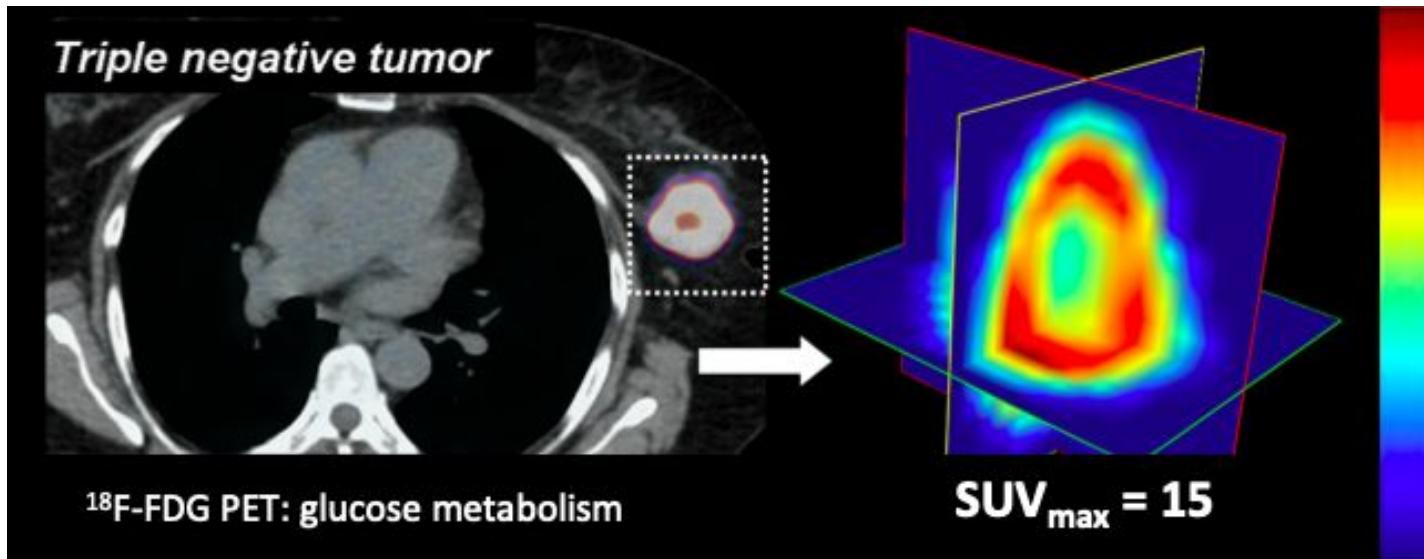




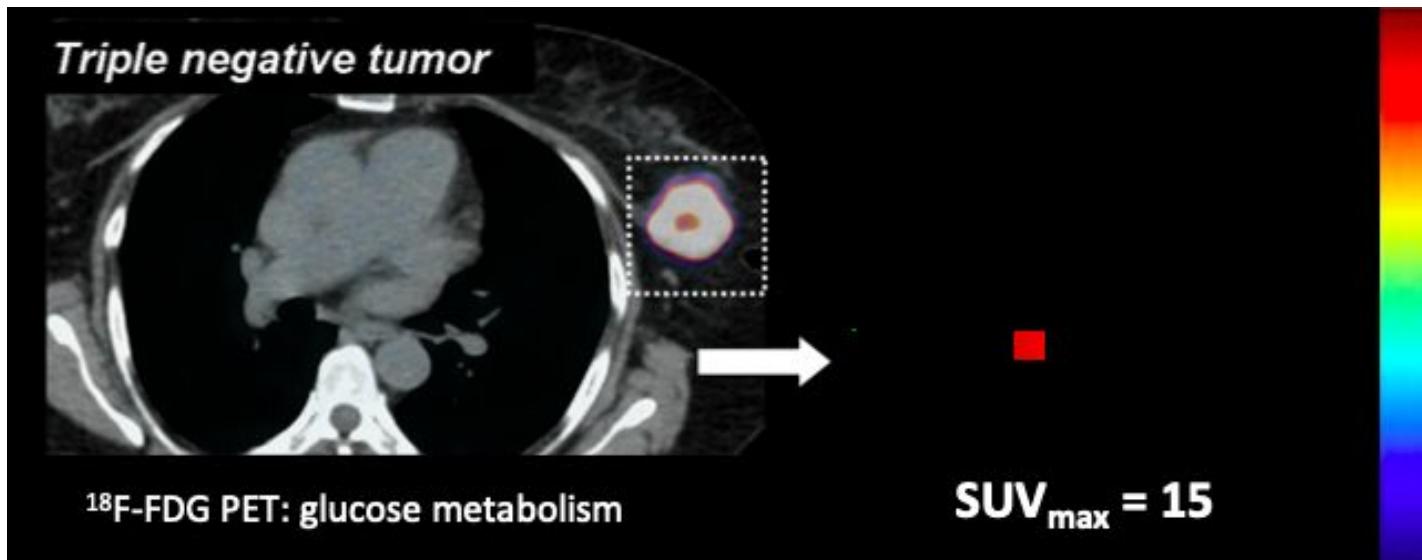
Qu'est ce qu'un modèle radiomique ?

Doctorants Orsay

Contenu des images actuellement largement sous-exploité



Contenu des images actuellement largement sous-exploité



Contenu des images actuellement largement sous-exploité

Radiology



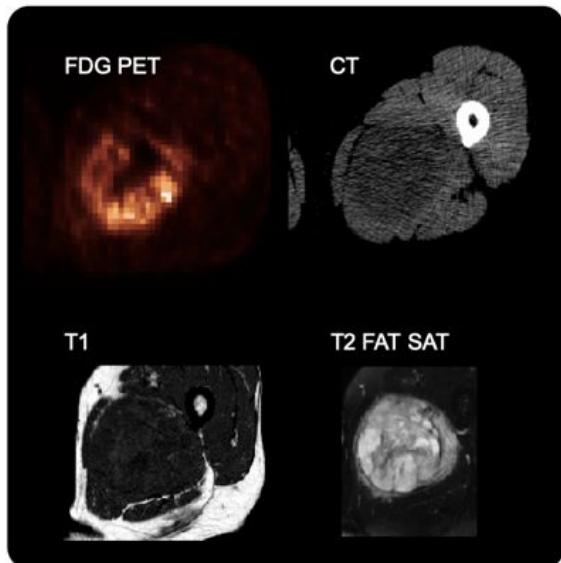
Robert J. Gillies, PhD
Paul E. Kinahan, PhD
Hedvig Hricak, MD, PhD, Dr(hc)

Radiomics: Images Are More than Pictures, They Are Data¹

In the past decade, the field of medical image analysis has grown exponentially, with an increased number of pattern recognition tools and an increase in data set sizes. These

“High-throughput extraction of quantitative features that result in the conversion of images into mineable data and the subsequent analysis of these data for decision support”

Contenu des images actuellement largement sous-exploité



RECHERCHE DE BIOMARQUEURS CONDUITE PAR LES DONNEES : **RADIOMIQUE** :
(# conduite par l'hypothèse)



Question d'intérêt => collecte des données (images)



Extraction d'un très grand nombre de descripteurs mathématiques



Analyse statistique multivariée **supervisée** => signature radiomique

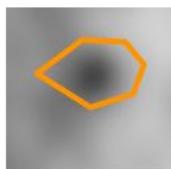
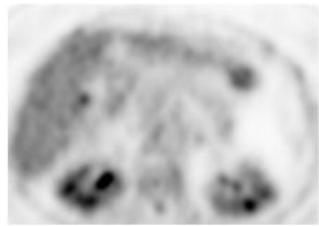


Interprétation => formulation de nouvelles hypothèses

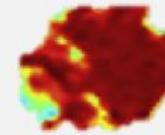
Définition de nouveaux biomarqueurs potentiels de façon optimisée

Que sont ces caractéristiques radiomiques (features) ?

Handcrafted
→



Histogram-based
(**SUVmax**, **SUVmean**)



Texture
(**entropy**)



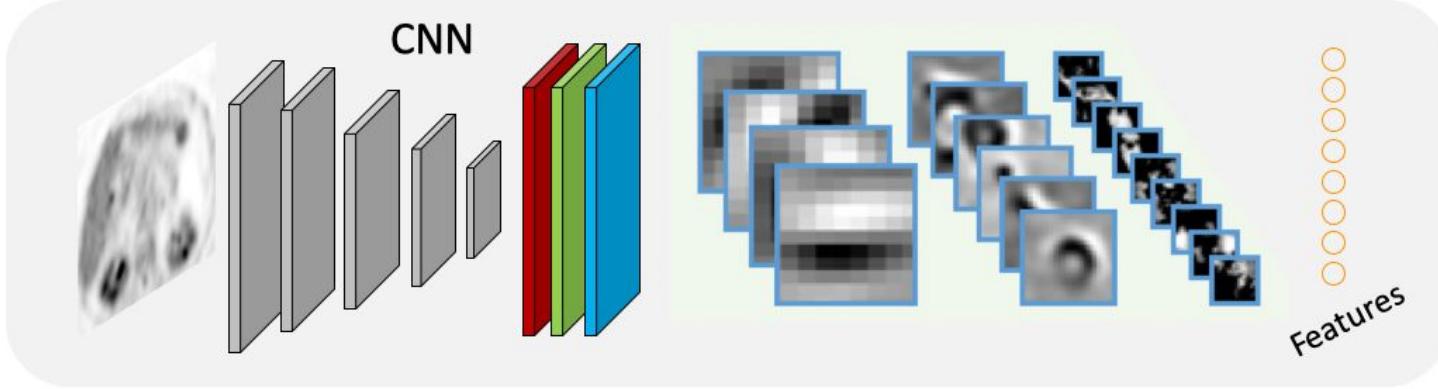
Shape
(**sphericity**)



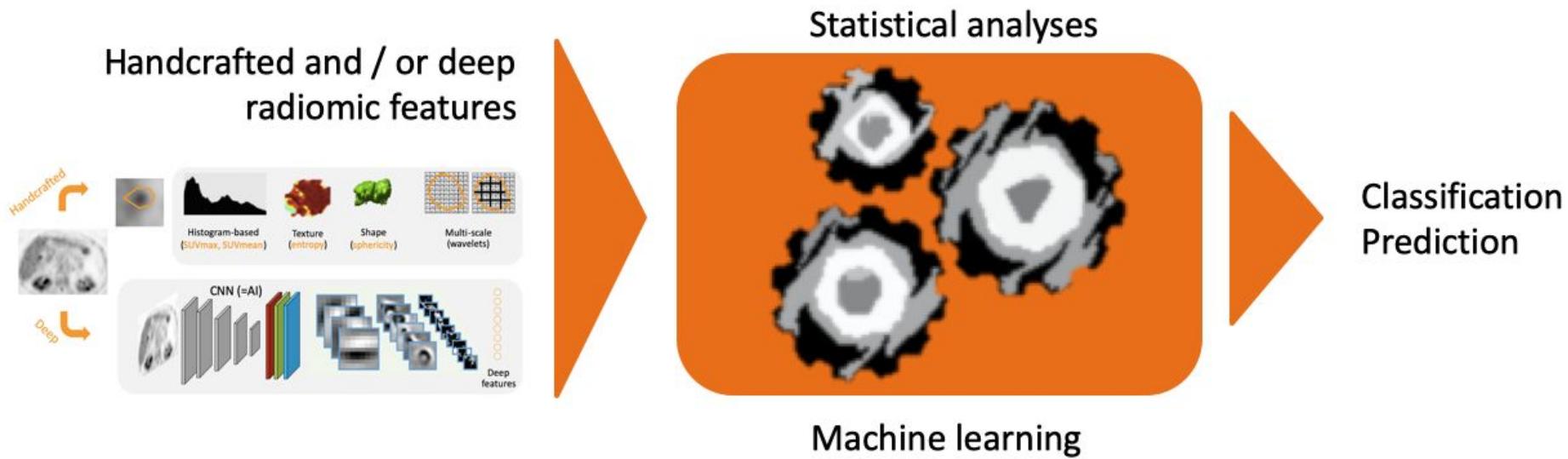
Multi-scale
(**wavelets**)



Deep
→

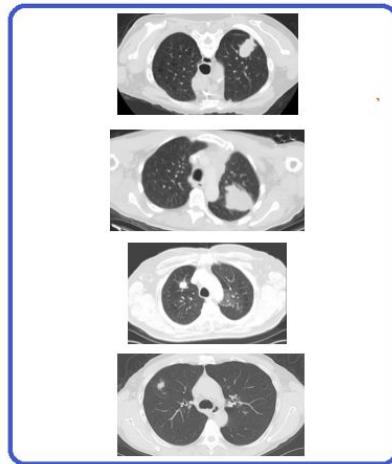


Que sont ces caractéristiques radiomiques (features) ?

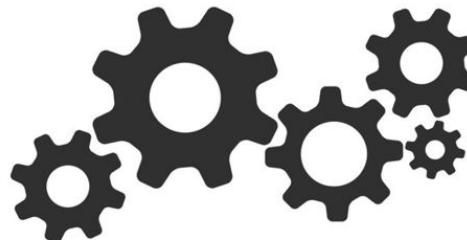


Objectifs pour la construction d'un modèle radiomique

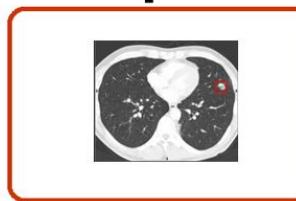
Données d'entraînement



Entraînement



Adénocarcinome !



Donnée indépendante

Sélectionner un modèle pertinent et **l'entraîner** sur un jeu de données dans l'espoir qu'il soit **performant** sur de nouvelles données **indépendantes**.

Quels sont les concepts et considérations de base ?

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	af	Antidiabétique	réduire_réstan	texture_mean	percentile_restan	area_mean	readmittence_mean	compactness_mean	mean_circularity_mean	mean_perimeter_mean	fractal_dimension_mean		
2	845200 M	17,99	20,56	22,08	1000	0,1203	0,2776	0,8025	0,5471	0,3435	0,0769		
3	845211 M	20,52	21,75	21,75	130	1,202	0,1204	0,2776	0,8025	0,5471	0,3435	0,0769	
4	8452003 M	18,49	21,25	21,25	130	1,202	0,1204	0,2776	0,8025	0,5471	0,3435	0,0769	
5	8452004 M	18,49	21,25	21,25	130	1,202	0,1204	0,2776	0,8025	0,5471	0,3435	0,0769	
6	8452002 M	20,29	24,54	23,51	1397	0,3095	0,3238	0,1508	0,5093	0,3085	0,0593		
7	8452006 M	12,45	15,7	8,37	477,1	0,1276	0,17	0,2579	0,8025	0,5471	0,3435	0,0769	
8	8452007 M	18,45	20,88	20,88	1000	0,1203	0,2776	0,8025	0,5471	0,3435	0,0769		
9	8452001 M	13,71	20,08	9,52	571,9	0,1309	0,1645	0,0946	0,8085	0,5194	0,3453	0,0781	
10	8452005 M	12,45	15,7	7,29	477,1	0,1276	0,17	0,2579	0,8025	0,5471	0,3435	0,0769	
11	8452003 M	12,48	24,04	8,37	475,9	0,1200	0,2296	0,2273	0,0945	0,203	0,0843		
12	8452008 M	18,42	23,28	24,27	792,8	0,1206	0,2296	0,2273	0,0945	0,203	0,0843		
13	8452009 M	19,39	17,09	19,39	391	0,8971	0,2296	0,2273	0,0945	0,203	0,0843		
14	8452010 M	15,17	24,8	15,14	1123	0,0919	0,2296	0,2273	0,0945	0,203	0,0843		
15	8452011 M	13,82	23,03	13,82	792,7	0,1203	0,2296	0,2273	0,0945	0,203	0,0843		
16	8452012 M	13,73	22,61	9,46	570,8	0,1143	0,2296	0,2128	0,0805	0,206	0,0760		
17	8452013 M	14,44	22,92	9,46	570,7	0,1143	0,2296	0,2128	0,0805	0,206	0,0760		
18	8452005 M	14,68	20,15	9,46	686,5	0,09887	0,0779	0,0729	0,0525	0,1598	0,0702		
19	8452001 M	18,13	20,08	20,11	798,8	0,1117	0,2013	0,1712	0,0519	0,2164	0,0726		
20	8452004 M	18,13	20,08	20,11	1400	0,1117	0,2013	0,1712	0,0519	0,2164	0,0726		
21	8452006 R	13,54	14,06	9,46	346,3	0,09779	0,08124	0,06464	0,04791	0,1065	0,05764		
22	8452005 R	13,08	25,71	9,47	520	0,10795	0,1217	0,04908	0,0311	0,2907	0,04013		



Comment apprêhender les données et évaluer leur capacité à construire un modèle pertinent ?

Comment sélectionner un modèle performant, adapté aux données et à la question posée ?

Comment entraîner ce modèle efficacement et tirer le meilleur parti des données d'entraînement ?

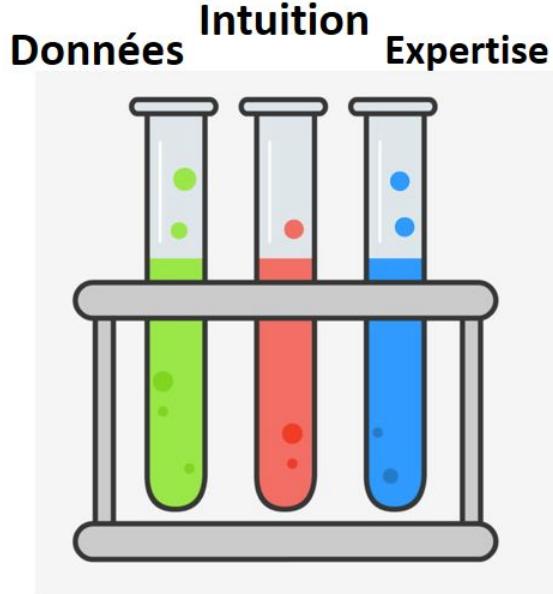
Comment garantir aux utilisateurs la performance de ce modèle pour de nouvelles données ?

Toutes ces questions cachent un nombre très divers de méthodes. L'objectif est de sélectionner et imbriquer les plus adaptées pour répondre aux questions, **dans la limite de l'information contenue dans les données.**

Tout est une histoire de données !

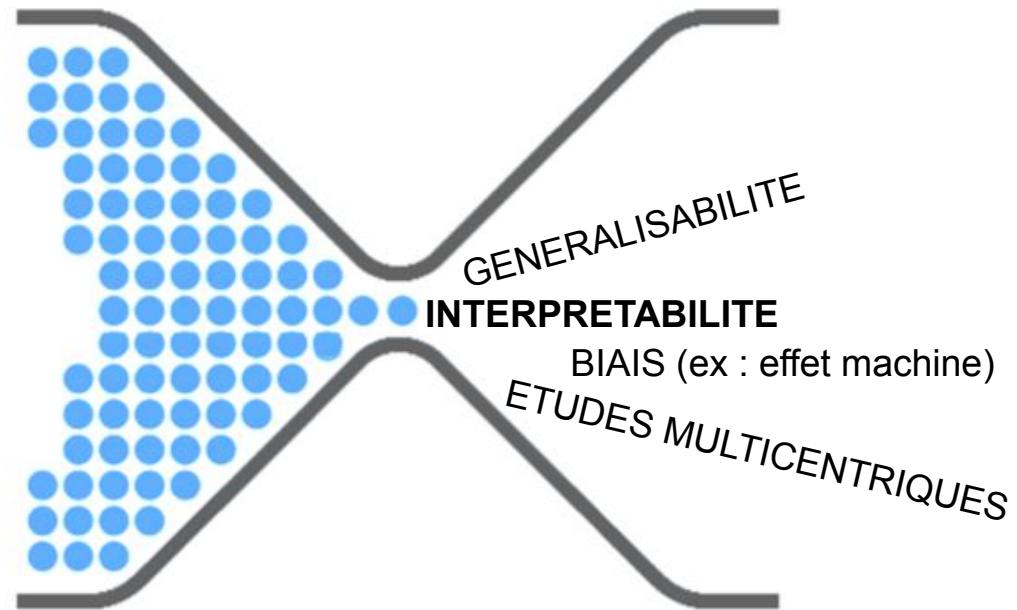


La construction d'un modèle radiomique résulte d'un **démarche expérimentale**. On **observe des données** censées rendre compte de la réalité et on cherche à en déduire une loi qui soit **toujours performante**.



- Tous nos résultats et conclusions doivent être mis en perspective par rapport à nos données initiales !
- **Plus on a de données mieux c'est !**

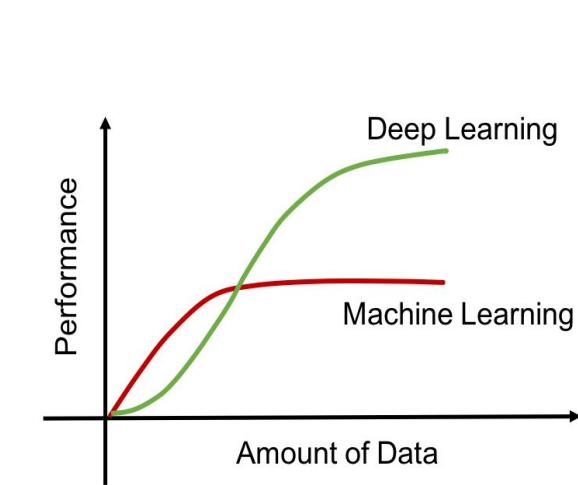
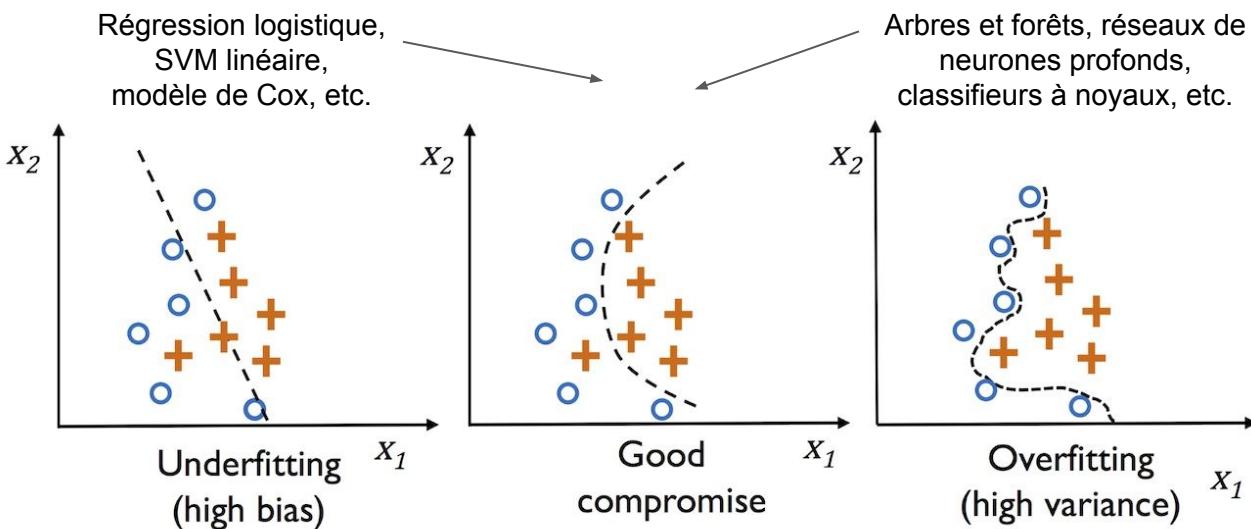
De nombreux challenges



Un très grand nombre de **résultats prometteurs**.

Aucun modèle déployé en clinique pour utilisation prédictive ou étude conduite par l'hypothèse.

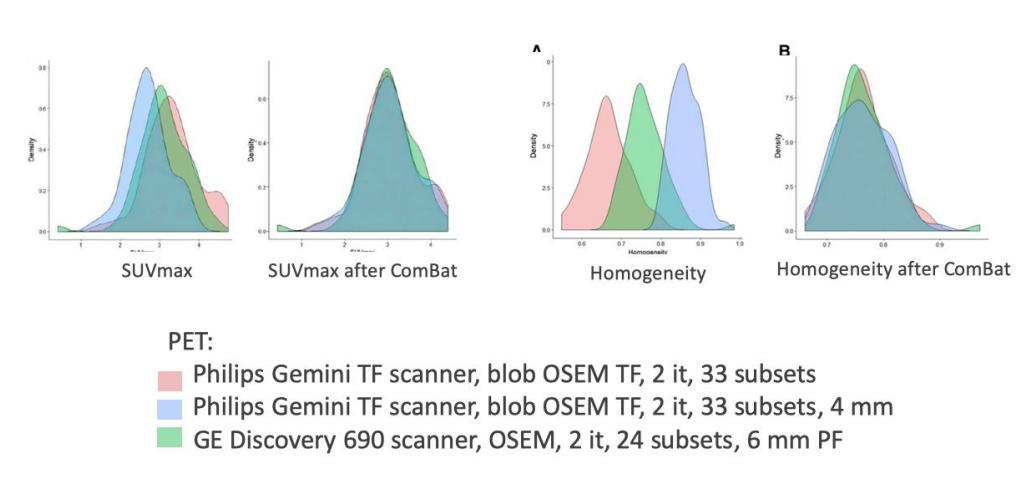
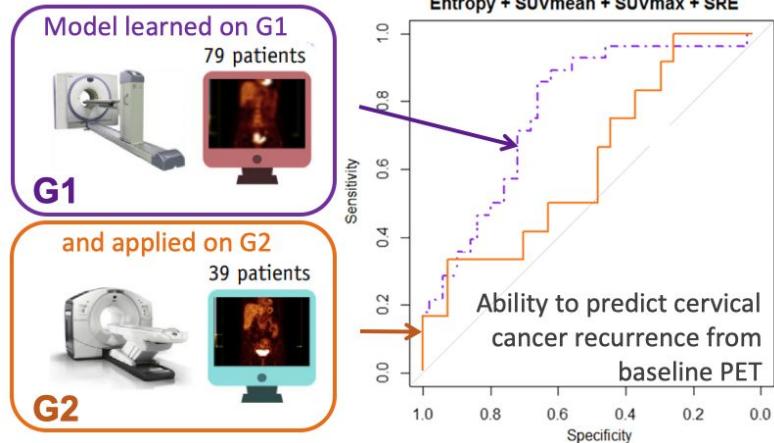
Généralisabilité



Bias-variance trade-off !

Dépend fortement du **nombre d'individus** et de la méthode utilisée !

Biais



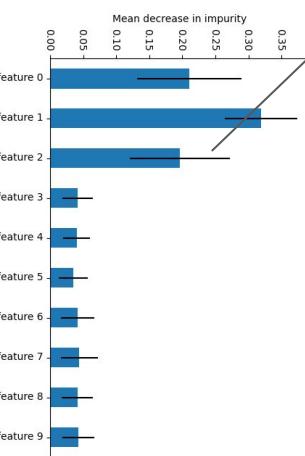
Les modèles sont sensibles aux **configurations d'acquisition des images** !
Diminution possible grâce à l'**harmonisation** (ex : ComBat, standardisation par lots).

Attention ! Certains biais sont plus insidieux !

Interprétabilité

Interprétabilité globale (modèle)

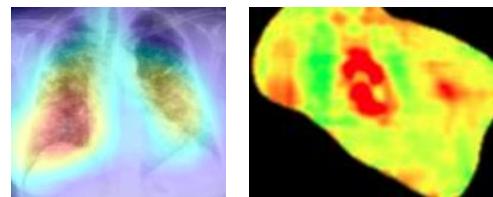
- 0.627 (± 0.601) * PET_original_glcn_ClusterShade
- 0.548 (± 0.546) * CT_original_glcn_Correlation
- 0.550 (± 0.782) * PET_original_glcn_InverseVariance
+ 2.76 (± 0.895) * PET_original_fo_Skewness
- 1.58 (± 0.690) * PET_original_gldm_SmallDepLowGrayLevelEmphasis
- 1.31 (± 0.771) * PET_original_gldm_GrayLevelNonUniformity
- 0.667 (± 0.539)



Interprétabilité locale (individu)

QUANTITATIVE

SPATIALE (images)



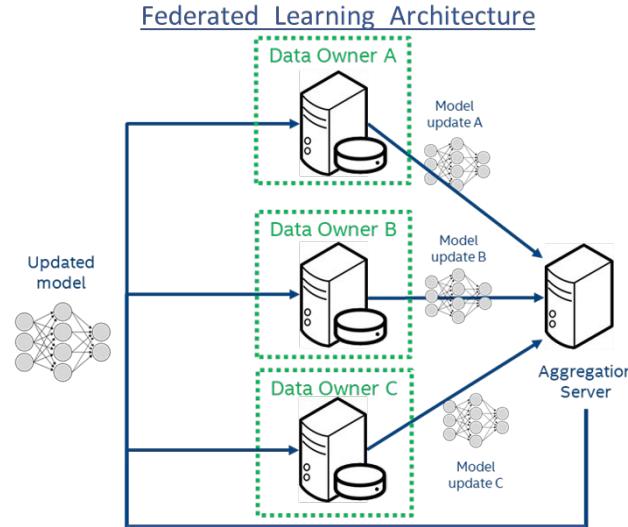
Transparence algorithmique (technique)

MAITRISE ET CONNAISSANCE DES TENANTS ET ABOUTISSANTS TECHNIQUES DE LA METHODE UTILISEE

Si nous comprenons précisément ce sur quoi se base les modèles :

1. Nous pouvons faire **émerger des hypothèses**,
2. Nous pouvons savoir si nous pouvons leur **faire confiance**,
3. Il nous est plus facile de **déceler certains biais potentiels**.

Etudes multicentriques



Le nombre d'individus est **crucial** ! Un base de données **représentative du contexte applicatif aussi** !

1. Collaboration et **partage de données (RGPD)**
2. Collaboration et **apprentissage fédéré**

Merci pour votre attention !

