université de BORDEAUX

Inria

Introduction to modeling, simulation and data science in oncology

S. Benzekry

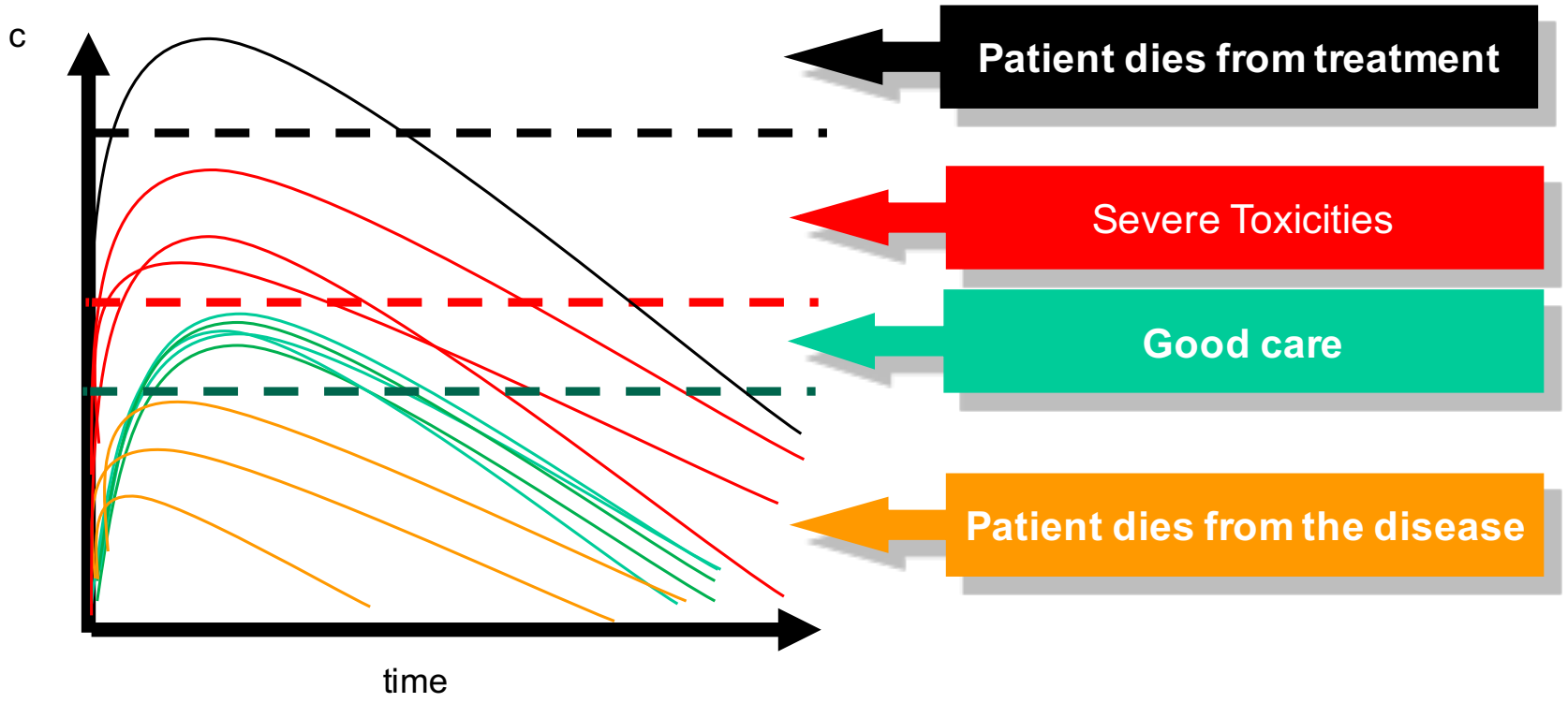
Formation « Ecole doctorale Mathématiques et Informatique »

MONC

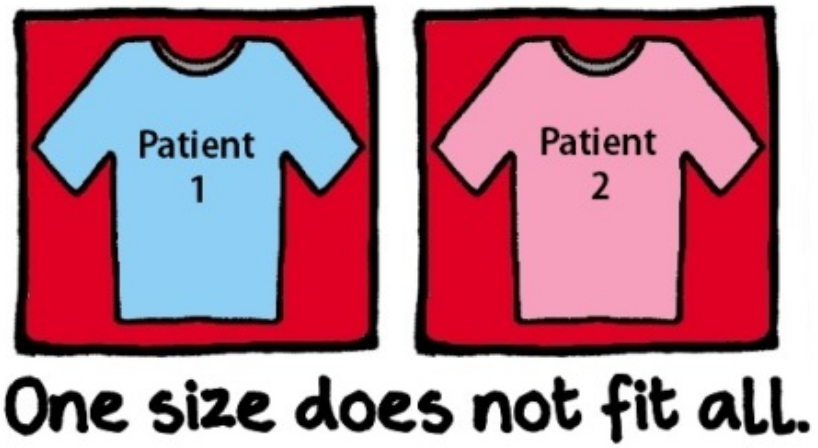
Modeling in ONCology

2. Population modeling

Define the dose and scheduling of the drug administration



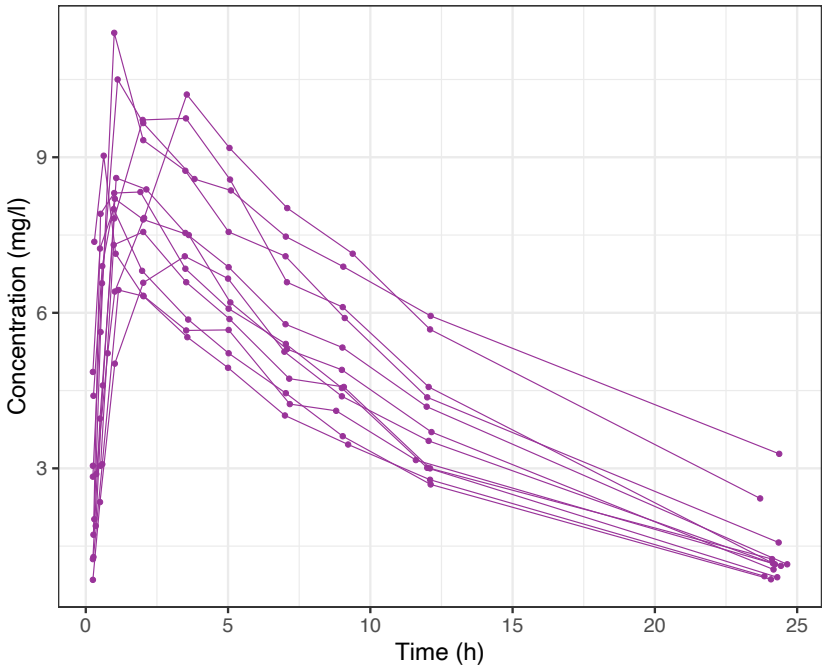
Personalize the treatment

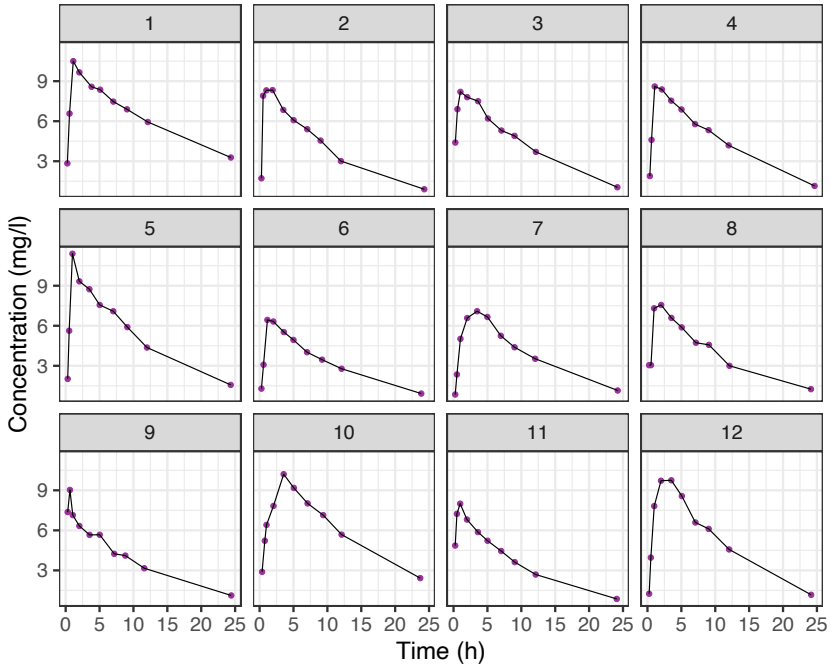


⇒ **adaptive** dosing strategies

Population data: inter-individual variability

Theophylline concentration





y_j^i

individual i

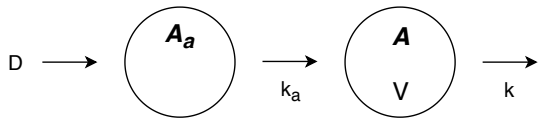
time t_j

Part of these slides are adapted from Lavielle's course Statistics in Action

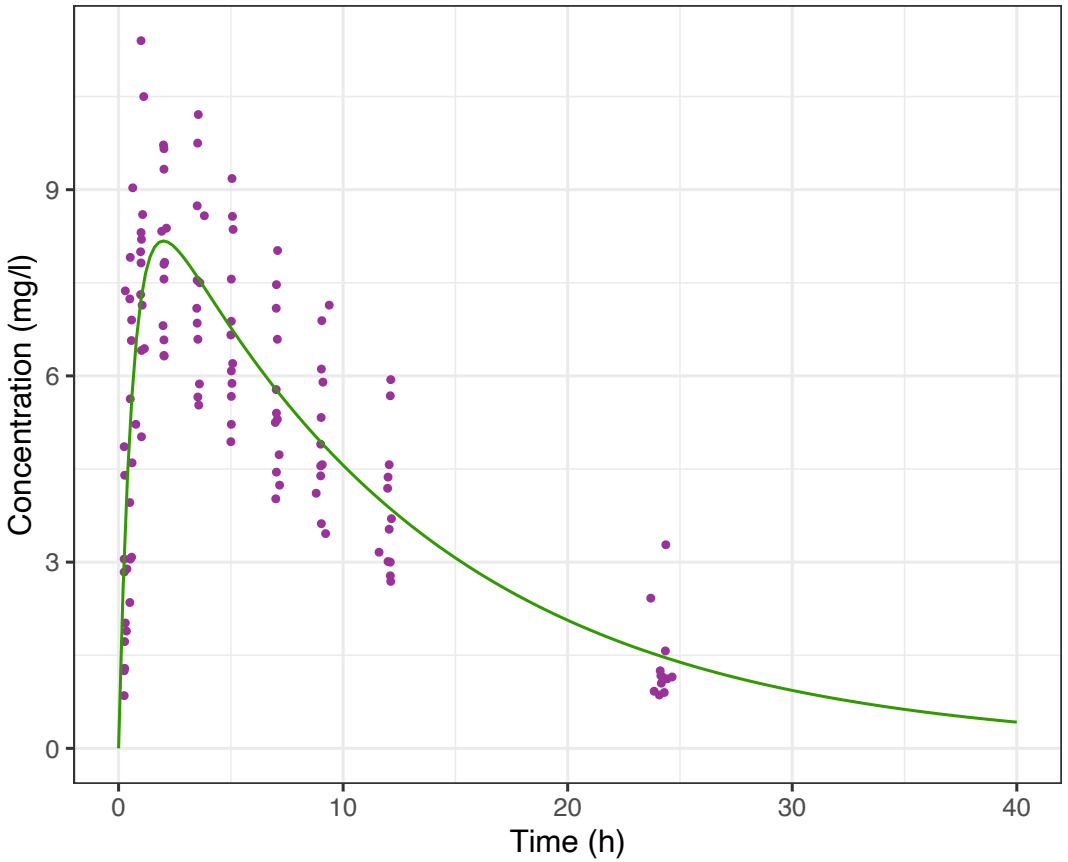
Fit a single pharmacokinetics model

$$y_j^i = M(t_j^i; \theta) + \varepsilon_j^i$$

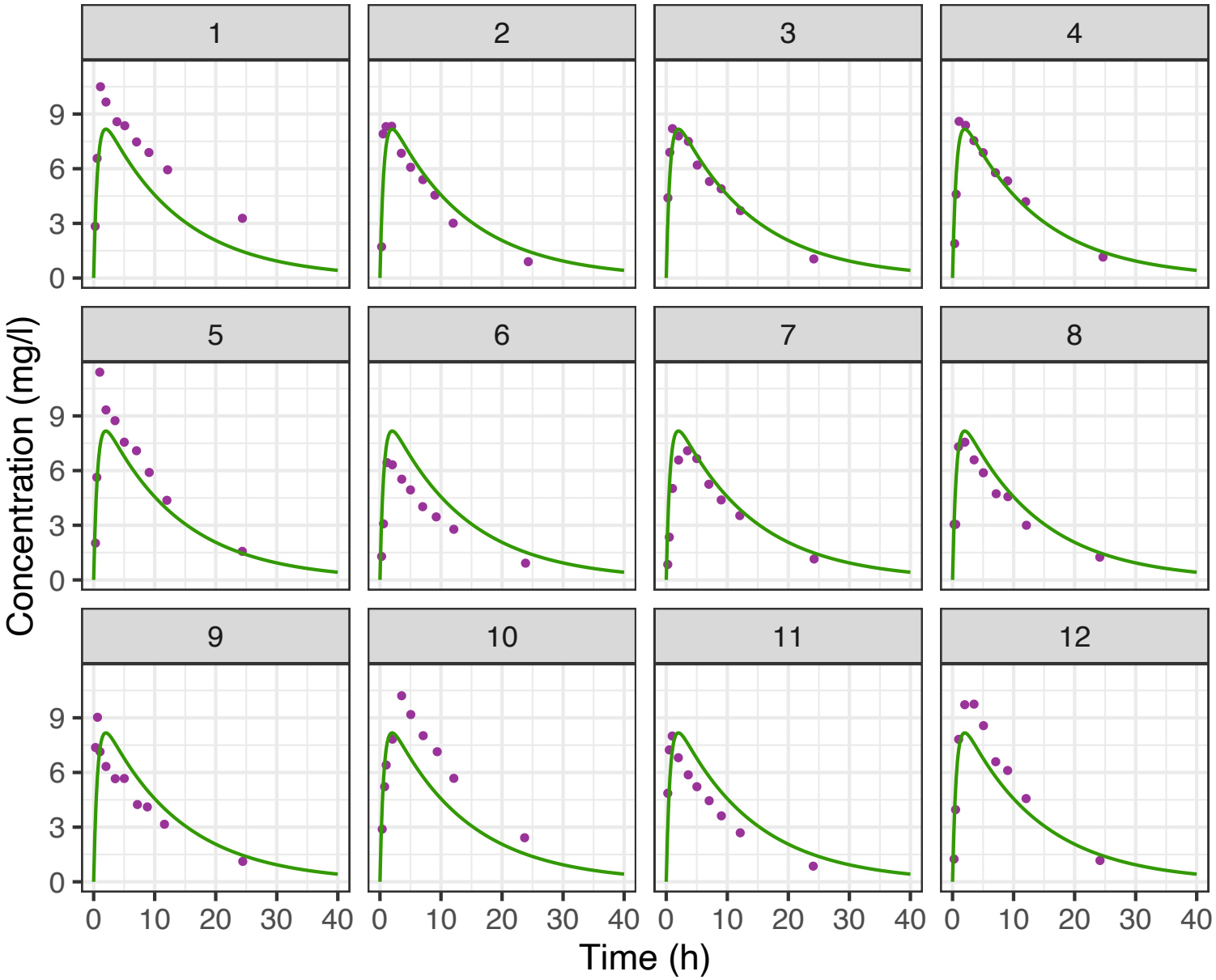
$$\hat{\theta} = \underset{\theta}{\text{argmin}} \sum_{i=1}^{N} \sum_{j=1}^{n_j} \left(y_j^i - M\left(t_j^i; \theta\right) \right)^2$$



$$\begin{cases} \frac{dA_a}{dt} = -k_a A_a \\ \frac{dA}{dt} = k_a A_a - kA \\ A_a(t=0) = D, \quad (t=0) = 0 \end{cases} \qquad C(t) = \frac{A(t)}{V}.$$



Individual comparison



Mixed-effects modeling

$$y_j^i = M\left(t_j^i; \psi^i\right) + \varepsilon_j^i, \quad \varepsilon_j^i \sim \mathcal{N}\left(0, \sigma^2\right)$$

Instead of fitting a set of parameters $\hat{\psi}^i$ for each individual (i.e. $p \times N$ parameters)….

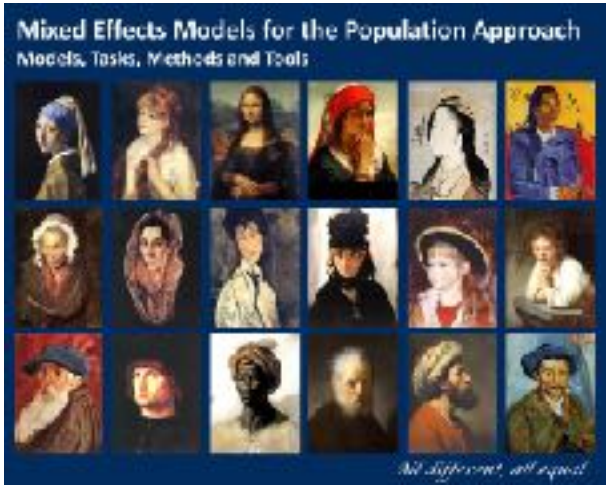
Assume a **statistical population distribution** on the individual parameters ψ^i

$$\psi^i \sim \mathcal{N}\left(\psi_{pop}, \Omega\right)$$

or

$$\psi^i = \psi_{pop} + \eta^i, \quad \eta^i \sim \mathcal{N}\left(0, \Omega\right)$$

fixed effects **ramdom** effects



Lavielle,CRC press, 2014

only $p + \dfrac{p(p+1)}{2}$ parameters to estimate (if full covariance matrix)

Transformation of the parameters

- To ensure **positivity** of the parameters: Lognormal distribution

$$\log\left(\psi^i\right) = \log\left(\psi_{pop}\right) + \eta^i, \quad \eta^i \sim \mathcal{N}\left(0, \Omega\right)$$

(in Monolix, ω = *std* of random effect)

- To ensure **bounds** on the parameters: Logit function $(0,1) \to \mathbb{R}$

$$\text{logit}(\psi^i) = \ln\left(\frac{\psi^i}{1-\psi^i}\right) = \text{logit}(\psi_{pop}) + \eta^i, \quad \eta^i \sim \mathcal{N}\left(0, \Omega\right)$$

- Also: Probit function $(0,1) \to \mathbb{R}$

$$\text{probit}(x) = z \Leftrightarrow \mathbb{P}\left(\mathcal{N}(0,1) \leq z\right) = x$$

$$\text{probit}(\psi^i) = \text{probit}(\psi_{pop}) + \eta^i, \quad \eta^i \sim \mathcal{N}\left(0, \Omega\right)$$

Likelihood, parameters estimation

The (population) parameters to be estimated are $\theta = \left(\psi_{pop}, \Omega, \sigma\right)$

Likelihood

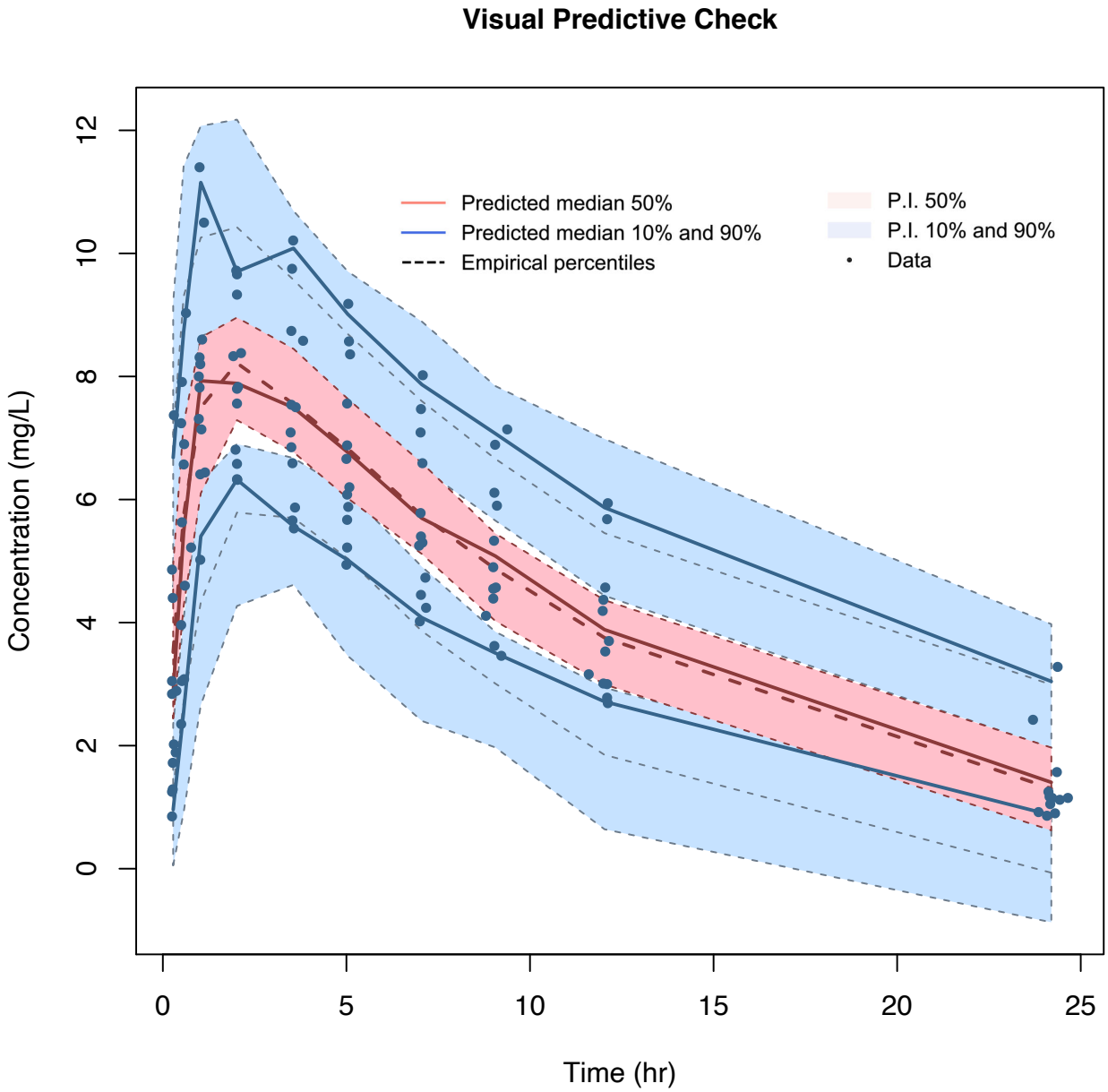
$$L(\theta) = p(y;\theta) = \int p\left(y,\psi;\theta\right) d\psi = \int p\left(y\,|\,\psi\right) p(\psi) d\psi = \prod_{i=1}^{N} \int p\left(y^i\,|\,\psi^i\right) p(\psi^i) d\psi^i$$

No analytical formula if model is nonlinear.

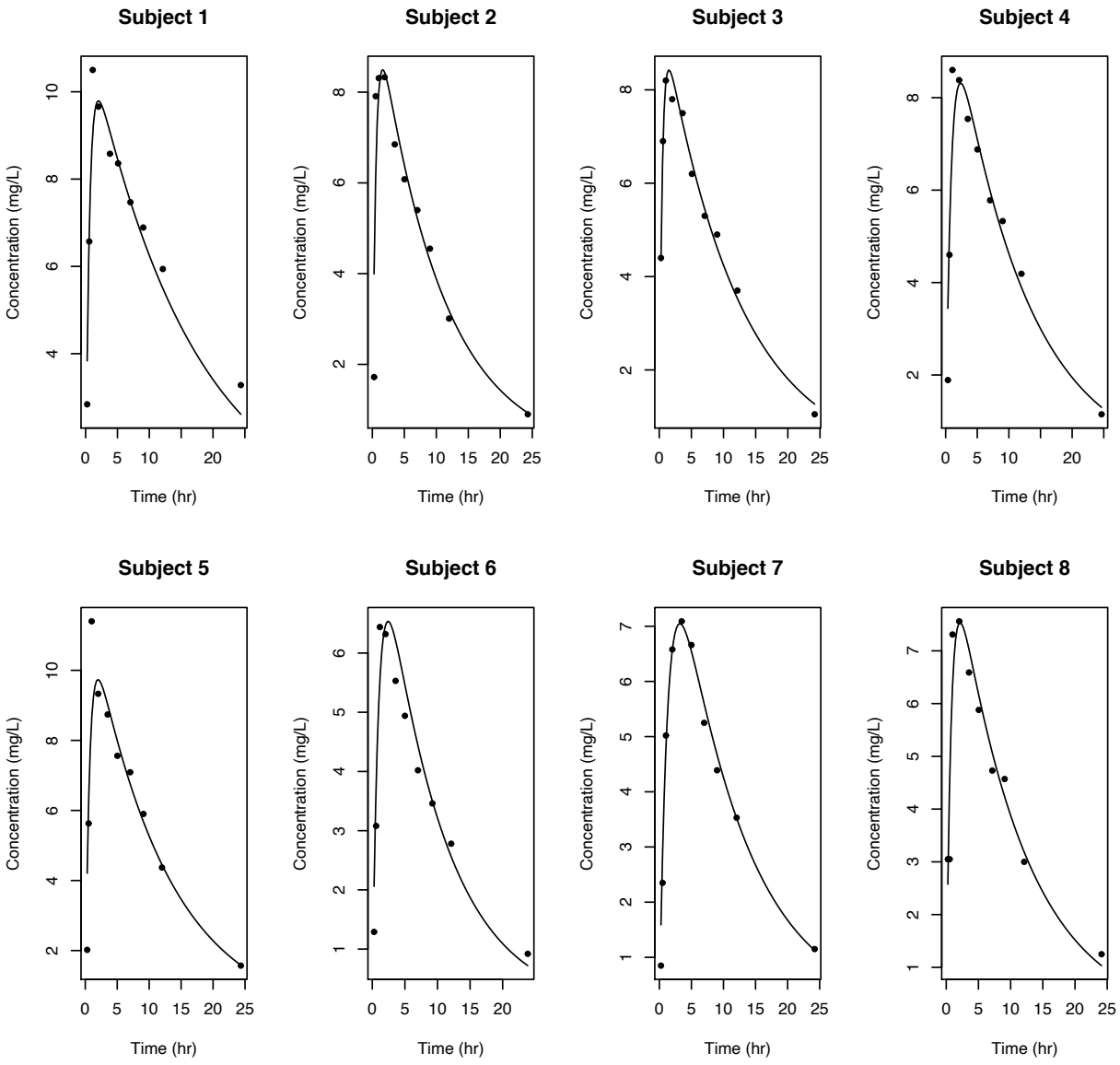
Specific algorithms : first-order (FO, 1970's), FO conditional estimation (FOCE, 1992), Expectation-Maximization (EM), **Stochastic Approximation of EM** (2000's)

Two main softwares: NONMEM and **Monolix**

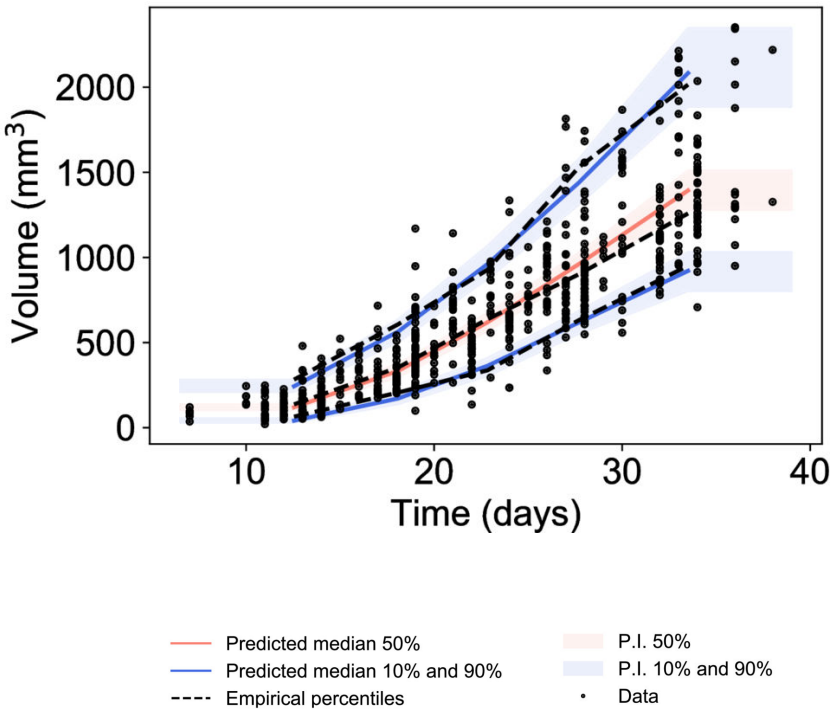
Application: Theophylline data

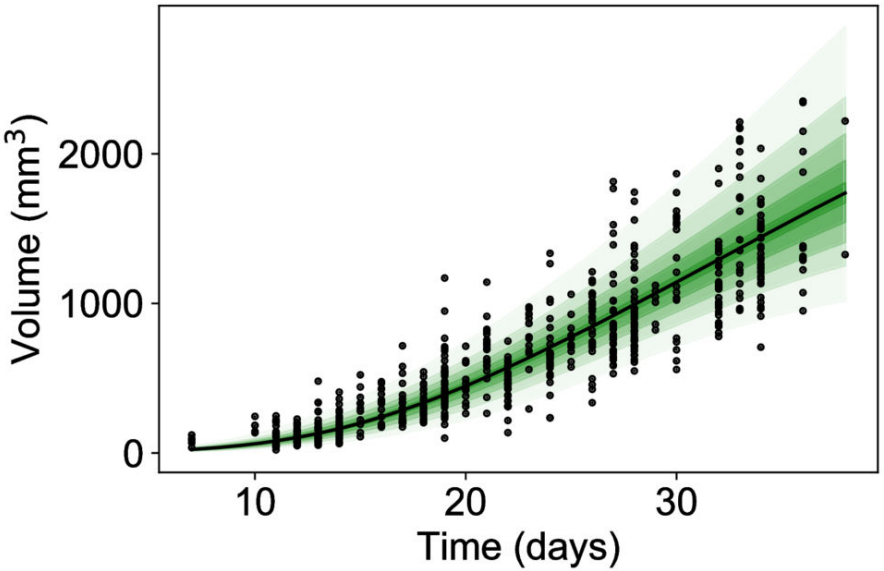


Application: Theophylline data



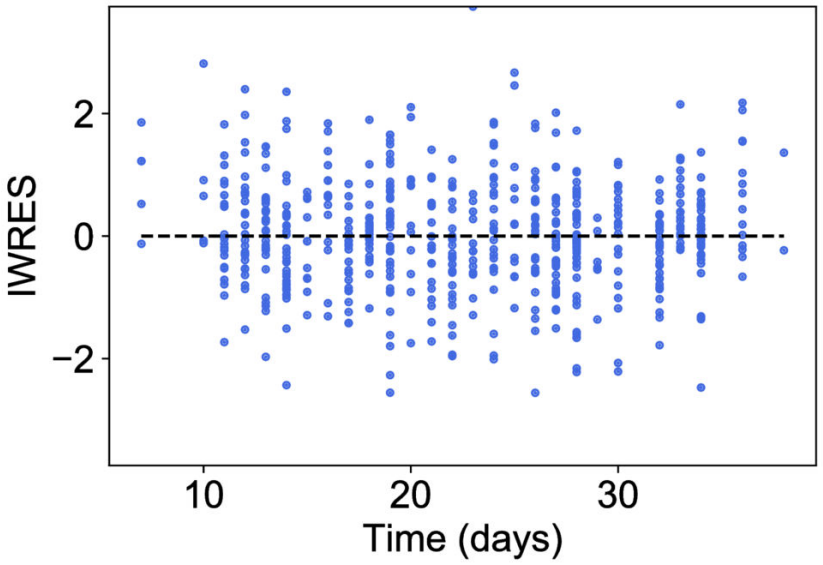
Application: tumor growth data

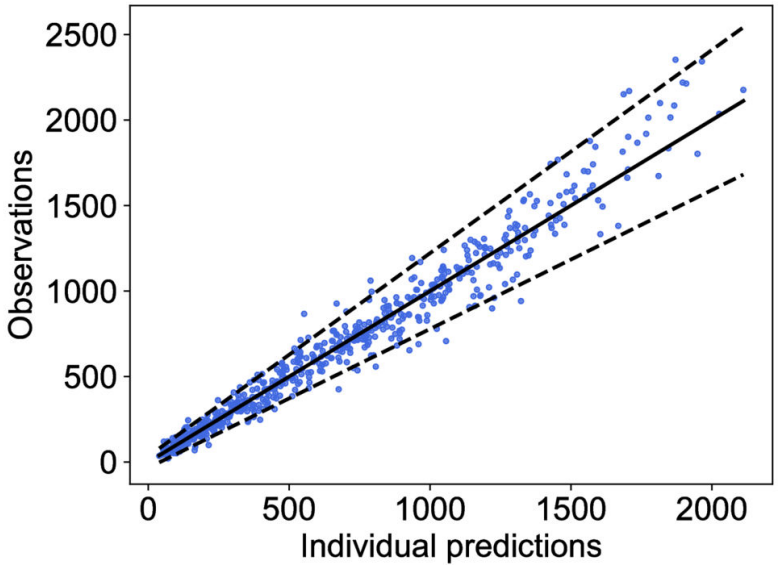


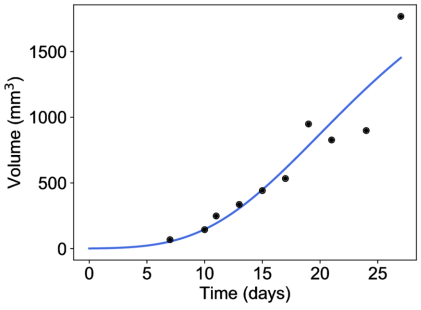


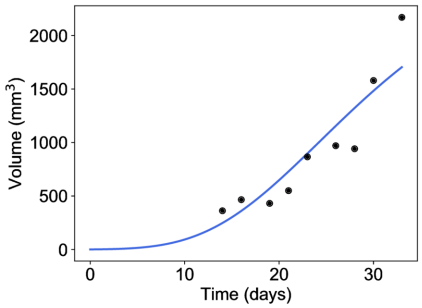
$$\frac{dV}{dt} = \alpha e^{-\beta t} V$$ Gompertz

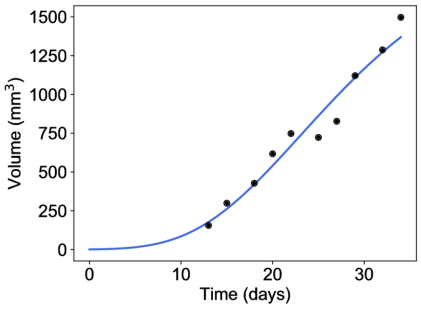
Residuals and individual predictions











Bayesian estimation of the individual parameters
Maximum A Posteriori

- **Bayes**' formula $$p\left(\psi^i \middle| y^i;\hat{\theta}\right) = \frac{p\left(y^i \middle| \psi^i;\hat{\theta}\right) p\left(\psi^i;\hat{\theta}\right)}{p\left(y^i;\hat{\theta}\right)}$$

- **Maximum a posteriori** estimator of ψ^i = **mode**

$$\hat{\psi}^i = \underset{\psi}{\text{argmax}} \left\{ p\left(\psi \middle| y^i\right) \right\} = \underset{\psi}{\text{argmax}} \left\{ \frac{p\left(y^i \middle| \psi\right) p\left(\psi\right)}{p\left(y^i\right)} \right\} = \underset{\psi}{\text{argmax}} \left\{ p\left(y^i \middle| \psi\right) p\left(\psi\right) \right\}$$

$$= \underset{\psi}{\text{argmin}} \left\{ -2\log\left(p\left(y^i \middle| \psi;\hat{\theta}\right)\right) - 2\log\left(p\left(\psi;\hat{\theta}\right)\right) \right\}$$

$$\varepsilon_j^i \sim \mathcal{N}(0,\sigma)$$

and

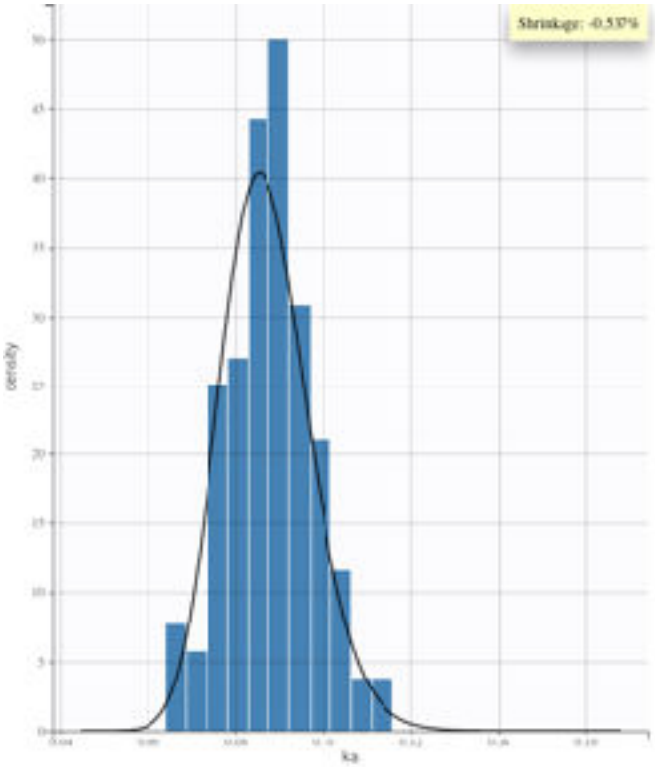
$$\psi^i \sim \mathcal{N}\left(\psi_{pop}, \omega^2 I_p\right)$$

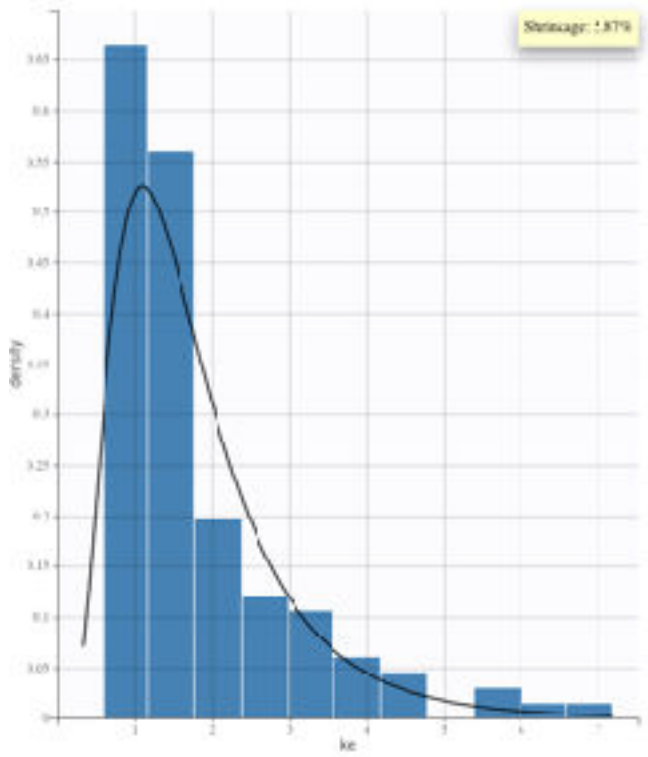
$$\rightarrow = \underset{\psi}{\text{argmin}} \left\{ \sum_{j=1}^{n^i} \frac{\left(y_j^i - M\left(t^i j;\hat{\theta}\right)\right)^2}{\sigma^2} + \sum_{p=1}^{P} \frac{\left(\psi_p - \hat{\psi}_{pop,p}\right)^2}{\omega^2} \right\}$$

Shrinkage

We assumed $\psi_p^i \sim \mathcal{N}\left(0, \omega_p^2\right)$ and computed $\hat{\psi}^i$ thus we should have

$$\eta\text{- shrinkage} = 1 - \frac{\text{Var}\left(\hat{\psi}_p^i\right)}{\omega_p^2} \simeq 0$$



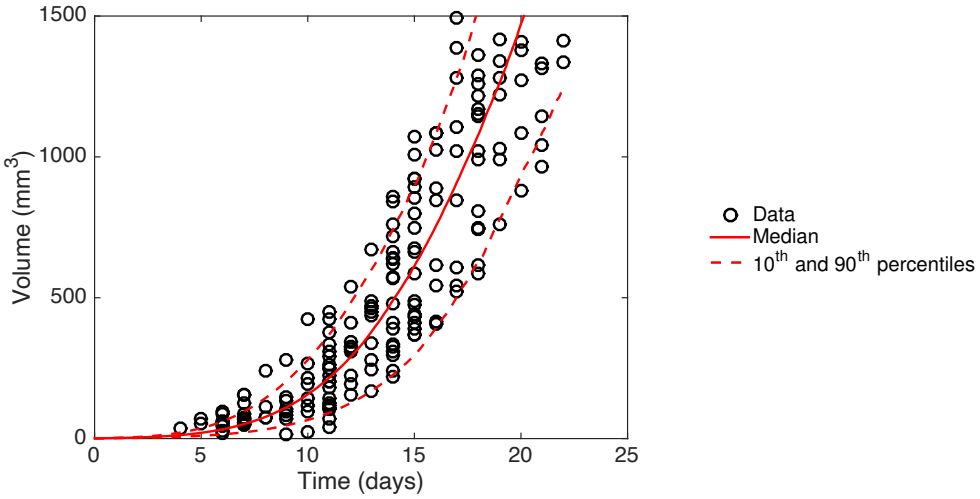


Population approach and its use for prediction

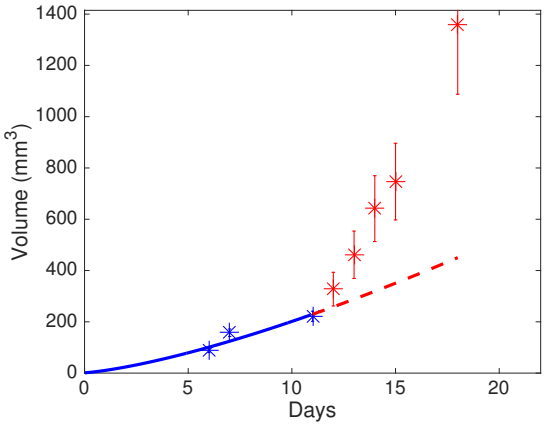
Nonlinear mixed-effects statistical modeling

$$y_j^i = M(t_j^i, \theta^i) + \varepsilon_j^i, \quad \varepsilon_j^i \sim \mathcal{N}(0, \sigma_j^i)$$

$$\theta^1, \cdots, \theta^N \sim \mathcal{LN}(\theta_{pop}, \theta_\omega), \quad \theta_{pop} \in \mathbb{R}^P, \theta_\omega \in \mathbb{R}^{P \times P}$$

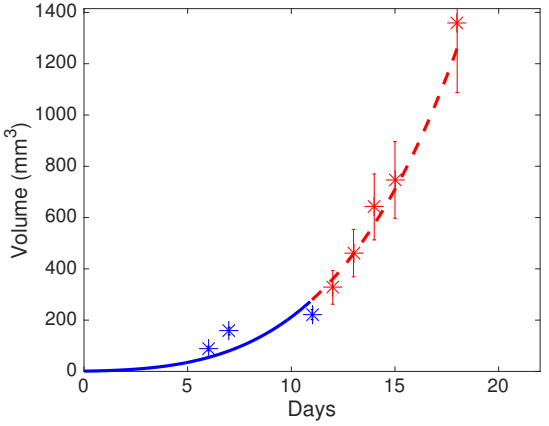


No a priori



With a priori (bayesian)

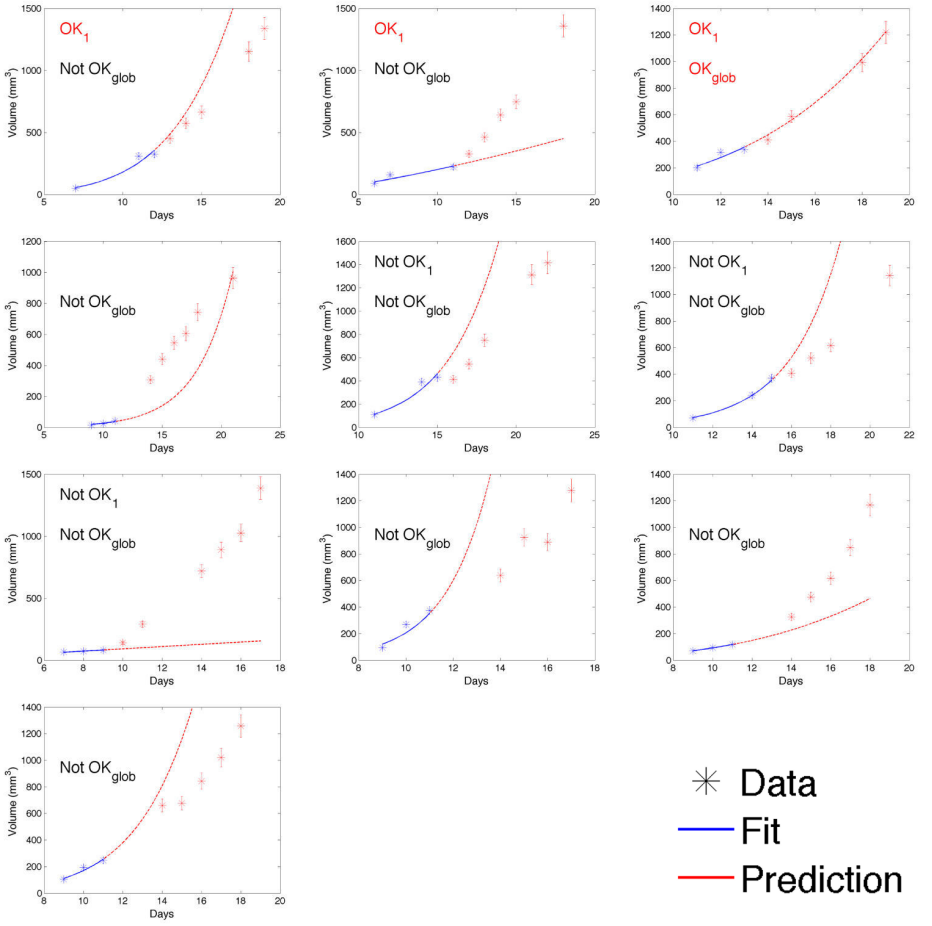
$$l(\theta^i) = \sum_{j=1}^{J^i} \frac{\left(y_j^i - M(t_j^i, \theta^i)\right)^2}{2\sigma^2} + \sum_{p=1}^{P} \frac{\left(\theta_p^i - \theta_{pop,p}\right)^2}{\theta_{\omega,p}^2} + C$$



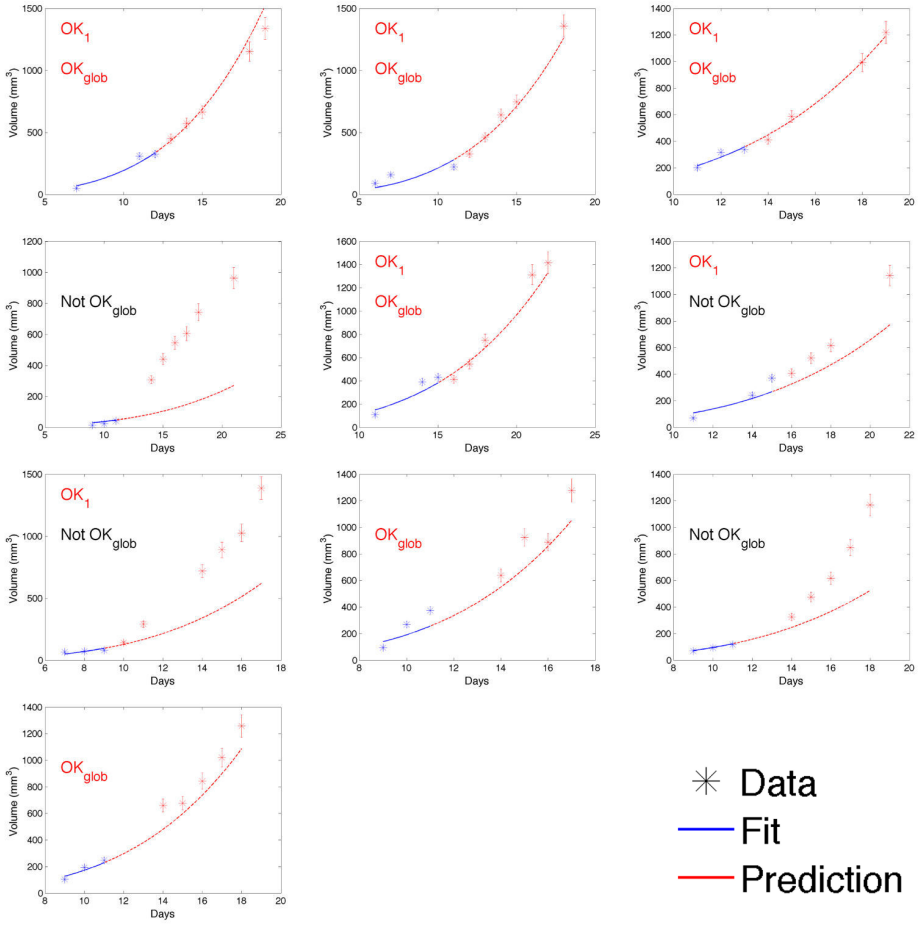
Benzekry et al., PLoS Comp Biol, 2014

Power law model: all animals

No a priori

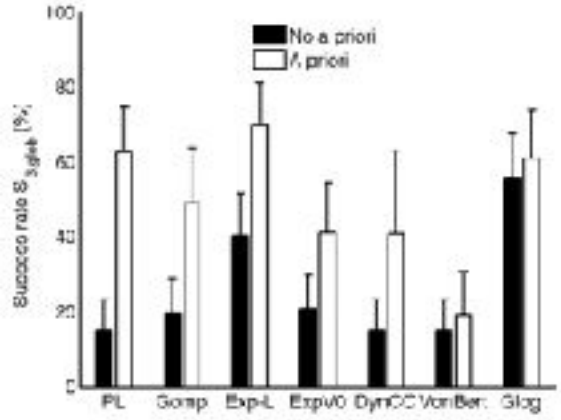


A priori

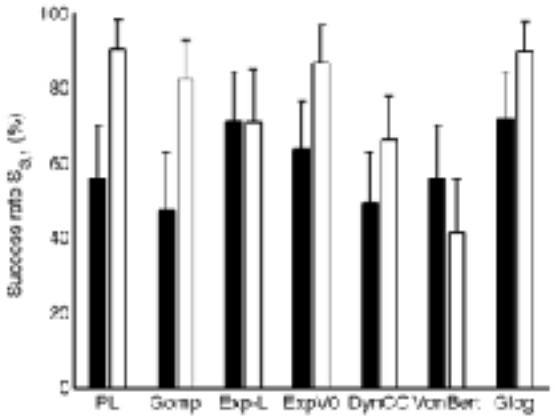


Prediction improvement for all models

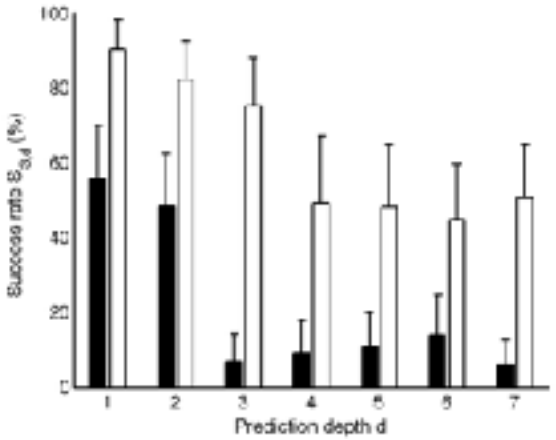
Full future curve



Next data point

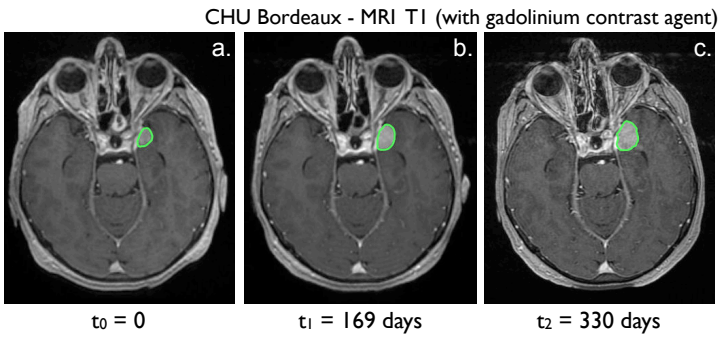


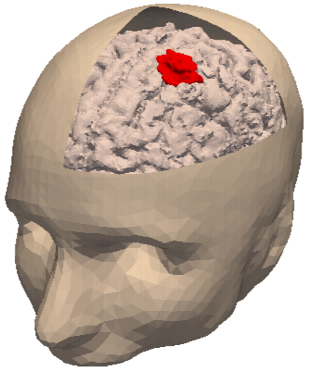
All prediction depths
Power law



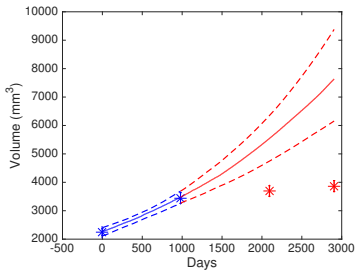
Randomly assign (100 replicates) half of the animals to the **« learning group »** and the other half to the **« predict » group**

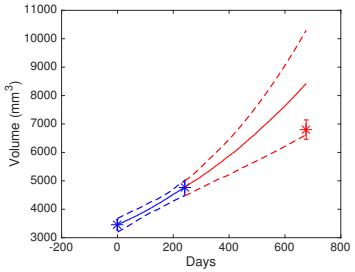
Limits of the bayesian approach: meningioma human data



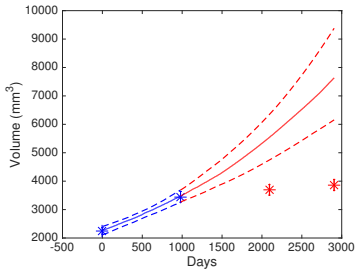


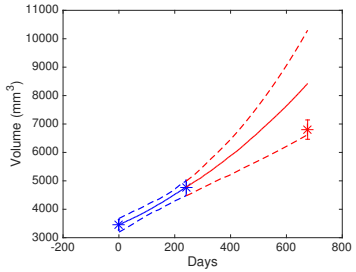
No a priori

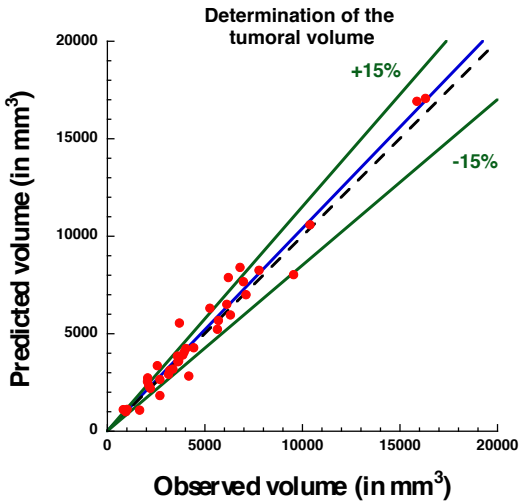




A priori



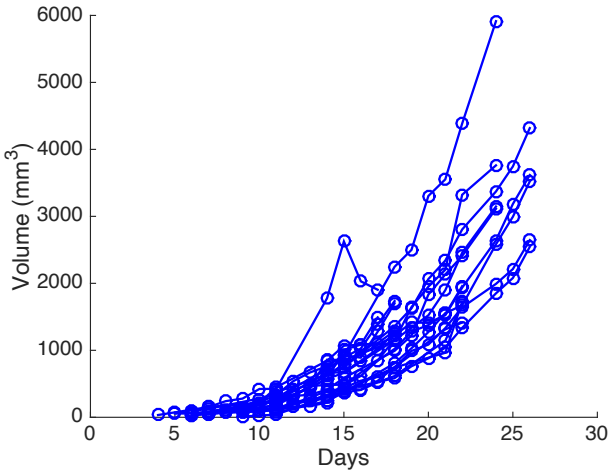




It depends on the degree of heterogeneity of the population

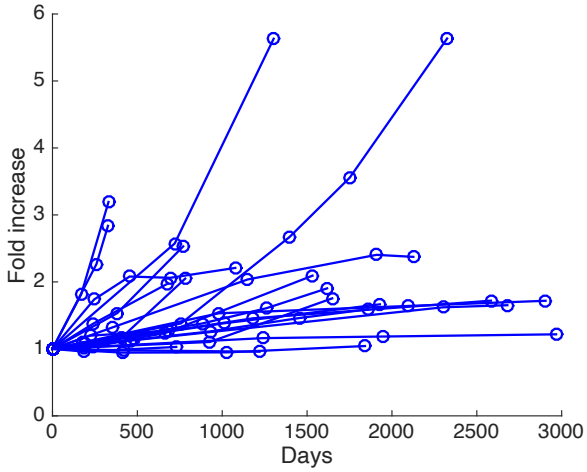
Mice data (LLC s.c.) = homogeneous

Model	Par.	Unit	Median value (CV)	NSE (%) (CV)
Power law	α	$\mathrm{mm}^{3(1-\gamma)}\cdot\mathrm{day}^{-1}$	0.886 (30.8)	8.17 (52.5)
	γ	-	0.788 (7.56)	2.28 (58.6)



Human data (meningioma) = heterogeneous

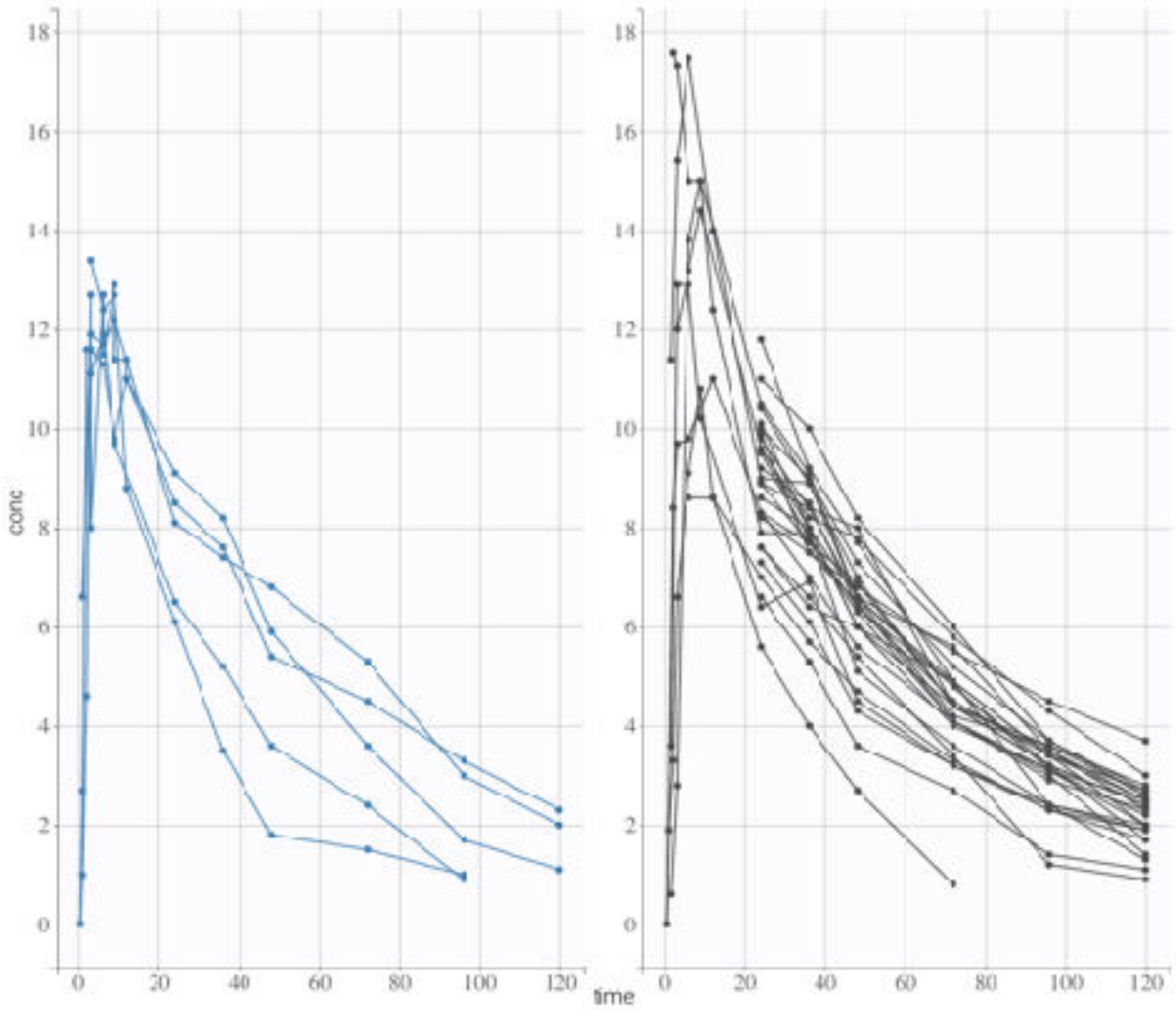
Model	Par.	Unit	Median value (CV)	NSE (%) (CV)
Gompertz V_0	α_0	-	0.00251 (2.5e+03)	38.5 (692)
	β	-	2.85e-05 (1e+04)	35.8 (1.17e+03)
	V_0	-	2.31e-03 (134)	3.09 (3.98e+03)



Covariates

Females

Males



Covariates

- Let $x^i = \left(x_1^i, \cdots, x_L^i\right)$ be a vector of individual covariates (age, sex, weight,…)

- We can try to explain the variability by the covariates

$$\psi^i = \psi_{pop} + \sum_{l=1}^{L} \beta_l x_l^i + \eta^i$$

fixed effects

random effects

Covariates

No covariate

Population parameter estimates

	VALUE	STOCH. APPROX. S.E.	STOCH. APPROX. R.S.E.(%)
Fixed Effects			
ka_pop	0.0121	0.000657	5.47
ke_pop	0.475	0.0353	7.42
V_pop	1.36	0.0947	6.97
Standard Deviation of the Random Effects			
omega_ka	0.168	0.0267	15.1
omega_ke	0.166	0.0218	13.1
omega_V	0.151	0.0485	32.2
Error Model Parameters			
a	0.563	0.124	22
b	0.118	0.0245	20.8

Covariate

Population parameter estimates

	VALUE	STOCH. APPROX. S.E.	STOCH. APPROX. R.S.E.(%)
Fixed Effects			
ka_pop	0.0216	0.00172	7.96
beta_ka_sex_1	-0.295	0.091	30.9
ke_pop	0.454	0.029	6.4
V_pop	1.49	0.228	15.2
beta_V_wt	-0.00813	0.00304	37.4
Standard Deviation of the Random Effects			
omega_ka	0.141	0.0355	25.2
omega_ke	0.167	0.0397	23.7
omega_V	0.123	0.0451	36.7
Error Model Parameters			
a	0.526	0.122	23.1
b	0.125	0.0242	19.3

$$k_a = k_{a,pop} + \beta_{ka,sex}\left[sex = 1\right]^i + \eta^i$$

$$V^i = V_{pop} + \beta_{V,wt} wt^i + \eta_V^i$$

Wald test (stochastic approximation):

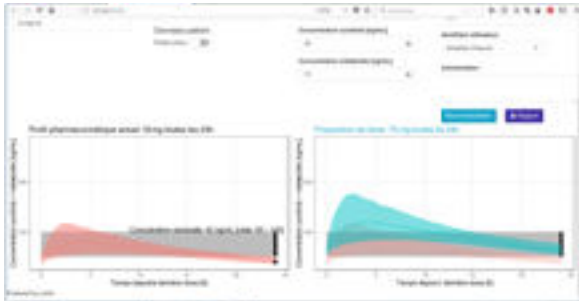
ka

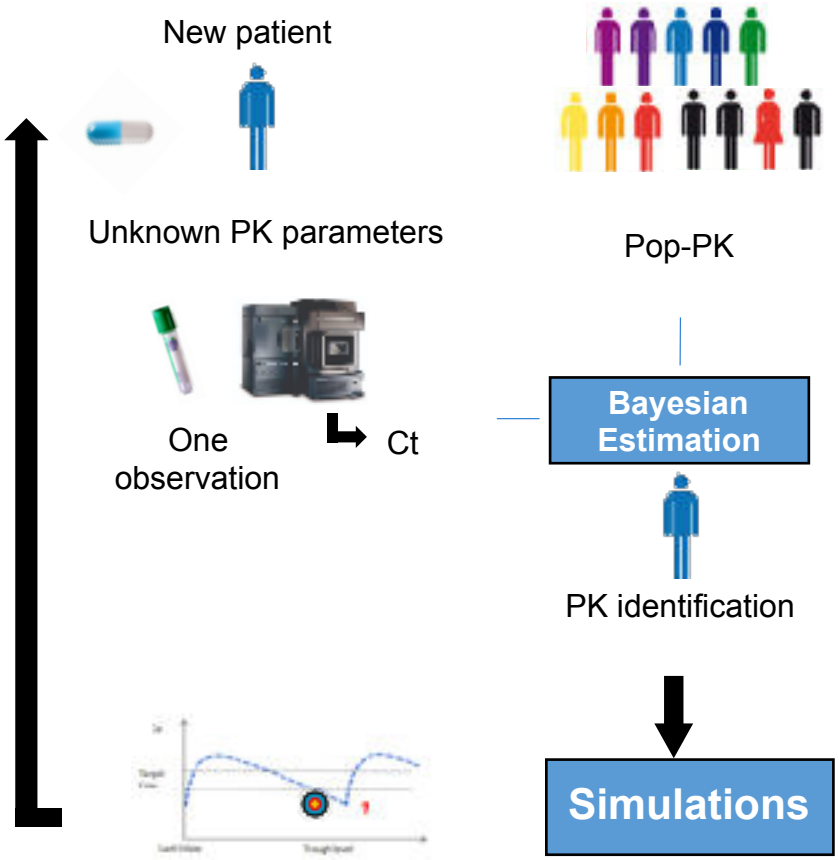
	STATISTICS	P-VALUE
beta_ka_sex_1	9.23686	0.00219588

	STATISTICS	P-VALUE
beta_V_wt	3.62474	0.0571955

Precision medicine and biotargeting of TKIs

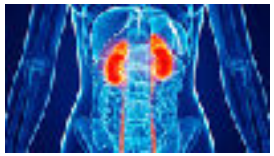
Therapeutic drug monitoring of Tyrosine Kinase Inhibitors (imatinib, sunitinib, dasatinib, cabozantinib, sorafenib, ibrutinib…).





Sunitinib in metastatic kidney cancer

Patient #	Starting Dose (mg)	Total Su + met (ng/ml)	Sampling Time	Simulated Trough Level (ng/ml)	Proposed Dose (mg)	% change
1	50	195	5H30	161	25	-50
2	50	55	23H00	56	62,5	25
3	50	37,4	24H15	40	87,5	75
4	50	40	23h45	42	75	50
5	50	166	22H20	158	25	-50
6	50	161	4H45	136	25	-50
7	50	70	24H00	73	50	no change
8	50	161	4h45	136	25	-50
9	50	17,1	24H00	18	100	**100**
10	50	170	12H30	149	25	-50
11	50	90	24H00	90	37,5	-25
12	50	44,3	24H00	47	75	50
13	50	88	2H15	76	50	no change
14	50	106	19H00	100	37,5	-25
15	50	54,2	6H00	42	87,5	75
16	50	141	1H30	81	37,5	-25
17	50	128	24H00	106	37,5	-25
18	50	118,9	1H00	81	50	no change
19	50	145	19H00	115	37,5	-25
20	50	87	9H30	72	50	no change
21	50	104	3H20	90	37,5	-25
22	50	125	24h00	112	37,5	-25
23	50	62	19H00	58	62,5	25
24	50	246	24H00	231	12,5	**-75**
25	50	150	24H00	143	25	-50
26	50	83	12h00	71	50	no change
27	50	216	24h00	204	12,5	-75
28	50	197	24h00	192	25	-50
29	50	116	8H30	97	37,5	-25
30	50	78	24H00	71	50	no change



Standard dose: 50 mg

80% of APHM patient have dose modification of Sutent®
12.5 <>100 mg
(-75% ⇨ + 100%!)

References

- Course « Statistics in Action with R » by Marc Lavielle

http://sia.webpopix.org/index.html

- Lavielle, M. (2014). Mixed Effects Models for the Population Approach. CRC Press.

- Case studies and documentation on practical use of mixed-effects modeling with Monolix:

http://lixoft.com/lixoft-university/