Figures and Tables

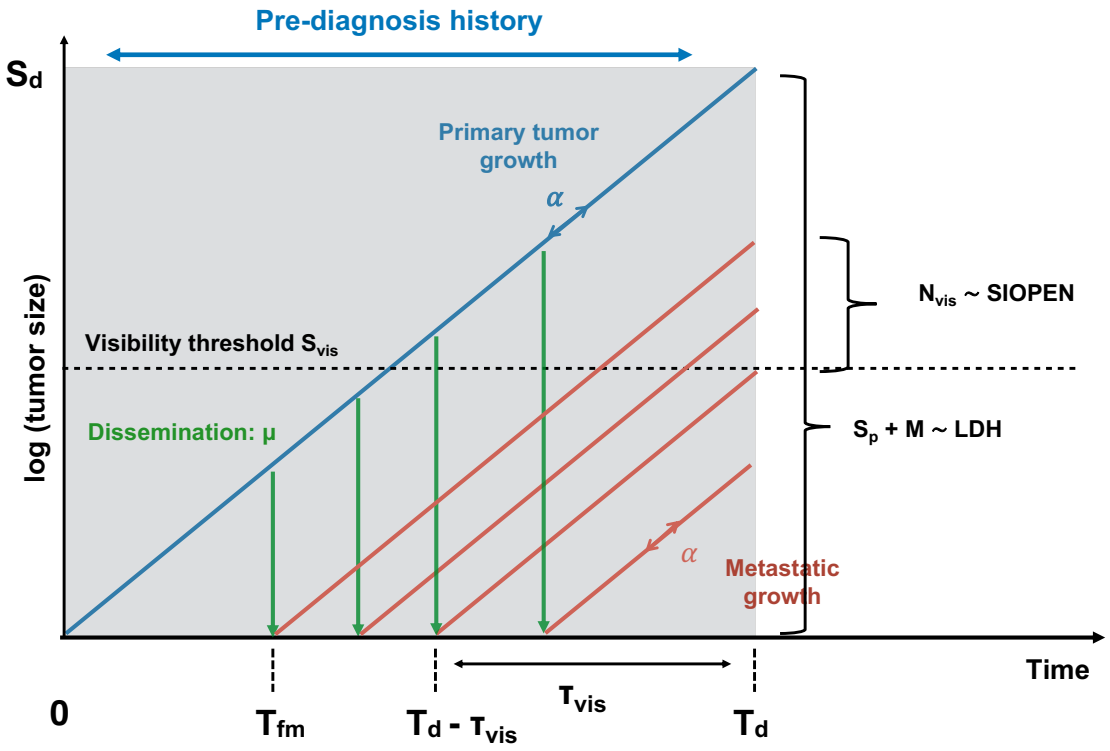
Table 1: Patients characteristics

Sex	Male	n=28 (57.2%)
	Female	n=21 (42.8%)
Age	≥ 18 months	n=43 (87.8%)
	< 18 months	n=6 (12.2%)
LDH rates	≥ 1250 UI/L	n=18 (36.7%)
	< 1250 UI/L	n=31 (63.3%)
MYCN	Amplified	n=23 (46.9%)
	Non amplified	n=26 (53.1%)
SIOPEN	≥ 4	n=30 (65.2%)
Only if MIBG was positive	< 4	n=16 (34.8%)
Metastasis	Present	n=43 (87.8%)
	Absent	n=6 (12.2%)

Table 2: Multivariable Cox analysis of overall survival

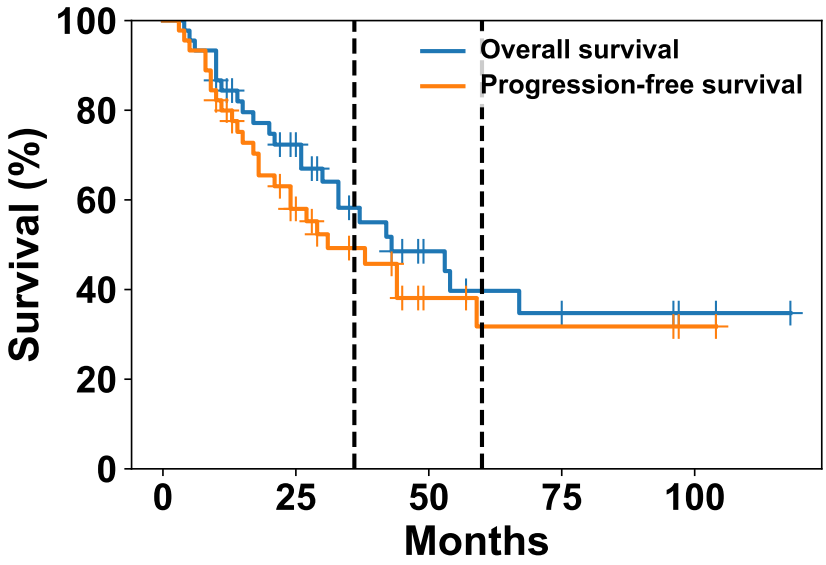
covariate	Hazard ratio	p	coef lower 95%	coef upper 95%
age	1	0.722	0.985	1.02
sex	1.44	0.463	0.54	3.86
log(LDH)	4.01	0.00562	1.5	10.7
SIOPEN	0.999	0.963	0.966	1.03
MYCN	0.477	0.346	0.102	2.22
tumor size	1	0.99	0.999	1
$\log(\mu)$	0.837	0.0739	0.689	1.02
visible threshold	0.997	0.549	0.988	1.01

Figure 1: Schematic of the mathematical model



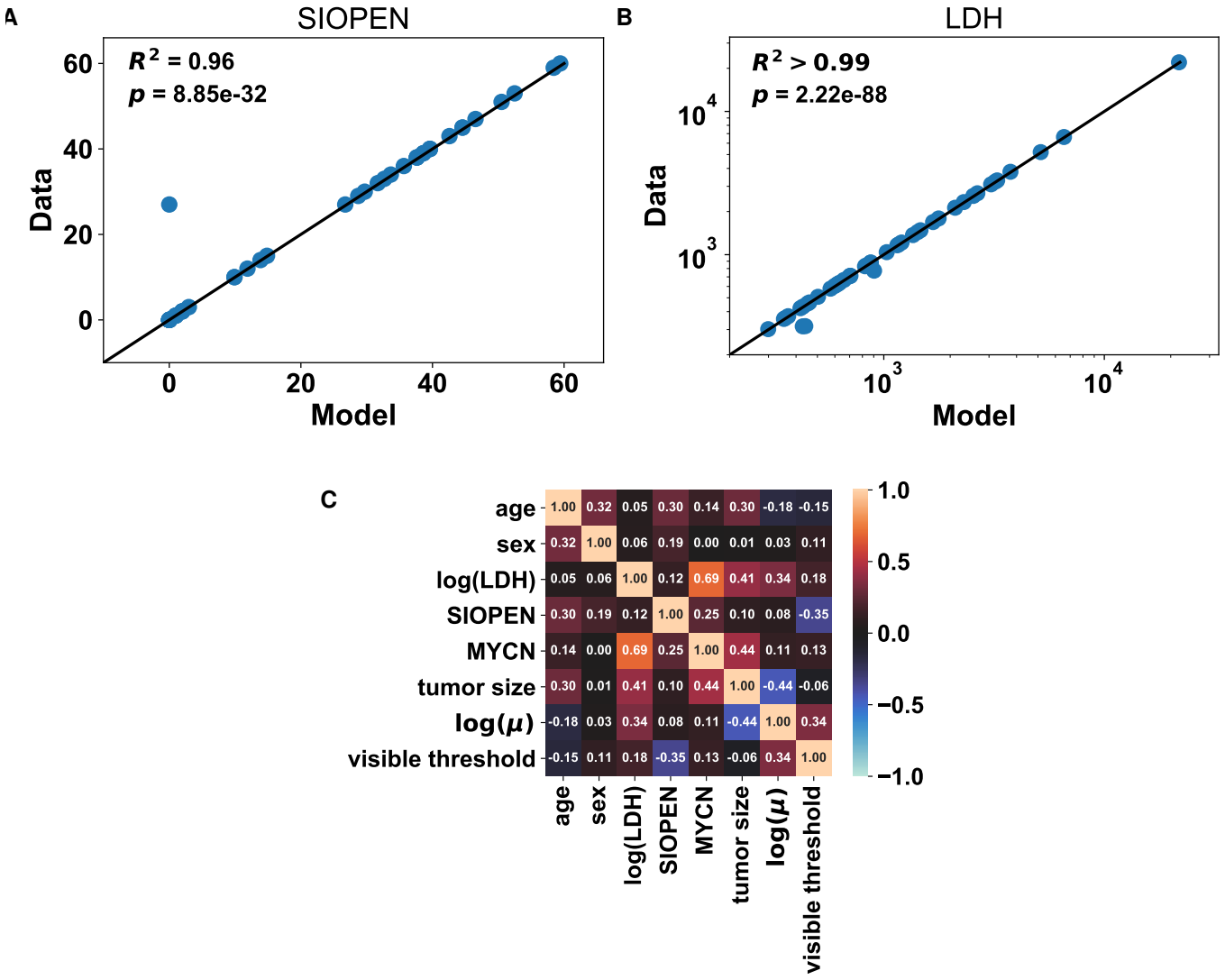
Primary and secondary tumors are assumed to have exponential growth kinetics governed by a proliferation rate α. Dissemination of metastasis is controlled by the parameter μ. From these and primary tumor size at diagnosis S_d, the primary tumor age T_d can be computed and simulations of the natural history can be performed. Adjunction of a visibility threshold S_{vis} results in predictions of the number of visible metastases N_{vis} and total cancer mass (primary + secondary tumors) $S_p + M$. These are respectively compared to the SIOPEN score and lactate dehydrogenase level. The time of birth of the first metastasis is denoted T_{fm} and the time to reach S_{vis} from one cell τ_{vis}. Note that the number of visible metastases at time T_d is the total number of metastases at time $T_d - \tau_{vis}$.

Figure 2: Overall and progression-free survival



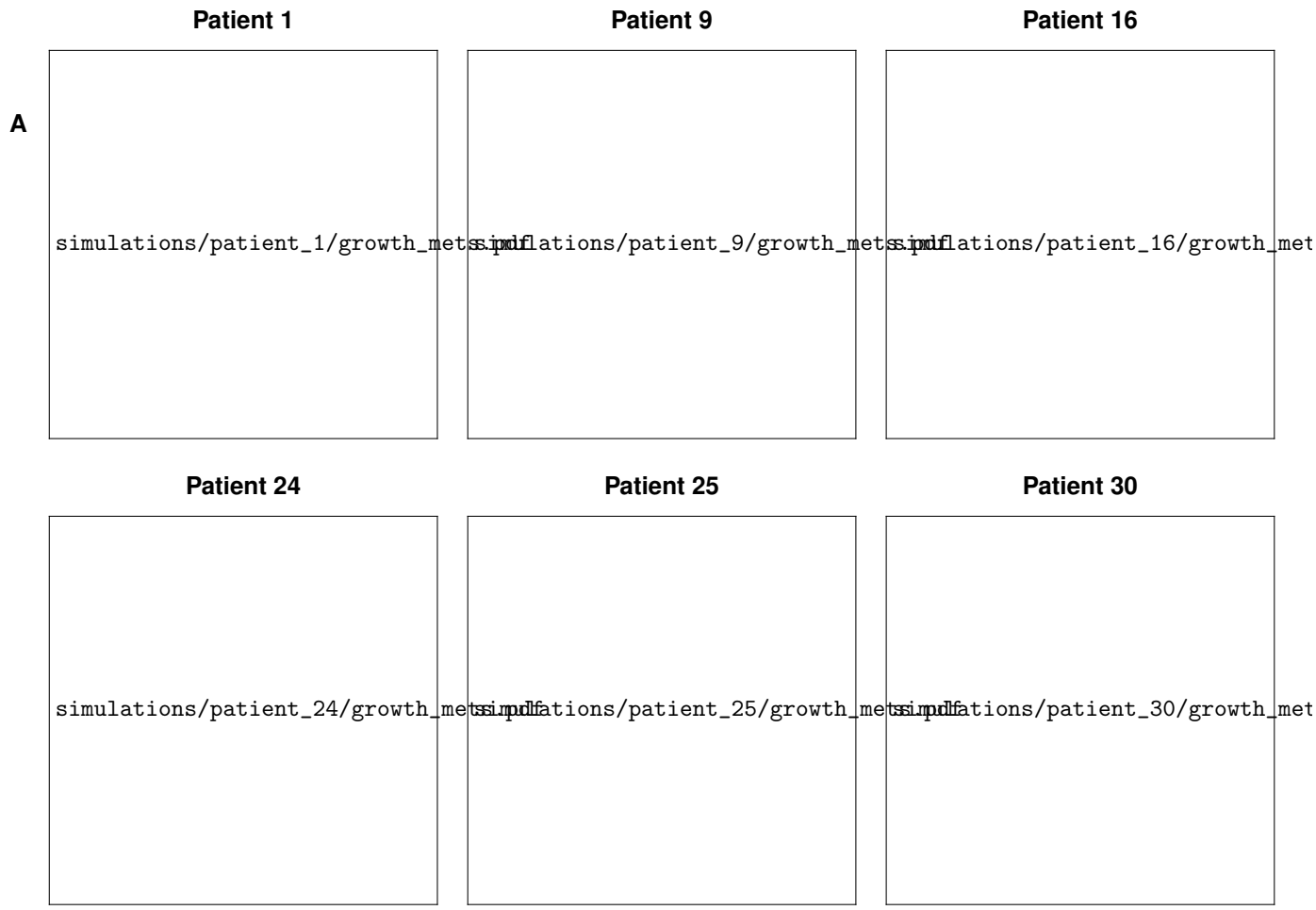
Kaplan-Meier estimates of overall and progression-free survival. Dotted lines show survival rates at 3 years (PFS=48.1%, OS=56.6%) and 5 years (PFS=31.8%, OS=38.2%). Median PFS = 31 months, median OS = 43 months.

Figure 3: Descriptive power of the mathematical model



A. Fit of the SIOPEN data. Solid line is the identity line.
B. Fit of the LDH data. Solid line is the identity line.
C. Correlation matrix of all features including clinical variables and (log) of the mathematical parameter μ. Level of darkness indicates positive correlation whereas brightness indicates negative correlation

Figure 4: Mechanistic simulations of the pre-diagnosis history of high-risk neuroblastoma patients

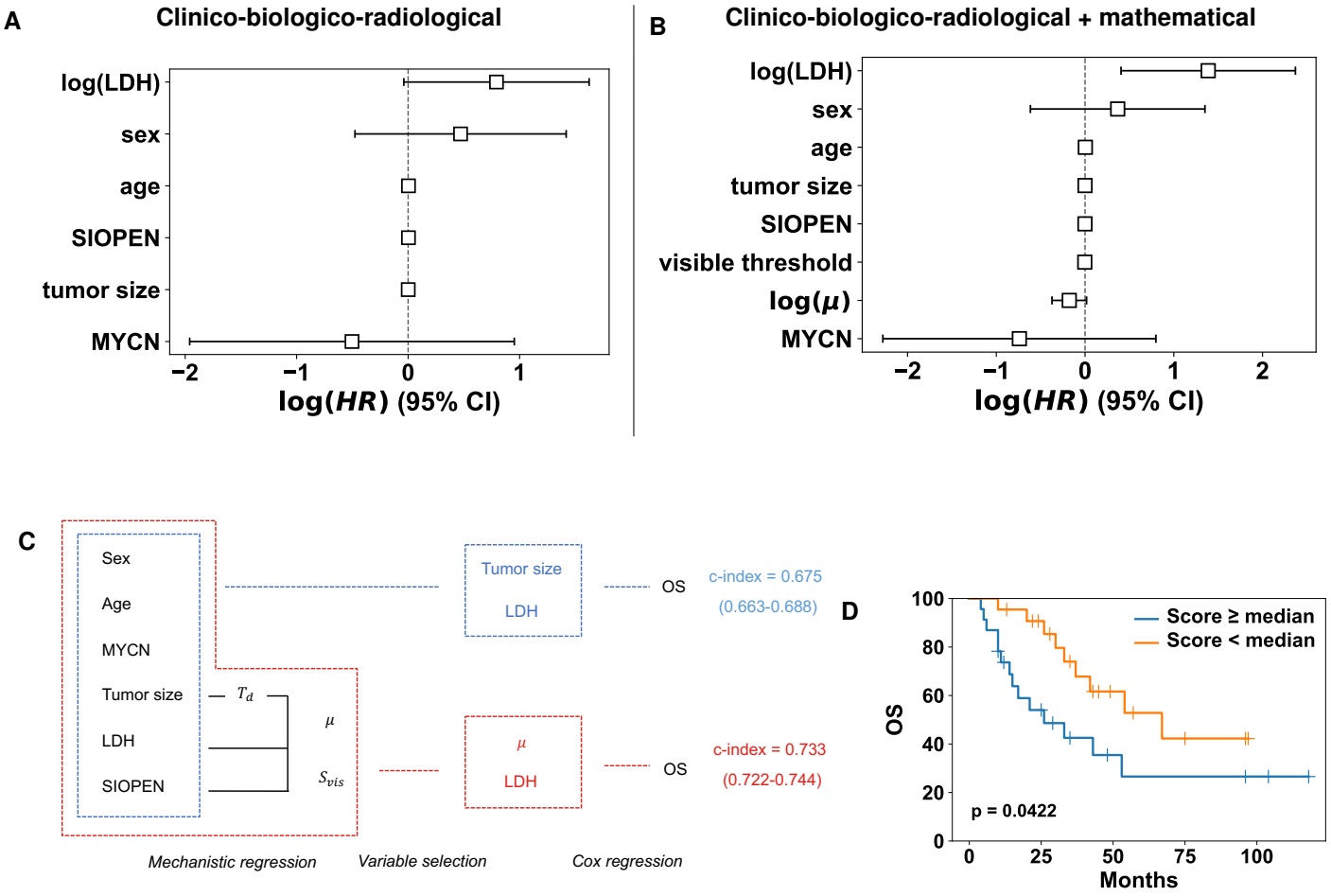


A mechanistic model was calibrated using patient data of SIOPEN score (number of visible metastases) and LDH levels (total cancer burden), which resulted in individual values of the model parameters μ and S_{vis}. These values are reported in the supplementary file S1 and were used to simulate the pre-diagnosis natural history of the disease.

A. Simulations of the primary tumor (blue) and metastases (red) growth kinetics, for representative patients

B. Distributions of the times of the primary tumor (PT) diagnosis (age of the PT, T_d) and birth of the first metastasis (T_{fm})

Figure 5: Prognosis of overall survival



A. Hazard ratios and 95% confidence intervals of the clinical variables in multivariable Cox regression.
B. Same as A. with $\log(\mu)$ and visible threshold S_{vis} as additional variables.
C. Regression analysis for survival prediction. To predict overall survival, two Cox regression-based models were compared: using clinico-biologico-radiological variables only (in blue) versus these augmented with two computational biomarkers derived from fitting the mechanistic model to the quantitative data: μ and S_{vis}. These were determined by maximum likelihood estimation from the LDH and SIOPEN data, using in addition the tumor size S_d data to determine the age of the primary tumor T_d. Variable selection was performed in both cases from multivariable Cox regression and c-indices were computed from the final Cox models using 100 replicates of 5-fold cross-validations.
D. Separation of patients according to the Cox score predicted from the selected variables in the final model (log(LDH) and $\log(\mu)$). p-value is from a log-rank test.