Accounting for the recruitment process into Bayesian modeling of vaccine data

Mauro Gasparini, **Marco Ratta**
Politecnico di Torino

Presentation Outline

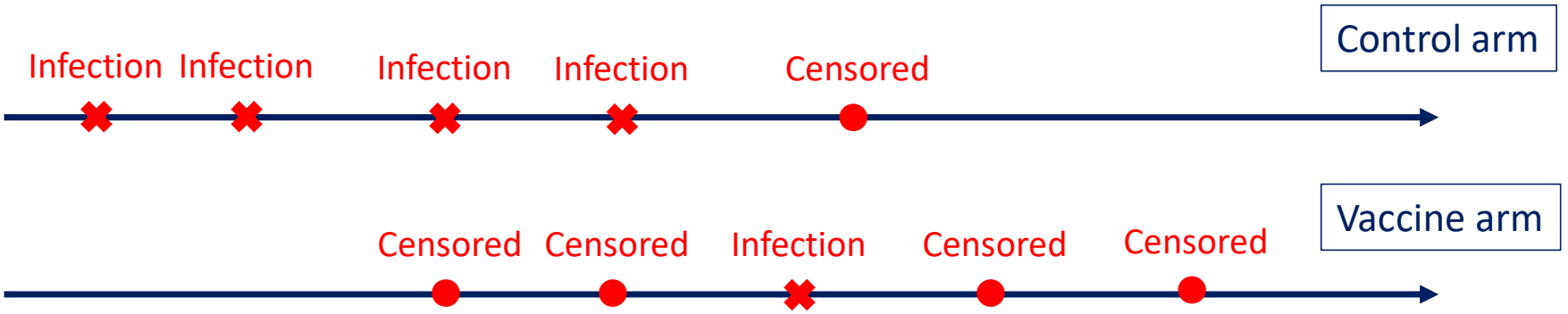
- Introduction to Vaccine Efficacy (VE)

- Bayesian Modeling of Vaccine Efficacy:
  - Exact method conditional on the total number of cases
  - Full Likelihood Bayesian model for Vaccine Efficacy

- Simulation Study

- Case study

Introduction to Vaccine Efficacy



$$x_c \sim \text{Poisson}(\lambda_c s_c)$$

$$x_v \sim \text{Poisson}(\lambda_v s_v)$$

$$VE = 1 - \frac{\lambda_v}{\lambda_c}$$

x_c = number of infections in control arm
x_v = number of infections in vaccine arm
s_c = Total surveillance time in control arm
s_v = Total surveillance time in vaccine arm
λ_c = Incidence rate in control arm
λ_v = Incidence rate in vaccine arm

Bayesian modeling of Vaccine Efficacy

The likelihood of the unknown parameters λ_c and λ_v can be expressed as:

$$f(s_v, s_c, x_v + x_c, x_v \mid \lambda_v, \lambda_c) = f_{S_v, S_c}(s_v, s_c \mid \lambda_v, \lambda_c) \times f_{x_v + x_c \mid s_v, s_c}(x_v + x_c \mid s_v, s_c, \lambda_v, \lambda_c) \times f_{x_v \mid x_v + x_c, s_v, s_c}(x_v \mid x_v + x_c, s_v, s_c, \lambda_v, \lambda_c)$$

Which is composed by three contributions:

- Marginal density of the Surveillance times
- Conditional density of the total infections given the surveillance times
- Conditional density of the vaccine infections given the total infections and the surveillance times

Exact Method Conditional on the total number of cases

In the current practice the surveillance times and the total number of infections are considered **data** rather than statistics.
This means that the first two factors of the full likelihood are considered independent from the vaccine efficacy and the likelihood is reduced to:

$$f(s_v, s_c, x_v + x_c, x_v \mid \lambda_v, \lambda_c) = f_{X_v \mid X_v + X_c, s_v, s_c}(x_v \mid x_v + x_c, s_v, s_c, \lambda_v, \lambda_c)$$

Which can be proven to be the density of a Binomial distribution

$$X_v \mid x_v + x_c, s_v, s_c \sim \text{Binomial}\left(x_v + x_c, \frac{s_v \lambda_v}{s_v \lambda_v + s_c \lambda_c}\right)$$

Motivation for the Full Likelihood

Imagine a study of infinite length (no censoring possible), then the ratio between the surveillance times approximates 1 - VE, in fact

$$S_c^i \sim \text{Exp}(\lambda_c)$$
$$S_v^i \sim \text{Exp}(\lambda_v)$$

$$\longrightarrow$$

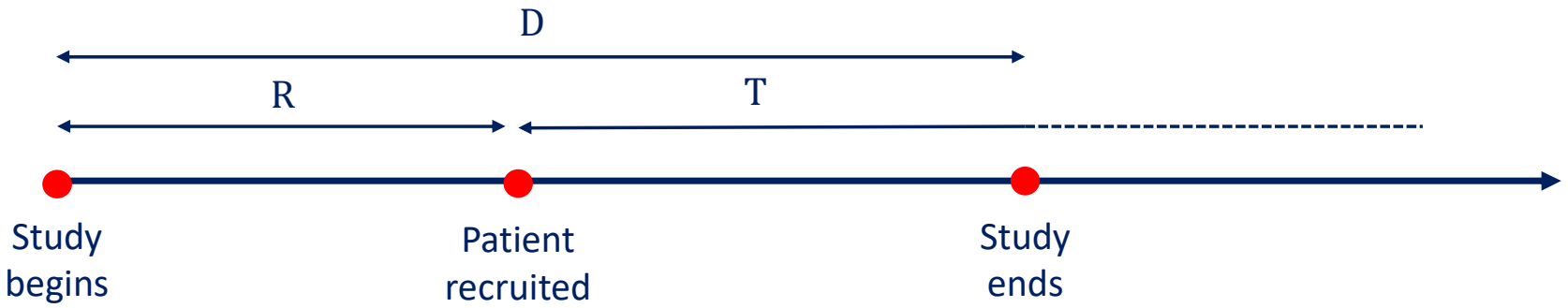
$$S_c \sim \text{Normal}\left(\frac{n_c}{\lambda_c}, \frac{n_c}{\lambda_c^2}\right)$$
$$S_v \sim \text{Normal}\left(\frac{n_v}{\lambda_v}, \frac{n_v}{\lambda_v^2}\right)$$

Which demonstrates that the surveillance times are statistics which depend on VE, hence should be included in the likelihood, but ….

In practice the study duration is limited, so the above does not hold and censoring process must be taken into account !!

Accounting for the recruitment process

For a generic patient let R be the random recruitment time, T the random time to infection, D the study duration and $C = D - R$ the random censoring time.



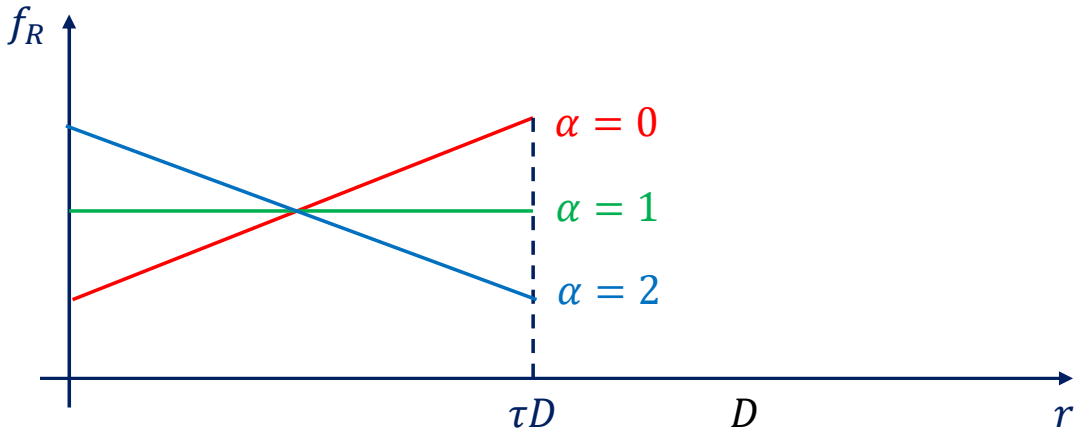
Then for the central limit theorem (TCL), then the total surveillance times can be expressed as

$$S_c \mid \lambda_c \sim \text{Normal}\left(n_c \mathbb{E}[\min(T_c, C)], n_c \text{Var}[\min(T_c, C)]\right)$$
$$S_v \mid \lambda_v \sim \text{Normal}\left(n_v \mathbb{E}[\min(T_v, C)], n_v \text{Var}[\min(T_v, C)]\right)$$

A class of recruitment densities

In order to exploit the previous result we need to make some assumptions on the recruitment process. In our context we use a parametric linear density recruitment:

$$f_R(r;\alpha,\tau) = \frac{2(1-\alpha)}{(\tau D)^2} r + \frac{\alpha}{\tau D} \qquad (0 \le r \le \tau D)$$



Where $\boldsymbol{\tau}$ represents the **truncation parameter** (the fraction of the study used for recruitment) and $\boldsymbol{\alpha}$ is the related to the **rate of accrual intensity.**

Simulation study

We conduct a 3 simulation studies testing our approach versus the standard one:

- **SIMULATION 1:** Fixing $\tau = 0.7$ and making VE vary in the set $(0.3, 0.5, 0.7, 0.9)$ and α in the set $(0.2, 0.6, 1, 1.4, 1.8)$.

- **SIMULATION 2:** Fixing $\alpha = 1$ and making VE vary in the set $(0.3, 0.5, 0.7, 0.9)$ and τ in the set $(0.1, 0.3, 0.5, 0.7, 0.9)$

- **SIMULATION 3:** Fixing $\mathrm{VE} = (0.3, 0.9)$ and making vary α in the set $(0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8)$ and τ in the set $(0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$.

Simulation study

Percentage reduction in the 95% credibility interval (CI) length

VE=0.3

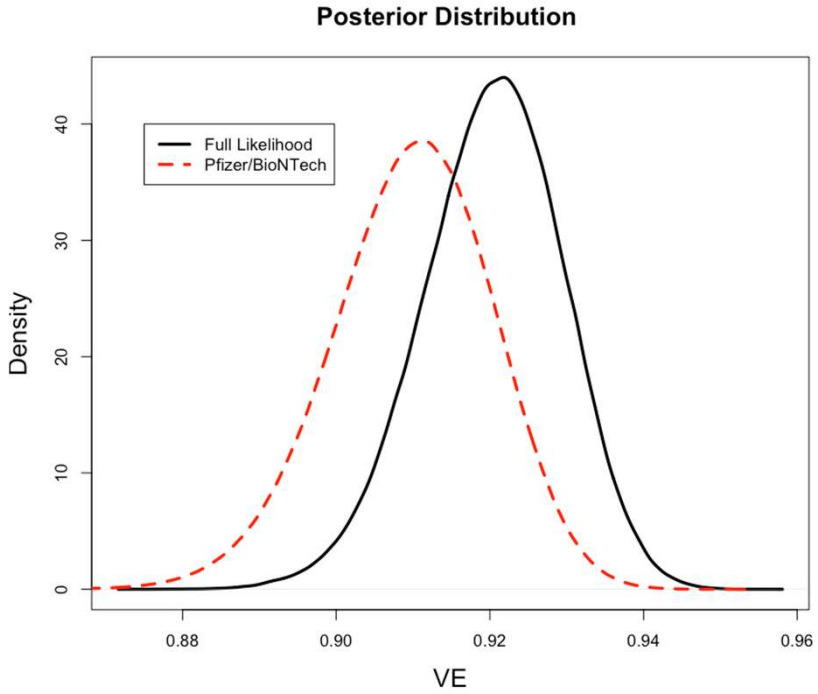
τ \ α	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8
0.9	2.51	2.67	2.9	3.14	3.47	3.86	4.34	5	5.83
	3.7	3.84	4.06	4.33	4.72	5.17	5.72	6.47	7.43
0.7	5.51	5.6	5.78	6.05	6.45	6.94	7.58	8.38	9.44
	8.2	8.13	8.23	8.49	8.87	9.35	10.01	10.88	11.92
0.5	11.7	11.49	11.48	11.66	11.96	12.43	13.03	13.83	14.82
	15.76	15.45	15.35	15.42	15.6	15.94	16.45	17.06	17.84
0.3	19.68	19.38	19.23	19.21	19.3	19.51	19.79	20.2	20.69
	22.62	22.44	22.34	22.31	22.34	22.4	22.52	22.71	22.94
0.1	24.25	24.2	24.16	24.14	24.16	24.16	24.19	24.24	24.29

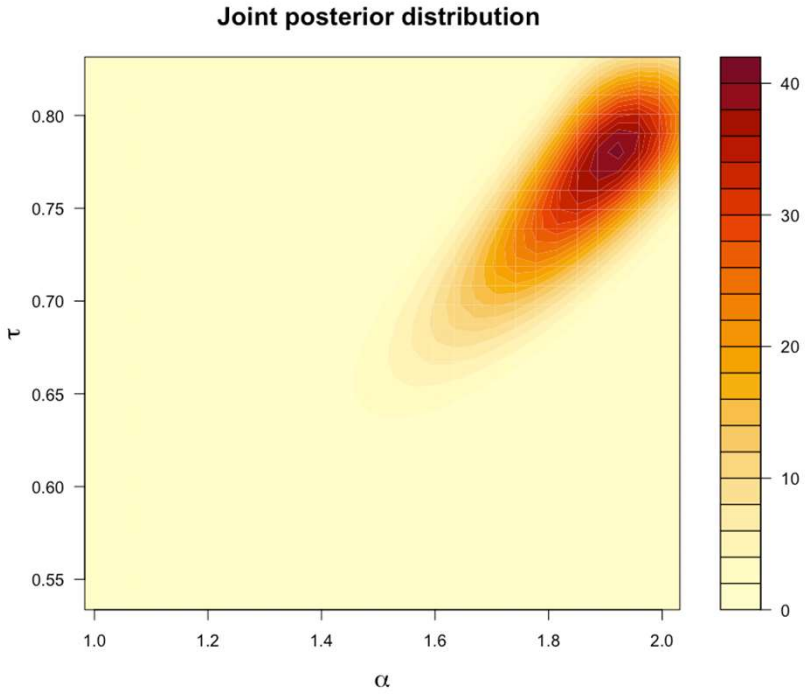
VE=0.9

τ \ α	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8
0.9	0.53	0.59	0.65	0.69	0.79	0.9	1	1.18	1.41
	0.85	0.86	0.94	1.01	1.1	1.21	1.38	1.59	1.88
0.7	1.31	1.33	1.4	1.46	1.6	1.73	1.93	2.19	2.55
	2.1	2.08	2.13	2.2	2.34	2.52	2.75	3.05	3.49
0.5	3.37	3.3	3.29	3.34	3.48	3.69	3.98	4.36	4.9
	5.48	5.26	5.2	5.24	5.36	5.61	5.95	6.39	7.07
0.3	9.05	8.65	8.49	8.43	8.57	8.8	9.2	9.76	10.53
	14.75	14.24	13.96	13.87	13.92	14.11	14.44	14.98	15.71
0.1	21.47	21.19	21.02	20.93	20.9	20.99	21.13	21.36	21.68

- The gain is higher for **larger values of α**
- The gain is higher for **lower values of τ**
- The gain is higher for **lower values of VE**

Case study: Pfizer/BioNTech trial





- **More precise point estimation**
  92.03 [90.1, 93.7] VS 90.98 [88.8, 92.9]
  12.7% gain in 95% CI length

- **Good identifiability in estimation of α and τ**
  92.03 [90.1, 93.7] VS 90.98 [88.8, 92.9]
  12.7% gain in 95% CI length

Discussion

Using the Full Likelihood in the estimation of Vaccine Efficacy (VE) is an improvement over the currently used exact method conditional on the total number of cases, in fact:

- It improves the point estimation both in terms of median MSE and mean MSE
- It provides shorter credibility intervals (CI)

In particular the gain is higher when:

- Recruitment process is fast (low τ, large α)
- Vaccine Efficacy (VE) is low

References

- M. Ewell. Comparing methods for calculating confidence intervals for vaccine efficacy. *Statistics in Medicine*, 15:2379–2392, 1996. 2
- F.P. Polack et al. Safety and efficacy of the bnt162b2 mrna covid-19 vaccine. *The New England Journal of Medicine*, 383:2603–2615, 2020. 1, 2, 12, 13
- S. Senn. The design and analysis of vaccine trials for covid-19 for the purpose of estimating efficacy. *Pharmaceutical Statistics*, 21(4):790–807, 2022. 2
- S.J. Thomas et al. Safety and efficacy of the bnt162b2 mrna covid-19 vaccine through 6 months. *The New England Journal of Medicine*, 385:1761–1773, 2021. 1, 12, 13, 14

Backup slides

Simulation 1

	Percentage in 95% CI length reduction				
	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
VE=0.3	24.16	19.32	11.97	6.48	3.49
VE=0.5	23.95	18.02	10.44	5.43	2.90
VE=0.7	23.39	15.36	7.92	3.90	2.06
VE=0.9	20.94	8.59	3.50	1.60	0.80
	% means MSE reduction				
	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
VE=0.3	44.86	37.79	24.92	13.63	7.59
VE=0.5	44.34	35.21	21.70	11.72	7.32
VE=0.7	44.28	31.43	17.07	8.65	5.54
VE=0.9	38.67	17.85	7.24	3.94	2.23
	% medians MSE reduction				
	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
VE=0.3	44.81	37.73	24.89	13.61	7.57
VE=0.5	44.28	35.16	21.64	11.71	7.33
VE=0.7	44.22	31.37	17.03	8.66	5.59
VE=0.9	38.59	17.78	7.21	4.03	2.39

Simulation 2

	Percentage in 95% CI length reduction				
	$\alpha = 0.2$	$\alpha = 0.6$	$\alpha = 0.1$	$\alpha = 1.4$	$\alpha = 1.8$
VE=0.3	5.55	5.81	6.50	7.62	9.46
VE=0.5	4.65	4.88	5.46	6.44	8.11
VE=0.7	3.37	3.51	3.94	4.71	6.01
VE=0.9	1.37	1.43	1.64	1.98	2.57
	% means MSE reduction				
	$\alpha = 0.2$	$\alpha = 0.6$	$\alpha = 0.1$	$\alpha = 1.4$	$\alpha = 1.8$
VE=0.3	11.85	12.51	13.63	16.02	19.86
VE=0.5	10.65	11.33	11.72	14.15	17.17
VE=0.7	8.19	7.77	8.63	10.68	13.27
VE=0.9	3.20	3.86	3.89	4.04	4.83
	% medians MSE reduction				
	$\alpha = 0.2$	$\alpha = 0.6$	$\alpha = 0.1$	$\alpha = 1.4$	$\alpha = 1.8$
VE=0.3	11.84	12.48	13.64	16.00	19.84
VE=0.5	10.61	11.31	11.70	14.15	17.14
VE=0.7	8.20	7.78	8.64	10.65	13.24
VE=0.9	3.30	3.93	3.95	4.06	4.85