# Exploring gene expression as individual-level and trial-level surrogates of vaccine response

## M2 Internship

### Background

Direct measurement of the primary outcome in a clinical trial is often time-consuming, expensive, or infeasible. For instance, when the outcome is rare, it may take many years to observe enough events to observe a statistically significant difference between groups, even if the treatment is effective. Consequently, there is considerable interest in identifying *surrogate markers*—variables that can be observed more quickly, at lower cost, or with greater ease, and that can serve as reliable substitutes for the primary outcome to accurately infer the treatment effect(s) [1].

Surrogate markers may be classed into two categories. Trial-level surrogate markers are predictive of the average effect of a given treatment, whereas individual-level surrogate markers are predictive of an individual's treatment effect [2]. However, despite the differences in both concept and clinical application, the distinction between these classes is often unclear in the literature, and the relationship between the two is poorly understood.

The identification and subsequent validation of surrogate markers is particularly relevant in the context of vaccine development [3]. Trial-level surrogate markers could provide evidence for the rapid selection of candidate vaccines in experimental adaptive trials, and allow for vaccine evaluation when a phase 3 trial is not feasible due to the absence of endemic disease in the target population [4]. Individual-level surrogates can allow for the identification of high-risk individuals who do not respond to a given vaccine. Both types of surrogates can provide insight into the complex mechanisms which drive vaccine effectiveness [5].

High-throughput technologies, such as transcriptomics data, hold great promise for informing effective vaccine design and aiding precision medicine [6]. Transcriptomic data describe gene expression—the process by which information encoded in DNA is transformed into proteins that shape phenotypes. Changes in gene expression occur in the first hours and days after vaccination and are likely to reflect early biological perturbations leading to protective immunity. Prior research has shown that these markers can be predictive of individual responses to certain vaccines [7, 8, 9]. However, the potential of these markers to act as trial- or individual-level surrogates has only started to be explored, and requires specialist methods [10]. The aim of this internship project will be to apply diverse statistical learning approaches to identify both individual- and trial-level surrogate markers in context of high-dimensional vaccine experiments.

## Objectives

1. Characterize theoretical relations between individual and trial level surrogates (e.g. with toy counterexamples)
2. Benchmark various prediction models and machine learning approaches for the identification of individual-level surrogates in high-dimensional context, and compare univariate and multivariate strategies for individual-level surrogate identification from high-dimensional datasets.
3. Apply the select approach to public molecular vaccinology datasets [11], and compare its results to those of the RISE method.
4. Evaluate the robustness of any derived signatures across multiple studies & populations

## Required Skills

- Master 2/Bachelor/Engineering school with a major in biostatistics or similar
- Programming proficiency with R
- Interest in biomedical research and scientific curiosity
- Written and spoken English proficiency

**Hosting laboratory**
SISTM Team
INSERM U1219 and INRIA Sud-Ouest

**Location**
Bordeaux Population Health Research Center
Université de Bordeaux – ISPED
146, Rue Léo Saignat
33076 Bordeaux

**Duration:**
Internship of 4 to 6 month available starting from January 2025.

**Compensation:**
Intern gratification according to government recommendations (15% of social security ceiling, i.e. around 660€/month).

**Contact:**
Send a detailed CV and a motivation letter to Boris Hejblum (boris.hejblum@u-bordeaux.fr) and Arthur Hughes (arthur.hughes@u-bordeaux.fr).

# References


[1] Christensen, R., Ciani, O., Manyara, A. M., and Taylor, R. S. (March, 2024) Surrogate endpoints: a key concept in clinical epidemiology. *Journal of Clinical Epidemiology,* **167**, 111242.

[2] Gabriel, E. E., Daniels, M. J., and Halloran, M. E. (April, 2016) Comparing Biomarkers as Trial Level General Surrogates. *Biometrics,* **72**(4), 1046–1054.

[3] Gilbert, P. B., Fong, Y., Hejazi, N. S., Kenny, A., Huang, Y., Carone, M., Benkeser, D., and Follmann, D. (April, 2024) Four statistical frameworks for assessing an immune correlate of protection (surrogate endpoint) from a randomized, controlled, vaccine efficacy trial. *Vaccine,* **42**(9), 2181–2190.

[4] Plotkin, S. A. (July, 2010) Correlates of Protection Induced by Vaccination. *Clinical and Vaccine Immunology,* **17**(7), 1055–1065.

[5] Johnston, M. I. and Fauci, A. S. (August, 2008) An HIV Vaccine — Challenges and Prospects. *New England Journal of Medicine,* **359**(9), 888–890.

[6] Weiner, J., Kaufmann, S. H., and Maertzdorf, J. (September, 2015) High-throughput data analysis and data integration for vaccine trials. *Vaccine,* **33**(40), 5249–5255.

[7] Querec, T. D., Akondy, R. S., Lee, E. K., Cao, W., Nakaya, H. I., Teuwen, D., Pirani, A., Gernert, K., Deng, J., Marzolf, B., Kennedy, K., Wu, H., Bennouna, S., Oluoch, H., Miller, J., Vencio, R. Z., Mulligan, M., Aderem, A., Ahmed, R., and Pulendran, B. (November, 2008) Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nature Immunology,* **10**(1), 116–125.

[8] Kazmin, D., Nakaya, H. I., Lee, E. K., Johnson, M. J., van der Most, R., van den Berg, R. A., Ballou, W. R., Jongert, E., Wille-Reece, U., Ockenhouse, C., Aderem, A., Zak, D. E., Sadoff, J., Hendriks, J., Wrammert, J., Ahmed, R., and Pulendran, B. (February, 2017) Systems analysis of protective immune responses to RTS,S malaria vaccination in humans. *Proceedings of the National Academy of Sciences,* **114**(9), 2425–2430.

[9] Nakaya, H. I., Wrammert, J., Lee, E. K., Racioppi, L., Marie-Kunze, S., Haining, W. N., Means, A. R., Kasturi, S. P., Khan, N., Li, G.-M., McCausland, M., Kanchan, V., Kokko, K. E., Li, S., Elbein, R., Mehta, A. K., Aderem, A., Subbarao, K., Ahmed, R., and Pulendran, B. (July, 2011) Systems biology of vaccination for seasonal influenza in humans. *Nature Immunology,* **12**(8), 786–795.

[10] Hughes, A., Parast, L., Thiébaut, R., and Hejblum, B. P. (September, 2025) RISE: Two-Stage Rank-Based Identification of High-Dimensional Surrogate Markers Applied to Vaccinology. *Statistics in Medicine,* **44**(20–22).

[11] Hagan, T., Gerritsen, B., Tomalin, L. E., Fourati, S., Mulè, M. P., Chawla, D. G., Rychkov, D., Henrich, E., Miller, H. E. R., Diray-Arce, J., Dunn, P., Lee, A., Deckhut-Augustine, A., Gottardo, R., Haddad, E. K., Hafler, D. A., Harris, E., Farber, D., Kleinstein, S. H., Levy, O., McElrath, J., Montgomery, R. R., Peters, B., Pulendran, B., Rahman, A., Reed, E. F., Rouphael, N., Sarwal, M. M., Sékaly, R. P., Fernandez-Sesma, A., Sette, A., Stuart, K., Togias, A., Tsang, J. S., Levy, O., Gottardo, R., Sarwal, M. M., Tsang, J. S., Suárez-Fariñas, M., Sékaly, R.-P., Kleinstein, S. H., and Pulendran, B. (October, 2022) Transcriptional atlas of the human immune response to 13 vaccines reveals a common predictor of vaccine-induced antibody responses. *Nature Immunology,* **23**(12), 1788–1798.