



HAL
open science

Mechanistic modeling of metastatic relapse in early breast cancer to investigate the biological impact of prognostic biomarkers

Célestin Bigarré, François Bertucci, Pascal Finetti, Gaëtan Macgrogan, Xavier Muracciole, Sébastien Benzekry

► **To cite this version:**

Célestin Bigarré, François Bertucci, Pascal Finetti, Gaëtan Macgrogan, Xavier Muracciole, et al.. Mechanistic modeling of metastatic relapse in early breast cancer to investigate the biological impact of prognostic biomarkers. *Computer Methods and Programs in Biomedicine*, 2023, 231, pp.107401. 10.1016/j.cmpb.2023.107401 . hal-04008520

HAL Id: hal-04008520

<https://inria.hal.science/hal-04008520v1>

Submitted on 9 Jan 2024


HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

Mechanistic modeling of metastatic relapse in early breast cancer to investigate the biological impact of prognostic biomarkers

Célestin Bigarré¹ ✉ , François Bertucci^{2,3}, Pascal Finetti², Gaëtan MacGrogan^{4,5}, Xavier Muracciole^{1,6} †, Sébastien Benzekry¹ †

¹COMPUTational pharmacology and clinical Oncology Department, Inria Sophia Antipolis – Méditerranée; ²Predictive Oncology laboratory, Marseille Cancer Research Centre (CRCM), Inserm U1068, CNRS UMR7258, Institut Paoli-Calmettes, Aix-Marseille University, Marseille, France. Equipe labellisée Ligue Nationale Contre Le Cancer.; ³Department of Medical Oncology, CRCM, Institut Paoli-Calmettes, Aix-Marseille University, CNRS, Inserm, Marseille, France; ⁴Department of Biopathology, Institut Bergonié, Regional Comprehensive Cancer Centre, Bordeaux, France; ⁵Inserm U1218, Bordeaux Public Health, University of Bordeaux, Bordeaux, France; ⁶Department of Radiotherapy, Assistance Publique - Hôpitaux de Marseille, Aix Marseille University, Marseille, France.

✉ **For correspondence:**
celestin.bigarre@inria.fr

Contributions: †Senior authors have contributed equally to this work

Present address: COMPO team, Faculté de pharmacie, 27 bd Jean Moulin, 13005 Marseille, FRANCE

Funding: Inria - Inserm “digital health” Phd Grant (C. BIGARRÉ and S. BENZEKRY)

Competing interests: The authors declare no competing interests.

Abstract

Background and Objective

Estimating the risk of metastatic relapse is a major challenge to decide adjuvant treatment options in early-stage breast cancer (eBC). To date, distant metastasis-free survival (DMFS) analysis mainly relies on classical, agnostic, statistical models (e.g., Cox regression). Instead, we propose here to derive mechanistic models of DMFS.

Methods

The present series consisted of eBC patients who did not receive adjuvant systemic therapy from three datasets, composed respectively of 692 (Bergonié Institute), 591 (Paoli-Calmettes Institute, IPC), and 163 (Public Hospital Marseille, AP-HM) patients with routine clinical annotations. The last dataset also contained expression of three non-routine biomarkers. Our mechanistic model of DMFS relies on two mathematical parameters that represent growth (α) and dissemination (μ). We identified their population distributions using the mixed-effects modeling. Critically, we propose a novel variable selection procedure allowing to: (i) identify the association of biological

parameters with either α , μ or both and (ii) generate an optimal candidate model for DMFS prediction.

Results

We found that Ki67 and Thymidine Kinase-1 were associated with α and nodal status and Plasminogen Activator Inhibitor-1 with μ . The predictive performances of the model were excellent in calibration but moderate in discrimination, with c-indices of 0.72 (95% confidence interval [0.48, 0.95], AP-HM), 0.63 ([0.44, 0.83], Bergonié) and 0.60 (95% CI [0.54, 0.80], IPC).

Conclusions

Overall, we demonstrate that our novel method combining mechanistic and advanced statistical modeling is able to unravel the biological roles of clinicopathological parameters from DMFS data.

Introduction

Breast cancer is the most common cancer amongst women and has a high survival probability at 5 years¹. However, 15% of the patients diagnosed with early-stage breast cancer (eBC) will suffer from distant metastatic relapse after surgery, with limited treatment options^{2,3}. Prevention of metastatic relapse is the purpose of adjuvant (post-operative) systemic therapies designed to eradicate the minimal residual disease. Such therapies, which include chemotherapy, and/or hormone therapy in hormone receptor-positive tumor and/or trastuzumab in human epidermal growth factor 2 (HER2)-positive tumors⁴, have substantially improved the metastasis-free and overall survivals⁵⁻⁷. Nevertheless, the clinical outcome of eBC patients is heterogeneous. Current routine prognostic features are mainly age, lymph node status, tumor size and grade, and HR and HER2 statuses.

However, several critical issues remain such as the identification of patients who would have been cured by surgery and radiotherapy alone, thus avoiding the use of toxic chemotherapy⁸. The current relapse risk assessment models are simple regressions based on the above-cited biological parameters (BP). Examples are the Nottingham Prognostic Index⁹ and the PREDICT score^{10,11}. As of today, multiparameter genomic tests with elaborate gene expression signatures (e.g. MammaPrint^{12,13}, Oncotype DX¹⁴ or Endopredict¹⁵) can be used in clinical practice for predicting the clinical course. However, their use remains limited, in part due to their expensive price. More recently, machine learning algorithms have been developed for prognosis¹⁶, but few of them focused on the prediction of breast cancer relapse¹⁷. In addition, these approaches are agnostic and do not rely on biological knowledge.

In contrast, mechanistic models of metastatic development have been developed during the last decades, integrating the pathophysiology of the metastatic process¹⁸. They have been used to estimate the occult metastatic burden at diagnosis after the resection of the primary tumor¹⁹⁻²¹, to predict the impact of individual treatments in pancreatic cancer²² or to describe brain metastasis in lung cancer²³. Following pre-clinical validation^{24,25}, we showed in a previous work²⁶, that a mechanistic approach based on simulation of the metastatic disease could be used to create a predictive tool of breast can-

cer relapse. Building upon this work, we here tested our mechanistic models on three datasets of eBC patients who did not receive any adjuvant systemic therapy. This allowed us to calibrate the models using mixed-effects modeling from data of the natural history of the disease. We show — through a careful univariate analysis — that our model is able to describe the biological links between BP and processes of the metastatic disease. Then we propose a model selection procedure to establish the best covariate structures to use in prediction. Eventually, we establish the predictive performances of the selected models for each dataset.

Materials and methods

Patient datasets

The data consisted of distant metastasis-free survival (DMFS) information and main prognostic variables for patients with operated eBC from three databases. Inclusion criteria were: invasive breast carcinoma, early-stage — here defined as localized (non-metastatic) at the time of diagnosis —, treated with primary surgery followed or not by adjuvant radiotherapy, without any adjuvant systemic therapy (hormone therapy, chemotherapy, trastuzumab), with clinicopathological data and follow-up available for DMFS. The patients who did not experience distant metastatic relapse were censored at the time of death or last follow-up.

The first dataset contained data from 591 women who were treated at the Bergonié institute (Bordeaux, France) between 1989 and 1993. The clinicopathological parameters were: age at the time of diagnosis, pathological tumor size, axillary lymph node status, tumor grade, and expression of estrogen (ER) and progesterone (PR) receptors, HER2 and Ki67, based on immunohistochemistry (IHC) assays. The tumors were considered ER- or PR-positive when more than 1% of the cells showed expression of the corresponding receptor in IHC, Ki67 high when 14% or more of the cells expressed the marker (and Ki67 low otherwise), and HER2-positive when the IHC score was 3+ or 2+ with 60% or more of the cells expressing HER2.

The second dataset included data from 676 patients extracted from our clinically annotated database (8,982 invasive breast cancer samples) made from aggregation of 36 public gene expression datasets²⁷. This dataset included the same clinicopathological annotations as the Bergonié dataset.

The third dataset was composed of 167 patients treated between 1980 and 1990 at the public hospital of Marseille (AP-HM), France. Information on individual DMFS, age status, pathological tumor size, axillary lymph node status and grade were available with the same definition as for the other data sets. Protein concentration information, based on biochemical assays, was available for ER, PR, Urokinase Plasminogen Activator (UPA) and Plasminogen Activator Inhibitor-1 (PAI-1) as well as the enzymatic activity for Thymidine Kinase 1 (TK). Tumors were considered ER- or PR-positive when the concentration of respective proteins was greater than 15 femtomoles per milligram of total proteins.

Missing values could not be imputed uniformly across all three datasets and were thus removed from each dataset: 55 patients were removed from the Bergonié dataset (646 patients initially), 16 from the IPC dataset (692 patients initially), and 7 from the AP HM dataset (174 patients initially).

Mechanistic modeling of the metastatic process

The model has been introduced and extensively described previously²⁶ but we include here a brief description for self-consistency. We consider a simple description of the natural history of eBC, starting at time $t = 0$ with one cell. The primary tumor (PT) grows following a Gompertz law²⁸:

$$V_p(t) = \exp\left(\frac{\alpha}{b}(1 - e^{-bt})\right)$$

where V_p is the number of cells in the tumor at time t , α is the specific growth rate (i.e. $V_p^{-1} \cdot dV_p/dt$, expressed in d^{-1}) at 1 cell and b is the exponential decay parameter of the initial growth rate (unitless)²⁹. These assumptions implicate that the tumor size converges to a theoretical limit $K = \exp(\alpha b)$ cells when $t \rightarrow +\infty$. To avoid over-parametrization and based on the assumption of a lethal tumor burden of 1 kg, we fixed K to 10^{12} cells³⁰ and considered α as the only free parameter for growth (b being computed using the previous equation). At the time of diagnosis (t_{diag}), the patient undergoes a surgery (possibly combined with local adjuvant radiotherapy) considered to entirely remove the PT. To reflect this, the size of the PT for times $t \geq t_{diag}$ is set to zero cells. During the course of the pre-surgical period, we assume that all cells from the primary tumor have an instantaneous probability of dissemination of μ (expressed in $cell^{-1} d^{-1}$). This gives the following continuous expression for the total number of metastases at time t :

$$N_{com}(t) = \mu \int_0^t V_p(s) ds$$

To be consistent with the biological reality, the numerical implementation considers that the number of metastases is an integer, given by:

$$N(t) = \lfloor N_{com}(t) \rfloor,$$

where $\lfloor x \rfloor$ is the integer part of x . New metastases then appear when $N(t)$ increases. The (integer part) of the increment gives the number of metastases that are seeded. Metastases start at a size of one cell and are assumed to follow a Gompertz growth law with the same parameters as for the PT²⁵.

To define the time to distant metastatic relapse (TTR), we considered that metastases are detected as soon as they reach a detectability threshold V_{detect} taken to correspond to a tumor of 5 mm (detection limit in imaging)^{31,32}. From the size V_{diag} of the PT at diagnosis and the growth parameters, we can compute t_{diag} , the time between the initiation of the disease and the diagnosis and τ_{vis} , the time needed for a metastasis to reach V_{detect} ²⁶:

$$t_{diag} = -\frac{\log K}{\alpha} \left(1 - \frac{\log V_{diag}}{\log K}\right) \quad \tau_{vis} = -\frac{\log K}{\alpha} \left(1 - \frac{\log V_{detect}}{\log K}\right)$$

Since the first metastasis emitted will be the first to reach the visibility threshold, the time to relapse is given as a function of V_{diag} and the two mathematical parameters (MP) α and μ (see Figure A).

$$TTR(\alpha, \mu, V_{diag}) = \begin{cases} \tau_{vis} + \arg \min_t \{N(t) \geq 1\} - t_{diag}, & \text{when } N(t_{diag}) \geq 1 \\ +\infty, & \text{otherwise} \end{cases}$$

Statistical mixed-effects population model

Given individual values α^i and μ^i for the i -th patient with observed size of the PT at diagnosis V_{diag}^i , we assumed a log-normal observation error model for the time to metastatic relapse T^i , to ensure positivity,

$$\log(T^i) = \log\left(TTR(\alpha^i, \mu^i; V_{diag}^i)\right) + \varepsilon^i,$$

where $\varepsilon^i \sim \mathcal{N}(0, \sigma^2)$ is the residual error with standard deviation σ . In this expression, T^i is the observed time of metastatic relapse for patient i and $TTR(\alpha^i, \mu^i; V_{diag}^i)$ is the model prediction of the time of metastatic relapse for patient i .

We also assumed a log-normal distribution of the individual MPs, in the population, with a linear effect of the covariates (BP at diagnosis), denoted by the vector C^i :

$$\begin{cases} \log \alpha^i = \log \alpha_{pop} + \beta_\alpha \cdot C^i + \eta_\alpha^i \\ \log \mu^i = \log \mu_{pop} + \beta_\mu \cdot C^i + \eta_\mu^i \end{cases}$$

where α_{pop} and μ_{pop} are the typical values of α and μ in the population, β_α and β_μ are the vectors of the covariate effects, and $\eta^i = (\eta_\alpha^i, \eta_\mu^i)$ are the random-effects, i.e., independent identically distributed random variables with distribution $\mathcal{N}_2(0, \Omega)$. The variance matrix $\Omega = \begin{pmatrix} \omega_\alpha^2 & 0 \\ 0 & \omega_\mu^2 \end{pmatrix}$ quantifies inter-individual variability. Due to identifiability issues in the AP-HM dataset, we removed the inter-individual variability on α (i.e., we fixed $\omega_\alpha = 0$).

From the survival function implicitly defined by the structural error model and using a likelihood definition compatible with censored data, we defined a maximum likelihood estimator for the mixed-effect model (see previous work²⁶ and supplementary text for technical details).

The relative standard errors of the population-level MPs of the models — α_{pop} , μ_{pop} , Ω , and σ^2 , β_α , and β_μ — were obtained by 100-replicates bootstrap and used to assess parametric identifiability.

Variable selection

The definition of the covariate effects allows for each BP to potentially influence the distribution of α , μ , both, or none, depending on the corresponding coefficients β_α and β_μ . To identify the impact of the BPs on the MPs, we used a two-step approach. First, we performed a univariate analysis in which we tested for significant effects in either α or μ , using models including only one covariate on one MP. Specifically, for a BP C_k , we tested for the null hypotheses $H_0 : \beta_{k,\alpha} = 0$ or $\beta_{k,\mu} = 0$. The univariate models were assessed in 100-samples bootstraps. The standard deviation of the bootstrap distributions was used to evaluate the precision of the MP estimation. For each covariate, we tested if the corresponding coefficient was significantly non null with a Wald test using the bootstrap estimate of the coefficient standard deviation.

In the second step, the optimal covariate model for each dataset was selected using a backward elimination procedure based on the Bayesian information criteria (BIC) adapted for the selection of covariates in mixed-effects models³³. Specifically, we started from a model containing all statistically significant covariates for both α and μ . Then we iteratively generated all possible nested models with one covariate less and selected the model with the minimal BIC.

To verify that each β was correctly identifiable, the significance of each covariate coefficient in the selected multivariate model was then re-assessed in the multivariate model with a Wald test, based on 100-samples bootstrap estimation of the standard deviation.

Individual predictions

Individual predictions of survival curves \hat{S}^i were obtained by taking the empirical expectation of the survival function with respect to the inter-individual variability over N_{sim} replicates,

$$\hat{S}^i(t) = \frac{1}{N_{sim}} \sum_{j=1}^{N_{sim}} S\left(t \mid \alpha_j^i, \mu_j^i; V_{diag}^i\right),$$

with α_j^i and μ_j^i sampled from the distribution of α^i and μ^i .

Prediction performance metrics

To assess the prediction performance of the models at a fixed time point, we used calibration plots. These were obtained by predicting the survival probabilities at landmark times t_l for all patients $\hat{S}^i(t_l)$, binning these into 5 quantiles groups and computing the median prediction in each bin. For each bin, this prediction was plotted against the actual data group DMFS at time t_l estimated by the Kaplan-Meier method^{34,35}.

Performance of the models for prediction was also assessed by the concordance index using the individual predicted DMFS probability at 5 years to order the comparable pairs³⁶.

Numerical implementation of the model

The mechanistic model was implemented as an R package with high performance simulation code in C++. All model simulations used a time step of 20 days and were performed with tumor size expressed in number of cells. Tumor diameter data were converted into numbers of cells assuming spherical shape and a cell density of 106 cells per mm³^{37,38}.

Parameter identification was performed using the stochastic approximation of expectation maximization algorithm³⁹ implemented in the saemix R package version 3.0⁴⁰. All computations were performed with R version 4.0.4⁴¹.

Data availability

The data from Institut Bergonié and AP-HM analyzed in this study are not publicly available due to patient privacy requirements but are available upon reasonable request from the corresponding author. The data from IPC is available online and has been previously described²⁷.

Results

Biological parameters discrepancies between the datasets

The distributions of the BPs were different in the three datasets (Table 1). There was significantly more Ki67 high patients in the IPC dataset (chi-squared test, $p < 0.001$) and more node-positive patients in the Bergonié dataset when compared to the IPC dataset ($p < 0.001$) and the AP-HM dataset ($p < 0.001$). The distribution of the grade values was significantly different between the Bergonié and IPC datasets ($p < 0.001$), with more

	Bergonié, N = 591	IPC, N = 676	APHM, N = 167	p-value
Pathological tumor size	15 (12, 20)	20 (15, 25)	20 (15, 26)	<0.001
Age				<0.001
<50	147 (25%)	338 (71%)	35 (21%)	
≥50	444 (75%)	138 (29%)	132 (79%)	
Unknown	0	200	0	
Pathological lymph nodes				<0.001
nodes-	364 (62%)	648 (96%)	147 (88%)	
nodes+	227 (38%)	28 (4.1%)	20 (12%)	
ER status				0.038
ER-	134 (23%)	140 (21%)	50 (30%)	
ER+	457 (77%)	536 (79%)	117 (70%)	
PR status				<0.001
PR-	189 (32%)	271 (40%)	80 (48%)	
PR+	402 (68%)	405 (60%)	87 (52%)	
HER2 status				0.7
HER2-	523 (88%)	594 (88%)		
HER2+	68 (12%)	82 (12%)		
Grade				<0.001
1	182 (31%)	131 (19%)		
2	267 (45%)	298 (44%)		
3	142 (24%)	247 (37%)		
Ki67 status				<0.001
low	390 (66%)	273 (40%)		
high	201 (34%)	403 (60%)		
TK			60 (31, 154)	
UPA			0.79 (0.38, 1.33)	
PAI-1			3.4 (2.0, 5.3)	

Table 1. Patient and disease characteristics. For categorical variables (age, presence of invaded lymph nodes, ER, PR, HER2 and Ki67 statuses, grade), number of patients in each dataset (and proportion of the dataset's values (%)). For continuous variables (pathological tumor size and PAI-1), median value (first quartile – third quartile) in each dataset. P-values correspond to the adequate test of identical distributions in all datasets with available information (chi-squared test for categorical variables, Kruskal-Wallis rank sum test for continuous variables).

low-grade tumors in the Bergonié data and more high-grade tumors in the IPC data. The proportion of ER-positive and PR-positive patients across the three datasets was also significantly different and no statistical difference was found for the distribution of HER2 in the Bergonié and IPC data.

The distribution of primary tumor size appeared log-normal for the three datasets, with a smaller median in the Bergonié dataset than in the IPC set ($p < 0.001$, Brown-Mood median test) whereas no statistical differences in median could be found between the IPC and AP-HM datasets (Supplementary Figure 1). The three datasets also exhibited differences in DMFS (Supplementary Figure 2), with lower DMFS in the IPC dataset than in the Bergonié dataset (Cox regression hazard ratio $HR = 2.2$, $p < 0.001$) or in the AP-HM

dataset ($HR = 1.8, p < 0.001$). No significant difference in DMFS was found between the Bergonié and AP-HM datasets ($HR = 1.3, p = 0.157$).

Inter-individual variability of the mathematical parameters accurately describes distant metastasis-free survival curves

We first used our mechanistic model without covariates (base model) to see if individual variability in the pathological tumor size at diagnosis (included as a direct parameter of the model), associated to log-normal inter-individual variability of the MPs α and μ , was able to describe the observed TTR in the three datasets (1). Figures 1B, 1C, 1D show the descriptive performances of the mechanistic models. For all three models, the model-based population survival curves correctly described the observed DMFS data and remained within the confidence intervals of the Kaplan-Meier estimators, except for a slight underestimation of the DMFS for small times. On the IPC dataset (figure 1D), for times larger than 16 years, the model prediction, while still in the confidence interval of the data, is quite different from the Kaplan-Meier estimator of survival. This can be explained by the small number of patients (19 patients) still at risk after 16 years.

The identifiability of the population parameters for the base models was assessed in a 100-replicate bootstrap assay. Normalized bootstrap densities are shown in Figure 2. Figures 2A-2C depict univariable densities for $\log \alpha_{pop}$, while Figure 2D-2F depict univariable densities for $\log \mu_{pop}$. Joint bootstrap densities for the fixed-effects are displayed in Figure 2G-2I. Similar univariable plots are shown for the residual error σ (Figure 2J-2L), and standard deviation of the random effects ω_α (Figure 2M-2N) and ω_μ (Figure 2N). (2O-2Q). Figure 2R-2S present the joint bootstrap distribution for ω_α and ω_μ . Aside from the slightly bimodal distribution of ω_μ in the Bergonié dataset (Figure 2O and 2R), the bootstrap distributions are unimodal indicating no practical identifiability misspecification in the base models for the IPC and AP-HM datasets and all but one parameter in the Bergonié dataset. Even with the bimodal distribution on ω_μ in the Bergonié dataset, the corresponding relative standard error (RSE) is 50.0% (Table 2) which is classically considered acceptable. The rest of the parameters of the base models were estimated with good accuracy (Table 2). Relative standard errors (RSE) for population values $\log \alpha_{pop}$ and $\log \mu_{pop}$ were all below 15%, and RSE for inter-individual variability parameters ω_α and ω_μ were below 30%. For the residual error σ , RSE were below 15% for the Bergonié and IPC datasets, and below 50% for the AP-HM dataset.

Mechanistic modelling yields biological insight on the impact of biological parameters on metastasis

To study the association of the BPs with either growth or dissemination, we tested all univariate models with effect of a BP as a covariate on either α or μ (Table 3-??). On the Bergonié and IPC datasets, the ER, PR, HER2, and Ki67 statuses had a significant effect on α and μ , and the lymph node status had a significant effect on μ . The grade of the tumor was also significantly associated with both α and μ on the IPC dataset. On the AP-HM dataset, only PAI 1 had a significant effect on μ .

For each dataset, an optimal BP set was selected using a backward elimination procedure starting from the model including all significant BPs (Table 3-??). Model selection was performed based on the BIC to compromise between model performances and number of covariates. For each dataset, the covariate model with the lowest BIC amongst the

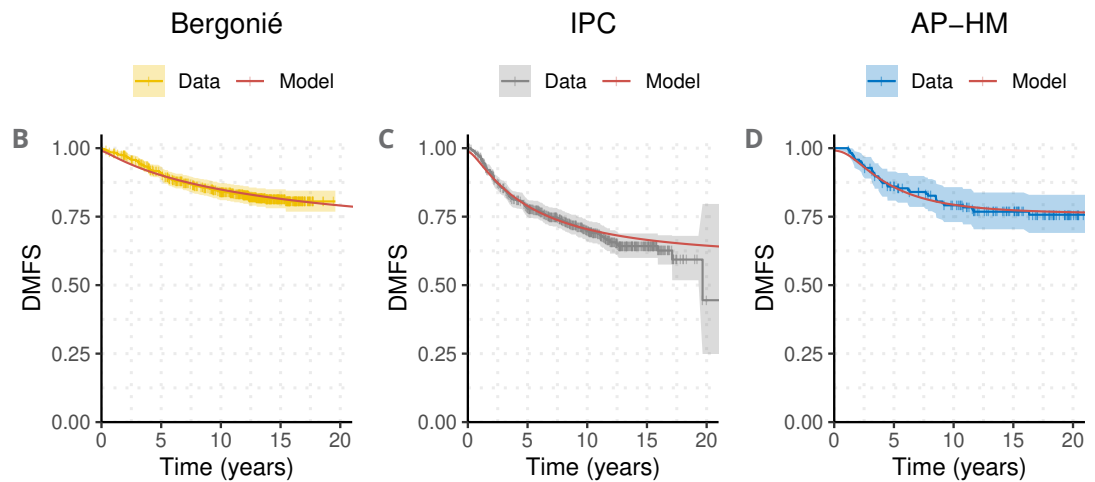
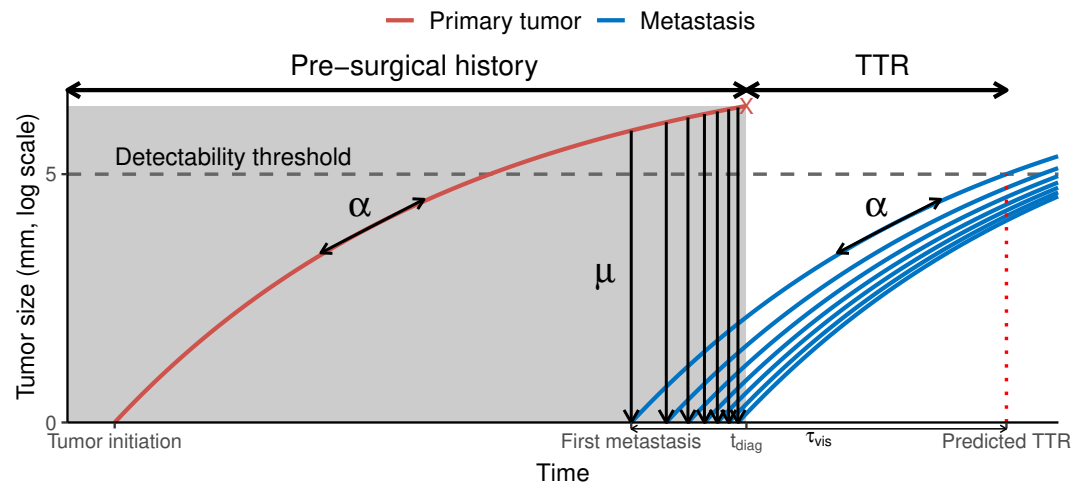


Figure 1. Mechanistic model of the time metastatic relapse. A. Overview of the mechanistic model. The model prediction of the time to metastatic relapse (TTR) is computed from the size of the primary tumor at diagnosis and two mathematical mechanistic parameters, α controlling the growth rate of the primary tumor and metastases, and μ controlling the seeding of new metastases. The scheme shows a unitless simulation of the model. B – D. Fits of the mechanistic model (without covariate effects) on the Bergonié (B), IPC (C) and AP-HM (D) datasets. Each panel presents for one dataset, the DMFS and the model’s prediction of the survival function in the population. The model was trained and evaluated on the full dataset.

nested models was kept as the best model (Figure 3). Ties on the minimum (differences on the BIC < 4) were resolved by choosing the larger model, as the BIC is known to be favor smaller models when compared to other criteria⁴². The final models contained the ER and Ki67 statuses on α for the Bergonié dataset (Figure 3A), a more complex model with effects of the grade and PR on α and the HER2 and nodes statuses on μ for the IPC dataset (Figure 3B), and only PAI 1 on μ for the AP-HM dataset (Figure 3C).

Estimation of the population parameters for the selected models is presented in Table 2. The effect of grade 3 for the IPC dataset were estimated with good precision (RSE 14.9%), whereas the effects of the other variables were estimated with larger but still ac-

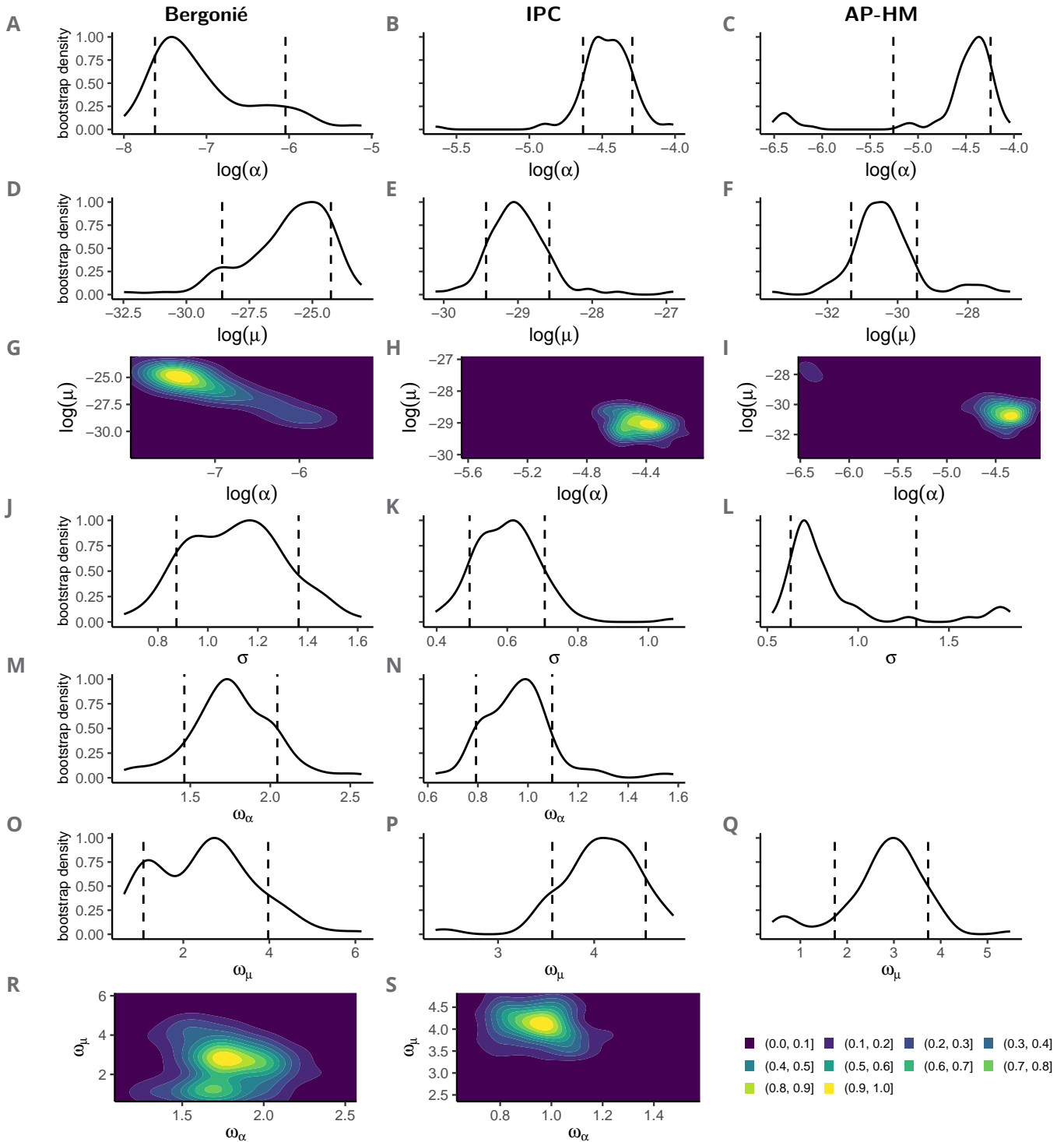


Figure 2. Bootstrap distribution of the population parameters for the base models. Plain lines show the normalized bootstrap densities (scaled to a maximum of 1) obtained by kernel smoothing of the bootstrap histogram. The dashed lines correspond to the bootstrap 5- and 95-percentiles. (A-C) Distribution of $\log \alpha$ for Bergonié (A), IPC (B), and AP-HM (C). (D-F) Distribution of $\log \mu$ for Bergonié (D), IPC (E), and AP-HM (F). (G-I) Joint distribution of $\log \alpha$ and $\log \mu$ for Bergonié (G), IPC (H), and AP-HM (I). (J-L) Distribution of σ for Bergonié (J), IPC (K), and AP-HM (L). (M-N) Distribution of ω_α for Bergonié (M), and IPC (N). (O-Q) Distribution of ω_μ for Bergonié (O), IPC (P), and AP-HM (Q). (R-S) Joint distribution of ω_α and ω_μ for Bergonié (R), and IPC (S).

Dataset		Base model estimation (RSE)	Full model estimation (RSE)
Bergonié	$\log \alpha_{pop}$	-6.85 (9.0%)	-6.7 (19.9%)
	$\log \mu_{pop}$	-25.91 (6.4%)	-29.44 (10.2%)
	σ	1.44 (13.4%)	0.79 (10.5%)
	ω_α	1.19 (21.9%)	1.79 (35.2%)
	ω_μ	2.30 (50.0%)	3.23 (46.4%)
	$\beta_{\alpha,ER}$		-0.89 (58.9%)
	$\beta_{\alpha,KI67}$		1.49 (33.2%)
	IPC	$\log \alpha_{pop}$	-4.59 (4.2%)
$\log \mu_{pop}$		-29.21 (1.5%)	-29.38 (1.3%)
σ		0.96 (10.1%)	0.77 (7.2%)
ω_α		0.51 (28.5%)	0.32 (30.4%)
ω_μ		4.17 (10.2%)	4.20 (7.8%)
$\beta_{\alpha,PR}$			-0.46 (40.6%)
$\beta_{\alpha,grade\ 2}$			0.58 (46.1%)
$\beta_{\alpha,grade\ 3}$			1.66 (14.9%)
$\beta_{\mu,HER2}$			2.06 (30.6%)
$\beta_{\mu,nodes}$			2.81 (45.8%)
AP-HM	$\log \alpha_{pop}$	-4.34 (14.1%)	-4.39 (16.4%)
	$\log \mu_{pop}$	-30.53 (3.4%)	-32.24 (4.0%)
	σ	0.74 (44.1%)	0.67 (63.7%)
	ω_μ	3.05 (29.1%)	3.31 (30.3%)
	$\beta_{\mu,PAI-1}$		0.30 (38.4%)

Table 2. Values of the Parameters. The mathematical mechanistic parameters (α and μ) were assumed to follow a log-normal distribution such that $\log \alpha^i$ and $\log \mu^i$ are gaussian with respective mean $\log \alpha_{pop}$ and $\log \mu_{pop}$, and respective standard deviation ω_α and ω_μ . For each dataset, the base models correspond to the case with no BP effect (aside from the pathological tumor size). The full models correspond to the best models from the selection procedure, where conditionally to the vector of covariates (C^i) included in the model, the mathematical parameters followed a log-normal distribution such that, $\log \alpha^i$ and $\log \mu^i$ are gaussian with respective mean $\log \alpha_{pop} + \beta_\alpha \cdot C^i$ and $\log \mu_{pop} + \beta_\mu \cdot C^i$ (where β_α and β_μ are vectors of the BP specific coefficients) and respective standard deviation ω_α and ω_μ . For all models, the log-residual error on time to distant metastatic relapse was assumed to follow a centered gaussian distribution with variance σ^2 . Estimation was performed using the stochastic approximation of expectation maximization algorithm. Relative standard errors (RSE) were computed from a 100 replicates bootstrap.

ceptable uncertainty ($RSE < 50\%$). Only the effect of the ER status on α for the Bergonié data was estimated with high RSE (58.9%). In the final models, all BPs were significant in multivariate analysis as covariates (Wald test). The coefficients of all BPs were in the same order of magnitude as the corresponding inter-individual variability standard deviation, confirming that the estimated coefficients have significant impact on the individual parameter distributions.

Covariate models accurately describe DMFS curves

To assess the ability of the models to describe different subpopulations of patients, we stratified in each dataset by the BP included in the complete model and we analyzed

	Bergonié			IPC			AP-HM			
	β	R.S.E.	p-value	β	R.S.E.	p-value	β	R.S.E.	p-value	
α	Age \geq 50	-0.02	1	0.5			0.01	90	1	
	nodes+	-0.5	0.2	0.2	0.4	1	0.3	3	0.8	
	ER+	-2	0.3	0.0002	-0.8	0.3	0.0003	-0.6	0.9	0.3
	PR+	-2	0.3	8×10^{-5}	-1	0.3	0.01	-0.4	2	0.6
	HER2+	2	0.3	0.0009	-0.9	0.5	0.05			
	Grade 2	-0.4	0.9	0.2	-0.8	0.3	0.001			
	Grade 3	0.7	0.6	0.1	1	0.2	5×10^{-10}			
	Ki67 high	2	0.2	3×10^{-7}	0.9	0.4	0.006			
	TK							0.08	0.9	0.3
	UPA							0.2	0.9	0.3
Pai-1							0.08	1	0.4	
μ	Age \geq 50	-0.02	2	0.5			-0.05	0.7	0.2	
	nodes+	-2	0.4	0.03	3	0.4	0.02	-0.7	1	0.5
	ER+	-4	0.2	2×10^{-6}	-1	0.4	0.01	-0.5	2	0.5
	PR+	-3	0.3	8×10^{-5}	-2	0.2	1×10^{-5}	-1	0.6	0.08
	HER2+	4	0.3	0.0006	3	0.2	1×10^{-6}			
	Grade 2	-1	0.6	0.08	-0.9	0.5	0.06			
	Grade 3	0.7	1	0.4	2	0.2	5×10^{-5}			
	Ki67 high	3	0.2	2×10^{-5}	1	0.4	0.007			
	TK							-0.0004	4	0.8
	UPA							0.5	0.6	0.1
Pai-1							0.3	0.4	0.006	

Table 3. Univariate effects of the BP The univariate models assumed a log-normal distribution independent of the covariate for one of the two computational biomarkers (α or μ), and a log-normal conditional distribution with respect for the covariate for the other computational biomarker, with a median equal to the sum of a typical population value, and of the weighted covariate value. The estimation of the weight coefficient (beta), the relative standard error (R.S.E.) obtained by bootstrap with 100 repetitions, as well as the p-values for the corresponding Wald test, are presented in the table for all possible univariate models.

the model predicted DMFS for each subgroup of patients (Figure 4). On the Bergonié dataset, the model captured well the DMFS difference between the ER-positive and ER-negative patients and between the Ki67 high and Ki67 low patients (Figure 4A-4B). On the IPC dataset, the model was able to correctly describe the differences of DMFS when comparing patients with different grades, HER2, based on the presence of invaded lymph nodes, or between PR-positive and PR-negative patients (Figure 4C-4F). For the AP-HM dataset, the model captured very well the DMFS for patient in the upper PAI 1 tercile, was adequate for the second tercile but seemed to slightly overestimate DMFS for larger times, and the DMFS for the lower tercile of PAI 1 was underestimated by the model (Figure 4G).

Predictive performances

Next, we evaluated the prediction performances of the models. Calibration curves indicated excellent individual predictive power for all three models at the 5-, 10- and 15-years landmark times (Figure 5). Calibrations at 5 years (Figure 5A-5C) and 10 years (Figure 5D-5F) were good for all three models and for all DMFS probabilities, with a slight trend

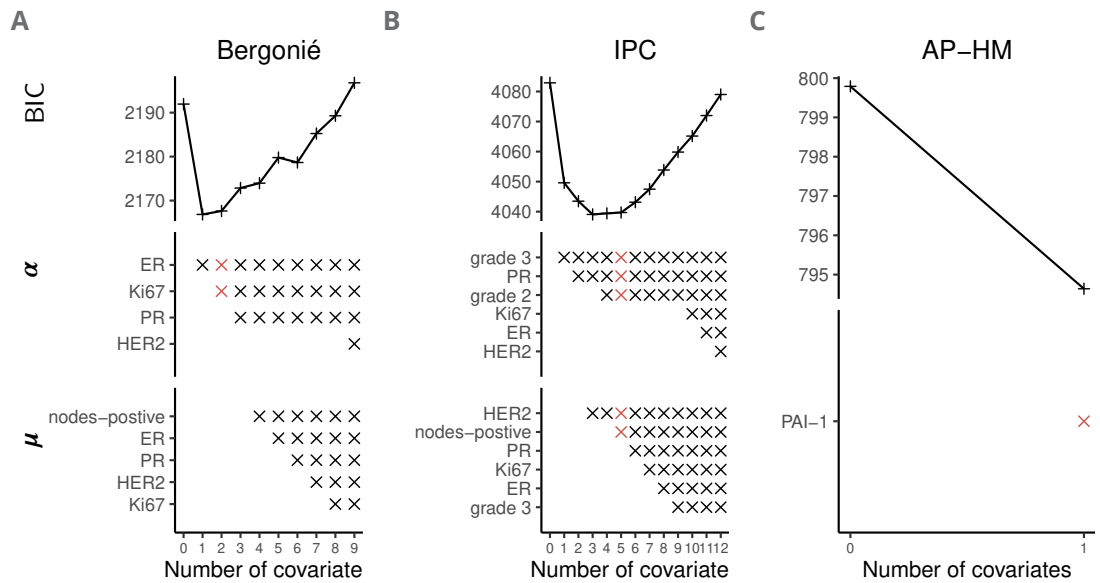


Figure 3. Biological parameters selection For each panel, the top line presents the evolution of the Bayesian information criteria (BIC) as a function of the number of parameters during the backward elimination procedure. Starting from the model containing all the covariates with a significant effect in univariate analysis, the selection procedure iteratively tested all models with one parameter less, keeping at each step the one with minimal BIC. The bottom lines show the evolution of the covariate model during the selection process. Backward selection on the Bergonié (A), IPC (B) and AP-HM, (C) data.

to overestimate DMFS probability at 10 years for the higher probability group. In the Bergonié dataset, the model slightly underestimated DMFS at 15 years for smaller probabilities while the over-estimation of higher probabilities was more pronounced (Figure 5G). Whereas for the IPC (5H) and AP-HM (Figure 5I) datasets, the calibration at 15 years was still good.

We then computed Harrell's concordance index (c-index) as another measure of the prediction performance³⁶. The standard deviation for the c-index was computed in 10-folds (8-folds for the AP-HM dataset) cross-validation. The performances were modest with a c-index of 0.63 (95% confidence interval (CI) [0.44–0.83]) in the Bergonié dataset, 0.71 (95% CI [0.42, 0.99]), 0.67 (95% CI [0.54, 0.80]) in the IPC dataset, and 0.72 (95% CI [0.48, 0.95]) in the AP-HM dataset.

Discussion

Classical statistical models of metastatic risk, although able to detect correlations between biomarkers and outcome, fail to give causal insights about the mechanisms at stake. For prediction, several genomics-based prognostic tools are commercially available (Oncotype DX Recurrence Score, Prosigna Risk of Recurrence score, EndoPredict, Breast Cancer Index) for estimation of the recurrence risk in HR-positive and HER2-negative eBC. However, the cost of these tests limits their clinical use and, again, phenotypic insights are limited. Our approach helps to provide mechanistic information from routine clinical markers available at diagnosis.

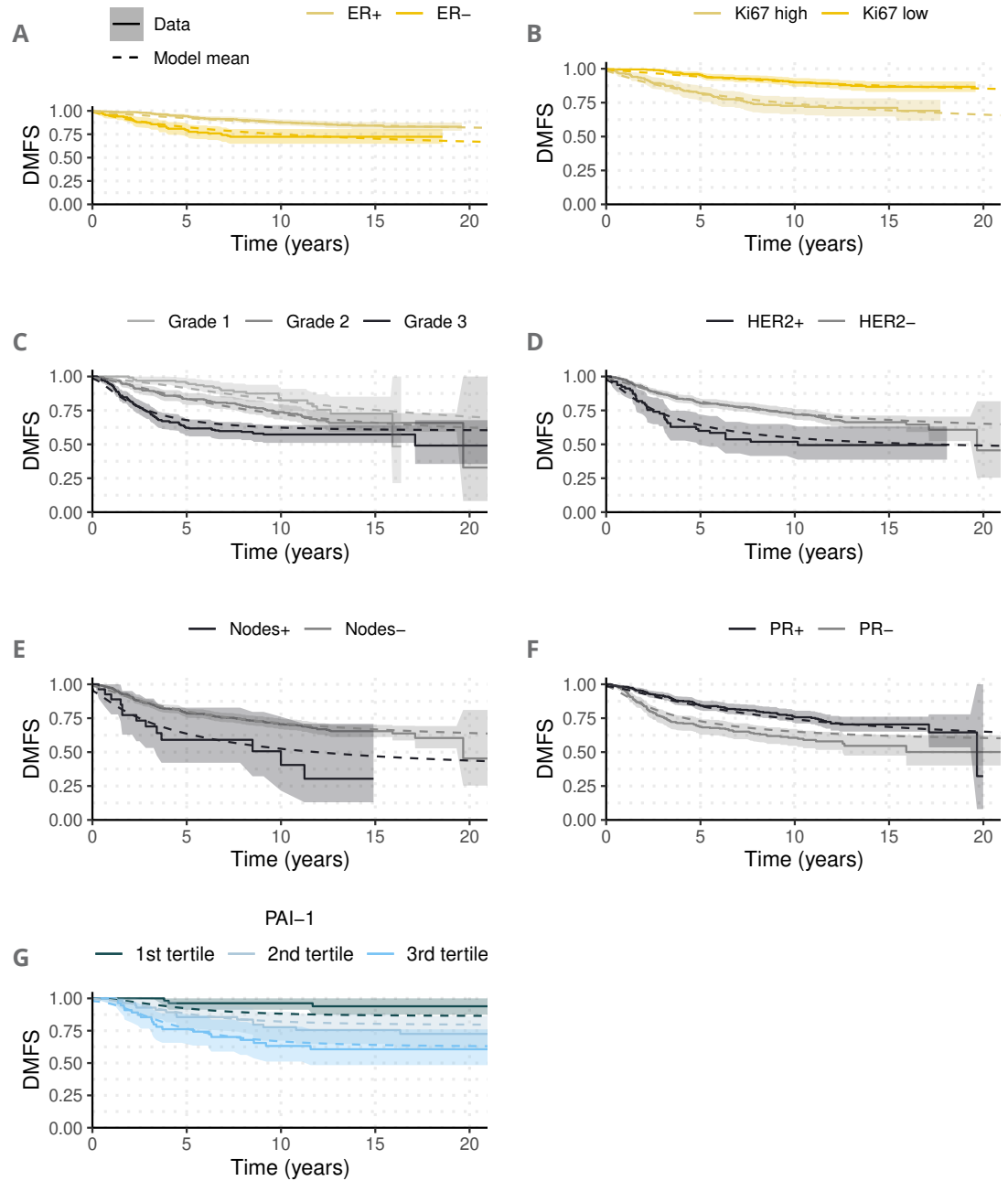


Figure 4. Model predictions in stratified groups. Group-comparison of the distant metastasis-free survival data (Kaplan-Meier estimate, solid line and 95% confidence interval, colored band) and the model mean prediction of the metastasis-free survival function. Bergonié dataset, patients stratified on the ER (A) or Ki67 (B) status. Paoli-Calmettes Institutes (IPC) dataset, patients stratified by grade (C), HER2 (D) status, invaded lymph nodes (E) or PR (F) status. G. APMH dataset, patients stratified by PAI-1 tertiles.

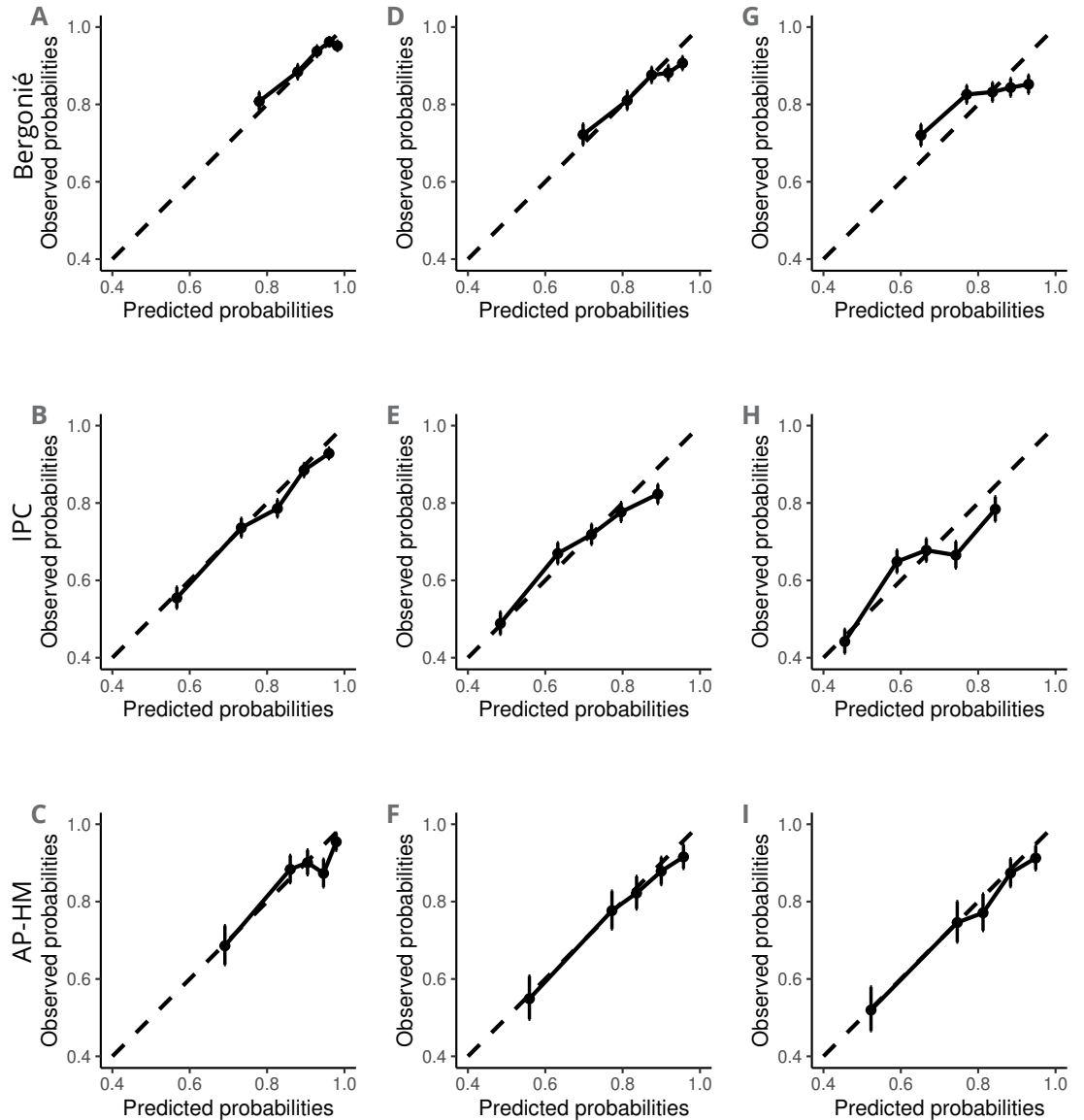


Figure 5. Calibration curves. At a fixed time-point, cross-validations predictions of the distant metastasis-free survival (DMFS) were binned into 5 quantile groups. The median prediction and 95% confidence interval of each group is compared to the Kaplan-Meier estimate of the group DMFS at the specified time. The identity (dashed line) is indicated for comparison to perfect prediction. (A-C): calibration curves at 5 years for the selected model on the respectively the Bergonié, IPC and AP-HM datasets (D-F): calibration curves at 10 years for the selected models (G-I), calibration curves at 15 years for the selected models.

We used a simple mechanistic model of metastatic development based on two processes – growth and dissemination – to analyze DMFS data from three different datasets of eBC patients, two of which contained only routinely available data, whereas the third contained also non-routine markers (UPA, PAI-1 and TK), known to have biological roles in the metastatic process⁴³. We not only correctly described the DMFS in the three populations, but also showed that our method could be used to link biological features with specific parts of the metastatic process.

First, primary tumor size is directly incorporated as a central variable in the mathematical model. Then, studying the effect of the BPs on the population distribution of the growth parameter α and the dissemination parameter μ allowed us to associate each BP with one (or both) aspect(s) of the metastatic process. In order to gather as many patients as possible and to closely match information easily available in routine clinical care, we focused on basic clinical and biological markers. For μ based on the data from the AP-HM cohort, our model supports the association of protease inhibitor PAI 1 with metastatic dissemination potential, consistent with previous preclinical⁴⁴ and clinical⁴³ studies. On the two other datasets (Bergonié and IPC), the model specifically associated the presence of invaded lymph nodes with μ . For α the model pinpointed the Ki-67 marker as an important predictor of the growth parameter α and as the least important predictor of the dissemination parameter μ (Bergonié dataset), in accordance with the established biological role of Ki-67 as a marker of cell division and tumor proliferation⁴⁵. The hormone receptor statuses were relevant in α for both the IPC (where PR status was selected) and the Bergonié (where ER status was selected) datasets. Their effect on μ was less clear, since no hormone receptor status was selected on μ but PR status persisted in μ up until late steps of the elimination process. Overall, these findings support the ability of our mechanistic model to identify the biological role of specific markers.

In this regard, our work had to account for the following limitations. The differences in DMFS within each dataset (Figure 1B), the measurement methods (IHC staining for Bergonié, mRNA expression for IPC, and protein concentration for AP-HM) or variable availability prevented the possibility to properly compare the results across datasets. In addition, in order to ensure identifiability of the model from the available data, multiple simplifying hypotheses had to be assumed. First, we assumed that the initial local treatment (surgery \pm adjuvant radiotherapy) left no remaining cell in the tumor bed, thus preventing post-surgery emergence of new metastasis. Second, based on multiple biological findings stating that primary and secondary tumors have comparable growth rates^{30,46,47}, we assumed that all metastases follow the same growth law. Relaxing this assumption would require one or two additional MP which would impair practical identifiability of the model. Third, given the limited data on the tumor sizes (only the PT diameter), we had to assume spherical shape and uniform cell density in the tumors. While this is a common assumption in the modeling literature, integrating imaging (e.g., MRI) data could provide quantitative data about intra-tumor cell densities and be used to refine the modeling⁴⁸.

The prediction performances were mitigated, with very good performances in calibration at various time points, but surprisingly low c-index values. The c-index computation may not be very accurate considering that the cross-validation sampling was not stratified on the event indicator variable, meaning that the number of events and thus the number of comparable pairs for the c-index computation was probably different be-

tween the cross-validation folds. Individual-level predictive power still needs further investigation to be up-part with the existing agnostic models. In particular a larger and more homogenous dataset with more patients and raw values of the markers instead of dichotomized categories could give better results. However, prediction was not the main aim of our study. We rather wanted to develop biologically-based, mechanistic models of metastasis tailored for DMFS data and assess their ability to bring causal insights between biomarkers at surgery and DMFS.

Having established the relevance of our model to identify causal and mechanistic relationships between BPs and DMFS, future perspectives are to integrate more complex data into the model, possibly allowing to refine the mechanisms included in the current version (parsimonious here to ensure identifiability). This includes large-dimensional genomic⁴⁹, transcriptomic data and/or established gene expression signatures (e.g., Onco-type)¹³ imaging data⁴⁷ or digital pathology⁵⁰. In addition, the dynamic structure of our model is particularly well tailored to account for systemic therapy, either neo-adjuvant, adjuvant or in the metastatic setting. For the first, our model could integrate on-treatment data (from, e.g., liquid biopsies) that could offer surrogate markers to pathologic complete response in order to predict long-term outcome⁵¹, or potentiate these “window-of-opportunity” trials to identify predictive biomarkers of response⁵². In the adjuvant and metastatic settings, further modeling efforts are needed to account for the impact of systemic therapy on (possibly occult) metastases. Identification of treatment related, patient-specific MPs would give a relevant tool to personalize the treatment, e.g., to determine and reduce the number of chemotherapy cycles necessary to avoid distant relapse. Further evaluation in model-based prospective trials based on specific guidelines⁵³ and including comparisons with transcriptomic tests will be required before translation to clinical care⁵⁴.

References

1. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer Statistics, 2022. *CA: A Cancer Journal for Clinicians*, 7–33. ISSN: 1542-4863. doi:[10.3322 / caac . 21708](https://doi.org/10.3322/caac.21708). <https://onlinelibrary.wiley.com/doi/abs/10.3322/caac.21708> (2022).
2. Holleczeck, B., Stegmaier, C., Radosa, J. C., Solomayer, E.-F. & Brenner, H. Risk of Local-Regional Recurrence and Distant Metastases of Patients with Invasive Breast Cancer up to Ten Years after Diagnosis – Results from a Registry-Based Study from Germany. *BMC Cancer*, 520. doi:[10.1186/s12885-019-5710-5](https://doi.org/10.1186/s12885-019-5710-5). pmid: [31146706](https://pubmed.ncbi.nlm.nih.gov/31146706/) (2019).
3. Steeg, P. S. Targeting Metastasis. *Nature Reviews Cancer*, 201–218. ISSN: 1474-175X. doi:[10.1038/nrc.2016.25](https://doi.org/10.1038/nrc.2016.25). pmid: [27009393](https://pubmed.ncbi.nlm.nih.gov/27009393/) (2016).
4. Cardoso, F., Kyriakides, S., *et al.* Early Breast Cancer: ESMO Clinical Practice Guidelines for Diagnosis, Treatment and Follow-Up. *Annals of Oncology*, 1194–1220. ISSN: 0923-7534. doi:[10.1093/annonc/mdz173](https://doi.org/10.1093/annonc/mdz173). pmid: [31161190](https://pubmed.ncbi.nlm.nih.gov/31161190/) (2019).
5. Cameron, D., Piccart-Gebhart, M. J., *et al.* 11 Years' Follow-up of Trastuzumab after Adjuvant Chemotherapy in HER2-positive Early Breast Cancer: Final Analysis of the HERceptin Adjuvant (HERA) Trial. *The Lancet*, 1195–1205. ISSN: 0140-6736. doi:[10.1016/s0140-6736\(16\)32616-2](https://doi.org/10.1016/s0140-6736(16)32616-2). pmid: [28215665](https://pubmed.ncbi.nlm.nih.gov/28215665/) (2017).

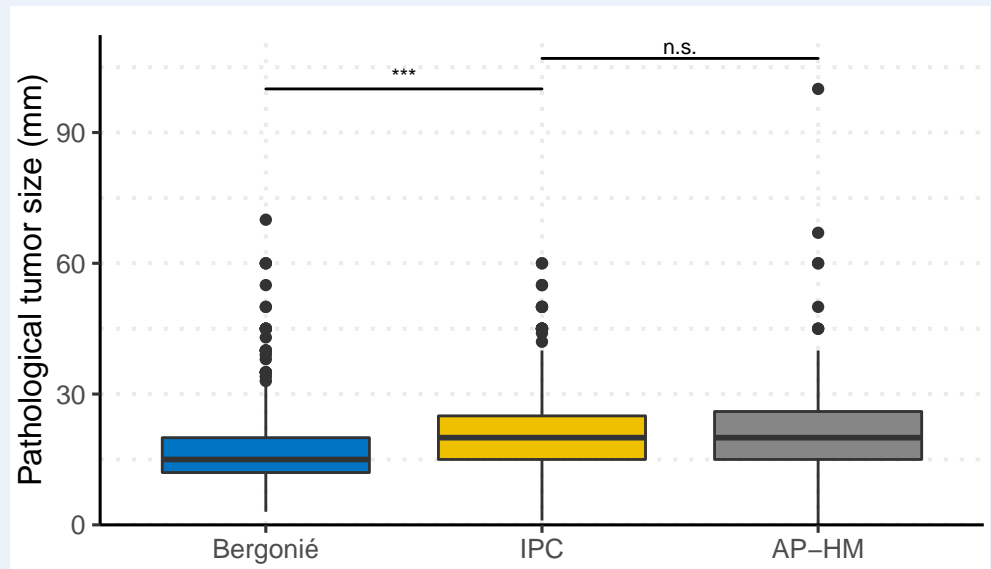
6. (EBCTCG), E. B. C. T. C. G., Peto, R., *et al.* Comparisons between Different Polychemotherapy Regimens for Early Breast Cancer: Meta-Analyses of Long-Term Outcome among 100 000 Women in 123 Randomised Trials. *Lancet*, 432–444. ISSN: 0140-6736. doi:[10.1016/s0140-6736\(11\)61625-5](https://doi.org/10.1016/s0140-6736(11)61625-5). pmid: [22152853](https://pubmed.ncbi.nlm.nih.gov/22152853/) (2012).
7. Organisation, N. A. T. Controlled Trial of Tamoxifen as Adjuvant Agent in Management of Early Breast Cancer Interim Analysis at Four Years. *The Lancet*, 257–261. ISSN: 0140-6736. doi:[10.1016/s0140-6736\(83\)91683-5](https://doi.org/10.1016/s0140-6736(83)91683-5) (1983).
8. Pondé, N. F., Zardavas, D. & Piccart, M. Progress in Adjuvant Systemic Therapy for Breast Cancer. *Nature Reviews Clinical Oncology*, 27–44. ISSN: 1759-4774. doi:[10.1038/s41571-018-0089-9](https://doi.org/10.1038/s41571-018-0089-9). pmid: [30206303](https://pubmed.ncbi.nlm.nih.gov/30206303/) (2019).
9. Galea, M. H., Blamey, R. W., Elston, C. E. & Ellis, I. O. The Nottingham Prognostic Index in Primary Breast Cancer. *Breast Cancer Research and Treatment*, 207–219. ISSN: 0167-6806. doi:[10.1007/bf01840834](https://doi.org/10.1007/bf01840834). pmid: [1391987](https://pubmed.ncbi.nlm.nih.gov/1391987/) (1992).
10. Wishart, G., Bajdik, C., *et al.* A Population-Based Validation of the Prognostic Model PREDICT for Early Breast Cancer. *European Journal of Surgical Oncology (EJSO)*, 411–417. ISSN: 0748-7983. doi:[10.1016/j.ejso.2011.02.001](https://doi.org/10.1016/j.ejso.2011.02.001). pmid: [21371853](https://pubmed.ncbi.nlm.nih.gov/21371853/) (2011).
11. Wishart, G. C., Azzato, E. M., *et al.* PREDICT: A New UK Prognostic Model That Predicts Survival Following Surgery for Invasive Breast Cancer. *Breast Cancer Research*, R1. ISSN: 1465-5411. doi:[10.1186/bcr2464](https://doi.org/10.1186/bcr2464). pmid: [20053270](https://pubmed.ncbi.nlm.nih.gov/20053270/) (2010).
12. Buyse, M., Loi, S., *et al.* Validation and Clinical Utility of a 70-Gene Prognostic Signature for Women With Node-Negative Breast Cancer. *JNCI: Journal of the National Cancer Institute*, 1183–1192. ISSN: 0027-8874. doi:[10.1093/jnci/djj329](https://doi.org/10.1093/jnci/djj329). pmid: [16954471](https://pubmed.ncbi.nlm.nih.gov/16954471/) (2006).
13. Van 't Veer, L. J., Dai, H., *et al.* Gene Expression Profiling Predicts Clinical Outcome of Breast Cancer. *Nature*, 530–536. ISSN: 0028-0836. doi:[10.1038/415530a](https://doi.org/10.1038/415530a). pmid: [11823860](https://pubmed.ncbi.nlm.nih.gov/11823860/) (2002).
14. Sparano, J. A. Gene Expression Assays in Early-Stage Breast Cancer. *Oncology (Williston Park, N.Y.)* ISSN: 0890-9091. pmid: [30334241](https://pubmed.ncbi.nlm.nih.gov/30334241/) (2018).
15. Gradishar, W. J., Robson, M. E., *et al.* Predicting Expected Absolute Chemotherapy Treatment Benefit in Women with Early-Stage Breast Cancer Using a 12-Gene Expression Assay. *Journal of Clinical Oncology*, 525–525. ISSN: 0732-183X. doi:[10.1200/jco.2018.36.15_suppl.525](https://doi.org/10.1200/jco.2018.36.15_suppl.525) (15_suppl 2018).
16. Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V. & Fotiadis, D. I. Machine Learning Applications in Cancer Prognosis and Prediction. *Computational and Structural Biotechnology Journal*, 8–17. ISSN: 2001-0370. doi:[10.1016/j.csbj.2014.11.005](https://doi.org/10.1016/j.csbj.2014.11.005). pmid: [25750696](https://pubmed.ncbi.nlm.nih.gov/25750696/) (2015).
17. Kim, W., Kim, K. S., *et al.* Development of Novel Breast Cancer Recurrence Prediction Model Using Support Vector Machine. *Journal of Breast Cancer*, 230–238. ISSN: 1738-6756. doi:[10.4048/jbc.2012.15.2.230](https://doi.org/10.4048/jbc.2012.15.2.230). pmid: [22807942](https://pubmed.ncbi.nlm.nih.gov/22807942/) (2012).
18. Benzekry, S. Artificial Intelligence and Mechanistic Modeling for Clinical Decision Making in Oncology. *Clinical Pharmacology & Therapeutics*, 471–486. ISSN: 0009-9236. doi:[10.1002/cpt.1951](https://doi.org/10.1002/cpt.1951). pmid: [32557598](https://pubmed.ncbi.nlm.nih.gov/32557598/) (2020).

19. Hanin, L. & Rose, J. Suppression of Metastasis by Primary Tumor and Acceleration of Metastasis Following Primary Tumor Resection: A Natural Law? *Bulletin of Mathematical Biology*, 519–539. ISSN: 0092-8240. doi:[10.1007/s11538-017-0388-9](https://doi.org/10.1007/s11538-017-0388-9). pmid: [29302774](https://pubmed.ncbi.nlm.nih.gov/29302774/) (2018).
20. Iwata, K., Kawasaki, K. & Shigesada, N. A Dynamical Model for the Growth and Size Distribution of Multiple Metastatic Tumors. *Journal of Theoretical Biology*, 177–186. ISSN: 0022-5193. doi:[10.1006/jtbi.2000.1075](https://doi.org/10.1006/jtbi.2000.1075). pmid: [10704301](https://pubmed.ncbi.nlm.nih.gov/10704301/) (2000).
21. Koscielny, S., Tubiana, M. & Valleron, A. J. A Simulation Model of the Natural History of Human Breast Cancer. *British Journal of Cancer*, 515–524. ISSN: 1532-1827. doi:[10.1038/bjc.1985.222](https://doi.org/10.1038/bjc.1985.222). <https://doi.org/10.1038/bjc.1985.222> (1985).
22. Haeno, H., Gonen, M., *et al.* Computational Modeling of Pancreatic Cancer Reveals Kinetics of Metastasis Suggesting Optimum Treatment Strategies. *Cell*, 362–375. ISSN: 0092-8674. doi:[10.1016/j.cell.2011.11.060](https://doi.org/10.1016/j.cell.2011.11.060). pmid: [22265421](https://pubmed.ncbi.nlm.nih.gov/22265421/) (2012).
23. Bilous, M., Serdjebi, C., *et al.* Quantitative Mathematical Modeling of Clinical Brain Metastasis Dynamics in Non-Small Cell Lung Cancer. *Scientific Reports*, 13018. ISSN: 2045-2322. doi:[10.1038/s41598-019-49407-3](https://doi.org/10.1038/s41598-019-49407-3). <https://www.nature.com/articles/s41598-019-49407-3> (1 2019).
24. Baratchart, E., Benzekry, S., *et al.* Computational Modelling of Metastasis Development in Renal Cell Carcinoma. *PLoS Computational Biology*, e1004626. ISSN: 1553-734X. doi:[10.1371/journal.pcbi.1004626](https://doi.org/10.1371/journal.pcbi.1004626). pmid: [26599078](https://pubmed.ncbi.nlm.nih.gov/26599078/) (2015).
25. Benzekry, S., Tracz, A., *et al.* Modeling Spontaneous Metastasis Following Surgery: An In Vivo-In Silico Approach. *Cancer Research*, 535–547. ISSN: 0008-5472. doi:[10.1158/0008-5472.can-15-1389](https://doi.org/10.1158/0008-5472.can-15-1389). pmid: [26511632](https://pubmed.ncbi.nlm.nih.gov/26511632/) (2016).
26. Nicolò, C., Périer, C., *et al.* Machine Learning and Mechanistic Modeling for Prediction of Metastatic Relapse in Early-Stage Breast Cancer. *JCO Clinical Cancer Informatics*, 259–274. doi:[10.1200/cci.19.00133](https://doi.org/10.1200/cci.19.00133). pmid: [32213092](https://pubmed.ncbi.nlm.nih.gov/32213092/) (2020).
27. De Nonneville, A., Finetti, P., Mamessier, E. & Bertucci, F. RE: NDRG1 in Aggressive Breast Cancer Progression and Brain Metastasis. *JNCI: Journal of the National Cancer Institute*, djac031–. ISSN: 0027-8874. doi:[10.1093/jnci/djac031](https://doi.org/10.1093/jnci/djac031). pmid: [35148398](https://pubmed.ncbi.nlm.nih.gov/35148398/) (2022).
28. Norton, L. A Gompertzian Model of Human Breast Cancer Growth1. *Cancer Research*, 7067–7071. ISSN: 0008-5472 (24_Part_1 1988).
29. Vaghi, C., Rodallec, A., *et al.* Population Modeling of Tumor Growth Curves and the Reduced Gompertz Model Improve Prediction of the Age of Experimental Tumors. *PLoS Computational Biology*, e1007178. ISSN: 1553-734X. doi:[10.1371/journal.pcbi.1007178](https://doi.org/10.1371/journal.pcbi.1007178). pmid: [32097421](https://pubmed.ncbi.nlm.nih.gov/32097421/) (2020).
30. Klein, C. A. Parallel Progression of Primary Tumours and Metastases. *Nature Reviews Cancer*, 302–312. ISSN: 1474-175X. doi:[10.1038/nrc2627](https://doi.org/10.1038/nrc2627). pmid: [19308069](https://pubmed.ncbi.nlm.nih.gov/19308069/) (2009).
31. Kundel, H. L. Predictive Value and Threshold Detectability of Lung Tumors. *Radiology*, 25–29. ISSN: 0033-8419. doi:[10.1148/radiology.139.1.7208937](https://doi.org/10.1148/radiology.139.1.7208937). pmid: [7208937](https://pubmed.ncbi.nlm.nih.gov/7208937/) (1981).
32. MacMahon, H., Austin, J. H. M., *et al.* Guidelines for Management of Small Pulmonary Nodules Detected on CT Scans: A Statement from the Fleischner Society. *Radiology*, 395–400. ISSN: 0033-8419. doi:[10.1148/radiol.2372041887](https://doi.org/10.1148/radiol.2372041887). <https://pubs.rsna.org/doi/10.1148/radiol.2372041887> (2005).

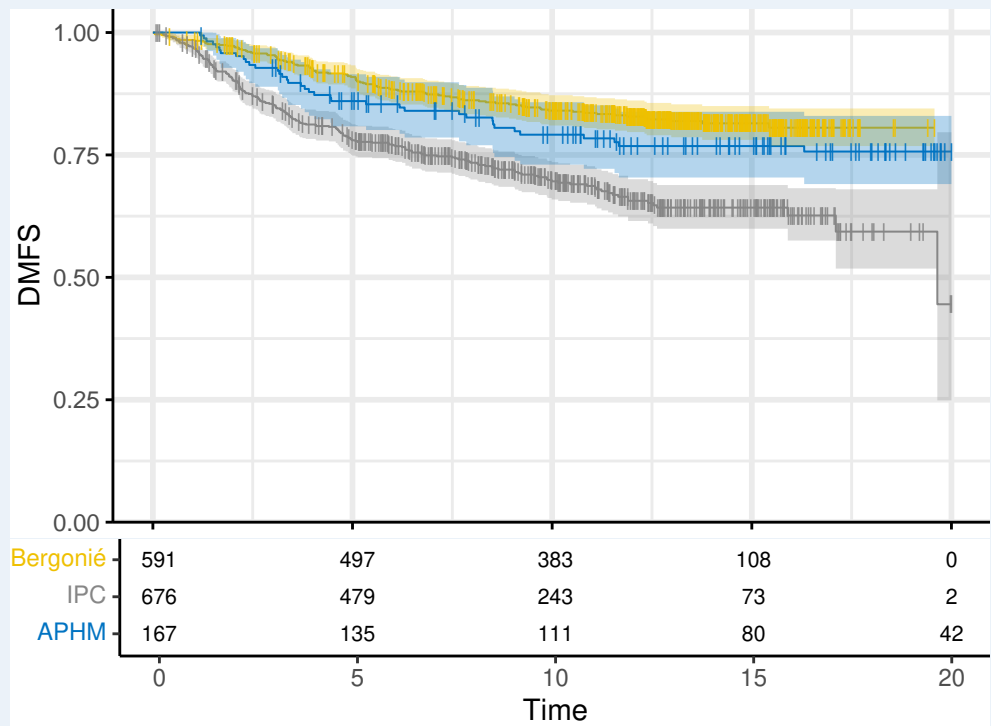
33. Delattre, M., Lavielle, M. & Poursat, M.-A. A Note on BIC in Mixed-Effects Models. *Electronic Journal of Statistics*. ISSN: 1935-7524. doi:[10.1214/14-ejs890](https://doi.org/10.1214/14-ejs890) (2014).
34. Austin, P. C., Harrell, F. E. & Klaveren, D. Graphical Calibration Curves and the Integrated Calibration Index (ICI) for Survival Models. *Statistics in Medicine*, 2714–2742. ISSN: 0277-6715. doi:[10.1002/sim.8570](https://doi.org/10.1002/sim.8570). pmid: [32548928](https://pubmed.ncbi.nlm.nih.gov/32548928/) (2020).
35. Kaplan, E. L. & Meier, P. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, 457–481. ISSN: 0162-1459. doi:[10.2307/2281868](https://doi.org/10.2307/2281868). JSTOR: [2281868](https://www.jstor.org/stable/2281868) (1958).
36. Harrell, F. E., Califf, R. M., Pryor, D. B., Lee, K. L. & Rosati, R. A. Evaluating the Yield of Medical Tests. *JAMA*, 2543–2546. ISSN: 0098-7484. doi:[10.1001/jama.1982.03320430047030](https://doi.org/10.1001/jama.1982.03320430047030). pmid: [7069920](https://pubmed.ncbi.nlm.nih.gov/7069920/) (1982).
37. Spratt, J. A., von Fournier, D., Spratt, J. S. & Weber, E. E. Decelerating Growth and Human Breast Cancer. *Cancer*, 2013–2019. ISSN: 0008-543X. doi:[10.1002/1097-0142\(19930315\)71:6<2013::AID-CNCR2820710615>3.0.CO;2-V](https://doi.org/10.1002/1097-0142(19930315)71:6<2013::AID-CNCR2820710615>3.0.CO;2-V). [https://doi.org/10.1002/1097-0142\(19930315\)71:6%3C2013::AID-CNCR2820710615%3E3.0.CO;2-V](https://doi.org/10.1002/1097-0142(19930315)71:6%3C2013::AID-CNCR2820710615%3E3.0.CO;2-V) (1993).
38. Spratt, J. S., Meyer, J. S. & Spratt, J. A. Rates of Growth of Human Solid Neoplasms: Part I. *Journal of Surgical Oncology*, 137–146. ISSN: 0022-4790. doi:[10.1002/jso.2930600216](https://doi.org/10.1002/jso.2930600216). <https://doi.org/10.1002/jso.2930600216> (1995).
39. Lavielle, M. *Mixed Effects Models for the Population Approach* (2014).
40. Comets, E., Lavenu, A. & Lavielle, M. Parameter Estimation in Nonlinear Mixed Effect Models Using Saemix, an R Implementation of the SAEM Algorithm. *Journal of Statistical Software*, 1–41. doi:[10.18637/jss.v080.i03](https://doi.org/10.18637/jss.v080.i03) (2017).
41. R Core Team. *R: A Language and Environment for Statistical Computing* <https://www.R-project.org/> (R Foundation for Statistical Computing, Vienna, Austria, 2021).
42. Stoica, P. & Selen, Y. Model-Order Selection: A Review of Information Criterion Rules. *IEEE Signal Processing Magazine*, 36–47. ISSN: 1558-0792. doi:[10.1109/MSP.2004.1311138](https://doi.org/10.1109/MSP.2004.1311138) (2004).
43. Jänicke, F., Prechtel, A., *et al.* Randomized Adjuvant Chemotherapy Trial in High-Risk, Lymph Node-Negative Breast Cancer Patients Identified by Urokinase-Type Plasminogen Activator and Plasminogen Activator Inhibitor Type 1. *JNCI: Journal of the National Cancer Institute*, 913–920. ISSN: 0027-8874. doi:[10.1093/jnci/93.12.913](https://doi.org/10.1093/jnci/93.12.913). pmid: [11416112](https://pubmed.ncbi.nlm.nih.gov/11416112/) (2001).
44. Duffy, M. J., McGowan, P. M., Harbeck, N., Thomssen, C. & Schmitt, M. uPA and PAI-1 as Biomarkers in Breast Cancer: Validated for Clinical Use in Level-of-Evidence-1 Studies. *Breast Cancer Research*, 428. doi:[10.1186/s13058-014-0428-4](https://doi.org/10.1186/s13058-014-0428-4). pmid: [25677449](https://pubmed.ncbi.nlm.nih.gov/25677449/) (2014).
45. Dowsett, M., Nielsen, T. O., *et al.* Assessment of Ki67 in Breast Cancer: Recommendations from the International Ki67 in Breast Cancer Working Group. *JNCI: Journal of the National Cancer Institute*, 1656–1664. ISSN: 0027-8874. doi:[10.1093/jnci/djr393](https://doi.org/10.1093/jnci/djr393). pmid: [21960707](https://pubmed.ncbi.nlm.nih.gov/21960707/) (2011).
46. Steel, G. G. & Lamerton, L. F. The Growth Rate of Human Tumours. *British Journal of Cancer*, 74–86 (1966).

47. Friberg, S. & Mattson, S. On the Growth Rates of Human Malignant Tumors: Implications for Medical Decision Making. *Journal of Surgical Oncology*, 284–297. ISSN: 0022-4790. doi:[10.1002/\(sici\)1096-9098\(199708\)65:4<284::aid-jso11>3.0.co;2-2](https://doi.org/10.1002/(sici)1096-9098(199708)65:4<284::aid-jso11>3.0.co;2-2). pmid: [9274795](https://pubmed.ncbi.nlm.nih.gov/9274795/) (1997).
48. Jarrett, A. M., Kazerouni, A. S., *et al.* Quantitative Magnetic Resonance Imaging and Tumor Forecasting of Breast Cancer Patients in the Community Setting. *Nature Protocols*, 5309–5338. ISSN: 1754-2189. doi:[10.1038/s41596-021-00617-y](https://doi.org/10.1038/s41596-021-00617-y). pmid: [34552262](https://pubmed.ncbi.nlm.nih.gov/34552262/) (2021).
49. Bertucci, F., Ng, C. K. Y., *et al.* Genomic Characterization of Metastatic Breast Cancers. *Nature*, 560–564. ISSN: 0028-0836. doi:[10.1038/s41586-019-1056-z](https://doi.org/10.1038/s41586-019-1056-z). pmid: [31118521](https://pubmed.ncbi.nlm.nih.gov/31118521/) (2019).
50. Niazi, M. K. K., Parwani, A. V. & Gurcan, M. N. Digital Pathology and Artificial Intelligence. *The Lancet Oncology*, e253–e261. ISSN: 1470-2045. doi:[10.1016/s1470-2045\(19\)30154-8](https://doi.org/10.1016/s1470-2045(19)30154-8). pmid: [31044723](https://pubmed.ncbi.nlm.nih.gov/31044723/). <https://www.sciencedirect.com/science/article/pii/S1470204519301548?via%3Dihub> (2019).
51. Conforti, F., Pala, L., *et al.* Surrogacy of Pathologic Complete Response in Trials of Neoadjuvant Therapy for Early Breast Cancer. *JAMA Oncology*, 1668–1675. ISSN: 2374-2437. doi:[10.1001/jamaoncol.2022.3755](https://doi.org/10.1001/jamaoncol.2022.3755). pmid: [36201176](https://pubmed.ncbi.nlm.nih.gov/36201176/) (2022).
52. Marron, T. U., Galsky, M. D., *et al.* Neoadjuvant Clinical Trials Provide a Window of Opportunity for Cancer Drug Discovery. *Nature Medicine*, 626–629. ISSN: 1078-8956. doi:[10.1038/s41591-022-01681-x](https://doi.org/10.1038/s41591-022-01681-x). pmid: [35347282](https://pubmed.ncbi.nlm.nih.gov/35347282/) (2022).
53. Setting Guidelines to Report the Use of AI in Clinical Trials. *Nature Medicine*, 1311–1311. ISSN: 1078-8956. doi:[10.1038/s41591-020-1069-z](https://doi.org/10.1038/s41591-020-1069-z). pmid: [32908276](https://pubmed.ncbi.nlm.nih.gov/32908276/) (2020).
54. Topol, E. J. High-Performance Medicine: The Convergence of Human and Artificial Intelligence. *Nature Medicine*, 44–56. doi:[10.1038/s41591-018-0300-7](https://doi.org/10.1038/s41591-018-0300-7) (2019).
55. Commenges, D. & Jacqmin-Gadda, H. Dynamical Biostatistical Models, 174–209. doi:[10.1201/b19109-15](https://doi.org/10.1201/b19109-15) (2015).

Supplementary Figures



Supplementary Figure 1. Distribution of the pathological tumor sizes (mm) in each of the three datasets. Statistical analysis was performed with Brown-Mood Median Test (***) $p < 0.001$, n.s. non-significant)



Supplementary Figure 2. Kaplan-Meier estimate of the distant metastasis-free survival in the datasets

Supplementary Text

Derivation of the likelihood for the model

Likelihood for censored observations

The contribution of a single individual (potentially censored) observation (T^i, δ^i) , with T^i the reported time and δ^i the censoring indicator, to the likelihood is given by⁵⁵,

$$L^i = h(T^i)^{\delta^i} S(T^i)$$

Where S and h are respectively the survival function and hazard function of the model.

Likelihood of the model for finite TTR

In the case of our mechanistic model, given individual values α^i and μ^i with observed size of the PT at diagnosis V_{diag}^i , we can compute the predicted DMFS, $TTR(\alpha^i, \mu^i; V_{diag}^i)$. When $TTR(\alpha^i, \mu^i; V_{diag}^i) < \infty$, the statistical model associated with this prediction is a log-normal distribution around this median value:

$$\log(T^i) = \log\left(TTR(\alpha^i, \mu^i; V_{diag}^i)\right) + \epsilon^i,$$

where $\epsilon^i \sim \mathcal{N}(0, \sigma^2)$. This equation implicitly defines a survival function for T^i :

$$T^i \sim \log \mathcal{N}\left(\log TTR(\alpha^i, \mu^i; V_{diag}^i), \sigma^2\right).$$

The corresponding survival and hazard function are then:

$$S_{TTR}(t) = 1 - \Phi\left(\frac{\log t - \log TTR(\alpha^i, \mu^i; V_{diag}^i)}{\sigma}\right),$$

$$h_{TTR}(t) = \frac{1}{\sigma t S_{TTR}(t)} \varphi\left(\frac{\log t - \log TTR(\alpha^i, \mu^i; V_{diag}^i)}{\sigma}\right),$$

Where Φ and φ are respectively the cumulative distribution function and the probability density function of a standard gaussian distribution.

Likelihood of the model for infinite TTR

In the case where $TTR(\alpha^i, \mu^i; V_{diag}^i) = \infty$, we used the natural extension of the likelihood:

$$L^i = \begin{cases} 1, & \delta^i = 0 \\ 0, & \delta^i = 1 \end{cases}$$