Radiomics for outcome modeling: state-of-the-art and challenges

Mathieu Hatt, PhD, HDR – CR INSERM

mathieu.hatt@inserm.fr

Laboratoire du Traitement de l'Information Médicale

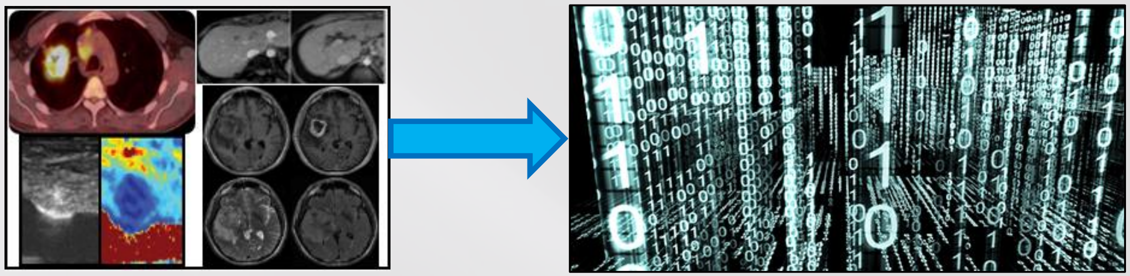
LaTIM, UMR INSERM-UBO 1101, Brest

Orsay, 23 Mars 2018

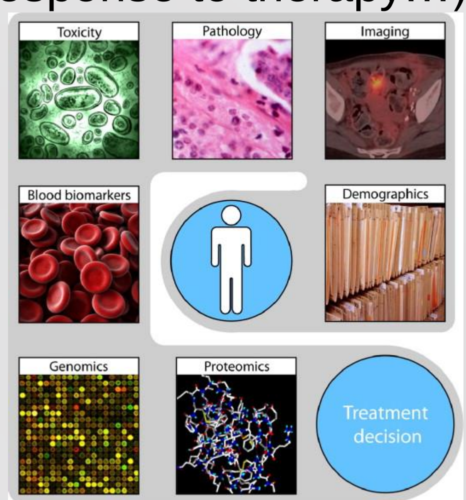
Introduction

Radiomics: definition

- Radiomics is the high-throughput extraction of quantitative features from medical images[1]
- The approach considers « pictures » as « minable » data[2]



- Radiomics aims at building models that are predictive of some patient outcome (e.g. survival, response to therapy…) or characteristic (tumor type, phenotype, genotype…)

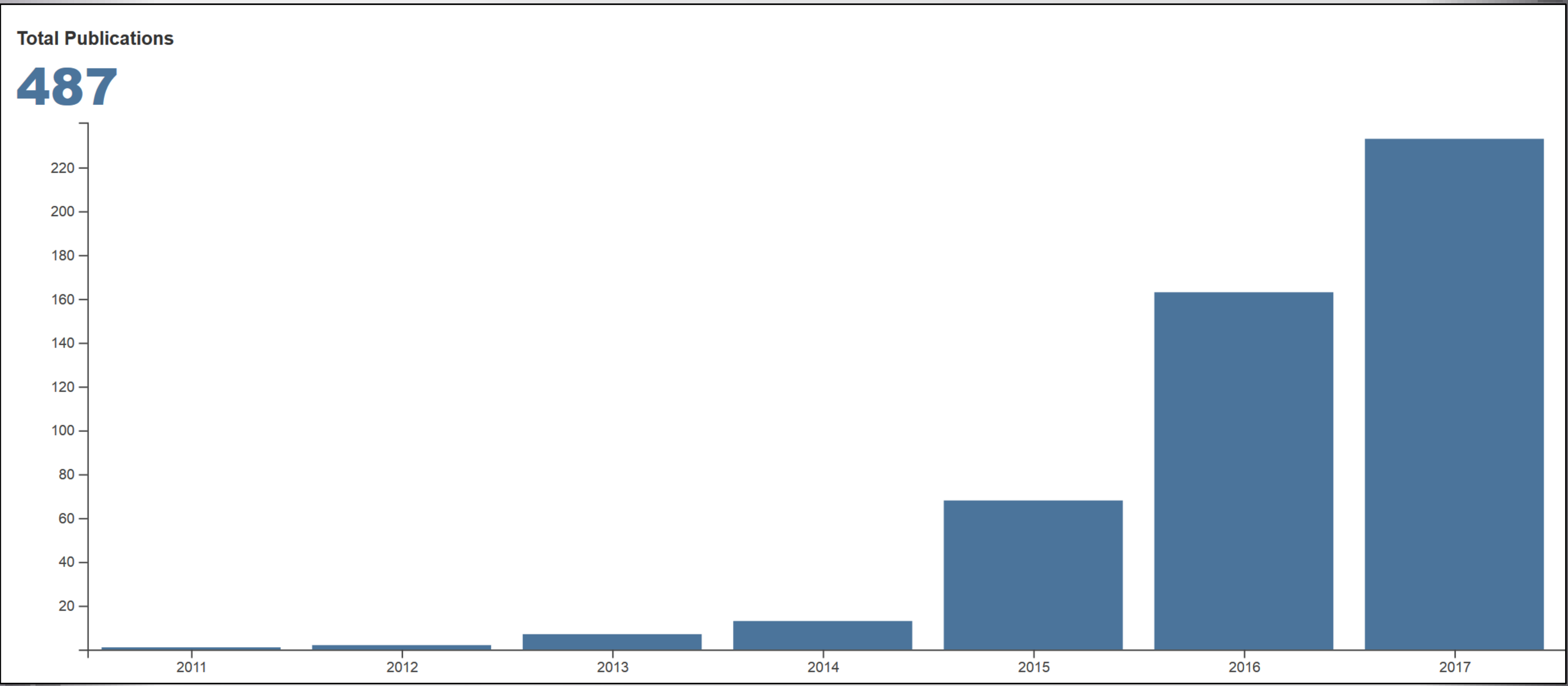


1. Lambin, *et al*. **Radiomics: extracting more information from medical images using advanced feature analysis**. *Eur J Cancer* 2012
2. Gillies, *et al*. **Radiomics: Images Are More than Pictures, They Are Data**. *Radiology* 2016

Introduction

Radiomics: exponential growth

Radiomics: ~500 publications



Source: web of science

Introduction

History of radiomics

The terms “radiomics” and “radiogenomics” were first employed in 2010 to describe how imaging features can reflect gene expression:

single diffusion weighting. Although quantitative values of diffusion are not derived, the data are nonetheless very amenable to pixelwise analysis of heterogeneity.

NIH-PA Author Manuscript

Anatomic Imaging and gene expression patterns: Radiomics

Referring again to figure 1, the physiology and anatomy of organs and tumors is driven by gene expression patterns which are a product of cellular genetics interfacing with the microenvironment. Over the last few years, it has become clear that distinct sub-regions of tumors, identifiable by MR imaging, have distinct gene expression patterns (31, 42–44). This indicates that underlying molecular biology can affect the “anatome”. Recently, there have been attempts to determine if quantitative analysis of the anatome can be used to infer an underlying molecular gene expression pattern. This involves “radiomics” which is the extraction of quantitative features from radiographic images. Relating these to gene expression patterns using sophisticated bioinformatic approaches is sometimes termed

Clin Radiol. Author manuscript; available in PMC 2014 May 04.

Gillies et al. Page 5

“radiogenomics”. The central hypothesis of cancer radiomics is that tumor imaging features

Gillies, *et al*. **The biology underlying molecular imaging in oncology: from genome to anatome and back again.** *Clin Radiol* 2010

Introduction

Radiomics: is this really new?

- The term radiomics has become popular since 2012
- Textural features (a large chunck of radiomics features) exist since the 70’s and have been used in medical imaging since the 90’s [1-3]
- Numerous publications before 2012 (quantification) could be categorized as « radiomics studies »

- Some « new » elements of radiomics:
  - Larger number of features (>hundreds) / « high-throughput »
  - Relying on machine learning (selection/classifier)
  - Link with biology (including genetics)

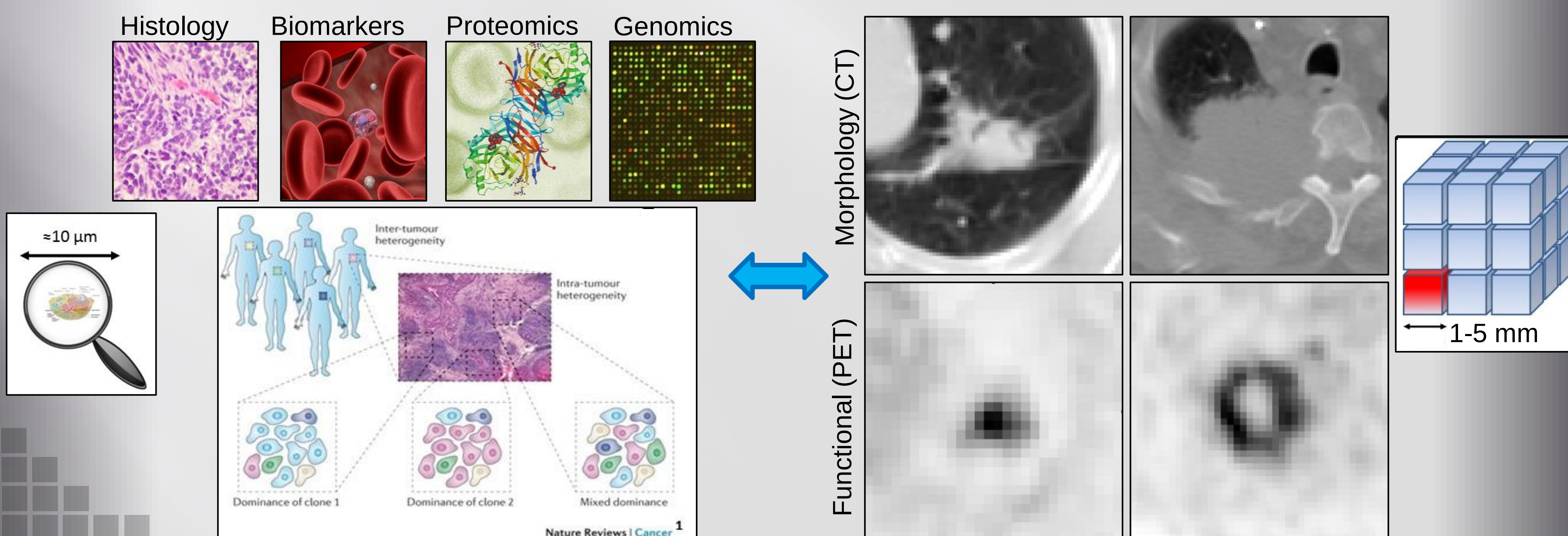
1. Schad, *et al*. **MR tissue characterization of intracranial tumors by means of texture analysis**. *Magn Reson Imaging* 1993
2. Mir, *et al*. **Texture analysis of CT-images for early detection of liver malignancy**. *Biomed Sci Instrum.* 1995
3. El Naqa, *et al*. **Exploring feature-based approaches in PET images for predicting cancer treatment outcomes**. *Pattern Recognit*. 2009

Introduction

Radiomics: rationale

Macroscopic/microscopic heterogenity

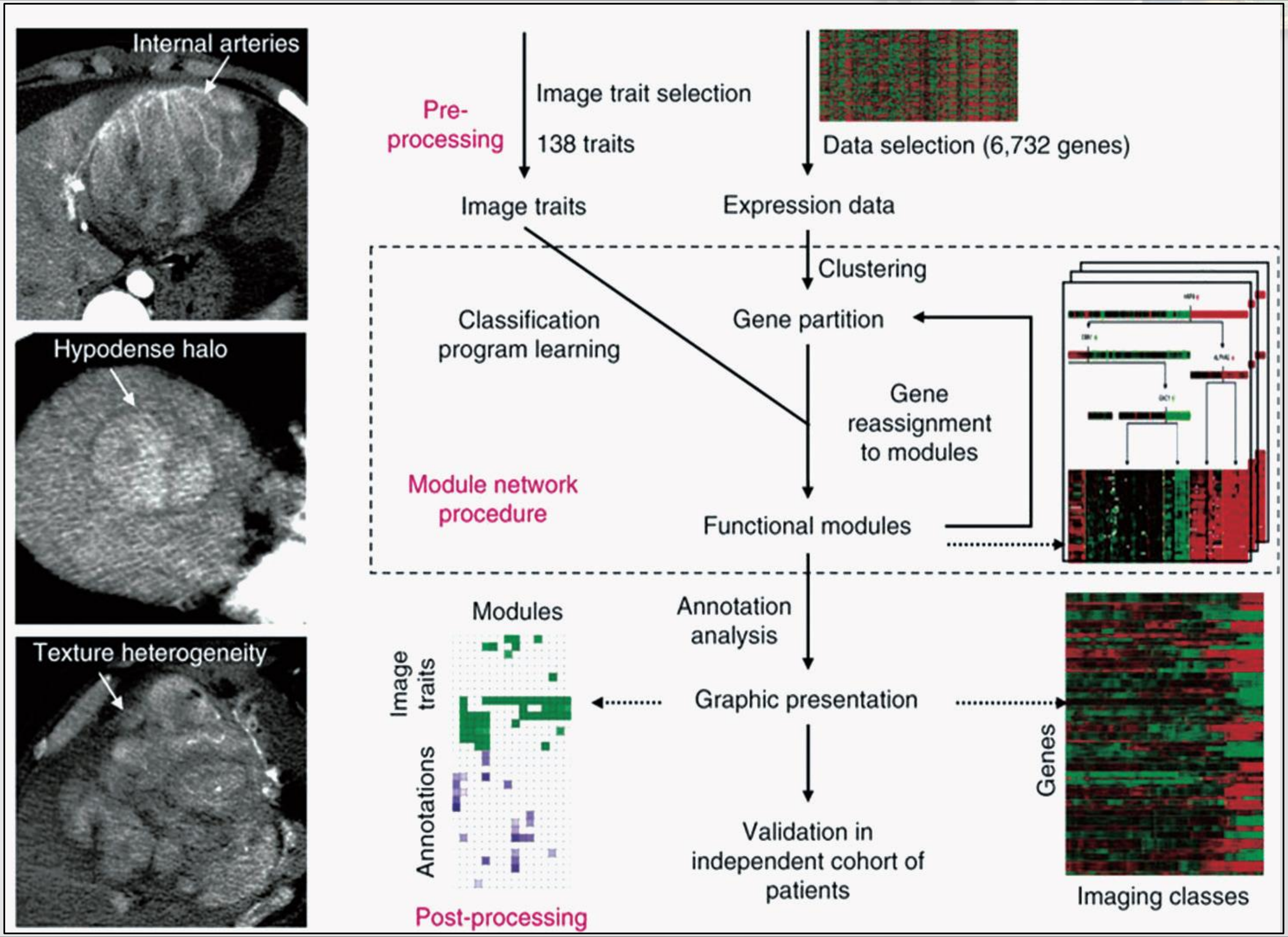
- Tumours are heterogeneous entities [1]
  - Genetic, cellular, tissular
  - <u>Hypothesis</u>: caracteristics in images (macro scale) reflect at least partly caracteristics in smaller scales (including genetic) [2]



1. Gerlinger, *et al*. **Intratumor heterogeneity and branched evolution revealed by multiregion sequencing**. *N Engl J Med.* 2012
2. Segal, *et al*. **Decoding global gene expression programs in liver cancer by noninvasive imaging**. *Nat Biotechnol*. 2007

Introduction

Early works (example)

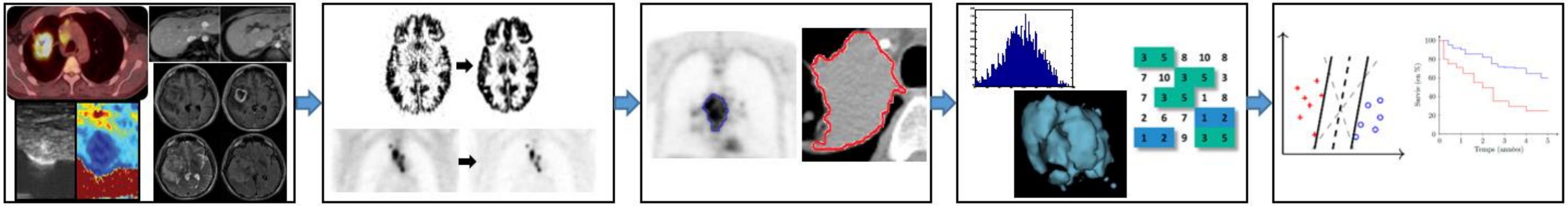


Segal, *et al*. **Decoding global gene expression programs in liver cancer by noninvasive imaging**. *Nat Biotechnol*. 2007

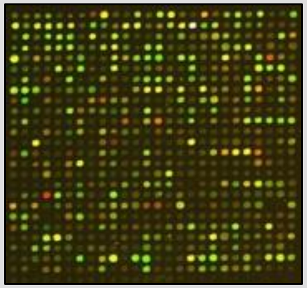
Radiomics

Standard worklow

Clinical data



Images | Pre-processing | Segmentation | Features extraction | Features selection Training & validation



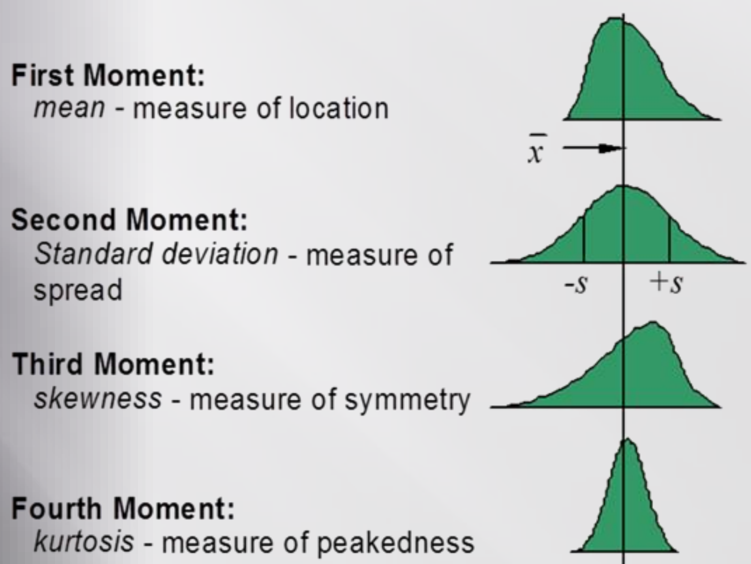
Genomics (and other –omics)

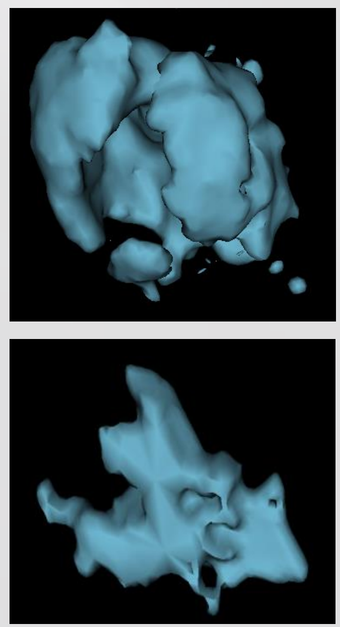
Radiomics

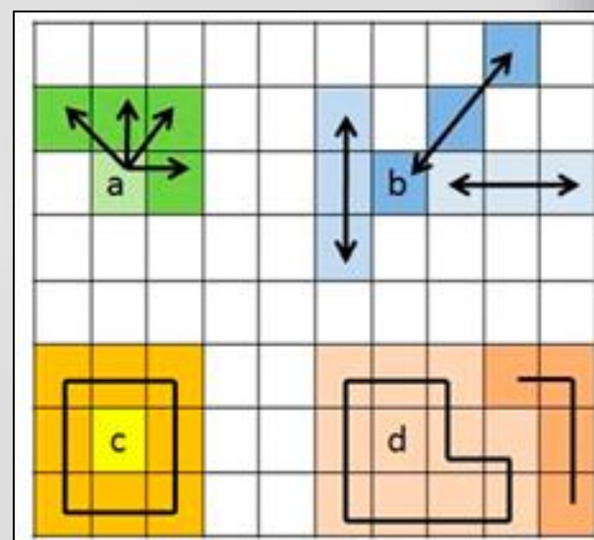
Features

« Usual » radiomics:

- Intensity-based (e.g. histogram)
- Shape descriptors (e.g. sphericity)
- Texture analysis 2nd or higher order (e.g. GLCM)







Lambin, *et al*. **Radiomics: extracting more information from medical images using advanced feature analysis**. *Eur J Cancer* 2012

Radiomics

Features

Less frequently used / more recent:

- Fractal analysis[1]
- Filter-based (e.g. Law's, Riesz[2] …)
- Others (metabolic gradient[3], CoLIAGE[4] …)

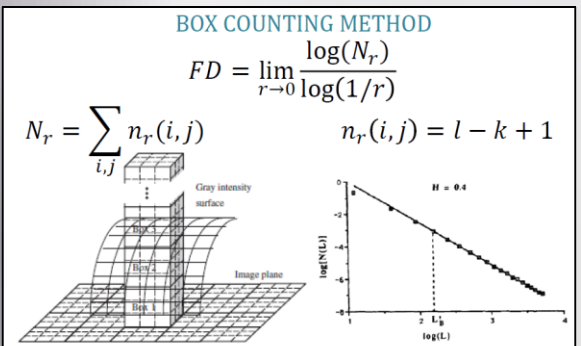
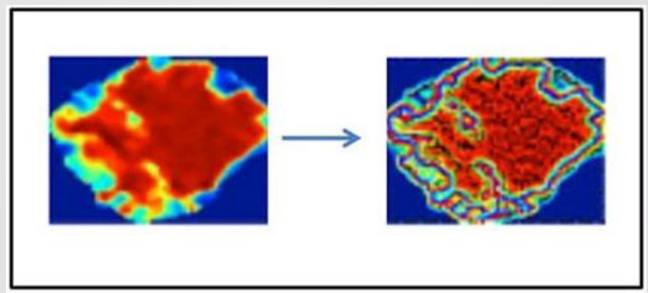
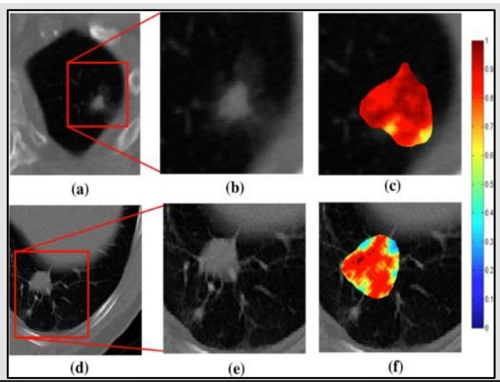
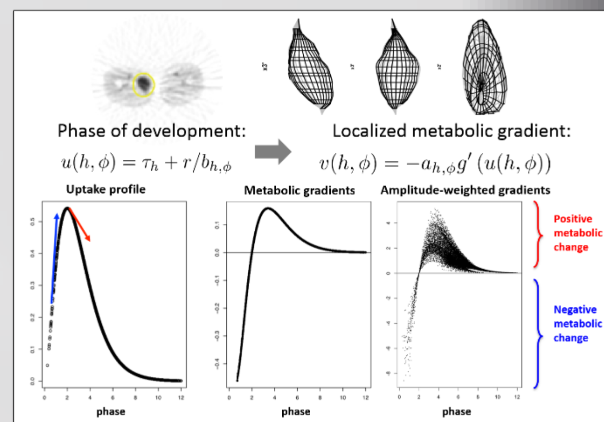


Image from Li et al., Pattern Recognition 2009; 420 (11): 2460–2469.





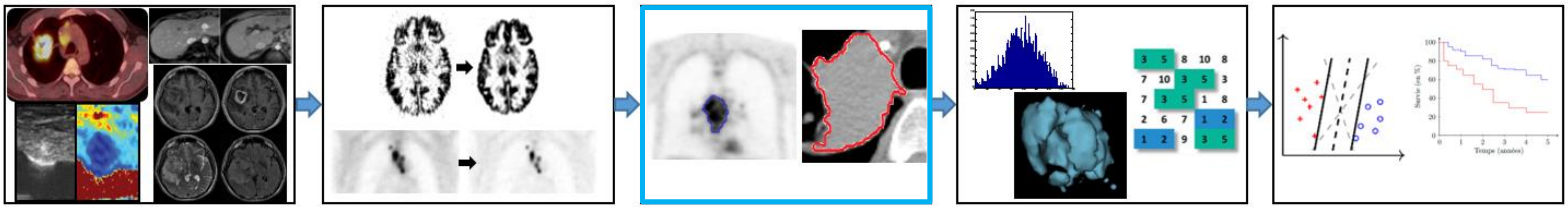


1. Michallek, *et al*. **Fractal analysis in radiological and nuclear medicine perfusion imaging: a systematic review**. *Eur Radiol*. 2014
2. Cirujeda, *et al*. **A 3-D Riesz-Covariance Texture Model for Prediction of Nodule Recurrence in Lung CT**. *IEEE Trans Med Imaging*. 2016
3. Wolsztynski, *et al*. **Localized metabolic gradient as an independent prognostic variable from FDG-PET in sarcoma**. *SNMMI*. 2017
4. Prasanna, *et al*. **Co-occurrence of Local Anisotropic Gradient Orientations (CoLIAGe): A new radiomics descriptor**. *Sci Rep*. 2016

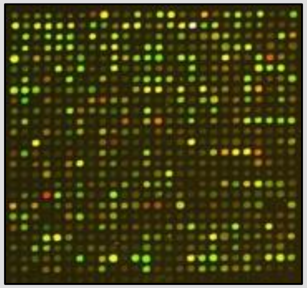
Radiomics

Standard workflow

Clinical data



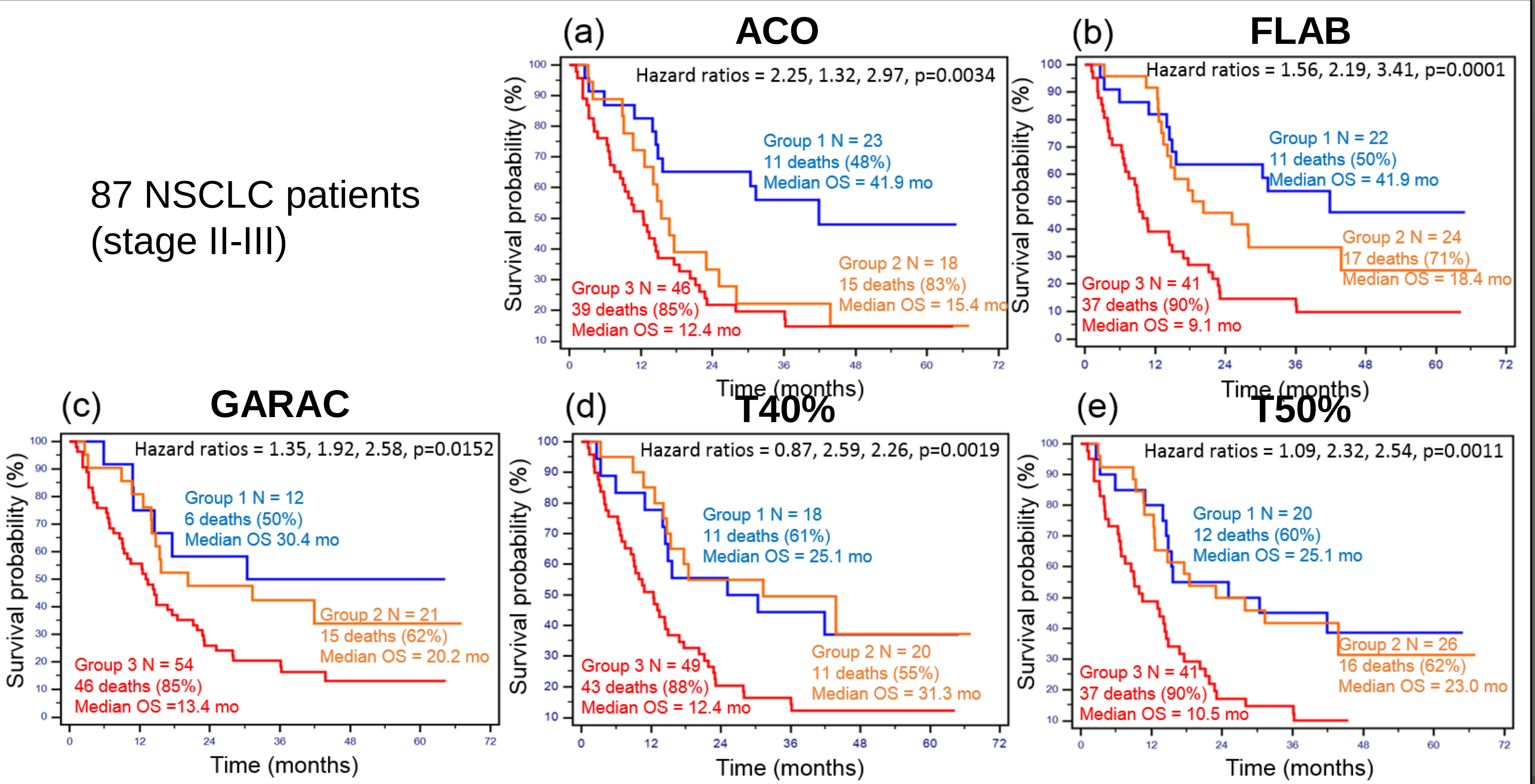
Images | Pre-processing | Segmentation | Features extraction | Features selection Training & validation



Genomics (and other –omics)

Segmentation step: how critical for radiomics?

87 NSCLC patients
(stage II-III)

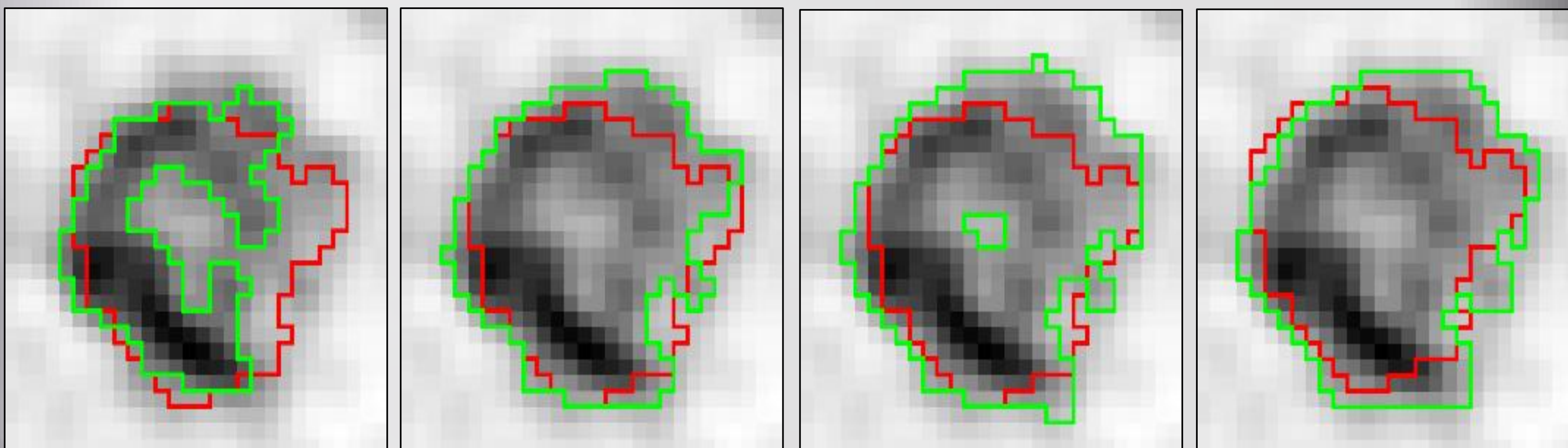


Hatt, *al*. **Tumour functional sphericity from PET images: prognostic value in NSCLC and impact of delineation method.** *Eur J Nucl Med Mol Imaging* 2018

Radiomics

Segmentation: impact?

- Segmentation step: how critical for radiomics?
  - Potential solutions:
    - Use ensemble / consensus methods (e.g. STAPLE)[1]

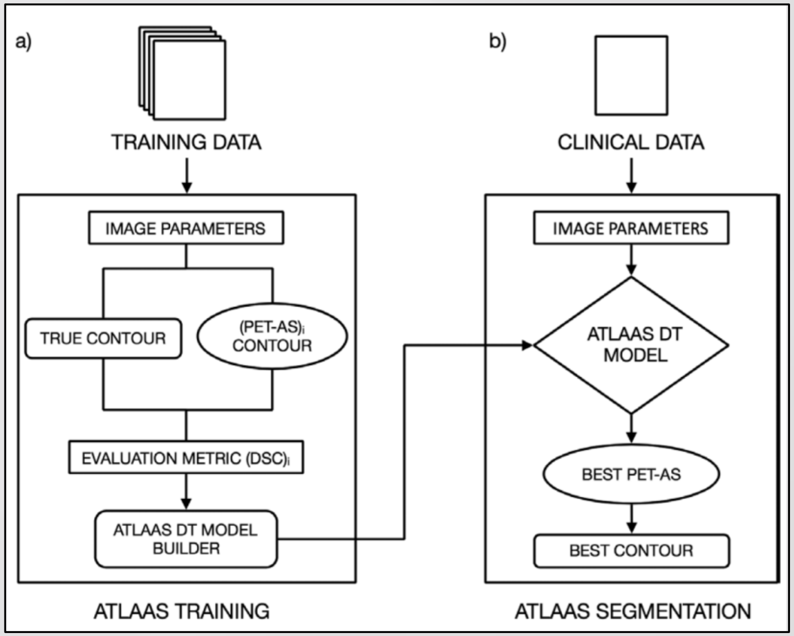


T40 FLAB CNN STAPLE consensus

1. Hatt, *et al*. **The first MICCAI challenge on PET tumor segmentation**. *Med Image Anal.* 2018
2. Berthon, *et al*. **ATLAAS: an automatic decision tree-based learning algorithm for advanced image segmentation in positron emission tomography.** *Phys Med Biol* 2016

Segmentation step: how critical for radiomics?

- Potential solutions:
  - Use ensemble / consensus methods (e.g. STAPLE)[1]
  - Use machine learning models that select the best method for a given configuration (e.g. ATLAAS)[2]

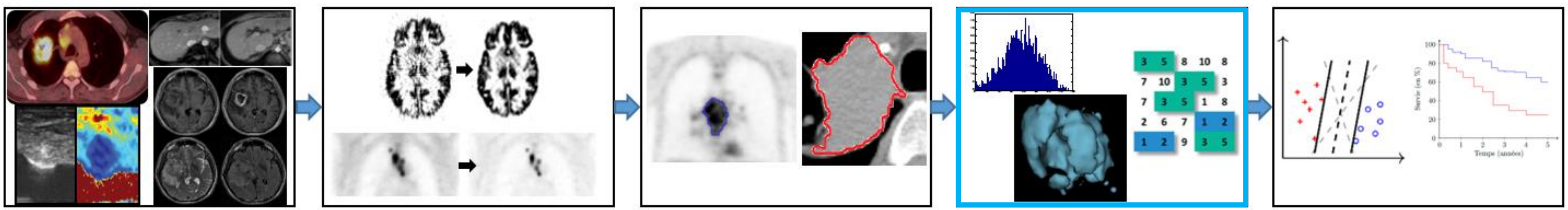


1. Hatt, *et al*. **The first MICCAI challenge on PET tumor segmentation**. *Med Image Anal.* 2018
2. Berthon, *et al*. **ATLAAS: an automatic decision tree-based learning algorithm for advanced image segmentation in positron emission tomography.** *Phys Med Biol* 2016

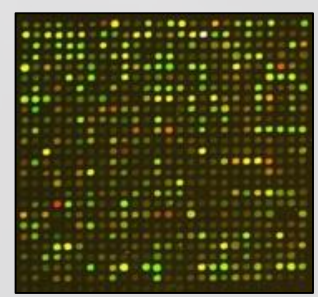
Radiomics

Worfklow

Clinical data



Images | Pre-processing | Segmentation | Features extraction | Features selection Training & validation



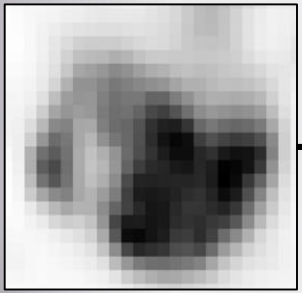
Genomics (and other –omics)

Radiomics

Challenges and issues: the complexity of textural features

- Workflow complexity

PET image



« Let's compute some textural features!»

Useful quantification of heterogeneity

Hatt, *et al.* **Characterization of PET/CT images using texture analysis: the past, the present… any future?** *Eur J Nucl Med Mol Imaging* 2017

Radiomics

Challenges and issues: the complexity of textural features

Reconstructed PET image → **1. Segmentation** → **Segmented tumor volume** → **2. Interpolation** → **Image with cubic voxels** → **3. Quantization** → **Quantized image** → **4. Texture matrices design** → **Texture matrices** → **5. Parameters calculation** → **Textural features** → **6. Statistical analysis** → **Selected features**

1. Segmentation
- Overall analysis
  - 2D or 3D
  - Segmentation tool (gradient, clustering…)
    - Including low/no uptake areas
    - Excluding low/no uptake areas

2. Interpolation
- Method
  - Nearest neighbors, B-Spline…

3. Quantization
- Method
  - Uniform, equal, Max-Lloyd, fixed bin size…
    - Quantization value
    - Size of bin

4. Texture matrices design
- Co-occurrence matrices
  - Number of directions
  - Voxels distance
  - 1 matrix per direction + averaging / no averaging

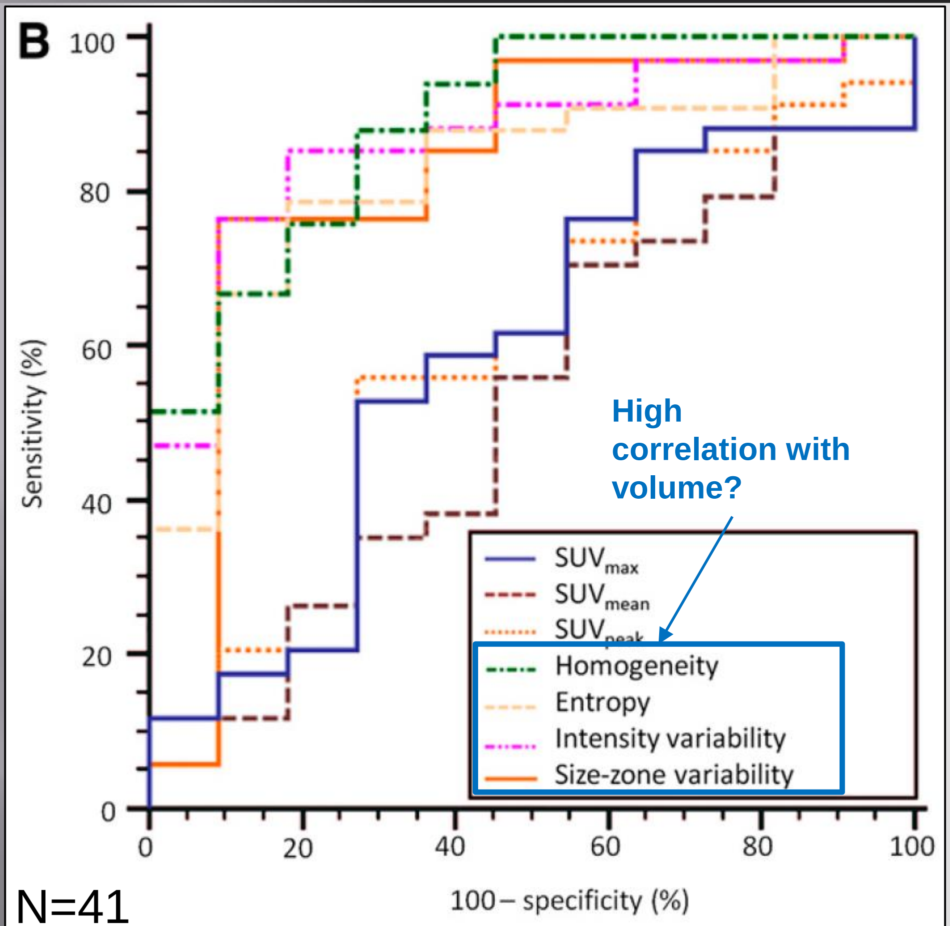
5. Parameters calculation
- Implementation
  - Actual formula used

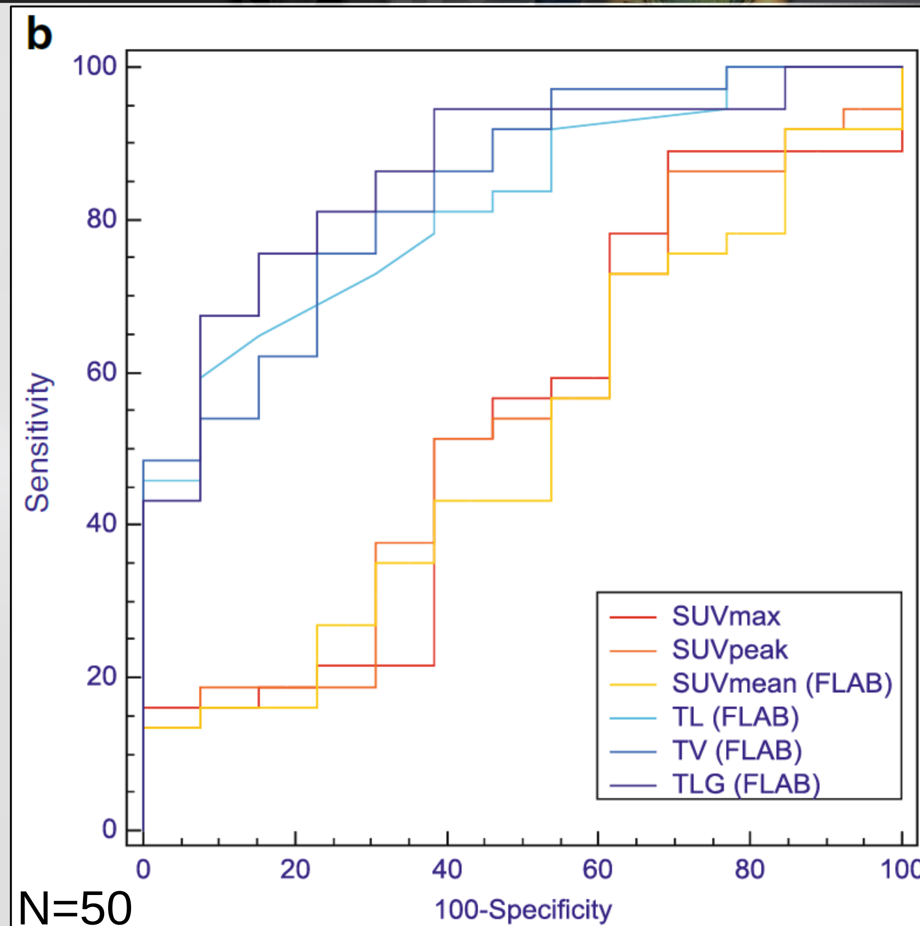
6. Statistical analysis
- Machine learning
  - Features selection method
  - Classifier
- Selection based on:
  - Robustness, reproducibility, redundancy, clinical value w.r.t. clinical endpoint

Hatt, *et al.* **Characterization of PET/CT images using texture analysis: the past, the present… any future?** *Eur J Nucl Med Mol Imaging* 2017

Radiomics

Challenges and issues: the volume/intensity confounding issue





FDG PET, esophageal cancer patients

Tixier, *et al*. **Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer**. *J Nucl Med* 2011

Hatt, *et al*. **Baseline ^{18}F-FDG PET image-derived parameters for therapy response prediction in oesophageal cancer**. *Eur J Nucl Med Mol Imaging* 2011

Radiomics

Challenges and issues: the volume/intensity confounding issue

Table 2 Correlations (Pearson coefficients) between parameters derived from FLAB delineations on noncorrected PET images. Significant correlation are shown in bold

Parameter	SUV_{mean}	MATV	Entropy	Homogeneity	Dissimilarity	Intensity variability	Size-zone variability	Zone percentage	High intensity emphasis	Area under the curve of the cumulative histogram
SUV_{mean}	**1.00**	0.20	**0.30**	−0.10	−0.02	0.08	0.09	**−0.40**	**0.40**	**−0.50**
MATV		**1.00**	**0.82**	**0.69**	**−0.77**	**0.97**	−0.16	**−0.70**	−0.22	0.07
Entropy			**1.00**	**0.60**	**−0.80**	**0.77**	−0.25	**−0.90**	−0.08	−0.07
Homogeneity				**1.00**	**−0.93**	**0.80**	**−0.36**	**−0.42**	**−0.67**	**0.59**
Dissimilarity					**1.00**	**−0.83**	**0.41**	**0.60**	**0.58**	**−0.45**
Intensity variability						**1.00**	−0.25	**−0.62**	**−0.41**	**0.28**
Size-zone variability							**1.00**	0.24	**0.43**	**−0.32**
Zone percentage								**1.00**	−0.18	**0.32**
High intensity emphasis									**1.00**	**−0.97**
Area under the curve of the cumulative histogram										**1.00**

Hatt, *et al*. **Robustness of intratumour ^{18}F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma**. *Eur J Nucl Med Mol Imaging* 2013

Radiomics

Challenges and issues: the volume/intensity confounding issue

Parameter	FT42%	AT	FLAB	FLAB PVC
SUV_{mean}	0.66	0.65	0.64	0.60
MATV	0.84	0.87	0.89	0.87
Entropy	0.84	0.86	0.88	0.85
Homogeneity	0.74[a]	0.82	0.86[a]	0.87
Dissimilarity	0.74[a]	0.81	0.85[a]	0.88
Intensity variability	0.85	0.87	0.90	0.88
Size-zone variability	0.66	0.70	0.72	0.86
Zone percentage	0.74	0.78	0.81	0.82
High intensity emphasis	0.59	0.65	0.65[a]	0.83[a]
Area under the curve of the cumulative histogram	0.56	0.60	0.60[a]	0.77[a]

Hatt, *et al*. **Robustness of intratumour ^{18}F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma**. *Eur J Nucl Med Mol Imaging* 2013

Radiomics

Challenges and issues: the volume/intensity confounding issue

CONCLUSION

Each PET-imaged tumor is a single sampling of all radioactivities that are physically and biologically permissible for that particular scanner–tumor combination. Because image heterogeneity statistics accrue manifestations of possibilities, it is the very nature of these statistics to reflect small sample sizes. Thus, inclusion of small tumor volumes necessarily biases tracer uptake heterogeneity studies toward statistically significant differences even when no difference in uptake exists. We have argued that this bias is lessened if all ROIs included in comparative heterogeneity analyses are above a minimum number of voxels. We have described a technique for computing this number that, when applied to our specific ^{18}F-FDG PET image data, $\boxed{\text{yields a minimum comparison volume of 45 cm}^3.}$

Instituts thématiques Inserm Institut national de la santé et de la recherche médicale

Brooks, *et al*. **The effect of small tumor volumes on studies of intratumoral heterogeneity of tracer uptake**. *J Nucl Med* 2014

Radiomics

Challenges and issues: the volume/intensity confounding issue

Example Heterogeneity Statistic

We computed the local information entropy of a 2-dimensional image as described by Haralick et al. (*13*). In brief, the cooccurrence matrix describes the probability p that a pixel of a shade i occurs next to a pixel of shade j. This matrix can be computed for various directions, pixel separations, and bit depths. We computed the horizontal and vertical cooccurrence matrices for the nearest pixel neighbors of 8-bit gray-scale images. From each of these matrices, the local entropy

$$h = \boxed{-\sum_{j=103}^{255} \sum_{i=103}^{255}} p(i,j) \ \ln p(i,j) \qquad \text{Eq. 1}$$

was computed for each direction and then root-mean-square–averaged to obtain a single local entropy value. The limits on the summations reflect the 40% clinical threshold within the 8-bit (0–255) color scale.

- A single texture: $\text{entropy}_{\text{GLCM}}$
- Calculated following one single workflow:
  - Linear discretization into 152 bins
  - 2 GLCM matrices for 2 directions (vertical+horizontal) followed by averaging

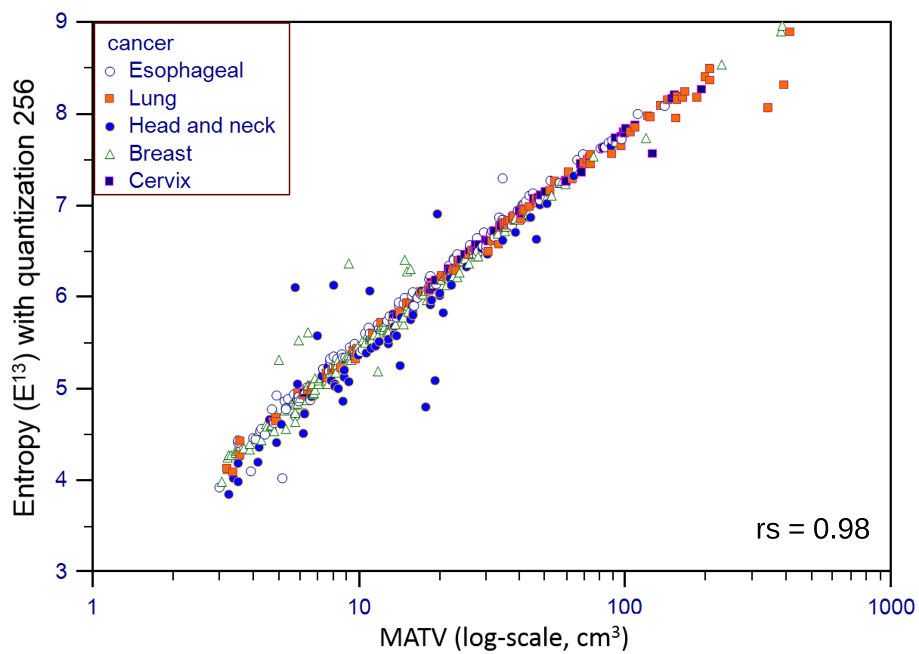
Instituts thématiques Inserm Institut national de la santé et de la recherche médicale

Brooks, *et al*. **The effect of small tumor volumes on studies of intratumoral heterogeneity of tracer uptake**. *J Nucl Med* 2014

Radiomics

Challenges and issues: the volume/intensity confounding issue

256 → 64 grey-levels



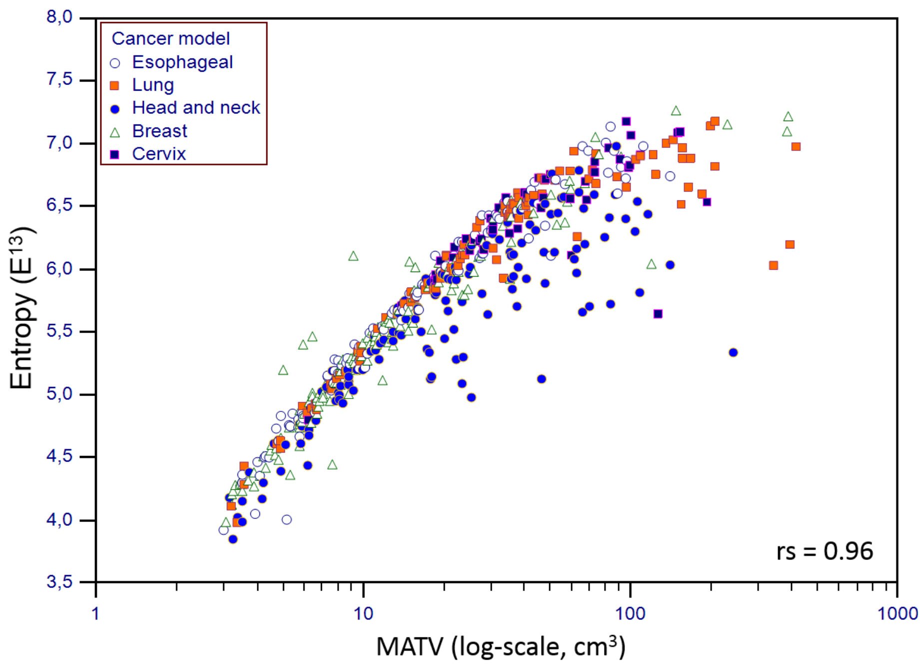
Hatt, *et al.* **^{18}F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort**. *J Nucl Med* 2015

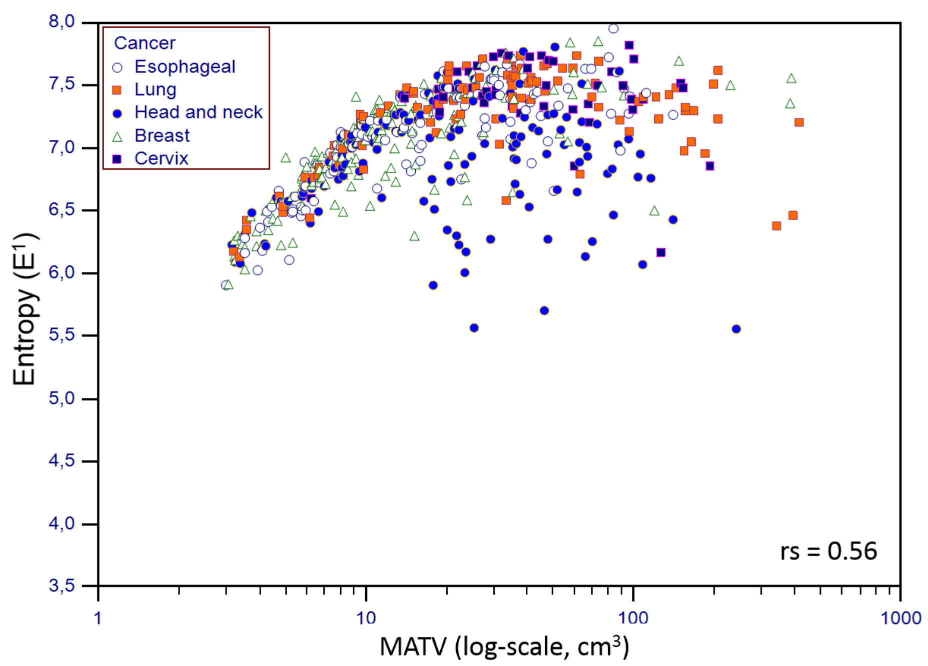
Radiomics

Challenges and issues: the volume/intensity confounding issue

256 → 64 grey-levels

13 GLCMs followed by averaging → 1 GLCM (13 directions)

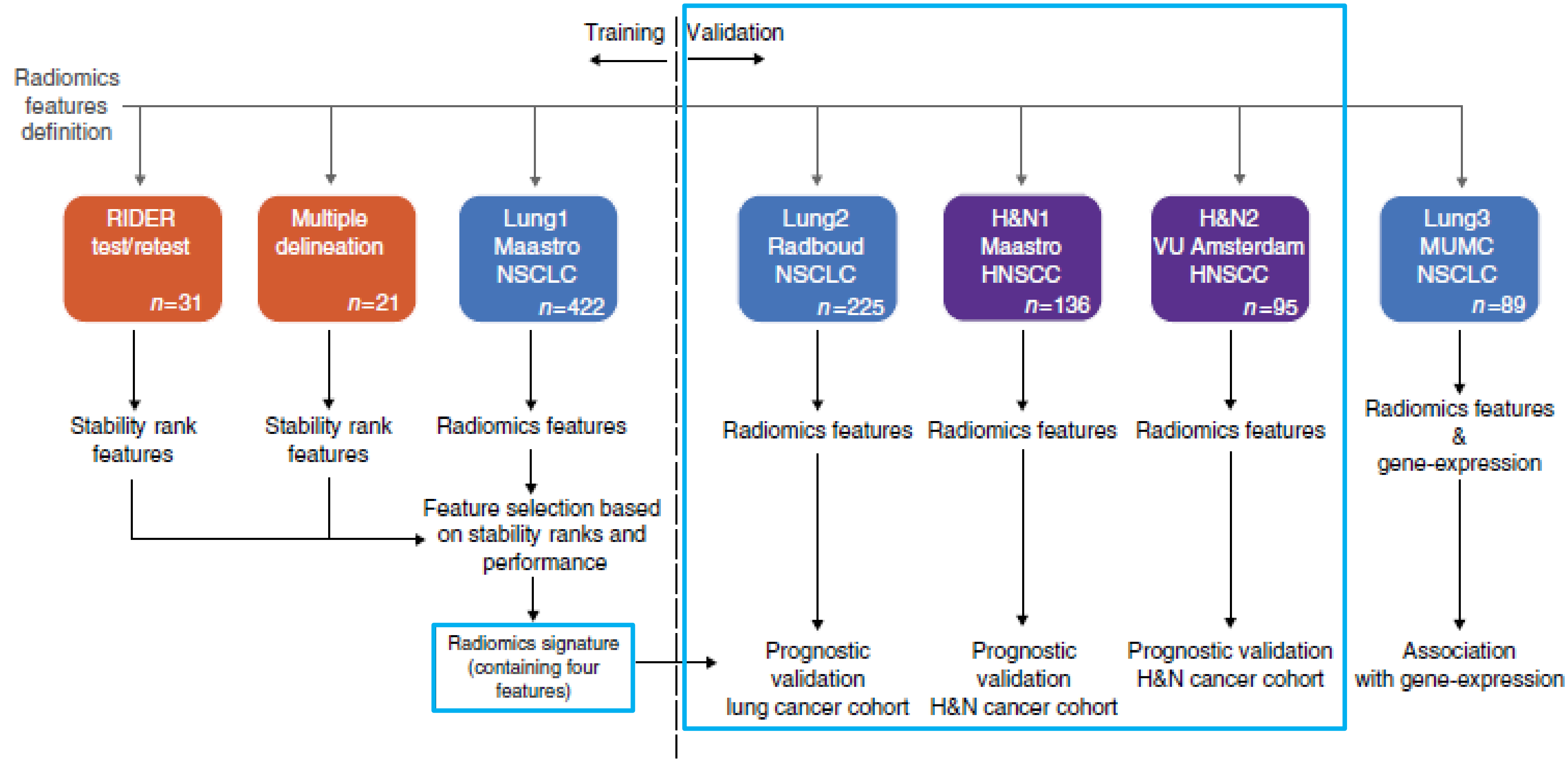




Hatt, *et al.* **^{18}F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort.** *J Nucl Med* 2015

Radiomics

Challenges and issues: the volume/intensity confounding issue

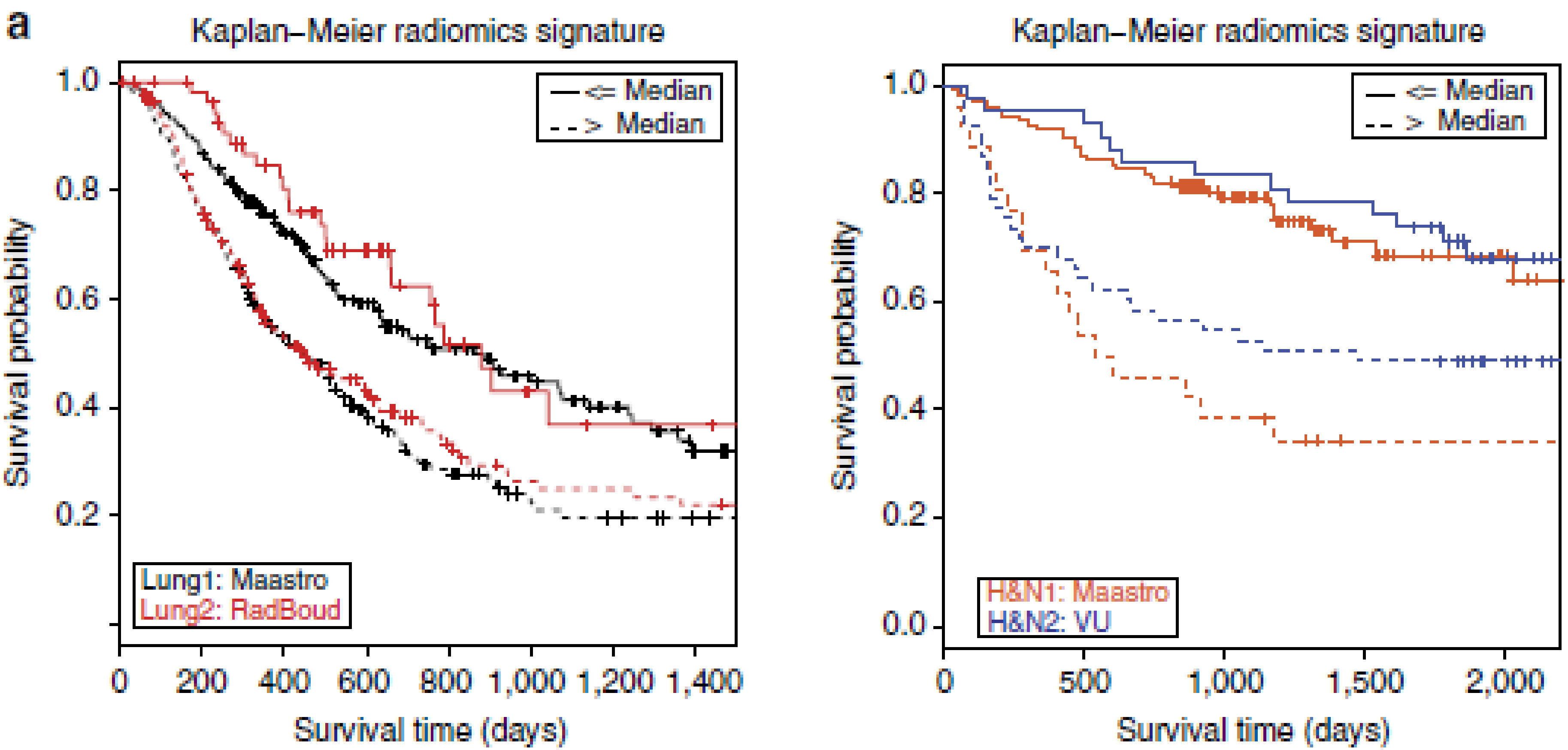


Aerts, *et al*. **Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach**. *Nat Commun*. 2014

Radiomics

Challenges and issues: the volume/intensity confounding issue

Intensity	Shape	Textural	Textural (wavelet)
Energy	Compactness	Grey-level non-uniformity	GLNU in HLH subband



Instituts thématiques Inserm Institut national de la santé et de la recherche médicale

Aerts, *et al*. **Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach**. *Nat Commun*. 2014

Radiomics

Challenges and issues: the volume/intensity confounding issue

Intensity	Shape	Textural	Textural (wavelet)
Energy	Compactness	Grey-level non-uniformity	GLNU in HLH subband

"(…) shown for the first time the translational capability of radiomics in two cancer types (…) radiomics quantifies a general prognostic cancer phenotype that likely can broadly be applied to other cancer types"

Instituts thématiques Inserm Institut national de la santé et de la recherche médicale

Aerts, *et al*. **Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach**. *Nat Commun*. 2014

Radiomics

Challenges and issues: the volume/intensity confounding issue

Supplemental table (C-index)

Dataset	TNM	Volume	Radiomics	TNM-Radiomics	Volume-Radiomics
Lung2	0.60	0.63	0.65	0.64	0.65
H&N1	0.69	0.68	0.69	0.70	0.69
H&N2	0.66	0.65	0.69	0.69	0.68

Spearman rank correlation with volume (N=300 H&N cancer patients):
PET: energy: 0.73, compactness: 0.98, GLNU: 0.99, GLNU_HLH: 0.89
CT: energy: 0.71, compactness: 0.94, GLNU: 0.98, GLNU_HLH: 0.95

Aerts, *et al*. **Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach.** *Nat Commun*. 2014
Vallières, et al. **Dependency of a validated radiomics signature on tumour volume and potential corrections.** (submitted)

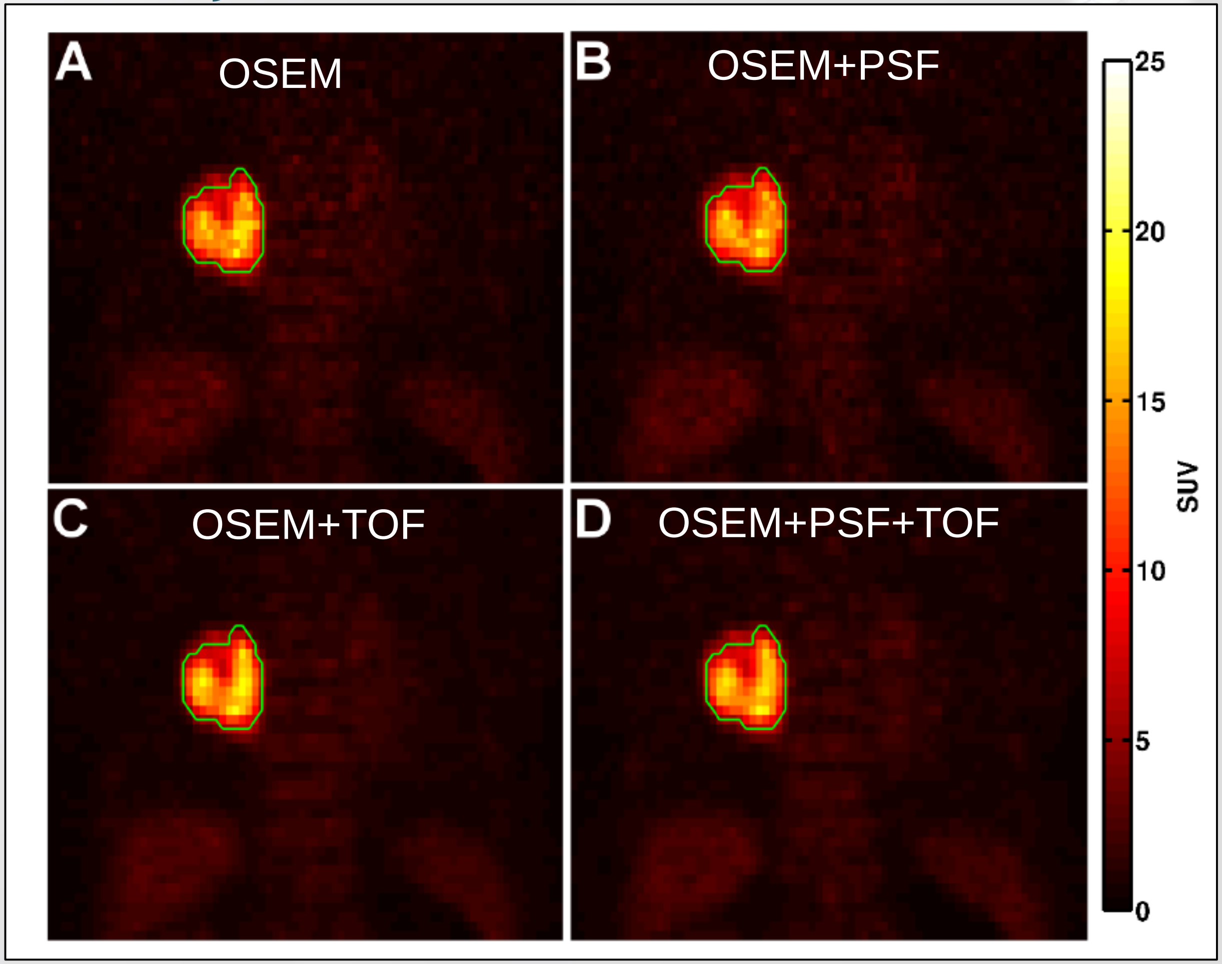
Dependency on reconstruction: PET

Image #	Acq. Mode	Grid-Size	Recon. Alg	Iter. number	Post-filter width (mm)	Legend
1	2D	128×128	OSEM	2	3	2D-128-OSEM2-3mm
2	2D	128×128	OSEM	2	5	2D-128-OSEM2-5mm
3	2D	128×128	OSEM	4	5	2D-128-OSEM4-5mm
4	2D	256×256	OSEM	2	3	2D-256-OSEM2-3mm
5	2D	256×256	OSEM	2	5	2D-256-OSEM2-5mm
6	3D	128×128	ITER	2	3	3D-128-ITER2-3mm
7	3D	128×128	ITER	2	6	3D-128-ITER2-6mm
8	3D	128×128	ITER	4	6	3D-128-ITER4-6mm
9	3D	256×256	ITER	2	3	3D-256-ITER2-3mm
10	3D	256×256	ITER	2	6	3D-256-ITER2-6mm

Acq. Mode = acquisition mode; Recon. Alg = reconstruction algorithm; Iter = iteration.

Gavalis, *et al*. **Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters**. *Acta Oncol*. 2010
Yan, *et al*. **Impact of Image Reconstruction Settings on Texture Features in 18F-FDG PET.** *J Nucl Med* 2015

Dependency on reconstruction: PET



Gavalis, *et al*. **Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters**. *Acta Oncol*. 2010
Yan, *et al*. **Impact of Image Reconstruction Settings on Texture Features in 18F-FDG PET.** *J Nucl Med* 2015

Radiomics

Challenges and issues: robustness/reliability

Dependency on reconstruction: PET

First-order features:
1: Mean
2: Median
3: Variance
4: Coefficient of variance
5: Skewness
6: Kurtosis
7: Energy
8: Entropy

Second-order features:
9: Homogeneity
10: Contrast-GLCM
11: Correlation
12: Sum of squares
14: Sum average
15: Sum variance
16: Sum entropy
17: Entropy-GLCM
18: Difference variance
19: Difference entropy
20: Information measure of correlation
21-22: Maximal correlation coefficient
23: Maximal probability
24: Diagonal moment
25: Dissimilarity
26: Difference energy
27: Inertia
28: Inverse difference moment
29: Sum energy
30: Cluster shade
31: Cluster prominence

Multicentric data !

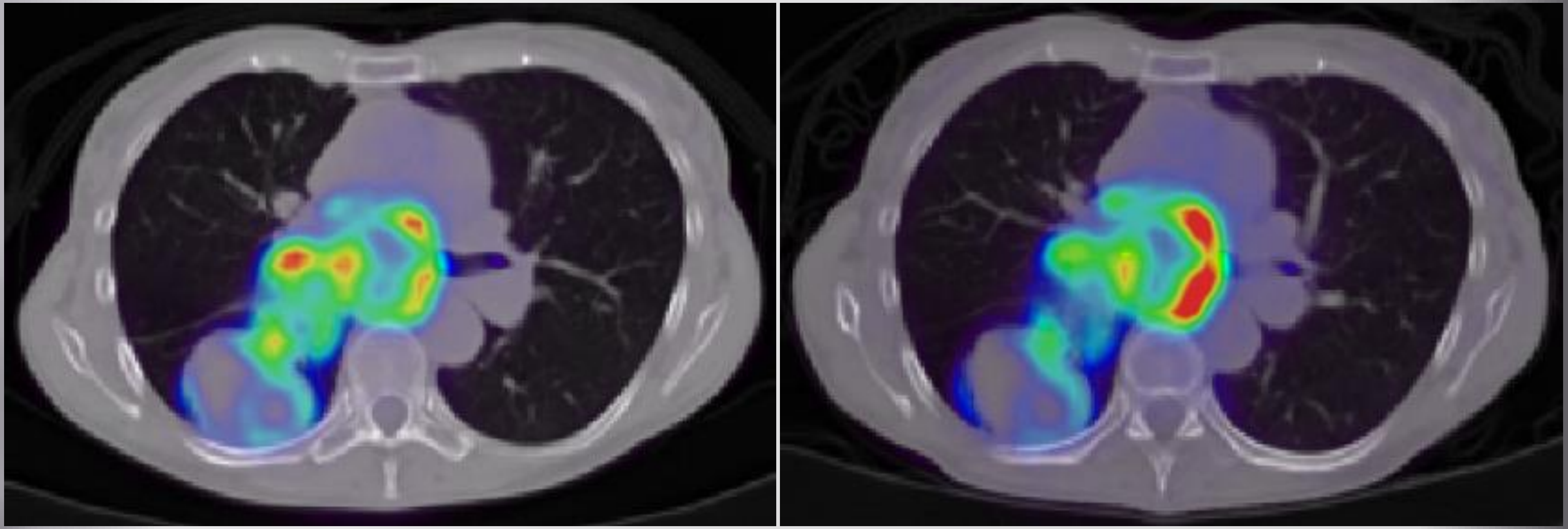
Gavalis, *et al*. **Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters**. *Acta Oncol*. 2010

Yan, *et al*. **Impact of Image Reconstruction Settings on Texture Features in 18F-FDG PET.** *J Nucl Med* 2015

Radiomics

Challenges and issues: robustness/reliability

Test-retest



Test PET/CT

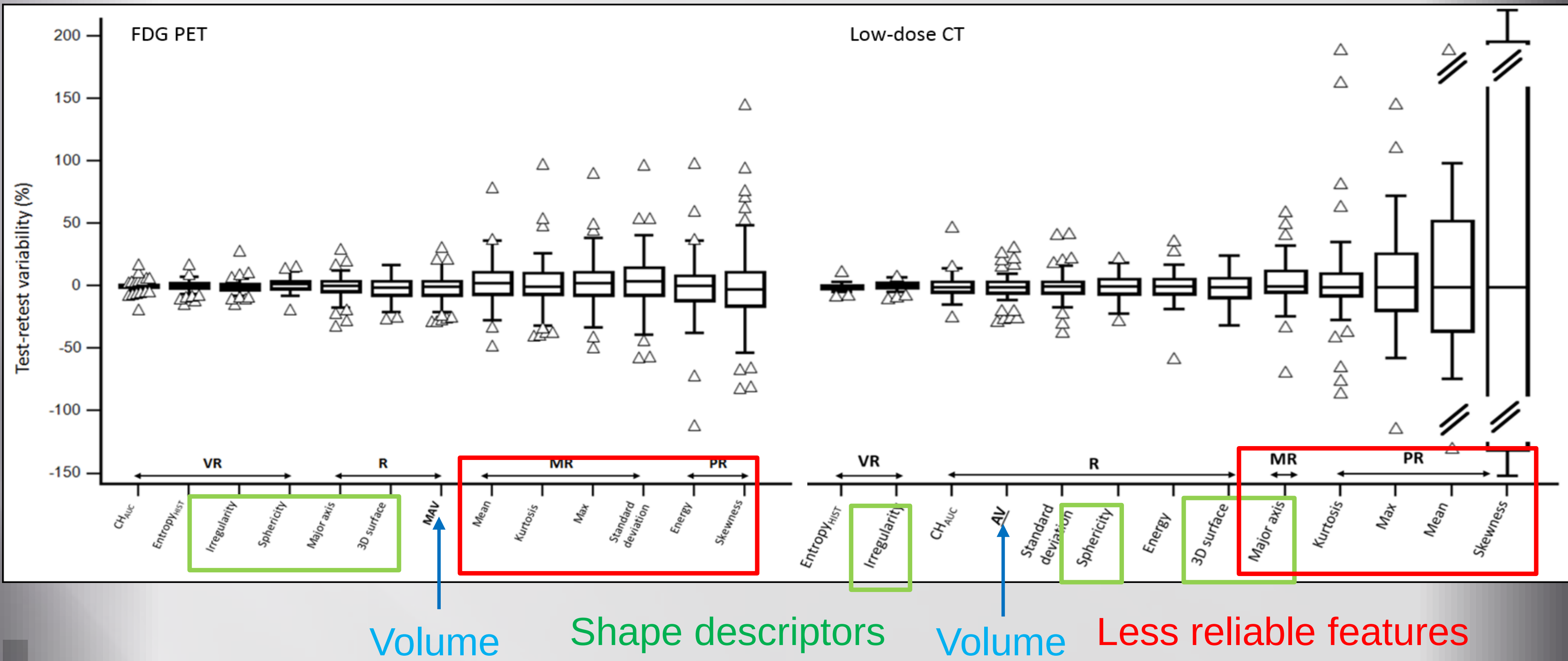
Re-test PET/CT

Tixier, *et al*. **Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET**. *J Nucl Med*. 2012
Desseroit, *et al*. **Reliability of PET/CT shape and heterogeneity features in functional and morphological components of Non-Small Cell Lung Cancer tumors: a repeatability analysis in a prospective multi-center cohort.** *J Nucl Med* 2016

Radiomics

Challenges and issues: robustness/reliability

Test-retest



Tixier, *et al*. **Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET**. *J Nucl Med*. 2012

Desseroit, *et al*. **Reliability of PET/CT shape and heterogeneity features in functional and morphological components of Non-Small Cell Lung Cancer tumors: a repeatability analysis in a prospective multi-center cohort.** *J Nucl Med* 2016

Radiomics

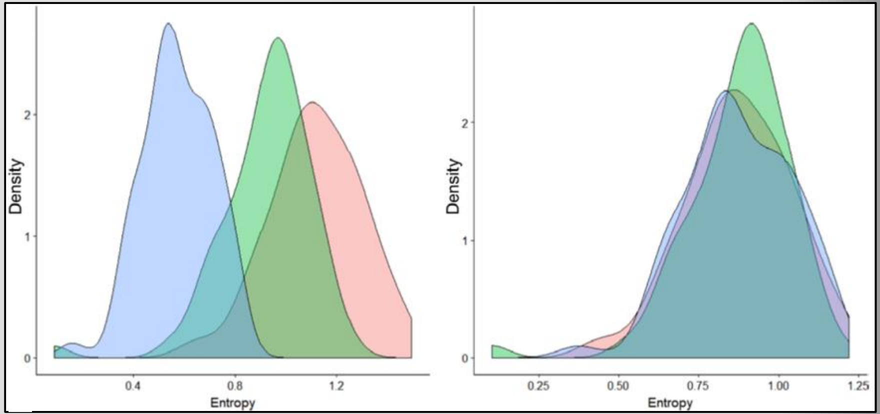
Challenges and issues: robustness/reliability

- Identify a compromise between:
  - Non reliable features (small pattern changes lead to large variability of the feature values)

  and

  - Perfectly « robust » features (always give the same value, unable to capture patterns or changes)

- Solutions for multi-centric data:
  - Use robust features only[1]
  - Pre-process images[2]
  - Post-process features[3]



Instituts thématiques Inserm
Institut national de la santé et de la recherche médicale

1. Upadhaya, *et al*. **Prognosis classification in glioblastoma multiforme using multimodal MRI derived heterogeneity textural features: impact of pre-processing choices**. *SPIE Medical Imaging* 2016
2. Vallières, *et al*. **A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities**. *Phys Med Biol*. 2018
3. Orlhac, *et al*. **A post-reconstruction harmonization method for multicenter radiomic studies in PET**. *J Nucl Med*. 2018

Radiomics

Challenges and issues: lack of standardisation

- Lack of standardisation and reproducibility of results
  - Different definitions / nomenclature
  - Missing implementation details
  - Different implementations / software (black boxes)

This results in:

- Sometimes confusing literature
- Meta-analysis impossible (e.g. entropy in paper 1 may not be the same entropy as in paper 2 !)
- Difficult or even impossible to reproduce / confirm the results

Vallières, *et al*. **Radiomics: Responsible Research For Faster Clinical Translation.** *J Nucl Med* 2018

Nomenclature

Textural Parameters of Tumor Heterogeneity in ^{18}F-FDG PET/CT for Therapy Response Assessment and Prognosis in Patients with Locally Advanced Rectal Cancer

Ralph A. Bundschuh[1–3], Julia Dinges[1], Larissa Neumann[1], Martin Seyfried[1], Norbert Zsótér[4], Laszló Papp[4], Robert Rosenberg[5], Karen Becker[6], Sabrina T. Astner[7], Martin Henninger[8], Ken Herrmann[2], Sibylle I. Ziegler[1], Markus Schwaiger[1], and Markus Essler[1,3]

[1]Nuklearmedizinische Klinik und Poliklinik, Klinikum rechts der Isar der Technischen Universität München, Munich, Germany; [2]Klinik und Poliklinik für Nuklearmedizin, Universitätsklinikum Würzburg, Wuerzburg, Germany; [3]Klinik und Poliklinik für Nuklearmedizin, Universitätsklinikum Bonn, Bonn, Germany; [4]Mediso Medical Imaging Systems Ltd., Budapest, Hungary; [5]Chirurgische Klinik, Kantonsspital Baden, Baden, Switzerland; [6]Institut für Pathologie, Klinikum rechts der Isar der Technischen Universität München, Munich, Germany; [7]Klinik und Poliklinik für Radioonkolgie und Strahlentherapie, Klinikum rechts der Isar der Technischen Universität München, Munich, Germany; and [8]Institut für Röntgendiagnostik, Klinikum rechts der Isar der Technischen Universität München, Munich, Germany

Bundschuh, *et al*. **Textural Parameters of Tumor Heterogeneity in ^{18}F-FDG PET/CT for Therapy Response Assessment and Prognosis in Patients with Locally Advanced Rectal Cancer**. *J Nucl Med*. 2014

Radiomics

Challenges and issues: nomenclature

Nomenclature

Parameter	AUC	95% confidence interval
SUV_{max}	0.52	0.32–0.71
Skewness	0.55	0.33–0.75
Kurtosis	0.61	0.39–0.81
SUV_{mean}	0.68	0.48–0.85
Diameter	0.68	0.48–0.85
COV	0.73	0.53–0.88
Volume	0.75	0.55–0.90
TLG	0.79	0.59–0.92

1st order features ≠ textural features !

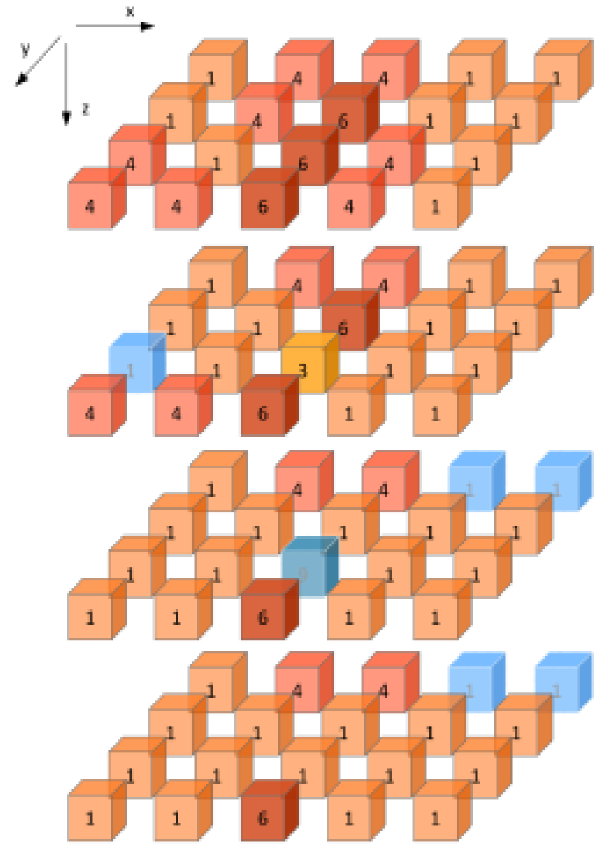
Bundschuh, *et al*. **Textural Parameters of Tumor Heterogeneity in ^{18}F-FDG PET/CT for Therapy Response Assessment and Prognosis in Patients with Locally Advanced Rectal Cancer**. *J Nucl Med*. 2014

Radiomics

Challenges and issues: lack of standardization

Imaging biomarkers standardisation initiative

- 06/2016-02/2018
- 20 research groups, 8 countries:
  - USA
  - Germany
  - The Netherlands
  - France
  - Canada
  - United Kingdom
  - Italy
  - Switzerland



Digital phantom. Blue voxels lie outside of the region of interest.

Participants

Study leader: *Alex Zwanenburg*

Cardiff University
Philip Whybra, Emiliano Spezi

Dana Farber Cancer Institute and Brigham and Women's Hospital, Harvard University
Andriy Fedorov, Hugo Aerts

Gemelli ART, Università Cattolica del Sacro Cuore
Jacopo Lenkowicz, Luca Boldrini, Nicola Dinapoli, Vincenzo Valentini

German Cancer Research Center (DKFZ)
Michael Götz, Nils Gählert, Fabian Isensee, Klaus H. Maier-Hein

INSERM Brest, University of Brest
Marie-Charlotte Desseroit, Taman Upadhaya, Mathieu Hatt

Leiden University Medical Center
Floris H.P. van Velden

MAASTRO clinic, Maastricht University
Ralph T.H. Leijenaar, Philippe Lambin

McGill University
Martin Vallières, Issam El Naqa

Memorial Sloan Kettering Cancer Center
Aditya Apte

Moffitt Cancer Center
Mahmoud A. Abdalah, Robert Gillies

OncoRay – National Center for Radiation Research in Oncology and NCT Dresden
Alex Zwanenburg, Stefan Leger, Esther Troost, Christian Richter, Steffen Löck

The Netherlands Cancer Institute (NKI)
Joost van Griethuysen, Cuong Viet Dinh, Uulke van der Heide

Universitätsklinikum Tübingen, Eberhard Karls University Tübingen
Jairo Socarras Fernandez, Daniela Thorwarth

University Hospital Zürich, University of Zürich
Marta Bogowicz, Stephanie Tanadini-Lang, Matthias Guckenberger

University of Bergen
Are Losnegård

University of California, San Francisco
Olivier Morin

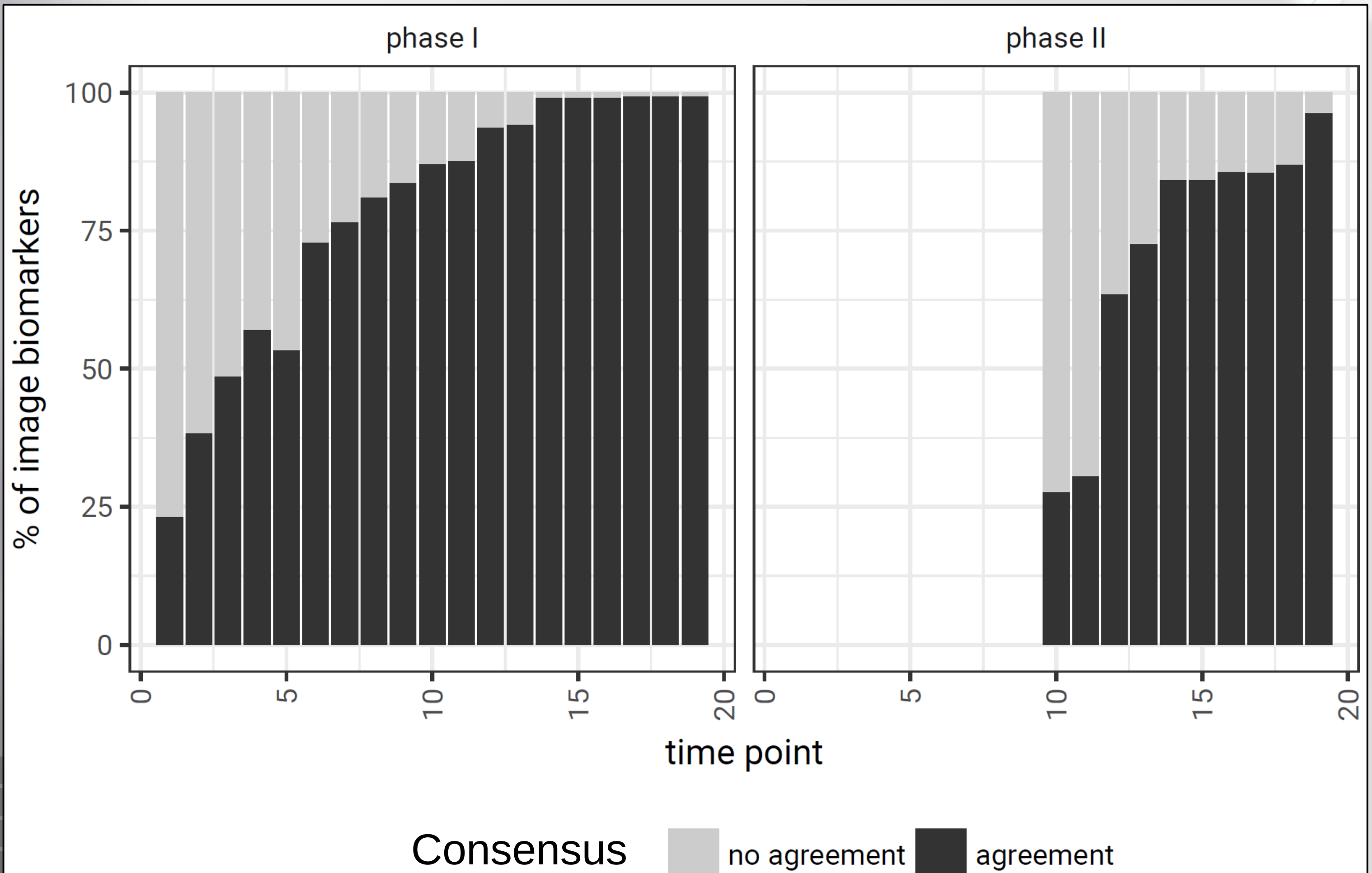
University of Groningen, University Medical Center Groningen
Lisanne V. van Dijk, Jorn Beukinga, Nanna M. Sijtsema, Roel J.H.M. Steenbakkers, Ronald Boellaard

Image Biomarker Standardisation Initiative. **Multicentre initiative for standardization of image biomarkers**. https://arxiv.org/abs/1612.07003

Zwanenburg, *et al*. **Standardized image biomarkers for high-throughput extraction of features from images**, *Nature Communications* (under review) 2018

Challenges and issues: lack of standardization

Imaging biomarkers standardisation initiative



Participants

Study leader: *Alex Zwanenburg*

Cardiff University
Philip Whybra, Emiliano Spezi

Dana Farber Cancer Institute and Brigham and Women's Hospital, Harvard University
Andriy Fedorov, Hugo Aerts

Gemelli ART, Università Cattolica del Sacro Cuore
Jacopo Lenkowicz, Luca Boldrini, Nicola Dinapoli, Vincenzo Valentini

German Cancer Research Center (DKFZ)
Michael Götz, Nils Gählert, Fabian Isensee, Klaus H. Maier-Hein

INSERM Brest, University of Brest
Marie-Charlotte Desseroit, Taman Upadhaya, Mathieu Hatt

Leiden University Medical Center
Floris H.P. van Velden

MAASTRO clinic, Maastricht University
Ralph T.H. Leijenaar, Philippe Lambin

McGill University
Martin Vallières, Issam El Naqa

Memorial Sloan Kettering Cancer Center
Aditya Apte

Moffitt Cancer Center
Mahmoud A. Abdalah, Robert Gillies

OncoRay – National Center for Radiation Research in Oncology and NCT Dresden
Alex Zwanenburg, Stefan Leger, Esther Troost, Christian Richter, Steffen Löck

The Netherlands Cancer Institute (NKI)
Joost van Griethuysen, Cuong Viet Dinh, Uulke van der Heide

Universitätsklinikum Tübingen, Eberhard Karls University Tübingen
Jairo Socarras Fernandez, Daniela Thorwarth

University Hospital Zürich, University of Zürich
Marta Bogowicz, Stephanie Tanadini-Lang, Matthias Guckenberger

University of Bergen
Are Losnegård

University of California, San Francisco
Olivier Morin

University of Groningen, University Medical Center Groningen
Lisanne V. van Dijk, Jorn Beukinga, Nanna M. Sijtsema, Roel J.H.M. Steenbakkers, Ronald Boellaard

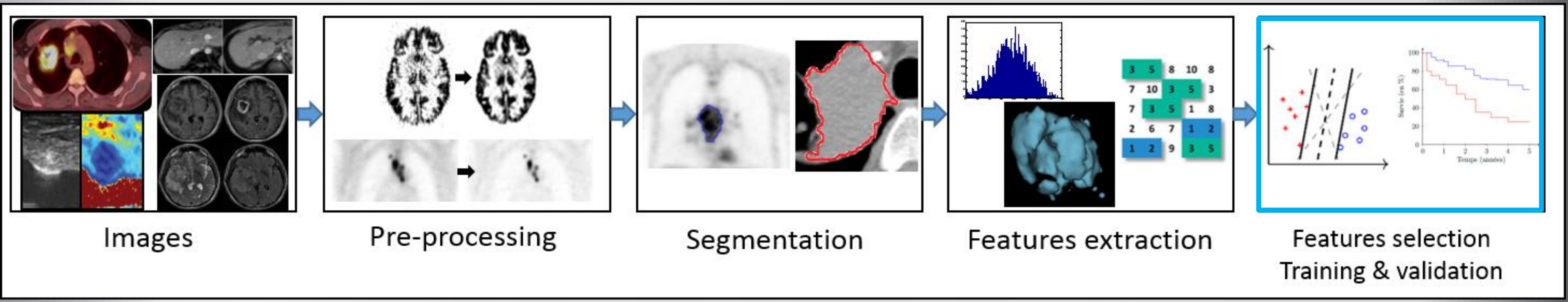
Image Biomarker Standardisation Initiative. **Multicentre initiative for standardization of image biomarkers.** **https://arxiv.org/abs/1612.07003**

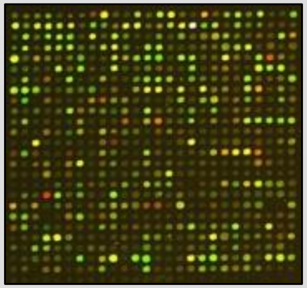
Zwanenburg, *et al*. **Standardized image biomarkers for high-throughput extraction of features from images**, *Nature Communications* (under review) 2018

Radiomics

Worfklow

Clinical data





Genomics (and other –omics)

Inappropriate statistical analysis

Table 1. Statistical characteristics of the selected studies divided in three categories: A) Studies with multiple hypotheses testing only, B) studies employing both multiple hypothesis testing and the optimum cut-off approach and C) studies with multiple hypothesis testing, with or without the optimum cut-off approach, but with validation analysis.

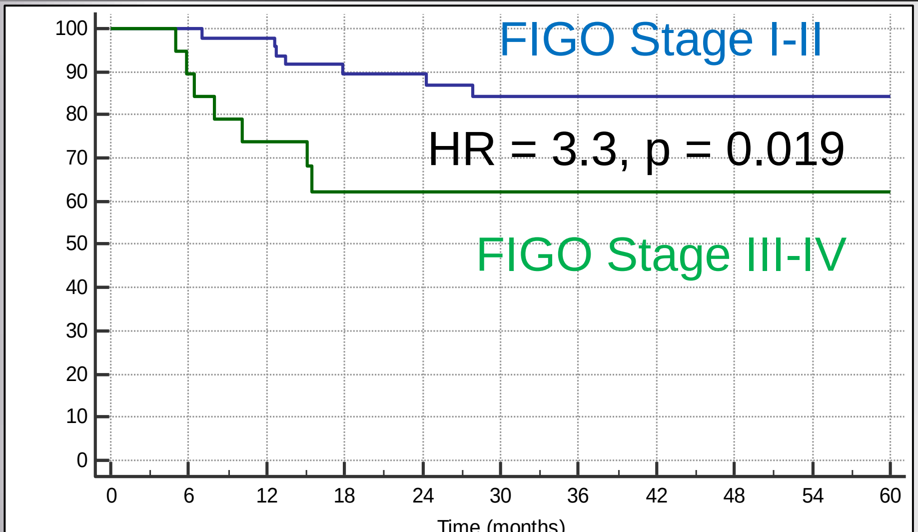
Category	Study	Multivariate analysis included volume	Optimum cut-off	Type I error adjustment	Validation dataset	cross correlation reported	Sample size	Hypotheses tested
A	Willaime [19]	Not applicable	No/Mean	No	No	Yes	12	68
	El Naqa [31]	NI*	Not clear	No	No	No	14/9	19
	Tixier [33]	NI	Not clear	No	No	Yes	41	54
	Yip [41]	No	No/Median	Yes#	No	No	36	90
B	Miles [30]	No	Yes	No	No	No	48	10
	Goh [32]	No	Yes	No	No	No	39	24
	Cook [29]	No	Yes	No	No	Yes	53	30
	Ganeshan [28]	No	Yes	No	No	Yes	21	15
	Ganeshan [34]	No	Yes	No	No	No	54	8
	Ng [36]	No	Yes	No	No	Yes	55	25
	Zhang [40]	Yes	Yes	No	No	No	72	40
	Cheng [39]	Yes	Yes	No	No	Yes	70	59‡
C	Vaidya [35]	Yes	No	No	LOOCV†	No	27	102
	Win [37]	No	Yes	No	Yes	No	66	12
	Ravanelli [38]	No	No/Median	No	LOOCV	No	53	16

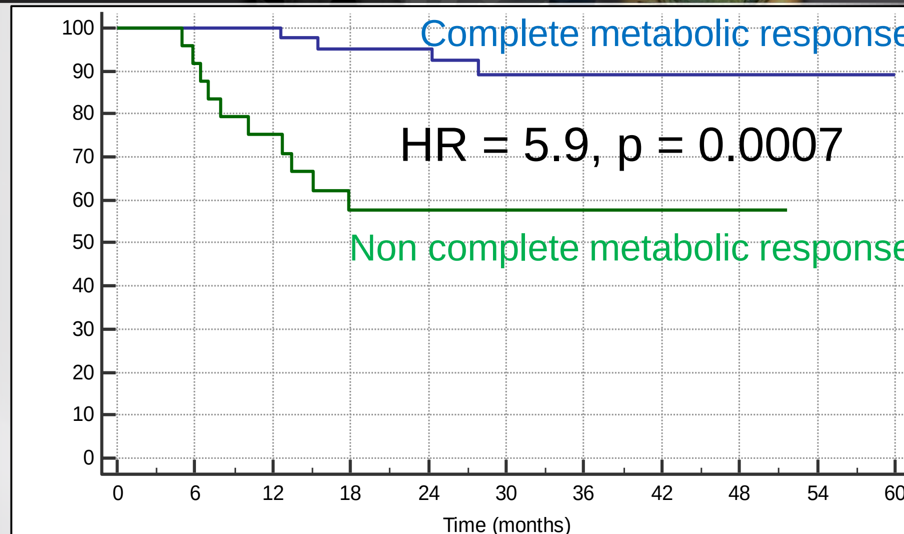
* No information provided
#For multiple hypotheses tested
†Leave one out cross validation
‡ Number is a conservative approximation due to the difficulty establishing the exact number of hypotheses tested

Chalkidou, *et al*. **False Discovery Rates in PET and CT Studies with Texture Features: A Systematic Review**. *PLoS One*. 2015

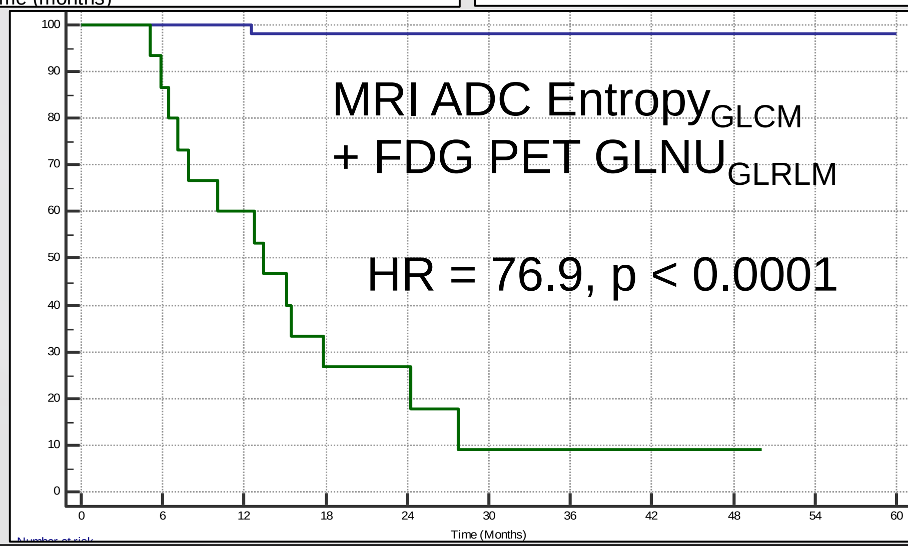
Radiomics

Challenges and issues: statistical analysis





Cervical cancer
Chemoradiotherapy
Loco-regional control



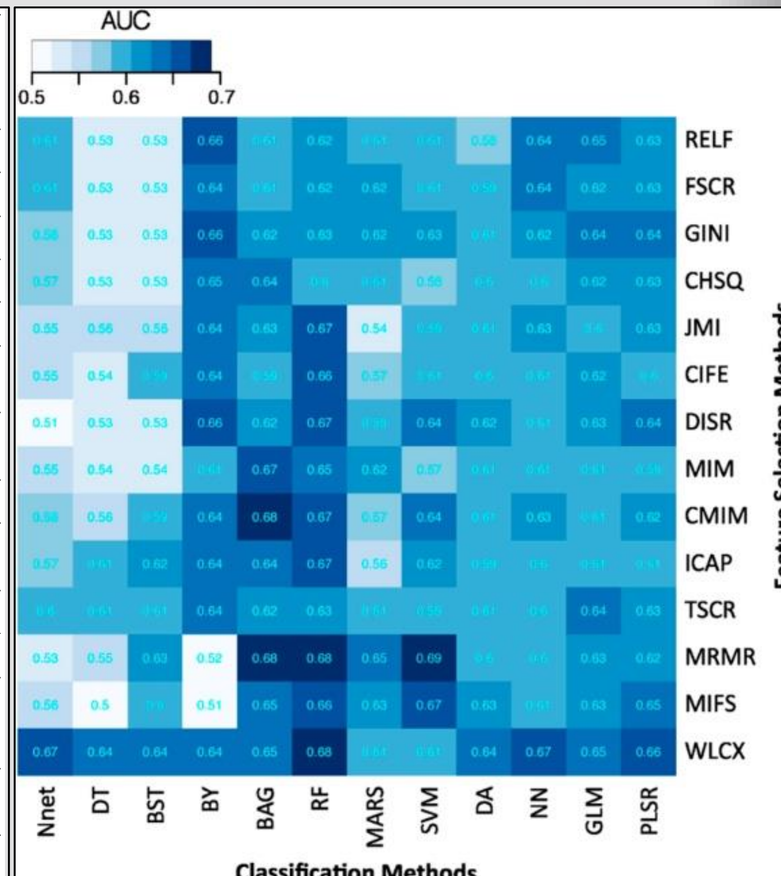
Lucia, *et al*. **Prediction of outcome using pretreatment ^{18}F-FDG PET/CT and MRI radiomics in locally advanced cervical cancer treated with chemoradiotherapy**. *Eur J Nucl Med Mol Imaging*. 2018

Challenges and issues: how to use machine learning?

Machine learning

- Choosing a classifier/feature selection method?

Classification method acronym	Classification method name	Feature Selection method acronym	Feature selection method name
Nnet	Neural network	RELF	Relief
DT	Decision Tree	FSCR	Fisher score
BST	Boosting	GINI	Gini index
BY	Bayesian	CHSQ	Chi-square score
BAG	Bagging	JMI	Joint mutual information
RF	Random Forset	CIFE	Conditional infomax feature extraction
MARS	Multi adaptive regression splines	DISR	Double input symmetric relevance
SVM	Support vector machines	MIM	Mutual information maximization
DA	Discriminant analysis	CMIM	Conditional mutual information maximization
NN	Neirest neighbour	ICAP	Interaction capping
GLM	Generalized linear models	TSCR	T-test score
PLSR	Partial least squares and prinicipal componenet regression	MRMR	Minimum redundancy maximum relevance
—	—	MIFS	Mutual information feature selection
—	—	WLCX	Wilcoxon



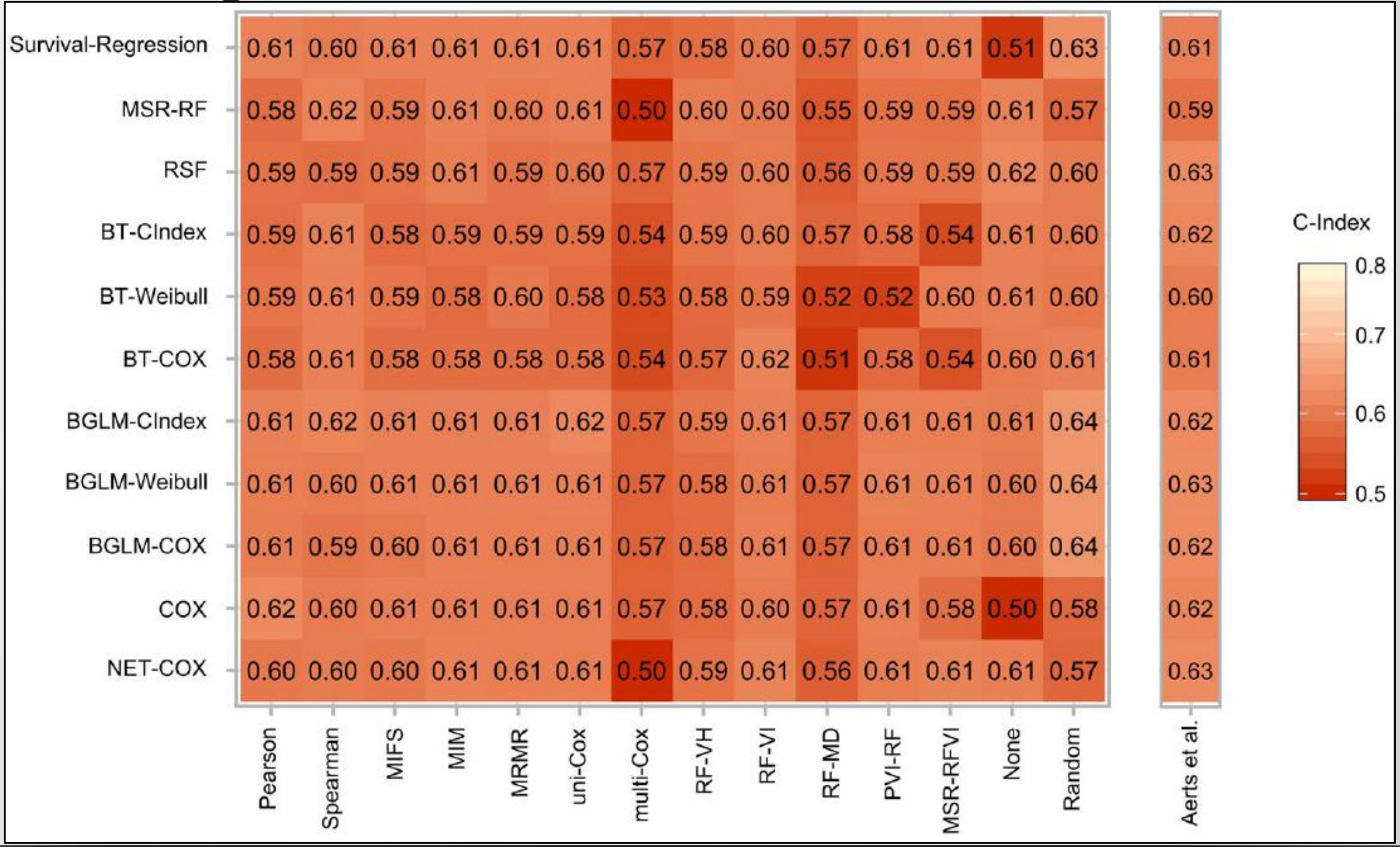
Parmar, *et al*. **Machine Learning methods for Quantitative Radiomic Biomarkers**. *Sci Rep.* 2015
Keger, *et al*. **A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling**. *Sci Rep* 2017

Radiomics

Challenges and issues: how to use machine learning?

Machine learning

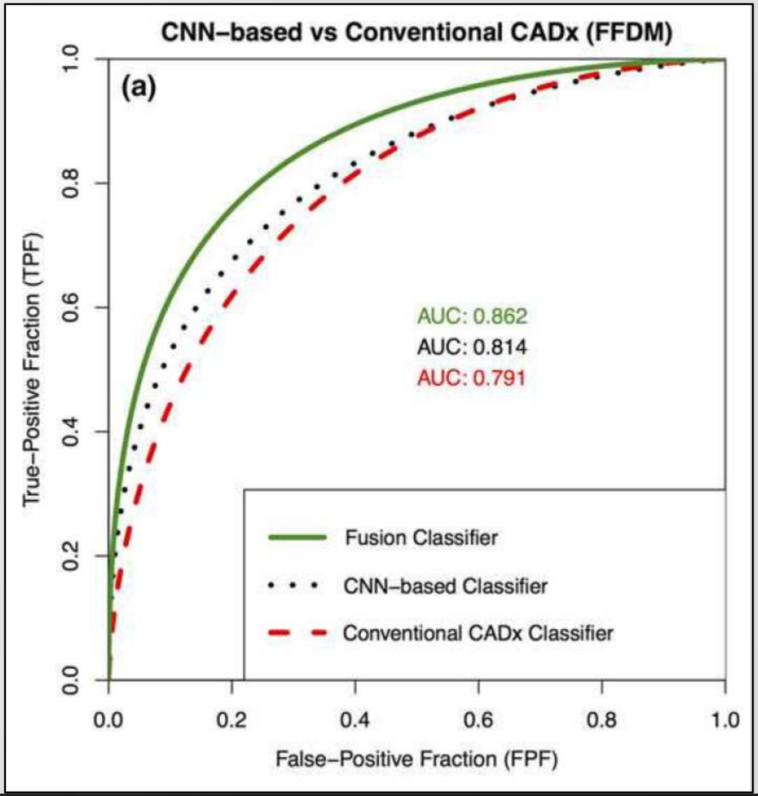
- Choosing a classifier/feature selection method?



Parmar, *et al*. **Machine Learning methods for Quantitative Radiomic Biomarkers**. *Sci Rep.* 2015
Keger, *et al*. **A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling**. *Sci Rep* 2017

Machine learning

- Choosing a classifier/feature selection method?
  - Potential solution: ensemble methods

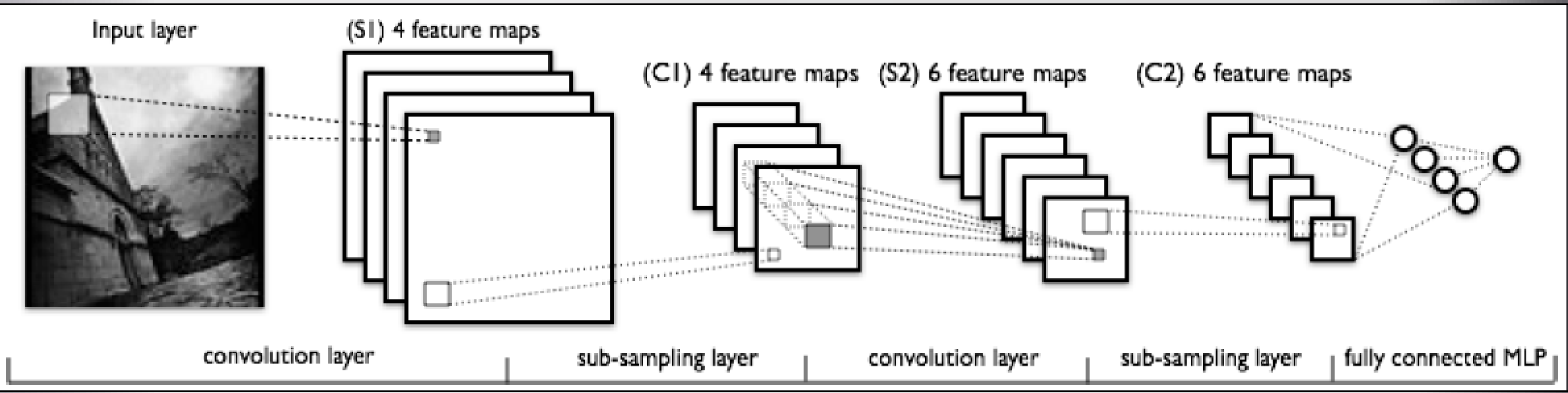


Antropova, *et al*. **A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets**. *Med Phys* 2017

Perspectives: potential of deep learning?

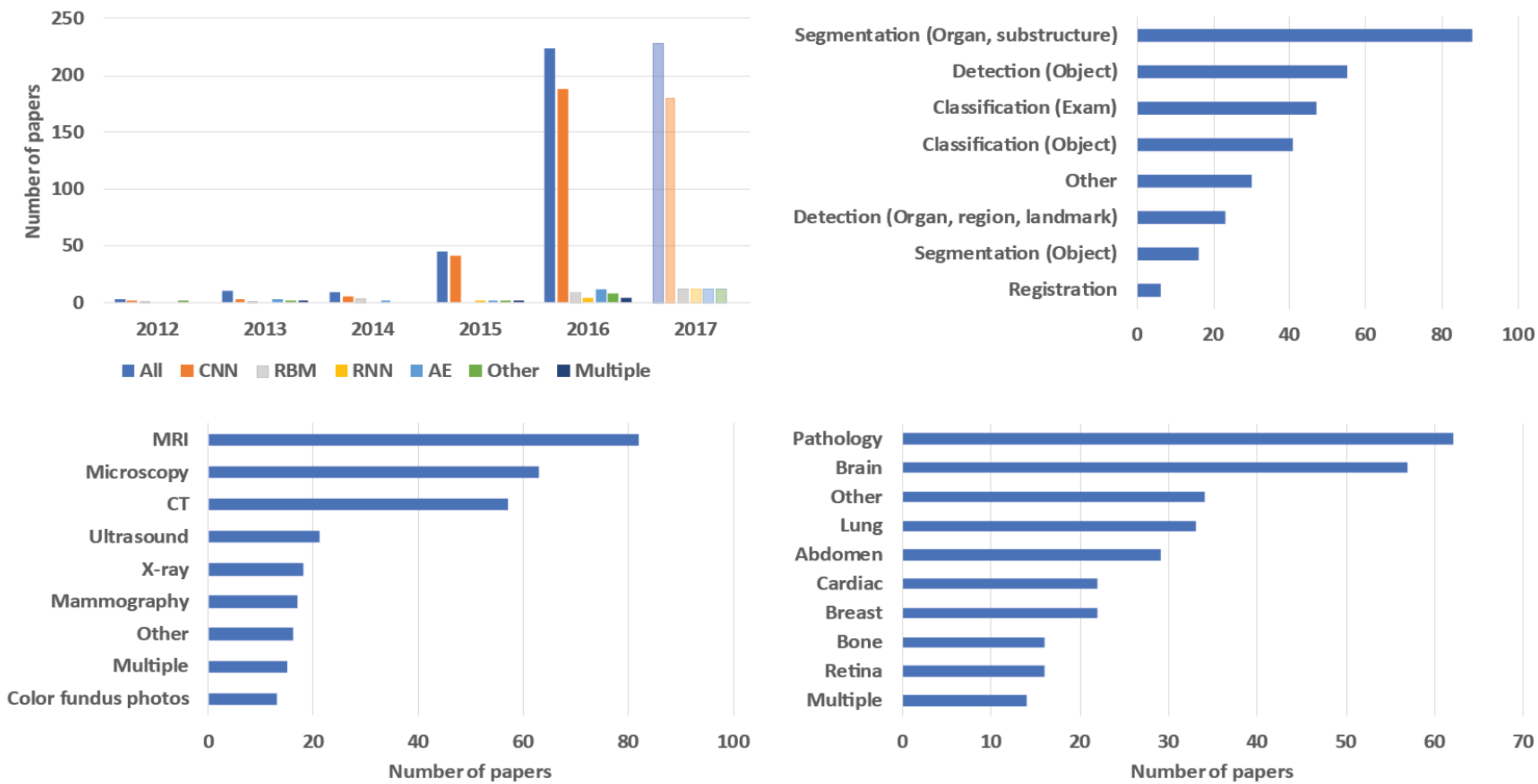
Deep learning

- Convolutional Neural Networks (CNN)



- Limitations (*a priori*)
  - Need (very) large datasets for efficient training
  - Black boxes that do not generate knowledge

Deep learning in medical imaging

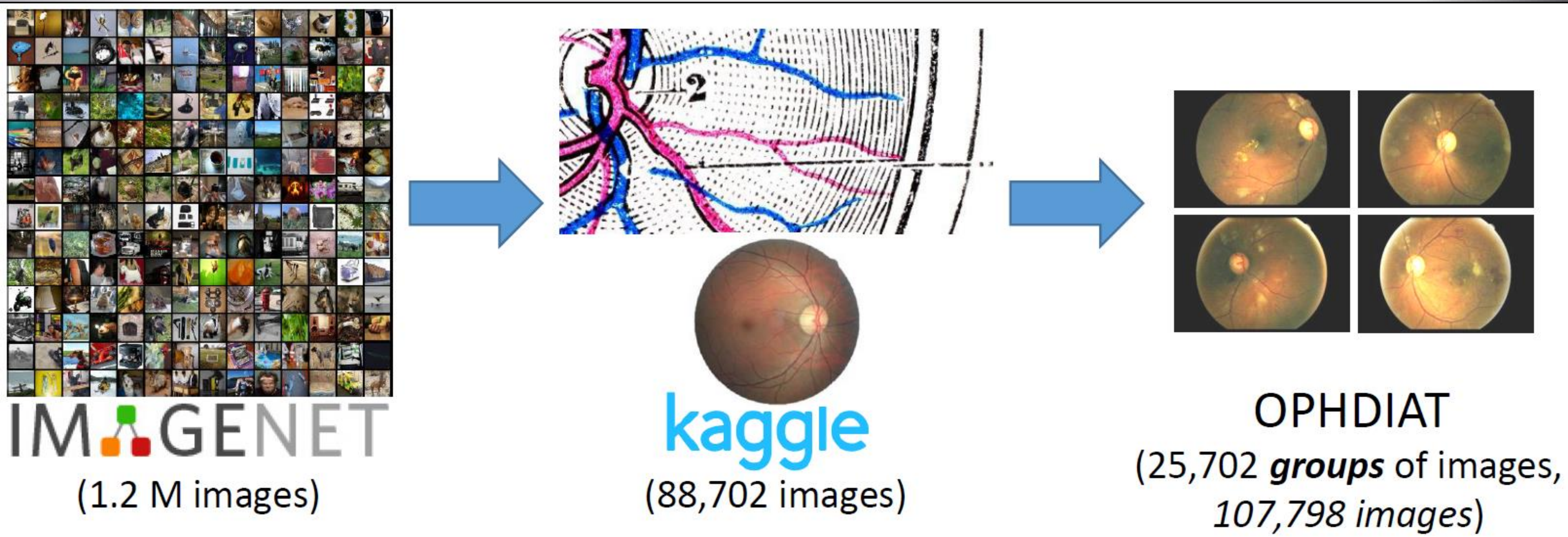


Litjens, *et al*. **a survey on deep learning in medical image analysis**. *Med Image Anal.* 2017

Radiomics

Perspectives: potential of deep learning?

- Deep learning limitations?
  - Need for large datasets
    - Data augmentation
    - Transfer learning / fine-tuning



Quellec, *et al*. **Deep image mining for diabetic retinopathy screening**. *Med Image Anal*. 2017

Radiomics

Perspectives: potential of deep learning?

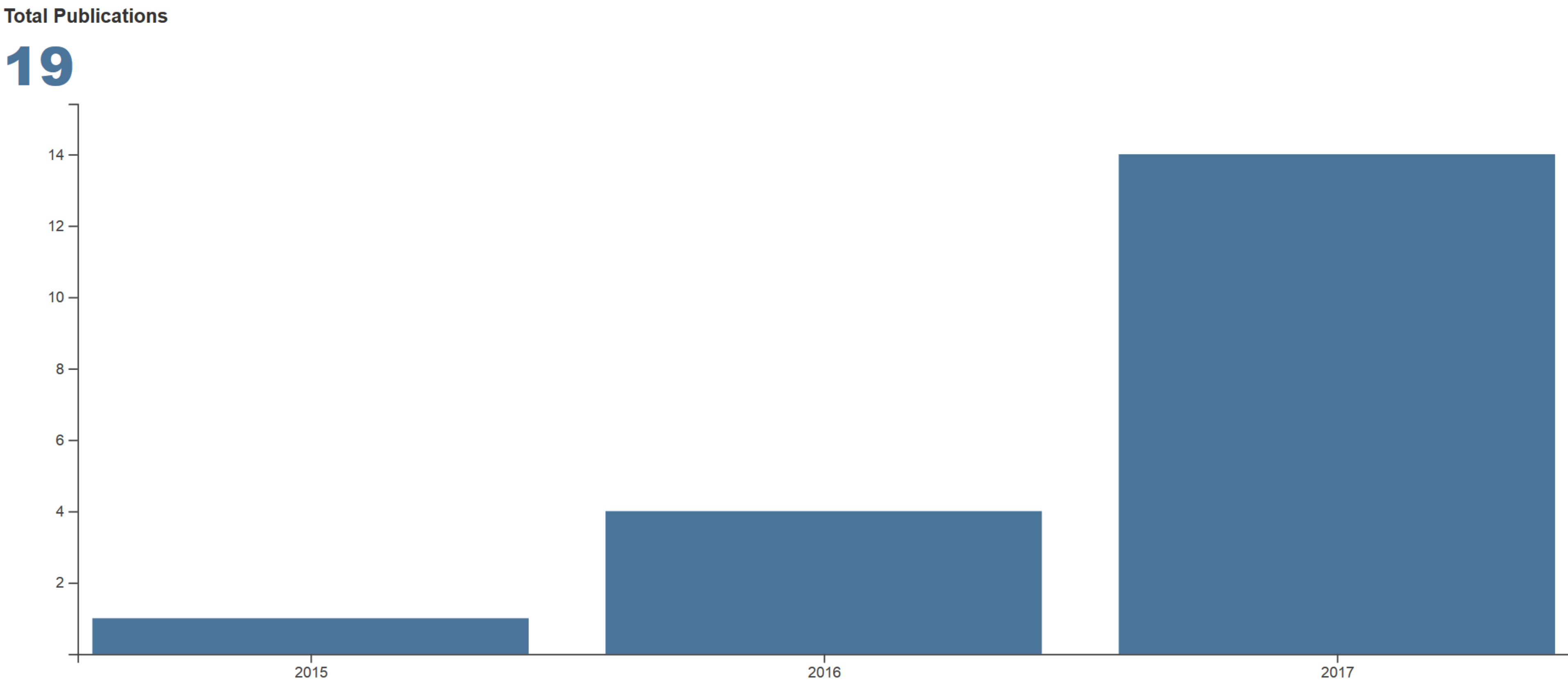
Deep learning limitations?

- Need for large datasets
  - Data augmentation
  - Transfer learning / fine-tuning
- Black boxes / knowledge generation
  - Networks visualization
  - Back propagation to exploit networks

Quellec, *et al*. **Deep image mining for diabetic retinopathy screening**. *Med Image Anal*. 2017

Perspectives: potential of deep learning?

Deep learning / CNN + radiomics

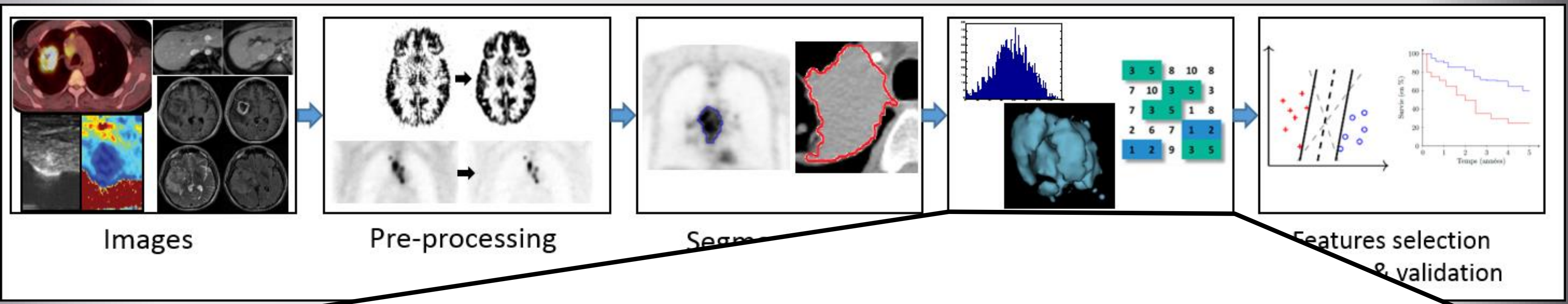


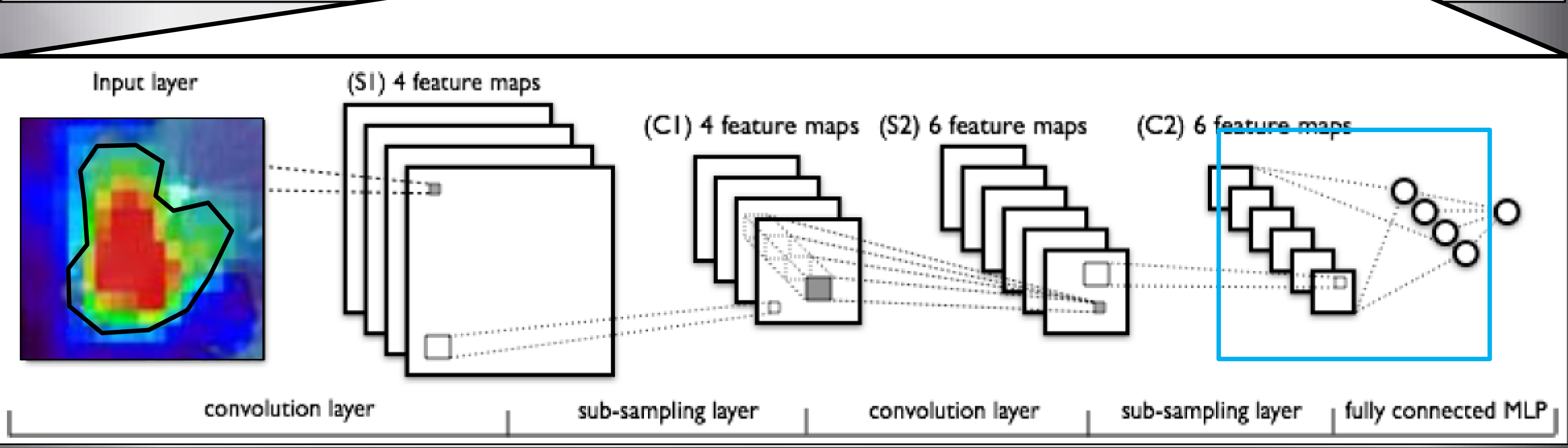
Source: web of science

Radiomics

Perspectives: potential of deep learning?

- Deep learning / CNN + radiomics

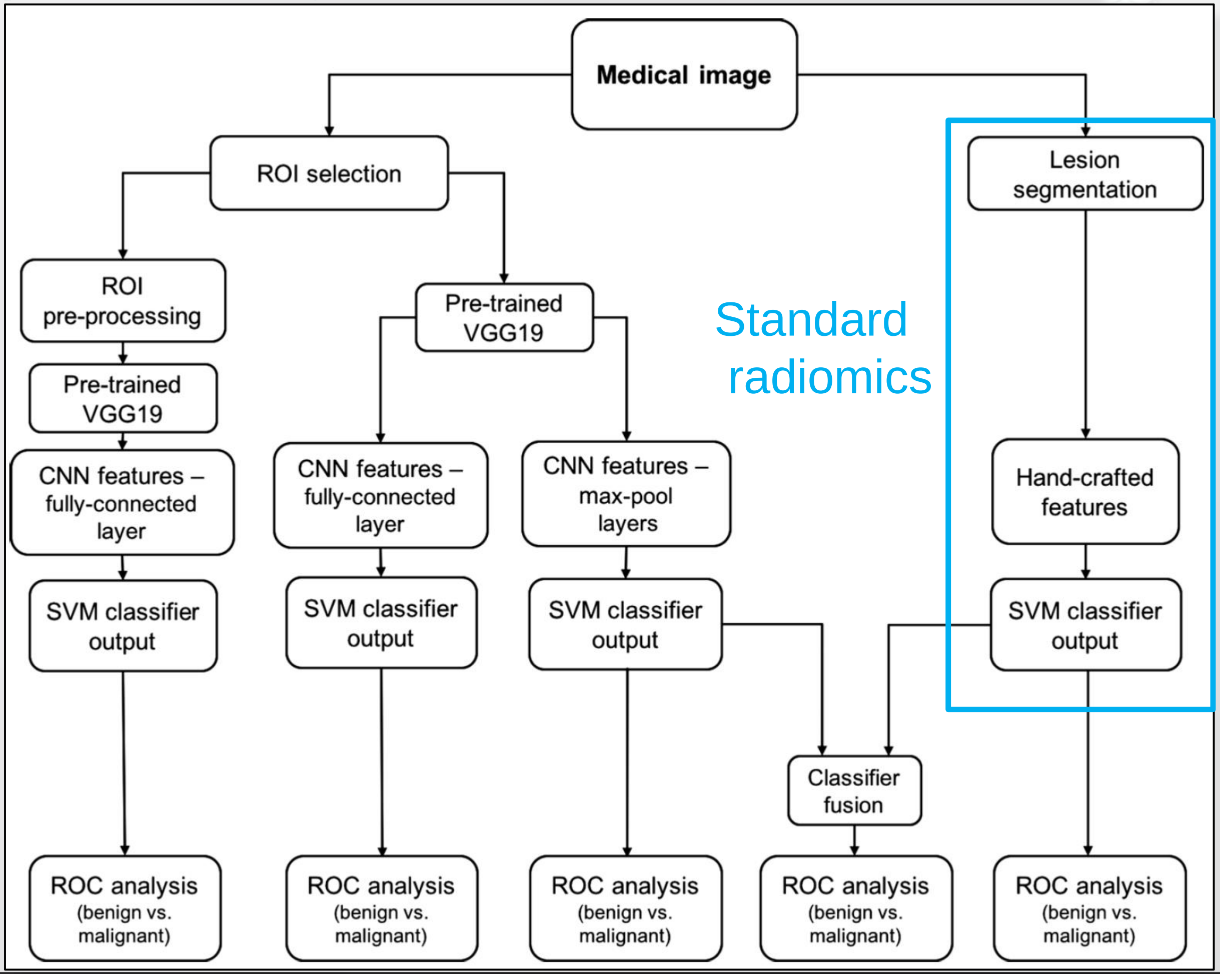




Radiomics

Perspectives: potential of deep learning?

- Deep learning / CNN + radiomics



Antropova, *et al*. **A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets**. *Med Phys* 2017

Radiomics

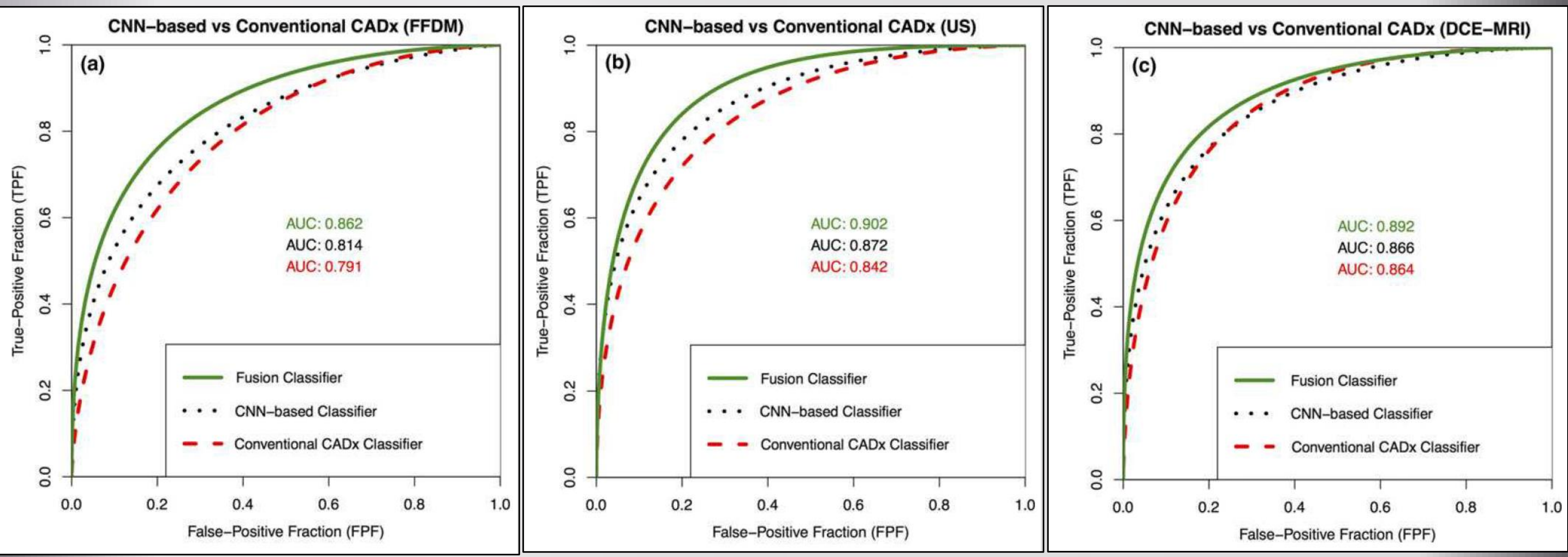
Perspectives: potential of deep learning?

Deep learning / CNN + radiomics

Full field digital mamography (FFDM)
N=245

Ultrasound (US)
N=1125

DCE-MRI
N=690

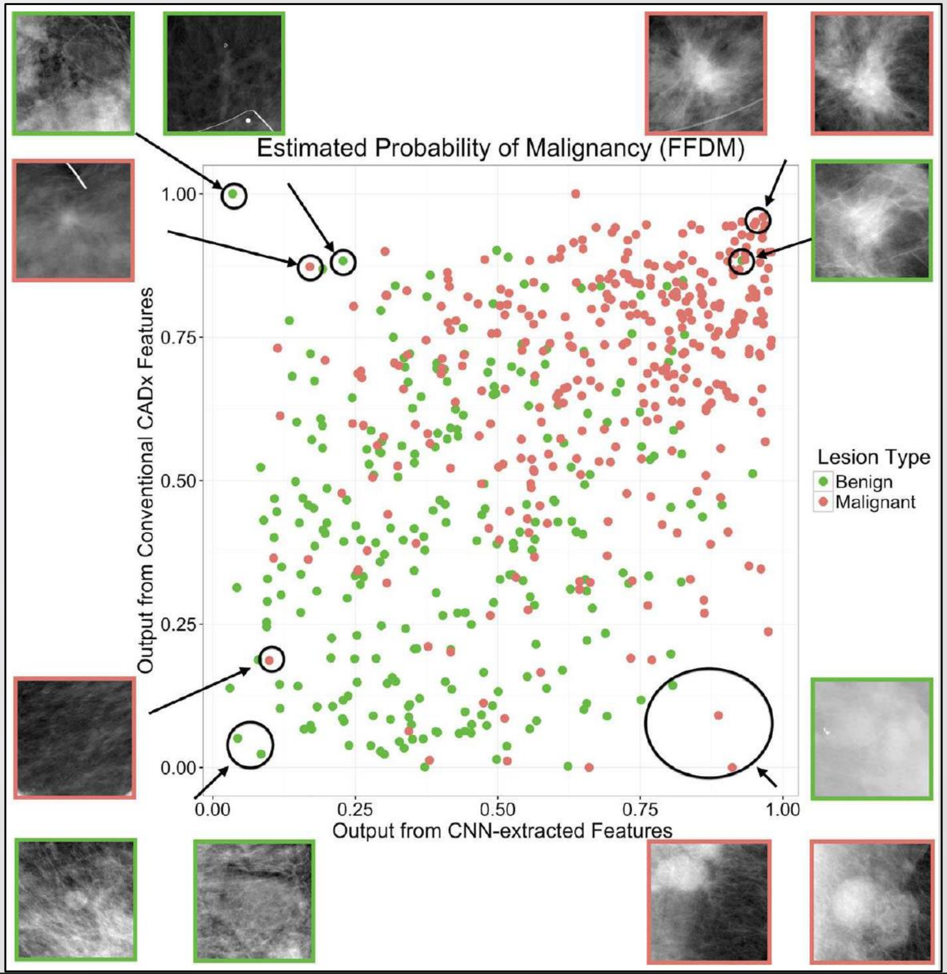


Antropova, *et al*. **A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets**. *Med Phys* 2017

Radiomics

Perspectives: potential of deep learning?

- Deep learning / CNN + radiomics

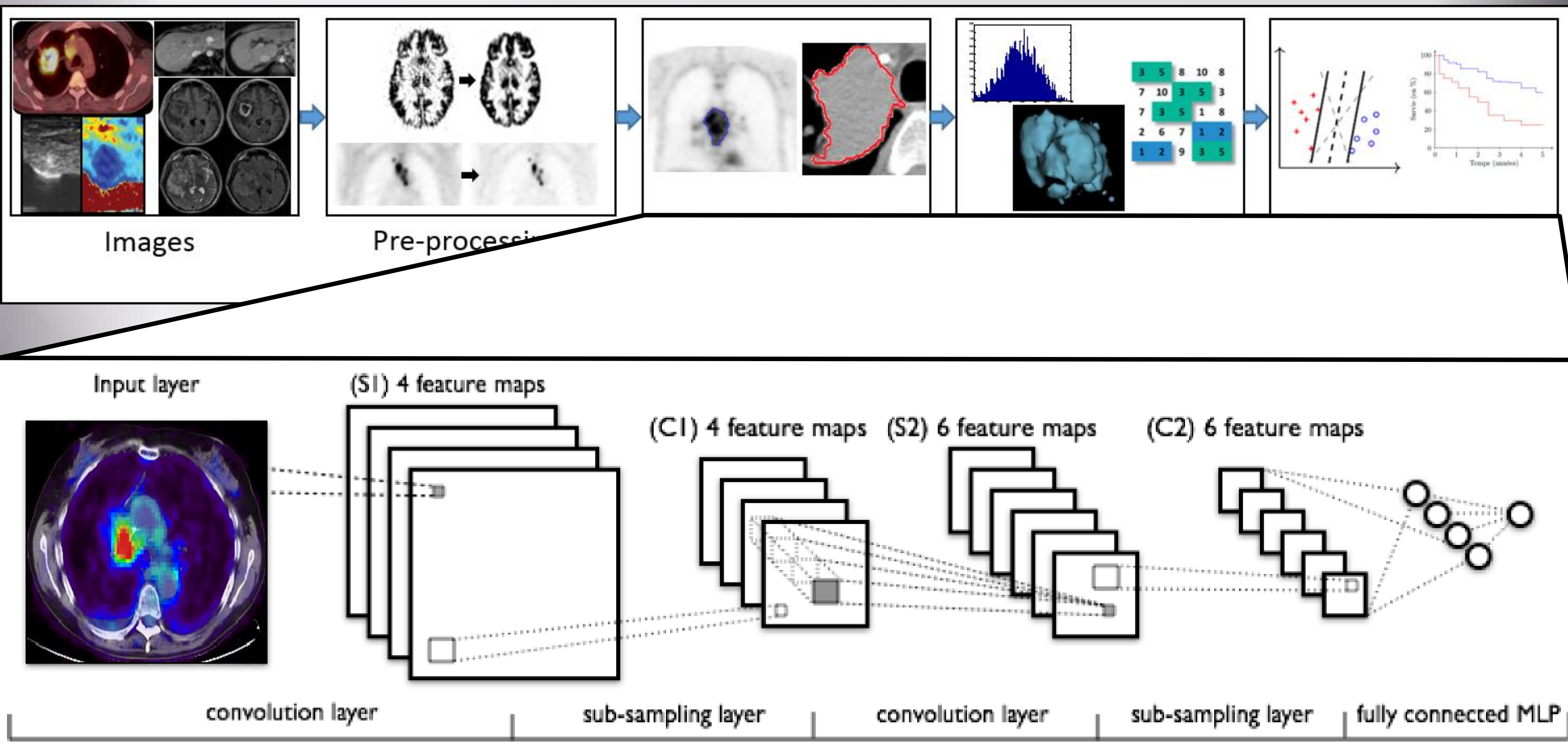


Antropova, *et al*. **A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets**. *Med Phys* 2017

Radiomics

Perspectives: potential of deep learning?

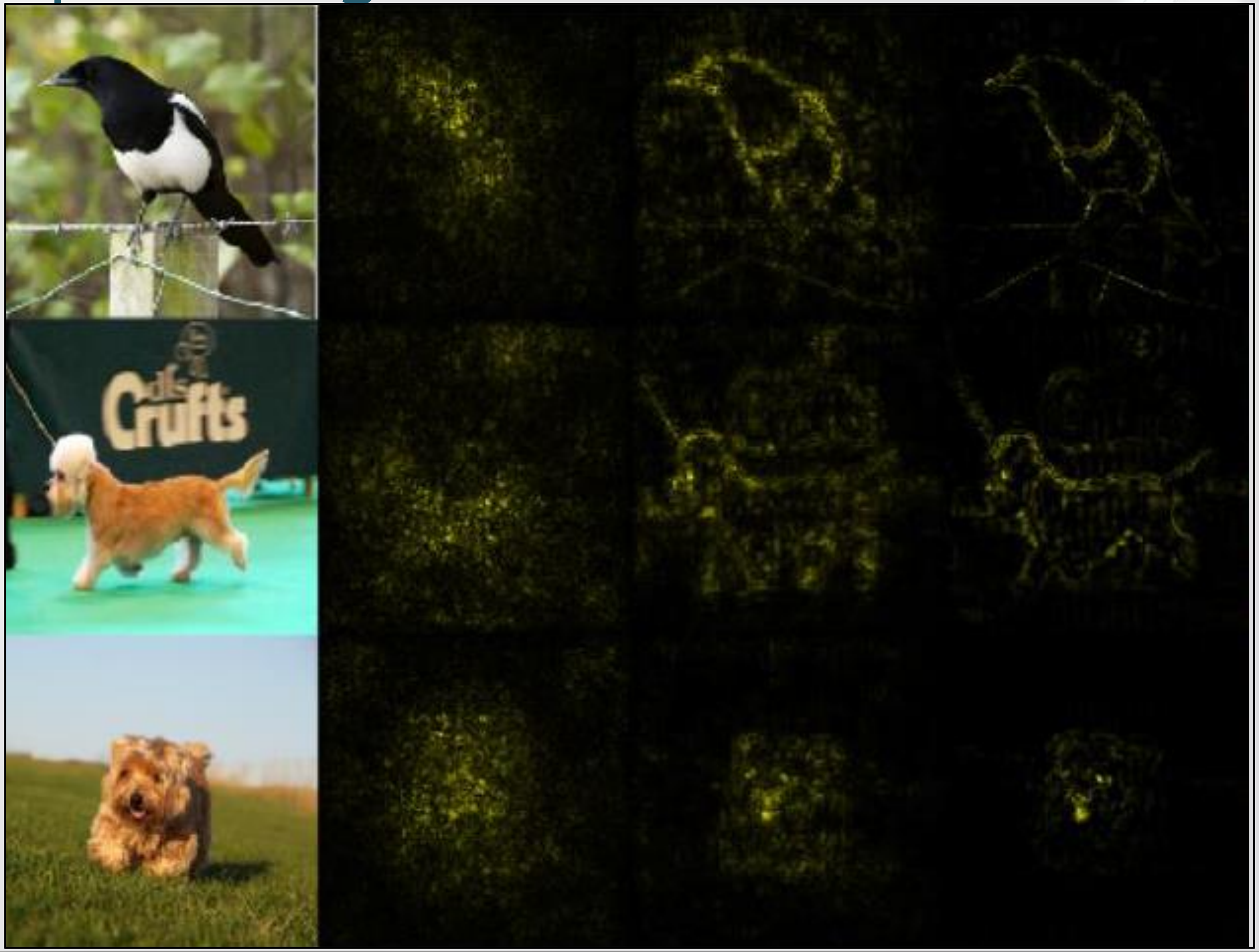
- Deep learning / CNN + radiomics



Radiomics

Perspectives: potential of deep learning?

- Deep learning / CNN + radiomics

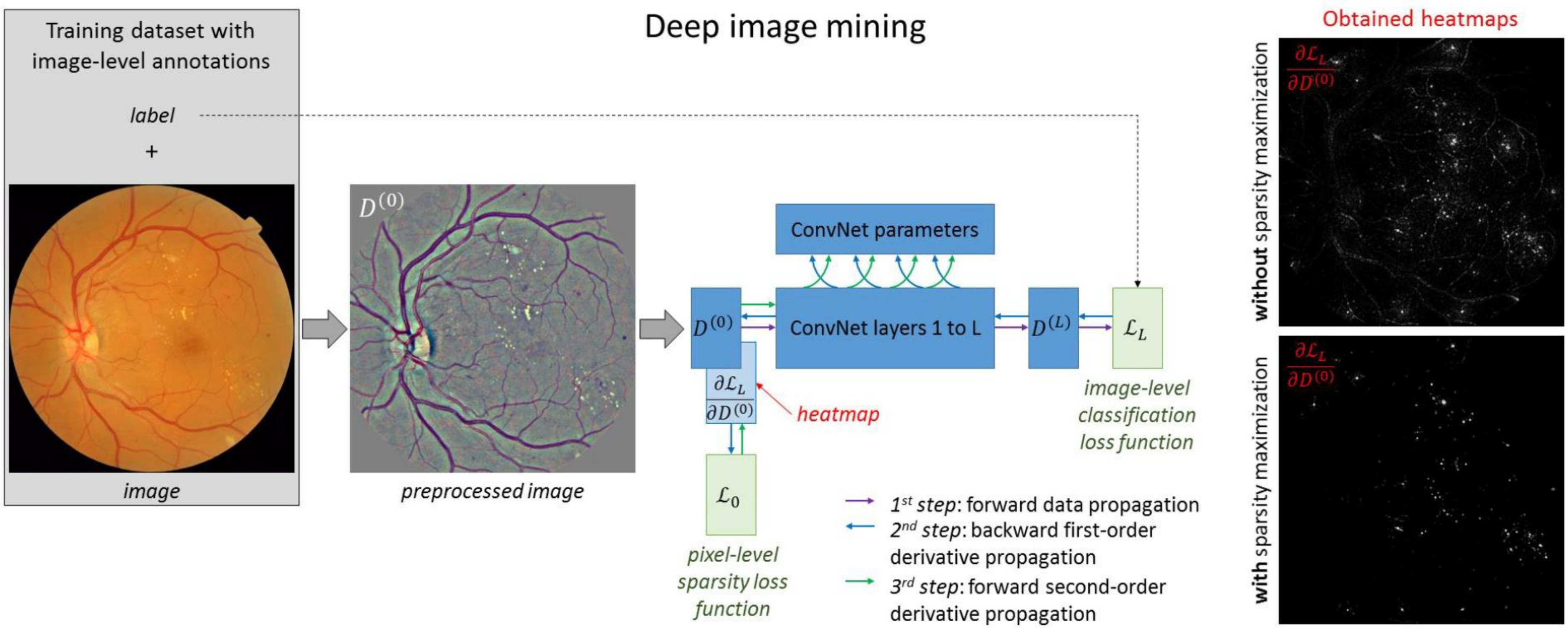


Samek, *et al*. **Evaluating the Visualization of What a Deep Neural Network Has Learned**. *IEEE Trans Neural Networks and Learning Systems*. 2017

Radiomics

Perspectives: potential of deep learning?

- Deep learning / CNN + radiomics
  - « Endpoint-guided » segmentation

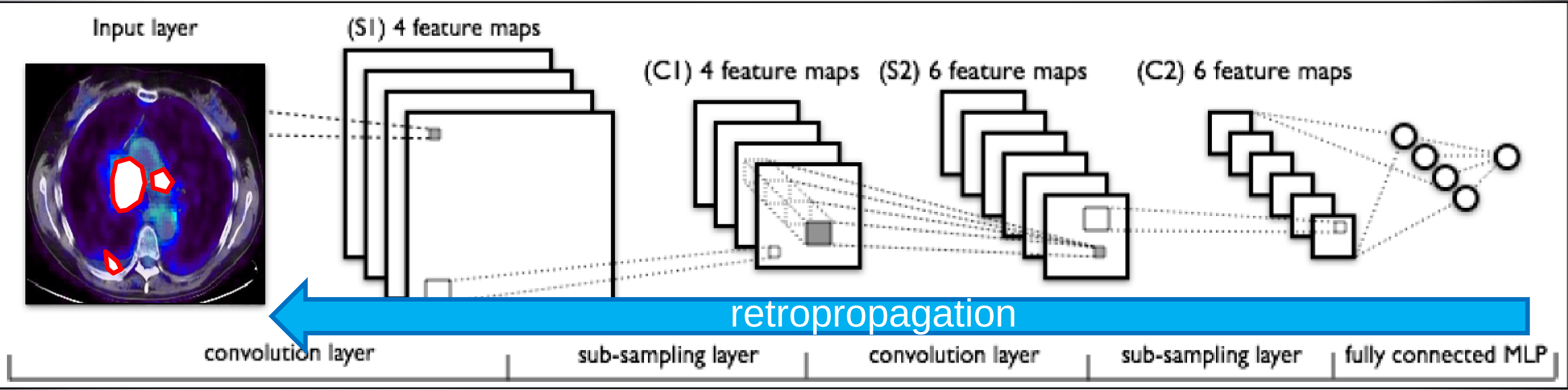


Quellec, *et al*. **Deep image mining for diabetic retinopathy screening**. *Med Image Anal*. 2017

Radiomics

Perspectives: potential of deep learning?

- Deep learning / CNN + radiomics
  - « Endpoint-guided » segmentation



Quellec, *et al.* **Deep image mining for diabetic retinopathy screening**. *Med Image Anal*. 2017

- Radiomics
  - Dynamic field of research
  - Numerous challenges and methodological issues
  - Lack of standardization (workflow, features)
  - Difficult statistical validation
- Potential solutions, perspectives
  - Larger, prospective, multicentric studies
  - Use robust machine learning methods (deep learning?)
  - Standardization of radiomics (ongoing)
  - Responsible research (share methods & data)

Thanks for your attention

Instituts thématiques Inserm

Institut national de la santé et de la recherche médicale