TIPIT project

Prediction of the patient's outcome in Non-small cell lung cancer from baseline 18F-FDG PET scans

institut Curie

Inserm
La science pour la santé
From science to health

LITO

Outlines

1. Introduction

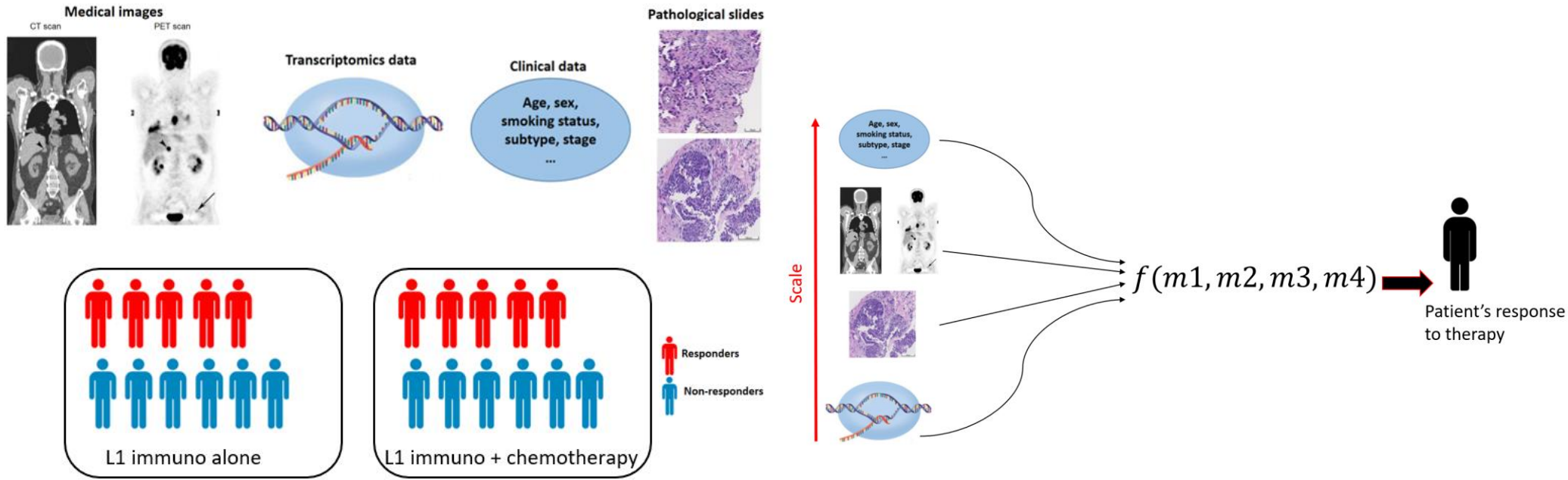
2. A basic example of multimodal Machine Learning with TCGA

3. Simple ML strategies for baseline PET predictions

4. An interpretation tool at the lesion level based on Shapley values

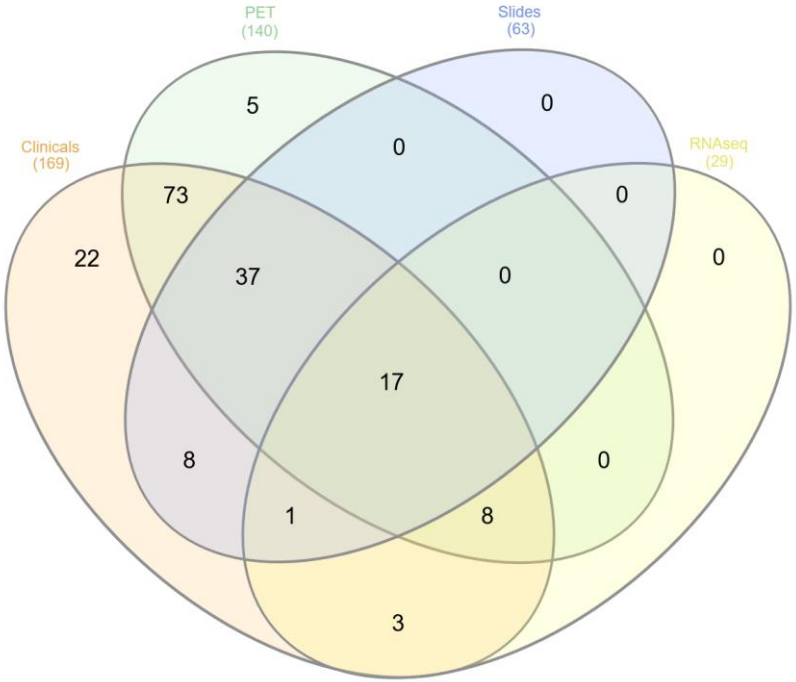
TIPIT project – multimodal predictions in NSCLC

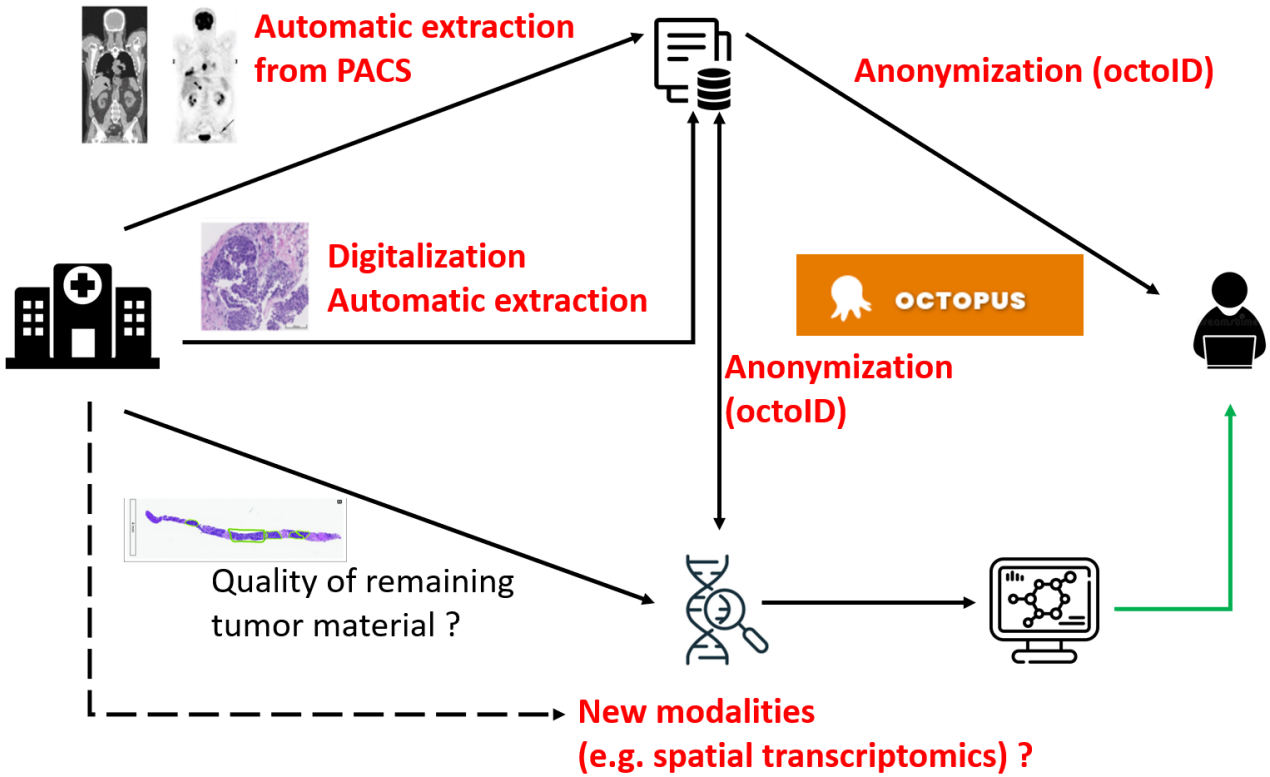
Can we use multimodal supervised Machine Learning to build new powerful and interpretable biomarkers for the response to immunotherapy for metastatic NSCLC patients ?



*Like many terms in Machine Learning "multimodality" is used for a lot of different things and by a lot of different people

The difficult challenge of building a multimodal cohort



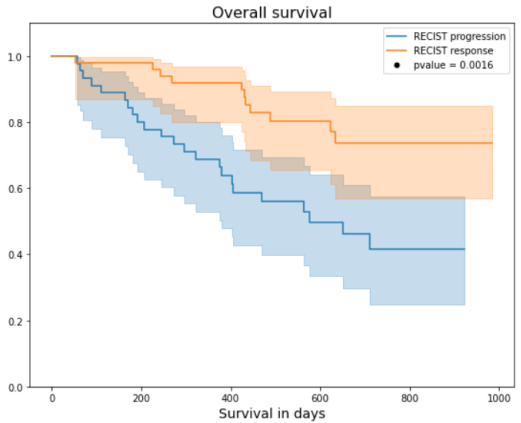


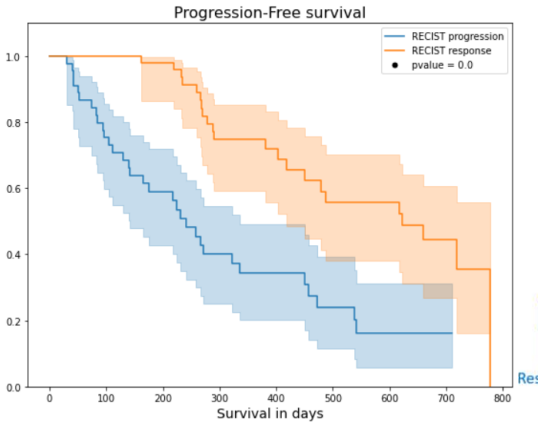
The extraction pipeline is composed of many steps and obstacles we do not have much impact on them

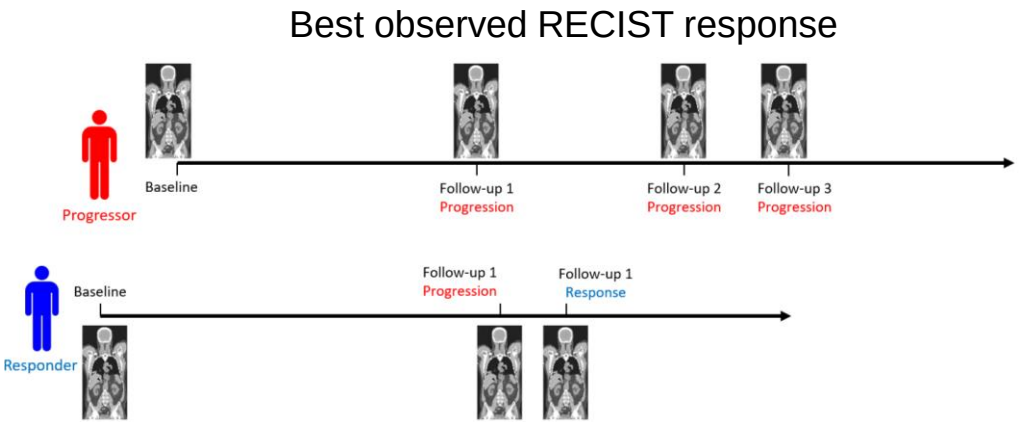
A more realistic use of Machine Learning in the TIPIT project

- With all the patients treated with immunotherapy, we cannot assess whether each patient benefitted from the treatment or not.

- We are more likely to build **prognostic biomarkers** (independent from the treatment) than **predictive biomarkers**







The predictiveness of our models should be assessed in further experiments (e.g clinical trials)

Outlines

1. Introduction

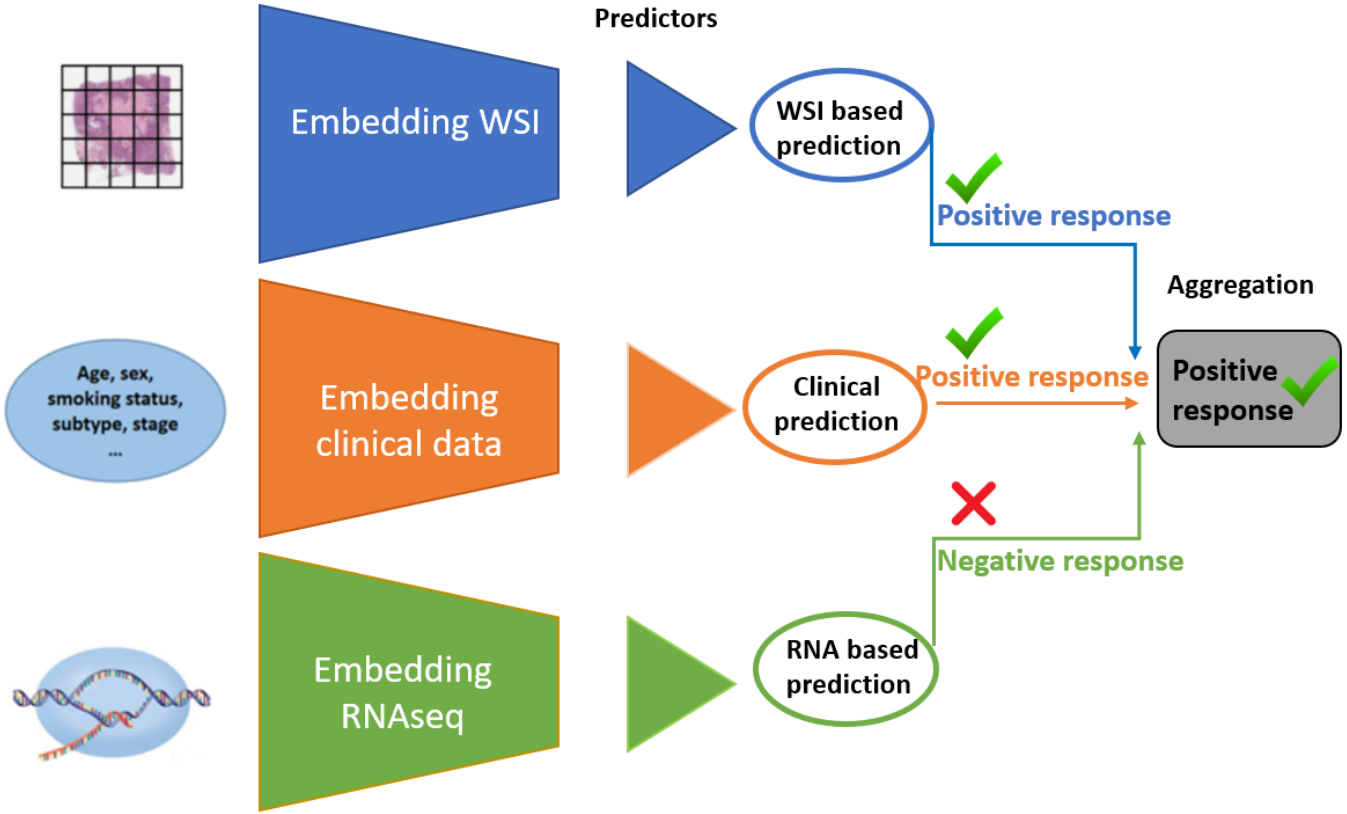
2. A basic example of multimodal Machine Learning with TCGA

3. Simple ML strategies for baseline PET predictions

4. An interpretation tool at the lesion level based on Shapley values

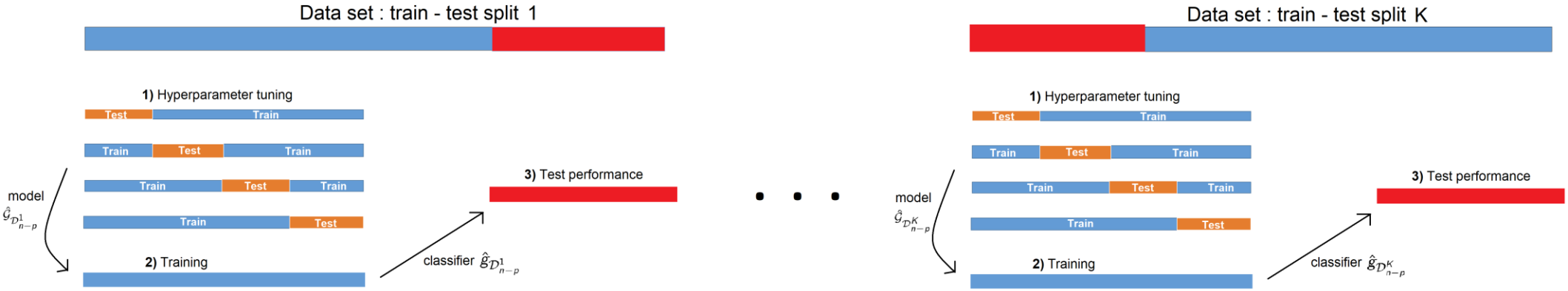
A late fusion scheme for TCGA adenocarcinoma patients

Can we predict the overall survival (OS) of **460 TCGA patients** from transcriptomic data, pathological slides (both extracted from solid biopsies) and clinical data ?



- Each modality is trained **independently** from the others
- Their predictions are **aggregated**

Combine hyperparameter tuning and cross-validation

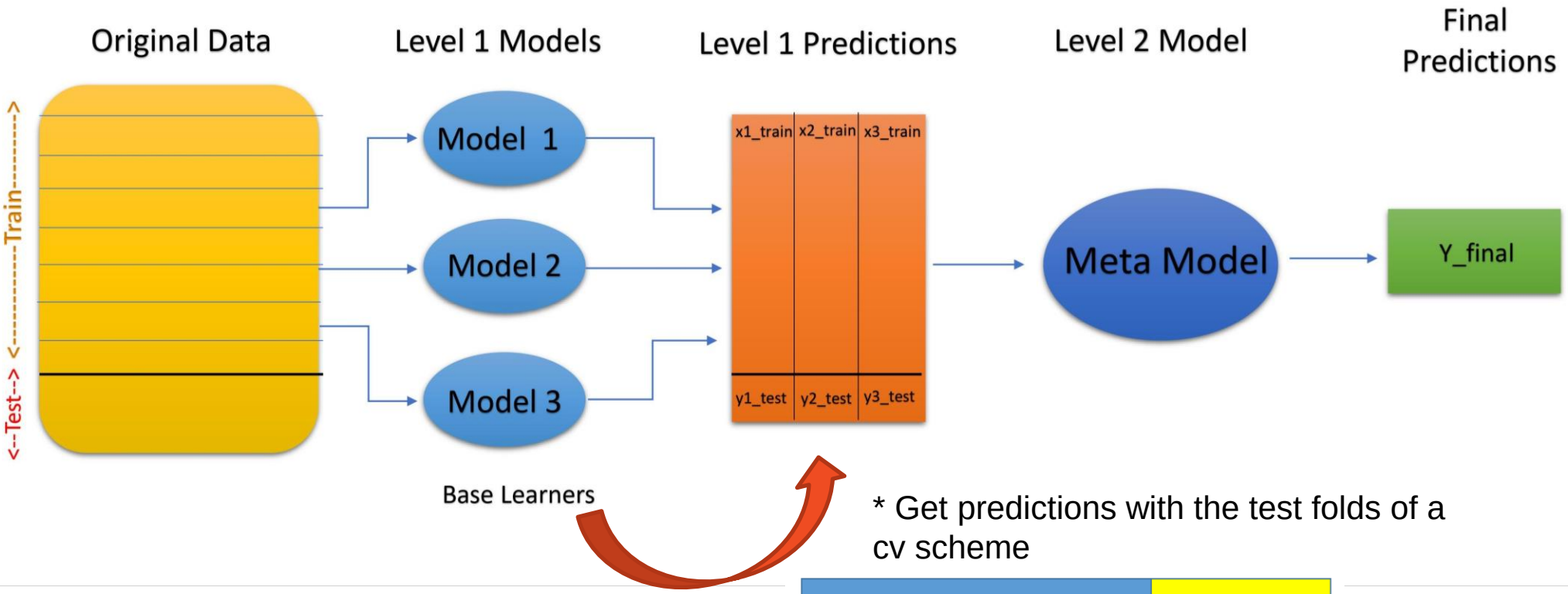


- The final result gives us and estimate of the performance we could get if we perform **hyperparameter tuning + training on the data set with the optimal hyperparameters** on the **full data set**

- It allows us to perform **stability analysis** (e.g feature selection)

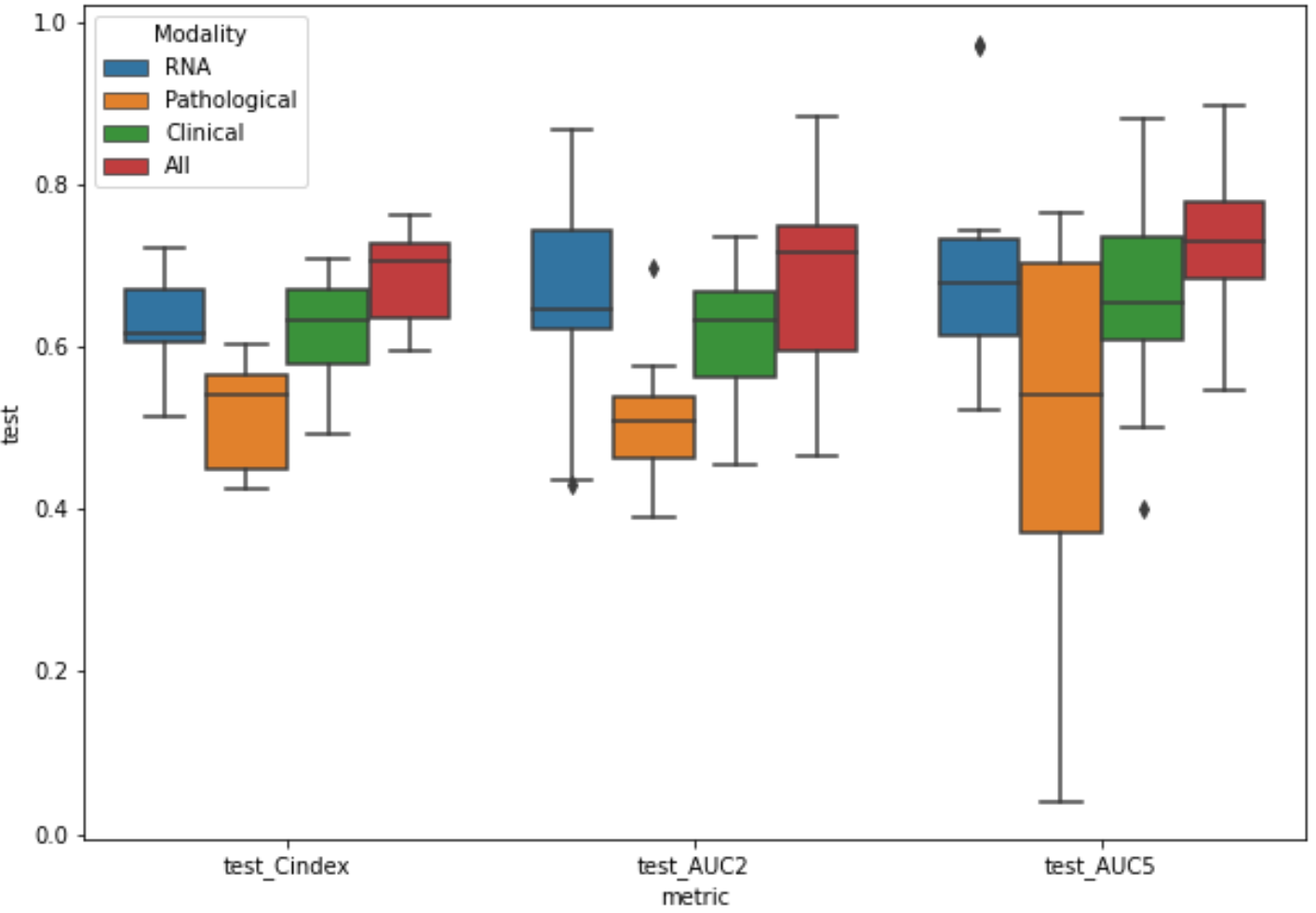
- We can apply more complex schemes than simple CV (e.g Monte-Carlo CV, repeated K-fold CV…)

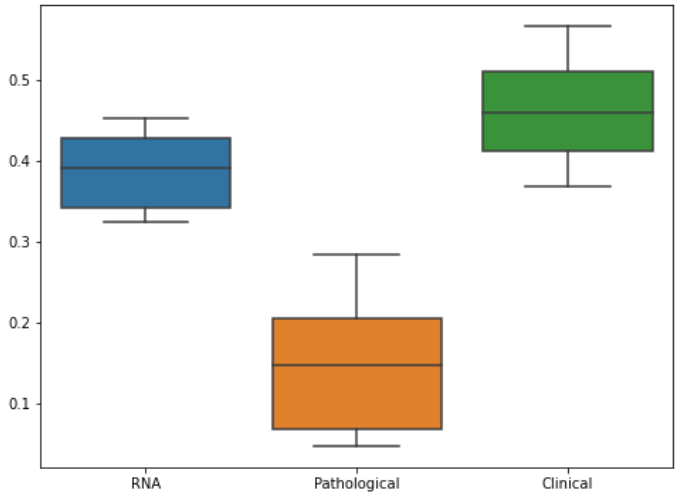
Stacking/Meta Learning for weighted aggregation



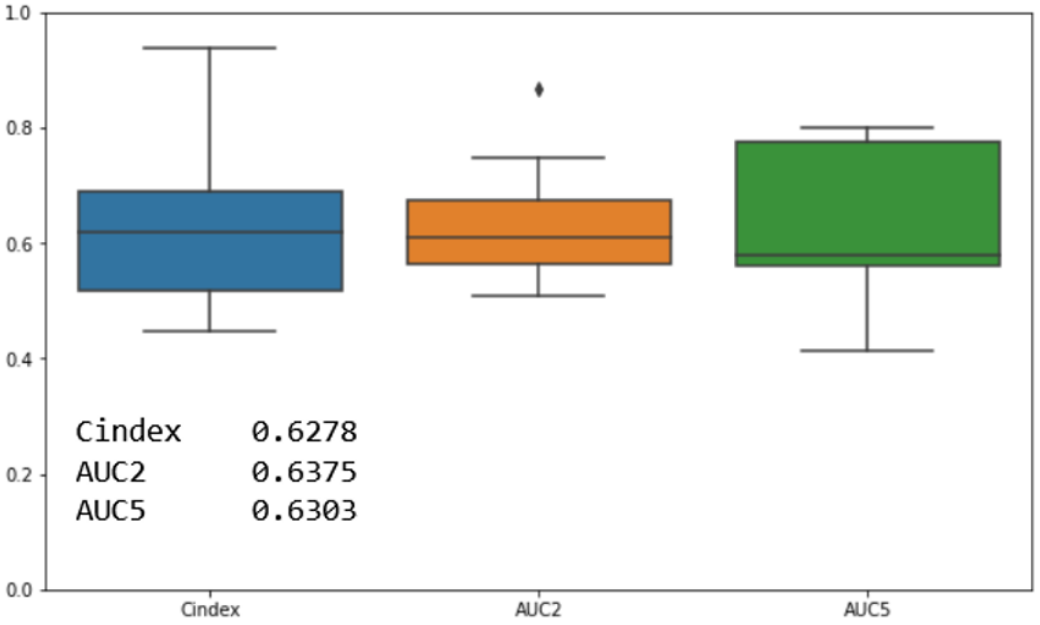
* Get predictions with the test folds of a cv scheme

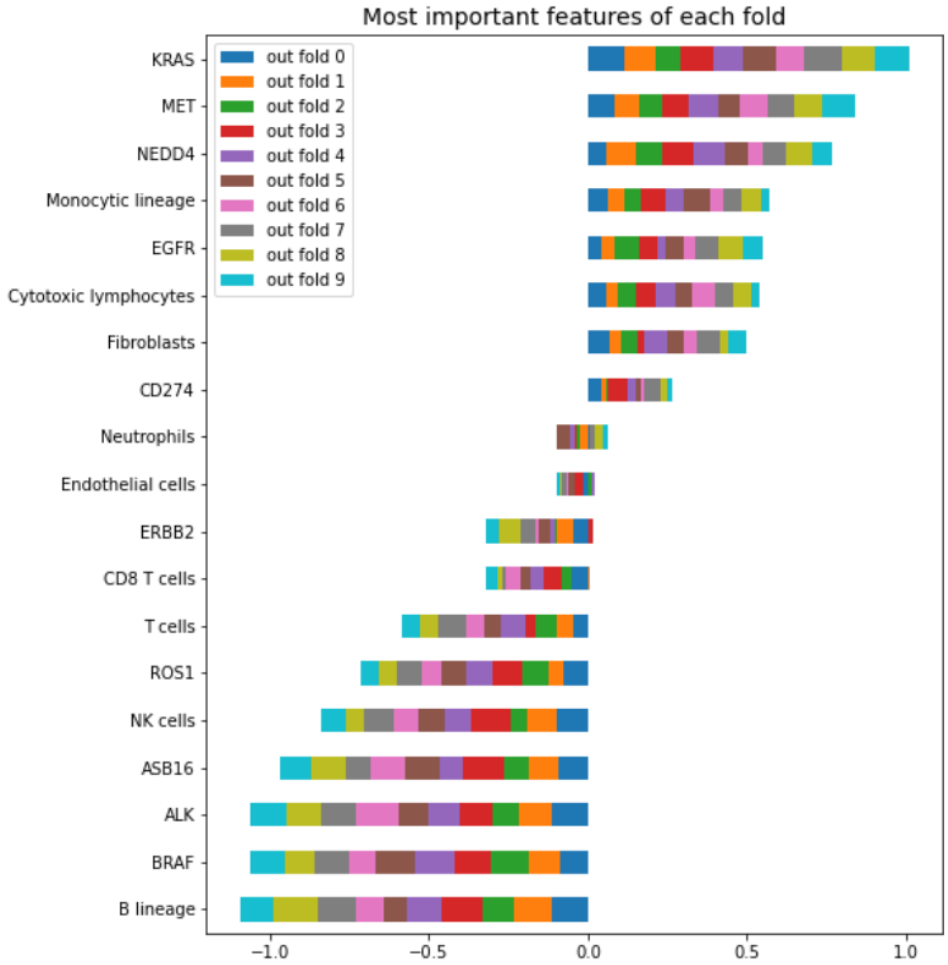
Results – Late fusion improves performances





Results – Interpretation of the transcriptomic model





Outlines

1. Introduction

2. A basic example of multimodal Machine Learning with TCGA

3. Simple ML strategies for baseline PET predictions

4. An interpretation tool at the lesion level based on Shapley values

A brief overview of our images

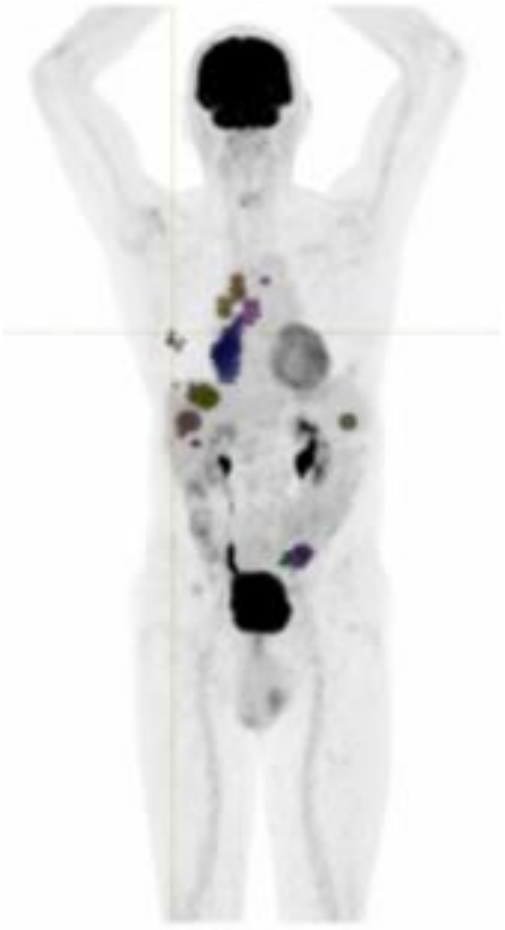
100 TIPIT patients with baseline PET

Manual segmentation and anatomical annotation by Marie

+

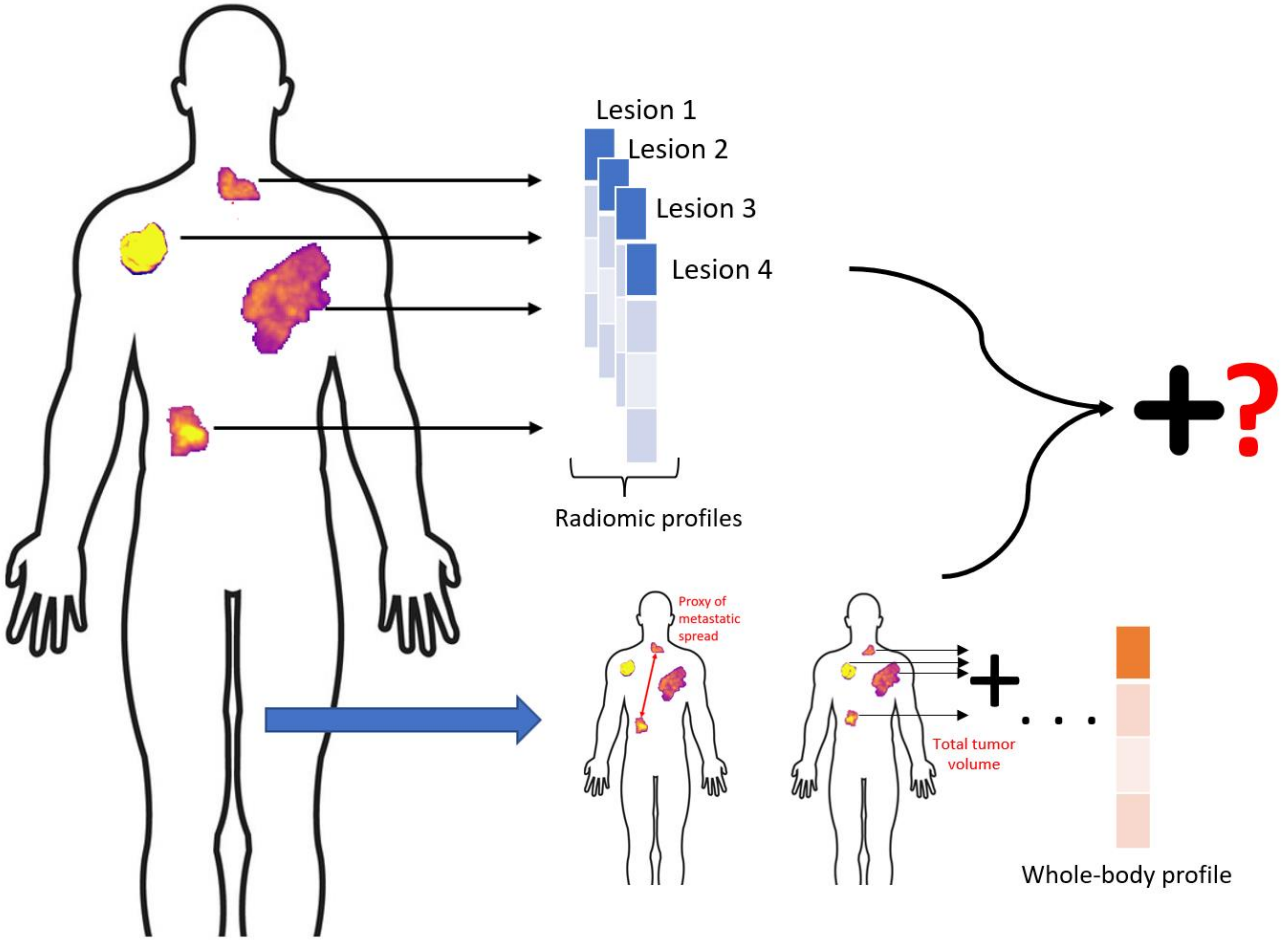
Manual corrections by Narinée and I

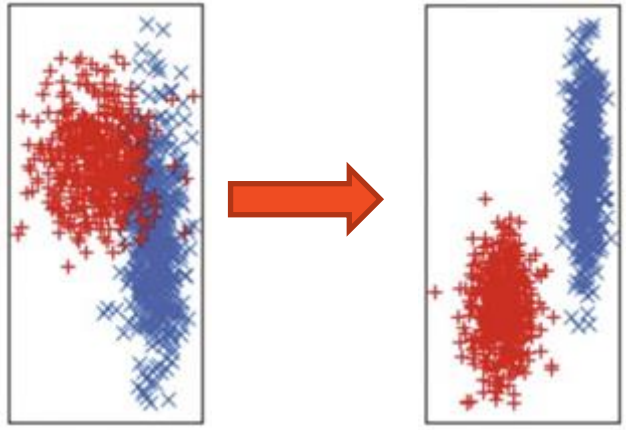
- Remove outlier voxels
- Add Necrosis
- Separate spatially disjoint areas
- Correct and harmonize annotations
- Redo some ROIs from scratch



* For all the lesions of all the patients we extracted radiomics features using a custom wrapper of Pyradiomics **(available on LITO GitHub)**

Leverage all the metastases of each patient





Improve sample representation by leveraging richer information

Different classes of hand-crafted aggregated features

Binary variables for the presence of metastases in different anatomical sites

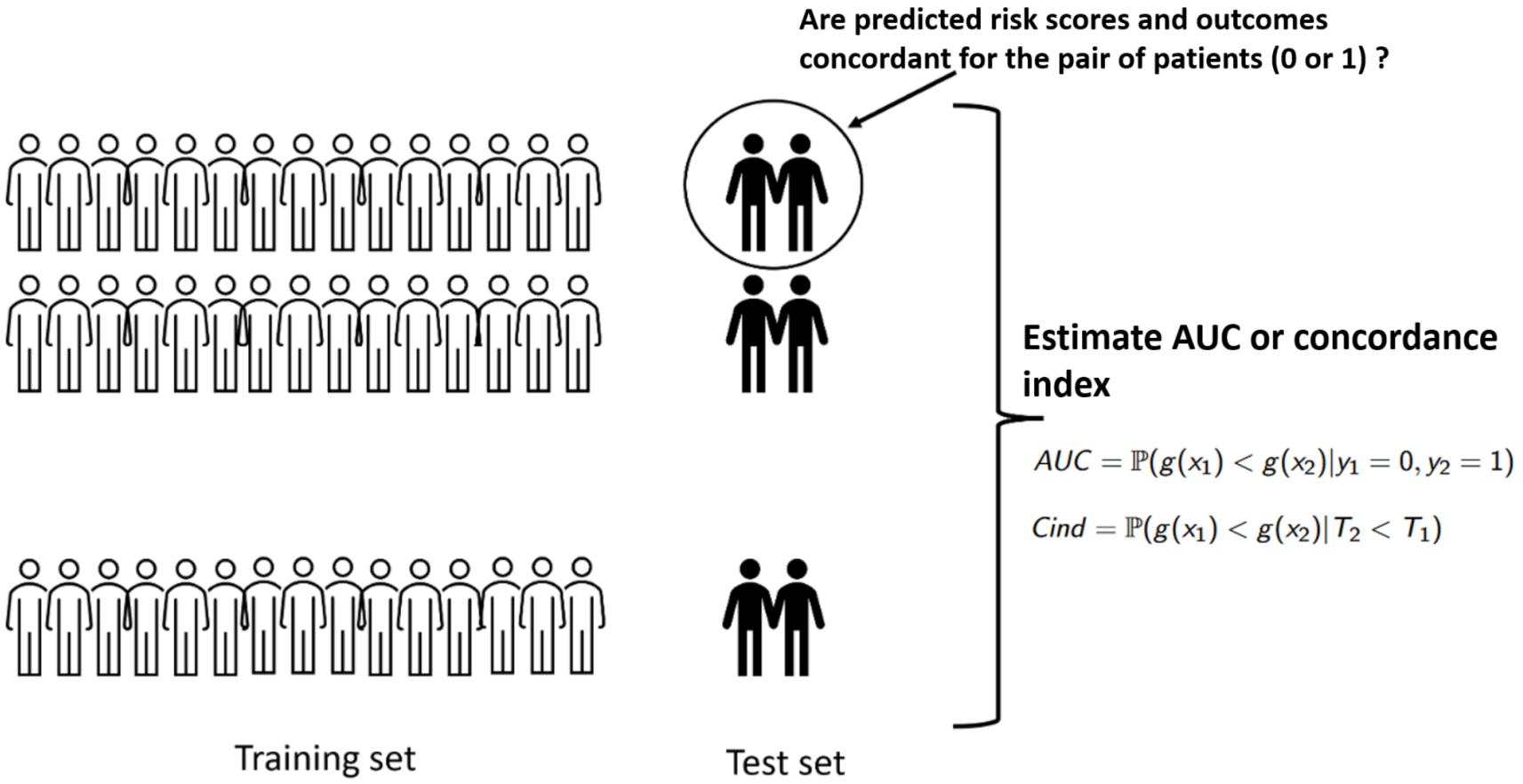
Base whole-body model (Total volume, Dmax, liver and spleen metabolism…)

Clinical variables (sex, age, BMI, Pack-year, PDL1, oncogene mutations, histology…)

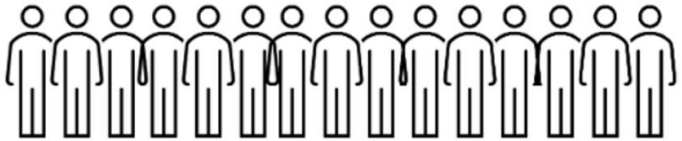
Volume and SUVmax aggregated at the TNM localizations level

Aggregation of radiomic features measured on each lesion (mean, max, std) – Multiple Instance Learning

LPOCV scheme to train an test with a maximum of patients



Late fusion strategy to keep things simple



Model 1

Model 2

Model 3

Model 4

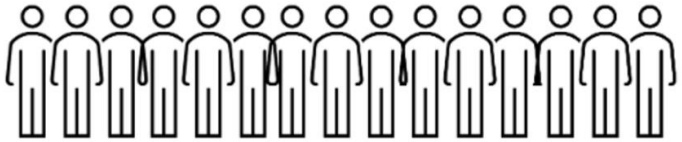
Model 5

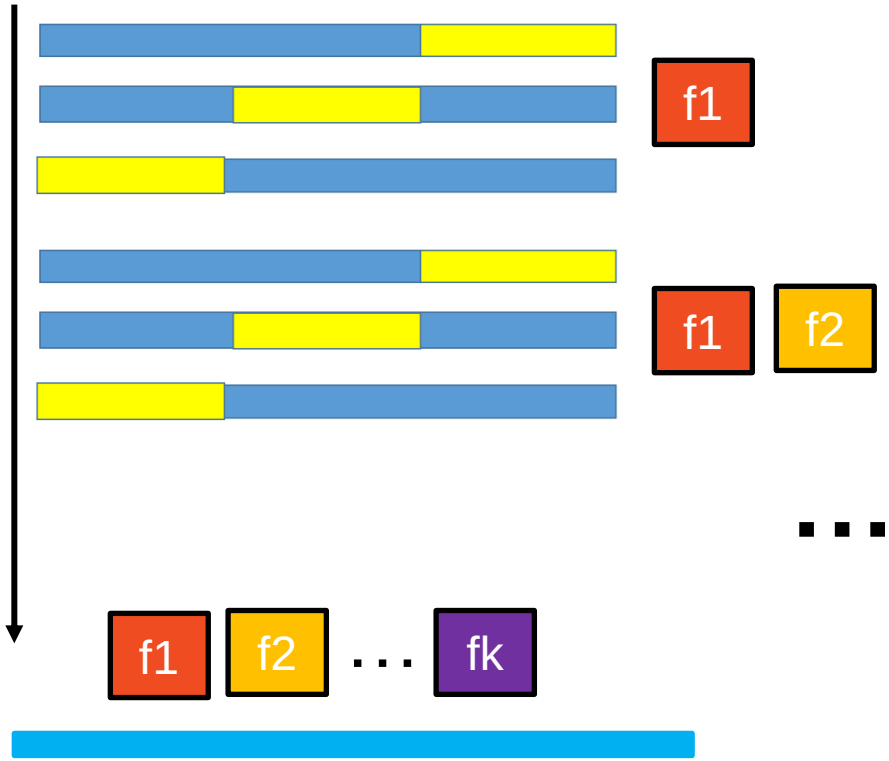
Cox's model + learn **mean** and **std** of the risk scores of each model

Sum the standardized risk scores of each model

* Weights can be learnt by measuring the performance of each model on the training set with a cv scheme

Recursive Feature Addition strategy to keep things simple





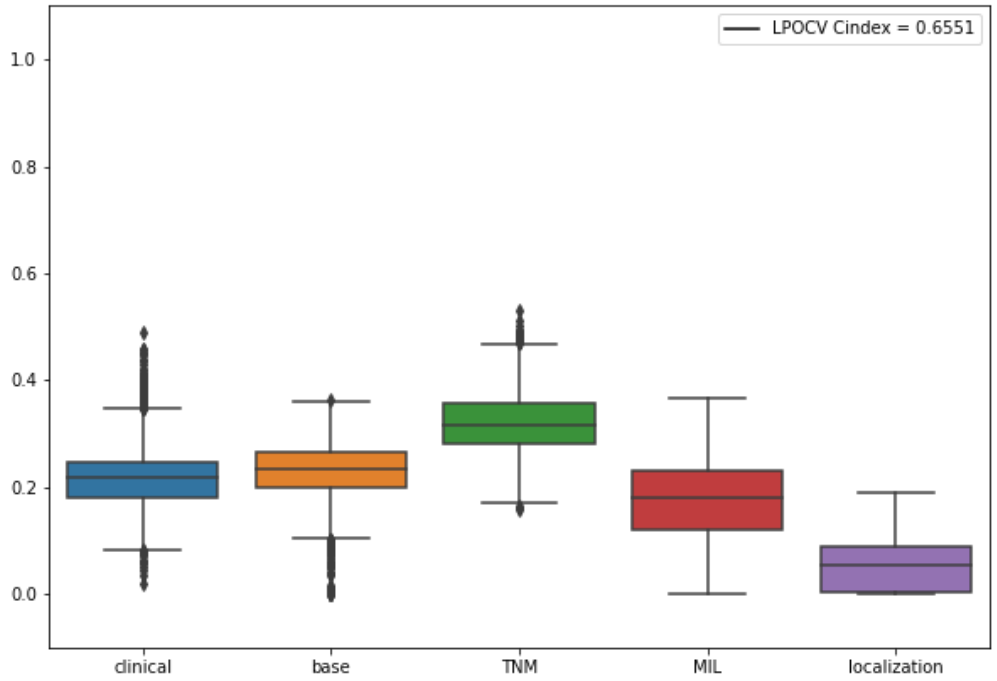
Select best feature with CV

Select best feature with CV if the gain in performance exceed a given threshold

Predict with the final model

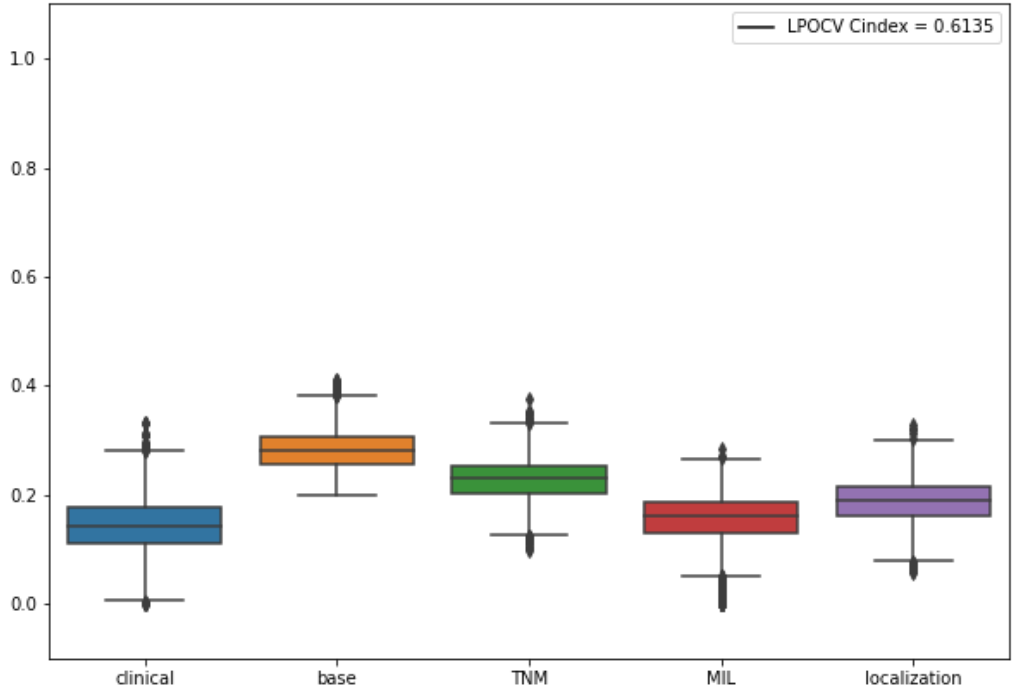
Train the final model on the full training set

Results - late fusion OS and PFS



Overall Survival

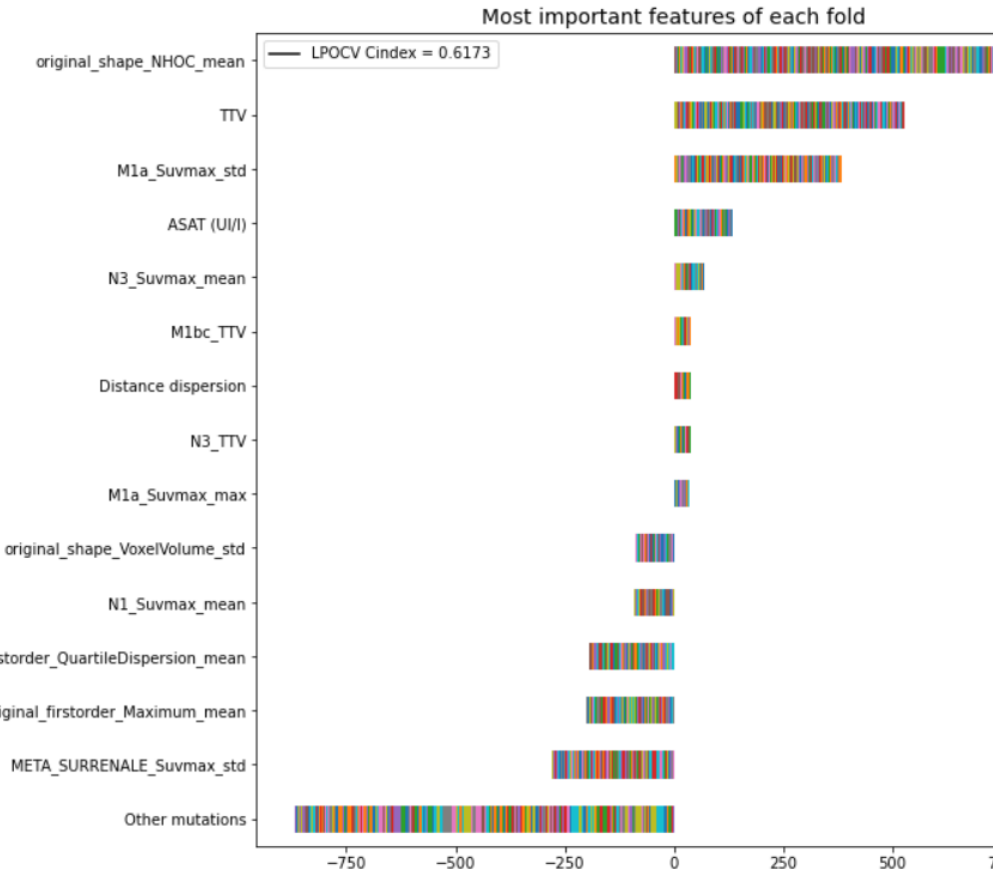
Without weights: 0.6703



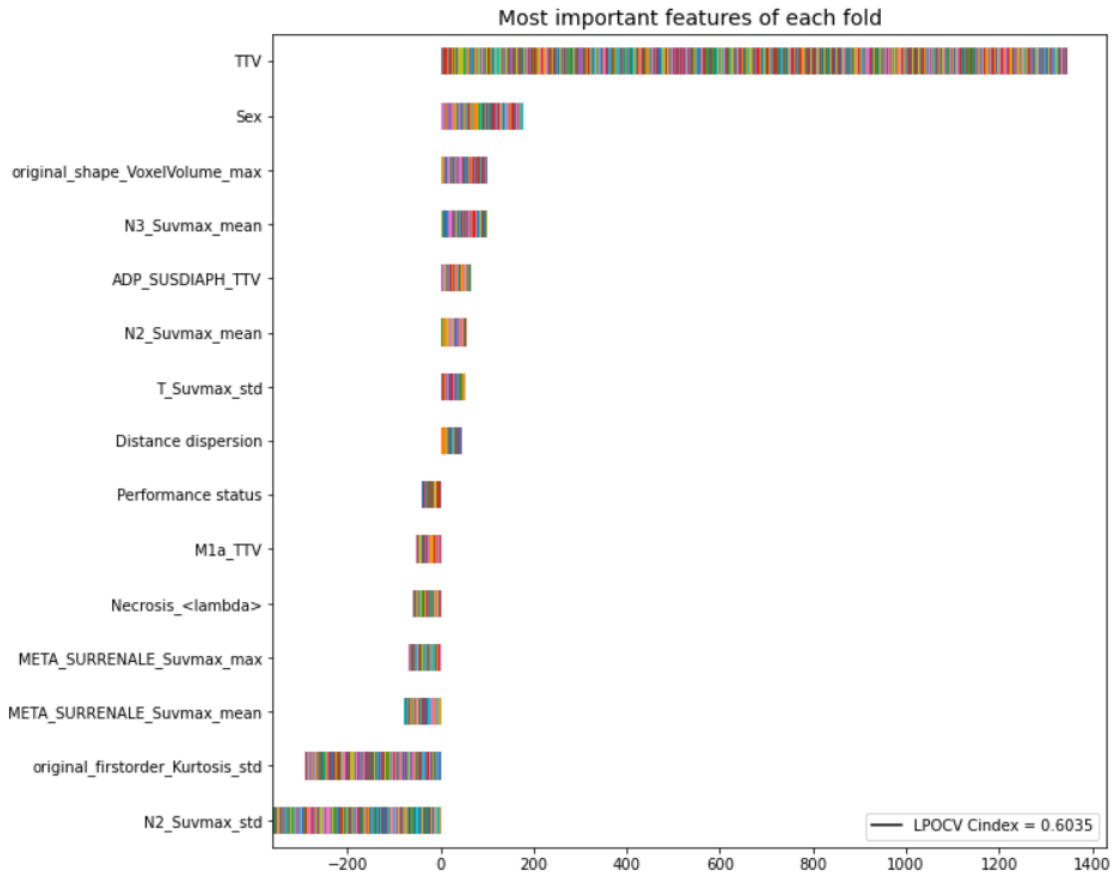
Progression-Free Survival

Without weights: 0.6346

Results – RFA OS and PFS

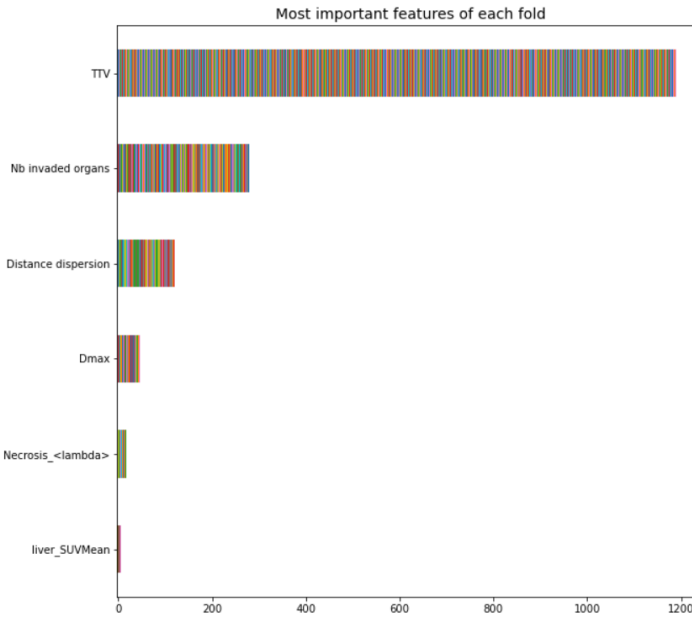


Overall Survival

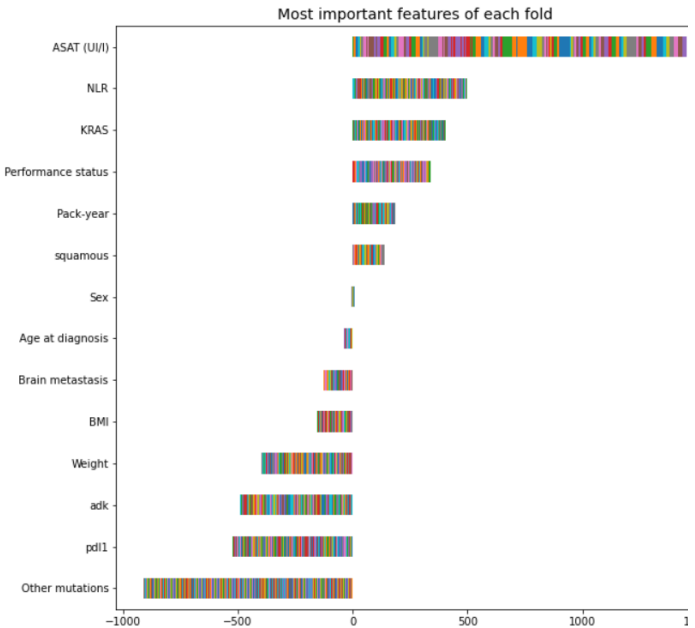


Progression-Free Survival

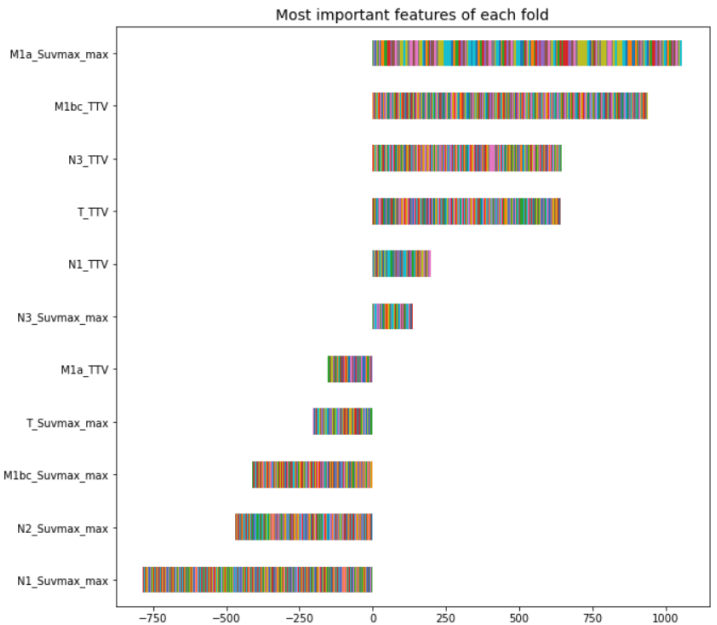
Results – feature importance late fusion OS



Base model



Clinical model



TNM model

Outlines

1. Introduction

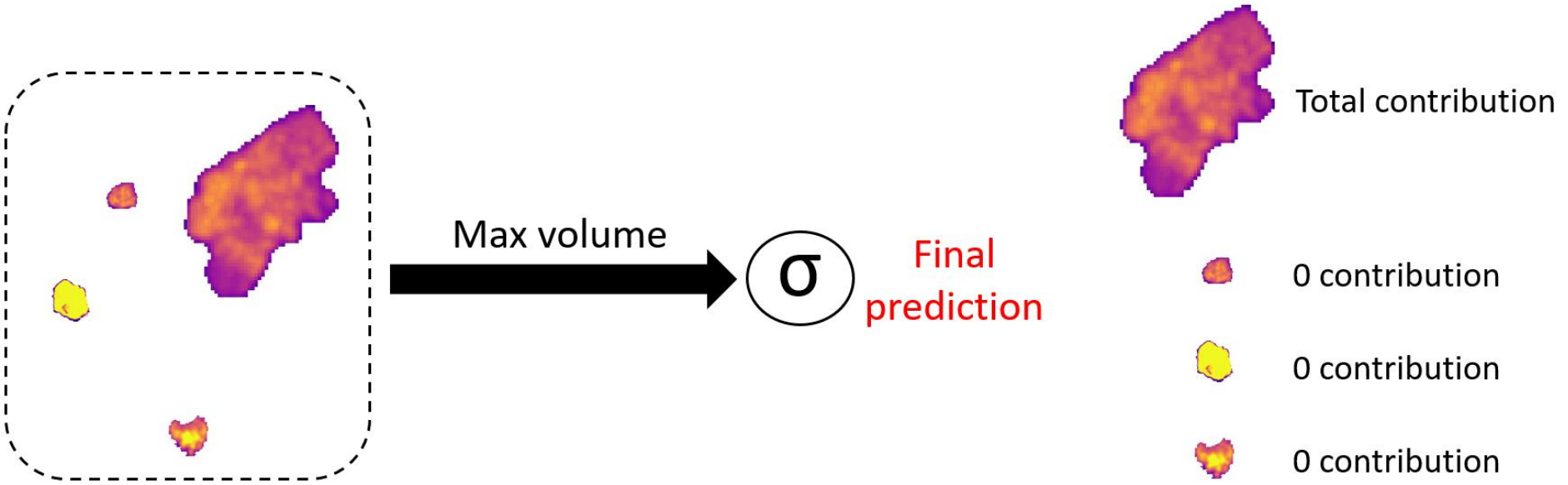
2. A basic example of multimodal Machine Learning with TCGA

3. Simple ML strategies for baseline PET predictions

4. An interpretation tool at the lesion level based on Shapley values

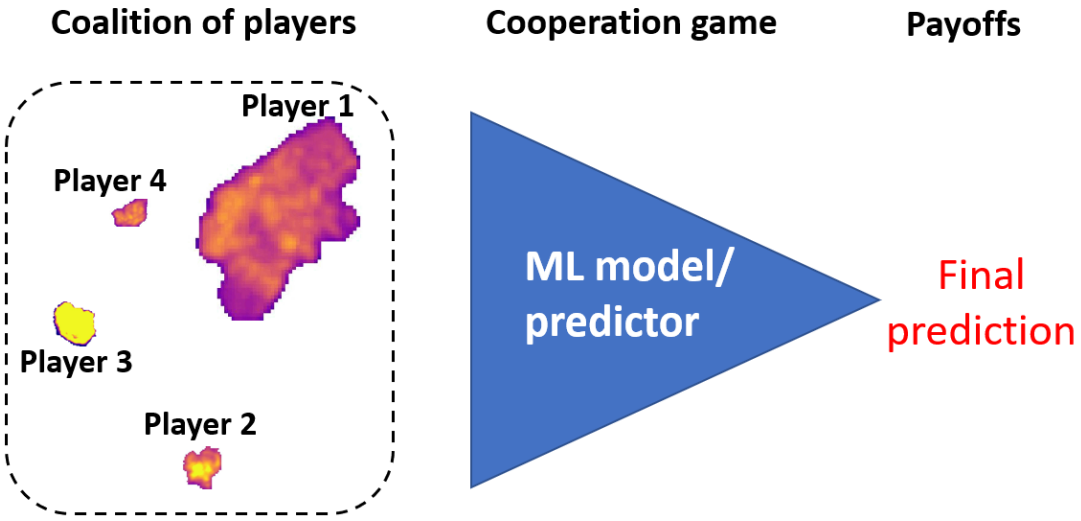
An interpretation tool at the lesion level

For a given aggregated model and for one patient, how to measure the impact of each of his lesions on the final prediction ?



How to generalize such and approach to any kind of aggregated predictive model ?

Shapley values to measure contributions



- All metastases (i.e players) cooperate to obtain the final prediction.
- We can borrow Shapley values from cooperative game theory for measuring the impact of each lesion to the final prediction.

$$\phi_i = \frac{1}{n} \sum_{S \subseteq N - \{i\}} \binom{n-1}{|S|}^{-1} \underbrace{(pred(S \cup \{i\}) - pred(S))}_{\text{marginal contribution of i}}$$

* It is the only payment rule satisfying Efficiency, Symmetry, Linearity and Null-player properties.

Independent contributions are not Shapley values !

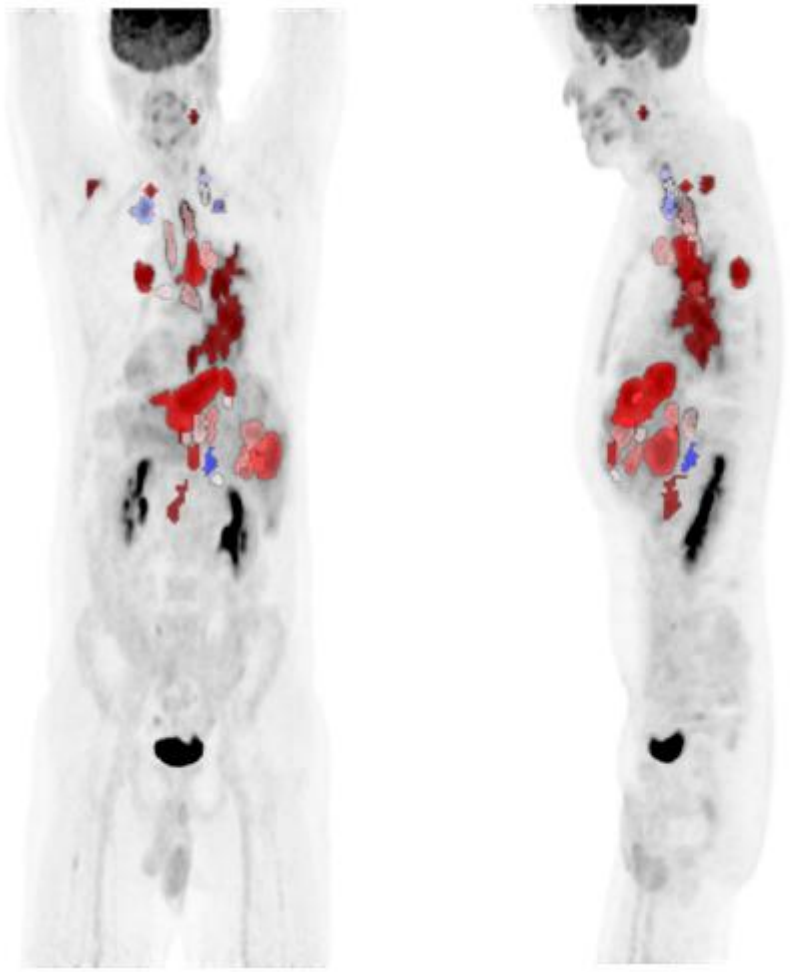
Let us consider a game g with two players x1, x2 such that the result obtained by each player separately is null and the result obtained from their cooperation is maximal:

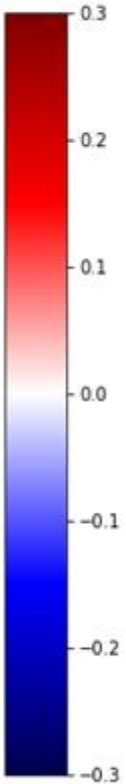
$$g(x_1) = g(x_2) = 0 \quad , \quad g(x_1, x_2) = 1 \quad , \quad g(\emptyset) = 0 \,(\text{no player involved})$$

$$\phi_1 = \frac{1}{2}(g(x_1) - g(\emptyset)) + \frac{1}{2}(g(x_1, x_2) - g(x_2)) = \frac{1}{2} \times 0 + \frac{1}{2} \times 1 = 0.5$$

$$\phi_2 = \ldots = 0.5$$

Application of Shapley values to OS late fusion model

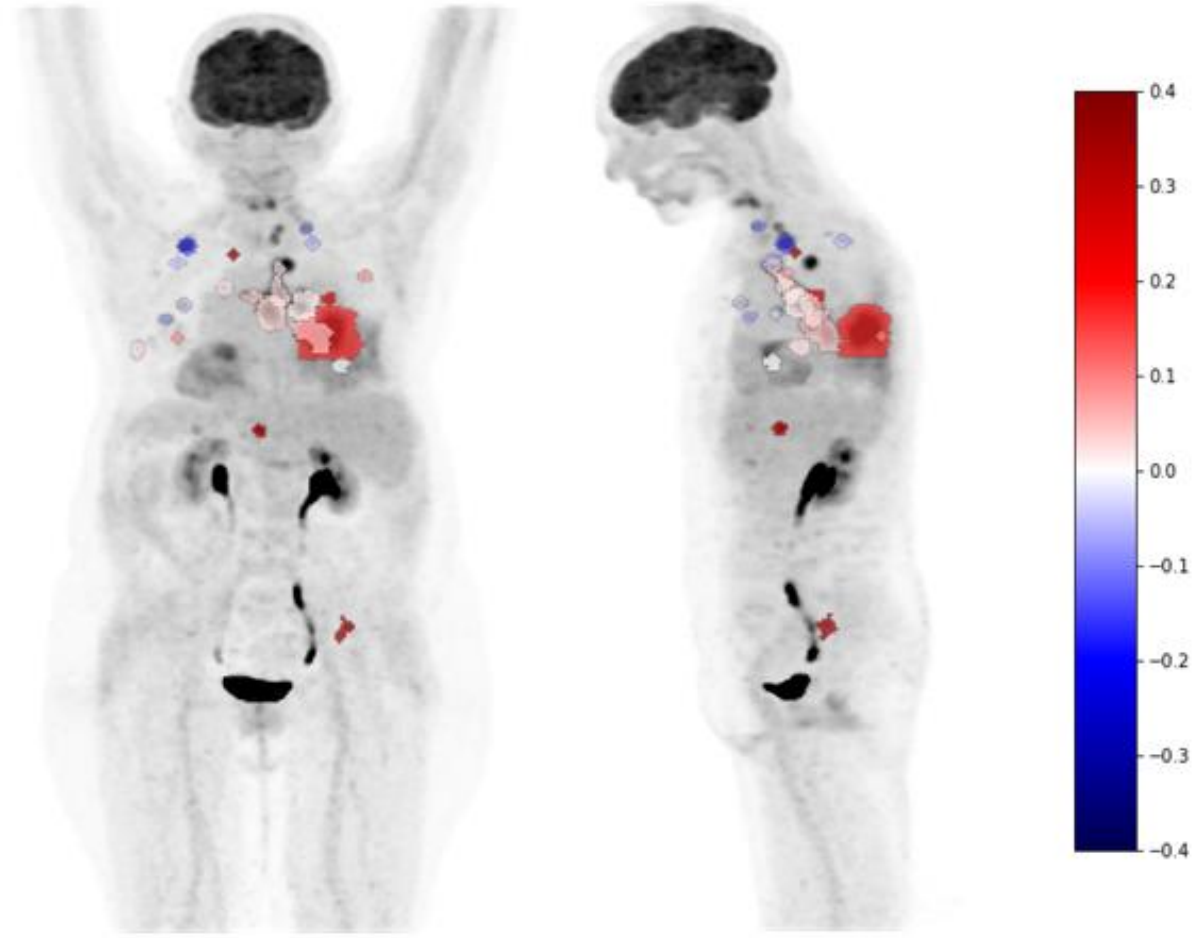




ADP_SOUSDIAPH_1	0.271069
ADP_CERV_HL_1	0.265905
META_OS	0.262040
PRIMITIF_PULM	0.257341
ADP_SOUSDIAPH_2	0.247039
ADP_SUSCLAV_CL	0.213111
META_FOIE_9	0.212359
META_PULM_CL	0.199501
META_FOIE_6	0.156821
ADP_MED_HL_2	0.132304
META_FOIE_8	0.109982
META_FOIE_10	0.079329
META_FOIE_4	0.067177
META_FOIE_5	0.055741
ADP_MED_HL_3	0.051112

Patient with highest predicted risk score

Application of Shapley values to OS late fusion model



META_OS	0.359036
META_PULM_CL_8	0.356151
ADP_SOUSDIAPH	0.290743
META_PULM_HL_1	0.244817
PRIMITIF_PULM	0.163355
META_PULM_CL_7	0.096740
META_PULM_HL_6	0.061479
ADP_MED_HL_4	0.037290
ADP_MED_CL	0.035760
ADP_MED_HL_2	0.029251
META_PULM_CL_1	0.022470
META_PULM_CL_5	0.018182
ADP_MED_HL_1	0.016128
ADP_MED_HL_3	0.001359
META_PULM_HL_5	-0.001487

Patient with high predicted risk score

Thank you for your attention !