



How to deal with the Poisson-gamma model to forecast patients' recruitment in clinical trials when there are pauses in recruitment dynamic?



Nathan Minois^{a, b, **}, Stéphanie Savy^{a, b}, Valérie Lauwers-Cances^c, Sandrine Andrieu^{a, b, c}, Nicolas Savy^{a, d, *}

^a University of Toulouse III, F-31073, Toulouse, France

^b INSERM, UMR 1027, F-31073, Toulouse, France

^c Epidemiology Unit, CHU of Toulouse, F-31062, Toulouse, France

^d Toulouse Institute of Mathematics, UMR 5219, CNRS, F-31062, Toulouse, France

ARTICLE INFO

Article history:

Received 30 August 2016

Received in revised form

9 December 2016

Accepted 3 January 2017

Available online 6 January 2017

Keywords:

Bias analysis

Clinical trial

Cox process

Empirical Bayesian statistics

Predictive modelling

ABSTRACT

Recruiting patients is a crucial step of a clinical trial. Estimation of the trial duration is a question of paramount interest. Most techniques are based on deterministic models and various *ad hoc* methods neglecting the variability in the recruitment process. To overpass this difficulty the so-called Poisson-gamma model has been introduced involving, for each centre, a recruitment process modelled by a Poisson process whose rate is assumed constant in time and gamma-distributed. The relevancy of this model has been widely investigated. In practice, rates are rarely constant in time, there are breaks in recruitment (for instance week-ends or holidays). Such information can be collected and included in a model considering piecewise constant rate functions yielding to an inhomogeneous Cox model. The estimation of the trial duration is much more difficult. Three strategies of computation of the expected trial duration are proposed considering all the breaks, considering only large breaks and without considering breaks. The bias of these estimations procedure are assessed by means of simulation studies considering three scenarios of breaks simulation. These strategies yield to estimations with a very small bias. Moreover, the strategy with the best performances in terms of prediction and with the smallest bias is the one which does not take into account of breaks. This result is important as, in practice, collecting breaks data is pretty hard to manage.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

In order to get marketing authorization, a new product has to succeed in clinical trials. A clinical trial is based on statistical considerations in order to show the product efficacy, taking into account the variability of the environment. It is a well known fact that the power of this test is linked to the number of patients we deal with. If an inadequate number of enrolled patients is used, then the study may fail to reject the null hypothesis due to lack of power. In many cases the goal number of patients to include is thus a fixed parameter of the trial. There has been much effort in computing the

sample size for clinical trials. Its computation is now standard in trial protocols (see ICH E9 guidance) and mandatory for most of the publications (see Consort Group works Schulz et al. [20]). On the other side relatively little attention is focused on improving the prediction of the recruitment process. Indeed, till now the most of techniques used by pharmaceutical companies are based on deterministic models and various *ad hoc* techniques. Rojavin [19] says "Patient recruitment and retention remains until now more of an art rather than a science".

The problem of predicting patients recruitment and evaluating the recruitment time in clinical trials is of paramount interest for planning trials because of scientific concern, economic and ethical reasons.

- **Ethical concern**, because it is not satisfactory to continue a study in vain.

* Corresponding author. Toulouse Institute of Mathematics, UMR 5219, CNRS, F-31062, Toulouse, France.

** Corresponding author. University of Toulouse III, F-31073, Toulouse, France.
E-mail address: nicolas.savy@math.univ-toulouse.fr (N. Savy).

- **Economical concern**, a clinical trial is an expensive study in itself and, as the duration of the trials is included in the duration of the exclusive right to exploit the drug (20 years) [22], a delay generates an enormous loss of income. Moreover, an improvement of the planning and monitoring of a trial may reduce costs.
- **Scientific concern**, because new drugs are increasingly developed and approved by regulatory agencies. When accrual rates are too low, there may be new information available during the enrolment period such as the results of other trials or a change in the understanding of the underlying biology.

For these reasons, stopping or continuing a trial is a decision with huge consequences and it will be useful to have some objective tools based on scientific criteria to take it.

Few authors have considered the problem of patients' recruitment. The reader can refer to Barnard et al. [11] for a systematic review of the existing models for recruitment and Anisimov [7], Bakhshi et al. [10], Gajewski et al. [13], Heitjan et al. [14] for a presentation and/or a discussion of methods. As far as we know, the pioneer work is the one of Morgan [18] where an estimation of the total study duration is proposed as a function of inclusion duration and based on data from previous clinical trials. Let us cite Lee [15] for a model of recruitment by Poisson processes. In Senn [21] a model for multicentric trials based on Poisson process is introduced. Poisson process appears as a natural assumption in literature [21]. However, Poisson processes depends on only one parameter, which is the rate of enrolment in our setting and Carter [12] have noticed that the use of the historic mean is a too simple model and highlighted the necessity to take into consideration the variability of the rates.

Inspired by queueing theory, Anisimov and Fedorov [9] proposed to use a doubly stochastic Poisson process to take into consideration the variation in recruitment rates between different centres. This model, called as a Poisson-gamma model, assumes that the patients arrive at different centres according to Poisson processes with the rates viewed as independent gamma distributed random variables. In Anisimov and Fedorov [9] the procedure of parameters estimation from data collected up to an interim time using empirical Bayesian technique have been suggested. The model has been validated using data from a large number of real trials [3]. The Poisson-gamma model was developed further for predicting recruitment process at initial and interim stages [1], to account for the situations when the centres opening dates may not be known and assumed to be uniformly distributed in some intervals [2,17] or when some centres can be closed or opened in the future [6]. Sensitivity analysis to errors in parameters' estimation has been investigated in Mijoule et al. [17]. Poisson-gamma model allows to develop techniques for analysing the effects of unstratified and centre-stratified randomization [4], for predictive event modelling [5], for predict randomization process [6] and for modelling clinical trials' cost [16].

The Poisson-gamma approach on the modelling of patients' recruitment is more and more popular. However, the assumption on the rates, which are assumed to be constant in time, has to be discussed. Indeed, in practice, there are breaks in recruitment process. A break is defined as a period during which a centre does not recruit any patient (holidays, week-end,...). This information is observed and can be collected but, in practice, it is a huge and complicated process. The Poisson-gamma model can be enriched considering piecewise constant rate function yielding to inhomogeneous Cox models. In this case the estimation of the trial duration is much more difficult. Three strategies of estimation of recruitment duration are proposed in this paper: one taking account of breaks, one not taking account of breaks and a third taking into account of only large breaks. These estimations are

biased. The bias is assessed by means of simulation study and the strategies are compared in terms of bias in estimation and in terms of predictive performances. Finally, these investigations allow us to deal with a question of paramount interest: Is it really useful to enrich the model in order to take into account breaks in recruitments?

The paper is organized as follows. Section 2 describes the Poisson-gamma model. The procedure of parameters estimation and the computation of the expected duration of the trial are given. Section 3 presents strategies which can be used to take into account of breaks in the recruitment model. In Section 4 authors explain the data generation procedure which allows to perform the simulation studies. These studies investigate the bias and the predictive performances of the strategies introduced in Section 3 together with the comparison of these strategies. The results are presented in Section 5. Finally the paper ends with a concluding Section 6.

2. The Poisson-gamma model without breaks in recruitment

2.1. Notations

Consider a multicentric clinical trial where C centres are involved to recruit n patients. Denote u_i the opening date of the i -th centre which is assumed to be observed. The recruitment process of centre i is denoted $\{N_t^i, t \geq u_i\}$ and is modelled by a Poisson process with intensity λ_i . The global inclusion process is $N = \sum_{i=1}^C N^i$. The parameter of paramount interest is the stopping time:

$$T = \left\{ \inf_{t \geq 0} : N_t \geq n \right\}.$$

In the sequel we consider, for any i , λ_i as a random variable which is gamma-distributed with parameters (α, β) whose probability density function is:

$$p_{\alpha, \beta}(x) = x^{\alpha-1} \frac{\beta^\alpha e^{-\beta x}}{\Gamma(\alpha)} 1_{\{x > 0\}},$$

$$\text{where } \Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt.$$

2.2. Estimation

In most settings, parameter (α, β) is unknown. To estimate this parameter, an empirical Bayesian strategy is used. Fix an interim time of analysis t_1 . Data collected on $[0, t_1]$ are used to calibrate the model. For any centre i , denote $\tau_{1,i} = (t_1 - u_i) \vee 0$ the duration of activity up to t_1 of centre i and $k_{1,i} = N_{t_1}^i$ the number of patients recruited by centre i up to t_1 . Notice that $\{(\tau_{1,i}, k_{1,i}), i = 1, \dots, C\}$ are observed data.

Theorem 1. (Anisimov et al. [8]). *Maximum likelihood estimation $(\hat{\alpha}, \hat{\beta})$ of the parameter (α, β) is obtained by maximization of the function:*

$$M_C^T(\alpha, \beta) = \alpha \ln(\beta) - \ln \Gamma(\alpha) + \frac{1}{C} \sum_{i=1}^C [\ln \Gamma(\alpha + k_{1,i}) - (\alpha + k_{1,i}) \ln(\beta + \tau_{1,i})].$$

2.3. Prediction

If centre i is initiated at time 0 ($u_i = 0$), the so-called forward rate $\hat{\lambda}_{1,i}$ knowing $\{N_{t_1}^i = k_{1,i}\}$ is Gamma-distributed with parameters $(\hat{\alpha} + k_{1,i}, \hat{\beta} + \tau_{1,i})$. If centre i is not initiated at time 0, the forward instantaneous intensity is time dependent and expressed

as:

$$t \rightarrow \hat{\lambda}_{1,i} \mathbf{1}_{\max(t_1, u_i) \leq t},$$

where $\hat{\lambda}_{1,i}$ is a predictive posterior rate in center i with distribution $\Gamma(\hat{\alpha} + k_{1,i}, \hat{\beta} + \tau_{1,i})$. The prediction process is now written as:

Theorem 2. (Anisimov et al. [8]). Given, $n_1 = \sum_{i=1}^C k_{1,i}$, the predictive recruitment process N expressed, for any $t > t_1$ as:

$$N_t = n_1 + \bar{N}_t^1, \quad \text{where} \quad \bar{N}_t^1 = \sum_{i=1}^C \bar{N}_t^{1,i},$$

and $\bar{N}^{1,i}$ is a Cox process whose predictive rate is $\Gamma(\hat{\alpha} + k_{1,i}, \hat{\beta} + \tau_{1,i})$ -distributed starting at time t_1 .

2.4. Expected duration

Consider $\bar{n}_1 = n - n_1$ the number of patients remaining to recruit after t_1 and $\bar{T}^1 = \{\inf_{t \geq t_1} : \bar{N}_t^1 \geq \bar{n}_1\}$ the remaining inclusion time. Denote $\hat{\Lambda}_1 = \sum_{i=1}^C \hat{\lambda}_{1,i}$.

Theorem 3. (Anisimov [6]). • Assume that all centres are initiated at the same time ($u_i = u > 0$, for any i) and denote $\tau = (t_1 - u) \vee 0$. Then \bar{T}^1 is $P_{VI}(\bar{n}_1, \hat{\alpha}C + n_1, \hat{\beta} + \tau)$ -distributed where $P_{VI}(n, a, b)$ denotes the Pearson VI distribution whose probability density function is:

$$p_{n,a,b}(x) = \frac{1}{\mathcal{B}(n, a)} \frac{x^{n-1} b^a}{(x + b)^{n+a}},$$

where $\mathcal{B}(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$ is the beta function.

- In practice, the $\tau_{1,i}$'s may be different. The distribution of \bar{T}^1 can be approximated for large n by a $P_{VI}(\bar{n}_1, \hat{A}_1, \hat{B}_1)$ distribution with

$$\hat{A}_1 = \frac{\left(\sum_{i=1}^C \hat{m}_{1,i}\right)^2}{\sum_{i=1}^C \hat{v}_{1,i}}, \quad \hat{B}_1 = \frac{\sum_{i=1}^C \hat{m}_{1,i}}{\sum_{i=1}^C \hat{v}_{1,i}}$$

where

$$\hat{m}_{1,i} = \frac{\hat{\alpha} + k_{1,i}}{\hat{\beta} + \tau_{1,i}}, \quad \hat{v}_{1,i} = \frac{\hat{\alpha} + k_{1,i}}{(\hat{\beta} + \tau_{1,i})^2}$$

Recall that for X a random variable $P_{VI}(n, a, b)$ -distributed with $a > 2$, the expectation and the variance are:

$$\mathbf{E}[X] = \frac{bn}{a-1}, \quad \mathbf{V}[X] = \frac{b^2n(n+a-1)}{(a-1)^2(a-2)}.$$

As a consequence, by Theorem 2 the expression of the expected duration of the trial is:

$$\mathbf{E}[T] = \begin{cases} t_1 + \bar{n}_1 \frac{\hat{B}_1}{\hat{A}_1 - 1} & \text{if } \hat{A}_1 > 2, \\ +\infty & \text{if } 0 < \hat{A}_1 \leq 2. \end{cases}$$

3. The Poisson-gamma model with breaks in recruitment

Assume that the recruitment process for centre i stops at some time denoted $b_{i,j}$ for periods denoted $d_{i,j}$. As in previous section, fix an interim time of analysis t_1 . Data collected on $[0, t_1]$ will be used to calibrate the model. For centre i , the data collected are:

- the number of recruited patients up to t_1 denoted $k_{1,i}$,
- the number of breaks up to t_1 :

$$j_{1,i} = \inf\{j : b_{i,j} < t_1, \quad \text{and} \quad b_{i,j} + d_{i,j} \geq t_1\},$$

- $(b_{1,i,j} = 1, \dots, j_{1,i})$ (resp. $(d_{1,i,j} = 1, \dots, j_{1,i})$) the breaks times (resp. durations) up to t_1 .

The duration of activity for centre i up to t_1 is thus $\tau_{1,i} = (t_1 - u_i - \sum_{j=1}^{j_{1,i}} d_{1,i,j}) \vee 0$. The recruitment dynamic for centre i , still denoted as N^i , is a non-homogeneous Cox process with intensity governed by λ_i which is $\Gamma(\alpha, \beta)$ distributed. Indeed, the instantaneous intensity for centre i is time dependent and expressed as:

$$t \rightarrow A(\lambda_i, t) = \lambda_i \mathbf{1}_{\{t \notin \mathcal{D}_{1,i}\}} \mathbf{1}_{\{t > u_i\}}, \tag{1}$$

with $\mathcal{D}_{1,i} = \{t : \exists j \in \{1, 2, \dots, j_{1,i}\}; t \geq b_{1,i,j} \text{ and } t \leq b_{1,i,j} + d_{1,i,j}\}$.

3.1. Estimation

The following theorem insures that (α, β) are estimated following the same strategy as in Section 2. Only the definition of $\tau_{1,i}$ differ.

Theorem 4. Maximum likelihood estimation $(\hat{\alpha}, \hat{\beta})$ of the parameter (α, β) is obtained by maximization of the function:

$$M_C^\Gamma(\alpha, \beta) = \alpha \ln(\beta) - \ln \Gamma(\alpha) + \frac{1}{C} \sum_{i=1}^C [\ln \Gamma(\alpha + k_{1,i}) - (\alpha + k_{1,i}) \ln(\beta + \tau_{1,i})].$$

3.2. Prediction

When considering breaks, the difficulty comes from the Cox process N modelling the recruitment which is non-homogeneous because of potential breaks in the dynamic. To overpass this difficulty, consider \tilde{N}^1 a homogeneous Cox process starting at t_1 with intensity $\sum_{i=1}^C \hat{\lambda}_{1,i}$. Recall \bar{N}^1 the remaining recruitment process defined in Theorem 2. Processes \bar{N}^1 and \tilde{N}^1 have the same intensities but the first one allows breaks while the second one does not (see Fig. 1).

Consider $T = \{\inf_{t \geq 0} : N_t \geq n\}$ the “true” duration of the clinical trial and $\bar{T}^1 = \{\inf_{t \geq t_1} : \bar{N}_t^1 \geq n\}$ (respectively $\tilde{T}^1 = \{\inf_{t \geq t_1} : \tilde{N}_t^1 \geq n\}$) the remaining time after t_1 considering there are (respectively there are no) breaks in the recruitment. Obviously $\tilde{T}^1 \leq \bar{T}^1$. The duration denoted by (B) on Fig. 1 is thus interpreted as the cumulated breaks duration.

3.3. Expected duration

The expected duration cannot be estimated directly but, as $\mathbf{E}[T] = t_1 + \mathbf{E}[\bar{T}^1] + \mathbf{E}[\bar{T}^1 - \tilde{T}^1]$, an estimation can be proposed.

Indeed, on the one hand, Theorem 3 gives us an estimation of $\mathbb{E}[\tilde{T}^1]$ since \tilde{N}^1 is a homogeneous Cox process and, on the other hand, assuming that breaks behave after t_1 as before t_1 and that the cumulated breaks duration is proportional to the duration of the follow up, $\mathbb{E}[\tilde{T}^1 - \tilde{T}^1]$ can be estimated by

$$\overline{BC}^1 = \frac{\mathbb{E}[\tilde{T}^1]}{\sum_{i=1}^C (t_1 - u_i)} \sum_{i=1}^C \sum_{j=1}^{j_{i,i}} d_{1,i,j}, \tag{2}$$

since $\sum_{i=1}^C \sum_{j=1}^{j_{i,i}} d_{1,i,j}$ is the cumulated breaks duration on $[0, t_1]$ and $\sum_{i=1}^C (t_1 - u_i)$ is the cumulated duration of activity of the centres on $[0, t_1]$. Replacing $\mathbb{E}[\tilde{T}^1]$, which cannot be computed, by $\mathbb{E}[\tilde{T}^1]$ yield to the following estimation of $\mathbb{E}[\tilde{T}^1 - \tilde{T}^1]$.

$$\widetilde{BC}^1 = \frac{\mathbb{E}[\tilde{T}^1]}{\sum_{i=1}^C (t_1 - u_i)} \sum_{i=1}^C \sum_{j=1}^{j_{i,i}} d_{1,i,j}. \tag{3}$$

The simulation study of Section 4 aims to quantify the bias when $\mathbb{E}[\tilde{T}^1 - \tilde{T}^1]$ is estimated by \widetilde{BC}^1 .

3.4. Another strategy

In practice, it may be useful to consider only sufficiently large breaks (greater than a predefined value denoted d_{\max}). Consider t_1 an interim time of analysis of the recruitment process. For the center i , the data collected on $[0, t_1]$ are:

- the number of recruited patients up to t_1 denoted $k_{1,i}$,
- the number of breaks which duration is greater than d_{\max} up to t_1

$$j_{1,i} = \inf \{j : b_{i,j} < t_1, \text{ and } b_{i,j} + d_{i,j} \geq t_1 \text{ and } d_{i,j} \geq d_{\max}\},$$

- the breaks times and durations up to t_1 greater than d_{\max} denoted $(b_{1,i,j} \in J)$ and $(d_{1,i,j} \in J)$, with

$$J = \{j = 1, \dots, j_{1,i}; d_{1,i,j} \geq d\}.$$

The duration of activity up to t_1 expressed as $\tau_{1,i} = (t_1 - u_i - \sum_{j \in J} d_{1,i,j}) \vee 0$. The rate of recruitment for centre i is governed by λ_i which is $\Gamma(\alpha, \beta)$ -distributed. The instantaneous intensities are still time dependent and expressed, for any centre i , as:

$$t \rightarrow \lambda_i \mathbf{1}_{\{t \notin \mathcal{D}_{1,i}\}} \mathbf{1}_{\{t > u_i\}},$$

with $\mathcal{D}_{1,i} = \{t; \exists j \in J; t \geq b_{1,i,j} \text{ and } t \leq b_{1,i,j} + d_{1,i,j}\}$. Considering these new notations and definitions, the estimation of the parameters of the model, the predictive process and expected duration of the trial can be computed as explained in sections 3.1–3.3.

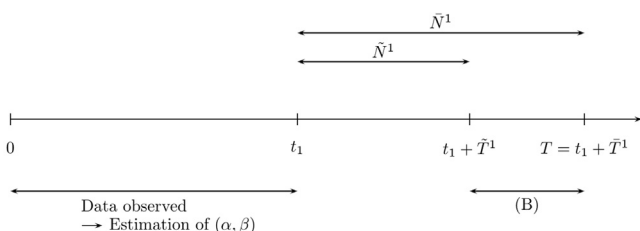


Fig. 1. Processes involved in the prediction of T .

4. Data generation procedure

Consider a multicentric trial involving $C = 60$ centres. We aim to recruit $n = 720$ patients in 365 days.

Remark 1. The whole numerical values introduced in this section to perform the data generation procedure are chosen in such a way that the recruitment dynamic is consistent with what is observed on real trial. To do so, authors have taken the parameters of the real case study of Anisimov and Fedorov [9].

In order to investigate different approaches of the breaks dynamic, different scenarios are proposed. Scenarios differ by the breaks generation procedure (times of breaks and durations of breaks). The scenarios are:

- **Scenario 1: Exponential generation.** The instants and durations of breaks are generated according to exponential distributions. The breaks times are exponentially distributed with intensity $1/60$ this means a break appears on average every 60 days. The breaks durations are exponentially distributed with intensity $1/14$ this means the average break duration is 14 days.
- **Scenario 2: Multinomial generation.** The instants of breaks are generated according to an exponential distribution with parameter $1/60$ meanwhile the durations are generated according to a multinomial probability. Five levels of duration (2, 4, 8, 16 and 32 days) are involved. The corresponding probability vector is built in such a way that it ensures that the total durations for each level are the same, (the breaks of 2 days happen twice more than the one of 4 days).
- **Scenario 3: Deterministic generation.** The instants of the breaks are the Saturdays and Sundays and a complete week every two months for holidays.

Recruitment dynamic is generated involving breaks themselves generated according to the scenarios defined above. The whole of the dynamic is known thus the true duration denoted T_0 of the trial is known. The study consists in considering the data collected on $[0, t_1]$ and to make use of the results of sections 2 and 3 to estimate the duration of the trial. The three strategies of estimation of the trial duration introduced in sections 2.4, 3.3 and 3.4 are considered:

- **Strategy 1: not taking into account of breaks.** The breaks times and durations are not collected. The parameters of the Poisson-gamma model for recruitment are estimated following results of Section 2. The expected duration is denoted by T_1 and is computed by means of Theorem 3.
- **Strategy 2: Taking into account of all the breaks.** The breaks times and durations are collected. This allow us to make use of the estimation of the trial duration as explained in Section 3.3. The expected duration is denoted T_2 .
- **Strategy 3: Taking into account of only large breaks.** Only the breaks that have a duration greater than a predefined value of 14 days are collected and used as explained in Section 3.4. This strategy leads to the expected duration denoted T_3 .

For a sake of simplicity, all centres are initiated at $t = 0$ ($u_i = 0$ for all i). The data generation procedure splits in two steps:

Step 1: Generate $R = 1000$ recruitment processes according to scenario 1, 2 or 3. Consider, for any $1 \leq r \leq R$, $\{N^r(t), 0 \leq t \leq T_0^r\}$, where T_0^r denotes the first time verifying $N^r(t) = 720$. The generation procedure is the following one:

1. Generate the breaks according to the scenario 1, 2, and 3 considered for a period of 730 days. The duration of 730 days is

Table 1
Average duration of the trial (in days) as a function of p , the strategy, together with its 95% confidence Interval. Absolute error between expected duration and true duration ($E_{0,p}$) as a function of the strategy together with its 95% confidence interval and the proportion of over-estimation ($S_{0,p}$).

		Real duration	Strategy 1	Strategy 2	Strategy 3
Scenario 1	Mean	384.37	381.92	374.66	377.73
	CI (95%)	[351,415]	[340,431]	[336,421]	[337,426]
	$E_{0,p}$	–	13.81	15.42	14.41
	CI (95%)	–	[2,13]	[1,40]	[0,38]
	$S_{0,p}$	–	0.41	0.27	0.32
Scenario 2	Mean	337.92	337.8	336.41	337.16
	CI (95%)	[303,373]	[304,374]	[303,372]	[353,406]
	$E_{0,p}$	–	9.78	9.88	9.82
	CI (95%)	–	[0,29]	[1,29]	[0,29]
	$S_{0,p}$	–	0.44	0.40	0.42
Scenario 3	Mean	369.17	367.95	356.94	–
	CI (95%)	[343,398]	[330,412]	[321,398]	–
	$E_{0,p}$	–	11.04	14.88	–
	CI (95%)	–	[1,32]	[1,38]	–
	$S_{0,p}$	–	0.44	0.16	–

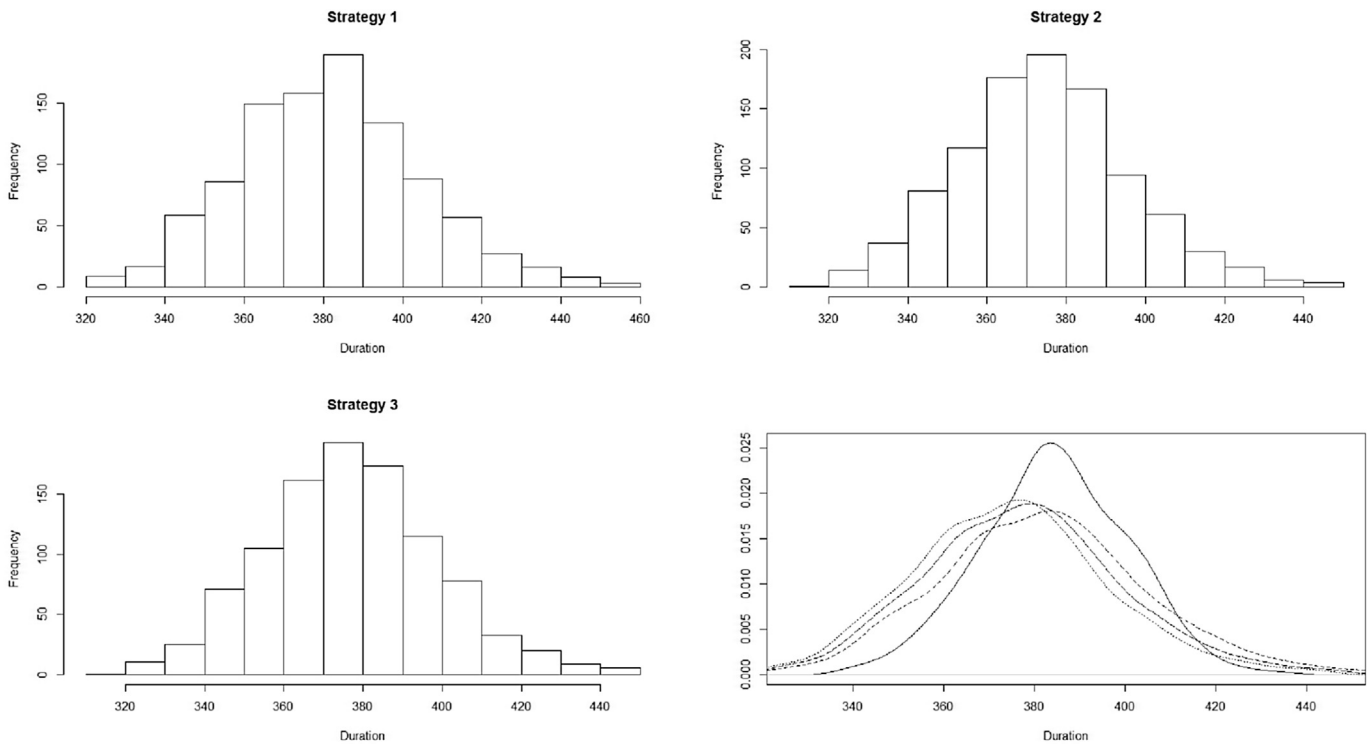


Fig. 2. Scenario 1. Histograms of expected duration for all strategies and comparison of the densities (solid line: T_0 , dashed line: T_1 , dotted line: T_2 , longdash line: T_3).

here arbitrary and chosen in order to be sure to catch the true duration of the trial.

2. Generate the rates according to a $\Gamma(2, 60.8)$ distribution.
3. Consider the modified rate function as defined in Equation (1).
4. Generate the recruitment process up to 730 days.
5. Identify T_0^r and shrink the recruitment process to $[0, T_0^r]$.

Step 2: Consider an interim time at $t_1 = 182$ days. For each simulation run $r = 1, \dots, R$ and each strategy $s = 1, 2, 3$,

1. Estimate parameters (α_s^r, β_s^r) of the gamma distribution applying Theorem 1 for $s = 1$ or Theorem 4 for $s = 2$ and $s = 3$ from data collected on $[0, t_1]$.
2. Compute the expected duration of the trial T_s^r through the application of Theorem 2 for $s = 1$ or following strategies explained in Section 3.3 for $s = 2$ and Section 3.4 for $s = 3$.

The performances of the model at interim time t_1 are measured by means of the absolute error defined by:

$$E_{s,s'} = \frac{1}{R} \sum_{r=1}^R |T_s^r - T_{s'}^r|, \quad \text{for } s, s' \in \{0, 1, 2, 3\}, \text{ and } s \neq s'.$$

Remark 2. The value chosen for the minimal break duration in analysis strategy 3 (14 days) is not coherent with the values chosen in scenario 3 (2 and 7 days). There is thus no difference between the first and the third scenario in this setting.

5. Results and discussion

5.1. Results on the bias

The first point investigated by the simulation study is the

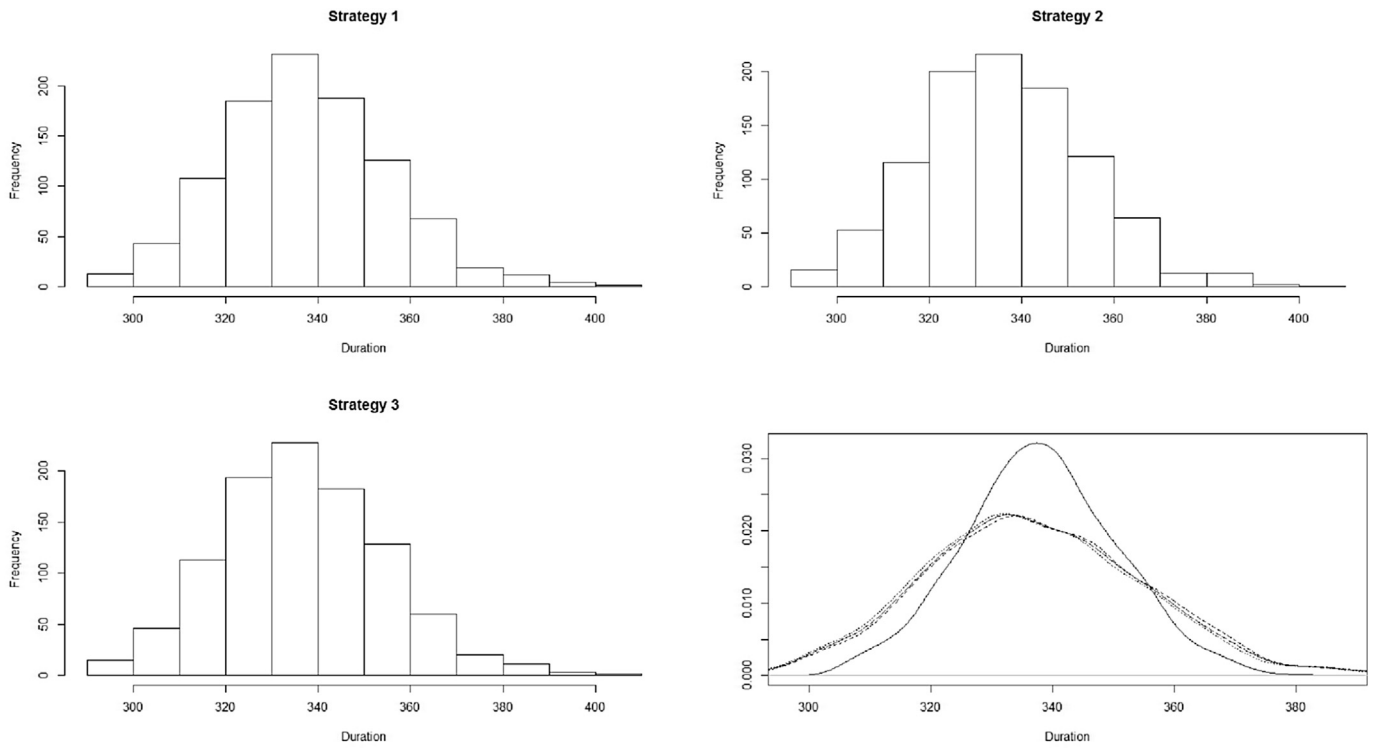


Fig. 3. Scenario 2. Histograms of expected duration for all strategies and comparison of the densities (solid line: T_0 , dashed line: T_1 , dotted line: T_2 , longdash line: T_3).

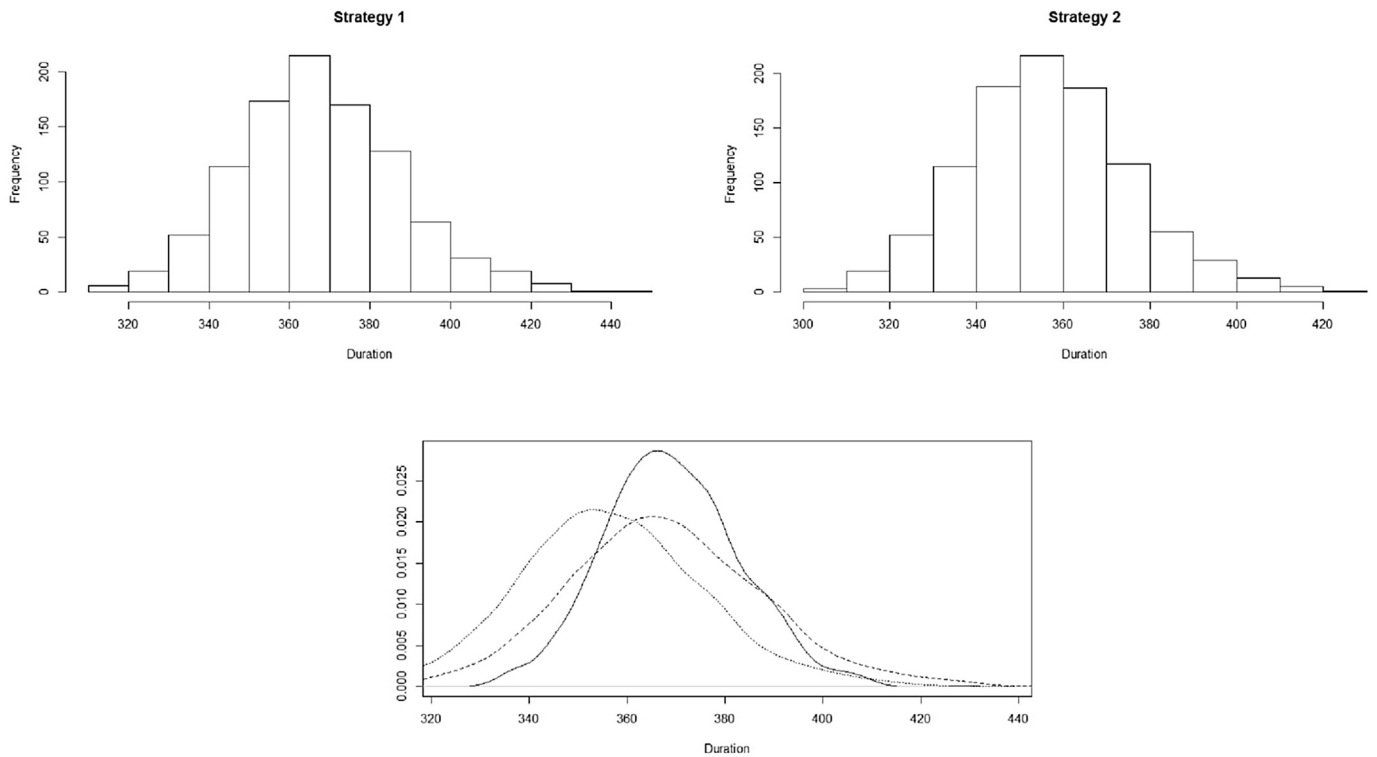


Fig. 4. Scenario 3. Histograms of expected duration for all strategies and comparison of the densities (solid line: T_0 , dashed line: T_1 , dotted line: T_2).

assessment of the bias embedded by the strategy of estimation of the expected duration for each scenario. The results are collected in Table 1. For each scenario, the mean duration (over the simulation runs) of the simulated recruitment dynamic together with its 95% confidence interval (which corresponds to the 25-th and 975-th values of the sorted samples) are identified. For Strategy p ($p = 1,2,3$), the average expected trial duration and its 95% confidence interval are computed. The bias is assessed by the average (over r) of the absolute errors made between the expected durations (T_p^r) and the true value of the trial duration (T_0^r) denoted $E_{0,p}$. Table 1 presents the value of $E_{0,p}$, its 95% confidence interval and the parameter denoted $S_{0,p}$ which measure the proportion of over-estimation defined as the proportion of runs for which the expected duration computed by means of Strategy p (T_p^r) is greater than the true duration (T_0^r).

These results are enriched by Fig. 2 (resp. 3 and 4) which refers to Scenario 1 (resp. 2, 3). Each figure is composed of the plots of the histograms of the ($R = 1000$) expected durations for each strategy together with the plot of the empirical density curves for each strategy completed by the one of the true durations. Keep in mind that, in accordance with Remark 2, Strategy 3 is not investigated for Scenario 3.

First, the computation strategies are efficient. Indeed, on Table 1, the mean value of the real duration is close to the mean durations estimated by the different strategies. This is confirmed by the values of $E_{0,p}$. Moreover, the width of the confidence intervals of $E_{0,p}$ is not large whatever the setting.

Second, for any scenario, the three strategies under-evaluate the trial duration. This is not surprising since the estimation (3) of $\mathbb{E}[\bar{T}^1 - \bar{T}^1]$ involves $\mathbb{E}[\bar{T}^1]$ instead of $\mathbb{E}[\bar{T}^1]$ and $\mathbb{E}[\bar{T}^1] \leq \mathbb{E}[\bar{T}^1]$. The

values of $S_{0,p}$ on Table 1 and Figs. 2–4 confirmed this results. Indeed, the histograms are uni-modal, not asymmetric and shift to the left comparing with the one of the real duration. The strategies are thus moderately biased.

5.2. Comparison of the strategies

Table 1 informs us that the error generated by Strategy 1 is minimal whatever the setting and that the confidence intervals for that strategy are the smallest. Strategy 1 is the less biased strategy with values of $S_{0,p}$ close to 0.5. Figs. 2–4 highlight this result.

To specify the comparisons of the three strategies, results analogous to those of Table 1 are collected in Table 2 for Scenario 1, Table 3 for Scenario 2 and Table 4 for Scenario 3. Each block of the table corresponds to a comparison of two strategies. These comparisons are given in terms of absolute error $E_{s,s'}$ together with its 95% confidence interval (which corresponds to the 25-th and 975-th values of the sorted samples) and the proportion $S_{s,s'}$ of over-estimations, this means the proportion of runs for which the expected duration computed by strategy s' is greater than the one computed by strategy s .

Whatever the scenario, Tables 2–4 indicate that the results obtained by the different strategies are more or less the same. Notice that, for Scenario 1, the largest difference is around 7 days for a study of 365 days which is very small. The confidence intervals are not large and the width is more or less the same whatever the strategy of computation. Finally the proportion of over-estimation S indicates that results from Strategy 1 over-estimate those from Strategy 3 which over-estimate those from Strategy 2.

5.3. Predictive performance of the strategies

On contrary to bias analysis, predictive behaviour analysis of a strategy is an individual concept. To investigate the predictive performances of a strategy, Fig. 5 related to Scenario 1, Fig. 6 to Scenario 2 and Fig. 7 to Scenario 3 are considered. Each figure splits in sub-plots which are the plots of the 1000 values of the real duration in abscissa and the expected duration in ordinate for each strategy of computation. The plots are enriched with the regression line and the straight line $y = x$.

Whatever the scenario and whatever the strategy, the regression analysis yields to the same conclusions: the quality of the regression model is good (quality of the fitting, normality, homoscedasticity and autocorrelation of the residuals values). It is thus possible to compare the slope to 1 and the intercept to 0 by means of Wald's tests and allow us to conclude that the differences observed on the different plots are not significant. The predictive performances are thus very good.

Differences between strategies and between scenarios are not significant but exist. These differences are connected to results on bias and illustrate the over-estimation issue already discussed. For instance, for Scenario 1, Fig. 5 shows that Strategy 1 is closest to the line $y = x$ and thus appears to be the best strategy (the slope are 0.94 for the first strategy, 0.87 for the second and 0.89 for the third). The under-estimation of the strategies regardless to the real duration is observable, the regression line is under the line $y = x$. It is easily seen on the plots that the difference between the regression line and the $y = x$ line increases with the true duration. The bias in the estimation is thus time dependent unlike what is observed for Scenario 3 with Strategy 2 (subplot on the right of Fig. 7).

6. Conclusions and recommendations

To conclude, results of Section 5 allow us to claim that the

Table 2
Results for Scenario 1. Comparison of the strategies of analysis in terms of absolute error together with its 95% confidence interval and the proportion of over-estimation.

		Strategy 1	Strategy 2
Strategy 2	$E_{2,p}$	7.27	–
	CI (95%)	[2,13]	–
	$S_{2,p}$	0.99	–
Strategy 3	$E_{3,p}$	4.25	3.07
	CI (95%)	[0,10]	[1,5]
	$S_{3,p}$	0.94	0.01

Table 3
Results for scenario 2. Comparison of the strategies of analysis in terms of absolute error together with its 95% confidence interval and the proportion of over-estimation.

		Strategy 1	Strategy 2
Strategy 2	$E_{2,p}$	1.44	–
	CI (95%)	[0,3]	–
	$S_{2,p}$	0.82	–
Strategy 3	$E_{3,p}$	0.82	0.83
	CI (95%)	[0,2]	[0,2]
	$S_{3,p}$	0.55	0.04

Table 4
Results for scenario 3. Comparison of the strategies of analysis in terms of absolute error together with its 95% confidence interval and the proportion of over-estimation.

		Strategy 1
Strategy 2	$E_{2,1}$	11.01
	CI (95%)	[9,14]
	$S_{2,1}$	1

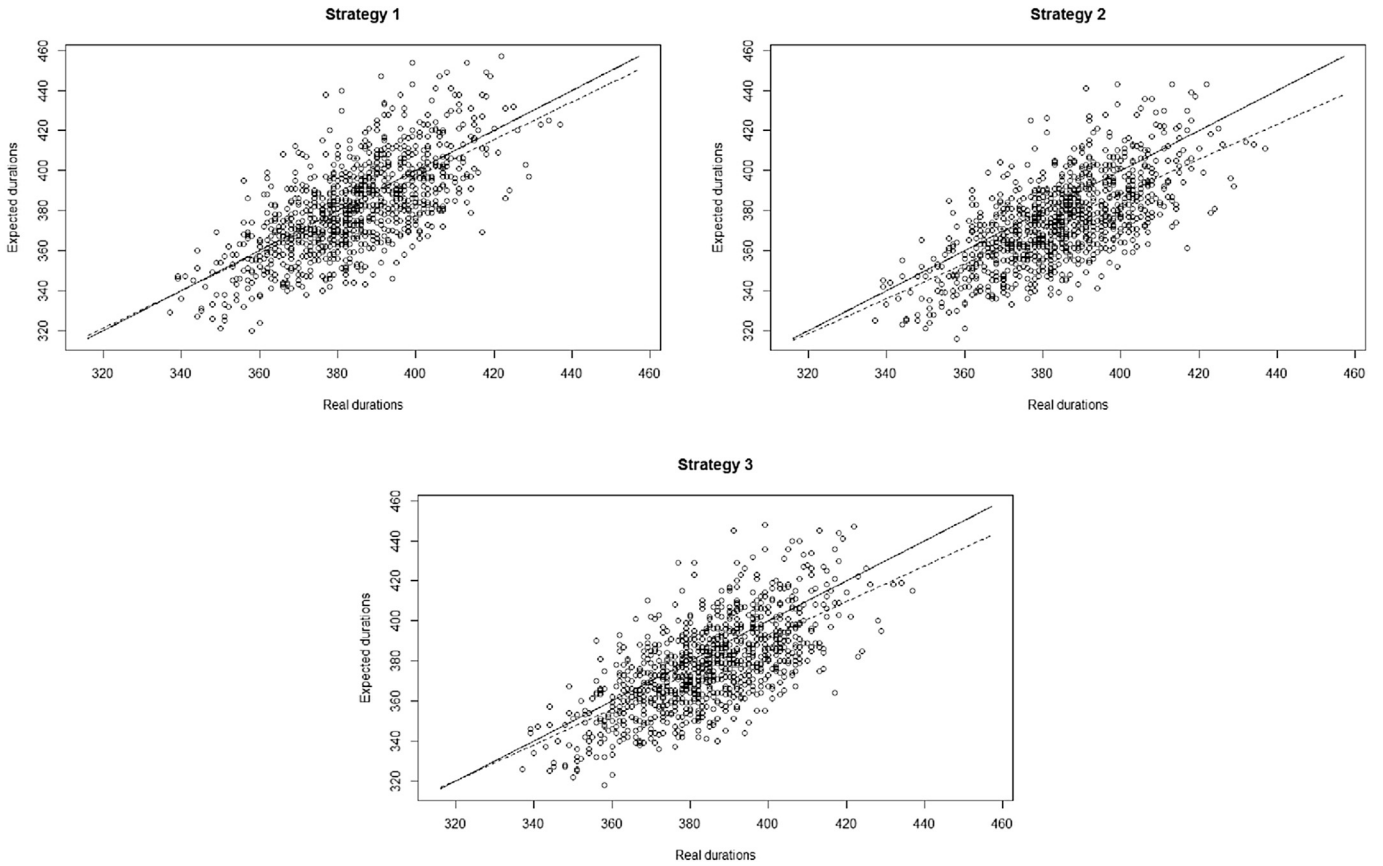


Fig. 5. Scenario 1. Linear Regression of the expected duration as a function of the real duration for each strategy. Solid line: $y = x$ line, dashed line: regression line.

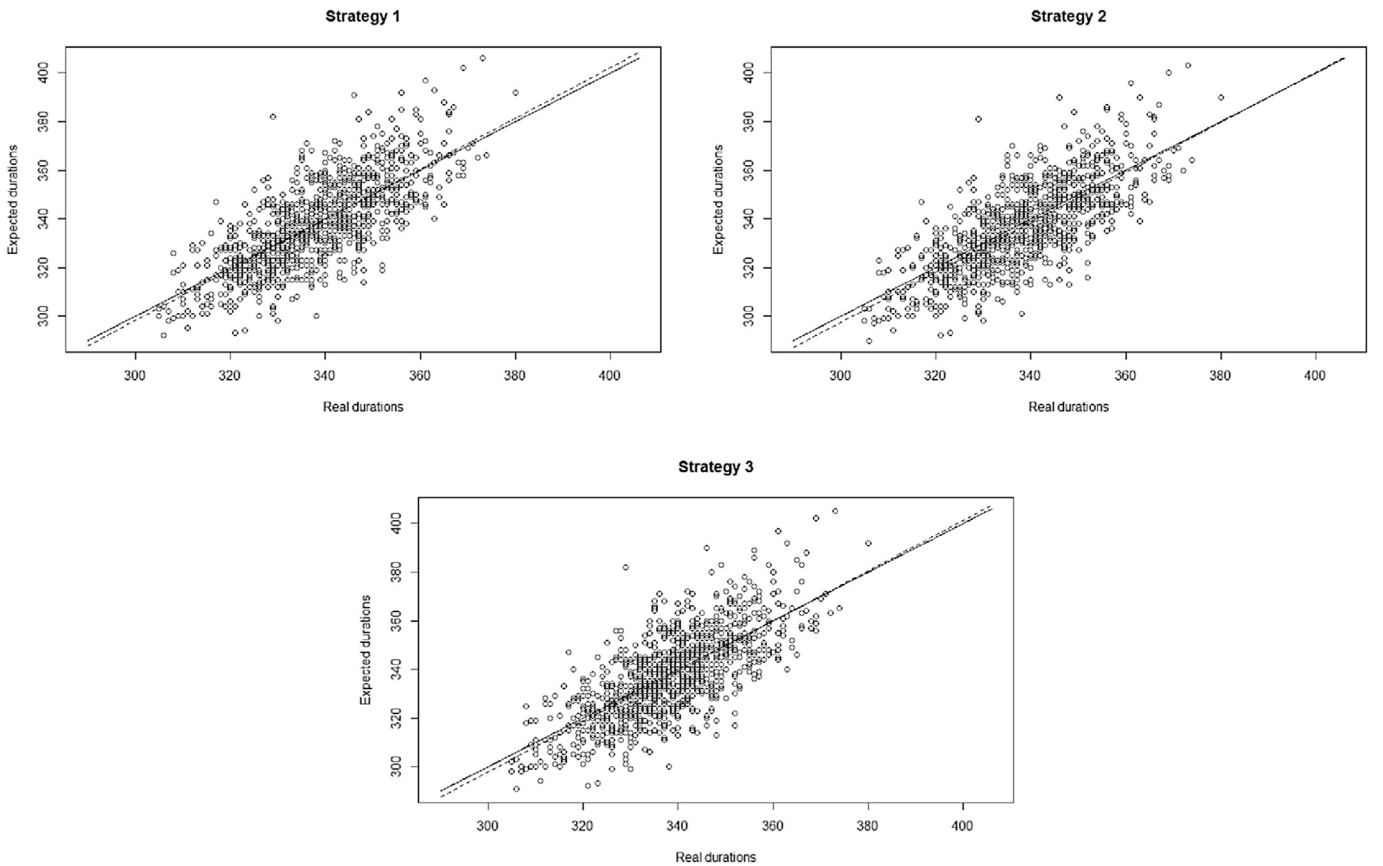


Fig. 6. Scenario 2. Linear Regression of the expected duration as a function of the real duration for each strategy. Solid line: $y = x$ line, dashed line: regression line.

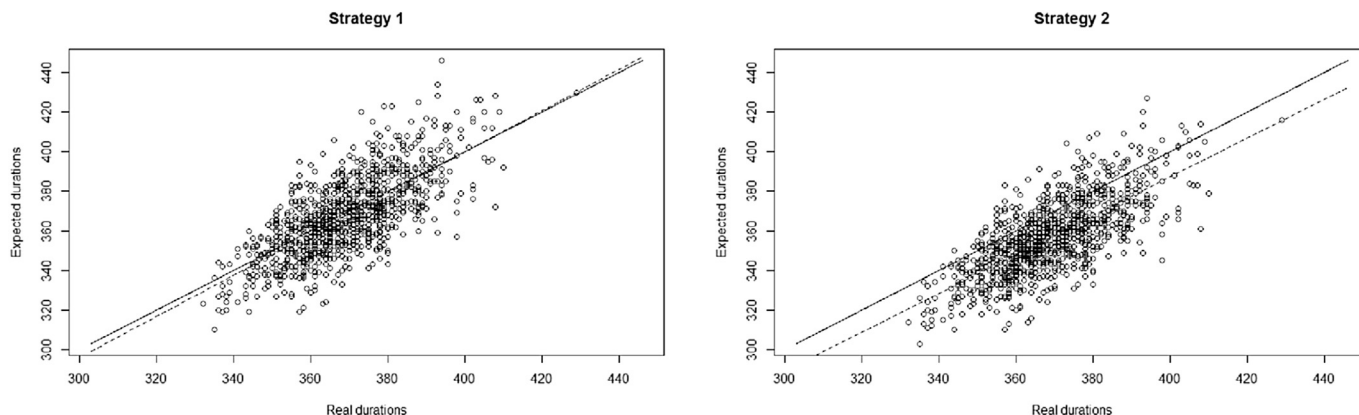


Fig. 7. Scenario 3. Linear Regression of the expected duration as a function of the real duration for each strategy. Solid line: $y = x$ line, dashed line: regression line.

procedures proposed in this paper to estimate the expected duration of the recruitment accounting of breaks in the recruitment process are moderately biased and have very good predictive performances. Notice that the assumption that the breaks behave after the interim time as they behave before the interim time is reasonable.

The main result states that the strategy with the best results is the one ignoring breaks. This is not surprising since, for this strategy, estimations (2) and (3) are the same and the induced bias is null. That strategy applies directly the Poisson-gamma model on the data collected. The consequence of the breaks is balanced by the estimation of the parameters of the gamma distribution.

Thus, finally, as a recommendation one suggests to not worry with breaks and to make use of a standard Poisson-gamma model. Indeed, to collect breaks data during recruitment process is a difficult process in practice and the impact on the recruitment process modelling is really moderate.

References

- [1] V.V. Anisimov, Using mixed Poisson models in patient recruitment in multicentre clinical trials, in: Proceedings of the World Congress on Engineering, vol. II, 2008, pp. 1046–1049. London, United Kingdom.
- [2] V.V. Anisimov, Predictive modelling of recruitment and drug supply in multicenter clinical trials, in: Proceedings of the Joint Statistical Meeting, ASA, Washington, USA, 2009, August, pp. 1248–1259.
- [3] V.V. Anisimov, Recruitment modeling and predicting in clinical trials, *Pharm. Outsourcing* 10 (1) (2009) 44–48.
- [4] V.V. Anisimov, Effects of unstratified and centre-stratified randomization in multi-centre clinical trials, *Pharm. Stat.* 10 (1) (2011) 50–59.
- [5] V.V. Anisimov, Predictive event modelling in multicentre clinical trials with waiting time to response, *Pharm. Stat.* 10 (6) (2011) 517–522.
- [6] V.V. Anisimov, Statistical modeling of clinical trials (recruitment and randomization), *Comm. Stat. Theory Methods* 40 (19–20) (2011) 3684–3699.
- [7] V.V. Anisimov, Discussion on the paper “Real-Time prediction of clinical trial enrollment and event counts: a review”, by DF Heitjan, Z Ge, and GS Ying, *Contemp. Clin. Trials* 46 (2016, Jan) 7–10.
- [8] V.V. Anisimov, D. Downing, V.V. Fedorov, Recruitment in multicentre trials: prediction and adjustment, in: 8th International Workshop in Model-oriented Design and Analysis, Almagro, Spain, Physica-Verlag/Springer, Heidelberg, 2007, June, pp. 1–8.
- [9] V.V. Anisimov, V.V. Fedorov, Modelling, prediction and adaptive adjustment of recruitment in multicentre trials, *Stat. Med.* 26 (27) (2007) 4958–4975.
- [10] A. Bakhshi, S. Senn, A. Phillips, Some issues in predicting patient recruitment in multi-centre clinical trials, *Stat. Med.* 32 (30) (2013, Dec) 5458–5468.
- [11] K.D. Barnard, L. Dent, A. Cook, A systematic review of models to predict recruitment to multicentre trials, *BMC Med. Res. Methodol.* 63 (10) (2010).
- [12] R.E. Carter, Application of stochastic processes to participant recruitment in clinical trials, *Control. Clin. Trials* 25 (5) (2004) 429–436.
- [13] B.J. Gajewski, S.D. Simon, S.E. Carlson, Predicting accrual in clinical trials with Bayesian posterior predictive distributions, *Stat. Med.* 27 (13) (2008, Jun) 2328–2340.
- [14] D.F. Heitjan, Z. Ge, G.S. Ying, Real-time prediction of clinical trial enrollment and event counts: a review, *Contemp. Clin. Trials* 45 (Pt A) (2015, Nov) 26–33.
- [15] Y.J. Lee, Interim recruitment goals in clinical trials, *J. Chronic Dis.* 36 (5) (1983) 379–389.
- [16] G. Mijoule, N. Minois, V. Anisimov, N. Savy, Additive model for cost modelling in clinical trial, in: Proceedings of the 7th International Workshop on Simulation, vol. 114, Springer Proceedings in Mathematics & Statistics, 2014, pp. 371–379.
- [17] G. Mijoule, S. Savy, N. Savy, Models for patients' recruitment in clinical trials and sensitivity analysis, *Stat. Med.* 31 (16) (2012) 1655–1674.
- [18] T.M. Morgan, Nonparametric estimation of duration of accrual and total study length for clinical trials, *Biometrics* 43 (4) (1987) 903–912.
- [19] M. Rojavin, Patient recruitment and retention: from art to science, *Contemp. Clin. Trials* 30 (5) (2009), 387–387.
- [20] K.F. Schulz, D.G. Altman, D. Moher, D.G. Altman, V. Barbour, J.A. Berlin, P.J. Devereaux, K. Dickersin, D. Elbourne, S. Ellenberg, V. Gebski, S. Goodman, P.C. Götzsche, T. Groves, S. Grunberg, B. Haynes, S. Hopewell, A. James, P. Juhn, P. Middleton, D. Minckler, D. Moher, V.M. Montori, C. Mulrow, S. Pocock, D.L. Schriger, K.F. Schulz, I. Simera, E. Wager, M. Clarke, G. Guyatt, Consort 2010 statement: updated guidelines for reporting parallel group randomised trials, *J. Clin. Epidemiol.* 63 (8) (Aug 2010) 834–840.
- [21] S. Senn, Some controversies in planning and analysing multi-centre trials, *Stat. Med.* 17 (1998) 1753–1765.
- [22] A.-P. Vlasto, Brevets et médicament en France, pourquoi l'application du droit des brevets au médicament est-elle autant critiquée? *Méd. Droit* 2007 (82) (2007) 25–32.