



**HAL**  
open science

## Supporting decision making and early prediction of survival for oncology drug development using a pharmacometrics-machine learning based model

Sébastien Benzekry, Mélanie Karlsen, Abdessamad Elkaoutari, René Bruno, Aleš Neubert, François Mercier, Bruno Gomes, Suresh Vatakuti, Peter Curle, Candice Jamois

### ► To cite this version:

Sébastien Benzekry, Mélanie Karlsen, Abdessamad Elkaoutari, René Bruno, Aleš Neubert, et al.. Supporting decision making and early prediction of survival for oncology drug development using a pharmacometrics-machine learning based model. PAGE - Population Approach Group Europe, Jun 2022, Ljubljana, Slovenia. hal-03921394

**HAL Id: hal-03921394**

**<https://inria.hal.science/hal-03921394>**

Submitted on 3 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.

# Supporting decision making and early prediction of survival for oncology drug development using a pharmacometrics-machine learning based model.

Sébastien Benzekry<sup>1</sup>, Mélanie Karlsen<sup>1</sup>, Abdessamad El Kaoutari<sup>1</sup>, René Bruno<sup>2</sup>, Ales Neubert<sup>3</sup>, François Mercier<sup>3</sup>, Martin Stern<sup>3</sup>, Bruno Gomes<sup>3</sup>, Suresh Vatakuti<sup>3</sup>, Peter Curle<sup>3</sup> and Candice Jamois<sup>3</sup>

(1) COMPO Inria – Inserm, Marseille, France; (2) Genentech-Roche, Marseille, France; (3) Roche pRED, Basel, Switzerland

**Thirtieth PAGE meeting, Ljubljana, Slovenia**  
**30 June 2022**

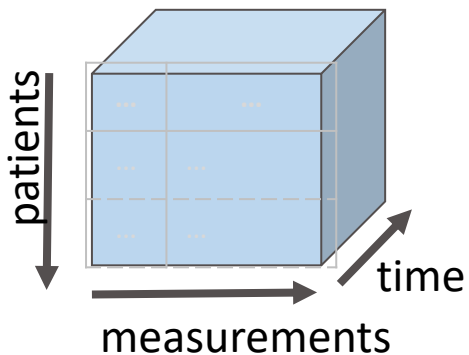
## Table of contents

1. Project schematic and objectives
2. Methodology
  - a. Data
  - b. Pharmacometrics model development
  - c. Machine Learning model development (Train set)
3. Results
  - a. Minimal signature model
  - b. Performance metrics by features category
  - c. External model validation (Test set)
4. Summary and Perspectives

### Raw longitudinal data

Tumor Kinetics (TK)

Pharmacodynamic markers (PD)  
Albumin, CRP, neutrophils, LDH



### Raw baseline data

Clinical characteristics

Genomic and Transcriptomic data

Feature selection

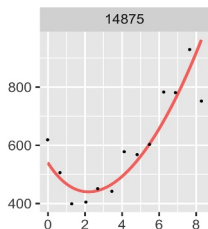
### Model-based individual parameters

TK

PD

patients

parameters



### Minimal signature

Clinical characteristics

Genomic and Transcriptomic data

Machine Learning

### Outputs

Individual prognosis

- C-index
- 12-months survival

Study level predictions

- Survival curve
- Hazard ratio

- **Baseline data**

- Patients' and disease characteristics  
**p = 73 parameters**
- Transcriptomic and mutational data  
**p = 58,311 and 395**

- **Longitudinal data**

- Tumor kinetics (TK, SLD)  
**5,570/3,065 observations**
- 4 PD markers (CRP, LDH, Albumin and Neutrophils)  
**61,296/47,255 observations**

Study	Description	N
<b>FIR GO28625</b>	Phase 2 study for the efficacy and safety of anti-programmed death-ligand 1 (PD-L1) atezolizumab (ATZ) in advanced NSCLC selected by tumor cell (TC) or tumor-infiltrating immune cell (IC) PD-L1 expression	133
<b>POPLAR GO28753</b>	Phase 2 randomised controlled trial (RCT) of ATZ versus docetaxel for patients with previously treated NSCLC (locally advanced or metastatic NSCLC who failed to platinum therapy)	134
<b>BIRCH GO28754</b>	Phase 2 study of ATZ in patients with PD-L1 positive locally advanced or metastatic NSCLC	595
<b>Train</b>		<b>862</b>
<b>Test - OAK GO28915</b>	Phase 3 RCT of ATZ versus docetaxel (DTX) in patients with previously treated NSCLC	<b>553</b>
<b>Train + Test</b>		<b>1415</b>

Train

Test

NSCLC: Non-Small Cell Lung Cancer; p = number of parameters, N: number of patients treated with atezolizumab (patients from French centers were excluded for legal reasons (N=118)); In total, data from 1074 patients from OAK were used as Test set (553 from the ATZ arm, 521 from the DTX arm); PD: Pharmacodynamic; SLD: Sum of the Largest Diameters. CRP: C Reactive Protein; LDH: Lactate Dehydrogenase.

1. Fehrenbacher L *et al.* Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet* (2016)
2. Fehrenbacher L *et al.* Updated Efficacy Analysis Including Secondary Population Results for OAK: A Randomized Phase III Study of Atezolizumab versus Docetaxel in Patients with Previously Treated Advanced Non-Small Cell Lung Cancer. *Journal of Thoracic Oncology* (2018)
3. Solange Peters *et al.* Phase II Trial of Atezolizumab As First-Line or Subsequent Therapy for Patients With Programmed Death-Ligand 1-Selected Advanced Non-Small-Cell Lung Cancer (BIRCH). *JCO* (2017)
4. Spigel D.R *et al.* FIR: Efficacy, Safety, and Biomarker Analysis of a Phase II Open-Label Study of Atezolizumab in PD-L1-Selected Patients With NSCLC. *Journal of Thoracic Oncology* (2018)

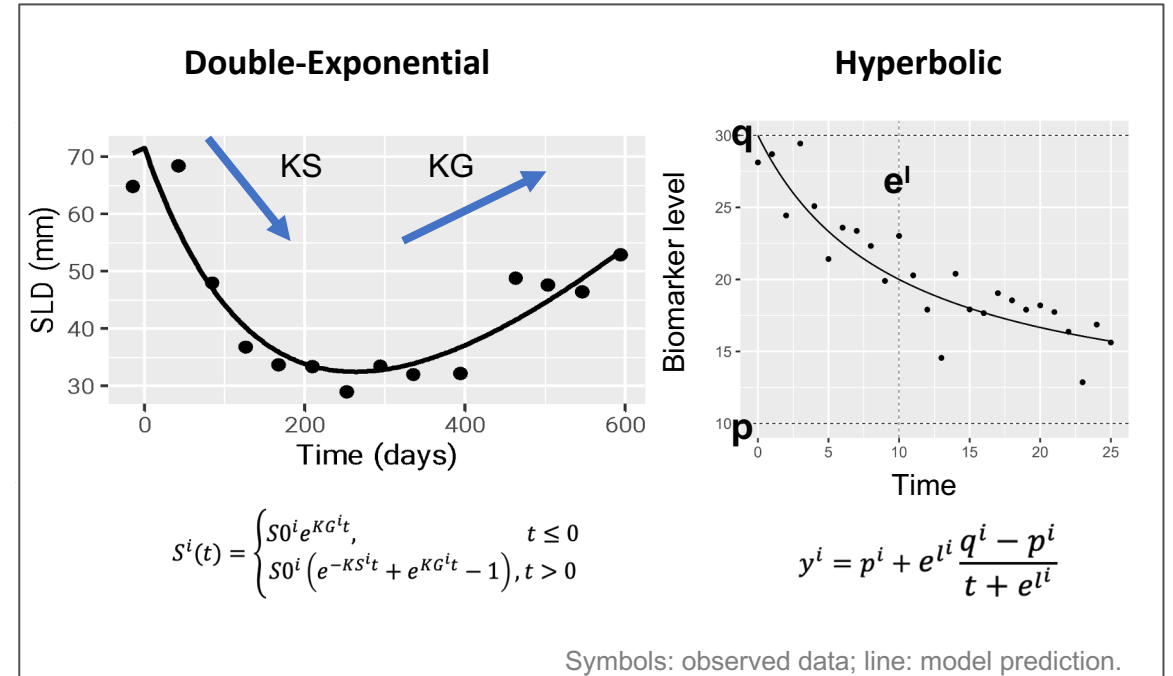
- Tumor kinetics (TK) : **double-exponential** model<sup>1,2</sup>
- PD time courses : **empirical models**<sup>3</sup>
  - constant? linear? hyperbolic? double-exponential?
- Statistical Nonlinear Mixed Effect (**NLME**) model
- Observation model : constant (TK) or proportional (PD)

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

- Inter-individual

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

- Population **parameters** : SAEM algorithm for likelihood maximization
- **Individual empirical Bayes estimates (EBEs)** from the maximum a posteriori estimator
- Fits performed using the R Monolix2020R1 API



- ★ **Individual TK and PD model parameters** = inputs of the Machine Learning (ML) algorithm
- ★ Model-derived **baseline parameters were excluded** (baseline markers already in clinical category)

Is there any kinetic pattern in the PD data?

TK

LDH

Neutrophils

CRP

Albumin

Best Overall Response

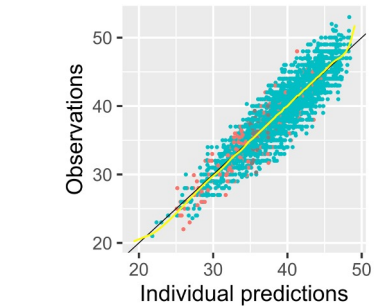
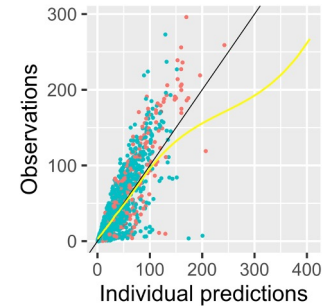
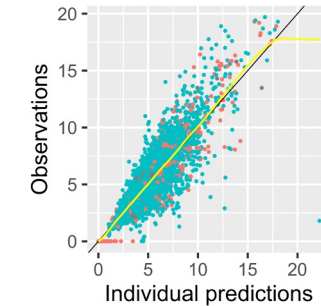
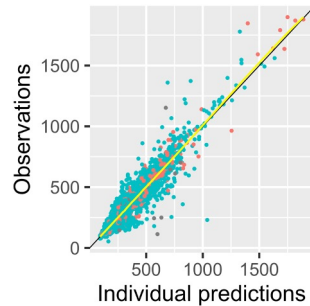
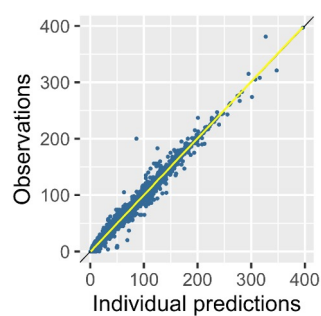
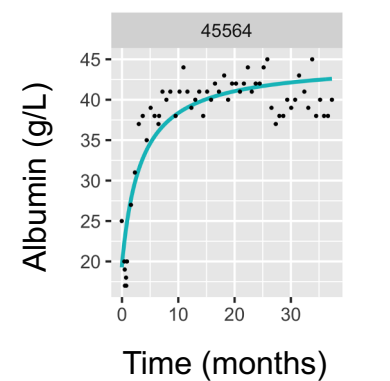
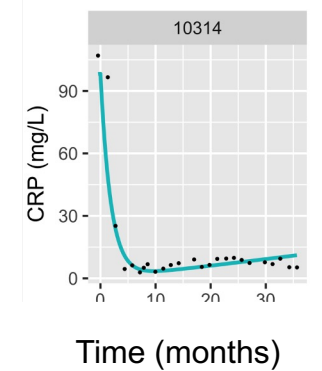
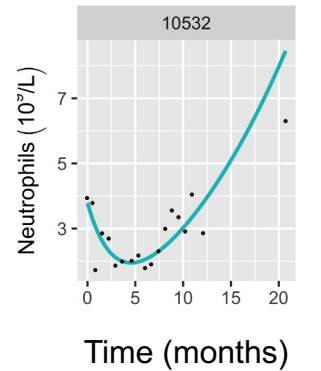
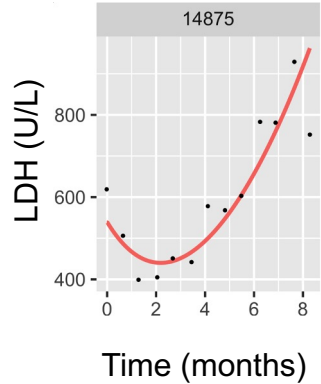
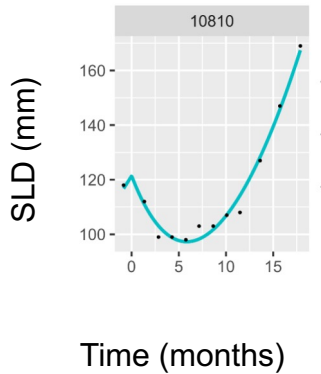
CR+PR+SD  
PD

model	BICc	b
<b>dexp</b>	39,886	0.56
hyperbolic	40,915	0.62
linear	42,462	0.70
constant	42,982	0.74

model	BICc	b
<b>dexp</b>	102,449	0.14
hyperbolic	102,943	0.14
linear	105,193	0.17
constant	106,249	0.18

model	BICc	b
<b>dexp</b>	28,764	0.21
hyperbolic	29,712	0.22
linear	30,020	0.23
constant	31,332	0.25

model	BICc	b
<b>hyperbolic</b>	48,007	0.056
dexp	48,395	0.058
linear	49,436	0.063
constant	49,724	0.065

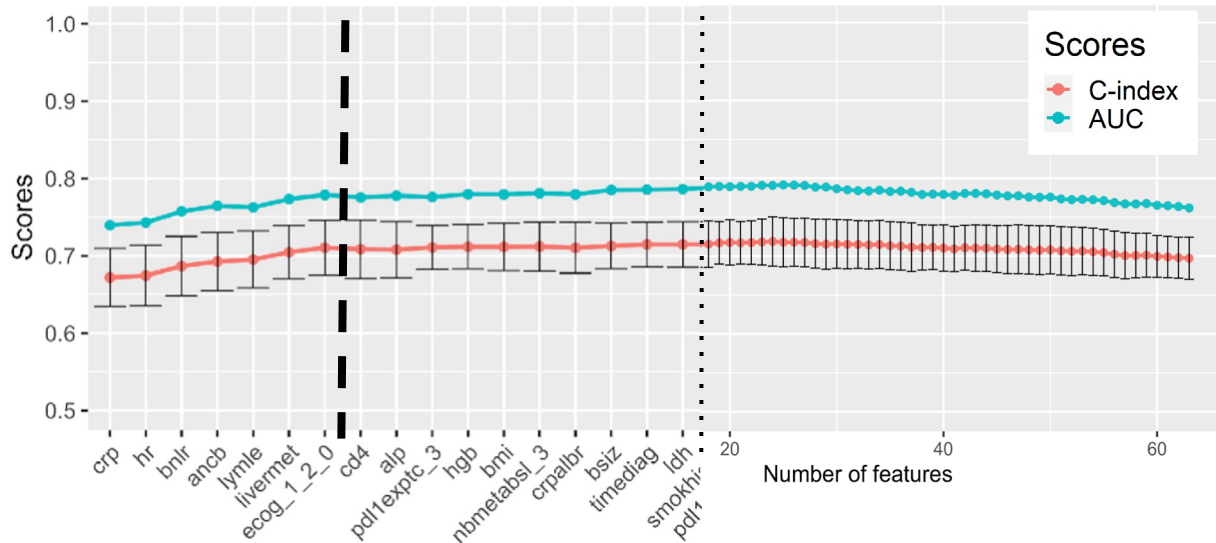


- **Preprocess** to handle missing data  
Drop zero-variance or >25% NA columns, dummification, **NA imputation**, scaling
- **Dimensionality reduction** for FMI and **RNAseq data** (bootstrap LASSO)
- **Features selection**:
  - 5 methods: **LASSO**, random survival forest (RSF) importance, Cox-based and stepwise forward/backward
  - 3 strategies : i) all variables, ii) per feature set and iii) pooled selected sets
- **4 survival algorithms tested**:  
Cox, Cox and accelerated failure time with gradient boosting and **random survival forest (RSF)**
- Evaluation of machine learning models
  - **Model development**: 10-fold cross-validation  
C-index, calibration curves and 12-months survival classification metrics
  - **Study-level** predictions: survival curves, **hazard ratios**

Working principle: need for a **minimal signature** model with limited number of easily measurable variables



### 7 features



- All features sorted using LASSO
- Incremental models with increasing number of features
- Minimal set of features that reaches the plateau

Minimal signature baseline characteristics (p = 11)
CRP
Heart rate
Neutrophils-to-lymphocytes ratio
Neutrophils
Lymphocytes-to-leukocytes ratio
Liver metastases
ECOG (0 vs 1)
PD-L1 (≥ 50% on tumor cells)
Hemoglobin
SLD
Lactate dehydrogenase

### 7 features

4 features added because of established prognostic/predictive value<sup>1,2,3</sup>

<sup>1</sup>Fangfang Wu *et al.* Prognostic value of baseline hemoglobin-to-red blood cell distribution width ratio in small cell lung cancer: A retrospective analysis. *Thoracic Cancer*. 2020 Apr; 11(4): 888–897

<sup>2</sup>Matthen Mathew, Rachael A. Safyan, Catherine A. Shu. PD-L1 as a biomarker in NSCLC: challenges and future directions - *Annals of Translational Medicine*. 2017 Vol 5, No 18.

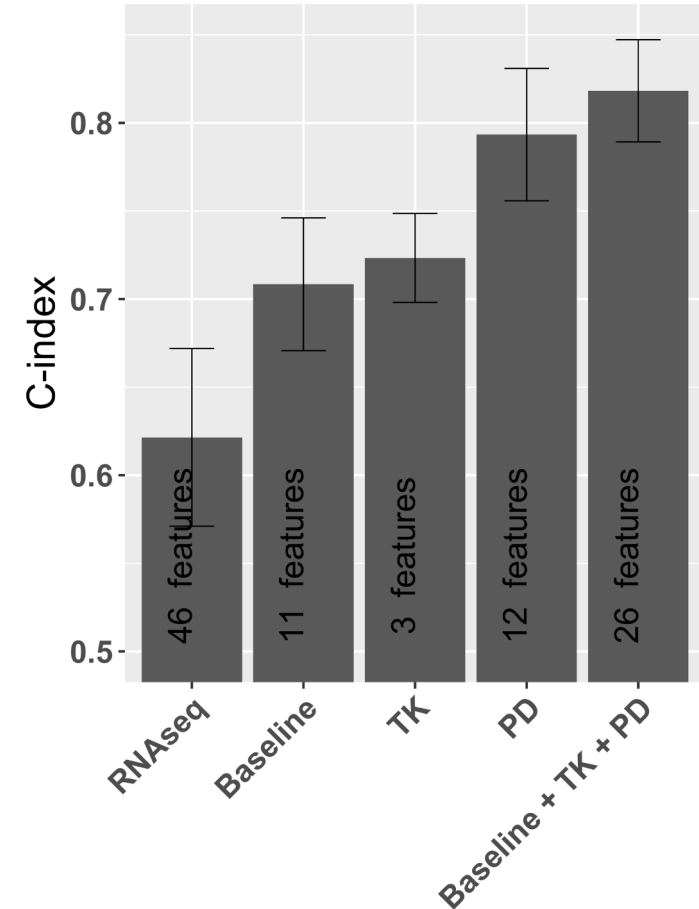
<sup>3</sup>Bernhard C D. *et al.* Long-term Survival Is Linked to Serum LDH and Partly to Tumour LDH-5 in NSCLC. *Anticancer Research* April 2010, 30 (4) 1347-1351.

# Inria Prediction metrics by features category



- Each feature set exhibits **differential predictive power**
- **RNAseq has low** individual predictive power.  
⇒ discarded
- **Model-based dynamic features** (TK, PD) outperform baseline clinical features, with much less variables
- **Model-based PD outperforms TK metrics**

Cross-validated C-indices by features category



From 10-fold cross-validation on FIR, BIRCH and POPLAR (train data set, N=559)

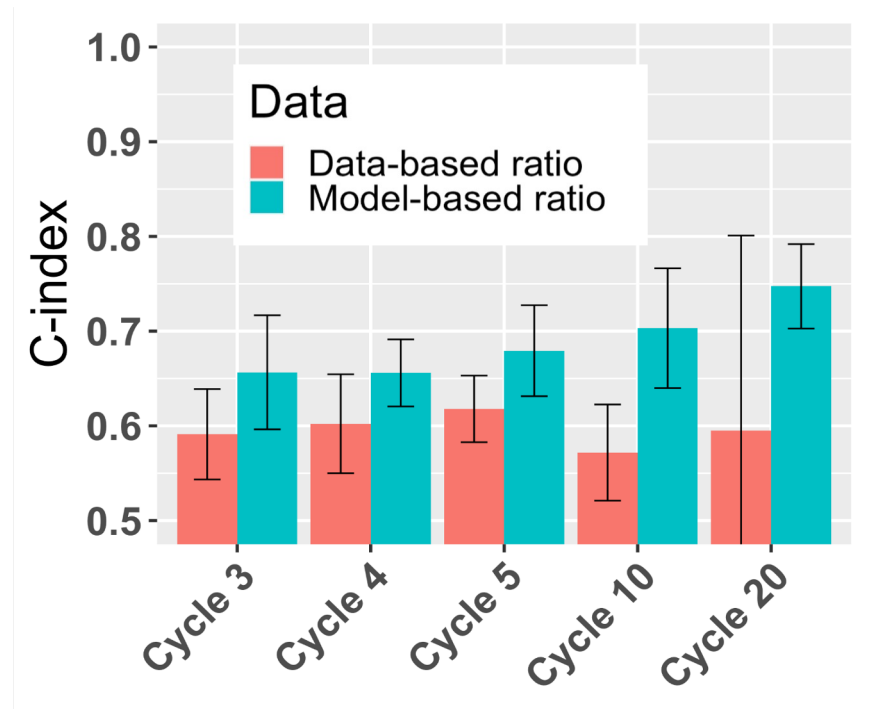
➤ A pooled model of 26 features has **very good predictive metrics**

## Comparison of predictive performance

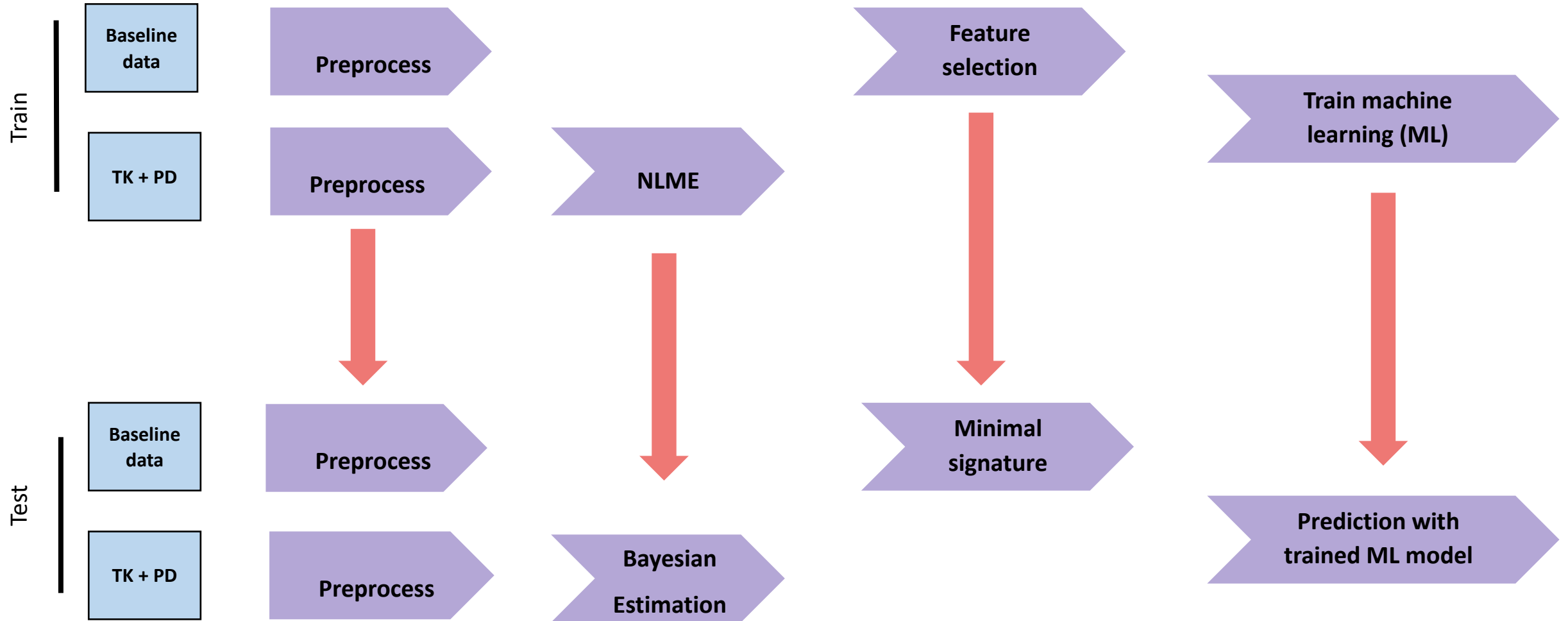
- Using SLD, CRP, Albumin, Neutrophils and LDH parameters
- Several cycle landmark times were used (Cycle 3 to Cycle 20)
- ML model learned from the train set truncated at Cycle X
- For all landmarks, the predictive power of the **data-based** versus **model-based** ratio from baseline at Cycle X pre-dose is compared

- Model-based metrics clearly have both **better predictive power** and **narrower uncertainty**
- This illustrates how dynamic modeling allows to **capture the kinetics and correct for intra-individual stochasticity (noise)**
- Model-based ratio from baseline is **more predictive with increasing number of cycles**

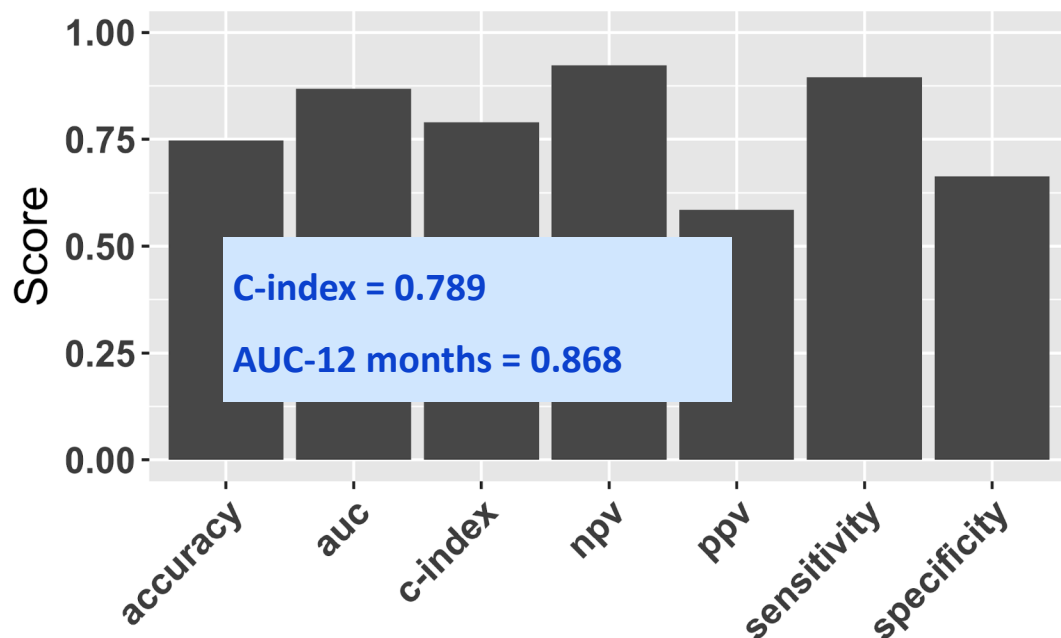
**Cross-validation c-indices for **model-** and **observed-** based ratio to baseline**



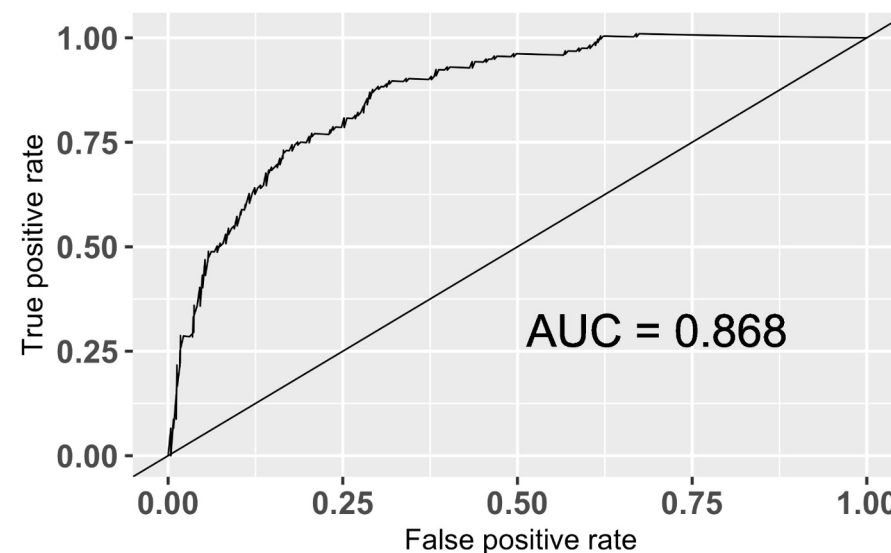
Note: 10-fold cross-validation on FIR, BIRCH and POPLAR (train data set). Only relative change from baseline used as model metric. Other model-based parameters ignored here for fair comparison



Performance metrics for **minimal signature model prediction**



Prediction of **12-months survival**



Results presented are based on full test set and are prone to immortal time bias

PDL1<sup>1</sup> : AUC = 0.65

RoPro<sup>2</sup>:

c-index = 0.69

AUC-3 months = 0.817

Train: 559 patients. Test: 396 patients. Performance metrics using full Test set (N= 391 patients). All metrics are computed at 12 months. Positive (1)= death, negative (0)= Alive; NPV: negative predictive value (NPV=TN/ (FN+TN); TN=true negative; FN=false negative; PPV: positive predictive value (PPV=TP/(TP+FP); TP= true positive; FP = False positive. To compute accuracy, censored patients were excluded (i.e., 17/396 patient at 12 months)

<sup>1</sup> Rizvi, H. et al., J Clin Oncol Molecular Determinants of Response to Anti-Programmed Cell Death (PD)-1 and Anti-Programmed Death-Ligand 1 (PD-L1) Blockade in Patients with Non-Small-Cell Lung Cancer Profiled with Targeted Next-Generation Sequencing. J. Clin. Oncol. 2018, 36, 633-641.

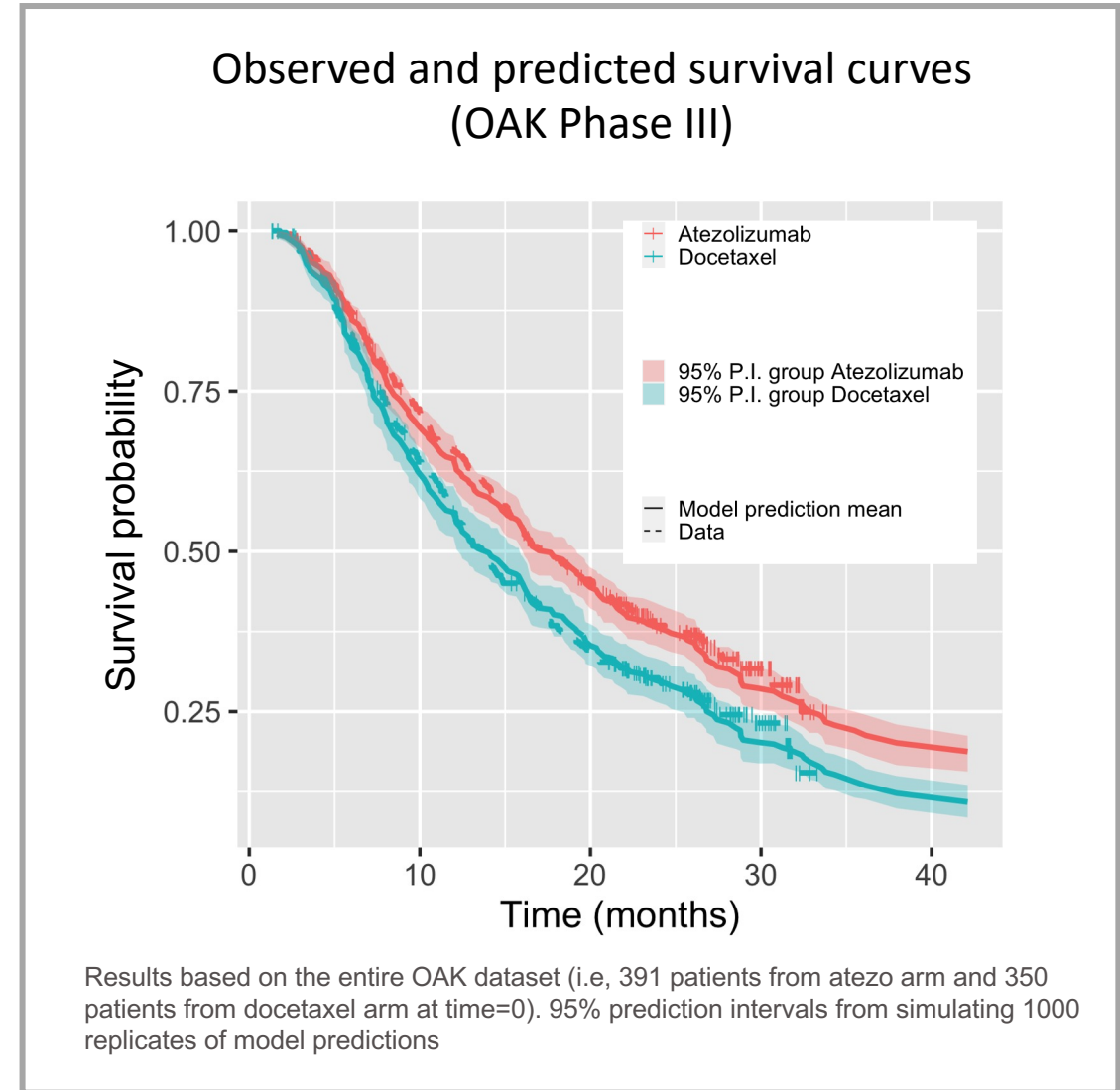
<sup>2</sup> Becker, T. et al. An enhanced prognostic score for overall survival of patients with cancer derived from a large real-world cohort. Ann Oncol 31, 1561-1568 (2020).

# The minimal signature model reproduced retrospectively the survival curves of atezolizumab and control arm in the OAK phase III trial

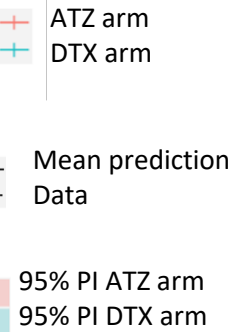
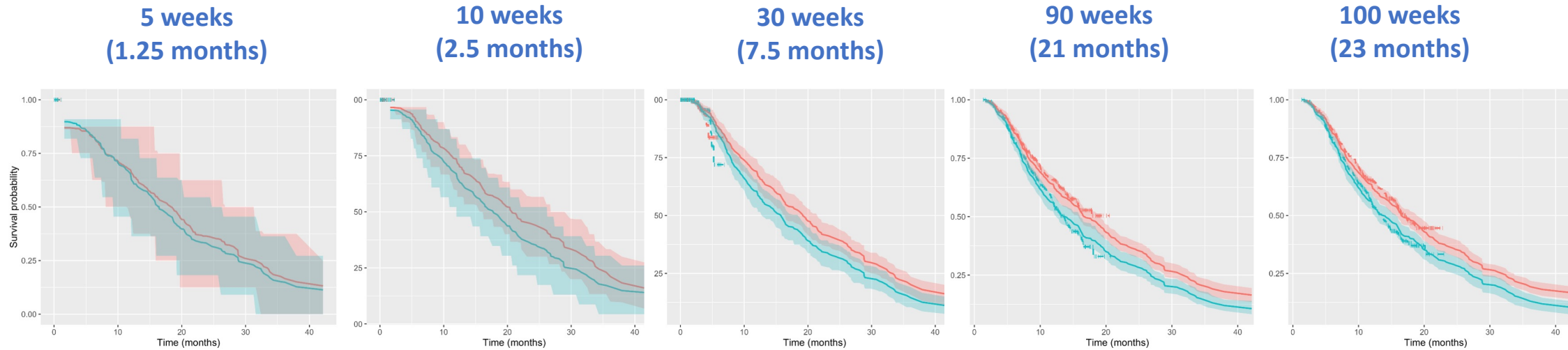
- The model **predicted well** survival curves of ATZ and control arm (docetaxel) from OAK
- The model was able to **predict ATZ survival benefit** over chemotherapy in OAK (HR < 1 with good match between observed and predicted HR)

Observed HR (95%CI)	D-Light Prediction HR (95%PI)
0.765 (0.64 - 0.913)	0.765 (0.692-0.829)

- The relationship between TK and PD metrics and survival is **not drug specific**



# The minimal signature model predicts OS benefit for atezolizumab over docetaxel in OAK 7.5 months after first patient randomized



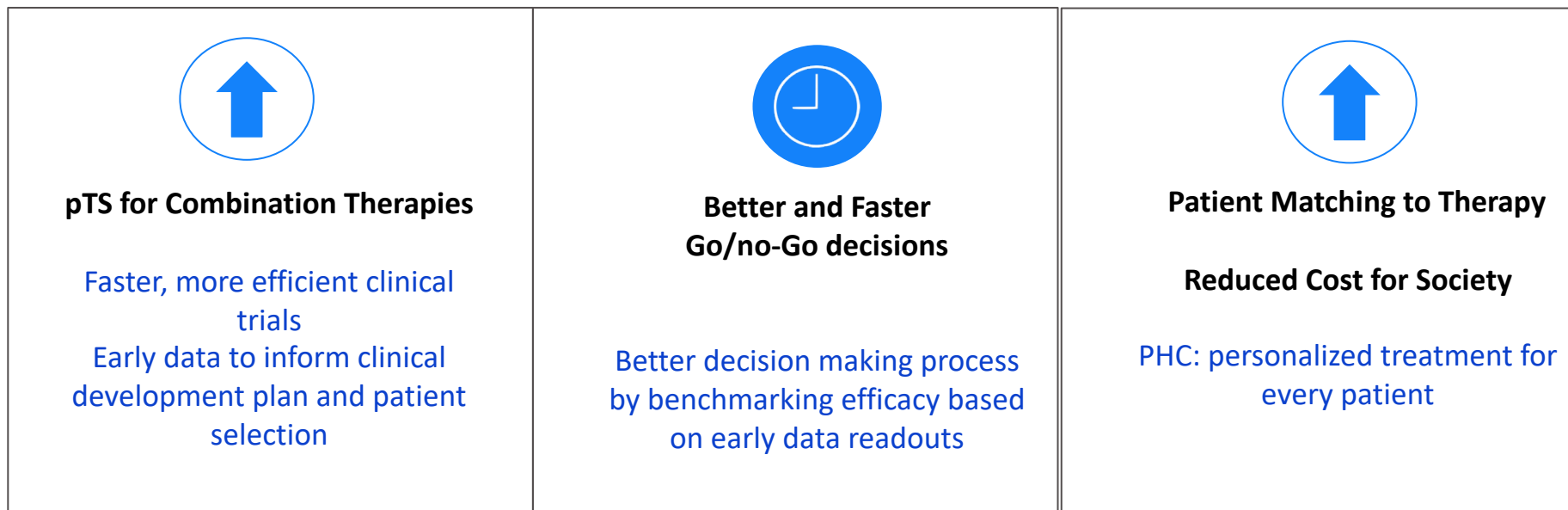
Data HR (95% CI)	NA (NA - NA)	<b>NA (NA - NA)</b>	<b>1.04 (0.386 - 2.79)</b>	<b>0.708 (0.564 - 0.887)</b>	0.748 (0.606 - 0.922)
Predicted HR (95% PI)	0.923 (0.369 - 2.5)	<b>0.836 (0.547 - 1.4)</b>	<b>0.802 (0.655 - 0.907)</b>	<b>0.809 (0.746 - 0.903)</b>	0.798 (0.715 - 0.89)
Number of patients (DTX - ATZ)	11 - 8	23 - 30	163 - 183	352 - 386	352 - 386
Median nb of data points (TK/Alb/CRP/LDH/Neutrophils)	1/2/2/2/2 - 1/2/1/2/1	1/3/3/2/2.5 - 1/2/2/2/2	2/4/4/4/5 - 2/4/4/4/4	4/8/6/7/8 - 4/10/9/9/10	4/8/6/7/8 - 4/10/9/9/10

Truncation of data is based on the randomization date of the first patient treated in OAK (e.g. at 5 weeks after randomization, 12 patients received atezolizumab (ATZ), 11 docetaxel (DTX); and median TK data points were 2 and 1 for docetaxel and atezo arms respectively)

# Summary and applications for early drug development



- A **minimal signature** NLME-ML model of **26 features** (clinical, TK and PD model features)
  - **Baseline**: CRP, Heart rate, Neutrophils-to-lymphocytes ratio, Neutrophils, Lymphocytes-to-leukocytes ratio, Liver metastases, ECOG (0 vs 1), PD-L1 ( $\geq 50\%$  on tumor cells), Hemoglobin, SLD, LDH
  - **Longitudinal**: tumor kinetics (SLD), albumin, CRP, LDH, neutrophils
- Analysis (preprocess, feature selection, CV, train, predict) fully automated in a **R package** (> 12,000 lines of code)
- Could be **applied to early phase data** to assess the decision to move an asset to a later phase of development
- Potential **extrapolation to other drugs** within the same disease setting;





Doing now what patients need next