Quantitative modeling of metastasis: cancer at the organism scale

S. Benzekry
Head of the Inria-Inserm team COMPO
Marseille, France

SMB meeting, 2021 June 14th

COMPO: COMPutational pharmacology and clinical Oncology

MATHEMATICAL MODELING

Mechanistic modeling

D. Barbolosi
S. Benzekry

Pharmacometrics

D. Barbolosi, S. Benzekry
F. Gattacceca

Statistics Data science

S. Benzekry

PHARMACOLOGY

Experimental

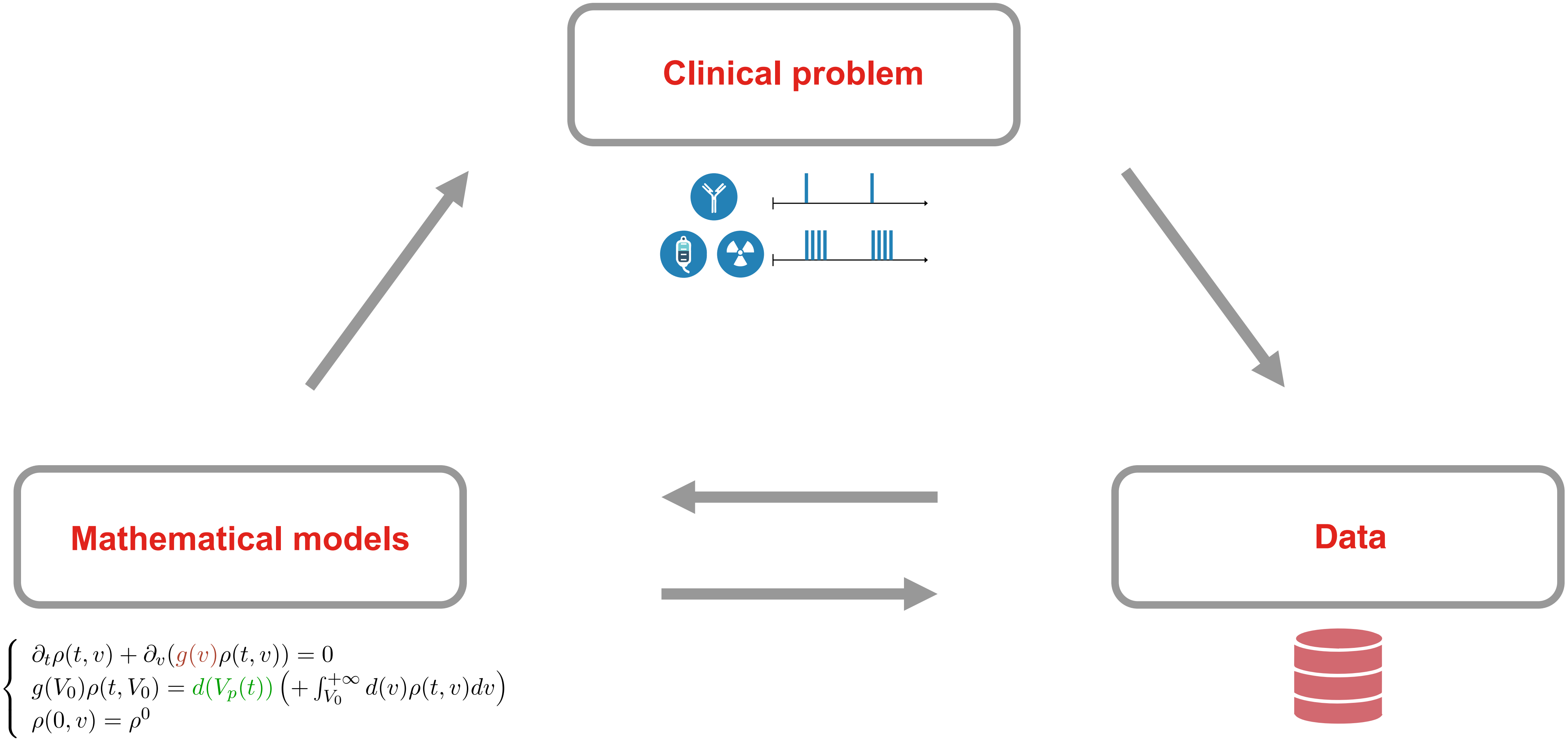
J. Ciccolini
R. Fanciullino
A. Rodallec

Clinical

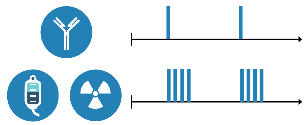
J. Ciccolini
R. Fanciullino
B. Lacarelle

MEDICINE

Pr L. Greillier Dr X. Muracciole
Pr S. Salas



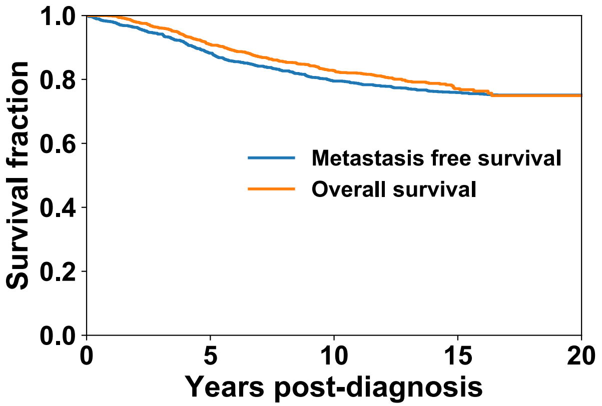
Clinical problem
Mathematical models
Data
$\partial_t \rho(t,v) + \partial_v(g(v)\rho(t,v)) = 0$
$g(V_0)\rho(t,V_0) = d(V_p(t)) \left(+ \int_{V_0}^{+\infty} d(v)\rho(t,v)dv\right)$
$\rho(0,v) = \rho^0$



Clinical problem

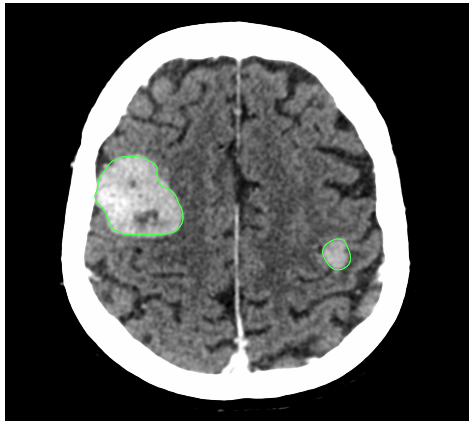
Early-stage breast cancer

- 94% of cases are local or regional at diagnosis but 30% will relapse
- Estimation of the metastatic risk is key to individualize adjuvant therapy
- Reduce the number of chemo cycles for patients with low risk

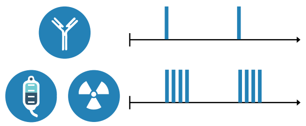


Brain metastases in non-small cell lung cancer (NSCLC)

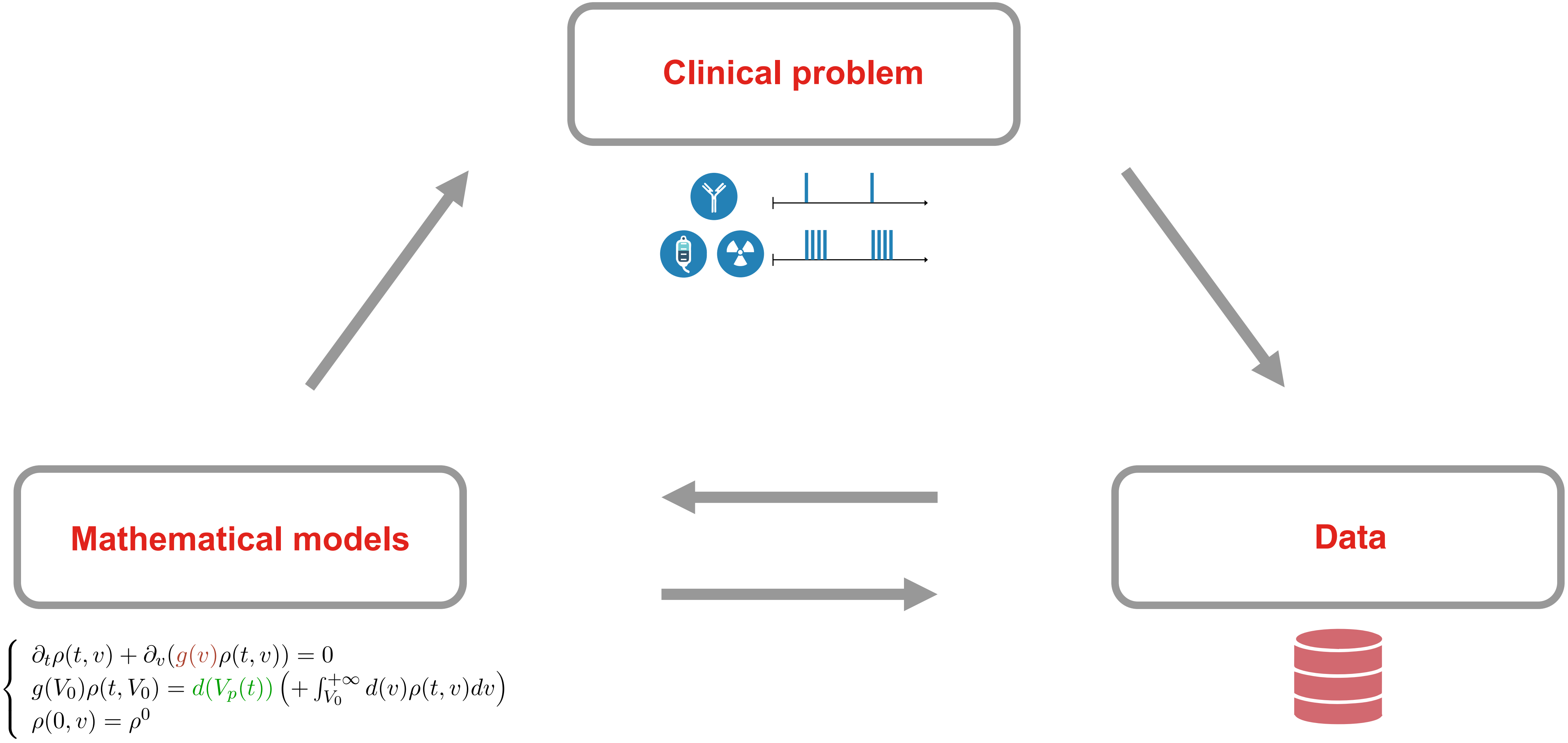
- Decide whether to use whole brain radiation therapy or just (stereotactic) surgery



Objectives

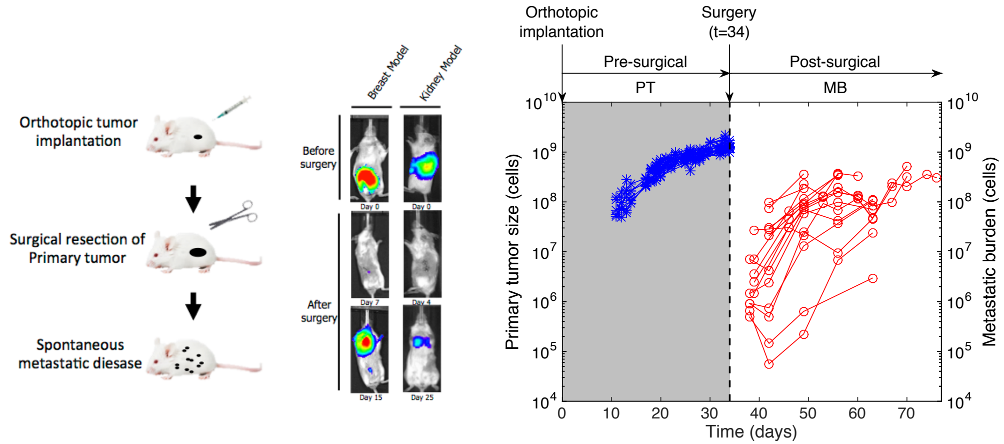


- Use a mechanistic model to predict metastasis
- Combine with machine learning algorithm to select features
- Benchmark predictive power to standard survival methods and machine learning algorithms


Clinical problem
Mathematical models
Data
$$\begin{cases} \partial_t \rho(t,v) + \partial_v(g(v)\rho(t,v)) = 0 \\ g(V_0)\rho(t,V_0) = d(V_p(t)) \left(+ \int_{V_0}^{+\infty} d(v)\rho(t,v)dv \right) \\ \rho(0,v) = \rho^0 \end{cases}$$

Data

Experimental

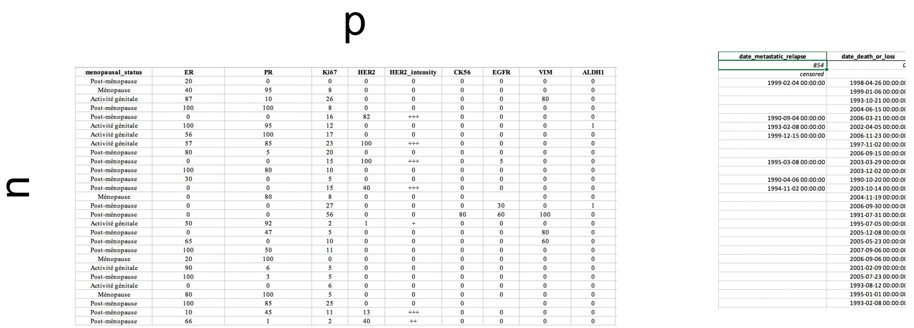


n > 400 animals with different treatments and schedules

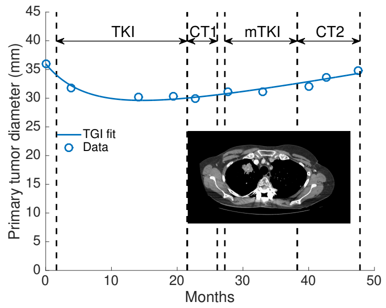
SMARTc wet lab
collaboration with J. Ebos, Roswell Park Cancer Institute, Buffalo, NY

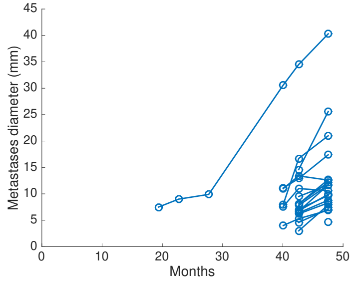
Clinical

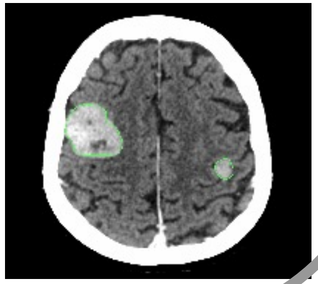
- Databases of metastatic relapse in breast cancer patients with no adjuvant therapy (n=642, p=21 and n=167, p=9)

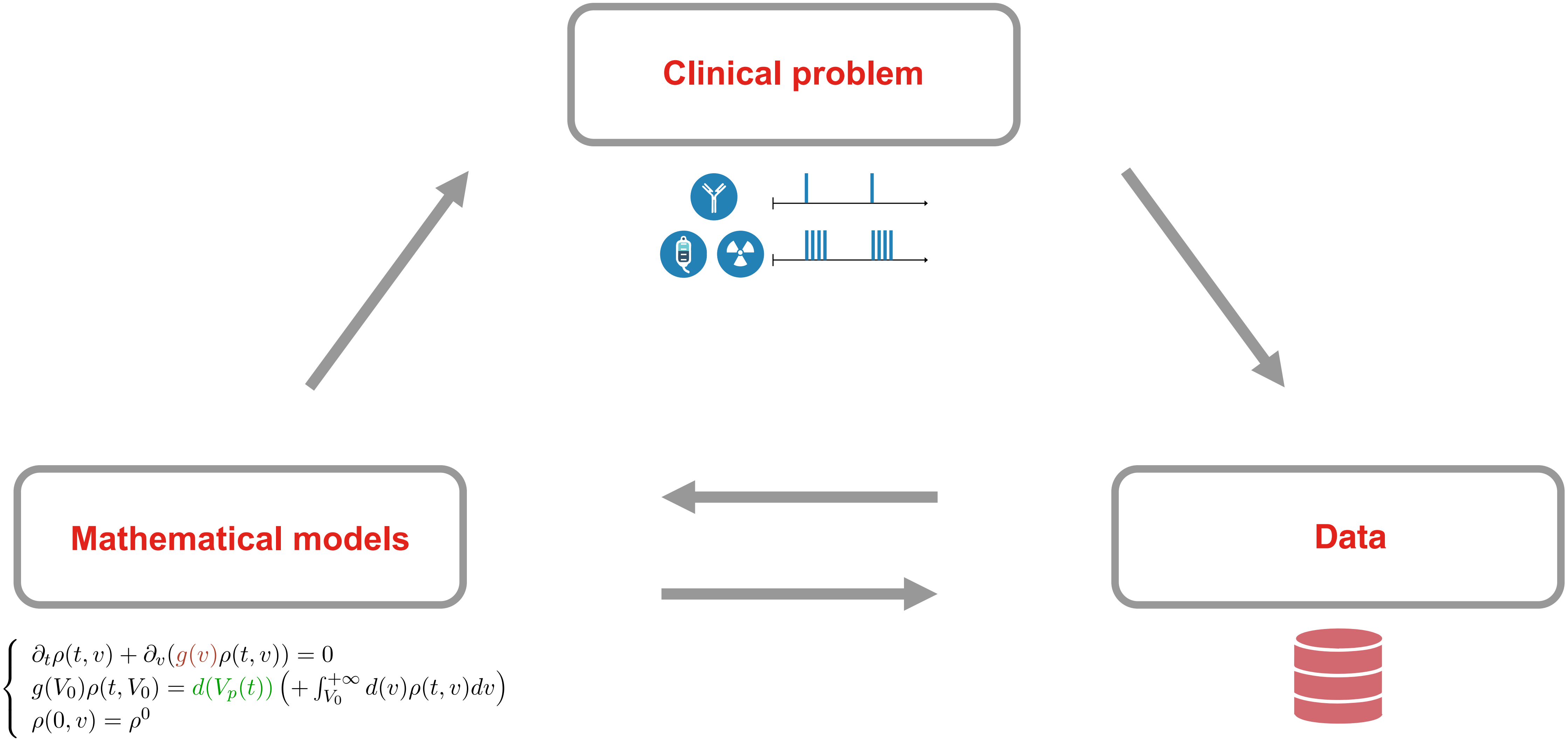


- Number and sizes of brain metastasis in individual NSCLC patients (n=31)

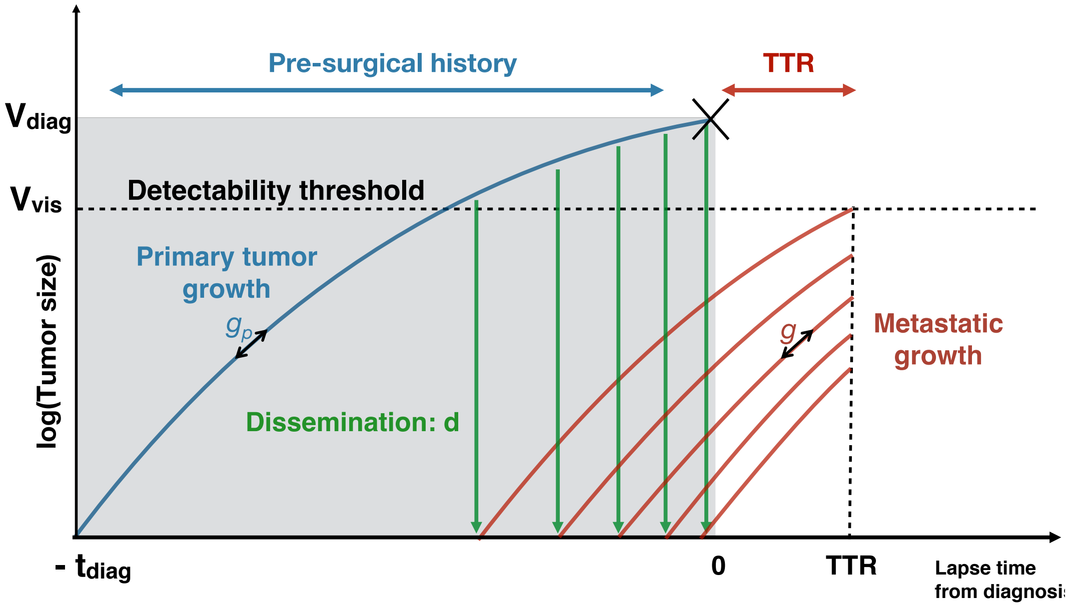







Clinical problem
Mathematical models
Data
$$\begin{cases} \partial_t \rho(t,v) + \partial_v (g(v)\rho(t,v)) = 0 \\ g(V_0)\rho(t,V_0) = d(V_p(t)) \left(+ \int_{V_0}^{+\infty} d(v)\rho(t,v)dv\right) \\ \rho(0,v) = \rho^0 \end{cases}$$

Mathematical models



Growth rates of primary and secondary tumors g_p and g

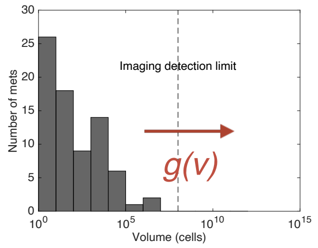
$$\frac{dV_p}{dt} = \mathbf{g_p(t, V_p)}$$

ex: Gompertz

Dissemination rate $d(V_p) = \mu V_p^\gamma$

Size distribution of the metastases $\rho(t,v)$ *[Iwata et al., 2000]*

$$\begin{cases} \partial_t \rho(t,v) + \partial_v (g(v)\rho(t,v)) = 0 \\ g(V_0)\rho(t,V_0) = d(V_p(t)) \left(+ \int_{V_0}^{+\infty} d(v)\rho(t,v)dv \right) \\ \rho(0,v) = \rho^0 \end{cases}$$



Metastatic burden (total number of metastatic cells)

$$M(t) = \int_{V_0}^{+\infty} v\rho(t,v)dv = \int_0^t d\left(V_p(t-s)\right) V(s)ds$$

Number of metastases with size larger than the visible size V_{vis}

$$N_{vis}(t) = \int_{V_{vis}}^{+\infty} \rho(t,v)dv = \int_0^{t-\tau_{vis}} d(V_p(t))dt$$

Time to relapse (TTR) = time from diagnosis to first visible met

$$TTR = \inf\{t > 0 : N_{vis}(t_{diag} + t) \geq 1\}$$

Statistical procedure for model calibration: nonlinear mixed effects modeling

- Classical approach considers each **subject independently**

Subject $1 \le i \le N$, Time t_j

Likelihood maximization

$$y_j^i = M(t_j^i, \theta^i) + \sigma\varepsilon_j^i, \quad \varepsilon_j^i \sim \mathcal{N}(0,1)$$

$$\hat{\theta}^i = \min_{\theta^i} \sum \left(y_j^i - M(t_i, \theta^i)\right)^2$$

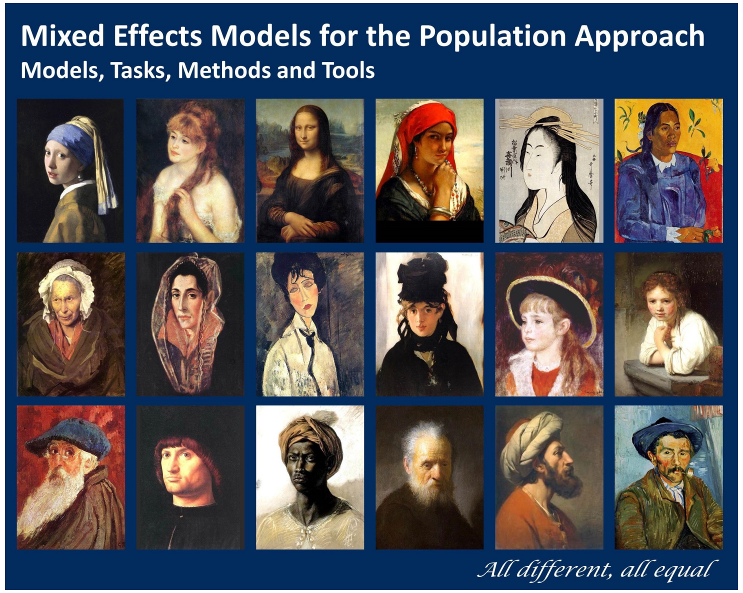
- **Population** approach

$$y_j^i = M(t_j^i, \theta^i) + \sigma\varepsilon_j^i, \quad \varepsilon_j^i \sim \mathcal{N}(0,1)$$

+

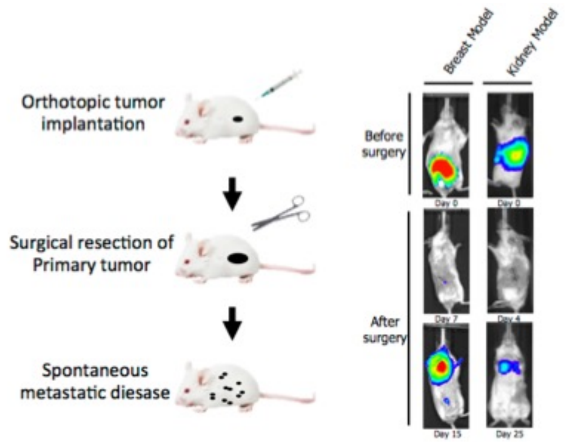
$$\ln\left(\theta^i\right) = \ln\left(\theta_{pop}\right) + \eta^i, \quad \eta_i \sim \mathcal{N}(0,\omega^2)$$

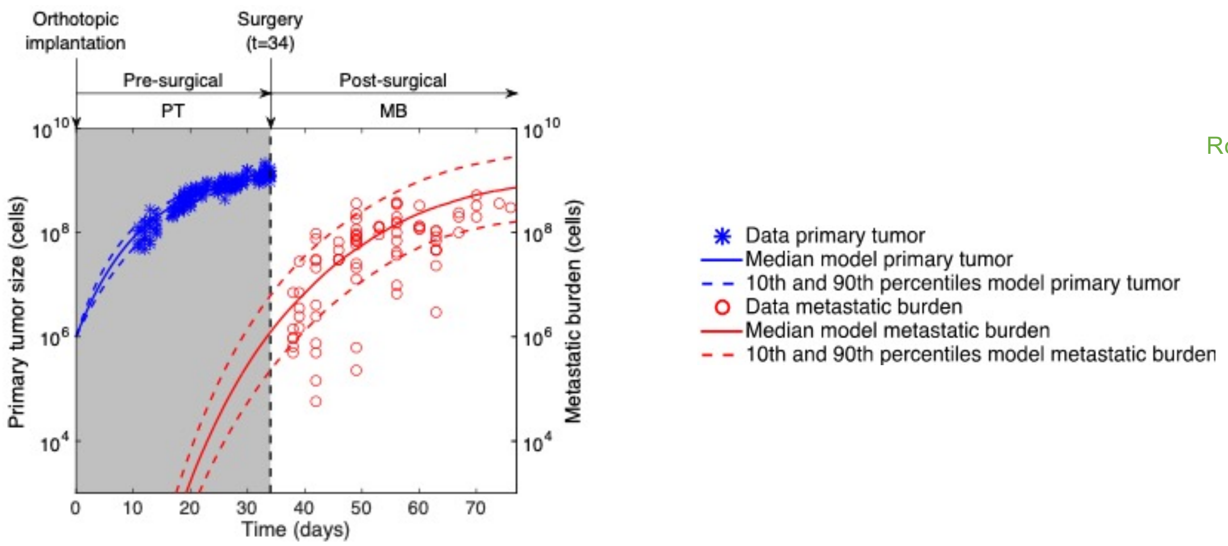
- Parameters to be estimated = θ_{pop} (p), ω $\left(\frac{p(p+1)}{2}\right)$ and σ



Lavielle,CRC press, 2014

Validation on animal data

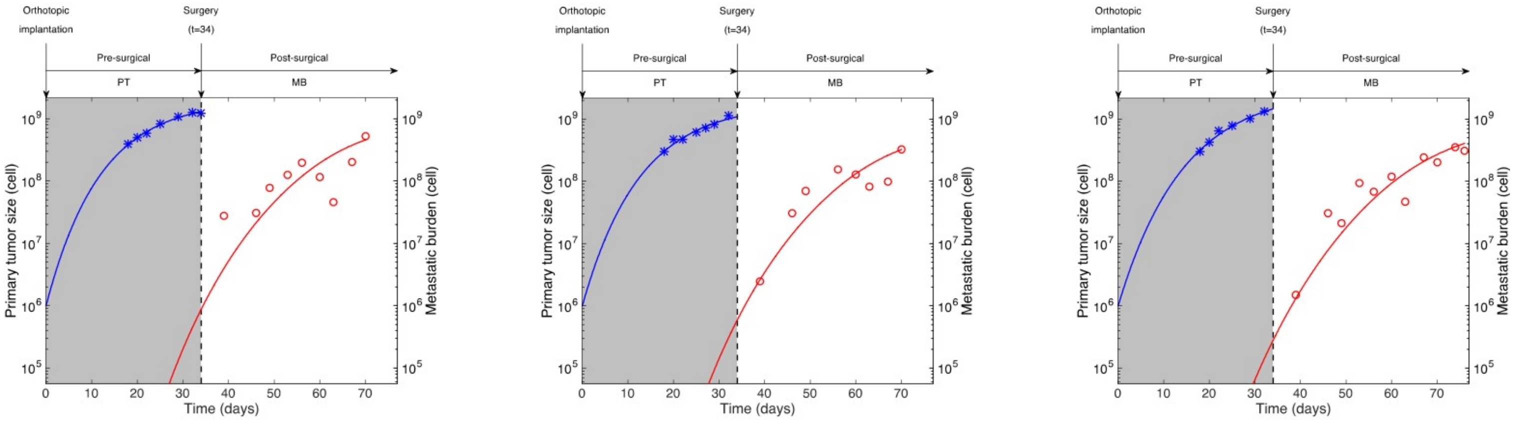




Nonlinear **mixed-effects** statistical model for inter-animal variability

$$\ln\left(\theta^i\right) = \ln\left(\theta_{pop}\right) + \eta^i, \quad \eta_i \sim \mathcal{N}(0, \omega^2)$$

- Main difficulty: PDE model
- Solution: fast Fourier transform for computation of metastatic burden



$\Rightarrow$ **same growth** for PT and mets: $\alpha_p = \alpha$, $\beta_p = \beta$

Benzekry et al. (Ebos), Cancer Res, 2016

Clinical data of individual breast metastatic relapse

K = 21 features

outcome

n =642 patients w/o adjuvant

menopausal_status	ER	PR	Ki67	HER2	HER2_intensity	CK56	EGFR	VIM	ALDH1
Post-ménopause	20	0	0	0	0	0	0	0	0
Ménopause	40	95	8	0	0	0	0	0	0
Activité génitale	87	10	26	0	0	0	0	80	0
Post-ménopause	100	100	8	0	0	0	0	0	0
Post-ménopause	0	0	16	82	+++	0	0	0	0
Activité génitale	100	95	12	0	0	0	0	0	1
Activité génitale	56	100	17	0	0	0	0	0	0
Activité génitale	57	85	23	100	+++	0	0	0	0
Post-ménopause	80	5	20	0	0	0	0	0	0
Post-ménopause	0	0	15	100	+++	0	5	0	0
Post-ménopause	100	80	10	0	0	0	0	0	0
Post-ménopause	30	0	5	0	0	0	0	0	0
Post-ménopause	0	0	15	40	+++	0	0	0	0
Ménopause	0	80	8	0	0	0		0	0
Post-ménopause	0	0	27	0	0	0	30	0	1
Post-ménopause	0	0	56	0	0	80	60	100	0
Activité génitale	50	92	2	1	+	0	0	0	0
Post-ménopause	0	47	5	0	0	0	0	80	0
Post-ménopause	65	0	10	0	0	0	0	60	0
Post-ménopause	100	50	11	0	0	0	0	0	0
Ménopause	20	100	0	0	0	0	0	0	0
Activité génitale	90	6	5	0	0	0	0	0	0
Post-ménopause	100	3	5	0	0	0	0	0	0
Activité génitale	0	0	6	0	0	0	0	0	0
Ménopause	80	100	5	0	0	0	0	0	0
Post-ménopause	100	85	25	0	0	0		0	0
Post-ménopause	10	45	11	13	+++	0	0	0	0
Post-ménopause	66	1	2	40	++	0	0	0	0

metastatic_relapse	date_metastatic_relapse
Yes	04/02/1999
No	
No	
No	
Yes	04/09/1990
Yes	08/02/1993
Yes	15/12/1999
No	
No	
Yes	08/03/1995
No	
Yes	06/04/1990
Yes	02/11/1994
No	
No	
No	
No	
No	
No	
No	
No	
No	
No	
No	
No	
No	
Yes	27/10/1999
No	
No	
No	
No	

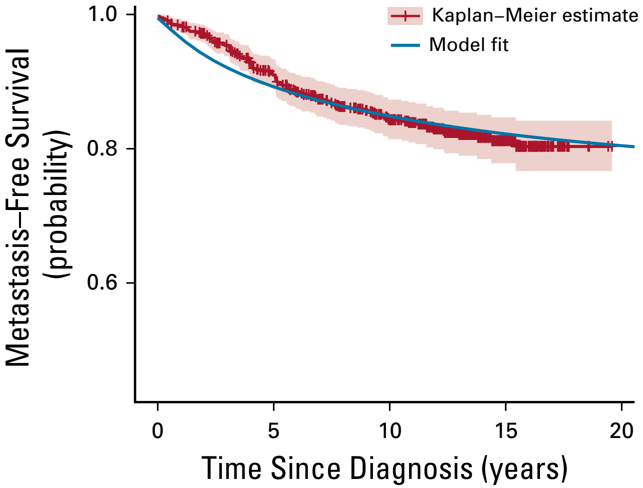
Institut Bergonié, Bordeaux, FR

Descriptive power: fit to the data

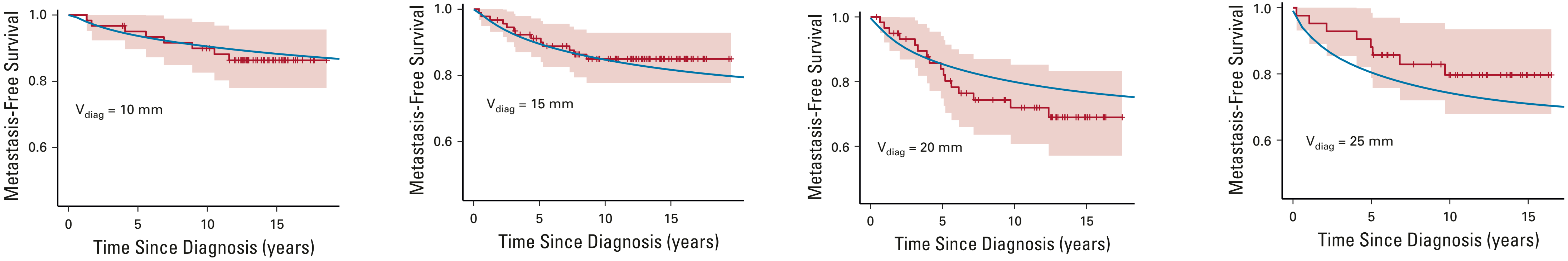
$$\ln\left(T^i\right) = \ln\left(TTR\left(V^i_{diag};\alpha^i,\mu^i\right)\right) + \varepsilon^i, \quad \varepsilon^i \sim \mathcal{N}(0,\sigma^2)$$

$$S\left(t|\alpha^i,\mu^i\right) = \mathbb{P}\left(T^i > t|\alpha^i,\mu^i\right)$$

$$\boxed{\begin{aligned} \ln\left(\alpha^i\right) &= \ln\left(\alpha_{pop}\right) + \eta^i_\alpha, \quad \eta^i_\alpha \sim \mathcal{N}(0,\omega^2_\alpha) \\ \ln\left(\mu^i\right) &= \ln\left(\mu_{pop}\right) + \eta^i_\mu, \quad \eta^i_\mu \sim \mathcal{N}(0,\omega^2_\mu) \end{aligned}}$$



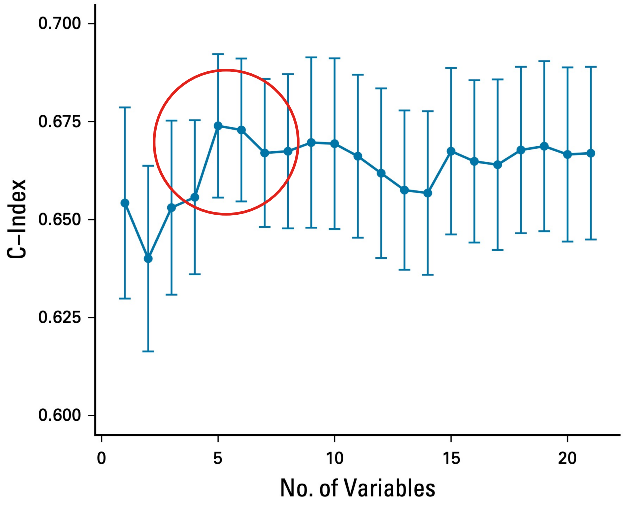
Parameter	Estimate	Relative Standard Error (%)
Model without covariates		
$\log \alpha_{pop}$	–6.34	12.6
$\log \mu_{pop}$	–26.8	3.68
σ	0.542	28.4
ω_α	3.37	36.4
ω_μ	3.78	15.9

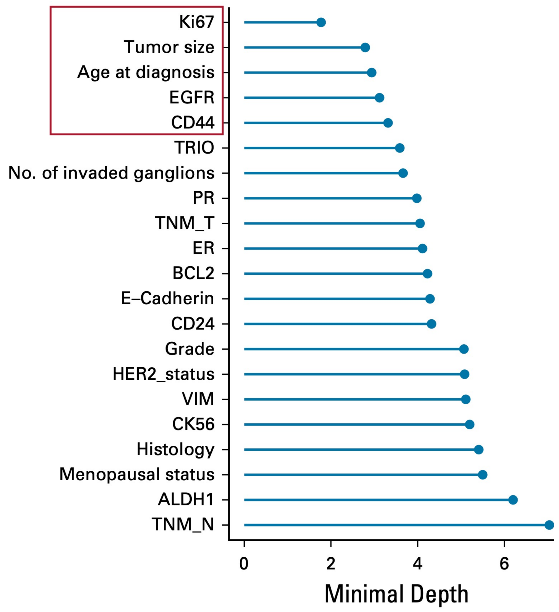


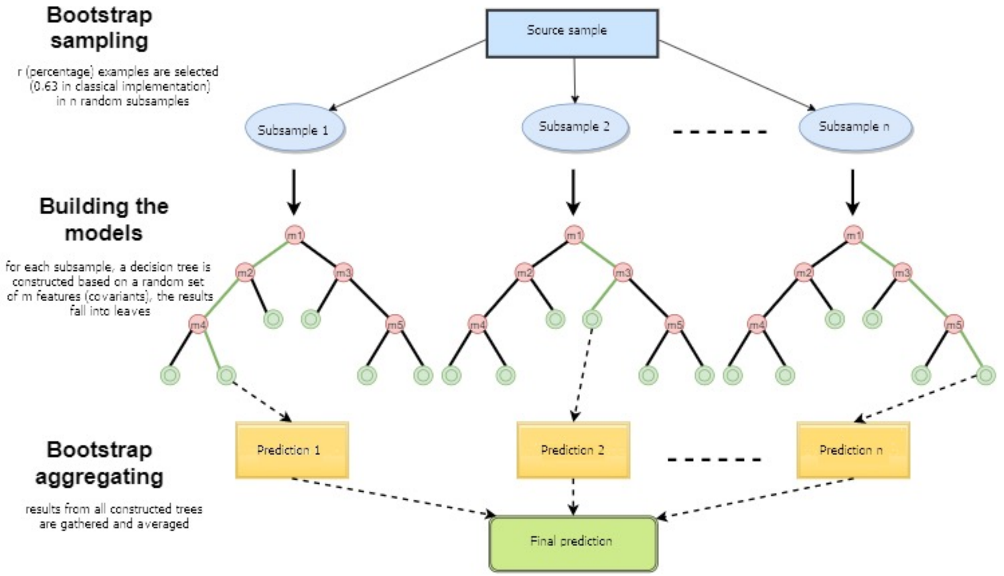
Nicolò et al., JCO: Clin Cancer Inform, 2020

Feature selection: random survival forests

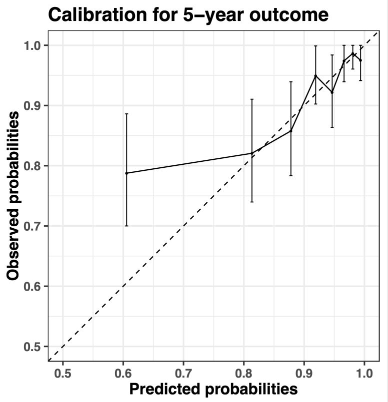
c-index = 0.69 (0.67 – 0.71)

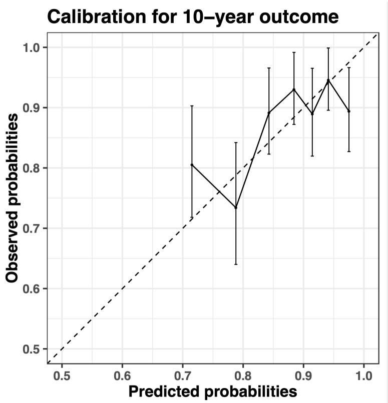


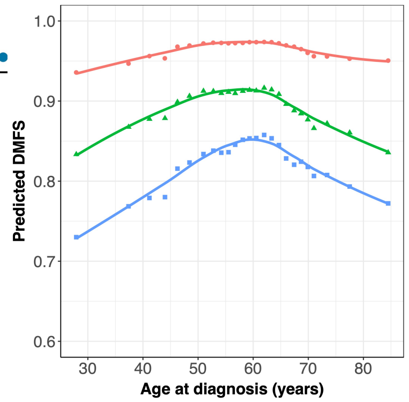


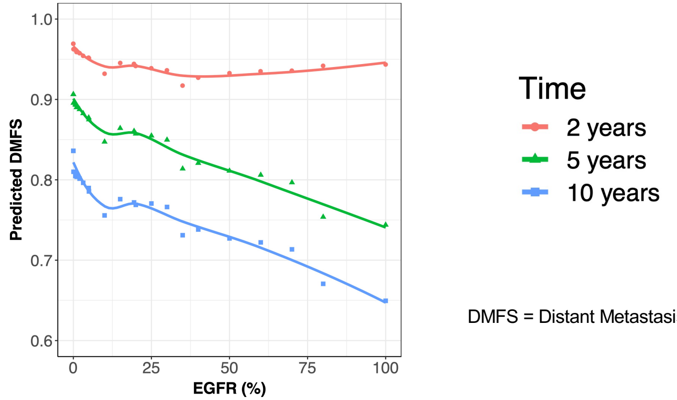


M. Dmitrievsky, https://www.mql5.com/en/articles/3856







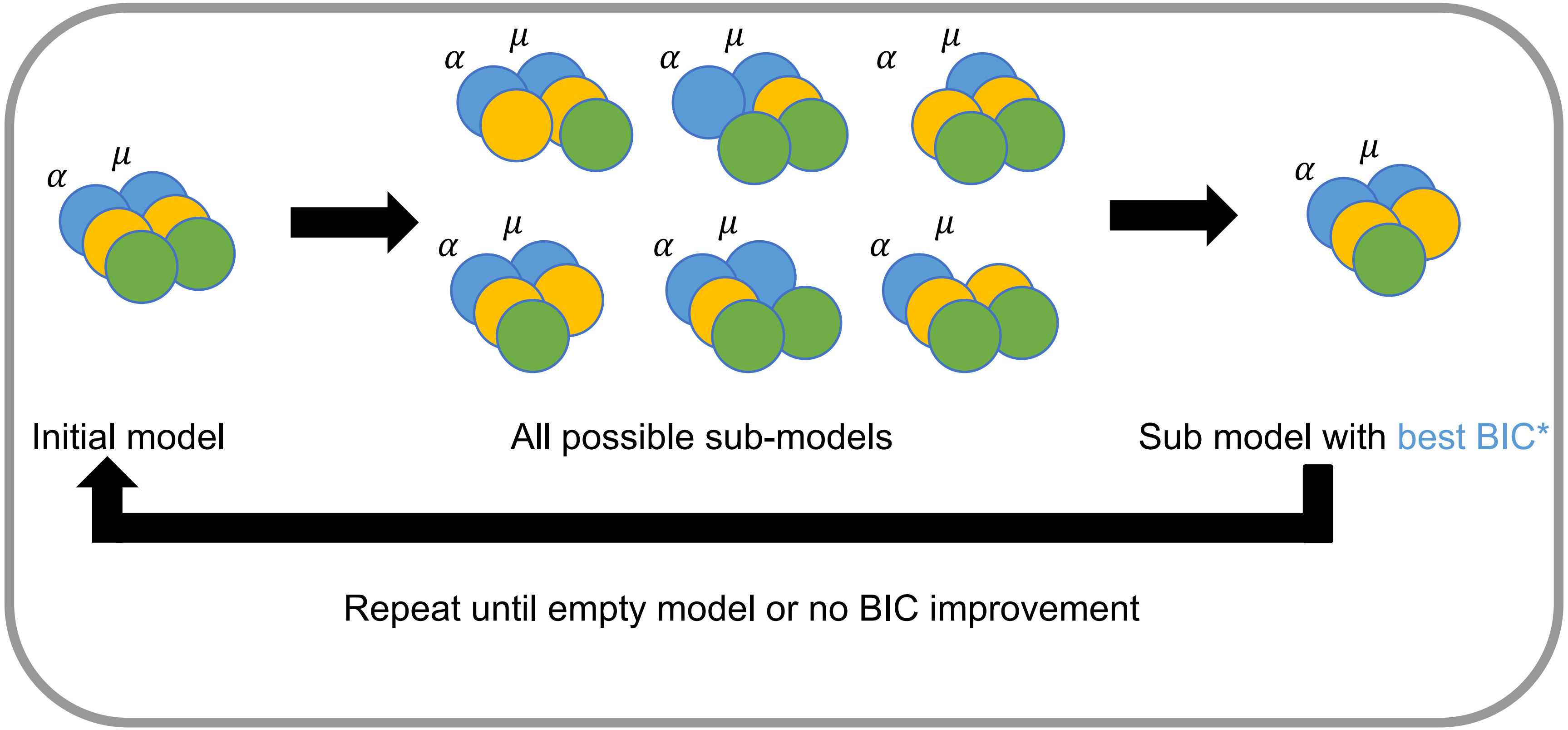


DMFS = Distant Metastasis-Free Survival

⇒ nonlinear effect of covariates and non proportional hazard

Ishwaran et al., Ann Appl Stat, 2008

Mechanistic selection: backward stepwise selection



$$\ln\left(\mu^i\right) = \ln\left(\mu_{pop}\right) + \beta_\mu^T \mathbf{x}_\mu^\mathbf{i} + \eta_\mu^i, \quad \eta_\mu^i \sim \mathcal{N}(0, \omega_\mu^2)$$

$$\ln\left(\alpha^i\right) = \ln\left(\alpha_{pop}\right) + \beta_\alpha^T \mathbf{x}_\alpha^\mathbf{i} + \eta_\alpha^i, \quad \eta_\alpha^i \sim \mathcal{N}(0, \omega_\alpha^2)$$

Parameter	Estimate	Relative Standard Error (%)	*P*
Model with covariates			
$\log \alpha_{pop}$	−9.01	10.8	
$\beta_{Ki67,\alpha}$	0.093	29.6	.001
$\beta_{CD44,\alpha}$	0.017	57.7	.083
$\log \mu_{pop}$	−25.9	4.4	
$\beta_{EGFR,\mu}$	0.053	38.1	.009
σ	0.606	24	
ω_α	2.75	22.1	
ω_μ	3.03	20.5	

Selected Covariate model:

α ← Ki67 (proliferation marker), CD44 (stem cell marker)

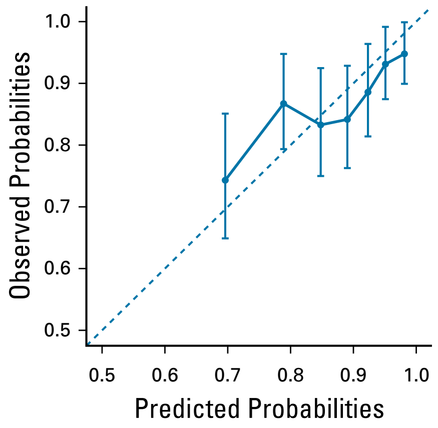
μ ← EGFR (basal marker)

***Bayesian Information Criteria**, classical criteria for model selection

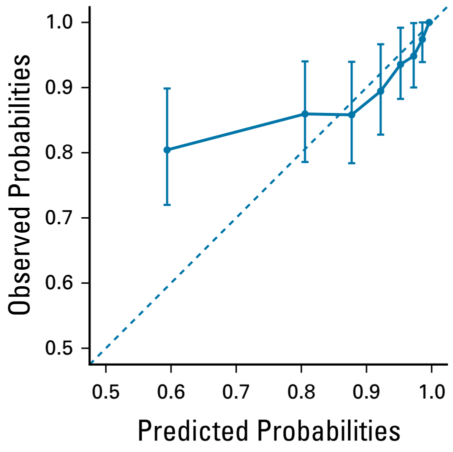
Predictive power

c-index = 0.67
(10-folds cross-validation)

Calibration for 10-Year Outcome



Calibration for 5-Year Outcome



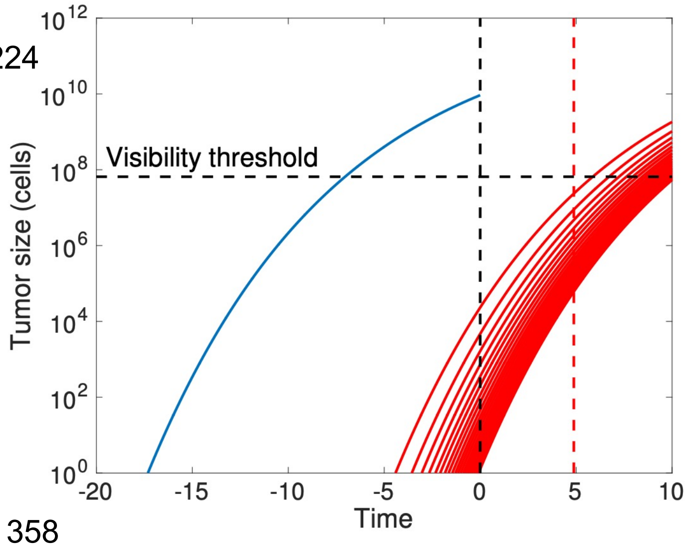
Algorithm	AUROC	Accuracy	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	F1
5-year							
Mechanistic model	0.73	0.68	**0.75**	0.67	0.19	**0.96**	0.30
Random survival forest	0.73	0.69	0.64	0.70	0.18	0.95	0.28
Random forest	**0.75**	0.66	0.71	0.66	0.18	**0.96**	0.28
Logistic regression	**0.75**	0.83	0.42	0.87	0.24	0.94	**0.31**
k-Nearest neighbor	0.62	**0.91**	0.02	**1.00**	**0.41**	0.91	0.05
Gradient boosting	0.71	0.90	0.11	0.98	0.36	0.92	0.17
Support vector machine	0.64	0.87	0.09	0.95	0.15	0.91	0.11
Cox	0.71	0.72	0.66	0.73	0.20	0.95	**0.31**
10-year							
Mechanistic model	0.67	**0.67**	0.62	**0.68**	**0.30**	0.89	**0.41**
Random survival forest	**0.69**	0.62	**0.71**	0.60	0.28	**0.90**	**0.41**
Cox	0.65	0.65	0.61	0.65	0.28	0.88	0.39

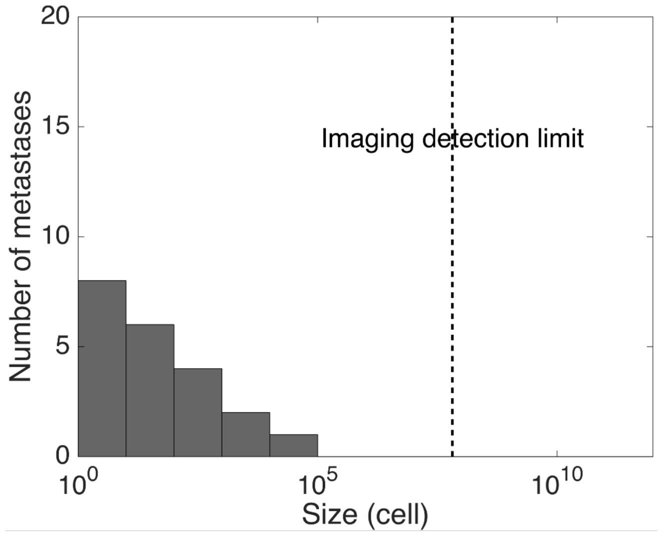
⇒ **Similar predictive power** as classical statistical Cox model or other machine learning algorithms

Nicolò et al., JCO: Clin Cancer Inform, 2020

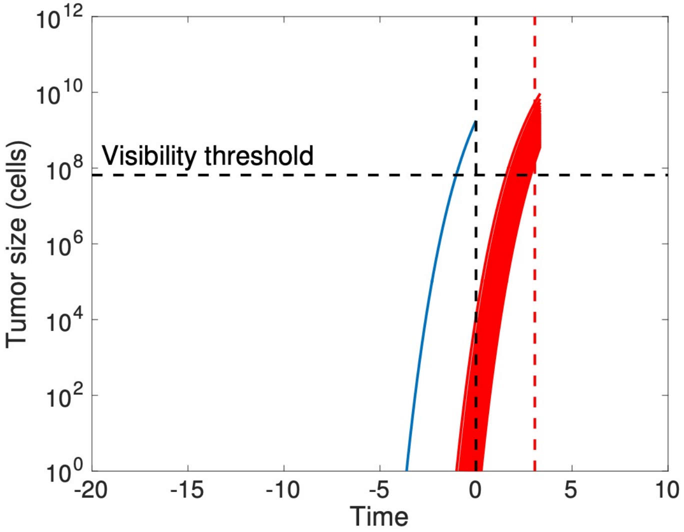
Predictive simulations of the mechanistic model

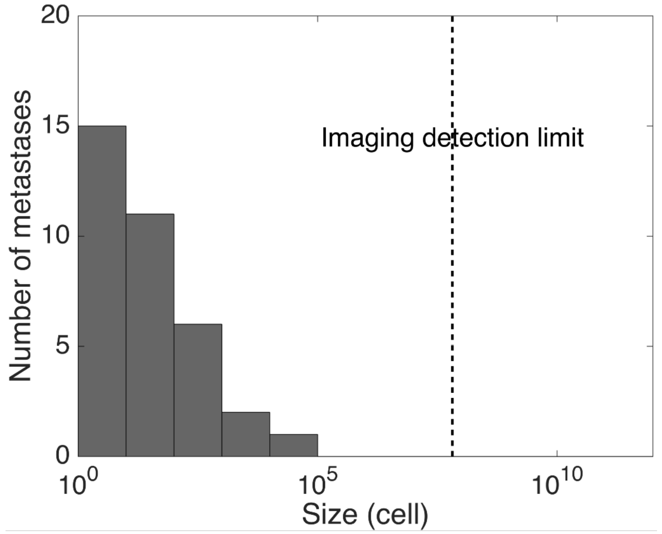
Patient 224





Patient 358





	Patient 224	Patient 358
Tumor size (mm)	26	15
Ki67 (%)	40	52
EGFR (%)	0	80
CD44 (%)	0	0
Observed TTR (years)	4.88	3.06
Predicted TTR (years)	6.18	1.97
Prediction error (years)	1.3	1.09
Pre-surgical history (years)	17.3	3.61
Number of mets	21	35

$$\log \hat{\alpha}^i = \log \alpha_{pop} + \beta_{Ki67,\alpha} \cdot Ki67^i + \beta_{CD44,\alpha} \cdot CD44^i$$

$$\log \hat{\mu}^i = \log \mu_{pop} + \beta_{EGFR,\mu} \cdot EGFR^i$$

Nicolò et al., JCO: Clin Cancer Inform, 2020

C. Bigarré
X. Muracciole

Second dataset: PAI-1

Data

N = 167 patients

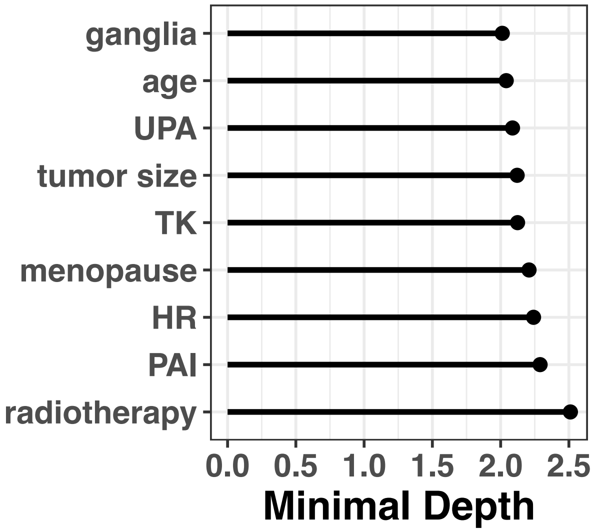
Outcome: metastasis free survival

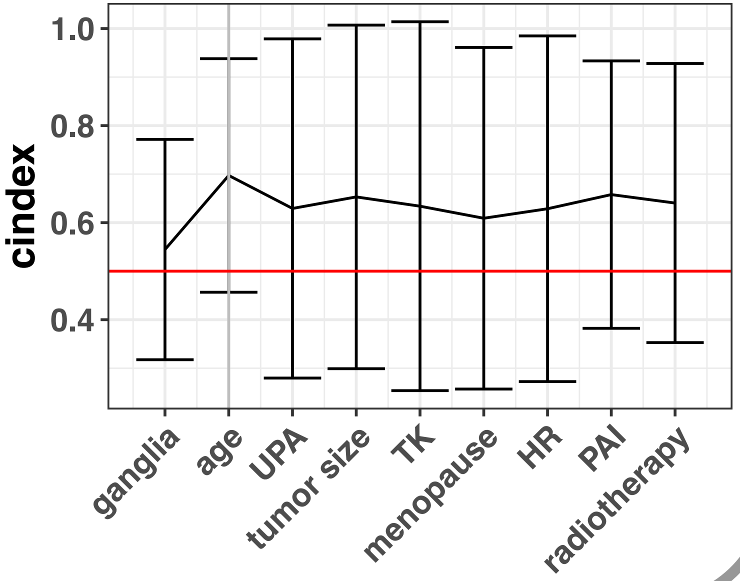
p = 9 covariates

- patient info
  - Age
  - Menopausal status
- tumor biomarkers
  - PAI-1
  - TK
  - UPA
  - HR status
- sugery info
  - Tumor size
  - Number of positives ganglia
- treatment
  - adjuvant radiotherapy

PAI-1, TK, UPA → validated predictors of metastatic relapse*

Random survival forest





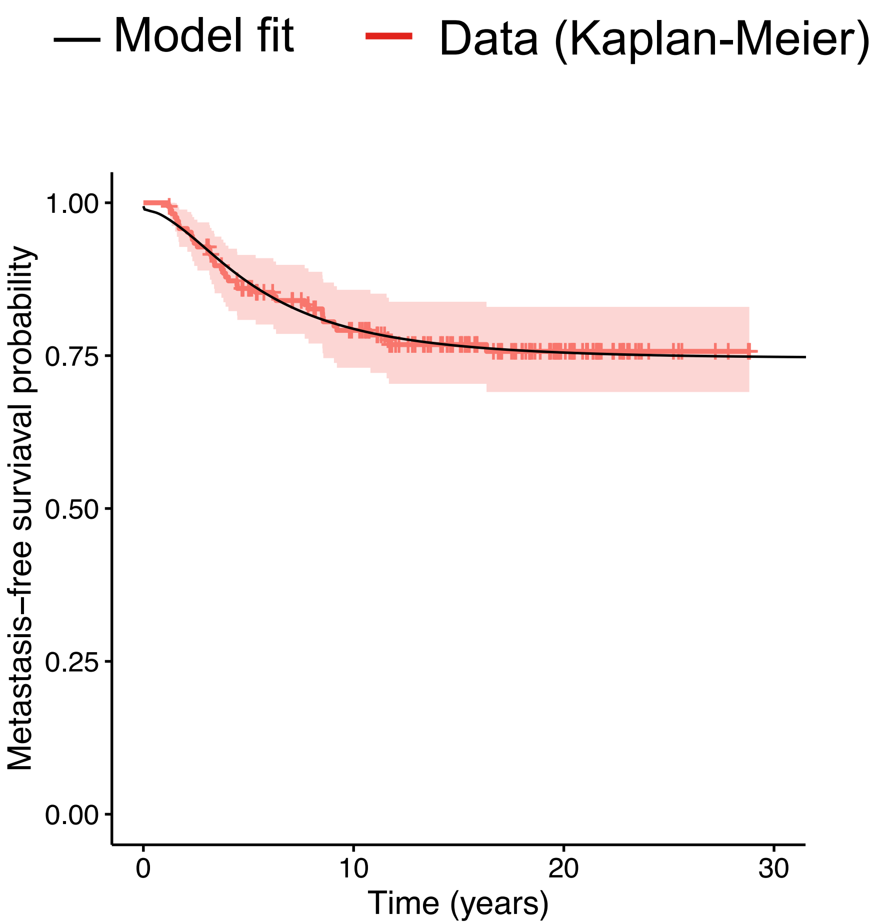
No clear results for variable importance → All covariates included in the next step

*Duffy, M. J. et al., M. uPA and PAI-1 as biomarkers in breast cancer: validated for clinical use in level-of-evidence-1 studies. *Breast Cancer Res* **16**, 428 (2014).

*Harbeck N. et al., Ten-year analysis of the prospective multicentre Chemo-N0 trial validates ASCO-recommended biomarkers uPA and PAI-1 for therapy decision making in n
ode-negative breast cancer patients, Eur J Cancer (2013)

Selected mechanistic model

C. Bigarré



$$\log T^i \sim \log\left(TTR(V^i,\alpha^i,\mu^i)\right) + \mathcal{N}(0,\sigma^2)$$
$$\log(\alpha^i) \sim \log(\alpha_{pop}) + \beta_{HR,\alpha} \cdot 1^i_{HR} + \mathcal{N}(0,\Omega^2_\alpha)$$
$$\log(\mu^i) \sim \log(\mu_{pop}) + \beta_{PAI,\mu} \cdot PAI^i + \mathcal{N}(0,\Omega^2_\mu)$$

Parameter	Estimate	R.S.E	p-value
α	0.019	18%	
$\beta_{HR,\alpha}$	-0.713	50.1%	**0.0459**
μ	5.13e-15	381%	
$\beta_{PAI,\mu}$	0.352	31.7%	**0.0016**
σ	0.377	16.9%	
Ω^2_α	0.308	109%	
Ω^2_μ	14.9	34.1%	

R.S.E = Relative Standard Error

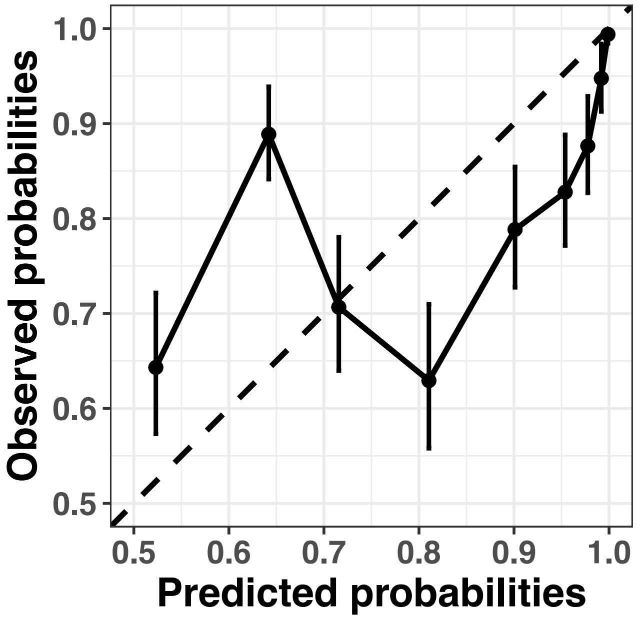
Predictive performances

C. Bigarré

RSF

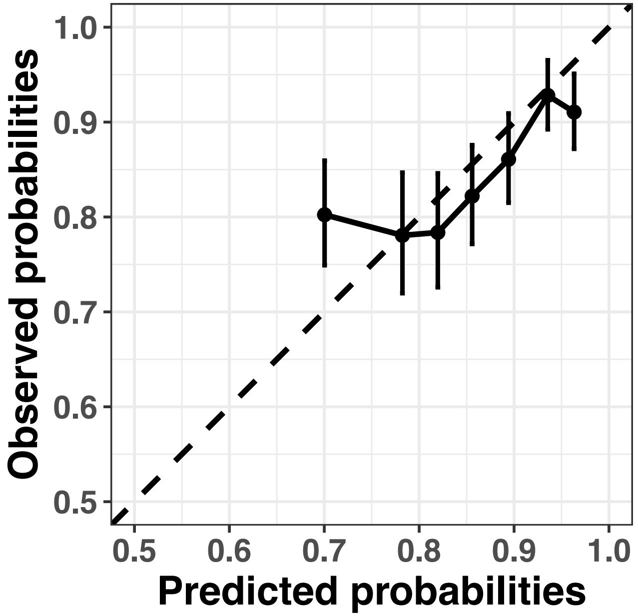
c-index: 0.70 (0.45-0.94)

Calibration at 10 years



Mechanistic model

c-index: 0.70 (0.68-0.72)



Cox

	c-index
Full model (p=9)	0.67 (0.33, 1.00)
Tumor size + HR + PAI-1	0.71 (0.45, 0.97)

Conclusions and perspectives

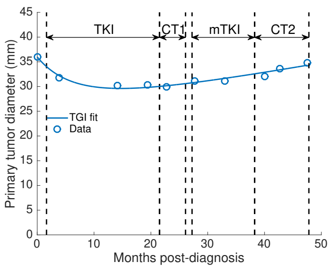
- Similar predictive performances of Cox regression (c-index 0.67 - 0.72), random survival forest (c-index 0.67-0.71) and a novel mechanistic model (c-index 0.63 - 0.70) for pure prediction
- Predictive power is confirmed (improved) in an external data set
- Mechanistic modeling provides biological and clinical insights that ML does not:
  - Ki67 correlates with proliferation rate α (expected but reassuring), also CD44 or hormonal status
  - EGFR and PAI1 correlate with μ (metastatic potential)
  - prediction of the **invisible metastatic state** at diagnosis $\Rightarrow$ potential for **personalized adjuvant therapy**
- Current/future avenues:
  - Further investigations to refine the modeling (dormancy, etc…)
  - Include **treatment**
  - Include (high-dimensional) transcriptomic data

Results: individual clinical data

Methodology

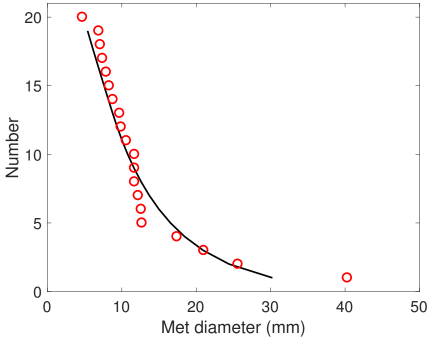
Primary tumor

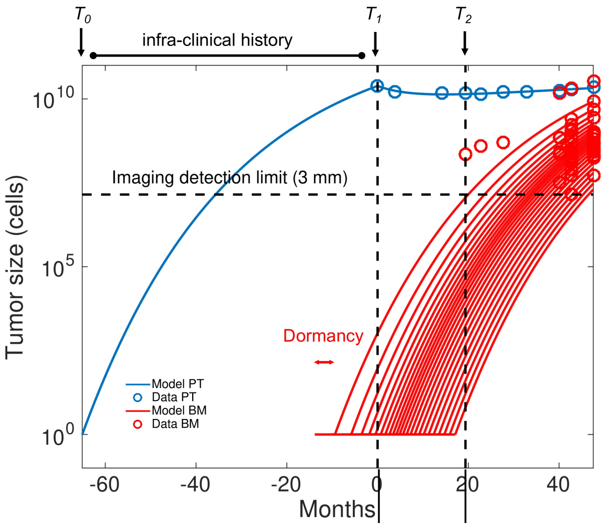
$$\frac{dS_p}{dt} = \alpha_1 S_p - \kappa e^{-\frac{t-T_d}{t_{res}}\ln 2} S_p, \quad \forall\, t \geq T_d.$$

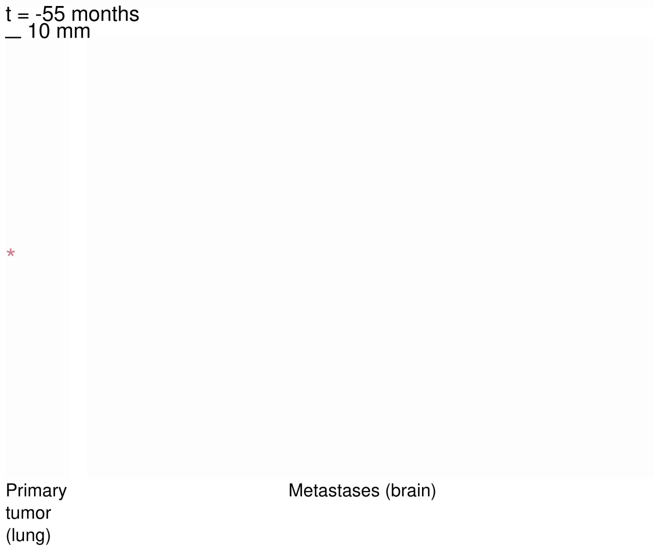


Metastases

$$f(t,v) = \int_v^{+\infty} \rho(t,u)du = N\left(t - \tau(v)\right)$$





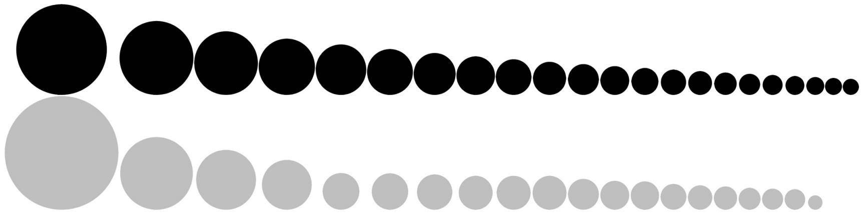


- Main difficulties:

no data on PT pre-treatment + fit PDE to discrete data

Model

Data



Results

- Dormancy, estimated to 133 days ± 4.2%
- Good fit of the entire longitudinal data (47 BM sizes over 6 time points) with only 3 parameters.
- Onset of BM 14-19 months before diagnosis

Ongoing (collab. P. Schlicke, TU München)

- Extend to larger cohort: complex treatment histories
- Include effect of treatment ⇒ $g(t, v)$ ⇒ nontrivial difficulties for param estimation

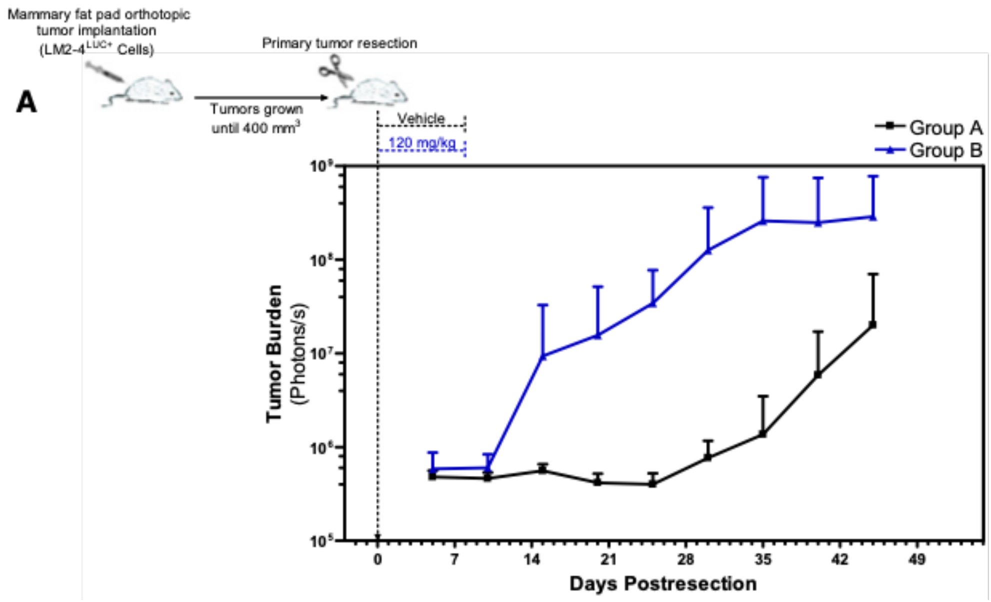
Bilous et al., Sci Rep, 2019

Differential effects of anti-angiogenic therapies between primary tumor and metastases

Cancer Cell
Report

Accelerated Metastasis after Short-Term Treatment with a Potent Inhibitor of Tumor Angiogenesis

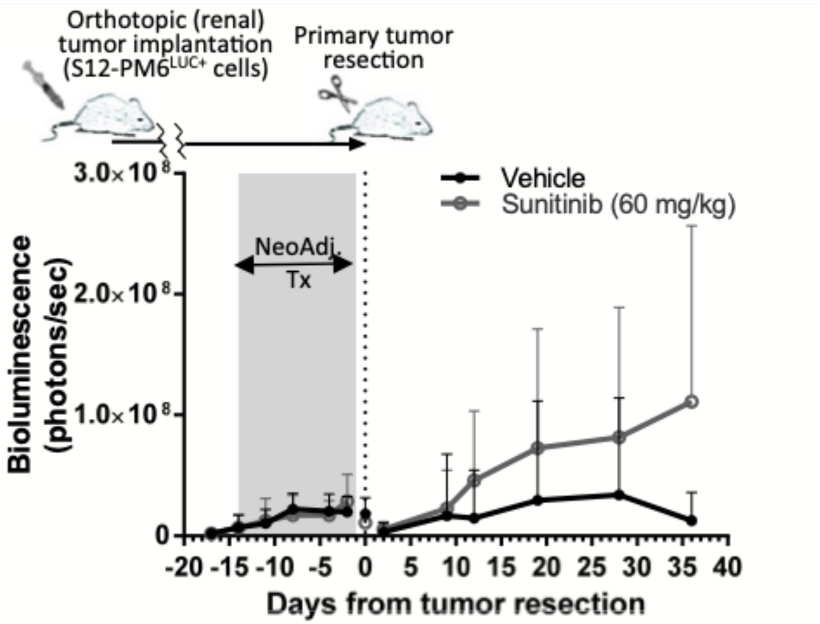
John M.L. Ebos,[1,2] Christina R. Lee,[1] William Cruz-Munoz,[1] Georg A. Bjarnason,[3] James G. Christensen,[4] and Robert S. Kerbel[1,2,*]



Published online: October 31, 2014
Research Article
TRANSPARENT PROCESS
OPEN ACCESS
EMBO Molecular Medicine

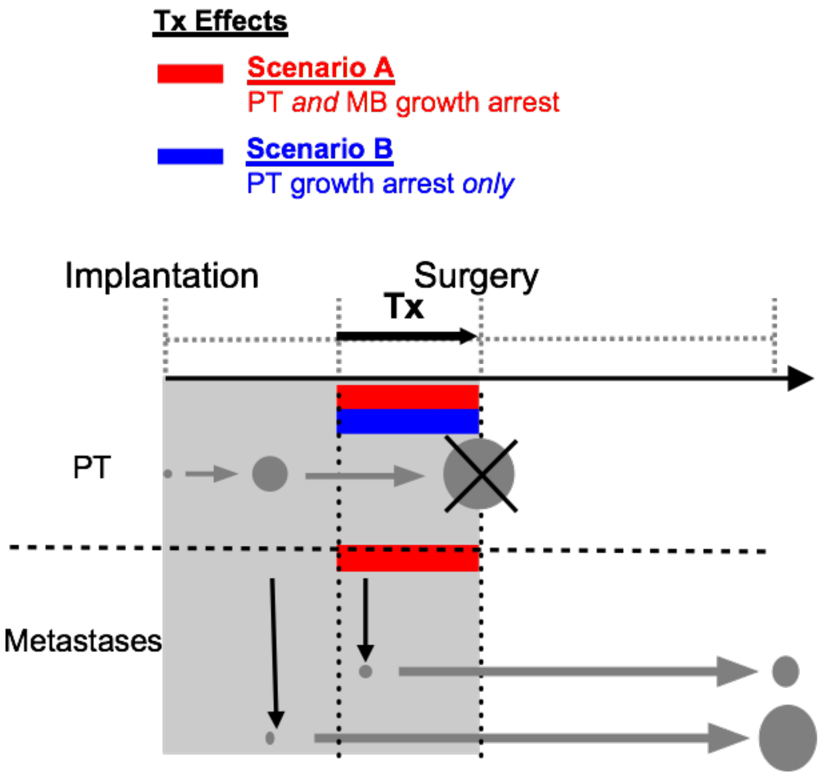
Neoadjuvant antiangiogenic therapy reveals contrasts in primary and metastatic tumor efficacy

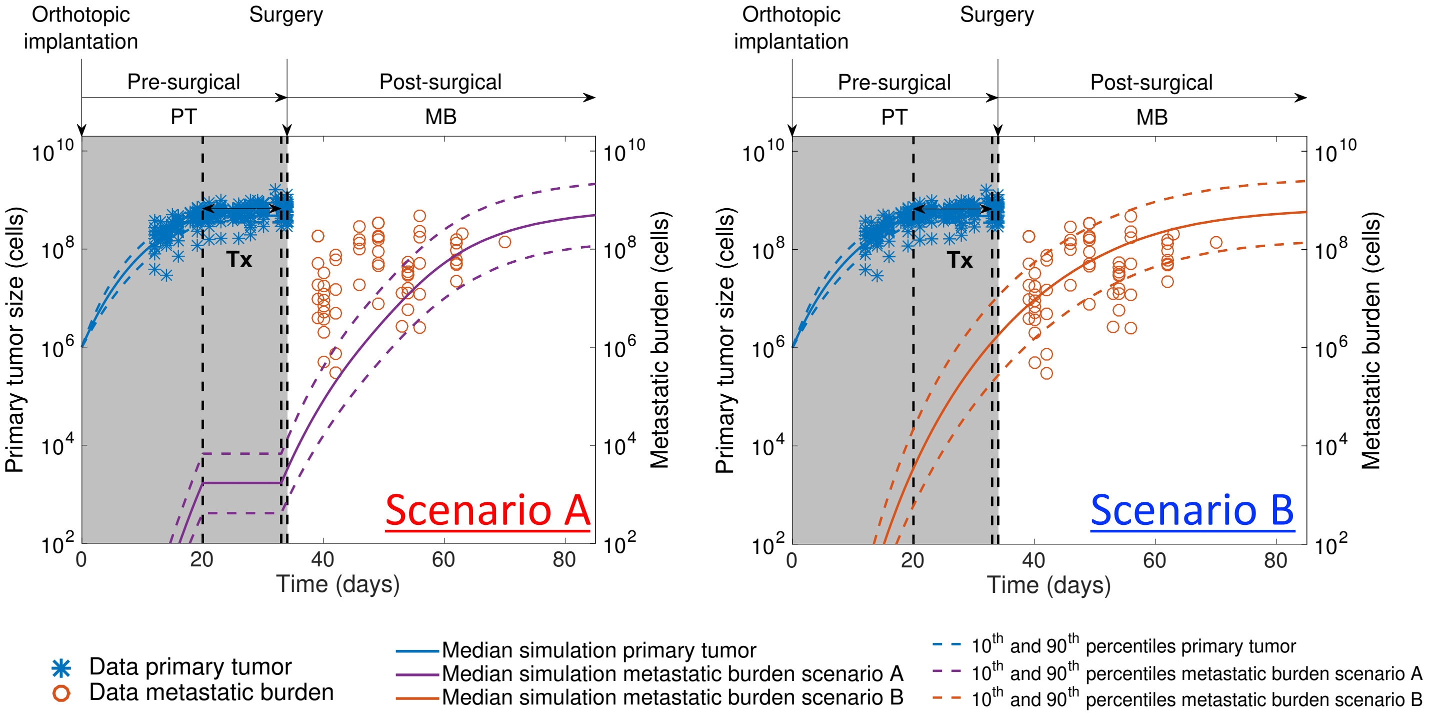
John M L Ebos[1,*], Michalis Mastri[1], Christina R Lee[2], Amanda Tracz[1], John M Hudson[2], Kristopher Attwood[3], William R Cruz-Munoz[2], Christopher Jedeszko[2], Peter Burns[2,4] & Robert S Kerbel[2,4]



Simulations of the effect of neoadjuvant sunitinib treatment on metastases suggest no effect on growth of metastases

- Parameter values from the previous study on control groups ⇒ simulations are pure mechanistic predictions
- In first approximation, the effect of the drug was modeled by setting the tumor growth rate to zero during the phase of treatment





Nicolo et al. (Ebos, Benzekry), in preparation, 2021

We have open positions!!

- **Full research tenure position**

sebastien.benzekry@inria.fr