon, G. Mulier and C. de Margerie-Mellon. ‘Efficacy of a deep learning-based software for chest X-ray analysis in an emergency department’. In: *Diagnostic and Interventional Imaging* 106.9 (Sept. 2025), pp. 299–311. DOI: [10.1016/j.diii.2025.03.007](https://doi.org/10.1016/j.diii.2025.03.007). URL: <https://hal.science/hal-05364852>.
- [57] A. Senellart, C. Chadebec and S. Allasonnière. ‘MultiVae: A Python package for Multimodal Variational Autoencoders on Partial Datasets.’ In: *Journal of Open Source Software* 10.110 (5th June 2025), p. 7996. DOI: [10.21105/joss.07996](https://doi.org/10.21105/joss.07996). URL: <https://hal.science/hal-05372674> (cit. on p. 11).
- [58] E. Shotar, B. Mathon, H. Salle, A. Rouchaud, C. Mounayer, N. Bricout, J.-P. Lejeune, K. Janot, A. Amelot, O. Naggara et al. ‘Meningeal Embolization for Preventing Chronic Subdural Hematoma Recurrence After Surgery’. In: *JAMA Cardiology* 334.2 (8th July 2025), p. 127. DOI: [10.1001/jama.2025.7583](https://doi.org/10.1001/jama.2025.7583). URL: <https://hal.science/hal-05158216>.
- [59] M. Strigalev, D. Fuks, S. Katsahian, L. Parlati, U. Marchese, M. Conticchio, C. Ronde-Roupie, A. Nassar, A. Dhote, V. Mallet and S. Tzedakis. ‘The ‘social gradient’ in primary liver cancer in France: A national observational study’. In: *JHEP Reports Innovation in Hepatology* 7 (5th Sept. 2025). DOI: [10.1016/j.jhepr.2025.101585](https://doi.org/10.1016/j.jhepr.2025.101585). URL: <https://hal.science/hal-05374726>.
- [60] S. Tzedakis, D. Berzan, U. Marchese, A. Challine, V. Mallet, A. Dohan, H. Jeddou, A. Nassar, S. Katsahian and D. Fuks. ‘Implementation and short-term outcomes of minimally invasive liver surgery in France’. In: *British Journal of Surgery* 112.4 (28th Apr. 2025), znaf017. DOI: [10.1093/bjs/znaf017](https://doi.org/10.1093/bjs/znaf017). URL: <https://hal.science/hal-05252413>.
- [61] M. Ursino, C. Alberti, G. Cambonie, R. Kemp, A. Vanhecke, L. Levoyer, A. Diallo, M. Hallman, J.-C. Rozé, R. Carbajal et al. ‘TRECOPA: prophylactic treatment of the ductus arteriosus in preterm infants by acetaminophen—statistical analysis plan for the randomized phase III group sequential trial’. In: *Trials* 26.1 (13th Feb. 2025), p. 52. DOI: [10.1186/s13063-025-08751-8](https://doi.org/10.1186/s13063-025-08751-8). URL: <https://inria.hal.science/hal-04954264>.
- [62] M. Ursino, G. Villacampa, J. Rekowski, M. Dimairo, O. Solovyeva, D. Ashby, J. Berlin, O. Boix, M. Calvert, A.-W. Chan, C. Coschi, T. Evans, E. Garrett-Mayer, R. Golub, C. Guo, K. Hayward, S. Hopewell, J. Isaacs, S. P. Ivy, T. Jaki, O. Kholmanskikh, A. Kightley, S. Lee, R. Liu, A. Mander, L. Marshall, J. Matcham, D. Patel, R. Peck, K. R. Rantell, D. Richards, M. Rouhifard, L. Seymour, Y. Tanaka, C. Weir, J. de Bono and C. Yap. ‘SPIRIT-DEFINE explanation and elaboration: recommendations for enhancing quality and impact of early phase dose-finding clinical trials protocols’. In: *EClinicalMedicine* 79 (Jan. 2025), p. 102988. DOI: [10.1016/j.eclinm.2024.102988](https://doi.org/10.1016/j.eclinm.2024.102988). URL: <https://inria.hal.science/hal-04954281>.
- [63] L. Vincent, A.-s. Jannot, H. Mechiche, U. Rodts and G. Désaméricq. ‘Use of Hospital-at-Home Services for Injectable Chemotherapy for Patients With Multiple Myeloma in France in 2019 and 2020: A Real-World Nationwide Study Based on the French Hospital Discharge Database’. In: *eJHaem* 6.5 (11th Sept. 2025), p. 106214. DOI: [10.1002/jha2.70144](https://doi.org/10.1002/jha2.70144). URL: <https://inria.hal.science/hal-05368969>.
- [64] A. Vuorinen, E. Comets and M. Ursino. ‘A comparative analysis of Phase I dose-finding designs incorporating pharmacokinetics information’. In: *The American Statistician* (12th Sept. 2025). DOI: [10.1080/00031305.2025.2560371](https://doi.org/10.1080/00031305.2025.2560371). URL: <https://hal.science/hal-05261453>.
- [65] M. Wack, A. Coulet, A. Burgun and B. Rance. ‘Enhancing clinical data warehousing with provenance data to support longitudinal analyses and large file management: The gitOmmix approach for genomic and image data’. In: *Journal of Biomedical Informatics* 163 (Mar. 2025), p. 104788. DOI: [10.1016/j.jbi.2025.104788](https://doi.org/10.1016/j.jbi.2025.104788). URL: <https://hal.science/hal-05064217>.

- [66] J.-B. Woillard, S. Benzekry, J. Josse, M. White-Koning, E. Chatelut, E. Comets, F. Lemaitre, B. Franck, M. Grégoire, F. Stanke-Labesque, S. Zohar, M. Ursino and C. Battail. ‘Digital Pharmacological Twins: Bridging Multi-scale Modelling and Artificial Intelligence for Precision Medicine : the DIGPHAT consortium’. In: *Therapies* (2025). DOI: [10.1016/j.therap.2025.09.006](https://doi.org/10.1016/j.therap.2025.09.006). URL: <https://hal.science/hal-05288607>.

Invited conferences

- [67] L. Ducrot, G. Fevry, C. Dano, R. Texier, J. Murriss and S. Katsahian. ‘SurvTreeSHAP(t) : scalable explanation method for tree-based survival models’. In: Workshop on Explainable Artificial Intelligence (XAI), IJCAI 2025. Montreal, Canada, 16th Aug. 2025. URL: <https://hal.science/hal-05108033>.

International peer-reviewed conferences

- [68] M. Abd Nikooie Pour, E. Blomqvist, P. Giesteira Cotovio, A. Coulet, L. Ferraz, S. Hertling, S. Jain, E. Jiménez-Ruiz, F. Kraus, P. Lambrix, H. Li, Y. Li, X. Liu, P. Monnin, H. Paulheim, C. Pesquita, A. Sharma, P. Shvaiko, M. Silva, G. Sousa, C. Trojahn, J. Vataščinová, B. Yaman, O. Zamazal and L. Zhou. ‘Results of the Ontology Alignment Evaluation Initiative 2025’. In: OM 2025 - Ontology Matching 2025. Nara, Japan, 2nd Nov. 2025. URL: <https://inria.hal.science/hal-05447839>.
- [69] J.-B. Baitairian, B. Sebastien, R. Jreich, S. Katsahian and A. Guilloux. ‘Sharp Bounds for Continuous-Valued Treatment Effects with Unobserved Confounders’. In: JdS 2025 - 56èmes Journées de Statistique de la SFdS. Marseille, France, 2nd June 2025. URL: <https://hal.science/hal-05371431>.
- [70] P. Chassat, V. T. Nguyen, L. Ducrot, E. Lanoy and A. Guilloux. ‘Toward Valid Generative Clinical Trial Data with Survival Endpoints’. In: ML4H - Machine Learning for Health Symposium. San Diego (CA), United States, 2025. URL: <https://hal.science/hal-05375245>.
- [71] M. I. E. Ghebrriout, G. Guibon, I. Lerner and E. Vincent. ‘QUARTZ: QA-based Unsupervised Abstractive Refinement for Task-oriented Dialogue Summarization’. In: *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*. EMNLP 2025 : The 2025 Conference on Empirical Methods in Natural Language Processing. Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing. Suzhou, China, 4th Nov. 2025. URL: <https://hal.science/hal-05300943>.
- [72] G. Houry, T. Boeken, S. Allasonnière and J. Feydy. ‘Untangling Vascular Trees for Surgery and Interventional Radiology’. In: MICCAI 2025 - 28th International Conference on Medical Image Computing and Computer Assisted Intervention. Vol. 15968. Lecture Notes in Computer Science. Daejeon, South Korea: Springer Nature Switzerland, 21st Sept. 2025, pp. 669–679. DOI: [10.1007/978-3-032-05114-1_64](https://doi.org/10.1007/978-3-032-05114-1_64). URL: <https://hal.science/hal-05271418> (cit. on p. 12).
- [73] J. H. Jhee, A. Megina, P. Constant Dit Beaufils, M. Karakachoff, R. Redon, A. Gaignard and A. Coulet. ‘Predicting clinical outcomes from patient care pathways represented with temporal knowledge graphs’. In: ESWC 2025 - 22nd European Semantic Web Conference. Vol. LNCS-15718. Lecture Notes in Computer Science. Portorož, Slovenia: Springer Nature Switzerland, 30th Apr. 2025, pp. 282–300. DOI: [10.1007/978-3-031-94575-5_16](https://doi.org/10.1007/978-3-031-94575-5_16). URL: <https://hal.science/hal-05136324>.
- [74] A. Marfoglio, J. H. Jhee and A. Coulet. ‘Clinical Data Goes MEDS? Let’s OWL make sense of it’. In: 17th International SWAT4HCLS Conference, Semantic Web Applications and Tools for Health Care and Life Sciences. Amsterdam, Netherlands: arXiv, 2026. DOI: [10.48550/arXiv.2601.04164](https://doi.org/10.48550/arXiv.2601.04164). URL: <https://inria.hal.science/hal-05537716>.
- [75] N. Mouadden, O. Laousy, R. Marini, V. Ong, M.-P. Revel, G. Chassignon, S. Christodoulidis and M. Vakalopoulou. ‘Conditional Latent Diffusion Models for Irregularly Spaced Longitudinal Radiological Data’. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2025. Vol. 15974. Lecture Notes in Computer Science LNCS-15974. Daejeon, South Korea: Springer Nature Switzerland, 18th Sept. 2026, pp. 106–115. DOI: [10.1007/978-3-032-05182-0_11](https://doi.org/10.1007/978-3-032-05182-0_11). URL: <https://hal.science/hal-05472492>.

- [76] A. Senellart, M. Solal, S. Allasonnière and N. Burgos. ‘Mitigating the reconstruction-detection trade-off in VAE-based unsupervised anomaly detection’. In: ISBI 2026 - International Symposium on Biomedical Imaging. London, United Kingdom, 8th Apr. 2026. URL: <https://hal.science/hal-05423947>.
- [77] M. Ursino, S. Boulet, C. Collignon, F. Francis-Oliviero, E. Lhomme, R. Porcher, F. Saillour, G. Varoquaux, V. Vercamer, R. Thiébaud and S. Zohar. ‘Innovative clinical trial approach for evaluating digital medical devices under European HTA fast-Track frameworks’. In: ISCB 2025 - 46th Annual Conference of the International Society for Clinical Biostatistics. Basel, Switzerland, 24th Aug. 2025. URL: <https://hal.science/hal-05364910>.

National peer-reviewed Conferences

- [78] J. H. Jhee, A. Megina, P. Constant Dit Beaufile, M. Karakachoff, R. Redon, A. Gaignard and A. Coulet. ‘Prédiction d’événements cliniques à partir des parcours de soins de patients représentés par des graphes de connaissances temporels’. In: IC 2025 - 36e Journées Francophones d’Ingénierie des Connaissances. Dijon, France, 2nd July 2025. URL: <https://inria.hal.science/hal-05375218>.

Conferences without proceedings

- [79] S. Alsaidi, M. Vincent, O. Boyer, N. Garcelon, M. Couceiro and A. Coulet. ‘Comparing representations of long clinical texts for the task of patient note-identification’. In: CL4Health@NAACL 2025: Second Workshop on Patient-oriented Language Processing (CL4Health), at NAACL 2025. Albuquerque, New Mexico, United States, 4th May 2025. URL: <https://hal.science/hal-05000565>.
- [80] S. Boulet, V. Thévenet, J. Abécassis, C. Collignon, B. Giraudeau, E. Lhomme, F. Petit, R. Porcher, L. Richert, F. Saillour, M. Ursino, G. Varoquaux, V. Vercamer, S. Zohar and R. Thiébaud. ‘Narrative review on the clinical evaluation of AI-based digital medical devices from a methodological perspective’. In: PEPR-SanteNum 2025 - Journées Annuelles du PEPR Santé Numérique. Lille, France, 14th Oct. 2025. URL: <https://hal.science/hal-05364958>.
- [81] L. Comin, S. Marie, M. Ursino, S. Zohar, N. Tournier and E. Comets. ‘Modelling the whole-body pharmacokinetics of radiolabeled glyburide in healthy volunteers’. In: EMIM 2025 - 20th European Molecular Imaging Meeting. Bilbao, Spain, 11th Mar. 2025. URL: <https://hal.science/hal-05065283>.
- [82] S. Kaisaridi, J. Ortholand, C. Tuna, N. Gensollen and S. Tezenas Du Montcel. ‘A mixture model for subtype identification: Application to CADASIL’. In: ISCB 46 - 46th Annual Conference of the International Society for Clinical Biostatistics. Basel, Switzerland, 24th Aug. 2025. URL: <https://hal.science/hal-05266776>.
- [83] S. Salas, C. Chevreau, E. Jouve, N. Penel, A. Le Cesne, E. Bompas, F. Bertucci, J. E. Kurtz, N. Isambert, L. Chaigneau, J. Micallef, E. Foinet, N. Andre, S. Zohar and J.-Y. Blay. ‘PROPAN: Bayesian dose-finding of propranolol combined with metronomic fixed-dose cyclophosphamide based on bivariate efficacy-tolerability outcomes in patients with locally advanced or metastatic angiosarcoma (LAMA): A phase I-II sequential trial by the French Sarcoma Group.’ In: American Society of Clinical Oncology (ASCO) Annual meeting 2025. Vol. 43. 16_suppl. Chicago, United States, June 2025. DOI: [10.1200/JCO.2025.43.16_suppl.e23542](https://doi.org/10.1200/JCO.2025.43.16_suppl.e23542). URL: <https://inria.hal.science/hal-05470756>.

Doctoral dissertations and habilitation theses

- [84] T. Fabacher. ‘Reproducible information extraction from French clinical texts : application to chronic diseases.’ Université de paris, 11th June 2025. URL: <https://hal.science/tel-05462856>.
- [85] L. Muyama. ‘Learning clinical diagnostic pathways from electronic health records using machine learning’. Université Paris Cité, 5th Mar. 2025. URL: <https://theses.hal.science/tel-05446161>.

- [86] S. Tzedakis. ‘Real-world hepatobiliary surgery : multiscale analysis of care pathways and postoperative complications’. Université Paris Cité, 18th Dec. 2025. URL: <https://theses.hal.science/tel-05534014>.

Reports & preprints

- [87] J.-B. Baitairian, B. Sebastien, R. Jreich, S. Katsahian and A. Guilloux. *Sensitivity Analysis to Unobserved Confounders: A Comparative Review to Estimate Confounding Strength in Sensitivity Models*. 2025. DOI: [10.48550/arXiv.2510.16560](https://doi.org/10.48550/arXiv.2510.16560). URL: <https://hal.science/hal-05330588>.
- [88] A. Bercu, A. Guilloux, C. Proust-Lima and H. Jacqmin-Gadda. *A regularized multi-state model for covariate selection with interval-censored survival data*. 2025. URL: <https://inria.hal.science/hal-05368412> (cit. on p. 11).
- [89] S. Boufadene, J. Feydy and F.-X. Vialard. *Fast Large Deformation Matching with the Energy Distance Kernel*. 5th May 2025. URL: <https://hal.science/hal-05035450>.
- [90] A. Favier, S. Chounta, A. Garcia, F. Jabot-Hanin, X. Chen, N. Garcelon, A. Burgun, M. Higuera, A. Guilloux, A. Benmerah, Y. Martin, K. Billot, J.-M. Rozet, I. Perrault, V. Cormier-Daire, C. Huber, M. Zaidan, T. Attie-Bitach, S. Saunier and A. Rausell. *COBT: A gene-based rare variant burden test for case-only study designs using aggregated genotypes from public reference cohorts*. 16th July 2025. DOI: [10.1101/2025.07.16.25331553](https://doi.org/10.1101/2025.07.16.25331553). URL: <https://inria.hal.science/hal-05368419>.
- [91] G. Houry, J. Feydy and F.-X. Vialard. *Gromov-Wasserstein at Scale, Beyond Squared Norms*. 5th Feb. 2026. URL: <https://hal.science/hal-05495620>.
- [92] R. Jaquet, A. Challine, S. Tzedakis, C. de Ponthaud, A. Dhote, S. Katsahian and A. Lazzati. *Impact of thresholds on travel time in digestive cancer surgery patients: A French nationwide cohort*. 29th Sept. 2025. URL: <https://hal.sorbonne-universite.fr/hal-05289114>.
- [93] J. Murriss, M. Bhan, L. Ducrot and S. Katsahian. *Bridging interpretability and survival endpoints in health technology assessment*. 26th Feb. 2025. URL: <https://hal.science/hal-04967719>.
- [94] J. Murriss, O. Bouaziz, M. Jakubczak, S. Katsahian and A. Lavenu. *Random survival forests for the analysis of recurrent events for right-censored data, with or without a terminal event*. 9th Apr. 2025. URL: <https://hal.science/hal-04612431>.
- [95] A. Senellart, S. Allasonnière and A. Senellart. *Bridging the inference gap in Multimodal Variational Autoencoders*. 15th Jan. 2025. URL: <https://hal.science/hal-04932313>.
- [96] J. Zhong, E. Bacry, A. Guilloux and J.-F. Muzy. *KANFormer for Predicting Fill Probabilities via Survival Analysis in Limit Order Books*. 5th Dec. 2025. URL: <https://hal.science/hal-05399393>.

Other scientific publications

- [97] S. Boulet, V. Thévenet, J. Abécassis, C. Collignon, B. Giraudeau, E. Lhomme, F. Petit, R. Porcher, L. Richert, F. Saillour, M. Ursino, G. Varoquaux, V. Vercamer, S. Zohar and R. Thiébaud. ‘Narrative review on the clinical evaluation of AI-based digital medical devices from a methodological perspective’. In: *SbN 2025 - 10th Statistics & Biopharmacy Conference*. Paris, France, 8th Oct. 2025. URL: <https://hal.science/hal-05364943>.
- [98] S. Boulet, V. Thévenet, J. Abécassis, C. Collignon, B. Giraudeau, E. Lhomme, F. Petit, R. Porcher, L. Richert, F. Saillour, M. Ursino, G. Varoquaux, V. Vercamer, S. Zohar and R. Thiébaud. ‘Narrative review on the clinical evaluation of AI-based digital medical devices from a methodological perspective’. In: *PEPR-SanteNum 2025 - Journées Annuelles du PEPR Santé Numérique*. Lille, France, 14th Oct. 2025. URL: <https://hal.science/hal-05347091>.
- [99] T. Desrozières, Y. Soreze, V. NAU, B. Crestani, D. Drummond, C. Dupin, R. Borie, M. Reynaud-Gaubert, K. Hurley, F. Dastot - Le Moal, S. Amselem, M. Legendre, N. Nathan and C. Louvrier. ‘SFTPA1 and SFTPA2 defects associated with interstitial lung disease are not always missense variations, a chILD EU study’. In: *ERS 2025 - European respiratory Society*. Amsterdam, Netherlands, 27th Sept. 2025. URL: <https://inserm.hal.science/inserm-05296112>.

- [100] F. Ghestem, G. Marenne, E. Genin, A.-S. Jannot, A. Baudot and A.-L. Leutenegger. ‘Knowledge graph to dissect genotype phenotype associations’. In: IGES 2025 - Annual Meeting on International Genetic Epidemiology Society. Cologne, Germany, Apr. 2025. URL: <https://hal.science/hal-05333821>.
- [101] A. Poirot-Bourdain, A. van Straaten, E. Bacry, A.-S. Jannot and J. Feydy. ‘SurvivalGPU: a GPU Powered Library for Scalable Survival Analysis in R and Python’. In: AI4Health Summer School. Paris, France, 30th June 2025. URL: <https://hal.science/hal-05390757>.
- [102] M. Ursino, S. Boulet, C. Collignon, F. Francis-Oliviero, E. Lhomme, R. Porcher, F. Saillour, G. Varoquaux, R. Thiébaud and S. Zohar. ‘Innovative Clinical Trial Approach for Evaluating Digital Medical Devices under European Fast-Track Regulatory Frameworks’. In: EPICLIN 2025 - Conférence francophone d’EPIdémologie CLINique. Bordeaux, France, 14th May 2025. URL: <https://hal.science/hal-05085050>.

12.3 Cited publications

- [103] J.-B. Baitairian, S. Bernard, J. Rana, S. Katsahian and A. Guilloux. ‘Sensitivity Analysis to Unobserved Confounders: A Comparative Review to Estimate Confounding Strength in Sensitivity Models’. Submitted. Preprint available on arXiv. 2025 (cit. on p. 11).
- [104] J.-B. Baitairian, B. Sebastien, R. Jreich, S. Katsahian and A. Guilloux. ‘Sharp Bounds for Continuous-Valued Treatment Effects with Unobserved Confounders’. In: *Biomedical Journal* (2025) (cit. on p. 11).
- [105] P. Chassat, V. T. Nguyen, L. Ducrot and A. Guilloux. ‘Toward Valid Generative Clinical Trial Data with Survival Endpoints’. In: *Proceedings of the ML4H Conference*. 2025 (cit. on p. 11).
- [106] E. Jemelen, F. Orchard, W. Madie, B. Valentin, J. Belin, E. Laas, G. Jeannerod, P. Mares, S. Katsahian and A. Guilloux. ‘Evaluating breast cancer screening performance without registries using medico-administrative data’. In: *Scientific Reports* 15.1 (2025), p. 25096 (cit. on p. 11).