University of Zurich[UZH]

Master Thesis Biostatistics

Frequentists and Bayesian methods to incorporate recruitment rate stochasticity at the design stage of a clinical trial

Pilar Pastor Martínez

Supervision by Malgorzata Roos

May 20, 2025

Index

Recruitment and Patient Leakage

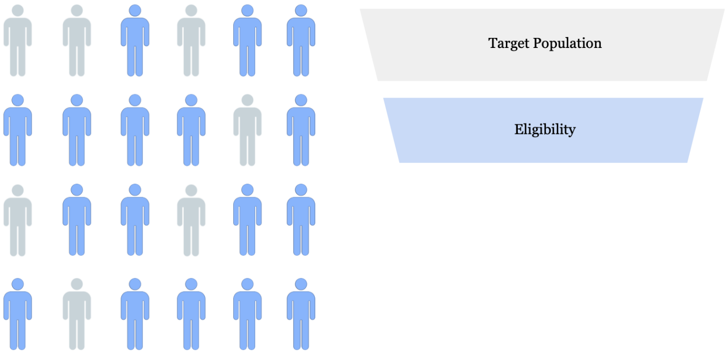
Why recruitment rates?

According to Carter (2004)

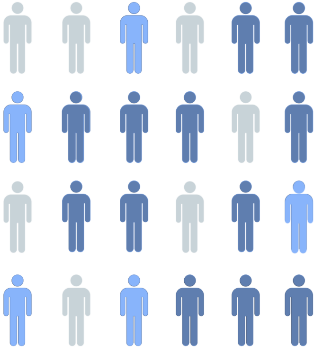
- Timely recruitment vital to the success of a clinical trial
- Inadequate number of subjects $\rightarrow$ lack of power
- Recruitment period too long $\rightarrow$ competing treatments
- Recruitment of patients varies at each stage
- Methods applicable to all the stages

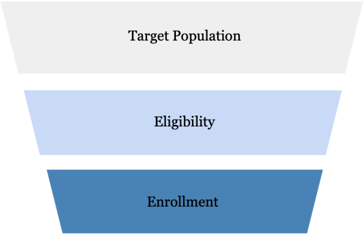
Target Population

Eligibility

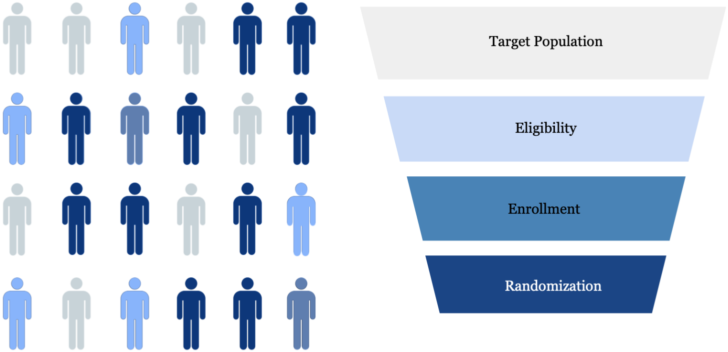


Enrollment

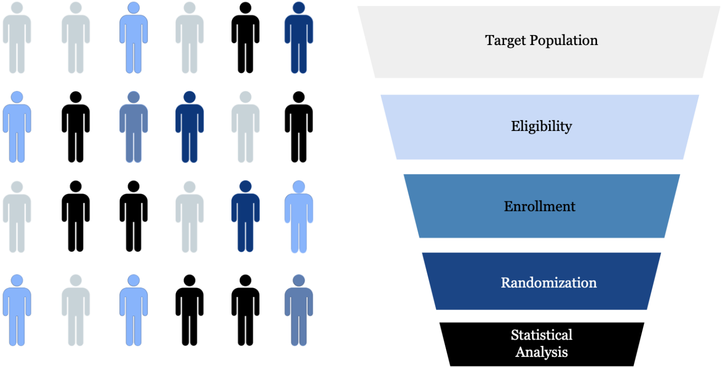




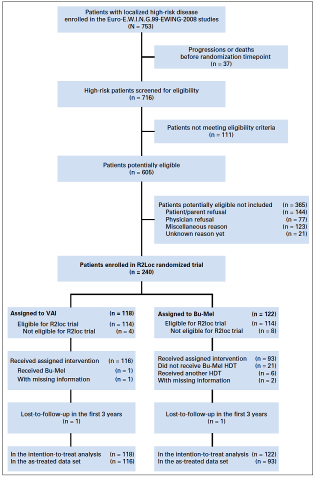
Randomization



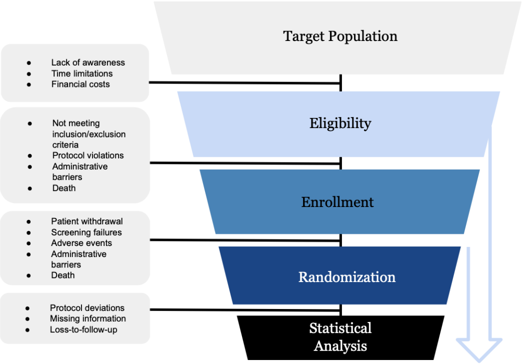
Statistical Analysis



CONSORT


Patients with localized high-risk disease enrolled in the Euro-E.W.I.N.G.99-EWING-2008 studies (N = 753)
Progressions or deaths before randomization timepoint (n = 37)
High-risk patients screened for eligibility (n = 716)
Patients not meeting eligibility criteria (n = 111)
Patients potentially eligible (n = 605)
Patients potentially eligible not included (n = 365)
Patient/parent refusal (n = 144)
Physician refusal (n = 77)
Miscellaneous reason (n = 123)
Unknown reason yet (n = 21)
Patients enrolled in R2Loc randomized trial (n = 240)
Assigned to VAI (n = 118)
Eligible for R2loc trial (n = 114)
Not eligible for R2loc trial (n = 4)
Assigned to Bu-Mel (n = 122)
Eligible for R2loc trial (n = 114)
Not eligible for R2loc trial (n = 8)
Received assigned intervention (n = 116)
Received Bu-Mel (n = 1)
With missing information (n = 1)
Received assigned intervention (n = 93)
Did not receive Bu-Mel HDT (n = 21)
Received another HDT (n = 6)
With missing information (n = 2)
Lost-to-follow-up in the first 3 years (n = 1)
Lost-to-follow-up in the first 3 years (n = 1)
In the intention-to-treat analysis (n = 118)
In the as-treated data set (n = 116)
In the intention-to-treat analysis (n = 122)
In the as-treated data set (n = 93)

Patient Leakage



Definitions

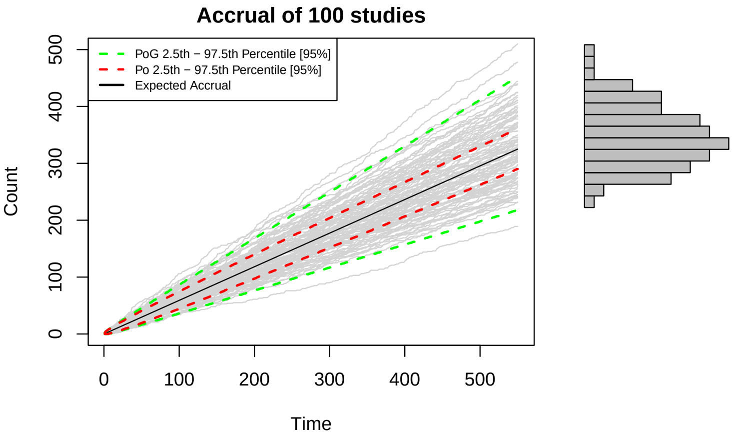
- **Recruitment rate**: Per time-unit (Piantadosi, 2024)

$$\lambda = \frac{\Delta C}{\Delta T} = \frac{C_1 - C_0}{T_1 - T_0} = \frac{C_1}{T_1}$$

- **Accrual**: Cumulative Recruitment
- **Aleatory uncertainty**: randomness inherent and unpredictable
- **Epistemic uncertainty**: arises from limited knowledge about parameters

Methods for Recruited Counts

Motivation Models for Counts



Models for Counts

Recruitment in unit of time (t=1):

Methods	Counts	Expectation	Variance	Aleatory	Epistemic
Expectation	$C = \lambda$	λ	0	No	No
Poisson	$C \sim \text{Po}\ (\lambda)$	λ	λ	Yes	No
Poisson - Gamma	$C \sim Po(\Lambda); \Lambda \sim G(\alpha, \beta)$	$\frac{\alpha}{\beta}$	$\frac{\alpha(\beta+1)}{\beta^2}$	Yes	Yes

Accrual for time t [0,t]:

Methods	Counts	Expectation	Variance	Aleatory	Epistemic
Expectation	$C(t) = \lambda t$	λt	0	No	No
Poisson	$C(t) \sim Po(\lambda t)$	λt	λt	Yes	No
Poisson - Gamma	$C(t) \sim Po(\Lambda t); \Lambda \sim G(\alpha, \beta)$	$t\frac{\alpha}{\beta}$	$t\frac{\alpha(\beta+t)}{\beta^2}$	Yes	Yes

Multicenter Trial on Palliation in Terminal Esophageal Cancer

Example from Carter (2004):

- Recruitment Rate $\lambda = \frac{\textit{Counts}}{\textit{Time}} = 0.591$ per day
- Time $t = 550$ days

Multicenter Trial on Palliation in Terminal Esophageal Cancer

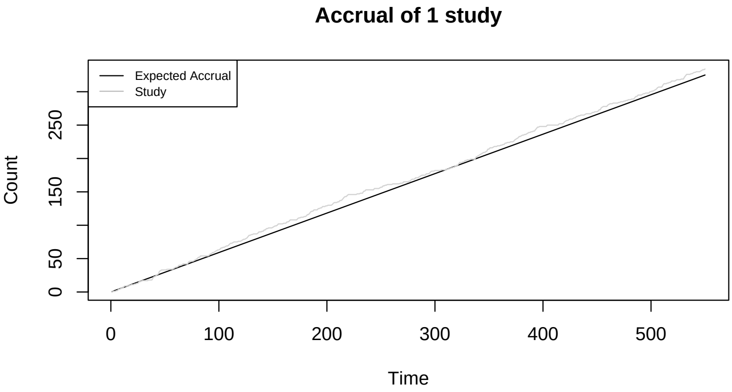
Example from Carter (2004):

- Recruitment Rate $\lambda = \frac{\textit{Counts}}{\textit{Time}} = 0.591$ per day
- Time $t = 550$ days
- Models for Counts at time point t:
  - **Expectation**: $EC(t) = \lambda t$
  - **Poisson**: $C(t) \sim Po(\lambda t)$
  - **Poisson - Gamma**: $C(t) \sim Po(\Lambda t)$; $\Lambda \sim G(\alpha, \beta)$

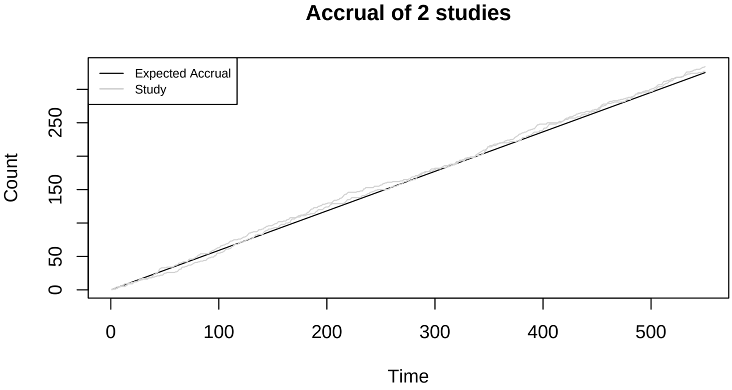
Accrual at time point t

- **Expectation**: $EC(t) = E\underbrace{(C + \ldots + C)}_{t \text{ times}} = tEC = \lambda t$
- **Poisson**: $\underbrace{Po(\lambda) + \ldots + Po(\lambda)}_{t \text{ times}} = Po(\lambda t)$

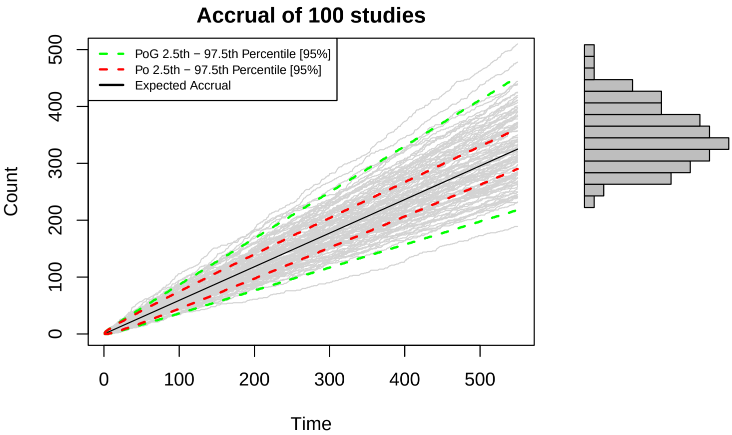
Accrual of 1 study



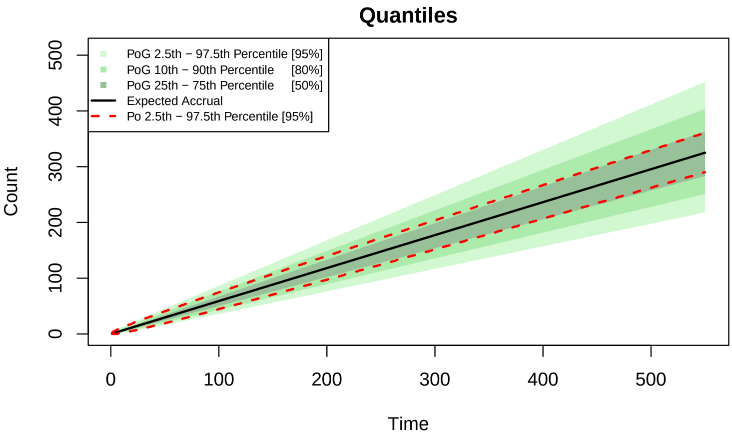
Accrual of 2 studies



Accrual of 100 studies



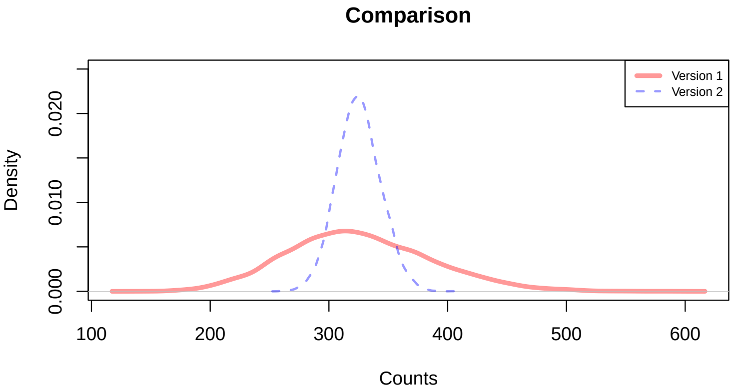
Poisson-Gamma’s uncertainty bands



Different Versions for generating λ

1. λ fixed over time and varies across studies
2. λ varies over time and across studies

Version 1 different from Version 2



Negative binomial derived from Poisson-Gamma model at time point t

Let $C(t)|\Lambda \sim Po(\Lambda t)$ and $\Lambda \sim G(\alpha, \beta)$

$$
\begin{aligned}
p(c) &= \int_0^\infty p(c|\lambda)p(\lambda)d\lambda \\
&= \int_0^\infty \frac{(\lambda t)^c \exp(-\lambda t)}{c!} \left[(\lambda)^{\alpha-1} \exp(-\beta\lambda) \frac{\beta^\alpha}{\Gamma(\alpha)}\right] d\lambda \\
&= \frac{\beta^\alpha t^c \Gamma(\alpha + c)}{c!\Gamma(\alpha)(\beta + t)^{\alpha+c}} \underbrace{\int_0^\infty \frac{(\beta + t)^{\alpha+c}}{\Gamma(\alpha + c)} \lambda^{\alpha+c-1} \exp(-(\beta + t)\lambda) d\lambda}_{=1} \\
&= \binom{\alpha + c - 1}{\alpha - 1} \left(\frac{t}{\beta + t}\right)^c \left(\frac{\beta}{\beta + t}\right)^\alpha,
\end{aligned}
$$

$$C(t) \sim NBin\left(\alpha, \frac{\beta}{\beta + t}\right)$$

Expectation and Variance for Counts

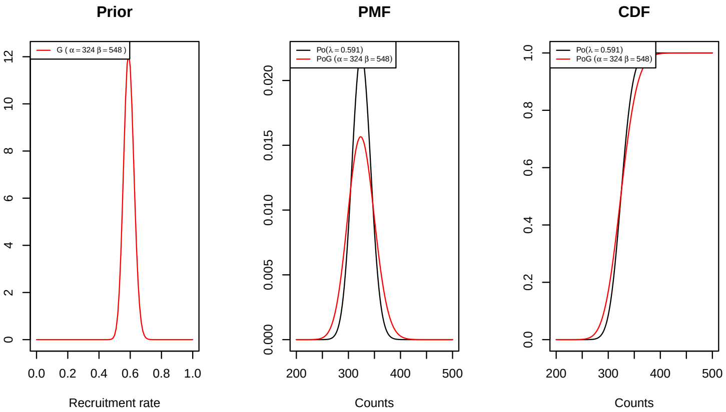
Using the expressions of iterated expectation and variance (Held and Bové, 2014)

$$E(C(t)) = E_\Lambda[E_{C(t)}(C(t)|\Lambda)] = E_\Lambda[\Lambda t] = t\alpha/\beta$$

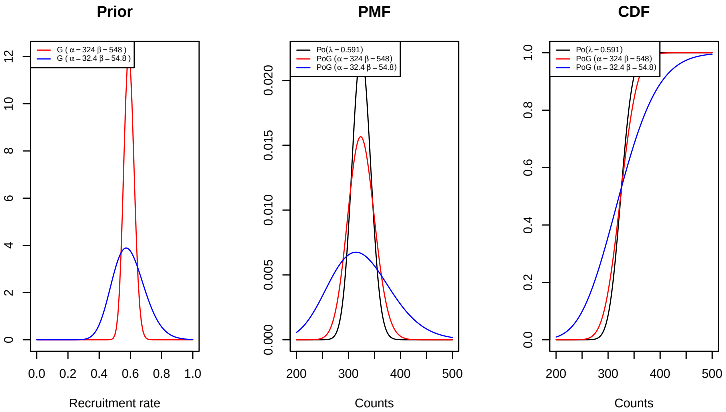
$$\begin{aligned} Var(C(t)) &= Var_\Lambda[E_{C(t)}(C(t)|\Lambda)] + E_\Lambda[Var_{C(t)}(C(t)|\Lambda)] \\ &= Var_\Lambda[\Lambda t] + E_\Lambda[\Lambda t] \\ &= t^2\alpha/\beta^2 + t\alpha/\beta = \frac{t\alpha(\beta + t)}{\beta^2} \end{aligned}$$

Sensitivity Analysis

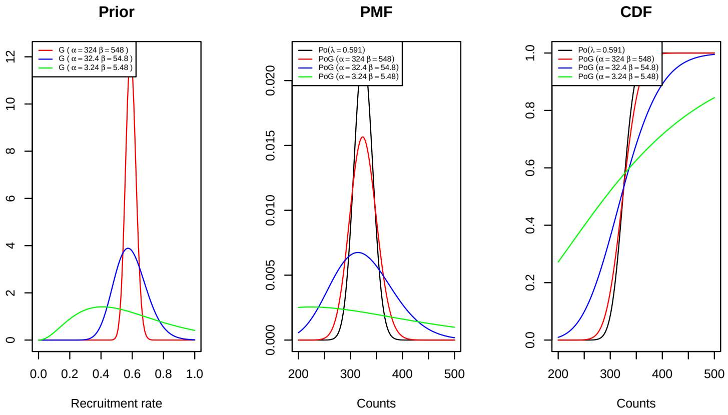
Sensitivity Analysis



Sensitivity Analysis

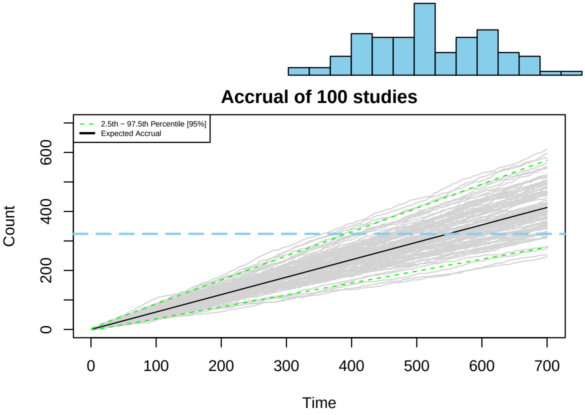


Sensitivity Analysis



Methods for Waiting Time

Motivation Models for Waiting Time



Models for Waiting Time with fixed sample sixe c

Methods	Time	Expectation	Variance	Aleatory	Epistemic
Expectation	$T(c) = c/\lambda$	c/λ	0	No	No
Erlang	$T(c) \sim \mathrm{G}(c, \lambda)$	c/λ	c/λ^2	Yes	No
Gamma-Gamma	$T(c) \sim \mathrm{G}(c, \Lambda); \Lambda \sim \mathrm{G}(\alpha, \beta)$	$c\frac{\beta}{\alpha-1}$	$\frac{c\beta^2(c+\alpha-1)}{(\alpha-1)^2(\alpha-2)}$	Yes	Yes

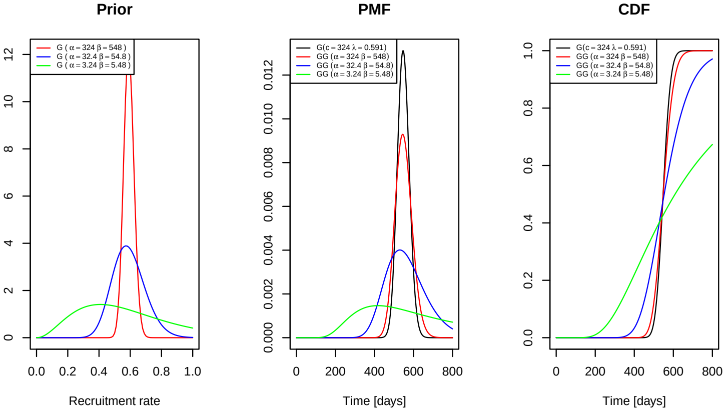
Expectation and Variance for Waiting Times

Using the expressions of iterated expectation and variance (Held and Bové, 2014) and that when $\Lambda \sim \mathrm{G}(\alpha, \beta)$ then, $\frac{1}{\Lambda} \sim \mathrm{IG}(\alpha, \beta)$

$$\mathrm{E}\,T(c) = \mathrm{E}_\Lambda[\mathrm{E}_{T(c)}(T(c)|\Lambda)] = \mathrm{E}_\Lambda\left[\frac{c}{\Lambda}\right] = c\mathrm{E}_\Lambda\left[\frac{1}{\Lambda}\right] = c\frac{\beta}{\alpha-1},\ \alpha > 1$$

$$\begin{aligned}
\mathrm{Var}(T(c)) &= \mathrm{Var}_\Lambda[\mathrm{E}_{T(c)}(T(c)|\Lambda)] + \mathrm{E}_\Lambda[\mathrm{Var}_{T(c)}(T(c)|\Lambda)] \\
&= \mathrm{Var}_\Lambda\left[\frac{c}{\Lambda}\right] + \mathrm{E}_\Lambda\left[\frac{c}{\Lambda^2}\right] \\
&= c^2\mathrm{Var}_\Lambda\left[\frac{1}{\Lambda}\right] + c\mathrm{E}_\Lambda\left[\frac{1}{\Lambda^2}\right] \\
&= \frac{c\beta^2(c+\alpha-1)}{(\alpha-1)^2(\alpha-2)},\ \alpha > 2
\end{aligned}$$

Sensitivity Analysis for Time



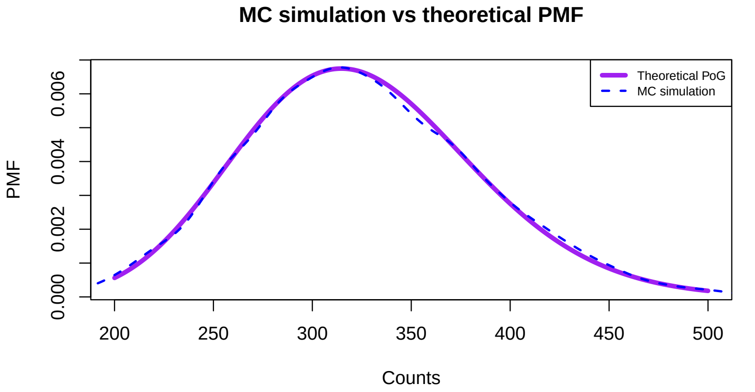
Comparison Exact Methods vs Monte Carlo

Comparison Exact Methods vs Monte Carlo

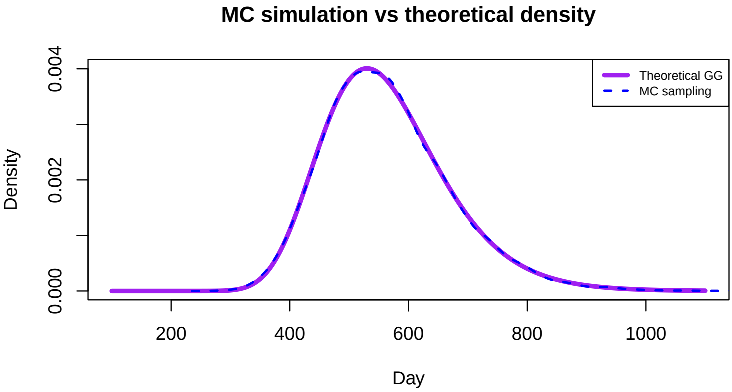
Model	Estimated Probabilty	MCse	Exact Probability
$C(T) \sim \mathrm{Po}(\lambda T)$	$\mathrm{P}(C(T) \geq 324) = 0.504$	0.005	0.508
$C(T) \sim \mathrm{PoG}(T, \alpha, \beta)$	$\mathrm{P}(C(T) \geq 324) = 0.48$	0.005	0.501

Model	Estimated Probabilty	MCse	Exact Probability
$T(C) \sim \mathrm{G}(C, \lambda)$	$\mathrm{P}(T(C) \geq 548) = 0.498$	0.005	0.496
$T(C) \sim \mathrm{GG}(C, \alpha, \beta)$	$\mathrm{P}(T(C) \geq 548) = 0.52$	0.005	0.52

Comparison Exact Methods vs Monte Carlo – Poisson-Gamma Counts



Comparison Exact Methods vs Monte Carlo – Gamma-Gamma Time



Carter's Questions

- $M = 10^3$ from Carter's $\rightarrow$ 580 days (innacurate)
- 90% chance of accruing $N = 324$ patients:
  - Erlang exact distribution $\rightarrow$ 588 days
  - Gamma-Gamma exact distribution $\rightarrow$ 707 days

Conclusions

- **Visual tools:** Graphs clarify recruitment flow, delays, and patient leakage in trials
- **Unified Notation:** Consistent math framework allows precise analysis of count and time models
- **Exact Methods:** Extended Monte Carlo methods to capture both aleatory and epistemic uncertainty
- **Flexible Recruitment:** Framework supports both fixed and time-varying recruitment rates
- **Practical Impact:** Exact methods aid trial design; open-source R code enables real-world use

References

Carter, R. E. (2004). Application of stochastic processes to participant recruitment in clinical trials. *Controlled Clinical Trials*, 25(5):429–436.

Held, L. and Bové, D. S. (2014). *Applied Statistical Inference*. Springer.

Piantadosi, S. (2024). *Clinical Trials: A Methodologic Perspective*. John Wiley & Sons.

Thank you for your attention