TIPIT
Towards an Integrative approach for Precision ImmunoTherapy

LITO lab meeting

Nicolas Captier

21/12/2023

Outlines

1. **Introduction:** build new biomarkers for immunotherapy outcome in Non-Small Cell Lung Cancer

2. **Materials and Methods:** A retrospective multimodal cohort to develop multimodal predictors

3. **Results:** New proofs of the benefits of multimodal machine learning to build accurate prognostic models

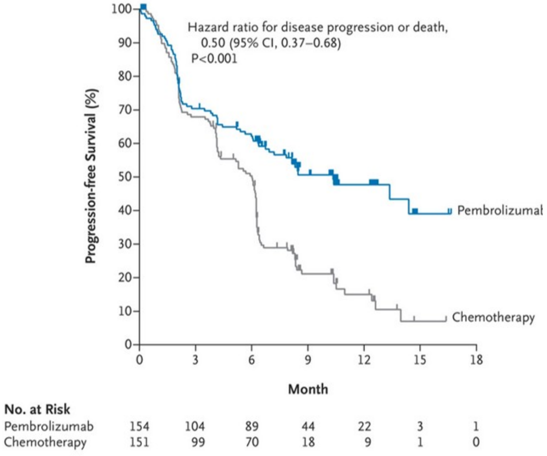
4. **Discussion:** What can we do next ?

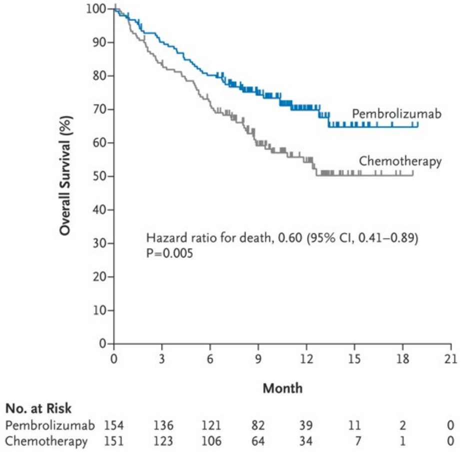
Multimodality to predict immunotherapy outcome in lung cancer

- Immunotherapy is the standard-of-care for metastatic Non-Small Cell Lung Cancer (NSCLC)

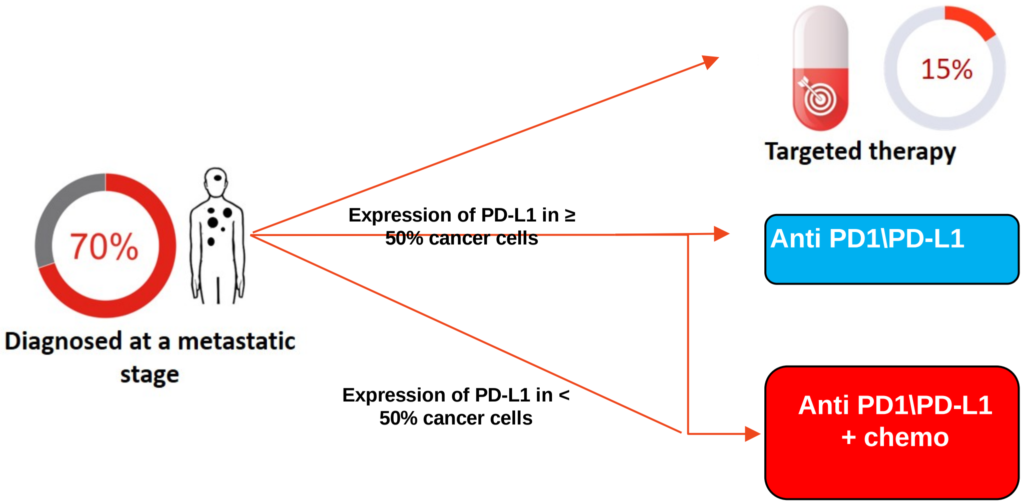
- Highly variable responses + only 40% of patients are alive at 2 years

- Established univariate biomarkers are very few with limited power.





Standard-of-care for advanced NSCLC in Europe



Multimodality to predict immunotherapy outcome in lung cancer

1. **Methodological question:** Can multimodal machine learning (& modelling) approaches build more accurate prognostic (potentially predictive) signatures ?

2. **Biological question:** Can we identify mechanisms associated with immunotherapy response with multimodal analyses ?

Outlines

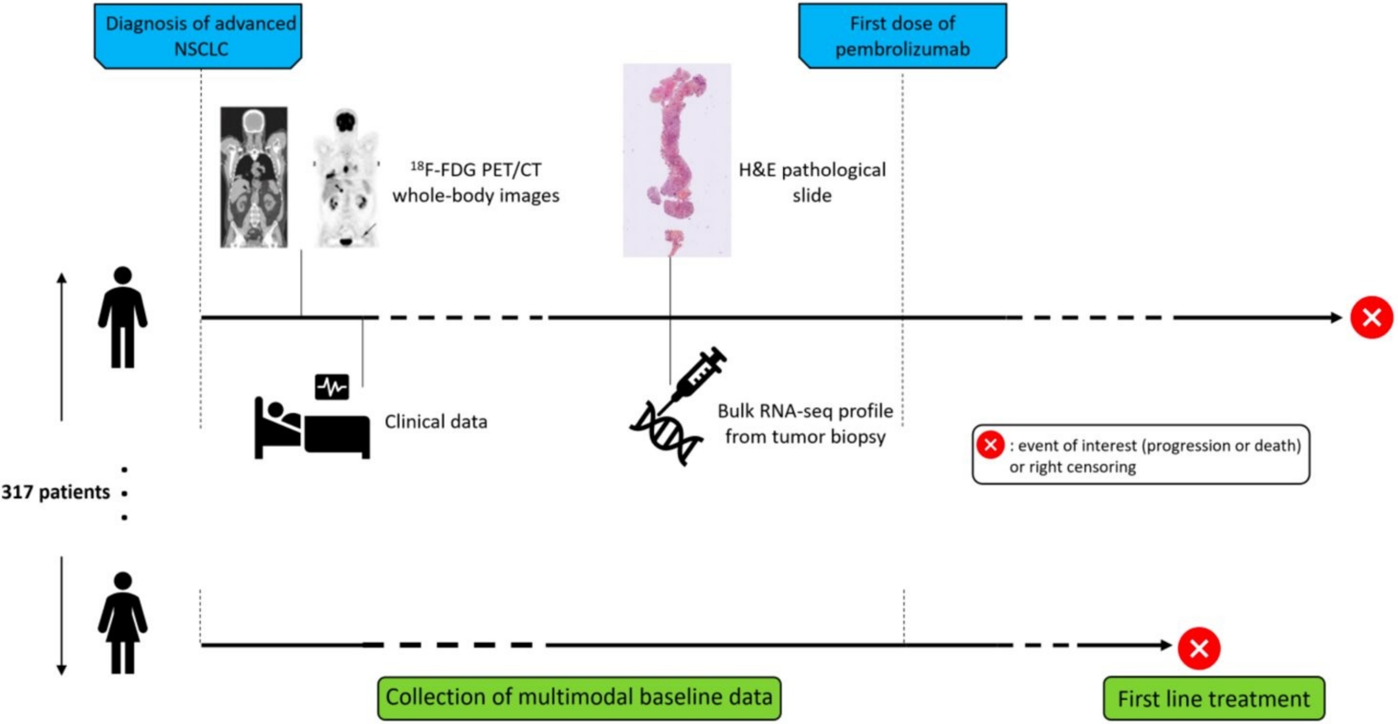
1. **Introduction:** build new biomarkers for immunotherapy outcome in Non-Small Cell Lung Cancer

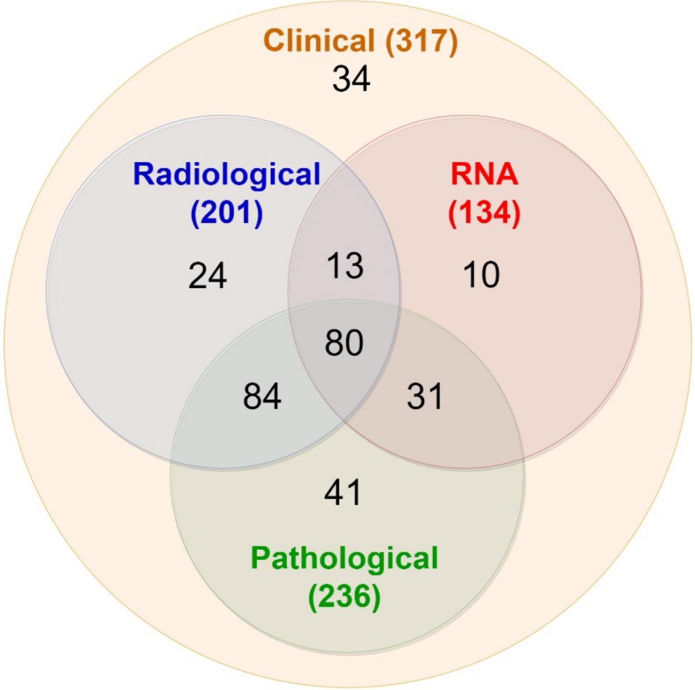
2. **Materials and Methods:** A retrospective multimodal cohort to develop multimodal predictors

3. **Results:** New proofs of the benefits of multimodal machine learning to build accurate prognostic models

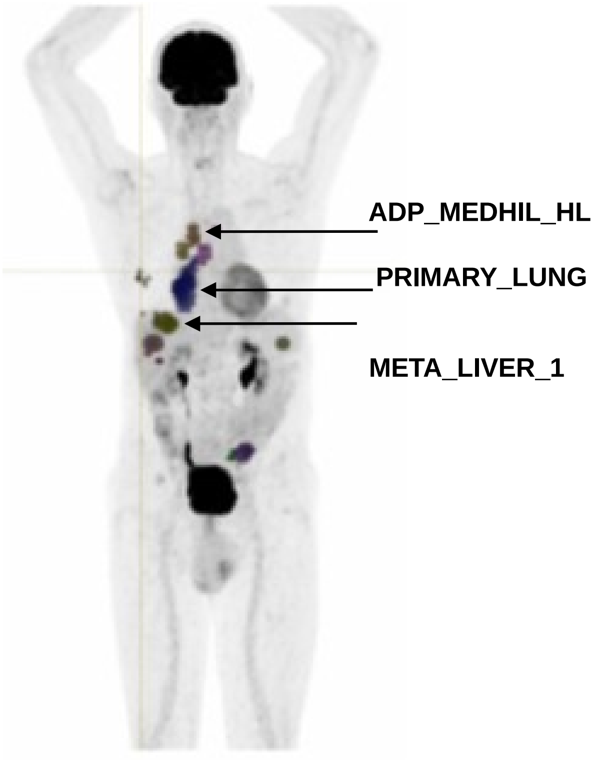
4. **Discussion:** What can we do next ?

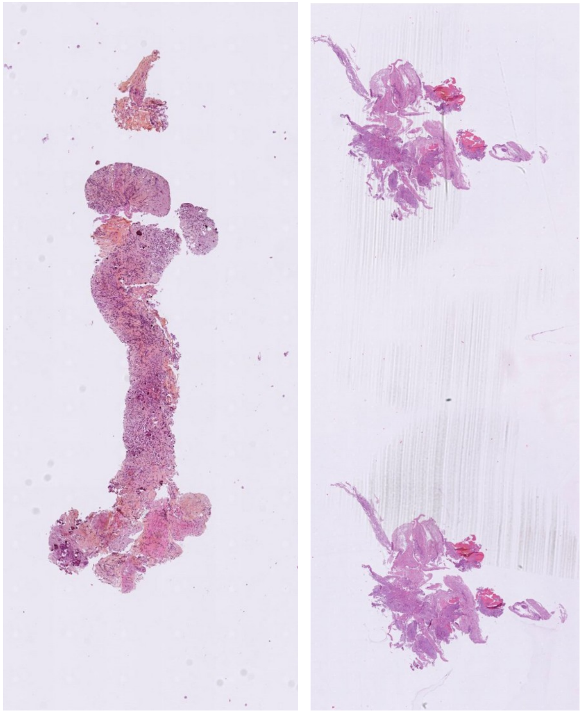
A retrospective multimodal cohort - I

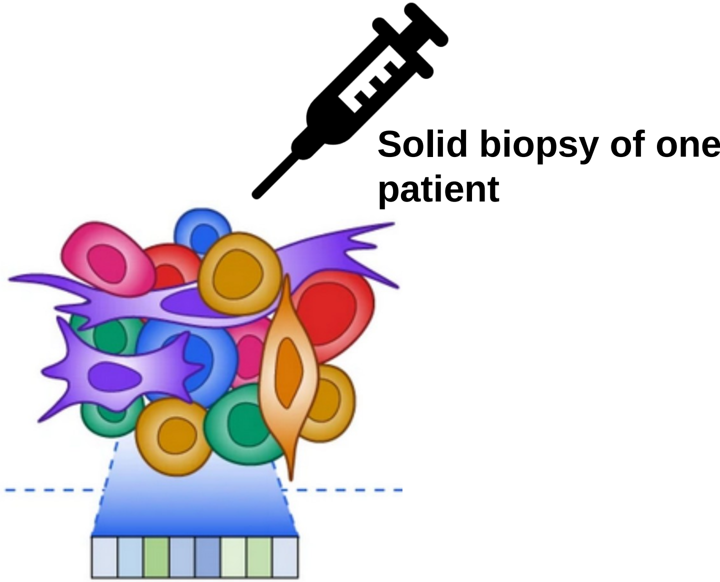




A retrospective multimodal cohort - II







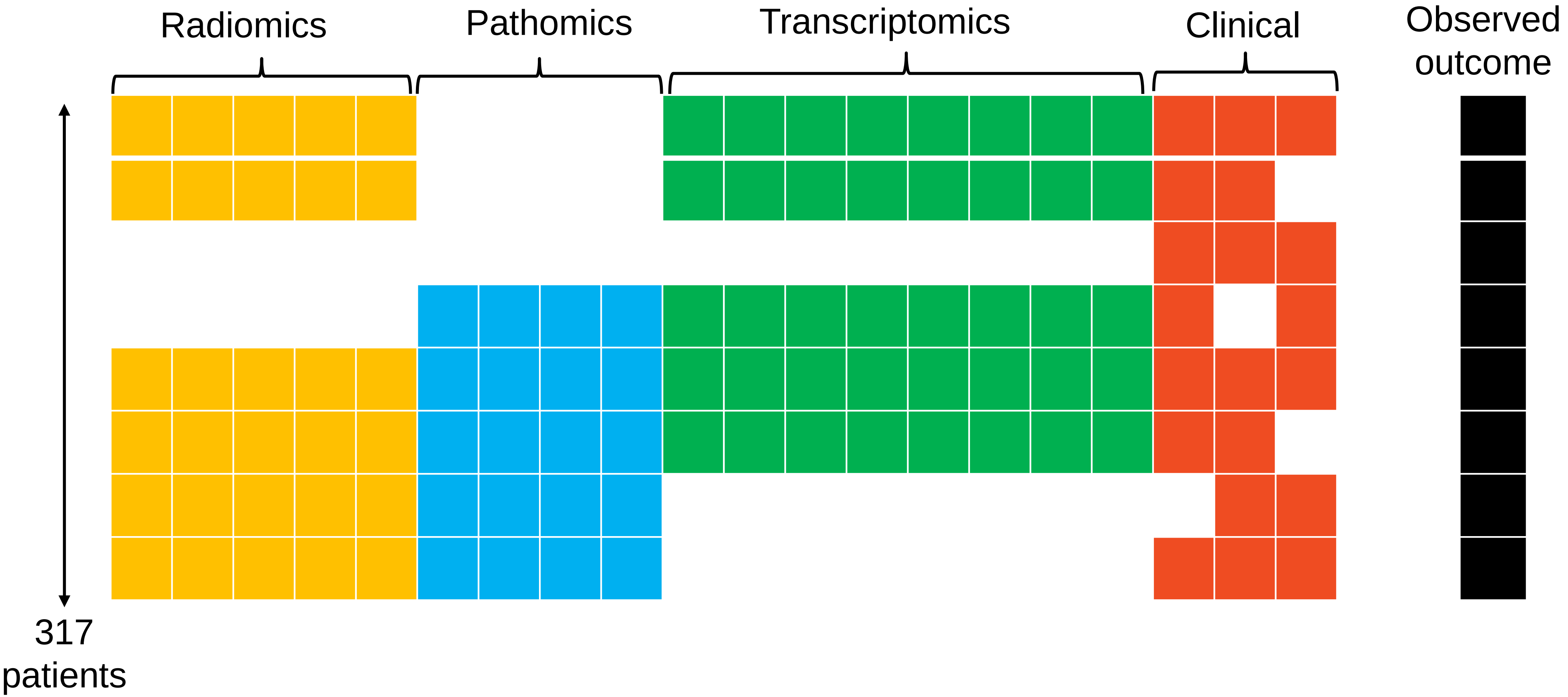
~20000 protein-coding genes

201 annotated/segmented 18F-FDG PET/CT

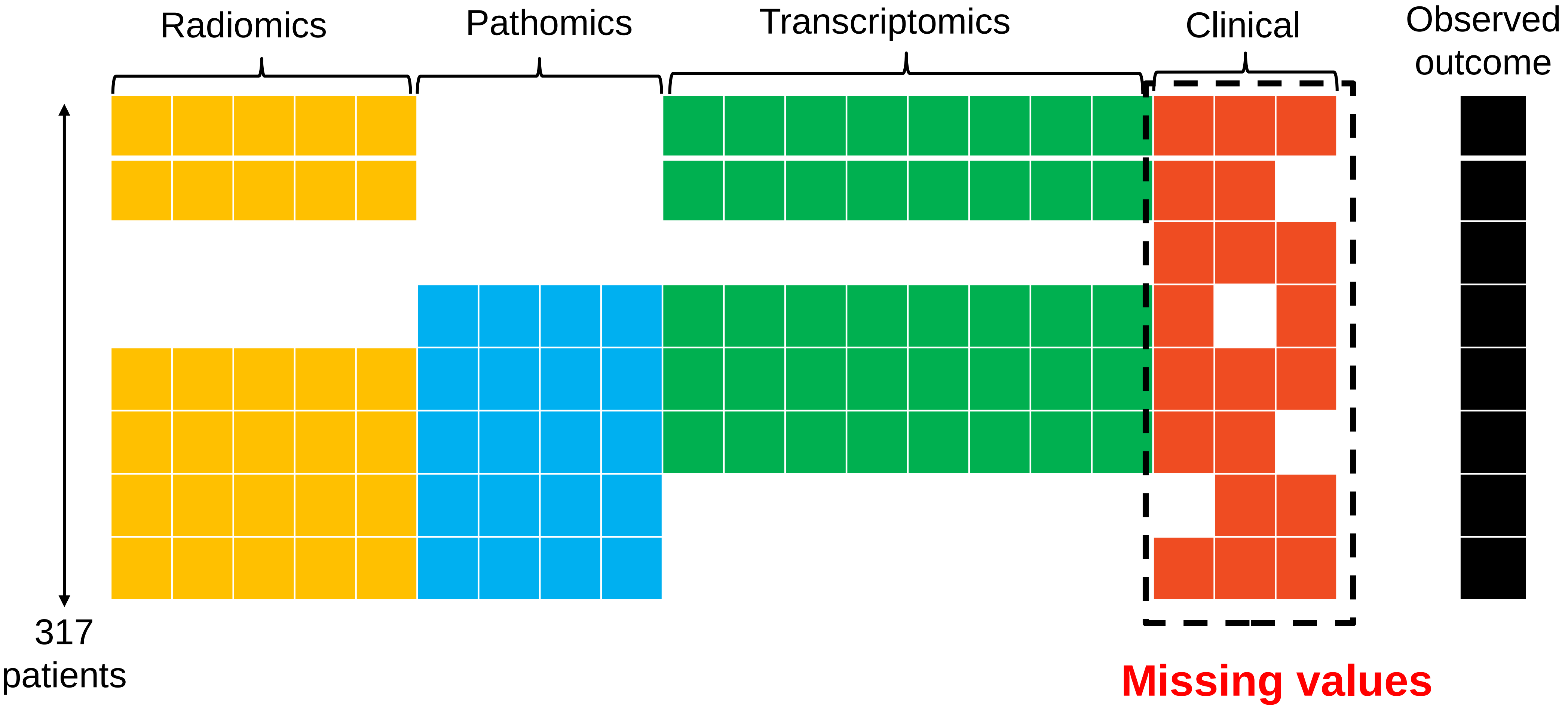
236 annotated/segmented digitized HE slides

134 bulk RNA-seq profiles from solid biopsy

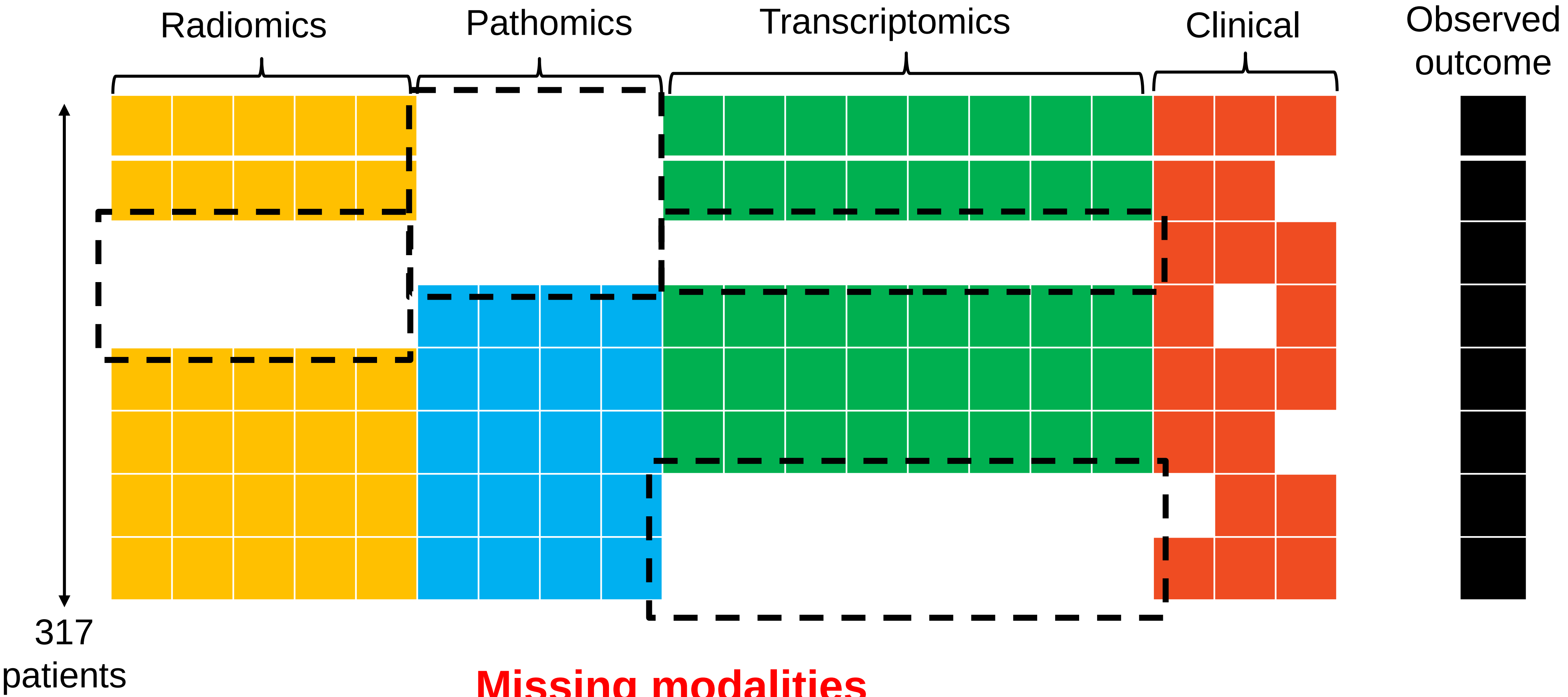
A retrospective multimodal cohort - III



A retrospective multimodal cohort - III



A retrospective multimodal cohort - III



Missing modalities

Investigate multiple learning tasks to extract consensus trends

Multiple outcomes

Survival outcomes:
Overall Survival (OS)

Progression-Free Survival (PFS)

Binary outcomes:
Death at 1 year

Progression at 6 months

Investigate multiple learning tasks to extract consensus trends

Multiple outcomes

Survival outcomes:
Overall Survival (OS)

Progression-Free Survival (PFS)

Binary outcomes:
Death at 1 year

Progression at 6 months

Multiple algorithms

Linear algorithms:
Logistic regression with elastic net penalty

Cox model with elastic net penalty

Tree ensemble algorithms:
XGBoost

Random Survival Forest

Investigate multiple learning tasks to extract consensus trends

Multiple outcomes

Survival outcomes:
Overall Survival (OS)

Progression-Free Survival (PFS)

Binary outcomes:
Death at 1 year

Progression at 6 months

Multiple algorithms

Linear algorithms:
Logistic regression with elastic net penalty

Cox model with elastic net penalty

Tree ensemble algorithms:
XGBoost

Random Survival Forest

Multiple fusion strategies

Late fusion

Early fusion (without or with univariate feature selection)

Fusion with attention weights (DyAM)

Outlines

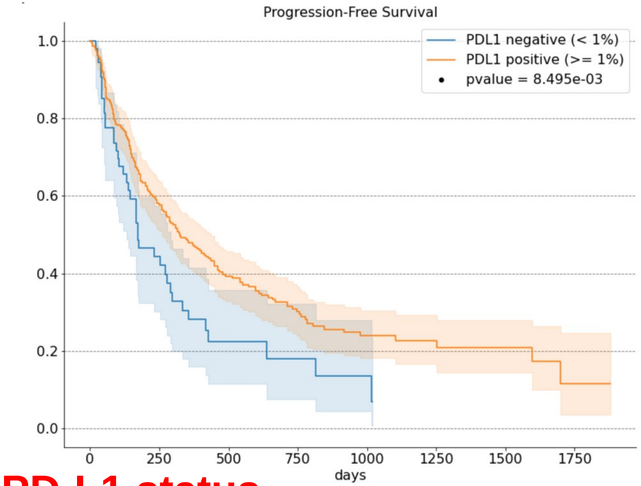
1. **Introduction:** build new biomarkers for immunotherapy outcome in Non-Small Cell Lung Cancer

2. **Materials and Methods:** A retrospective multimodal cohort to develop multimodal predictors

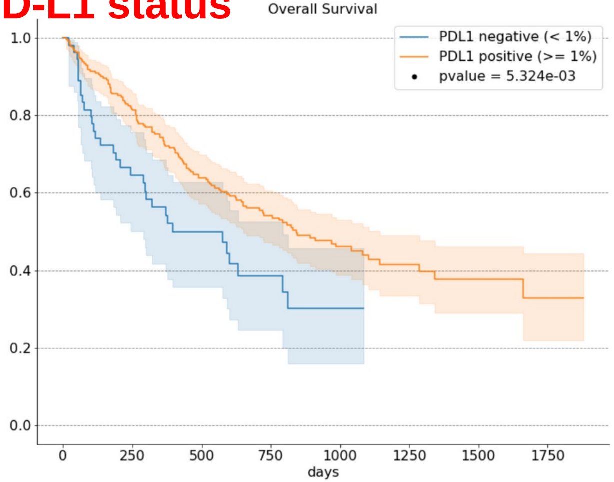
3. **Results:** New proofs of the benefits of multimodal machine learning to build accurate prognostic models

4. **Discussion:** What can we do next ?

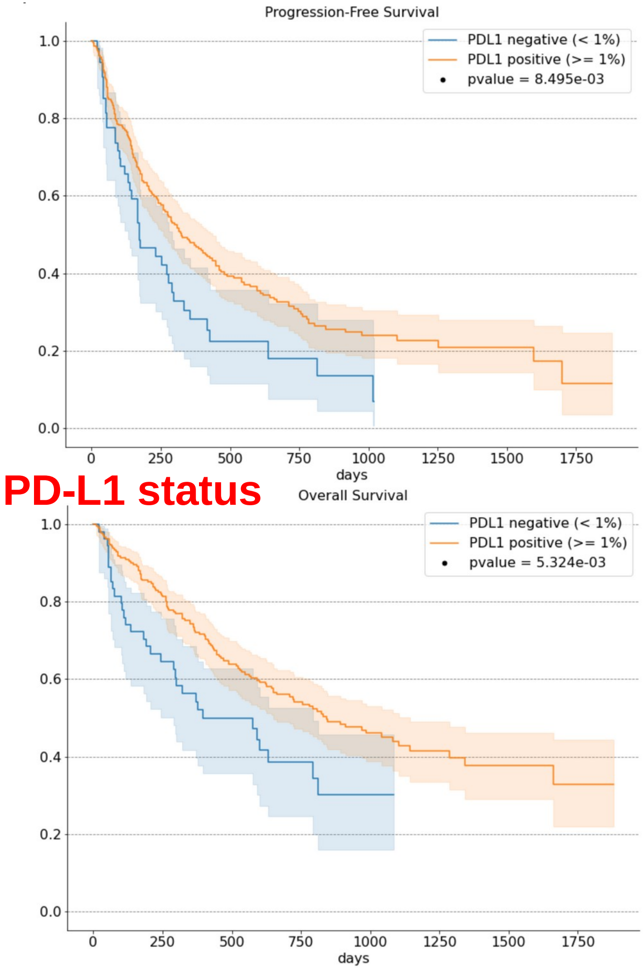
Univariate established biomarkers show limited performance



PD-L1 status

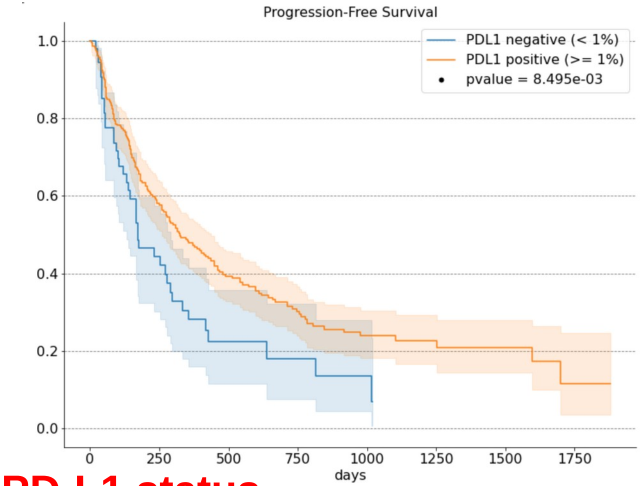


Univariate established biomarkers show limited performance

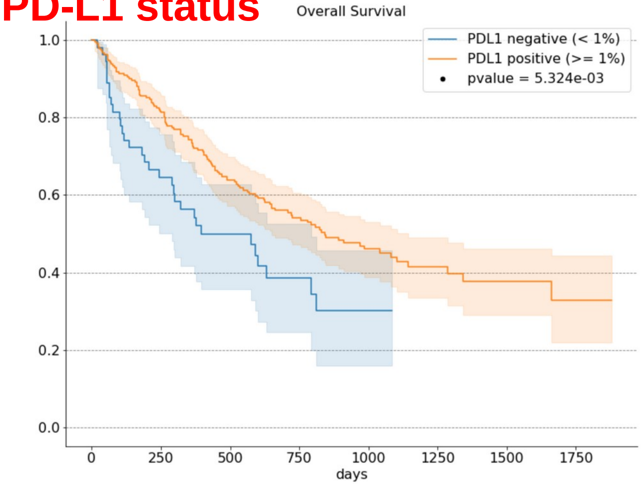


	C-index OS (whole cohort)
PD-L1 binary status	0.54 [0.51 – 0.57] pval=0.014
PD-L1 score (100 – TPS)	0.53 [0.48 – 0.58] pval=0.104

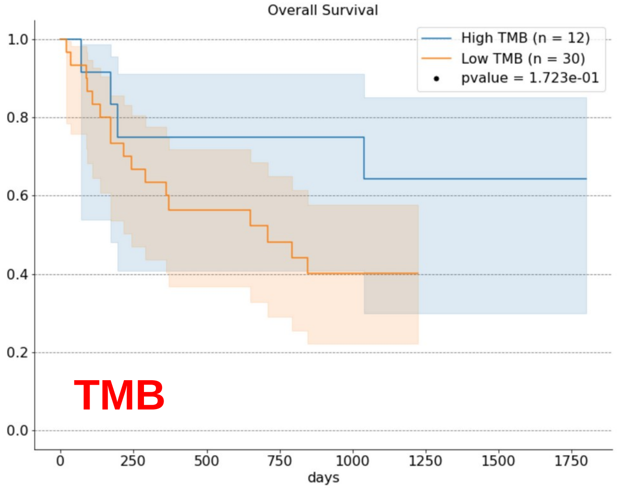
Univariate established biomarkers show limited performance

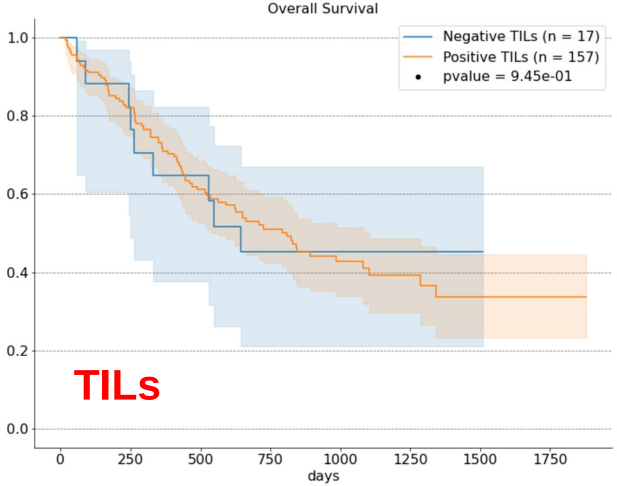


PD-L1 status



	C-index OS (whole cohort)
PD-L1 binary status	0.54 [0.51 – 0.57] pval=0.014
PD-L1 score (100 – TPS)	0.53 [0.48 – 0.58] pval=0.104





SysBio Curie

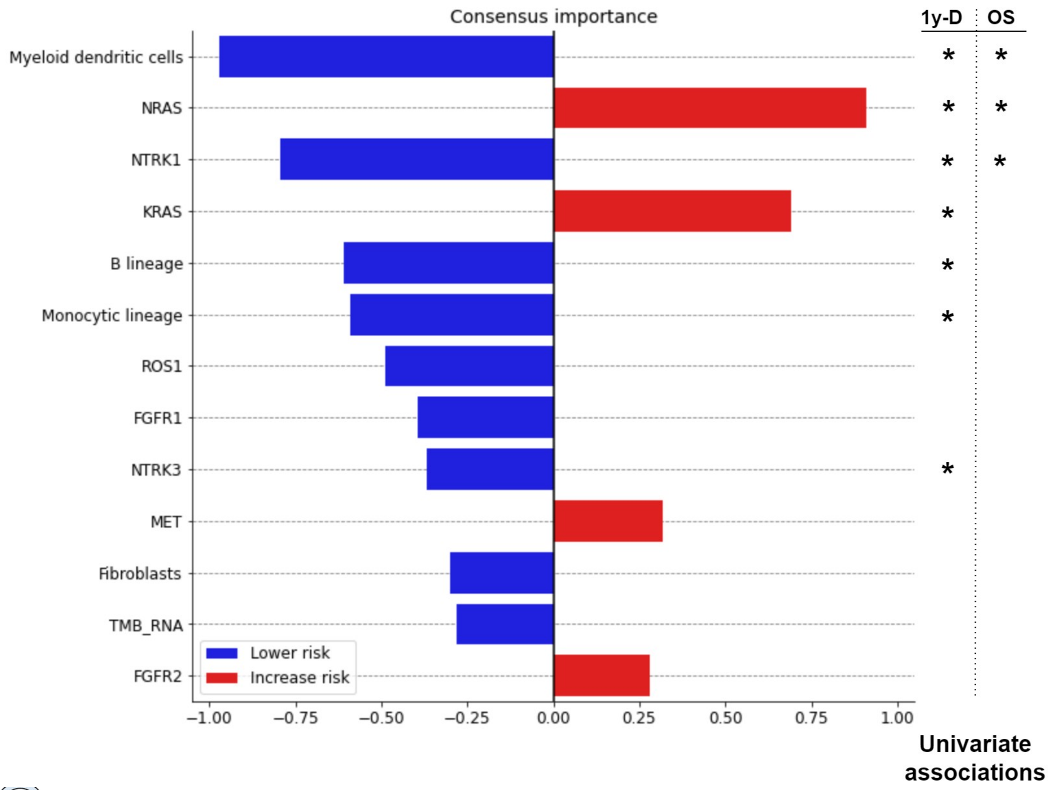
A benchmark of unimodal predictors

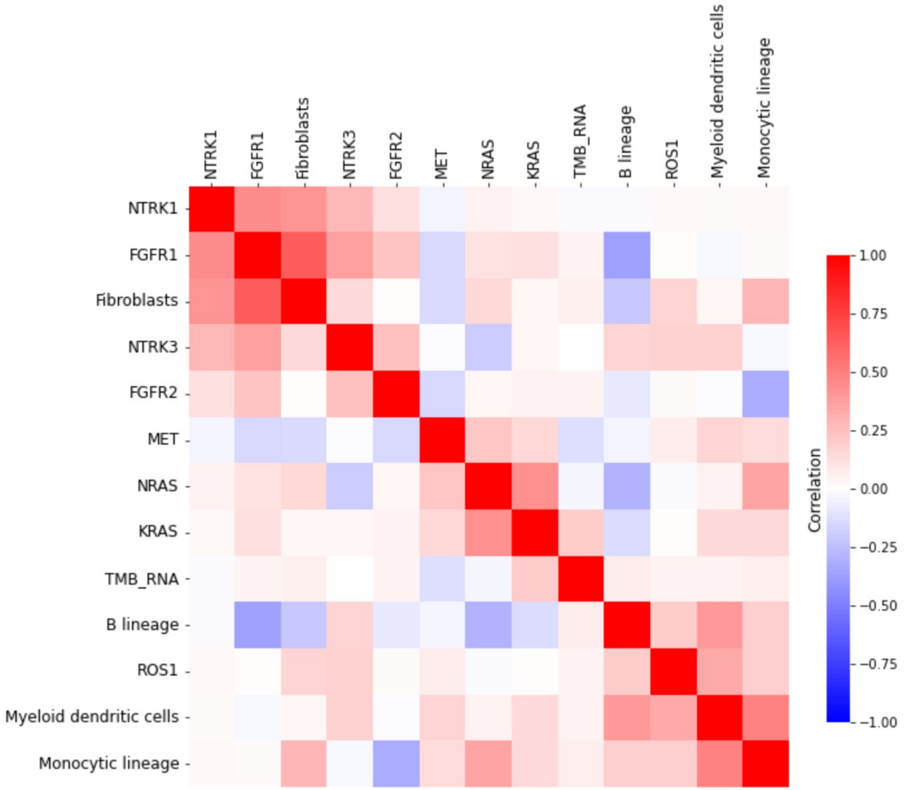
Target (number of patients)		OS (n=79)	1-year death (n=77)	PFS (n=80)	6-months progression (n=75)
Metric		**C-index**	**AUC**	**C-index**	**AUC**
Clinical	Tree ensembles	0.67±0.01 *	0.59±0.05	0.56±0.02	0.58±0.04
	Linear	0.60±0.02 *	0.73±0.02 *	0.53±0.03	**0.61±0.03 ***
Radiomics	Tree ensembles	0.61±0.02 *	0.62±0.04	0.57±0.01	0.56±0.05
	Linear	0.61±0.02 *	0.47±0.03	0.55±0.02	0.48±0.04
Pathomics	Tree ensembles	0.59±0.02	0.54±0.05	0.56±0.02	0.58±0.06 *
	Linear	0.58±0.02	0.56±0.03	0.51±0.02	**0.61±0.03 ***
RNA	Tree ensembles	**0.69±0.02 ***	**0.75±0.04 ***	0.57±0.02	0.60±0.04 *
	Linear	0.58±0.02	0.65±0.03	**0.59±0.02 ***	**0.61±0.03**

*: permutation p-value ≤ 0.05

Robust feature importance – RNA modality

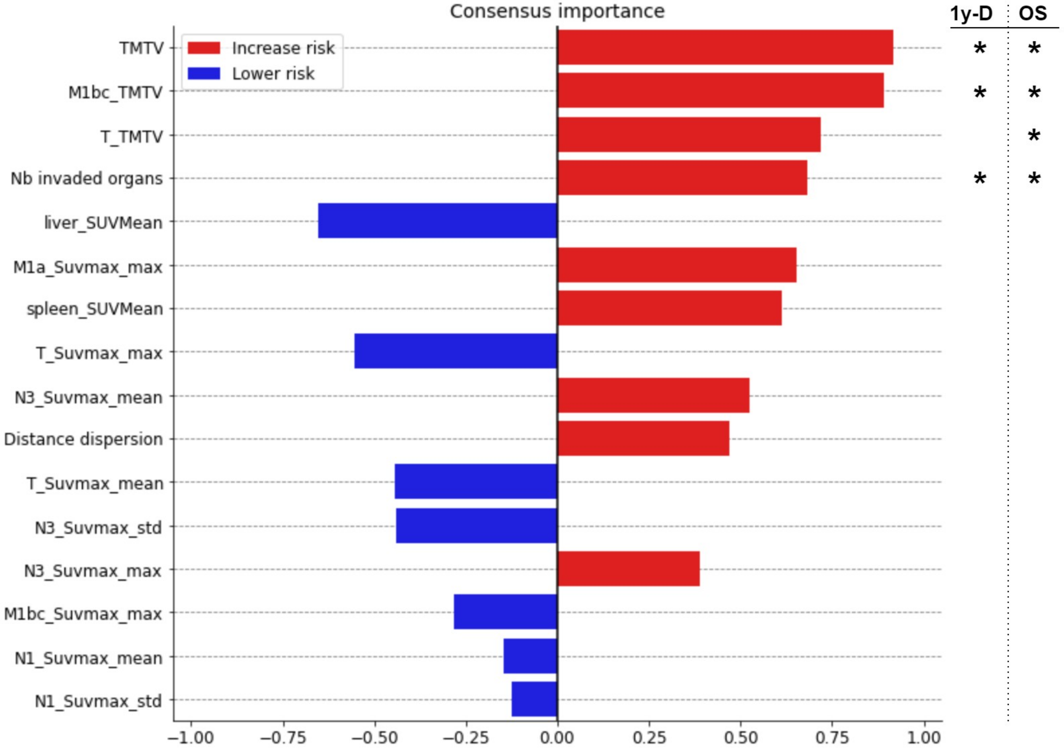
Feature importance ranking aggregated over both tasks (OS and 1 year death) and both algorithms (linear and tree ensemble)



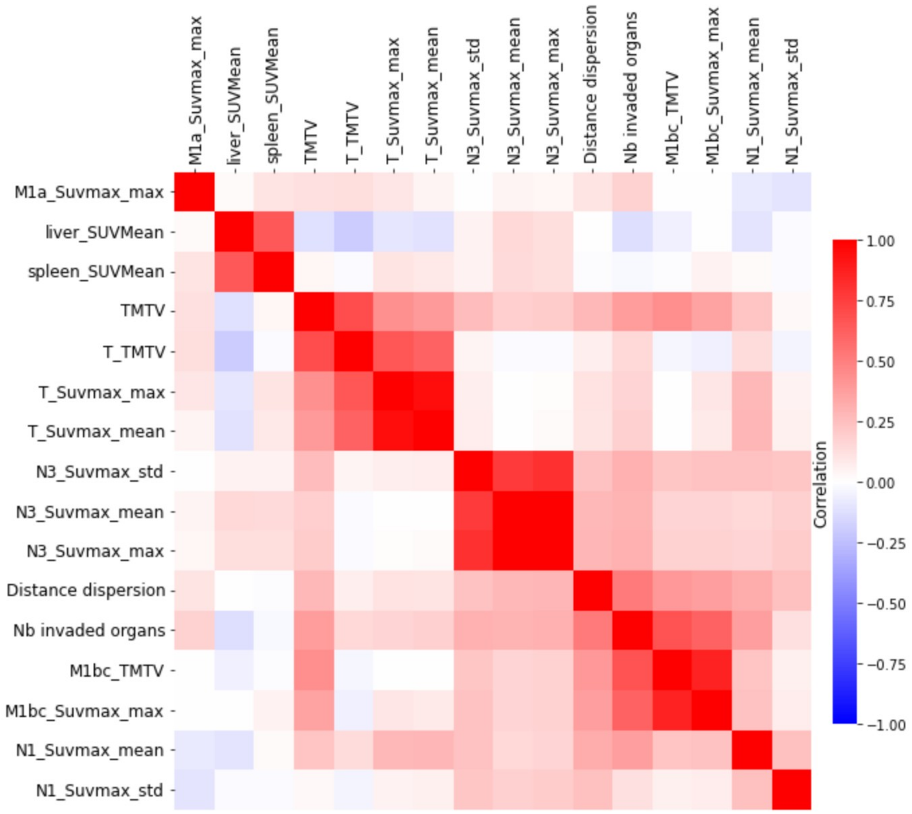


Robust feature importance – radiomic modality

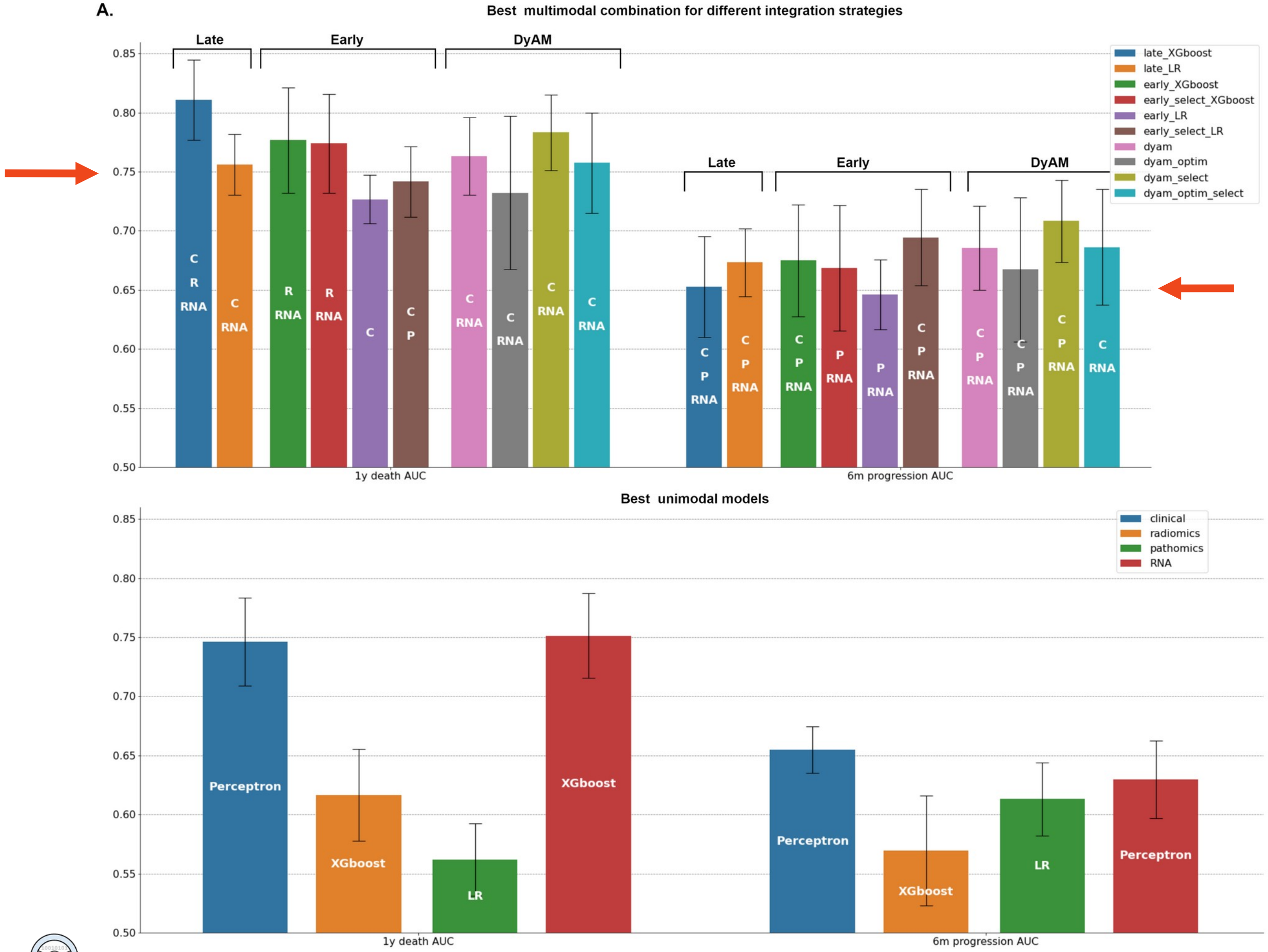
Feature importance ranking aggregated over both tasks (OS and 1 year death) and both algorithms (linear and tree ensemble)



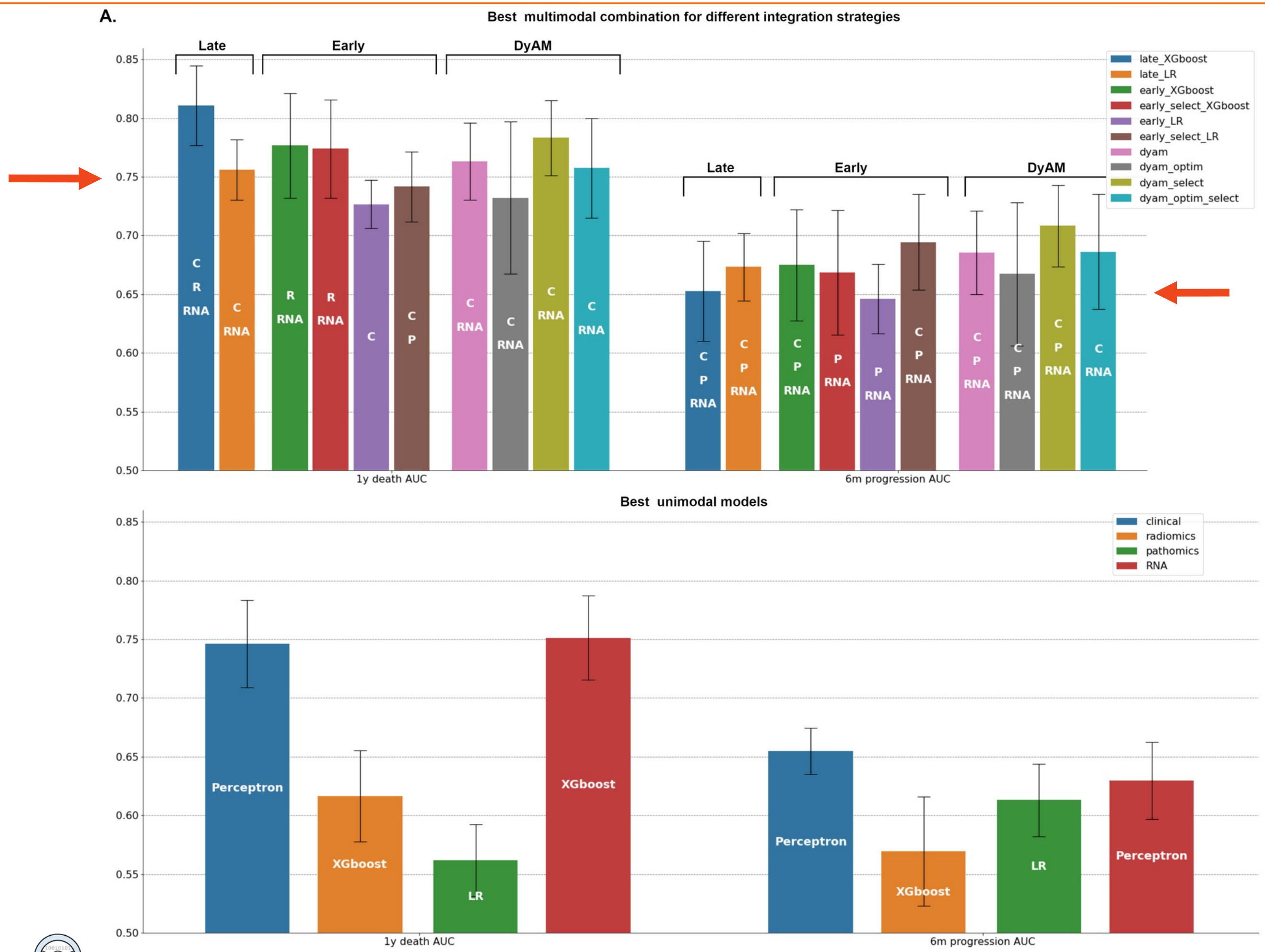
Univariate associations



Benchmark highlights benefits of multimodal approaches

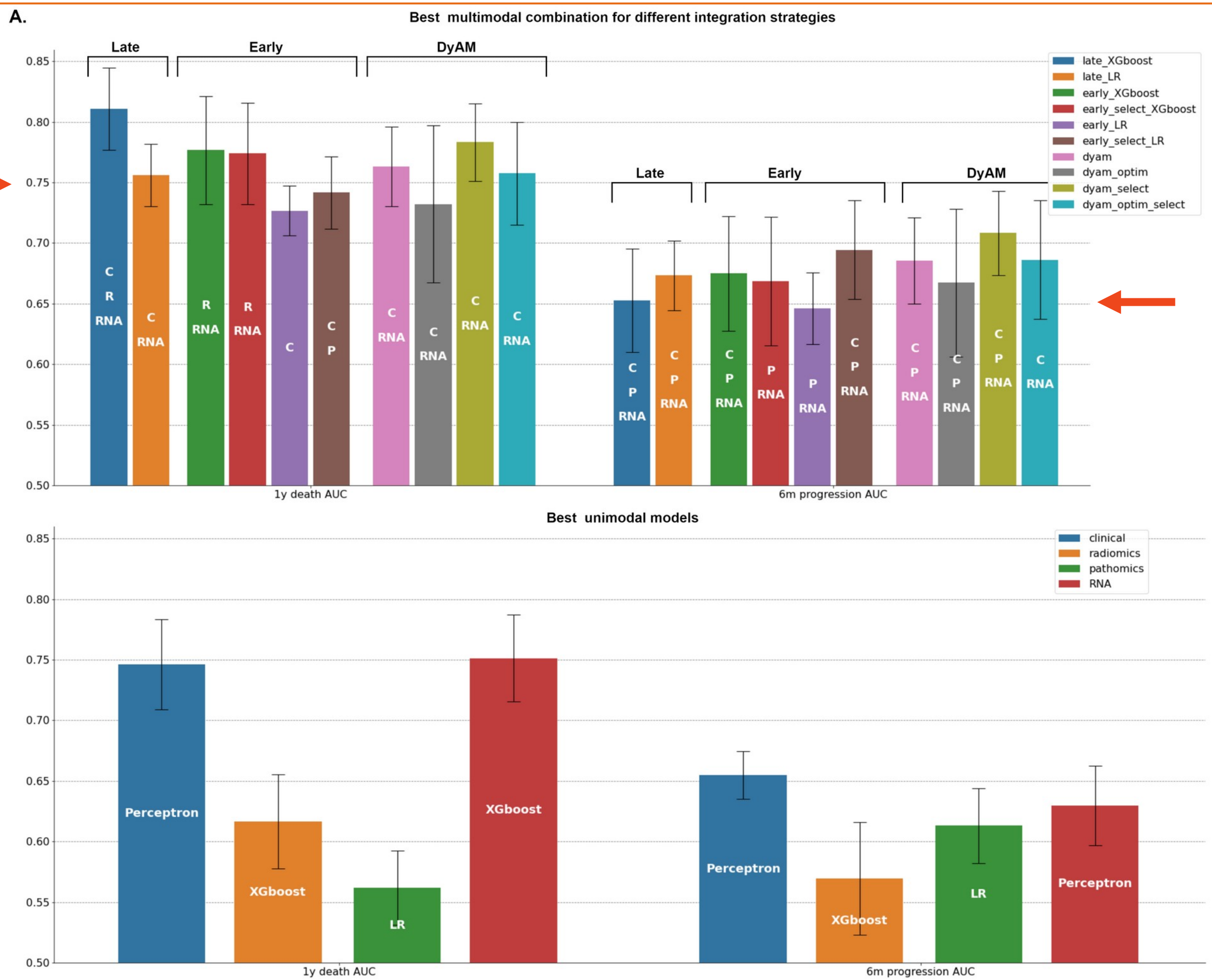


Benchmark highlights benefits of multimodal approaches



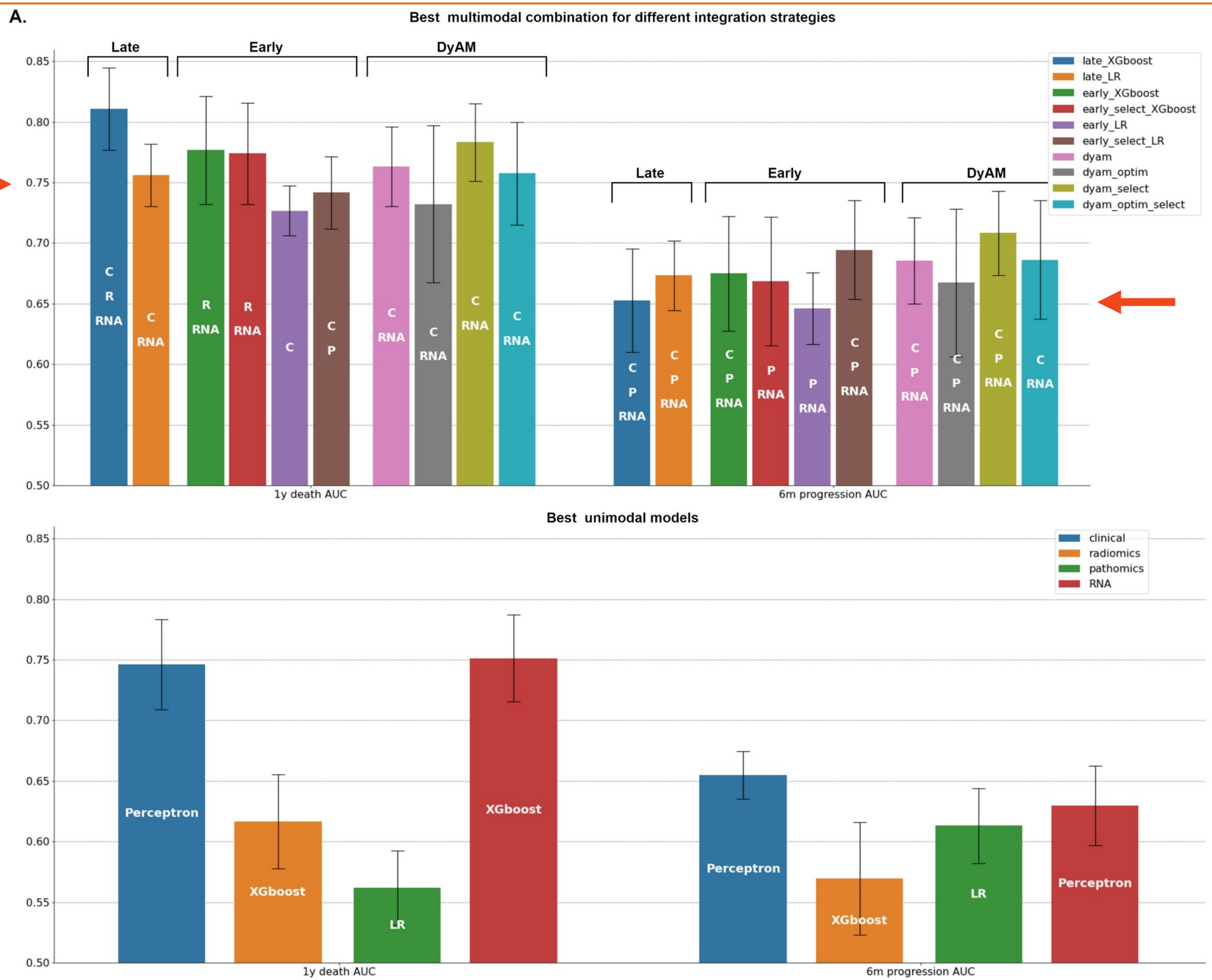
- The majority of multimodal models outperformed the best unimodal models

Benchmark highlights benefits of multimodal approaches



- The majority of multimodal models outperformed the best unimodal models
- Late fusion performed the best for 1 year death prediction

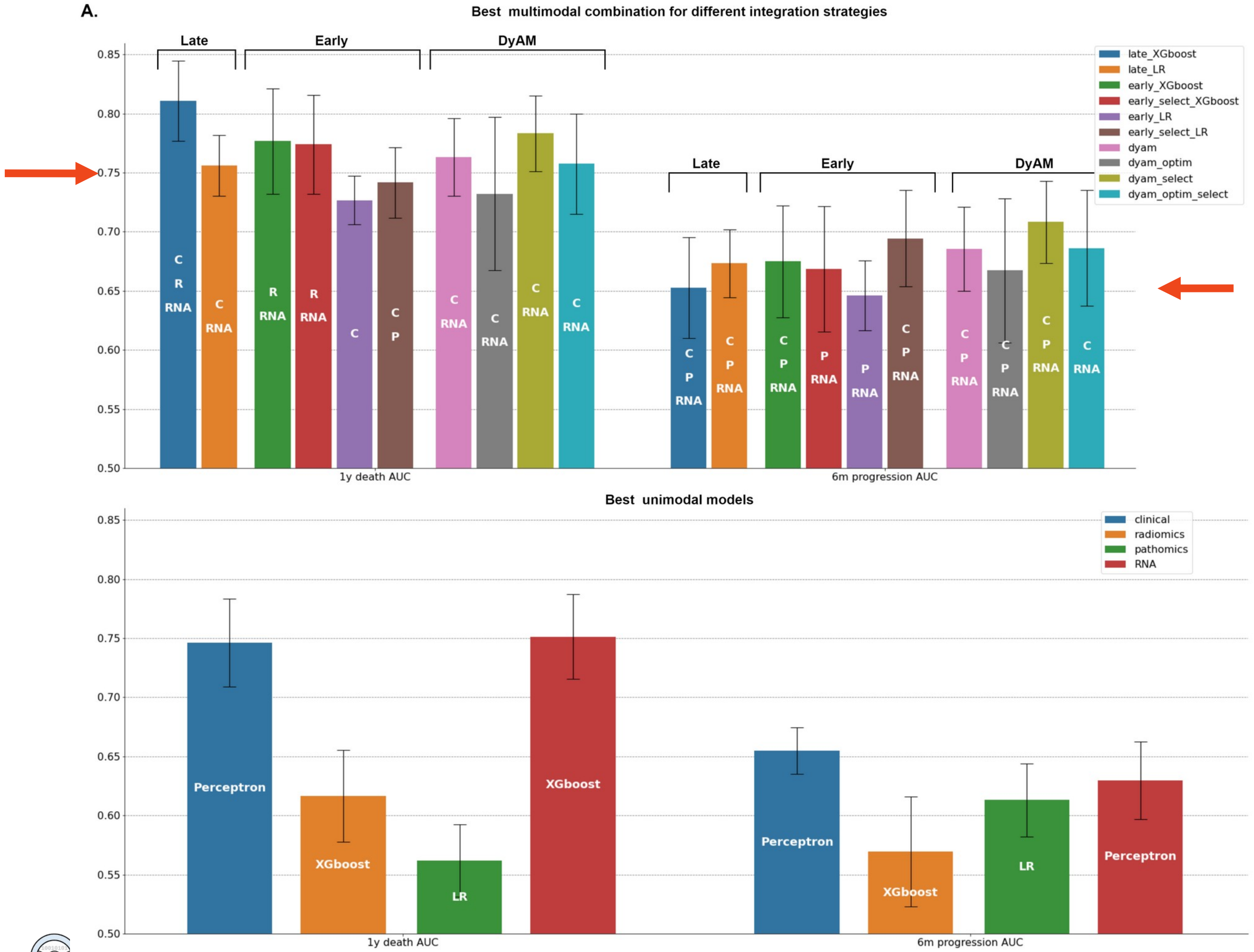
Benchmark highlights benefits of multimodal approaches



- The majority of multimodal models outperformed the best unimodal models
- Late fusion performed the best for 1 year death prediction
- Clinical, RNA and radiomic consistently involved in the best multimodal models for 1y death prediction

SysBio Curie

Benchmark highlights benefits of multimodal approaches



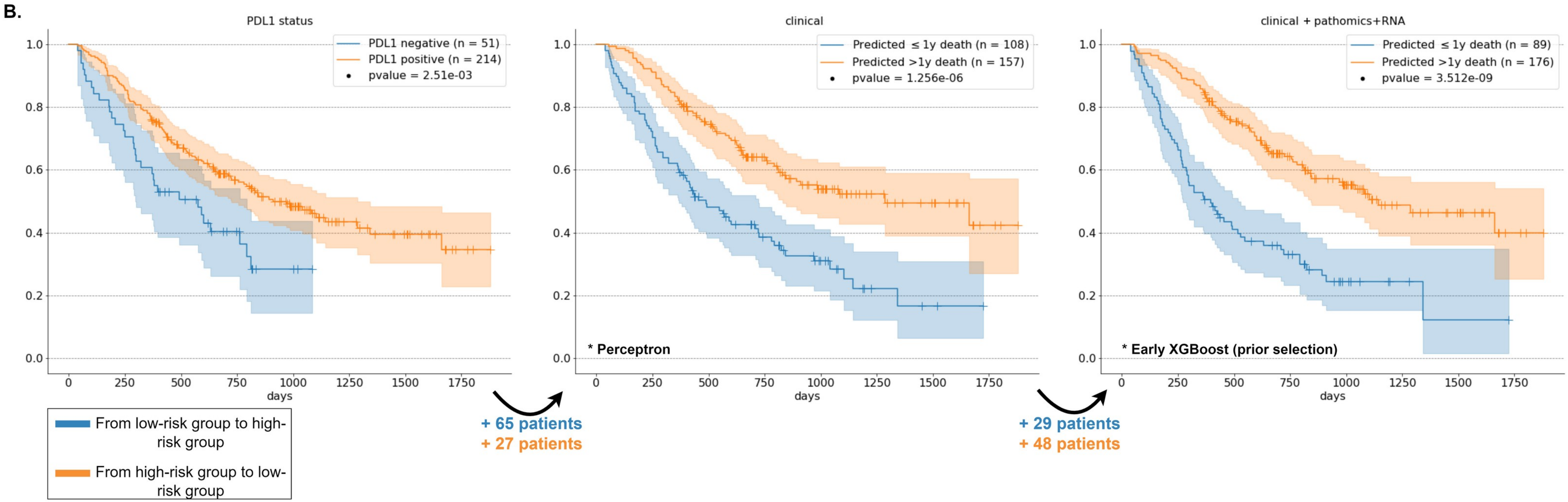
- The majority of multimodal models outperformed the best unimodal models
- Late fusion performed the best for 1 year death prediction
- Clinical, RNA and radiomic consistently involved in the best multimodal models for 1y death prediction

No integration strategy performed the best for all the tasks

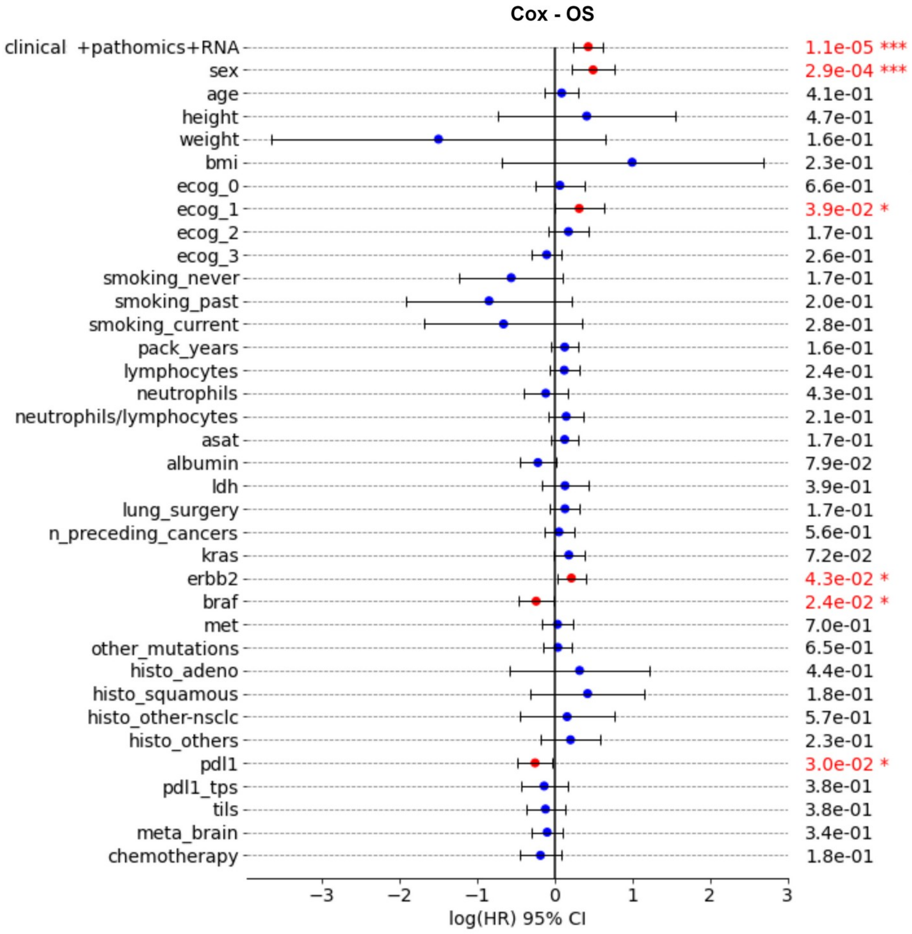
Multimodal score improves patient risk stratification

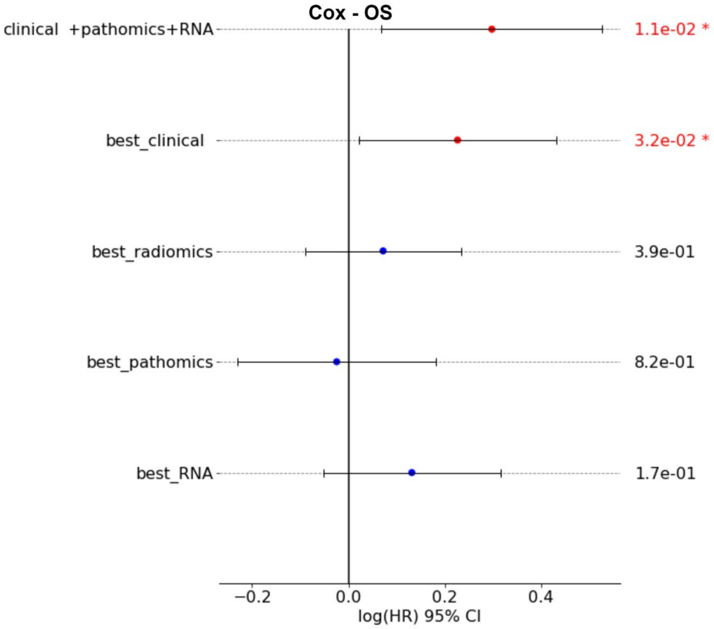
- A threshold is learnt on the training set of each fold of the cross-validation scheme and applied to the test set.

- Group membership is defined as the most frequently attributed group (low risk vs high risk) across the 100 repeats.



Multimodal score brings additional predictive information





Likelihood-ratio tests show a significant effect of the multimodal score wrt to routine clinical information

Conclusions

1. Late fusion is a relevant baseline strategy/starting point for further multimodal studies (simplicity, handle missing modalities easily…)

2. We provided new evidence of the relevance of multimodal approach for building powerful predictors should motivate others to collect new and larger multimodal cohorts.

3. Simple signatures of the Tumor MicroEnvironment seems to predict well NSLC outcome and lead to relevant hypotheses.

4. The results need to be further validated on larger an external data sets. More complex (e.g. end-to-end strategies) should be investigated.

Outlines

1. **Introduction:** build new biomarkers for immunotherapy outcome in Non-Small Cell Lung Cancer

2. **Materials and Methods:** A retrospective multimodal cohort to develop multimodal predictors

3. **Results:** New proofs of the benefits of multimodal machine learning to build accurate prognostic models

4. **Discussion:** What can we do next ?

Project 1: Study the link between radiomics and RNA expression

- 93+ patients with RNAseq and PET images (biopsy site is missing for some of them)

- Are different clusters based on radiomics/imaging characteristics associated with differentially expressed genes or biological pathways ?

- Can we identify/probe biological pathways deregulation with radiomic characteristics ?

- Can we use RNAseq data to understand better the radiomic phenotype of tumors ?

Project 2: Multiple Instance Learning for metastatic disease

- Use each metastasis as an instance with a hidden outcome to predict patient’s outcome

Project 3: Improve PET representation with supervised learning

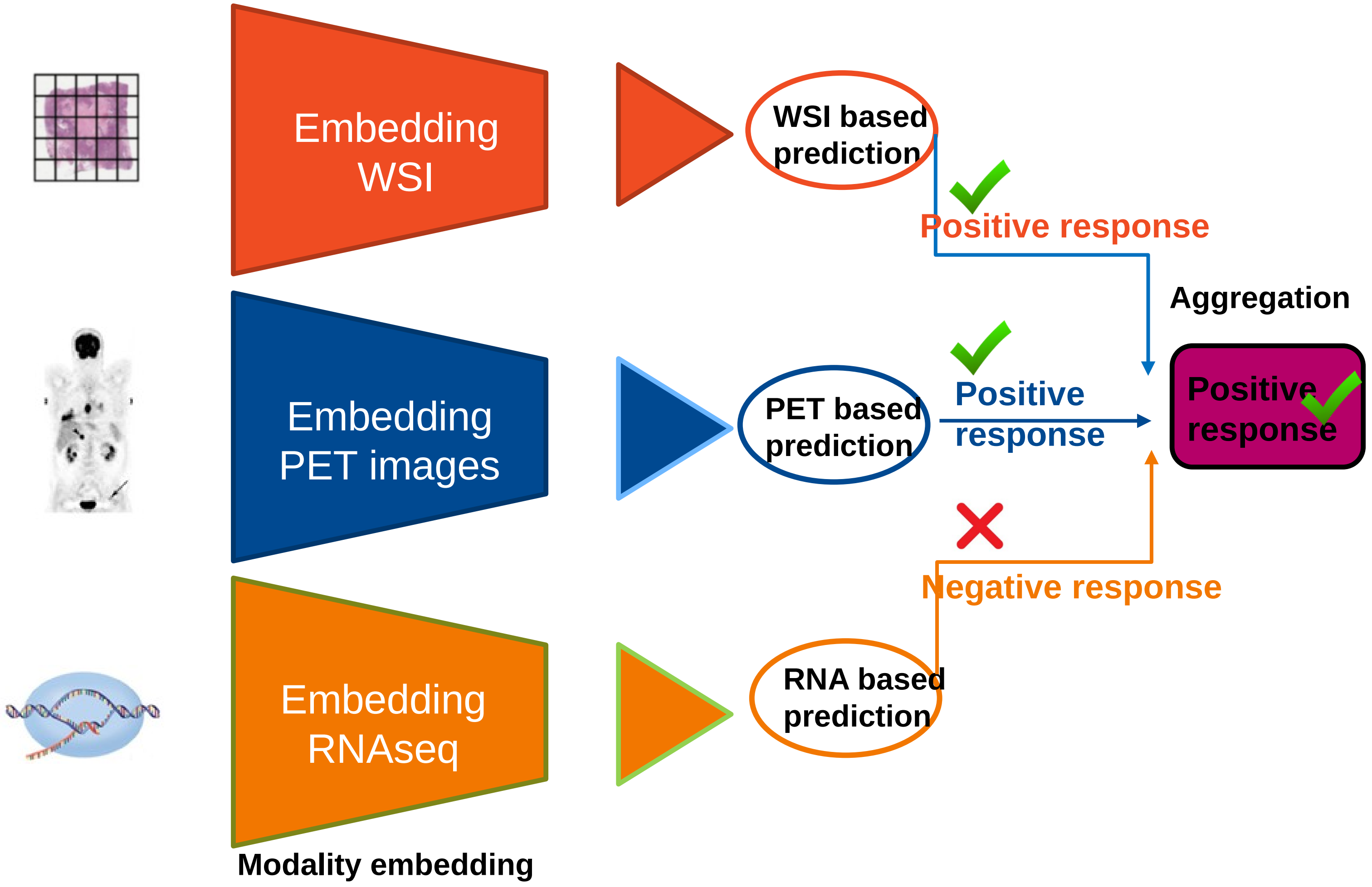
- Predict the total tumor volume in lymph nodes, distant regions... from unannotated MIP images

Acknowledgements

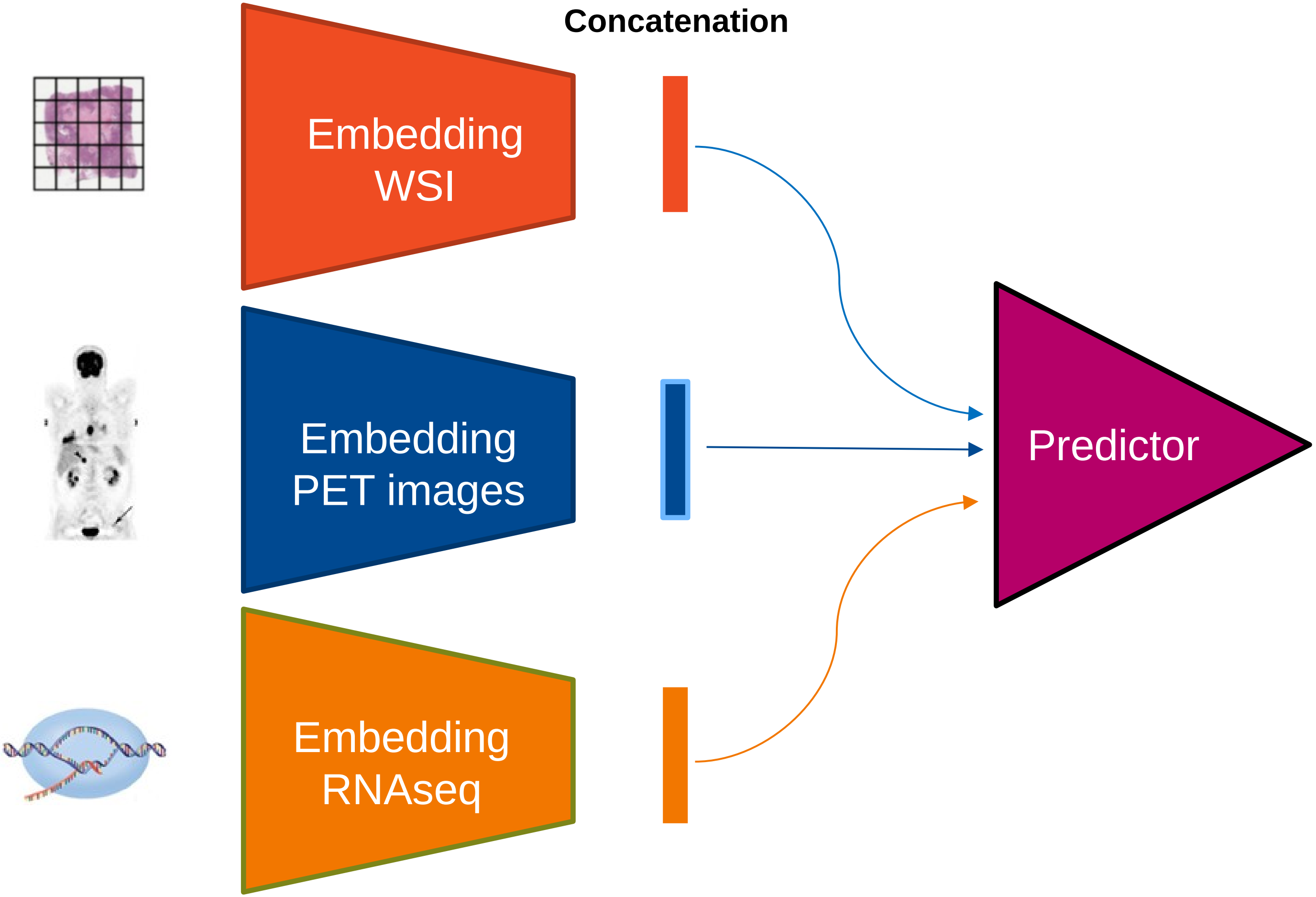
- **Curie-Montsouris Chest center:** Nicolas Girard, Sarah Lagha, Anne-Sophie Tedesco
- **Imaging department Curie:** Hervé Brisse, Marie Luporsi, Toulsie Ramtohul
- **Pathology department Curie:** Clément Beaulaton
- **Data department Curie:** Paulette Salamoun Feghali
- **LITO Curie:** Irène Buvat, Fanny Orlhac, Narinée Hovhannisyan, Nicolas Captier, Erwin Woff
- **SysBIO U900 Curie:** Emmanuel Barillot, Andrei Zinovyev, Christine Lonjou, Nicolas Captier
- **CBIO U900 Curie-Mines:** Thomas Walter, Marvin Lerousseau

* **Immunity and Cancer:** Hélène Salmon

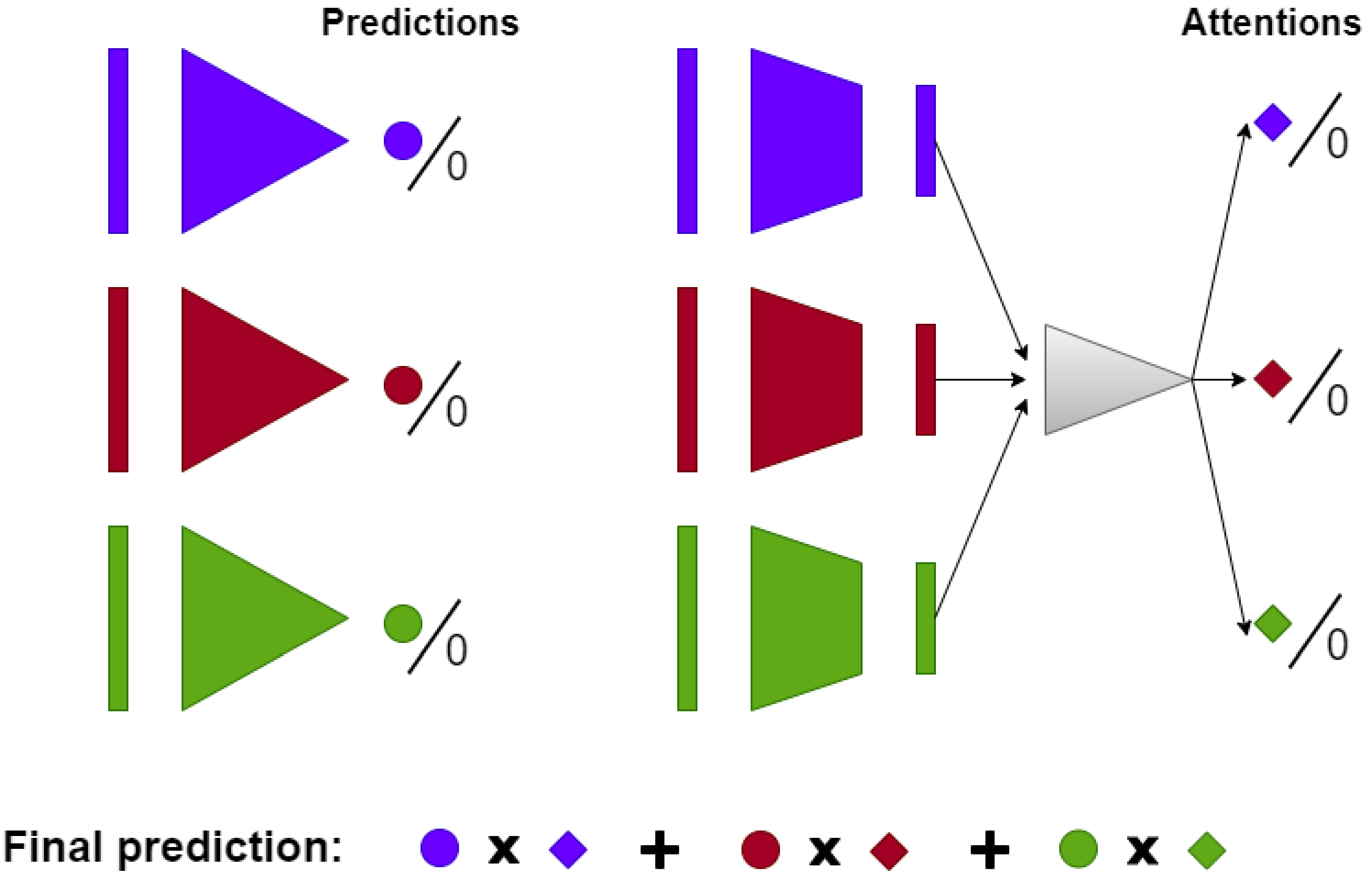
Late fusion strategy



Early fusion strategy



DyAM strategies



- Attention mechanism weighs each modality for each patient
- The model learns to combine different modalities through normalization and grey network