Résumé en Français (249 mots)

Un nombre croissant d'essais portant sur de nouvelles thérapies neuroprotectrices se concentre désormais sur les maladies neurologiques rares. Ce groupe, comprenant plus de 9000 entités, pose néanmoins des défis statistiques spécifiques pour la conception des essais cliniques, en raison de la rareté et de l'hétérogénéité des patients. Cette thèse explore l'usage des données de vie réelle, en particulier issues de registres, pour améliorer le design d'essais destinés aux maladies neurodégénératives ultra-rares, avec pour cas d'étude l'Ataxie Cérébelleuse Autosomique Récessive (ARCA).

Nous avons d'abord exploité les données du ARCA registry afin de modéliser l'histoire naturelle de la maladie par des modèles non-linéaires à effets mixtes, intégrant l'hétérogénéité interindividuelle et la gestion des covariables manquantes. Une méthode de sélection de modèles via imputation multiple a été développée.

Nous avons ensuite utilisé ce modèle pour simuler des essais cliniques randomisés et comparer divers scénarios de design. Nous avons évalué l'impact de plusieurs paramètres d'essais cliniques sur la puissance statistique et l'erreur de type I, afin d'identifier des configurations adaptées aux caractéristiques des ARCAs.

Enfin, nous avons étudié la problématique des très petits échantillons, des essais N-of-1, où les méthodes statistiques classiques ne sont pas applicables, par éloignement de l'asymptotique, en proposant des méthodes utilisant des les données de registres. Ces approches, bayésiennes (modèle de population) ou observationnelles (détection d'outliers), permettent d'estimer des effets individuels.

Ainsi, cette thèse contribue au développement de méthodologies pour les essais dans les maladies neurodégénératives ultra-rares, en tirant parti des données de vie réelle pour optimiser leur design.

Résumé en anglais (225 mots)

A growing number of trials focusing on new disease-modifying therapies are now concentrating on rare neurological diseases. However, this group, which comprises more than 9,000 entities, poses specific statistical challenges for clinical trial design due to the rarity of cases and the heterogeneity of patients. This thesis explores the use of real-world data, particularly from registries, to improve the design of trials for ultra-rare neurodegenerative diseases, using Autosomal Recessive Cerebellar Ataxia (ARCA) as a case study.

We first used data from the ARCA registry to model the natural history of the disease using non-linear mixed-effects models, incorporating inter-individual heterogeneity and missing covariate management. A model selection method using multiple imputation was developed.

We then used this model to simulate randomised clinical trials and compare various design scenarios. We evaluated the impact of several clinical trial parameters on statistical power and type I error risk in order to identify trial settings suited to the characteristics of ARCAs.

Finally, we studied the issue of very small samples, N-of-1 trials, where conventional statistical methods are not applicable, due to the distance from the asymptotic, by proposing methods using registry data. These approaches, Bayesian (population model) or observational (outlier detection), allow individual effects to be estimated.

This thesis therefore contributes to the development of methodologies for trials in ultra-rare neurodegenerative diseases, leveraging real-world data to optimise their design.