Mechanistic modeling of the natural history of cancer progression and metastatic dissemination in stage I non-small cell lung cancer patients after curative surgery

Romain Zakrajsek[1], Linh Nguyen Phuong[1], Célestin Bigarré[1], Andrea Vaglio[1], David Taieb[2], Xavier-Benoit D'journo[2], Paul Habert[2], Alexis Jacquier[2], Pascal-Alexandre Thomas[2], David Boulate[1,2,*], Sébastien Benzekry[1,*]

1 COMPutational pharmacology and clinical Oncology Team, Inria Sophia-Antipolis – Méditerranée, Cancer Research Center of Marseille, Inserm UMR1068, CNRS UMR7258, Aix-Marseille University UM105, Marseille, France
2 Assistance Publique - Hôpitaux de Marseille, France
* Joint senior co-authors

1. Background

LUCA-pi ("LUng Cancer – prevention and interception")

- Launched in 2024
- Aims to develop:
  - biological biomarkers
  - AI-based tools

  to improve lung cancer **prevention**, **early detection** and **recurrence**

Cancer recurrence

- 10-20% of early-stage cancer cases relapse after surgery[1]
- Hypothesized that micrometastases must exist in patients even at early-stages[2]
- Tumor cells disseminate through all of their natural history to create metastases[3]

EPITHOR ("Épidémiologie en chirurgie THORacique")

- French national thoracic surgery registry

LUCA-pi RETRO database

- Subset of EPITHOR from Hôpital Nord (Marseille)
- Patients who underwent thoracic surgery from 2011 to 2024

OBJECTIVES

- **Model the natural history of cancer progression and metastatic dissemination**
- **Explore covariates associated with the mechanistic parameters**
- **Create a "digital twin" for personalized individual risk prediction**

2. Data

Population

- Non-small cell lung cancer
- Stage 1 patients with no perioperative treatment

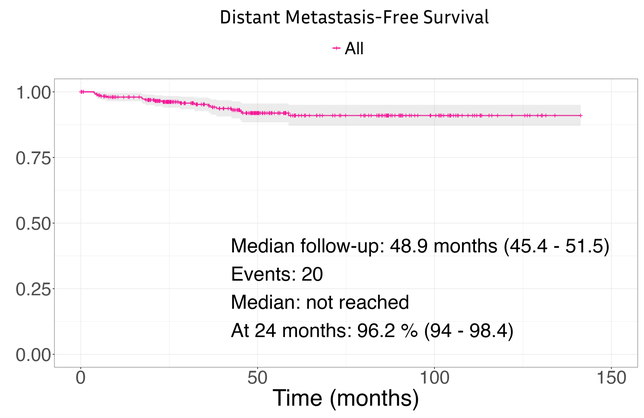
Features

- Type : clinical, anatomopathological, imaging (scanner , TEP scan), follow-up
- Volume and diameter of the primary tumor available
- Total : **41 features** included in this test (165 features available in total)

Patients		**N = 306**
Sex		
	Male	187 (61%)
	Female	119 (39%)
Age		66 (±9)
BMI		25.5 (±4.5)
Performance status		
	0	206 (67%)
	1	87 (28%)
	2	13 (5%)
Anatomical pathology		
	Adenocarcinoma	246 (80%)
	Squamous cell carcinoma	57 (19%)
	Large cell carcinoma	3 (1%)

Outcome

- Distant metastasis-free survival (DMFS)
  - Defined as **first occurrence of distant metastatic recurrence**
  - Censored at the time of death or last follow-up



3. Methods

Primary tumor growth

- Starts at $t = 0$ with one cell
- Grows following a Gompertz law : $V_p(t) = \exp\left(\frac{\alpha}{b}\left(1 - e^{-bt}\right)\right)$
  - V_p the number of cells in the tumor at time t
  - α the specific growth rate (i.e. $V_p^{-1} \cdot dV_p/dt$, expressed in d^{-1})
  - b the exponential decay parameter of the initial growth rate
- Tumor converges to theoretical limit $K = \exp(\alpha b)$ fixed to 10^{-12} cells

Metastatic dissemination

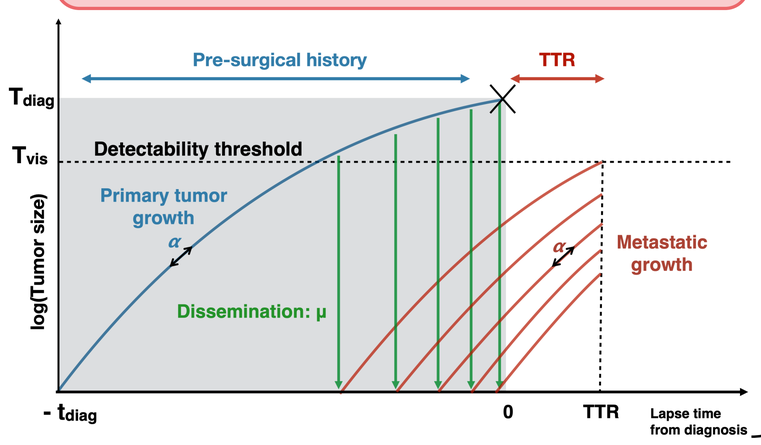
- Instantaneous probability of dissemination μ
- Expressed in $cell^{-1}d^{-1}$
- Assumed to occur at a PT volume-dependent rate $d(V_p)$
  - $d(V_p) = \mu V_p$

Time to metastatic relapse (TTR)

- Detectability threshold V_{detect} (5 mm)
- First metastasis emitted = first to reach visibility threshold
- TTR as a function of V_{diag}, α and μ

Non-linear mixed-effect model

- $\log(Y^i) = \log\left(T_{meta}(V^i_{diag}; \varphi^i)\right) + \varepsilon^i, \qquad \varepsilon^i \sim \mathcal{N}(0, \sigma^2)$
- $\begin{cases} \alpha^i = \alpha_{pop} + \beta_\alpha \cdot C^i + \eta^i_\alpha, & \eta^i_\alpha \sim \mathcal{N}(0, \omega^2_\alpha) \\ \log(\mu^i) = \log(\mu_{pop}) + \beta_\mu \cdot C^i + \eta^i_\mu, & \eta^i_\mu \sim \mathcal{N}(0, \omega^2_\mu) \end{cases}$
  - Y^i the observed (censored) time of metastasis-free survival for patient I
  - V^i_{diag} the size of the tumor at diagnosis
  - C^i the vector of patient covariates
  - $\varphi^i = (\alpha^i, \mu^i)$ the patient's individual parameters
  - η^i the individual random effects



Package METAMATS[2] (R & C++)

Step 1

Base model

- Fit on the data
- Estimation of the population and individual parameters

Likelihood maximization

- saemix package (version >= 3.1)

  OR
- Proprietary C++ optimized implementation of the model

Step 2

Univariate analyses

- Models with only one covariate on only one parameter
- 100 bootstraps
- Wald test (p-values)

Backward selection

- Model containing all covariates found significant in univariate analysis
- Iteratively evaluates the model with one fewer covariate

Covariate model selection

- Covariate model with lowest BIC amongst nested models is selected
- Larger model selected when ties on the minimum (BIC differences < 4)

Step 3

Full model

- Covariate model added
- Fit of the new "full" model
- Re-estimation of the population and individual parameters

Covariates impact

- Weight coefficients
- Relative Standard Errors
- Wald tests (p-values)

4. Results

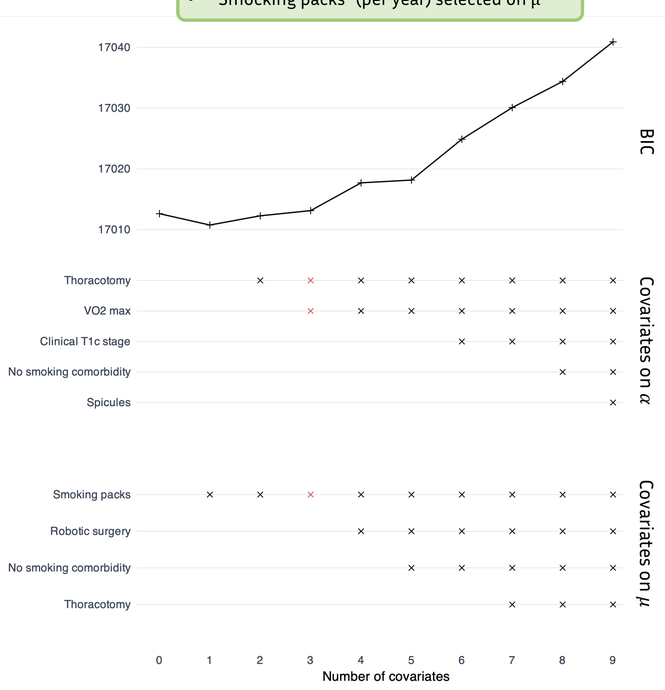
Univariate analysis on the mechanistic parameters

- Significant features
- Note **surgery related features are significant**

	Estimate	RSE (%)	p-value
α			
Clinical T1c stage	1.983	4.764	0.022
No smoking comorbidity	−3.567	3.228	0.035
Spicules	−2.262	7.416	0.039
Thoracotomy	−5.954	4.161	0.012
VO2 max	−5.641	5.602	0.050
μ			
No smoking comorbidity	−11.994	3.932	0.004
Robotic surgery	−8.369	3.679	0.026
Smoking packs	0.070	3.999	0.035
Thoracotomy	−13.060	5.224	0.042

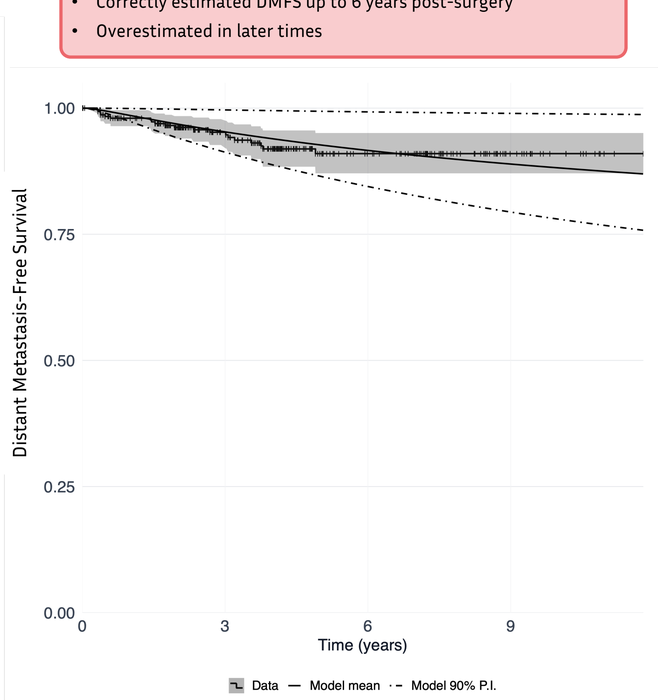
Covariate selection

- "Thoracotomy" and "VO2 max" selected on α
- "Smocking packs" (per year) selected on μ



Time to metastatic relapse (TTR)

- Kaplan-Meier estimate compared to the model predictions
- Correctly estimated DMFS up to 6 years post-surgery
- Overestimated in later times



5. Conclusion

Conclusions

- Metastatic dissemination can be estimated by an intrinsic parameter called μ
- Developed a tool to explore the covariates impact on the mechanistic parameters of the model
- Features such as the surgical approach ("thoracotomy" or the use of robotic surgery) could be **actionable** in practice by clinicians to improve DMFS

Limitations

- Growth law of the primary tumor and the metastasis are the same
- No distinction between the impact of the covariates on the primary tumor α or the metastatic α
- Does not take into account dormancy or post-surgery acceleration

Perspectives

- Explore different growth laws for primary tumor and metastasis
- Add new and more complex metastasis dissemination scheme
- **Individual predictions**
- **Causal inference**

1. Lou, F., Sima, C. S., Rusch, V. W., Jones, D. R., & Huang, J. (2014). Differences in patterns of recurrence in early-stage versus locally advanced non-small cell lung cancer. The Annals of thoracic surgery, 98(5), 1755–1761. https://doi.org/10.1016/j.athoracsur.2014.05.070
2. Klein, C. Parallel progression of primary tumours and metastases. Nat Rev Cancer 9, 302–312 (2009). https://doi.org/10.1038/nrc2627
3. Friberg, S., & Mattson, S. (1997). On the growth rates of human malignant tumors: implications for medical decision making. Journal of surgical oncology, 65(4), 284–297. https://doi.org/10.1002/(sici)1096-9098(199708)65:4<284::aid-jso11>3.0.co;2-2
4. Célestin Bigarré, Alice Daumas, Laurent Greillier, Xavier Muracciole, Laeticia Padovani, et al.. Metamats: A mechanistic software for the simulation, inference and prediction of clinical metastasis. 2025. (hal-04935804v2)