

## Research Internship (M2): Adaptive designs for early stage clinical trials in vaccinology

Emilie Kaufmann (CNRS & CRISTAL, Inria Lille (Scool team))

In a multi-armed bandit model, an agent sequential sample arms, that are unknown probability distributions, in order to achieve some goal [6]. The most studied goals are either to maximize the sum of the gathered samples (viewed as reward) or to solve a *pure exploration* task, such as identifying the arm with largest mean. Multi-armed bandits were originally motivated by (phase III) clinical trials in which arms model the efficacy of different treatments, but bandit designs have been seldom used in real clinical trials [8], notably because they are fully sequential: taking patients one by one and waiting to measure efficacy of the treatment would lead to very long trials. Moreover, the standard statistical methodology cannot be used when the data is collected adaptively.

Adaptive designs have been mainly used in the special case of phase I trials in oncology, in which bandit designs have also recently been suggested by [3]. In this context, the goal of a phase I trial is to find the maximal tolerated dose (MTD), and trials are performed on patients so there is also an incentive in treating them with the MTD during the trial, as it is the most efficient dose with acceptable toxicity. In contrast, for the development of prophylactic vaccines, early-stage trials are performed on healthy volunteers and the toxicity often depends less crucially on the dose level. Moreover, additionally to the side effects, the immune response is also measured a few weeks after receiving the vaccine. Several markers of the immune response are measured (typically 3 to 5) and the goal would be to determine a (non toxic) dose which is jointly optimizing these key markers.

Taking into accounts this *multi-objective* aspect calls for the development of a new pure exploration machinery. While pure-exploration with fixed confidence is now a well understood problem (see, e.g. [5, 4]) when each arm is a real-valued distributions, multi-dimensional outcomes in bandits have been studied less extensively [2, 1]. The goal of an adaptive early-stage trial may be phrased as adaptively finding doses on a Pareto front involving the different markers, possibly with toxicity constraints.

The objective of the internship is to propose new algorithms for multi-objective pure exploration, and adapt them to the use case of early-stage trials in vaccinology. During the internship we will be in contact with colleagues from the Inria/INSERM team SISTM in Bordeaux, who are currently collaborating in clinical trials on HIV vaccines. They have access to data from previous trials (see, e.g. [7]), and are working on building a simulator for the immune response that may be used in conjunction with our bandit algorithms.

**Practical information** The candidate should have a strong background in statistics and some coding skills (e.g. in Python or Julia). Some prior knowledge of multi-armed bandits is a bonus but not mandatory. The internship will take place at Inria Lille, in the **Scool team**, under the supervision of **Emilie Kaufmann**.

**Contact:** [emilie.kaufmann@univ-lille.fr](mailto:emilie.kaufmann@univ-lille.fr)



## References

- [1] Cagin Ararat and Cem Tekin. Vector optimization with stochastic bandit feedback. *arXiv 2110.12311*, 2021.
- [2] Peter Auer, Chao-Kai Chiang, Ronald Ortner, and Madalina M. Drugan. Pareto front identification from stochastic bandit feedback. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2016.
- [3] Maryam Aziz, Emilie Kaufmann, and Marie-Karelle Riviere. On multi-armed bandit designs for dose-finding clinical trials. *Journal of Machine Learning Research*, 22(14):1–38, 2021.
- [4] Rémy Degenne, Wouter M. Koolen, and Pierre Ménard. Non-asymptotic pure exploration by solving games. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.
- [5] Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Proceedings of the 29th Conference On Learning Theory*, 2016.
- [6] Tor Lattimore and Csaba Szepesvari. *Bandit Algorithms*. Cambridge University Press, 2019.
- [7] Laura Richert, Adélaïde Doussau, Jean-Daniel Lelièvre, Vincent Arnold, Véronique Rieux, Amel Bouakane, Yves Lévy, Geneviève Chêne, and Rodolphe Thiébaud. Accelerating clinical development of hiv vaccine strategies: methodological challenges and considerations in constructing an optimised multi-arm phase i/ii trial design. *Trials*, 15(68), 2014.
- [8] Clémence Réda, Emilie Kaufmann, and Andrée Delahaye-Duriez. Machine learning applications in drug development. *Computational and Structural Biotechnology Journal*, 18:241–252, 2020.