Poisson-gamma model for patients' recruitment in clinical trials: Investigations on boundaries of relevancy by simulation studies.

Nathan Minois
Guillaume Mijoule Stéphanie Savy Valérie Lauwers-Cances
Sandrine Andrieu Nicolas Savy

Workshop "Modélisation et simulation d'essais cliniques"

Inserm
INSTITUT de MATHEMATIQUES de TOULOUSE

1. Cohort study
2. Recruitment model
3. Robustness investigations

Inserm
INSTITUT de MATHEMATIQUES de TOULOUSE

1. Cohort study
   - Cohort definition
   - Recruitment process
2. Recruitment model
3. Robustness investigations

Definitions

Epidemiological cohort : Medical follow-up of a target population.

It is divided in three phases.

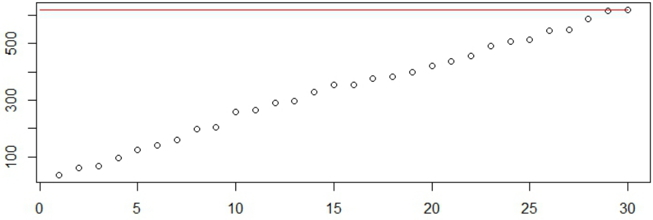
- **Recruitment :** By means of investigator centres, the necessary sample size of the study N is recruited.
- Follow-up : Monitoring of patients' state through medical visits.
- Analysis : Statistical test run on the data collected.

Introduction

Given the necessary sample size of the study N.

The recruitment process aims to reach *N* by means of several investigator centres.

The variable of interest is the duration of the recruitment.



1. Cohort study

2. Recruitment model
   - Notations
   - Poisson-gamma model

3. Robustness investigations

Notations

Global parameters :

- N : number of patient to recruit.
- C : number of investigator centres.

Model parameters :

- (α, β) : parameters of the gamma distribution.
- λ_i : the rate of the poisson process for centre i.
- t_1 : the interim time of observation.

Observations :

- k_i : recruitment of centre i at time t_1.
- τ_i : the duration of activity of centre i up to t_1.

Gamma distribution parameters

Theorem

At time t_1, the maximisation of the likelihood function gives us $(\hat{\alpha}, \hat{\beta} = 1/\hat{\mu})$.

$$M_C^{\Gamma}(\alpha,\mu) = \alpha \ln(\frac{\alpha}{\mu}) - \ln\Gamma(\alpha) + \frac{1}{C}\sum_{i=1}^{C}\left[\ln\Gamma(\alpha+k_i) - (\alpha+k_i)\ln(\frac{\alpha}{\mu}+\tau_i)\right].$$

They are the a-priori parameters of the bayesian estimation.

Inclusion process estimation

Bayesian reestimation

Theorem

The density of $\lambda_i | (N_i(t_1) = k_i)$ *is :*

$$p_\theta^{t_1}(x) = \frac{\mathbb{P}[N_i(t_1) = k_i \,|\, \lambda_i = x] p_\theta(x)}{\mathbb{P}[N_i(t_1) = k_i]}$$
$$= M e^{-(\beta+\tau_i)x} x^{k_i+\alpha-1} \mathbb{1}_{x>0}$$

This is a Gamma distribution with parameters $(\alpha + k_i, \beta + \tau_i)$.

Inclusion process estimation

Approximation of the global rate

We consider $\forall i = 1, \ldots, C$:

- $m_i = \mathbb{E}[\lambda_i] < \infty$
- $\sigma_i^2 = \mathbb{V}[\lambda_i] < \infty$

The gamma disribution parameters of the global rate Λ are :

$$A = \frac{m^2}{\sigma^2} \quad et \quad B = \frac{m}{\sigma^2};$$

with

$$m = \sum_{i=1}^{C} m_i, \quad \sigma^2 = \sum_{i=1}^{C} \sigma_i^2$$

Inclusion process estimation

Expected duration

Theorem

Given :

- $\tilde{N}$ *a doubly stochastic process with rate* $\Lambda = \sum_{i=1}^{C} \lambda_i$.
- *The expected duration* $\tilde{T} = \inf_{t \geq 0}\{\tilde{N}(t) = N\}$.

Conditionally to Λ, $\tilde{T}$ *follows a* $\Gamma(N, \Lambda)$ *distribution, therefore :*

$$\mathbb{E}[\tilde{T}] = N\frac{B}{A-1} \quad \text{si} \quad A > 1, \quad \mathbb{E}[\tilde{T}] = +\infty \quad \text{si} \quad 0 < A \leq 1$$

.

Inserm INSTITUT de MATHEMATIQUES de TOULOUSE

1 Cohort study

2 Recruitment model

3 Robustness investigations
- Motivations
- Issues
- Simulations
- Analysis of variance
- Results

Model hypothesis

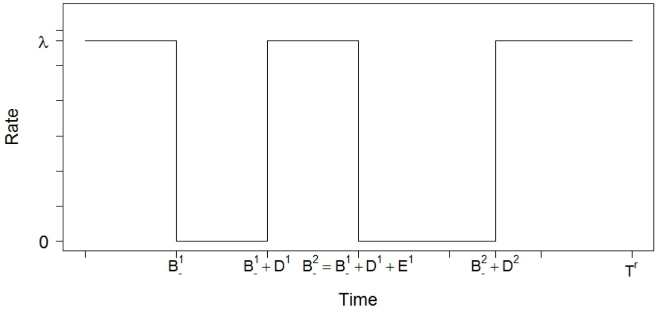
- C must be large enough (≥ 20).
- The recruitment rates must be constant over time.

On real data, the rates are rarely constant.

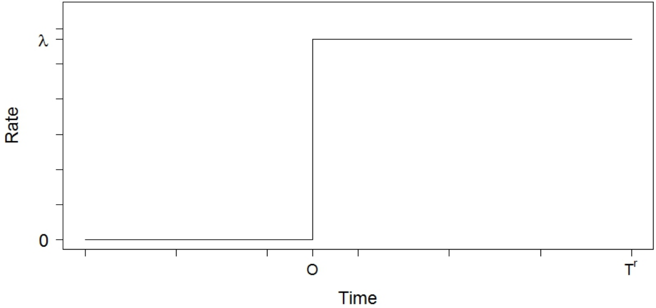
$\hookrightarrow$ cost/precision of its consideration.

Simulation study helps for the decision. It involves 20 centers to recruit 1000 patients.

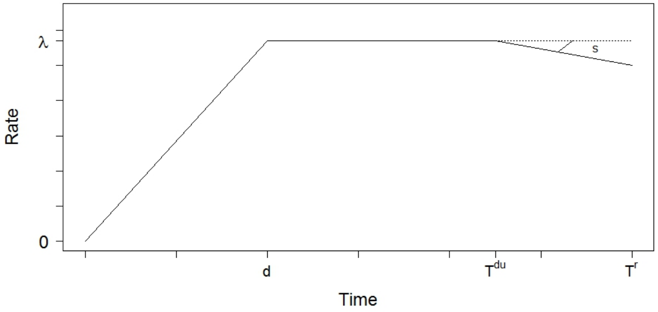
Scenario 1 : Breaks in recruitment process



Scenario 2 : Unknown opening dates



Scenario 3 : Rate changes over time



Inclusion process simulation

Data generation procedure :

- Generation of R global recruitment processes $\{N^r(t),\ 0 \leq t \leq T^r,\ 1 \leq r \leq R\}$.
- Given an interim time t_1 :
  - Estimation of (α, β) from data collected on $[0, t_1]$.
  - Calculation of the expected duration of the trial $T^r_{t_1}$ at t_1.
  - Measure of the performance of the model at interim time t_1 defined by :

$$E^r_{t_1} = \frac{T^r_{t_1} - T^r}{T^r}.$$

Processes generation

Simulation algorithm :

- Generation of the C rates $\sim$ Gamma(α, β) .
- Generation of the C recruitment dynamic.
- Aggregation of the C recruitment dynamics.
- Identification of the duration of the trial.
- Calculation of the relative error at $t_1 = 78$ and $t_1 = 104$ weeks.

Addition of the perturbations of the scenarios

Addition algorithm :

- Fix the modality of the factor condidered in the scenario.
- Perform 100 runs as follows :
  - Begin the processes generation.
  - Generate the factor's perturbations for each centre.
  - Shorten the global recruitments to 1000 patients.
  - Calculate the relative error at interim time $t_1 = 78$ and $t_1 = 104$ weeks.
- Calculate the average relative error at interim time t_1.

Scenario 1 : Breaks

Factors' modalities :

- **Number** : 1, 2 and 3 breaks.

  Simulation by an exponential random sampling.

- **Duration** : 0, 2, 4, 8, 12, and 24 weeks.

  Simulation by a multinomial random sampling with 7 different event probabilities.

Scenario 2 : Opening dates

Factor's modalities :

Values : 0, 1, 2, 4, 8, 12, 16, 20 and 24 weeks.

Simulation by a multinomial random sampling with 10 different event probabilities.

Scenario 3 : Rate changes

Factors' modalities :

- **Start-up :** 0, 2, 4 and 8 weeks.
  Simulation by a multinomial random sampling with 5 different event probabilities .
- **Drying-up start :** 108 and 120 weeks.
  Equiprobably shared between all the simulations.
- **Drying-up slope :** 0, 0.05, 0.1 and 0.2.
  Simulation by a multinomial random sampling with 5 different event probabilities .

Inserm INSTITUT de MATHEMATIQUES de TOULOUSE

Conducted analysis

The variable of interest is the relative error made defined as :

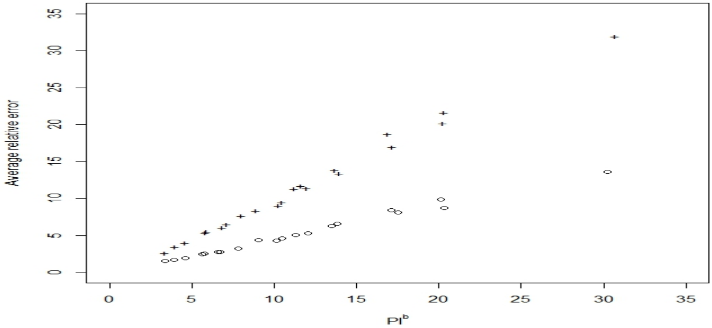
$$\mathrm{E} = \frac{\hat{T} - T_{\text{real}}}{T_{\text{real}}}$$

For each scenario, the impact of the factors is studied by an ANOVA.

Inserm
INSTITUT de MATHEMATIQUES de TOULOUSE

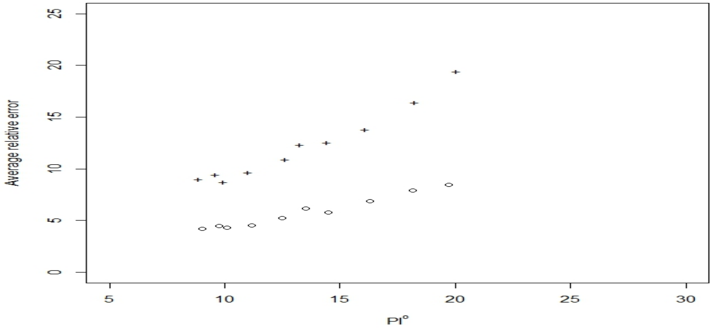
Breaks

Average relative error the expected trial duration from data collected at $t_1 = 78$ weeks, '**+**' and $t_1 = 104$ weeks, '**o**' as a function of PI^b.



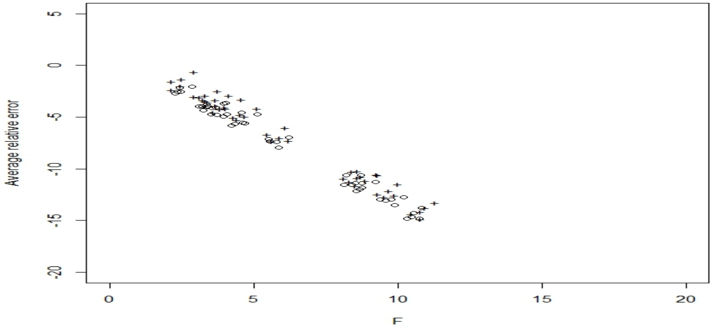
Opening dates

Average relative error the expected trial duration from data collected at t_1 = 78 weeks, '**+**' and t_1 = 104 weeks, '**o**' as a function of PI^o.



Rate changes

Average relative error the expected trial duration from data collected at t_1 = 78 weeks, '**+**' and t_1 = 104 weeks, '**o**' as a function of F.



Thank you for your attention

This research has benefited from the help of IRESP during the call for proposals launched in 2012 in the setting of Cancer Plan 2009-2013.