

Prédiction de la survie chez des patients atteints de lymphomes

Louis Rebaud

18/01/2024

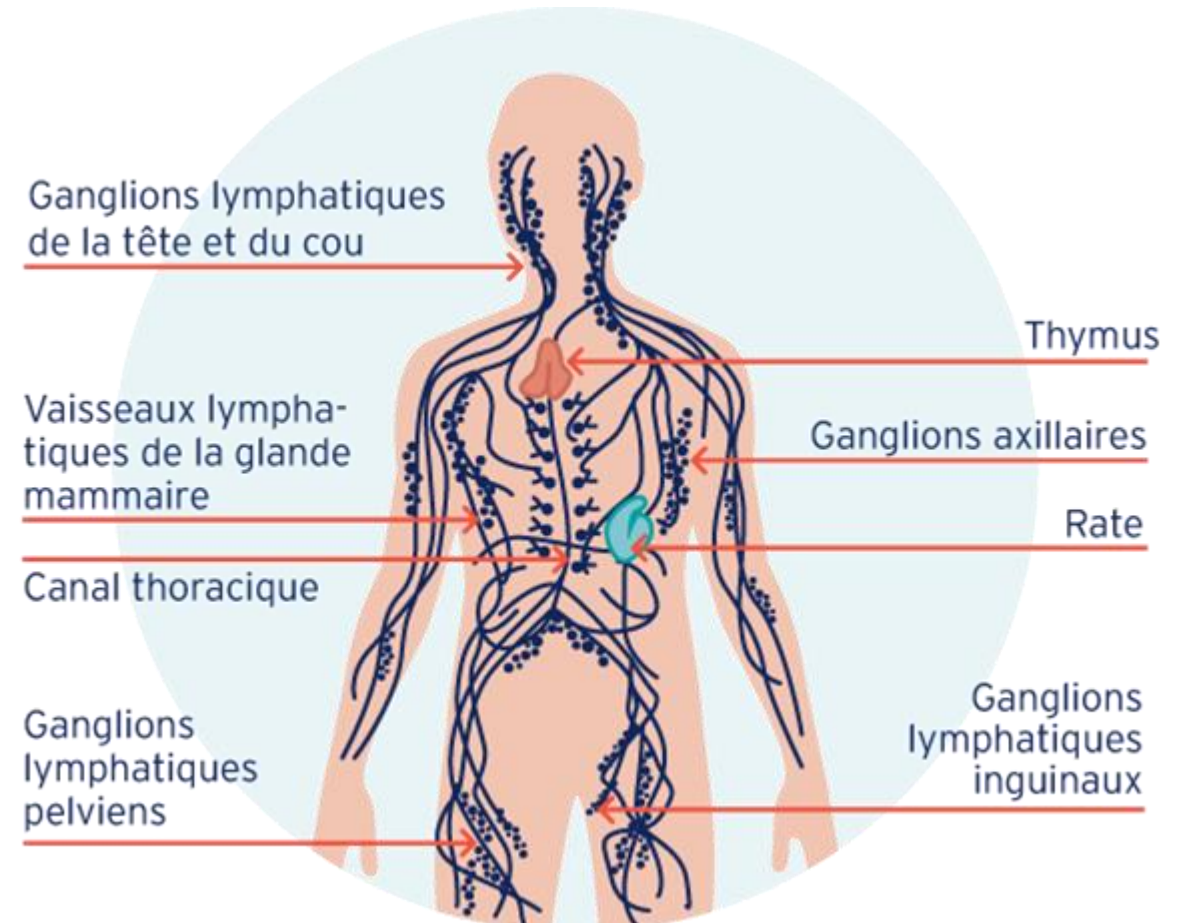


Sommaire

- Les lymphomes et le rôle de l'imagerie TEP
- Recherche semi-automatique de biomarqueurs
- Biomarqueurs identifiés
- Construction de signatures avec ICARE
- Extension de l'outil PARS de Siemens

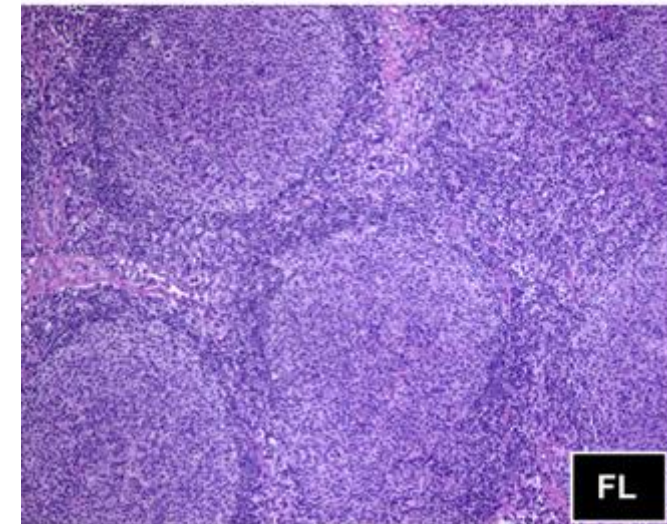
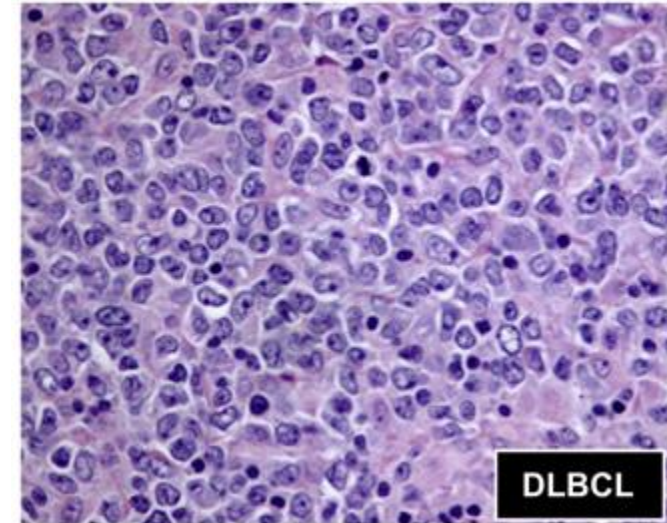
Les lymphomes

- Cancers du système lymphatique
- Traitement : combinaison de chimiothérapie, de radiothérapie et d'immunothérapie
- Multiples lésions dans différentes localisations
- Lymphomes **Non Hodgkiniens** :
 - 5e cancer en termes de fréquence (11 000 nouveaux cas/an en France)
 - Plus de 27 sous-types



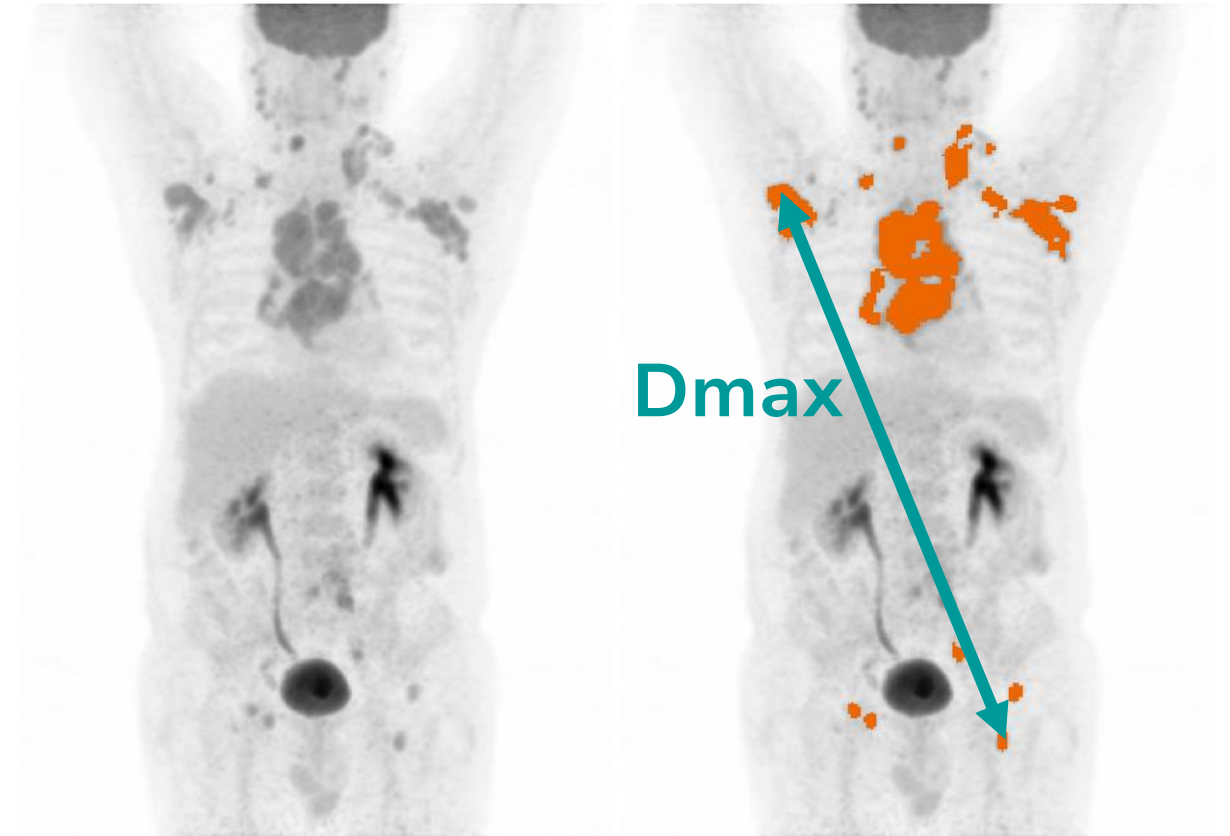
Les lymphomes

- Lymphome B diffus à grandes cellules (DLBCL)
 - Agressif
 - 65% de survie à 5 ans
 - Lymphomes Non Hodgkiniens le plus courant
- Lymphome folliculaire (FL)
 - Indolent
 - 90% de survie à 5 ans, 50% à 20 ans
 - Guérison complète rare
 - Peut muter en DLBCL

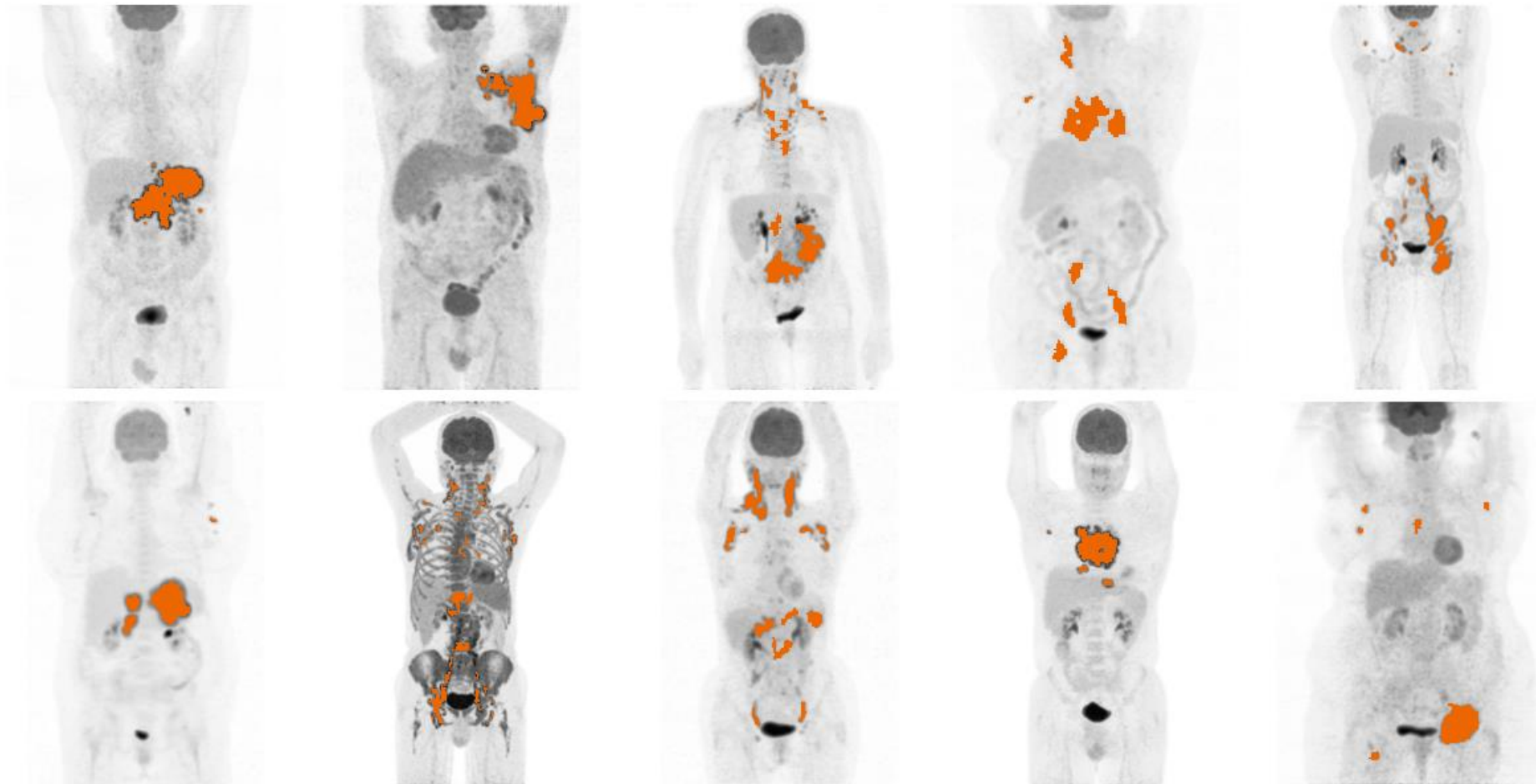


Les lymphomes

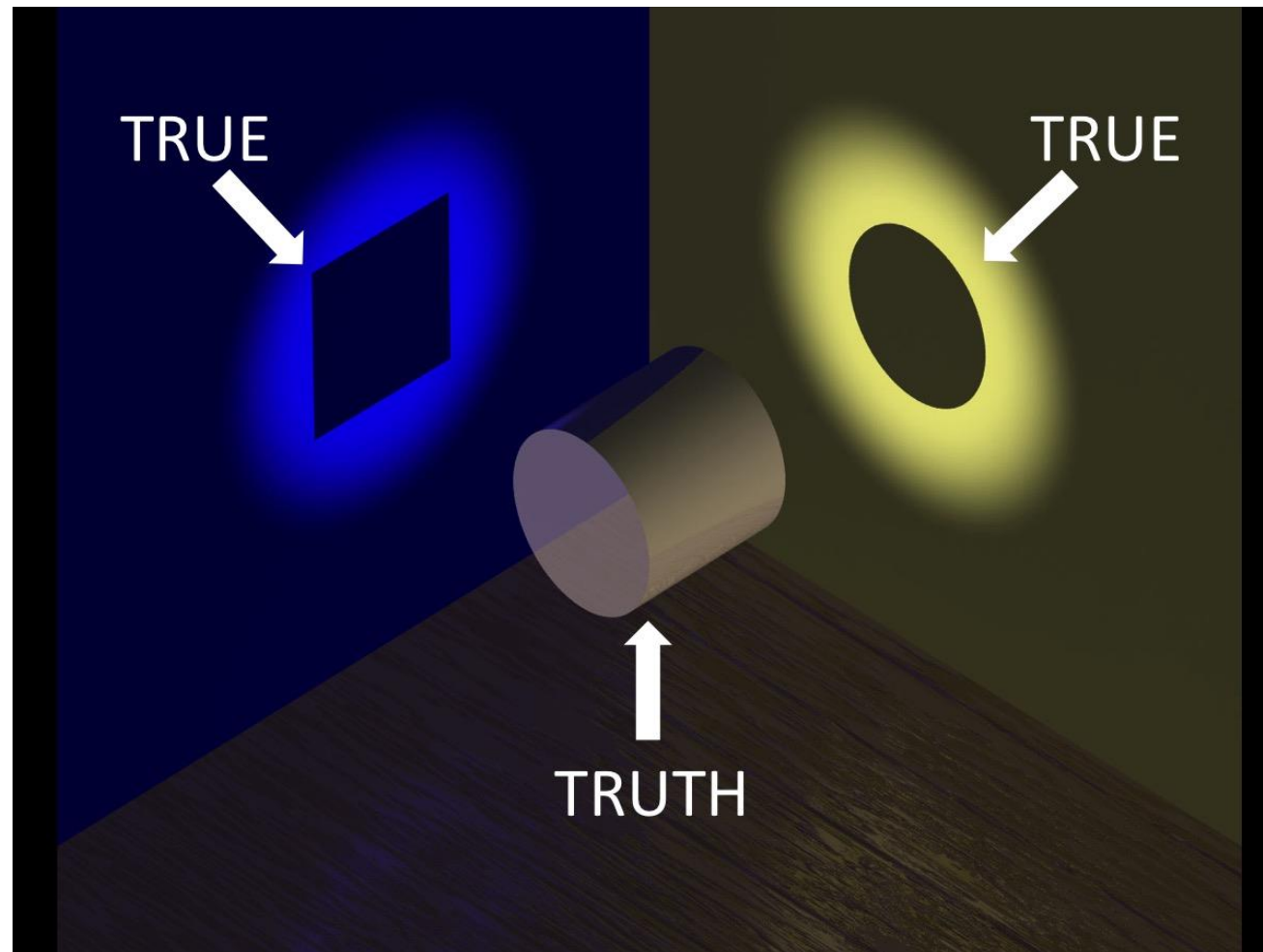
- TEP 18F-FDG/TDM
 - Diagnostic
 - Détection tumeurs
 - Evaluation de la réponse au traitement
- Volume métabolique tumoral total (TMTV):
seul facteur pronostique mesuré dans l'image
utilisé en clinique
- Le signal dans les régions saines n'est pas
interprété.



Le problème



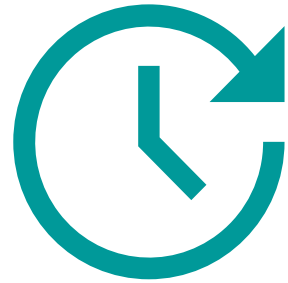
Le problème



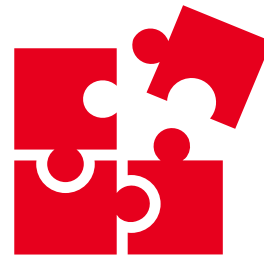
Le problème – Propriétés souhaitées



Nouveaux



Prédicatifs



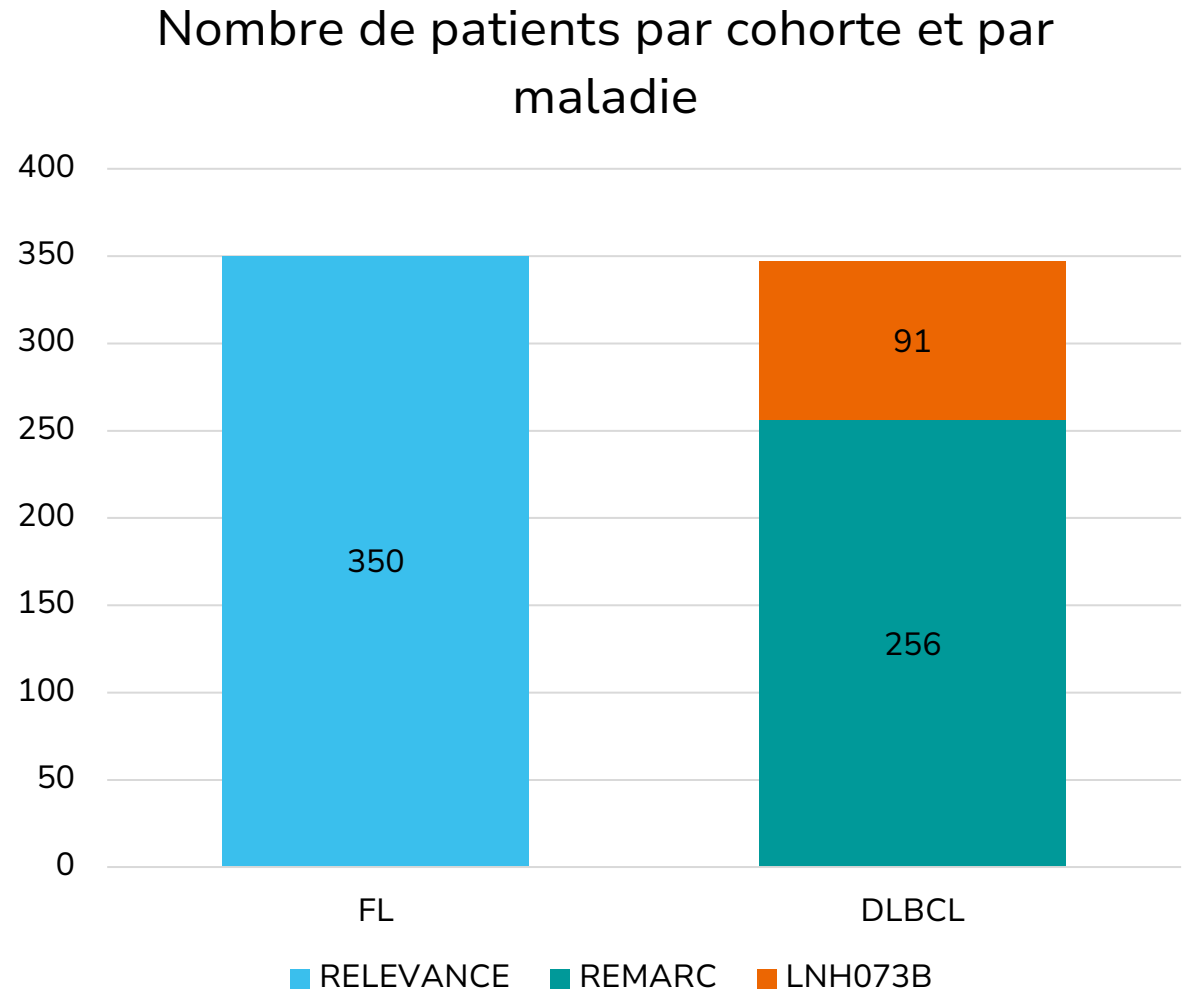
Complètent ce
qu'on sait déjà



Simples et
explicables

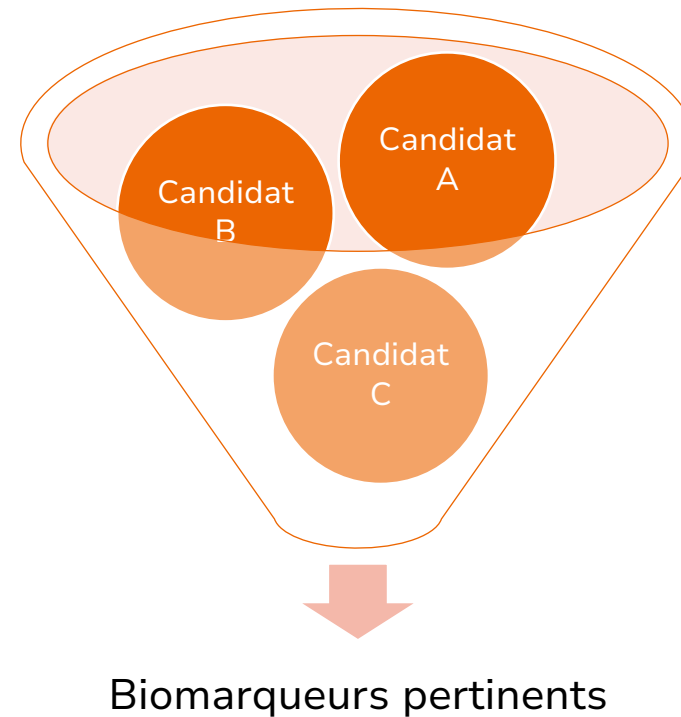
Les données

- 2 cohortes :
 - DLBCL (REMARC + LNH073B)
 - FL (RELEVANCE)
- TEP 18F-FDG/TDM baseline pour chaque patient
- Survie sans progression (PFS) et survie globale (OS) disponibles
- Traitement et comorbidités pour les FL



Recherche semi-automatique

- Construire automatiquement plusieurs biomarqueurs candidats
- Les tester un-à-un
- Ne garder que ceux qui ont les propriétés désirées
- Interpréter/simplifier les biomarqueurs sélectionnés



Avantages

Manuel	Semi-automatique
Lent	Rapide (milliers de candidats testés)
Peu exhaustif	Exhaustif (nombreuses mesures dans plusieurs endroits du corps)
Sujet aux faux positifs	Nombre d'expériences prédéfini -> taux de fausses découvertes contrôlé
Biaisé par l'observateur	Moins biaisé

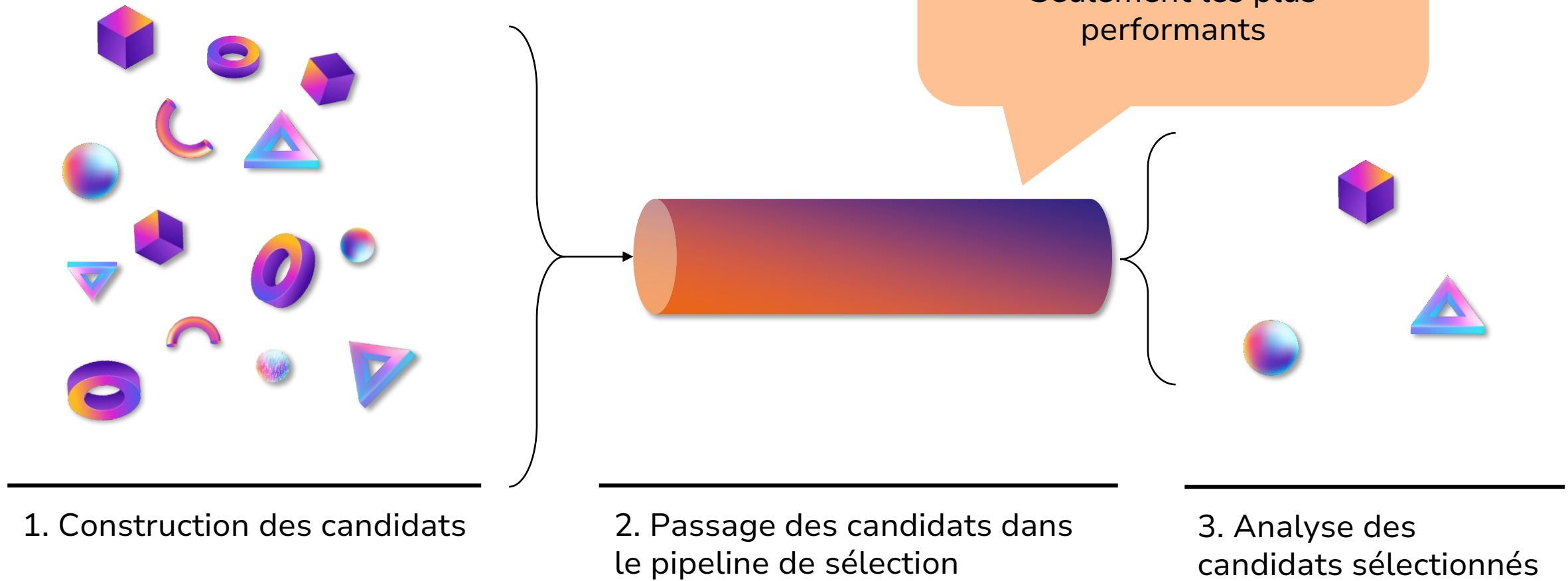
Recherche manuelle



Recherche semi-automatique



Processus de la recherche semi-automatique



Construction des candidats

- Candidats calculés sur la TEP et le scanner
- ROIs :
 - Lésions
 - Couronnes de tissus autour des lésions (shell)
 - Organes
- Organes segmentés par TotalSegmentator et MOOSE
- Features radiomiques extraites avec PyRadiomic pour chaque ROI dans chaque modalité
- Plusieurs agrégations à l'échelle du patient (min, max, median, mean, std, range)
- 7436 candidats créés



ROBI pipeline


- ROBI est le pipeline de sélection utilisé
- Composé de plusieurs filtres qui testent les candidats pour les propriétés désirées
- Les filtres sont arrangés dans un ordre optimisé pour maximiser le nombre de découvertes
- ROBI permet de contrôler le nombre de faux-positifs et estime leur nombre
- Publié sous la forme d'un package Python
- Son utilité va au-delà de la radiomique

README Apache-2.0 license

ROBI: Robust and Optimized Biomarker Identifier

ROBI is a selection pipeline that selects predictive biomarkers from any set of features. The selection is performed with a robust and adjustable control of the number of false positives as well as a control for confounders. ROBI can control for confounders and already known biomarkers in order to select only new and relevant information.

pypi package 0.0.16 License Apache 2.0 python 3.7-3.10



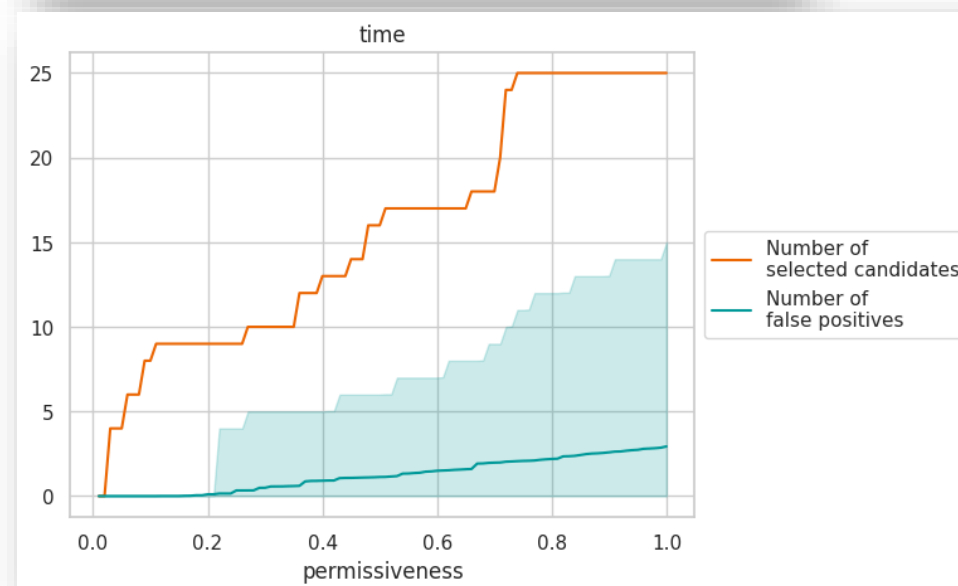
Key features:

- Robust control of the number of false positives by passing permuted datasets through the selection pipeline thousands of times. The proportion of false positive can be adjusted.
- Increased discovery rate via optimised feature selection.
- Reliable predictive power estimation through permutation tests instead of fixed thresholds.
- Selects only new information by controlling for confounders and correlations with known biomarkers.
- Fast parallelized implementation that can leverage both CPU and GPU for extensive tests: you can easily evaluate tens of thousands of potential biomarkers with millions of permutations in a few minutes.

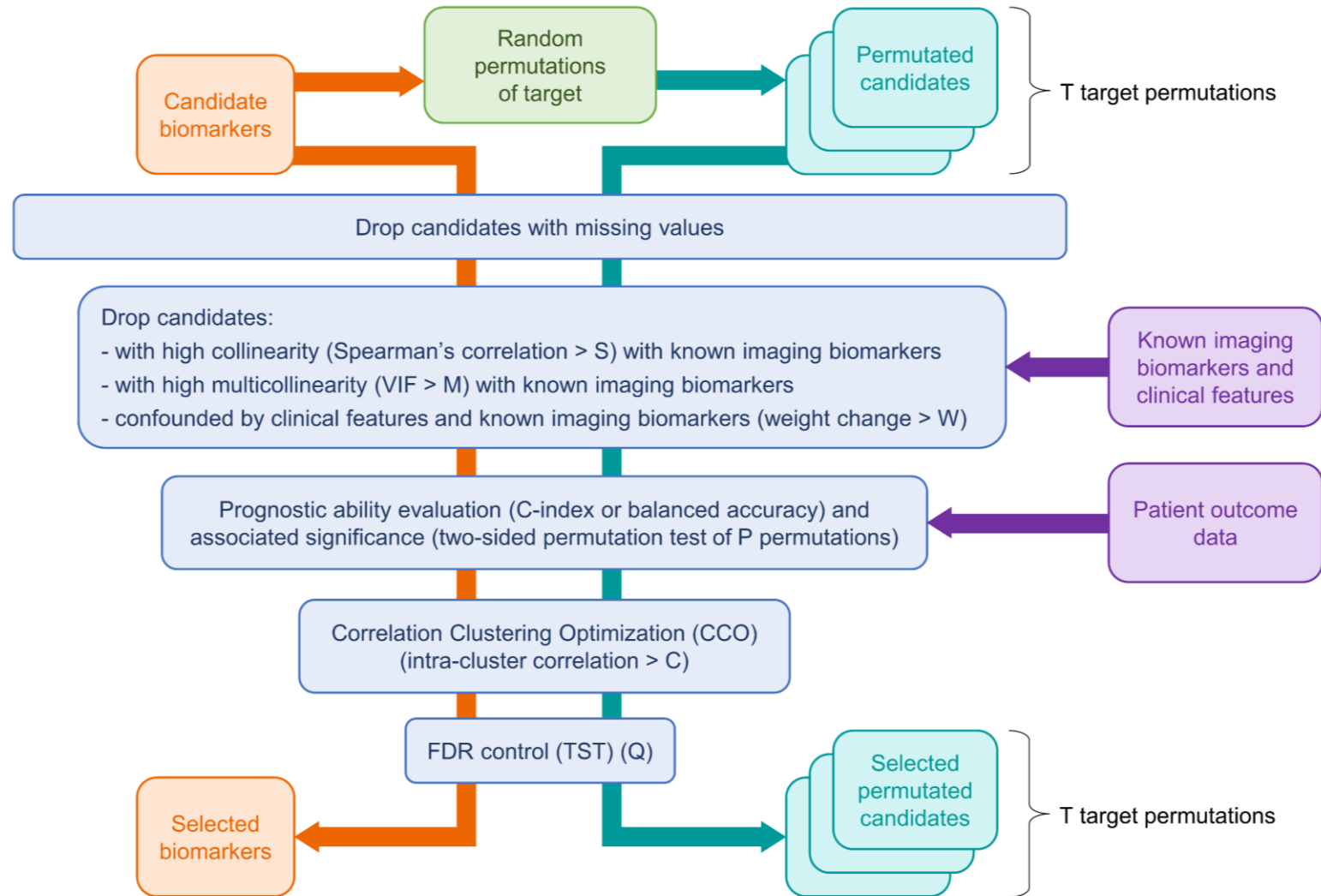
Installation

```
pip install robi
```

Although [PyTorch](#) is not required to use the package, ROBI runs much faster with its PyTorch implementation. The speed gain is great on CPU, and much greater on GPU. To use the PyTorch implementation, simply install PyTorch.



ROBI pipeline details

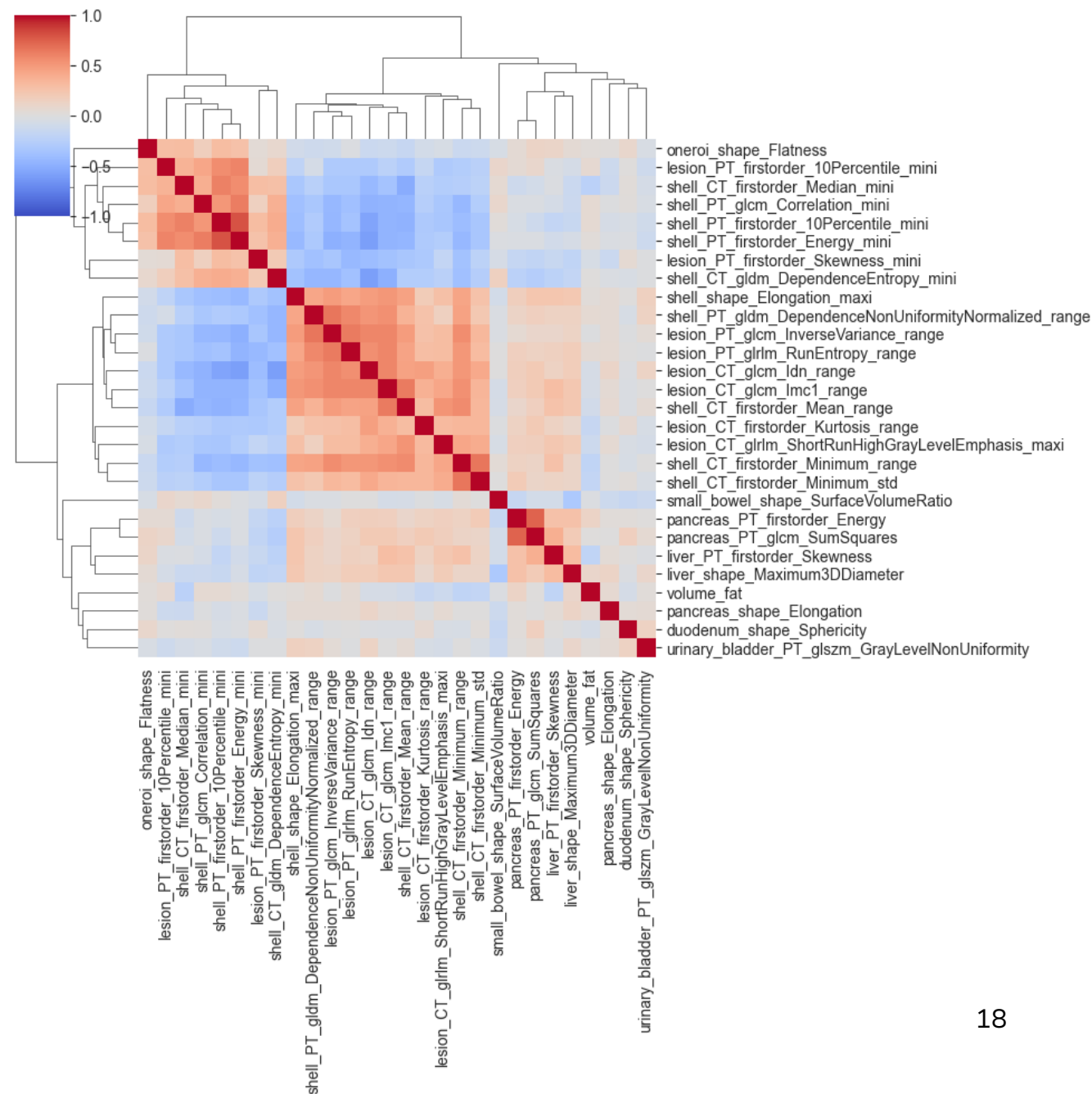


Candidats sélectionnés par ROBI (28)

Selected biomarker	C-index	Sign of the correlation with the risk
duodenum_shape_Sphericity	0.58 (p < 0.01)	+
lesion_CT_firstorder_Kurtosis_range	0.57 (p < 0.01)	+
lesion_CT_glcm_ldn_range	0.59 (p < 0.01)	+
lesion_CT_glcm_lmc1_range	0.59 (p < 0.01)	+
lesion_CT_glrlm_ShortRunHighGrayLevelEmphasis_maxi	0.59 (p < 0.01)	+
lesion_PT_firstorder_10Percentile_mini	0.56 (p < 0.02)	-
lesion_PT_firstorder_Skewness_mini	0.56 (p < 0.01)	-
lesion_PT_glcm_InverseVariance_range	0.57 (p < 0.01)	+
lesion_PT_glrlm_RunEntropy_range	0.57 (p < 0.01)	+
liver_PT_firstorder_Skewness	0.58 (p < 0.01)	+
liver_shape_Maximum3DDiameter	0.57 (p < 0.01)	+
oneroi_shape_Flatness	0.58 (p < 0.01)	+
pancreas_PT_firstorder_Energy	0.57 (p < 0.01)	+
pancreas_PT_glcm_SumSquares	0.57 (p < 0.01)	+
pancreas_shape_Elongation	0.58 (p < 0.01)	+
shell_CT_firstorder_Mean_range	0.57 (p < 0.01)	+
shell_CT_firstorder_Median_mini	0.59 (p < 0.01)	-
shell_CT_firstorder_Minimum_range	0.56 (p < 0.01)	+
shell_CT_firstorder_Minimum_std	0.58 (p < 0.01)	+
shell_CT_gldm_DependenceEntropy_mini	0.60 (p < 0.01)	-
shell_PT_firstorder_10Percentile_mini	0.59 (p < 0.01)	-
shell_PT_firstorder_Energy_mini	0.58 (p < 0.01)	-
shell_PT_glcm_Correlation_mini	0.59 (p < 0.01)	-
shell_PT_gldm_DependenceNonUniformityNormalized_range	0.57 (p < 0.01)	+
shell_shape_Elongation_maxi	0.59 (p < 0.01)	+
small_bowel_shape_SurfaceVolumeRatio	0.58 (p < 0.01)	-
urinary_bladder_PT_glszm_GrayLevelNonUniformity	0.57 (p < 0.01)	+
volume_fat	0.58 (p < 0.01)	-

Décodage des candidats sélectionnés

- ROBI ne sélectionne que les candidats les plus prédictifs
- Un candidat codant plusieurs informations biologiques pronostiques sera donc favorisé
- Beaucoup de candidats sélectionnés codent plusieurs informations pronostiques
- Certaines informations sont codées par plusieurs candidats

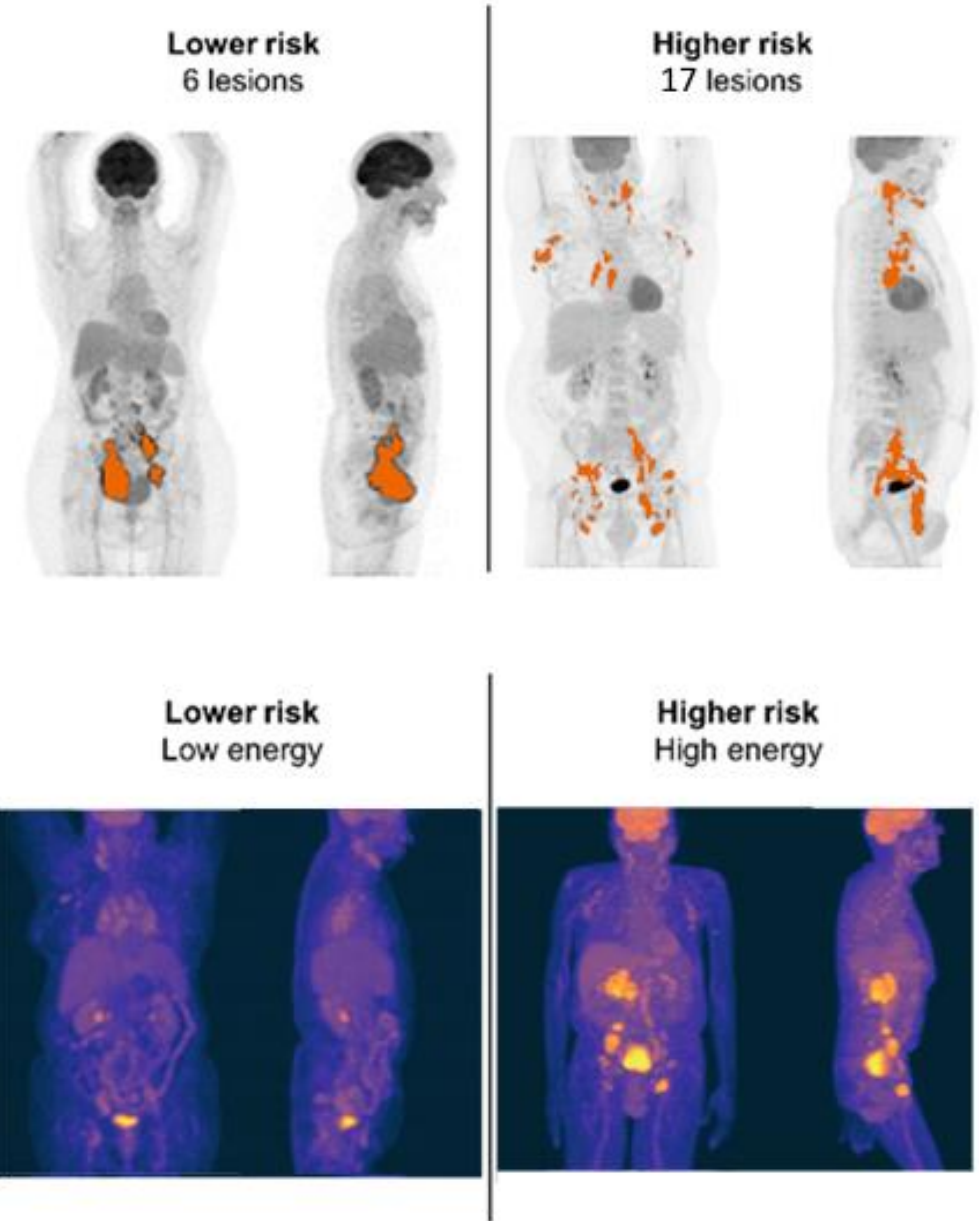


Informations réencodées manuellement (15)

Surrogate biomarkers associated with an increased risk	C-index (p-value)	Spearman's correlation with TMTV
Splenic involvement	0.54 (p < 0.01)	0.47
Liver involvement	0.56 (p < 0.01)	0.52
Respiratory system involvement	0.55 (p < 0.01)	0.43
High tumoral activity in the pancreas	0.57 (p < 0.01)	0.31
High Tumor Energy from the PET image	0.60 (p < 0.01)	0.79
High number of invaded organs	0.55 (p < 0.06)	0.79
High number of lesions	0.57 (p < 0.01)	0.58
Presence of small lesions	0.44 (p < 0.03)	-0.32
Presence of lesion with low activity	0.44 (p < 0.02)	-0.26
Presence of lesion in a region of homogeneous density	0.40 (p < 0.01)	-0.31
Low volume of subcutaneous fat	0.41 (p < 0.01)	-0.04
Large liver	0.60 (p < 0.01)	0.24
Low elongation of the pancreas	0.42 (p < 0.01)	-0.12
Presence of homogeneous density lesion	0.57 (p < 0.01)	0.39
High roundness of tumor burden	0.42 (p < 0.01)	-0.19

Description de la charge tumorale

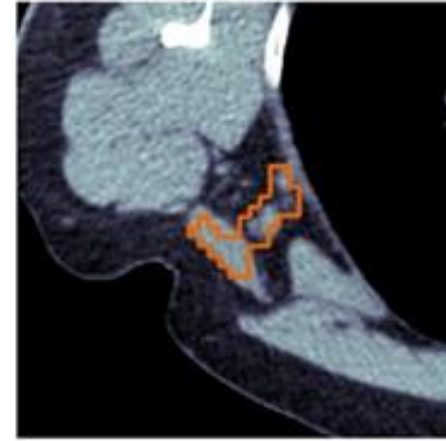
- Envahissement d'organes:
 - Rate
 - Foie
 - Poumons et trachée
 - Pancréas
- Plus il y a d'organes envahis, plus haut est le risque
- Plus il y a de lésions, plus haut est le risque
- Plus l'énergie tumorale totale ($\sum SUV^2$) est élevée, plus haut est le risque



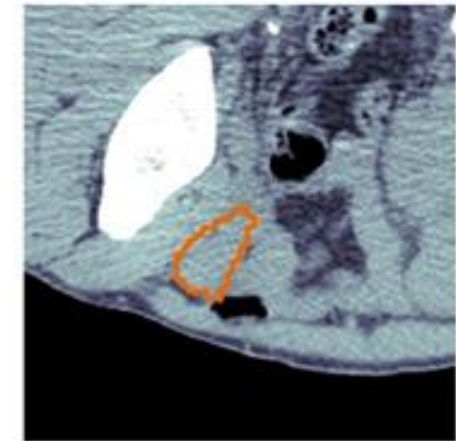
Présence de lésions particulières

- Plus la plus petite lésion est petite, plus le risque est élevé
- La présence de lésions très homogènes en densité et/ou dans une zone très homogène en densité augmente le risque
- La présence de lésions de faible activité augmente le risque

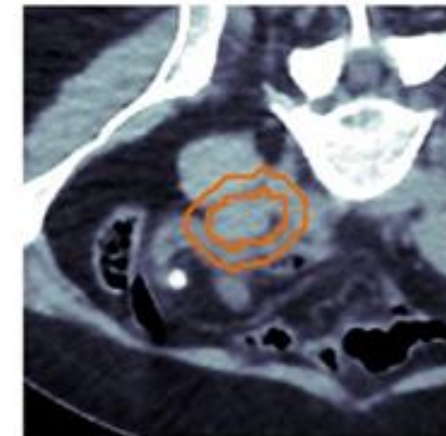
Lower risk
Heterogenous density
inside the lesion



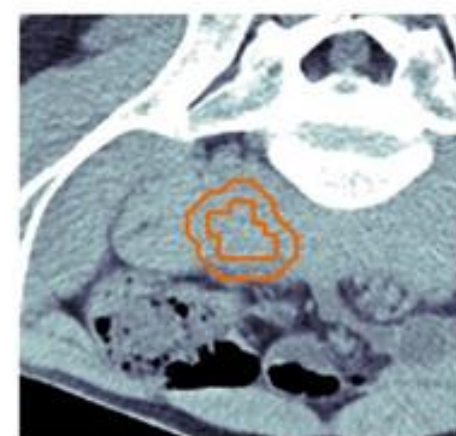
Higher risk
Homogeneous density
inside the lesion



Lower risk
Heterogenous density
around the lesion



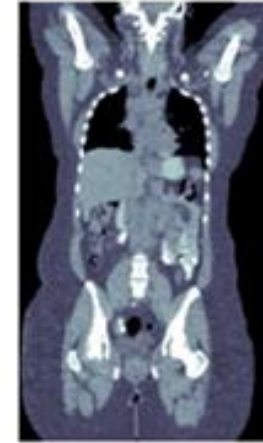
Higher risk
Homogeneous density
around the lesion



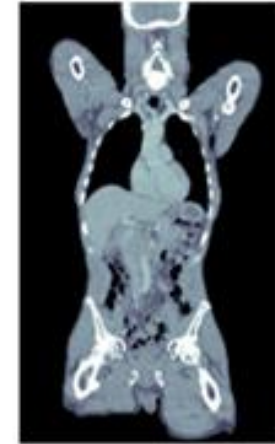
Comorbidités

- Plus le volume de graisse sous-cutanée est faible, plus haut est le risque (également lié à un ECOG élevé)
- Plus le foie est volumineux, plus le risque est élevé
- Moins le pancréas a une forme allongée, plus le risque est élevé
- Un pancréas peu allongé est aussi associé à du diabète (bACC 0.69 ($p < 0.001$)).

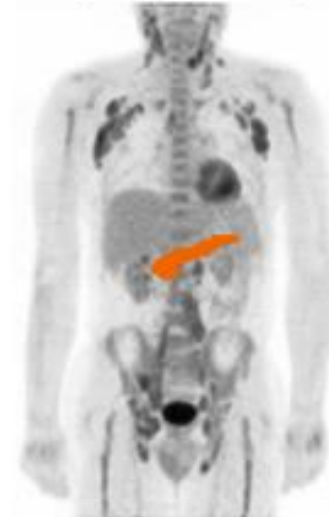
Lower risk
High volume of subcutaneous fat



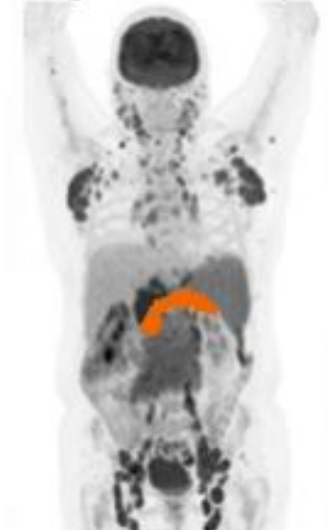
Higher risk
Low volume of subcutaneous fat



Lower risk
High elongation of the pancreas



Higher risk
Low elongation of the pancreas

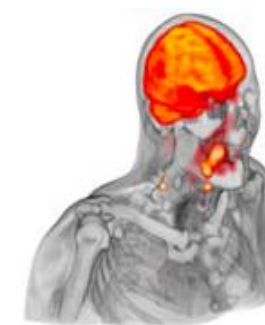
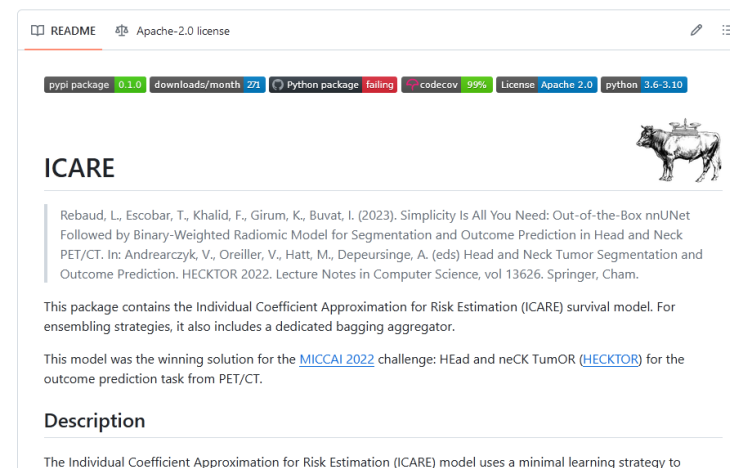


Le modèle ICARE

- Comment utiliser efficacement tous ces biomarqueurs ?
- Il faut construire des signatures (e.g. Cox model)
- Mais le risque d'overfitting est grand dans la prédiction de survie.
- ICARE réduit ce risque en apprenant seulement le strict minimum pour chaque feature:
 - Signe de la corrélation avec la target à prédire
 - Facteurs de normalisation (z-score)
- Avec ce modèle, notre équipe LITO a gagné le challenge HECKTOR en 2022
- Dans quelles conditions vaut-il mieux utiliser ICARE ?

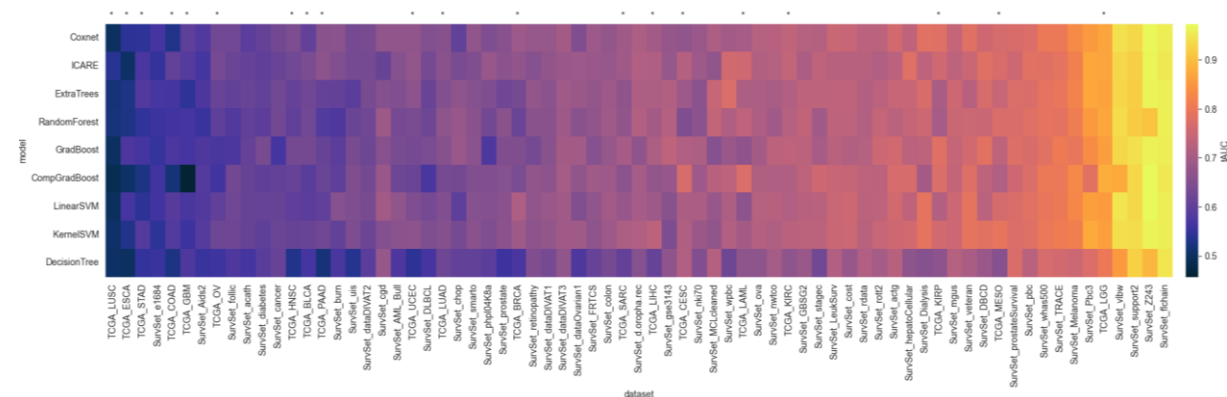
$$c_i = C_{index}(x_i, y)$$

$$s_i = \begin{cases} +1, & \text{if } c_i \geq 0.5 \\ -1, & \text{otherwise} \end{cases}$$

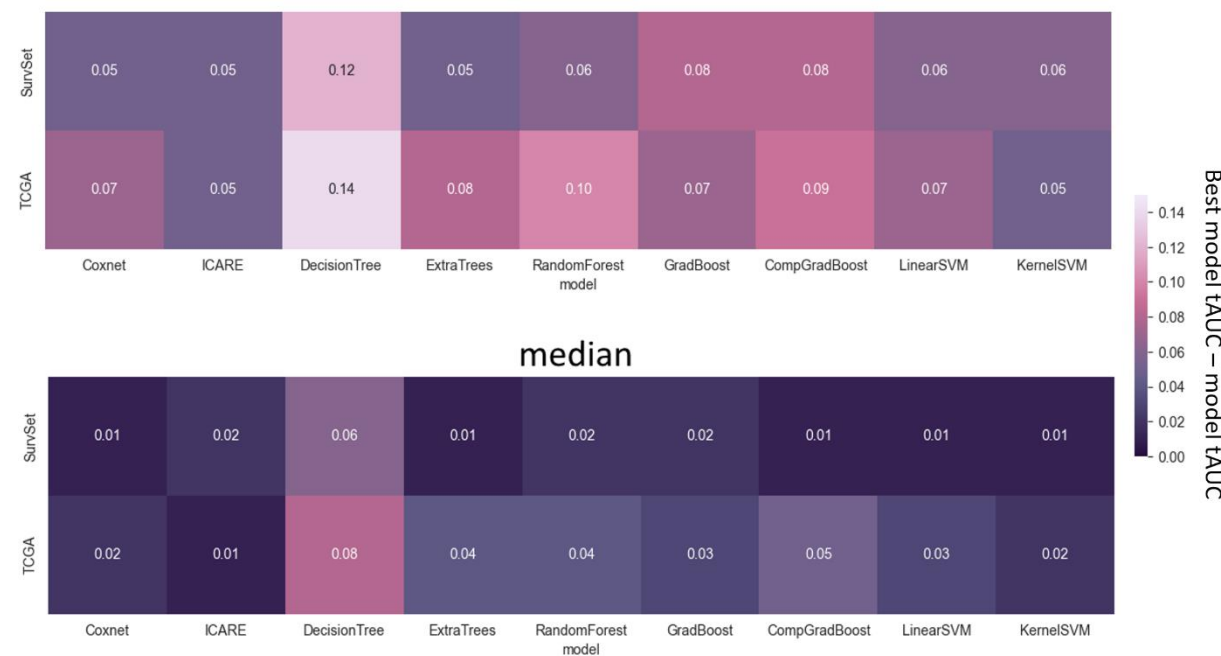


Evaluation de 9 modèles de machine learning sur 71 datasets médicaux

- 51 datasets proviennent de la collection SurvSet
- 20 datasets proviennent du TCGA
- 9 modèles (Cox, DecisionTree, Forests, SVMs, boosting, ICARE) ont été optimisés, entraînés et évalués sur tous les datasets
- Evalués en nested cross-validation avec le tAUC
- **Observation principale : le choix du modèle n'a pas un impact majeur sur les performances**

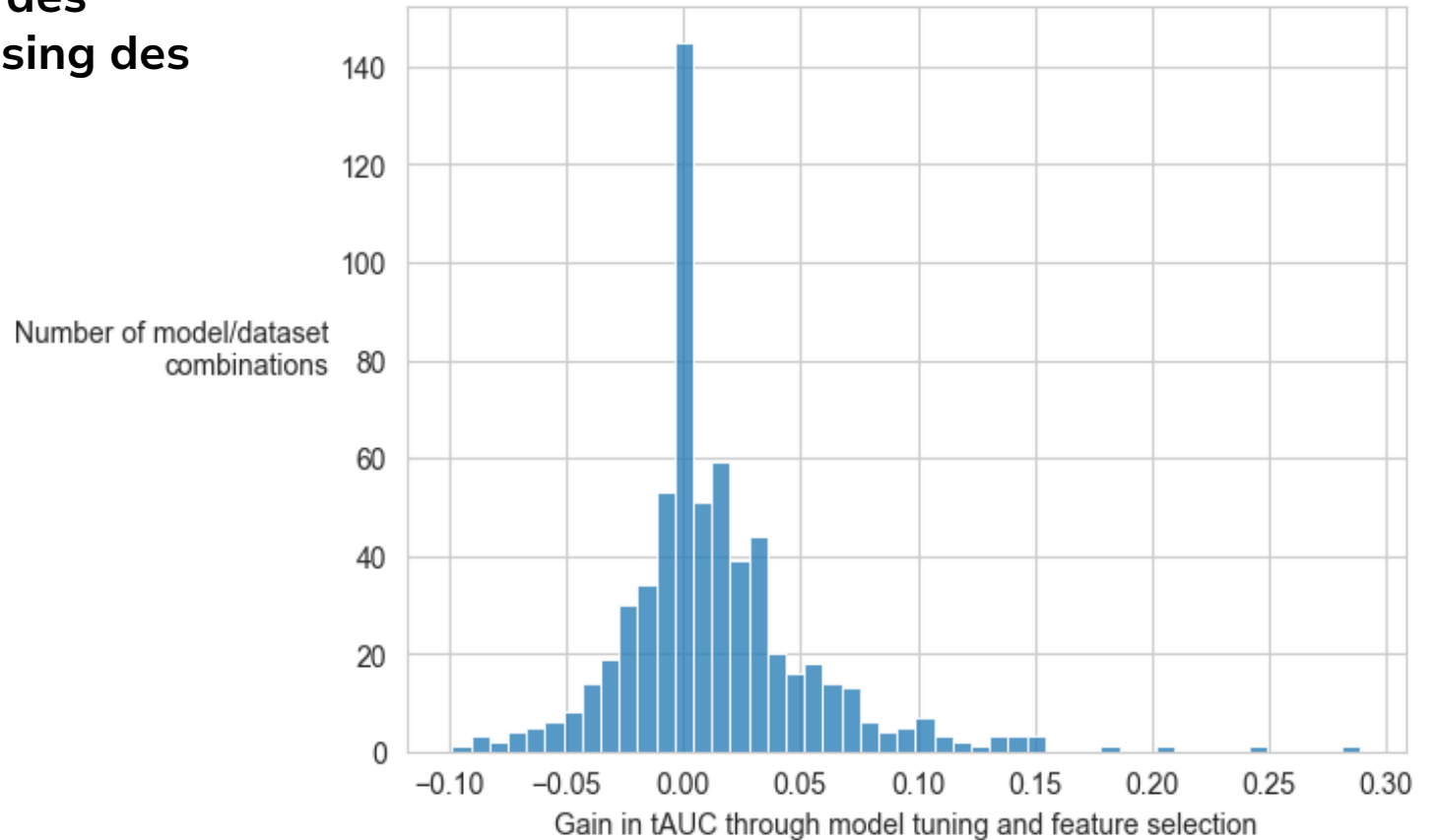


95th percentile



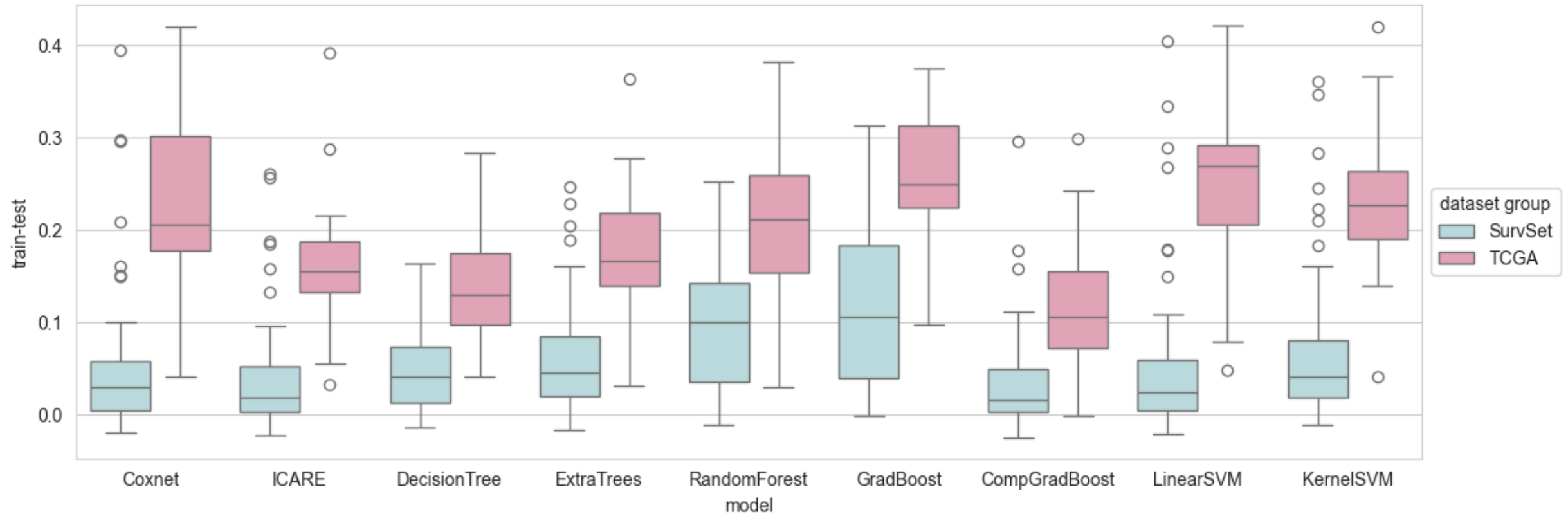
Evaluation de 9 modèles de machine learning sur 71 datasets médicaux

Seconde observation : l'optimisation des modèles, la sélection et le preprocessing des features n'ont pas d'impact majeur



Evaluation de 9 modèles de machine learning sur 71 datasets médicaux

- Troisième observation : certains modèles overfit moins que d'autres

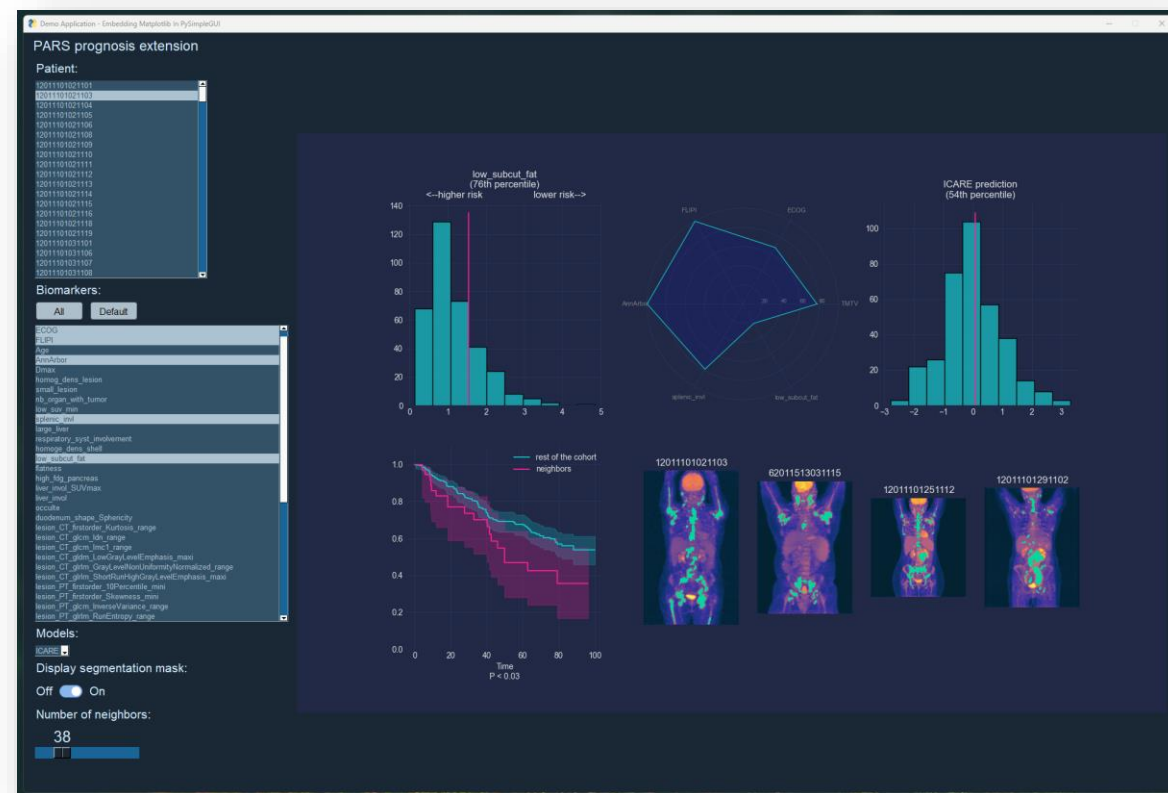


Evaluation de 9 modèles de machine learning sur 71 datasets médicaux

- Conclusions :
 - ICARE a des performances équivalentes aux autres modèles
 - ICARE overfit moins que d'autres modèles
 - ICARE est plus simple
- ICARE est un bon modèle pour partager de nouvelles signatures entre les centres

Extension PARS

- Je développe une extension de PARS qui intègre :
 - Les biomarqueurs identifiés
 - Différents modèles (dont ICARE)
 - Des visualisations pour donner plus d'intuition aux métriques
- L'outil est modulaire et de nouveaux modèles et features peuvent facilement être ajoutés
- Son but est de faciliter l'introduction et le test de nouvelles métriques pronostiques en clinique



Prochains objectifs

- Publication des différents articles
- Finir la rédaction de mon manuscrit de thèse
- Soutenance