



**HAL**  
open science

# Pharmacometrics modeling coupled with machine learning for early prediction of survival following atezolizumab monotherapy in non-small cell lung cancer

Sébastien Benzekry, Mélanie Karlsen, Abdessamad Elkaoutari, Suresh Vatakuti, Petter Curle, Candice Jamois

## ► To cite this version:

Sébastien Benzekry, Mélanie Karlsen, Abdessamad Elkaoutari, Suresh Vatakuti, Petter Curle, et al.. Pharmacometrics modeling coupled with machine learning for early prediction of survival following atezolizumab monotherapy in non-small cell lung cancer. *compsyscan22: A complex systems approach to cancer understanding*, Oct 2022, Lyon, France. hal-03935786

**HAL Id: hal-03935786**

<https://inria.hal.science/hal-03935786v1>

Submitted on 12 Jan 2023

**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.

S. Benzekry<sup>1</sup> M. Karlsen<sup>1</sup> A. El Kaoutari<sup>1</sup> S. Vatakuti<sup>2</sup> P. Curle<sup>3</sup> C. Jamois<sup>2</sup>

<sup>1</sup>COMPO, Inria Méditerranée, Centre de Recherche sur le Cancer de Marseille, Inserm, CNRS, IPC, Aix-Marseille University, Marseille, France; <sup>2</sup>Safety and Early Development Informatics and <sup>3</sup> Clinical Pharmacometrics, Pharmaceutical Sciences, Pharma Research and Early Development, Roche Innovation Center Basel, Switzerland

## BACKGROUND

- Treatment of advanced non-small cell lung cancer by atezolizumab-based therapy
- Prediction of survival from baseline or early on-treatment data could
  - guide treatment decision during **drug development**
  - inform **personalized** health care
- Current predictive biomarker: programmed death-ligand 1 (PD-L1)<sup>1</sup>
- State of the art from **baseline** clinical and biological data: ROPRO score<sup>2</sup>
- Predictive value of transcriptomic and mutation data is unclear
- Tumor kinetics (TK)** model parameter growth rate (KG) has important predictive power of hazard ratio (HR)<sup>3</sup>
- Predictive value of kinetics of **pharmacodynamic biomarkers** is unclear

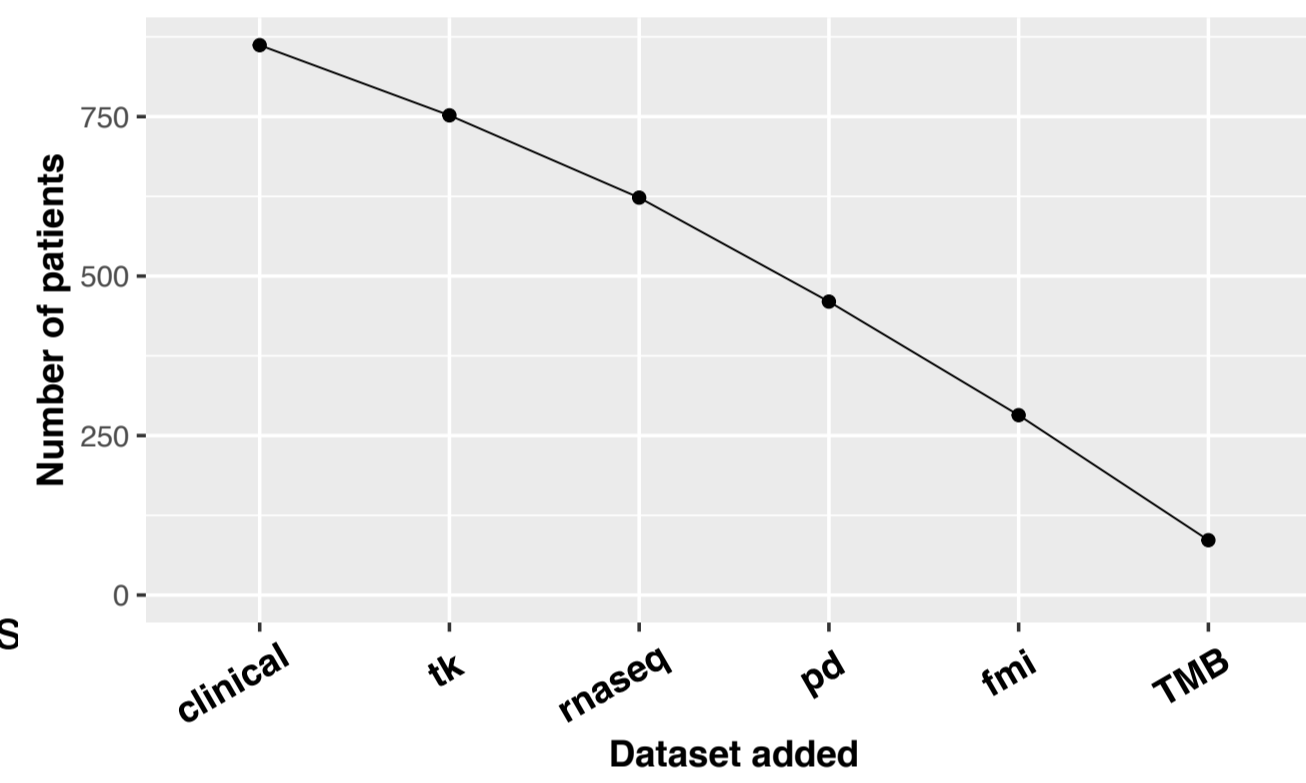
## STUDIES, PATIENTS AND DATA

Studies	Study Description	Population	Patients treated with atezolizumab
FIR GO28625	Phase 2 study that evaluated the efficacy and safety of anti-programmed death-ligand 1 (PD-L1) atezolizumab in advanced NSCLC selected by tumor cell (TC) or tumor-infiltrating immune cell (IC) PD-L1 expression	PD-L1 positive locally advanced or metastatic NSCLC (lines 1 and 2+)	133
POPLAR GO28753	Phase 2 randomised controlled trial (RCT) of atezolizumab versus docetaxel for patients with previously treated NSCLC	Locally advanced or metastatic NSCLC who failed platinum therapy	134
BIRCH GO28754	Phase 2 Study of Atezolizumab in participants with PD-L1 positive locally advanced or metastatic NSCLC	Locally advanced or metastatic NSCLC (lines 1, 2 or 3)	595
Total			862

5 sources of data:

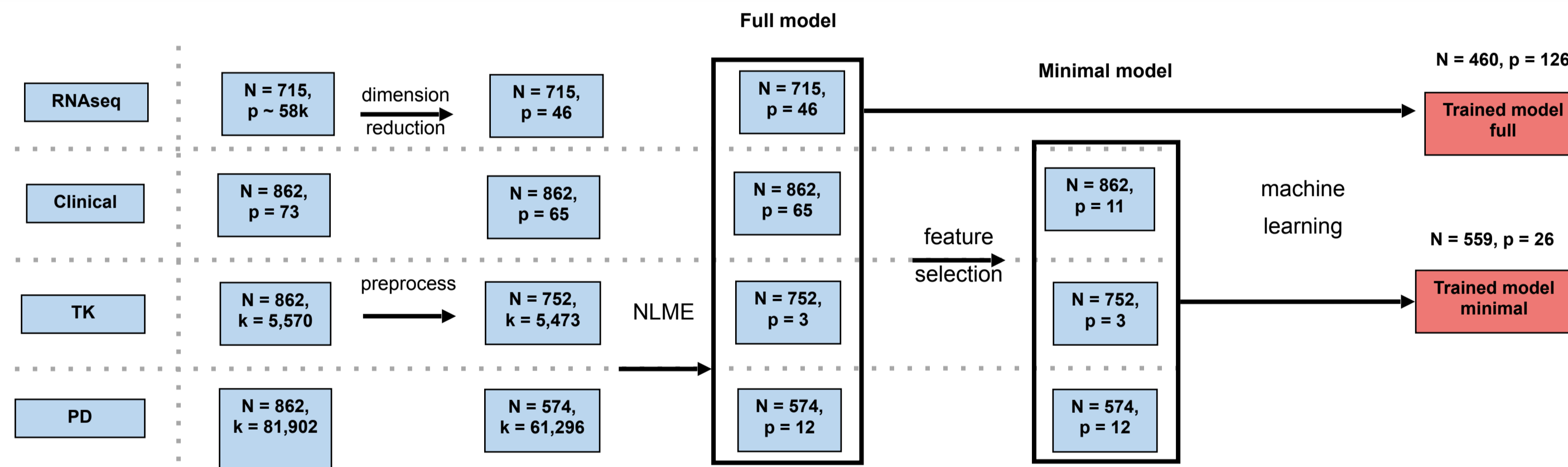
- Baseline**
  - clinical** and biological characteristics (73 variables)
  - RNAseq** (~ 58k variables)
  - FMI** (mutation data on 395 genes)
  - Tumor mutational burden (TMB)
- Longitudinal**
  - kinetics of tumor size** (TK, sum of largest diameters)
  - kinetics of 4 pharmacodynamic markers** (PD): albumin, C-reactive protein (CRP), lactate deshydrogenase (LDH), neutrophils

Number of patients with data available for each dataset



FMI and TMB disregarded because they would have highly reduced the number of patients

## METHODS

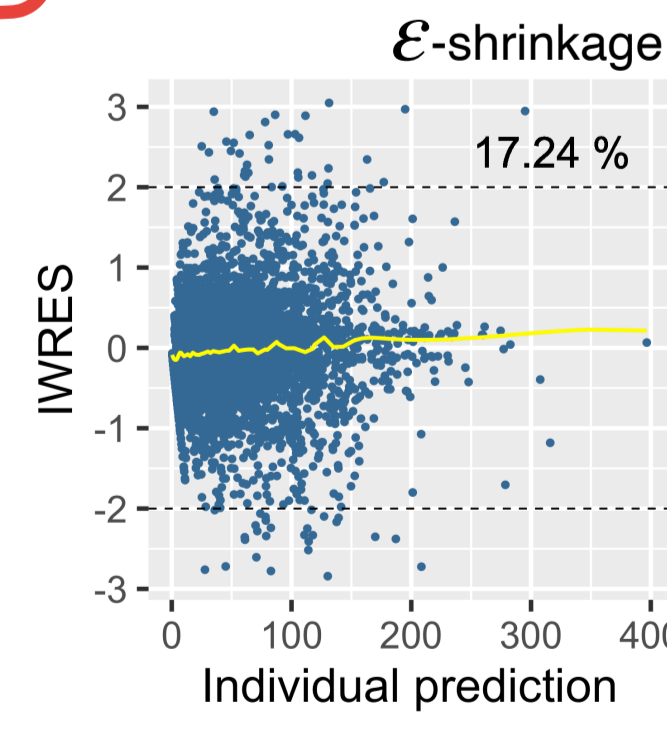
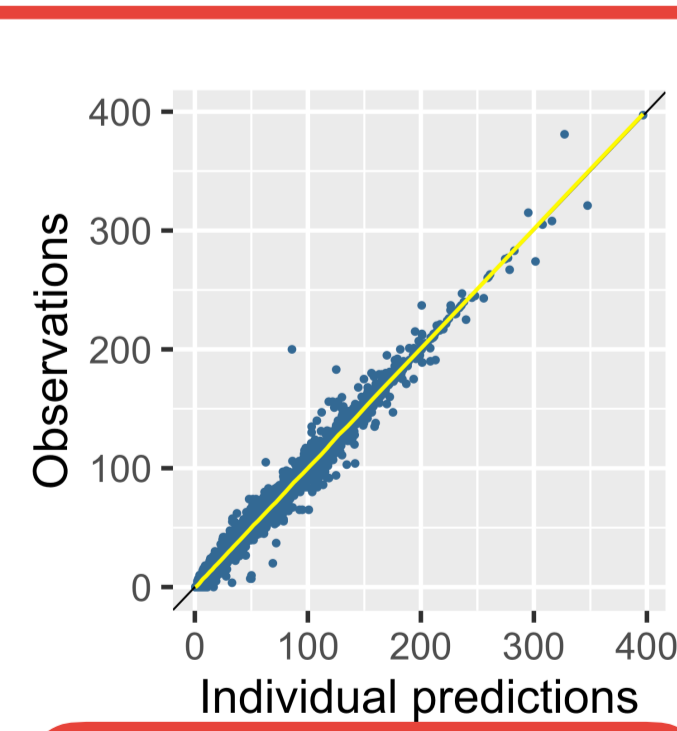
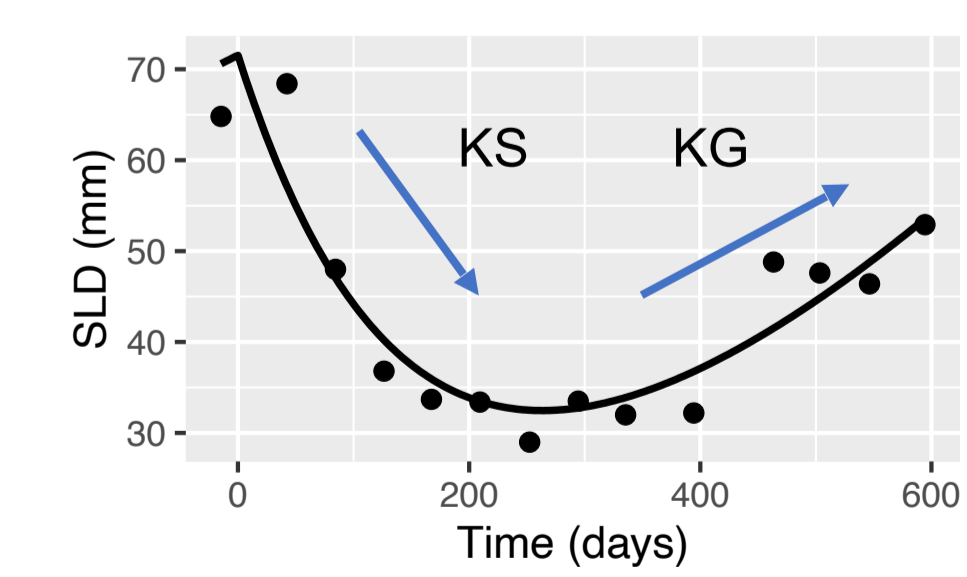


- Best survival model = **random survival forest**
- Multiple methods of **feature selection** were assessed
- Evaluation of performances with 10-fold **cross-validation**
- Results using full time courses suffer from **immortal time bias**
  - We used **truncated data sets**
    - At cycles 3, 5 and 10 pre-dose
    - i.e. after 6, 12 and 27 weeks of treatment

## RESULTS : NONLINEAR MIXED-EFFECTS MODELING (NLME)

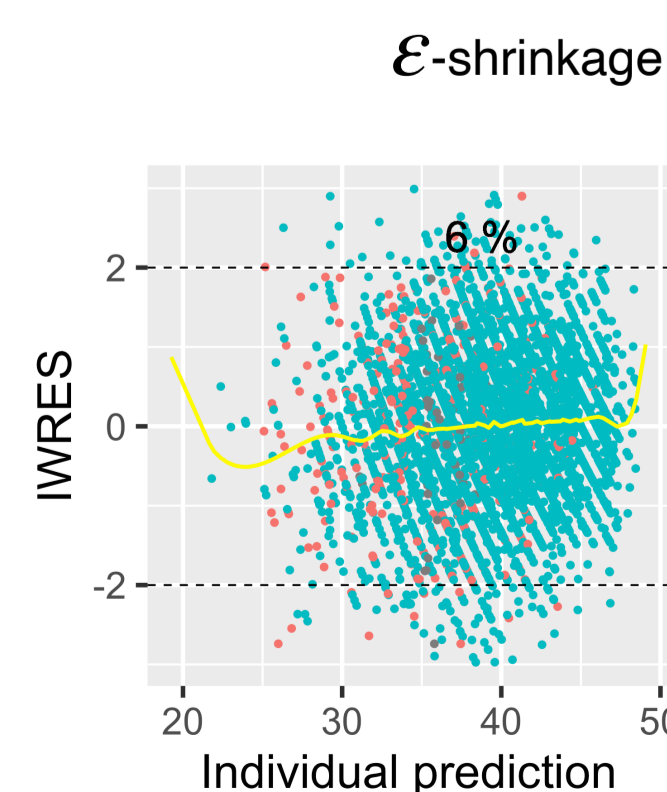
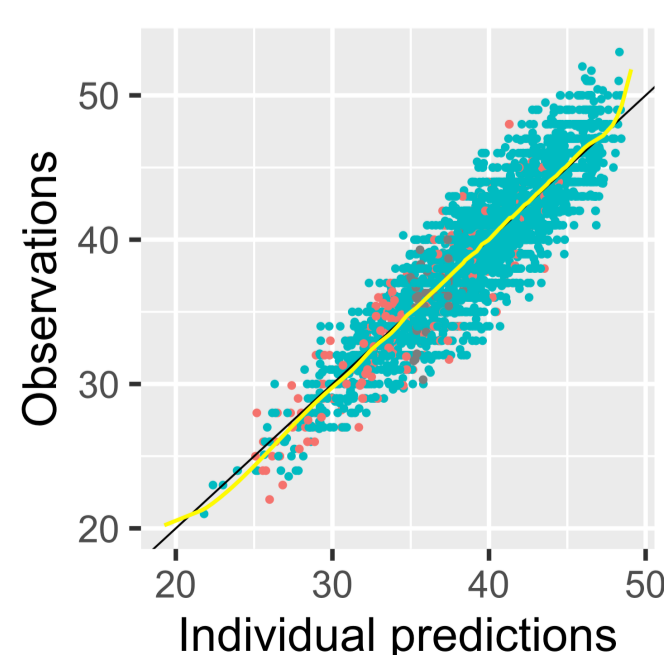
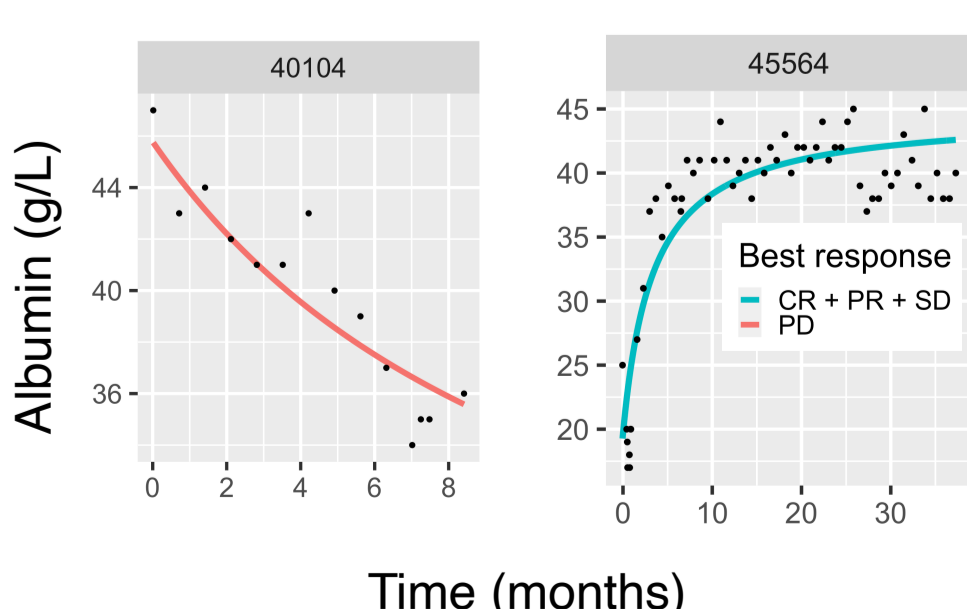
### Tumor kinetics

$$S(t) = S_0 (e^{-KS \cdot t} + e^{KG \cdot t} - 1)$$



### Albumin kinetics

$$y(t) = p + e^t \frac{q-p}{t+e^t}$$



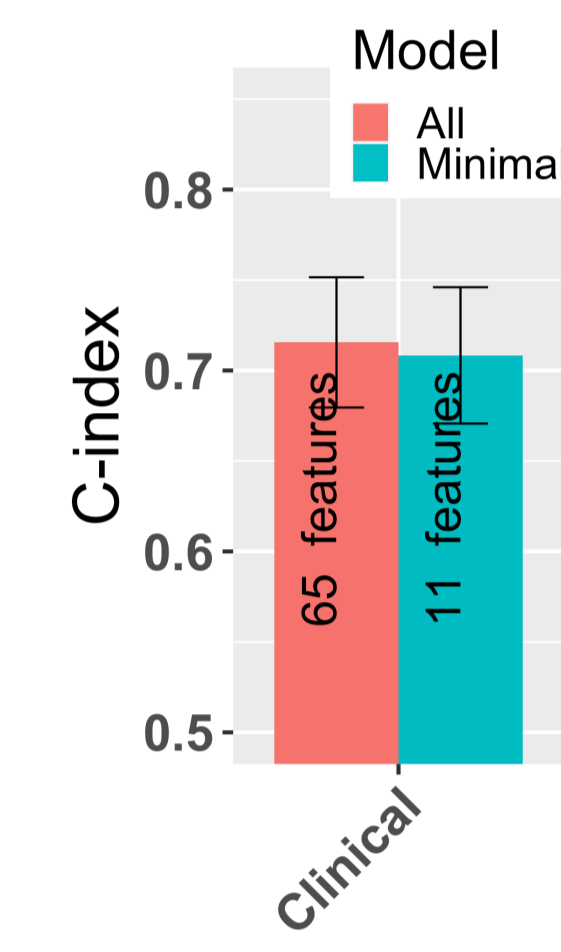
CRP, LDH and neutrophils kinetics were described using the above double-exponential model

## OBJECTIVE

To provide a digital **decision-enabling tool** by predicting **overall survival** based on early tumor size and longitudinal PD biomarker data using the strengths of **pharmacometrics** (PHMx) and machine learning (ML)

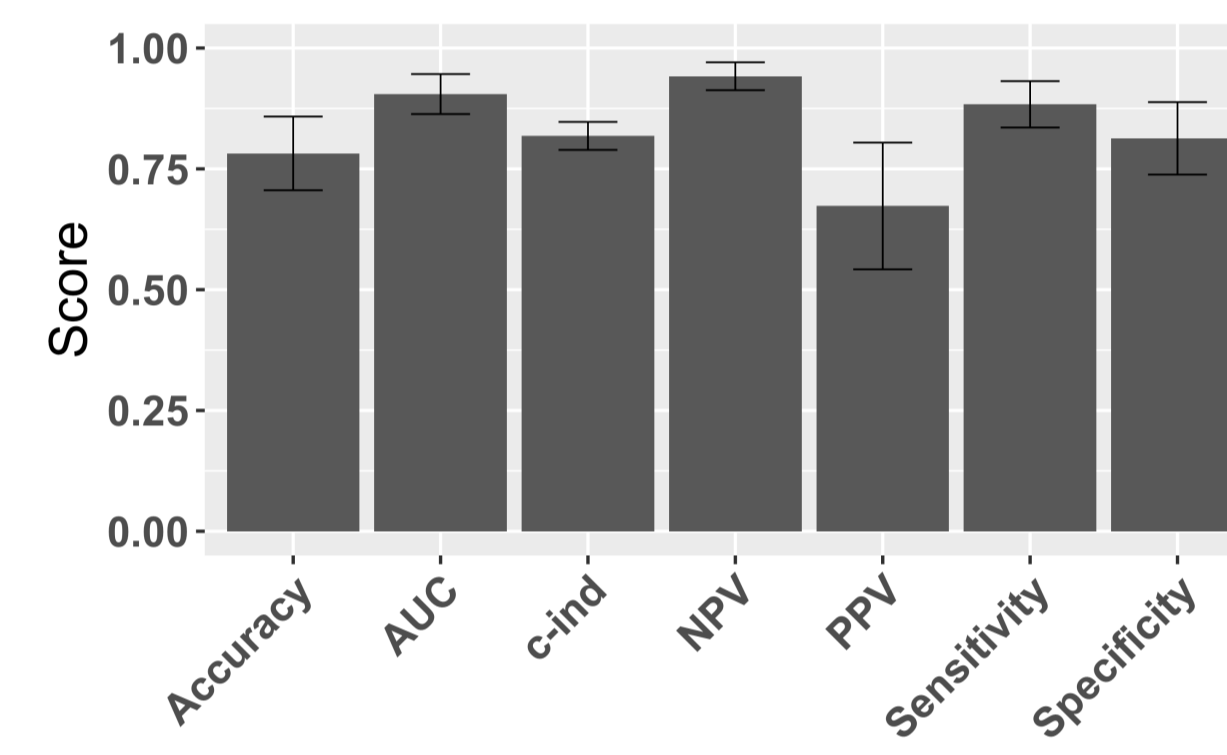
## RESULTS MACHINE LEARNING : FULL TIME COURSE

Minimal clinical signature
C-reactive protein
Heart rate
Neutrophils-to-lymphocytes ratio
Neutrophils
Lymphocytes-to-leukocytes ratio
Liver metastases
ECOG (0 vs 1)
PDL1 (≥ 50%)
Hemoglobin
Baseline sum of longest diameters
Lactate deshydrogenase

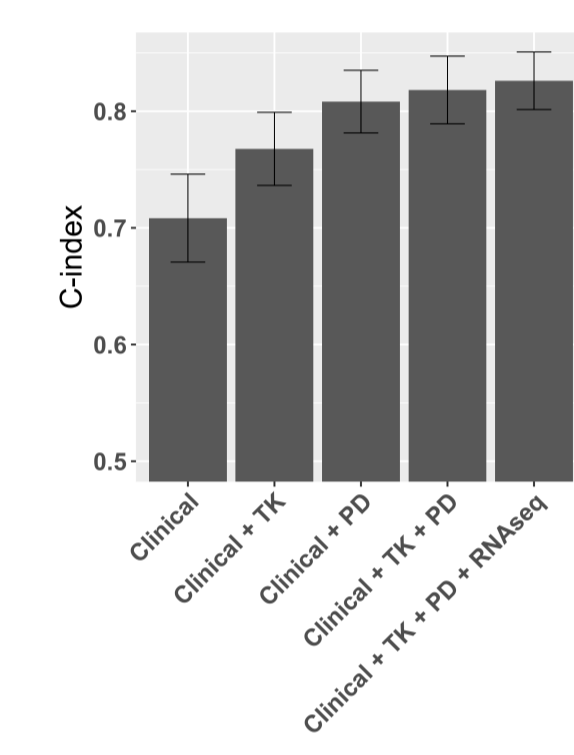


- A **minimal signature** achieves (almost) the same performances as the full model
- RNAseq** only brings **negligible** additional predictive power ⇒ dropped in the final model
- Substantial predictive power**, except for PPV (ability to **predict death**, 67.3% ± 13.1)

Evaluation of all predictive metrics in cross-validation

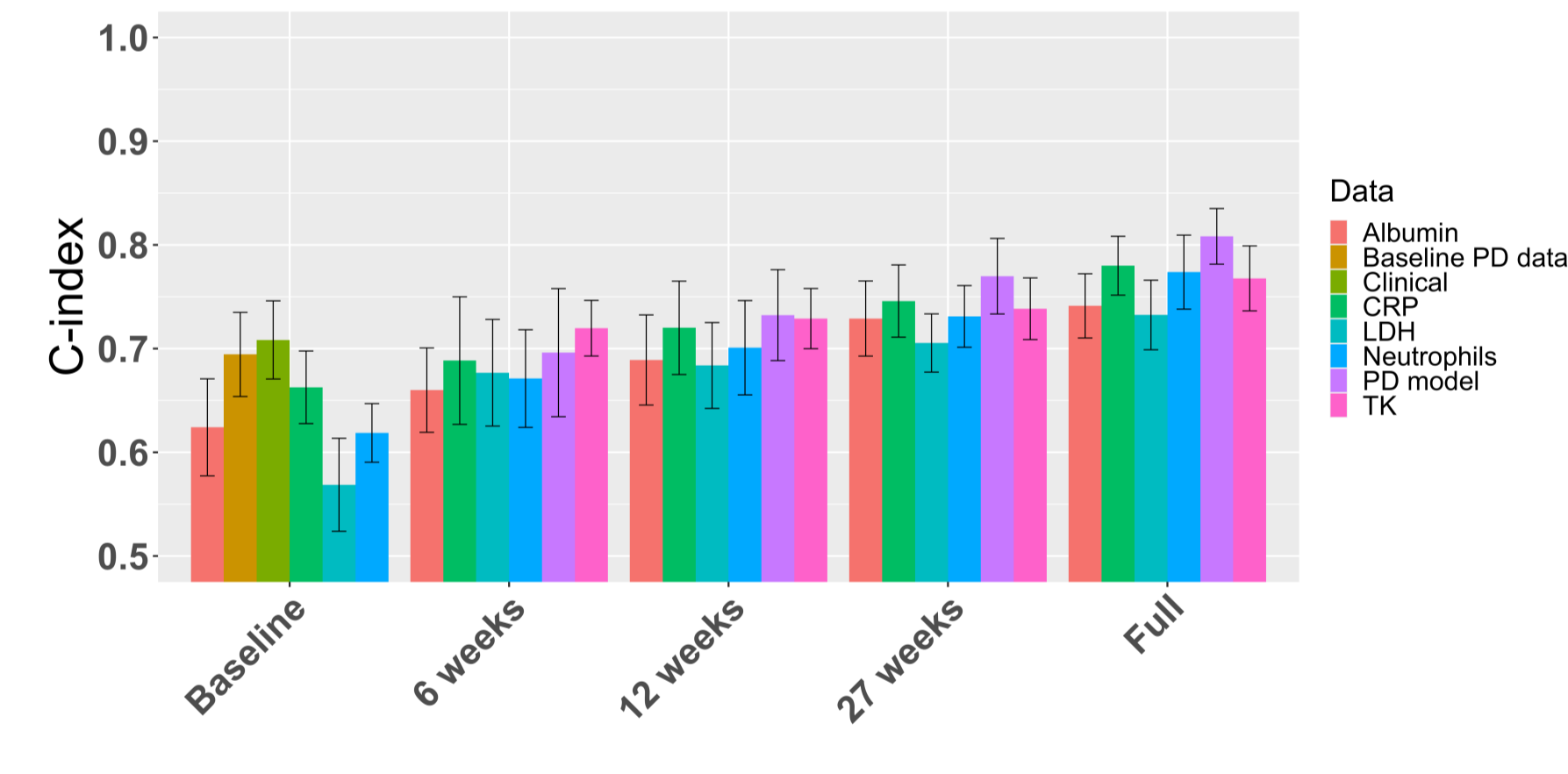
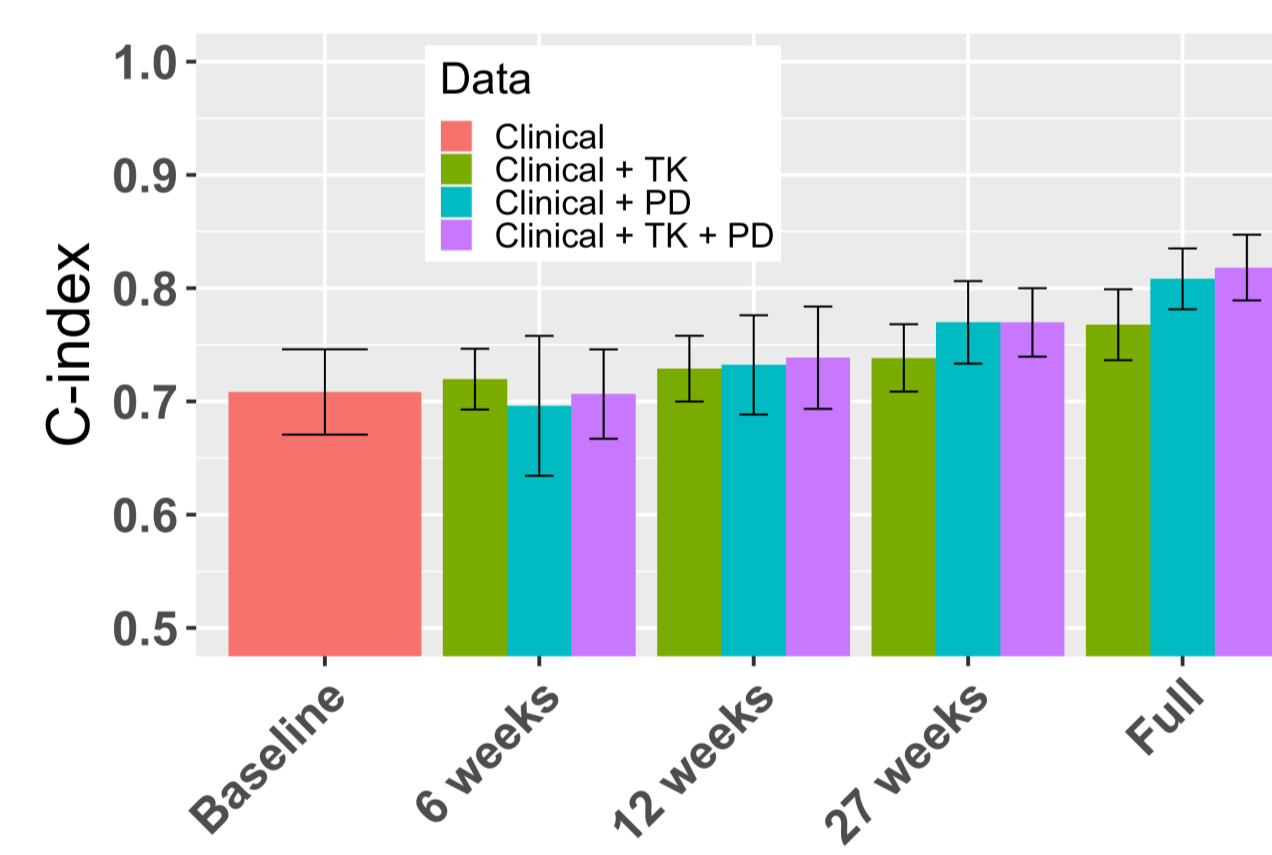


C-index as a function of feature sets



Positive (1) = death, Negative (0) = alive  
All classification metrics computed for survival at 12 months

## RESULTS MACHINE LEARNING : TRUNCATED TIME COURSE



- At least **4 completed cycles** of treatment (≥ 12 weeks) of data are required to achieve substantial individual predictive power
- PD model derived metrics are more informative of individual survival than TK metrics from 12 weeks onwards
- Best individual kinetic markers : **CRP and neutrophils**

## DISCUSSION

### Conclusion

- The **combination of NLME and ML** allowed to take the best of the two approaches in order to predict individual survival
  - NLME for **longitudinal TK and PD** data
  - ML to build **multivariable** models from a large number of features
- A model was established based on a 26 features **minimal signature**: 11 baseline clinical features + longitudinal TK (3 variables) + longitudinal PD (12 variables)
  - C-index = 0.818 ± 0.029, AUC = 0.905 ± 0.0414**
- RNAseq data did not yield substantial predictive power

### Perspectives

- External validation** on the phase 3 OAK trial
- Prediction of study-level overall survival in multiple arms, from early on-study data

1 Shukuya, T. & Carbone, D. P. Predictive Markers for the Efficacy of Anti-PD-1/PD-L1 Antibodies in Lung Cancer. Journal of Thoracic Oncology 11, 976–988 (2016).

2 Becker, T. et al. An enhanced prognostic score for overall survival of patients with cancer derived from a large real-world cohort. Ann Oncol 31, 1561–1568 (2020).

3 Claret, L. et al. A Model of Overall Survival Predicts Treatment Outcomes with Atezolizumab versus Chemotherapy in Non-Small Cell Lung Cancer Based on Early Tumor Kinetics. Clin Cancer Res 24, 3292–3298 (2018).