Predictive radiology: machine learning with medical images in clinical practice

Jonas Richiardi
jonas.richiardi@chuv.ch - http://richiardi.net

Department of Radiology

SIEMENS Healthineers

Advanced Clinical Imaging Technologies

Disclaimer

The information and views set out in this lecture are those of the author and do not necessarily reflect the official opinion of the CHUV or Siemens Healthineers.

Agenda

Context and introduction

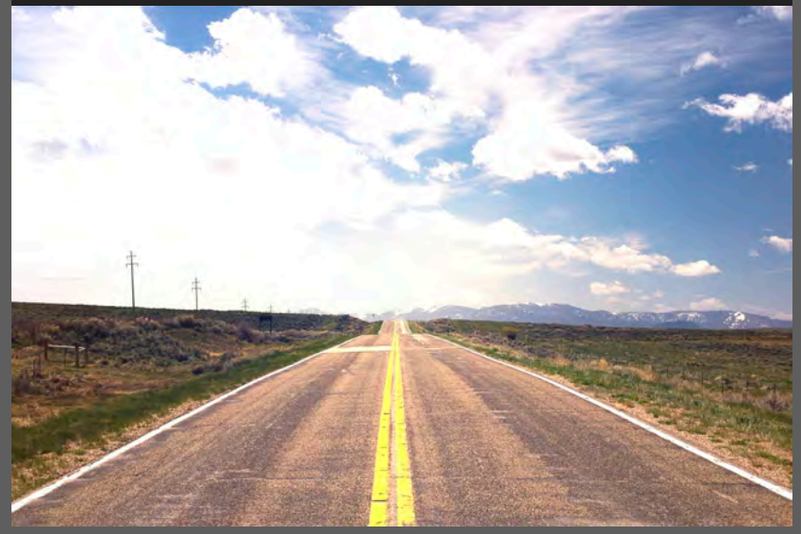
Methods and tools

Clinical applications

From radiology to precision medicine

Challenges

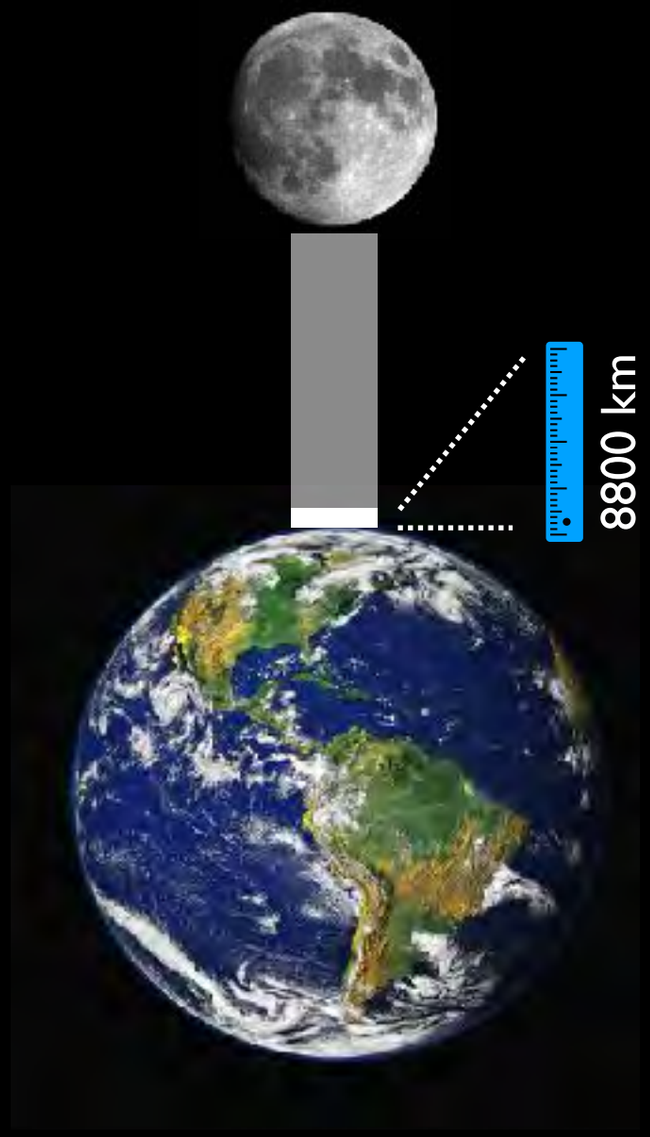
Conclusions



Context and introduction

Medical and imaging data are growing

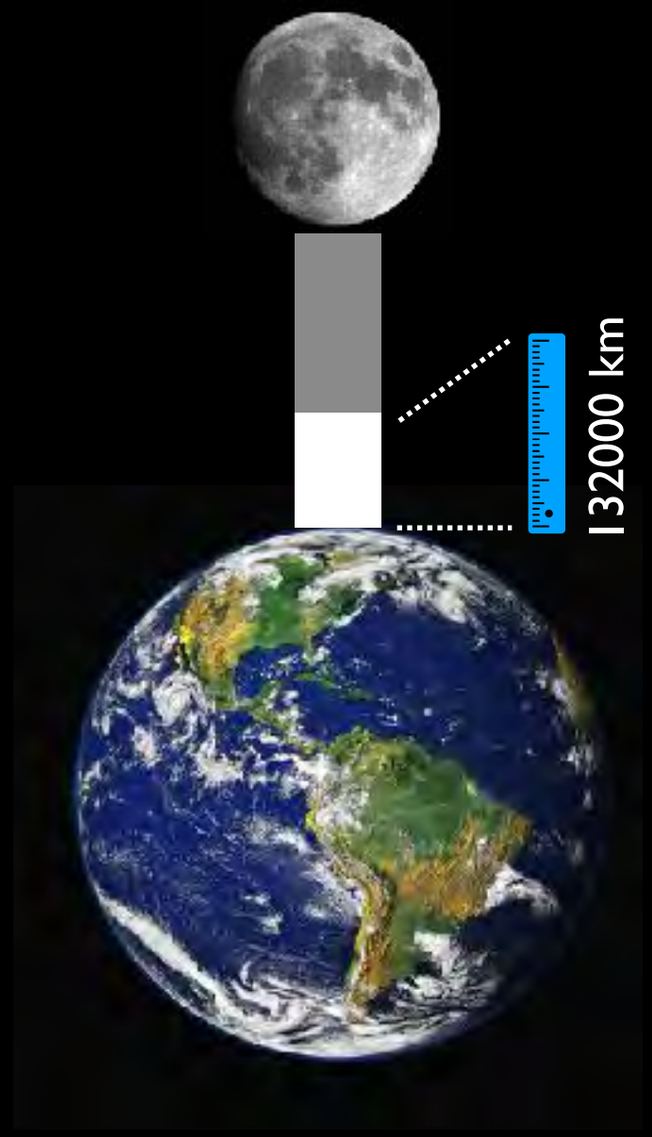
2013



153 ExaBytes

+48% p.a.

2020



2314 ExaBytes

inspired by [IDC/EMC, 2014]

Notable open data projects

NIH clinical center - chest CTs

30K+ subjects, 100K+ images

Medical ImageNet

1000 chest CTs, 831 bone tumor CTs,

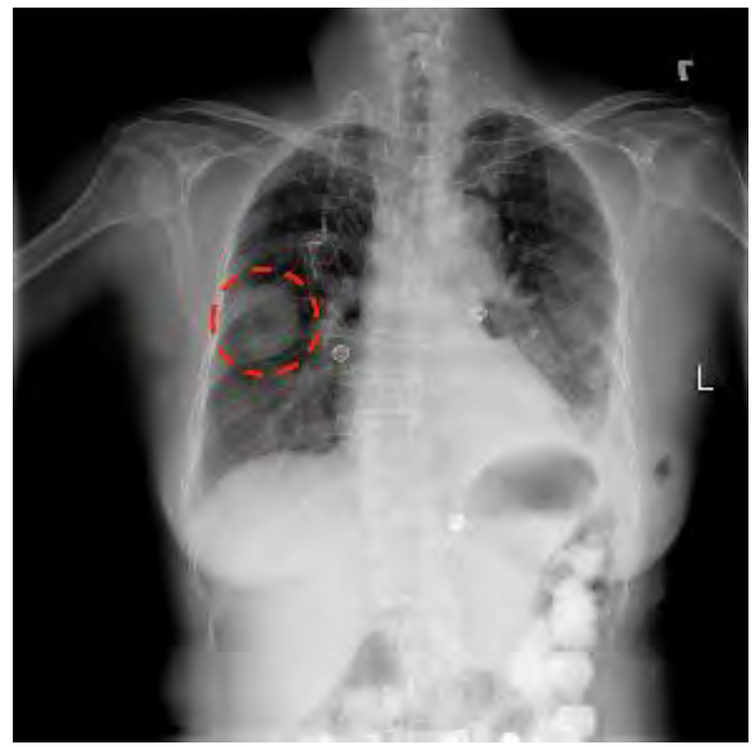
4K mammograms, 4K hand CTs

UK Biobank

Goal: 100K subjects with sMRI, rsfMRI, dwMRI, neuropsy, genotype, vitals etc.

Child Mind Institute Healthy brain network

Goal: 10K subjects with sMRI (T1, T2, qT1, qT2), rsfMRI, EEG, neuropsy, genotype, vitals etc.



[Wang et al-, CVPR 2017]

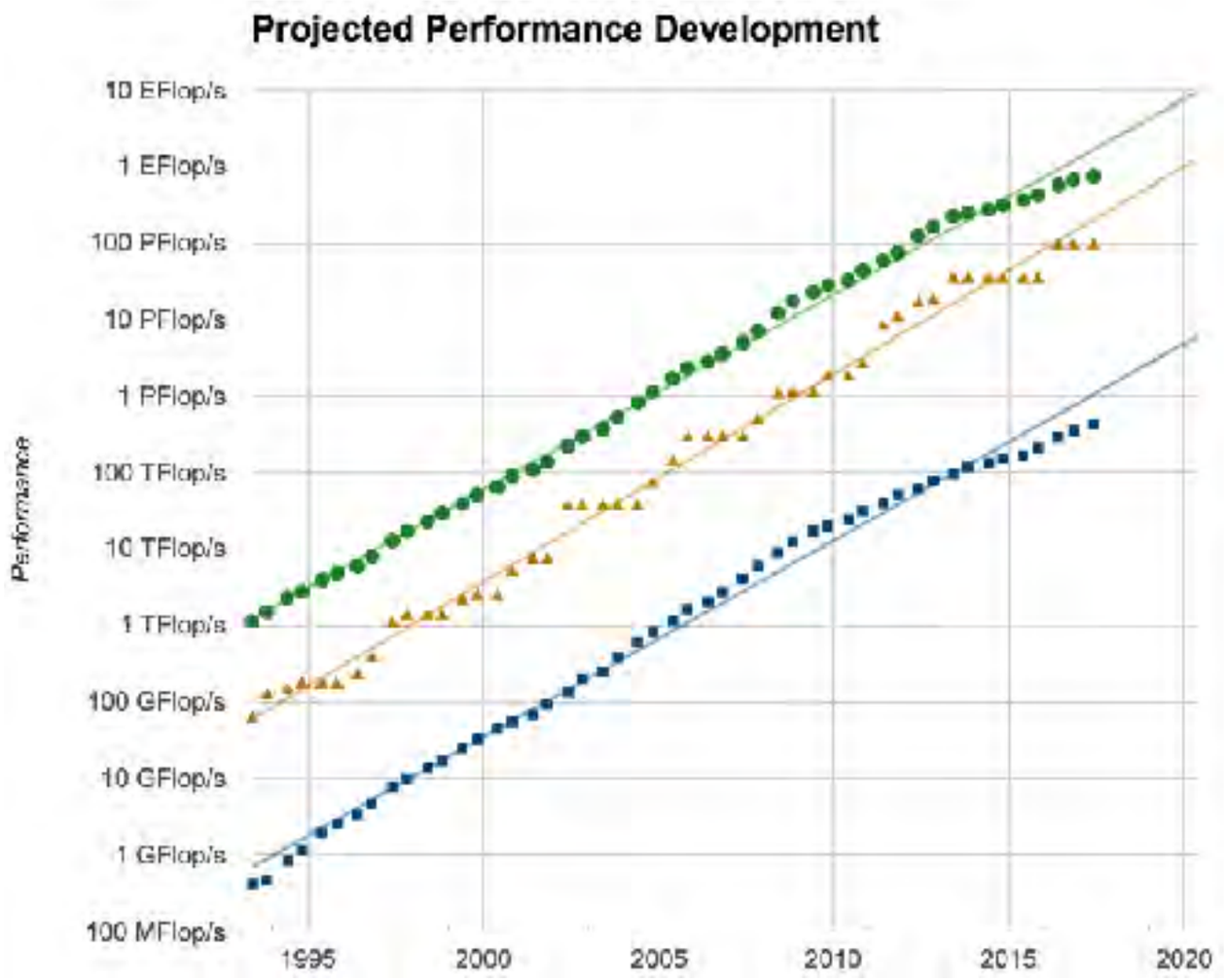
visceral

1609 CT/MRI annotated lesions

THE CANCER IMAGING ARCHIVE

31M images from 38K subjects

Computing power is constantly growing



← total
← #1

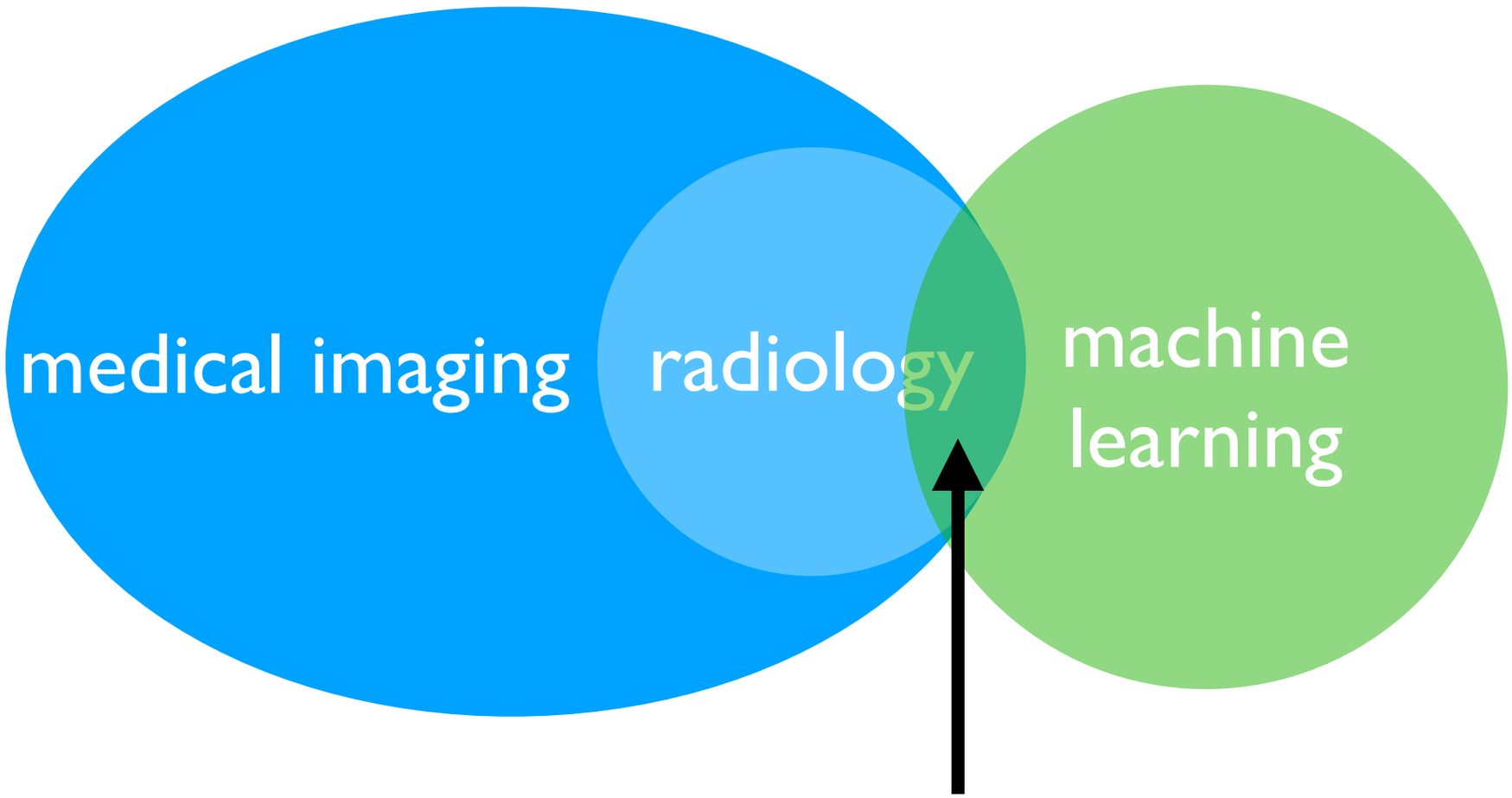
← #500

#1 Sunway TaihuLight, 93 PF
#3 Piz Daint, 19.6 PF
…
Radeon Vega (iMac Pro) 11 TF
PS4 1.8 TF
Intel Core i9 1 TF

Price is also going down constantly

<2 CHF/GFLOPS

Predictive radiology



predictive radiology

= data + algorithms + computing power + interpretability

Applications of predictive radiology

Diagnosis

CAD: abnormality detection, lesion segmentation etc

direct Dx

direct DDx

subtyping

Prognosis

clinical score change

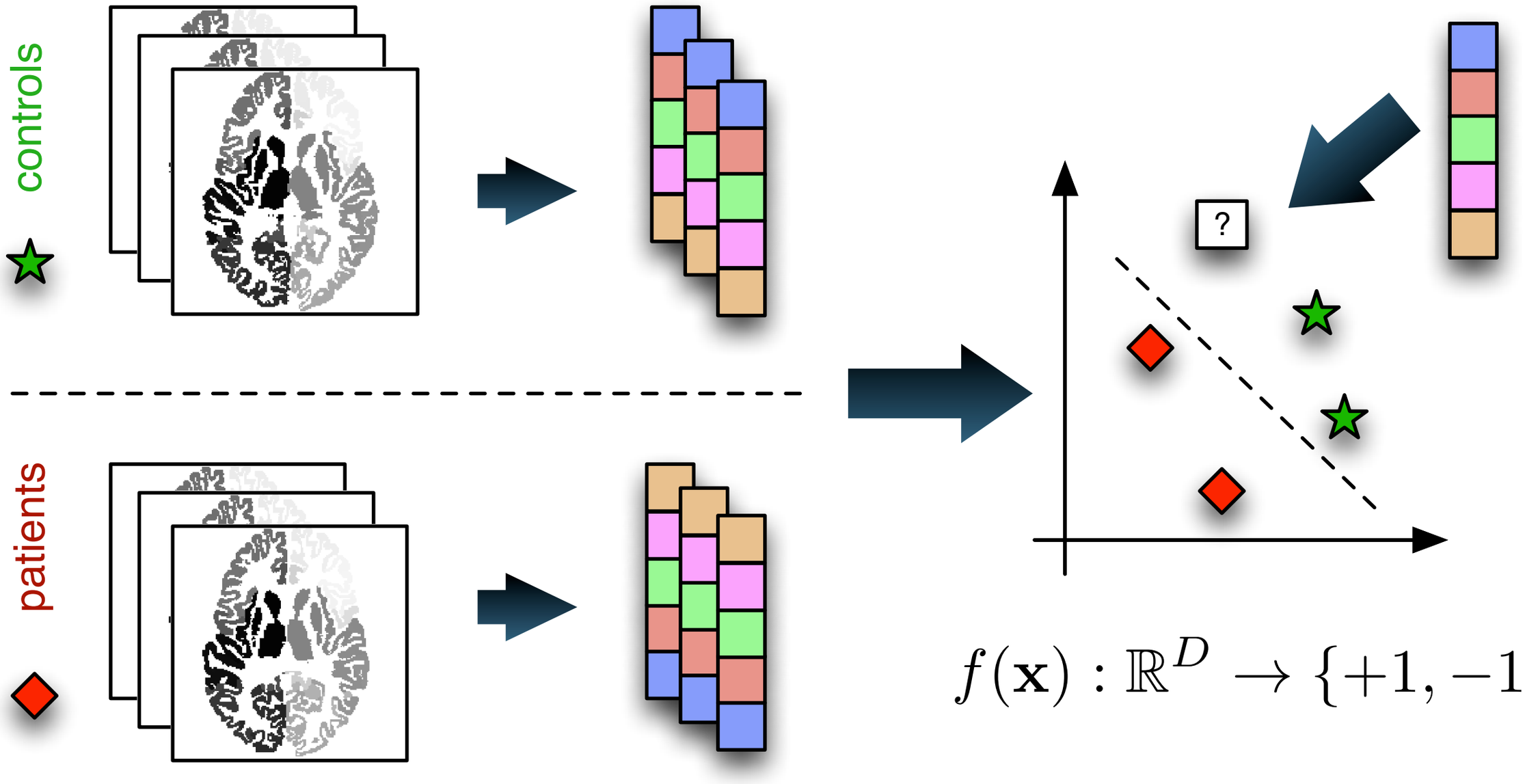
high-risk/low-risk stratification (care / trial enrichment)

Treatment planning

responder/non-responder

Methods and tools

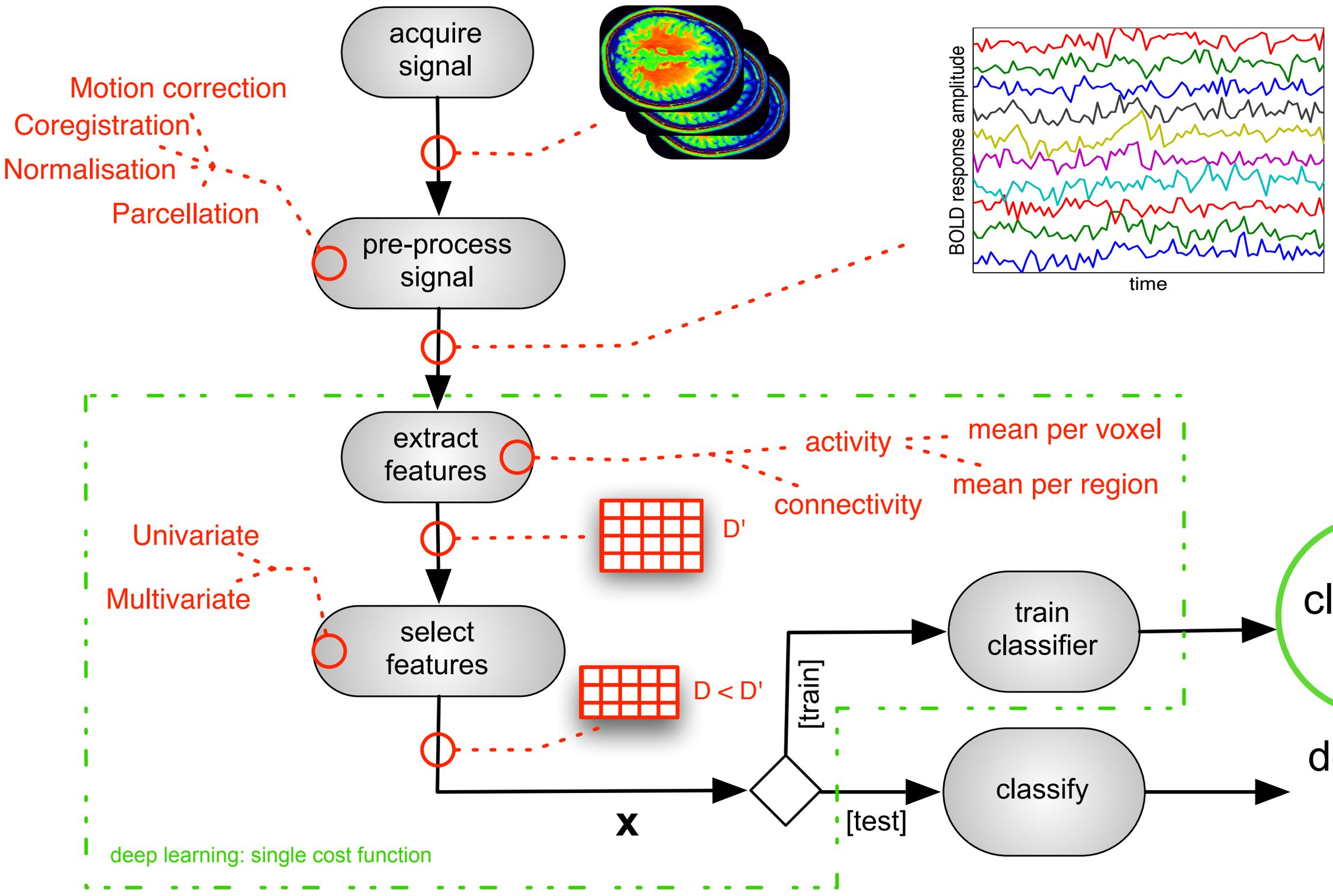
Overall principle



$$f(\mathbf{x}) : \mathbb{R}^D \rightarrow \{+1, -1\}$$

$\mathbf{y}$	$\tilde{\mathbf{x}}$	$\mathbf{x}$
$\{+1, -1\}$	$\mathbb{R}^V$	$\mathbb{R}^D,\ D < V$

Machine learning pipeline (fMRI)

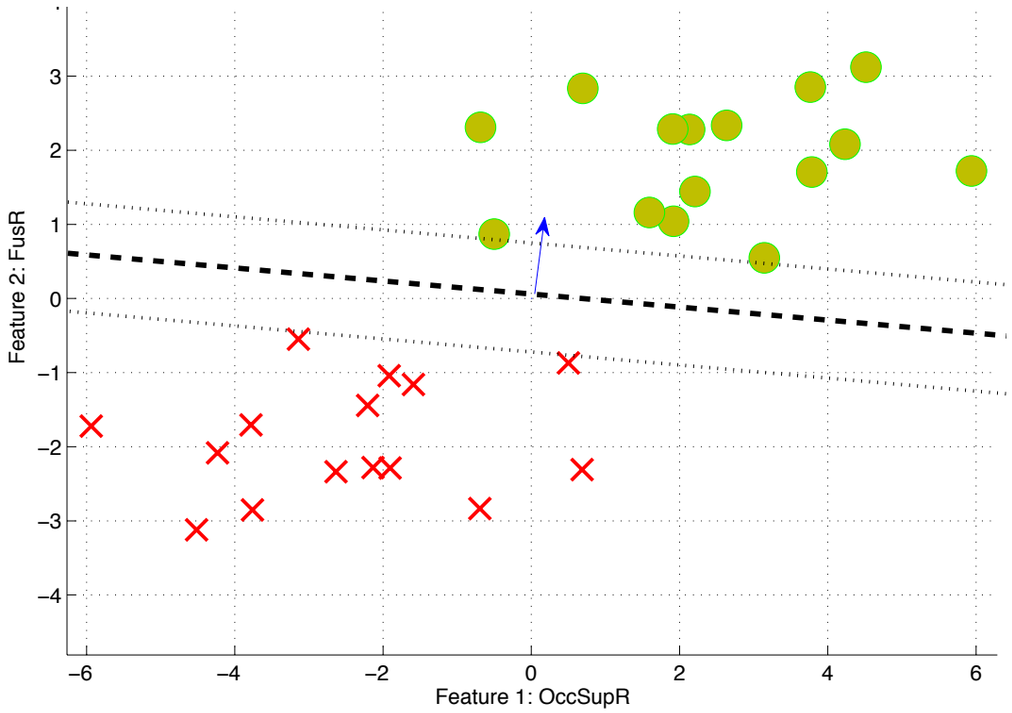


The linear support vector machine (SVM)

Parameters to optimise: **w** (normal vector to the separating hyperplane), b (offset)

Find: $\hat{\mathbf{w}}, \hat{b} = \underset{\mathbf{w},b}{\operatorname{argmin}} \frac{1}{2}||\mathbf{w}||^2$

With constraint: $\forall n \; \omega_n(\mathbf{x}_n^T\mathbf{w} + b) \geq 1$

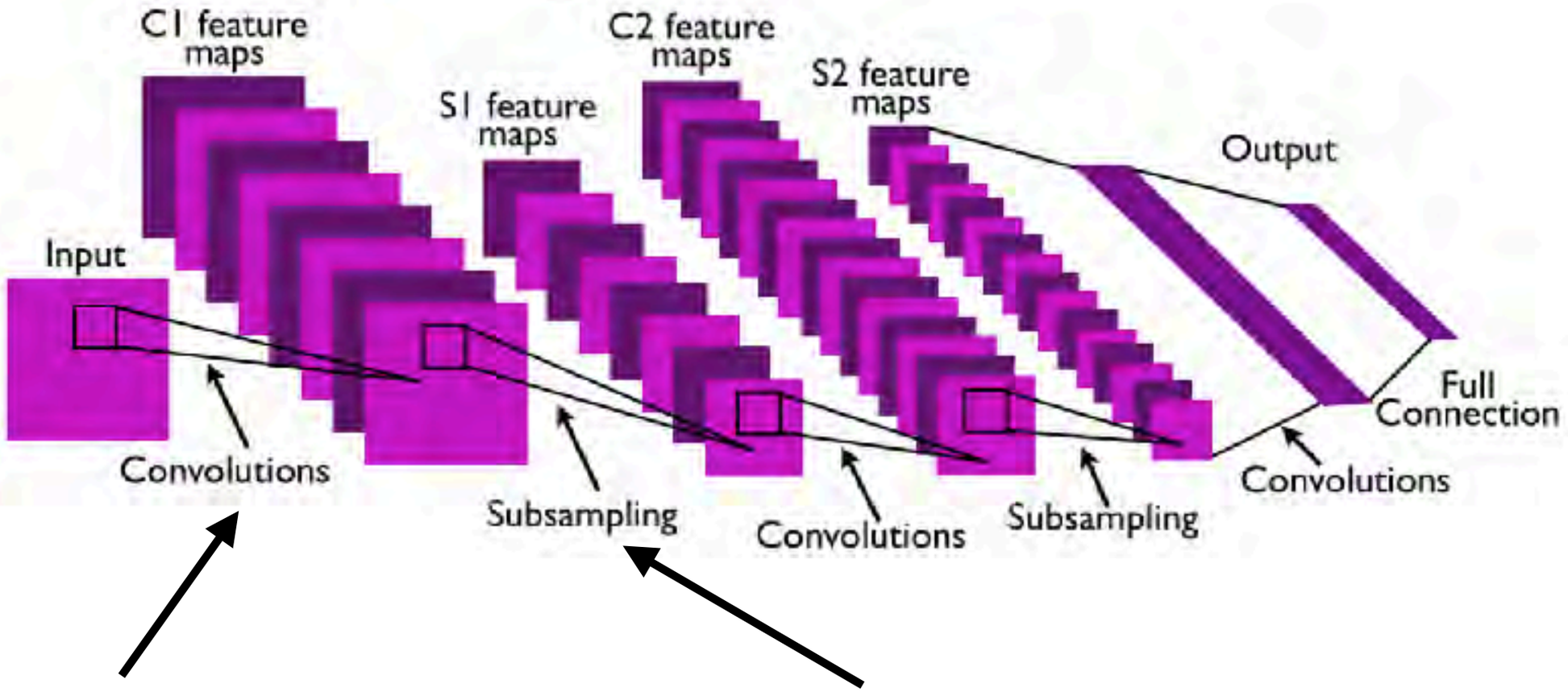


Yields a discriminant function $f(\mathbf{x}_m) = \mathbf{x}_m^T\hat{\mathbf{w}} + \hat{b}$

And a decision function $sgn(f(\mathbf{x}_m))$

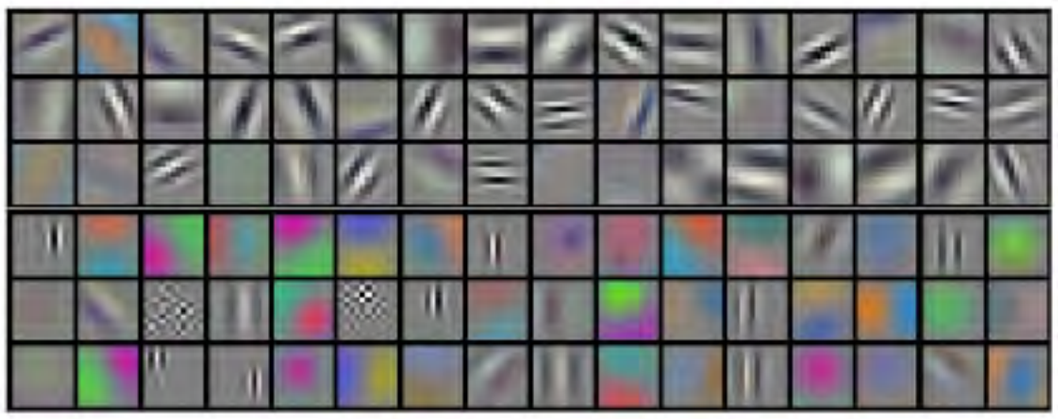
linSVM: a linear weighted sum - **w** tells you how to add / subtract features (voxels) to get **y** from **x.**

The convolutional neural network (CNN)



[LeCun et al., ISCAS, 2010]

uses learned image 'filters' 'pooling' reduces dimension

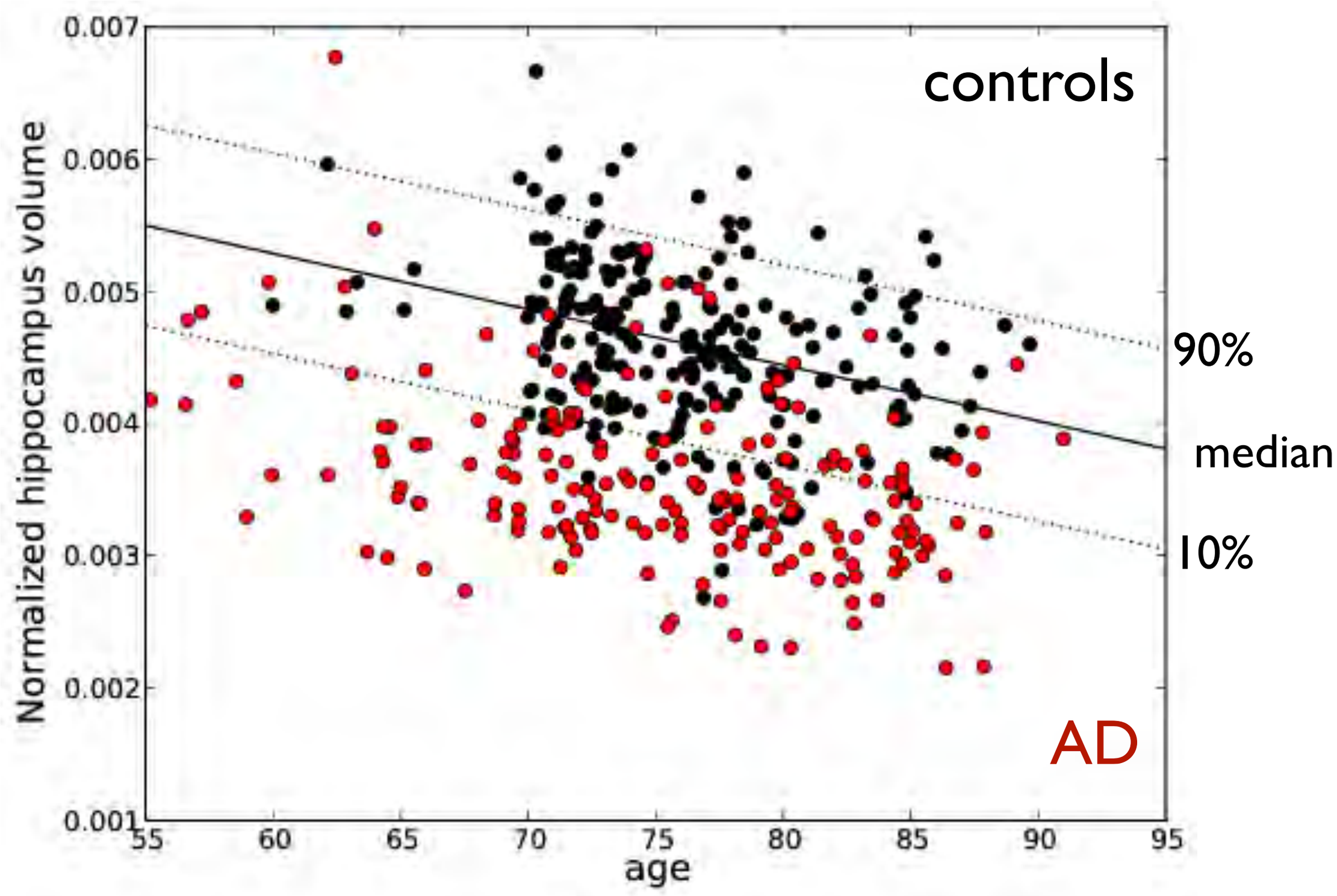


[Krizhevsky et al, NIPS 2012]

CNN: a non-linear method which learns its own features, needs "a lot" of data and can work very well

Clinical applications

CAD for dementia: volumetry



[Adapted from slides by Alexis Roche, 2016]

MorphoBox algorithm: [Schmitter et al, NeuroImage Clinical 2014]

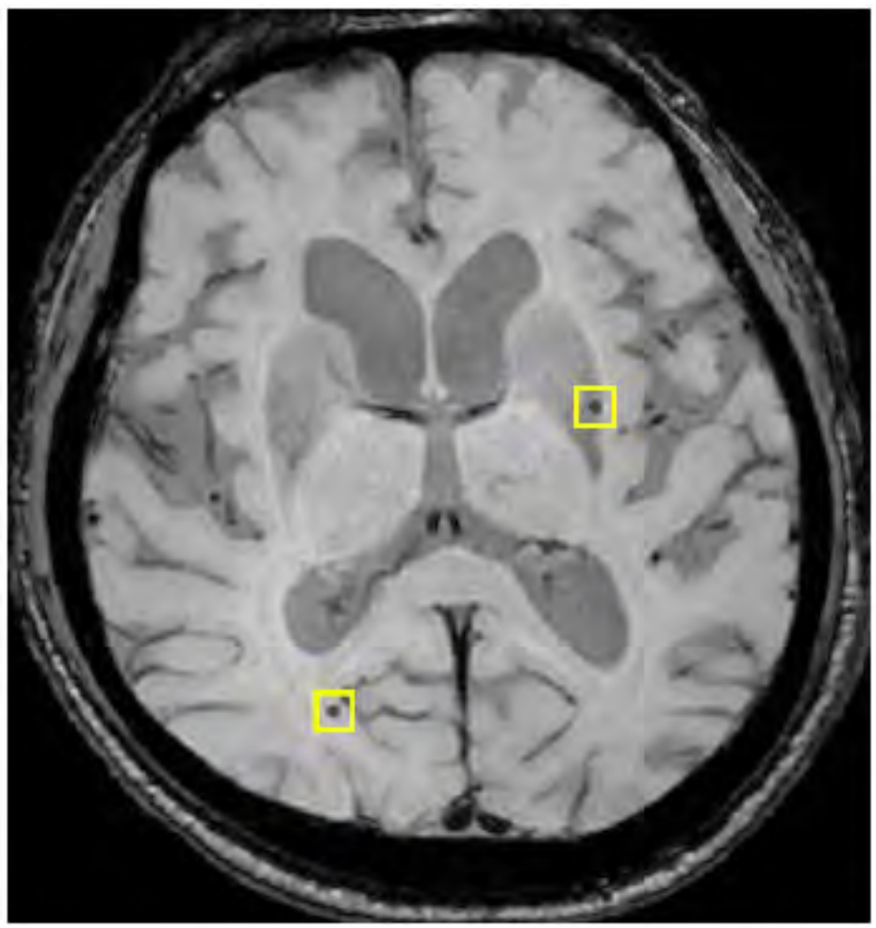
CAD for cerebrovascular disorders: microbleed detection

Data: SWI

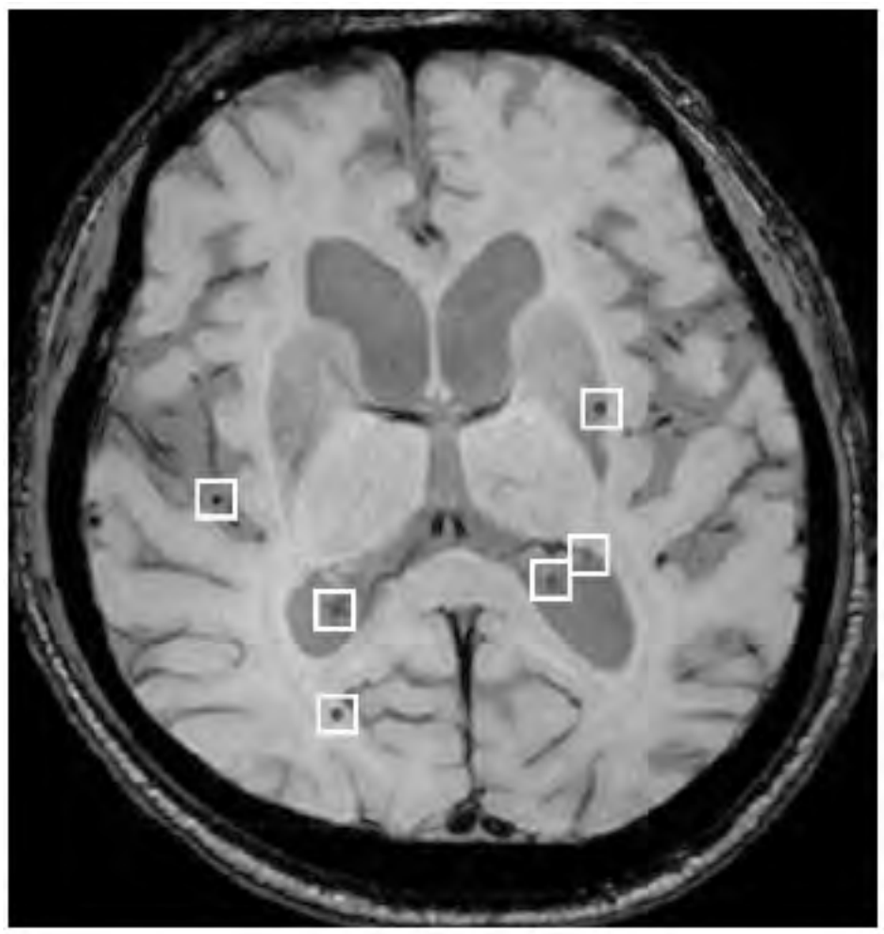
Classifier: 3DCNN

Subjects: 194 controls + 126 stroke (1149 CMBs)

Performance: sens 92%, 3 FP/subject (state of the art)



truth



detected

[Dou et al., IEEE TMI, 2016]

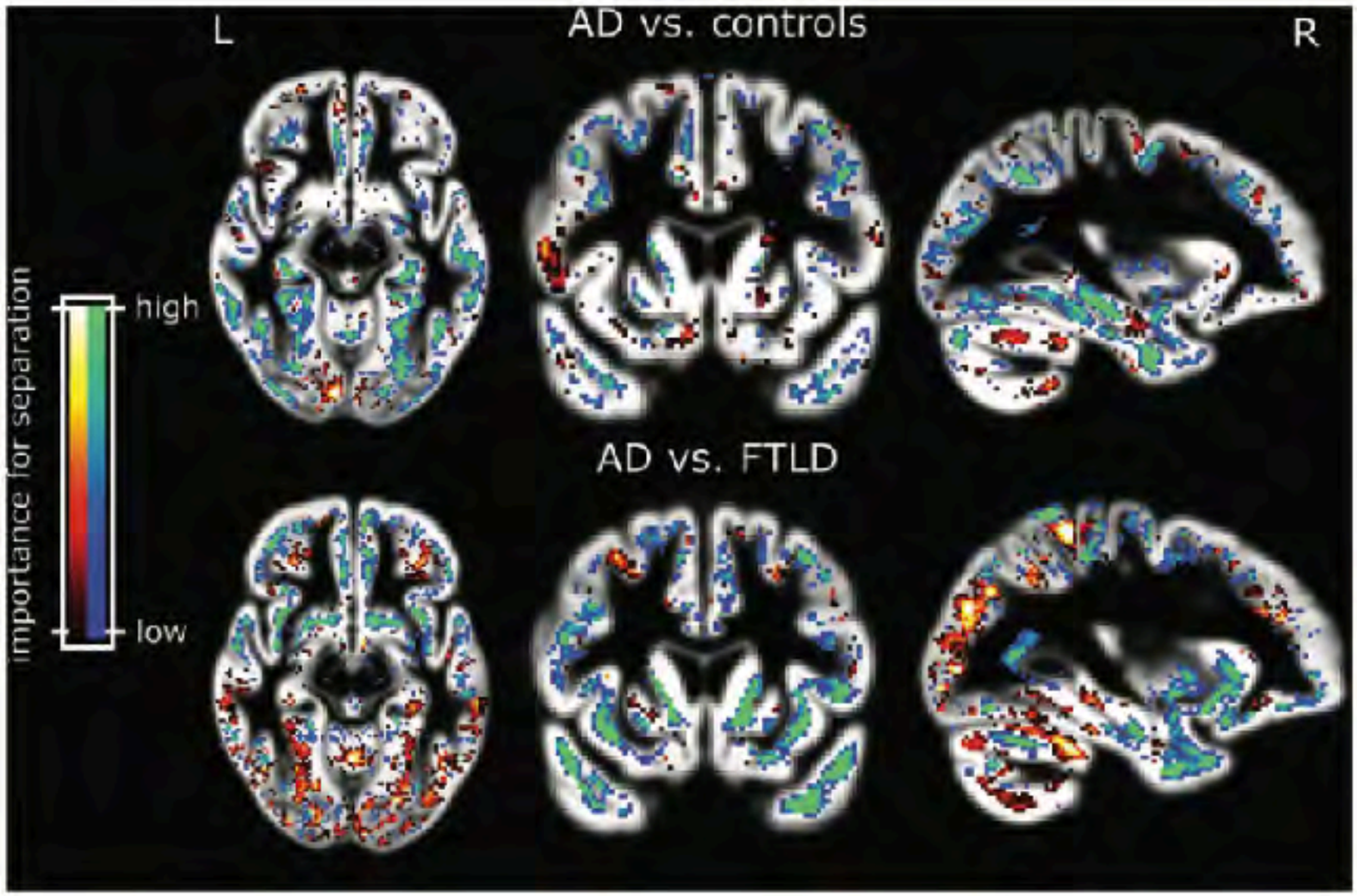
Automated multivariate dementia diagnosis

Features: modulated GM

Classifier: linSVM

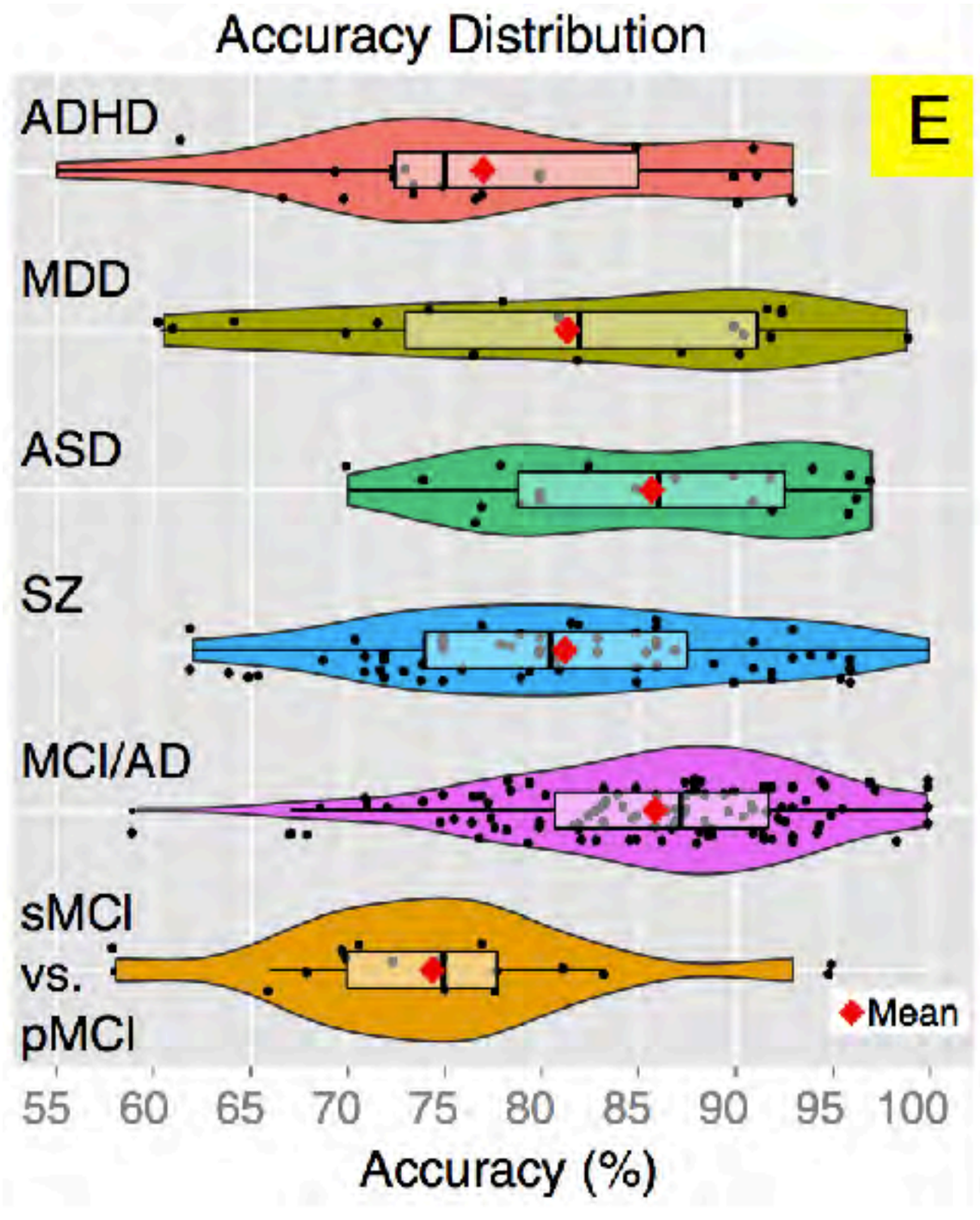
Subjects: 20+14 HC, 20+14 AD

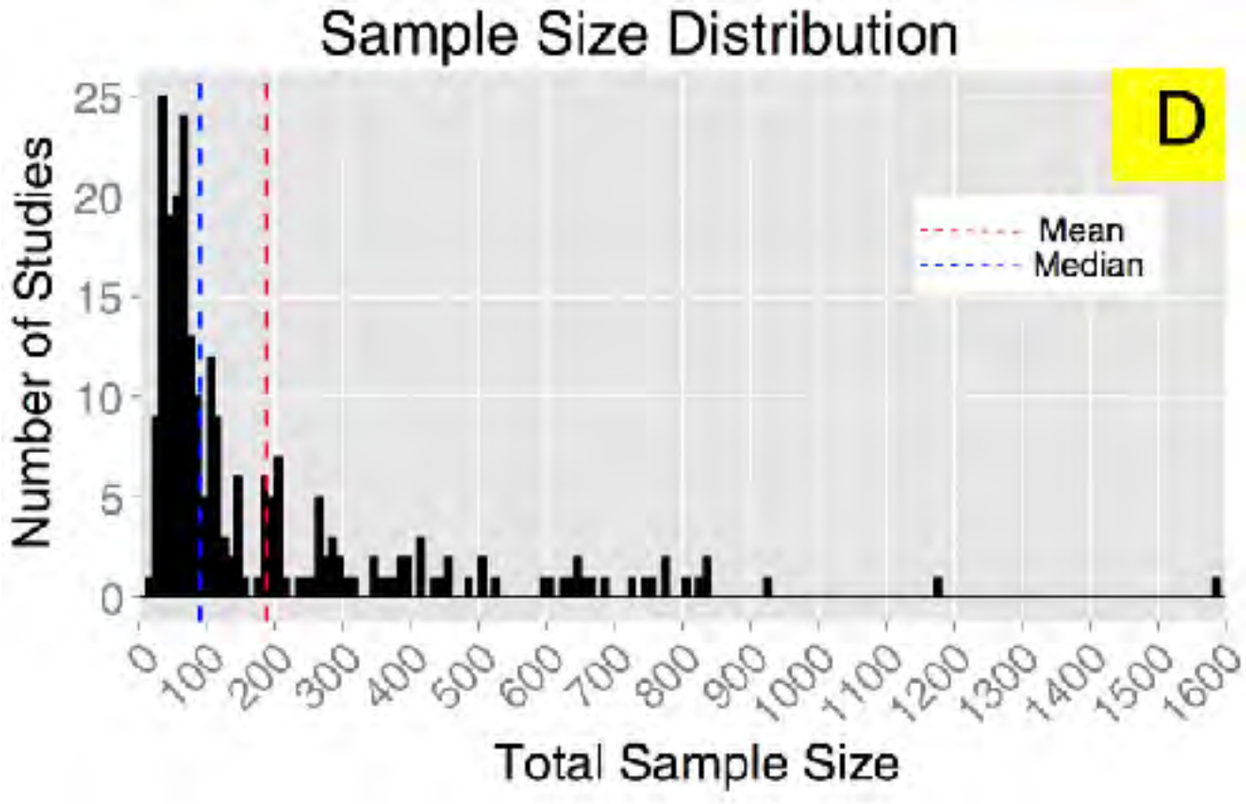
Performance: sens 97%, spec 94%



[Klöppel et al-, Brain, 2008]

Automated MR diagnosis in general



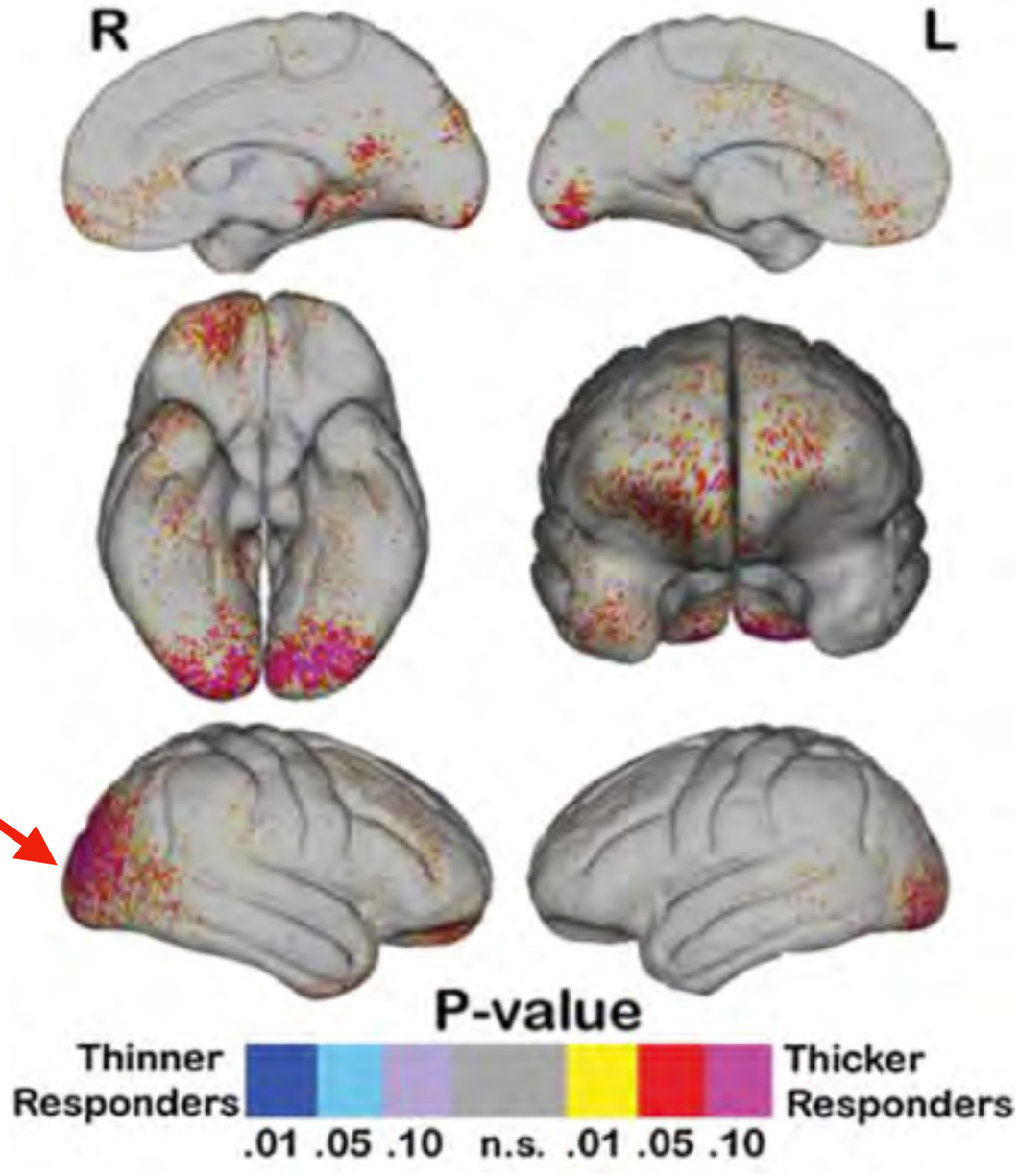


[Arbabshirani et al., NeuroImage 2017]

Schizophrenia treatment response prediction

Features: cortical thickness (GE 1.5T, SPGR)
Subjects: 39 SZ (25R/14 NR @ 16 wks) + 45 HC
Treatment: antipsychotics (olanzapine /risperidone)

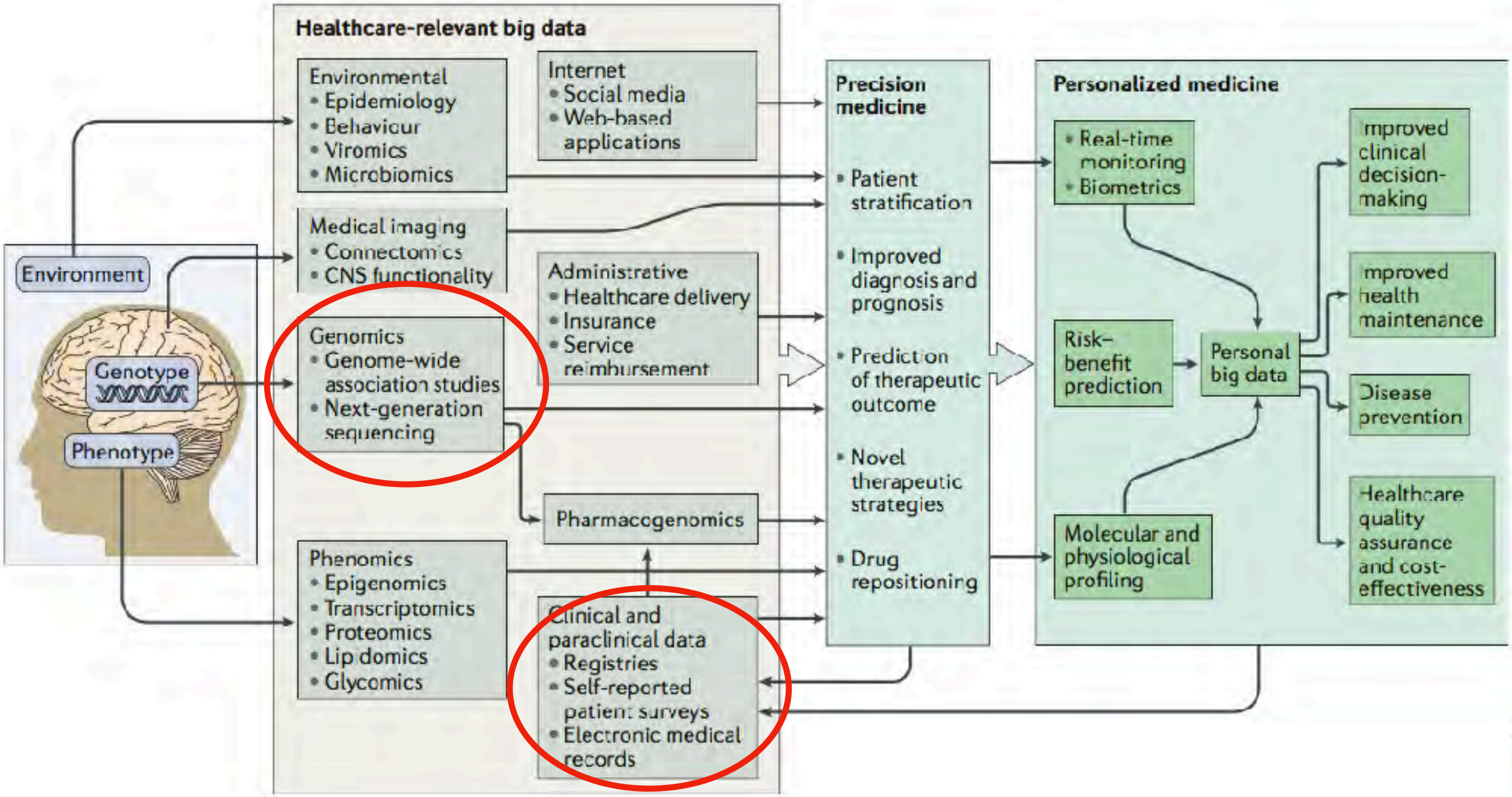
Responders have thicker OCx



[Szeszko et al., Schiz. Bull, 2012]

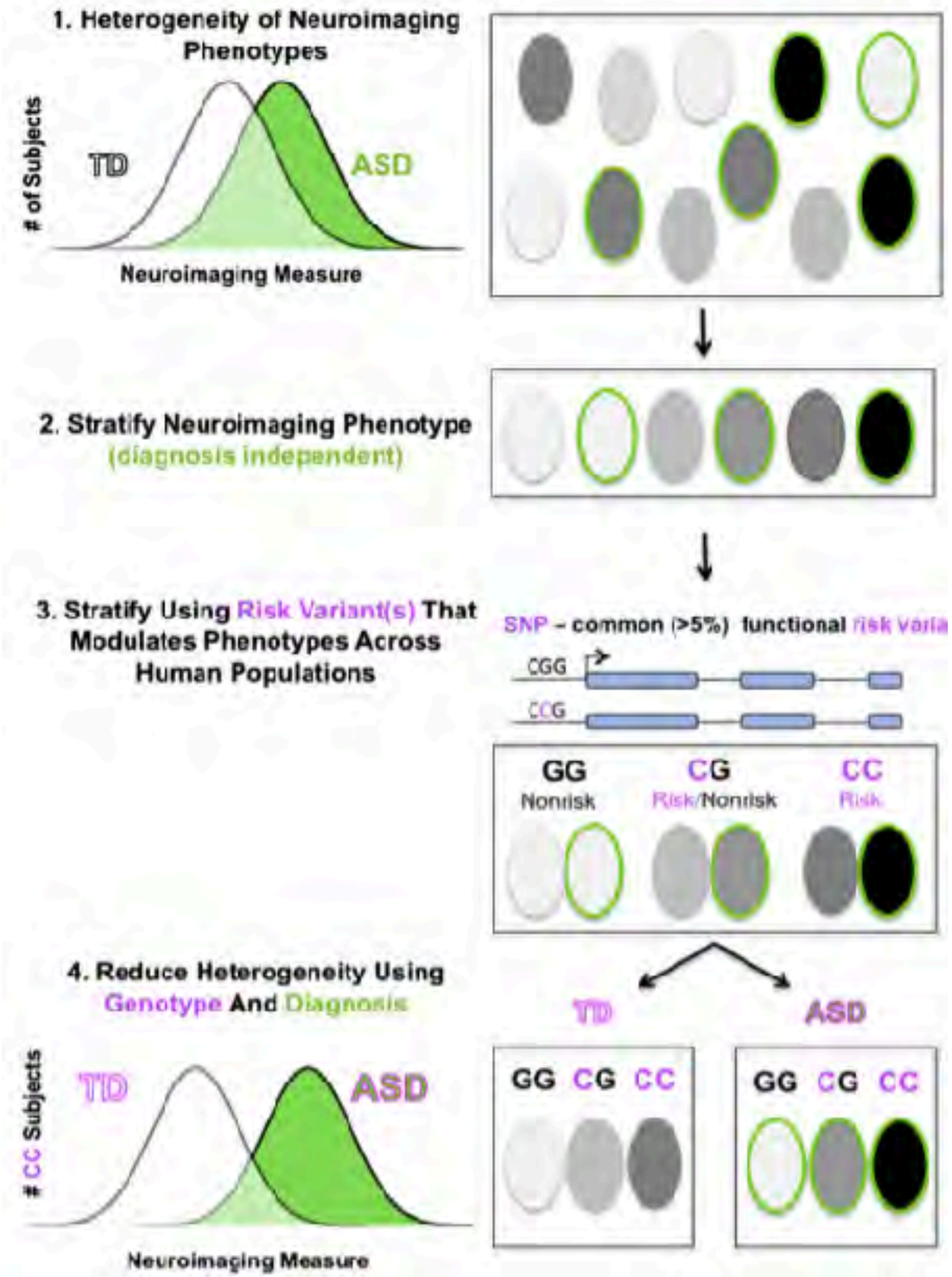
From radiology to precision medicine

Radiology data is only part of the picture



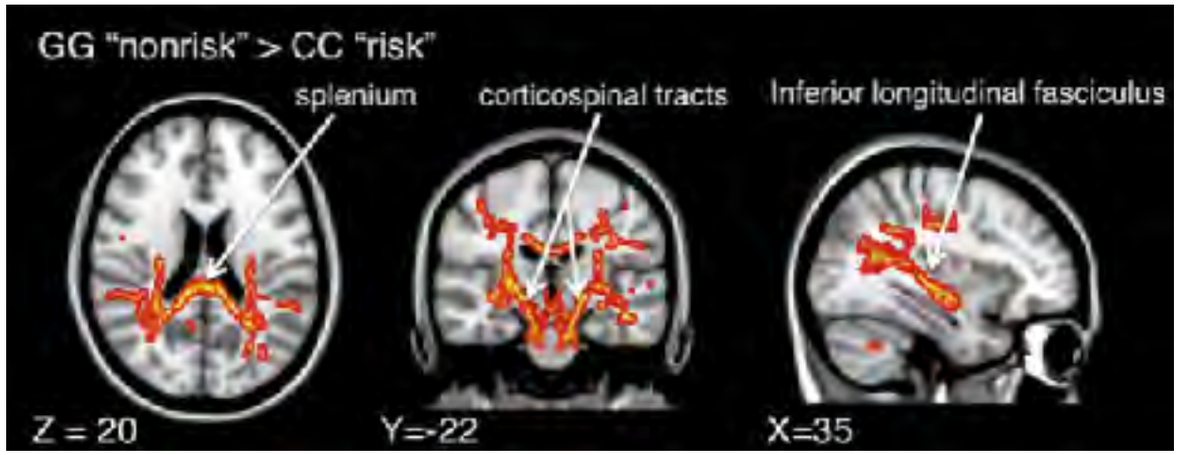
[Dendrou et al., Nat. Rev. Neurol. 2016]

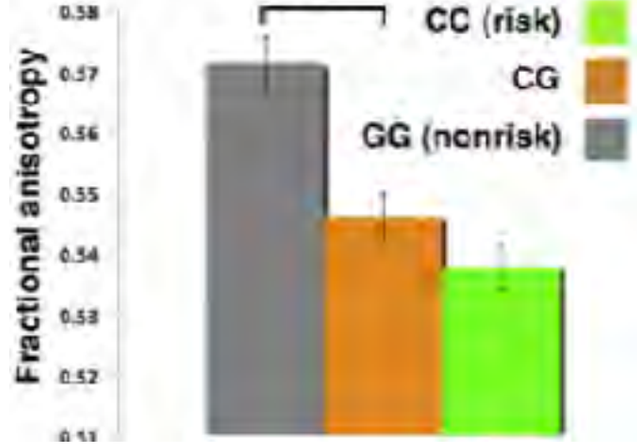
Imaging genomics for stratification

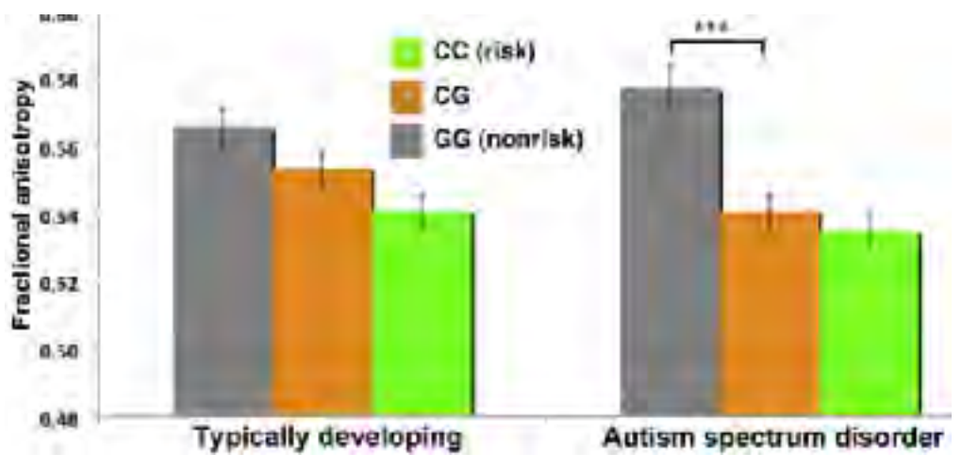


[Rudie et al.,, Neuron, 2012]

MET genotype: lower FA in TCx PCx OCx







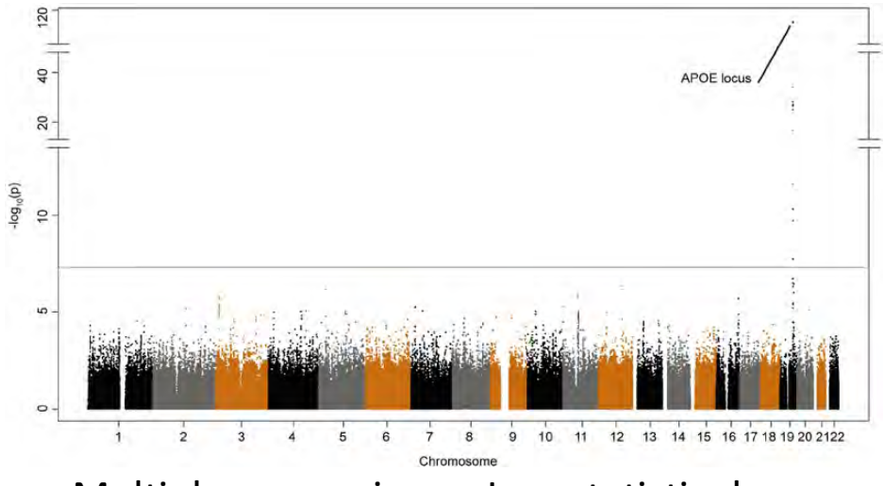
Imaging genomics for gene discovery

Problem finding genes for diseases using traditional techniques requires 1000s of subjects

10,000 patients

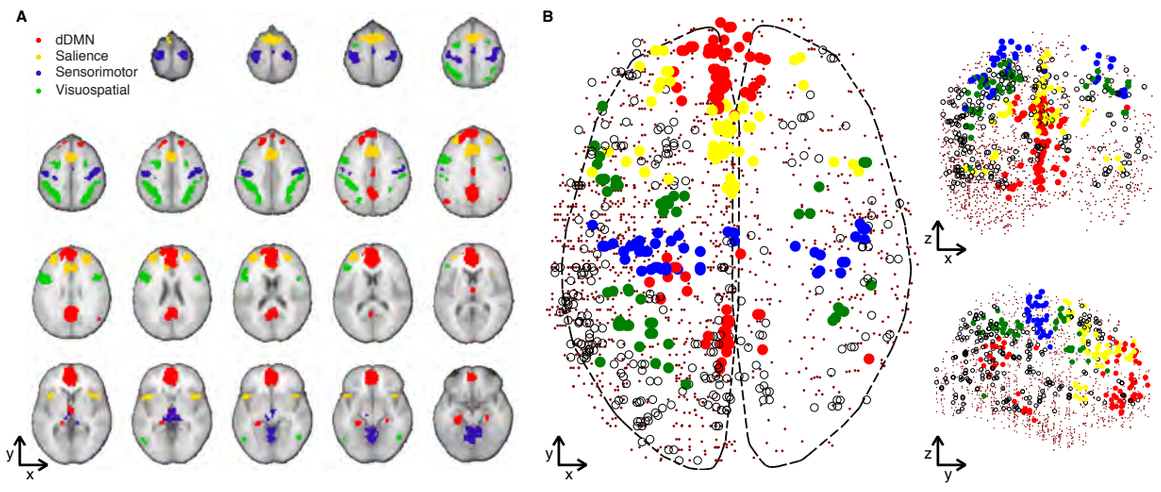
15,000,000 SNPs

1 diagnosis



Multiple comparisons: Low statistical power

Solution use brain imaging as intermediate phenotype



disease → imaging ← gene expression ← SNPs

[Richiardi, Altmann et al., Science, 2015]

Topic mapping and interpretation of radiology images

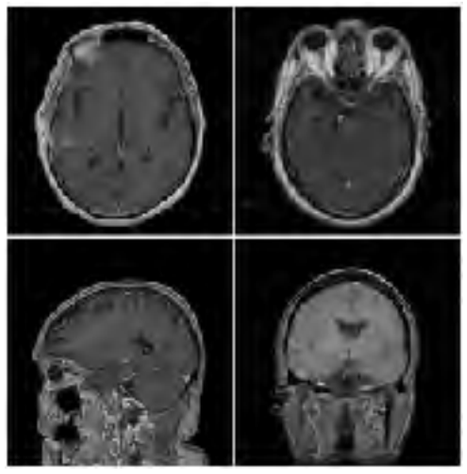
Imaging data: 216K 3D slices (1/4 MR, 3/4 CT)

Text data: 780K radiology reports

Performance:

- image → topic (60): 66% (top-1), 95% (top-5)
- image → disease (77): 71% (top-1), 88% (top-5)

example topic: "MRI of brain tumor"



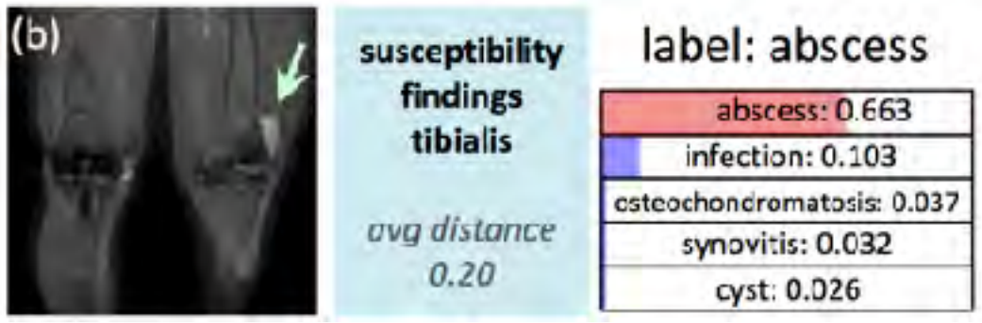
axial,contrast,mri,sagittal,post,flair,enhancement,blood,dynamic,brain,relative,volume,this,precontrast,from,tesla,fse,diffusion,gradient,resection,comparisons,maps,philips,progression,some,susceptibility,perfusion,stable,achieva,technique,echo,weighted,1.5,evidence,mass,findings,hemorrhage,enhanced,impression,frontal,signal,coronal,dti,tumor,t1-ffe,hydrocephalus,magnevist,reformations,bolus,lesion

automated image interpretation

input

... for example series 701 image 12 and series 401 image 27 with **findings** suggesting minimally enhancing rim laterally for example series 1101 image 21 may ... the **findings** suggest a fluid collection with ... the location suggests possibility of a **synovial** collection **synovial** thickening as the appearance is nonspecific correlation with clinical findings is recommended regarding the possibility of an **infection** **abscess**

output



[Shin et al., JMLR, 2016]

Challenges

Challenges

IT infrastructure

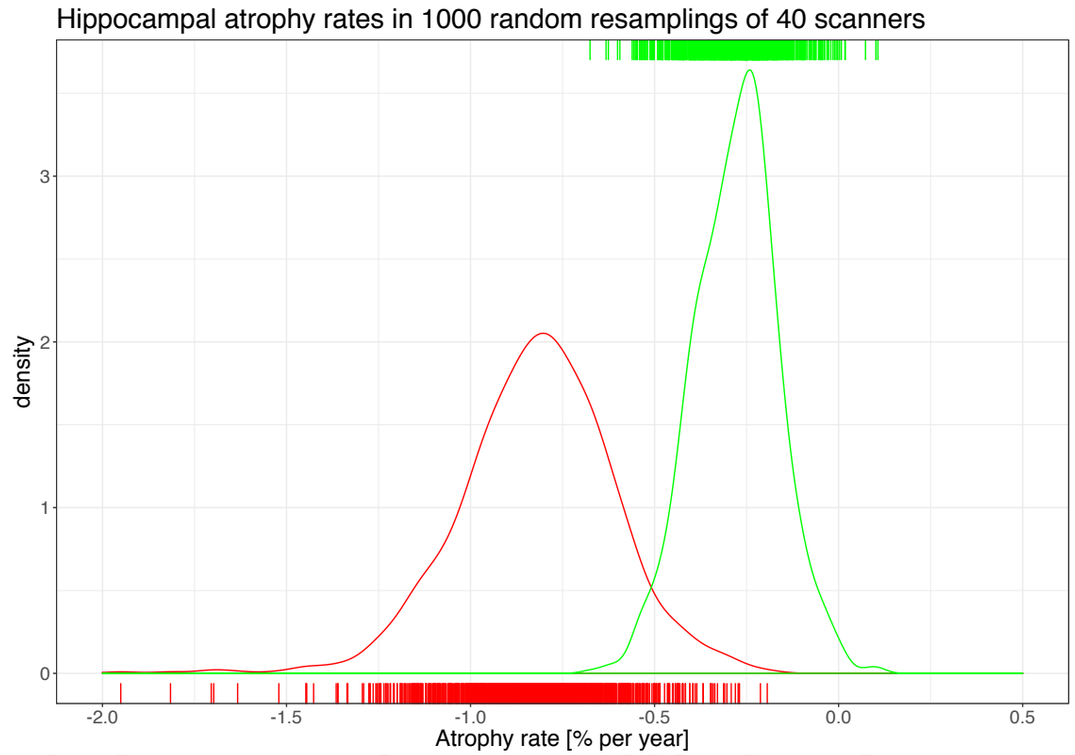
Move computation to data - how to standardise (SPHN)

Annotated data still small

Especially ICD-10 subcategories

How to use multi-site data ?

Most theory is built for IID data



How to decide between sensitivity and specificity?

Cost of false positive / false negatives differs widely across diseases

Conclusions

We can leverage large imaging datasets and computing power to improve healthcare: *hospitals are sitting on a goldmine.*

Methods are getting better very rapidly: *semi- and fully-automated tools will empower radiologists.*

Integrating imaging with -omics will lead to vastly improved understanding of disease: *radiology is central to precision medicine.*

Thanks

Collaborators

Allen Institute

C.K. Lee

IMAGEN consortium

FINDLab, Stanford

M. Greicius

TIG, UCL

A. Altmann

Berkeley

JB Poline

Siemens Healthineers

ACIT team

Kelp Lab, UCSC

C. Otter

Additional resources

Calhoun et al. (eds), *NeuroImage special issue on Individual Subject Prediction*, 2017

Greenspan et al., *Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique*, IEEE TMI, 2016

Rao et al., *Predictive modelling using neuroimaging data in the presence of confounds*, NeuroImage, 2017

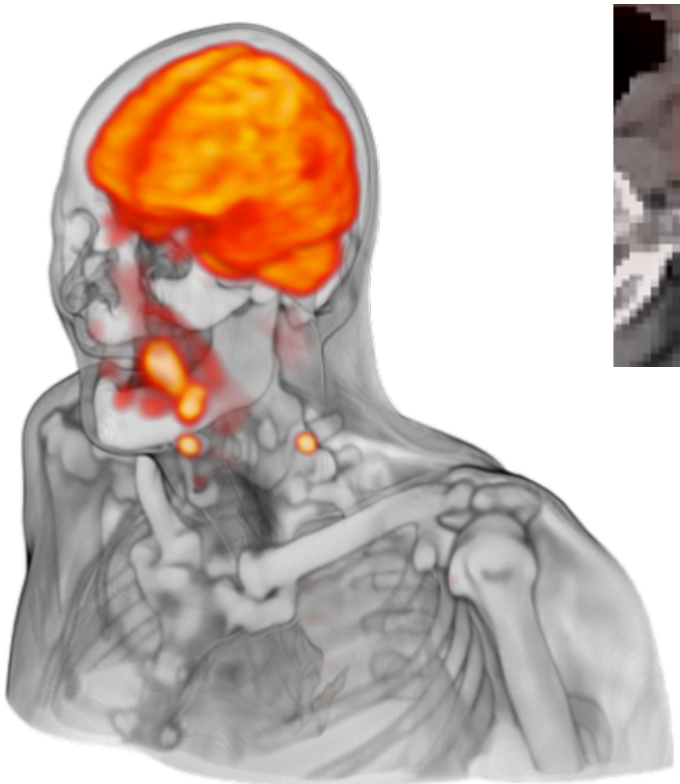
Shen et al., *Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: A study of the ADNI cohort*, NeuroImage, 2010

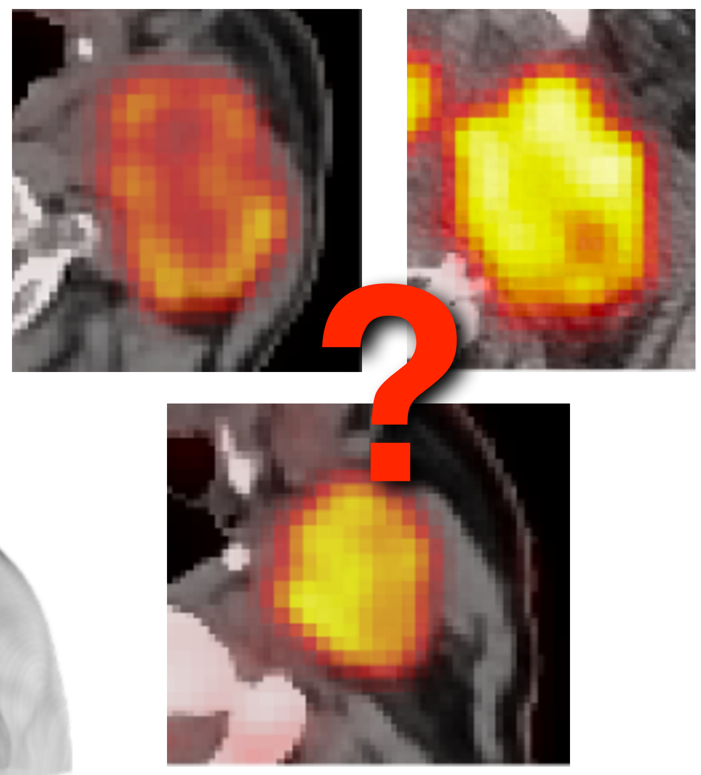
Hawrylycz et al., *Canonical genetic signatures of the adult human brain*, Nature Neuroscience, 2015

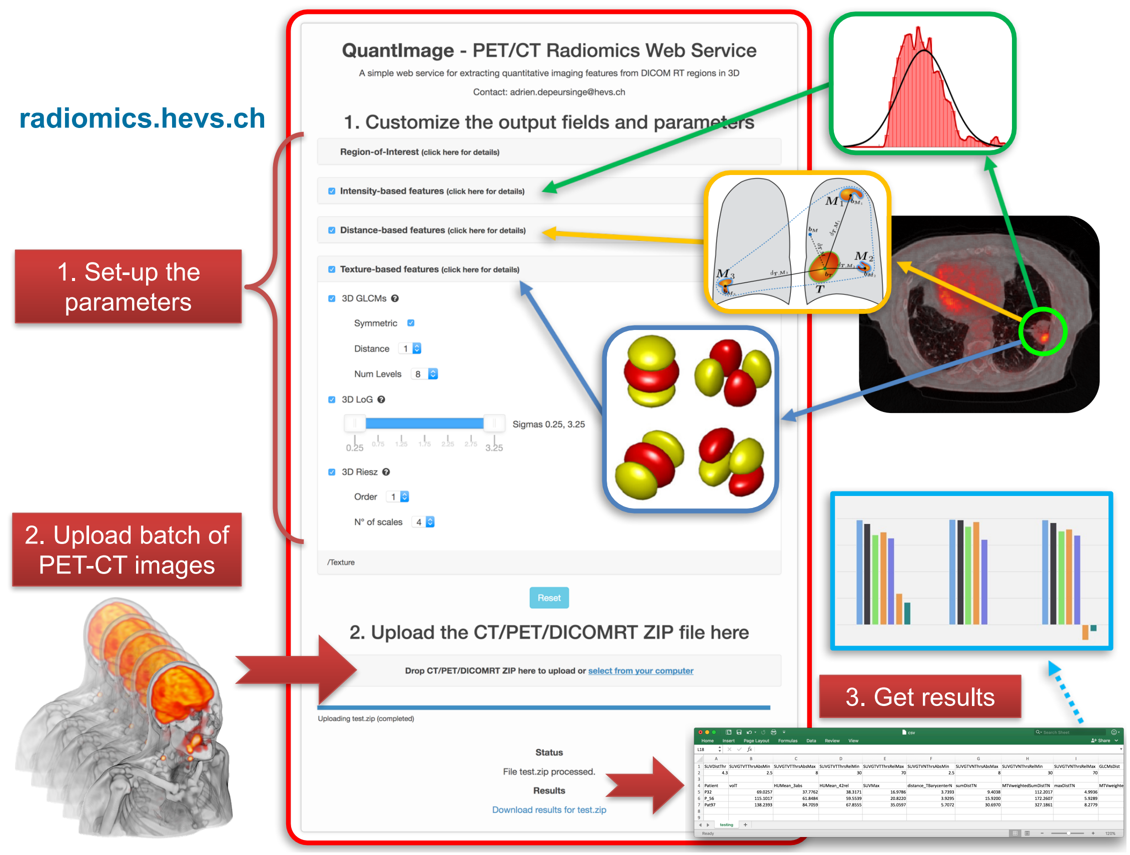
Miller et al., *Multimodal population brain imaging in the UK Biobank prospective epidemiological study*, Nature Neuroscience, 2016

LEVERAGING 3D TEXTURE INFORMATION IN PET AND CT IMAGES FOR PRECISION MEDICINE WITH THE QUANTIMAGE PLATFORM

A. Depeursinge, Y. Dicente, R. Schaer, J. Castelli, J. O. Prior





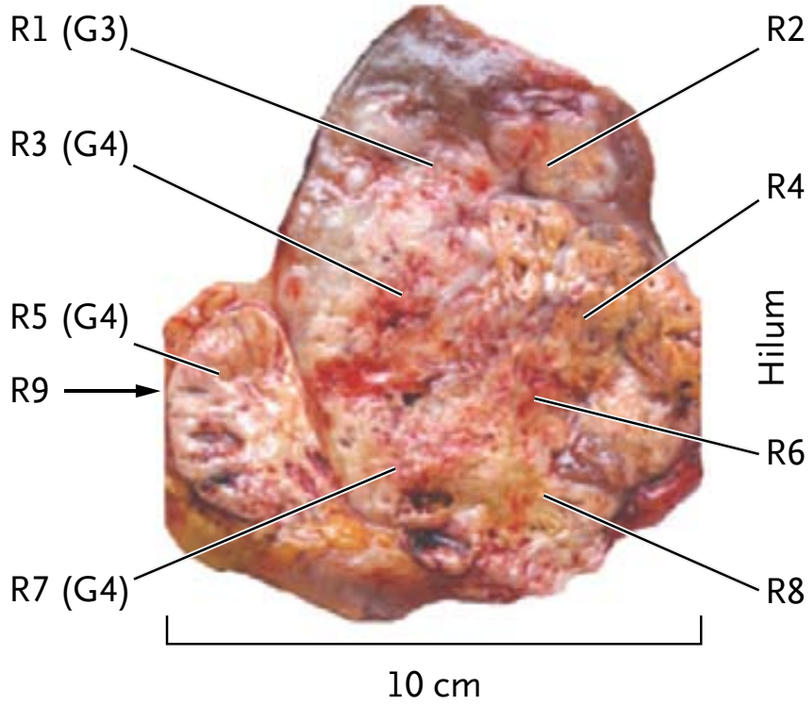


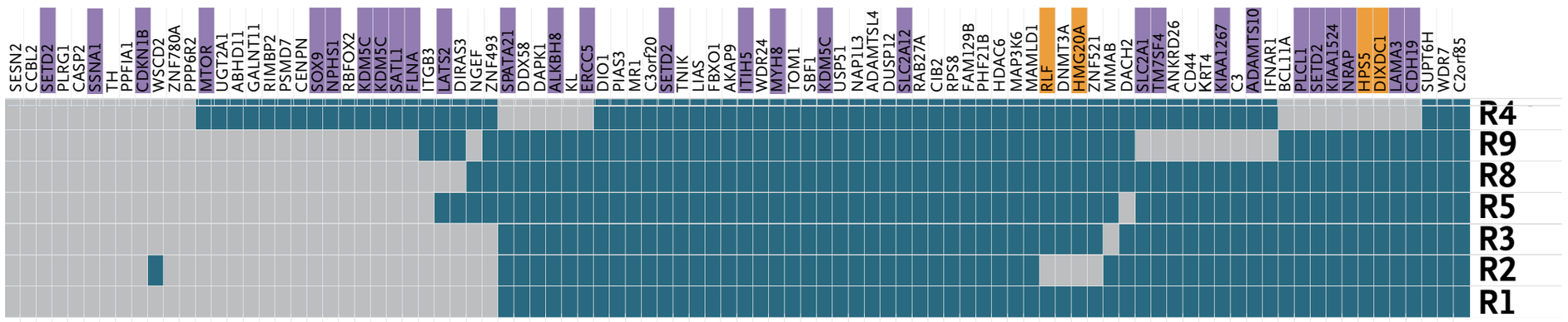
Outline

- Background - Radiomics
  - Personalized tumor phenotyping in PET - CT
- Methods
  - Image mining: intensity versus texture
  - Texture analysis: definition and qualitative review
  - The QuantImage platform
- Results
  - Head and Neck cancer: importance of geometry
  - Lung cancer: importance of aggregation
- Conclusions and perspectives

Personalized Tumor Phenotyping

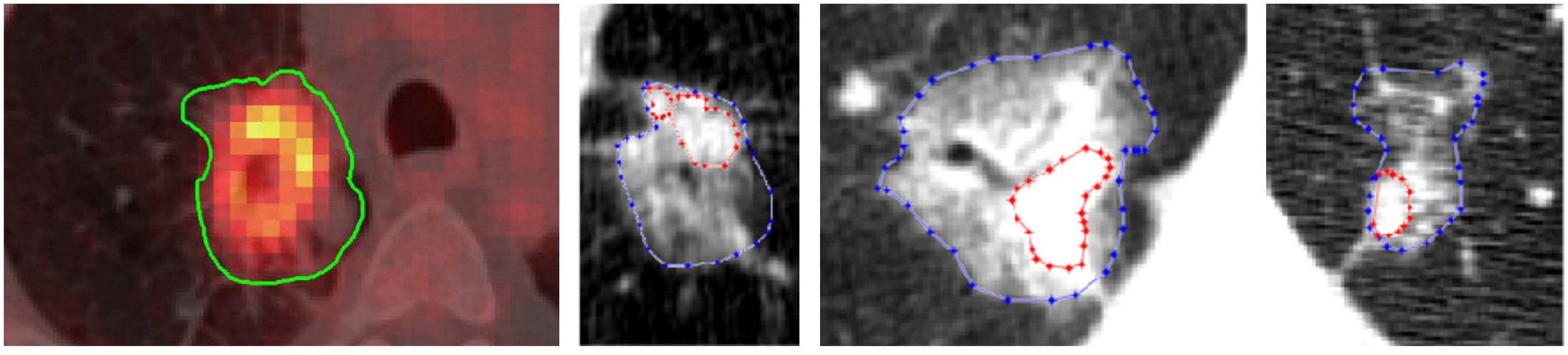
- Personalized medicine aims at enhancing the patient's quality of life and prognosis
  - Tailored treatment and medical decisions based on the molecular composition of diseased tissue
- Current limitations [Gerlinger2012]
  - Molecular analysis of tissue composition is invasive (biopsy), slow and costly
  - Cannot capture molecular heterogeneity





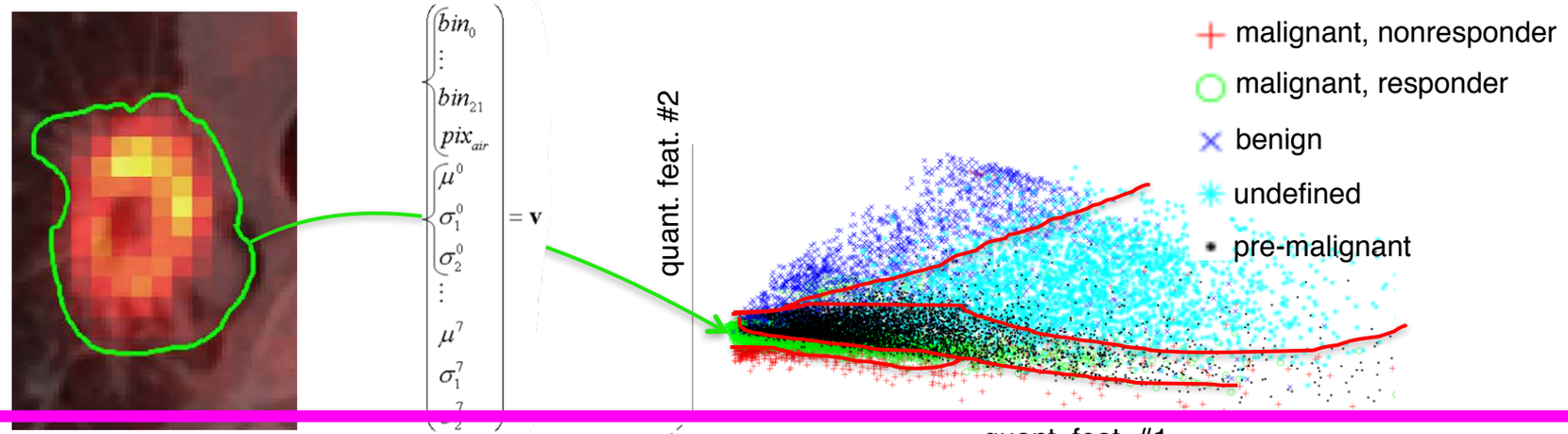
Personalized Tumor Phenotyping

- Huge potential for computerized medical image analysis
  - Explore tumor heterogeneity in existing diagnostic images
- Cancer ecosystem is composed of micro-habitats [Gatenby2013]
  - Relates to cancer subtype, patient survival, response to treatment
- The density, metabolism, and structure of tumor tissue observed in PET and CT images reflects their nature [Leijenaar2015]
  - E.g., active cancer cells, angiogenesis, necrosis [Aerts2014]
  - PET and CT axial views of non-small cell lung cancer:



Personalized Tumor Phenotyping

- Radiomics: image-based personalized phenotyping [Kumar2012]
  - Use image analysis to predict disease outcome
  - Surrogate slow, costly and invasive molecular analysis

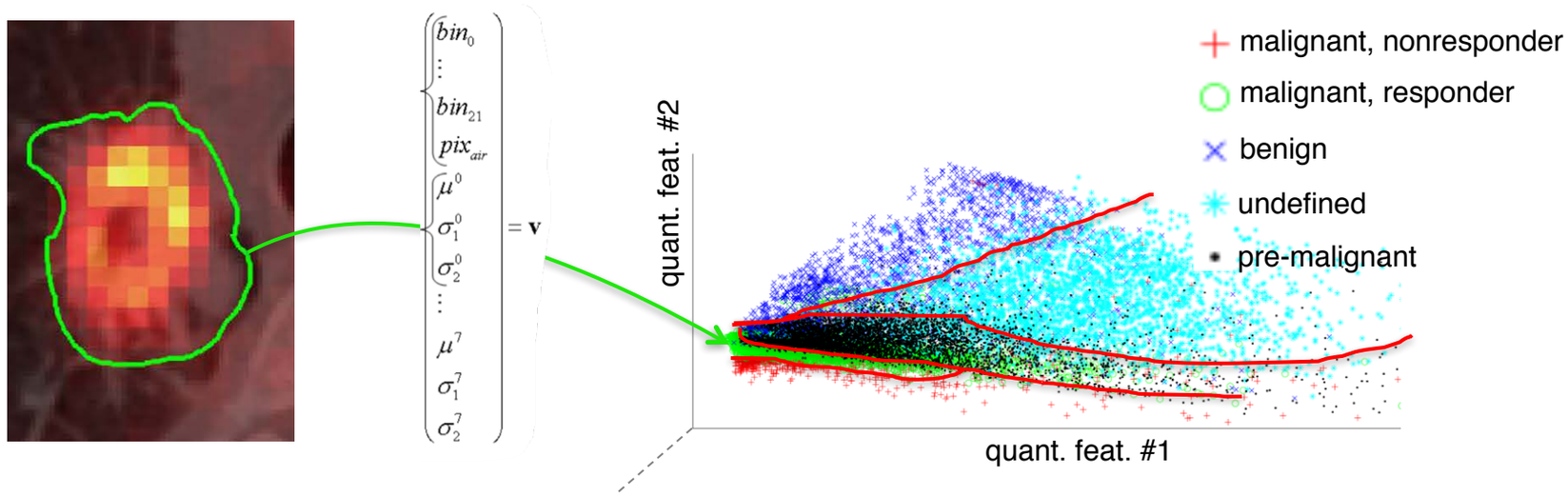


- Quantitative feature extraction
  - Intensity, shape, margin, texture
- Statistical and predictive models
  - Uni- and multi- variate

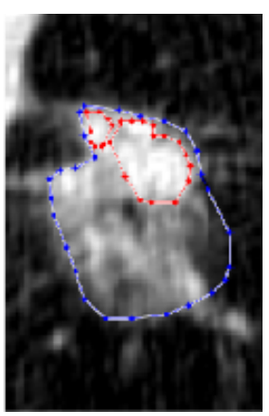
Deep learning

Personalized Tumor Phenotyping

- Radiomics: image-based personalized phenotyping [Kumar2012]
  - Use image analysis to predict disease outcome
  - Surrogate slow, costly and invasive molecular analysis



- Related work [Ganeshan2013, Ravanelli2013, Mattonen2014, Depeursinge2015, …]
  - ✗ No separate analysis of nodule components: mixing micro-habitats
  - ✗ Limited geometric specificity of current texture biomarkers

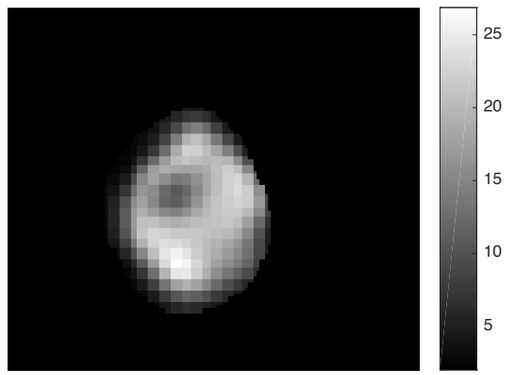


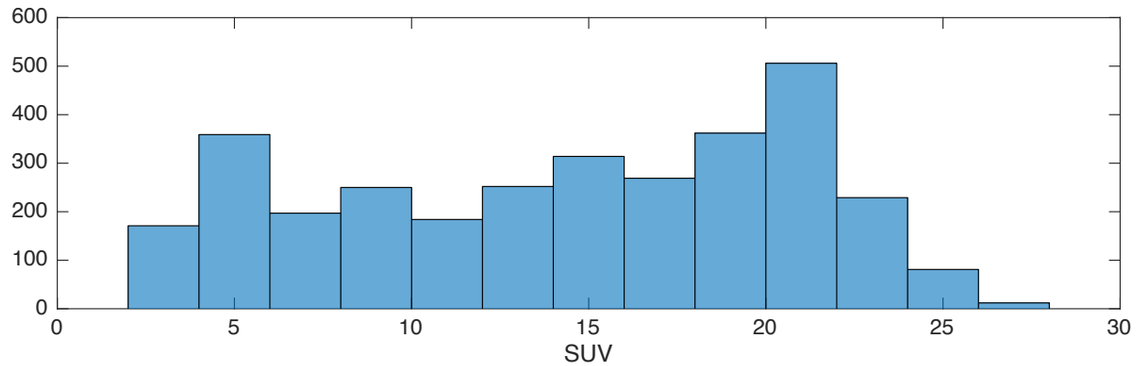
Outline

- Background - Radiomics
  - Personalized tumor phenotyping in PET - CT
- Methods
  - Image mining: intensity versus texture
  - Texture analysis: definition and qualitative review
  - The QuantImage platform
- Results
  - Head and Neck cancer: importance of geometry
  - Lung cancer: importance of aggregation
- Conclusions and perspectives

Intensity versus Texture

- Intensity features are based on statistics of voxel values
  - Standardized Uptake Values (SUV) in PET



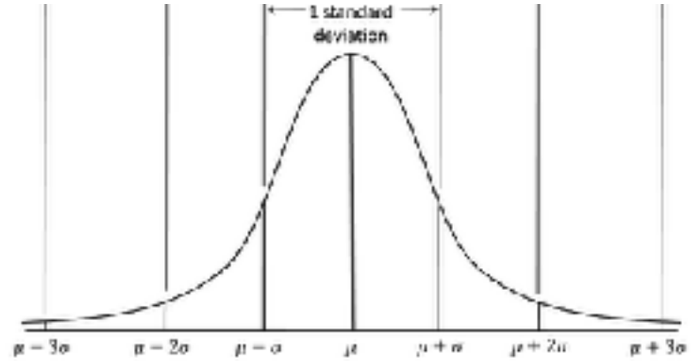


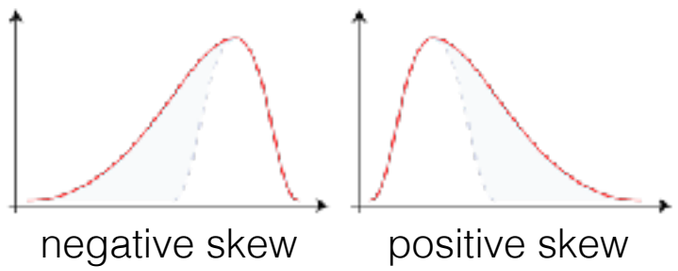
- Features:
  - mean $\mu_M = \frac{1}{N_M} \sum_M I(\boldsymbol{k})$
  - variance $\sigma_M = \frac{1}{N_M} \sum_M (I(\boldsymbol{k}) - \mu_M)^2$
  - skewness

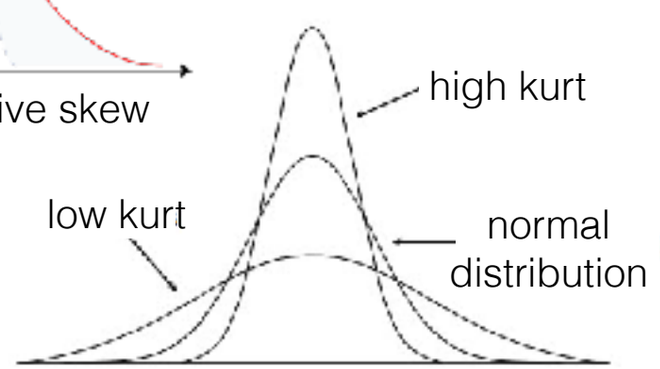
$$skew_M = \frac{\frac{1}{N_M} \sum_M (I(\boldsymbol{k}) - \mu_M)^3}{\sigma_M^3}$$

  - kurtosis

$$kurt_M = \frac{\frac{1}{N_M} \sum_M (I(\boldsymbol{k}) - \mu_M)^4}{\sigma_M^4}$$

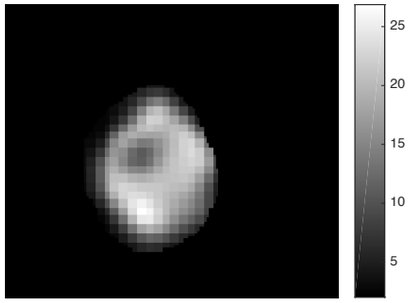


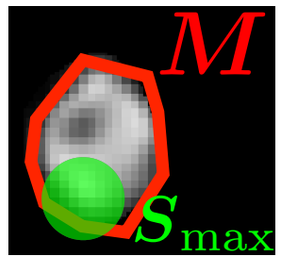




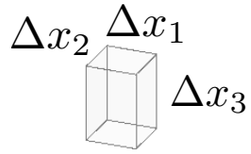
Intensity versus Texture

- Intensity features are based on statistics of voxel values
  - Specific to PET [Orlhac2014]:
    - $SUVmax = \max_{\boldsymbol{M}}(I(\boldsymbol{k}))$
    - $SUVpeak = \frac{1}{N_{\boldsymbol{M}\cap\boldsymbol{S}_{\max}}} \sum_{\boldsymbol{M}\cap\boldsymbol{S}_{\max}} I(\boldsymbol{k})$
      - $\boldsymbol{S}_{\max}$ is a spherical region of 1.2cm diameter centered at the position of $SUVmax$





  - Metabolic tumor volume (MTV): $MTV = vol_{voxel} \times N_{\boldsymbol{M}}$
    - $vol_{voxel} = \Delta x_1 \Delta x_2 \Delta x_3$ is the volume of one voxel in cm3

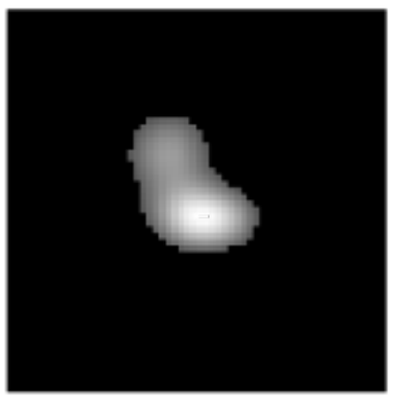


  - Total lesion glycolysis (TLG): $TLG = SUV_{mean} \times MTV$

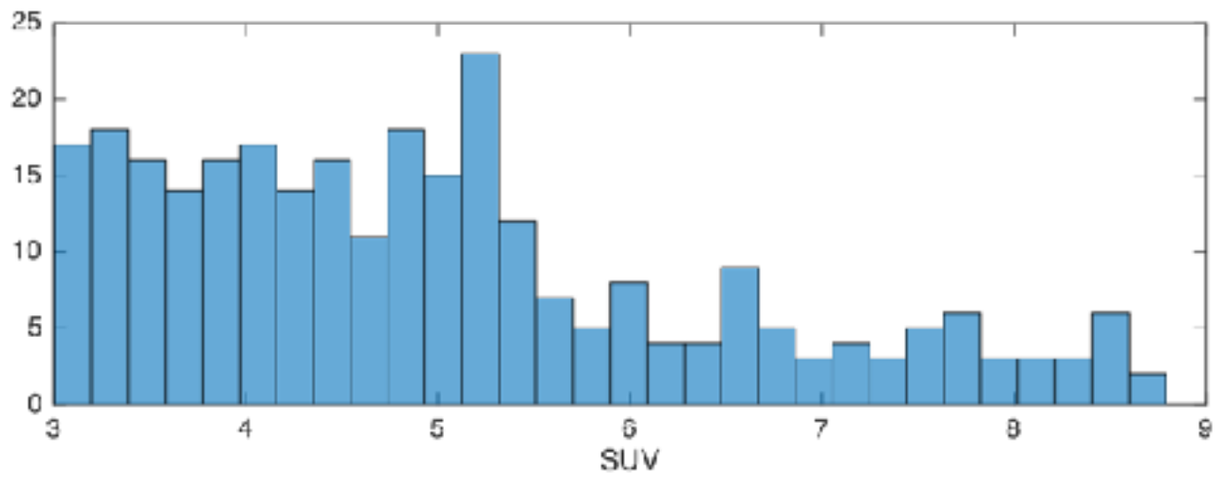
Intensity versus Texture

- Intensity features are insensitive to tissue morphology
  - I_A and I_B have identical intensity distributions

I_A



I_B



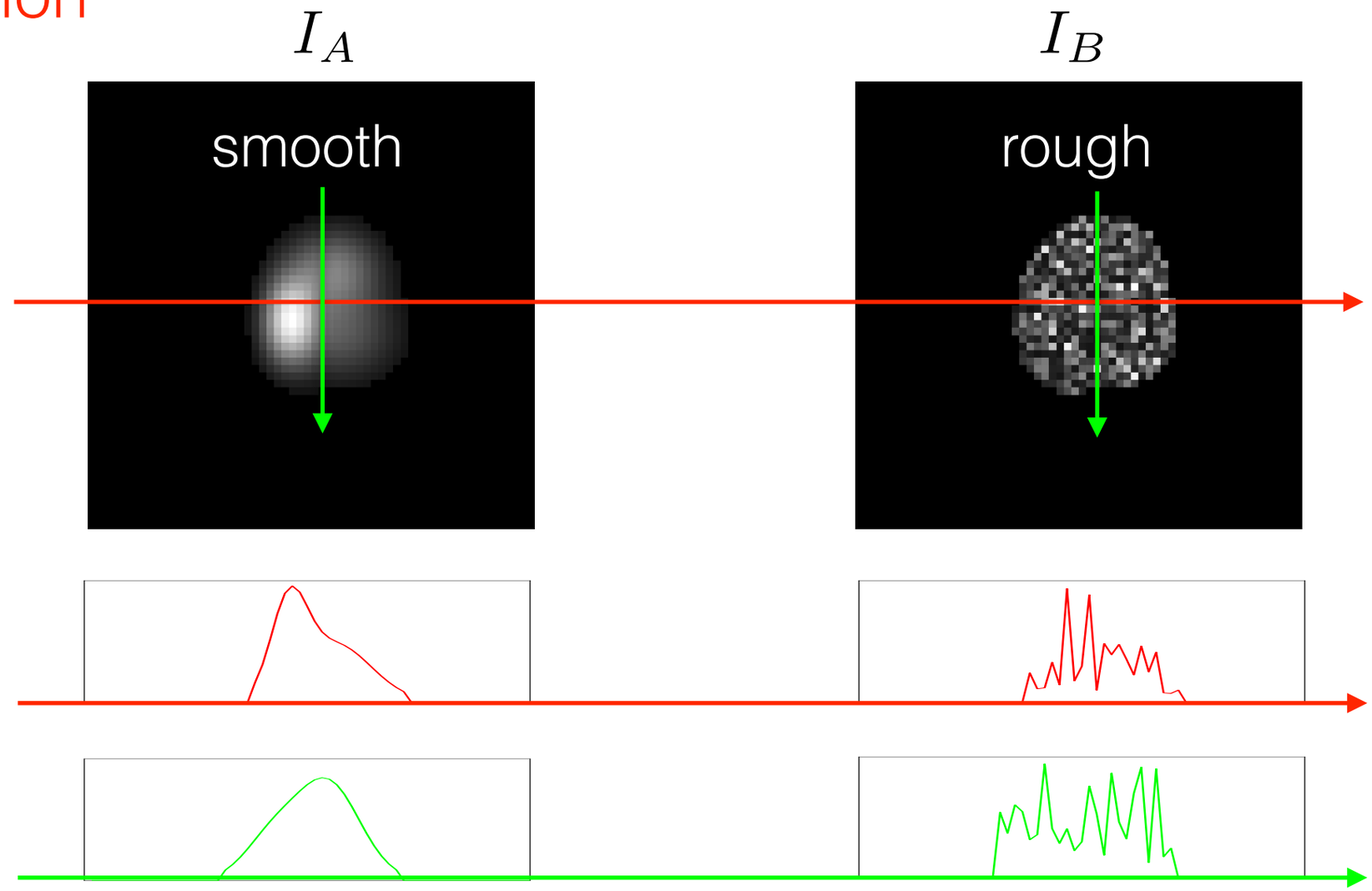
- Terms such as tumor “heterogeneity” [Kidd2008] are ambiguous in the context of imaging

Outline

- Background - Radiomics
  - Personalized tumor phenotyping in PET - CT
- Methods
  - Image mining: intensity versus texture
  - Texture analysis: definition and qualitative review
  - The QuantImage platform
- Results
  - Head and Neck cancer: importance of geometry
  - Lung cancer: importance of aggregation
- Conclusions and perspectives

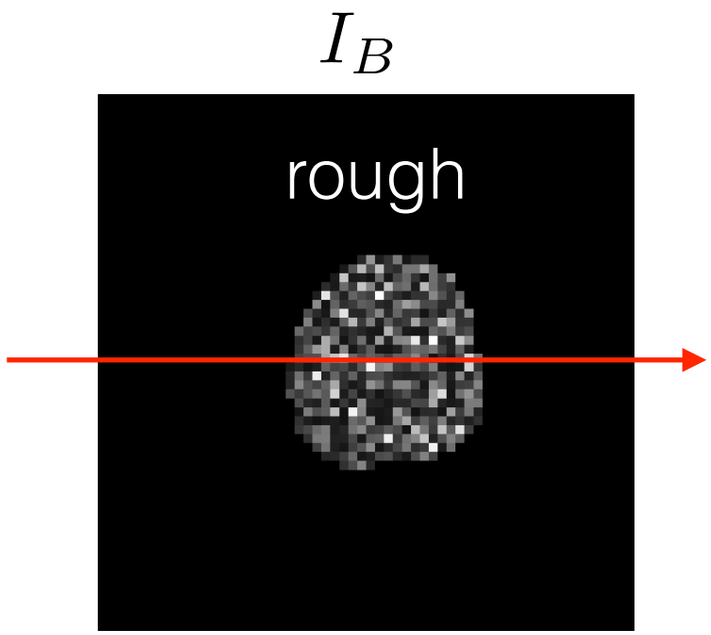
Texture analysis

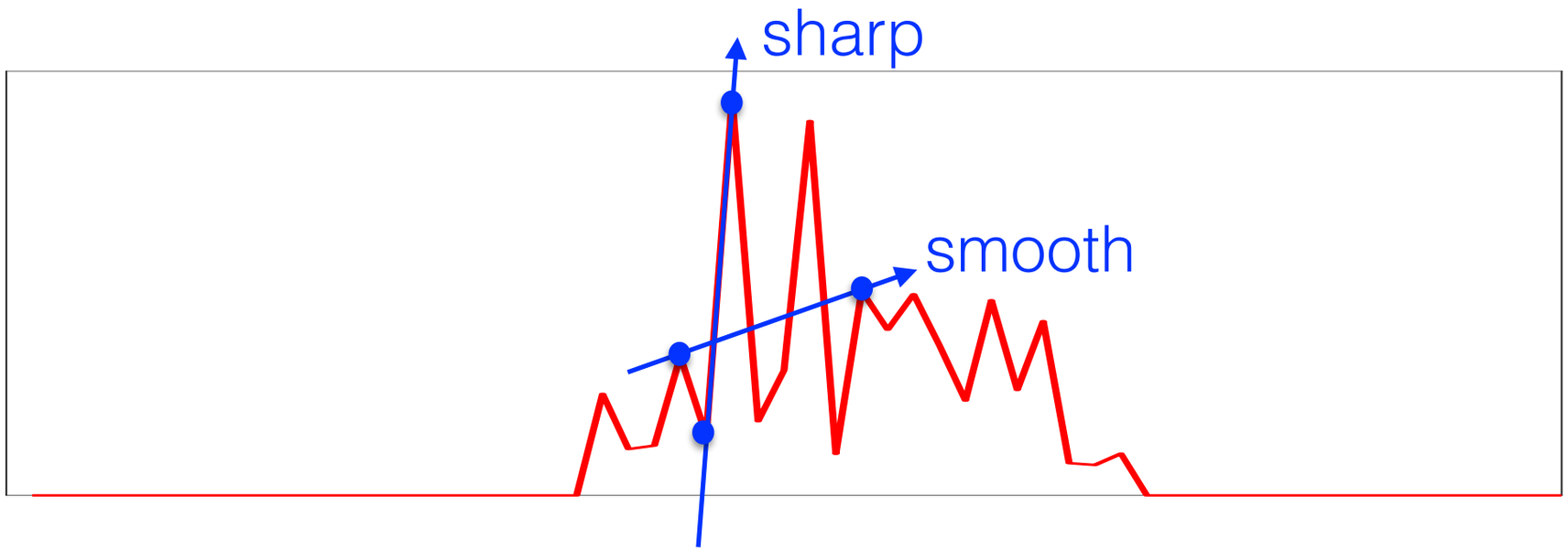
- Texture characterizes transitions between voxel values
  - direction



Texture analysis

- Texture characterizes transitions between voxel values
  - scale

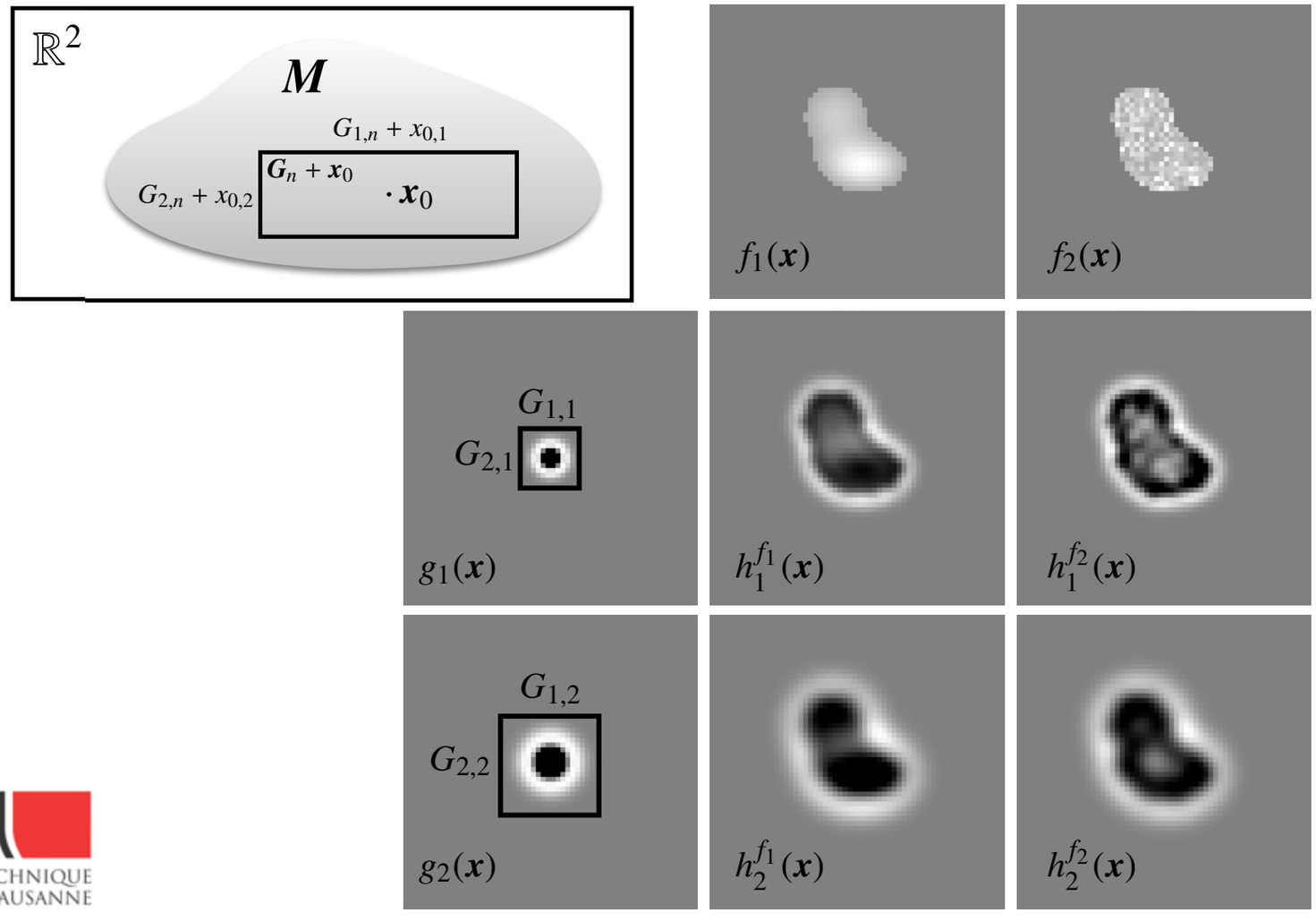


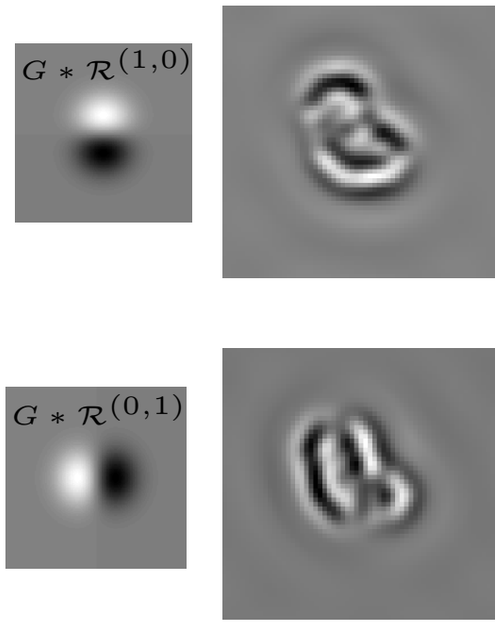


TEXTURE OPERATORS [DEPEURSINGE2017]

- Texture characterizes transitions between voxel values
  - A D-dimensional texture analysis approach is characterized by a set of N local operators quantifying transitions at the position $\boldsymbol{x}_0$

$$\mathcal{G}_n\{f\}(\boldsymbol{x}_0) \in \mathbb{R},\ n = 1, \ldots, N \qquad \text{with support} \qquad \boldsymbol{G}_n = G_{1,n} \times \cdots \times G_{D,n}$$

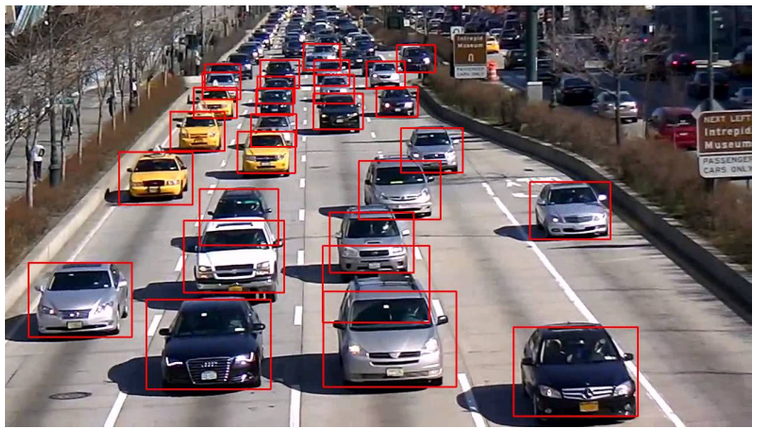




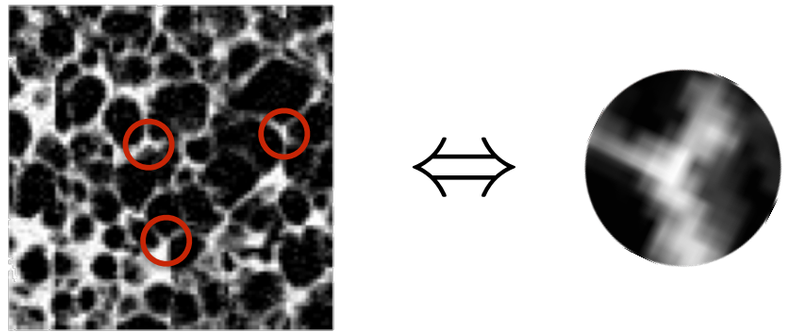
INVARIANCE OF TEXTURE OPERATORS [DEPEURSINGE2017]

- Global / local, invariances / equivariance of the texture operators to geometric transformations can be desirable
  - E.g., translation, rotation and scaling

cars in photographic imagery



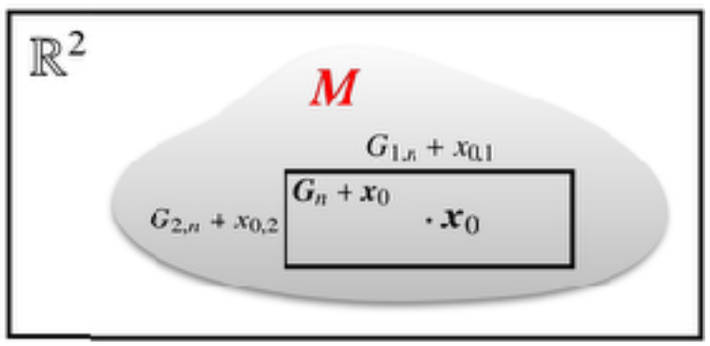
collagen junctions in lung CT

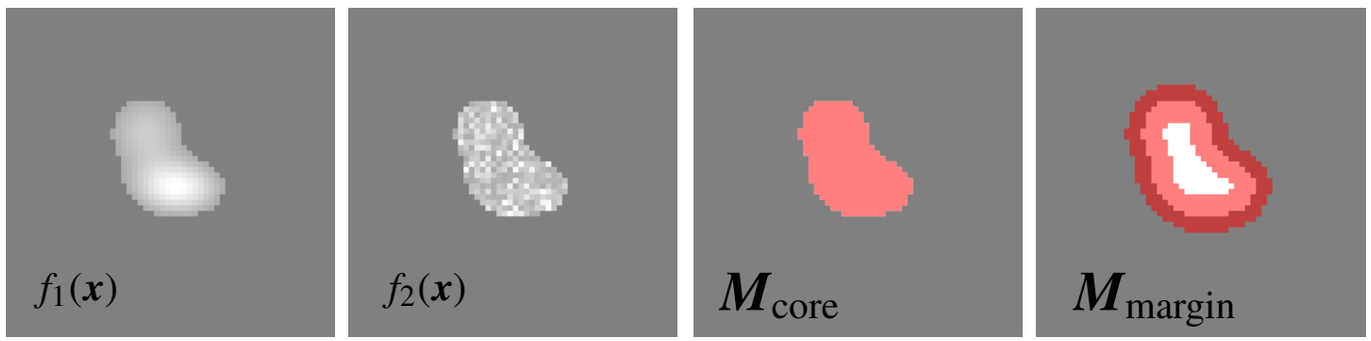


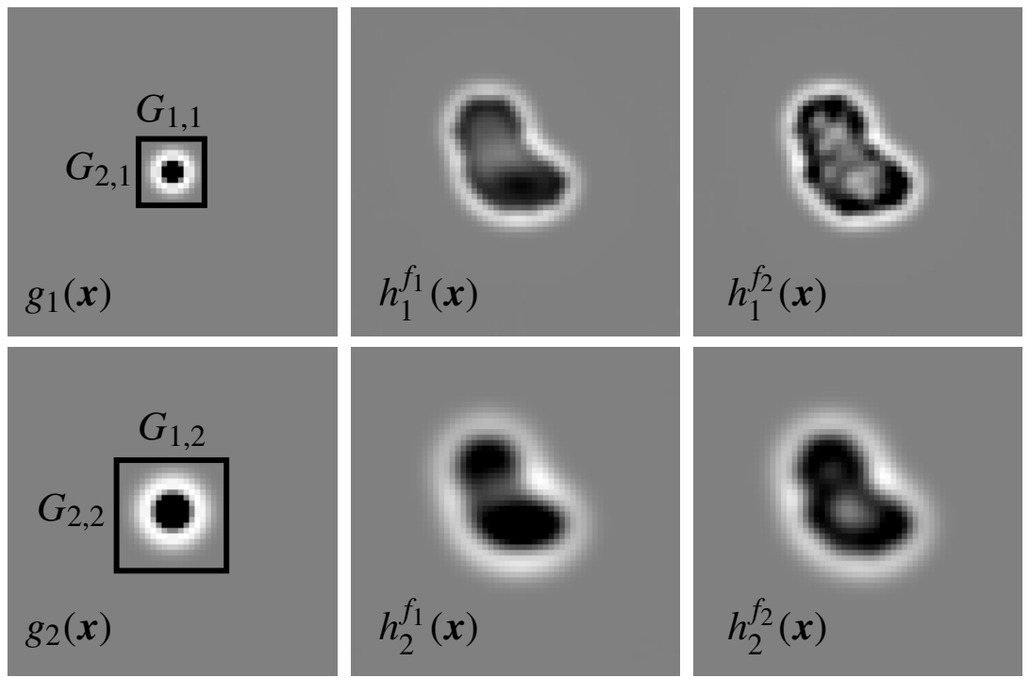
	Photographic image analysis	Biomedical image analysis
translation	global equivariance	global equivariance
rotation	no invariance	local invariance
scale	local invariance	no invariance

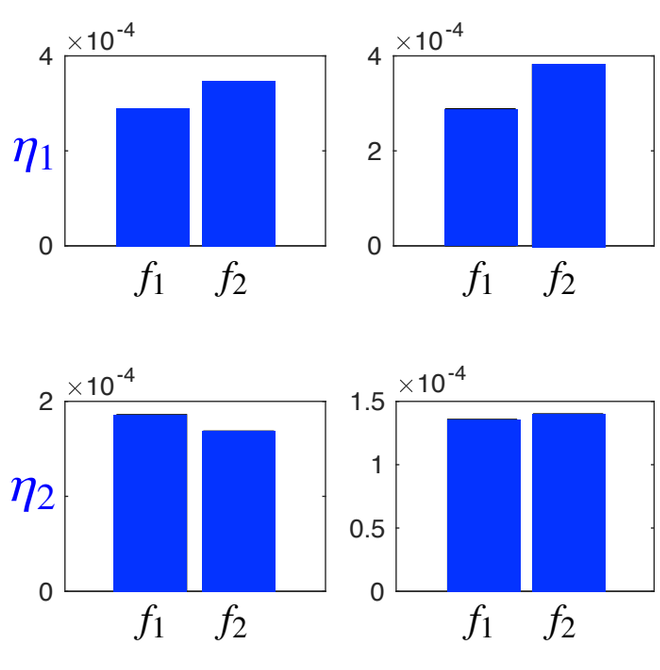
TEXTURE AGGREGATION [DEPEURSINGE2017]

- Texture characterizes transitions between voxel values
  - Operator's response maps $h_n(\boldsymbol{x}_0) = \mathcal{G}_n\{f\}(\boldsymbol{x}_0)$ must be aggregated over a region of interest (ROI) to get a scalar texture measurement η
    - E.g., provide estimates of features statistics



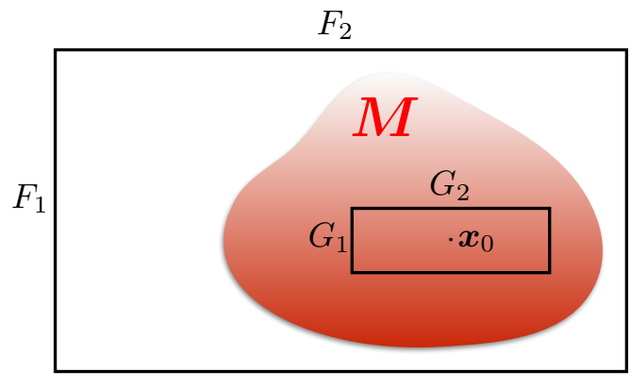




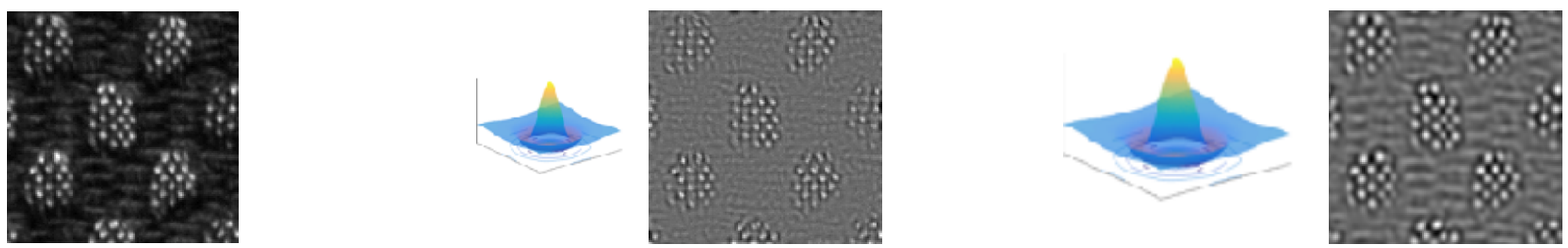


INTEGRATIVE AGGREGATION FUNCTIONS

- Undesirable effets of averaging
  - Example with operator: Laplacians of Gaussian (LoG) $\{g_1, g_2\}$ applied to all image positions $\boldsymbol{x}_0 \in F_1 \times F_2$

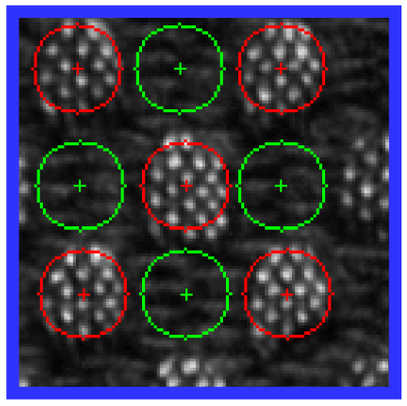


$f(\boldsymbol{x})$ $g_1(\boldsymbol{x})$ $h_1(\boldsymbol{x})$ $g_2(\boldsymbol{x})$ $h_2(\boldsymbol{x})$

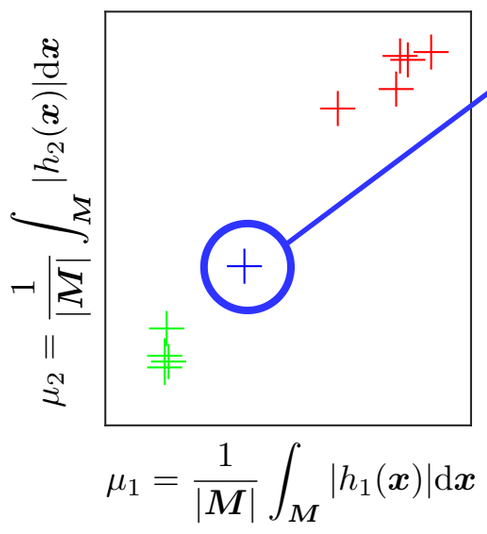


  - Operators' responses are averaged over $\boldsymbol{M}$

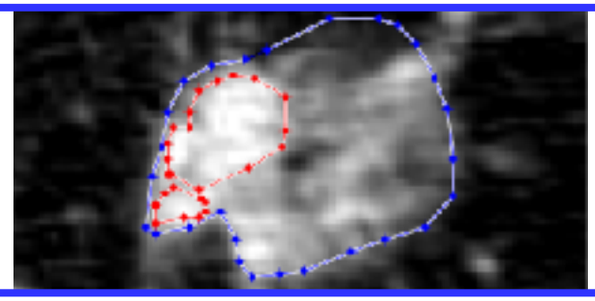
original image $f(\boldsymbol{x})$ with regions $\boldsymbol{M}_a, \boldsymbol{M}_b, \boldsymbol{M}_c$



feature space



The averaged responses over the entire image does not correspond to anything visually!



Nor biologically!

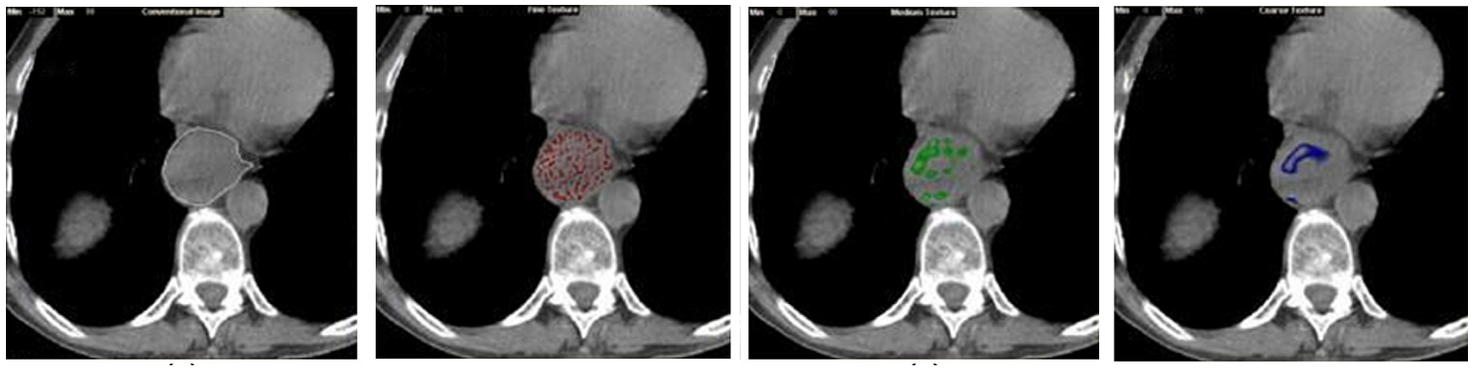
- Feature covariances can be better for aggregation [Cirujeda2016}

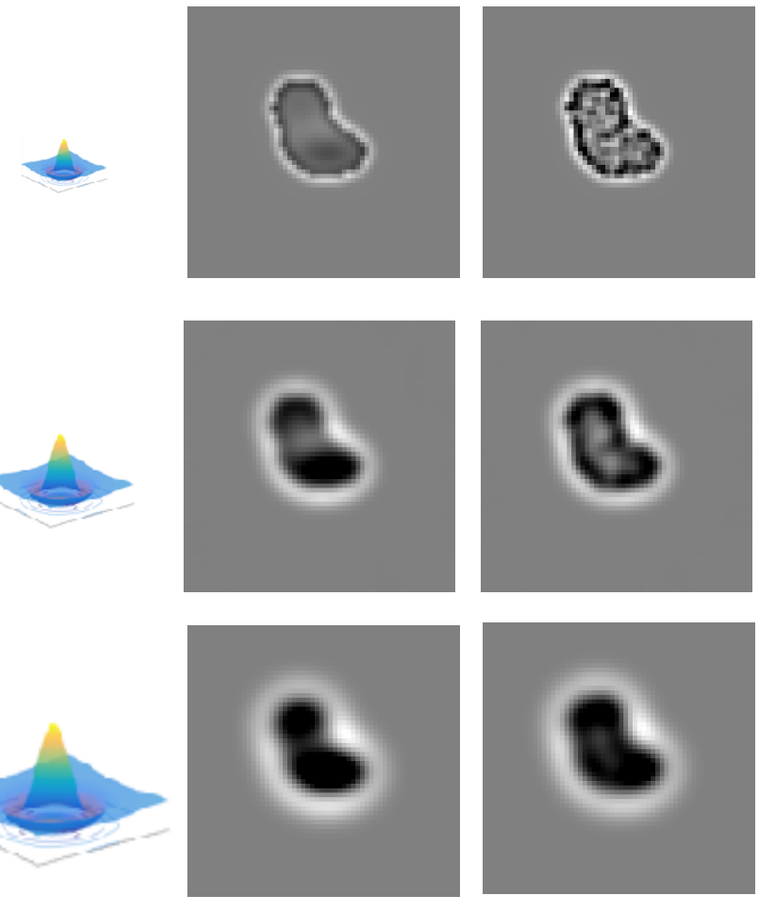
Outline

- Background - Radiomics
  - Personalized tumor phenotyping in PET - CT
- Methods
  - Image mining: intensity versus texture
  - Texture analysis: definition and qualitative review
  - The QuantImage platform
- Results
  - Head and Neck cancer: importance of geometry
  - Lung cancer: importance of aggregation
- Conclusions and perspectives

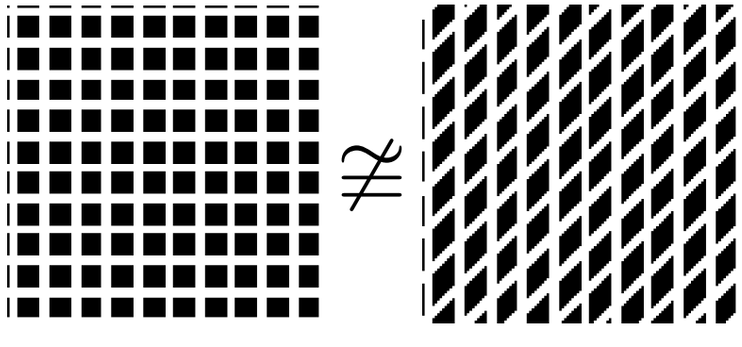
COMMON RADIOMICS TEXTURE FEATURES [DEPEURSINGE2017C]

- Laplacian of Gaussian (LoG) filters
  - Operator: second-order derivative of a Gaussian
  - Texture at multiple image scales (varying σ_{LoG})
  - Widely used (*e.g.*, TEXRAD [Ganeshan2012])



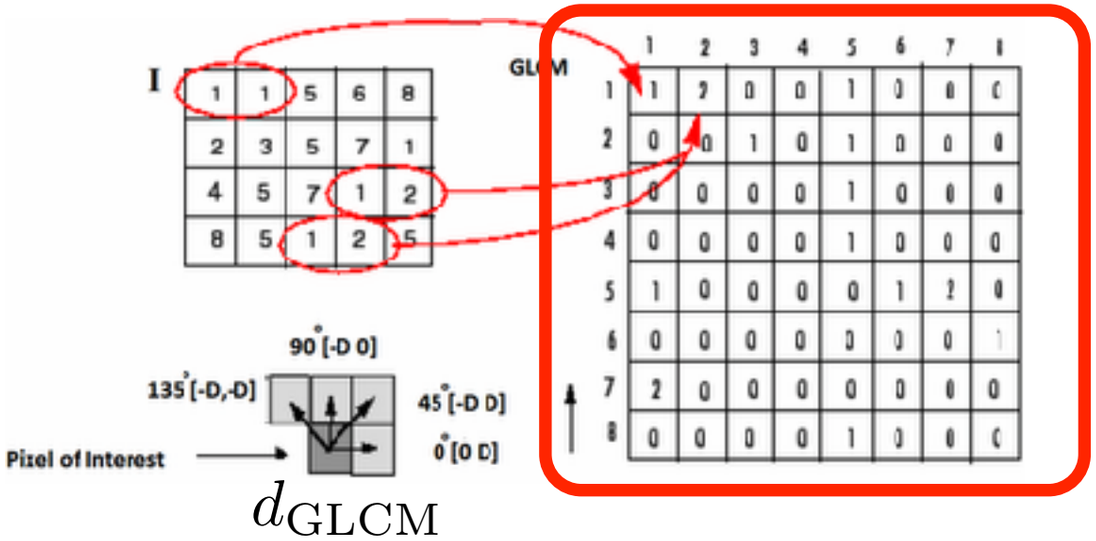


- Advantages
  - Translation-equivariant, locally rotation-invariant, easily extendable to 3D
- Drawbacks
  - Insensitive to directional texture patterns



COMMON RADIOMICS TEXTURE FEATURES [DEPEURSINGE2017C]

- Gray level co-occurence matrices (GLCM)
  - Operator: counting co-occurrence of pixel values
  - Widely used (*e.g.*, [Fried2016], LIFEx soft)

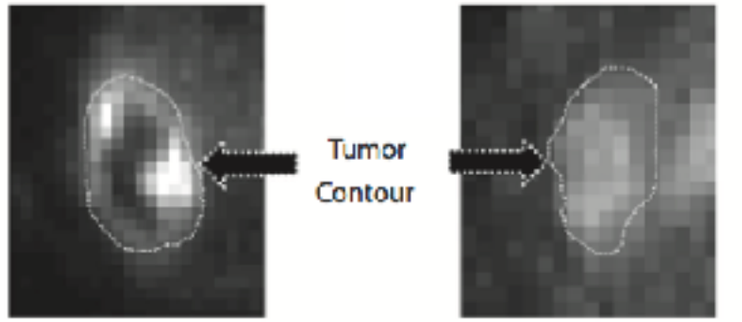


$$\mu_{\mathrm{energy}} = \sum_{i,j} p(i,j)^2$$

$$\mu_{\mathrm{contrast}} = \sum_{i,j} |i-j|^2 p(i,j)$$

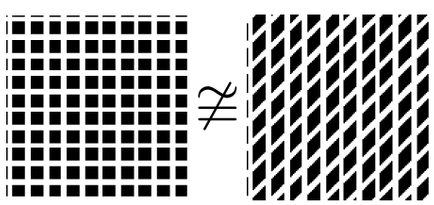
$$\vdots$$

$$\begin{pmatrix} \mu_1^1 \\ \vdots \\ \mu_P^1 \\ \mu_1^2 \\ \vdots \\ \mu_P^2 \\ \vdots \\ \mu_1^N \\ \vdots \\ \mu_P^N \end{pmatrix} = \boldsymbol{\mu}_M$$



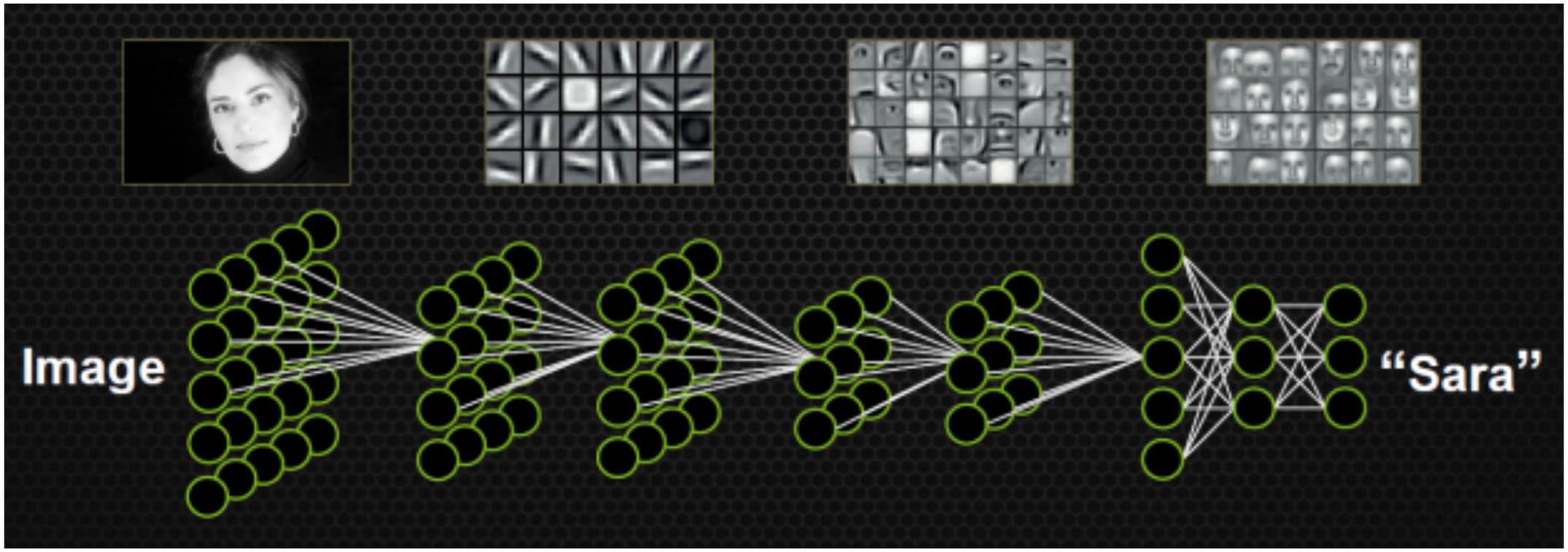
low μ_{energy} high μ_{energy}

- Advantages
  - Translation-equivariant, easily extendable to 3D
- Drawbacks
  - Not locally rotation-invariant (unless averaged over directions), poor multi-scale characterization, requires arbitrary gray-level reduction



Common radiomics texture features [Depeursinge2017c]

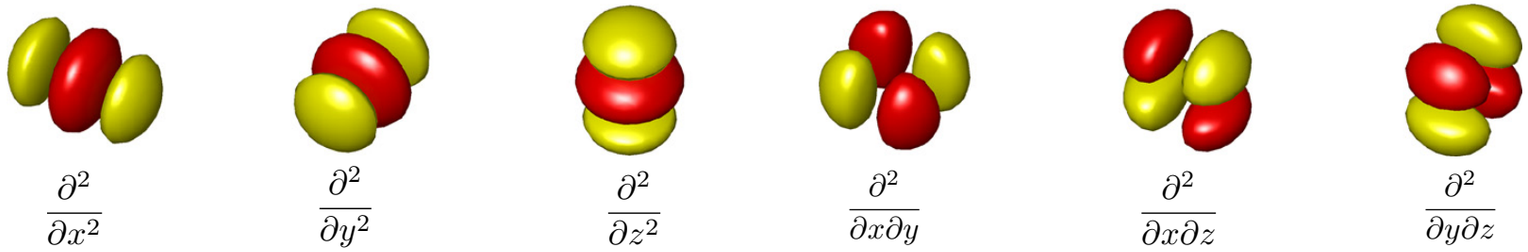
- Deep convolutional neural networks (CNNs) [Andrearczyk2017]



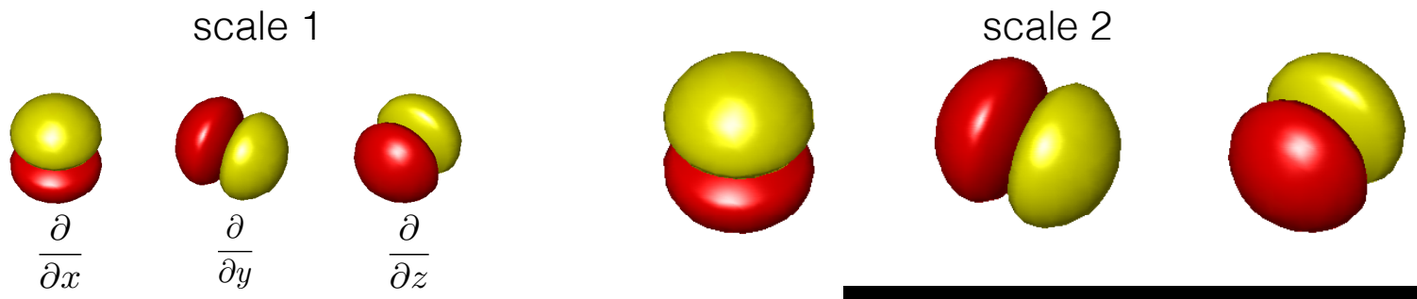
- Advantages
  - Translation-equivariant, extremely versatile (learns operators)
- Drawbacks
  - Not locally rotation-invariant, requires huge amount of training data in 3D

COMMON RADIOMICS TEXTURE FEATURES [DEPEURSINGE2017C]

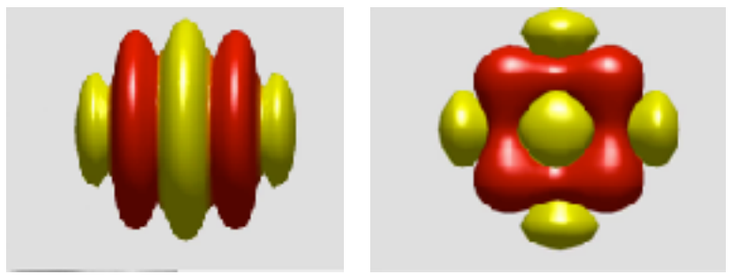
- Locally-oriented 3D Riesz wavelets [Chenouard2012, Dicente2017b]
  - Operator: directional filters behaving like local partial image derivatives
    - E.g., second-order:

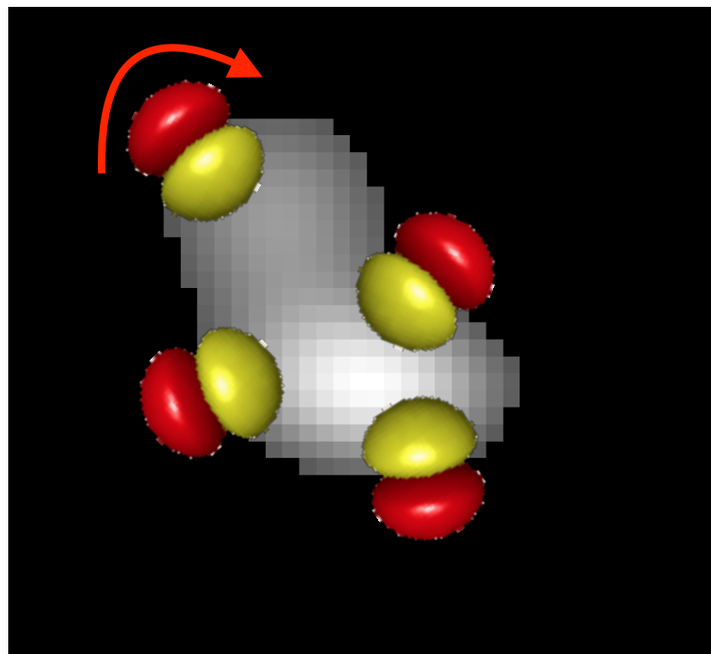


    - Suitable for exploring first- and higher-order transitions between voxel values
  - Multi-scale (wavelets)



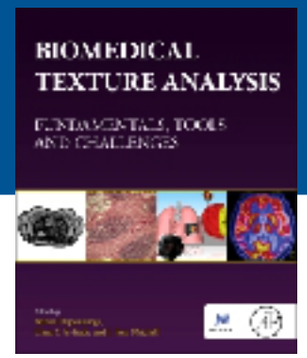
  - Steerable
    - Finds the 3D direction maximizing local image derivatives
    - Combines directional sensitivity with local rotation-invariance



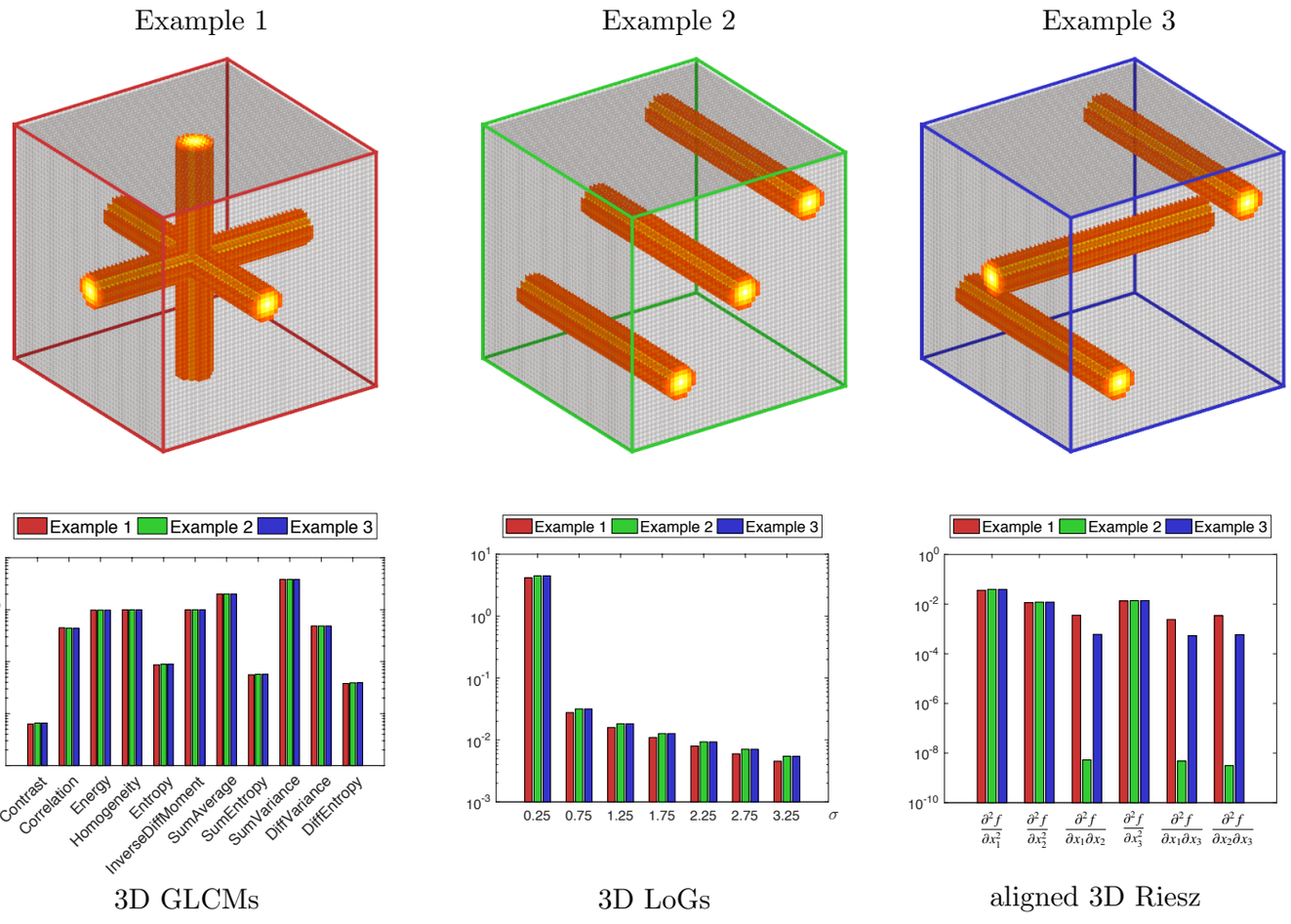


ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

INTERPRETATION OF GEOMETRY [DEPEURSINGE2017B]



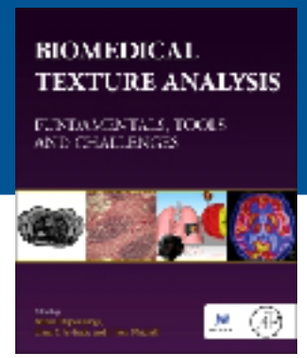
- Directionally-insensitive texture operators (e.g., locally rotation-invariant GLCMs or LoGs) cannot distinguish between the following patterns



3D GLCMs | 3D LoGs | aligned 3D Riesz

- Importance of directional sensitivity combined with local rotation invariance [Dicente2017, Depeursinge2017b]

INTERPRETATION OF GEOMETRY [DEPEURSINGE2017B]



- Importance of directional sensitivity combined with local rotation invariance [Dicente2017, Depeursinge2017b]

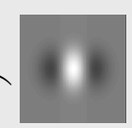
GLCM contrast
$d = 1 \quad (\Delta_{k_1} = 0, \Delta_{k_2} = 1)$

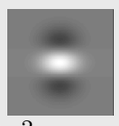
GLCMs

GLCM contrast
$d = 1 \quad (\Delta_{k_1} = 1, \Delta_{k_2} = 0)$

Riesz wavelets ($N = 2$)

$$\frac{1}{|\boldsymbol{M}|} \int_{\boldsymbol{M}} \left(g_{(2,0)}(f(\boldsymbol{x}), \boldsymbol{m}) \right)^2 \mathrm{d}\boldsymbol{m}$$

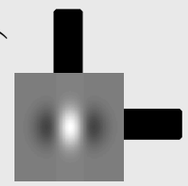


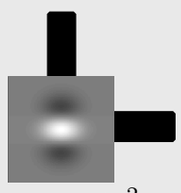


$$\frac{1}{|\boldsymbol{M}|} \int_{\boldsymbol{M}} \left(g_{(0,2)}(f(\boldsymbol{x}), \boldsymbol{m}) \right)^2 \mathrm{d}\boldsymbol{m}$$

aligned Riesz wavelets ($N = 2$)

$$\frac{1}{|\boldsymbol{M}|} \int_{\boldsymbol{M}} \left(g_{(2,0)}(f(\boldsymbol{R}_{\theta_{\max}(\boldsymbol{m})}\, \boldsymbol{x}), \boldsymbol{m}) \right)^2 \mathrm{d}\boldsymbol{m}$$





$$\frac{1}{|\boldsymbol{M}|} \int_{\boldsymbol{M}} \left(g_{(0,2)}(f(\boldsymbol{R}_{\theta_{\max}(\boldsymbol{m})}\, \boldsymbol{x}), \boldsymbol{m}) \right)^2 \mathrm{d}\boldsymbol{m}$$

Outline

- Background - Radiomics
  - Personalized tumor phenotyping in PET - CT
- Methods
  - Image mining: intensity versus texture
  - Texture analysis: definition and qualitative review
  - The QuantImage platform
- Results
  - Head and Neck cancer: importance of geometry
  - Lung cancer: importance of aggregation
- Conclusions and perspectives

QUANTIMAGE CLOUD PLATFORM [DICENTE2017]

- Free access, encrypted data transfer
- Image alignment and normalization
- State-of-the-art 3D radiomics features
  - Intensity (incl. PET specific)
  - 3D texture
- Part of the image biomarker standardisation initiative (IBSI) [Zwanenburg2017]

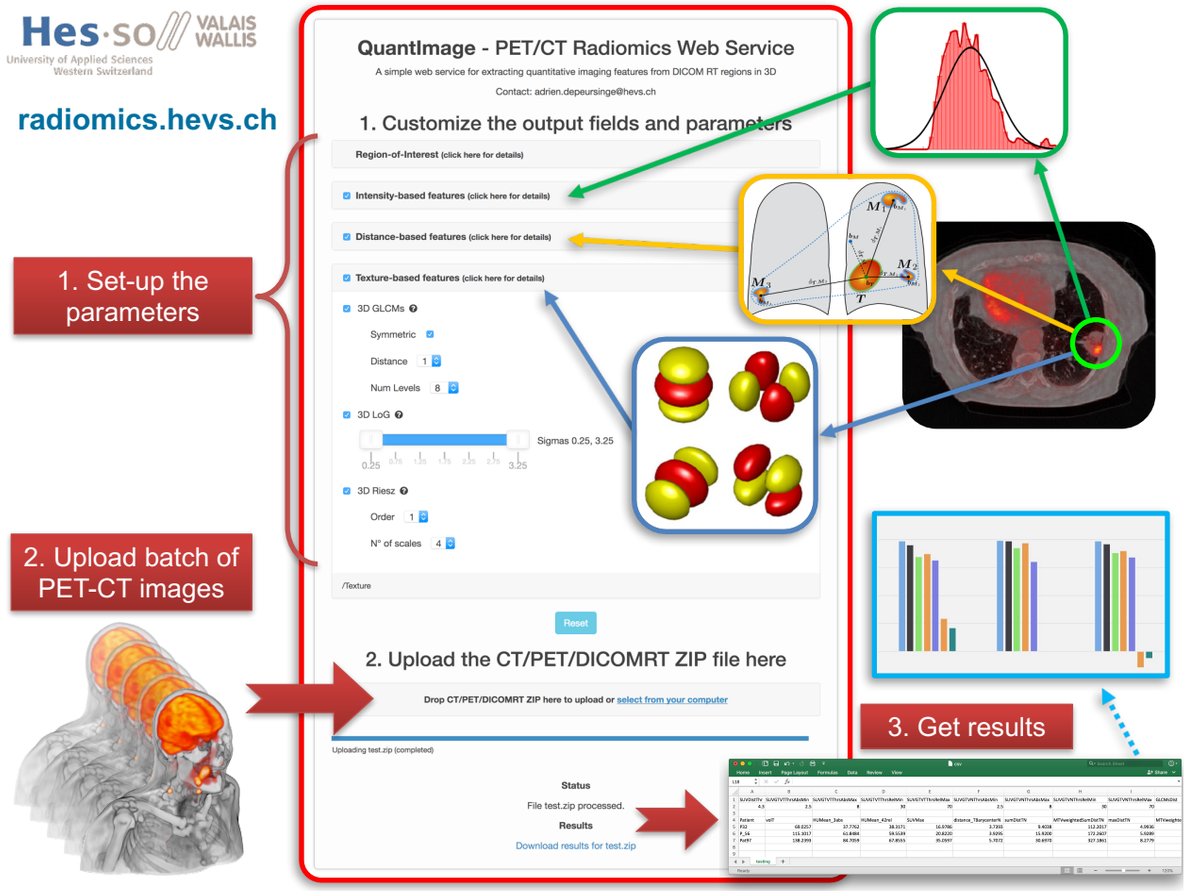
Hes·so VALAIS WALLIS
School of Management & Tourism

CHUV

EPFL
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Hes·so VALAIS WALLIS
University of Applied Sciences Western Switzerland

radiomics.hevs.ch



MOFFITT CANCER CENTER

Inserm

MGH 1811

University Hospital Zurich

MAASTRO

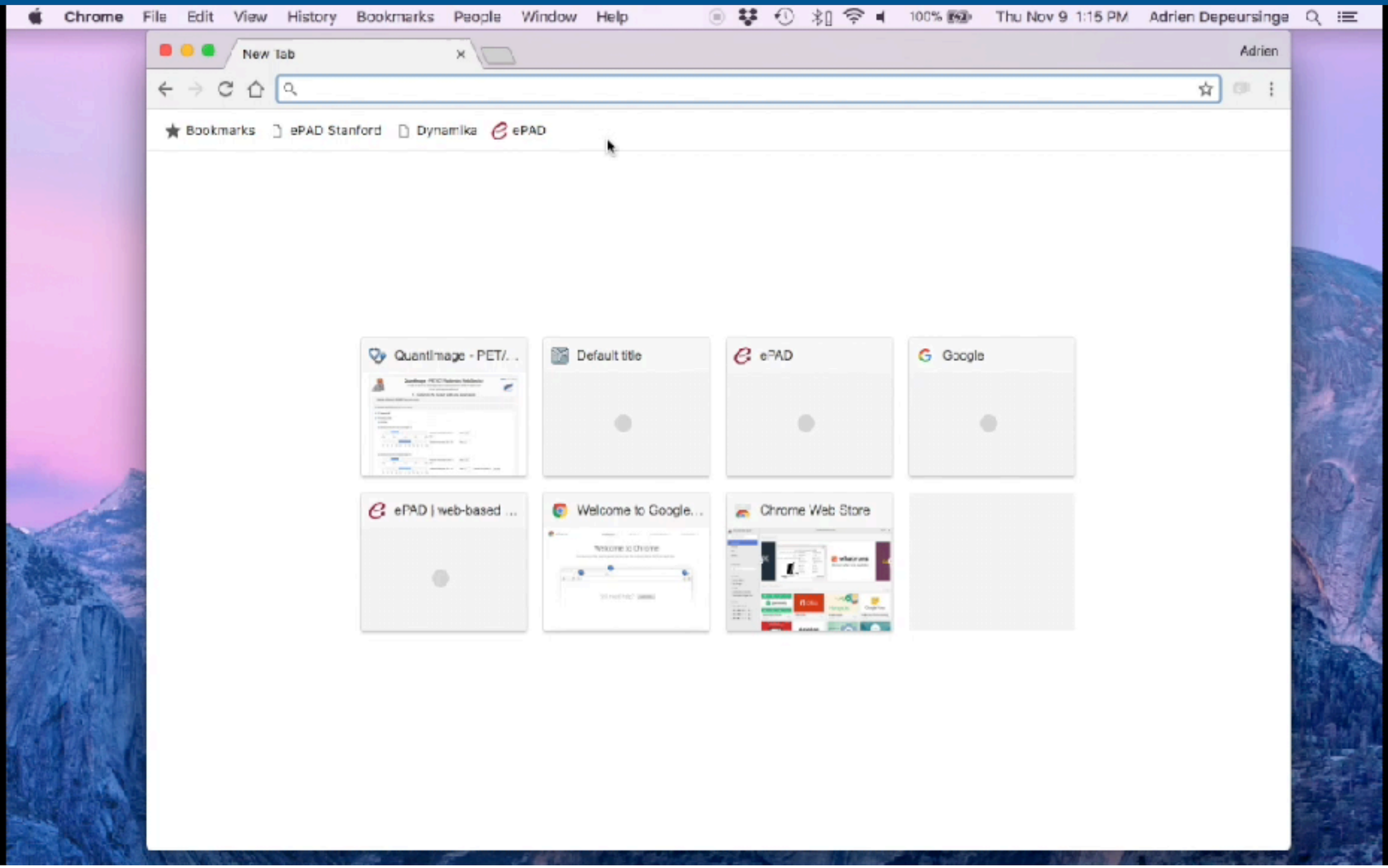
MEMORIAL SLOAN KETTERING CANCER CENTER

HARVARD MEDICAL SCHOOL

NCT NATIONAL CENTER FOR TUMOR DISEASES HEIDELBERG

Hes·so VALAIS WALLIS
School of Management & Tourism

QUANTIMAGE CLOUD PLATFORM [DICENTE2017]



https://radiomics.hevs.ch

Outline

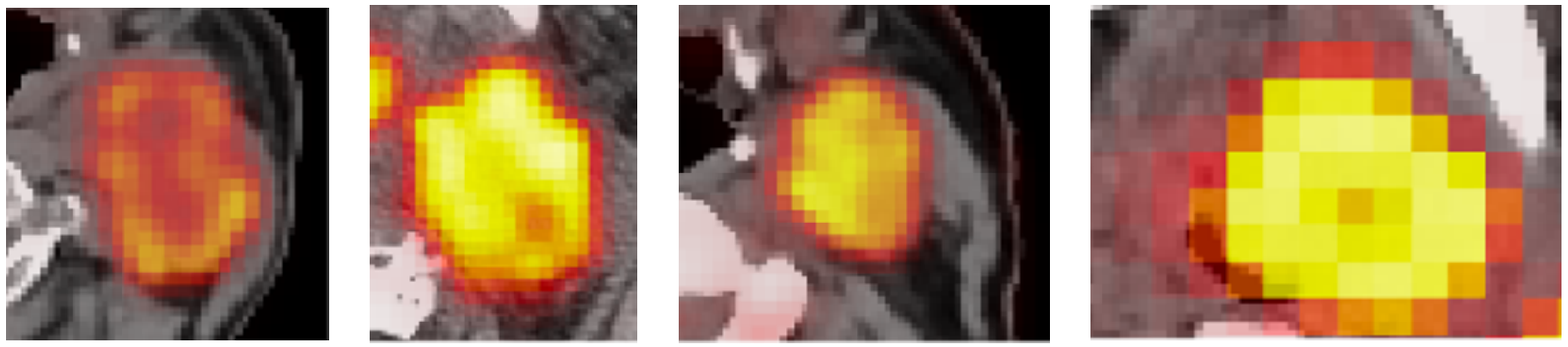
- Background - Radiomics
  - Personalized tumor phenotyping in PET - CT
- Methods
  - Image mining: intensity versus texture
  - Texture analysis: definition and qualitative review
  - The QuantImage platform
- Results
  - Head and Neck cancer: importance of geometry
  - Lung cancer: importance of aggregation
- Conclusions and perspectives

Head and Neck Cancer [Depeursinge2017b]

- Visual analysis of F-FDG PET images allows more accurate staging than American Joint Committee on Cancer (AJCC)[Yoo2013]
- Quantitative analyses of metabolic intensity could predict Overall Survival (OS) and Disease-Free Survival (DFS) [Castelli2017]
- The interleaved sub-tumoral regions of proliferating cancer cells and necrosis results in metabolic heterogeneity

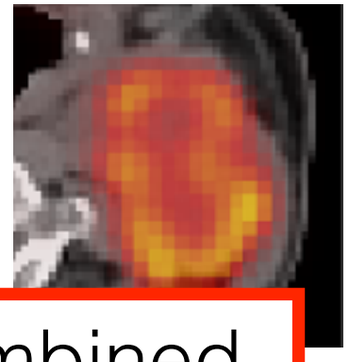
1. Use 3D texture analysis of PET images to quantify internal metabolism morphology ?
2. Use it to further predict OS and DFS ?
3. Differences between texture analysis approaches ?

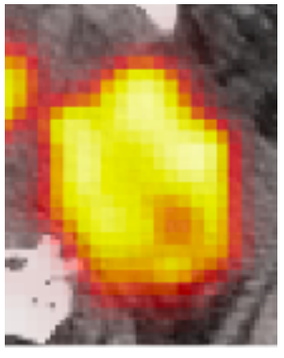
$\mathcal{H}_0$



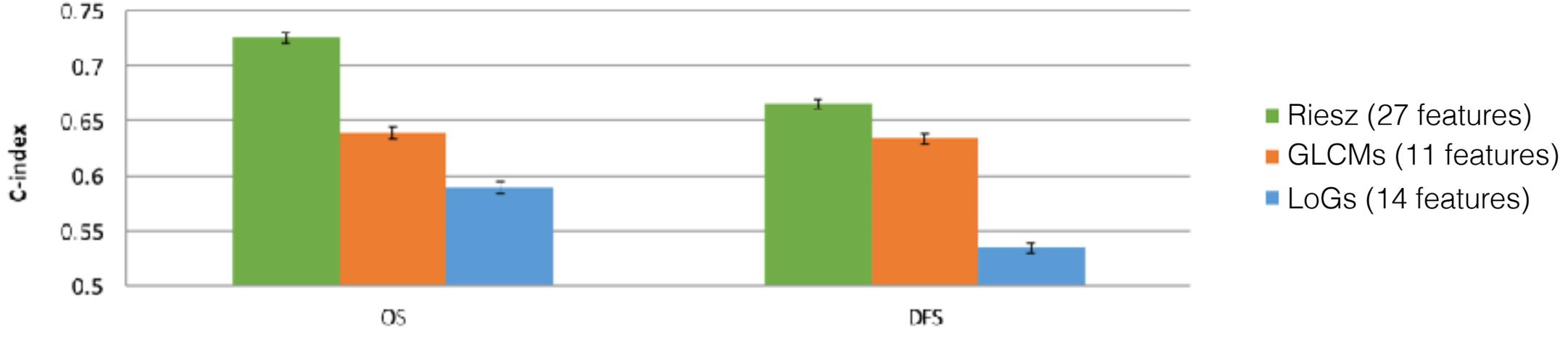
Head and Neck Cancer [Depeursinge2017b]

- 108 patients: 62 from Rennes and 46 from Lausanne
  - PET >8 weeks before RT, no metastasis at diagnosis
- Gross Tumor Volume (GTV) manually segmented on each PET/CT
- 6-months minimal follow-up (OS: 40 events, DFS: 47 events)
- Texture features parameters optimized
- Cox LASSO regression model [Simon2011]
- Stra





Importance of directional sensitivity combined with local rotation invariance

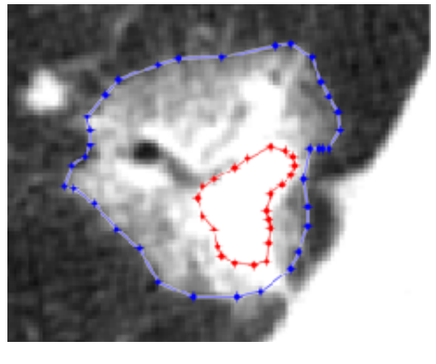


Outline

- Background - Radiomics
  - Personalized tumor phenotyping in PET - CT
- Methods
  - Image mining: intensity versus texture
  - Texture analysis: definition and qualitative review
  - The QuantImage platform
- Results
  - Head and Neck cancer: importance of geometry
  - Lung cancer: importance of aggregation
- Conclusions and perspectives

LUNG CANCER [CIRUJEDA2016]

- 92 lung adenocarcinoma from Stanford Hospital and Clinics
  - Gross Tumor Volume (GTV), Ground Glass Opacities (GGO) and solid regions $\boldsymbol{M}$ contoured in CT in 3D
  - DFS times available
    - 12 months cutoff (23 recurrences versus 69 remissions)

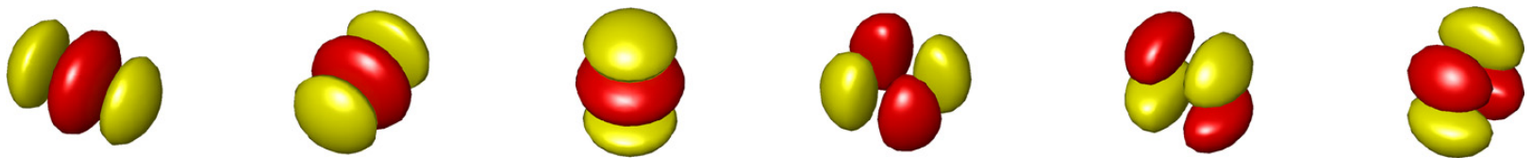


1. Difference when using GTV, GGO or solid ROIs for aggregation?
2. Difference between average and covariance for aggregation?

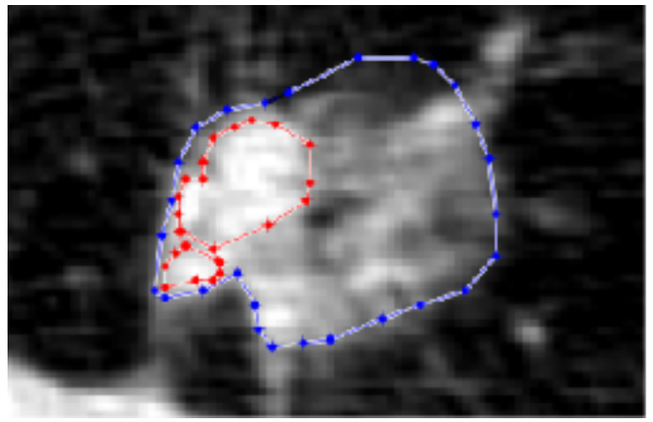
$\mathcal{H}_0$

LUNG CANCER [CIRUJEDA2016]

- 2nd-order aligned Riesz, 3 scales (18 features)



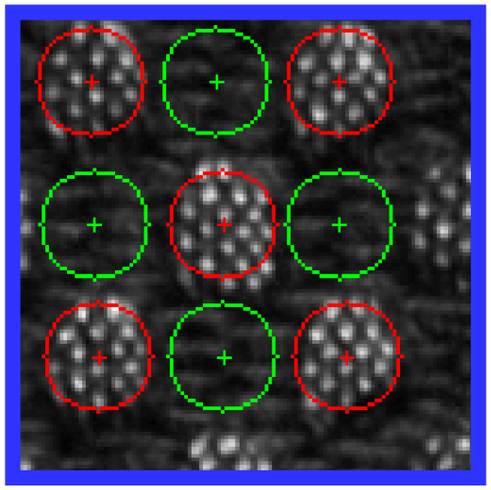
- Support vector machines (SVM)
  - Average versus covariance kernel
- 10-fold CV (5 repetitions)



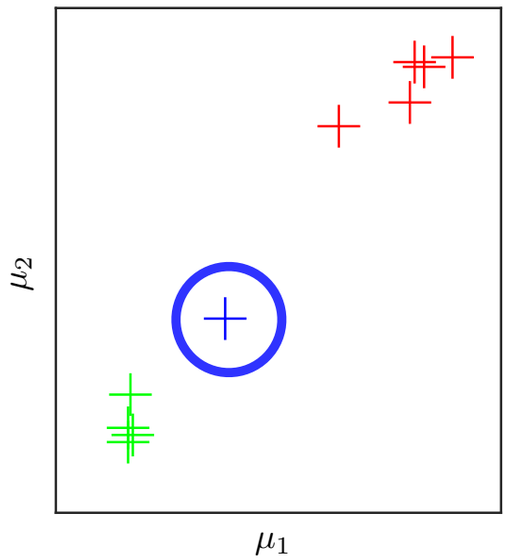
	AVERAGE-BASED SVMS		COVARIANCE-BASED SVMS	
	sensitivity	specificity	sensitivity	specificity
GGO	77.65 ± 0.14	68.97 ± 0.07	87.38 ± 0.05	78.33 ± 0.13
Solid	85.96 ± 0.09	76.45 ± 0.11	85.14 ± 0.13	76.67 ± 0.14
GTV	83.17 ± 0.15	70.24 ± 0.17	87.62 ± 0.05	78.33 ± 0.13

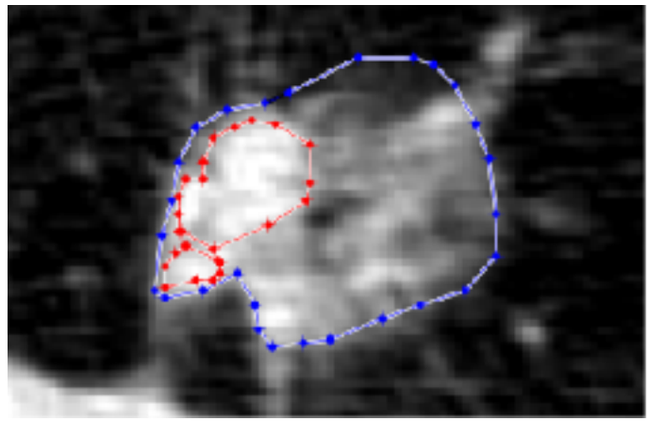
LUNG CANCER [CIRUJEDA2016]

original image $f(\boldsymbol{x})$ with regions $\boldsymbol{M}_a, \boldsymbol{M}_b, \boldsymbol{M}_c$



feature space





	AVERAGE-BASED SVMs		COVARIANCE-BASED SVMs	
	sensitivity	specificity	sensitivity	specificity
GGO	77.65 ± 0.14	68.97 ± 0.07	87.38 ± 0.05	78.33 ± 0.13
Solid	85.96 ± 0.09	76.45 ± 0.11	85.14 ± 0.13	76.67 ± 0.14
GTV	83.17 ± 0.15	70.24 ± 0.17	87.62 ± 0.05	78.33 ± 0.13

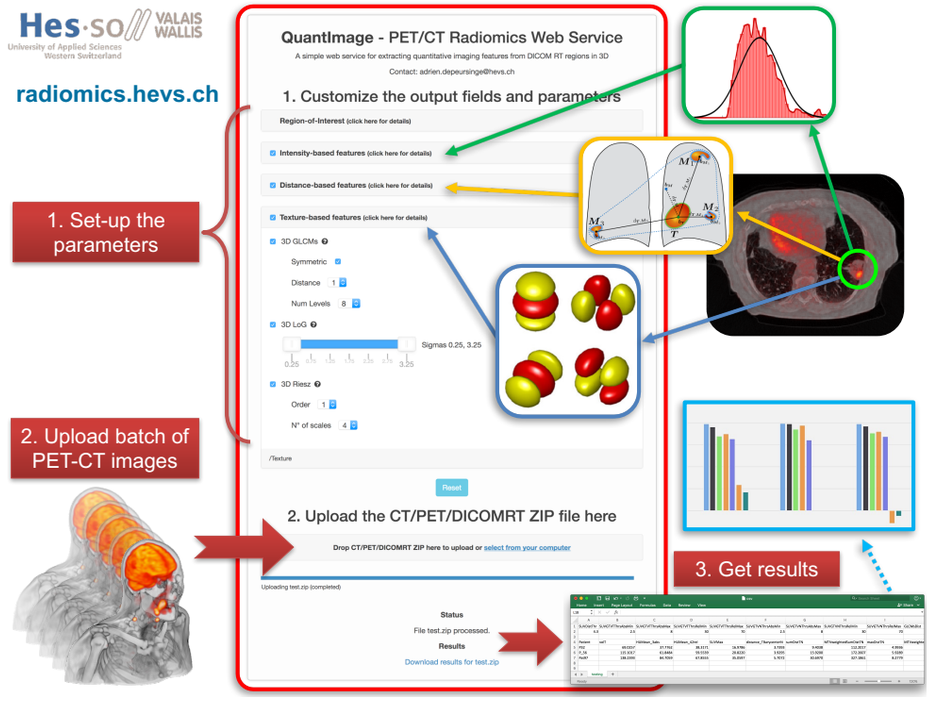
Importance of the feature aggregation function to avoid mixing tumor habitats

Outline

- Background - Radiomics
  - Personalized tumor phenotyping in PET - CT
- Methods
  - Image mining: intensity versus texture
  - Texture analysis: definition and qualitative review
  - The QuantImage platform
- Results
  - Head and Neck cancer: importance of geometry
  - Lung cancer: importance of aggregation
- Conclusions and perspectives

Conclusions & Perspectives

- Internal tumor structure (CT) and metabolism (PET) morphology could be modeled with 3D texture to predict clinical outcomes
  - OS, DFS in Head and Neck as well as Lung cancer
  - Large differences between texture analysis approaches
    - Importance directional sensitivity and local rotation invariance
    - Impact of the feature aggregation function
- Online tools available
  - The QuantImage platform for 3D PET/CT
- Limitations and perspectives
  - Validation on large and independent cohorts
  - Protocol and features standardization



- THANKS !

Hes·so VALAIS WALLIS
University of Applied Sciences Western Switzerland

Roger Schaer
Yashin Dicente
Pol Cirujeda
Henning Müller
MEDGIFT group

CHUV

John O. Prior
Niklaus Schaefer

UNICANCER
Centre Eugène Marquis

Joel Castelli

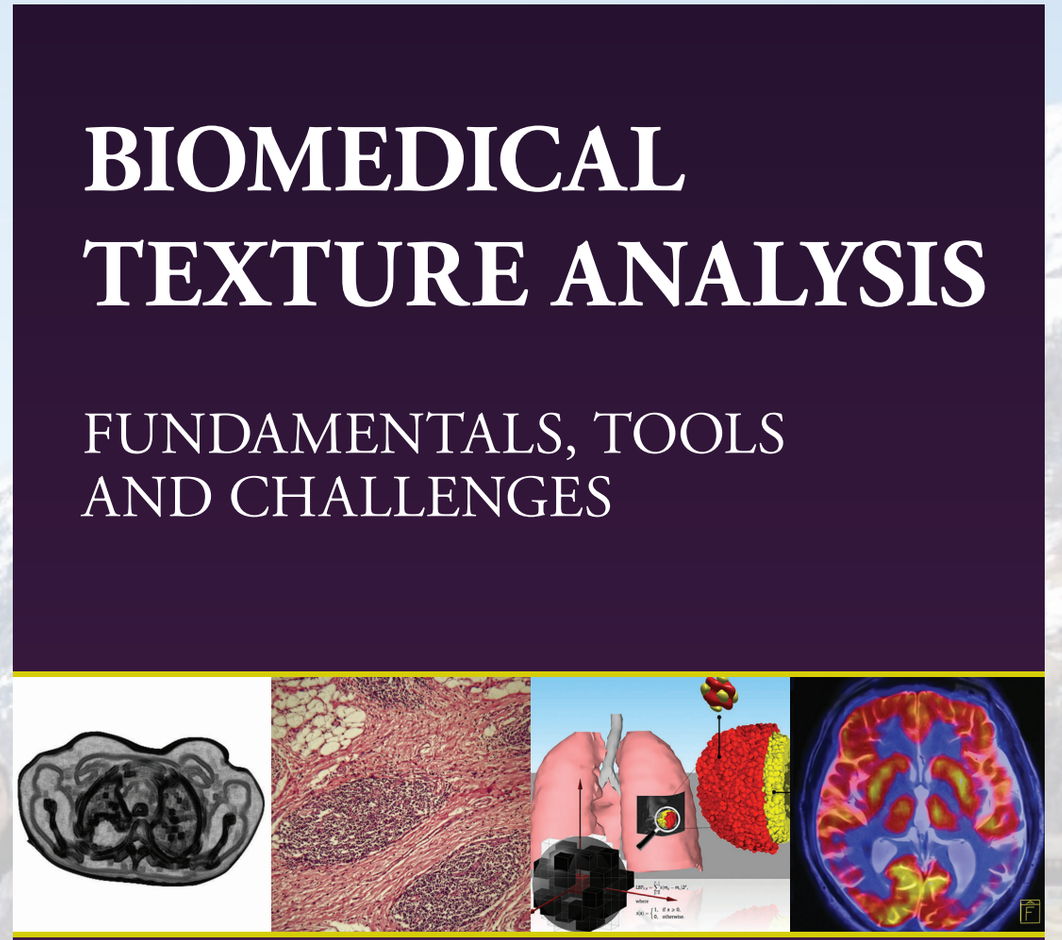
EPFL
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Julien Fageot
Michaël Unser
BIG group

STANFORD SCHOOL OF MEDICINE
Stanford University Medical Center

Daniel L. Rubin
Todd Aguilera
Sandy Napel
Olivier Gevaert

CIBM



EPFL
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

FNSNF
SWISS NATIONAL SCIENCE FOUNDATION

RCSO
réseau de compétences HES-SO
Information Services Network

Hes·so VALAIS WALLIS
School of Management & Tourism

References (sorting in alphabetical order)

[Aerts2014] Aerts, H. J. W. L.; Velazquez, E. R.; Leijenaar, R. T. H.; Parmar, C.; Grossmann, P.; Carvalho, S.; Bussink, J.; Monshouwer, R.; Haibe-Kains, B.; Rietveld, D.; Hoebers, F.; Rietbergen, M. M.; Leemans, C. R.; Dekker, A.; Quackenbush, J.; Gillies, R. J. & Lambin, P.
Decoding Tumour Phenotype by Noninvasive Imaging Using a Quantitative Radiomics Approach
Nature Communications, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved, 2014, 5

[Andrearczyk2017] Andrearczyk, V. & Whelan P. F.
Deep Learning in Texture Analysis and Its Application to Tissue Image Classification.
Biomedical Texture Analysis: Fundamentals, Applications and Tools, Elsevier-MICCAI Society Book series, Elsevier, 2017, 95-129

[Castelli2017] Castelli, J.; Depeursinge, A.; Ndoh, V.; Prior, J. O.; Ozsahin, M.; Devillers, A.; Bouchaab, H.; Chajon, E.; de Crevoisier, R.; Scher, N.; Jegoux, F.; Laguerre, B.; De Bari, B.; Bourhis, J.
A PET-based nomogram for oropharyngeal cancers
European Journal of Cancer, 2017, 75, 222-230

[Chenouard2012] Chenouard, N. & Unser, M.
3D Steerable Wavelets in Practice
IEEE Transactions on Image Processing, 2012, 21, 4522-4533

[Cirujeda2016] Cirujeda, P.; Müller, H.; Rubin, D.; Aguilera, T. A.; Jr., B. W. L.; Diehn, M.; Binefa, X. & Depeursinge, A.
A 3-D Riesz-Covariance Texture Model for Prediction of Nodule Recurrence in Lung CT
IEEE Transactions on Medical Imaging, 2016, 35(12), 2620-2630

[Depeursinge2014] Depeursinge, A.; Foncubierta-Rodrguez, A.; Van De Ville, D. & Müller, H.
Three-Dimensional Solid Texture Analysis and Retrieval in Biomedical Imaging: Review and Opportunities
Medical Image Analysis, 2014, 18, 176-196

[Depeursinge2015] Depeursinge, A.; Yanagawa, M.; Leung, A. N. & Rubin, D. L.
Predicting Adenocarcinoma Recurrence Using Computational Texture Models of Nodule Components in Lung CT
Medical Physics, 2015,42, 2054-2063

[Depeursinge2017] Depeursinge, A.; Fageot, J. & Al-Kadi, O. S.
Fundamentals of Texture Processing for Biomedical Image Analysis: A General Definition and Problem Formulation
Biomedical Texture Analysis: Fundamentals, Applications and Tools, Elsevier-MICCAI Society Book series, Elsevier, 2017, 1-27

[Depeursinge2017b] Depeursinge, A.
Multi-Scale and Multi-Directional Biomedical Texture Analysis: Finding the Needle in the Haystack
Biomedical Texture Analysis: Fundamentals, Applications and Tools, Elsevier-MICCAI Society Book series, Elsevier, 2017, 29-53

[Depeursinge2017c] Depeursinge, A. & Fageot, J.
Biomedical Texture Operators and Aggregation Functions: A Methodological Review and User's Guide
Biomedical Texture Analysis: Fundamentals, Applications and Tools, Elsevier-MICCAI Society Book series, Elsevier, 2017, 55-94

References (sorting in alphabetical order)

[Depeursinge2017d] Depeursinge, A.; Dicente Cid, Y.; Schaer, R.; de Crevoisier, R.; Prior, J. O.; Castelli, J.
Comparing 18-FDG PET 3D texture attributes for the prediction of survival and recurrence in oropharyngeal cancers treated with radiotherapy
Workshop on the Prediction and Modeling of Response to Molecular and External Beam Radiotherapies, Le Bono, France, 2017

[Dicente2017] Dicente Cid, Y.; Castelli, J.; Schaer, R.; Scher, N. ; Pomoni, A. ; Prior, J. & Depeursinge, A.
QuantImage: An Online Tool for High-Throughput 3D Radiomics Feature Extraction in PET-CT
Biomedical Texture Analysis: Fundamentals, Applications and Tools, Elsevier-MICCAI Society Book series, Elsevier, 2017, 349-377

[Dicente2017b] Dicente Cid, Y.; Müller, H.; Platon, A.; Poletti, P.-A. & Depeursinge, A.
3-D Solid Texture Classification Using Locally-Oriented Wavelet Transforms
IEEE Transactions on Image Processing, 2017, 26(4), 1899-1910

[Fried2016] Fried, D. V.; Mawlawi, O.; Zhang, L.; Fave, X.; Zhou, S.; Ibbott, G.; Liao, Z. & Court, L. E.
Potential Use of 18F-fluorodeoxyglucose Positron Emission Tomography–Based Quantitative Imaging Features for Guiding Dose Escalation in Stage III Non-Small Cell Lung Cancer
International Journal of Radiation Oncology, 2016, 94, 368-376

[Ganeshan2012] Ganeshan, B.; Skogen, K.; Pressney, I.; Coutroubis, D. & Miles, K.
Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: Preliminary evidence of an association with tumour metabolism, stage, and survival
Clinical Radiology , 2012, 67, 157-164

[Ganeshan2013] Ganeshan, B.; Goh, V.; Mandeville, H. C.; Ng, Q. S.; Hoskin, P. J. & Miles, K. A.
Non-Small Cell Lung Cancer: Histopathologic Correlates for Texture Parameters at CT
Radiology, 2013, 266(1), 326-336

[Gatenby2013] Gatenby, R. A.; Grove, O. & Gillies, R. J.
Quantitative Imaging in Cancer Evolution and Ecology
Radiology, 2013, 269, 8-14

[Gerlinger2012] Gerlinger, M.; Rowan, A. J.; Horswell, S.; Larkin, J.; Endesfelder, D.; Gronroos, E.; Martinez, P.; Matthews, N.; Stewart, A.; Tarpey, P.; Varela, I.; Phillimore, B.; Begum, S.; McDonald, N. Q.; Butler, A.; Jones, D.; Raine, K.; Latimer, C.; Santos, C. R.; Nohadani, M.; Eklund, A. C.; Spencer-Dene, B.; Clark, G.; Pickering, L.; Stamp, G.; Gore, M.; Szallasi, Z.; Downward, J.; Futreal, P. A. & Swanton, C.
Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing
New England Journal of Medicine, 2012, 366, 883-892

[Haralick1973] Haralick, R. M.; Shanmugam, K. & Dinstein, I.
Textural Features for Image Classification
IEEE Transactions on Systems, Man and Cybernetics, 1973, 3, 610-621

[Kidd2008] Kidd, E. A. & Grigsby, P. W.
Intratumoral Metabolic Heterogeneity of Cervical Cancer
Clinical Cancer Research, 2008, 14, 5236-5241

References (sorting in alphabetical order)

[Kumar2012] Kumar, V.; Gu, Y.; Basu, S.; Berglund, A.; Eschrich, S. A.; Schabath, M. B.; Forster, K.; Aerts, H. J. W. L.; Dekker, A.; Fenstermacher, D.; Goldgof, D. B.; Hall, L. O.; Lambin, P.; Balagurunathan, Y.; Gatenby, R. A. & Gillies, R. J.
Radiomics: The Process and the Challenges
Magnetic Resonance Imaging, 2012, 30, 1234-1248

[Leijenaar2015] Leijenaar, R. T.; Nalbantov, G.; Carvalho, S.; van Elmpt, W. J.; Troost, E. G.; Boellaard, R.; Aerts, H. J.; Gillies, R. J. & Lambin, P.
The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis
Scientific Reports, 2015, 5, 11075

[Mattonen2014] Mattonen, S. A.; Palma, D. A.; Haasbeek, C. J. A.; Senan, S.; & Ward, A. D.
Early prediction of tumor recurrence based on CT texture changes after stereotactic ablative radiotherapy (SABR) for lung cancer
Medical Physics, 2014, 41(3),1-14

[Orlhac2014] Orlhac, F.; Soussan, M.; Maisonobe, J.-A.; Garcia, C. A.; Vanderlinden, B. & Buvat, I.
Tumor Texture Analysis in 18F-FDG PET: Relationships Between Texture Parameters, Histogram Indices, Standardized Uptake Values, Metabolic Volumes, and Total Lesion Glycolysis
Journal of Nuclear Medicine, 2014, 55, 414-422

[Ravanelli2013] Ravanelli, M.; Farina, D.; Morassi, M.; Roca, E.; Cavalleri, G.; Tassi, G. & Maroldi, R.
Texture analysis of advanced non-small cell lung cancer (NSCLC) on contrast-enhanced computed tomography: prediction of the response to the first-line chemotherapy
European Radiology, 2013, 23(12), 3450-3455

[Simon2011] Simon, N.; Friedman, J; Hastie, T. & Tibshirani, R.
Regularization paths for Cox's proportional hazards model via coordinate descent
Journal of Statistical Software, 2011, 39(5), 1-13

[Yoo2013] Yoo, J.; Henderson, S. & Walker-Dilks, C.
Evidence-based Guideline Recommendations on the Use of Positron Emission Tomography Imaging in Head and Neck Cancer
Clinical Oncology, 2013, 25(4), e33-e66

[Zwanenburg2017] Zwanenburg, A.; Leger, S.; Vallières, M. & Löck, S.
Image biomarker standardisation initiative
CoRR. 2016;abs/1612.0.

Technische Universiteit
Eindhoven
University of Technology

Learning with less labels in medical image analysis

Veronika Cheplygina

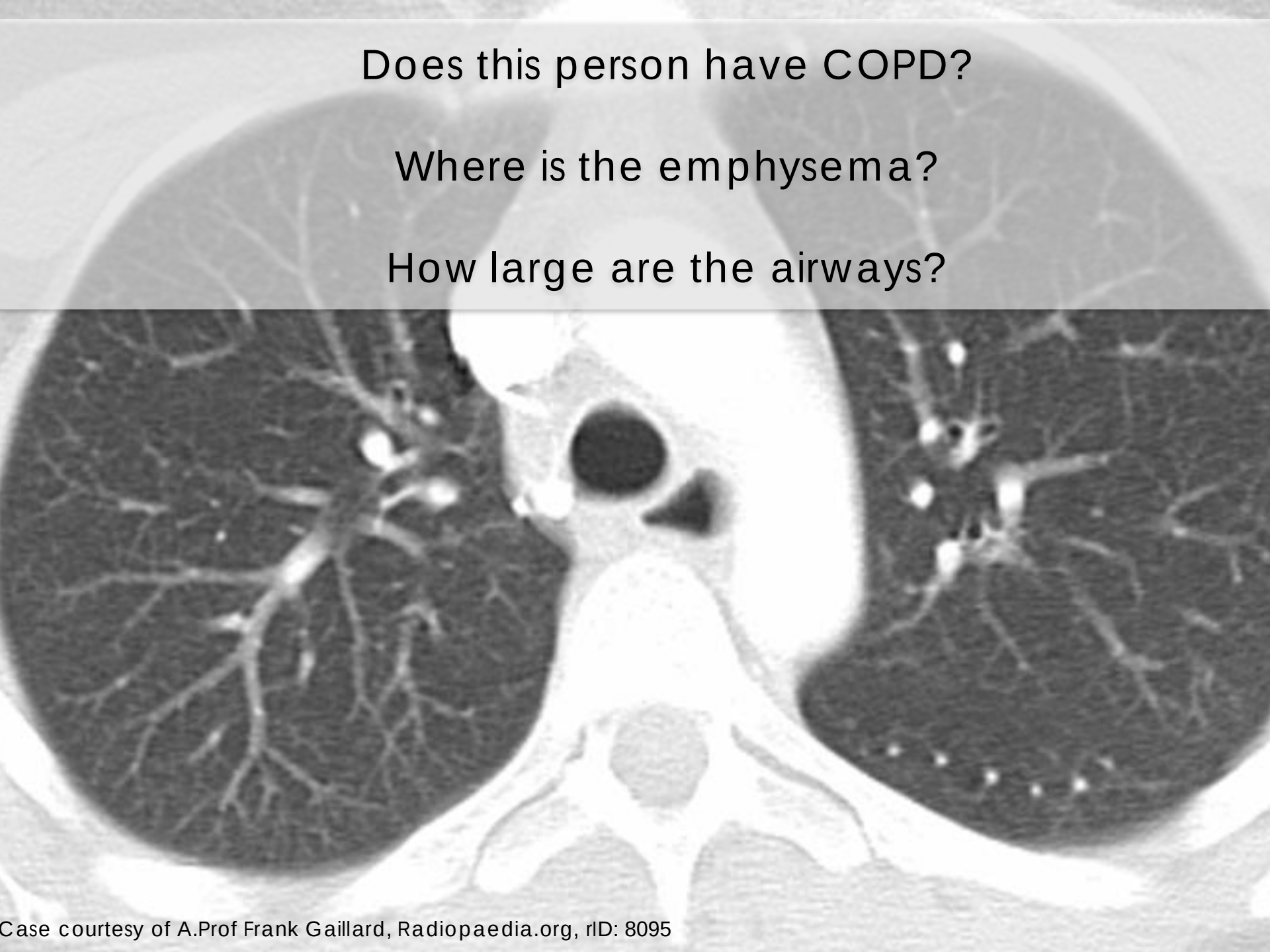
@vcheplygina

http://www.veronikach.com

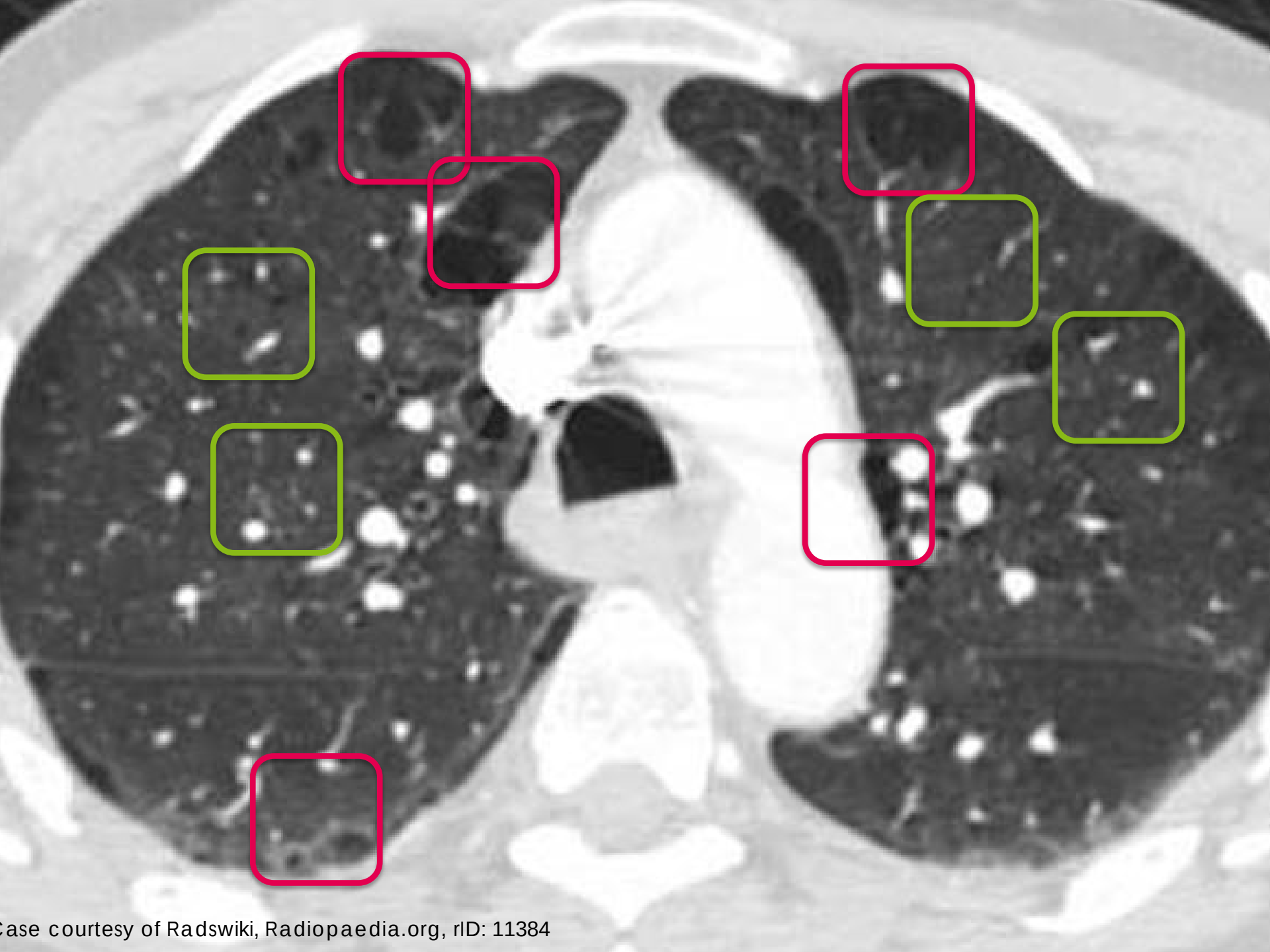
Does this person have COPD?

Where is the emphysema?

How large are the airways?

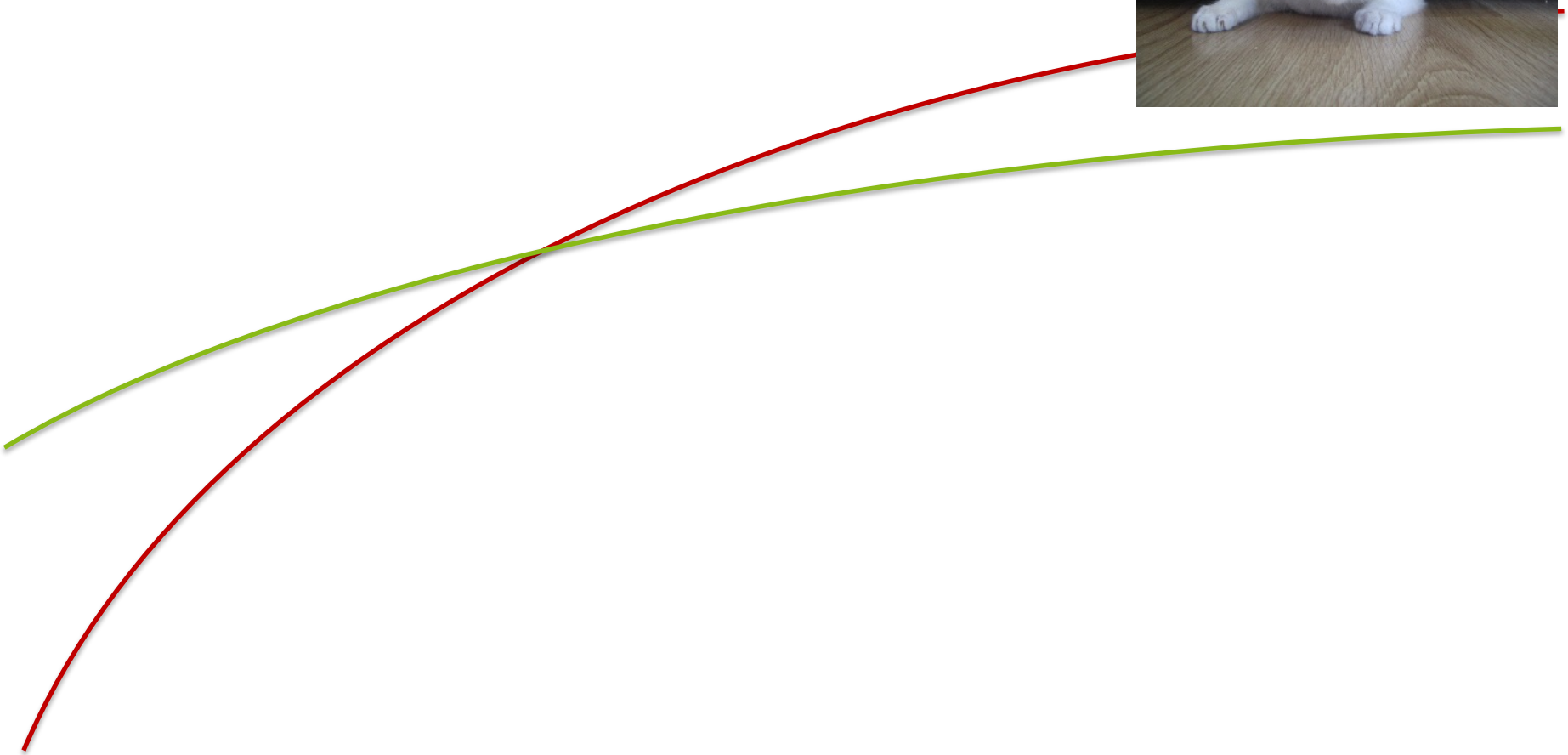


Case courtesy of A.Prof Frank Gaillard, Radiopaedia.org, rID: 8095



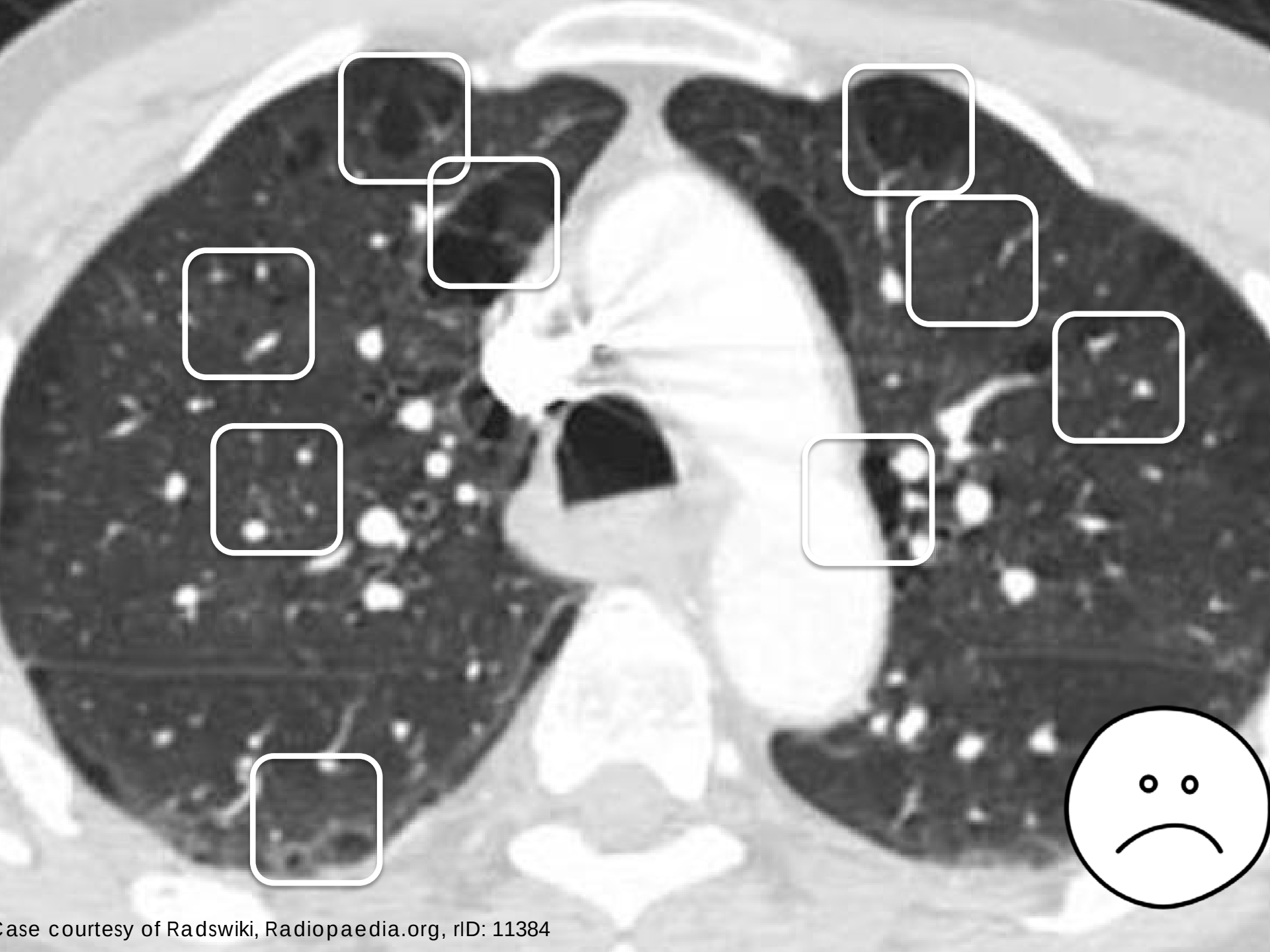
Case courtesy of Radswiki, Radiopaedia.org, rID: 11384

Learning curve




Big data


Representative & annotated data



Case courtesy of Radswiki, Radiopaedia.org, rID: 11384

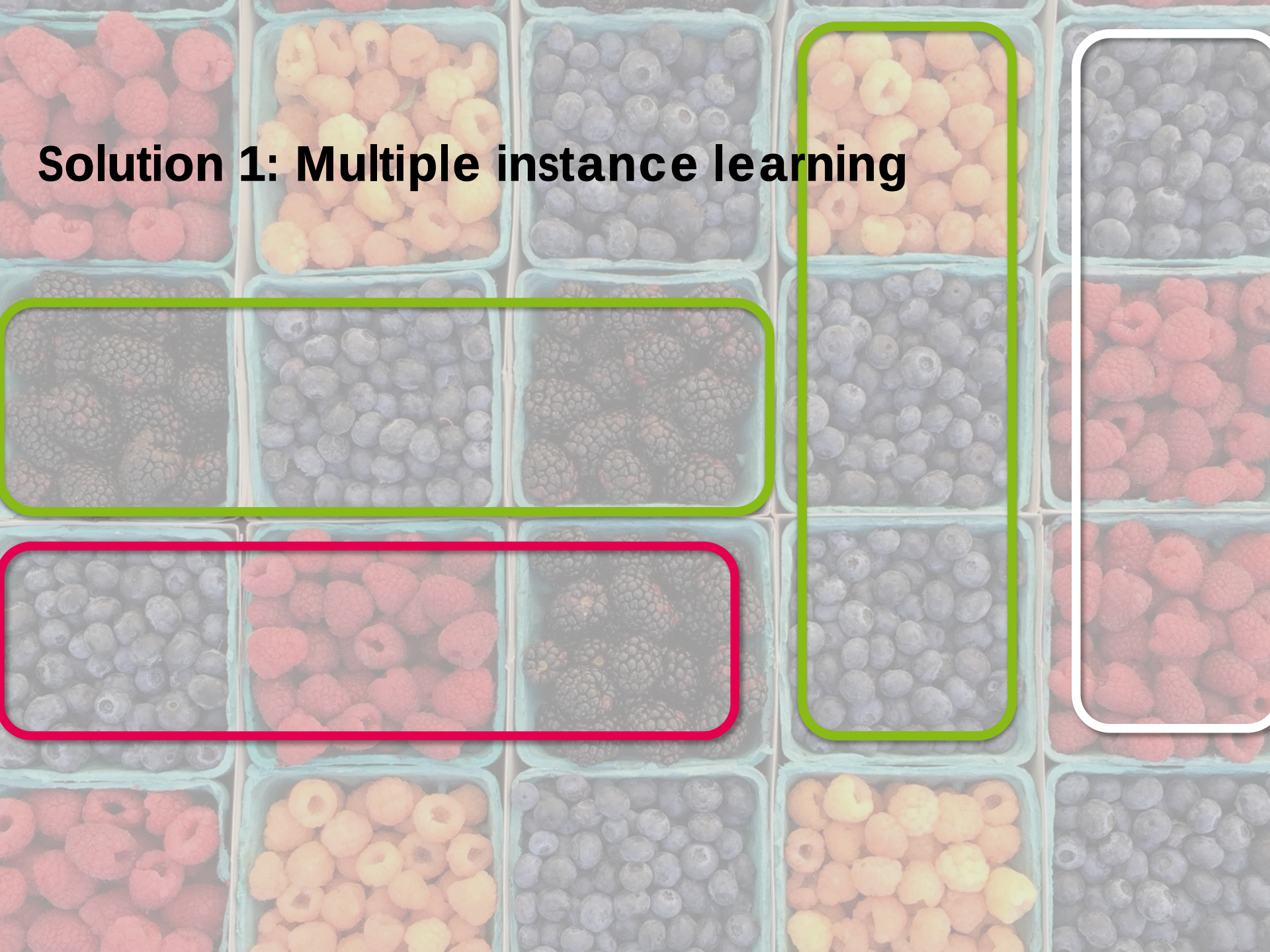
Cat(Dog)

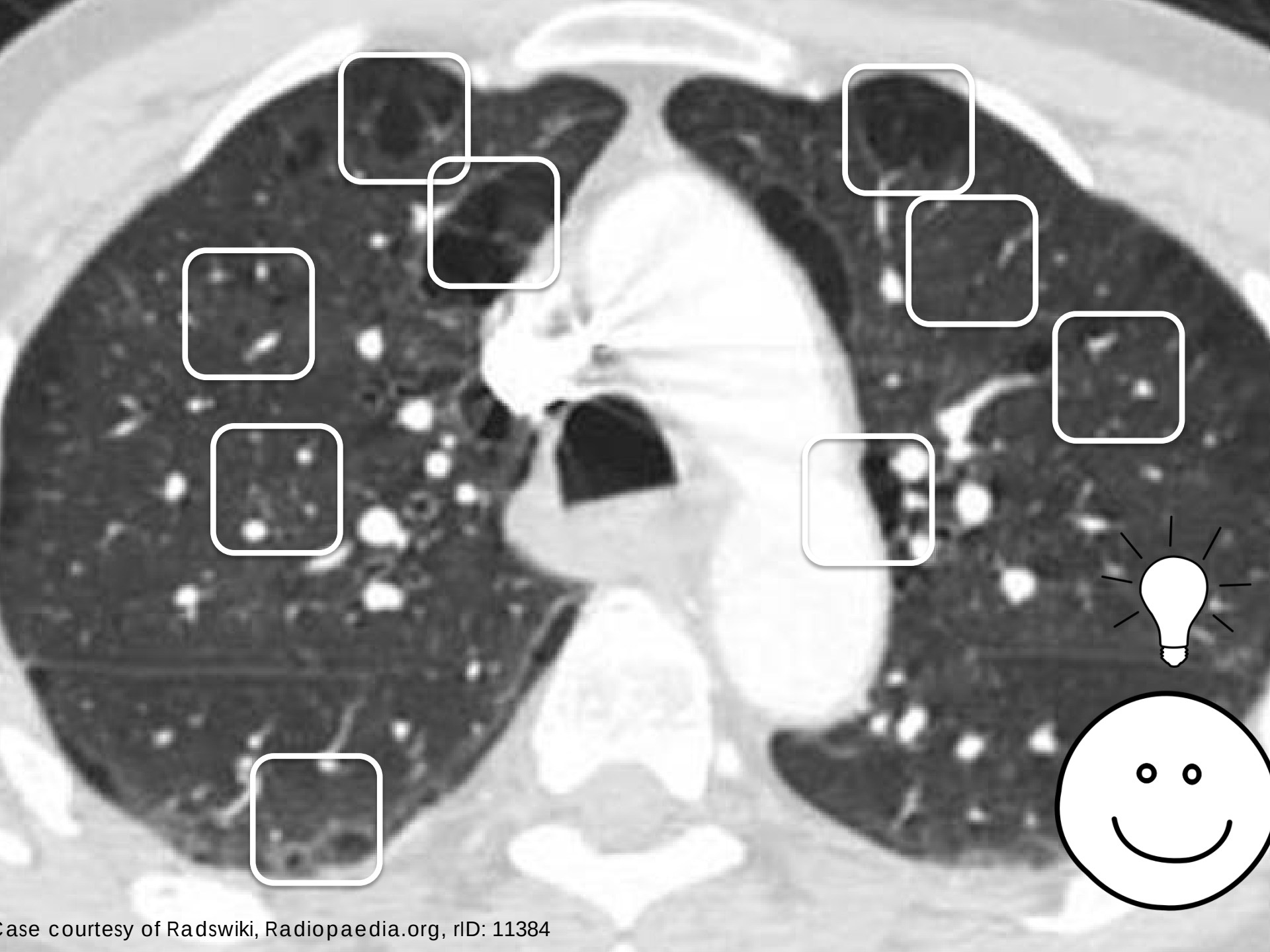
Dog(Cat)

Su, J., Vargas, D. V., & Kouichi, S. (2017). One pixel attack for fooling deep neural networks. *arXiv preprint arXiv:1710.08864*.

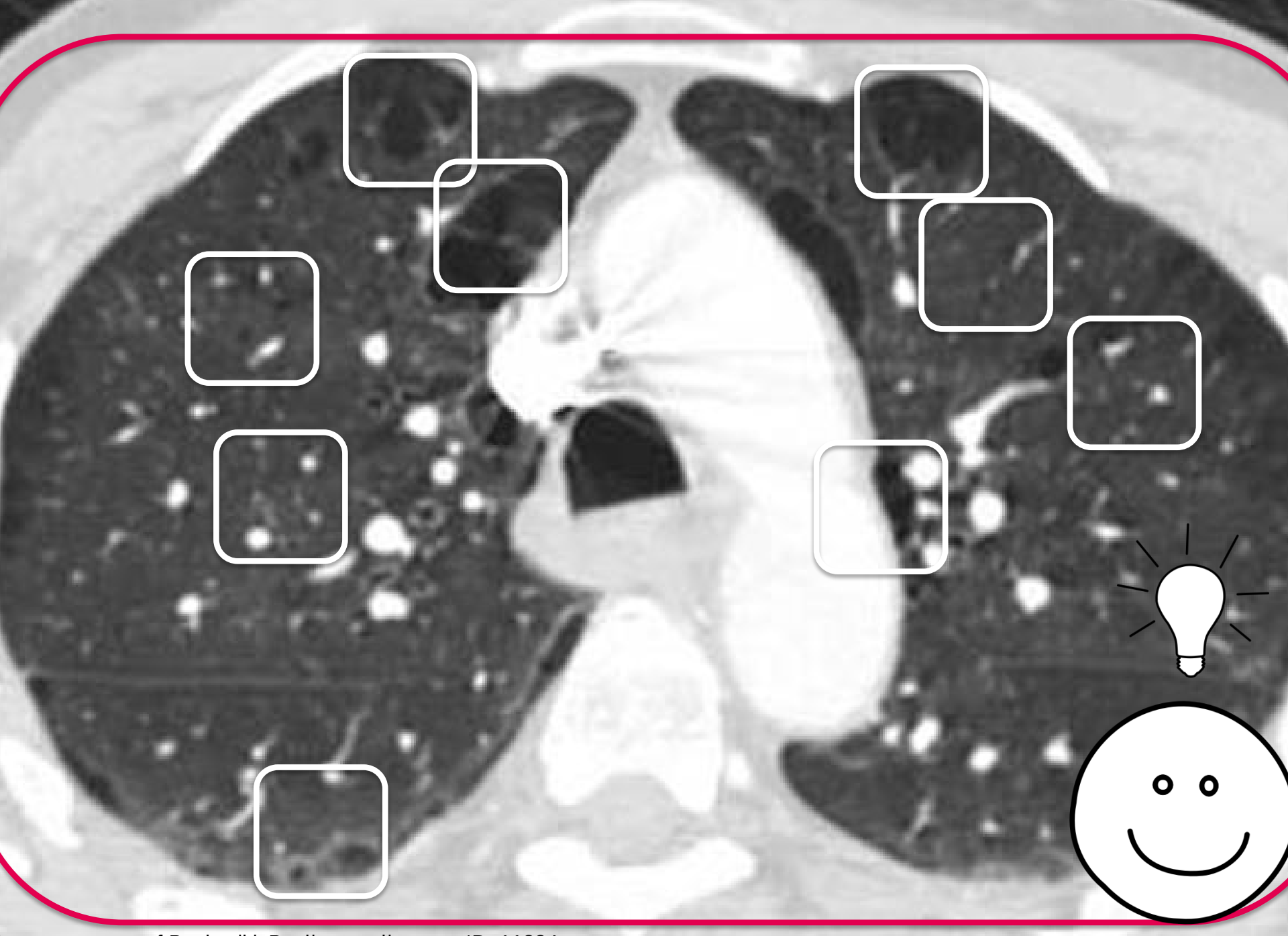
This talk: three “solutions”

Solution 1: Multiple instance learning



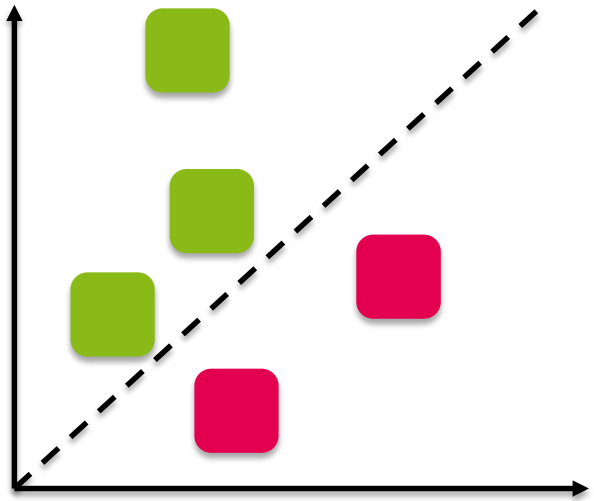


Case courtesy of Radswiki, Radiopaedia.org, rID: 11384

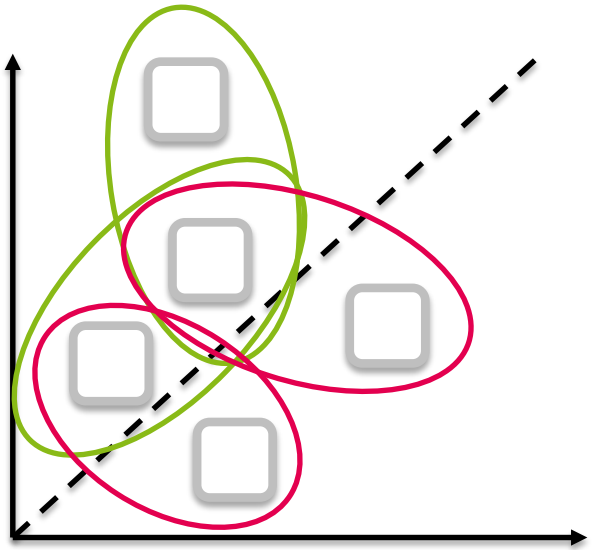


Case courtesy of Radswiki, Radiopaedia.org, rID: 11384

Supervised

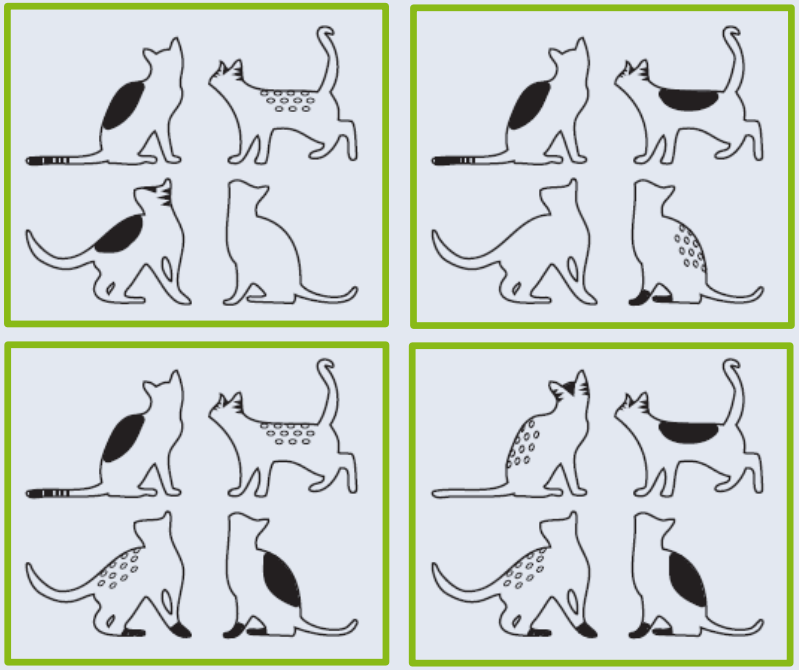


MIL



Dissimilarity-Based Multiple Instance Learning

What is different about the groups of cats on the front cover, from the groups on the back cover? If you can answer this question, you are probably also able to categorize another, previously unseen to you, group of cats. This thesis is about different applications where similar puzzles may occur, and how some machine learning algorithms approach such problems.



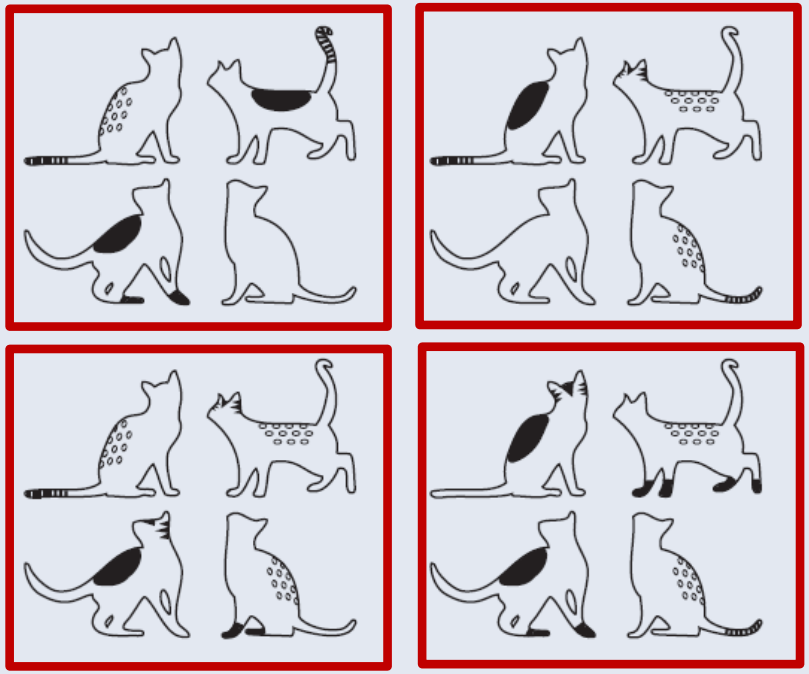
ISBN 978-94-6295-192-1

Dissimilarity-Based Multiple Instance Learning

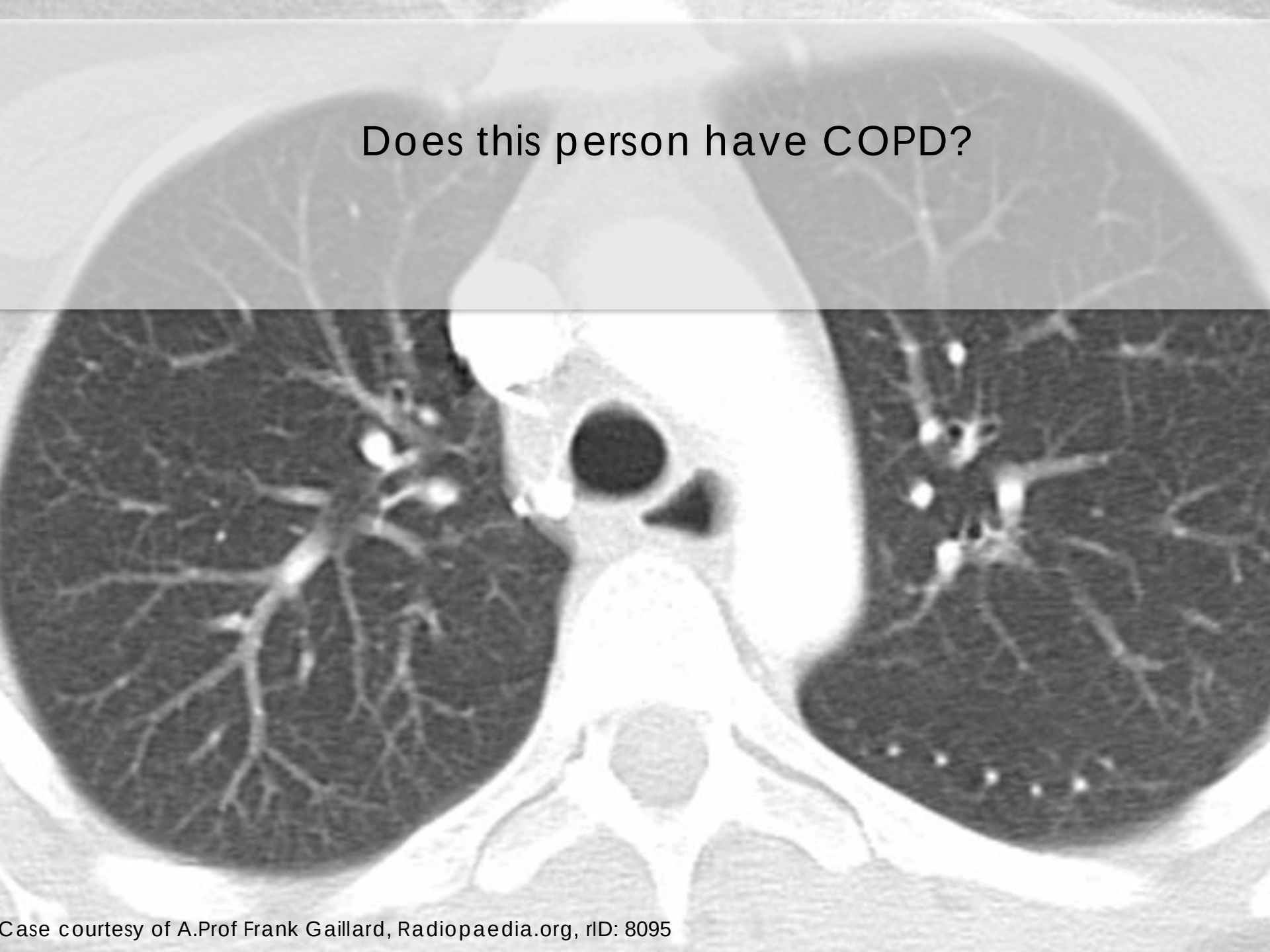
Veronika Cheplygina

Dissimilarity-Based Multiple Instance Learning

Veronika Cheplygina



Does this person have COPD?



Case courtesy of A.Prof Frank Gaillard, Radiopaedia.org, rID: 8095

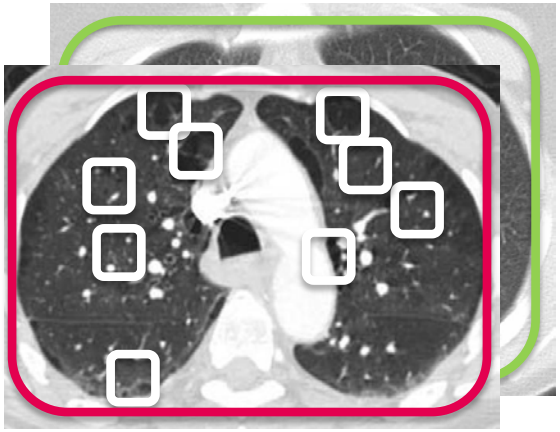
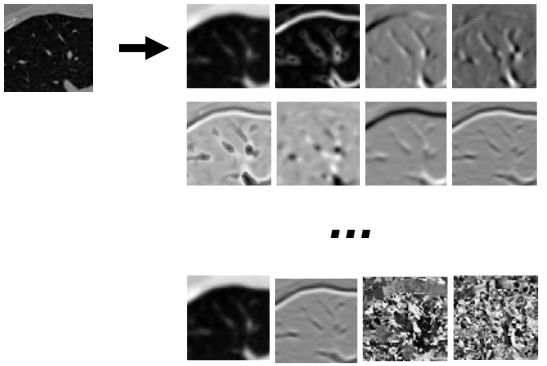
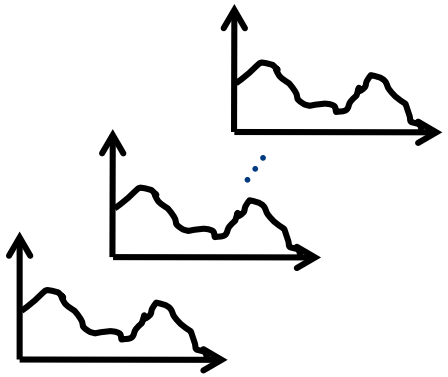


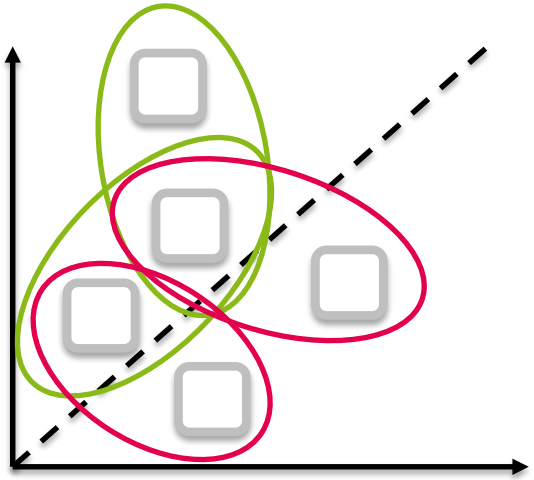
Image = COPD or not (lung function), 50 ROIs



Texture filters



Histograms



MIL

Search for “COPD-like” patch

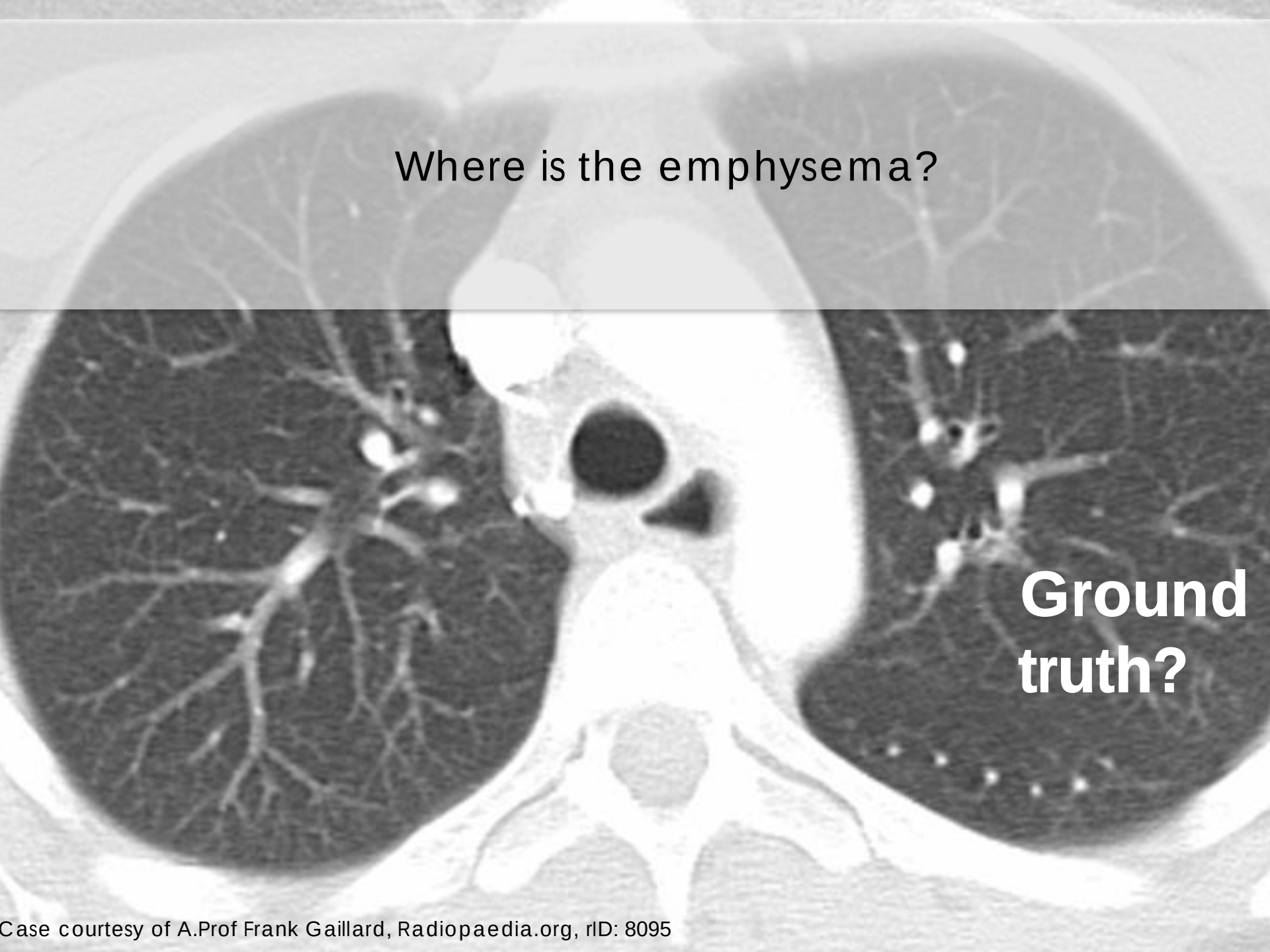
vs

Classify subject directly

Classifier	AUC $\mathcal{X}_{val}$	AUC $\mathcal{X}_{te}$
Simple logistic noisy	50.0	50.0
Simple logistic avg	**71.9**	**70.5**
Simple k-NN noisy	61.0	65.9
Simple k-NN avg	67.0	67.8
miSVM noisy	**69.7**	65.4
miSVM avg	**74.5**	**71.7**
MILBoost	55.8	61.4
Citation k-NN	65.2	61.5
mean-inst SVM	**74.0**	**74.2**
extremes SVM	**70.8**	68.6
BoW SVM	50.0	50.0
MILES	65.8	68.2
meanmin SVM	70.8	**71.3**
meanmin k-NN	65.0	**69.1**
emd SVM	**73.7**	**74.6**
emd k-NN	65.1	67.1

Cheplygina, V., Sorensen, L., Tax, D. M. J., Pedersen, J. H., Loog, M., & de Bruijne, M. (2014). Classification of COPD with multiple instance learning. In International Conference on Pattern Recognition (pp. 1508-1513).

Where is the emphysema?



Case courtesy of A.Prof Frank Gaillard, Radiopaedia.org, rID: 8095

Training

Test

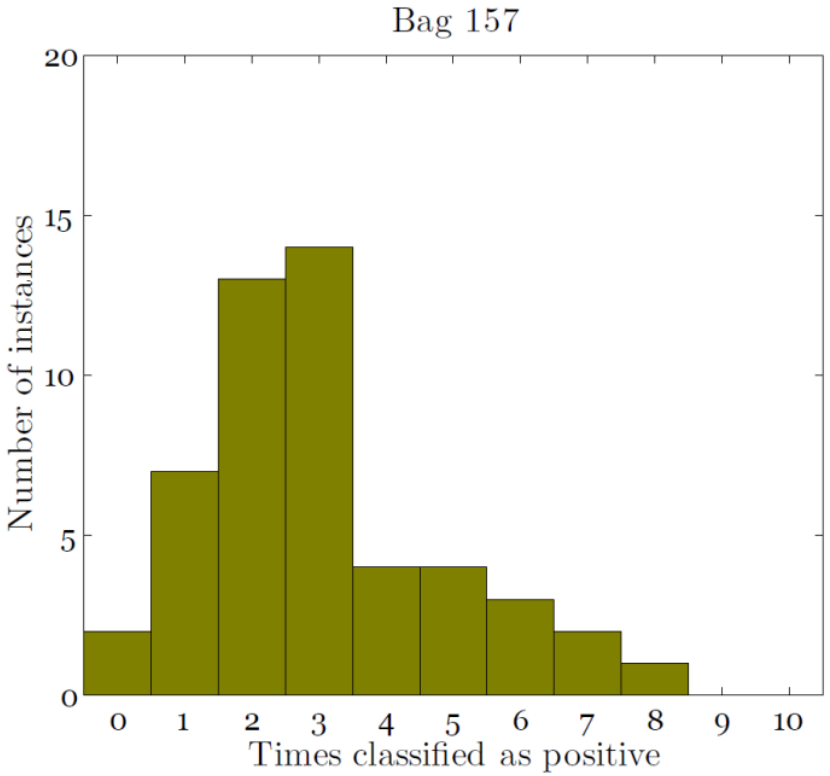
Evaluate stability

Fraction of agreement on positives

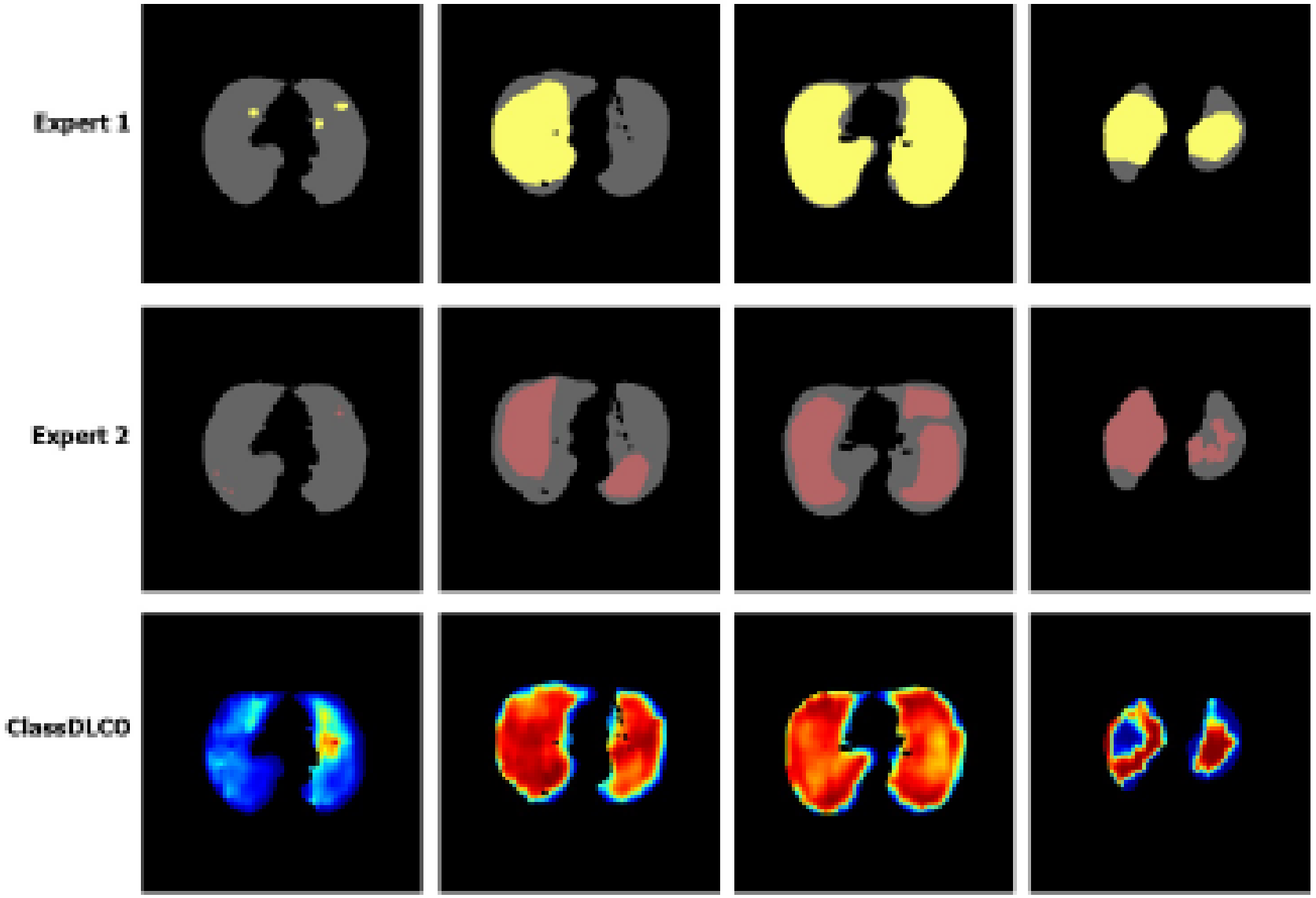
S_+

1.0	0.4	0.5	0.3	0.6	0.5	0.6	0.2	0.4	0.5
0.4	1.0	0.4	0.2	0.2	0.6	0.3	0.1	0.3	0.4
0.5	0.4	1.0	0.2	0.4	0.5	0.4	0.1	0.3	0.6
0.3	0.2	0.2	1.0	0.3	0.2	0.4	0.8	0.6	0.2
0.6	0.2	0.4	0.3	1.0	0.3	0.6	0.3	0.3	0.4
0.5	0.6	0.5	0.2	0.3	1.0	0.4	0.1	0.3	0.4
0.6	0.3	0.4	0.4	0.6	0.4	1.0	0.3	0.4	0.4
0.2	0.1	0.1	0.8	0.3	0.1	0.3	1.0	0.5	0.1
0.4	0.3	0.3	0.6	0.3	0.3	0.4	0.5	1.0	0.3
0.5	0.4	0.6	0.2	0.4	0.4	0.4	0.1	0.3	1.0

Any patches always positive?



Cheplygina, V., Sørensen, L., Tax, D. M. J., de Bruijne, M., & Loog, M. (2015) Label Stability in Multiple Instance Learning. *In Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 539-546*


Expert 1
Expert 2
ClassDLCO

Solution 2: Transfer learning

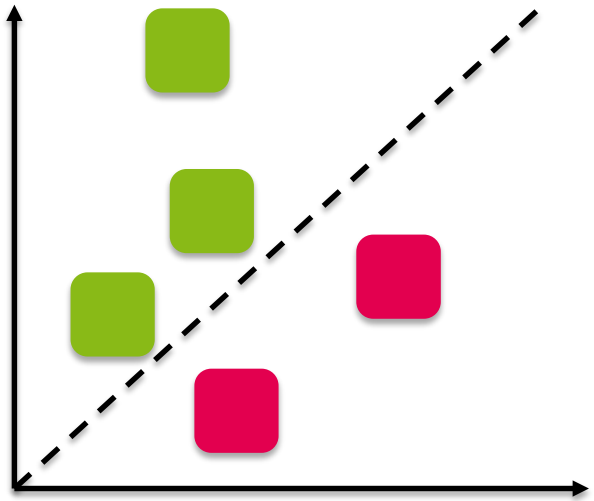
Not learning “from scratch”

- Use other labeled datasets

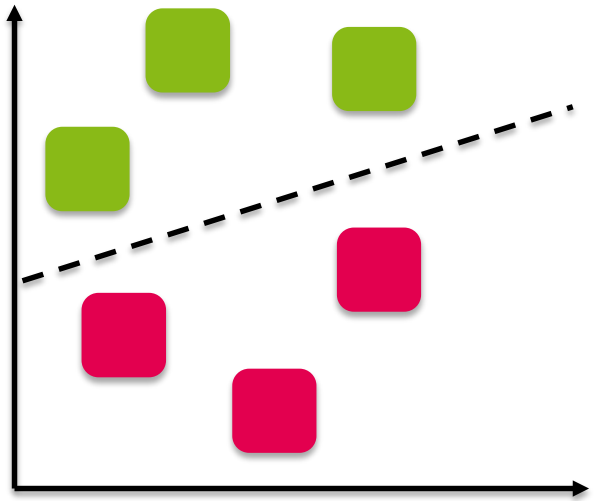
Dataset	Subjects	Age	GOLD (1/2/3/4)	Smoking (c/f/n)	Scanner	Resolution (mm)
DLCST	300 +	59 [50, 71]	69/28/2/0	77/23/0	Philips	0.72×0.72×1 to
	300 -	57 [49, 69]		74/26/0	16 rows Mx 8000	0.78×0.78×1
COPDGene1	74 +	64 [45, 80]	21/18/19/16	17/57/0	Siemens	0.65×0.65×0.75
	46 -	59 [45, 78]		23/20/3	Definition	
COPDGene2	42 +	65 [45, 78]	9/13/7/13	12/30/0	Siemens	0.65×0.65×0.75
	25 -	60 [47, 78]		9/11/5	Definition AS+	
Frederikshavn	8 +	66 [48, 77]	1/3/3/1	1/7/0	Siemens	0.58×0.58×0.6
	8 -	56 [25, 73]		1/2/5	Definition Flash	

...as training data

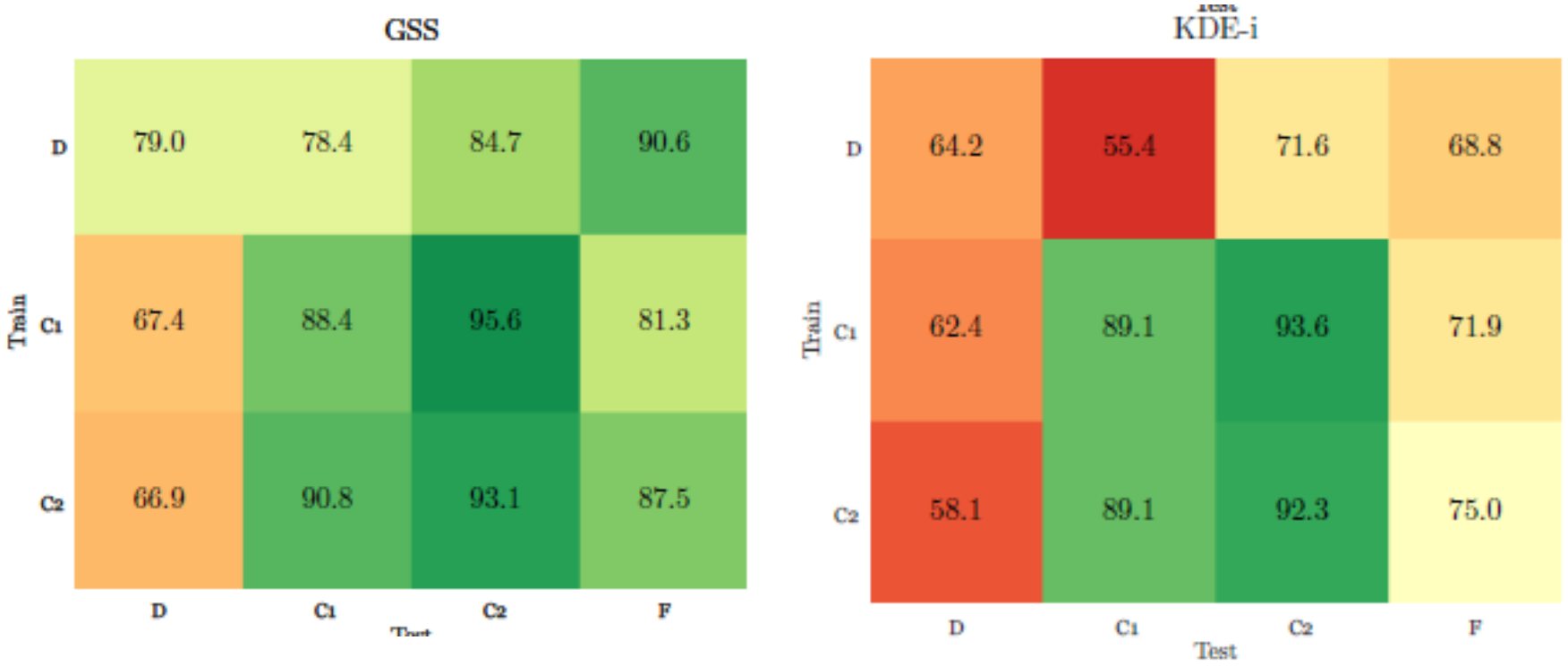
Source



Target

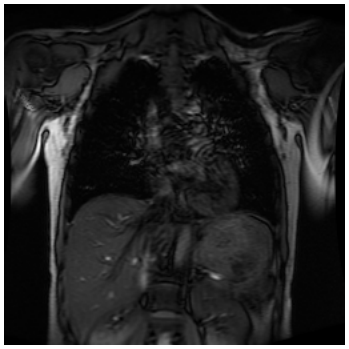


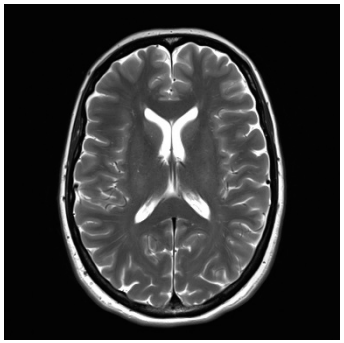
... to find good features

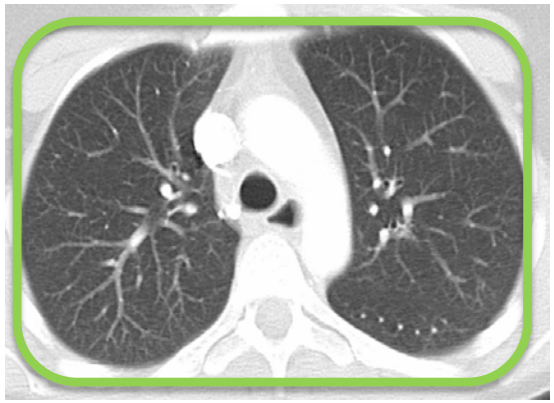


Cheplygina, V., Peña, I. P., Pedersen, J. H., Lynch, D. A., Sørensen, L., & de Bruijne, M. (2017). Transfer learning for multi-center classification of chronic obstructive pulmonary disease. In *Journal of Biomedical and Health Informatics*, to appear

... even if task and/or modality different

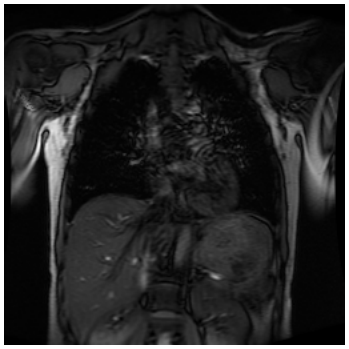


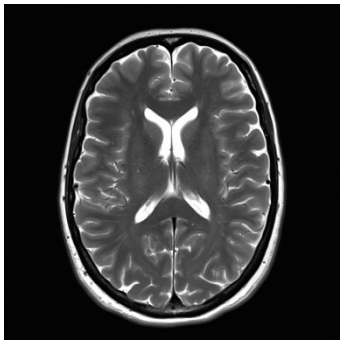


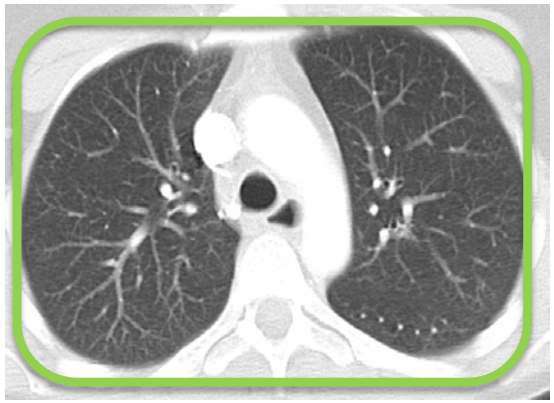


Case courtesy of A.Prof Frank Gaillard, Radiopaedia.org, rID: 51158

Which datasets to use? (Most similar? Most different?)

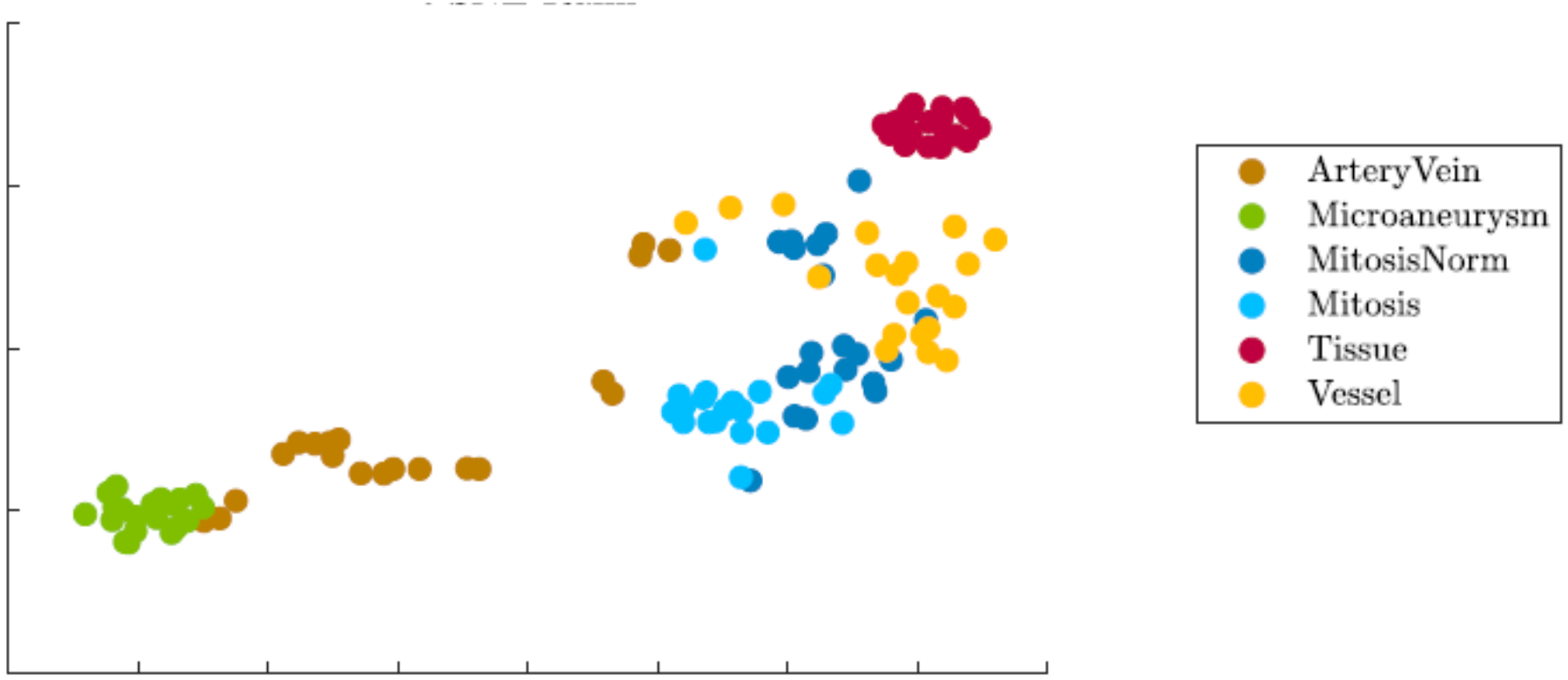






Case courtesy of A.Prof Frank Gaillard, Radiopaedia.org, rID: 51158

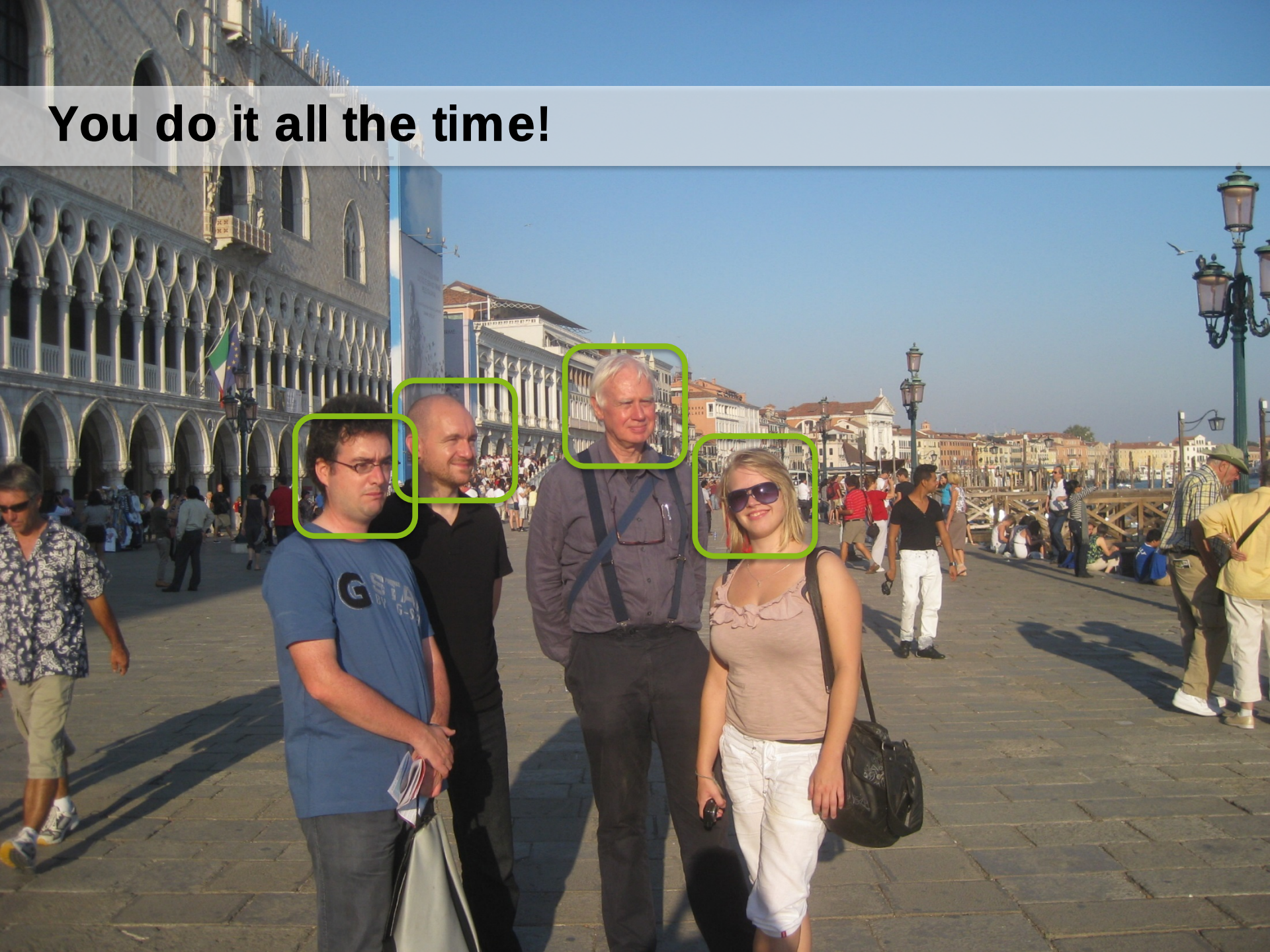
Meta-learning: how to quantify similarity of data?



Cheplygina, V., Moeskops, P., Veta, M., Bozorg, B. D., & Pluim, J. (2017). Exploring the similarity of medical imaging classification problems. In *Large-Scale Annotation of Biomedical Data and Expert Label Synthesis (MICCAI LABELS)* (pp. 59-66)

Solution 3: Crowdsourcing

You do it all the time!



tvst

DOI: 10.1167/tvst.5.5.6

Article

The Accuracy and Reliability of Crowdsource Annotations of Digital Retinal Images

Danny Mitry[1], Kris Zutis[2], Baljean Dhillon[3], Tunde Peto[1], Shabina Hayat[4], Kay-Tee Khaw[5], James E. Morgan[6], Wendy Moncur[7], Emanuele Trucco[2], and Paul J. Foster[1] for the UK Biobank Eye and Vision Consortium

IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 35, NO. 5, MAY 2016

AggNet: Deep Learning From Crowds for Mitosis Detection in Breast Cancer Histology Images

Shadi Albarqouni*, *Student Member, IEEE*, Christoph Baur, Felix Achilles, *Student Member, IEEE*, Vasileios Belagiannis, *Student Member, IEEE*, Stefanie Demirci, and Nassir Navab, *Member, IEEE*

Can Masses of Non-Experts Train Highly Accurate Image Classifiers?

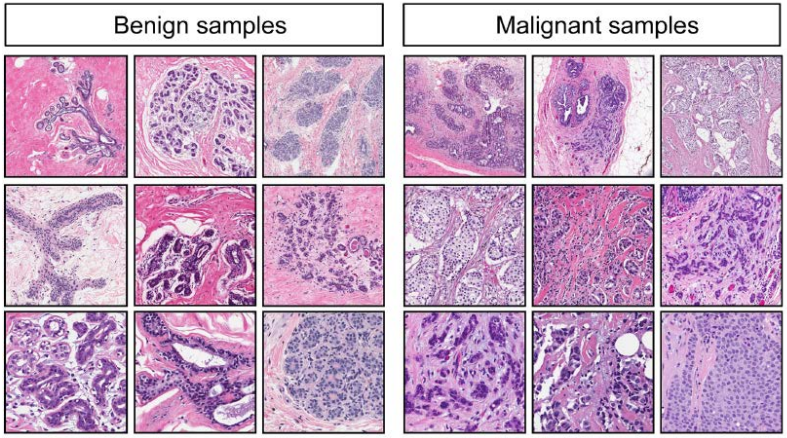
A Crowdsourcing Approach to Instrument Segmentation in Laparoscopic Images

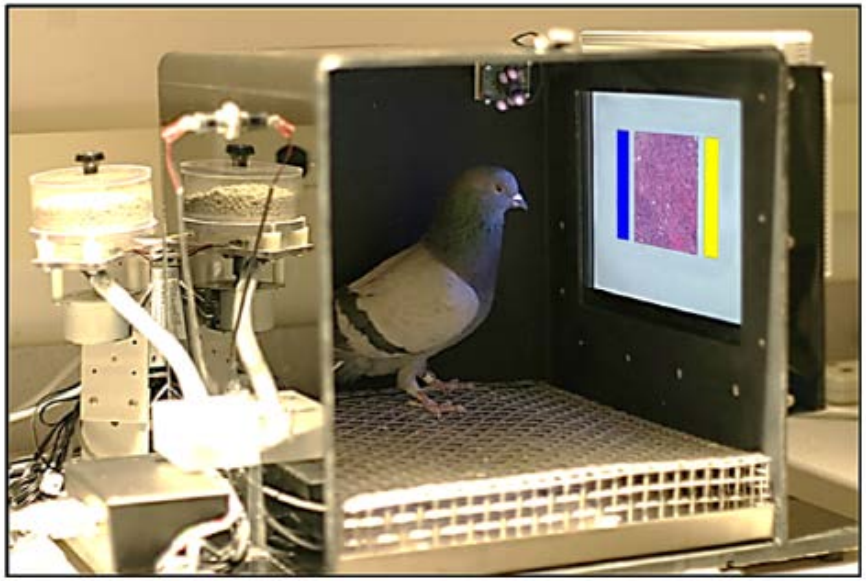
Lena Maier-Hein[1,⋆,⋆⋆], Sven Mersmann[1], Daniel Kondermann[2], Sebastian Bodenstedt[3], Alexandro Sanchez[2], Christian Stock[4], Hannes Gotz Kenngott[5], Mathias Eisenmann[3], and Stefanie Speidel[3]

RESEARCH ARTICLE

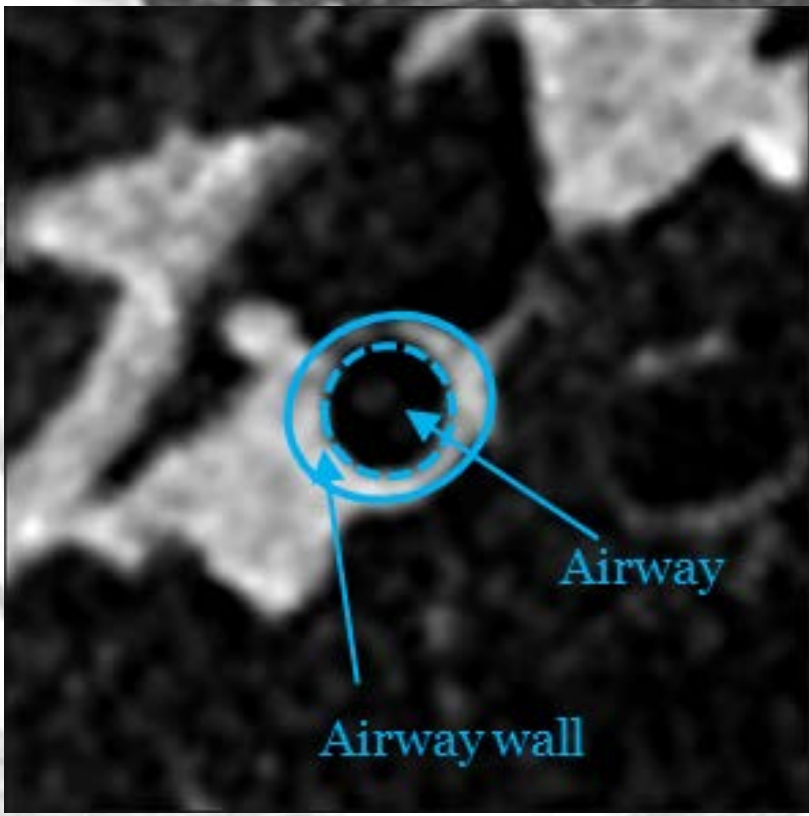
Pigeons (*Columba livia*) as Trainable Observers of Pathology and Radiology Breast Cancer Images

Richard M. Levenson[1]*, Elizabeth A. Krupinski[3], Victor M. Navarro[2], Edward A. Wasserman[2]*





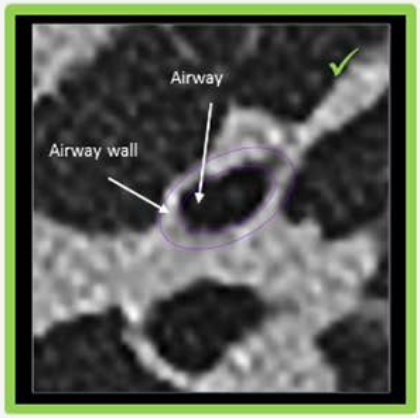
How large are the airways?

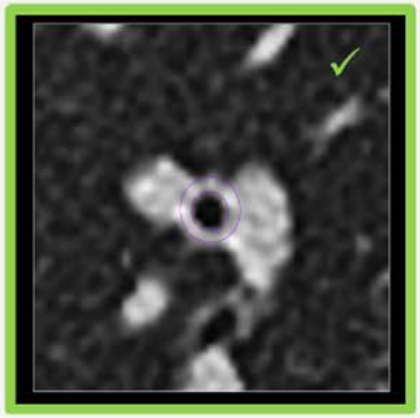


Case courtesy of A.Prof Frank Gaillard, Radiopaedia.org, rID: 8095

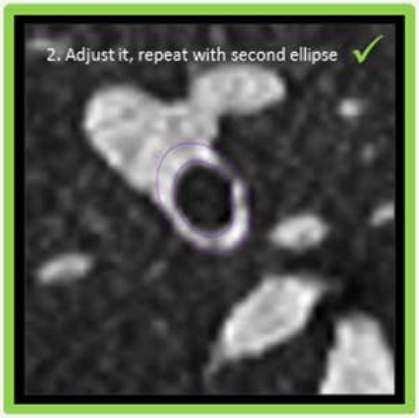
Welcome Veronika! Save lives by annotating airways!

Save lives by annotating airways!



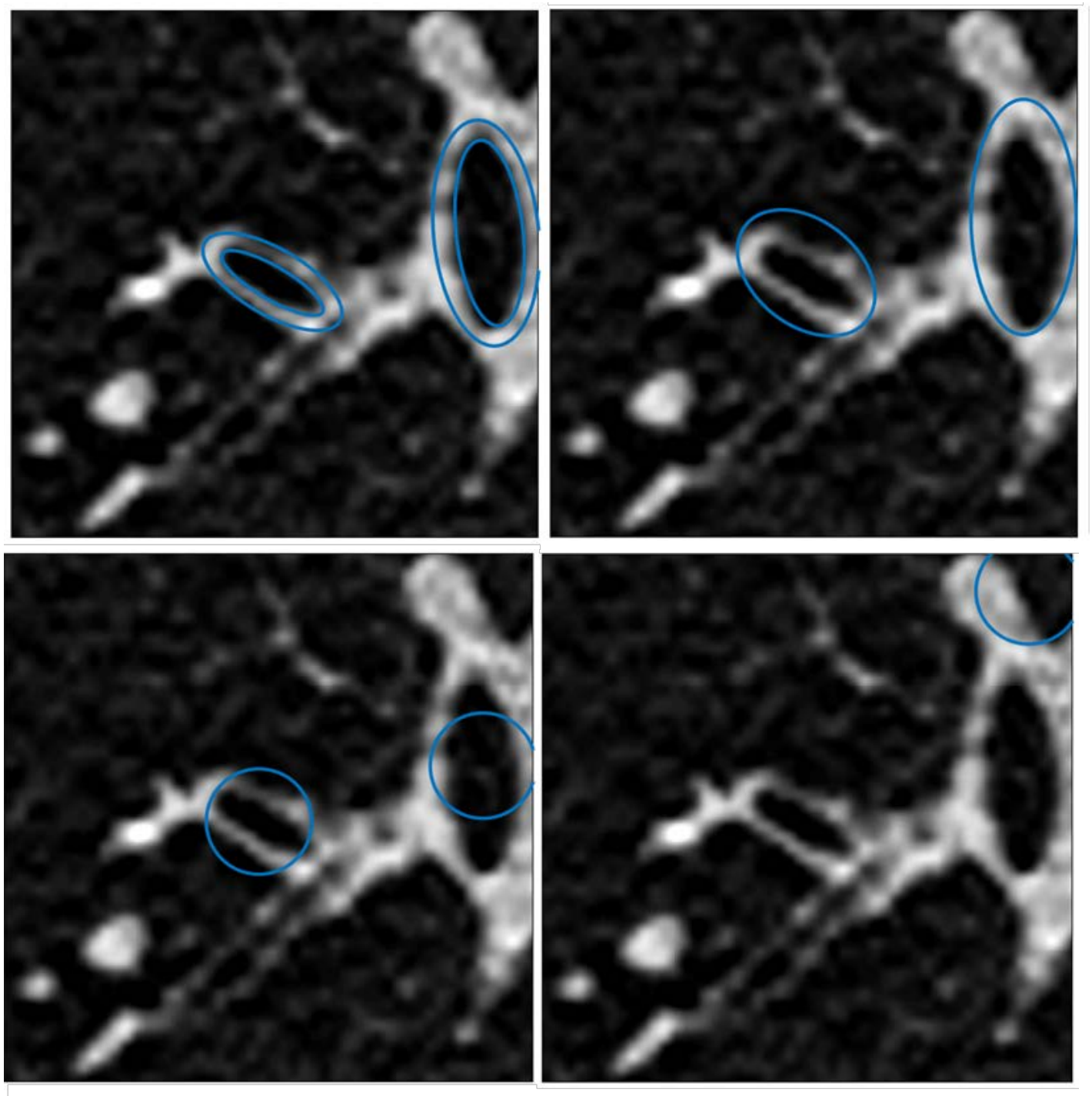


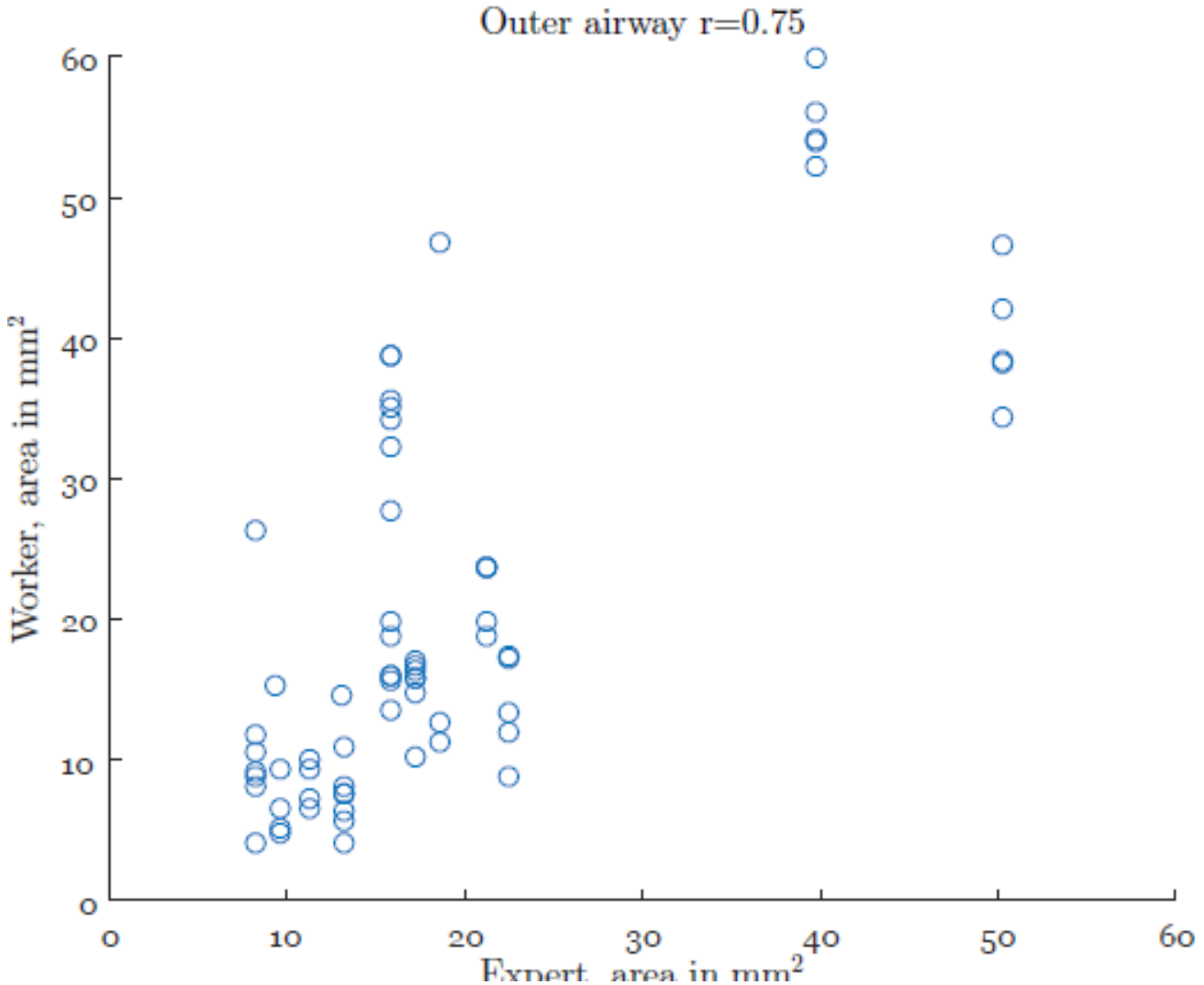
Help us find airways! We are researching how to detect lung diseases such as cystic fibrosis and COPD, and need help with measuring the airways inside the lungs. You will be looking at 2D slices from a 3D image of the lungs. If the slice crosses an airway, you should see a dark circle or



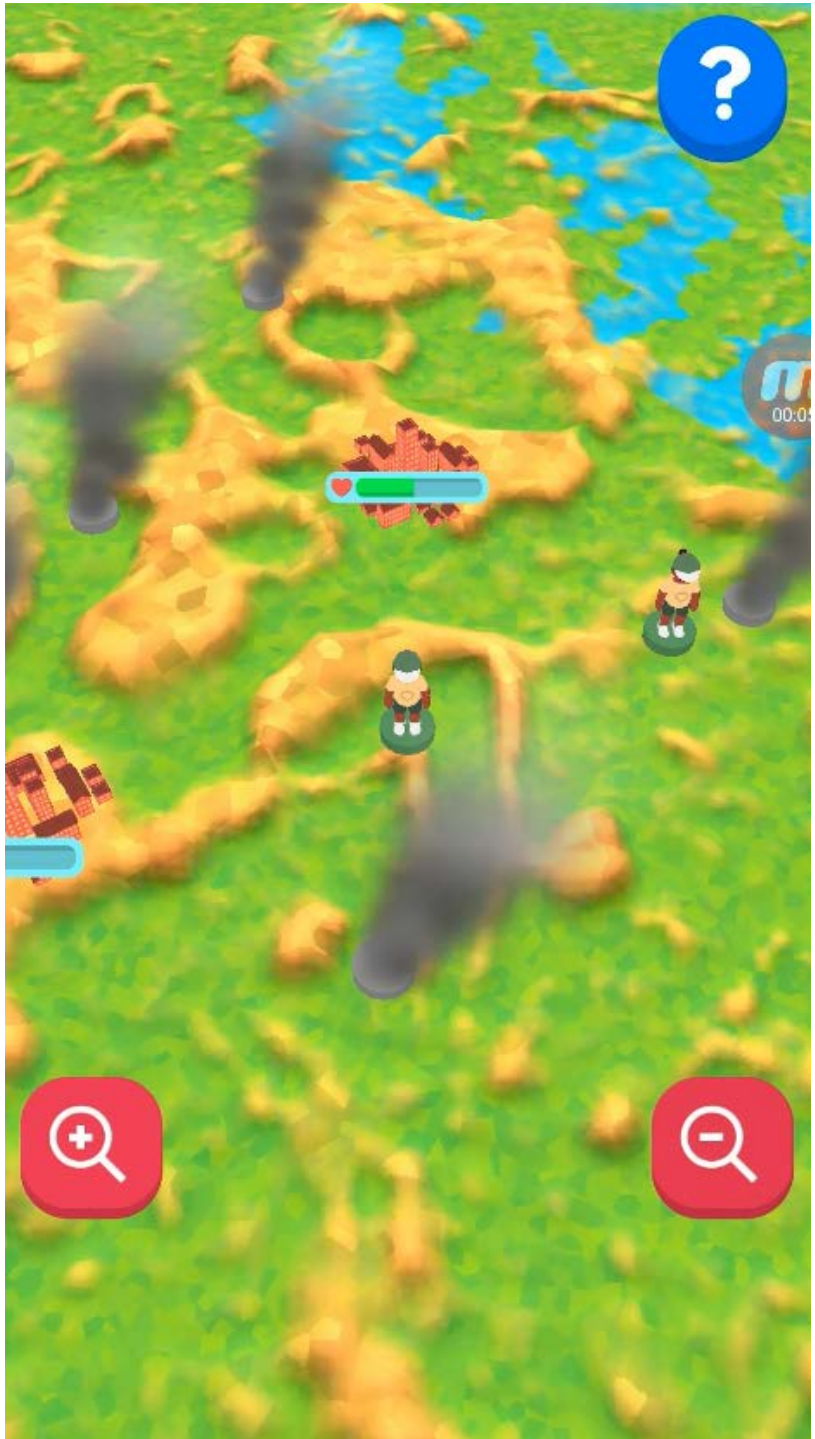
We want you to annotate BOTH the airway and the wall around it. You can do this by placing TWO ellipses at the center of the airway and adjusting them. One ellipse should be inside the other, and they should not cross.

Start





Cheplygina, V., Perez-Rovira, A., Kuo, W., Tiddens, H. A., & de Bruijne, M. (2016). Early Experiences with Crowdsourcing Airway Annotations in Chest CT. In Large-Scale Annotation of Biomedical Data and Expert Label Synthesis (MICCAI LABELS), pp. 209-218


?
OK

Learning with less labels

- Multiple instance learning
- Transfer learning
- Crowdsourcing

Thanks to:

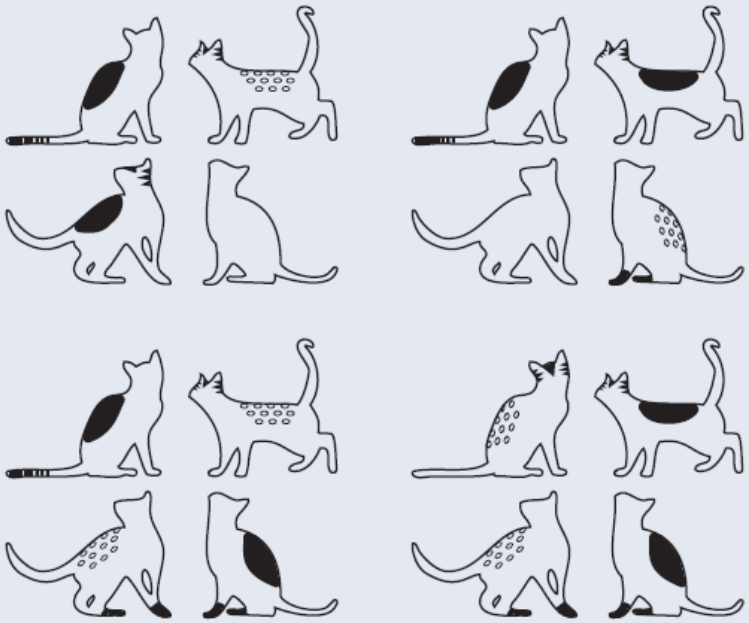
IMAG/e, Eindhoven University of Technology
BIGR, Erasmus MC Rotterdam
PRLab, Delft University of Technology

@vcheplygina

http://www.veronikach.com

Dissimilarity-Based Multiple Instance Learning

What is different about the groups of cats on the front cover, from the groups on the back cover? If you can answer this question, you are probably also able to categorize another, previously unseen to you, group of cats. This thesis is about different applications where similar puzzles may occur, and how some machine learning algorithms approach such problems.



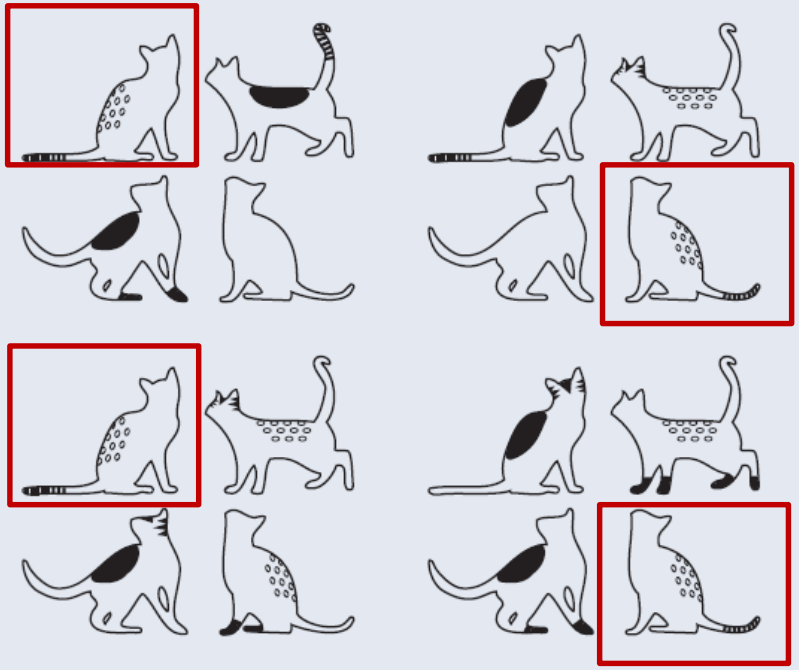
ISBN 978-94-6295-192-1

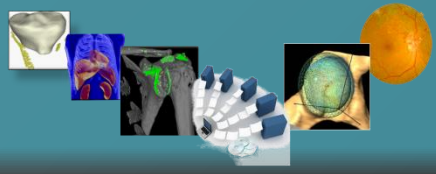
Dissimilarity-Based Multiple Instance Learning

Veronika Cheplygina

Dissimilarity-Based Multiple Instance Learning

Veronika Cheplygina





Radiomics for cancer outcome modeling: image analysis and machine learning challenges

Mathieu Hatt, PhD, HDR – CR INSERM

mathieu.hatt@inserm.fr

Laboratoire du Traitement de l'Information Médicale

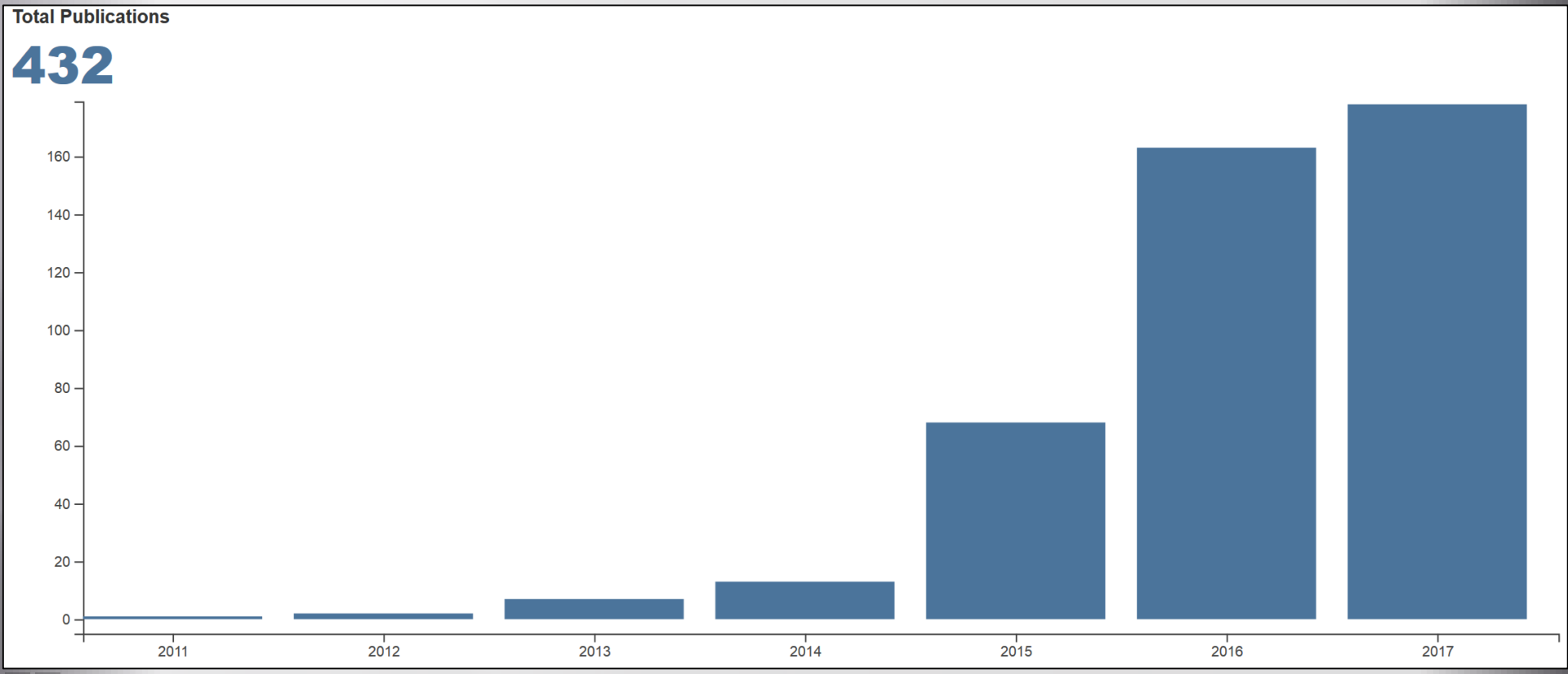
LaTIM, UMR INSERM-UBO 1101, Brest

Lausanne, November 13rd 2017

Introduction

Radiomics: exponential growth

Radiomics: ~430 publications (nov. 2017)



Source: web of science

Introduction

History of radiomics

The terms “radiomics” and “radiogenomics” were already 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.

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.

“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: evolution or revolution?

- Radiomics has become popular since 2012
- Merely a new incarnation of computer-aided diagnosis (CAD) systems (exist since the 80's)
- 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 could be categorized as « radiomics studies »

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

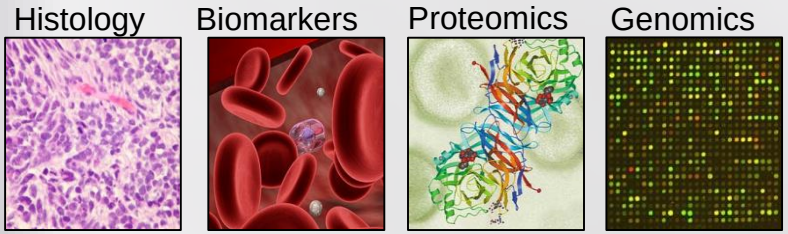
Introduction

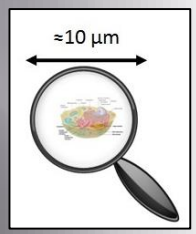
Radiomics: evolution or revolution?

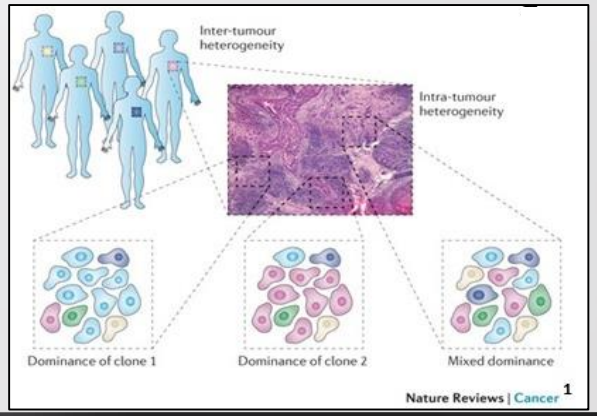
- What has changed?
  - New artificial intelligence techniques
    - e.g. neural networks, deep learning
  - Efficiency+cost of computing power
    - e.g. GPUs (graphical processing units)
  - Improvements of imaging devices (hardware+software)
    - e.g. PSF modeling and ToF in PET
  - Availability of data
    - e.g. images + clinical + histopathology + genetics/transcriptomics
  - Evolving applications for multimodal medical imaging
    - e.g. therapy follow-up, treatment planning

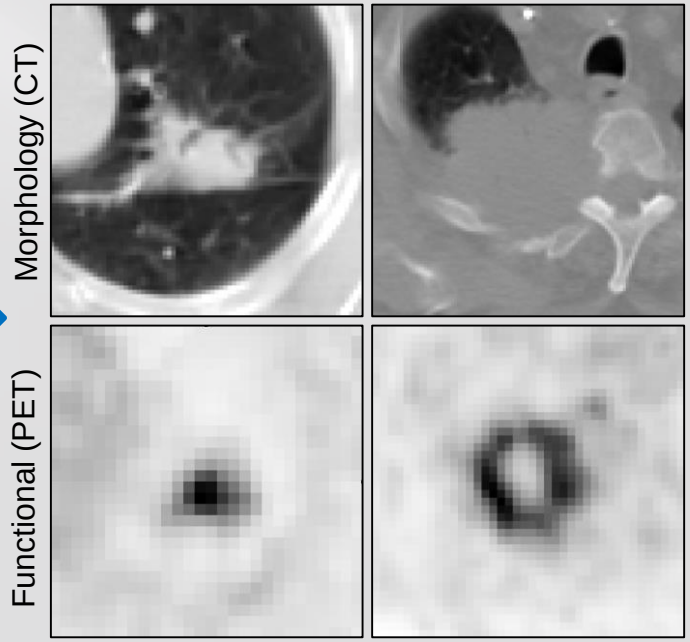
Introduction
Rationale

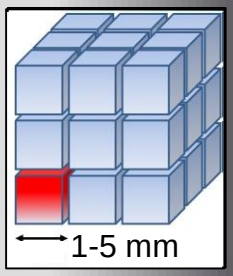
- Macroscopic/microscopic heterogenity
  - Tumours are heterogeneous entities [1]
    - Genetic, cellular, tissular
    - Hypothesis: caracteristics in images (macro scale) reflect at least partly caracteristics in smaller scales (including genetic) [2]







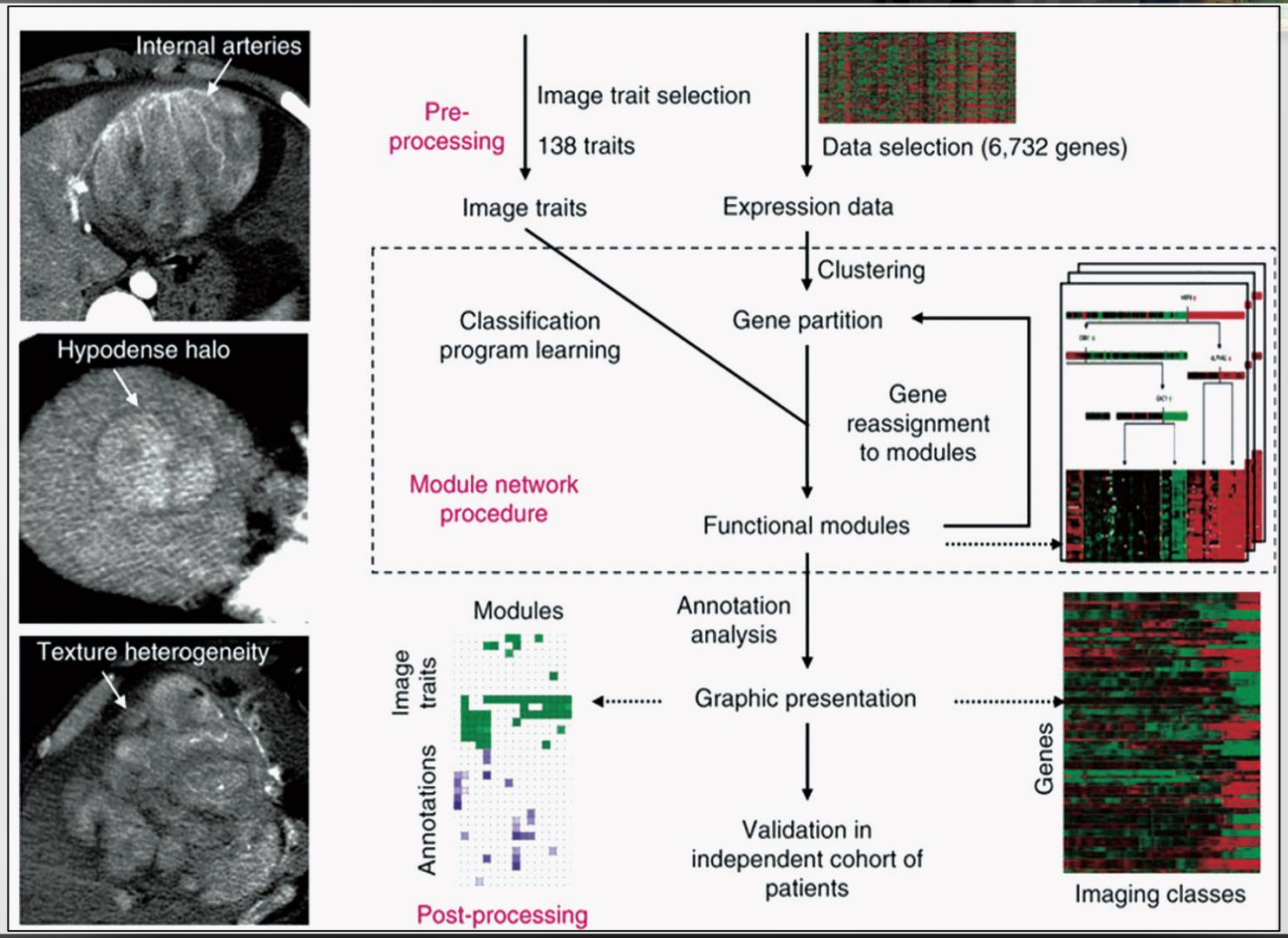




1. Gerlinger, *et al*. **Intratumor heterogeneity and branched evolution revealed by multiregion sequencing**. *N Engl J Med.* 2012
2. Segal, *et al*. **Decoding global gene expression programs in liver cancer by noninvasive imaging**. *Nat Biotechnol*. 2007

Introduction

Early works (example)

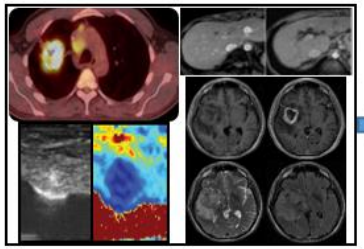


Segal, *et al*. **Decoding global gene expression programs in liver cancer by noninvasive imaging**. *Nat Biotechnol*. 2007

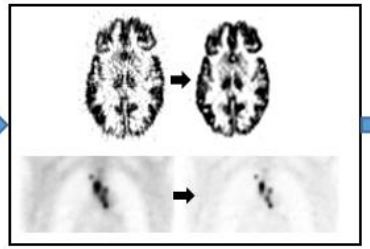
Radiomics

Standard worfklow

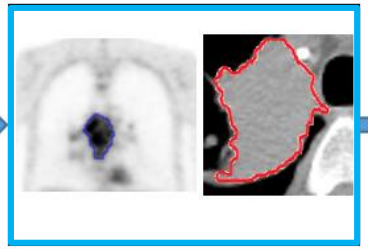
Clinical data



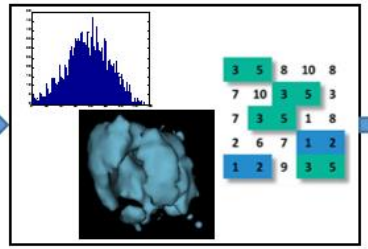
Images



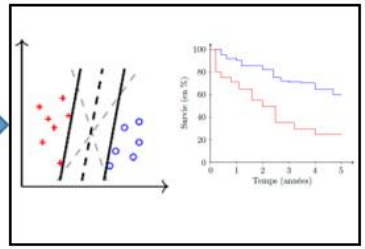
Pre-processing



Segmentation

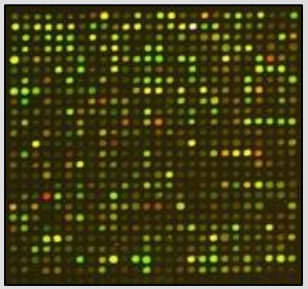


Features extraction



Features selection
Training & validation

Radiomics

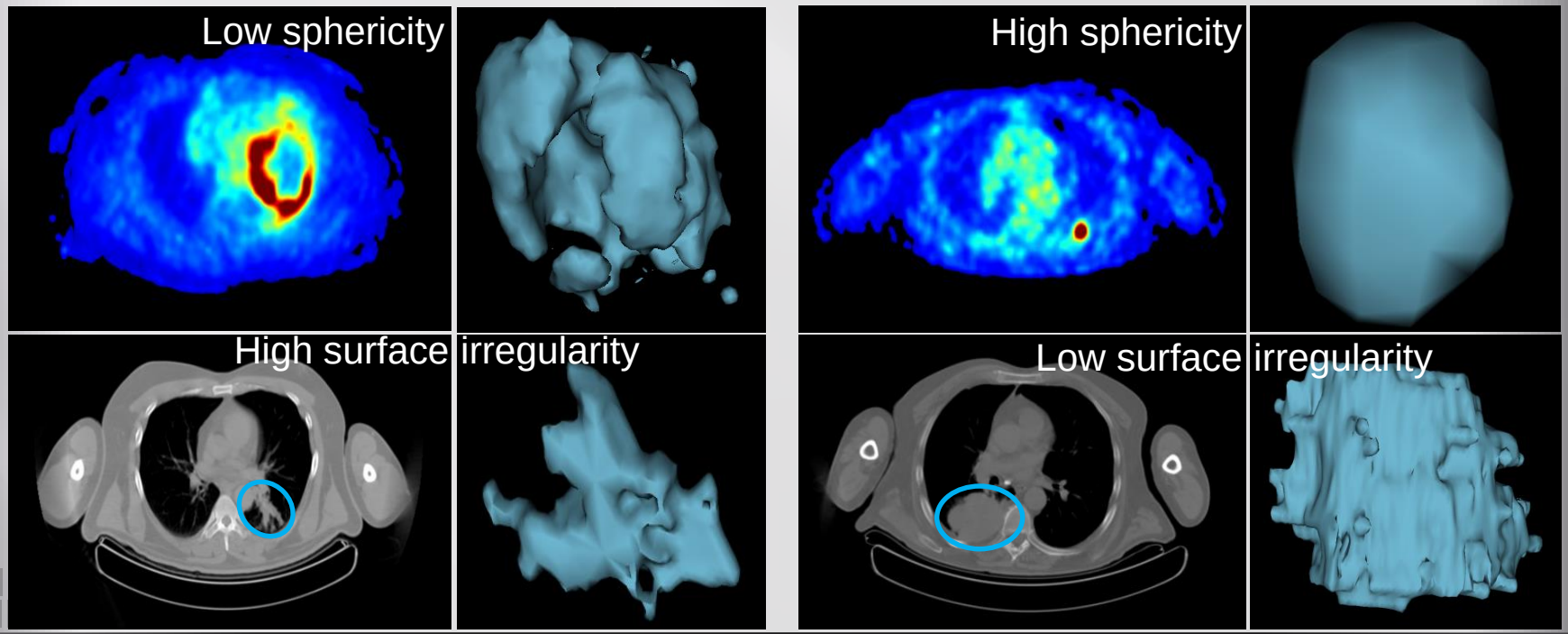


Genomics (and other –omics)

Radiomics

Segmentation: impact?

- Segmentation step: how critical for radiomics?
  - Accuracy very important
    - Especially for shape descriptors
    - Some textural features are also highly sensitive



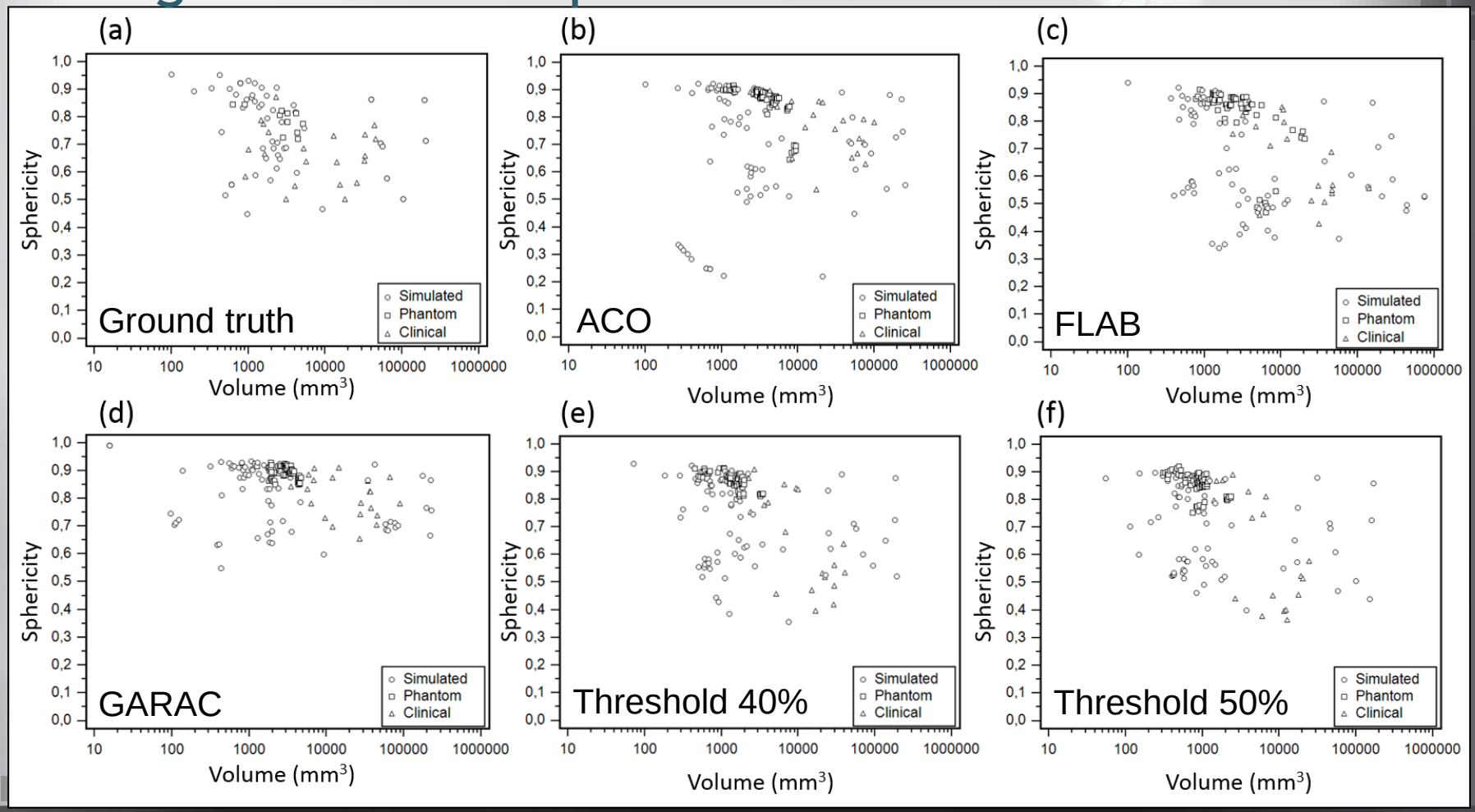
Hatt, *et al.* **Characterization of PET/CT images using texture analysis: the past, the present… any future?** *Eur J Nucl Med Mol Imaging* 2017

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

Segmentation: impact?

Segmentation step: how critical for radiomics?



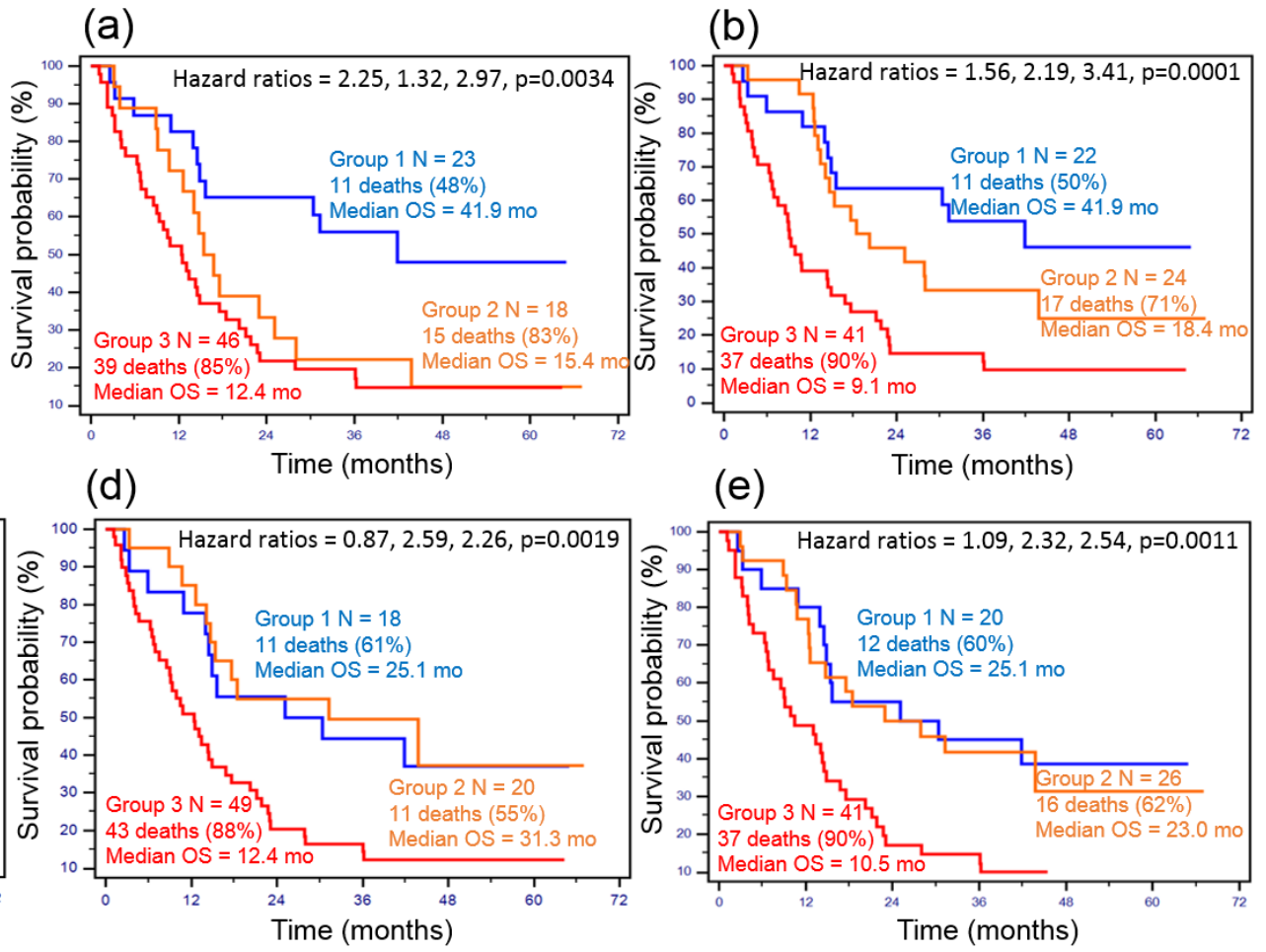
Hatt, et *al*. **Tumour functional sphericity from PET images: prognostic value in NSCLC and impact of delineation method**. *EJNMMI* 2017 (in press)

Radiomics

Segmentation: impact?

Segmentation step: how critical for radiomics?

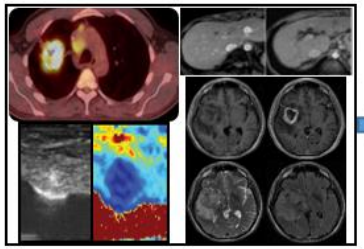
87 NSCLC patients
(stage II-III)



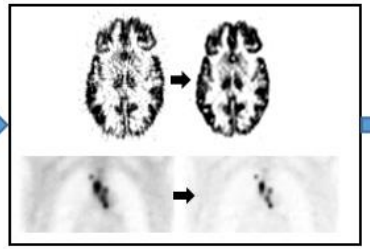
Hatt, *al*. **Tumour functional sphericity from PET images: prognostic value in NSCLC and impact of delineation method**. *EJNMMI* 2017 (in press)

Radiomics
Worfklow

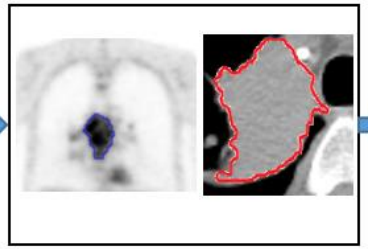
Clinical data



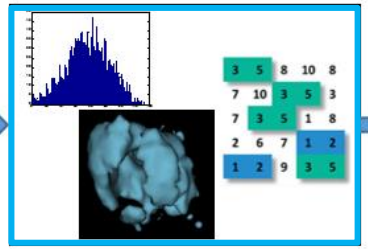
Images



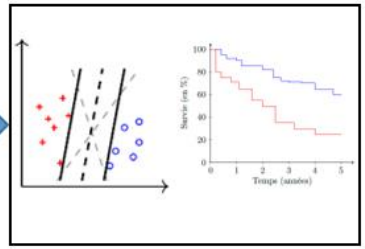
Pre-processing



Segmentation

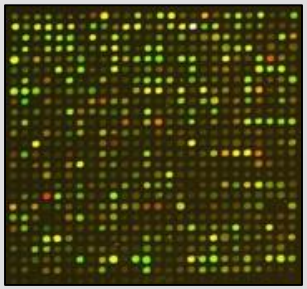


Features extraction



Features selection
Training & validation

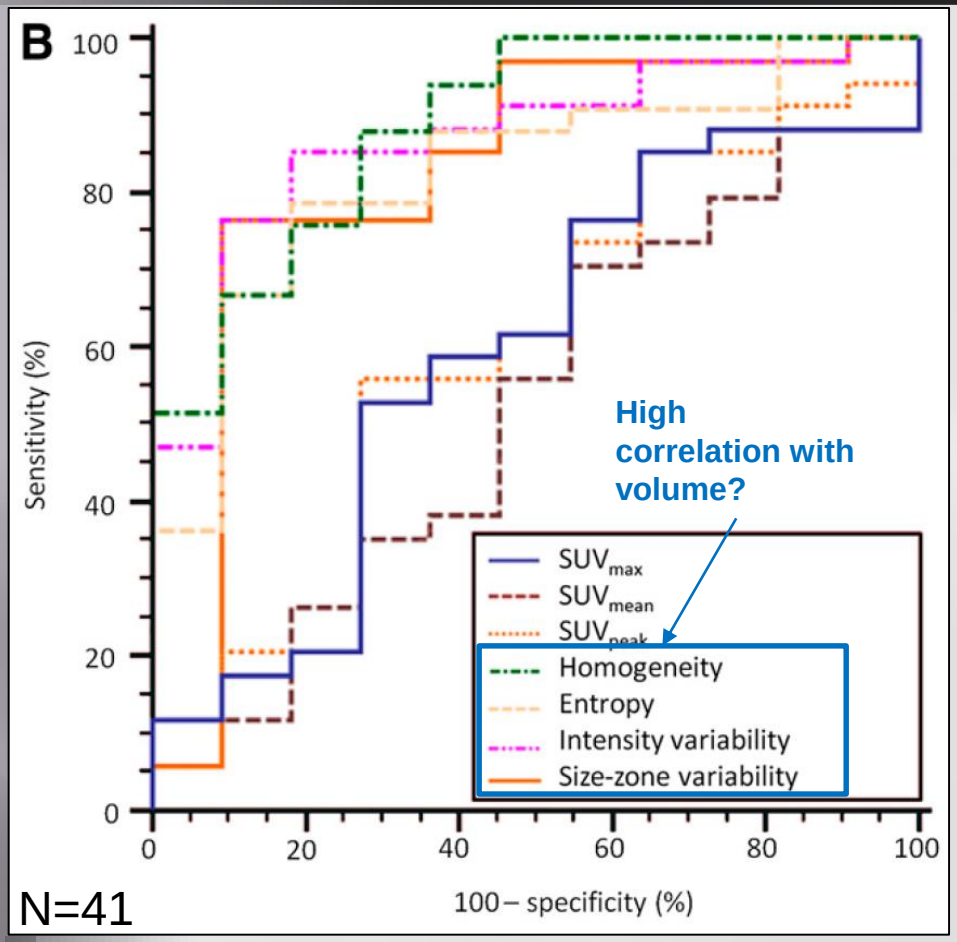
Radiomics

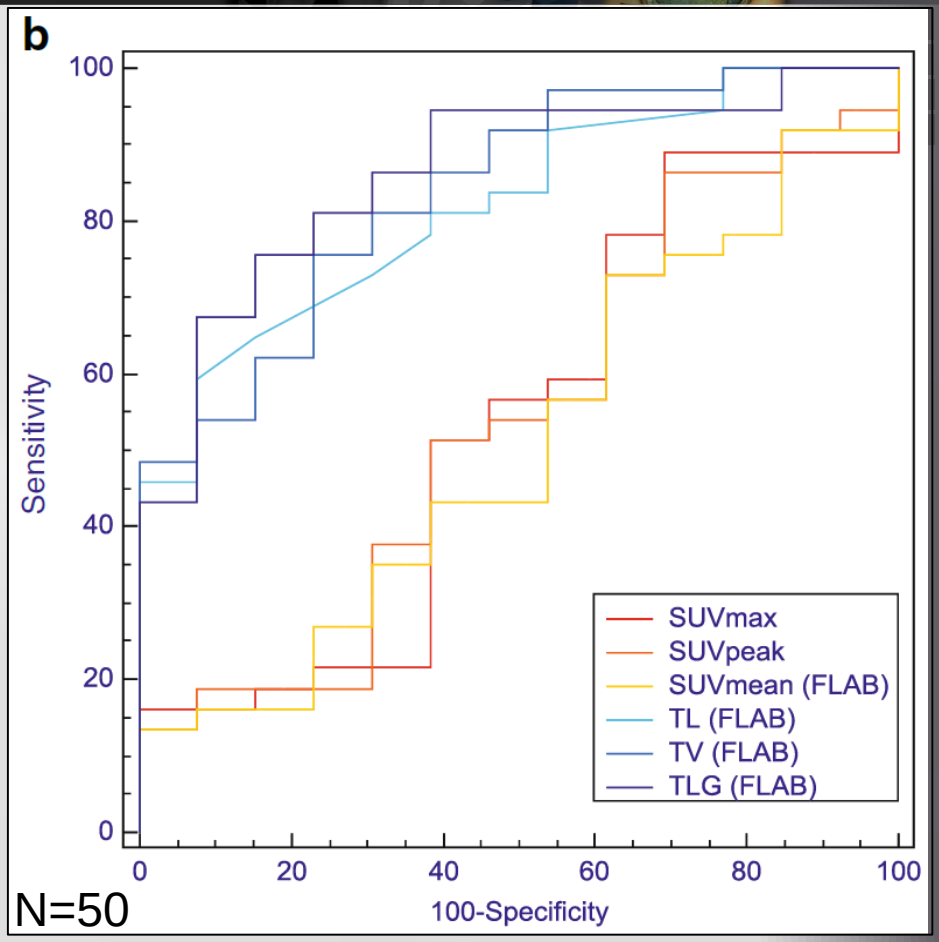


Genomics (and other –omics)

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

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, yields a minimum comparison volume of 45 cm^3.

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

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

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

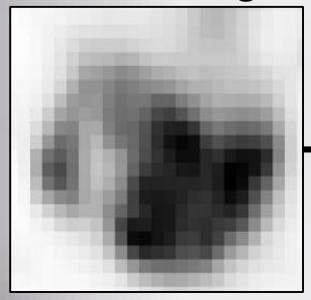
Brooks, *et al*. **The effect of small tumor volumes on studies of intratumoral heterogeneity of tracer uptake**. *J Nucl Med* 2014

Radiomics

Challenges and issues: the complexity of textural features

- Workflow complexity

PET image



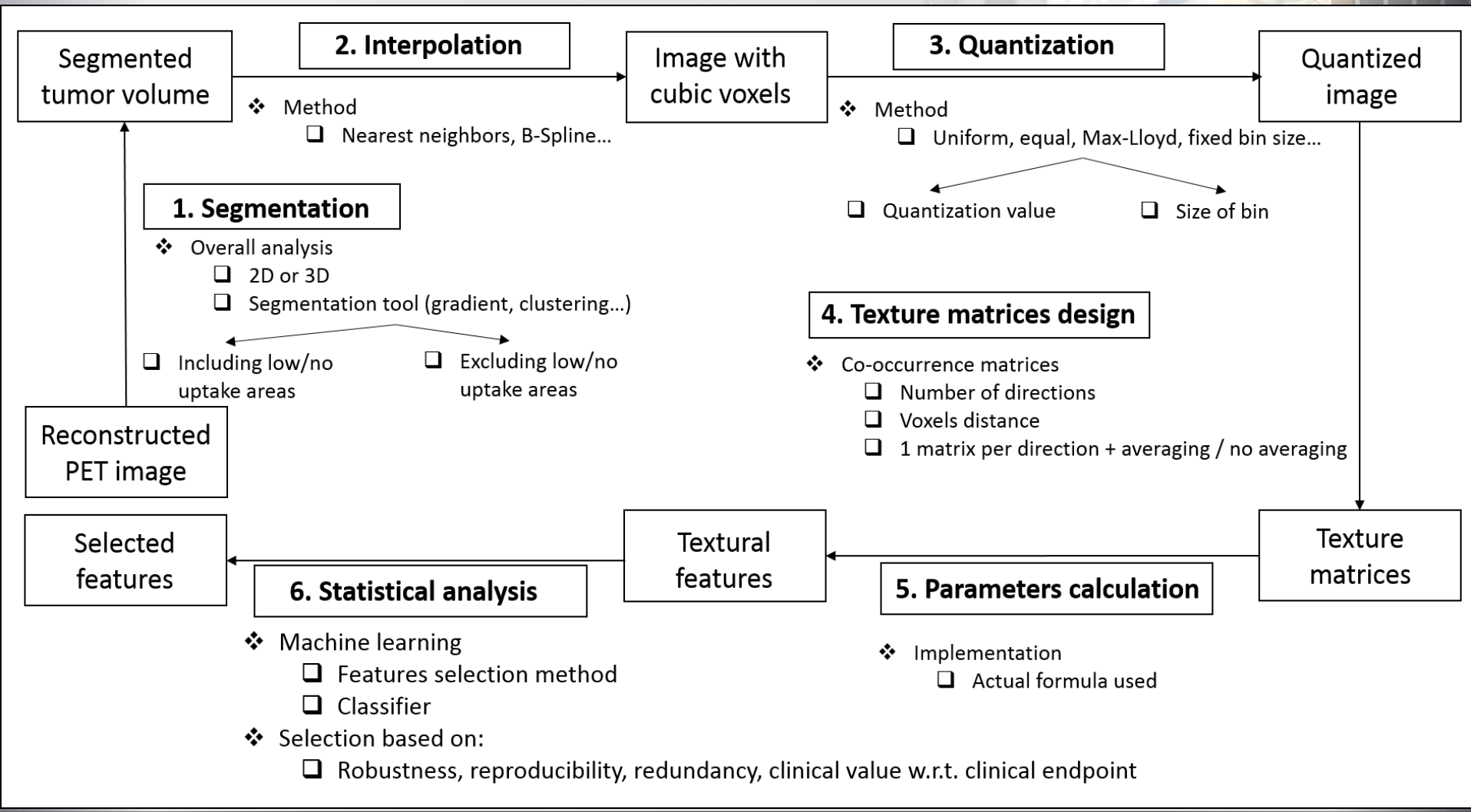
« Let's compute some textural features!»

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

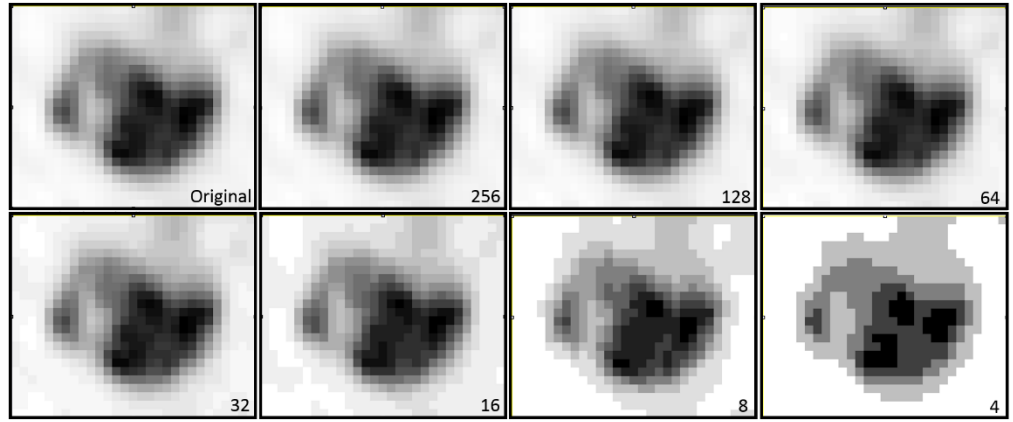


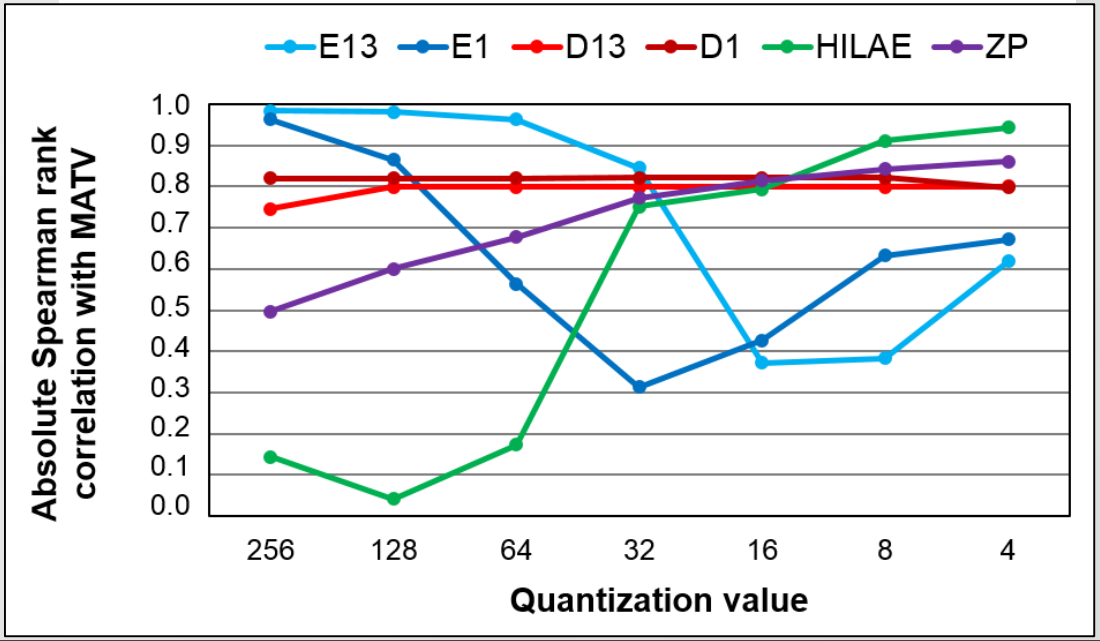
Hatt, *et al.* **Characterization of PET/CT images using texture analysis: the past, the present… any future?** *Eur J Nucl Med Mol Imaging* 2017

Radiomics

Challenges and issues: the volume/intensity confounding issue

N=555 tumors
5 cancer types
FDG PET images

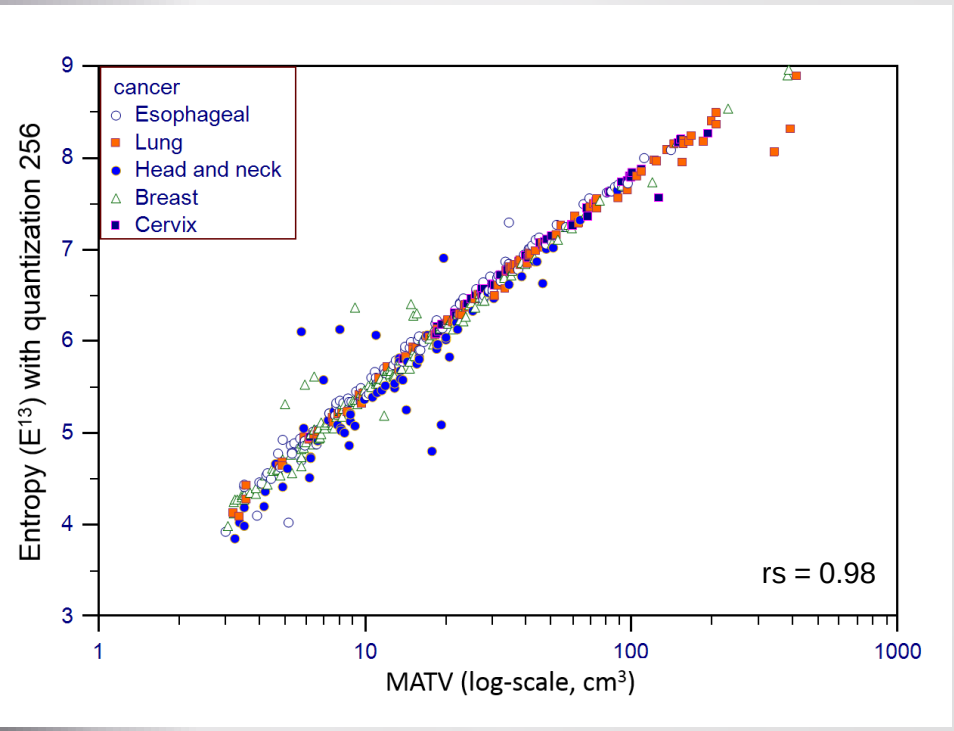




Hatt, *et al.* **^{18}F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort**. *J Nucl Med* 2015

Radiomics

Challenges and issues: the volume/intensity confounding issue

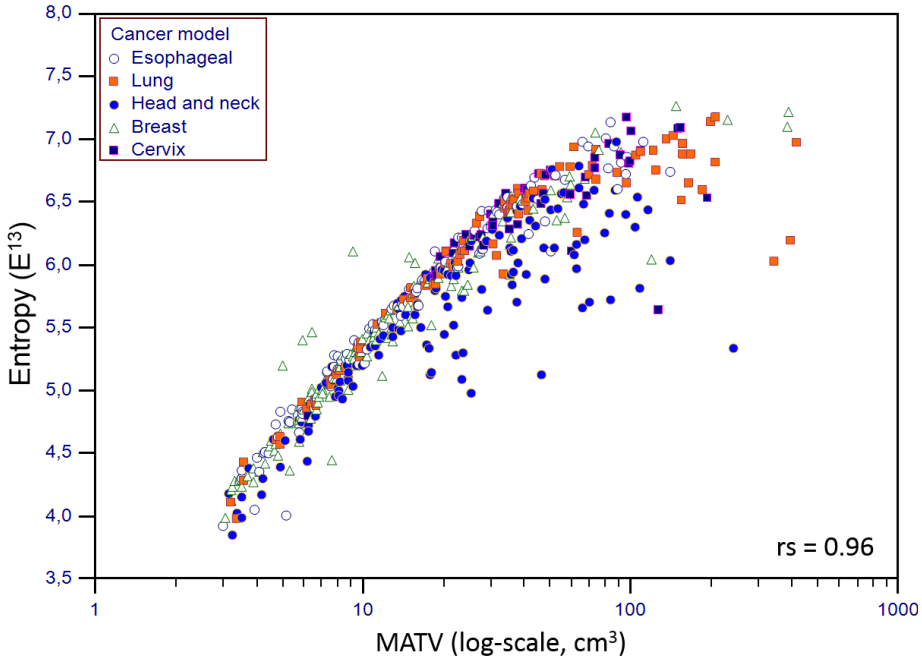


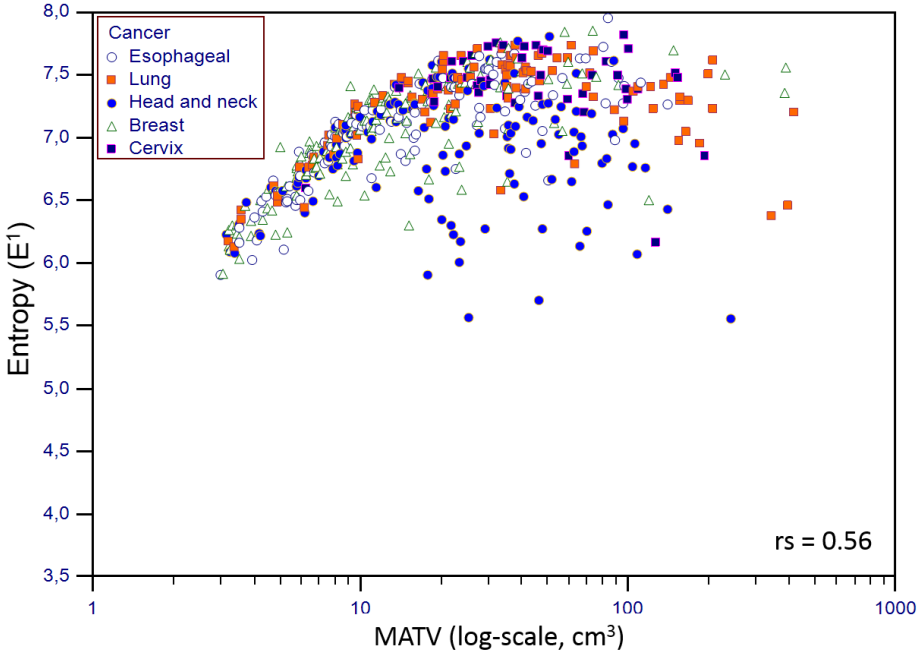
Hatt, *et al.* **^{18}F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort**. *J Nucl Med* 2015

Radiomics

Challenges and issues: the volume/intensity confounding issue

256 → 64 grey-levels

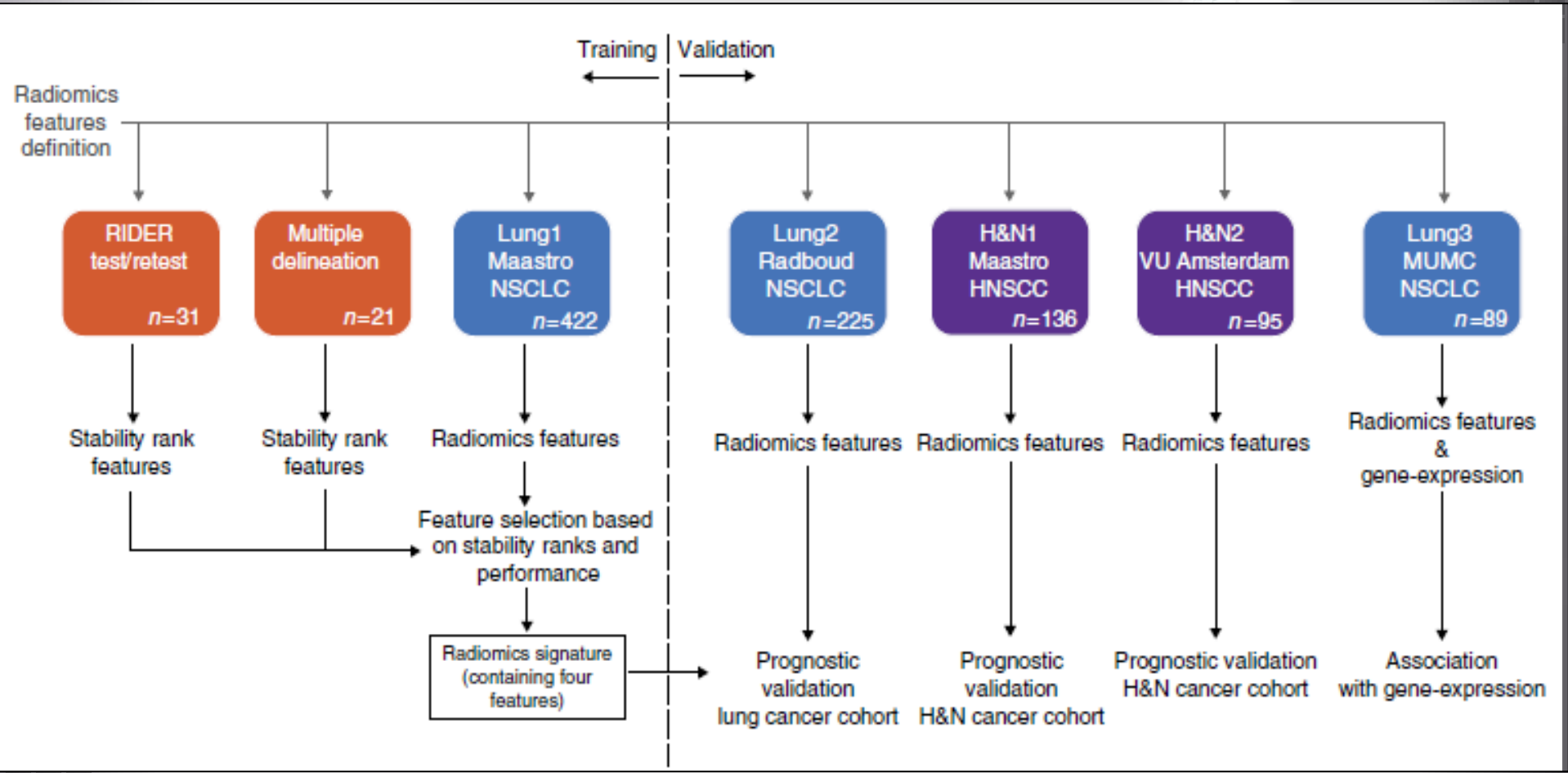




Hatt, *et al.* **^{18}F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort**. *J Nucl Med* 2015

Radiomics

Challenges and issues: the volume/intensity confounding issue

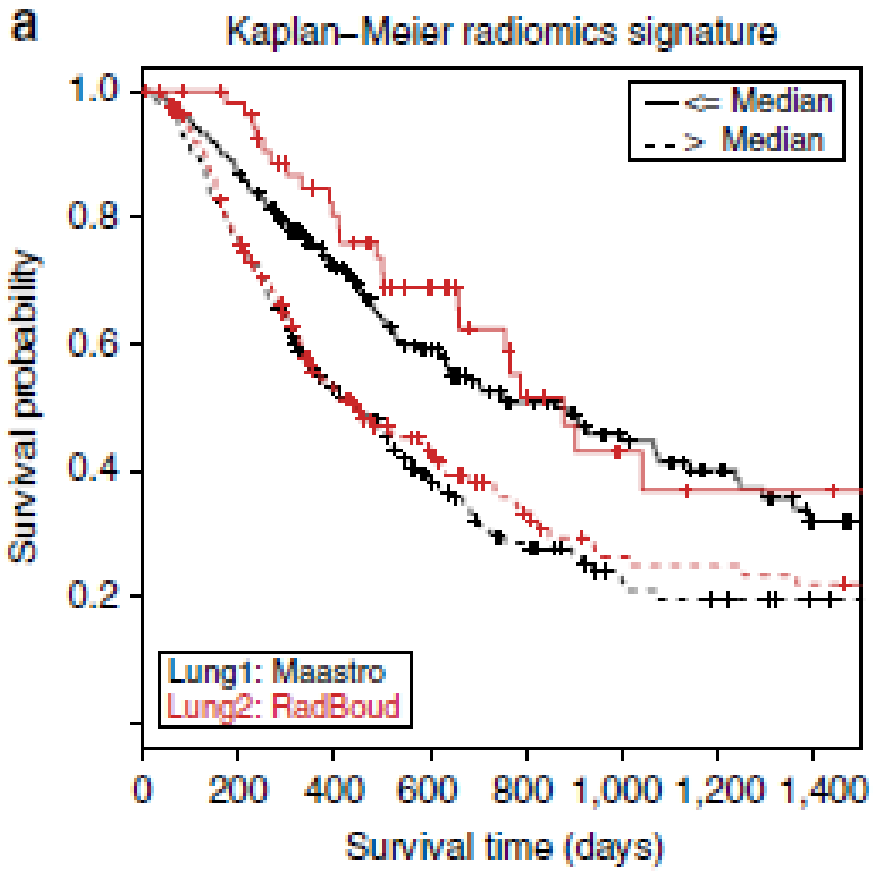


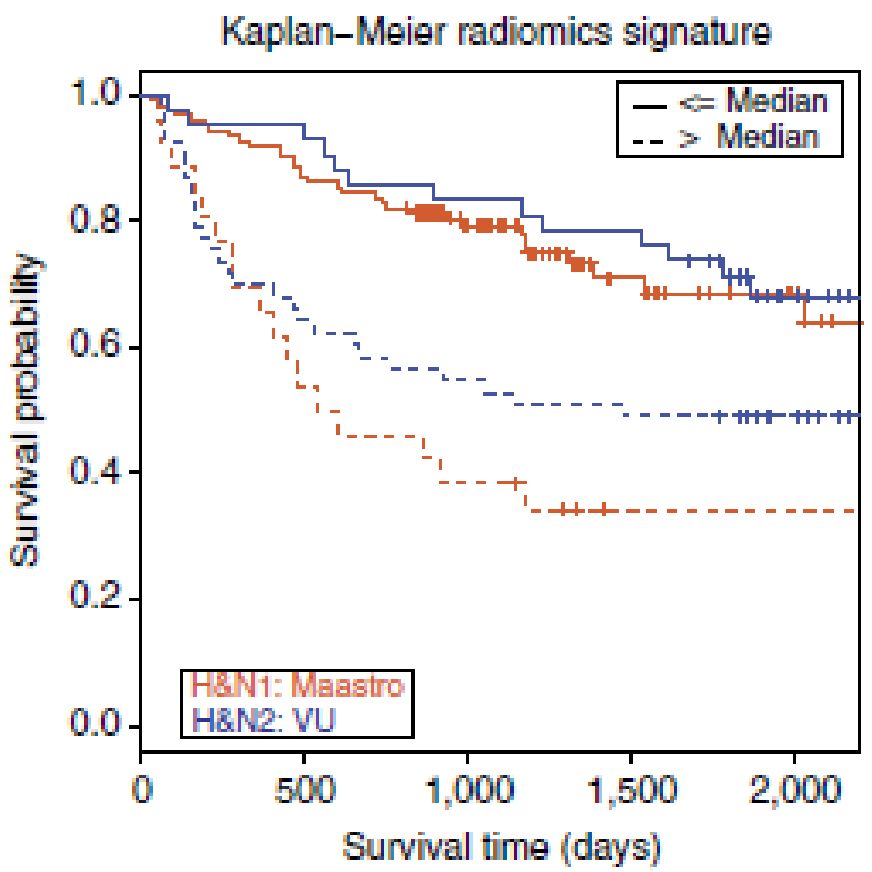
Aerts, *et al*. **Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach**. *Nat Commun*. 2014

Radiomics

Challenges and issues: the volume/intensity confounding issue

4-features signature:
intensity, shape, textural, textural in the wavelet domain
energy, compactness, grey-level non-uniformity (GLNU), GLNU high-low-high subband





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

4-features signature:

intensity, shape, textural, textural in the wavelet domain

energy, compactness, grey-level non-uniformity (GLNU), GLNU high-low-high subband

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

Rank Spearman correlations (N=300 head and neck cancer patients):
energy: 0.62, compactness: 0.80, GLNU: 0.99, GLNU_HLH: 0.94

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

Radiomics

Challenges and issues: dependency on reconstruction

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.

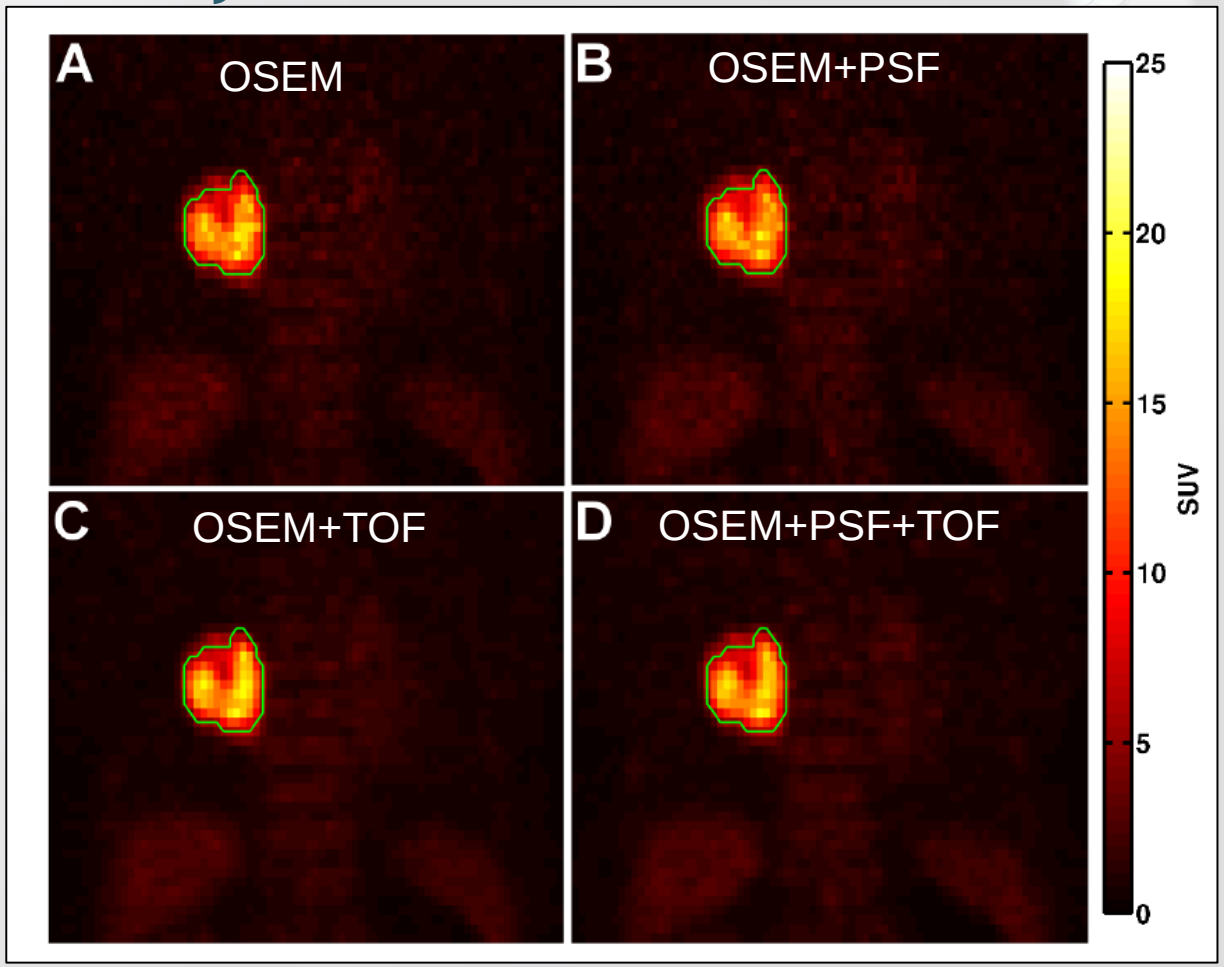
Gavalis, *et al*. **Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters**. *Acta Oncol*. 2010

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

Radiomics

Challenges and issues: dependency on reconstruction

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: dependency on reconstruction

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 !

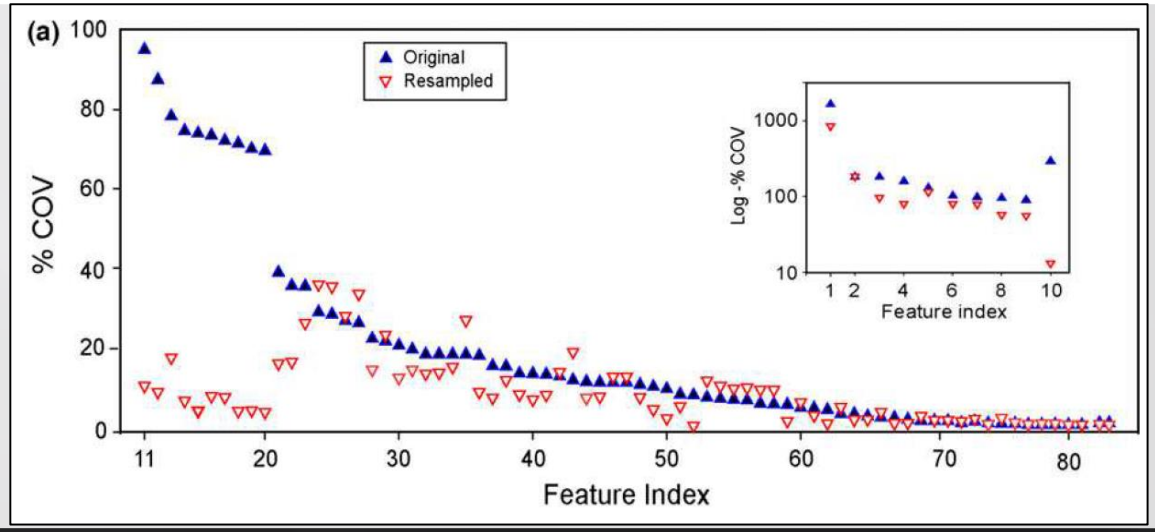
Gavalis, *et al*. **Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters**. *Acta Oncol*. 2010
Yan, *et al*. **Impact of Image Reconstruction Settings on Texture Features in 18F-FDG PET.** *J Nucl Med* 2015

Radiomics

Challenges and issues: dependency on reconstruction

Dependency on reconstruction: CT

CT Scanner	KVp	mAs	Scan Type	Pitch	Rotation time (Sec)	Reconstruction Kernel	Detector Configuration (mm)
GE Discovery STE (GE1)	120	250*	Helical	0.984	1.0	Standard	Det. Coverage = 40
GE Lightspeed 32 pro (GE2)	120	250*	Helical	0.984	1.0	Standard	Det. Coverage = 40
Philips Big Bore (P1)	120	250	Helical	1.024	1.0	Standard (B)	16 × 0.75
Philips Brilliance 64 (P2)	120	250	Helical	1.024	1.0	Standard (B)	64 × 0.625
Siemens Definition AS (S1)	120	250	Helical	1.0	1.0	I31f-2	64 × 0.625
Siemens Sensation 64 (S2)	120	250	Helical	1.0	1.0	B31f	64 × 0.625
Siemens Sensation 40 (S3)	120	250	Helical	1.0	1.0	B31f	40 × 0.625
Siemens Sensation 16 (S4)	120	250	Helical	1.0	1.0	B31f	16 × 0.75



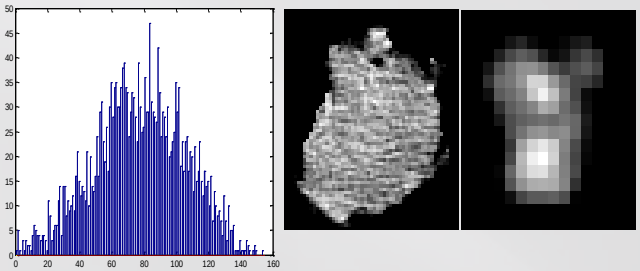
Shafiq-Ul-Hassan, *et al*. **Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels**. *Med Phys*. 2017

Radiomics

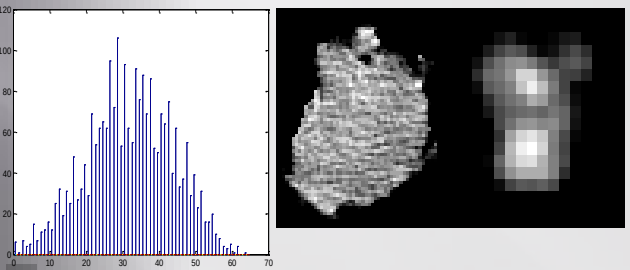
Challenges and issues: the complexity of textural features

- Dependency on pre-processing quantization
  - Quantization/discretization is required to build texture matrices

Original

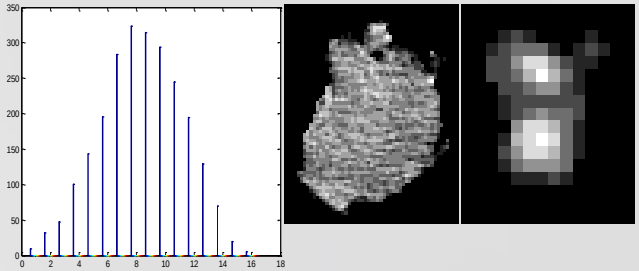


Linear transform[1]



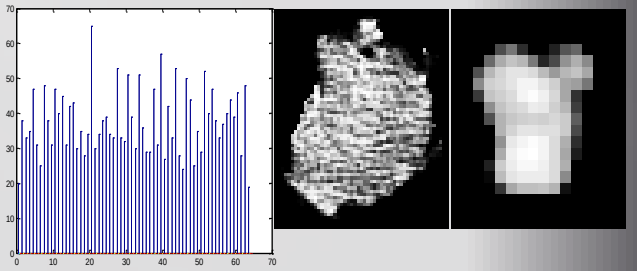
$$I_L(p) = 1 + (N-1) * \frac{I(p) - I_{min}}{I_{max} - I_{min}}$$

Fixed width bins[2]



$$I_F(p) = 1 + E\left[\frac{I(p) - I_{min}}{F}\right]$$

Histogram equalization[3]



$$I_E(p) = 1 + (N-1) * hic(p)$$

1. Tixier, *et al.* **Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in FDG PET**. *J Nucl Med* 2012
2. Leijenaar, *et al.* **The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis**. *Acta Oncol* 2013
3. Haralick, *et al.* **Textural Features for Image Classification**. *IEEE Transactions on Systems, Man, and Cybernetics* 1973

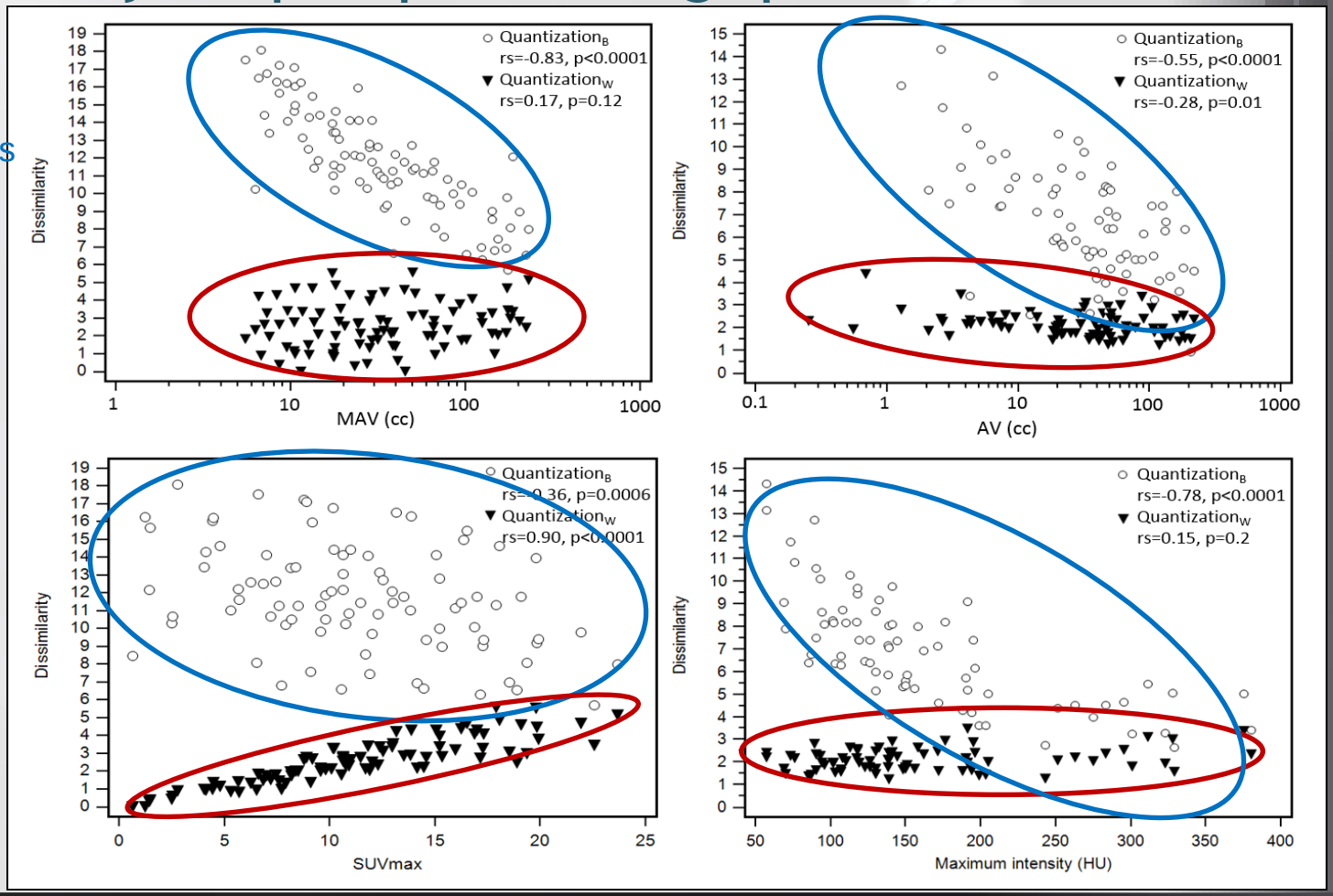
Radiomics

Challenges and issues: the complexity of textural features

Dependency on pre-processing quantization

Linear transform into 64 bins
$Quantization_B$

Fixed width bins
(0.5 SUV, 10 HU)
$Quantization_W$



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 2017*

Radiomics

Challenges and issues: lack of standardisation

Lack of standardisation

- Most papers do not provide enough details
  - Difficult/impossible to reproduce the results
  - Describe all choices and implementation details in appendix

- Some studies rely on available software (black boxes) that may not be properly validated (or at least that do not give the same results as other existing codes/software)
  - Authors may not know exactly how the features they use are actually implemented.

- → Meta-analysis impossible (entropy in paper 1 may not be the same entropy as in paper 2 !)
  - Sharing data and methods

Hatt, *et al.* **Characterization of PET/CT images using texture analysis: the past, the present… any future?** *Eur J Nucl Med Mol Imaging* 2017
Vallieres, *et al*. **Radiomics: Responsible Research For Faster Clinical Translation.** *J Nucl Med* 2017

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

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
  - 18 research groups fom 8 countries:
    - USA
    - Germany
    - The Netherlands
    - France
    - Canada
    - United Kingdom
    - Italy
    - Switzerland

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

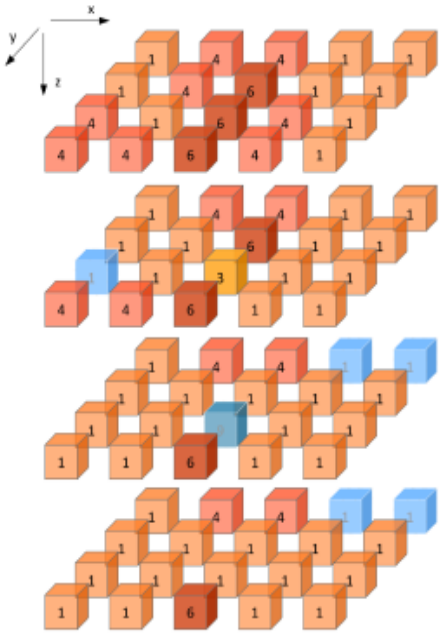
University of Groningen, University Medical Center Groningen
Lisanne V. van Dijk, Jorn Beukinga, Nanna M. Sijtsema, Roel J.H.M. Steenbakkers, Ronald Boellaard

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

Radiomics

Challenges and issues: lack of standardization

Imaging biomarkers standardisation initiative



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

Standardisation progress. Height indicates the number of features per family. LI: Local intensity; IH: intensity histogram; IVH: intensity-volume histogram; CM: co-occurrence matrix; RLM: run length matrix; SZM: size zone matrix; NGTDM: neighbourhood grey tone difference matrix; DZM: distance zone matrix; NGLDM: neighbouring grey level difference matrix

Current status:

		current	09-10-16
• no agreement	(< 3 institutions or < 50% identical)	52	192
• agreement	(> 50% identical)	59	85
• standardised	(> 80% identical)	240	25

Conclusions:

- Benchmarking of features is recommended: high initial differences
- Standard values found for most features

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

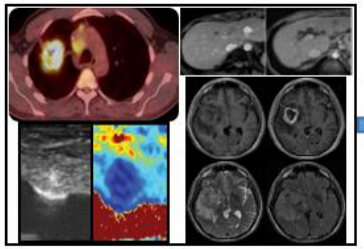
University of California, San Francisco
Olivier Morin

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

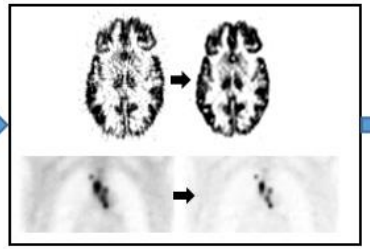
Image Biomarker Standardisation Initiative. **Multicentre initiative for standardization of image biomarkers**. **https://arxiv.org/abs/1612.07003**

Radiomics
Worfklow

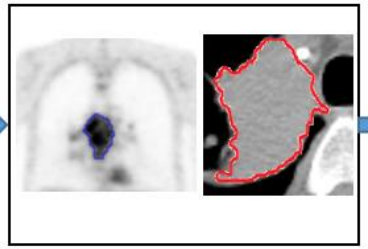
Clinical data



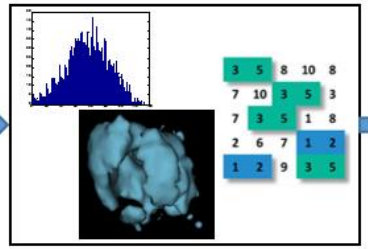
Images



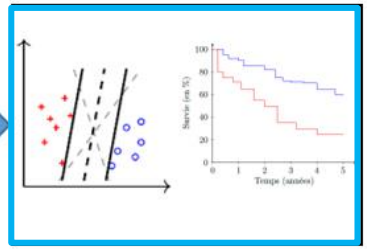
Pre-processing



Segmentation

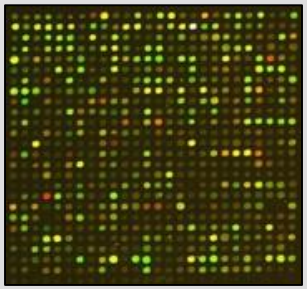


Features extraction



Features selection
Training & validation

Radiomics



Genomics (and other –omics)

Radiomics

Challenges and issues: statistical analysis

- Inapproriate statistical analyses

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: how to use machine learning?

- Machine learning
  - Choosing a classifier/feature selection method?

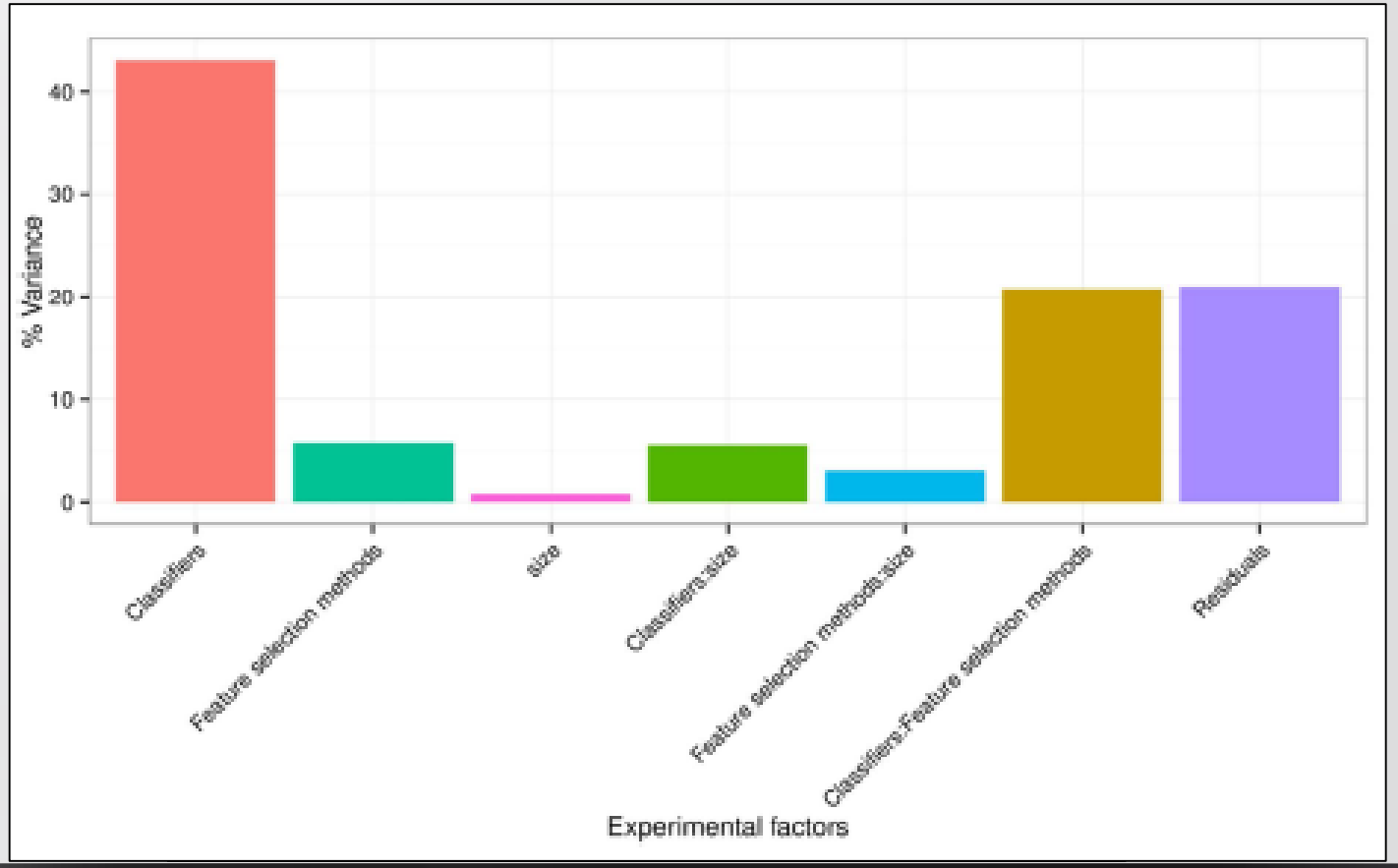
Classification method acronym	Classification method name	Feature Selection method acronym	Feature selection method name
Nnet	Neural network	RELF	Relief
DT	Decision Tree	FSCR	Fisher score
BST	Boosting	GINI	Gini index
BY	Bayesian	CHSQ	Chi-square score
BAG	Bagging	JMI	Joint mutual information
RF	Random 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

Radiomics

Challenges and issues: how to use machine learning?

- Machine learning
  - Choosing a classifier/feature selection method?

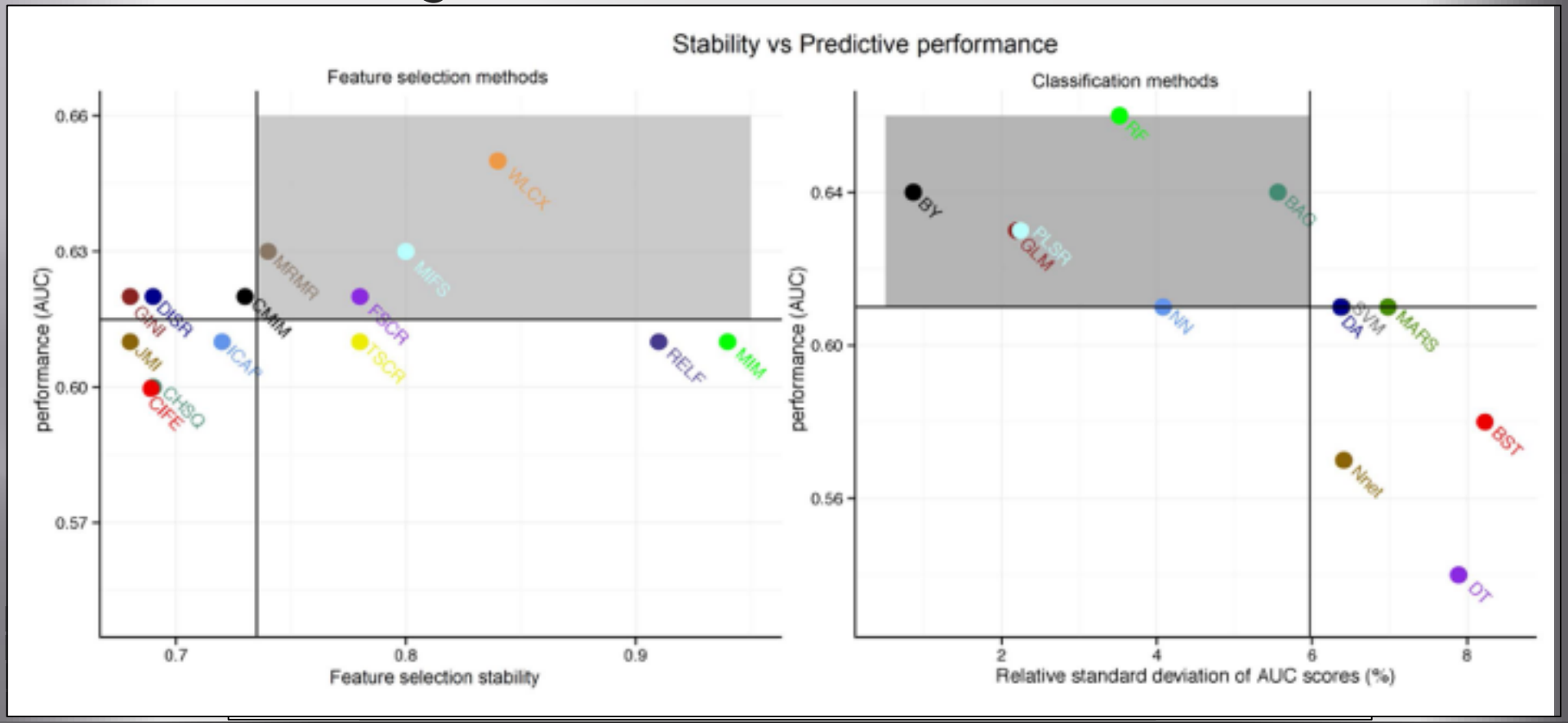


Parmar, *et al*. **Machine Learning methods for Quantitative Radiomic Biomarkers**. *Sci Rep.* 2015

Radiomics

Challenges and issues: how to use machine learning?

- Machine learning
  - Choosing a classifier/feature selection method?

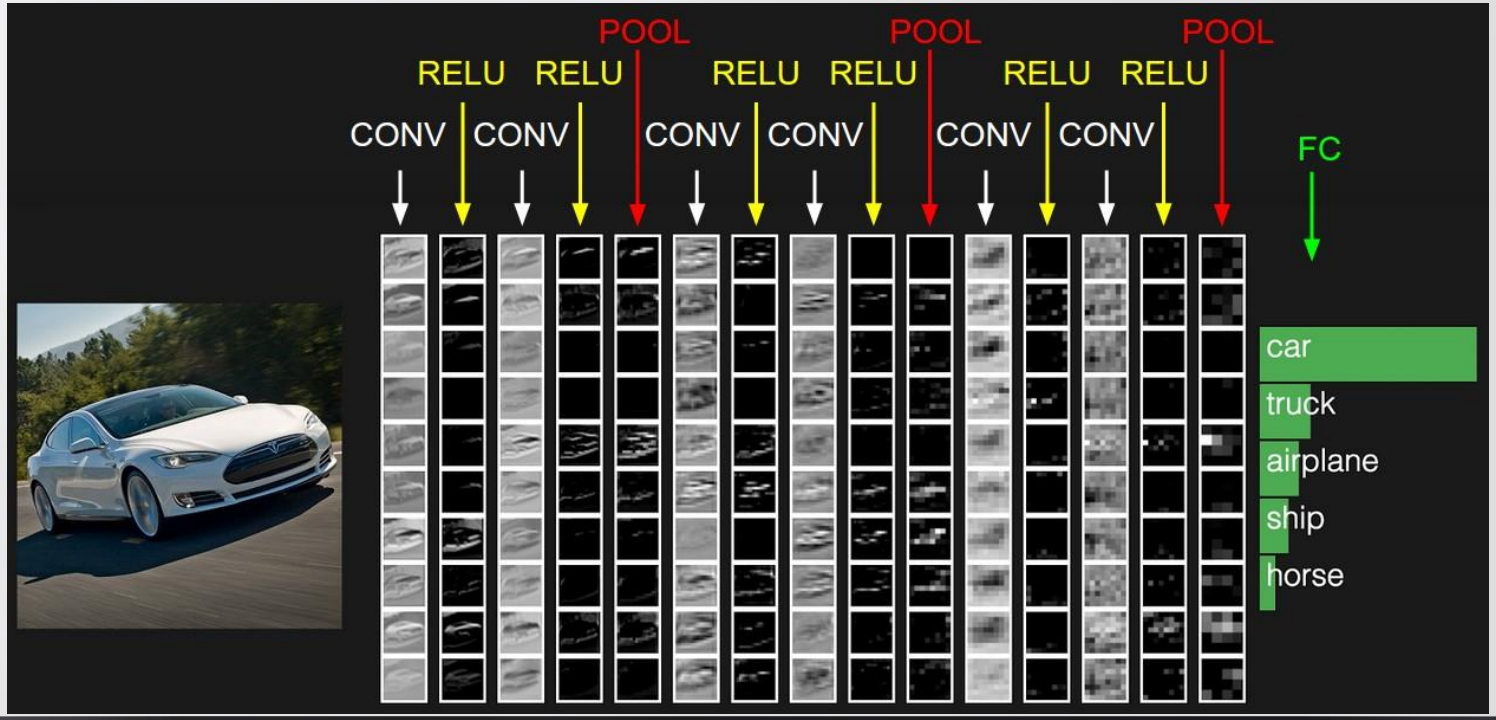


Parmar, *et al*. **Machine Learning methods for Quantitative Radiomic Biomarkers**. *Sci Rep.* 2015

Radiomics

Challenges and issues: potential of deep learning?

- Machine learning
  - Potential for deep learning approaches?
  - Convolutional Neural Networks = recent evolution of neural networks

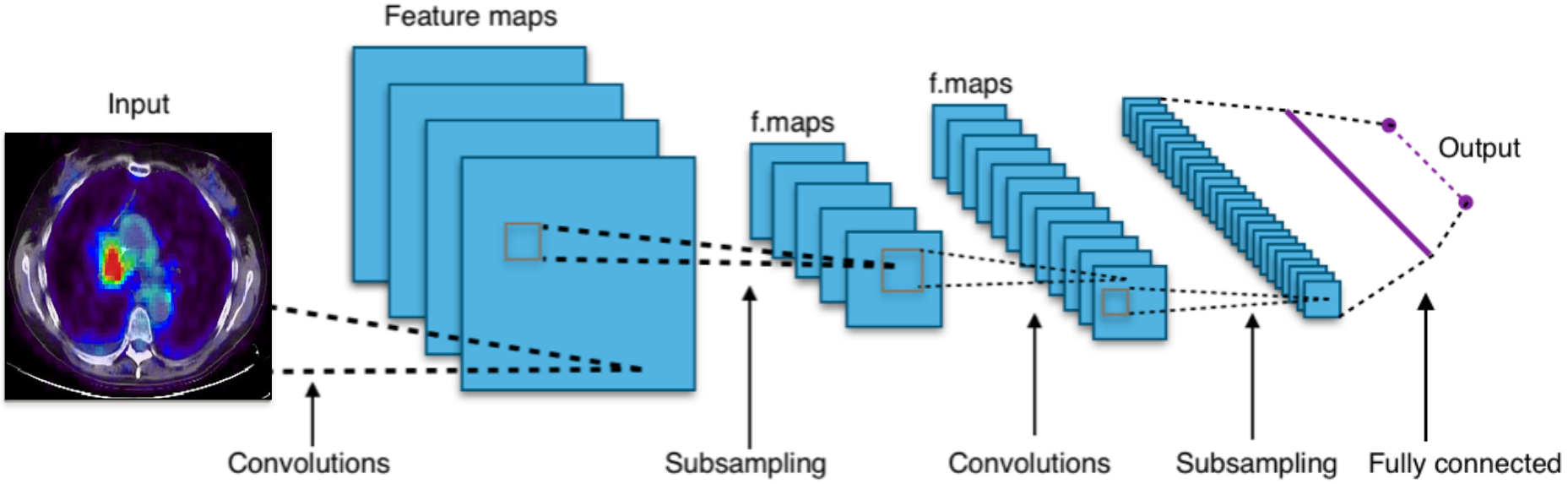


Lecun, *et al*. **Deep Learning**. *Nature.* 2015
Suzuki. **Overview of deep learning in medical imaging**. *Radiol Phys Technol.* 2017

Radiomics

Challenges and issues: potential of deep learning?

- Machine learning
  - Potential for deep learning approaches?
  - Convolutional Neural Networks = recent evolution of neural networks



Lecun, *et al*. **Deep Learning**. *Nature.* 2015
Suzuki. **Overview of deep learning in medical imaging**. *Radiol Phys Technol.* 2017

Radiomics

Challenges and issues: potential of deep learning?

- Deep learning applications

Journal of Medical Imaging 3(3), 034501 (Jul–Sep 2016)

Digital mammographic tumor classification using transfer learning from deep convolutional neural networks

Benjamin Q. Huynh, Hui Li, and Maryellen L. Giger*
University of Chicago, Department of Radiology, 5841 South Maryland Avenue, Chicago, Illinois 60637, United States

SCIENTIFIC REPORTS

OPEN

Bladder Cancer Treatment Response Assessment in CT using Radiomics with Deep-Learning

: 22 September 2016
: 18 July 2017

Kenny H. Cha[1], Lubomir Hadjiiski[1], Heang-Ping Chan[1], Alon Z. Weizer[2], Ajjai Alva[3], Richard H. Cohan[1], Elaine M. Caoili[1], Chintana Paramagul[1] & Ravi K. Samala[1]

SCIENTIFIC REPORTS

OPEN

A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme

: 24 May 2017
: 11 August 2017

Jiangwei Lao[1], Yinsheng Chen[2], Zhi-Cheng Li[3], Qihua Li[3], Ji Zhang[2], Jing Liu[4] & Guangtao Zhai[1]

SCIENTIFIC REPORTS

OPEN

Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma

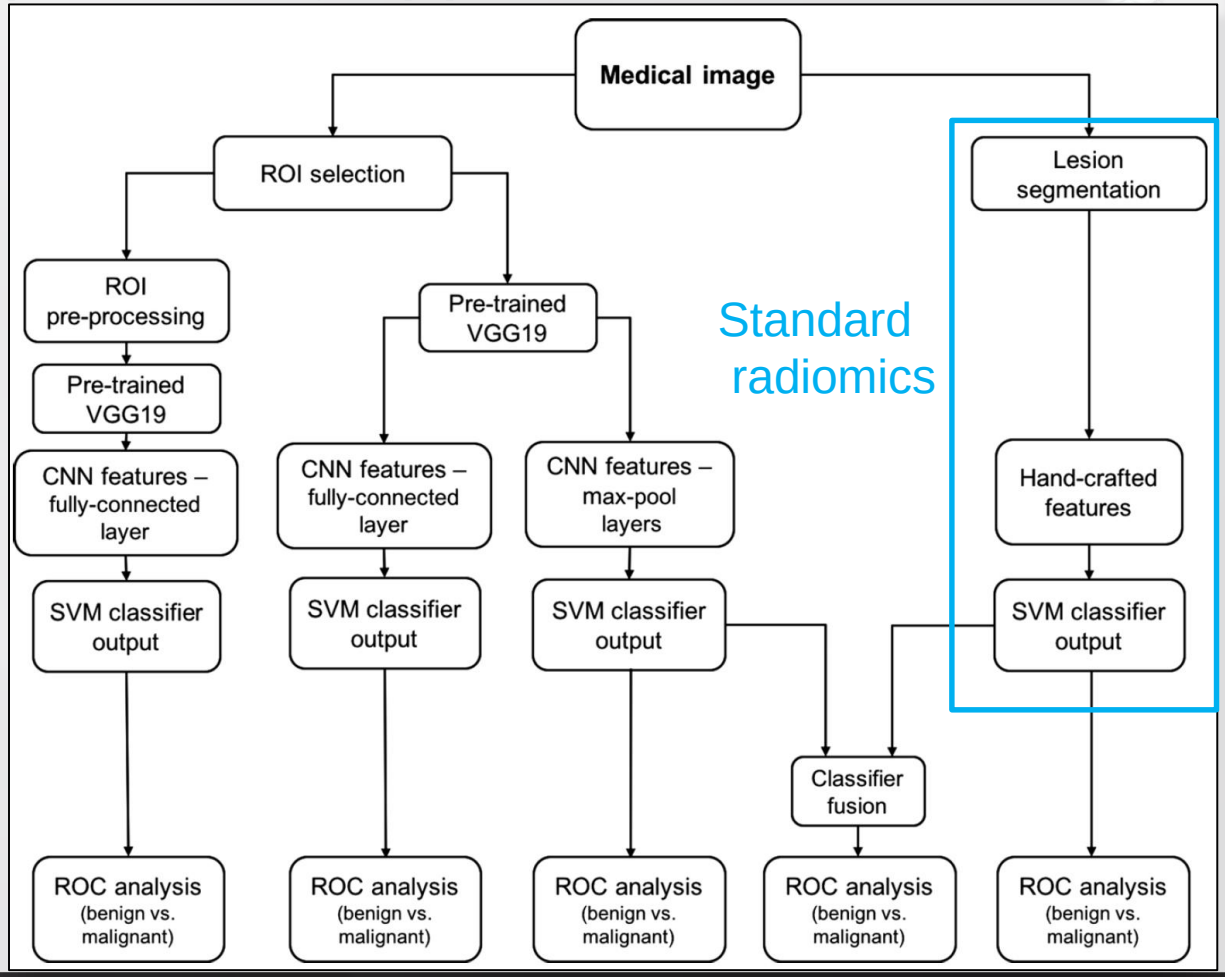
: 14 December 2016
: 25 May 2017
Online: 14 July 2017

Zeju Li[1], Yuanyuan Wang[1,2], Jinhua Yu[1,2], Yi Guo[1,2], Wei Cao[3]

Radiomics

Challenges and issues: potential of deep learning?

- Deep learning applications



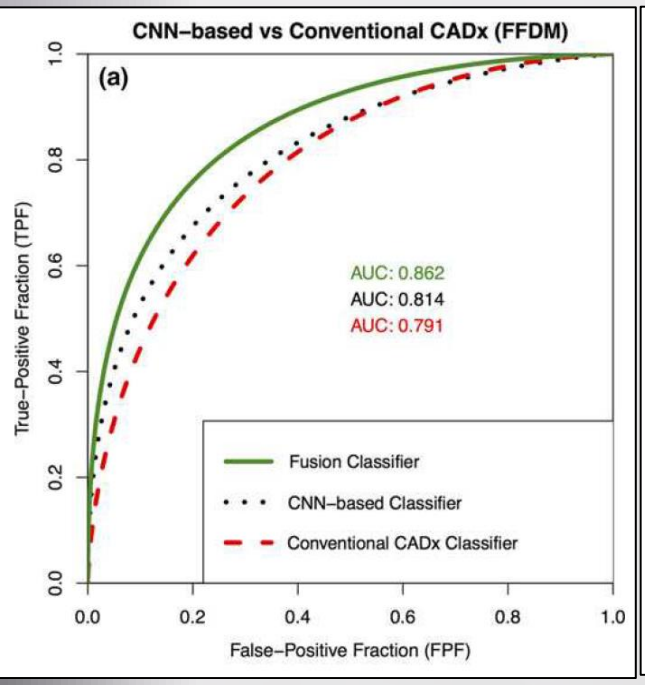
Antropova, *et al*. **A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets**. *Med Phys* 2017

Radiomics

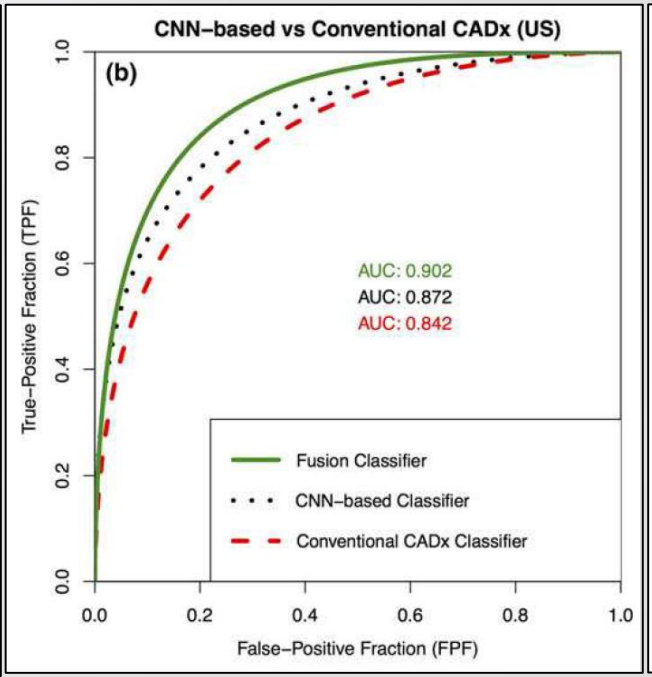
Challenges and issues: potential of deep learning?

Deep learning applications

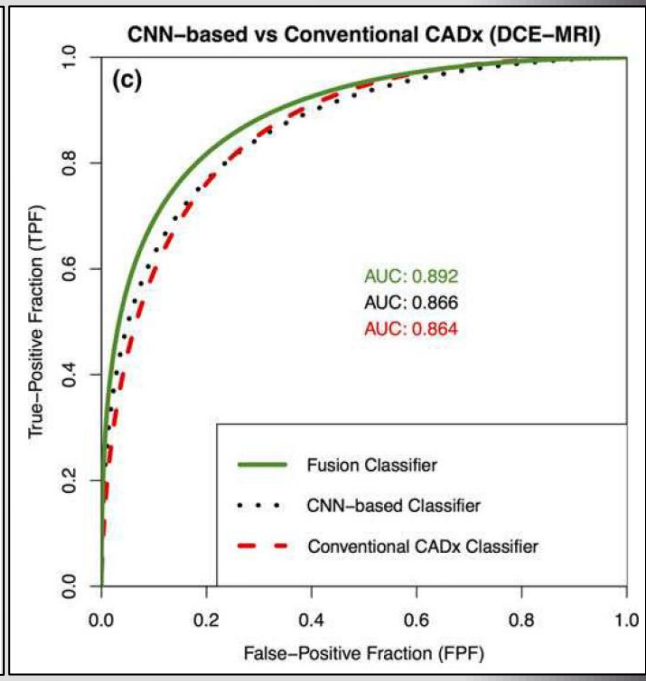
Full field digital mamography (FFDM)
N=245



Ultrasound (US)
N=1125



DCE-MRI
N=690

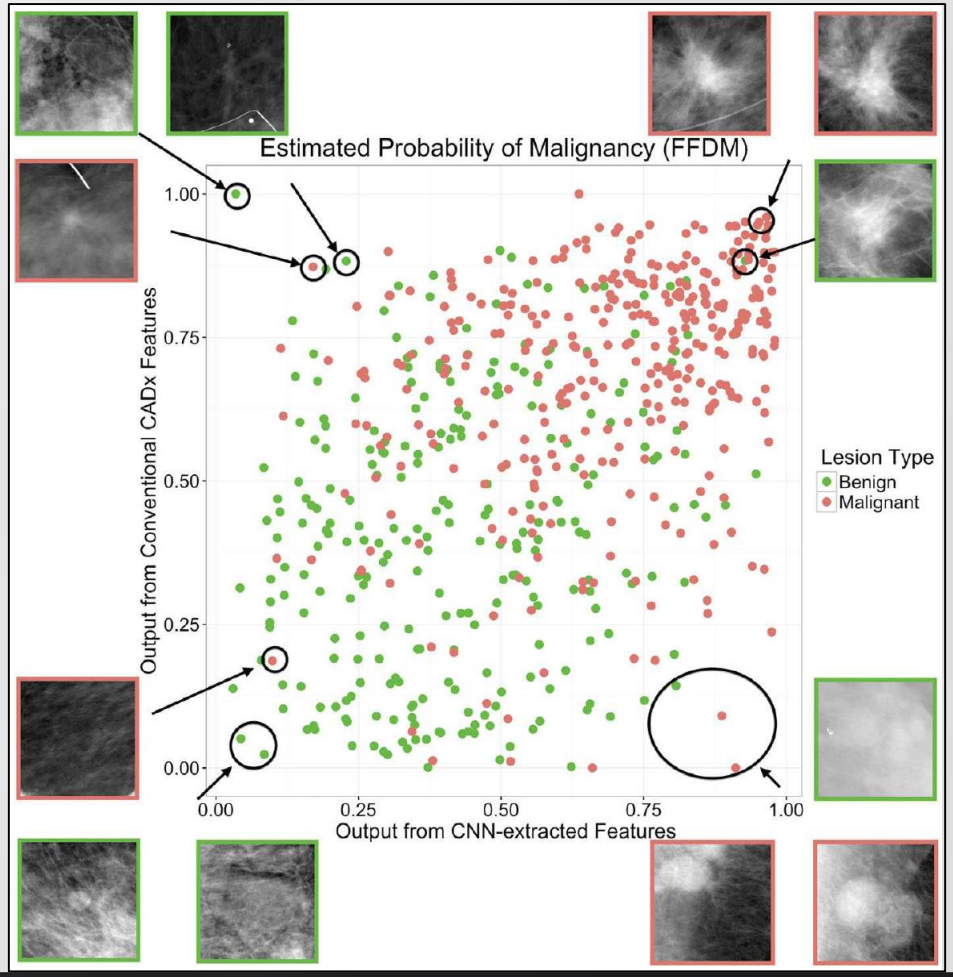


Antropova, *et al*. **A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets**. *Med Phys* 2017

Radiomics

Challenges and issues: potential of deep learning?

- Deep learning applications



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

Radiomics

Conclusions

- Radiomics
  - Very 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

DEVELOPING PREDICTIVE MARKERS FOR PAIN AND ANALGESIA

Eugene Duff

UNIVERSITY OF OXFORD

FMRIB

wellcome centre integrative neuroimaging

pipa group

DEPARTMENT OF PAEDIATRICS
Medical Sciences Division

Motivation: addressing translational challenges for FMRI

FMRI has provided extensive understanding of pain related neural dynamics associated with an enormous range of factors:

- Analgesics
- Placebo/nocebo
- Sensitization
- Attention
- Reward responsiveness
- Negative Emotions
- Relative Relief
- etc..

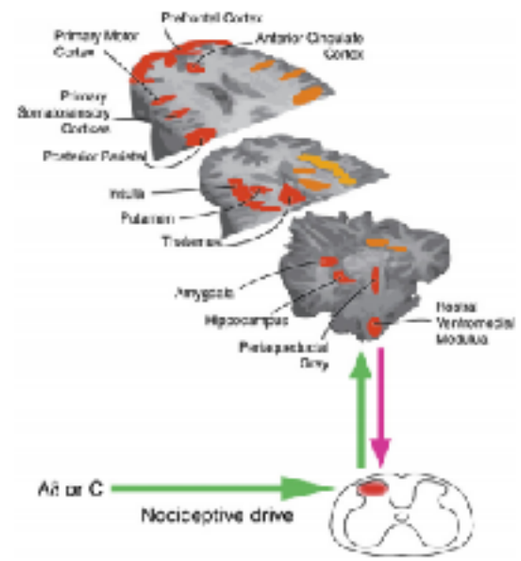
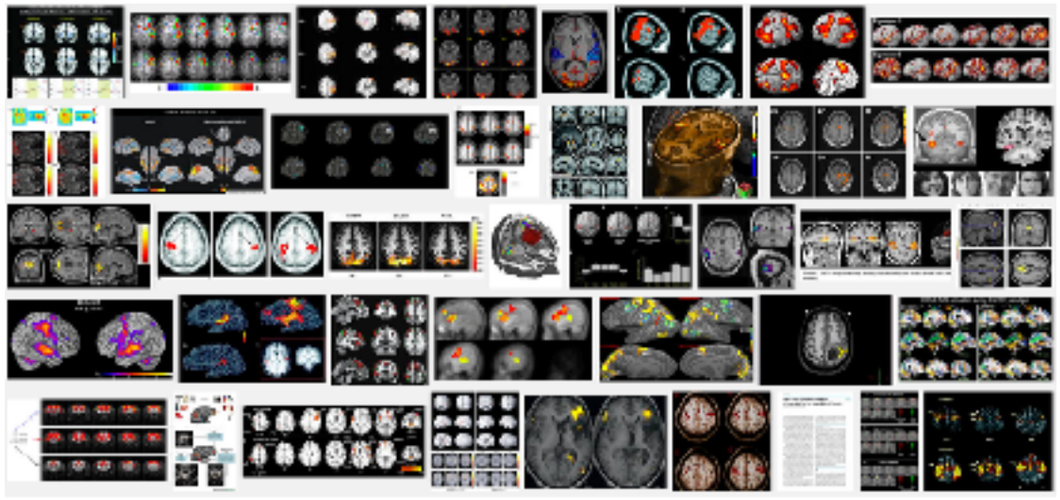


Figure 2. Neuroanatomy of Pain Processing
Main brain regions that activate during a painful experience. Highlighted as bilaterally active but with increased activation on the contralateral hemisphere (orange).

Yet, clinical translation of this understanding remains an ambition.



→ Meta-analyses
Improved analyses
New types of study?

An fMRI-Based Neurologic Signature of Physical Pain

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

An fMRI-Based Neurologic Signature of Physical Pain

Tor D. Wager, Ph.D., Lauren Y. Atlas, Ph.D., Martin A. Lindquist, Ph.D., Mathieu Roy, Ph.D., Choong-Wan Woo, M.A., and Ethan Kross, Ph.D.

METHODS

In four studies involving a total of 114 participants, we developed an fMRI-based measure that predicts pain intensity at the level of the individual person. In study 1, we used machine-learning analyses to identify a pattern of fMRI activity across brain regions — a neurologic signature — that was associated with heat-induced pain. The pattern included the thalamus, the posterior and anterior insulae, the secondary somatosensory cortex, the anterior cingulate cortex, the periaqueductal gray matter, and other regions. In study 2, we tested the sensitivity and specificity of the signature to pain versus warmth in a new sample. In study 3, we assessed specificity relative to social pain, which activates many of the same brain regions as physical pain. In study 4, we assessed the responsiveness of the measure to the analgesic agent remifentanil.

An fMRI-Based Neurologic Signature of Physical Pain

Table 1. Pain-Classification Performance, According to Study.*

Study	Discrimination between Pain and No Pain†				Effect Size‡		P Value	Performance on Forced-Choice Test§
	Signature-Response Threshold	Sensitivity	Specificity	Positive Predictive Value	AUC	Discriminability		
			percent (95% CI)					*percent (95% CI)*
Study 1								
Painful vs. warm¶	1.40	95 (86–100)	95 (86–100)	95 (85–100)	0.95	2.69	<0.001	100 (100–100)
Pain vs. pain anticipation	0.36	100 (100–100)	99 (96–100)	95 (86–100)	0.99	3.69	<0.001	100 (100–100)
Pain vs. pain recall	0.54	95 (85–100)	94 (89–98)	79 (64–92)	0.96	2.35	<0.001	100 (100–100)
Study 2								
Painful vs. warm‖**	1.32	93 (84–100)	93 (84–100)	93 (84–100)	0.92	1.54	<0.001	100 (100–100)
Painful vs. near pain threshold††	2.50	88 (77–97)	85 (72–95)	85 (73–96)	0.88	1.74	<0.001	100 (100–100)
High vs. low warmth	1.00	56 (36–75)	100 (100–100)	100 (100–100)	0.79	1.31	0.001	100 (100–100)
Study 3								
Painful vs. warm	1.40‡‡	85 (76–94)	78 (67–89)	80 (68–89)	0.86	1.64	<0.001	93 (86–98)
Painful vs. rejecter	1.40‡‡	85 (76–94)	73 (61–84)	76 (65–86)	0.88	1.83	<0.001	95 (89–100)
Rejecter vs. friend	1.40‡‡	27 (16–38)	88 (79–95)	69 (50–88)	0.57	0.31	0.22	56 (43–69)
Study 4								
Painful vs. warm, before drug treatment	1.40‡‡	90 (79–100)	81 (65–95)	83 (67–95)	0.89	1.61	<0.001	90 (79–100)
Painful vs. warm, during drug treatment	1.61	86 (73–96)	62 (42–80)	69 (52–84)	0.74	1.01	0.003	76 (61–90)
Painful before vs. during drug treatment	1.61	86 (72–96)	62 (43–79)	69 (54–83)	0.74	1.01	0.003	76 (60–92)

* Study 1 included 12 trials each in painful and warm conditions. Study 2 included a mean (±SD) of 24±13 trials for pain and 36±9 trials for warmth, depending on the ratings. Study 3 included 8 trials each in painful and warm conditions. Study 4 included 3 trials for pain and 3 for warmth in the before-drug-treatment condition and in the condition with peak drug concentration. CI denotes confidence interval.

† The tradeoff between sensitivity and specificity at different thresholds was assessed by means of receiver-operating-characteristic (ROC) plots; the signature-response threshold that minimized overall classification errors is reported here.

‡ For the area under the ROC curve, chance is 0.5. Discriminability is a measure of effect size under a gaussian model. Performance varied across studies, according to the number of trials averaged to form the condition maps.

§ For the two-choice (forced-choice) discrimination test, the classification threshold for the difference between paired observations is 0. The sensitivity, specificity, and positive predictive value are the same and are equal to the decision accuracy.

¶ Painful conditions were defined as temperatures greater than 44.5°C and as ratings of more than an average of 5.80 points on a visual-analogue scale (VAS), and warm conditions as temperatures of less than 44.5°C and ratings of less than 3.34 points on the VAS.

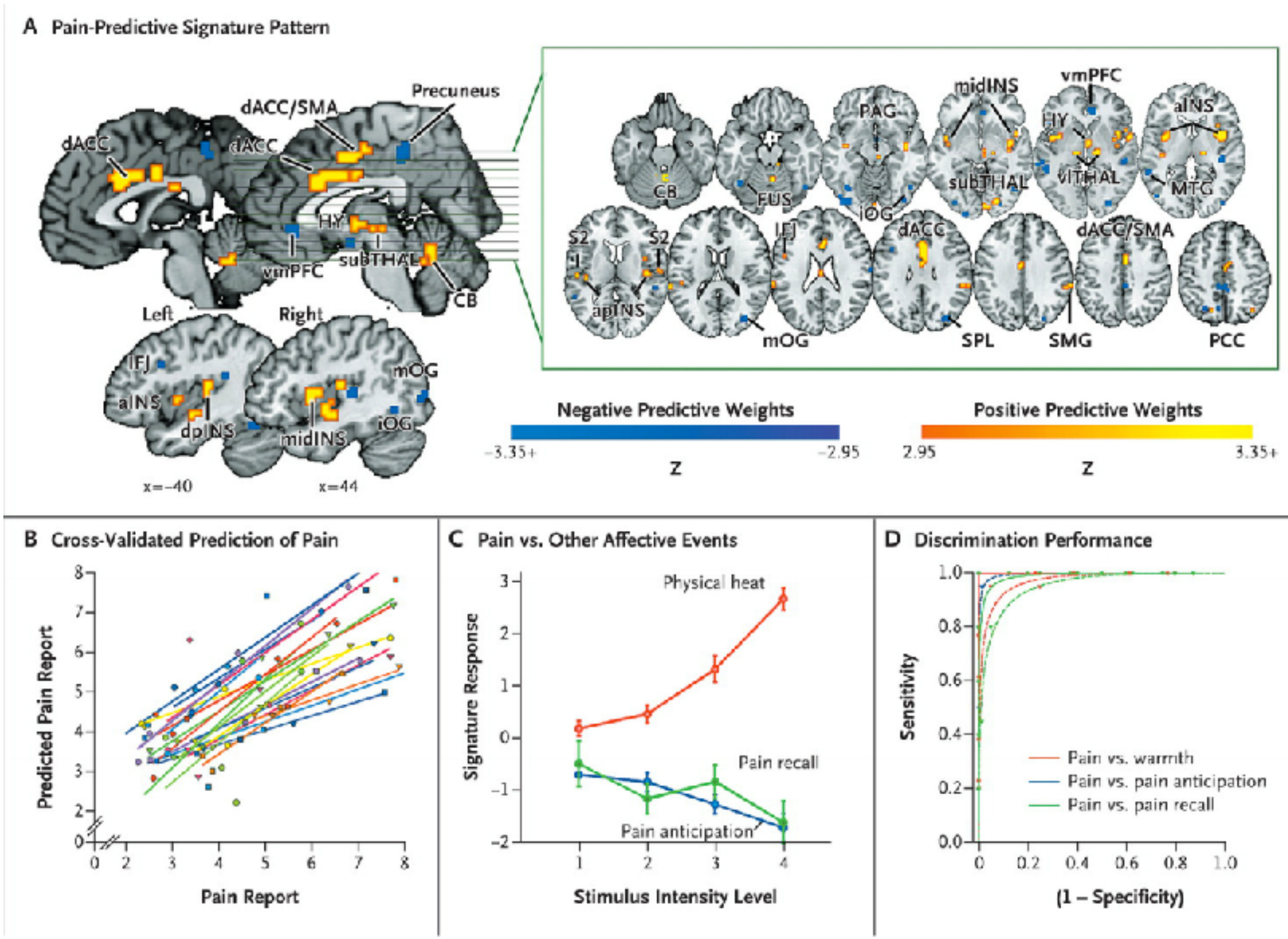
‖ Study 2 was conducted with the use of a scanner with a different field strength (3 T), so the threshold was reestimated.

** Participants made judgments of painful versus nonpainful conditions for each trial.

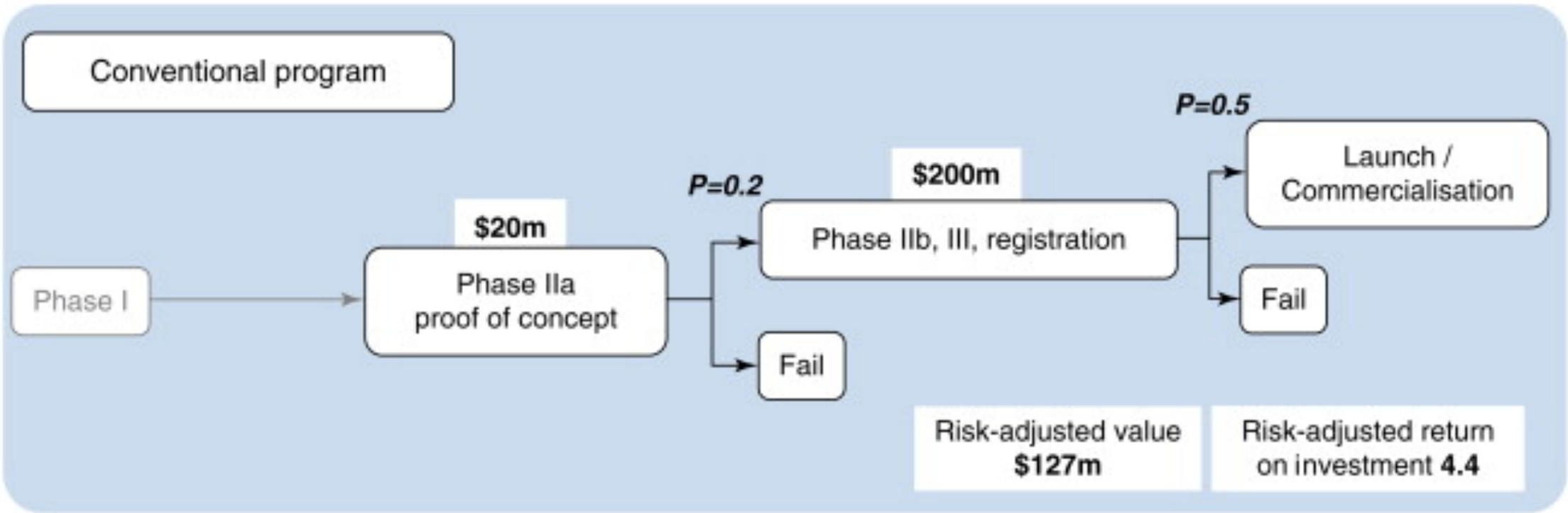
†† Participants rated pain or warmth intensity on a continuous VAS, with scores ranging from 0 to 99 points for warmth and from 100 to 200 points for pain. Pain was defined as a score of more than 125 points, near the pain threshold as a score of 75 to 125 points, high warmth as a score of 50 to 100 points, and low warmth as a score of 0 to 50 points.

‡‡ The threshold derived from study 1 was applied.

An fMRI-Based Neurologic Signature of Physical Pain

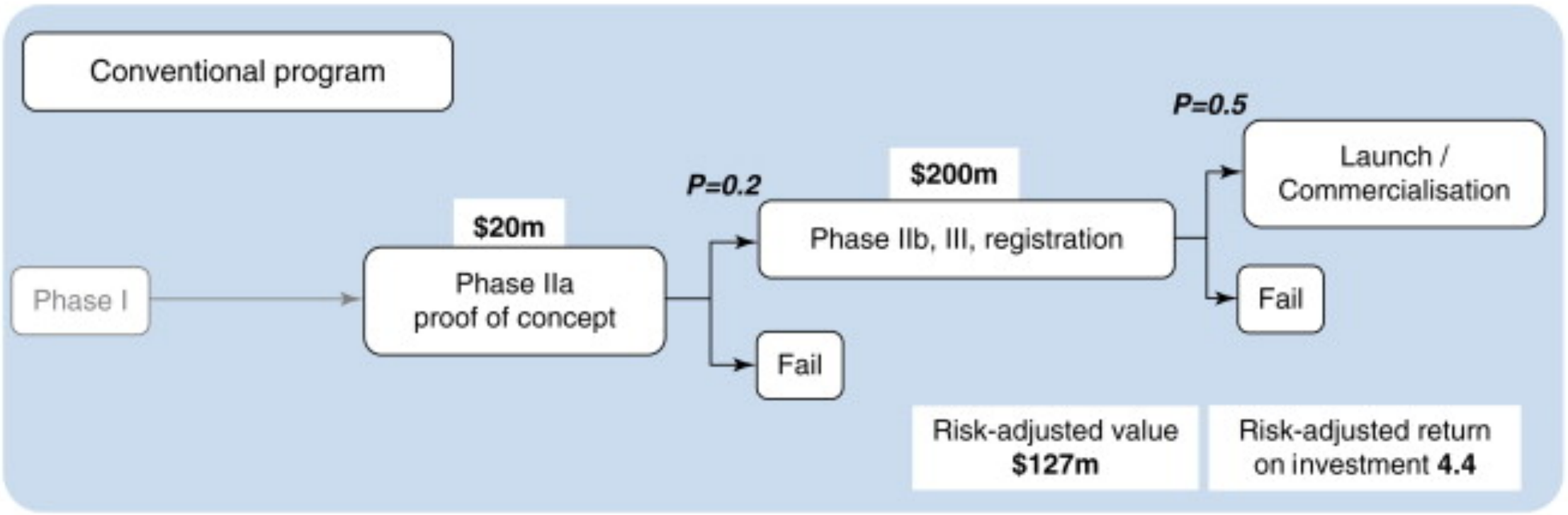


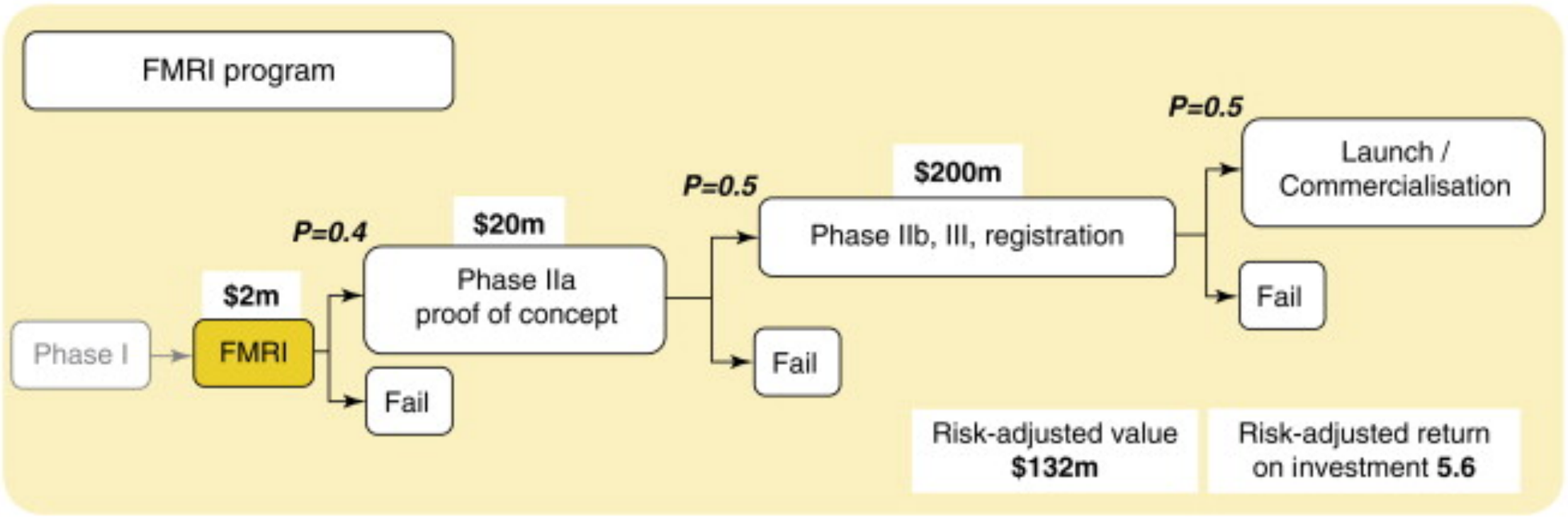
What is the value of human FMRI in CNS drug development?



Preston & Wise, 2010, Drug Discovery Today

What is the value of human FMRI in CNS drug development?

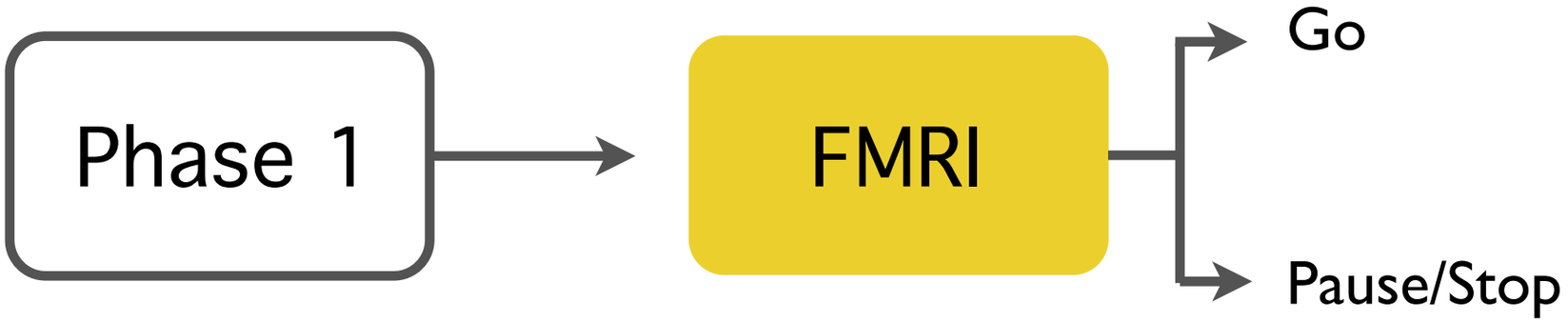




Drug Discovery Today

Preston & Wise, 2010, Drug Discovery Today

Decision making: Predictive imaging



Decision-making protocol ideally suited to multivariate prediction methods

But, what criteria/features do we use for Go / No Go decision?
How do we validate?

Three pillars of drug survival

Drug Discovery Today • Volume 17, Numbers 9/10 • May 2012

PERSPECTIVE

Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving Phase II survival

Paul Morgan[1], **Piet H. Van Der Graaf**[1,2], piet.van.der.graaf@pfizer.com, John Arrowsmith[3], Doug E. Feltner[4], Kira S. Drummond[5], Craig D. Wegner[6] and Steve D.A. Street[7]

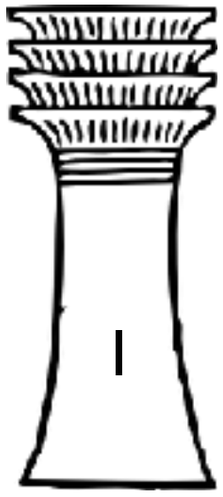
In an effort to uncover systematic learnings that can be applied to improve compound survival, an analysis was performed on data from Phase II decisions for 44 programs at Pfizer. It was found that not only were the majority of failures caused by lack of efficacy but also that, in a large number of cases (43%), it was not possible to conclude whether the mechanism had been tested adequately. A key finding was that an integrated understanding of the fundamental pharmacokinetic/pharmacodynamic principles of exposure at the site of action, target binding and expression of functional pharmacological activity (termed together as the 'three Pillars of survival') all determine the likelihood of candidate survival in Phase II trials and improve the chance of progression to Phase III.

Morgan et al 2013
Drug Discovery Today

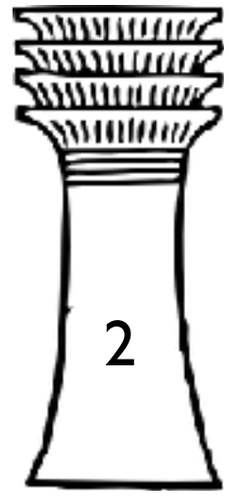
Pfizer

Three pillars of drug survival

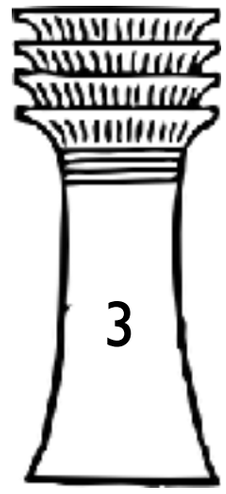
Target site exposure



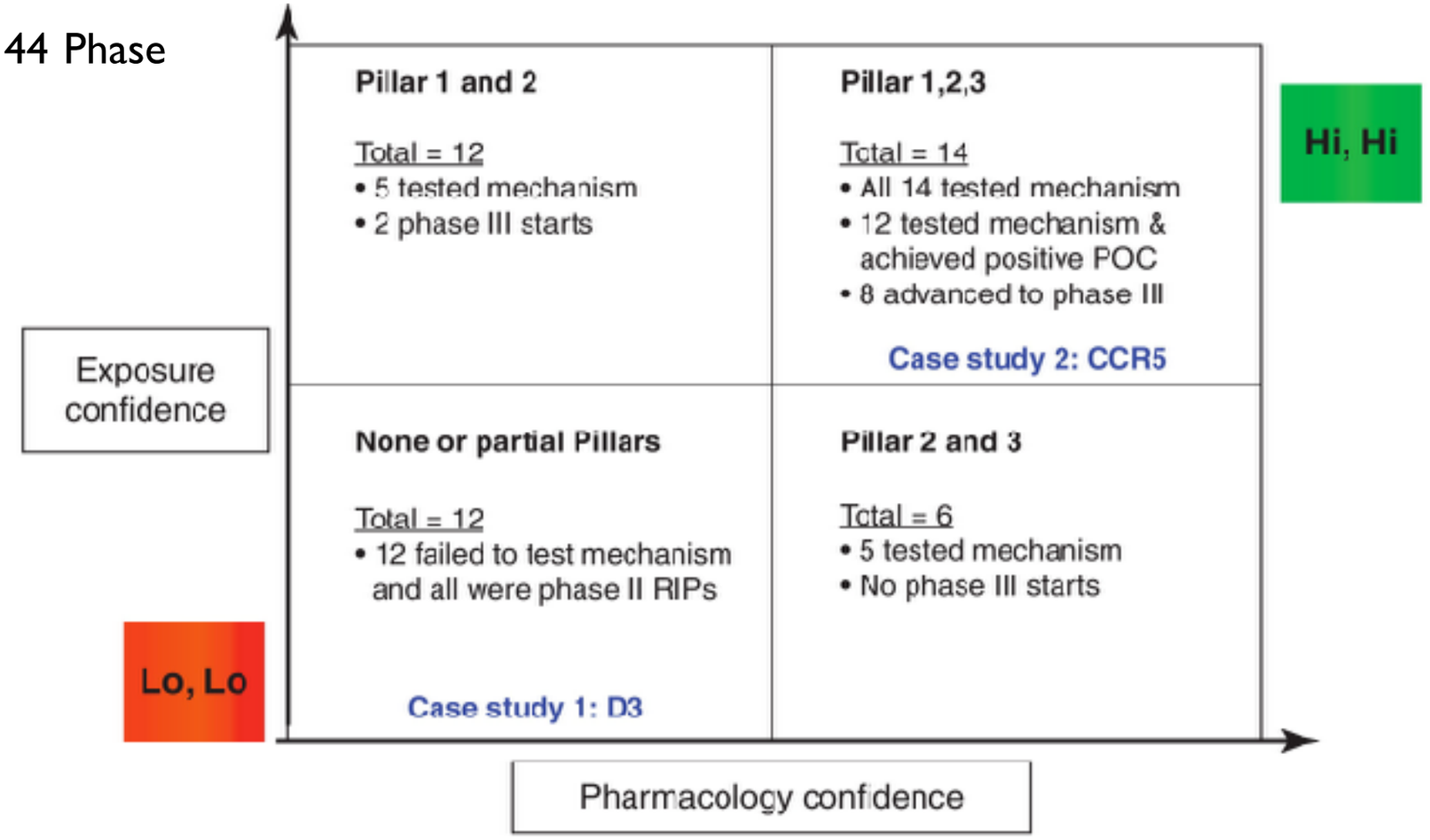
Binding to the target



Expression of functional activation



Meta-analysis of 44 Phase 2A Clinical trials

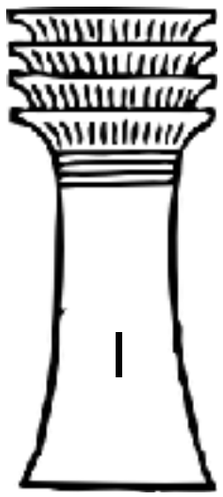


Morgan et al 2013 Drug Discovery Today

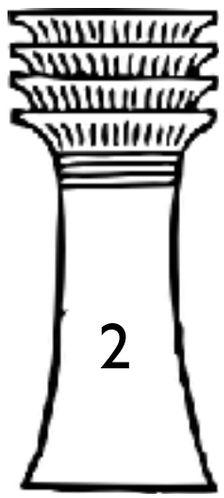
Pfizer

Three pillars of drug survival: FMRI surrogates

Target site exposure



Binding to the target



Expression of functional activation

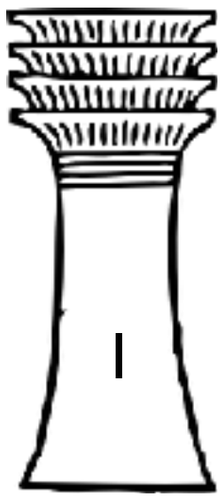
Pharmacodynamic effect

Morgan et al 2013
Drug Discovery Today

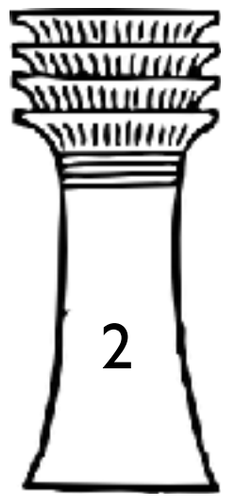
.

Three pillars of drug survival: FMRI surrogates

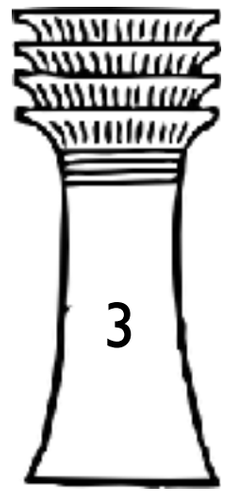
Target site exposure



Binding to the target



Expression of functional activation



Pharmacodynamic effect

Pharmacodynamic effect:

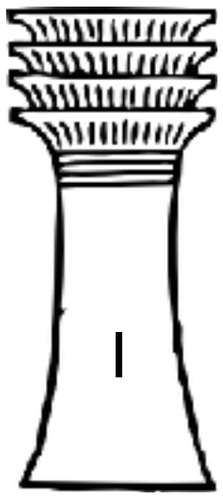
- does the drug alter brain responses in any way?
- can the drug condition be discriminated from placebo?

Morgan et al 2013
Drug Discovery Today

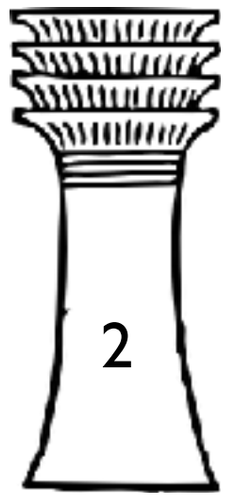
.

Three pillars of drug survival: FMRI surrogates

Target site exposure



Binding to the target



Expression of functional activation

Evidence of efficacy

Evidence of efficacy: does the modulation suggest efficacious action?

- e.g. reductions in established pain regions?
  - but will this enhance screening beyond pain ratings?
- changes common to existing efficacious compounds
  - could identify additional features predictive of analgesic action, not directly tied to pain relief
    - potentially identify signature effects earlier
    - identify signature effects in non-responders

RESEARCH ARTICLE | NEUROIMAGING

Learning to identify CNS drug action and efficacy using multistudy fMRI data

Eugene P. Duff[1,*], William Vennart[2], Richard G. Wise[3], Matthew A. Howard[4], Richard E. Harris[5], Michael Lee[1], Karolina Wartolowska[1], Vishvarani Wanigasekera[1], Frederick J. Wilson[2], Mark Whitlock[2], Irene Tracey[1], Mark W. Woolrich[1,6] and Stephen M. Smith[1]

Science Translational Medicine

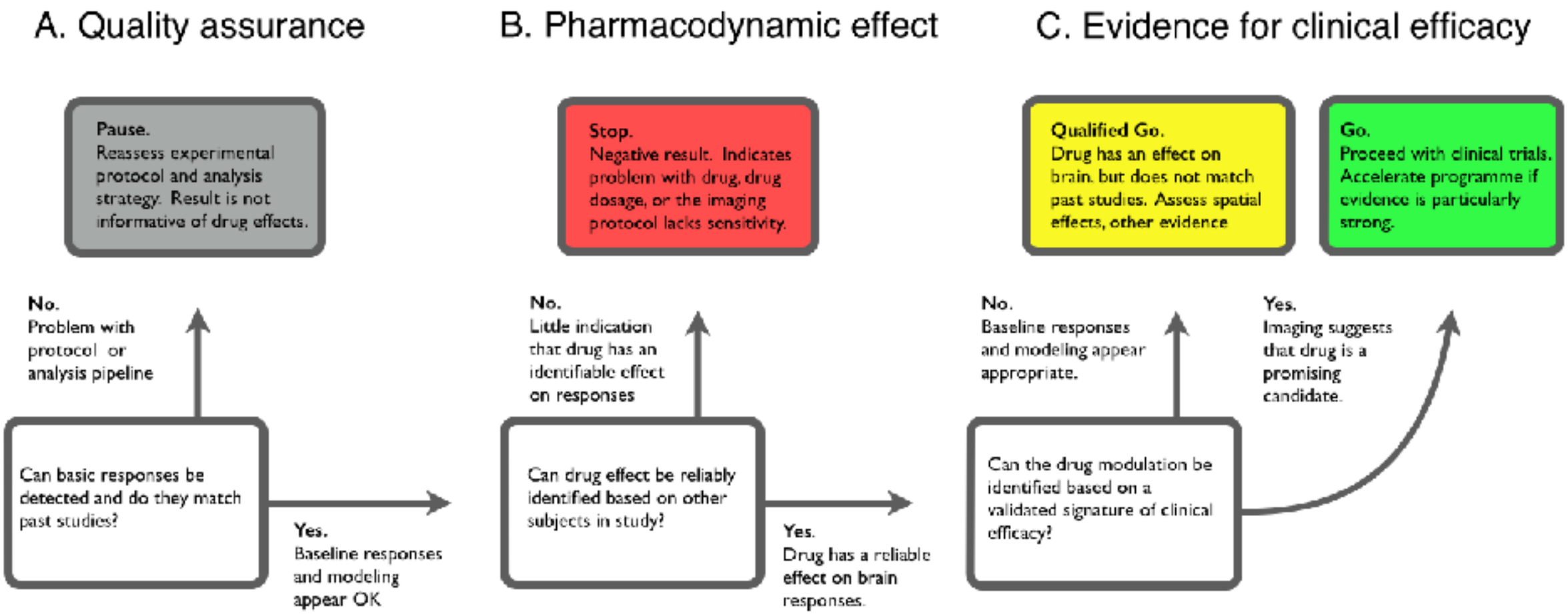
Vol 7, Issue 274
11 February 2015

Table of Contents

- Aim to learn to identify analgesic effects on pain responses from multistudy data

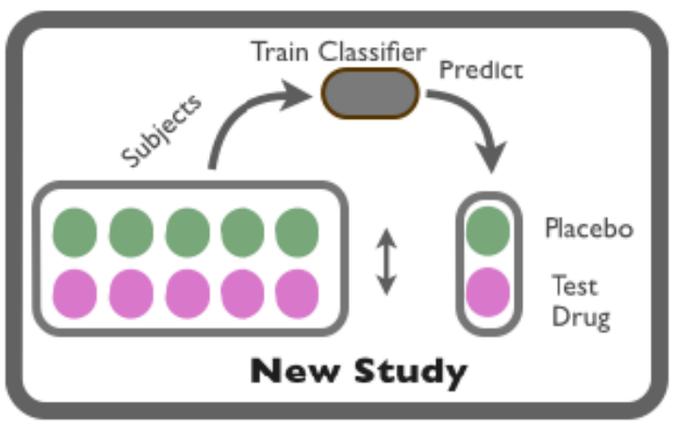
- Test our ability to identify these effects in new studies

FMRI CNS Drug Assessment Protocol



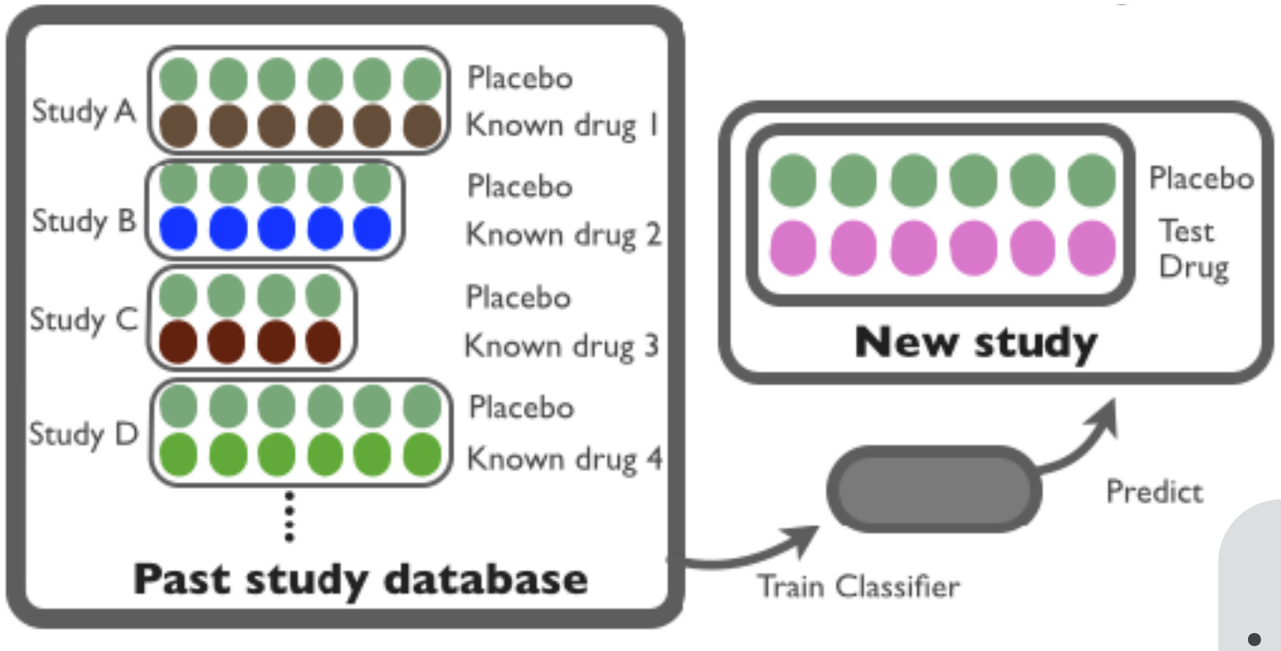
Implementing our decision procedures

Pharmacodynamic effect



Cross validated
discrimination accuracy

Signature of Clinical efficacy



Cross validated
discrimination accuracy

Multi-study signature of efficacy. Train MVPA algorithm to identify brain responses associated with established efficacious drugs, using a set of existing studies studies. Determine whether algorithm successfully identifies the presence of the test compound.

MVPA Analysis

- Whole-brain data, ICA-based features
- Forced-choice linear SVM
- Subject-wise drug vs placebo training and prediction

FMRI CNS Drug Assessment Protocol

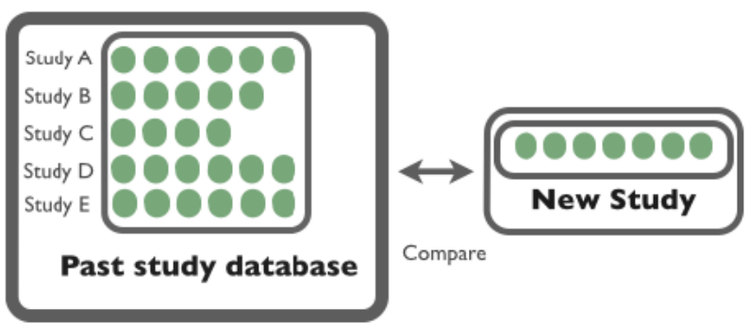
A. Quality assurance

Pause.
Reassess experimental protocol and analysis strategy. Result is not informative of drug effects.

No.
Problem with protocol or analysis pipeline

Can basic responses be detected and do they match past studies?

Yes.
Baseline responses and modeling appear OK



Stimulus validity assessment. Test for differences between responses of test study and responses elicited in validated existing studies.

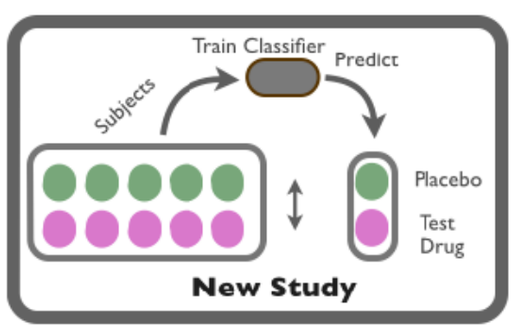
B. Pharmacodynamic effect

Stop.
Negative result. Indicates problem with drug, drug dosage, or the imaging protocol lacks sensitivity.

No.
Little indication that drug has an identifiable effect on responses

Can drug effect be reliably identified based on other subjects in study?

Yes.
Drug has a reliable effect on brain responses.



Cross-validated signature of pharmacodynamic effect. Train MVPA algorithm to discriminate drug from control session maps. Test on held-out subjects (i.e. leave-one-subject-out cross validation).

C. Evidence for clinical efficacy

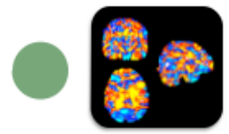
Qualified Go.
Drug has an effect on brain, but does not match past studies. Assess spatial effects, other evidence

Go.
Proceed with clinical trials. Accelerate programme if evidence is particularly strong.

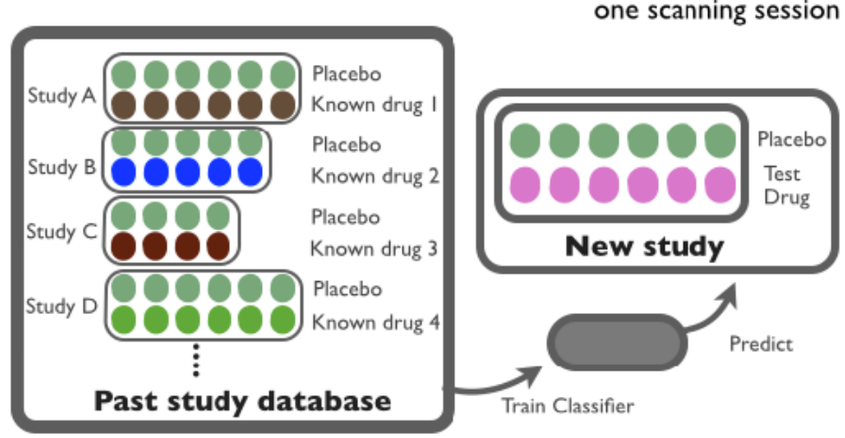
No.
Baseline responses and modeling appear appropriate.

Yes.
Imaging suggests that drug is a promising candidate.

Can the drug modulation be identified based on a validated signature of clinical efficacy?

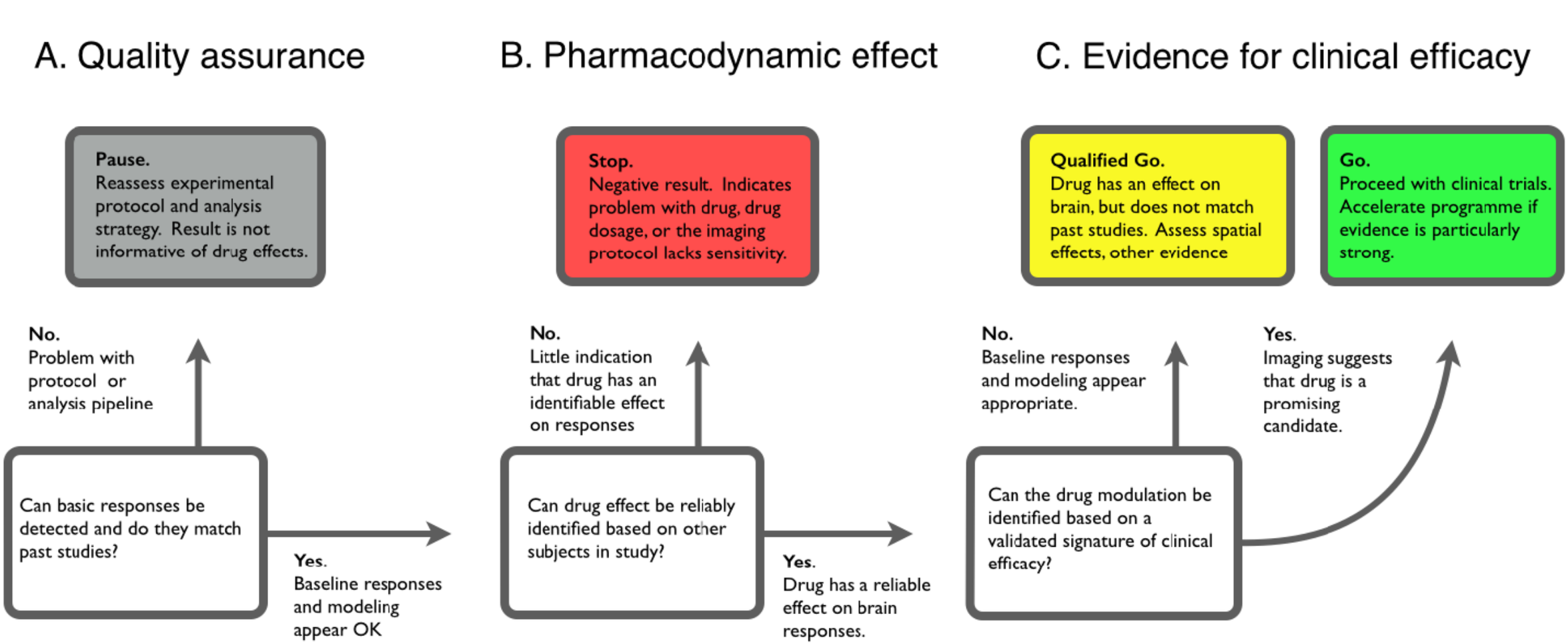


Points represent imaging data from one scanning session



Multi-study signature of efficacy. Train MVPA algorithm to identify brain responses associated with established efficacious drugs, using a set of existing studies studies. Determine whether algorithm successfully identifies the presence of the test compound.

FMRI CNS Drug Assessment Protocol



Go/No-Go decision rules must be predefined, and will be dependent on:

- strength of effects in existing efficacious compounds
- current confidence in the compound
- demand for a successful compound
- expected subsequent cost of development
- expected economics of a successful compound

Analgesic Datasets

Study	Drug (reference/ clinicaltrials.gov ID)	Patient condition	*n* subjects	Scanner	Dose	Stimuli	*n* trials	Pain score
Analgesic drug study assessments								
a	Gabapentin (*9*)	Healthy	12	3T Varian	1800 mg oral (taken 2 hours prior)	Punctate to hyperalgesic skin	20	N/R
b	Pregabalin (*24*)	Fibromyalgia	23	3T GE	225 mg/day oral (7 days daily dosing)	Thumb squeeze	6	Yes
c	Pregabalin (*25*) (NCT00610155)	PTNP	16	3T TIM Trio	150 mg oral BID (7 days daily dosing)	Brush-evoked allodynia	15	No
d	Tramadol (*25*) (NCT00610155)	PTNP	16	3T TIM Trio	200 mg oral BID (7 days daily dosing)	Brush-evoked allodynia	15	Yes
e	Remifentanil (*6*, *18*)	Healthy	22	3T Varian	2 ng/ml BPC i.v.	Punctate and thermal	10	Yes
f	Remifentanil (*21*)	Healthy	12	3T Varian	1.5 ng/ml BPC i.v.	Laser	50	No
g	THC (*23*)	Healthy	14	3T Varian	15 mg oral (taken 2 hours prior)	Punctate to hyperalgesic skin	20	Yes
h	Naproxen (*22*)	Osteoarthritis	19	3T GE HDx	220 mg oral (taken 1 hour prior)	Key turn	15	Yes

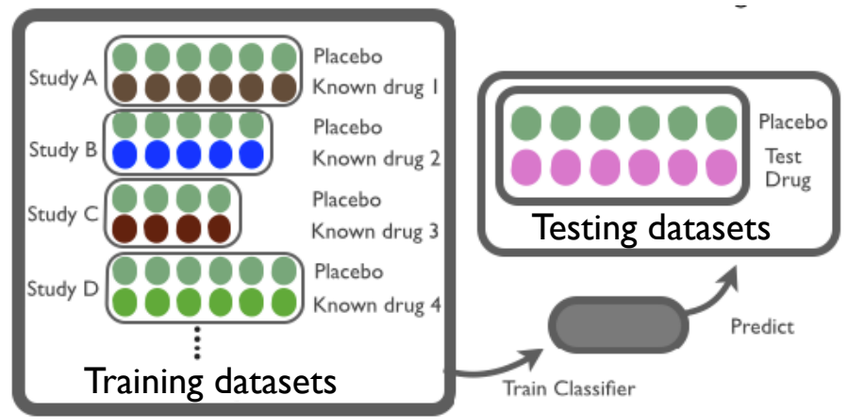
Control Datasets

Study	Drug (reference/ clinicaltrials.gov ID)	Patient condition	*n* subjects	Scanner	Dose	Stimuli	*n* trials	Pain score
Control study assessments								
i	2nd placebo (study b)	Fibromyalgia	23	3T GE	N/A	Thumb squeeze	6	No
j	2nd placebo (study e)	Healthy	22	3T Varian	N/A	Punctate and thermal	10	No
k	2nd placebo (study f)	Healthy	12	3T Varian	N/A	Laser	50	No
l	Remifentanil (study f)	Healthy	12	3T Varian	1.5 ng/ml BPC infusion	Flash	50	N/A
m	Remifentanil (study f)	Healthy	12	3T GE	1.5 ng/ml BPC infusion	Brief tone	50	N/A
n	Naproxen (study h)	Osteoarthritis	19	3T GE HDx	220 mg (1 hour)	Visual stimulus	15	N/A

Learning to identify CNS drug action and efficacy using multistudy fMRI data

Event related studies (Analgesics)

- Pregb (PTNP)
- Tram (PTNP)
- Pregb (Fibro)
- Naprox (OA)
- Remi (L)
- Remi (T/P)
- Gabap
- THC



Decision rules

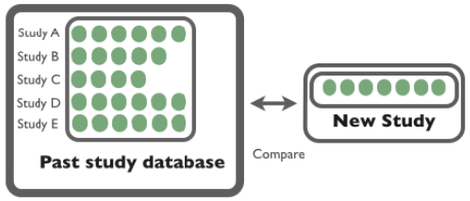
(QC) Greater than 5% of pain regions showing lower responses than database

Significant prediction accuracy

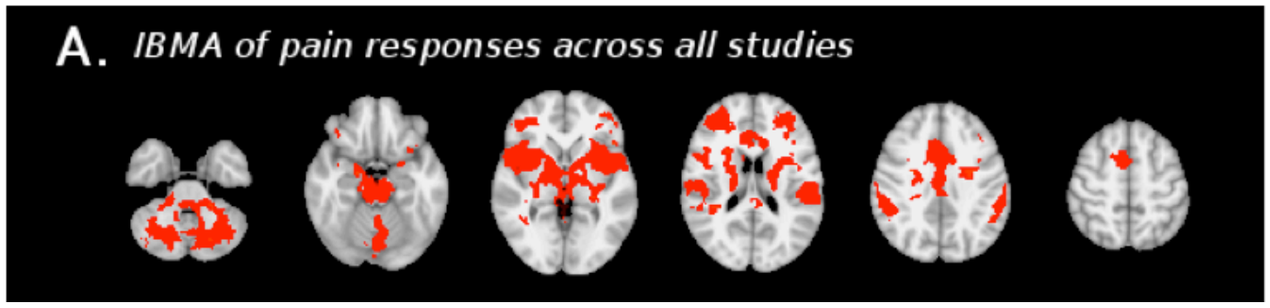
Non-significant drug placebo discrimination ($p<0.15$)

Non-significant prediction accuracy

Results



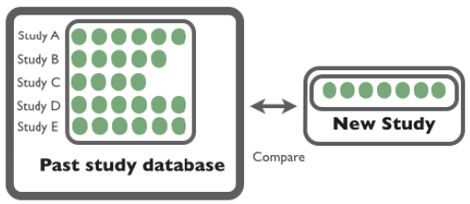
Study	Test compound	QA Area of reduced response (%)
Analgesic drug study assessments		
a	Gabapentin	0
b	Pregabalin	0
c	Pregabalin	0
d	Tramadol	0
e	Remifentanil	0
f	Remifentanil	0
g	THC	0
h	Naproxen	0
Control study assessments		
i	Placebo (study b)	0
j	Placebo (study e)	0
k	Baseline (study f)	0
l	Visual stimulus (study f)	79
m	Auditory stimulus (study f)	6
n	Visual stimulus (study h)	48

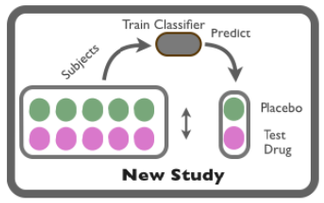


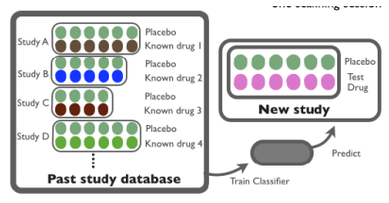
Results

Study	Test compound	QA	Pharmacodynamic effect	
		Area of reduced response (%)	Accuracy (range)	*P*
Analgesic drug study assessments				
a	Gabapentin	0	92% (70–100)	0.0002
b	Pregabalin	0	70% (52–83)	0.017
c	Pregabalin	0	81% (61–92)	0.002
d	Tramadol	0	56% (36–74)	0.22
e	Remifentanil	0	86% (70–94)	0.000
f	Remifentanil	0	83% (60–94)	0.003
g	THC	0	71% (49–86)	0.028
h	Naproxen	0	73% (55–83)	0.01
	Control study assessments			
i	Placebo (study b)	0	52% (22–54)	0.339
j	Placebo (study e)	0	27% (15–45)	0.97
k	Baseline (study f)	0	58% (36–78)	0.194
l	Visual stimulus (study f)	79	74% (51–90)	0.019
m	Auditory stimulus (study f)	6	100% (82–100)	0.000
n	Visual stimulus (study h)	48	73% (55–87)	0.009

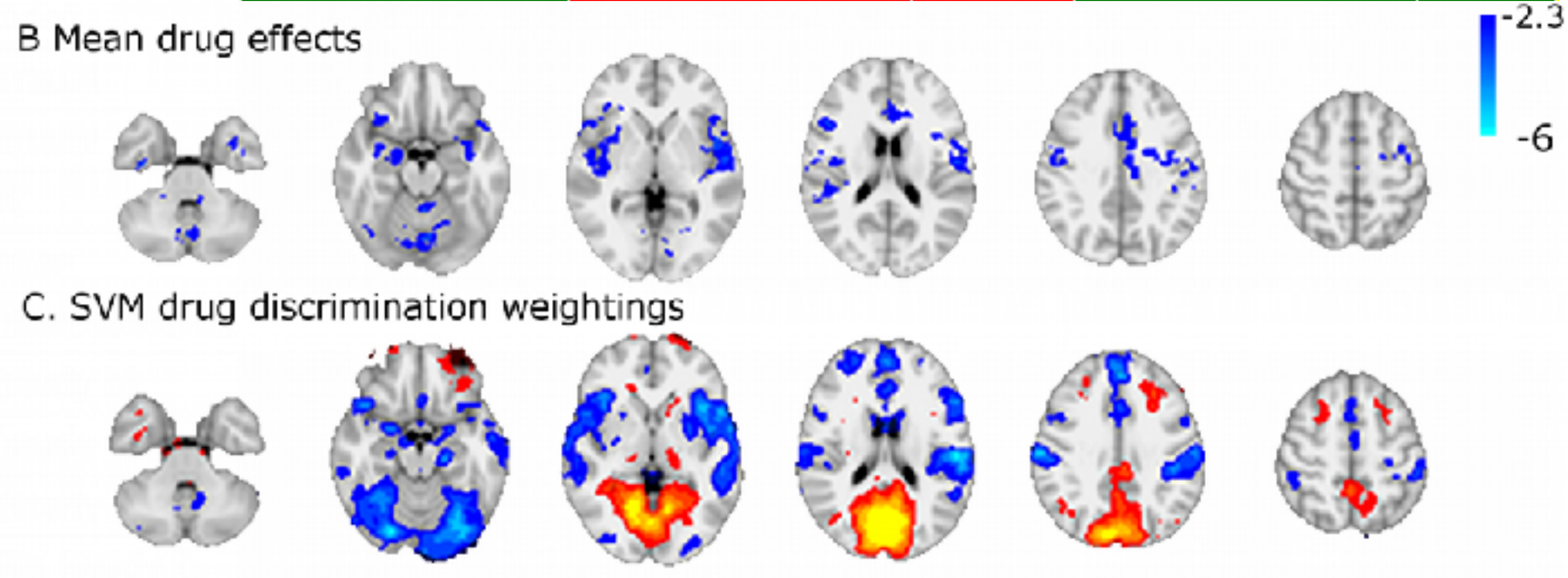
Results



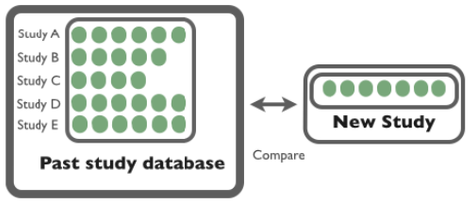


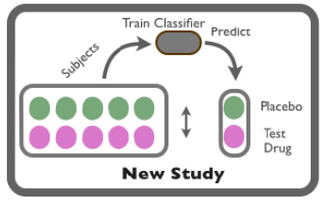


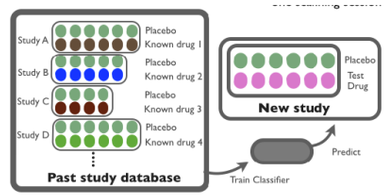
Study	Test compound	QA	Pharmacodynamic effect		Clinical efficacy	
		Area of reduced response (%)	Accuracy (range)	*P*	Accuracy (range)	*P*
Analgesic drug study assessments						
a	Gabapentin	0	92% (70–100)	0.0002	83% (60–94)	0.003
b	Pregabalin	0	70% (52–83)	0.017	61% (44–76)	0.105
c	Pregabalin	0	81% (61–92)	0.002	69% (42–79)	0.038
d	Tran					
e	Remif					
f	Remif					
g	Tl					
h	Napr					
	Control study					
i	Placebo					
j	Placebo					
k	Baseline					
l	Visual stimu					
m	Auditory stimulus (study f)	6	100% (82–100)	0.000	58% (36–78)	0.194
n	Visual stimulus (study h)	48	73% (55–87)	0.009	47% (30–65)	0.5



Results



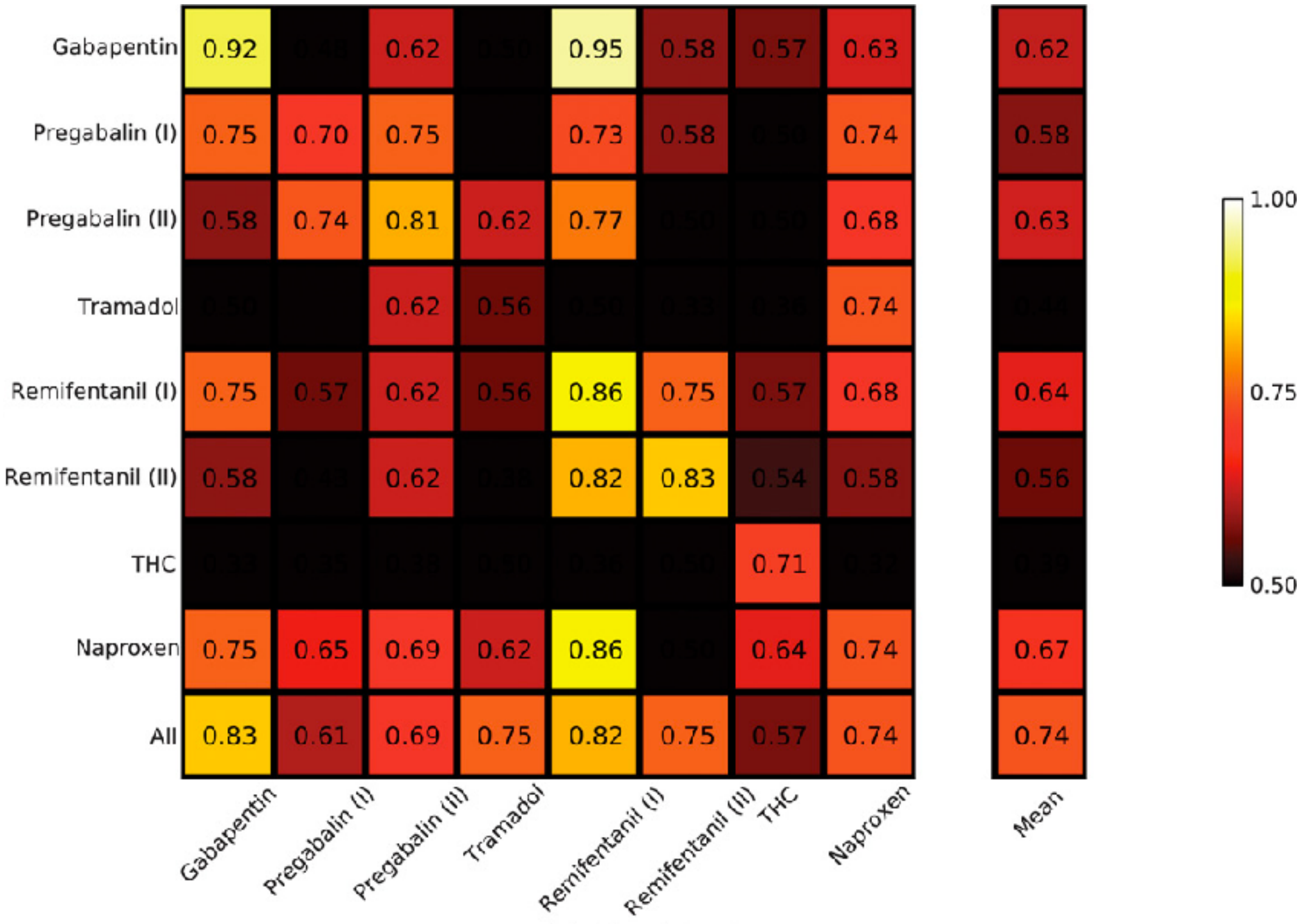




Study	Test compound	QA	Pharmacodynamic effect		Clinical efficacy		Decision
		Area of reduced response (%)	Accuracy (range)	*P*	Accuracy (range)	*P*	
Analgesic drug study assessments							
a	Gabapentin	0	92% (70–100)	0.0002	83% (60–94)	0.003	Go
b	Pregabalin	0	70% (52–83)	0.017	61% (44–76)	0.105	Go
c	Pregabalin	0	81% (61–92)	0.002	69% (42–79)	0.038	Go
d	Tramadol	0	56% (36–74)	0.22	75% (55–74)	0.01	Go (Q)
e	Remifentanil	0	86% (70–94)	0.000	82% (65–92)	0.000	Go
f	Remifentanil	0	83% (60–94)	0.003	75% (51–90)	0.003	Go
g	THC	0	71% (49–86)	0.028	57% (36–76)	0.22	Go (Q)
h	Naproxen	0	73% (55–83)	0.01	73% (55–83)	0.01	Go
Control study assessments							
i	Placebo (study b)	0	52% (22–54)	0.339	48% (32–64)	0.5	Stop
j	Placebo (study e)	0	27% (15–45)	0.97	45% (29–62)	0.584	Stop
k	Baseline (study f)	0	58% (36–78)	0.194	33% (16–57)	0.806	Stop
l	Visual stimulus (study f)	79	74% (51–90)	0.019	33% (16–57)	0.019	Reassess
m	Auditory stimulus (study f)	6	100% (82–100)	0.000	58% (36–78)	0.194	Reassess
n	Visual stimulus (study h)	48	73% (55–87)	0.009	47% (30–65)	0.5	Reassess

Duff et al, Learning to identify CNS drug action and efficacy using multistudy fMRI data. STM 2015

Between study prediction

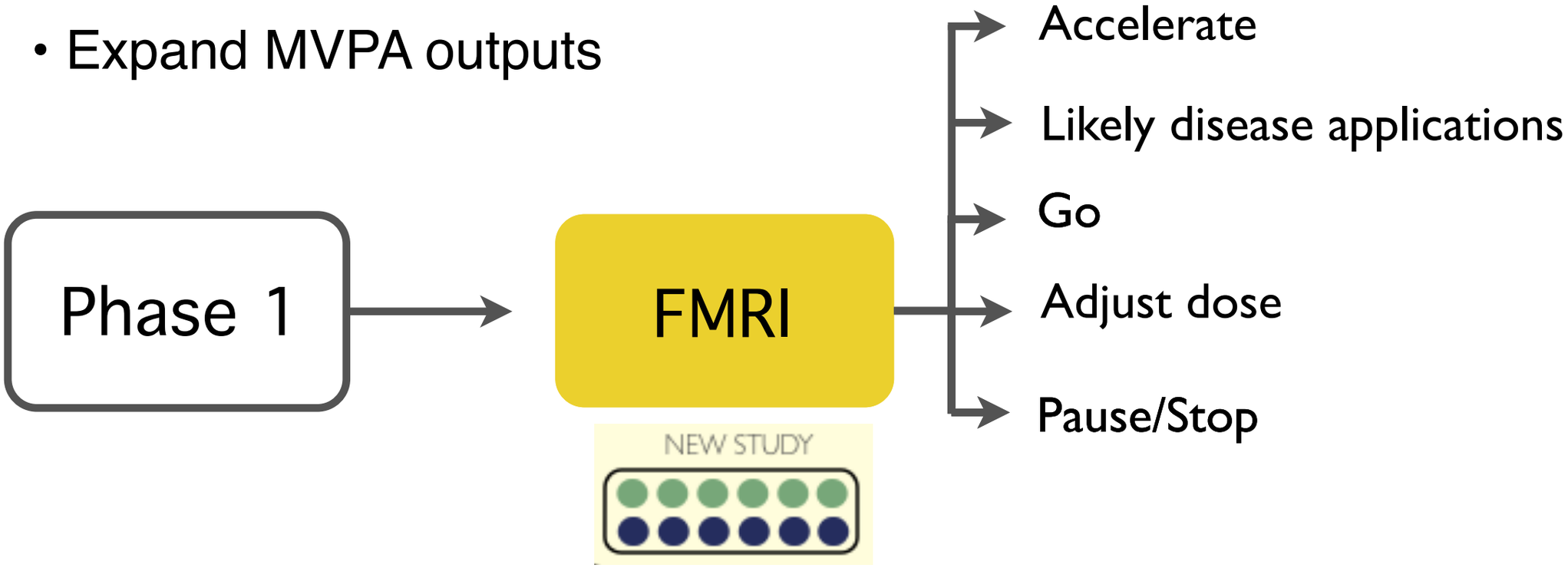


Conclusions

- Proof-of-concept that FMRI can play a role in clinical trials, and similar predictive clinical applications.
- It is possible to quantitatively integrate existing data into new studies, increasing the robustness and range of inferences.
- Standardised but evolving protocols can be important for clinical imaging and many other applications.
- Obtaining (and publishing) datasets to combine remains very challenging, particuarly commercial/clinical data.
- Immediate focus should be on protocols combining datasets within laboratories.
- Machine learning approaches are an effective way to boost sensitivity in these contexts

Refinements to prediction

- Expand MVPA outputs

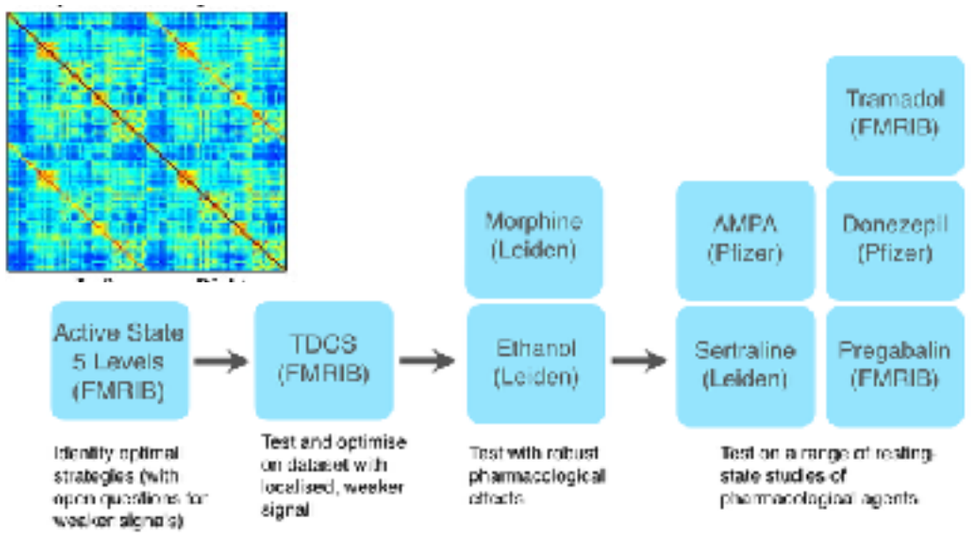


- Alternate approaches to generating signatures of efficacy that do not require existing efficacious compounds

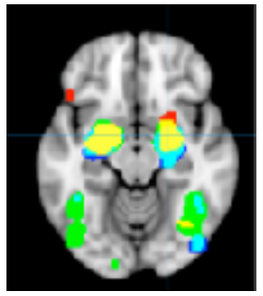
- Resting state

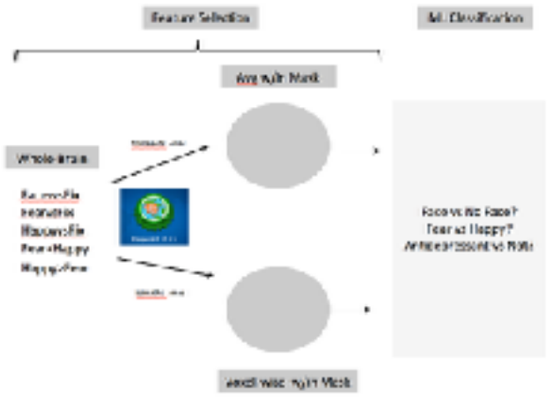
Further and ongoing work

Resting-state pharmacologic studies

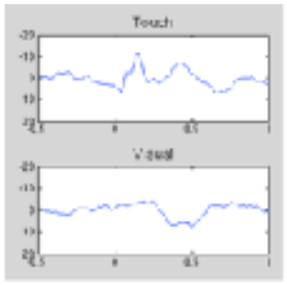


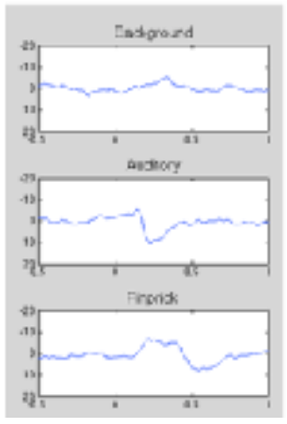
Biomarker for Antidepressant Action



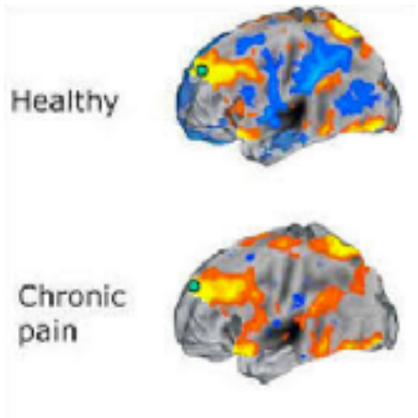


Pain Biomarker for Newborn babies (FMRI & EEG)

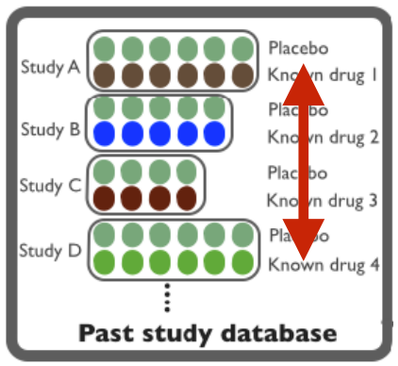




Biomarker for Chronic Pain



Dataset Harmonisation



Acknowledgements

Analysis group

Steve Smith

Mark Woolrich

Pain group

Irene Tracey

Karolina Wartolowska

Vishvarani Wanigasekera

Michael Lee (Cambridge)

Pfizer

William Vennart

Fred Wilson

Mark Whitlock

Cardiff University

Richard Wise

Kings College London

Matthew Howard

University of Michigan

Richard Harris

University of Leiden

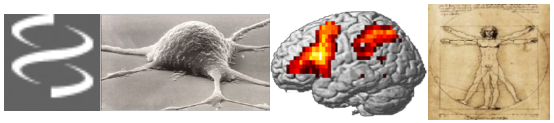
Serge Rombouts

Bernadet Klassens

University of Montreal

Najmeh Khalili-Mahani

Psychiatric Neuroscience Group



FONDAZIONE CON IL SUD

Translating molecular profiling into neuroimaging phenotypes

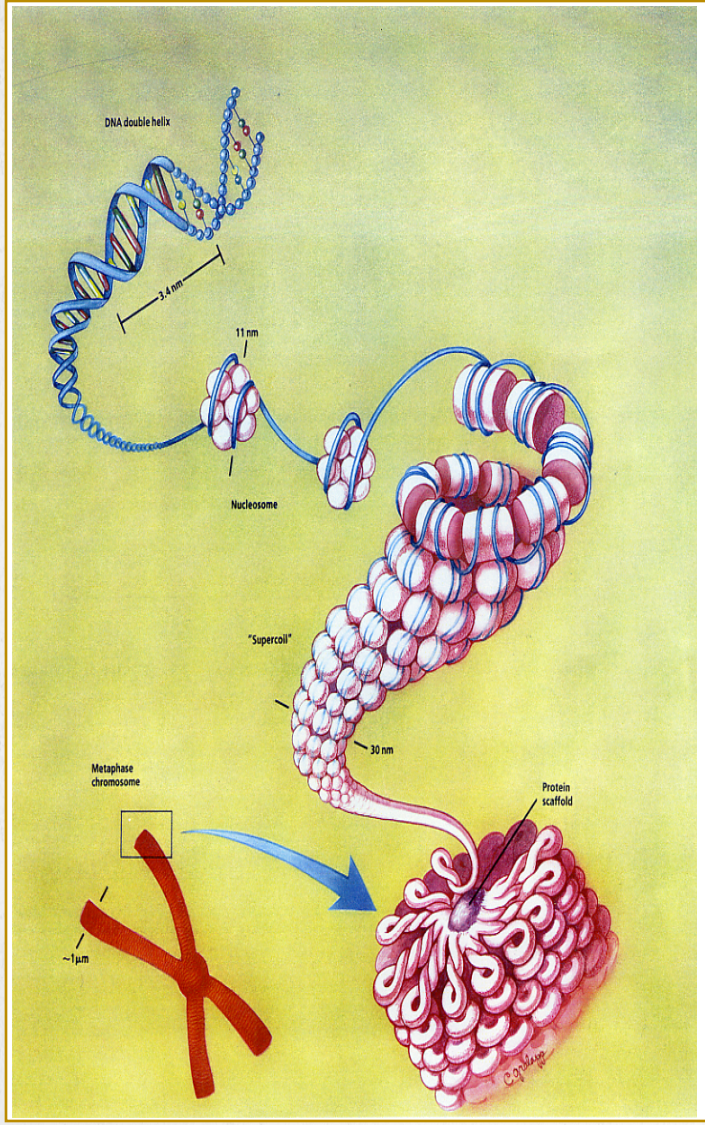
Giulio Pergola, PhD

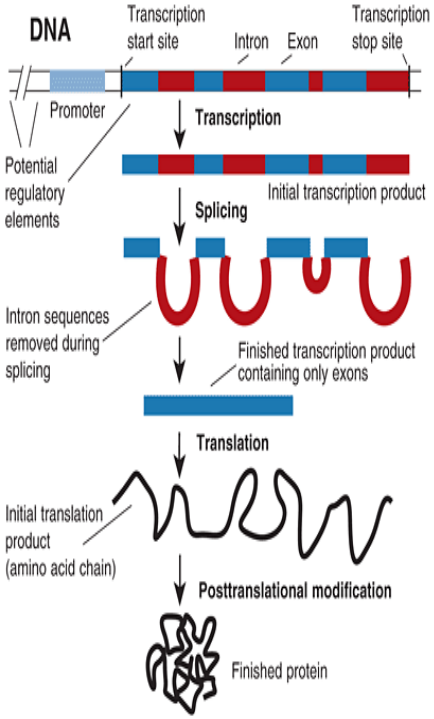
Brain Imaging, Networks and Data mining Lab

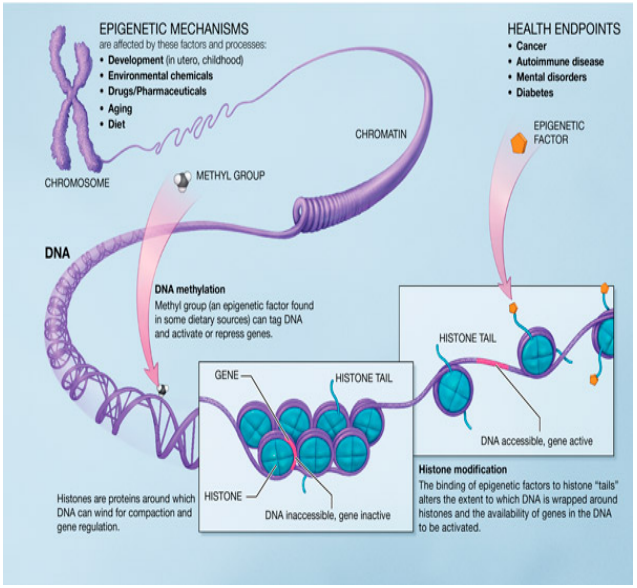
University of Bari Aldo Moro, Bari, Italy

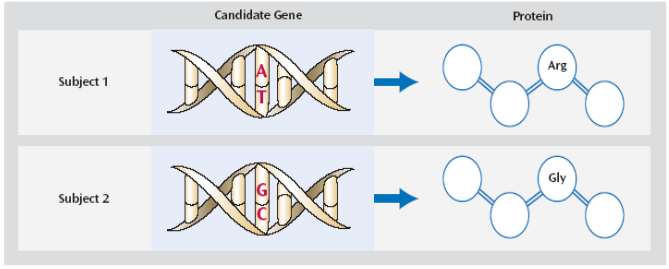
Giulio.Pergola@uniba.it

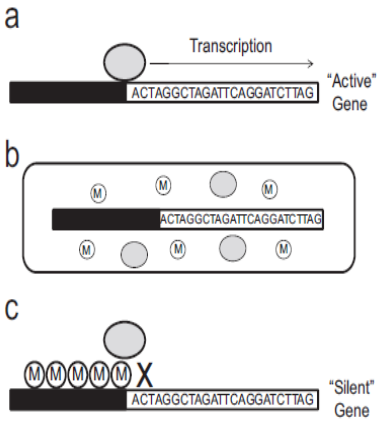
Primer – Genetic variation

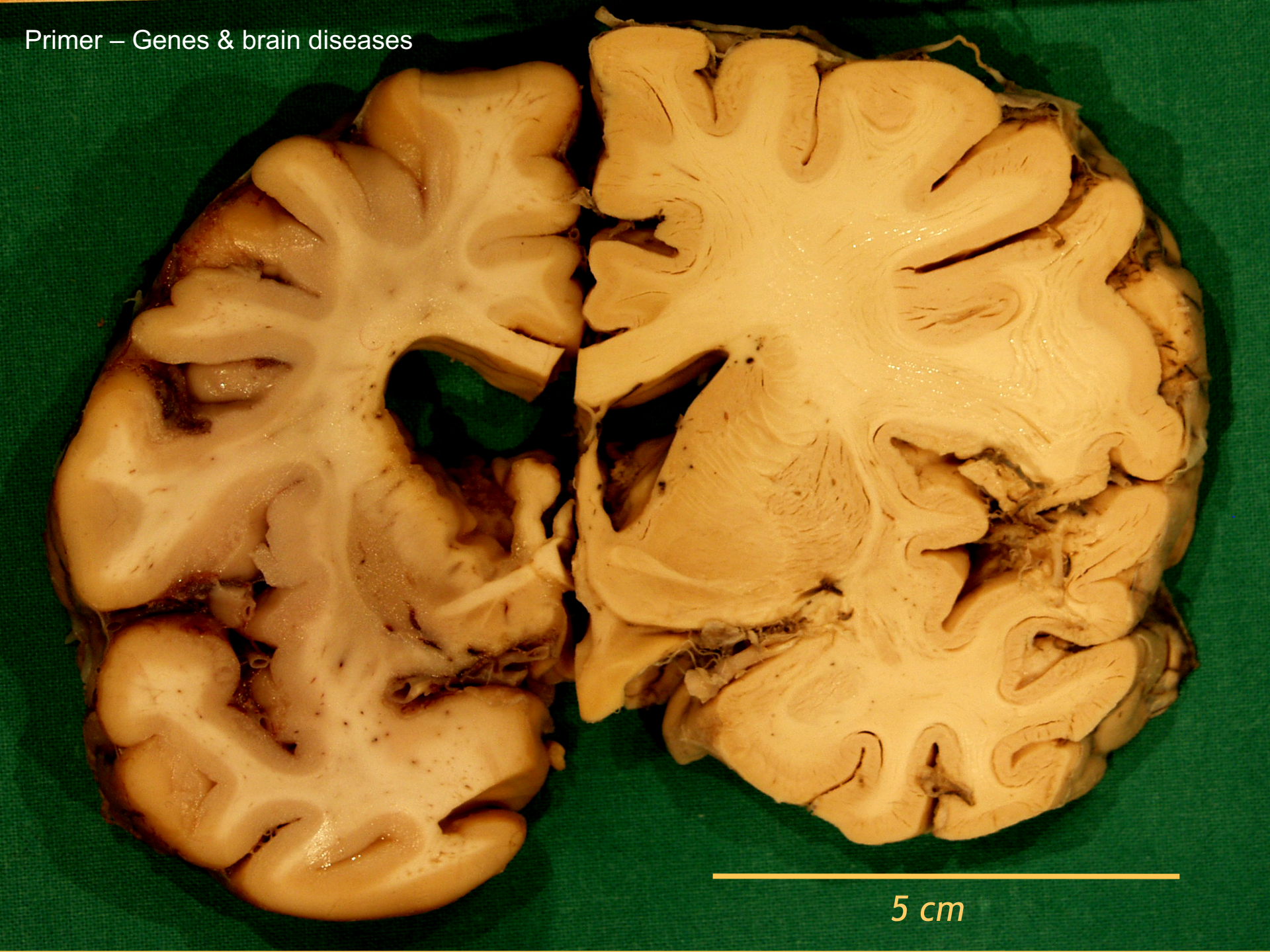




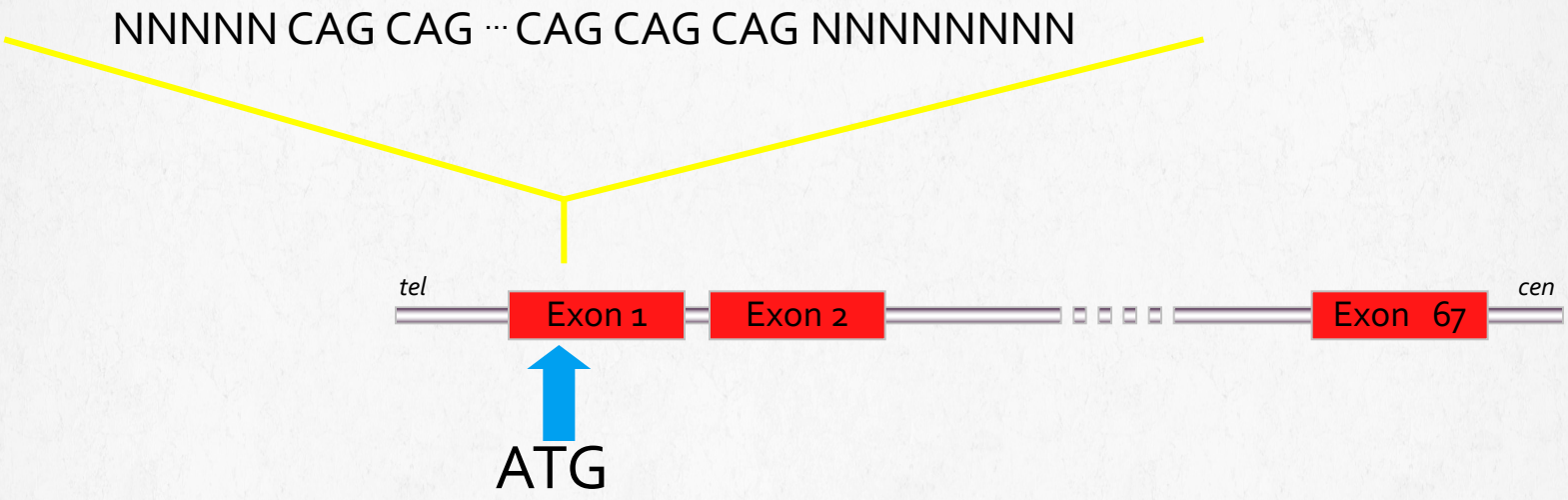







5 cm

Huntingtin gene

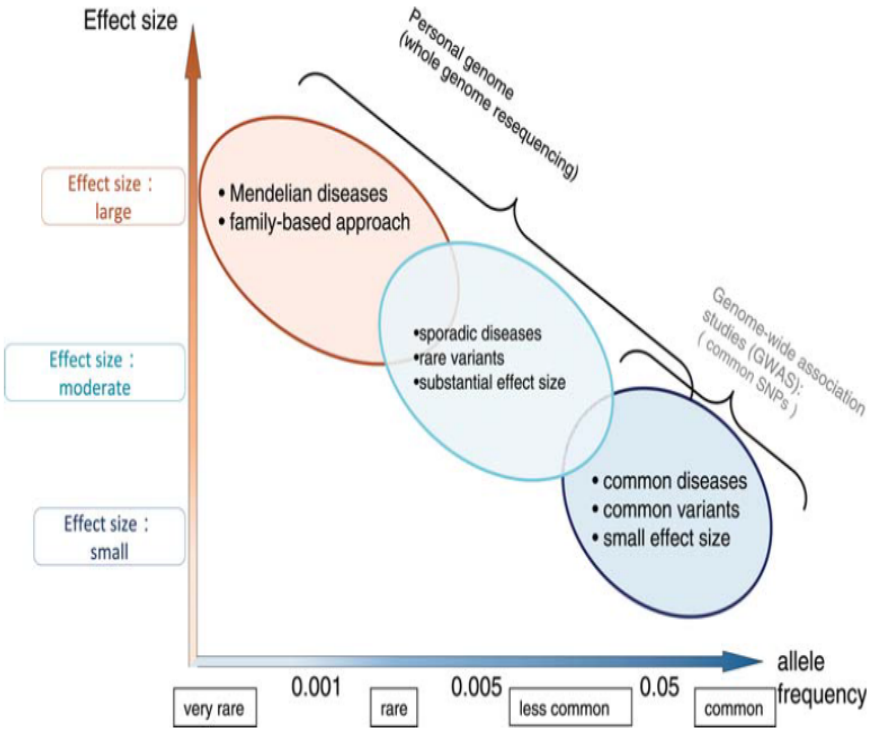


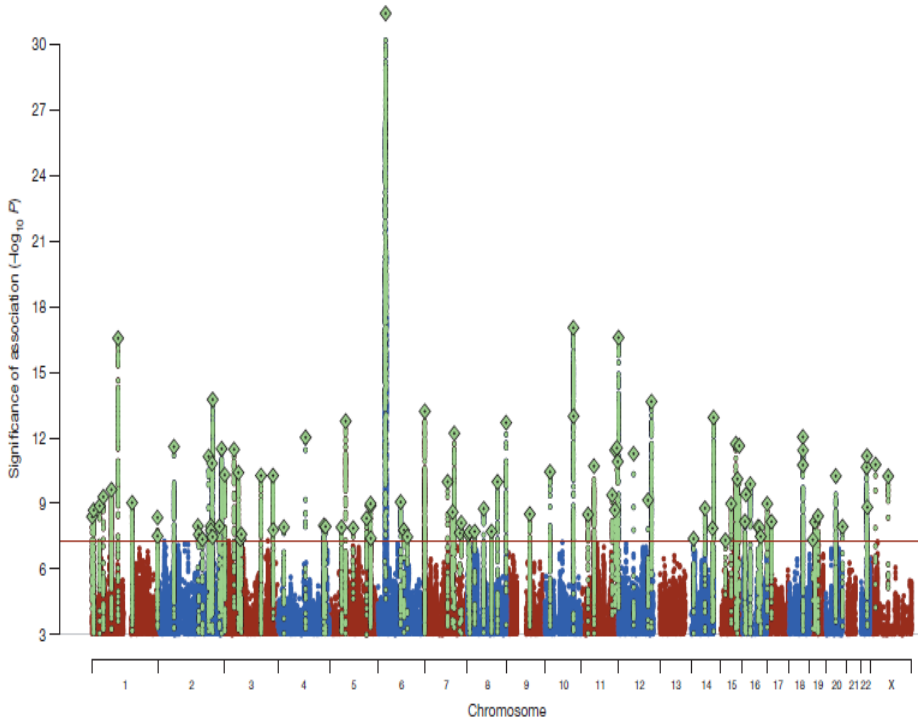
$(CAG)_n$ translation $(Gln)_n$

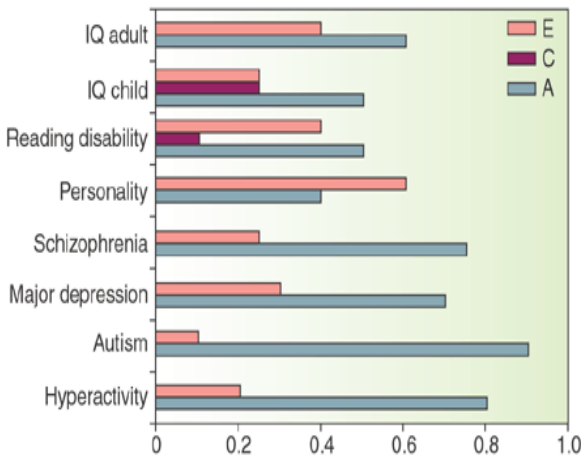
‚normal‘

‚pathogenic‘

Primer – Common genetic variation and the Psychiatric Genomic Consortium







Heritability is the proportion of phenotypic variation associated with genetic variation (as opposed to shared and non-shared environmental variation). Heritability of schizophrenia is estimated around 80% (McGuffin, Riley et al., 2001).

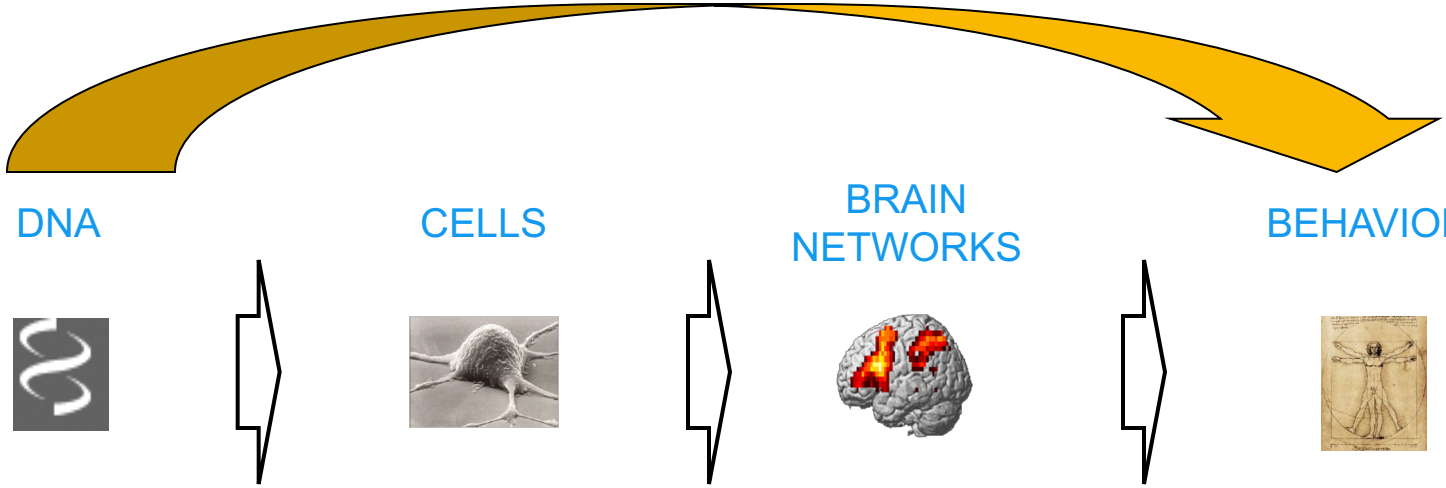
No genetic variation so far investigated explains more than a minimal proportion of risk for schizophrenia (Hirsch and Weinberger, 2003).

How can we study the influence of multiple genetic variants on the biology of schizophrenia?

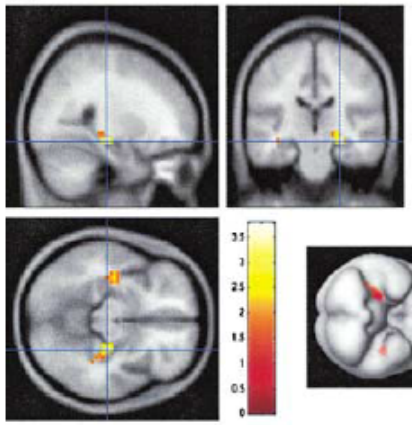
INTERMEDIATE PHENOTYPES

Any neurobiological measures associated with the genetic risk variants of a disease, e.g., biochemical, neurophysiological, neuroanatomical, neuropsychological (Bearden, 2006).

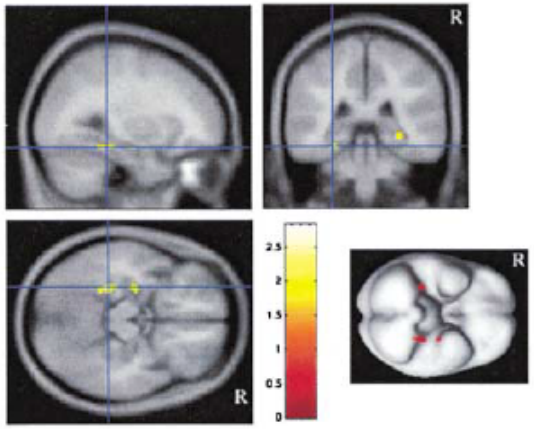
Intermediate phenotypes are heritable but independent of illness status, therefore can also be studied in healthy siblings of patients.

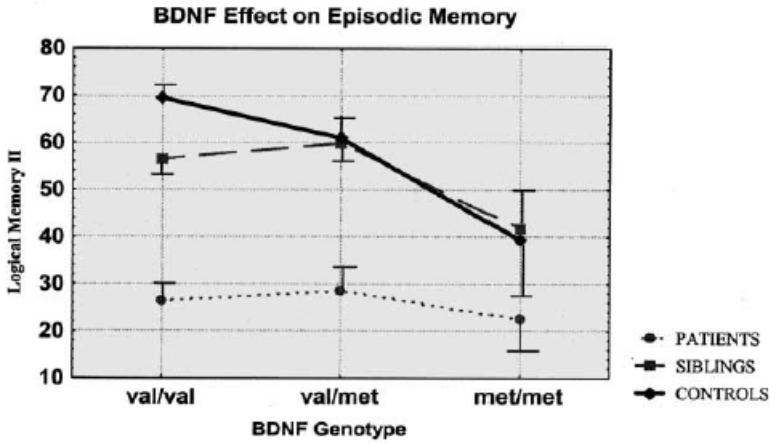


A



B





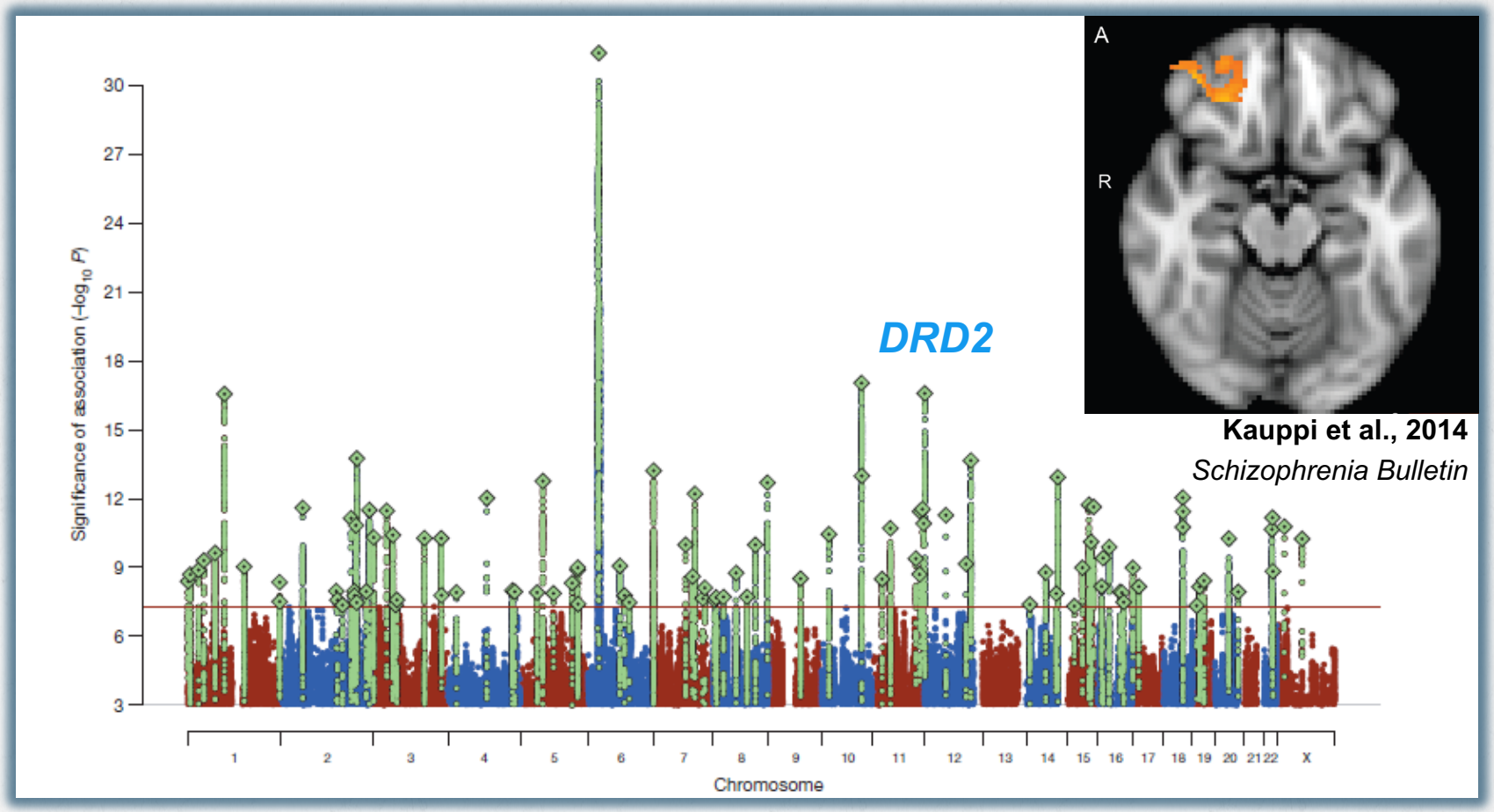
C

	Cohort 1		Cohort 2	
Genotype Group (number of subjects)	Val/val (8)	Val/met (5)	Val/val (12)	Val/met (5)
Age	38.8 (10.5)	38.8 (10.0)	28.8 (7.1)	32.8 (9.9)
Gender M/F	5/3	5/0	7/5	3/2
Education years	16.6 (3.7)	16.5 (1.0)	15.8 (2.1)	17 (6.3)
Reading Comprehension	109.5 (6.0)	109.8 (4.3)	107.9 (5.3)	114.6 (6.9)
IQ	112.8 (5.4)	109 (7.1)	100.2 (13.2)[a]	115.6 (6.2)[a]
Episodic Memory (Logical Memory II scores)	73.9 (18.4)	60.6 (25.0)	45.8 (26.3)[b]	68.4 (29.5)[b]
2 back performance				
%Correct responses	73.5 (28.0)	80.2 (17.0)	87.4 (13.4)	91.1 (13.1)
Reaction time millisec.	263.9 (103.6)	175.1(80.7)	100.7 (28.3)	103.5 (30.8)

- BDNF;
- Single exonic SNP;
- Note the sample sizes
- ***Clear molecular function, but limited clinical translation.***

Egan et al., 2003 *Cell*

Background – many players in schizophrenia risk



Kauppi et al., 2014
Schizophrenia Bulletin

Ripke et al., 2014 *Nature*

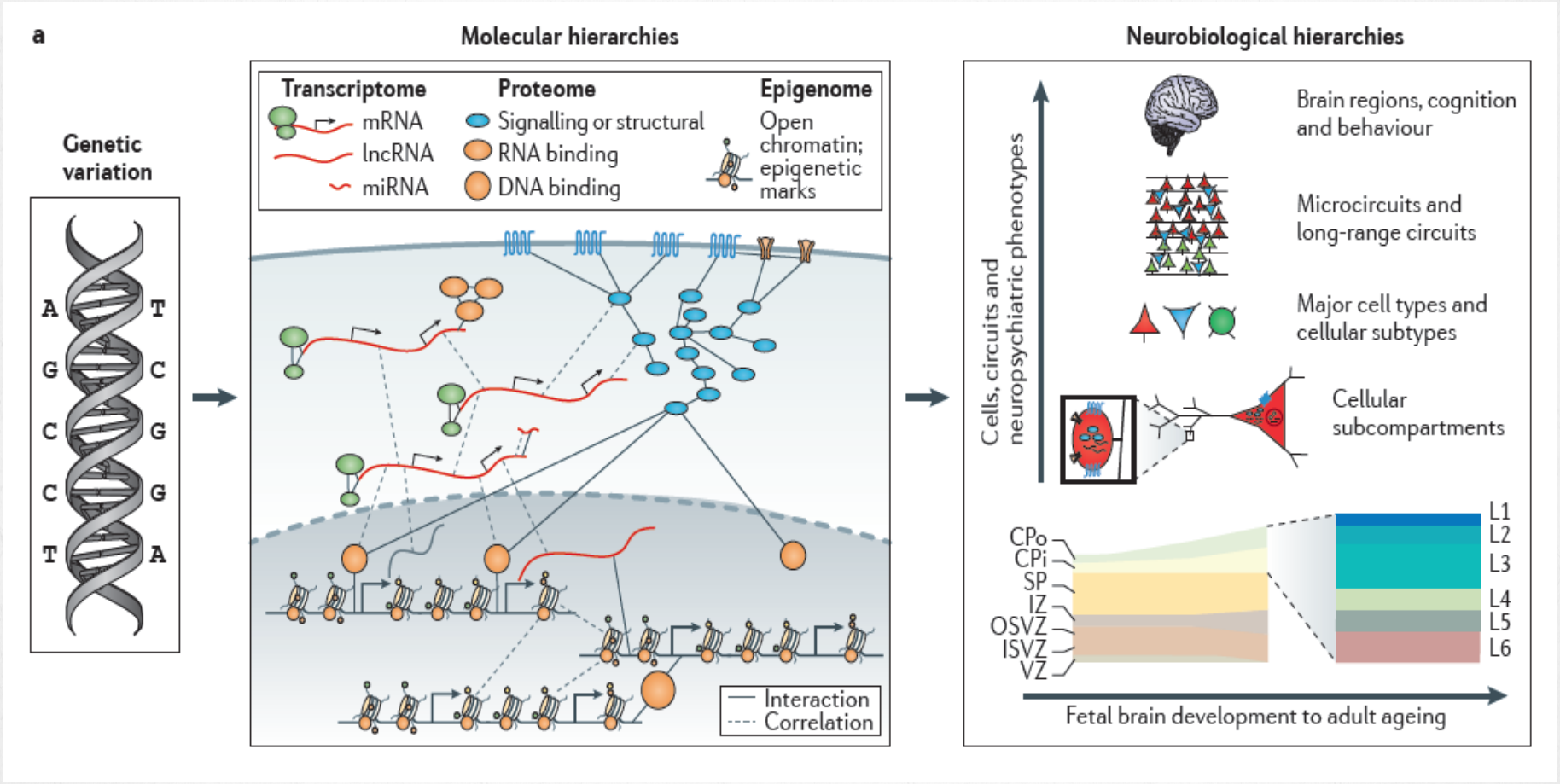
‘Guilty alleles’ or ‘guilty genes’?

A standard way to look at gene-trait associations is to search for risk alleles. Then, risk alleles are combined into ensembles reflecting the additive genetic component of a heritable trait.

We still don’t know which genes are impacted by these variants – but we know the polygenic risk score is not related with clinical outcome (Hettige et al., 2016 *Schizophrenia Research*; **Wimberley et al., 2017** *Schizophrenia Bulletin*).

Background – dealing with complexity

The central nervous system is hierarchically organized, and we can use this aspect of its organization to understand how it works.



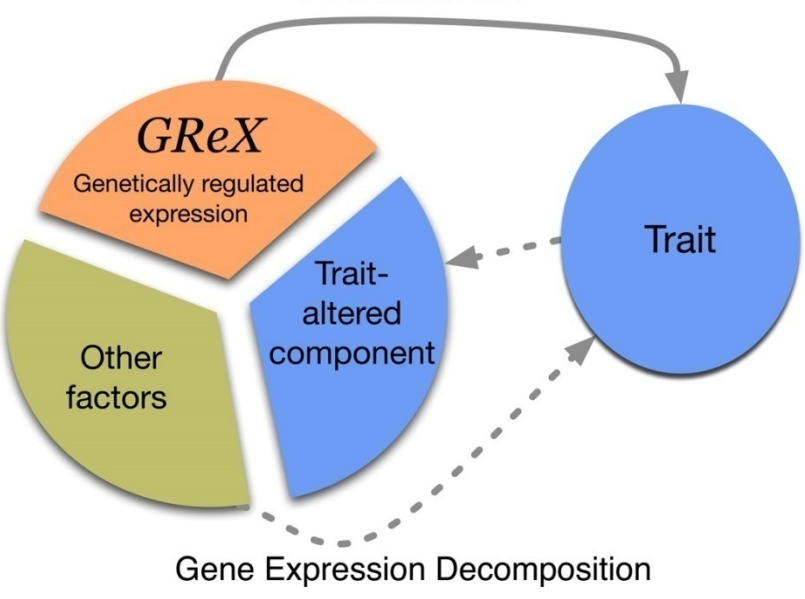
Parikshak and Geschwind, 2014 *Nature Rev Genetics*

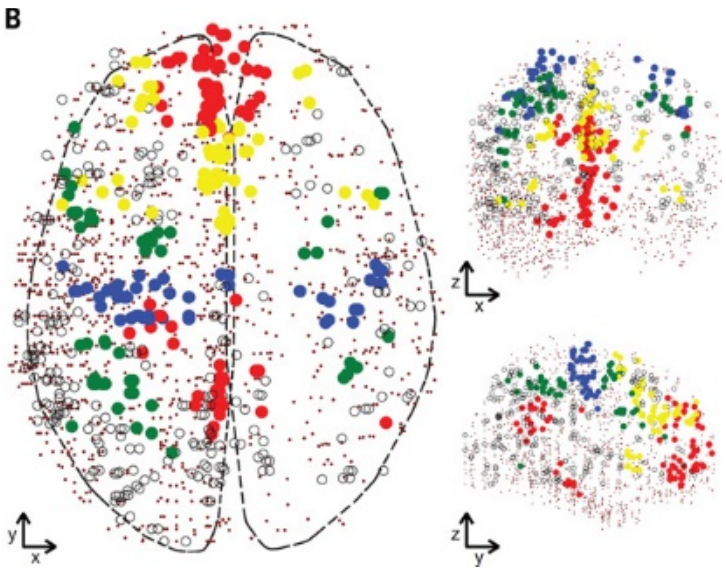
Background – gene expression → *function*

Idea

Gene expression may be considered as a molecular trait and can thus be predicted based on genetic variants.

In turn, gene expression is key to explain brain phenotypes.





Mechanisms

- **Genetic variants may affect gene expression by modulating the affinity of DNA binding factors.**
- **We can identify key genetic variants via brain post mortem studies.**

Gamazon et al., 2015 *Nature Genetics*

Richiardi et al., 2015 *Science*

Outline – functional translation of gene co-expression

Approach

Schizophrenia is associated with genetic factors, and many risk loci have been identified. Since the expression of individual genes risk is co-regulated and results in the co-expression of gene sets, we hypothesized that the transcription context of schizophrenia risk genes may be associated with schizophrenia phenotypes.

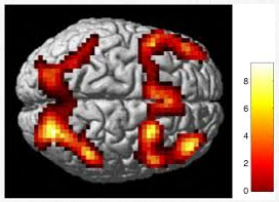
1 + 2 + 3

Translational genetics: *DRD2*

How do we translate genetic risk variants into biologically plausible mechanisms of risk and clinical translation?

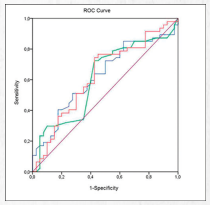
Insights from molecular profiling

Co-expression networks afford novel insights into how *DRD2*-related genes are associated with drug response.



Outlook: schizophrenia genes

Ongoing work on clinical predictions merging imaging & genetic data.

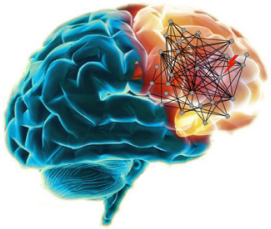


Genetic variants translate gene expression *into function*

Translational genetics: *DRD2* 1 → Insights from molecular profiling 2 → Outlook: Schizophrenia genes 3

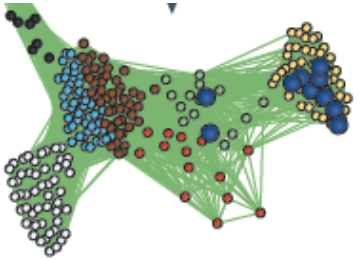
Translate knowledge derived from dead brains into models of living brains.
→ we need a *common language* available both *post mortem* and *in vivo*.

TRANSCRIPTOME



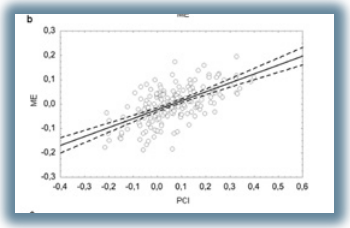
Preprocessing *post-mortem* gene expression data in the DLPFC of non-psychiatric individuals (Braincloud).

CANDIDATE NETWORK



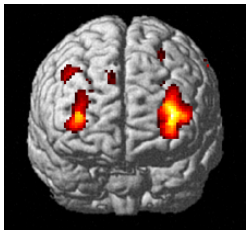
Identifying a genome-wide unsupervised network of genes to model gene-gene relationships.
Then, we can select gene sets of interest.

POLYGENIC CO-EXPRESSION INDEX



Co-eQTLs are detected and collapsed into continuous indices that approximate co-expression.

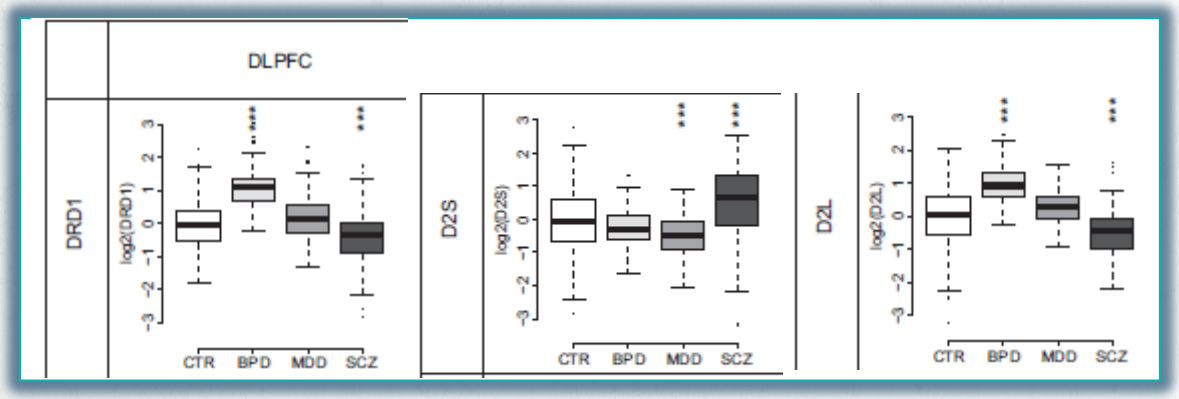
INTERMEDIATE PHENOTYPE



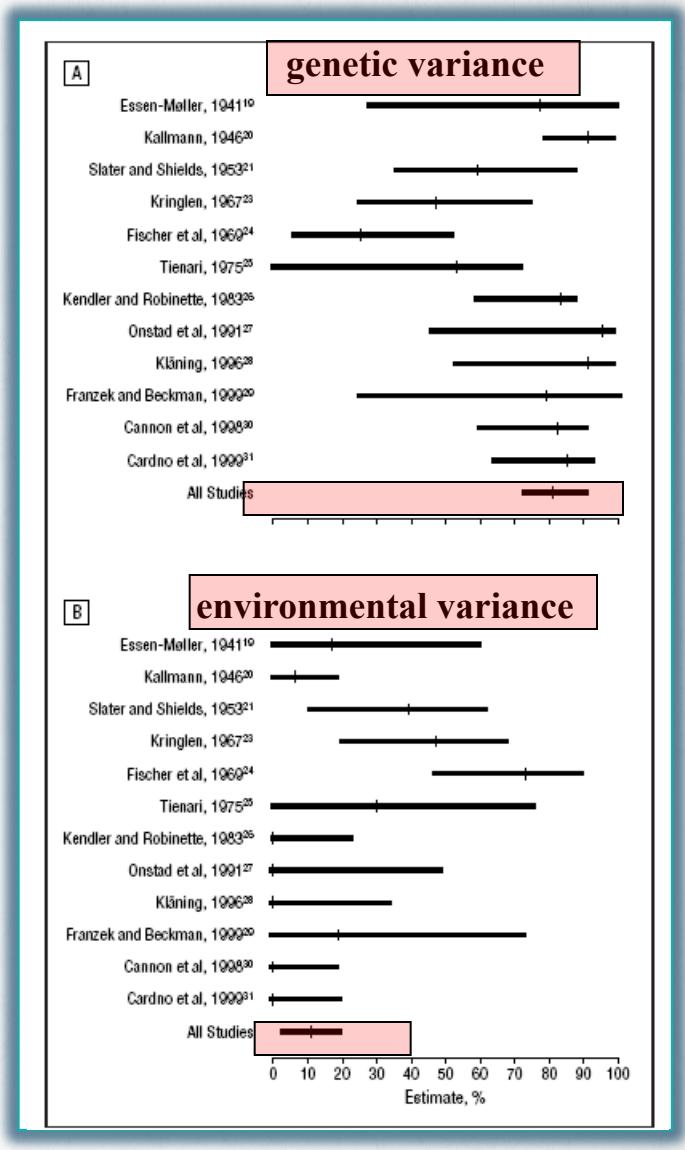
Polygenic co-expression indices are associated with intermediate phenotypes.

Gene of interest – schizophrenia, *DRD2*, and WM

- Twin and adoption studies established the strong heritability of schizophrenia;
- Patients with schizophrenia show altered dopaminergic neurotransmission and benefit from antipsychotics which target the D2 dopamine receptor;



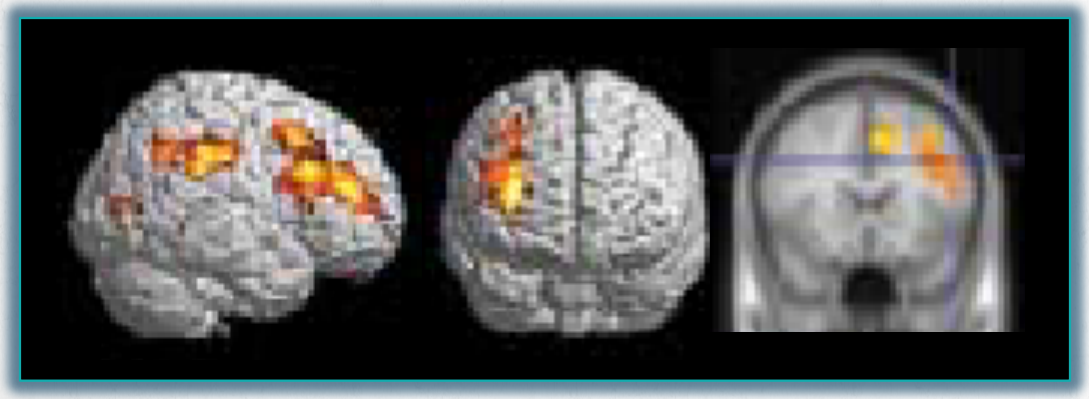
Kaalund et al., 2013 *Molecular Psychiatry*



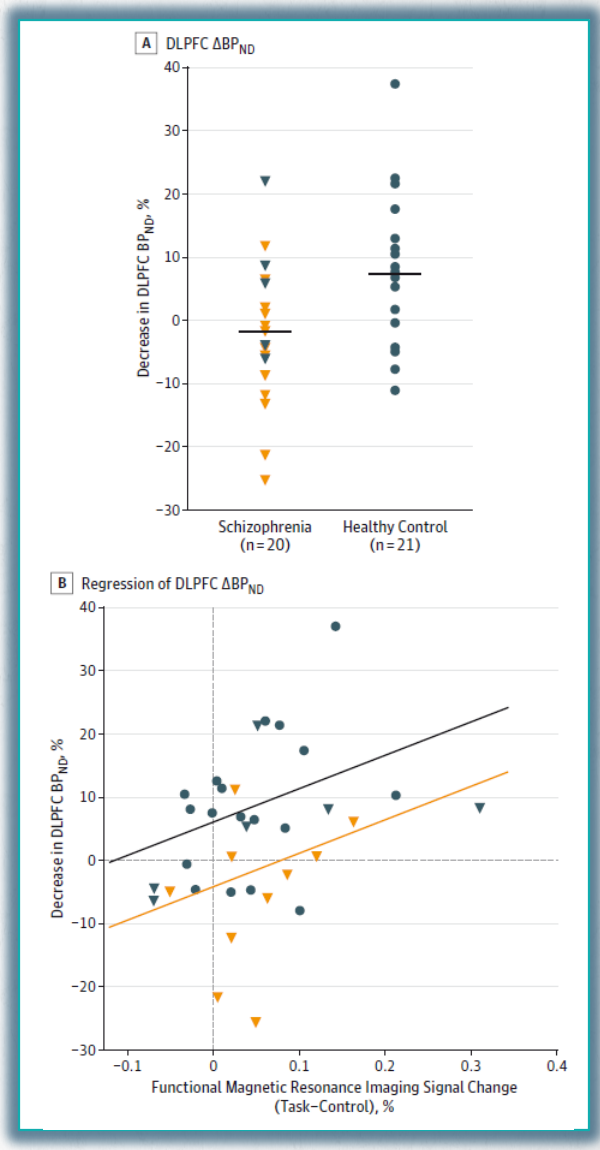
Sullivan et al., 2003 *Arch. Gen. Psychiatry*

Gene of interest – schizophrenia, *DRD2*, and WM

- Twin and adoption studies established the strong heritability of schizophrenia;
- Patients with schizophrenia show altered dopaminergic neurotransmission and benefit from antipsychotics which target the D2 dopamine receptor;
- Schizophrenia-specific functional brain alterations are related with D2 binding in the DLPFC;
- Similar alterations are found in unaffected siblings → genetic basis of schizophrenia brain phenotypes?

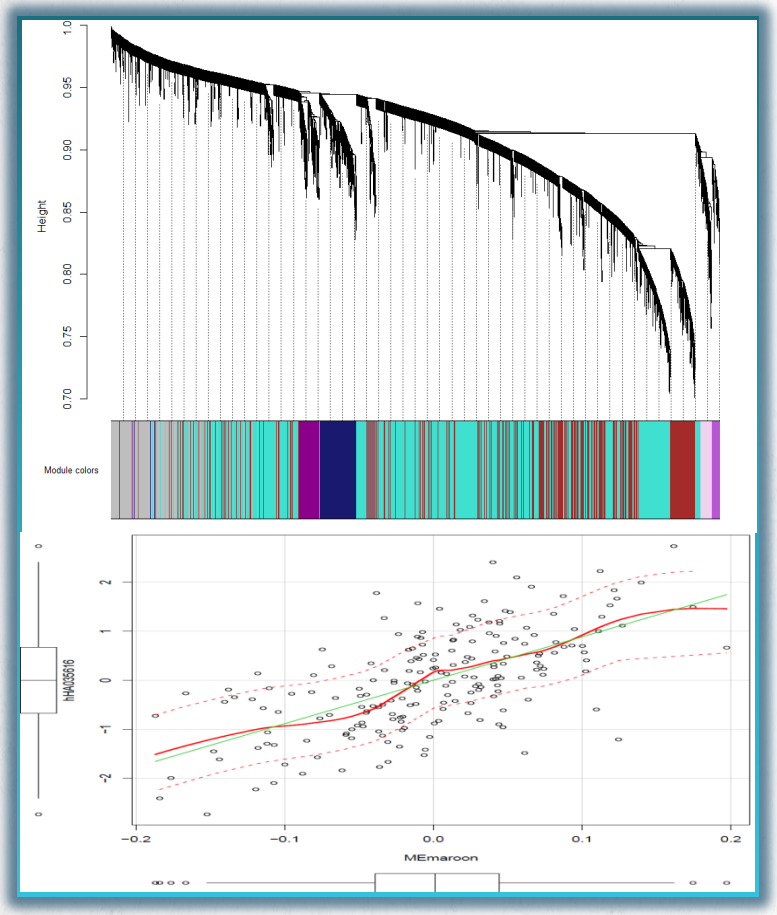


Callicott et al., 2003 *Am. J. of Psychiatry*



Slifstein et al., 2015 *JAMA Psychiatry*

Co-expression prediction of *DRD2* availability in DLPFC



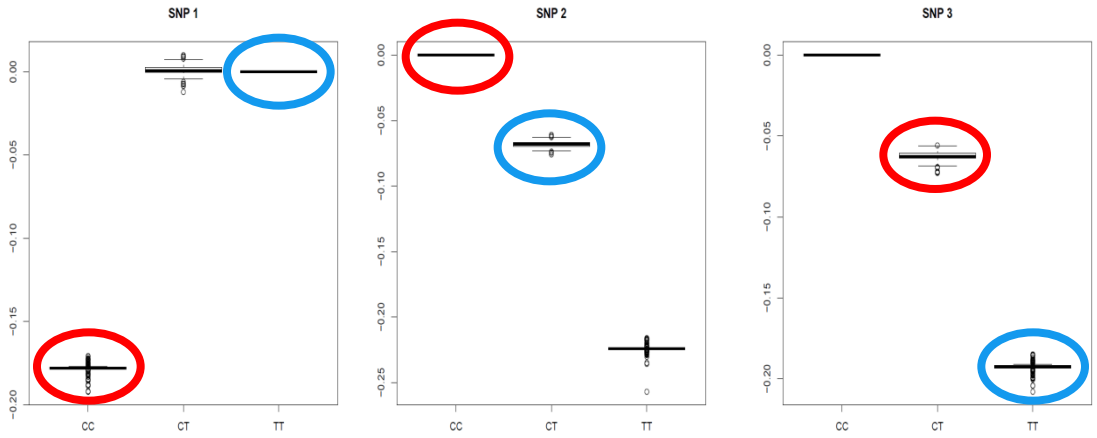
Rank	Marker	Gene	Gene Name	MAF
1	rs2486064	CHIT1	*Chitinase 1*	0,22
2	rs6902039	GPLD1	*Glycosylphosphatidylinositol Specific Phospholipase D1*	0,23
3	rs851436	OSR1	*Odd-Skipped Related 1*	0,48
4	rs9297283	POP1	*Processing Of Precursor 1, Ribonuclease P/MRP Subunit*	0,20
5	rs12940715	SDK2	*sidekick cell adhesion molecule 2*	0,12
6	rs1805453	DHX33	*DEAH (Asp-Glu-Ala-His) Box Polypeptide 33*	0,34
7	rs11213916	BTG4	*B-Cell Translocation Gene 4*	0,30
8	rs1037791	AGR2	*Anterior Gradient 2*	0,31

199 subjects (0-78 years, *Braincloud*)
23636 gene probes
D2L cluster included 85 genes

SNPs were significant at p < .005.

Pergola et al., 2017 *Translational Psychiatry*

Translational genetics – D' expression Index



TT = 0 **CT = -0.07** **CT = -0.19**

Subject 2 Genetic Score = -0.087

We computed D'-values (y-axis) of each SNP genotype population (x-axis) with reference to the major homozygous population.

Boxplots shows leave-one-out cross-validation of D'-values.

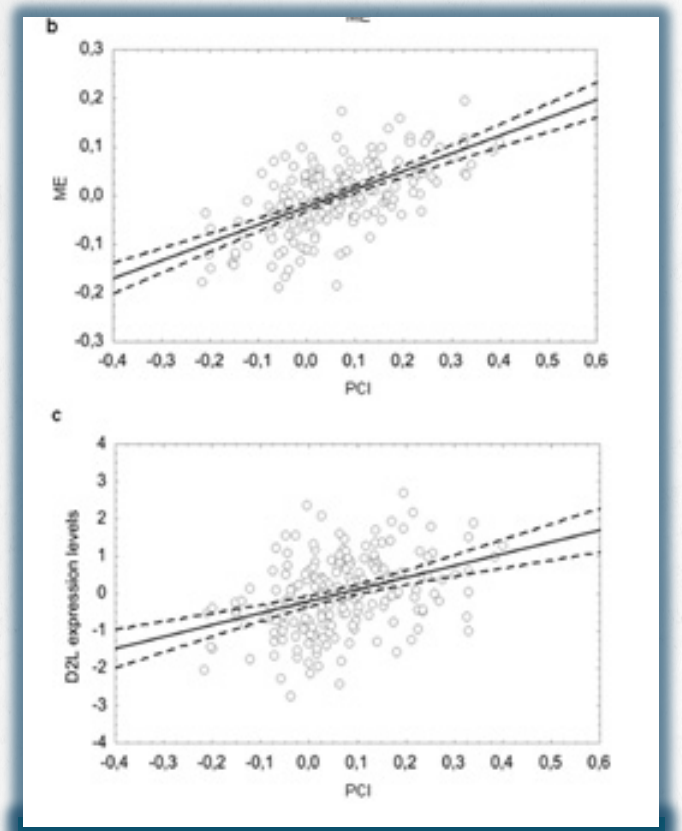
Genetic Score was defined as the mean of D'-values associated to subject's genotypes.

A strength of this approach is that it does not assume a linear influence of allelic dosage.

It is an explicit modeling of additive genetics (no machine learning).

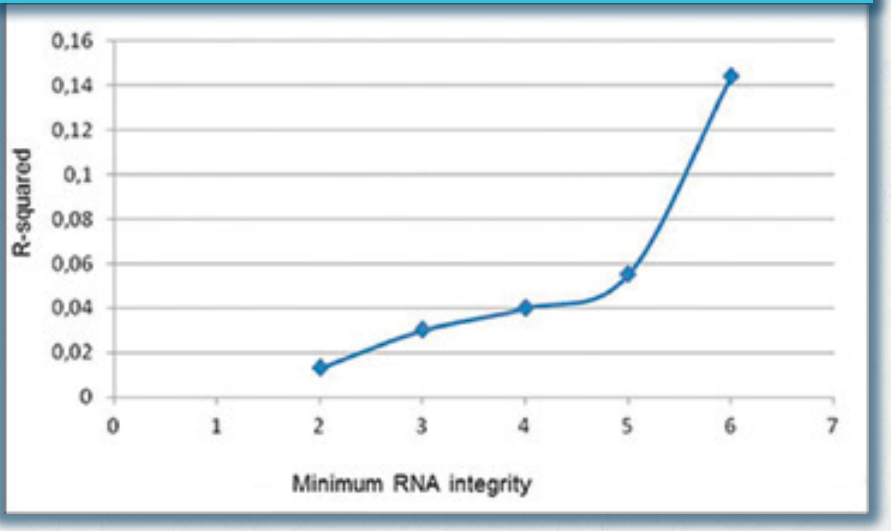
Pergola et al., 2016 *Psychological Medicine*

Translation – Polygenic Co-expression Index



The PCI is positively correlated with gene co-expression and thereby is also positively correlated with *DRD2*.

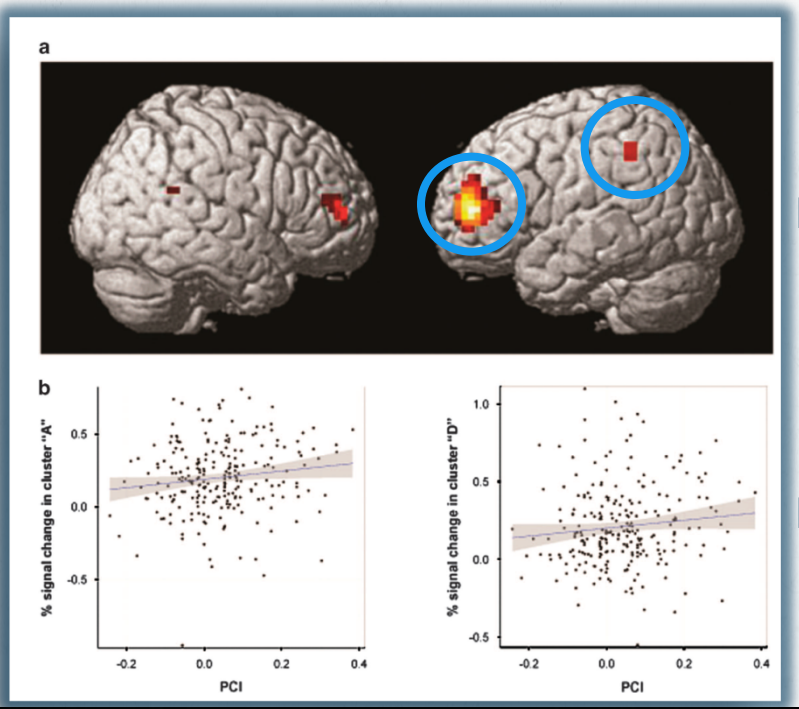
We replicated the correlation in an indipendent microarray dataset (BrainEAC, Trazbuni et al., 2011).
The replication was significant and its strength increased with data quality indexed by RIN.



Pergola et al., 2017 *Translational Psychiatry*

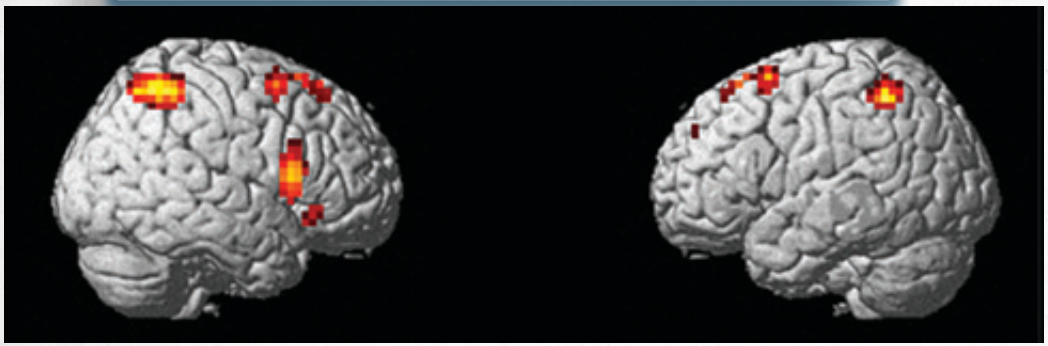
The case of *DRD2* – brain activity during n-back task

Voxel q (FDR) < .05, k = 6, masked for task activity



N = 124

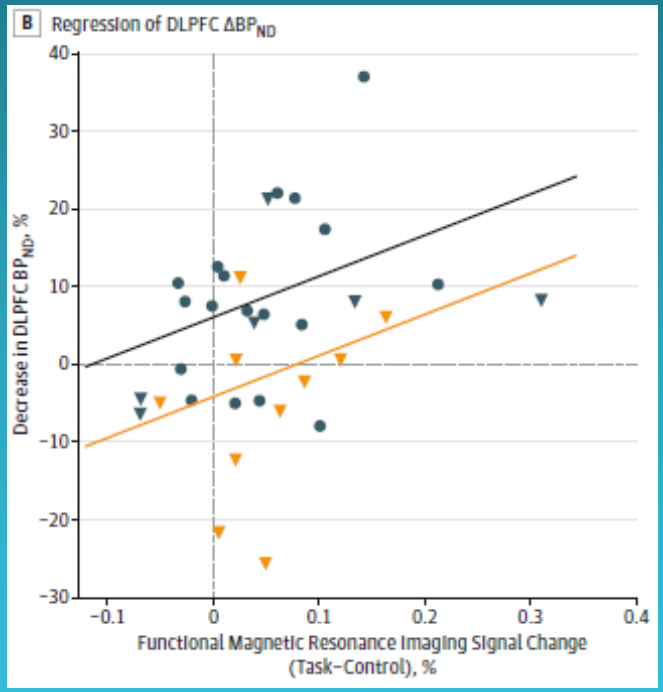
N = 244



N = 29 SCZ

↑ D_2 binding in the PFC

→ ↑ PFC BOLD during WM

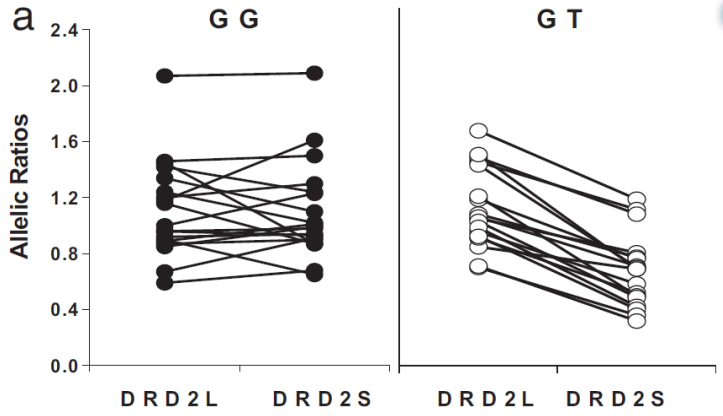


Slifstein et al., 2015 *JAMA Psychiatry*

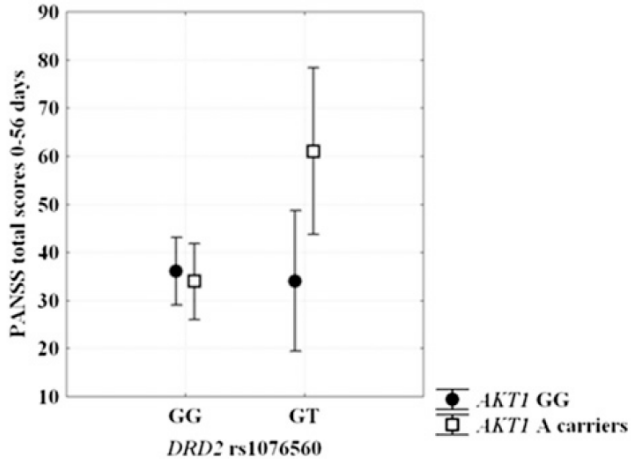
↑ predicted *DRD2* expression → ↑ PFC BOLD during WM in three independent samples of patients and controls.

Pergola et al., 2017 *Translational Psychiatry*

The case of *DRD2* – treatment response

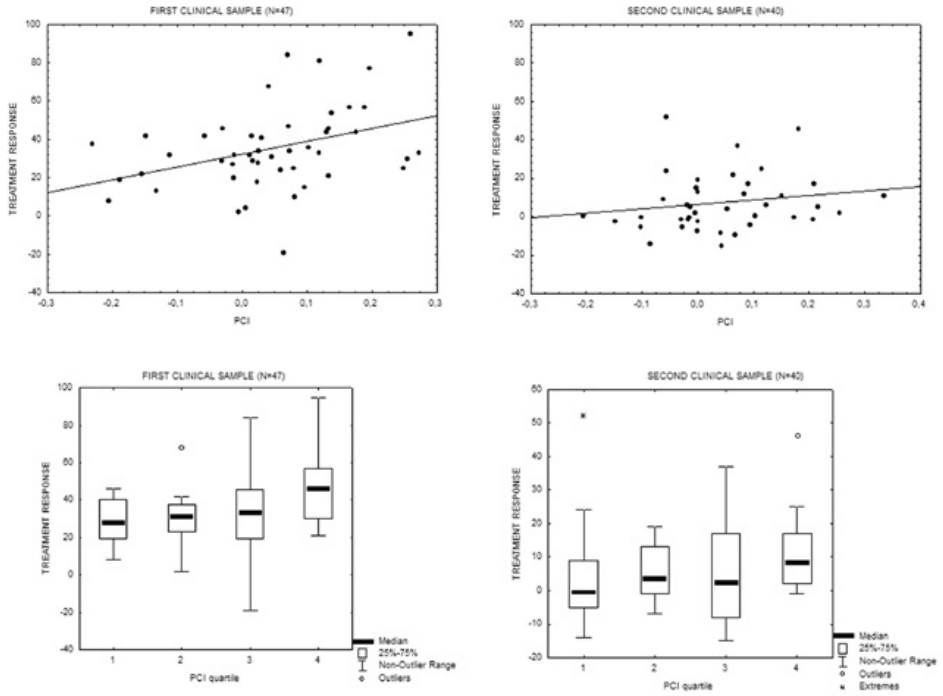


Zhang et al., 2007 *PNAS*



Blasi et al 2011 *PNAS*

↑ D2L → ↑ treatment response

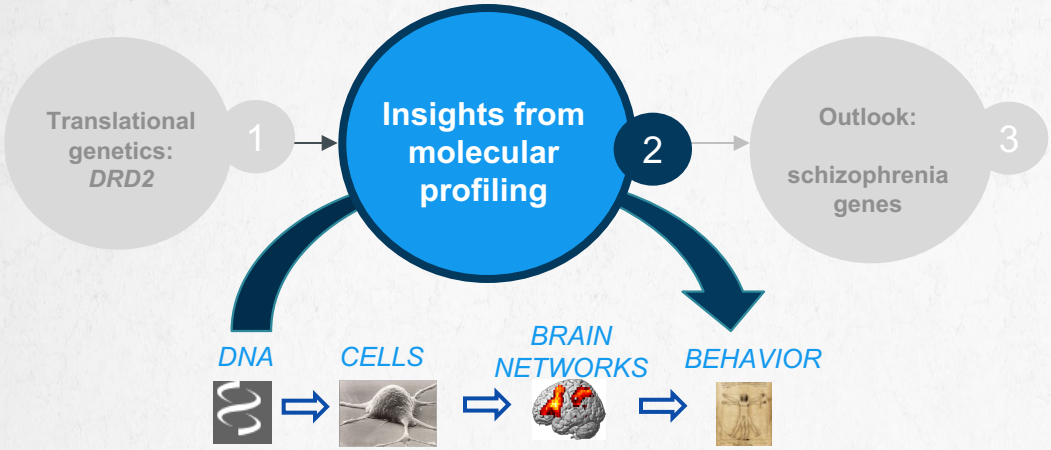


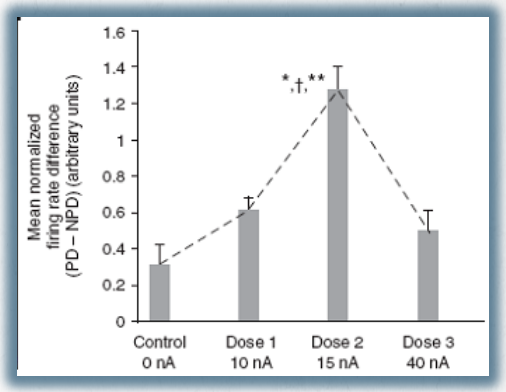
↑ predicted expression → ↑ treatment response to atypical antipsychotics in two independent samples of patients (total N=87).

Pergola et al., 2017 *Translational Psychiatry*

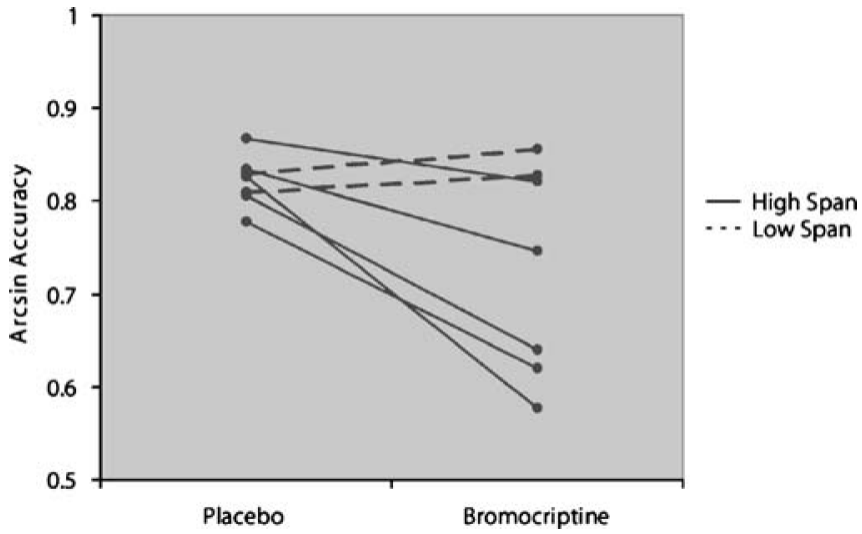
These results show the biological validity of the PCI identified via co-expression networks

Brain activity during working memory performance



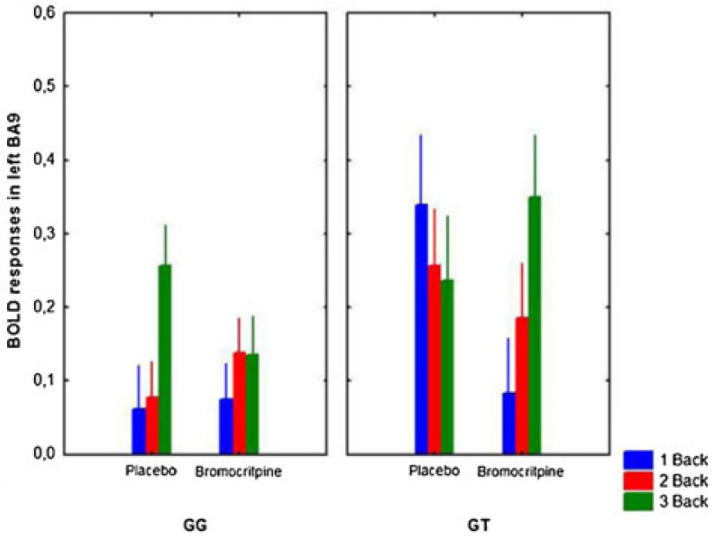


Vijayraghavan et al., 2007 *Nature Neuroscience*



Individuals with low WM capacity benefit from treatment with D_2 agonists like bromocriptine.

Gibbs and D'Esposito, 2005 *Psychopharmacology*

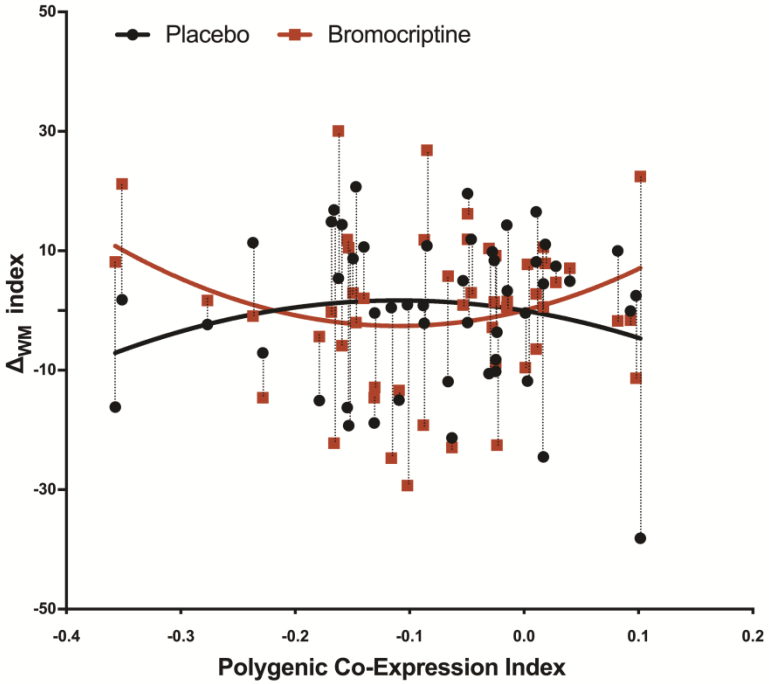


Load-dependent DLPFC activation is associated with *DRD2* rs1076560 genotype. Note that the effect *changes dramatically between loads.*

Gelao et al., 2014 *Psychopharmacology*

Bromocriptine response depends non-linearly on *DRD2*-PCI

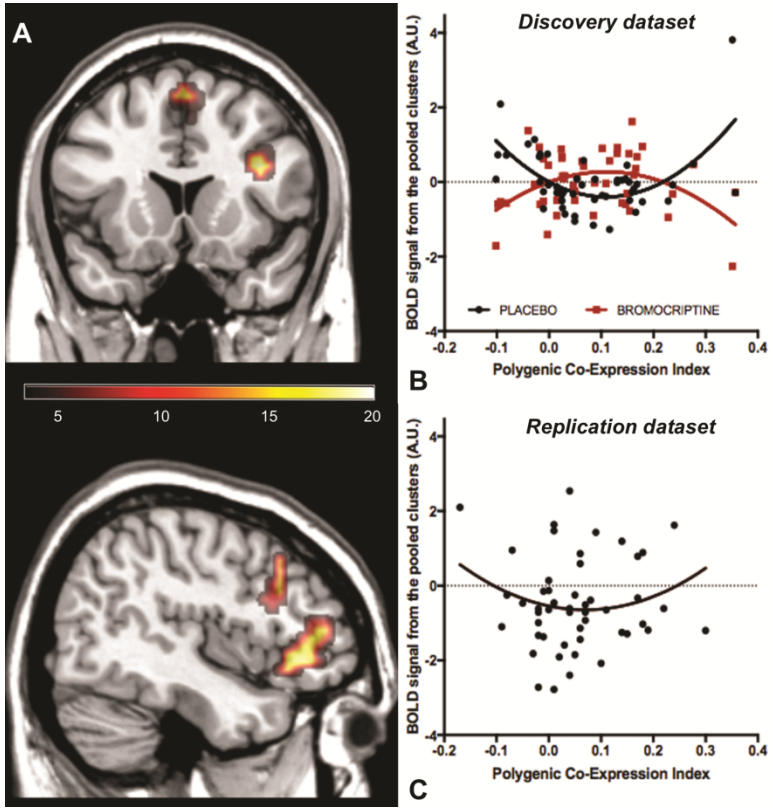
- **Discovery** N = 50, **Replication** N = 50 (previous study)
- Discovery: double-blind, crossover, randomized, placebo controlled trial with Bromocriptine 1.25 mg
- *DRD2*-PCI as linear & quadratic predictor of bromocriptine response
- **Differential accuracy** (3-2-back) during WM performance (Cassidy et al., 2016 *Journal of Neuroscience*) to index *individual WM capacity*
- **Differential PFC activation** between loads as the corresponding neural substrate (3-2-back)



Behavior: **DRUG × PCI² p = .045**

BOLD: **DRUG × PCI² cluster-level FWE p < .05**

Bootstrapped p < .05



Selvaggi, Pergola, et al., in revision

Discussion part 1 & 2

- The PCI approximates *DRD2* co-expression.

- **When assessing individual WM capacity and its neural correlates**, there was a **U-shaped relationship** between predicted *DRD2 co-expression* and brain/behavioral phenotypes.

- This relationship was **reverted by bromocriptine**, such that only ***participants with low WM capacity*** manifested a visible change.

- Critically, **individuals at the two extremes of the curve have *opposite allelic patterns***. Rather than risk alleles, **the ensemble of *DRD2*-related genetic variation** is associated with drug response based on its relationship with gene expression.

RELEVANCE

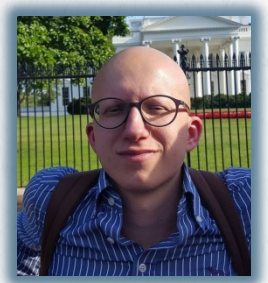
Gene co-expression networks reveal novel genetic players in the regulation of dopaminergic transmission.

Allelic variation *did not evolve to support drug response*, but to sustain molecular processes such as gene expression.

***Different alleles* can predict *the same outcome* when weighted for a cardinal principle of biological organization such as gene regulation.**

Thank you for your attention!!

Thanks to all the people who contributed to this work!



Thanks for supporting this work:

Prof. Alessandro Bertolino, MD, PhD

Fondazione CON IL SUD

Ministero della Salute

Brain& Behavior Foundation

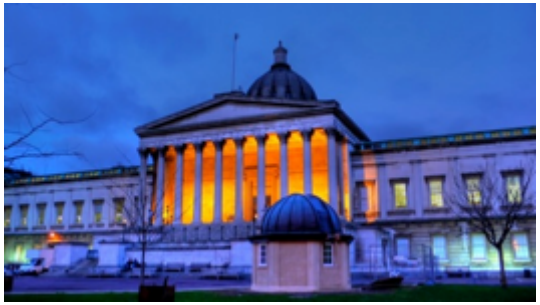
IMAGEMEND

Hoffmann-La Roche

Giuseppe Blasi, MD, PhD; Grazia Caforio, MD, PhD; **Pasquale Di Carlo, MD**; Leonardo Fazio, PhD; Barbara Gelao, PhD; Rita Masellis; **Marco Papalino, MD**; **Roberta Passiatore**; Tiziana Quarto; Antonio Rampino, MD, PhD; Raffaella Romano,PhD; **Pierluigi Selvaggi, MD**.

Giulio.Pergola@uniba.it

Partial Least Squares for large distributed datasets in imaging genetics



Andre Altmann

Department of Medical Physics and Biomedical Engineering

Translational Imaging Group (TIG)

Centre for Medical Image Computing (CMIC)

a.altmann@ucl.ac.uk

Predictive radiology for precision medicine
Lausanne, November 13th 2017

Outline

Imaging Genetics

Partial Least Squares (PLS)

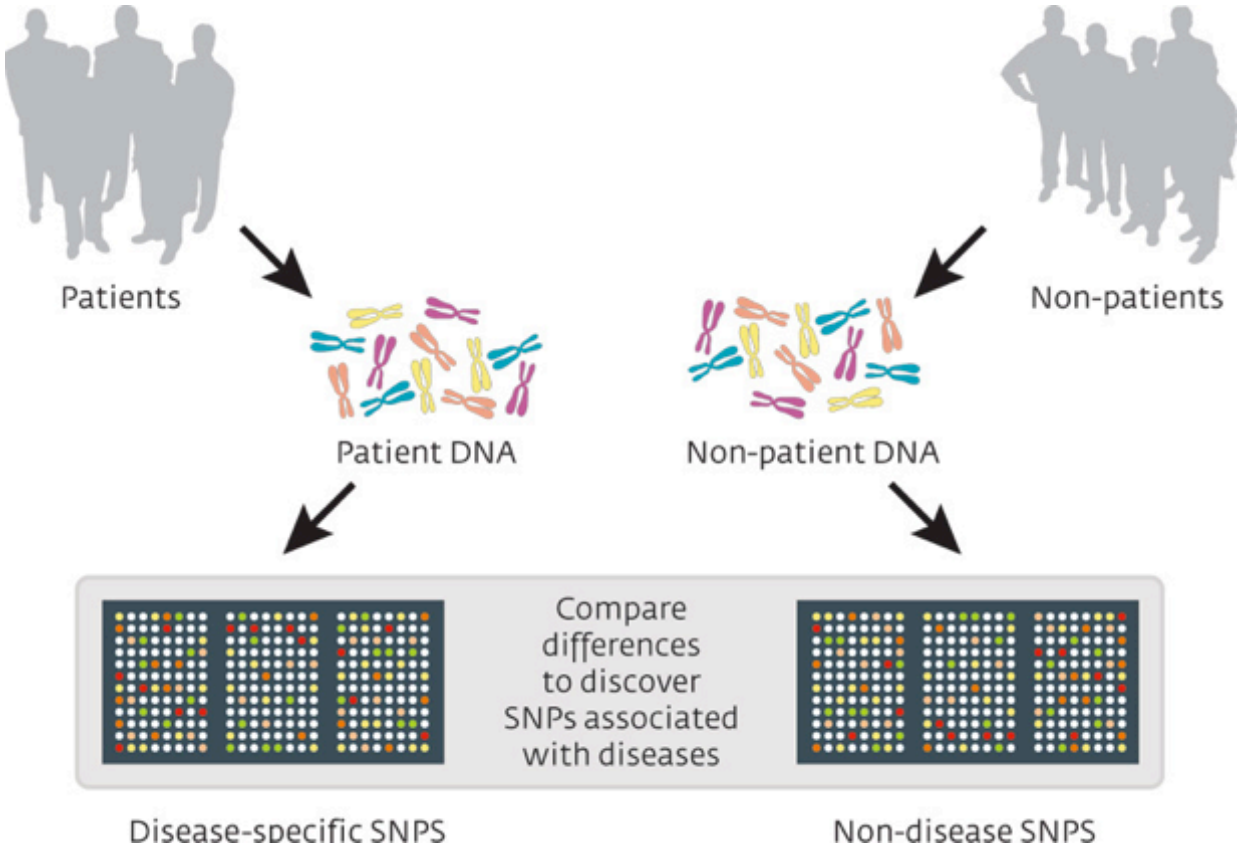
Application to AD

Distributed PLS

Imaging Genetics

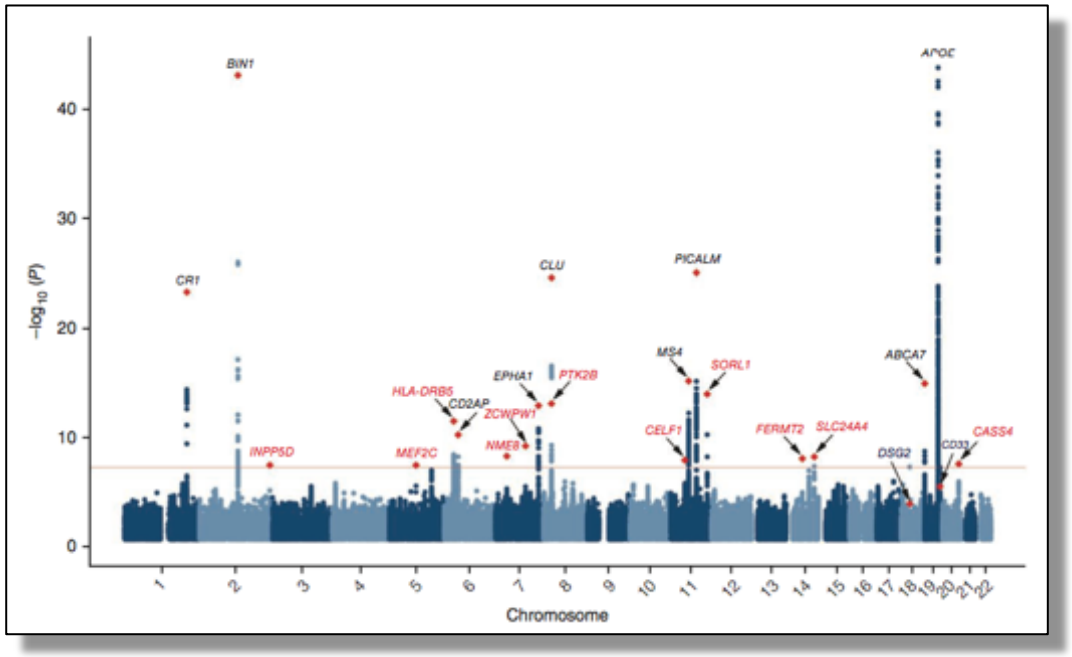
Genome-wide Association Study (GWAS)

- Single Nucleotide Polymorphisms (SNPs)
- ~ 1 million genetic markers measured
- Sample acquisition cheap (£50 per participant)



Genome-wide Association Study (GWAS)

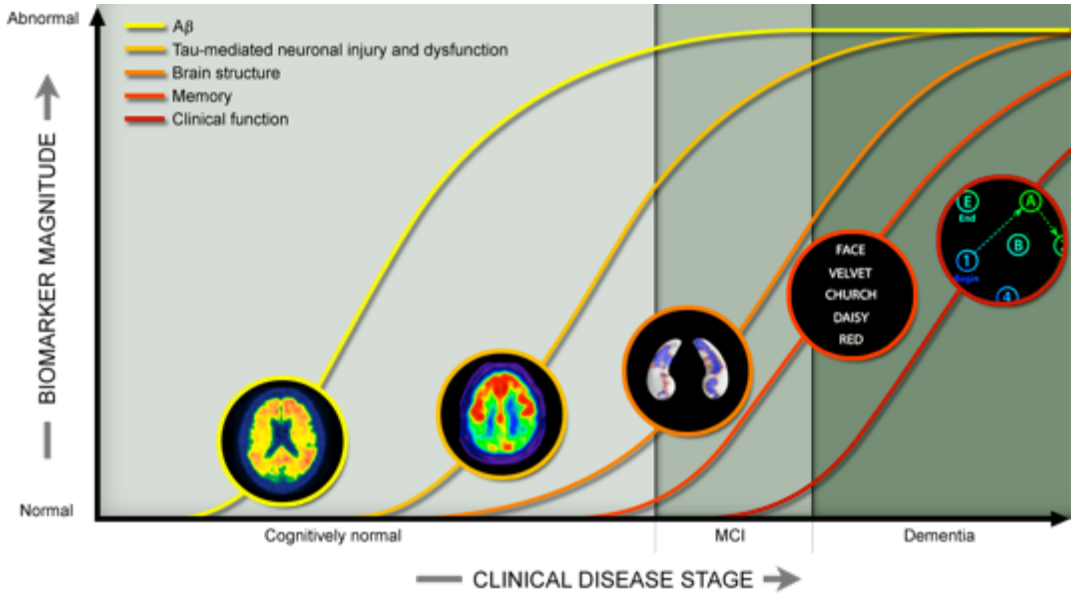
$$y = \beta_0 + \beta_1 \times \text{SNP} + \sum_{i=2}^{N} \beta_i x_i + \epsilon$$



- Mass univariate testing
- P<5x10^{-8} is ‘genome-wide significant’

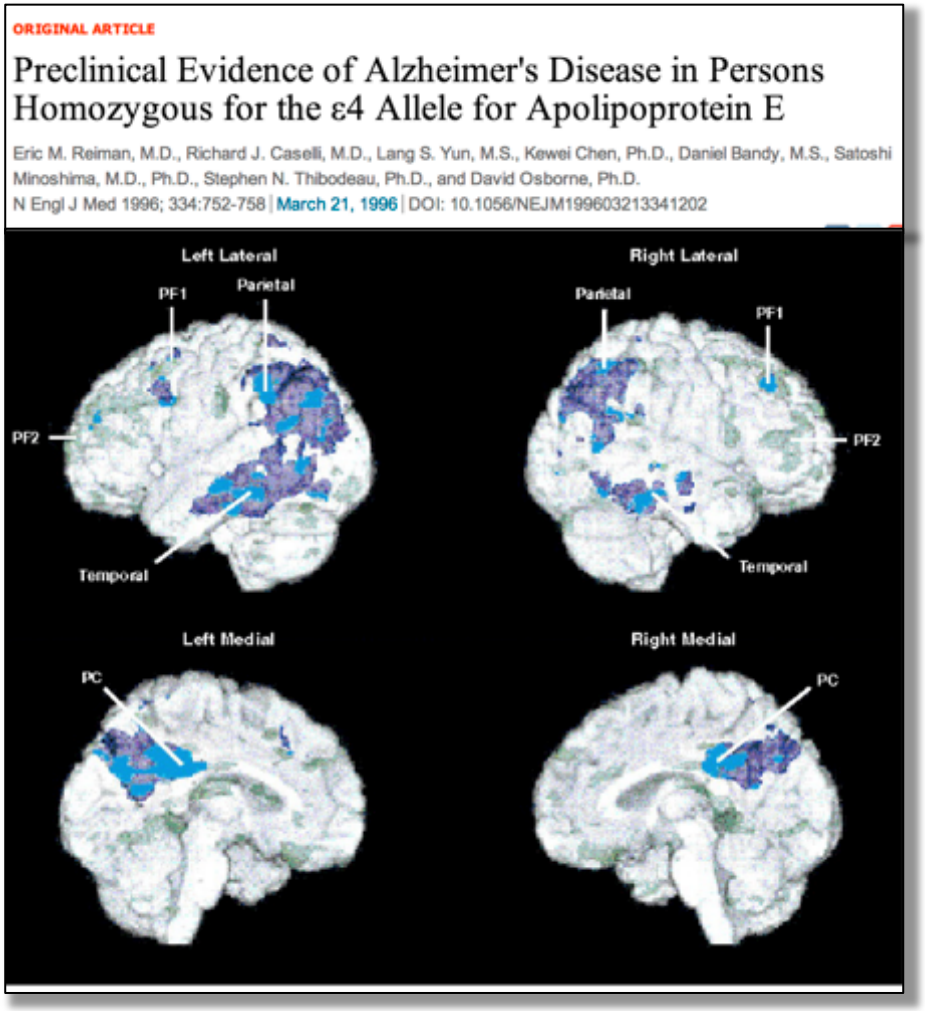
Why Imaging Genetics?

- Imaging genetics uses imaging phenotypes in genetic studies
- Imaging phenotypes are ‘closer’ to the disorder than diagnoses
  - Often diagnoses are ‘mixed bags’
  - Allows us to study how the disorder develops

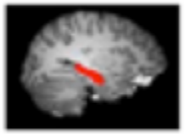
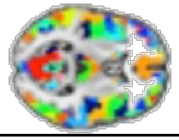
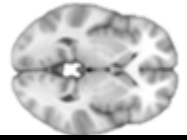
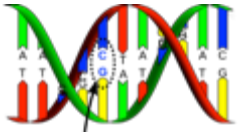
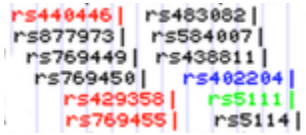


Early Imaging Genetics Example in AD

- APOE-e4 is a strong risk factor for AD
- Reiman et *al.* (1996) studied glucose metabolism
- APOE-e4 homozygotes showed reduced meta-bolism in typical AD regions

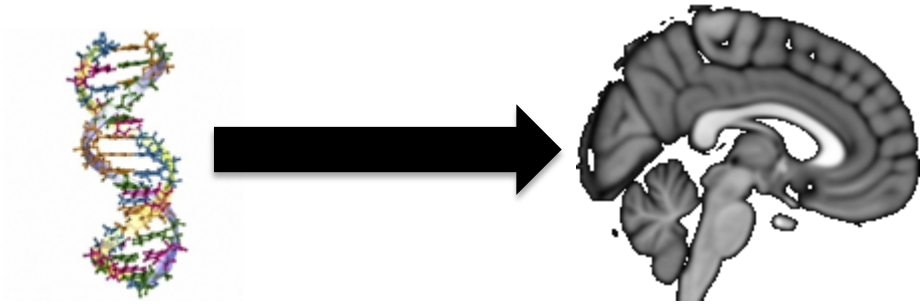


Imaging Genetics "Menu"

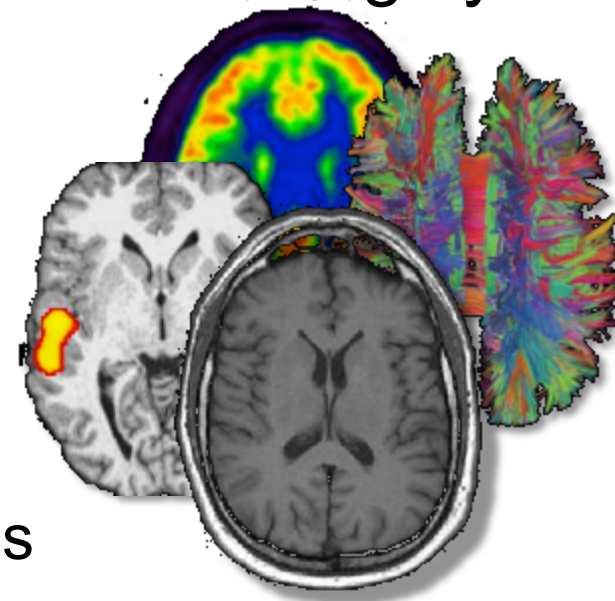
	Candidate ROI 	Many ROIs 	Voxelwise 
Candidate SNP 			[Filippini et al. 2009] 29,812 voxels 1 SNP
Candidate Gene 	[Joyner et al. 2011] 4 ROIs, 11 SNPs		
Genome-wide Gene 			[Hilbar et al. 2011] 31,622 voxels 18,044 Genes
Genome-wide SNP 	[Stein et al. 2012] 1 ROI, 1.1 mio SNPs [Potkin et al. 2009] 1 ROI, 317,503 SNPs		[Stein et al. 2010] 31,622 voxels 448,293 SNPs

Partial Least Squares

Imaging Genetics – Multiple Outputs

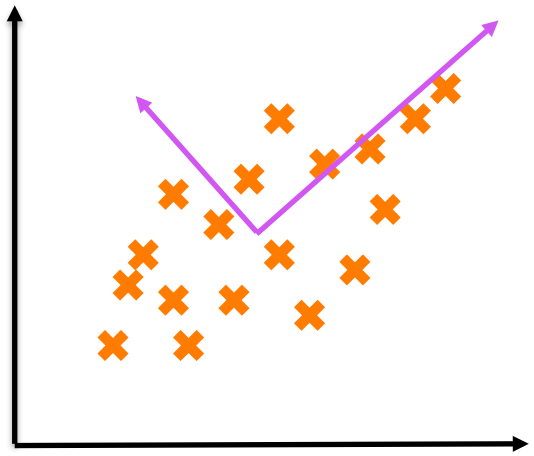


- Multivariate methods work well, however …
  … genetics data and imaging data are both highly multidimensional (P_1 and P_2 >> N)
  … multiple imaging modalities
- The 'easy' way out
  - Mass univariate
    - no machine learning, P_1 x P_2 tests
  - P_1 (independent) multivariate analyses
  - First identify patterns in one dataset (mainly imaging) followed by GWAS/multivariate analysis



Principle Component Analysis (PCA)

- Common method to identify patterns in data… or to reduce dimensionality



1. PCA for imaging data, PCA for genetics data
2. Find the latent ‘features’ that correlate
   - PCAs are computed independently
   - No information transfer between images and genetics

Methods for Multiple Outputs

- Multi-Task Learning (or Multi-Response Learning)
  - Instead of doing P_1 independent analyses
  - Also regularize across different tasks (brain regions)
- Parallel Independent Component Analysis (ICA)
- Canonical Correlation Analysis (CCA)
  - Close relative: reduced rank regression (RRR)
- Partial Least Squares (PLS)

RRR <–> PLS <–> CCA

UCL

Multi-output predictions from neuroimaging: assessing reduced-rank linear models

Mehdi Rahim, Bertrand Thirion, Gaël Varoquaux
Parietal - Inria / CEA. Paris-Saclay University. France
mehdi.rahim@inria.fr

PRNI, 2017

Foundation of PLS

- Related to CCA
  - Not maximizing correlation but covariance
- Two matched datasets **X** and **Y**, we aim to find a projection **u** and **v**, such that

$$\mathbf{P_x} = \mathbf{X}\vec{u}' \quad \mathbf{P_y} = \mathbf{Y}\vec{v}'$$

have the maximal covariance

$$\underset{\vec{u},\vec{v}}{\operatorname{argmax}} \operatorname{cov}(\mathbf{P_x}, \mathbf{P_y})$$

Foundation of PLS

- Related to CCA
  - Not maximizing correlation but covariance
- Two matched datasets **X** and **Y**, we aim to find a projection **u** and **v**, such that

$$\mathbf{P_x} = \mathbf{X}\vec{u}' \quad \mathbf{P_y} = \mathbf{Y}\vec{v}'$$

have the maximal covariance

$$\underset{\vec{u},\vec{v}}{\operatorname{argmax}} \operatorname{cov}(\mathbf{P_x}, \mathbf{P_y})$$

- **u** and **v** provide weights for each original feature
- Further projections identified iteratively
  - Deflating **X** and **Y** and computing the next projection
  - We get series of projections: $\mathbf{U}=[u_1,\ldots,u_N]$, $\mathbf{V}=[v_1,\ldots,v_N]$

SVD version of PLS

- Nowadays a popular used version of PLS is PLS-SVD (singular value decomposition)

$$\mathbf{M} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}'$$

Unitary matrix, right singular vectors

Unitary matrix, left singular vectors

Diagonal matrix, singular values

- Allows to compute **U** and **V** as

$$\mathbf{X}\mathbf{Y}' = \mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}'$$

cross-covariance matrix

- Solves PLS in 'one go', but …
  … **C** is pretty large (P_1 x P_2) feature x feature
  - 1 mio SNPs times 300k voxels

Memory efficient PLS-SVD

- 'Rephrasing' avoids computation of **C**

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{Y}'\mathbf{Y} \end{pmatrix} \mathbf{A} = \mathbf{A}L$$

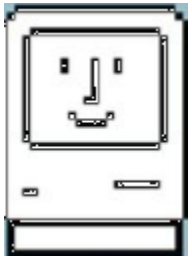
Diagonal matrix, eigenvalues

Eigenvectors

$$\mathbf{B} = \mathbf{A}\left(\mathbf{A}'\mathbf{Y}'\mathbf{Y}\mathbf{A}\right)^{-\frac{1}{2}}$$

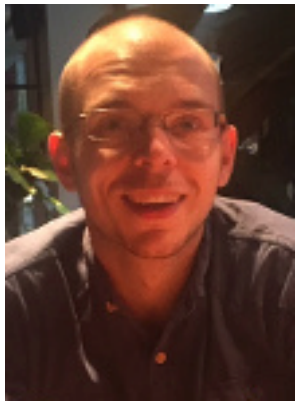
$$\mathbf{V} = \mathbf{Y}\mathbf{B} \quad \mathbf{U} = \mathbf{X}\left(\mathbf{Y}'\mathbf{Y}\mathbf{B}L^{-\frac{1}{2}}\right) \quad \mathbf{\Lambda} = L^{-\frac{1}{2}}$$

- Allows working with covariance matrices for **X** and **Y** of size *NxN* (subject x subject)
  … much more tractable!!



PLS in imaging genetics for AD

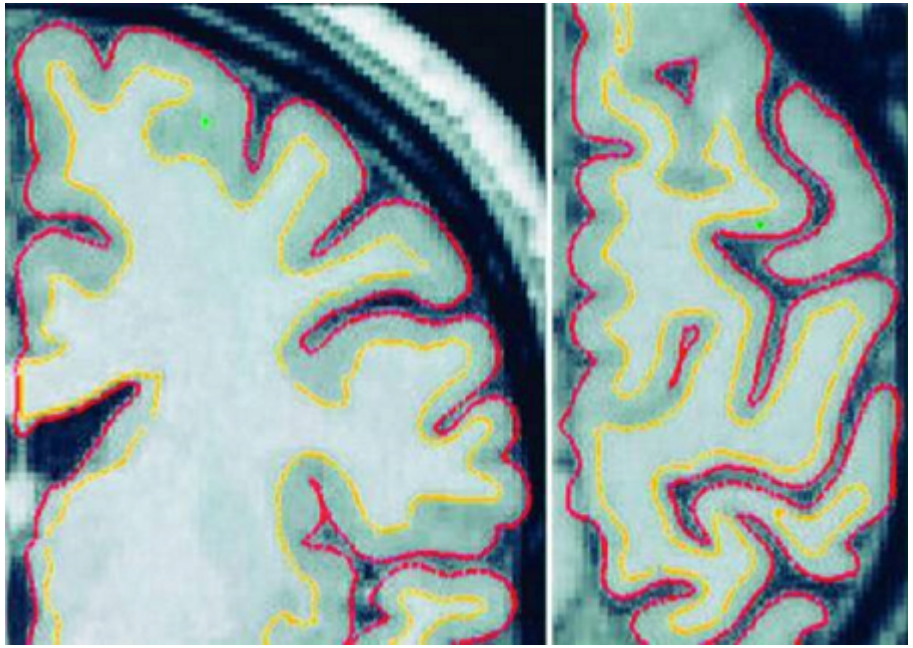
PLS applied: Cortical Thickness in AD

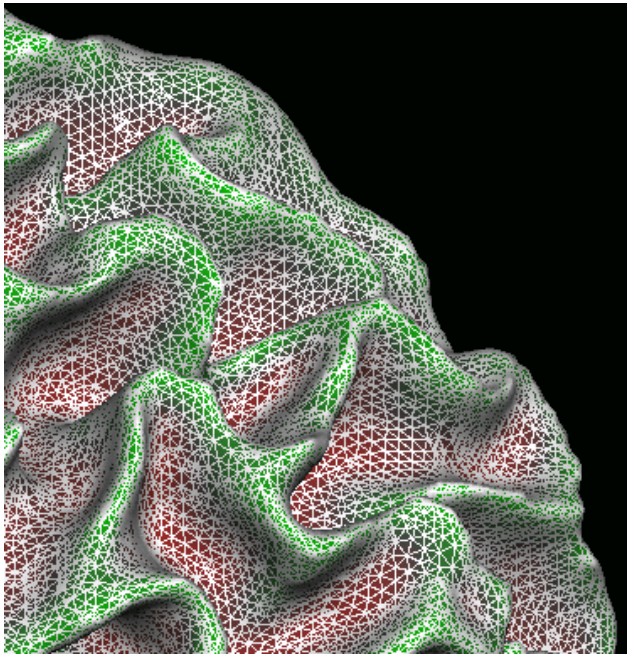


Marco Lorenzi

Genetics of cortical thickness in AD

- Loss of gray matter in AD
- Cortical thickness derived from structural T1 weighted MRI scans



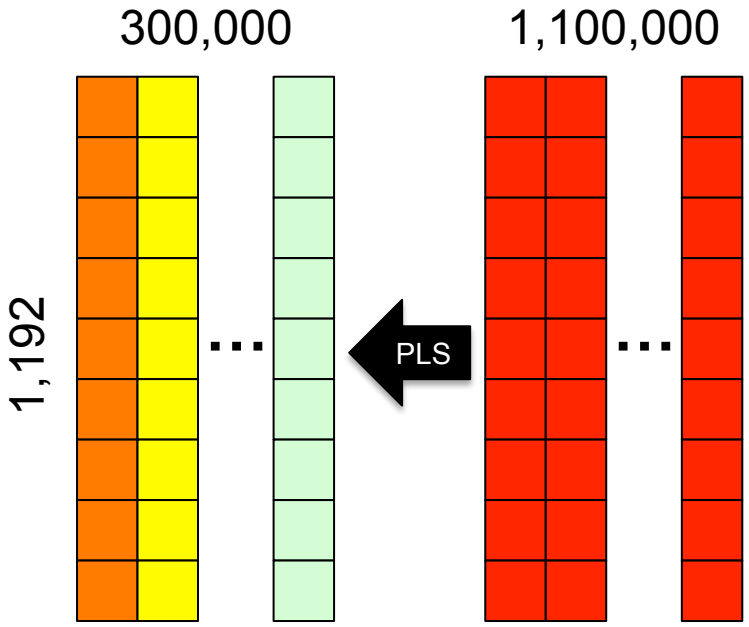


PLS applied: Cortical Thickness in AD

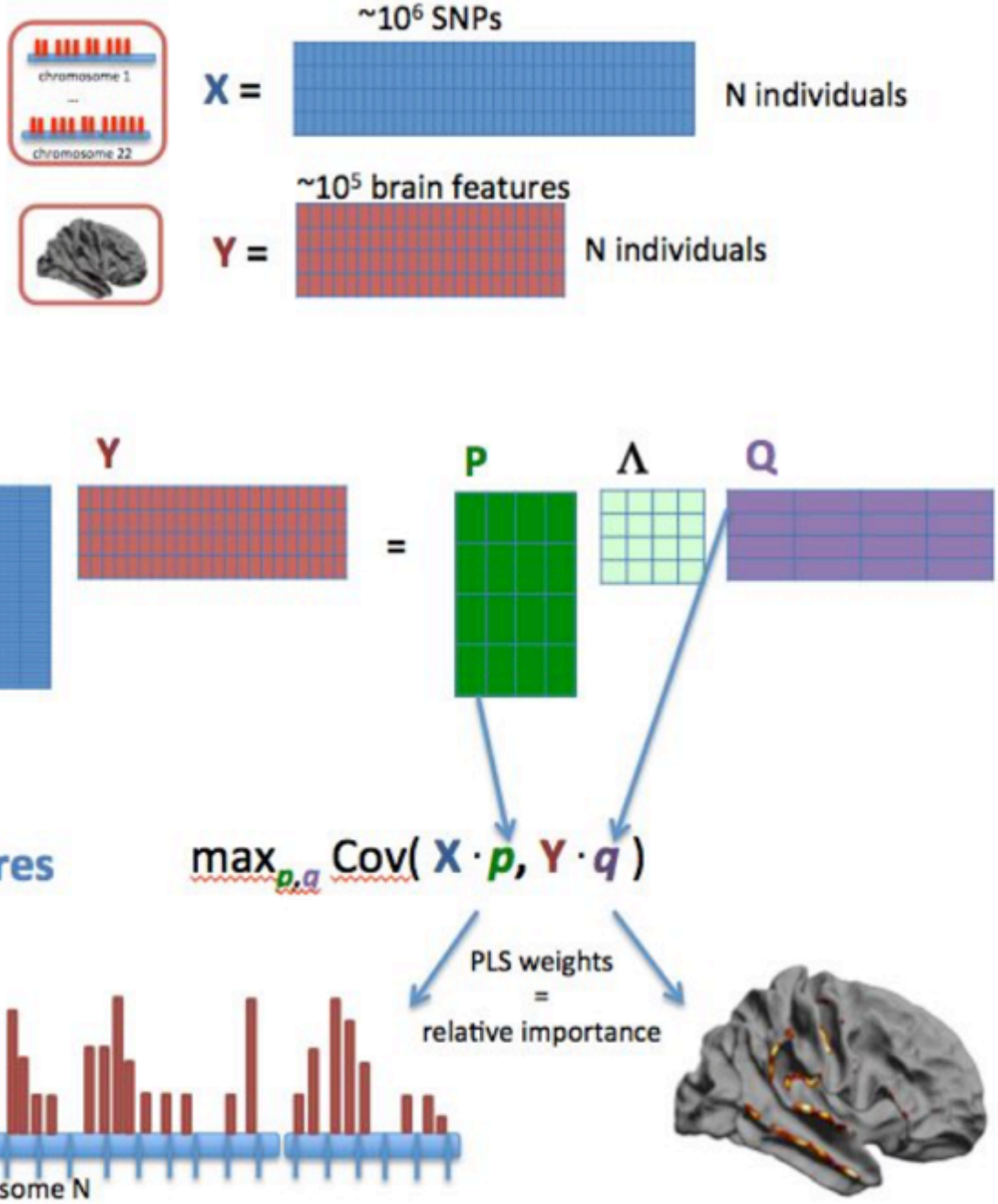
Genetics of cortical thickness in AD

- Loss of gray matter in AD
- Cortical thickness derived from structural T1 weighted MRI scans

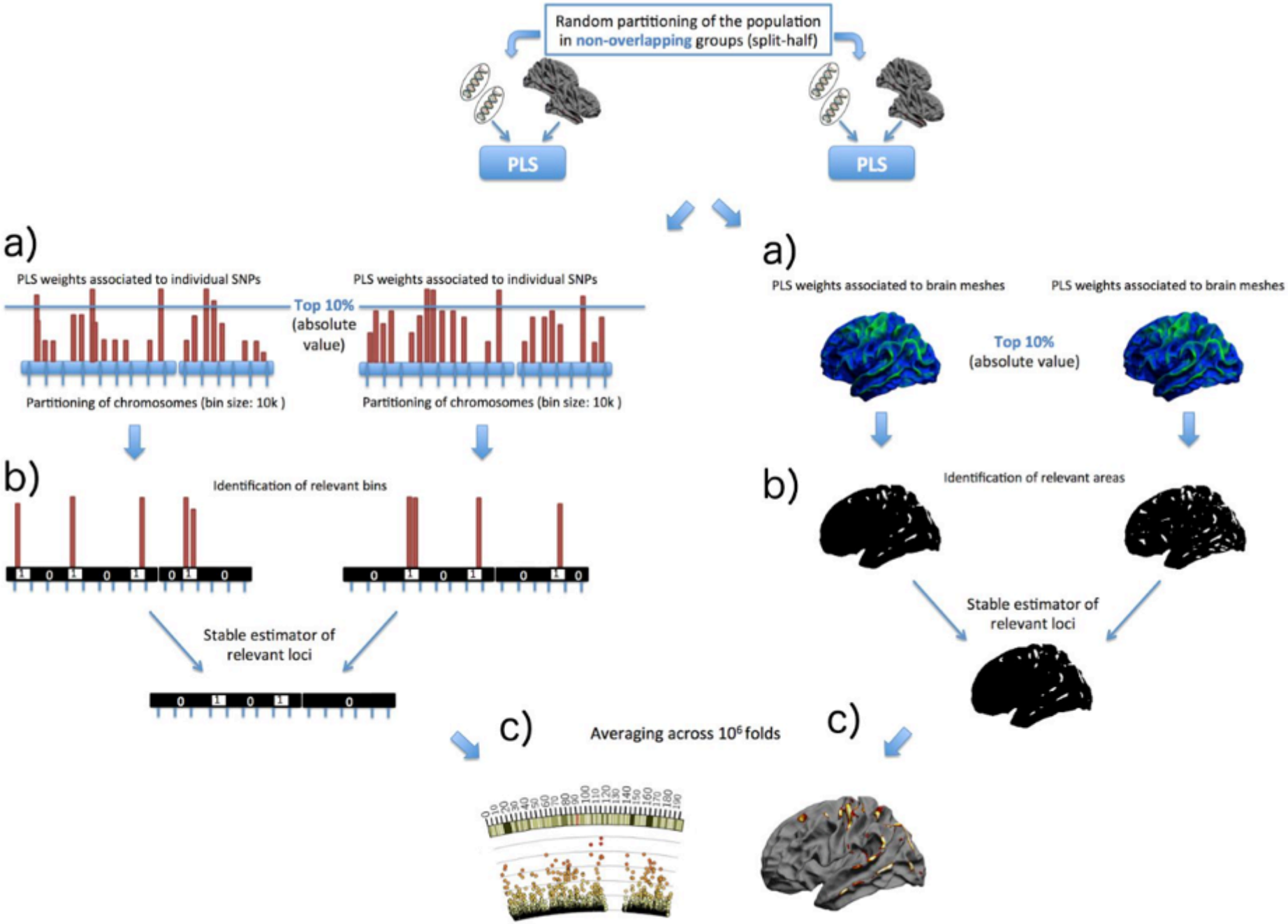
- ADNI database
  - 300,000 mesh points
  - 1.1 mio SNPs
  - 1,192 subjects (HC, MCI, AD)
    - 639 training (HC, AD)



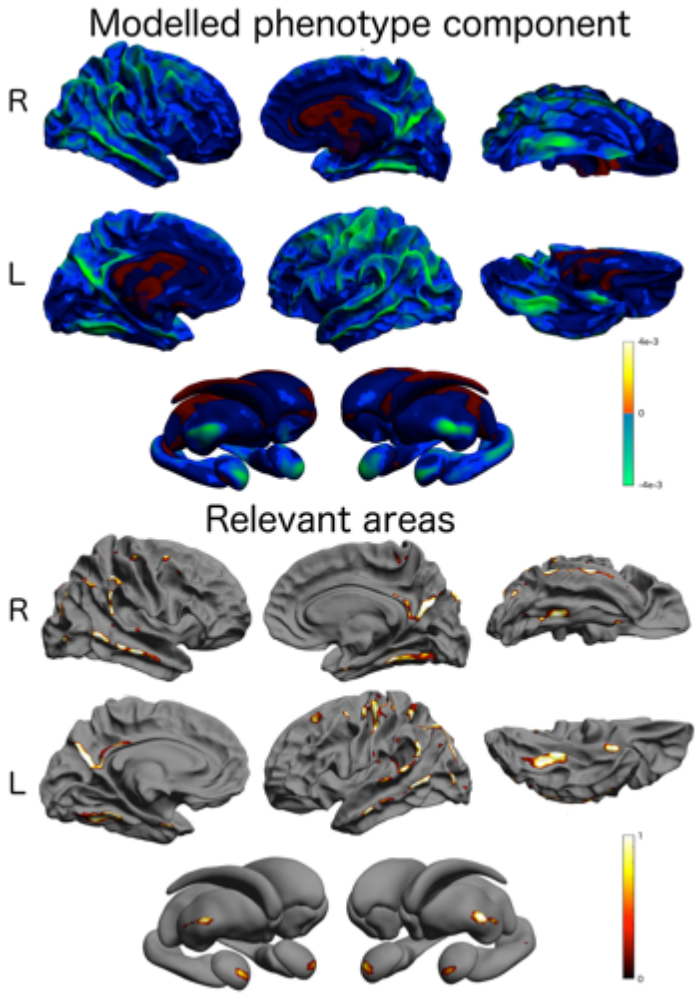
PLS applied: Cortical Thickness in AD

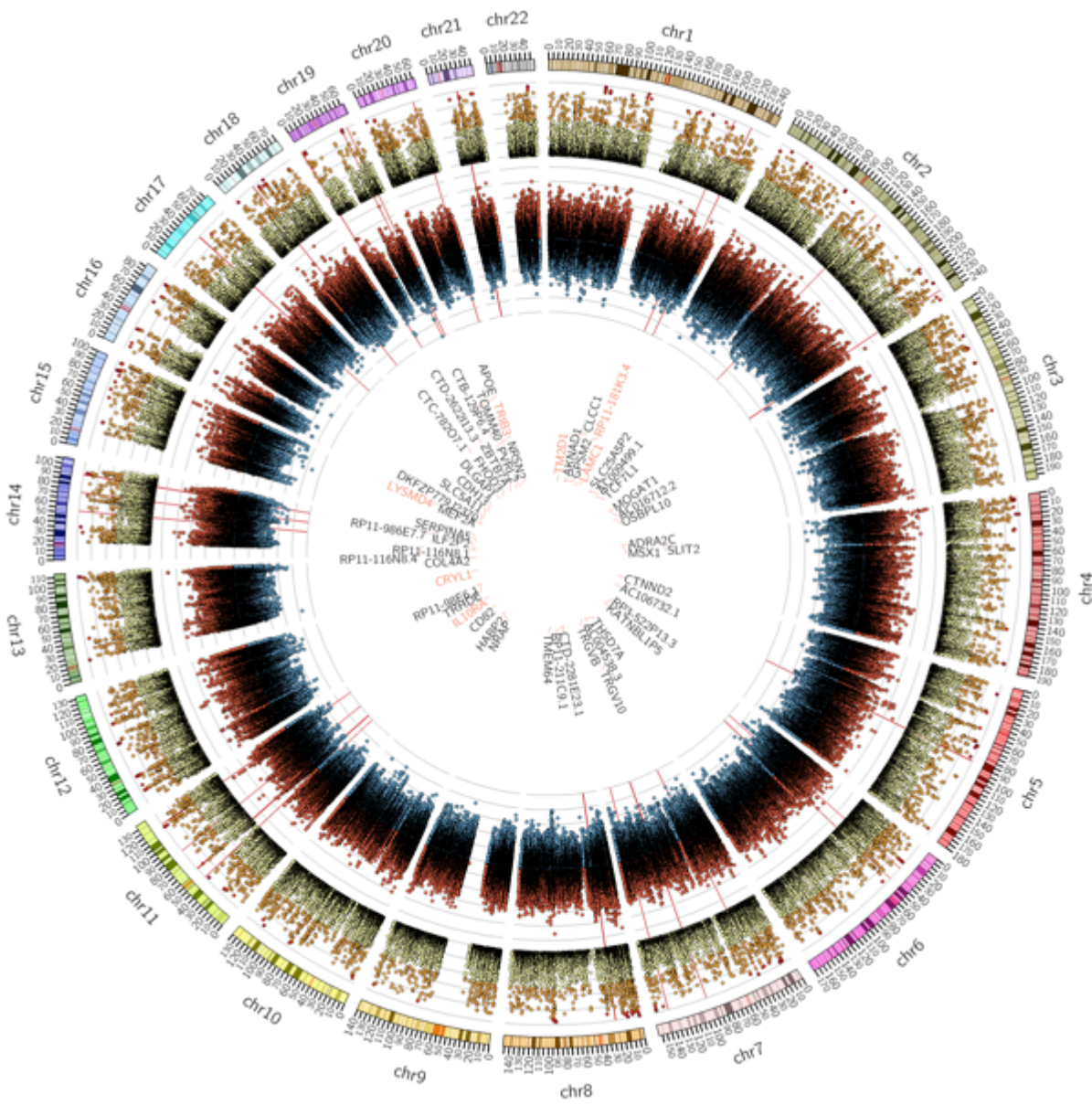


PLS stability with re-sampling



Results – Component 1





Lorenzi*,Altmann*, et al. *PNAS* (in revision)

Distributed PLS

Genetic Data are Sensitive Data!

RESEARCH

HUMAN GENETICS

Deriving genomic diagnoses without revealing patient genomes

Karthik A. Jagadeesh,[1*] David J. Wu,[1*] Johannes A. Birgmeier,[1]
Dan Boneh,[1,2†] Gill Bejerano[1,3,4†]

Patient genomes are interpretable only in the context of other genomes; however, genome sharing enables discrimination. Thousands of monogenic diseases have yielded definitive genomic diagnoses and potential gene therapy targets. Here we show how to provide such diagnoses while preserving participant privacy through the use of secure multiparty computation. In multiple real scenarios (small patient cohorts, trio analysis, two-hospital collaboration), we used our methods to identify the causal variant and discover previously unrecognized disease genes and variants while keeping up to 99.7% of all participants' most sensitive genomic information private.

Science, Aug 2017

Avoid Bureaucracy!

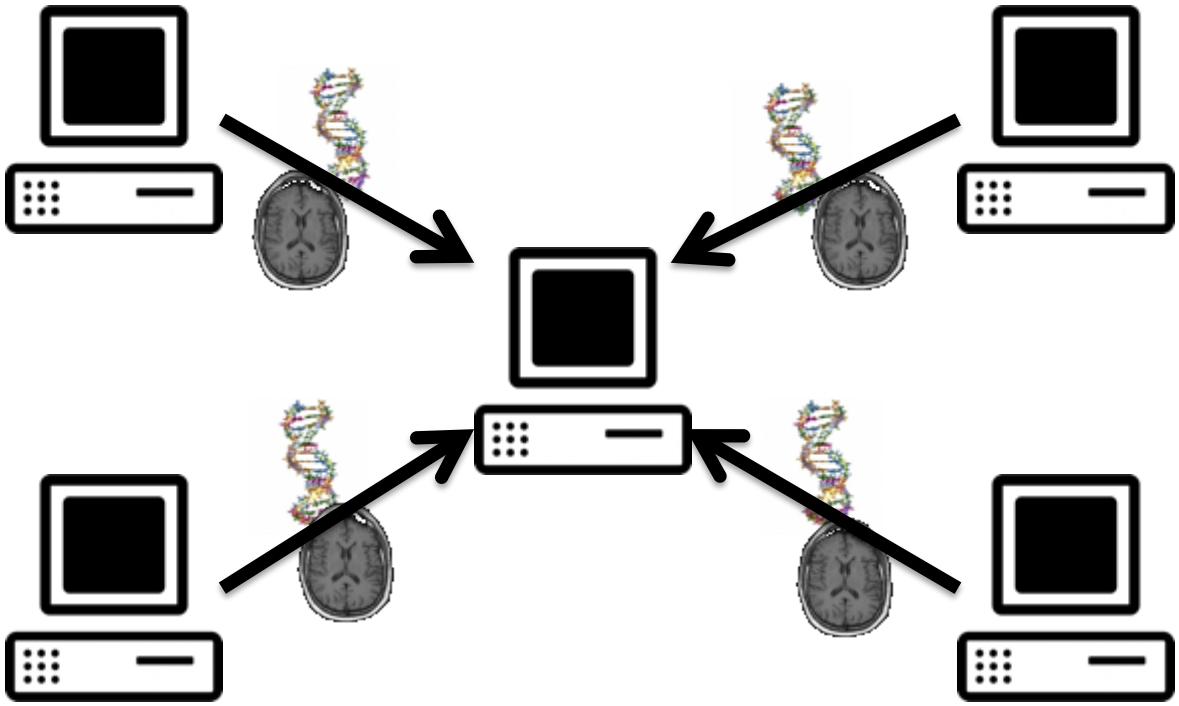
Meta analysis and online learning?

- Meta analysis
  - Every participants runs a local univariate analysis
  - Results (p-values, effect sizes) are shared and combined for a final result
  - Common in large genetic studies
  - Increasingly considered in imaging, e.g., ENIGMA

- Online learning
  - Machine learning method
  - Distributed data/data streams
  - Models are updated with new batches of data

Meta PLS

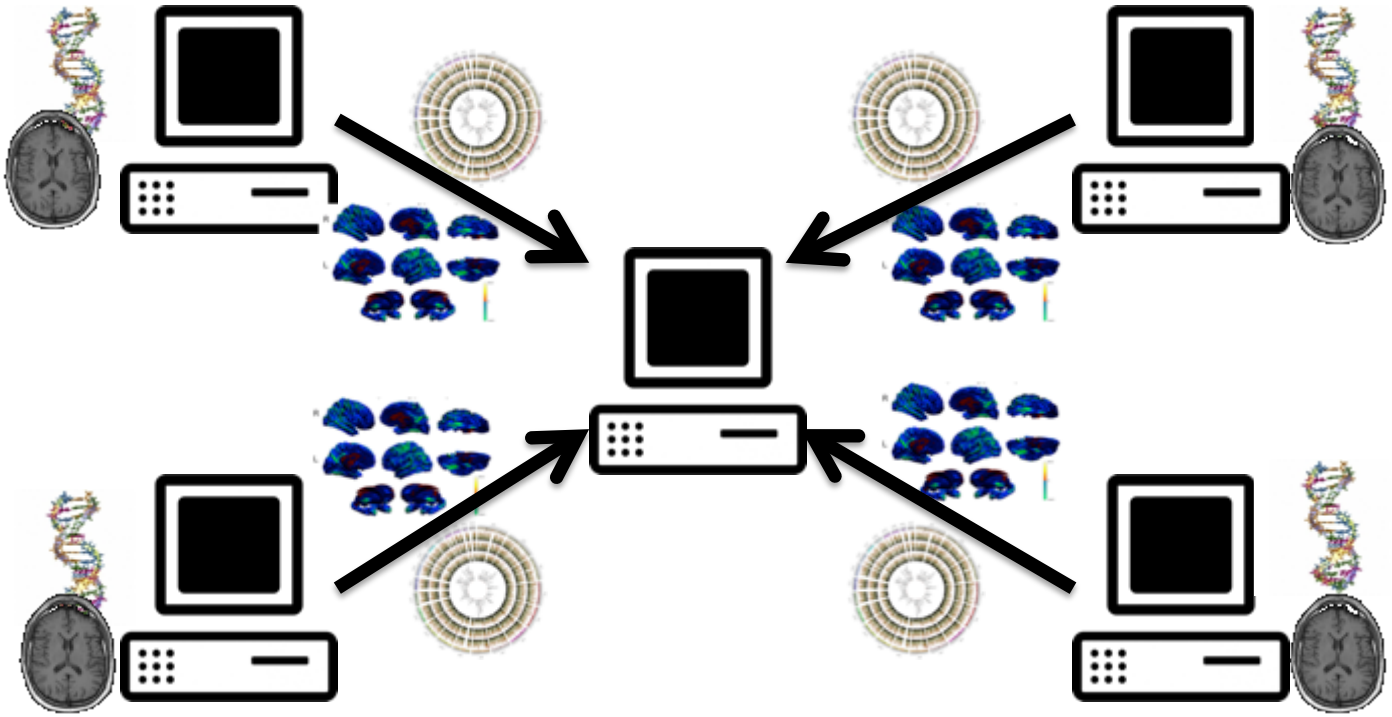
- Massively Multivariate Studies are small
- A “meta” version of PLS enables collaborative studies without the need to exchange individual level data



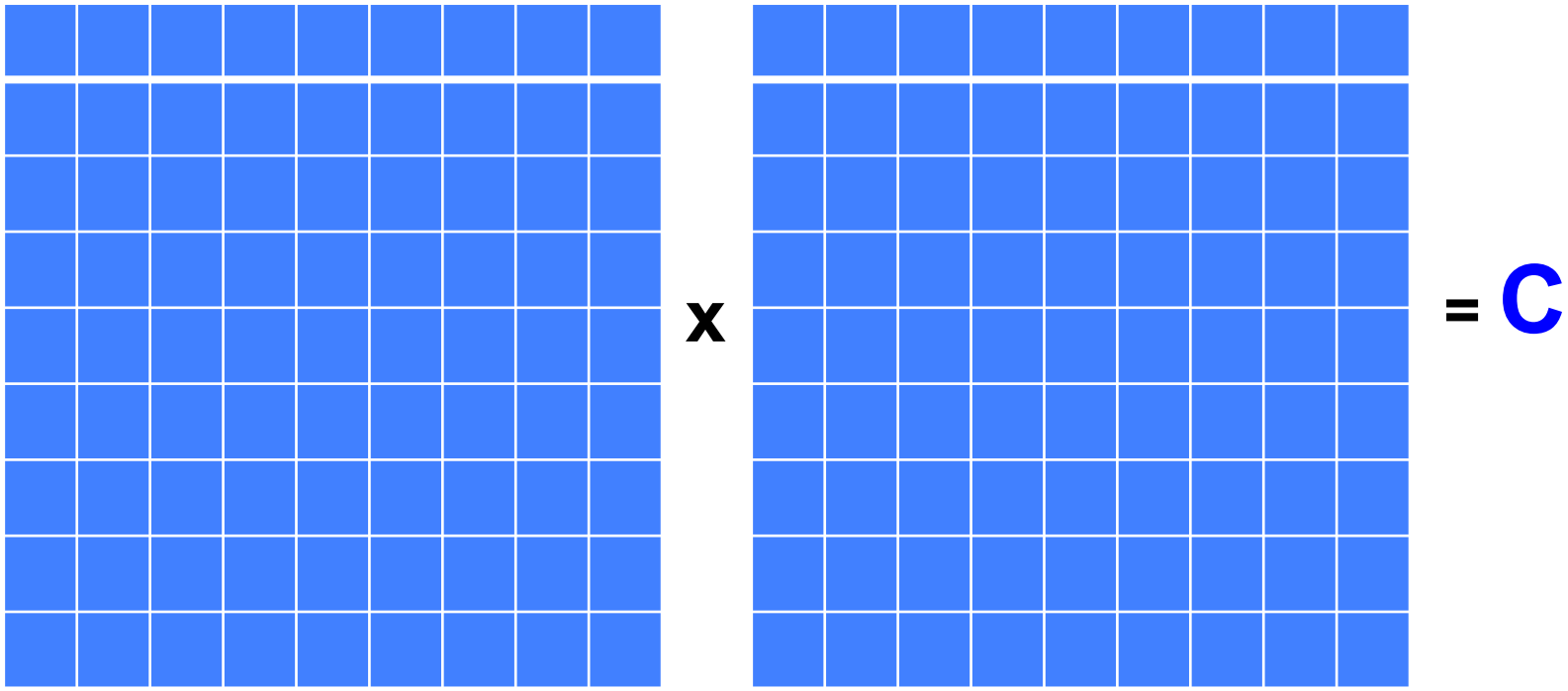
Secure multivariate large-scale multi-centric analysis through on-line learning: an imaging genetics case study

Marco Lorenzi[a], Boris Gutman[b], Paul M. Thompson[b], Daniel C. Alexander[a], Sebastien Ourselin[a], and Andre Altmann[a]

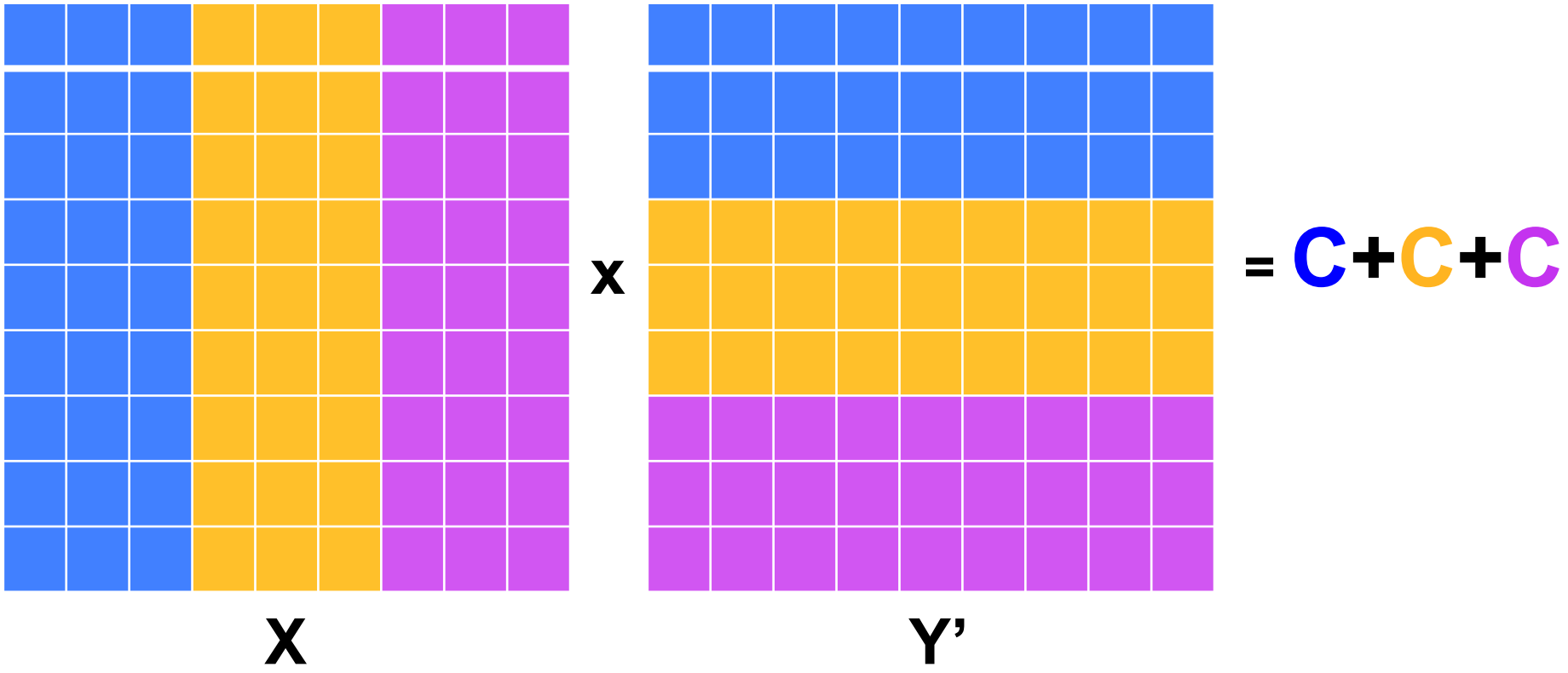
[SIPAIM'16]



- Distributed PLS:



- Distributed PLS:



- Instead of **C** we can share the decomposition of $\mathbf{C} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}$ with a reduced set of components

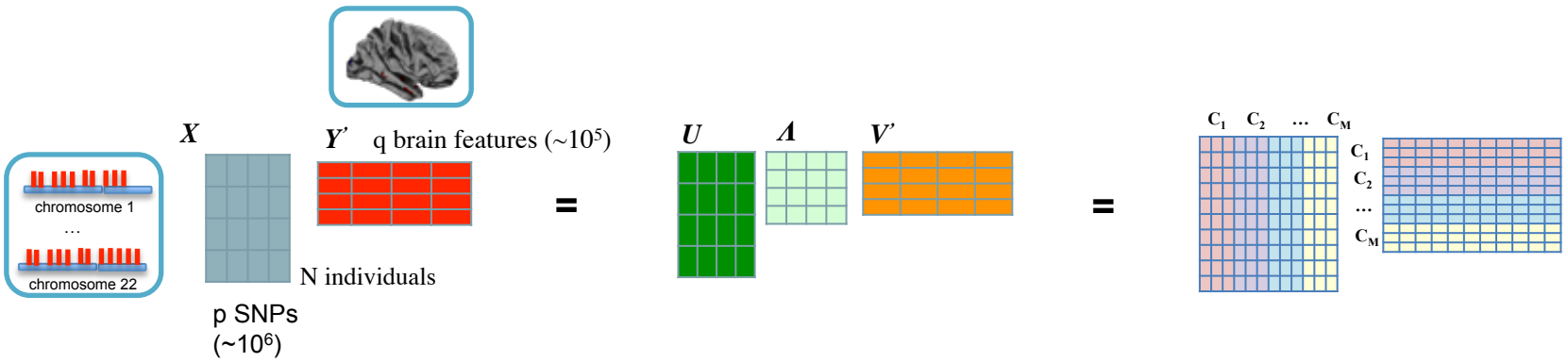
Meta PLS

- We compute the final result as PLS of

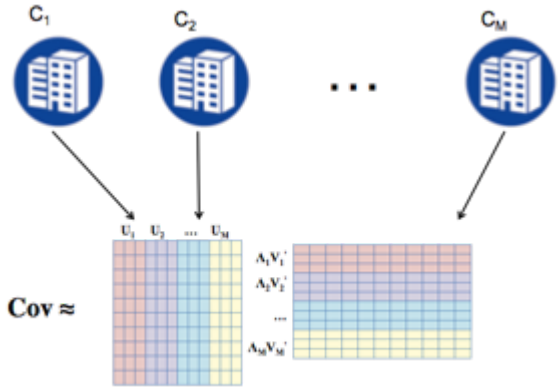
$$\tilde{\mathbf{X}} = [\mathbf{U}_1, \ldots, \mathbf{U}_d]$$
$$\tilde{\mathbf{Y}} = [\mathbf{V}_1\mathbf{\Lambda}_1, \ldots, \mathbf{V}_d\mathbf{\Lambda}_d]$$

- Approximation
- Depends on number of shared components
- Compare full PLS vs meta PLS (2-split)
  - 'component similarity': $\mathbf{u}_{full}$. $\mathbf{u'}_{meta}$
  - Feature-wise error: $\mathbf{u}_{full}$ - $\mathbf{u'}_{meta}$

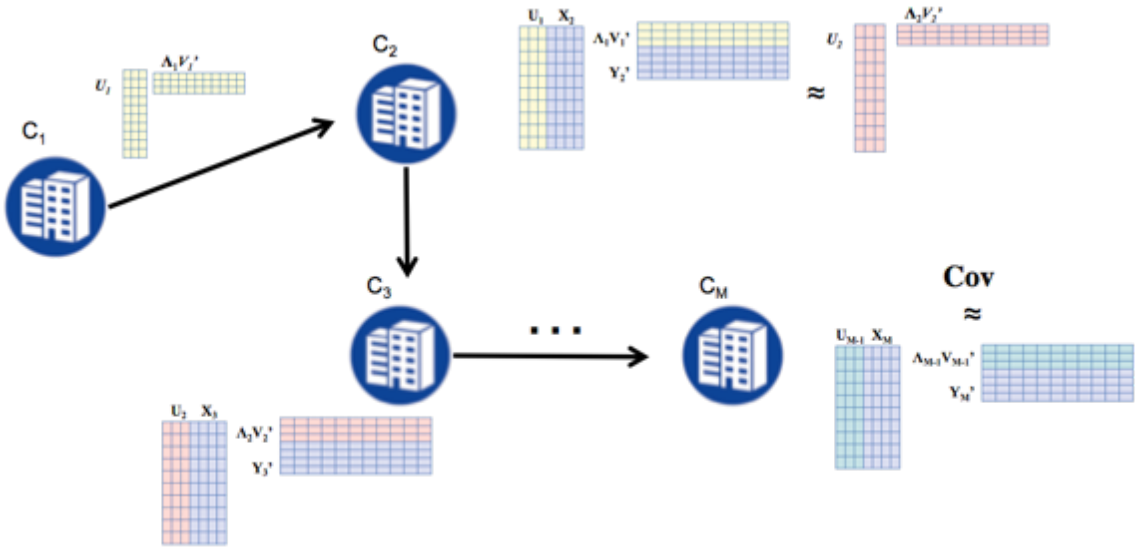
SVD - Partial least squares



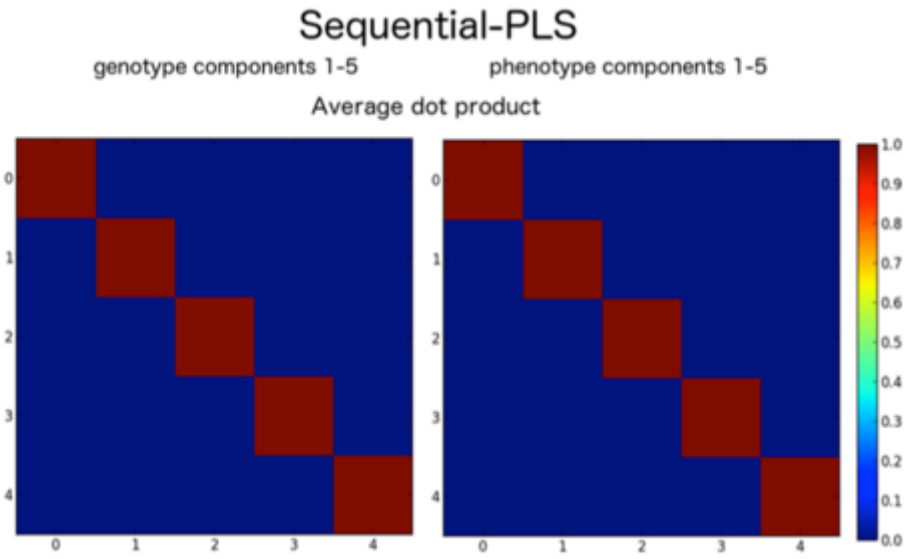
Meta PLS

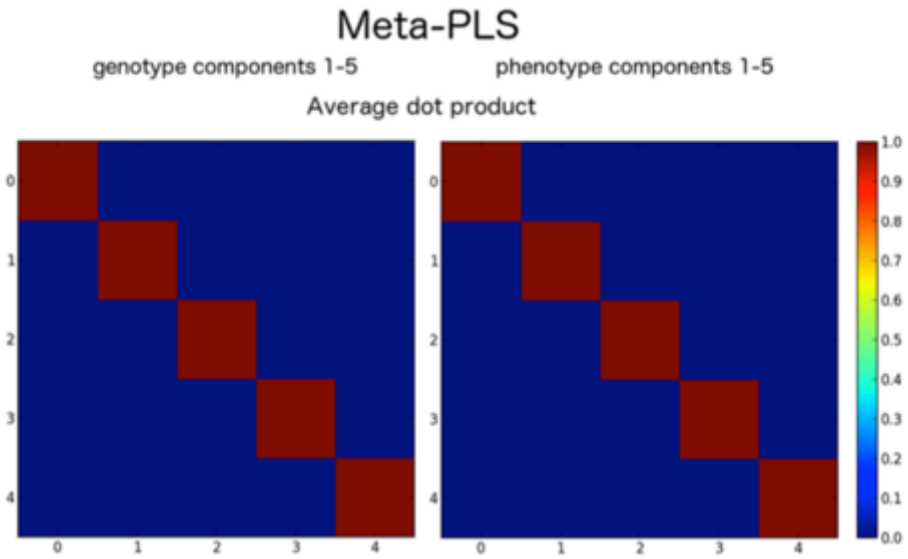


Sequential PLS



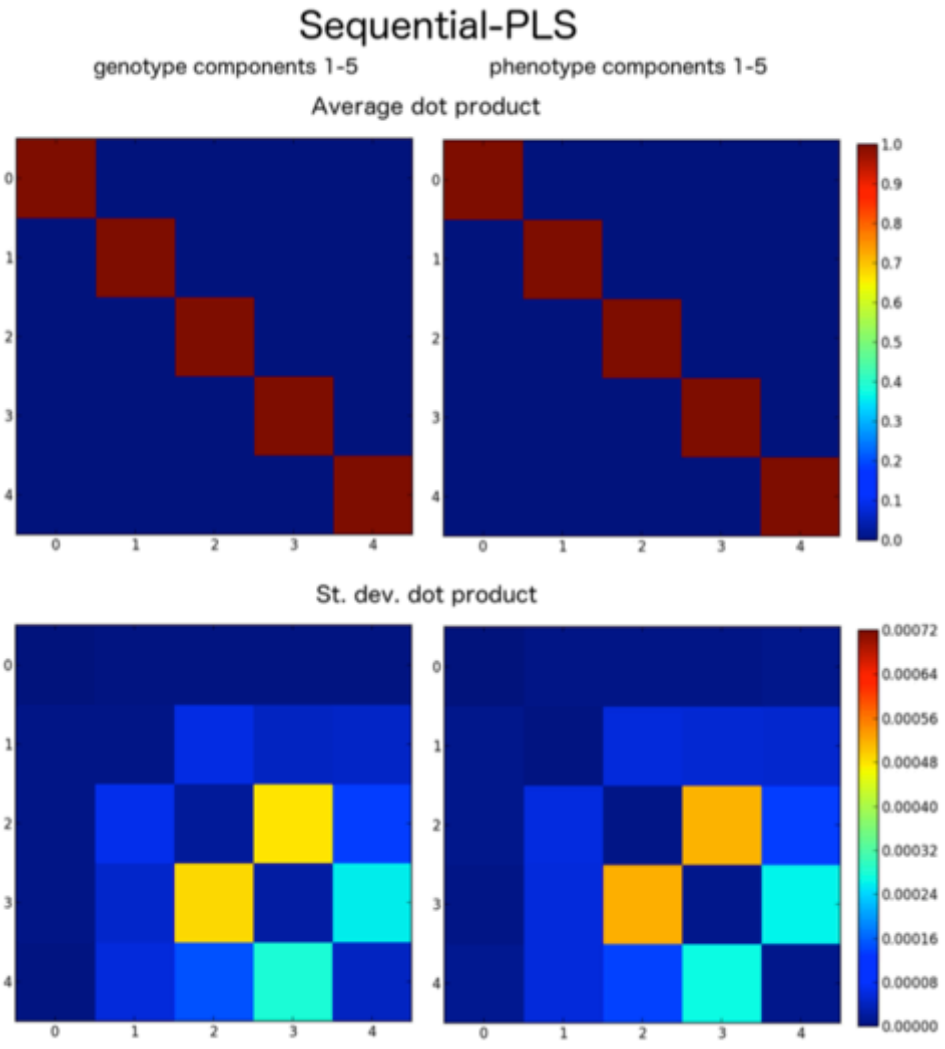
Meta-PLS vs Seq-PLS

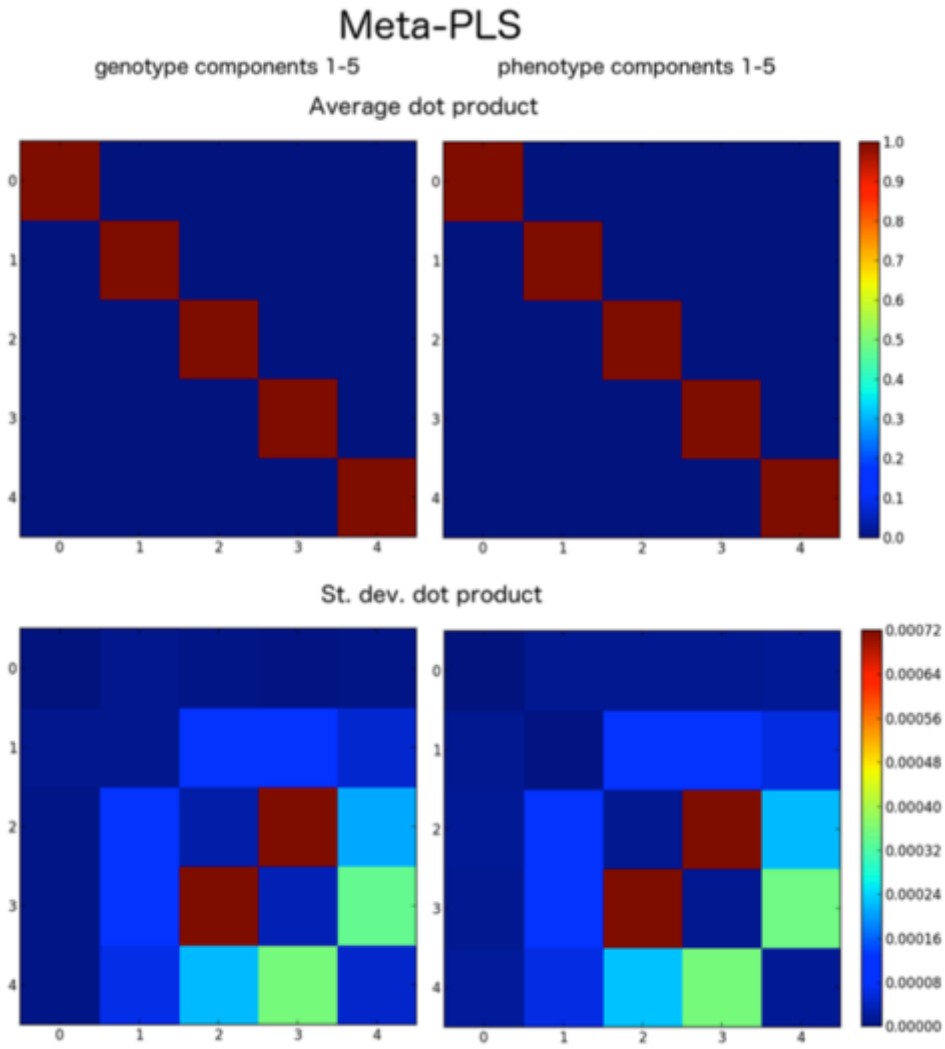




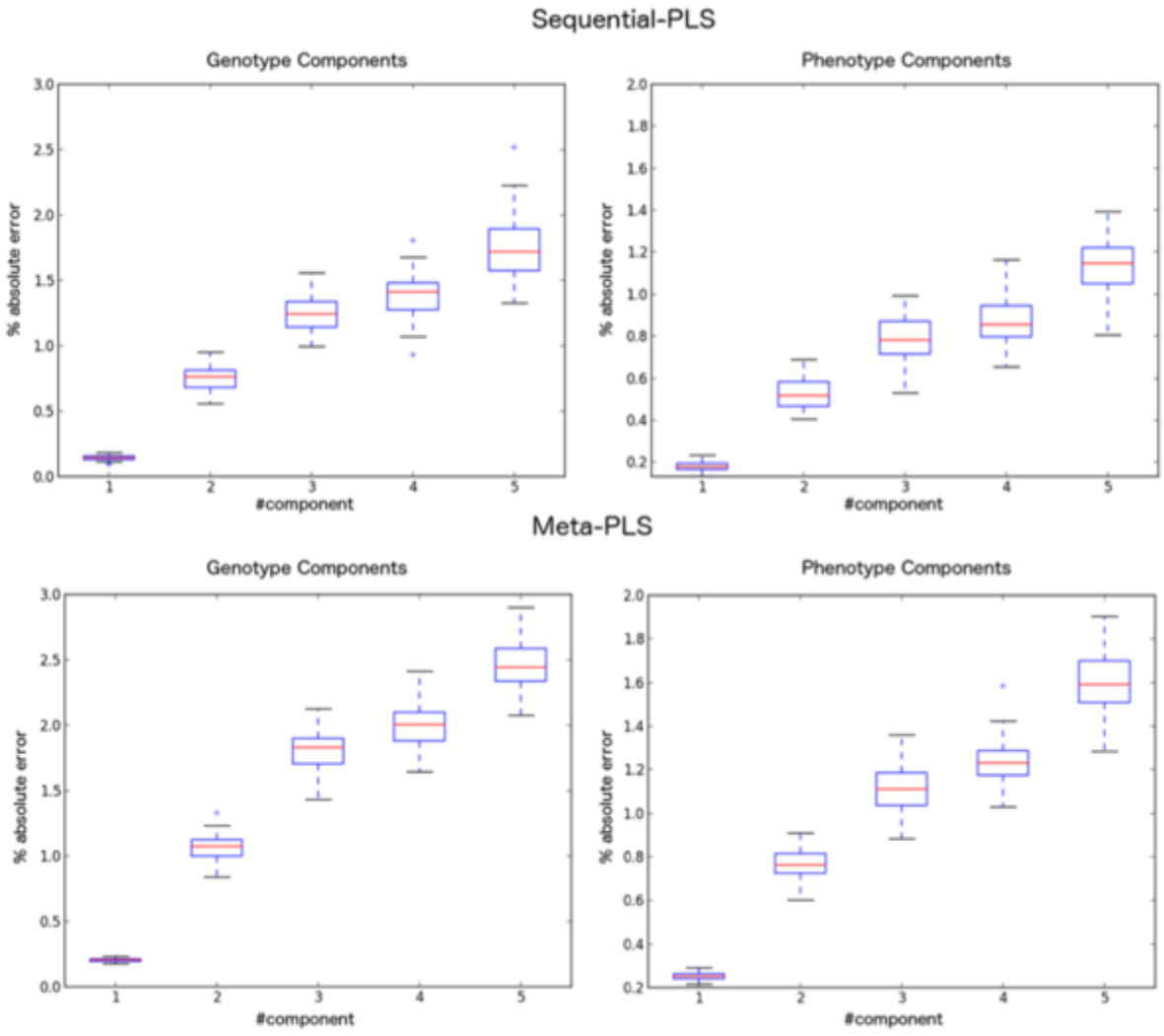
- 50 repetitions to compute mean and sd
- Shared components explain 90% of variability

Meta-PLS vs Seq-PLS

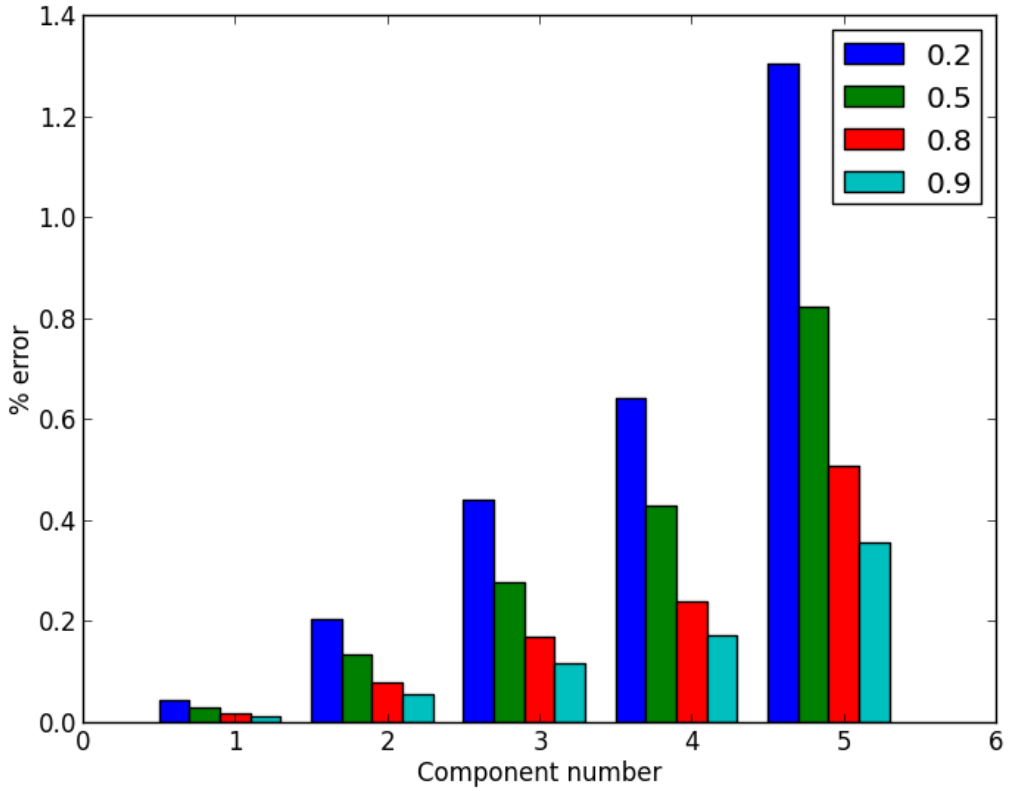




Meta-PLS vs Seq-PLS



Number of shared components



Practical considerations

- X and Y require centering and standardization before SVD
  - How best done in a distributed setting?
  - Effect of ‘late comers’
- Exploring real world application within

- Meta PLS allows
  - Processing large datasets with standard hardware
  - Processing large datasets across different sites

Acknowledgements

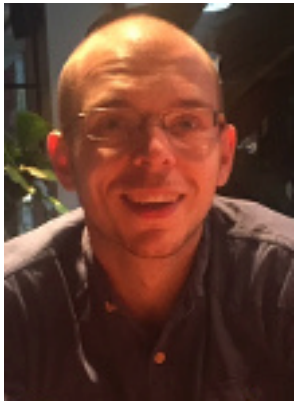
UCL

TRANSLATIONAL IMAGING GROUP

Sebastien Ourselin

DRC
DEMENTIA RESEARCH CENTRE

Nick C Fox
Jonathan Schott
Jonathan Rohrer



Marco Lorenzi

USC University of Southern California

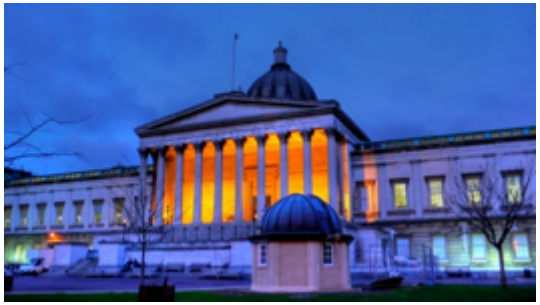
Paul M Thompson
Boris Gutman

Inria

cmic
Centre for Medical Image Computing

eMedLab

Partial Least Squares for large distributed datasets in imaging genetics



Andre Altmann

Department of Medical Physics and Biomedical Engineering

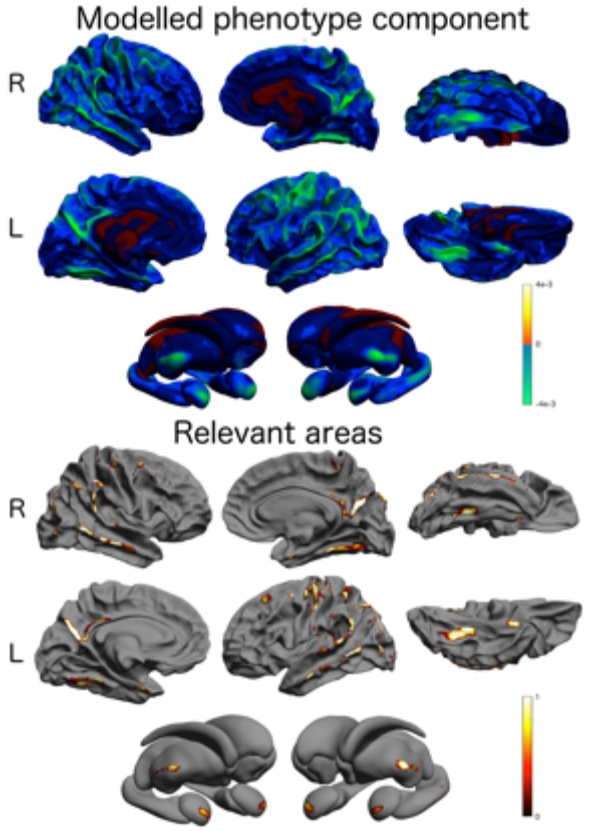
Translational Imaging Group (TIG)

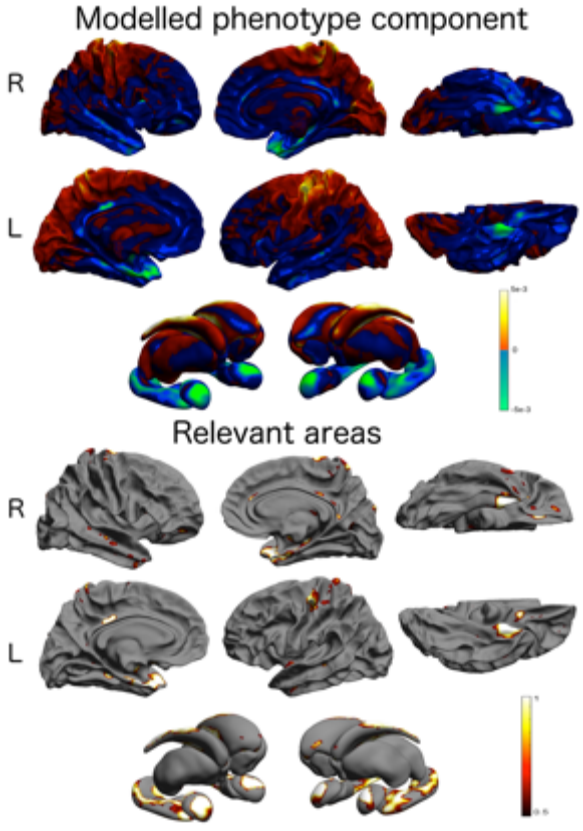
Centre for Medical Image Computing (CMIC)

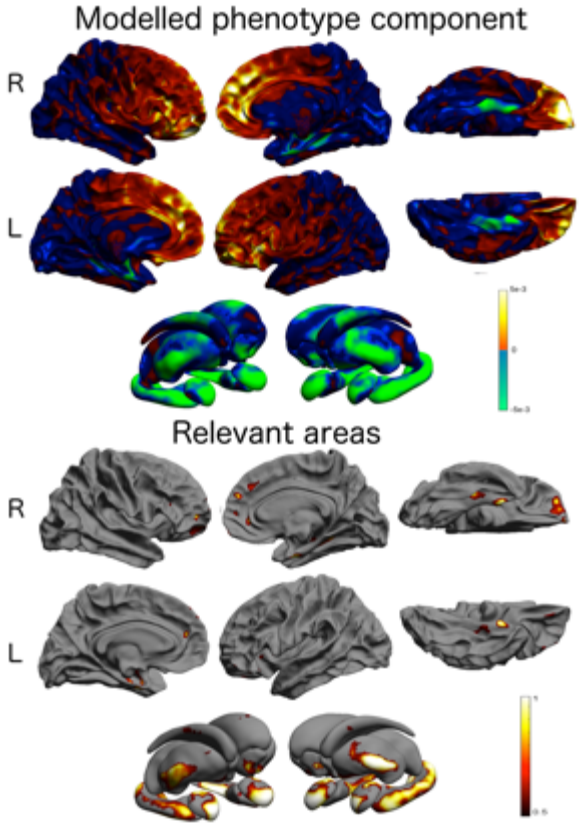
a.altmann@ucl.ac.uk

Predictive radiology for precision medicine
Lausanne, November 13th 2017

Results – Imaging







- PLS is CCA with “infinite regularization (L_2)”

Eigenproblems in Pattern Recognition

Tijl De Bie[1], Nello Cristianini[2], and Roman Rosipal[3]

rCCA:

$$\begin{pmatrix} 0 & \mathbf{S_{XY}} \\ \mathbf{S_{YX}} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{w_X} \\ \mathbf{w_Y} \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{S_{XX}} + \gamma \mathbf{I} & 0 \\ 0 & \mathbf{S_{YY}} + \gamma \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{w_X} \\ \mathbf{w_Y} \end{pmatrix}.$$

PLS:

$$\begin{pmatrix} 0 & \mathbf{S_{XY}} \\ \mathbf{S_{YX}} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{w_X} \\ \mathbf{w_Y} \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{w_X} \\ \mathbf{w_Y} \end{pmatrix}.$$