An important aspect of cancer research is the development of better prognostic tools for clinicians. These tools aim at predicting the survival time outcome of patients. Such tools are crucial as they assist clinicians in the choice of the best treatment strategy for each patient. An accurate prognostic model could help to save patients from unnecessary treatments.

This thesis is centred around the development of prognostic models in their three key aspects. These three aspects are the selection of relevant markers, the learning of the prognosis model itself and its validation.

We proposed the Coxlogit model for feature selection from survival and classification data. Classification and survival prediction are two common tasks in cancer research. With the Coxlogit model, we propose to model together these two tasks to improve the prediction and the feature selection. The Coxlogit model can be seen as a regularized mixture of a Cox and logistic models.

The relevance of a prognostic model is typically assessed with a hazard ratio between the predicted risk groups. We identified some limitations of the hazard ratio in this particular context. More precisely, it appears to be very sensitive to the choice of discretization of the risk scores and has extreme values with unbalanced risk groups. We investigate the effect of the discretization in risk groups for the hazard ratio and other related metrics. A new metric, the balanced hazard ratio, is also proposed to solve those issues.

The biomedical part of this thesis investigates the use of hypoxia related gene signatures as potential prognostic markers. In controlled experiments, cell lines were submitted to normoxia, hypoxia and cycling hypoxia and then used to reduce molecular signatures. Promising prognosis results were found on real breast cancer data. Moreover, these hypoxia related gene signatures turn out to be an added value to the standard clinical prognostic models.