

# APPLICATIONS RADIOMIQUES EN CANCÉROLOGIE

L Fournier, A Bouchouicha, J Deidier, L Duron, D Balvay, B Rance

Service de Radiologie, Service de Bioinformatique, Biostatistiques

Hôpital Européen Georges Pompidou

Team 2 "Imaging of angiogenesis" UMR 970, PARCC@HEGP

INSERM UMR\_S 1138 eq22 "Information Sciences to support Personalized Medicine"

# CONFLITS D'INTERETS

## Industriels

- ❑ Droits d'auteur de Merck et Pfizer
- ❑ Colloques de Sanofi and Novartis
- ❑ Financement recherche d'Invectys

## Financements publics

- ❑ Fondation pour la Recherche Médicale
- ❑ Fonds Unique Interministériel-Régions (BPI)

# ENJEUX DE L'IMAGERIE EN CANCÉROLOGIE

Prévention



Prédiction  
(facteurs de risque)

Surveillance

Diagnostic



Diagnostic  
précoce

Diagnostic  
différentiel

Pronostic



Gravité de la maladie  
(stratification)

Traitement



Sélection des  
patients

Efficacité des  
traitements

Toxicité des  
traitements

# DÉVELOPPEMENT DE NOUVEAUX BIOMARQUEURS

## Hypothesis Driven Research



**Nouveau  
biomarqueur  
potentiel**



Formuler une  
hypothèse

Collecter les  
données

Analyser &  
interpréter



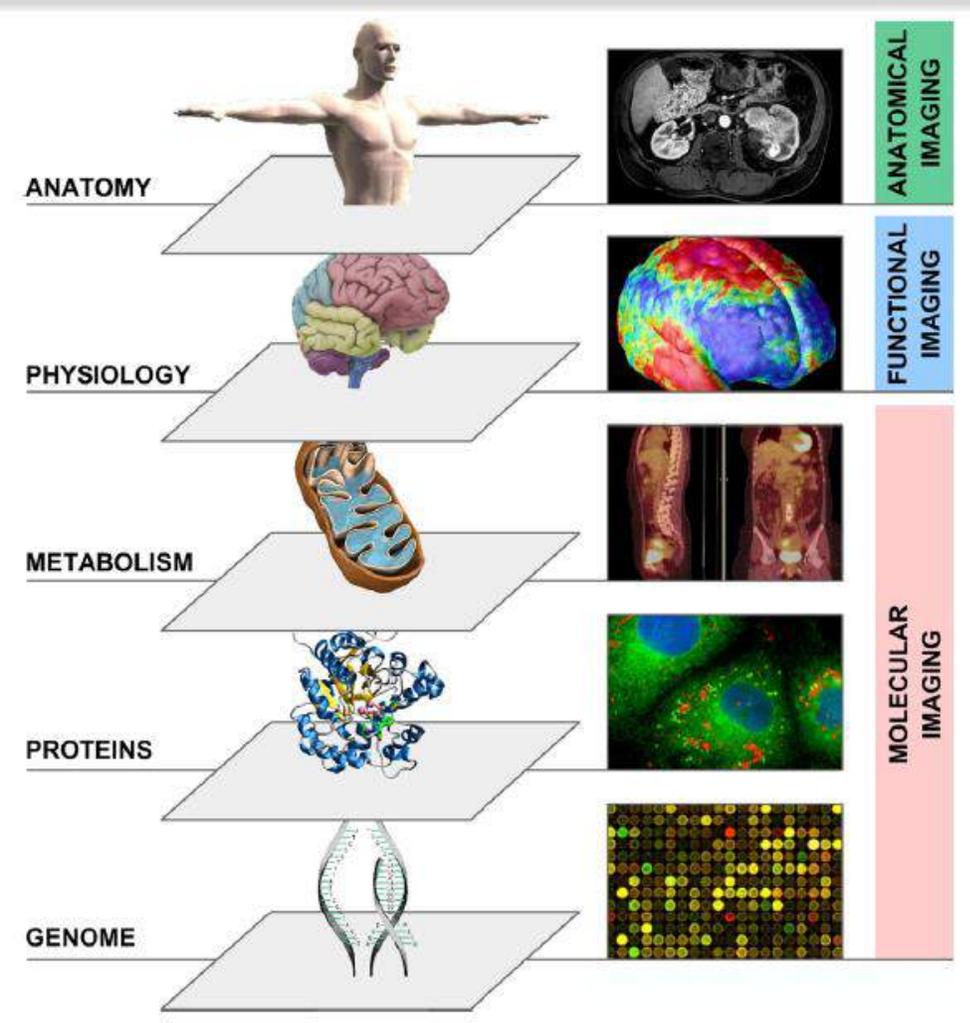
Validation  
technique

Validation  
clinique

# CHANGEMENT DE PARADIGMES

Deux concepts:

- 1) **Tester chaque biomarqueur = coûteux** en argent et en temps  
→ besoin d'optimiser la démarche de découverte de nouveaux biomarqueurs
- 2) L'image n'est qu'une **expression phénotypique** de la biologie sous-jacente  
→ probablement plus d'information dans l'image que ce qu'on extrait

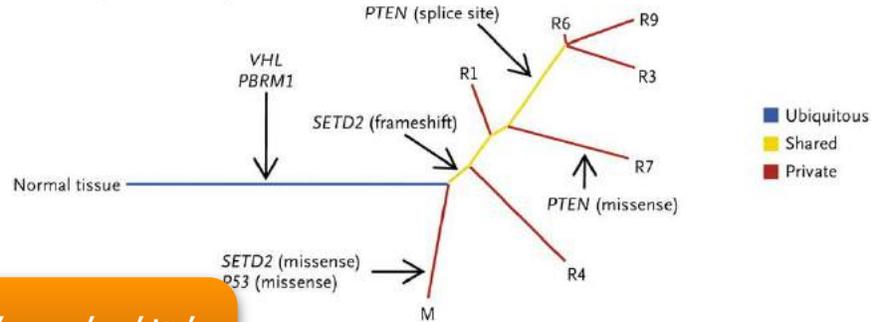


# L'IMAGERIE EST UN PHÉNOTYPE

*Lambin et al., Eur J Cancer 2012*

# HÉTÉROGÉNÉITÉ INTRA-LÉSIONNELLE

B Phylogenetic Relationships of Tumor Regions

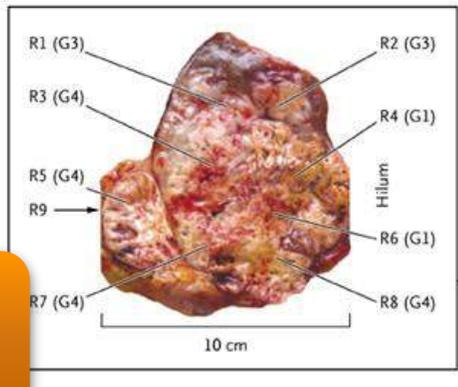


Hétérogénéité génétique

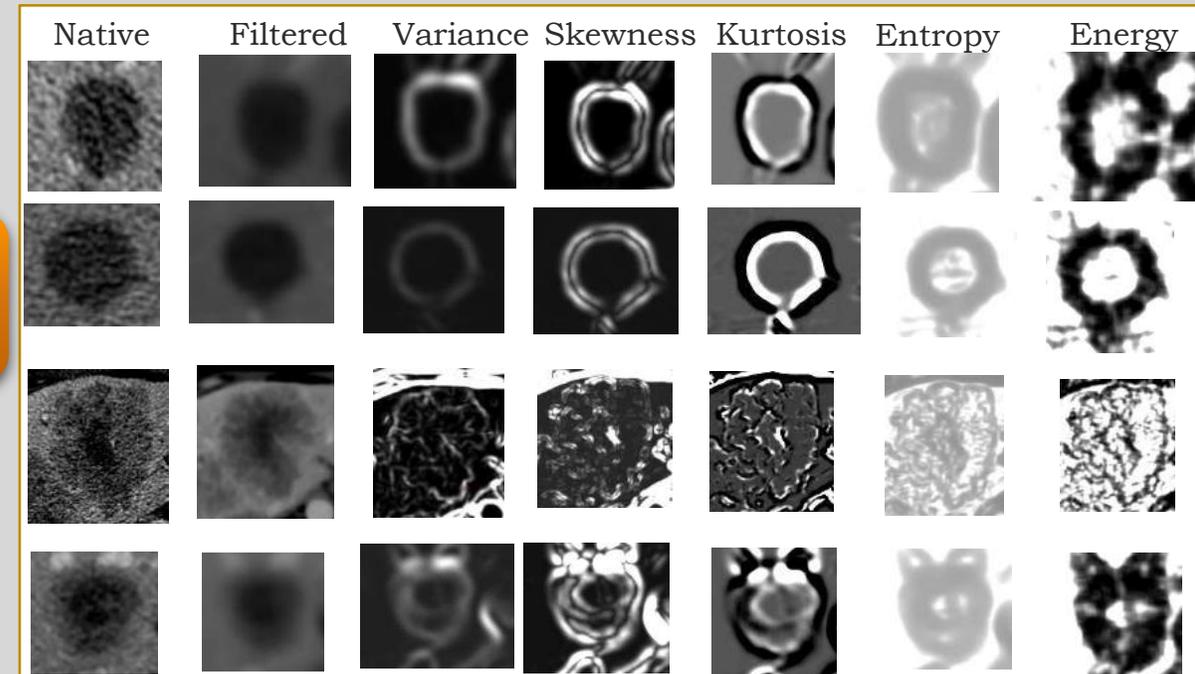
Gerlinger et al. N Engl J Med 2012

Hétérogénéité tissulaire

A Biopsy Sites



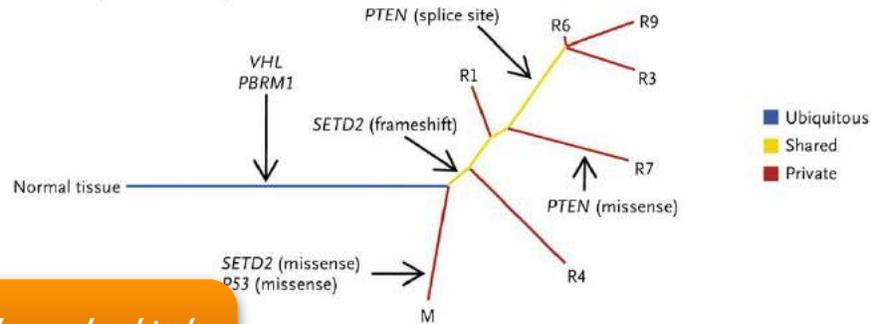
Hétérogénéité en imagerie



M Rozand, Master2

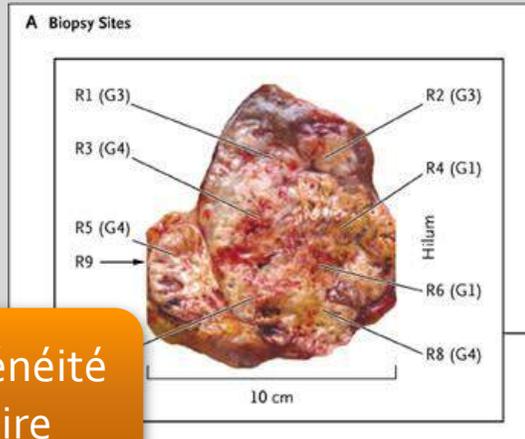
# HÉTÉROGÉNÉITÉ INTRA-LÉSIONNELLE

B Phylogenetic Relationships of Tumor Regions



Hétérogénéité  
génétique

Gerlinger et al. N Engl J Med 2012

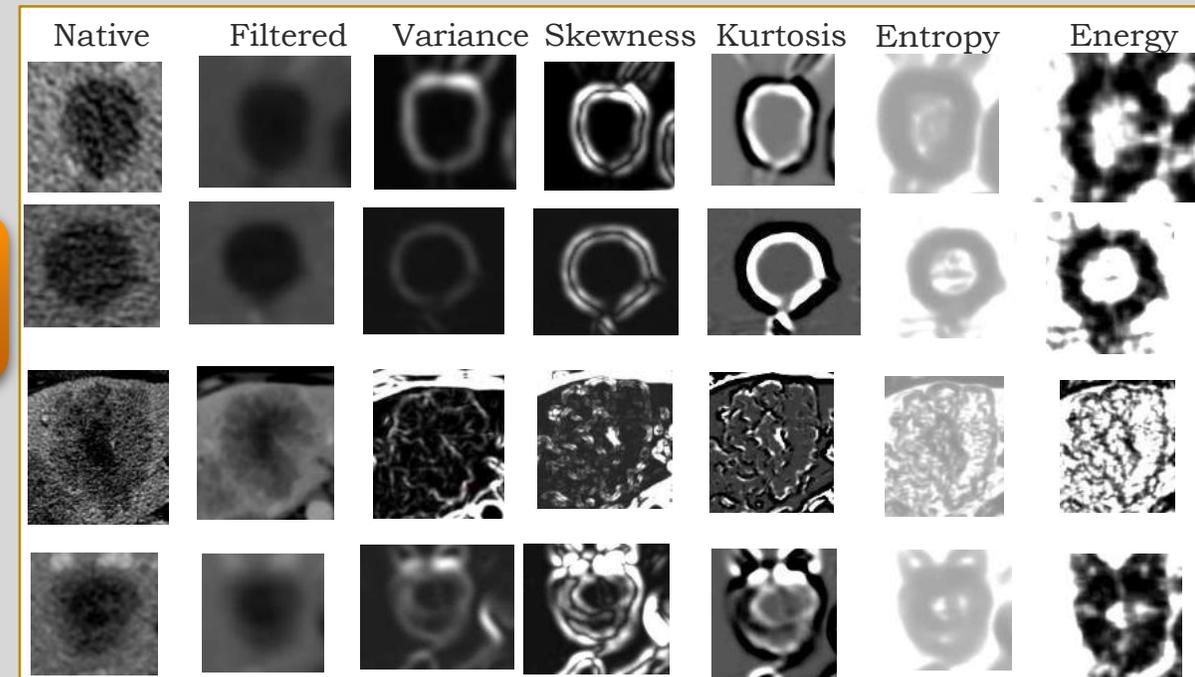


Hétérogénéité  
tissulaire

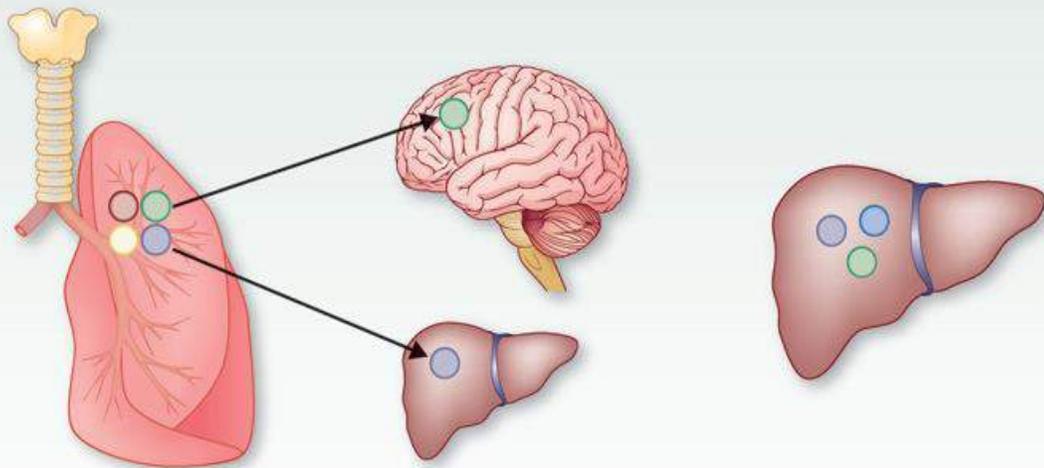
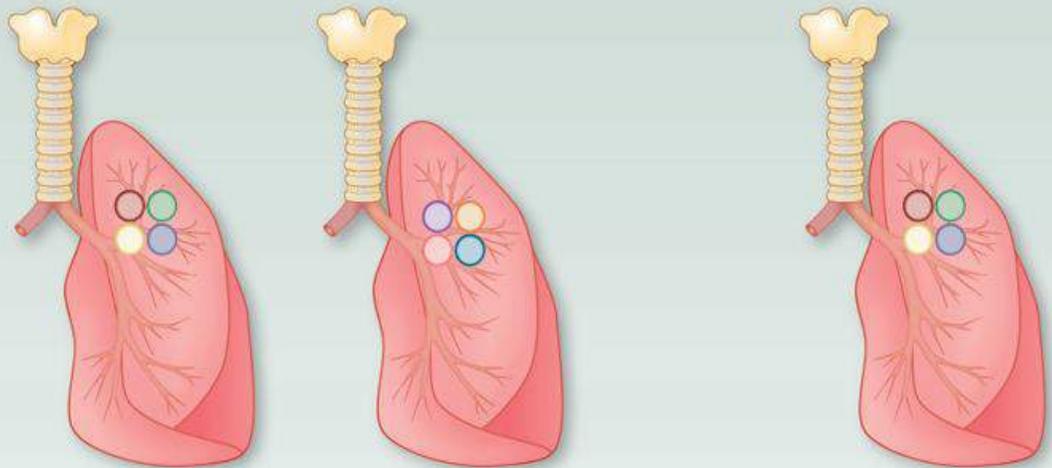
L'imagerie échantillonne  
l'ensemble de la tumeur

M Rozand, Master2

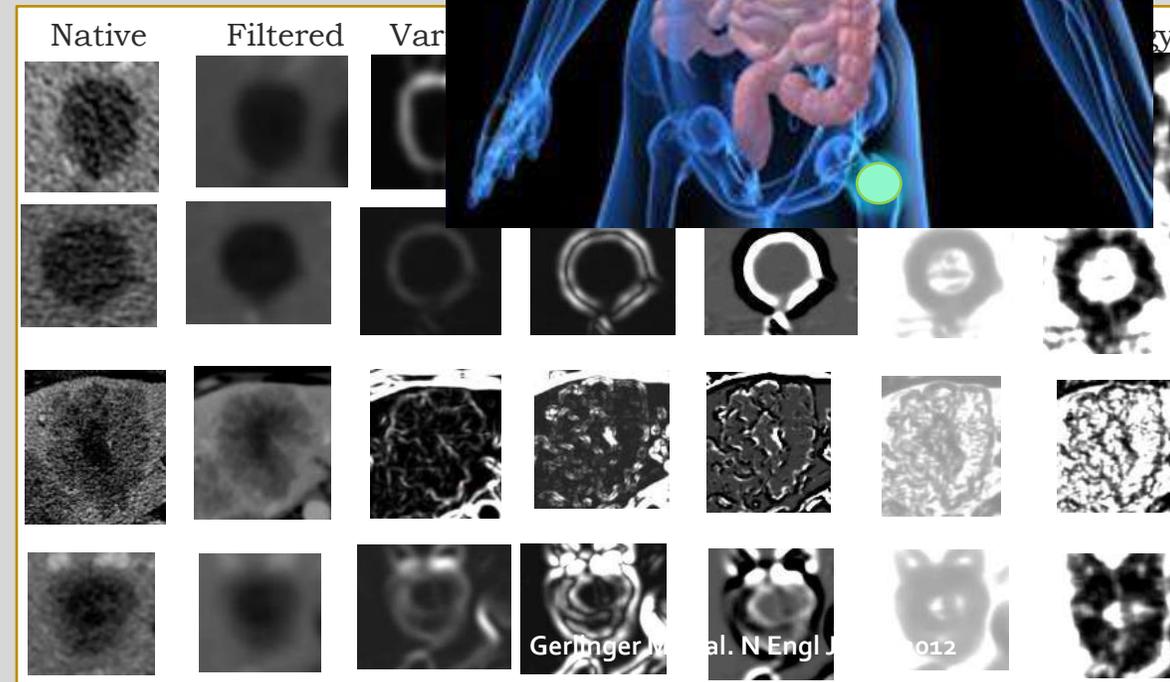
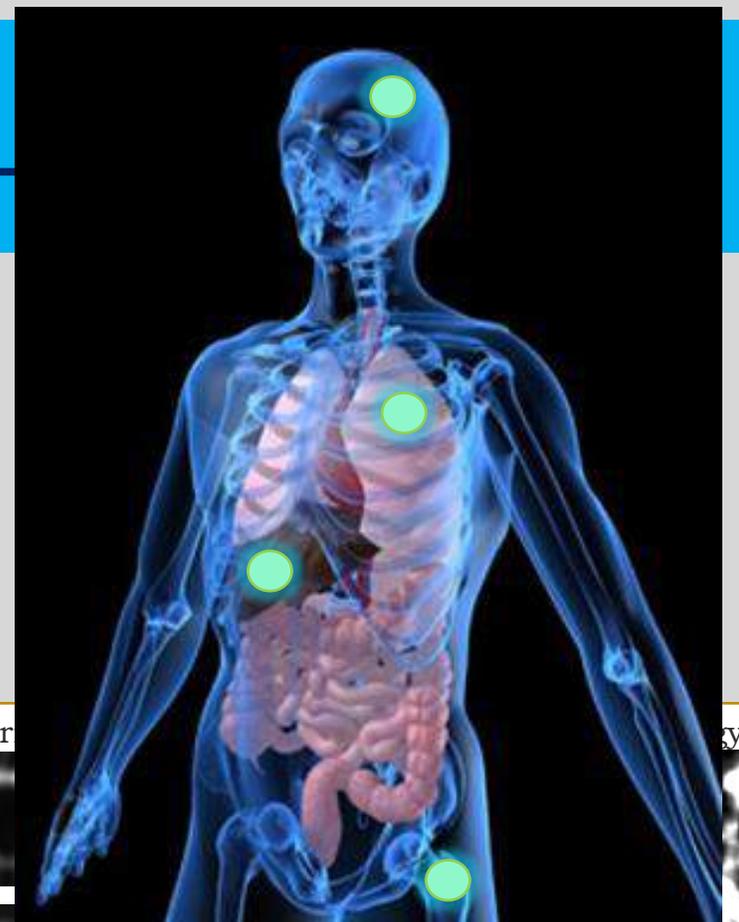
Hétérogénéité  
en imagerie



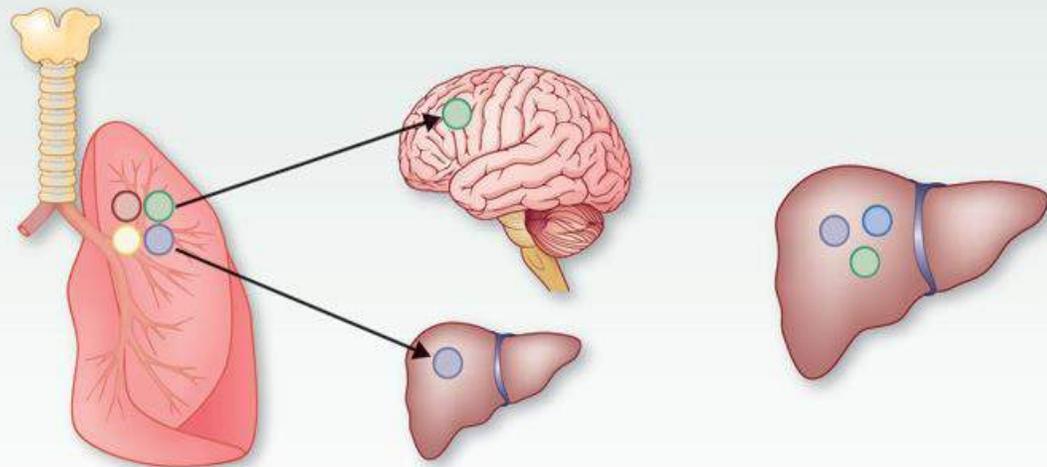
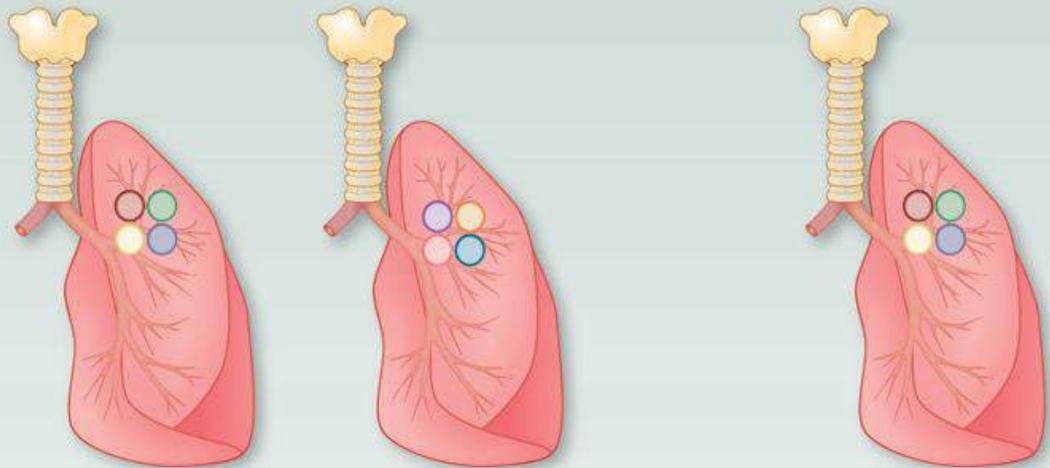
# HÉTÉROGÉNÉITÉ INTER-LESIONNELLE



Jamal-Hanjani et al. Clin Cancer Res 2015

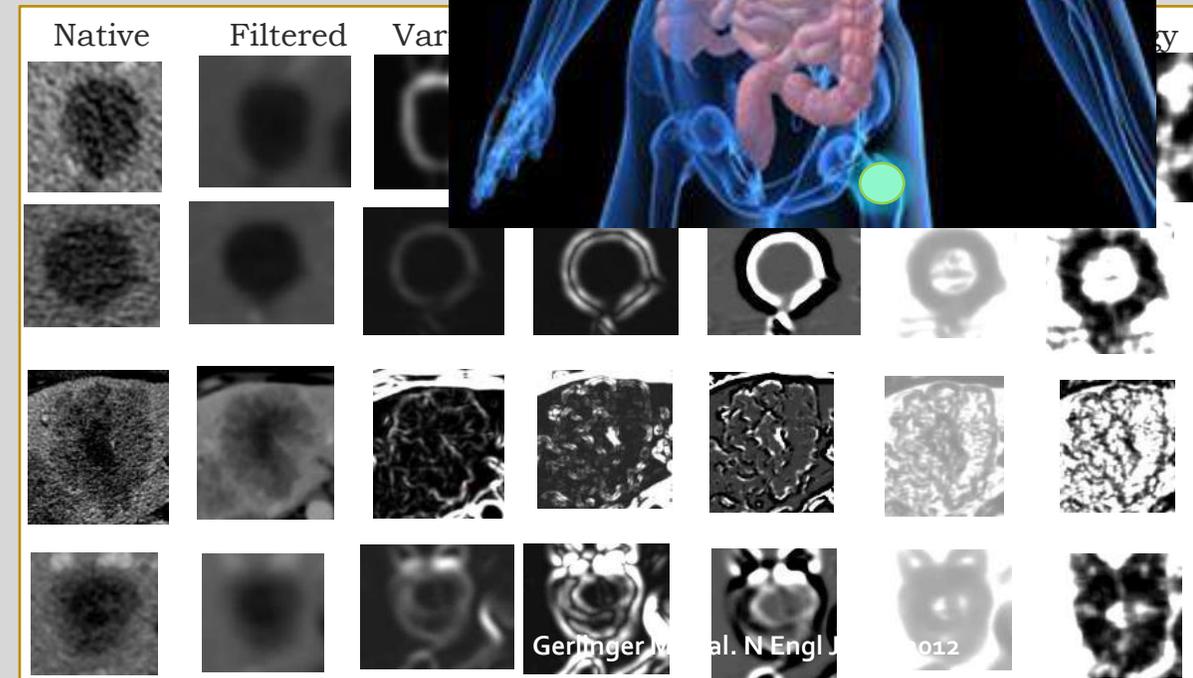


# HÉTÉROGÉNÉITÉ INTER-LESIONNELLE

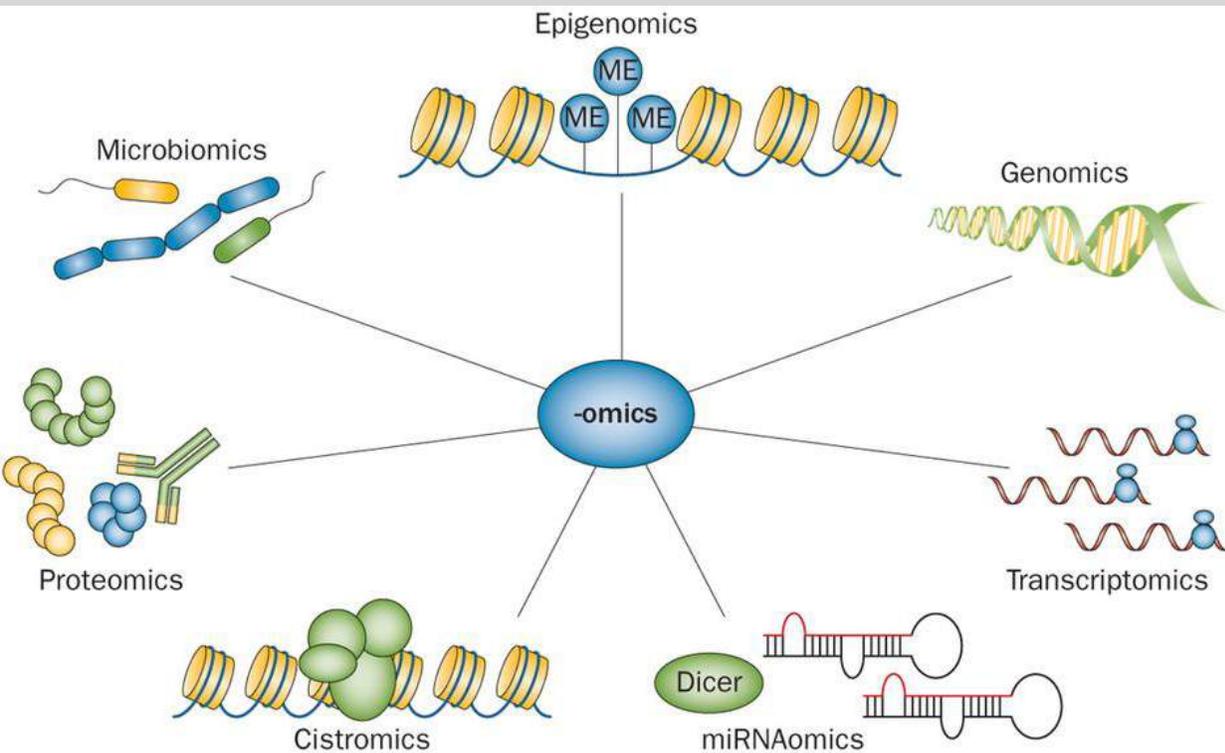


Jamal-Hanjani et al. Clin Cancer Res 2015

L'imagerie échantillonne toutes les tumeurs



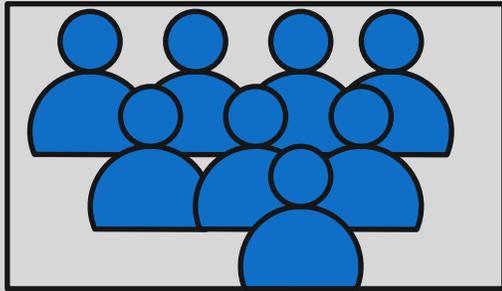
Gerlinger et al. N Engl J Med 2012



# OPTIMISER LE PROCESSUS DE DÉCOUVERTE

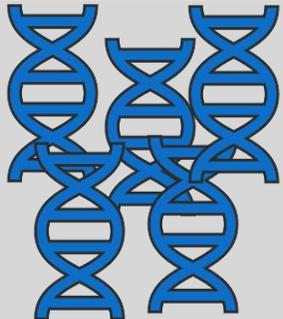
Weaver et al., Nat Rev Gastroenterol Hepatol 2014

# PRINCIPE DE LA GÉNOMIQUE

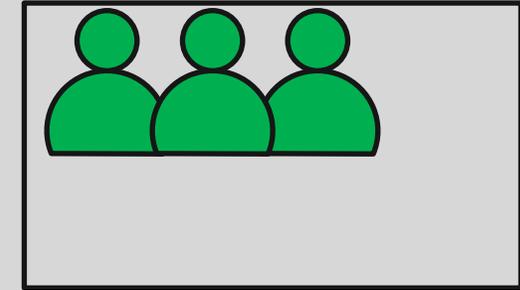
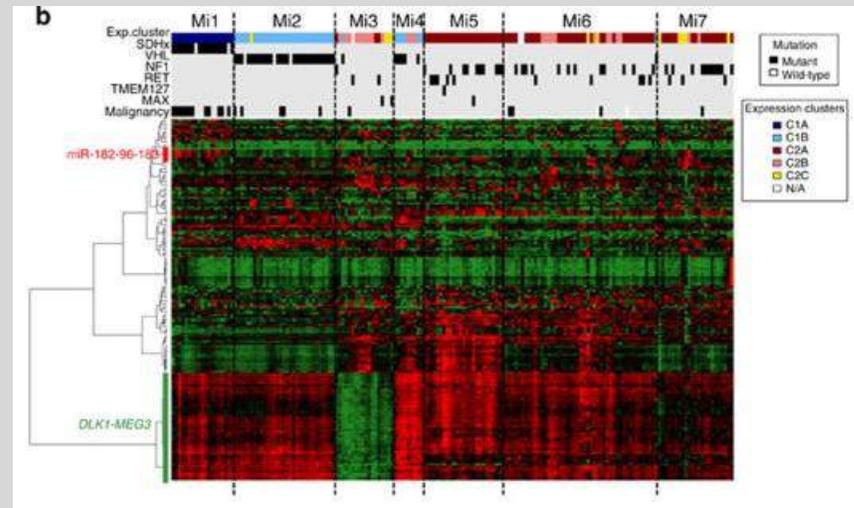


Population de patients

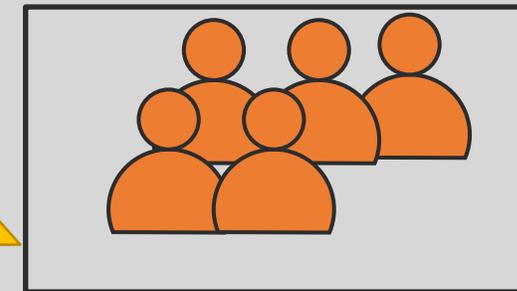
Séquençage complet du génome



*Castro-Vega, Nat Commun 2015*



Identifier des gènes qui différencient les patients en fonction de leur « devenir »



# APPLICATION À L'IMAGERIE

## Data Driven Research



Collecter un grand nombre de données



Qualification technique



Analyser & interpréter



Générer une hypothèse



**Nouveau biomarqueur potentiel**



Validation clinique



# DÉFINITION DE LA RADIOMIQUE

Extraction (+/- automatisée) d'un grand nombre de paramètres quantitatifs en imagerie sans hypothèse a priori



Corrélation au devenir

Pronostic

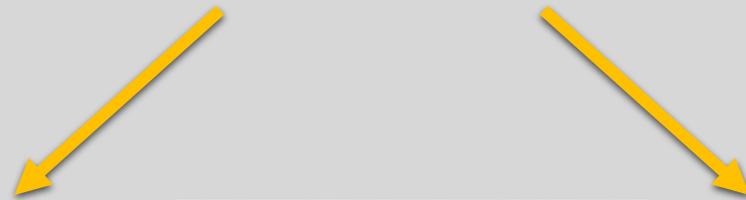
Réponse au traitement

«Radiogénomique»

Corrélation aux profils  
d'expression génique

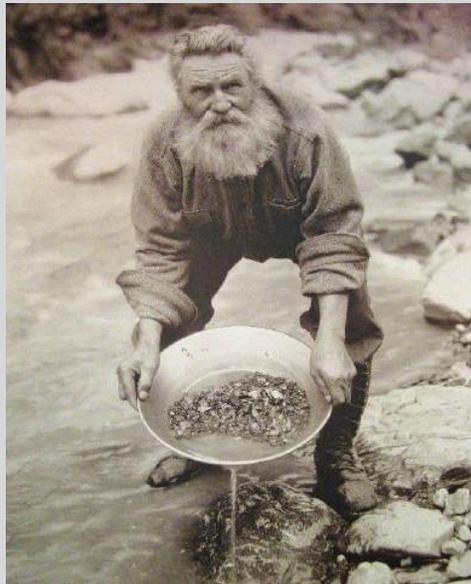
# DÉFINITION DE LA RADIOMIQUE

Extraction (+/- automatisée) d'un grand nombre de paramètres quantitatifs en imagerie sans hypothèse a priori



Corrélation au devenir

Recherche  
centrée sur la  
découverte



«Radiogénomique»

Corrélation aux profils  
d'expression génique

# RADIOMIQUE: PROCESSUS

1. Collecter les données (traitement d'image, **et/ou IA**)  
→ logiciel d'extraction d'un (très) grand nombre de descripteurs

À PARTIR D'IMAGES DE ROUTINE CLINIQUE

2. Nettoyer, uniformiser, enrichir les données  
→ "data curation"

SOUVENT L'ÉTAPE LIMITANTE

3. Traitement des données quantitatives  
→ méthodes bio-informatiques et bio-statistiques (**et/ou IA**)

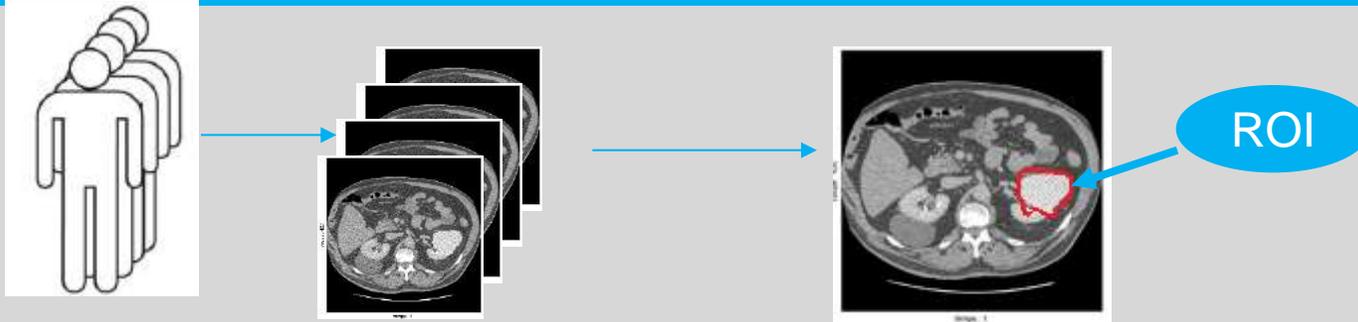
1. RÉDUCTION DU NOMBRE DE DESCRIPTEURS (basée sur leur qualité technique)
2. CORRÉLATION AU PARAMÈTRE D'INTÉRÊT CLINIQUE



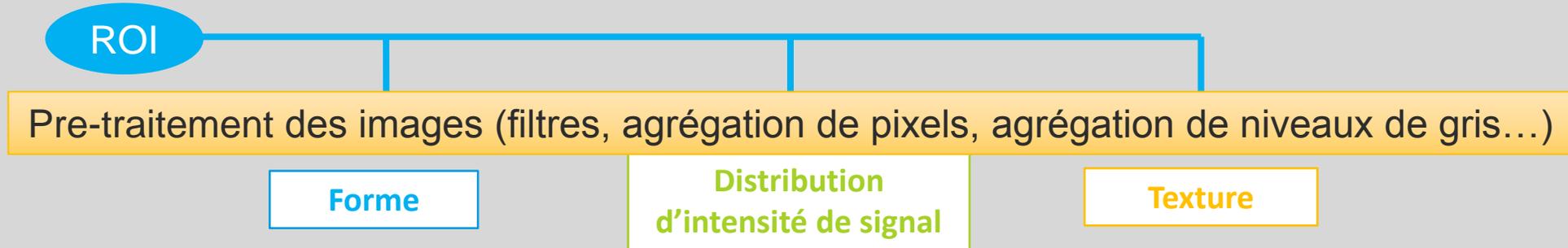
shutterstock.com - 793541692

# EXTRACTION DES DESCRIPTEURS

# SEGMENTATION (MANUELLE OU AUTOMATISÉE)



# EXTRACTION DES DESCRIPTEURS



# EXTRACTION DES DESCRIPTEURS

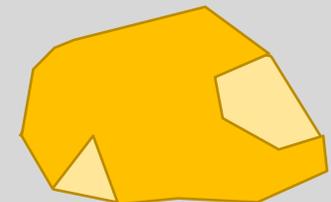
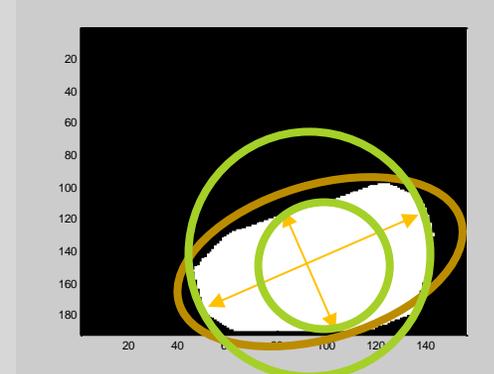
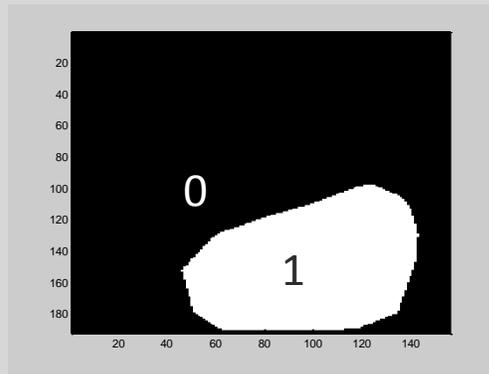
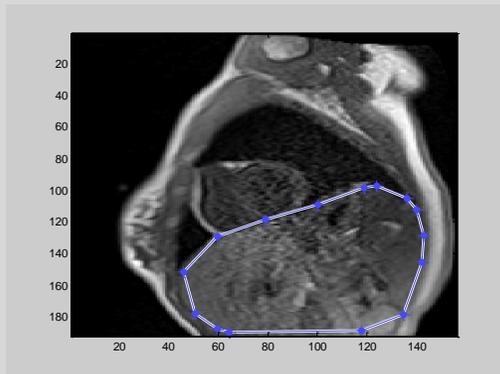
ROI

Pre-traitement des images (filtres, agrégation de pixels, agrégation de niveaux de gris...)

Forme



Analyse de contours  
Index de compacité  
Surface, volume, direction spatiale...

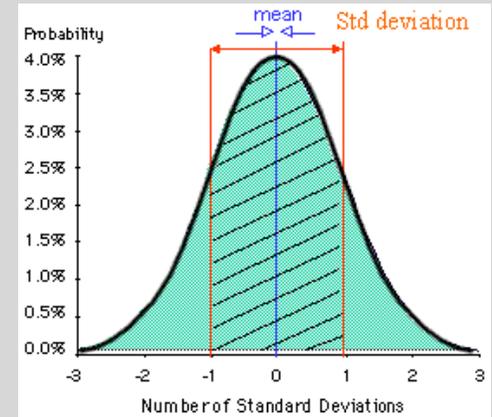


# EXTRACTION DES DESCRIPTEURS

ROI

Pre-traitement des images (filtres, agrégation de pixels, agrégation de niveaux de gris...)

Distribution  
d'intensité de signal



Histogramme de la distribution du signal

Moyenne → centre de la distribution

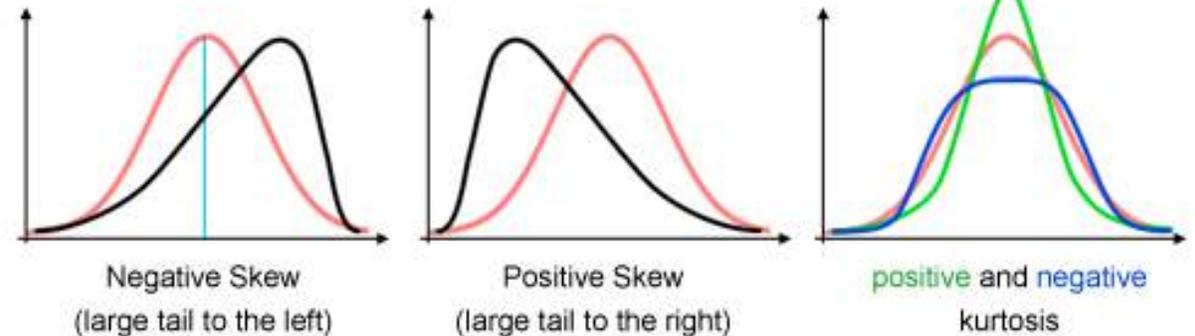
Déviati on standard ou Variance → echelle de la distribution

Skewness → symétrie de la distribution

Kurtosis → caractère pointu de la distribution

Energie, entropie...

© www.scratchapixel.com



# EXTRACTION DES DESCRIPTEURS

ROI

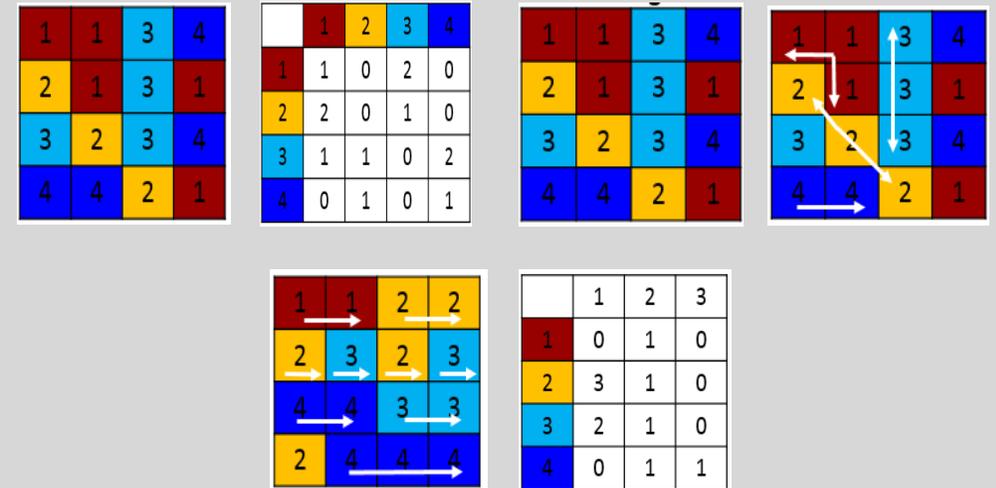
Pre-traitement des images (filtres, agrégation de pixels, agrégation de niveaux de gris...)

Texture

Matrices de distribution spatiale des pixels

**Gray level co-occurrence matrix** : changements de niveaux de gris par translation

**Gray level Run-Length matrix** : nombre de lignes de niveaux de gris équivalents de longueur ou direction variables



# EXTRACTION DES DESCRIPTEURS

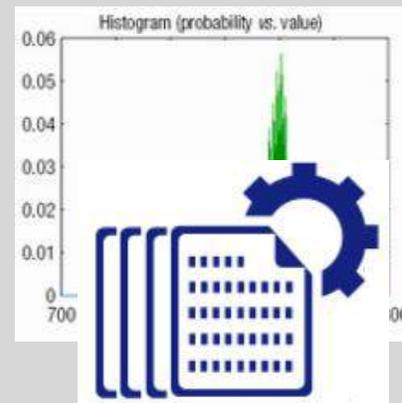
ROI

Pre-traitement des images (filtres, agrégation de pixels, agrégation de niveaux de gris...)

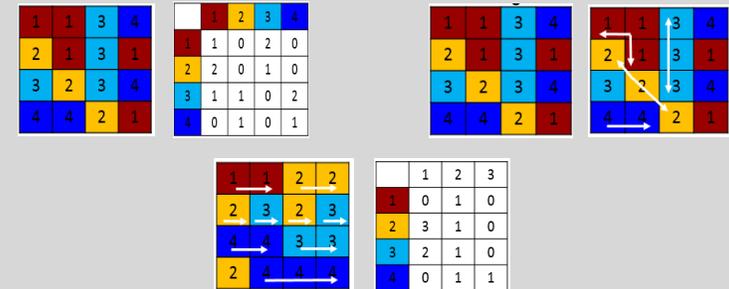
Forme



Distribution  
d'intensité de signal

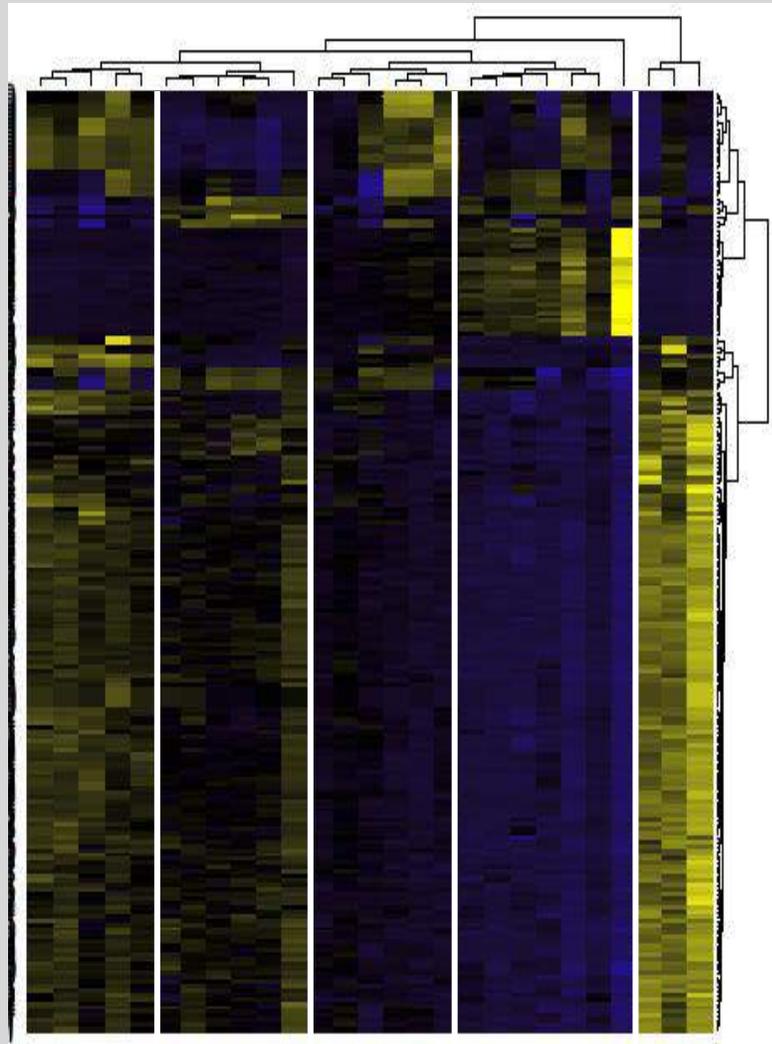


Texture



Grand nombre de  
descripteurs / patient

# DÉCOUVERTE DE PHENOTYPES EN IMAGERIE PAR CLUSTERING HIERARCHIQUE



Patients

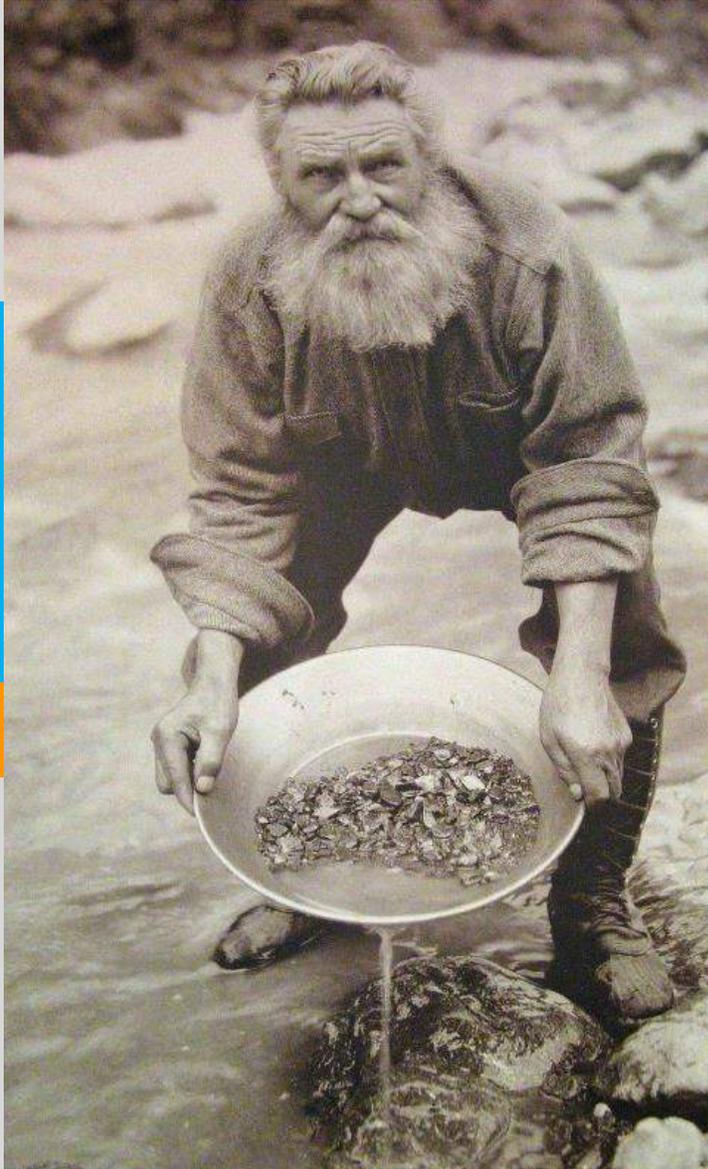
Descripteurs  
radiomiques

■ Values > médiane

■ Values < médiane

DEFINIR DES GROUPES DE  
PATIENTS AVEC DES  
PHENOTYPES COMMUNS

A Bouchouicha, PhD



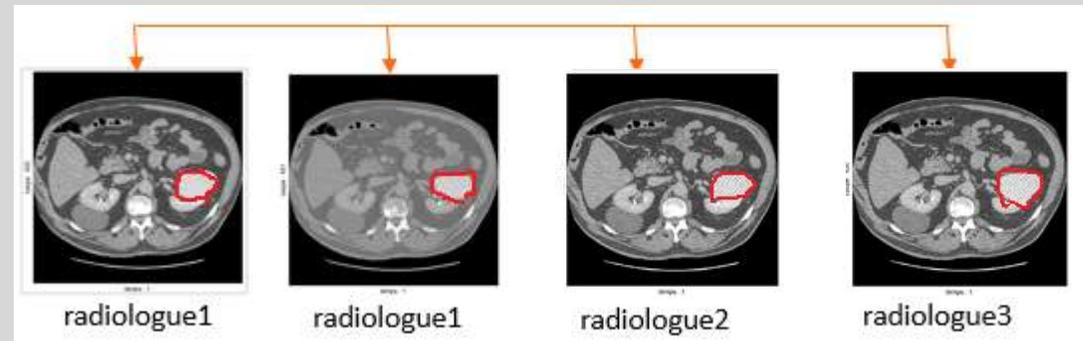
# RÉDUCTION DU NOMBRE DE DESCRIPTEURS

(exemple de stratégie)

# RÉDUCTION BASÉE SUR LA ROBUSTESSE



Objectif: éliminer  
les descripteurs  
non reproductibles



Reproductibilité intra- & inter-observateurs

**ICC**

(Intra-class Correlation Coefficient)

**CCC**

(Concordance Correlation Coefficient)

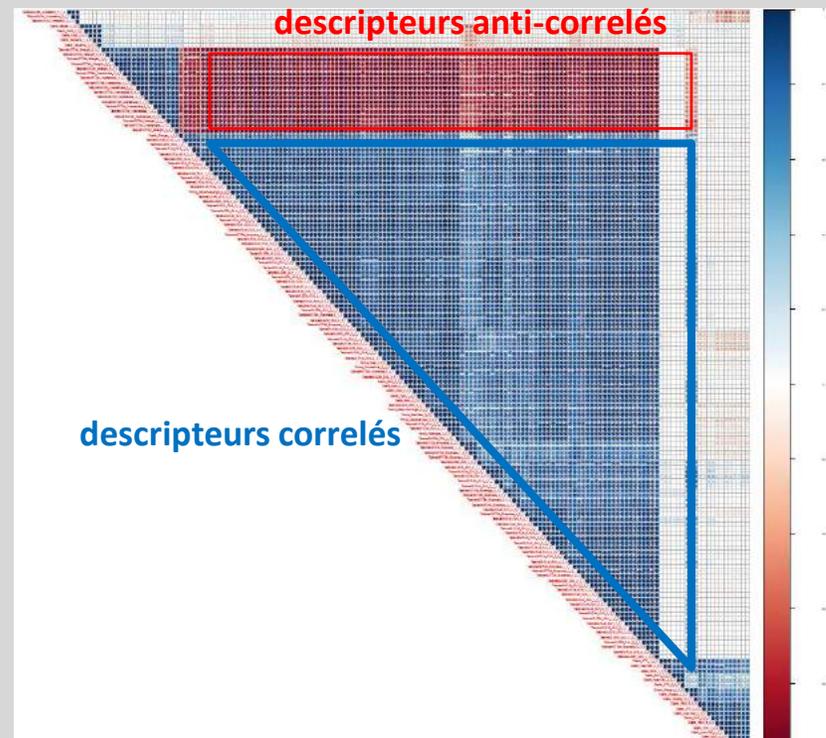
# RÉDUCTION BASÉE SUR L'INDÉPENDANCE

A Bouchouicha, PhD

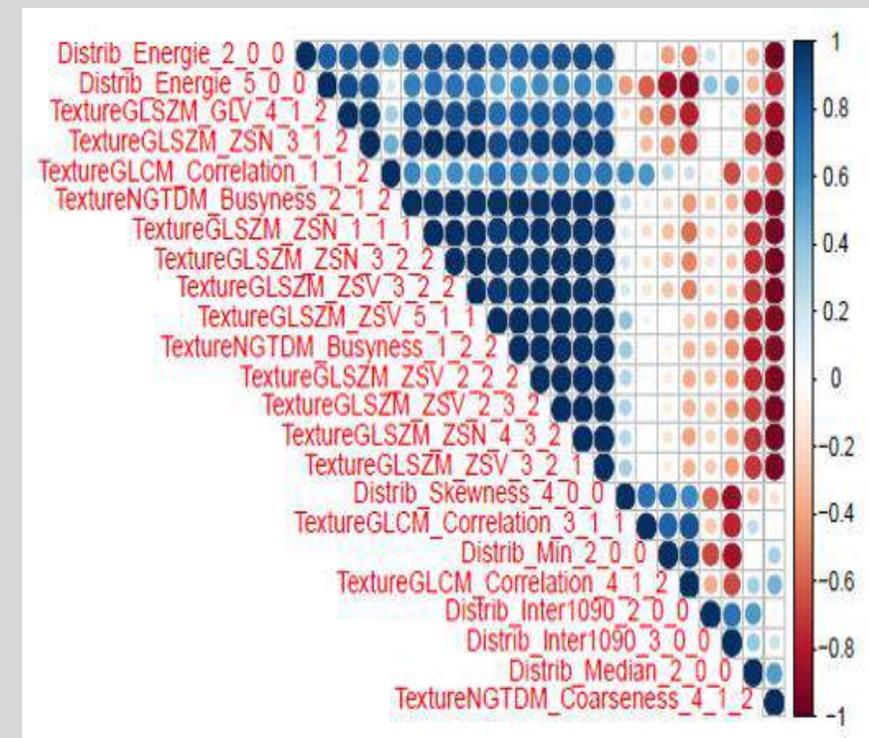
**Matrice de corrélation** : évaluer la redondance entre les descripteurs

Seuil de corrélation de Spearman= 0.9

Objectif: éliminer les descripteurs redondants (pas d'information supplémentaire)



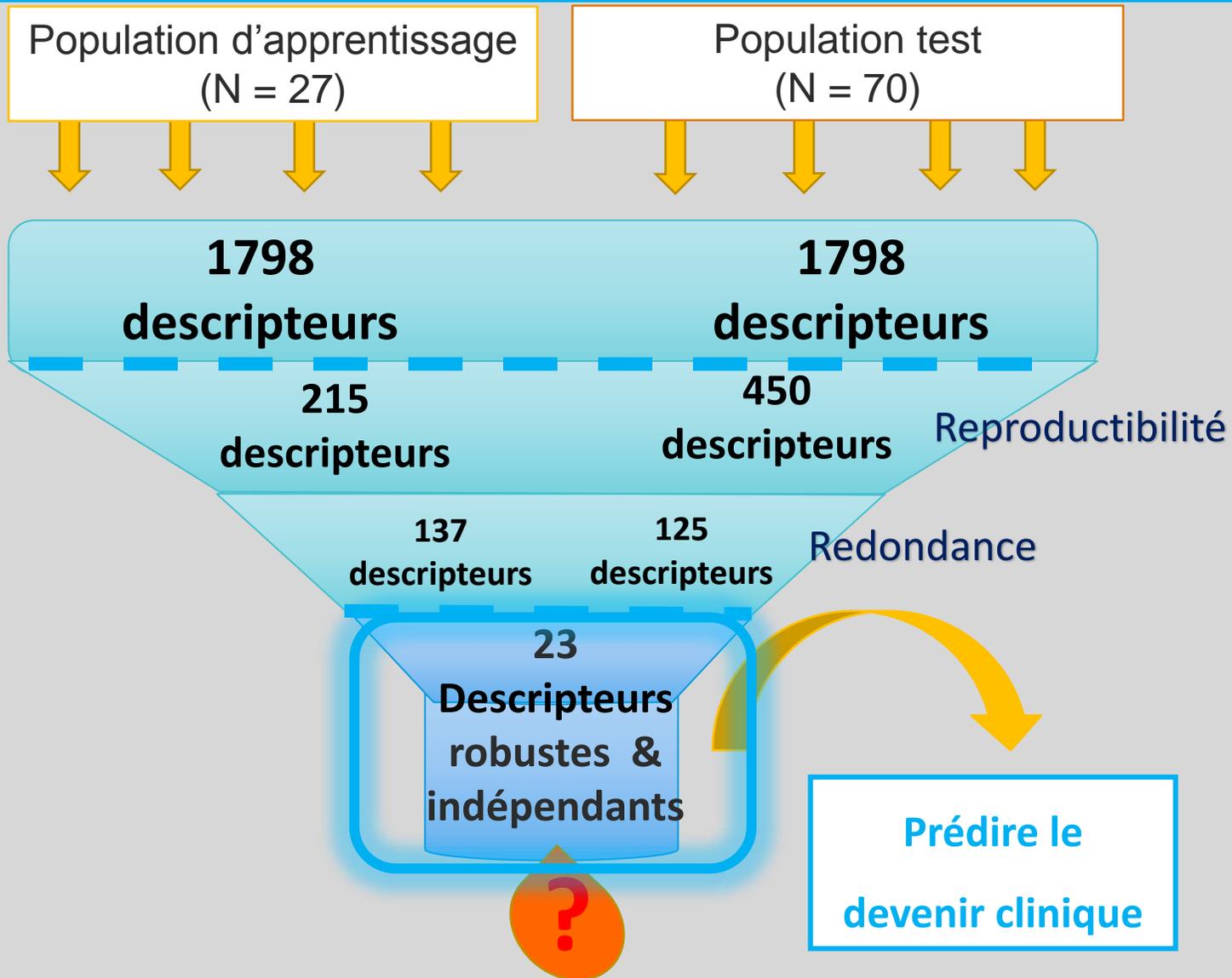
Corrélogramme des 158 descripteurs les plus robustes



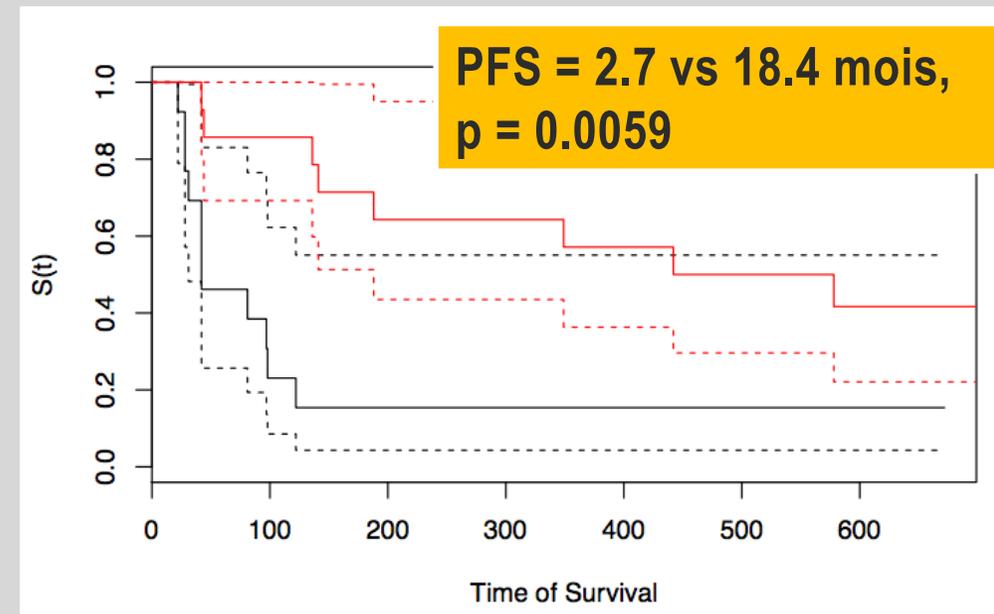
Corrélogramme des 23 descripteurs les plus robustes et informatifs

# RÉDUCTION DES DESCRIPTEURS

A Bouchouicha, PhD  
J Deidier, interne



Signature radiomique
Lengthening
Skewness
GLN (texture)



# APPLICATIONS

Deux exemples

# RADIOGÉNOMIQUE

nature  
biotechnology

Decoding global gene expression programs in liver cancer by noninvasive imaging

Eran Segal<sup>1</sup>, Claude B Sirlin<sup>2</sup>, Clara Ooi<sup>4</sup>, Adam S Adler<sup>5</sup>, Jeremy Gollub<sup>6</sup>, Xin Chen<sup>8</sup>, Bryan K Chan<sup>2</sup>, George R Matcuk<sup>7</sup>, Christopher T Barry<sup>3</sup>, Howard Y Chang<sup>5</sup> & Michael D Kuo<sup>2</sup>

Descripteurs quantitatifs  
ET qualitatifs  
(annotations sémantiques)

- Population d'apprentissage: 38 patients avec carcinoma hépatocellulaire (CHC)
  - 138 caractéristiques en imagerie sur des scanners
  - Analyse microarray de 6,732 genes
- Population test: 32 patients avec CHC

# RADIOGÉNOMIQUE : POPULATION D'APPRENTISSAGE

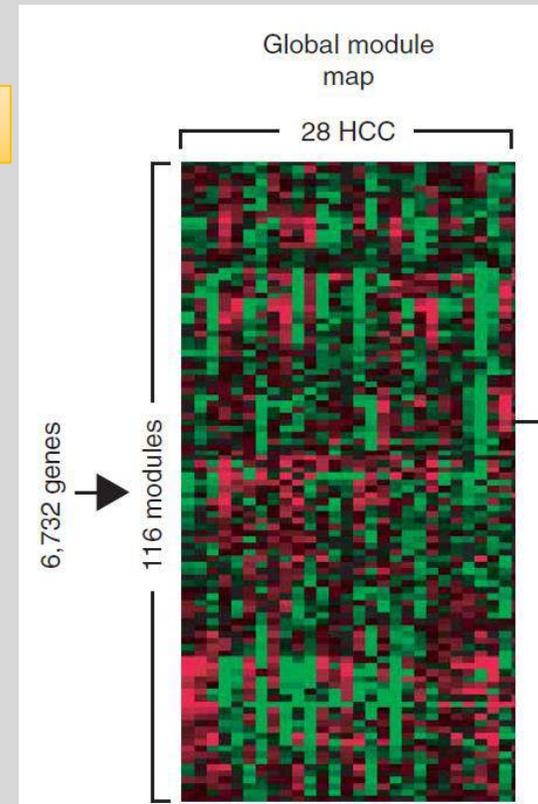
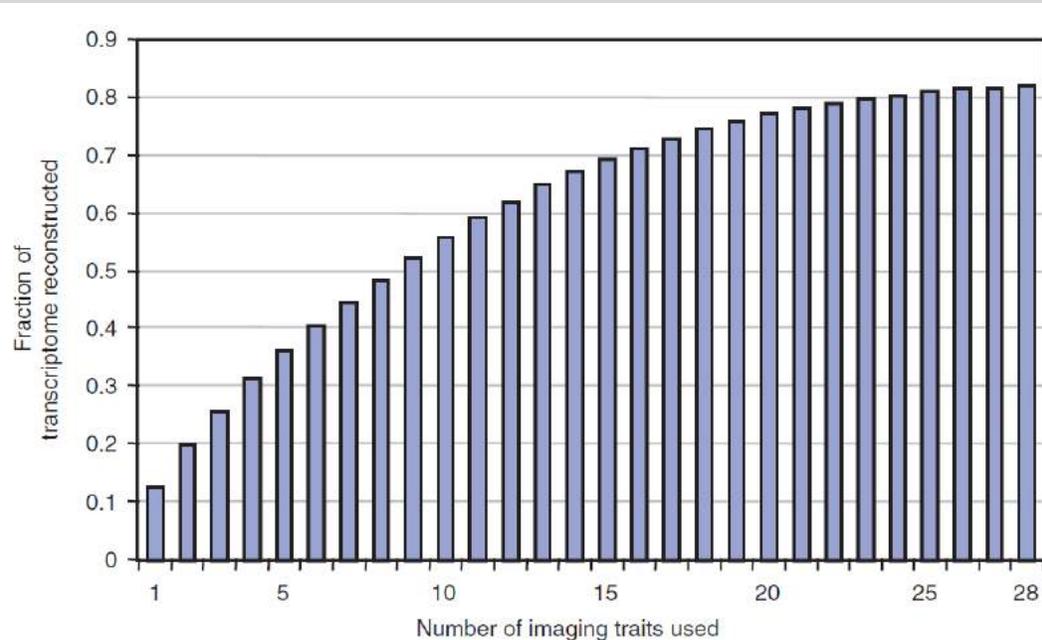
138 descripteurs

fréquence et prééminence,  
variabilité inter-observateurs,  
indépendance des autres  
paramètres

32 descripteurs

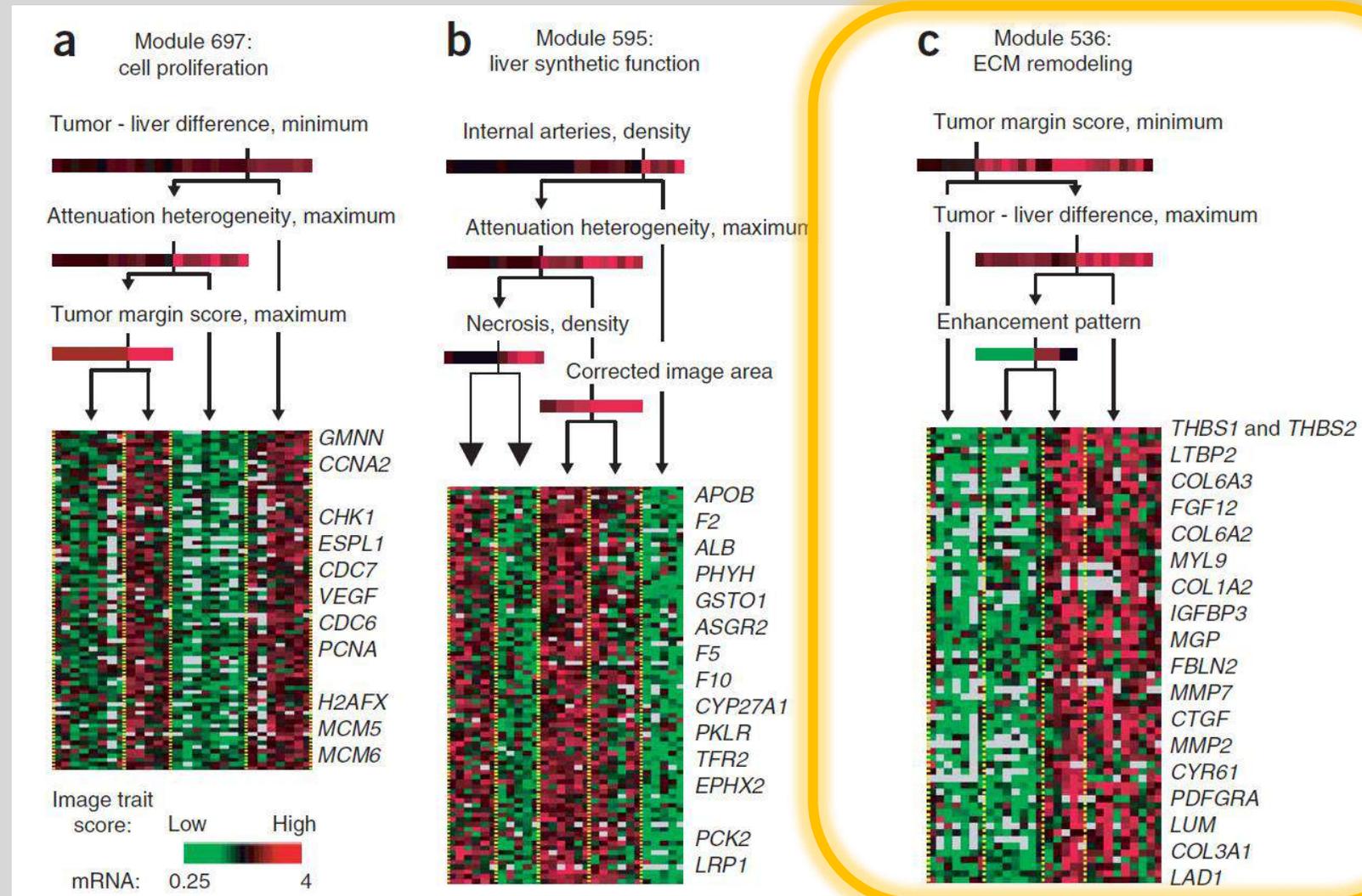
6,732 niveaux d'expression génique  
(microarray)

116 modules géniques

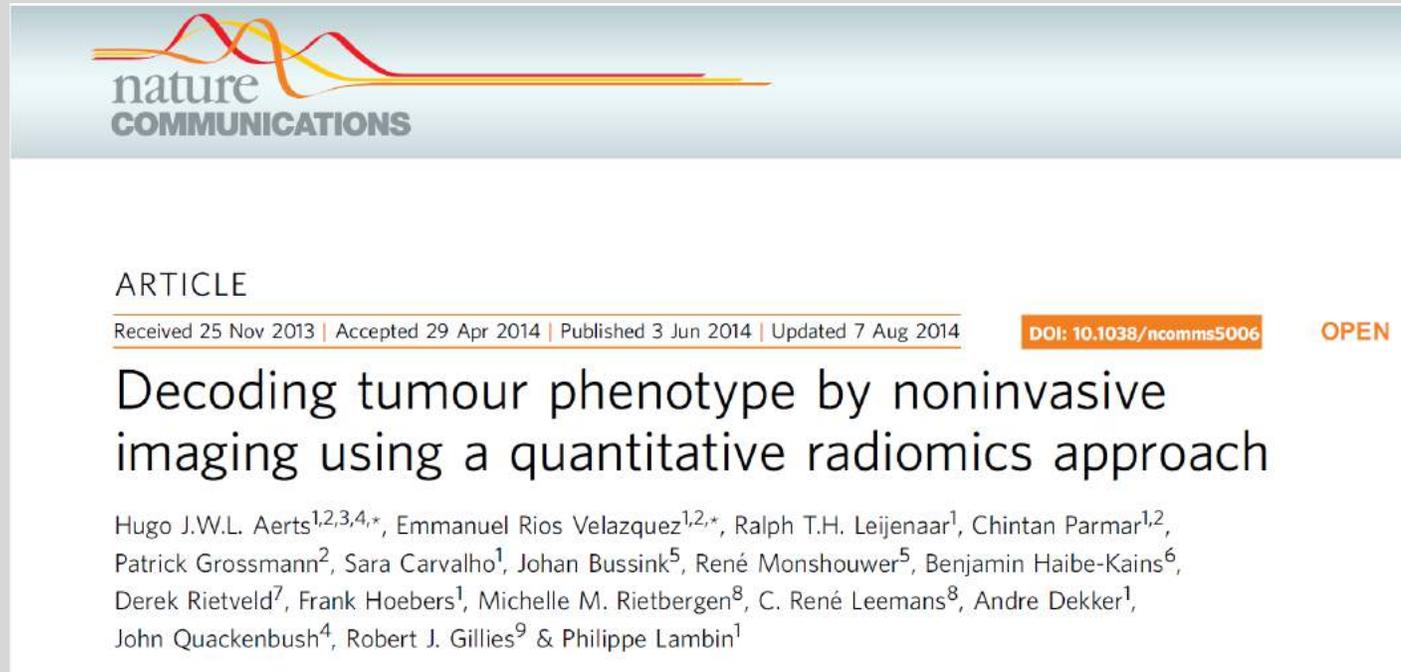


- Necrosis, percent
- Internal septa
- Texture heterogeneity, portal-venous phase
- Tumor margin score, minimum
- Enhancement pattern
- Internal arteries, rank
- Hypodense halo
- Wash-out, maximum
- Internal arteries, density
- Tumor - liver difference, maximum
- Corrected imaging area
- Necrosis, density
- Tumor margin score, maximum
- Texture heterogeneity, arterial phase
- Internal arteries, necrosis edge
- Capsule
- Wash-in, maximum
- Infiltration, percent
- Tumor - liver difference, minimum
- Attenuation - heterogeneity score, maximum

# RADIOGÉNOMIQUE: LIEN AVEC LA BIOLOGIE

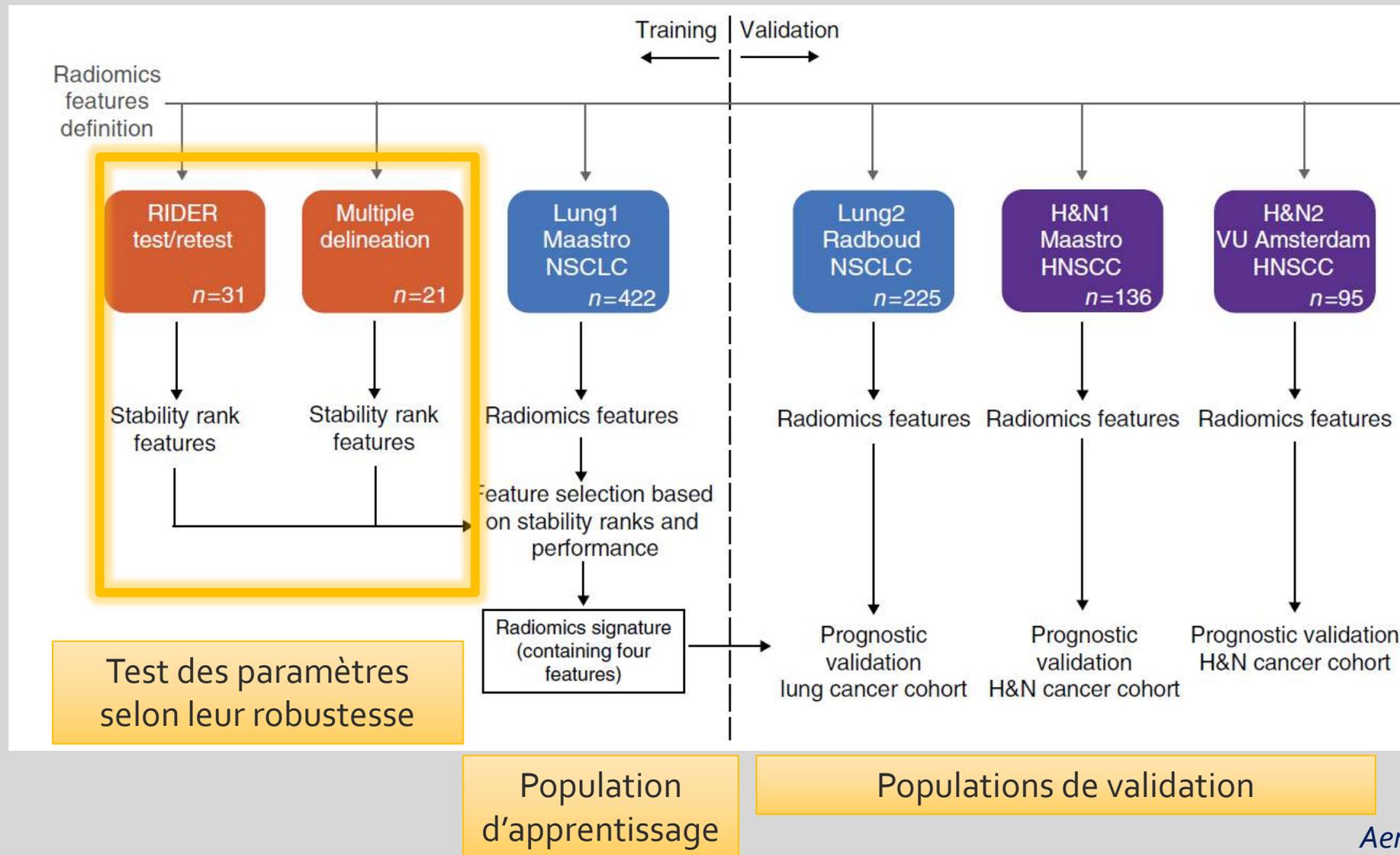


# RADIOMIQUE: PREDICTION DE LA SURVIE



- ❑ 7 sets de données : cancers poumon et ORL = 1,019 patients
- ❑ 440 caractéristiques de radiomique
- ❑ Corrélation à la survie et aux profils d'expression génique

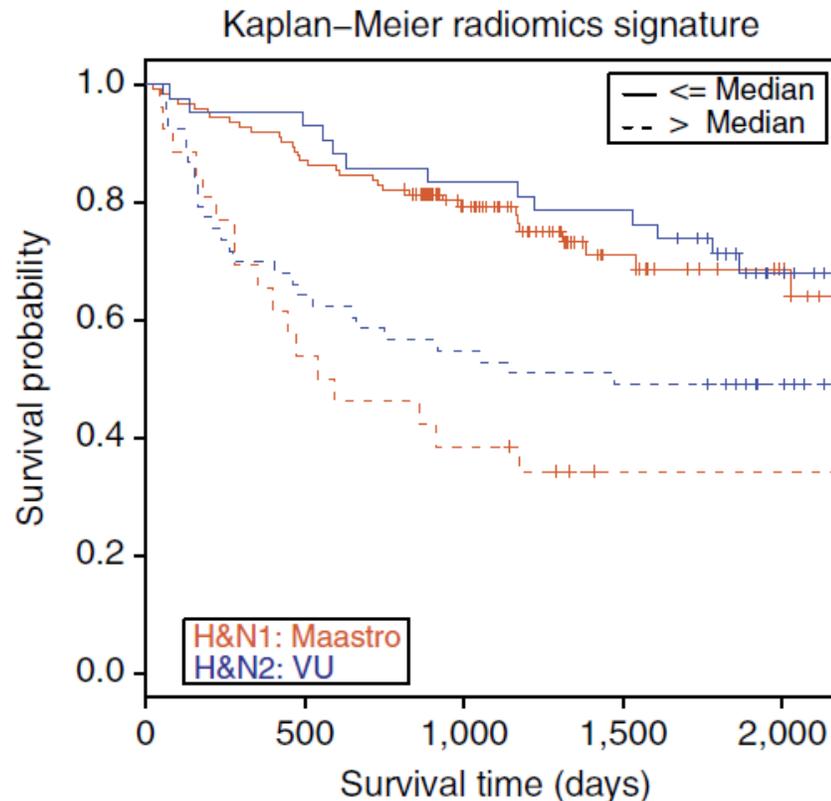
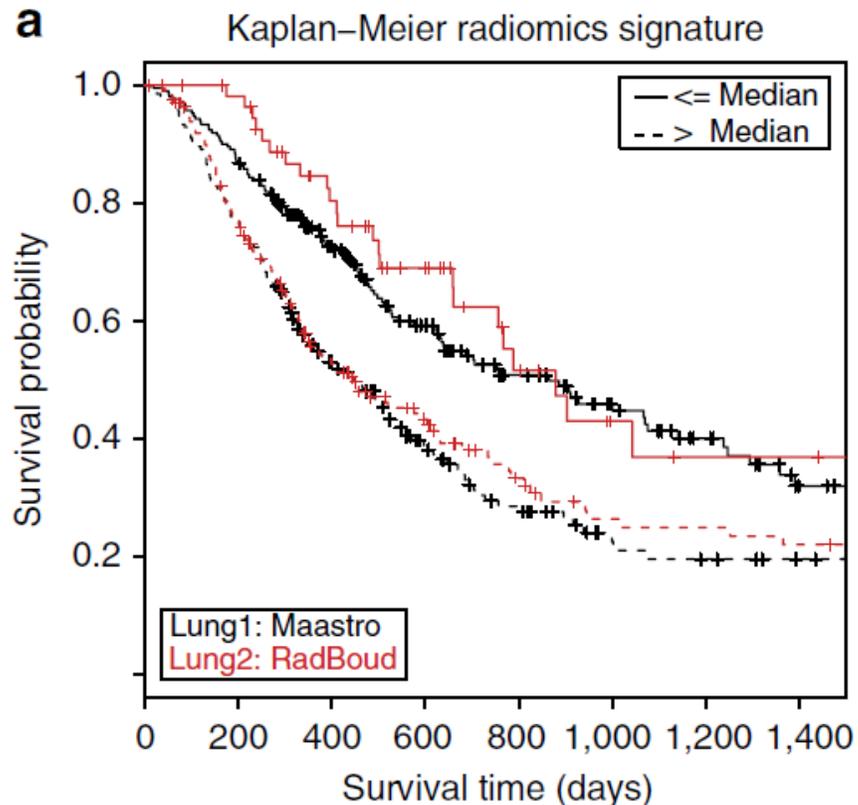
# RADIOMIQUE: PREDICTION DE LA SURVIE



# RADIOMICS: ANALYSIS WORKFLOW

Signature radiomique:

- (I) 'Statistics Energy
- (II) 'Shape Compactness'
- (III) 'Grey Level Nonuniformity'
- (IV) wavelet 'Grey Level Nonuniformity HLH'



- Les mêmes quatre paramètres prédisent la survie:
  - Localisations différentes
  - Traitement différents (chimo vs radiothérapie)
- Indépendants de TNM, volume tumoral
- Associés à un pattern d'expression génique

# PUBLICATIONS...

42 publications en mars 2016  
320 publications en mars 2018

- [Identifying relations between \*\*imaging\*\* phenotypes and molecular subtypes of breast cancer: Model discovery and external validation.](#)
    1. Wu J, Sun X, Wang J, Cui Y, Kato F, Shirato H, Ikeda DM, Li R. J Magn Reson **Imaging**. 2017 Feb 8. doi: 10.1002/jmri.25661. [Epub ahead of print] PMID: 28177554
  - [Radiomic analysis reveals DCE-MRI features for prediction of molecular subtypes of breast cancer.](#)
    2. Fan M, Li H, Wang S, Zheng B, Zhang J, Li L. PLoS One. 2017 Feb 6;12(2):e0171683. doi: 10.1371/journal.pone.0171683. PMID: 28166261
  - [Radiomics assessment of bladder cancer grade using \*\*texture\*\* features from diffusion-weighted \*\*imaging\*\*.](#)
    3. Zhang X, Xu X, Tian Q, Li B, Wu Y, Yang Z, Liang Z, Liu Y, Cui G, Lu H. J Magn Reson **Imaging**. 2017 Feb 15. doi: 10.1002/jmri.25669. [Epub ahead of print] PMID: 28199039
  - [Correlation between tumour characteristics, SUV measurements, metabolic tumour volume, TLG and textural features assessed with <sup>18</sup>F-FDG PET in a large cohort of oestrogen receptor-positive breast cancer patients.](#)
    4. Lemarignier C, Martineau A, Teixeira L, Vercellino L, Espié M, Merlet P, Groheux D. Eur J Nucl Med Mol **Imaging**. 2017 Feb 10. doi: 10.1007/s00259-017-3641-4. [Epub ahead of print] PMID: 28188325
- 
- [Non-invasive Identification of Vulnerable Atherosclerotic Plaques Using \*\*Texture\*\* Analysis in Ultrasound Carotid Elastography: An In Vivo Feasibility Study Validated by Magnetic Resonance \*\*Imaging\*\*.](#)
    5. Huang C, He Q, Huang M, Huang L, Zhao X, Yuan C, Luo J. Ultrasound Med Biol. 2017 Jan 30. pii: S0301-5629(16)30420-3. doi: 10.1016/j.ultrasmedbio.2016.12.003. [Epub ahead of print] PMID: 28153351
  - [Tissue classification for laparoscopic image understanding based on multispectral \*\*texture\*\* analysis.](#)
    6. Zhang Y, Wirkert SJ, Iszatt J, Kenngott H, Wagner M, Mayer B, Stock C, Clancy NT, Elson DS, Maier-Hein L. J Med **Imaging** (Bellingham). 2017 Jan;4(1):015001. doi: 10.1117/1.JMI.4.1.015001. PMID: 28149926



# CHALLENGES

- Imputable
- Reliable
- Meaningful
- Actionable

# ENJEU: COMPRENDRE (IMPUTABILITÉ)

## Hypothesis Driven Research



Les lumières

OU

## Data Driven Research



Le côté obscur

Que veulent  
dire les  
paramètres?

GLRLM-RLN

# ENJEU: COMPRENDRE (IMPUTABILITÉ)

## Hypothesis Driven Research



Les lumières

OU

## Data Driven Research

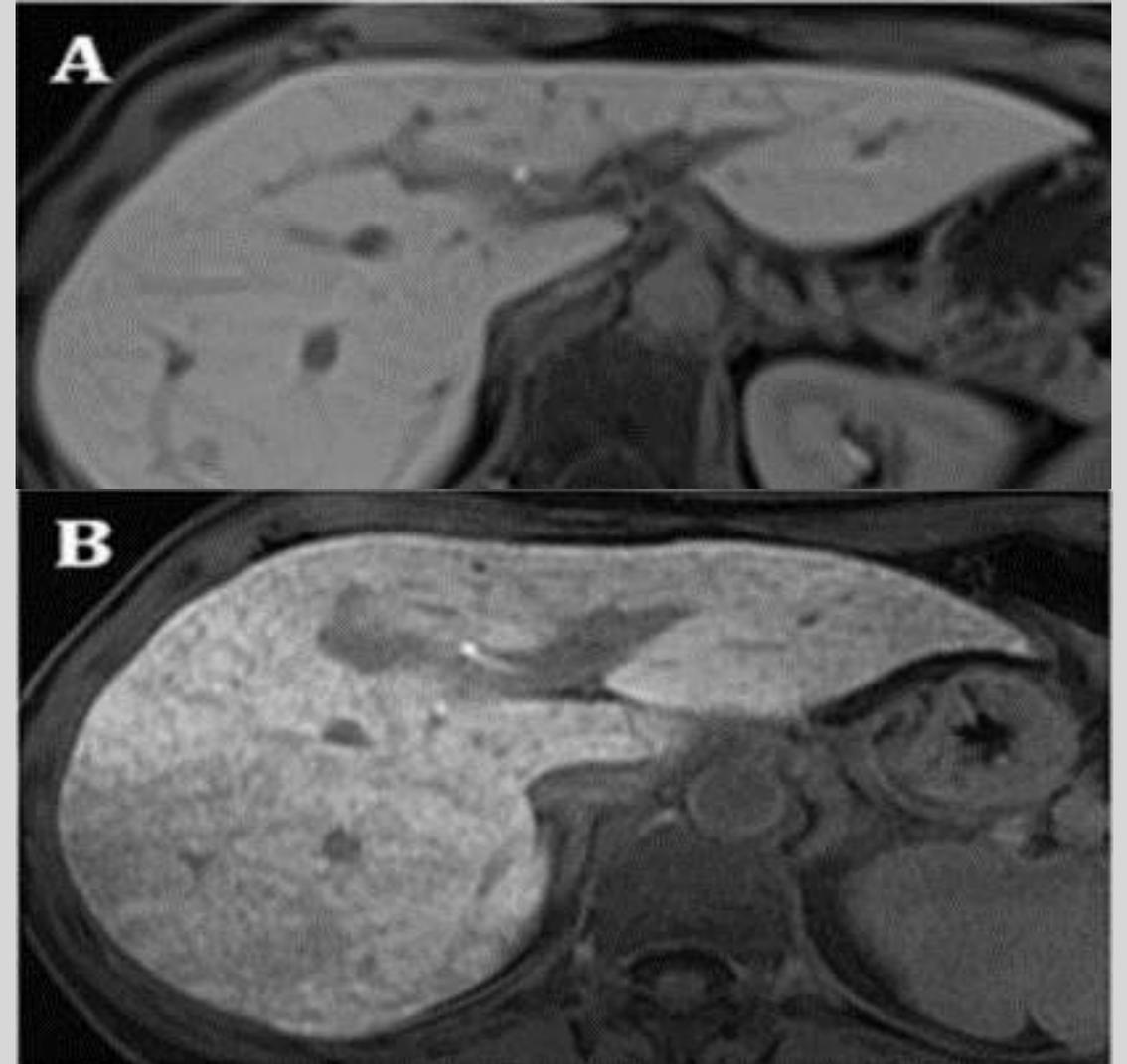


# ENJEU: COMPRENDRE (IMPUTABILITÉ)

- Détection d'une **différence physiologique** ou d'une **différence technique**?  
(Utiliser des cohortes avec une variété de machines / de protocoles d'acquisition / d'institutions...)

→ « **Sens radiologique** »

Paramètres de reconstruction différents  
→ paramètres de texture différents?



# ENJEU: COMPRENDRE (IMPUTABILITÉ)

- ❑ Résultats dépendants de la **population d'apprentissage**  
→ généralisation?

Donald

+

Houppes



# ENJEU: COHÉRENCE (CRÉDIBILITÉ)

- ❑ Pas de **standardisation** (difficile de reproduire ou comparer les résultats)

## **The image biomarker standardisation initiative**

The image biomarker standardisation initiative (IBSI) is an independent international collaboration which works towards standardisation of image biomarkers. Reproducibility and validation of studies in quantitative image analysis and radiomics is a major challenge for the field (Gillies et al., 2015; Yip and Aerts, 2016; Hatt et al., 2017). The IBSI therefore aims to provide a common nomenclature for image biomarkers, common biomarkers definitions, benchmarks for image processing and feature extraction, as well as reporting guidelines.

### **Copyright**

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

# ENJEU: VALIDATION

□ Processus de validation reste nécessaire

**Data Driven Research**



Collecter un grand nombre de données



Qualification technique



Analyser & interpréter



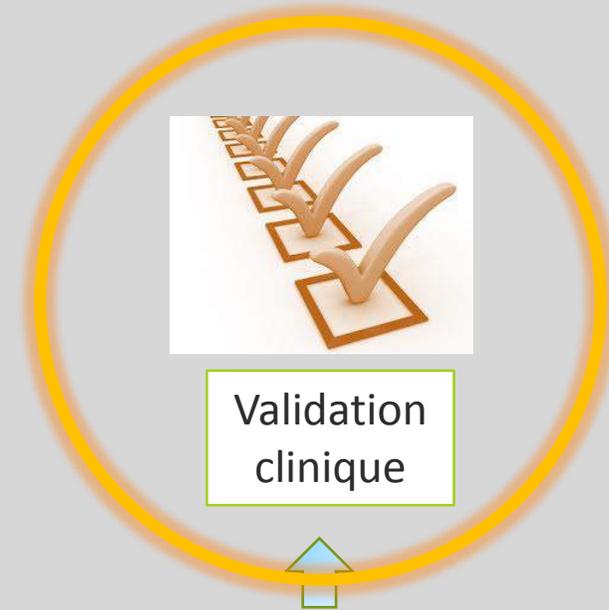
Générer une hypothèse



**Nouveau biomarqueur potentiel**



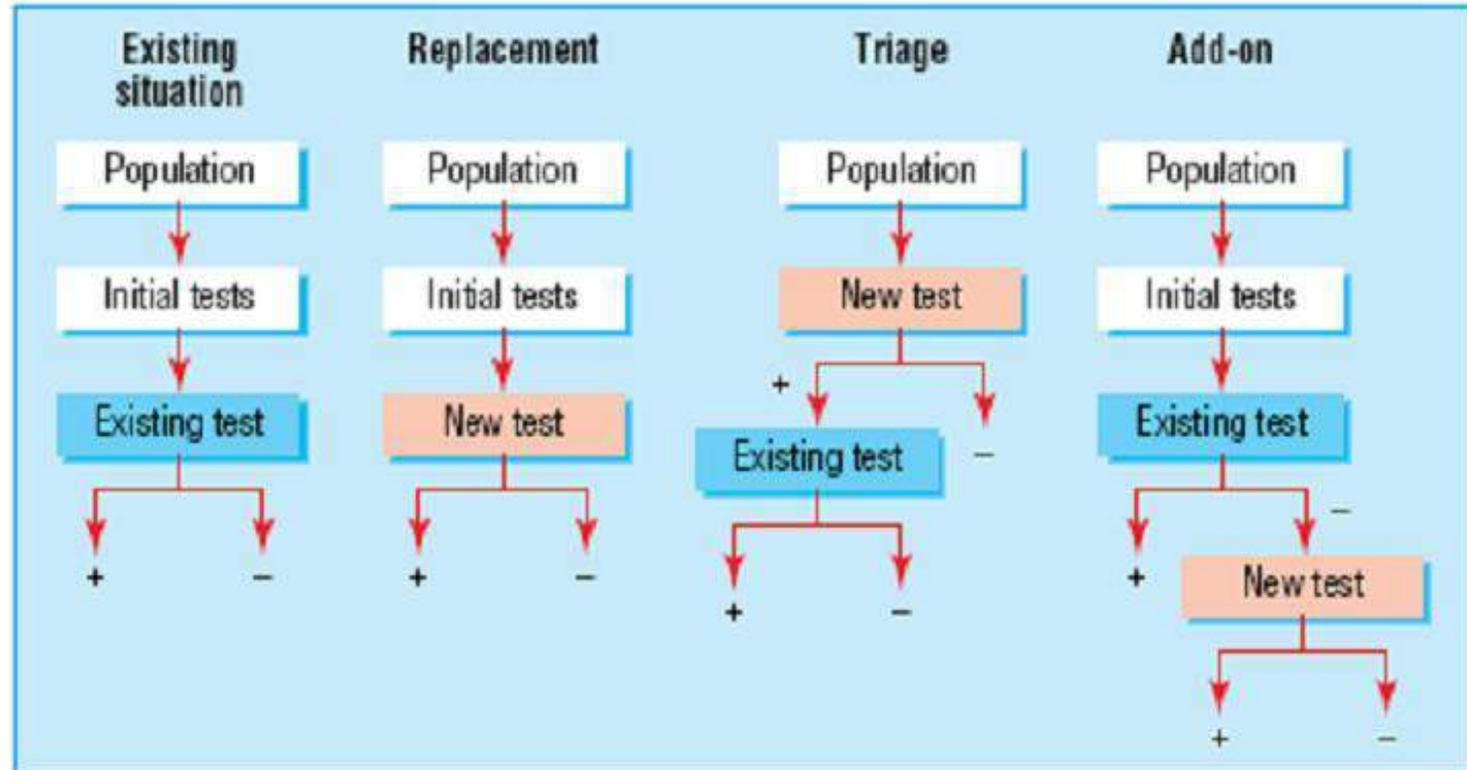
Validation clinique



# ENJEU: INTÉRÊT ?

## □ « Sens médical »

→ où se situe ce paramètre par rapport aux autres tests diagnostiques?



Roles of tests and positions in existing diagnostic pathways

# ENJEU: INTÉRÊT ?

## □ « Sens médical »

→ où se situe ce paramètre par rapport aux autres tests diagnostiques?

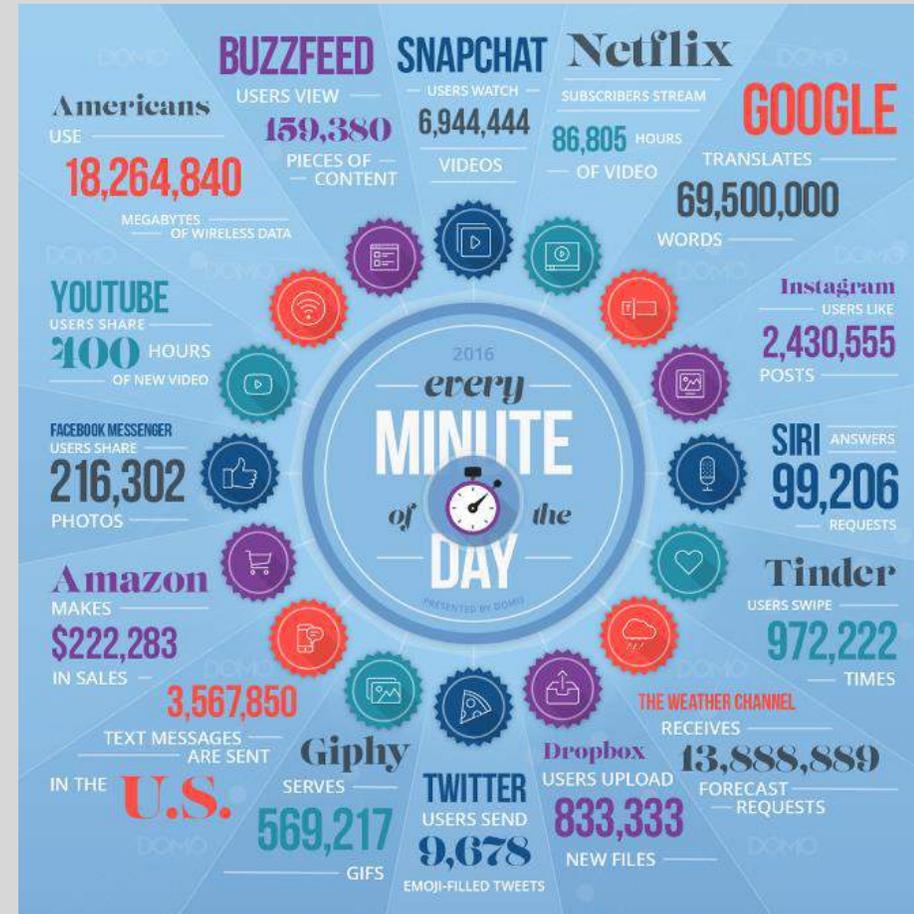
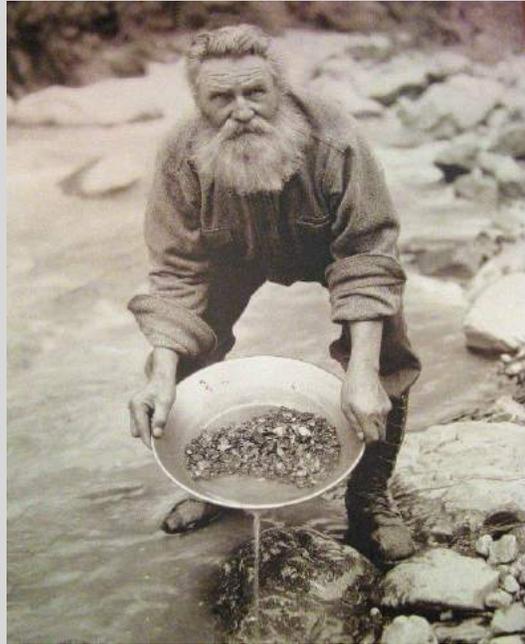
## □ « Actionabilité »: le paramètre que l'on prédit modifiera-t-il la prise en charge?

→ par rapport aux options thérapeutiques (ex: prédire la réponse à un traitement pour lequel il n'y a pas d'alternative)

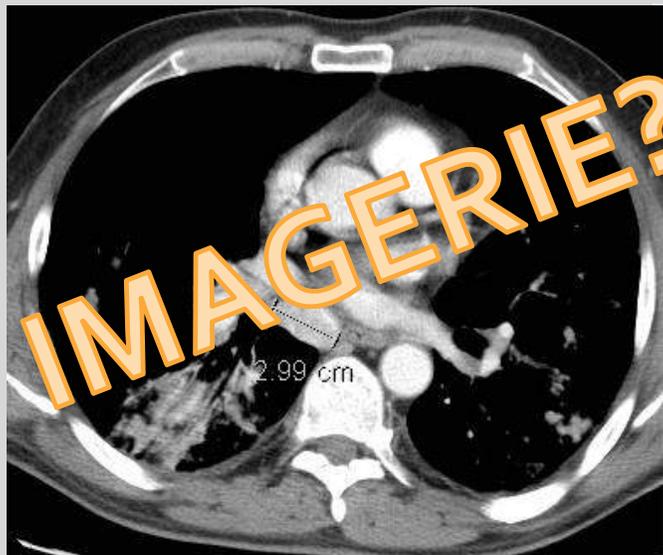
→ par rapport au bénéfice attendu

(Génomique < 1% découvertes ont mené à une mutation entraînant une modification de prise en charge)

# ENJEU: LA RUÉE VERS LES DONNÉES



# ENJEU: OÙ TROUVER LES DONNÉES?



Type	# instances	# patients
Demographics (age, sex, Hospital vital status)		742 487
Vital signs (temperature, blood pressure, weight, ...)	14 million	141 725
Diagnoses (ICD 10 codes)	6 million	338 871
Medical Procedures (French CCAM codes)	4 million	285 097
Clinical data (EHR DxCare forms)	70 million	468 057
Free Text reports (Hospitalization, Surgery, Imaging, Pathology ...)	3.5 million	359 547
Lab test results	103 million	379 483
Drug prescriptions	3.5 million	121 441
Omics data	80 million	173

Entrepôt de données  
HEGP



# ENJEU: OÙ TROUVER LES DONNÉES?

IMAGERIE?



Information  
dérivée de  
l'image

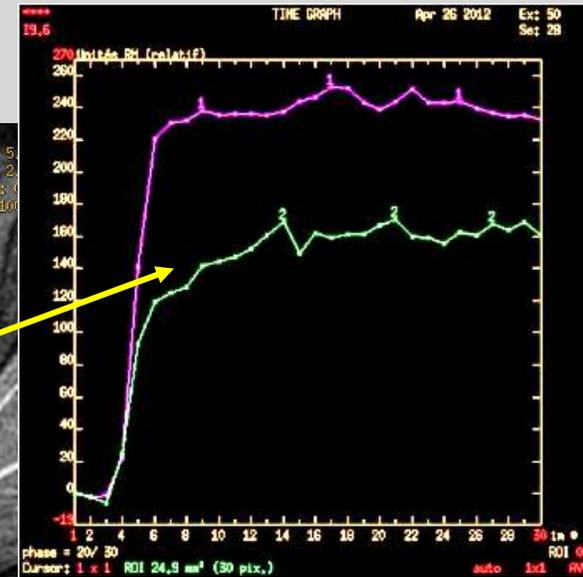
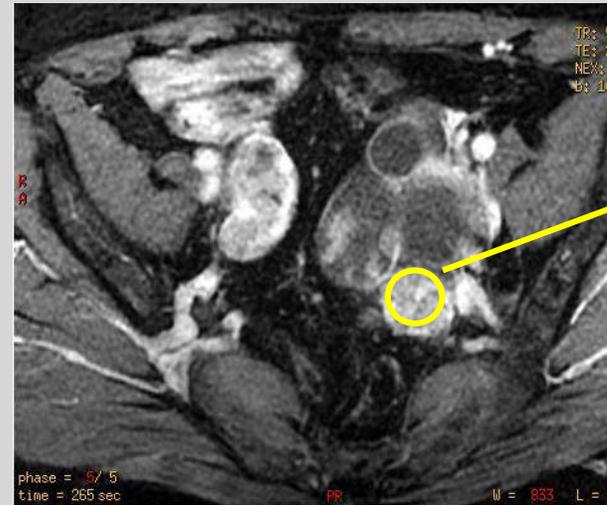
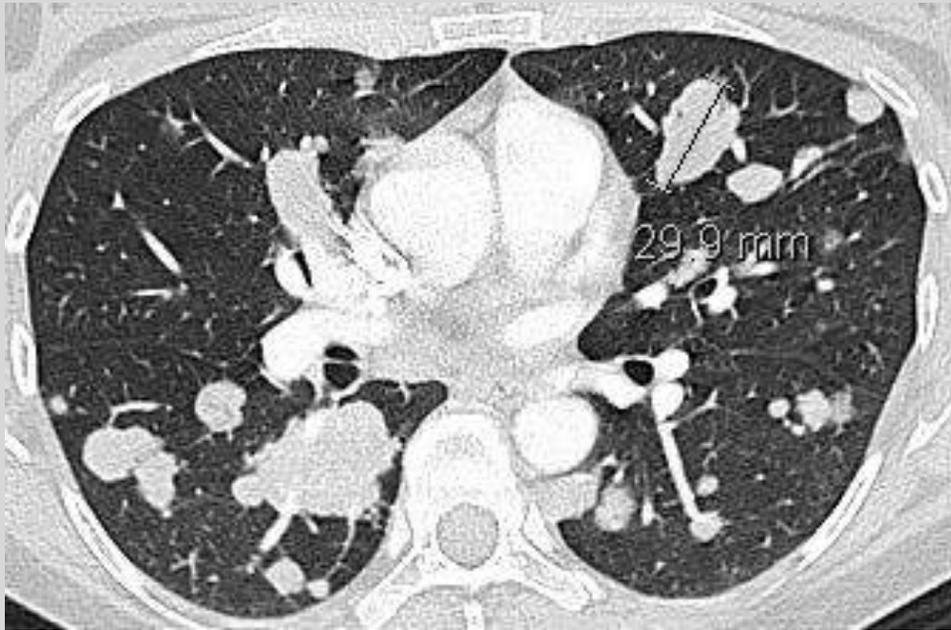
Type	# instances	# patients
Demographics (age, sex, Hospital vital status)		742 487
Signs (temperature, blood pressure, weight, ...)	14 million	141 725
Diseases (ICD 10 codes)	6 million	338 871
Medical Procedures (French CCAM codes)	4 million	285 097
Clinical data (EHR DxCare forms)	70 million	468 057
Free Text reports (Hospitalization, Surgery, Imaging, Pathology ...)	3.5 million	359 547
Lab test results	103 million	379 483
Drug prescriptions	3.5 million	121 441
Omics data	80 million	173

Entrepôt de données  
HEGP



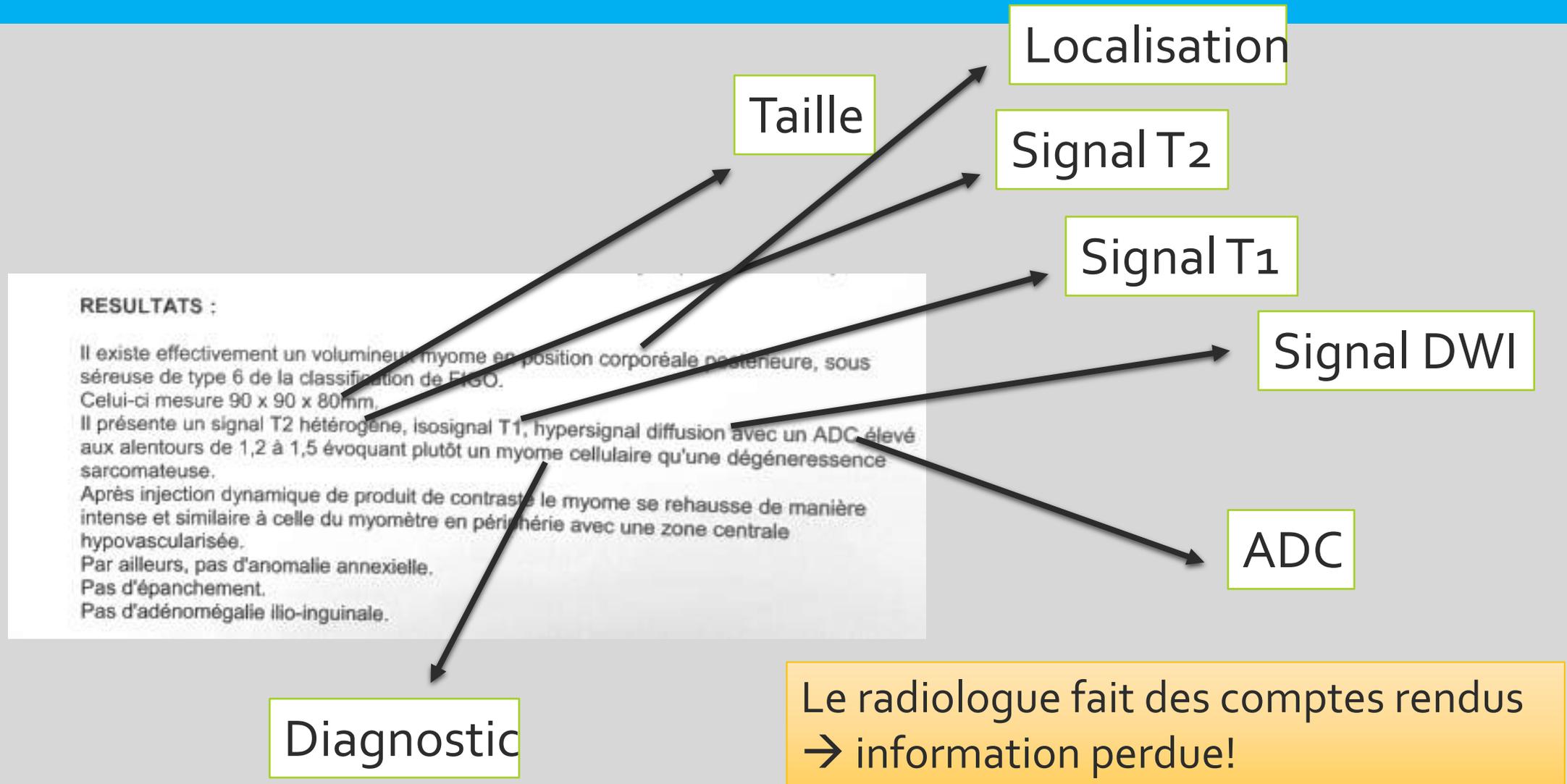
# ENJEU: OÙ TROUVER LES DONNÉES?

## Systeme d'archivage des images (PACS)



Le radiologue annote les images dans le PACS → information perdue!

# ENJEU: OÙ TROUVER LES DONNÉES?



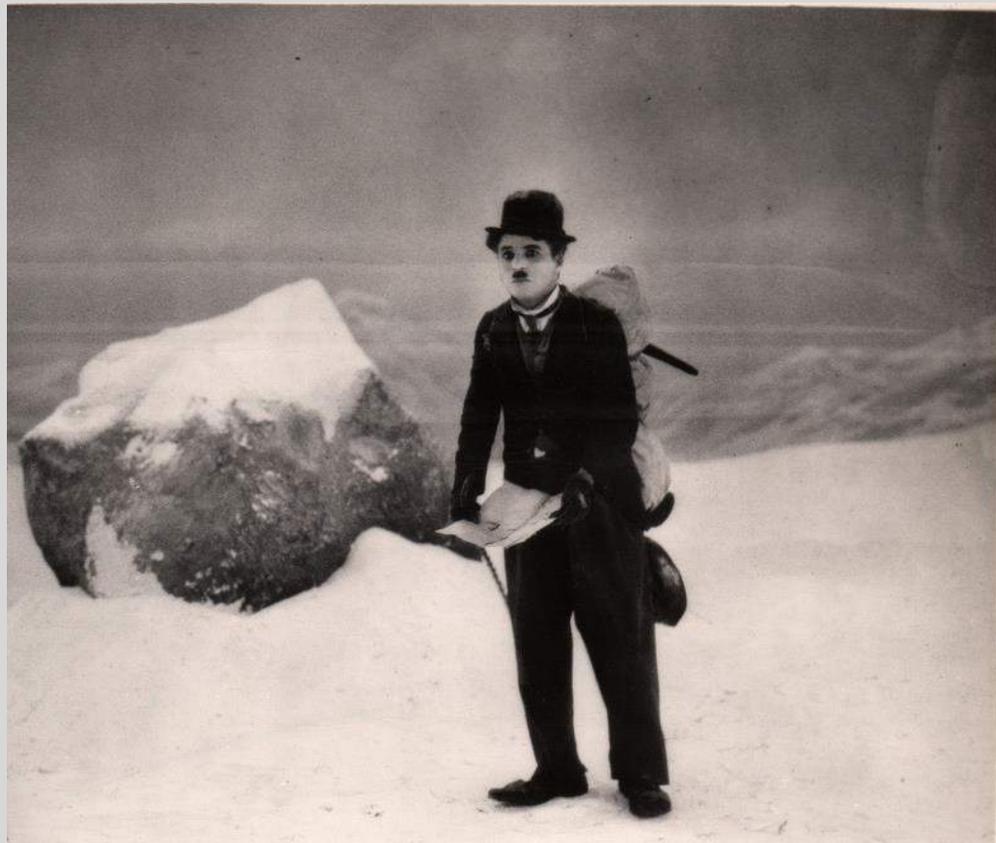
# ENJEU: OÙ TROUVER LES DONNÉES?

- ❑ Limites techniques: extraire les données
  - vision par ordinateur, traitement de la langue naturelle...
- ❑ Limites médicales
  - Nombre de patients limités par institution
  - Disparités de prise en charge entre les institutions, les pays...
  - « Moving target » : changements constants de traitements disponibles, de techniques d'imagerie

# CONCLUSION: RADIOMIQUE, POUR QUOI FAIRE?

- ❑ Pour **tester simultanément** un grand nombre de nouveaux paramètres extraits des images
  - Biomarqueurs potentiels qui doivent alors être compris et validés
- ❑ Pour générer de **nouvelles hypothèses**
  - Détecter des nouveaux phénotypes en imagerie pourrait permettre d'identifier des profils moléculaires pertinents
- ❑ Enjeux multiples
  - Obtenir des **cohortes** pertinentes
  - Avoir des résultats qui ont une **pertinence et conséquence médicale**

# RADIOMIQUE: UN TRAVAIL D'ÉQUIPE



Radiologue qui part chercher de l'or...

# RADIOMIQUE: UN TRAVAIL D'ÉQUIPE

- ❑ Service de radiologie, Hôpital Européen Georges Pompidou et équipe 2 "Imaging of angiogenesis" UMR 970, PARCC@HEGP
  - ❑ Daniel Balvay
  - ❑ Afef Bouchouicha
  - ❑ Jonas Deidier
  - ❑ Loïc Duron
  - ❑ Augustin Lecler
- ❑ Service de bio-informatique, Hôpital Européen Georges Pompidou et INSERM UMR\_S 1138 eq22 "Information Sciences to support Personalized Medicine"
  - ❑ Bastien Rance
  - ❑ Anita Burgun
- ❑ Service d'oncologie médicale, Hôpital Européen Georges Pompidou
  - ❑ Stéphane Oudard



Radiologue qui se fait des amis:  
Data scientists, bio-informaticiens, statisticiens...