

ML in healthcare

Christian Salvatore
Scuola Universitaria Superiore IUSS Pavia

christian.salvatore@iusspavia.it

Genomics can predict subtypes

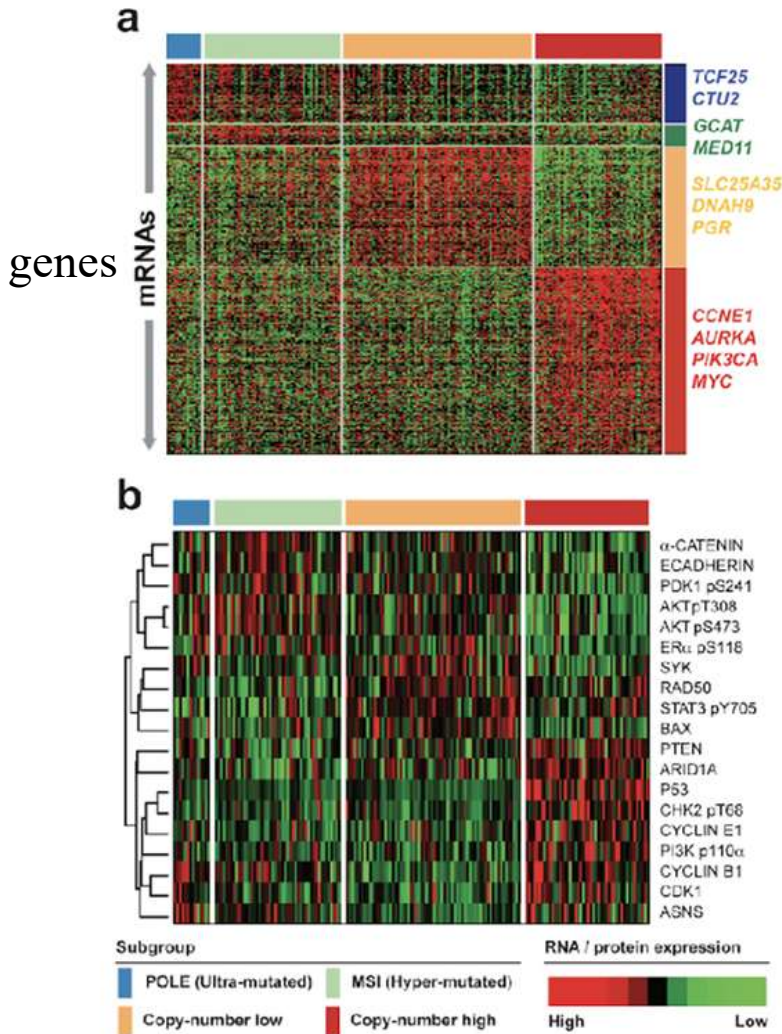
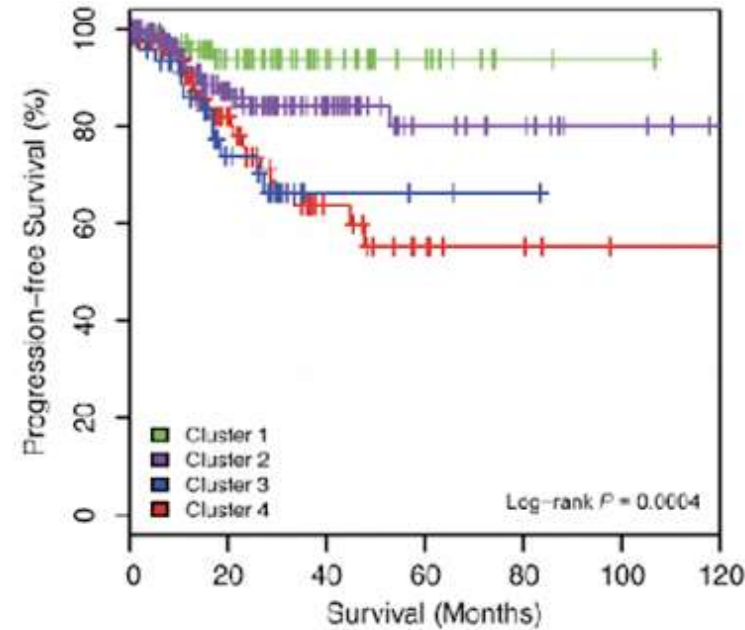


HHS Public Access
Author manuscript
Nature. Author manuscript; available in PMC 2013 November 02.

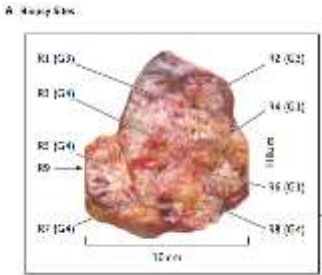
Published in final edited form as:
Nature. 2013 May 2; 497(7447): 67–73. doi:10.1038/nature12113.

Integrated Genomic Characterization of Endometrial Carcinoma

The Cancer Genome Atlas Research Network

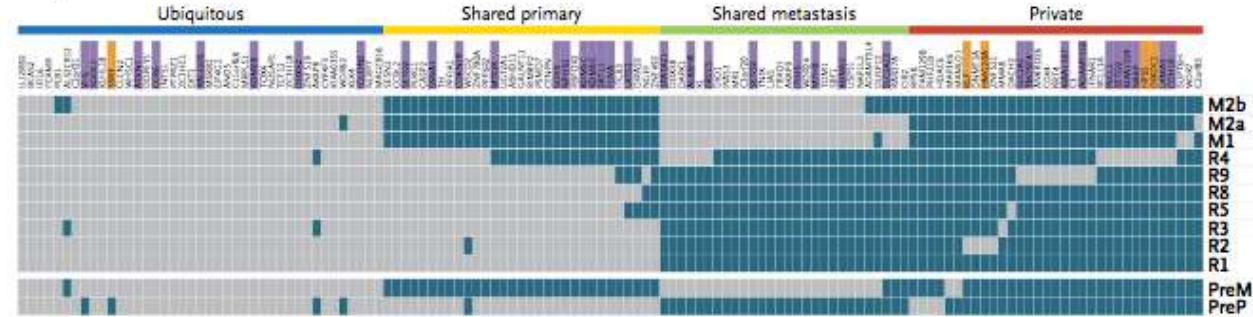


However... biopsy genomics can fail for INTER/INTRA-TUMOUR HETEROGENEITY



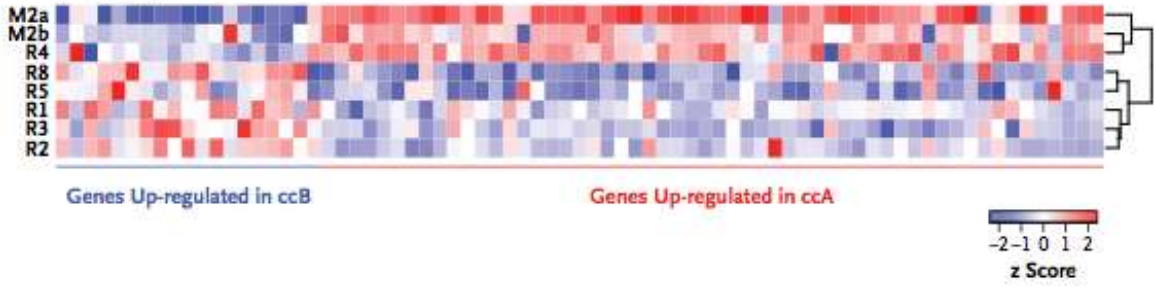
multiple spatially separated samples obtained from 4 primary renal carcinomas and associated metastatic sites

B Regional Distribution of Mutations



Gray = presence of a mutation
Dark blu = absence

C Prognostic Signature Genes



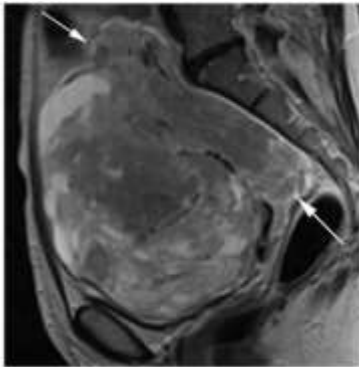
2 molecular subgroups:
clear-cell A (good prognosis)
clear-cell B (poor prognosis)

Can in vivo imaging
capture INTER/INTRA-LESION subtype heterogeneity
(US, CT, MRI, PET-CT)?

Can in vivo imaging of the lesion
be used as a digital biobank?

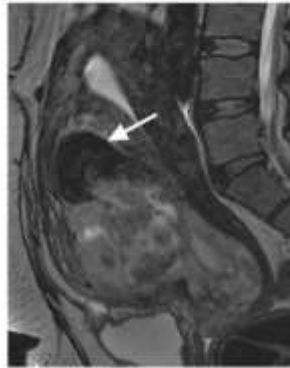
Image features (semantic)

a large uterine mass with nodular superior and posterior borders



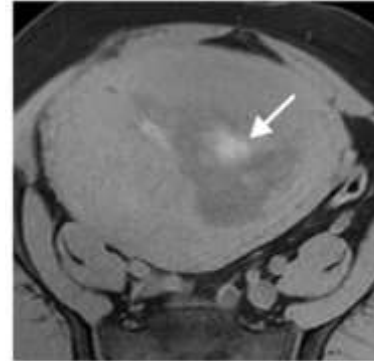
MRI T2-weighted

dark area in the myometrial mass



MRI T2-weighted

intra-lesional haemorrhage



Non contrast-enhanced MRI T1-weighted fat saturated

central unenhanced areas



Contrast-enhanced MRI T1-weighted fat saturated

Image features and genomics

IF 68.84

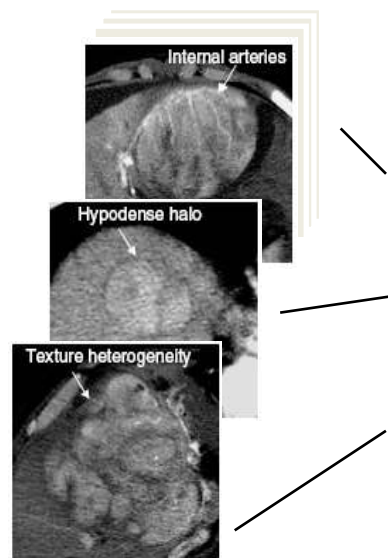
NATURE BIOTECHNOLOGY VOLUME 25 NUMBER 6 JUNE 2007

Liver cancer

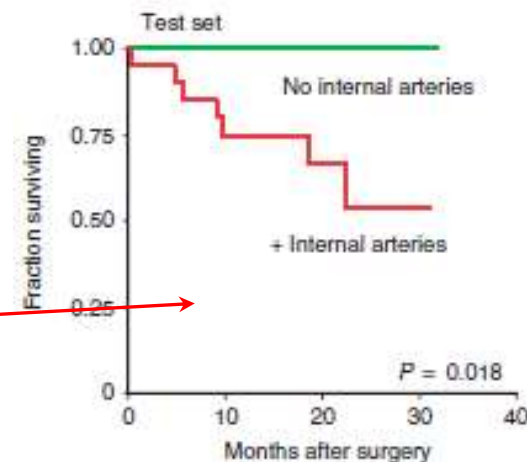
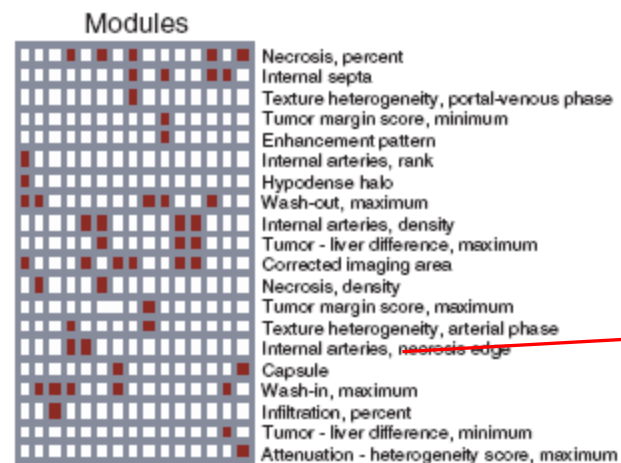
Decoding global gene expression programs in liver cancer by noninvasive imaging

Eran Segal¹, Claude B Sirlin², Clara Ooi¹, Adam S Adler⁵, Jeremy Gollub⁶, Xin Chen⁸, Bryan K Chan², George R Matcuk⁷, Christopher T Barry³, Howard Y Chang⁵ & Michael D Kuo²

32 CT imaging semantic features



116 gene "modules"
with coherent patterns of variations



Red=presence

Green=absence

IF 11.20



PNAS | April 8, 2008 | vol. 105 | no. 13 | 5213–5218

Brain cancer

Identification of noninvasive imaging surrogates for brain tumor gene-expression modules

Maximilian Diehn^{*†}, Christine Nardini^{*}, David S. Wang^{*}, Susan McGovern[‡], Mahesh Jayaraman[§], Yu Liang[¶], Kenneth Aldape[‡], Soonmee Cha^{||}, and Michael D. Kuo^{*,**††}

7 gene "modules"

with coherent patterns of variations

10 MRI imaging semantic features

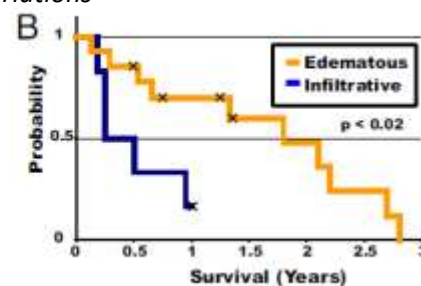
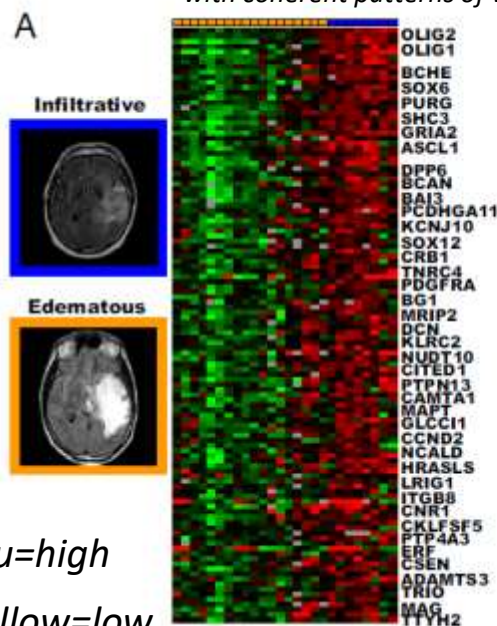


Image features and genomics

Radiology

Non-Small Cell Lung Cancer: Identifying Prognostic Imaging Biomarkers by Leveraging Public Gene Expression Microarray Data—Methods and Preliminary Results¹

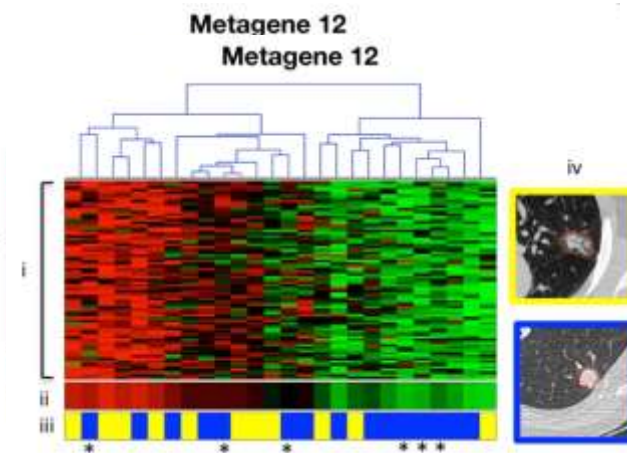
Oliver Gevaert, PhD
Jiajing Xu, MS
Chuong D. Hoang, MD
Ann N. Laung, MD
Yue Xu, PhD
Andrew Quon, MD
Daniel L. Rubin, MD, MS
Sandy Napel, PhD
Sylvia K. Plevritis, PhD

Purpose: To identify prognostic imaging biomarkers in non-small cell lung cancer (NSCLC) by means of a radiogenomics strategy that integrates gene expression and medical images in patients for whom survival outcomes are not available by leveraging survival data in public gene expression data sets.

Materials and Methods: A radiogenomics strategy for associating image features with clusters of coexpressed genes (metagenes) was defined. First, a radiogenomics correlation map is created

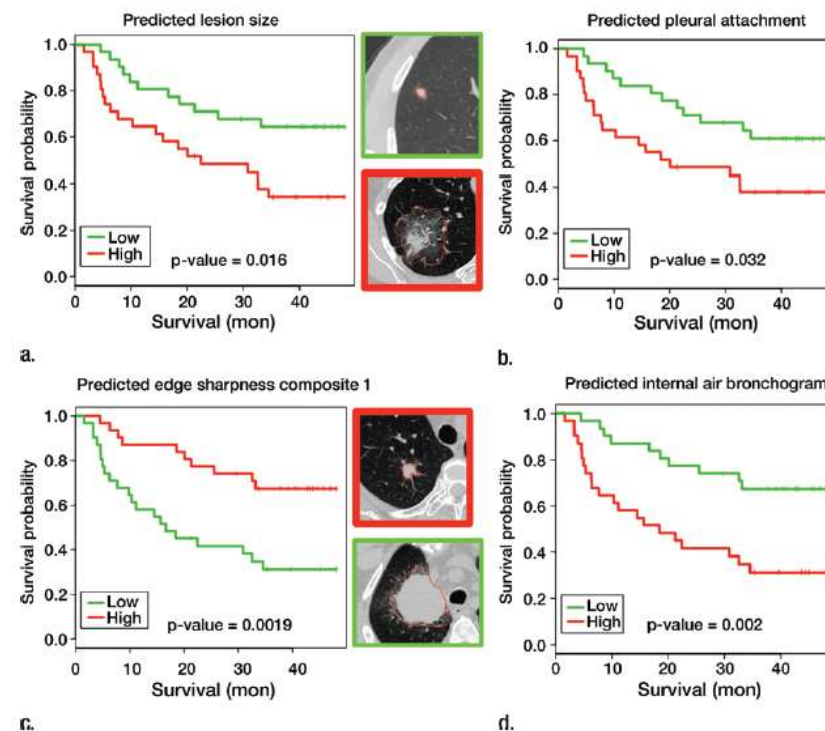
Radiology: Volume 264: Number 2—August 2012 ■ radiology.rsna.org

153 CT features

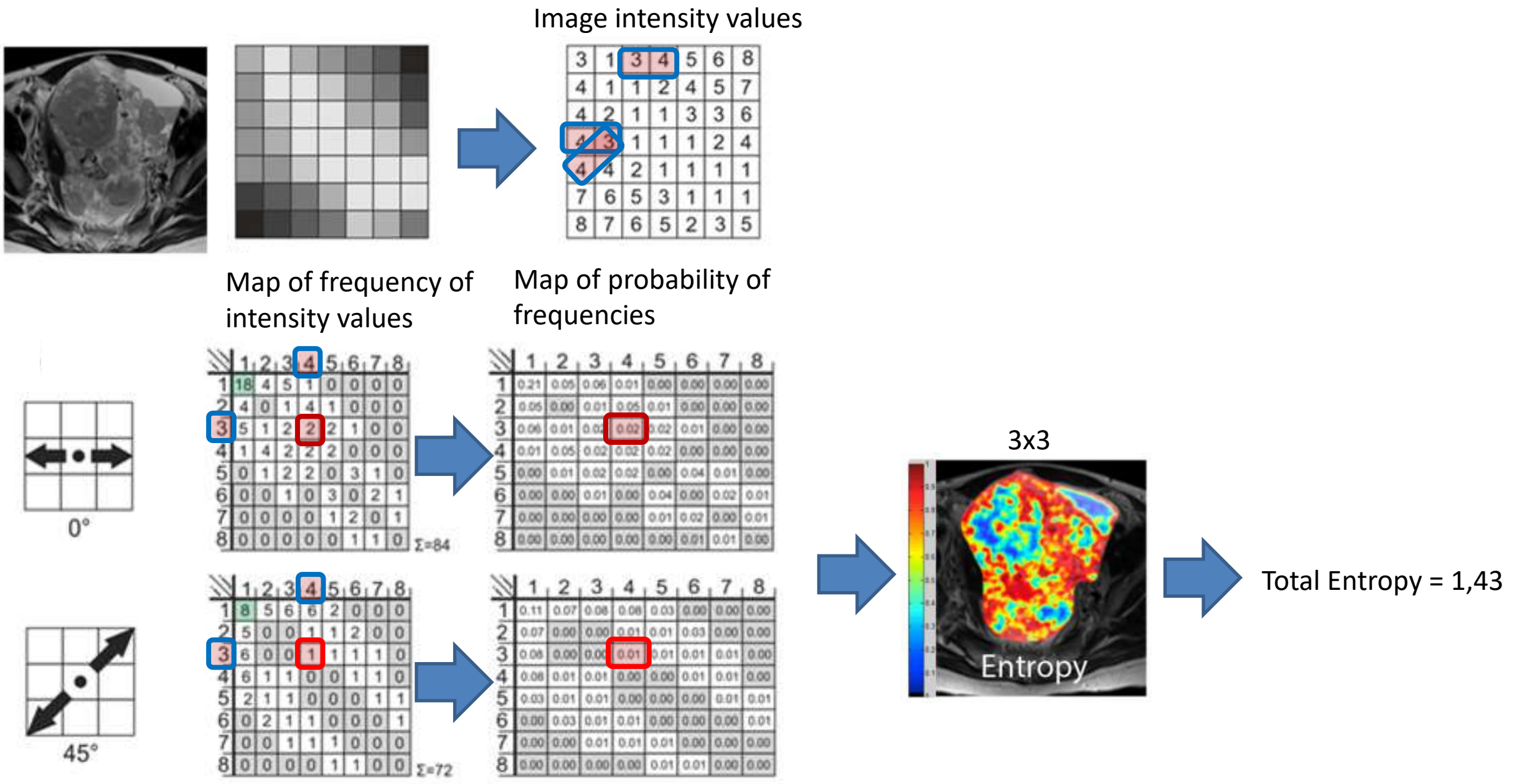


Non-Small Cell lung cancer

4 CT image features correlated with metagenes



Statistical image features computation example for RadiOmics



Radiomics: a new approach for the study of cancer



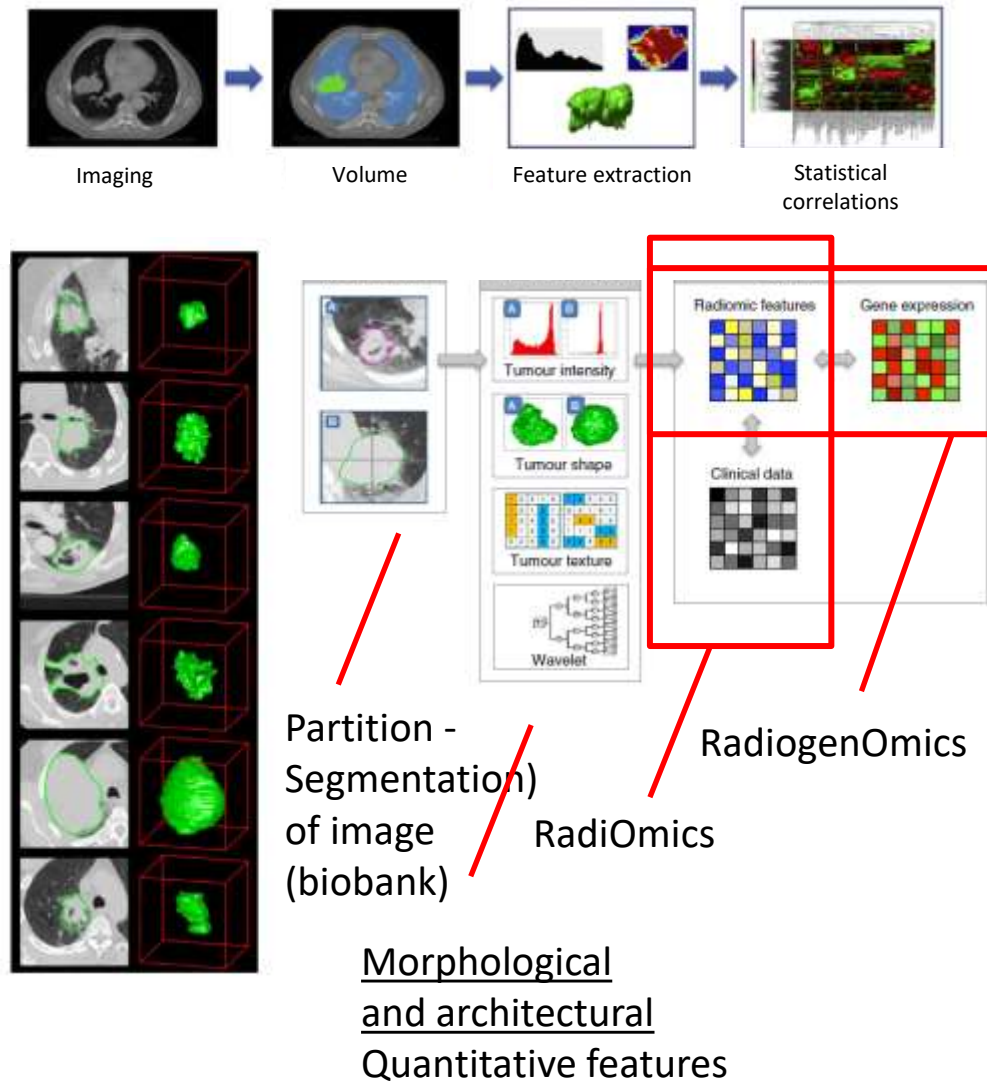
Published in final edited form as:
Eur J Cancer. 2012 March 1;48(4):441-446. doi:10.1016/j.ejca.2011.11.036.

Radiomics: Extracting more information from medical images using advanced feature analysis

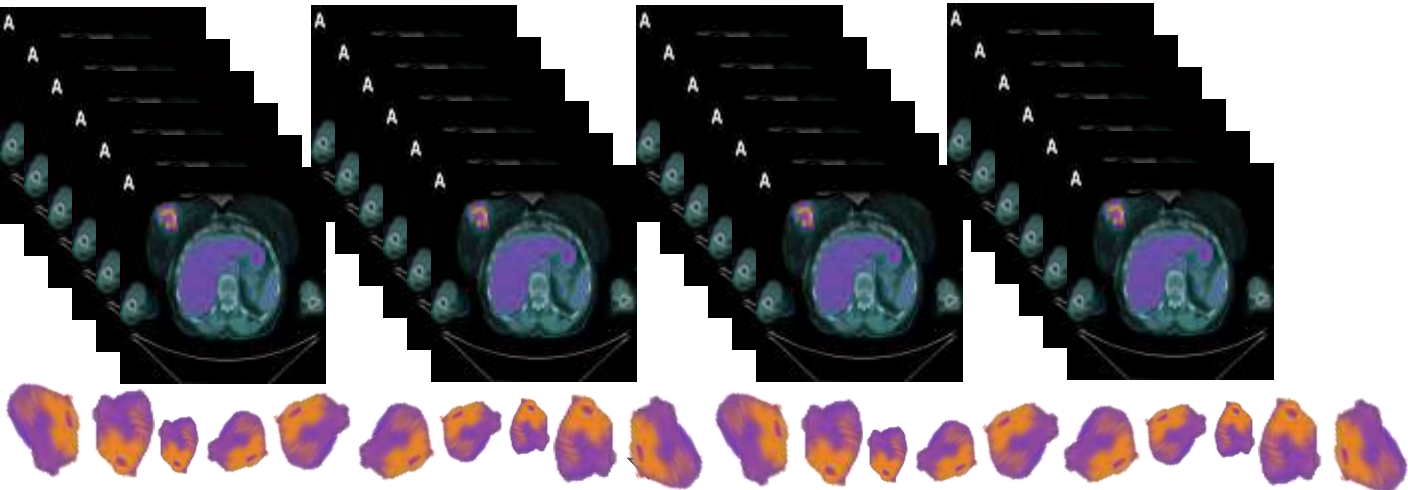
Philippe Lambin^{a,*}, Emmanuel Rios-Velazquez^{a,b}, Ralph Leijenaar^{a,c}, Sara Carvalho^{a,c}, Ruud G.P.M. van Stiphout^{a,d}, Patrick Granton^{a,e}, Catharina M.L. Zegers^{a,f}, Robert Gillies^{b,g}, Ronald Boellard^{c,h}, André Dekker^{a,i}, and Hugo J.W.L. Aerts^{a,d,e}

^aDepartment of Radiation Oncology (MAASTRO), GROW – School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, The Netherlands ^bH. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, USA ^cU University Medical Center, Department of Nuclear Medicine & PET Research, Amsterdam, The Netherlands ^dComputational Biology and Functional Genomics Laboratory, Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Harvard School of Public Health, USA

Comprehensive quantification of disease phenotypes by applying a large number of quantitative image features representing lesion heterogeneity and correlating with omics and clinical data

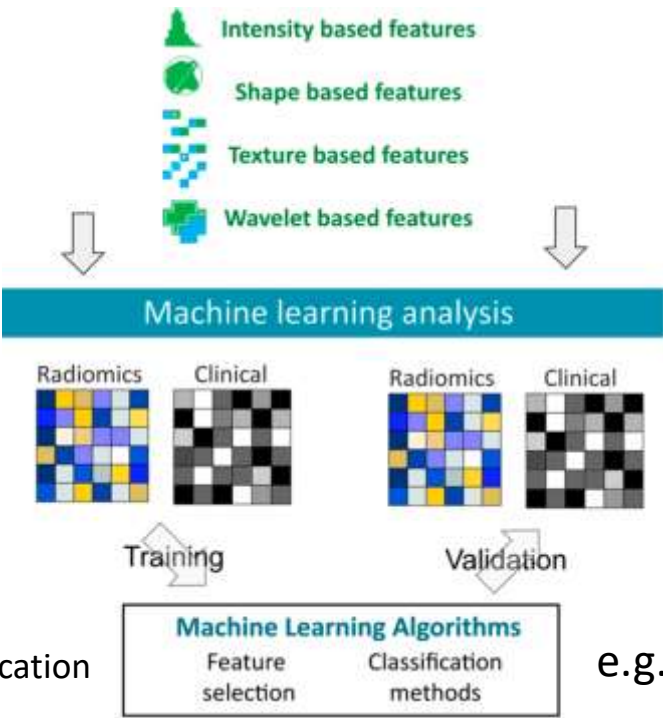


Predictive models based on radiomic signatures



Partitions
of image
(biobanks)

To automatically predict clinical outcome at the level of single patient by training and validating machine learning models with radiOmic features



*St Gallen classification

e.g.. Accuracy >0.85

PET radiomic predicts breast cancer subtypes

**patients with breast cancer prior biopsy*

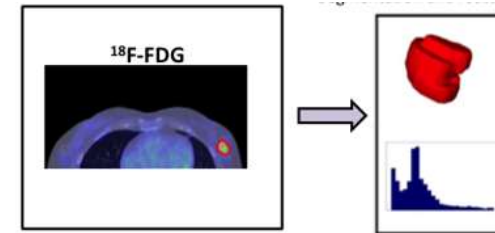
Eur J Nucl Med Mol Imaging
DOI 10.1007/s00259-017-3770-9



ORIGINAL ARTICLE

[¹⁸F]FDG PET/CT features for the molecular characterization of primary breast tumors

Lidija Antunovic¹ · Francesca Gallivanone² · Martina Sollini³ · Andrea Sagona⁴ ·
Alessandra Invento⁵ · Giulia Manfrinato⁶ · Margarita Kirienko³ · Corrado Tinterri⁴ ·
Arturo Chiti^{1,3} · Isabella Castiglioni²

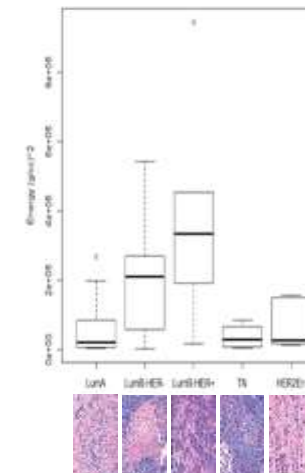
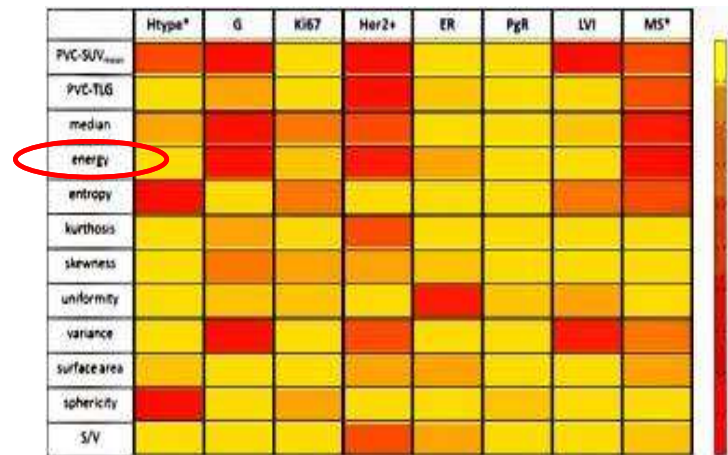


PET features

PVC-SUV_{mean}
PVC-TLG
Median
Energy
Entropy
Kurtosis
Skewness
Uniformity
Variance
Surface Area
Sphericity
Surface to
Volume ratio
.....

PET radiomic features vs himmunochemistry molecular factors

and molecular subtypes (5)*



*St Gallen
classification

MRI-DWI radiomic predicts breast cancer treatment response

Magn Reson Mater Phys (2017) 30:359–373
DOI 10.1007/s10334-017-0610-7

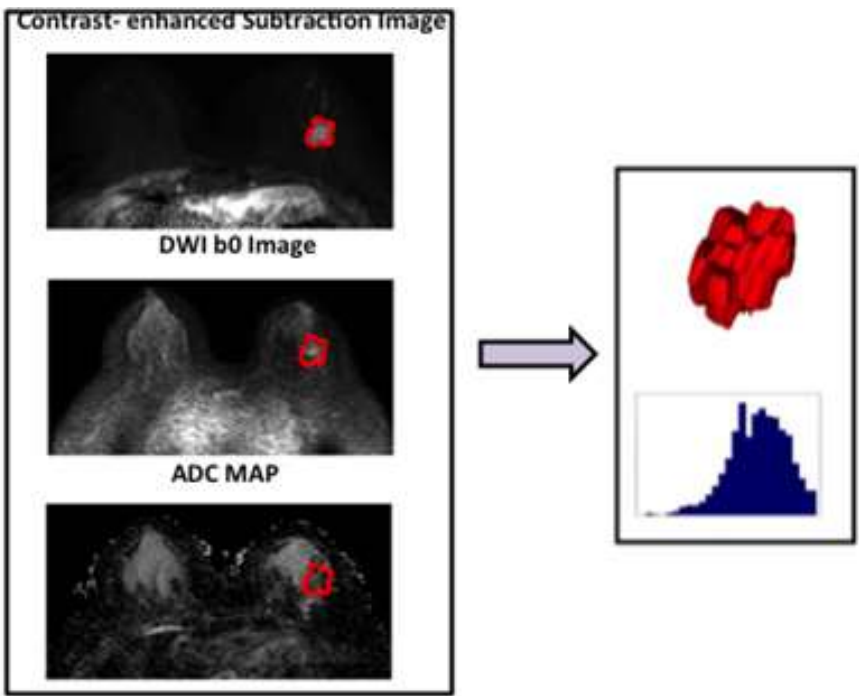


RESEARCH ARTICLE

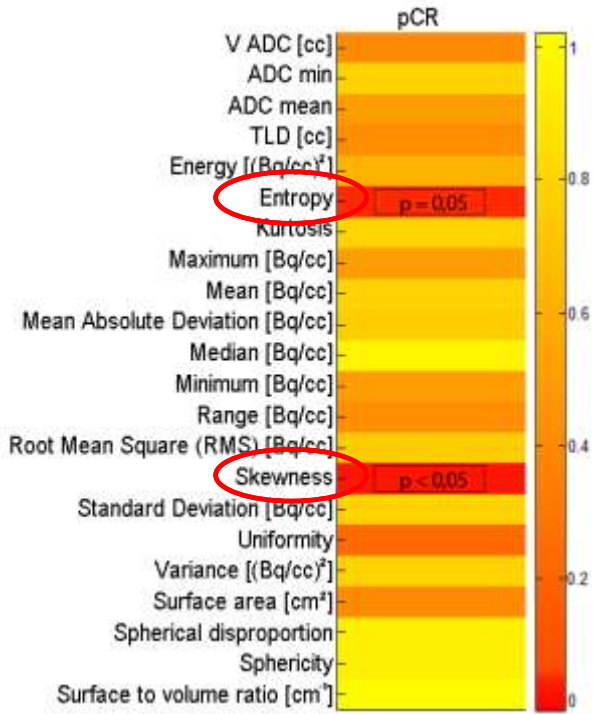
Biomarkers from in vivo molecular imaging of breast cancer: pretreatment ¹⁸F-FDG PET predicts patient prognosis, and pretreatment DWI-MR predicts response to neoadjuvant chemotherapy

Francesca Gallivanone¹ · Marta Maria Panzeri² · Carla Canevari³ · Claudio Losio² · Luigi Gianolli³ · Francesco De Cobelli^{2,4} · Isabella Castiglioni¹

**38 Breast cancer patients treated with NAC + surgery*



MRI-DWI radiomic features VS NAC pCR



PET radiomic predicts lung cancer treatment

[Frontiers in Bioscience, Landmark, 22, 1713-1723, June 1, 2017]

FDG PET/CT as theranostic imaging in diagnosis of non-small cell lung cancer

Margarita Kirienko¹, Francesca Gallivanone², Martina Sollini¹, Giulia Veronesi³, Emanuele Voulaz³, Lidjia Antunovic⁴, Lorenzo Leonardi⁴, Giorgio Testanera⁴, Isabella Castiglioni², Arturo Chiti^{1,4}

¹Department of Biomedical Sciences, Humanitas University, Rozzano, Milan, Italy, ²IBFM-CNR, Segrate, Milan, Italy, ³Thoracic Surgery, Humanitas Clinical and Research Center, Rozzano, Milan, Italy, ⁴Nuclear Medicine, Humanitas Clinical and Research Center, Rozzano, Milan, Italy

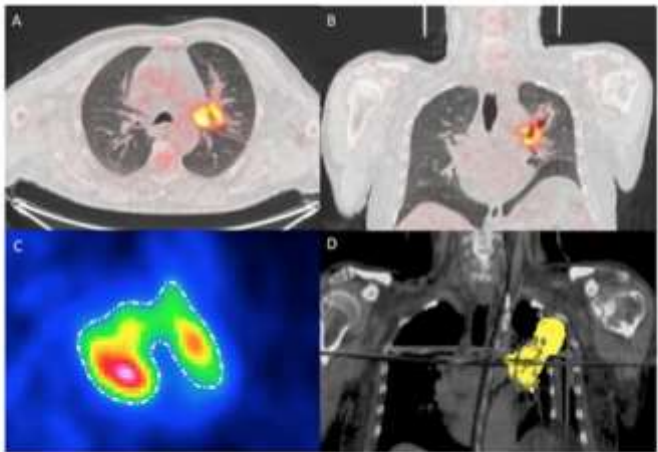
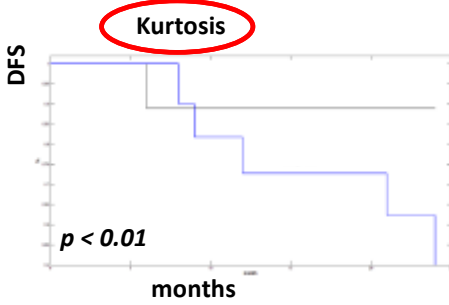
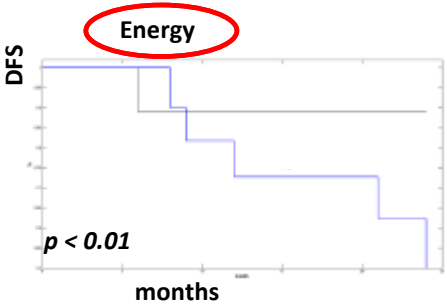


Figure 1. A 67-year-old male with squamous cell carcinoma, G3pT3N1 (stage IIIa) who underwent pneumonectomy and subsequent adjuvant chemotherapy with cisplatin and vinorelbine. After 36 months' follow-up the patient had no evidence of disease. $P_{VC-SUV} = 15.59$ g/cc, $SUV_{max} = 21.18$ g/cc, $Energy = 3.2656$ (g/cc)², $Entropy = 4.08$, $Kurtosis = 3.85$. The axial (A) and coronal (B) fused PET/CT images show the lesion in the upper left lobe; the axial PET (C) image shows the tumour segmentation operated by the algorithm; (D) metabolic tumour volume visualized on CT images.

**Non-Small Cell lung cancer patients treated with surgery + NAC*

Survival endpoint	PET feature	P-value	Cut-off	Sensitivity	Specificity
DSF	Energy[(MBq/cc) ²]	0.01	2.74	87.5%	75%
	Kurtosis	<0.05	2.71	75%	60%
DMFS	PVC-SUV _{mean}	<0.05	11.91	80%	80%
	SUV _{max}	<0.05	14.84	80%	80%
	Energy[(MBq/cc) ²]	<0.05	2.74	100%	80%
	Kurtosis	<0.05	2.9	100%	75%



Decision support systems in clinical medicine

- decision support systems, designed to increase the effectiveness of the medical analysis as it provides support to clinicians who need to make strategic medical decisions in the face of a medical problems that can not be solved with operational research models
- extract from a significant amount of data, in a short time and in a versatile way, new information useful to the clinical decision-making processes



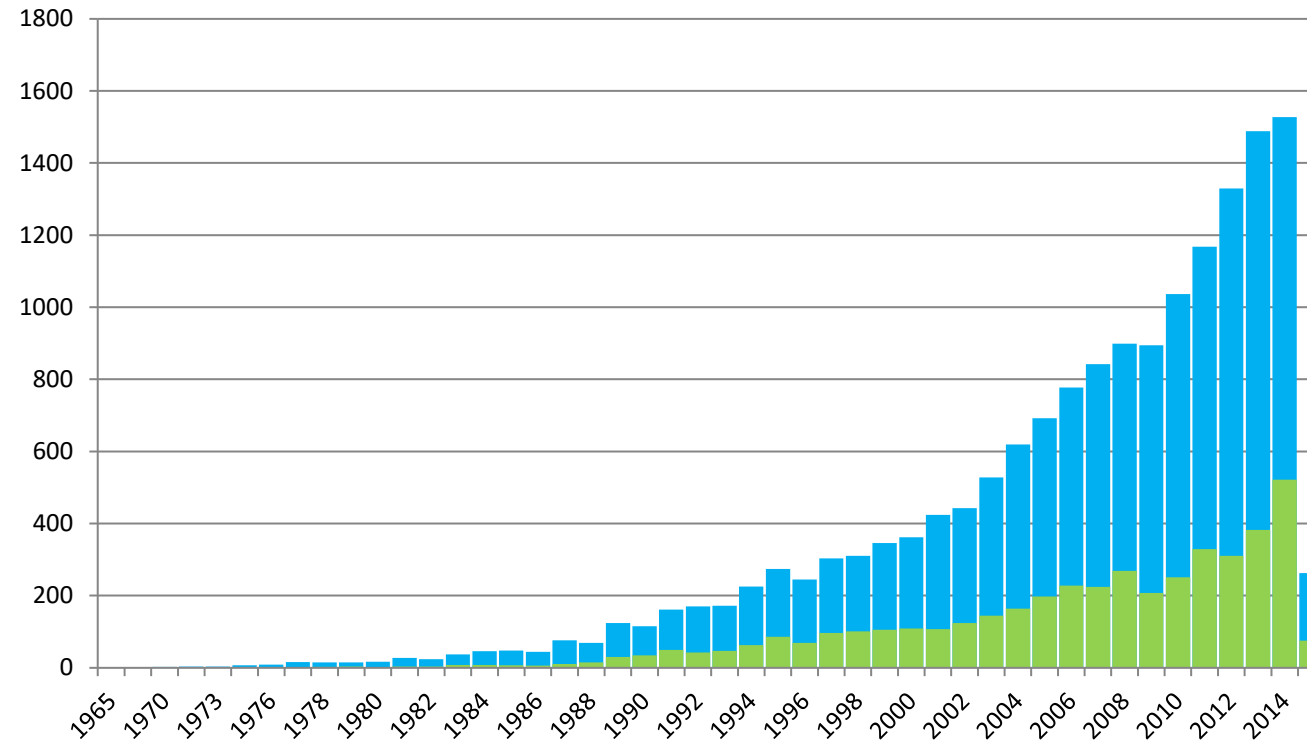
- Assisted Diagnosis
- Objective clinical assessment
- High diagnostic accuracy

In vivo imaging plays a new role in the diagnostic process of many diseases

[e.g. McKhann GM et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia 7(3): 263-69. (2011)]

Decision support systems in clinical medicine

publications per year from 1965 to 2015



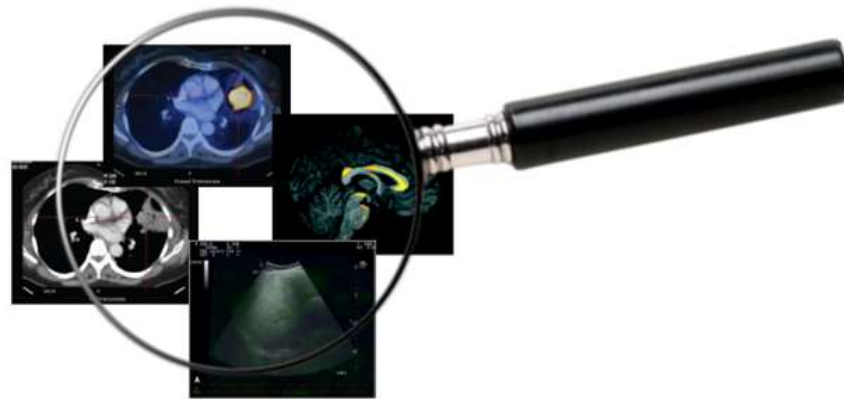
RESEARCH CRITERIA:

- "decision support system"
- "decision support system medicine"

source: pubmed.com

Machine learning applied to medical imaging

. how is this useful in medicine?

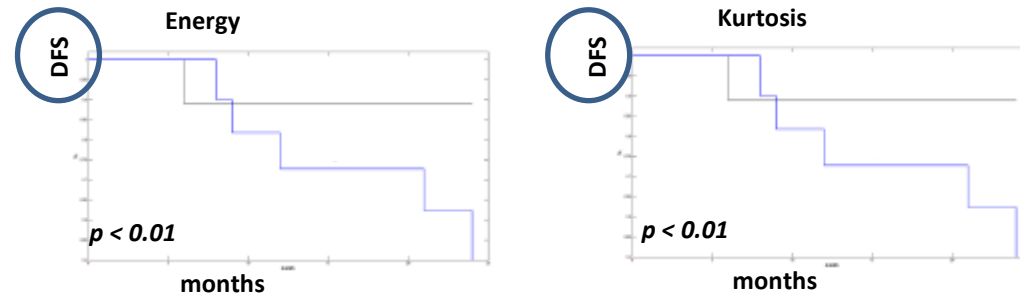


Reference standard for the AI model

- Reference Standard
- Gold Standard
- Ground truth

It defines the clinical condition to be predicted
(i.e. diagnostic subtype, prognostic subtype, treatment
response subtype)!

Reference standard for the training of the AI model



- Clinical follow-up endpoint (disease-free-survival DFS, overall survival OS, ...)
- Histological diagnosis from biopsy
- Histological diagnosis from biopsy and definite surgery
- Molecular diagnosis
- Radiological classification
- Treatment response
-

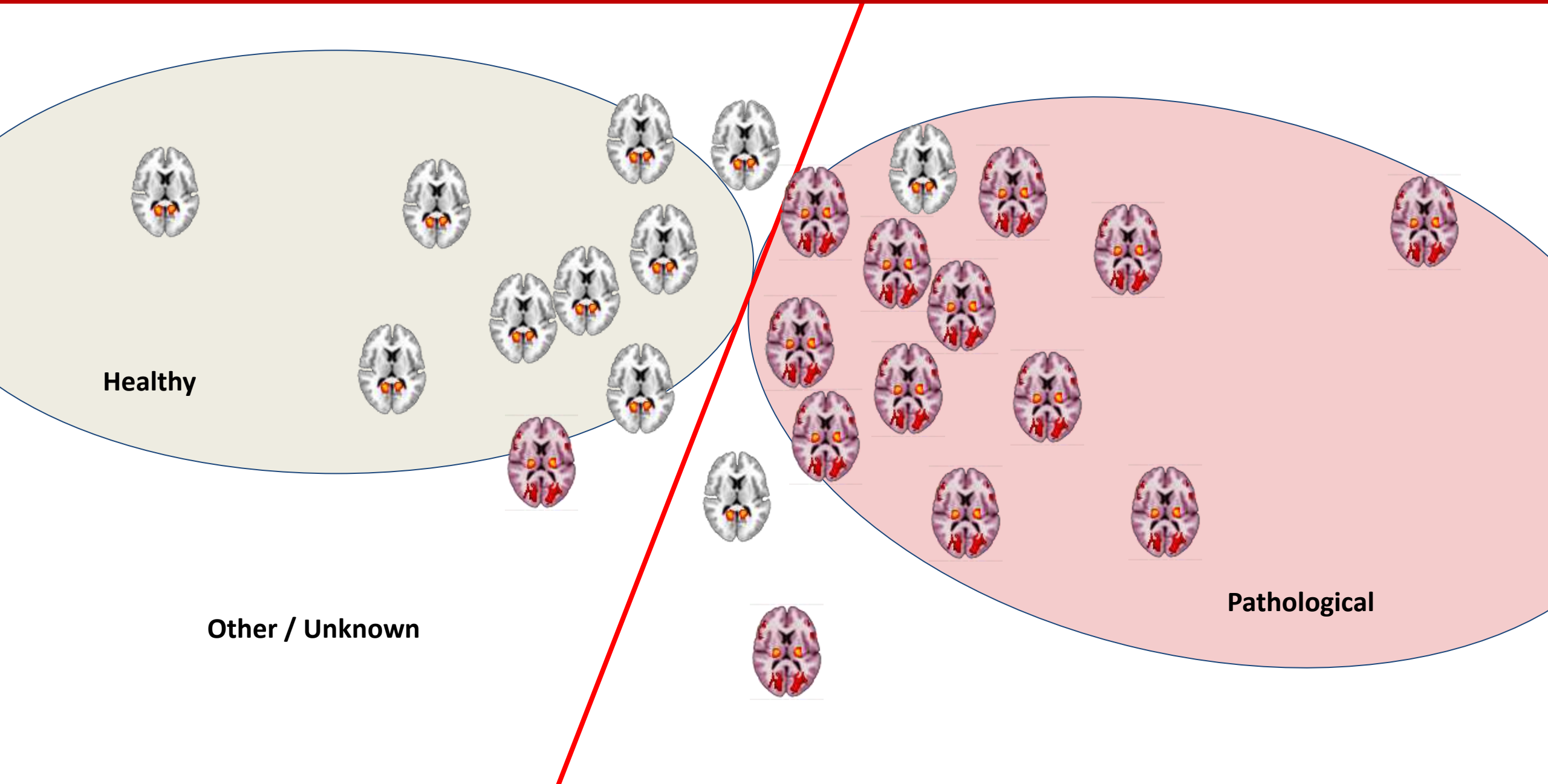
Machine learning applied to medical imaging



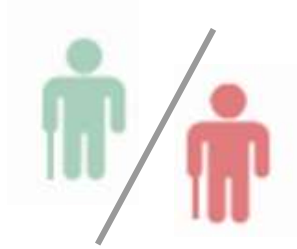
Machine learning applied to medical imaging



Machine learning applied to medical imaging



Machine learning applied to medical imaging



Diagnosis (early/differential)

is the patient healthy?



Prognosis

what will be the course of the disease?



Screening

*are the patients -within
a population- healthy?*



Treatment addressing

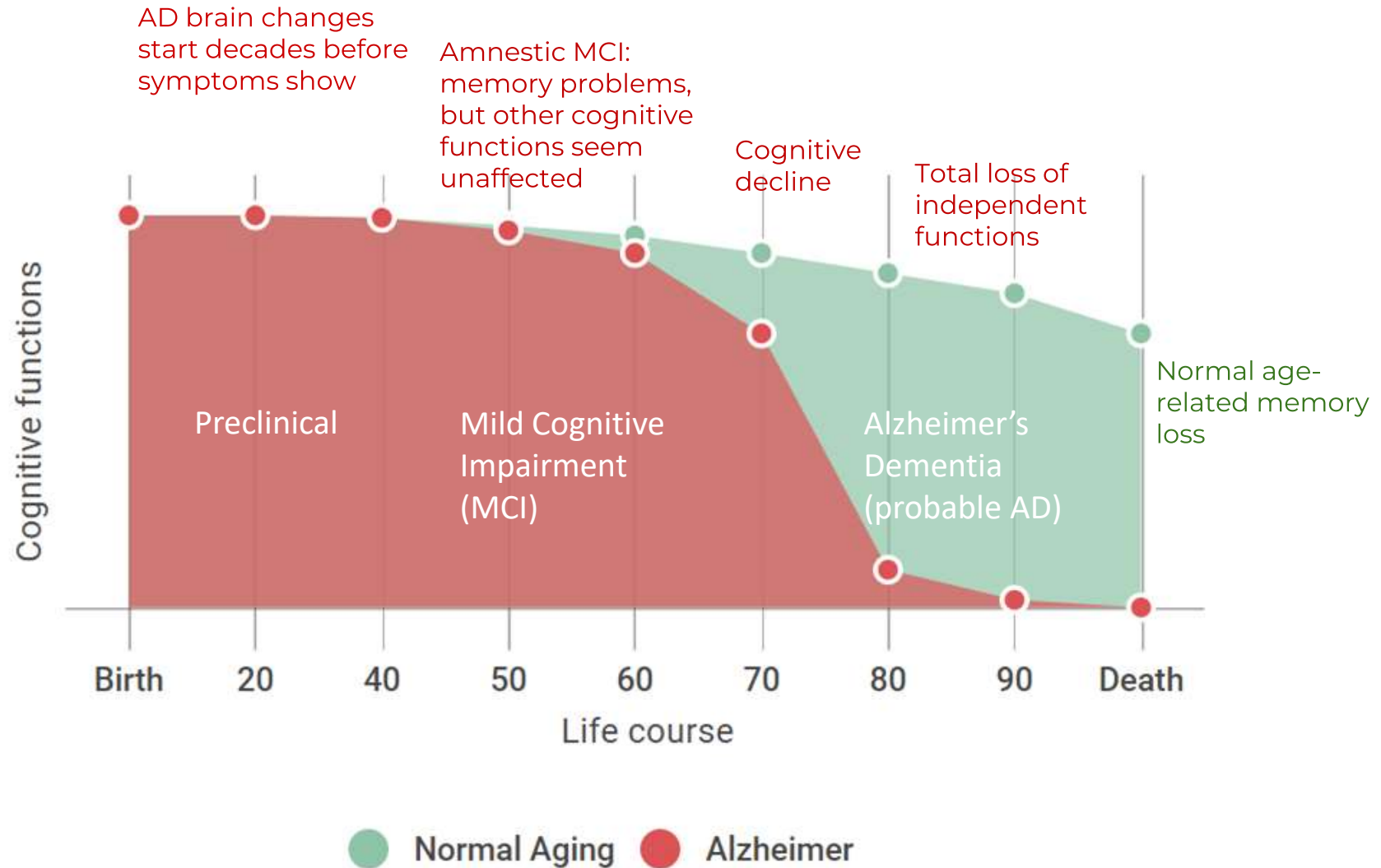
*will this therapy be effective
for this patient?*

ML Models | Examples

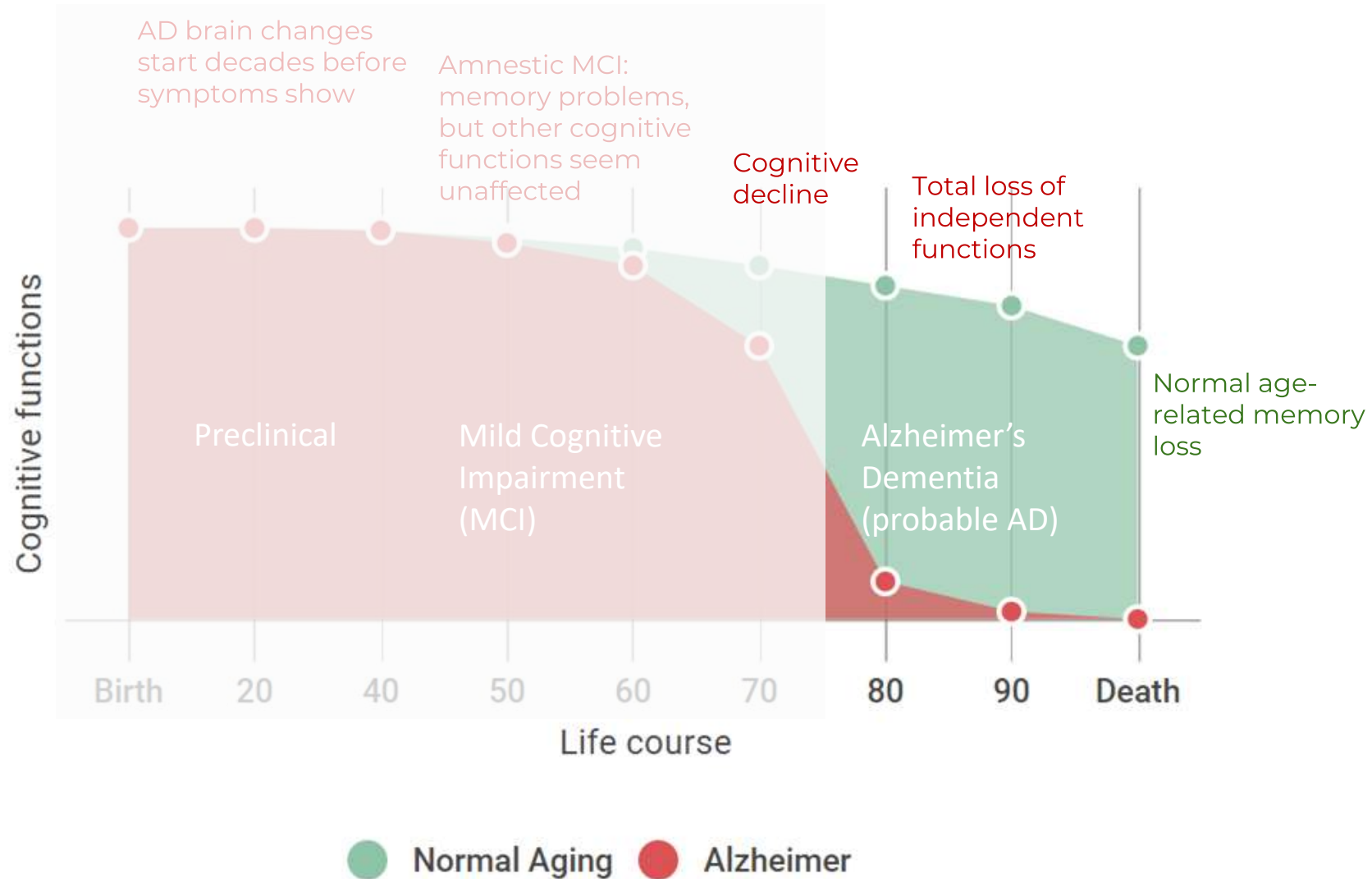
(using different types of data)

Alzheimer's Disease

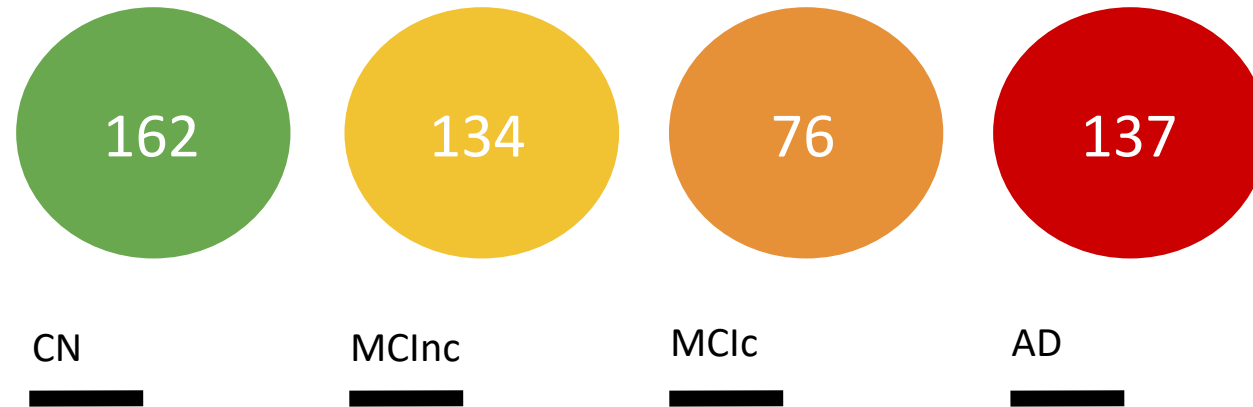
ML models | Examples



ML models | Examples



Clinical Diagnosis of Alzheimer's Disease



AD
MCIC
MCInc
CN

Alzheimer's Disease

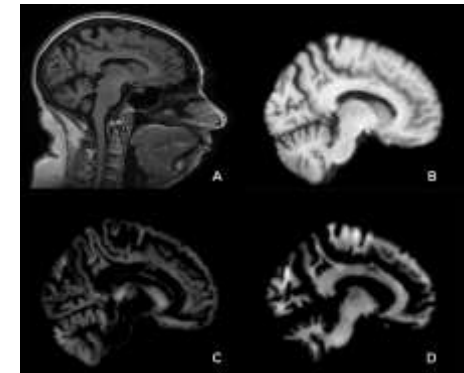
Mild Cognitive Impairment, converting to Alzheimer's Dementia

Mild Cognitive Impairment, not converting to Alzheimer's Dementia




Cognitively-Normal subjects

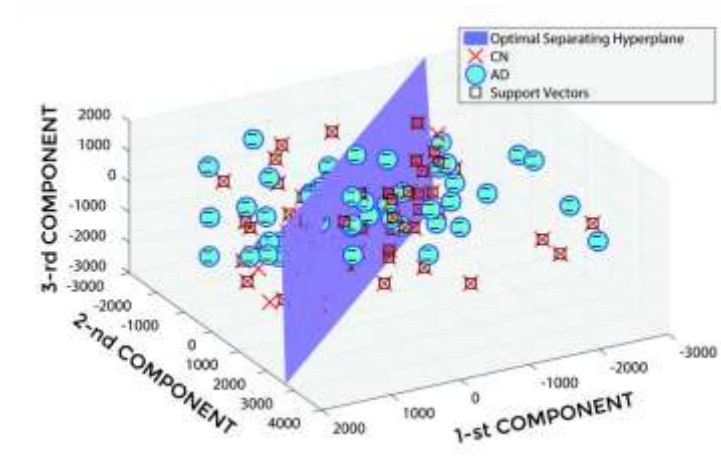
A dataset
of 509 subjects

Structural MRI
T1 weighted
1.5 Tesla

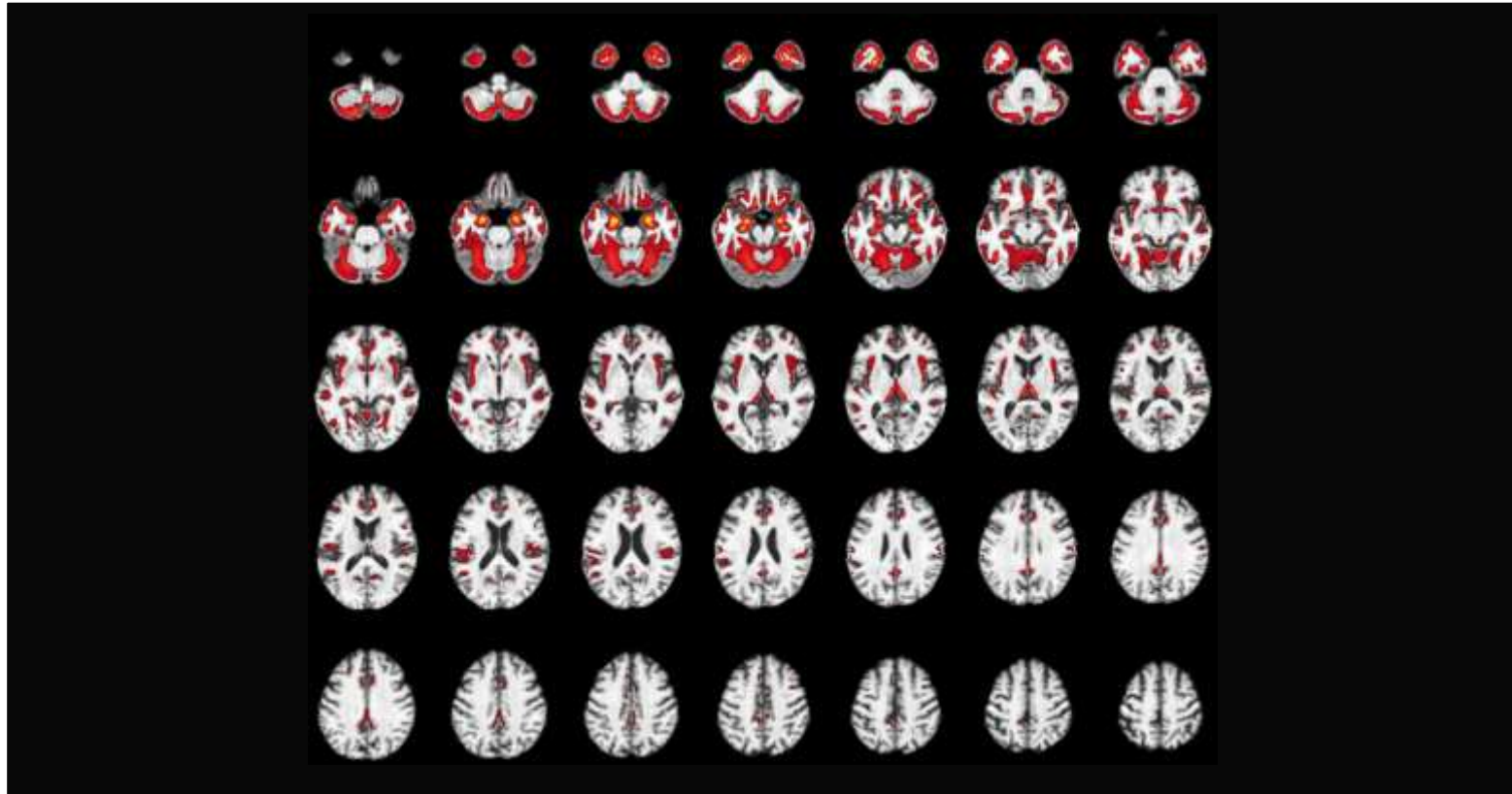


Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: A machine learning approach.
Salvatore et al. 2015, Frontiers in Neuroscience.

	Clinical Diagnosis			Accuracy (automatic classification)
	AD	vs	CN	76%
	MCIc	vs	CN	72%
	MCIc	vs	MCIc	66%



Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: A machine learning approach.
Salvatore et al. 2015, Frontiers in Neuroscience.

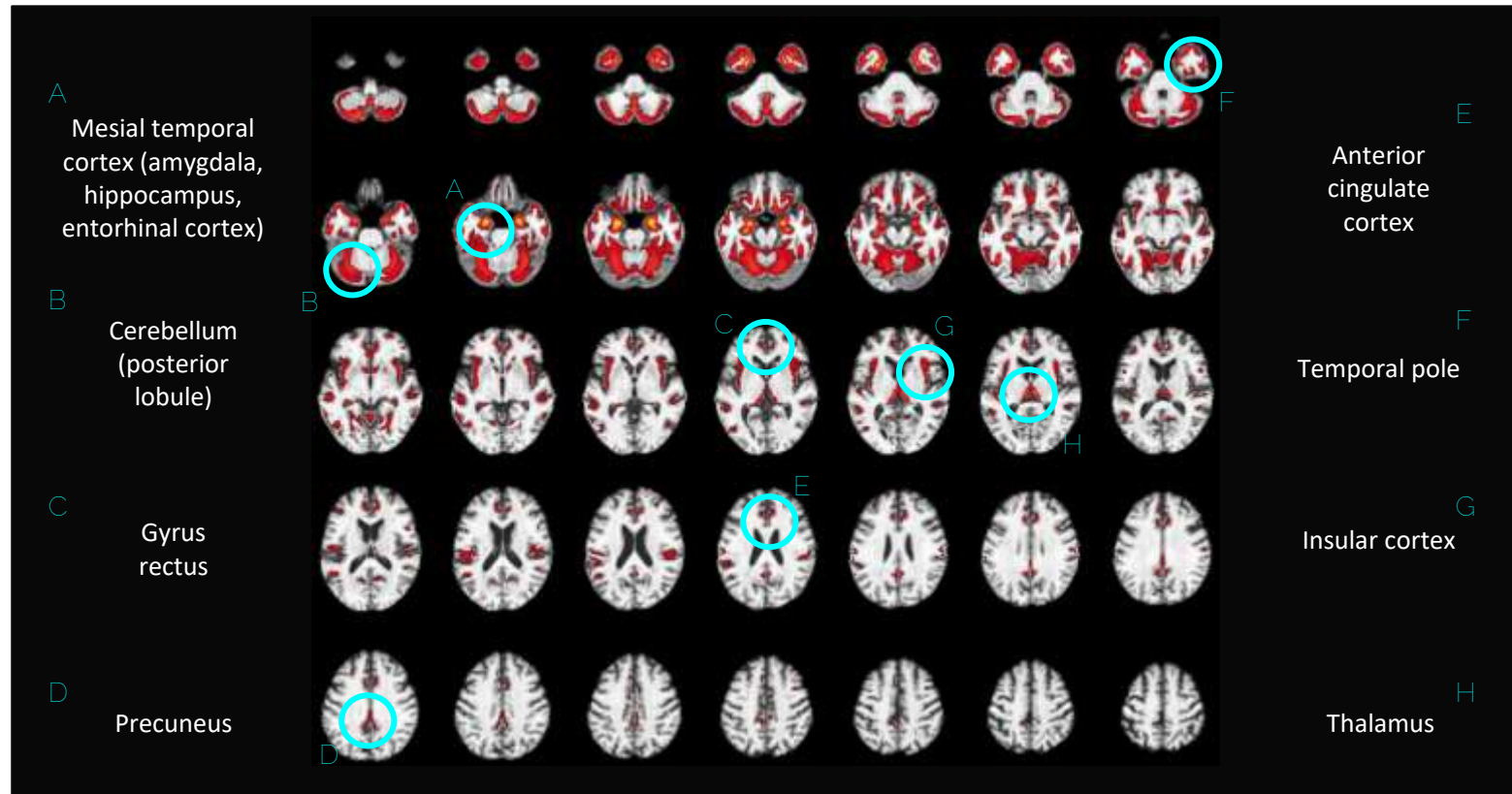


**Best
Structural-MRI
Predictors**

AD vs CN

Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: A machine learning approach.
Salvatore et al. 2015, Frontiers in Neuroscience.

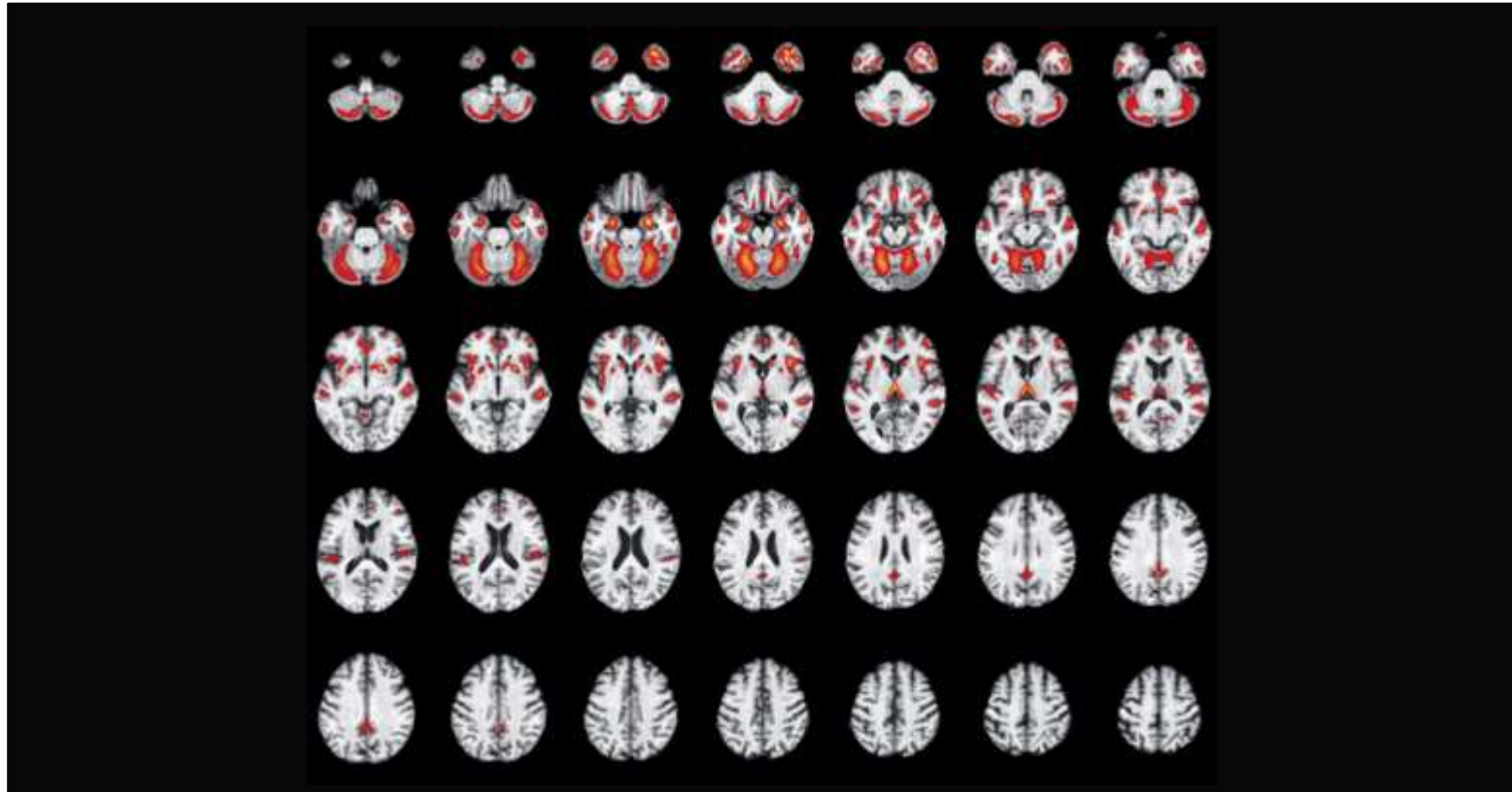
ML models | Examples



**Best
Structural-MRI
Predictors**

AD vs CN

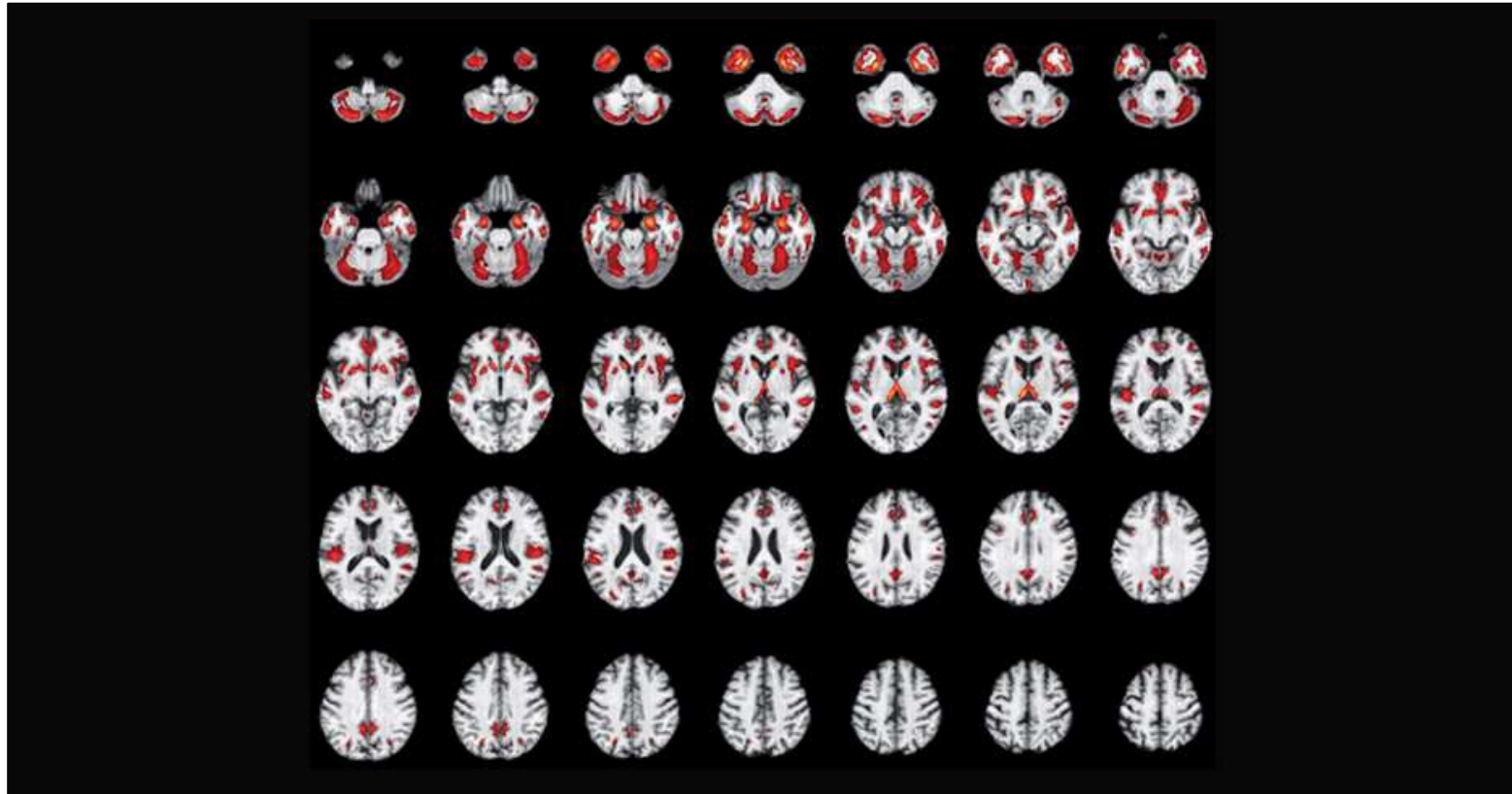
Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: A machine learning approach.
Salvatore et al. 2015, Frontiers in Neuroscience.



**Best
Structural-MRI
Predictors**

MCIc vs CN

Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: A machine learning approach.
Salvatore et al. 2015, Frontiers in Neuroscience.

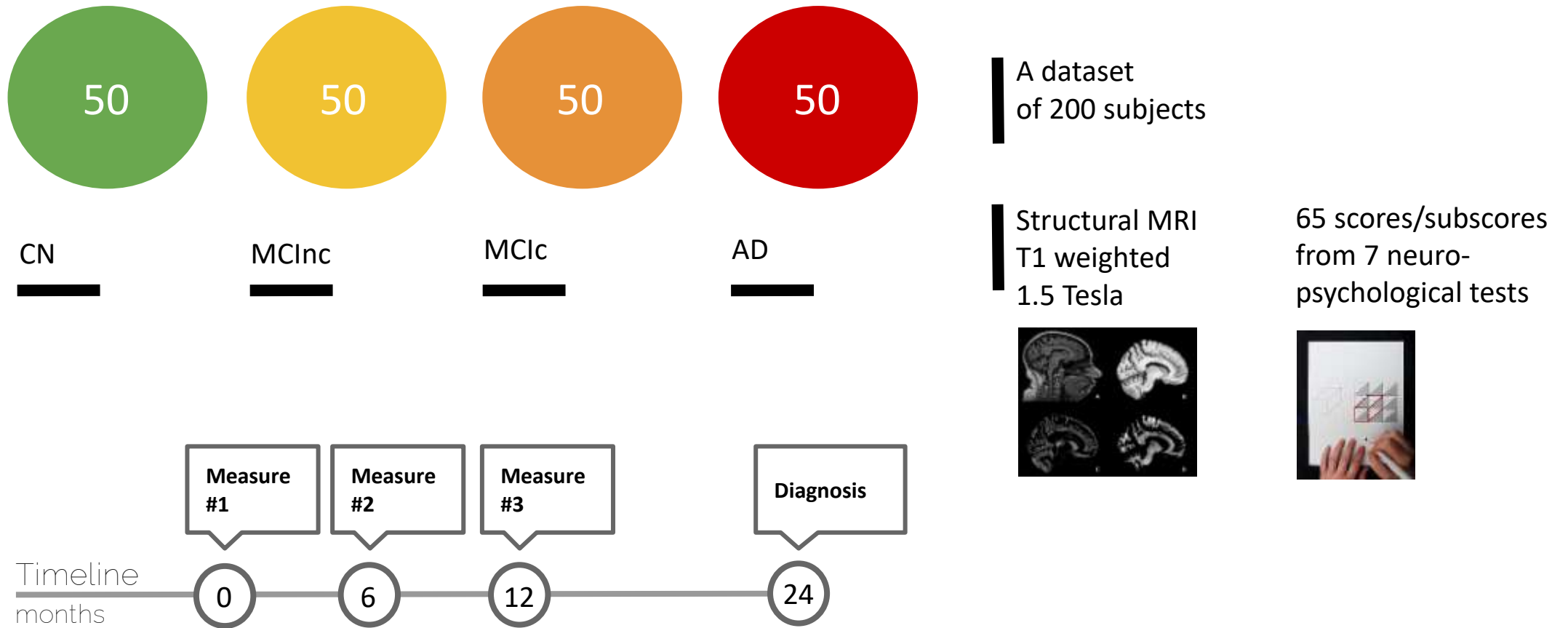


**Best
Structural-MRI
Predictors**

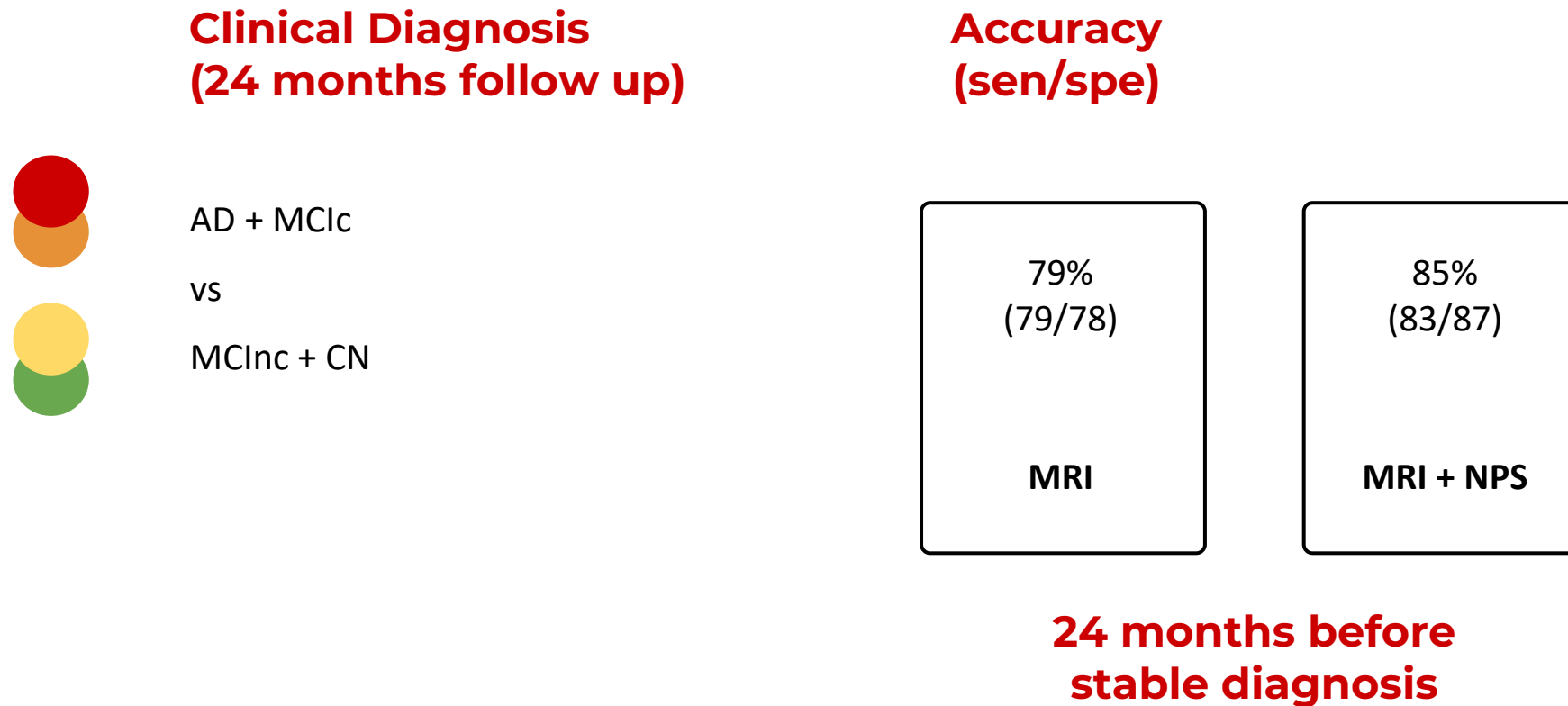
MCIc vs MCIc

Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: A machine learning approach.
Salvatore et al. 2015, Frontiers in Neuroscience.

Clinical Diagnosis of AD at 24 months follow up



MRI characterizes the progressive course of AD and predicts conversion to Alzheimer's dementia twenty-four months before probable diagnosis. Salvatore et al. 2018, *Frontiers in Aging Neuroscience*.



MRI characterizes the progressive course of AD and predicts conversion to Alzheimer's dementia twenty-four months before probable diagnosis. *Salvatore et al. 2018, Frontiers in Aging Neuroscience.*

ML models | Examples

Best Neuropsychological Predictors

24 months before stable diagnosis

Ability in remembering appointments, family occasions, holidays, medications in [FAQ](#)
Functional abilities

Ability in writing checks, paying bills, or balancing checkbook in [FAQ](#)
Functional abilities

Ability in assembling tax records, business affairs in [FAQ](#)
Functional abilities

Total score of trial 5 in [AVLT](#)
Memory and learning

Ability in keeping track of current events in [FAQ](#)
Functional abilities

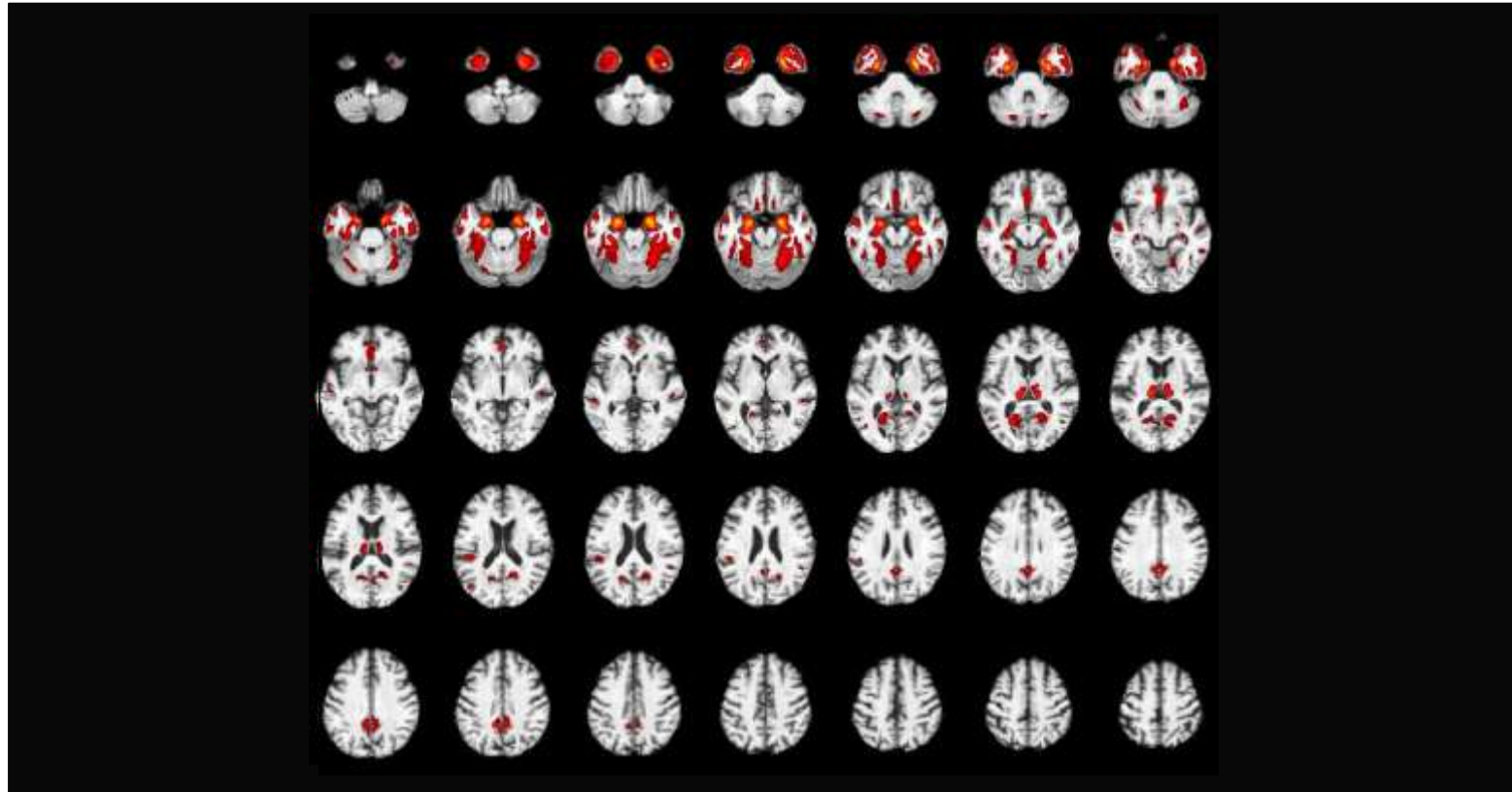
Total intrusions of trial 1 in [AVLT](#)
Memory and learning

Correct answers in the Backwards task in [Digit-Span Test](#)
Working memory

Correct answers in Vegetables task in [Category Fluency Test](#)
Language

Correct answers after a 30-min delay in [AVLT](#)
Memory and learning

MRI characterizes the progressive course of AD and predicts conversion to Alzheimer's dementia twenty-four months before probable diagnosis. Salvatore et al. 2018, *Frontiers in Aging Neuroscience*.

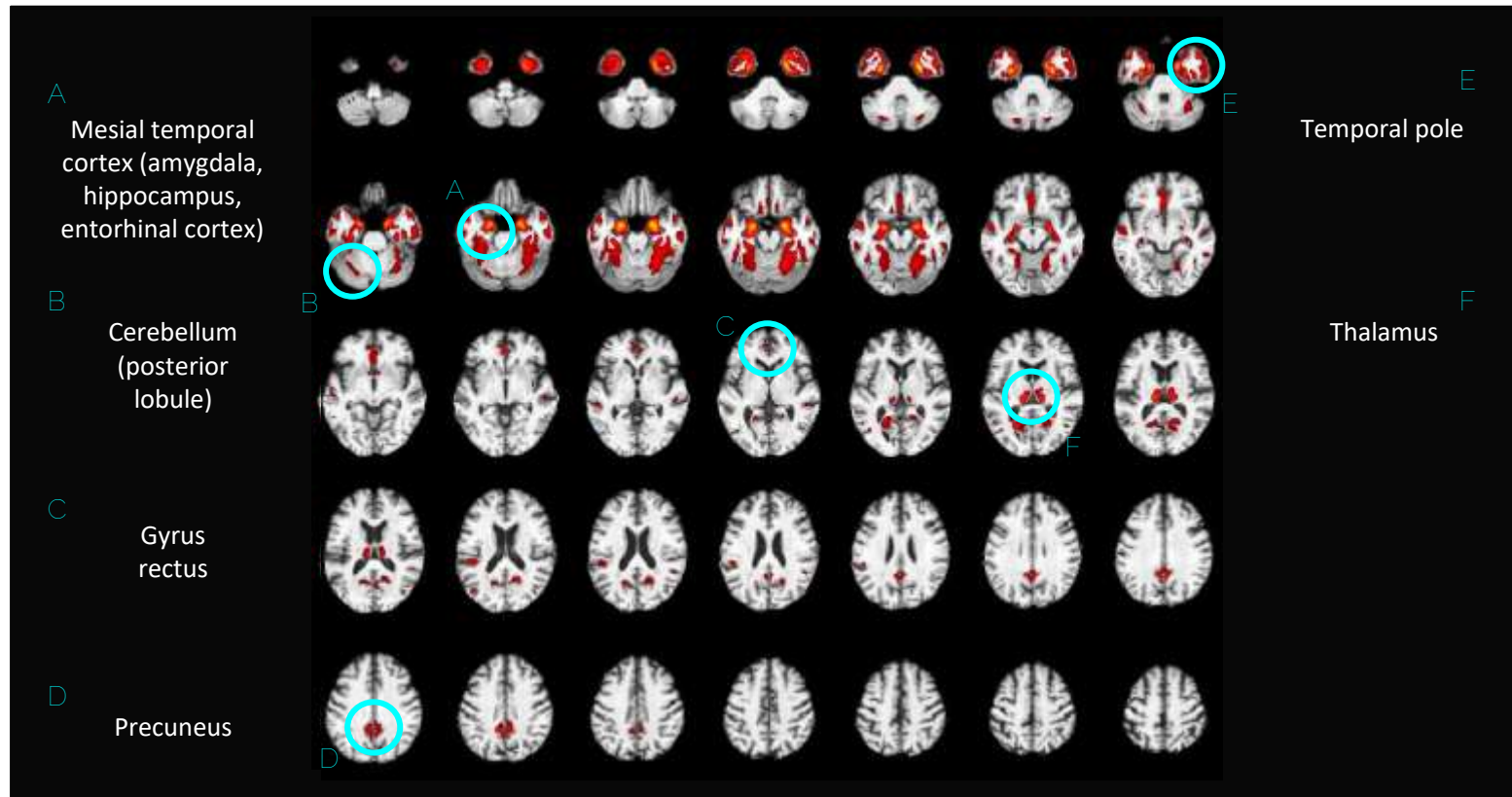


**Best
Structural-MRI
Predictors**

**24 months
before
stable
diagnosis**

MRI characterizes the progressive course of AD and predicts conversion to Alzheimer's dementia twenty-four months before probable diagnosis. *Salvatore et al. 2018, Frontiers in Aging Neuroscience.*

ML models | Examples



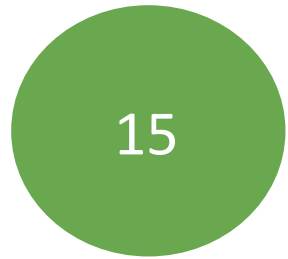
**Best
Structural-MRI
Predictors**

**24 months
before
stable
diagnosis**

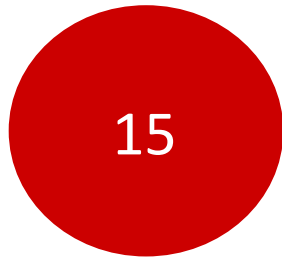
MRI characterizes the progressive course of AD and predicts conversion to Alzheimer's dementia twenty-four months before probable diagnosis. Salvatore et al. 2018, *Frontiers in Aging Neuroscience*.

Autism Spectrum Disorder

Confirm a motor signature of autism



TD



ASD



A dataset of 30
pre-school children (~3 years old)

17 kinematic features
collected during a reach-to-drop task



ASD
TD

Autism Spectrum Disorder
Typically-developing children

Use of machine learning to identify children with autism and their motor abnormalities. *Crippa, Salvatore et al. 2015, Journal of Autism and Developmental Disorders.*

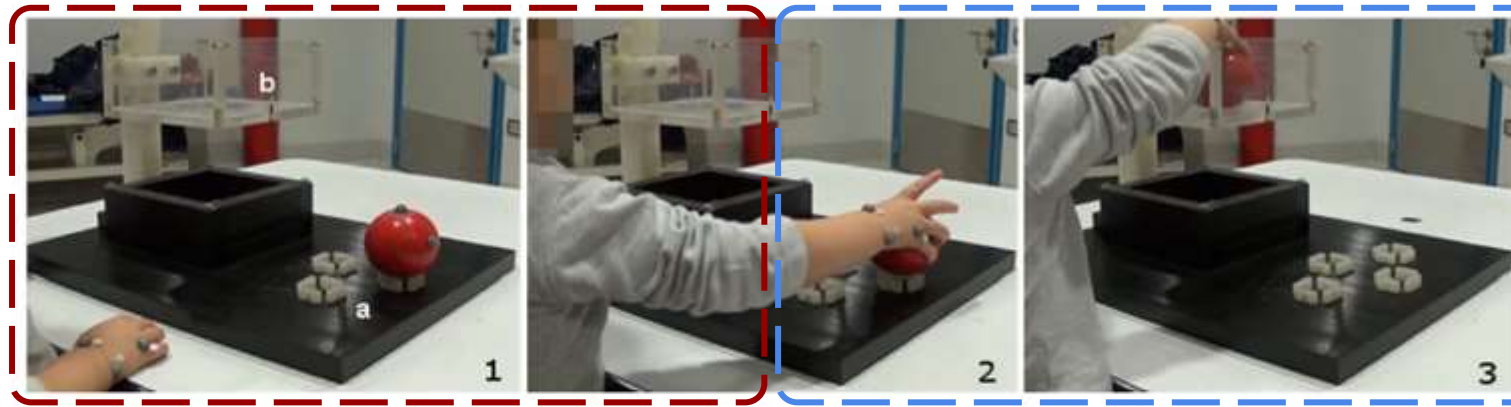


Fig. 1 The experimental task consisted of grasping a rubber ball (2) that was placed over a support (see 1, a); that is, a reach-to-grasp movement before they dropped it in a hole (3). The hole (1, c) was located inside a see-through *square box* (21 cm high, 20 cm wide) and was large enough not to require fine movements. The goal area is

transparent to allow seeing through. 4 markers are placed on the basket under the goal area, 2 on the ball and 3 on each hand (attached to the ulnar and radial surfaces of the participant's wrist and to the hand dorsum on the 4th and 5th metacarpals)

sub-movement 1

the movement necessary to reach the ball and place it on its support

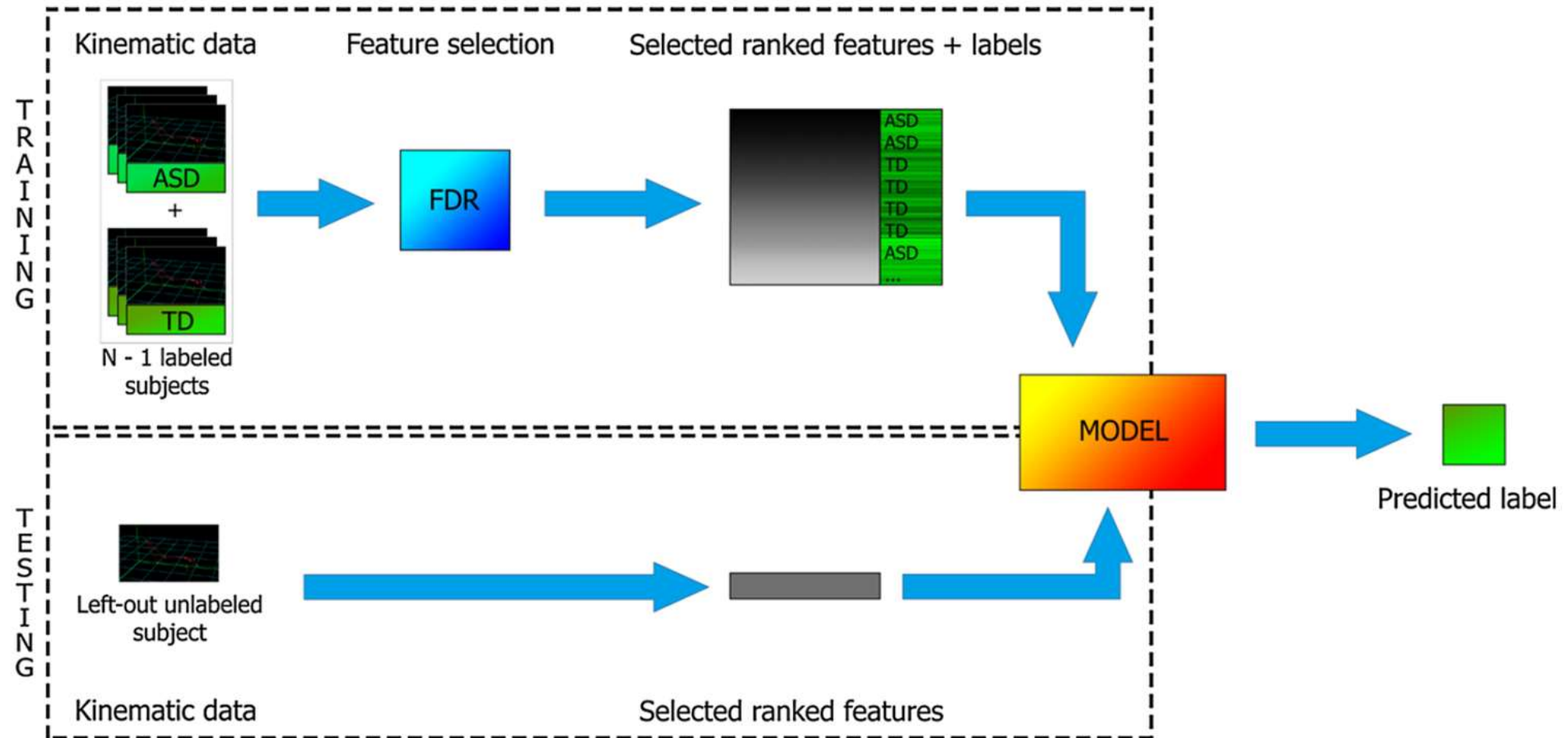
- # movement units
- total movement duration
- peak velocity
- peak acceleration
- time of peak acceleration
- peak deceleration
- time of peak deceleration

sub-movement 2

the movement to transport the ball from a support to the target hole

- # movement units
- total movement duration
- peak velocity
- peak acceleration
- time of peak acceleration
- peak deceleration
- time of peak deceleration
- wrist angle

ML models | Examples



Use of machine learning to identify children with autism and their motor abnormalities. Crippa, Salvatore et al. 2015, *Journal of Autism and Developmental Disorders*.

ML models | Examples

Diagnostic accuracy
(sensitivity / specificity)

Overall
mean

85
(82 / 89)%

Optimal
configuration

97
(100 / 94)%

Diagnostic accuracy
(sensitivity / specificity)

Overall
mean

85
(82 / 89)%

Optimal
configuration

97
(100 / 94)%

7 optimal features out of 17

sub-movement 2

the movement to transport the ball from a support to the target hole in which the ball was to be dropped

1. total movement duration
2. delta wrist angle
3. # movement units
4. time of peak deceleration
5. peak acceleration
6. time of peak velocity
7. peak velocity

Attention-Deficit/Hyperactivity Disorder

Diagnosis of ADHD & Identification of a signature



TD



ADHD



A dataset of 44
school-aged children (~11 years old)

Multidomain profile of measures

- . biological data
- . neuropsychological data
- . Near-InfraRed Spectroscopy (NIRS)

ADHD
TD

Attention Deficit / Hyperactivity Disorder
Typically-developing children

The Utility of a computerized algorithm Based on a Multi-Domain Profile of Measures for the Diagnosis of attention deficit/hyperactivity Disorder. *Crippa, Salvatore et al. 2017, Frontiers in Psychiatry.*

Multi-domain profile of measures



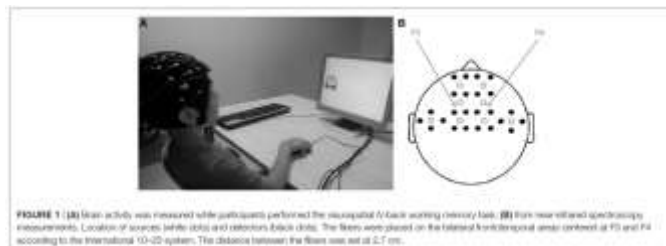
Biological data

10 features



Neuropsychological data

18 features

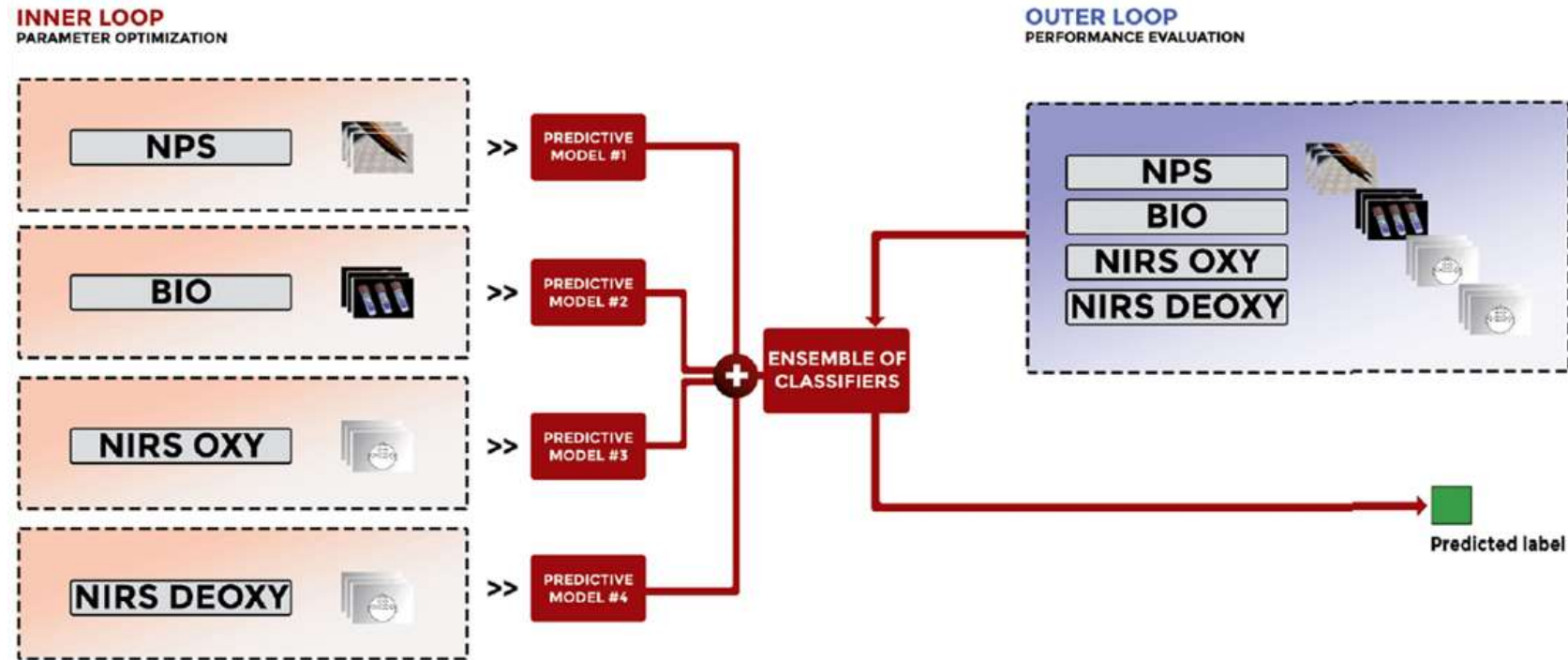


Near-InfraRed Spectroscopy (NIRS)

Oxy/Deoxy data from 32 channels

The Utility of a computerized algorithm Based on a Multi-Domain Profile of Measures for the Diagnosis of attention deficit/hyperactivity Disorder. *Crippa, Salvatore et al. 2017, Frontiers in Psychiatry.*

ENSEMBLE OF CLASSIFIERS



The Utility of a computerized algorithm Based on a Multi-Domain Profile of Measures for the Diagnosis of attention deficit/hyperactivity Disorder. Crippa, Salvatore et al. 2017, *Frontiers in Psychiatry*.

ML models | Examples

Measures	Accuracy (mean ± sd)	Sensitivity (mean ± sd)	Specificity (mean ± sd)
Neuropsychological	62 ± 17	70 ± 27	57 ± 24
Biological	66 ± 21	58 ± 40	73 ± 29
NIRS OXY	57 ± 27	48 ± 47	67 ± 33
NIRS DEOXY	78 ± 22	72 ± 34	82 ± 24
NIRS OXY + DEOXY	72 ± 32	73 ± 29	68 ± 43

biological features

.linoleic acid
.PUFA
.AA
.EPA
.omega-3 index
.AA/DHA
.AA/EPA
.MUFA

linoleic acid and total amount of
polyunsaturated fatty acids

The Utility of a computerized algorithm Based on a Multi-Domain Profile of Measures for the Diagnosis of attention deficit/hyperactivity Disorder. Crippa, Salvatore et al. 2017, *Frontiers in Psychiatry*.

ML models | Examples

Measures	Accuracy (mean ± sd)	Sensitivity (mean ± sd)	Specificity (mean ± sd)
Neuropsychological	62 ± 17	70 ± 27	57 ± 24
Biological	66 ± 21	58 ± 40	73 ± 29
NIRS OXY	57 ± 27	48 ± 47	67 ± 33
NIRS DEOXY	78 ± 22	72 ± 34	82 ± 24
NIRS OXY + DEOXY	72 ± 32	73 ± 29	68 ± 43

neuropsychological features

- .sustained attention**-false alarms
- .visual set**-shifting-RT inhibition
- .sustained attention-coefficient of variation
- .visual set-shifting-number of inhibition errors
- .focused attention**-RT correct responses
- .focused attention-correct rejections target non-relevant position
- .focused attention-SD of correct responses RT
- .focused attention-misses
- ...

measures of vigilance, focused and sustained attention, and cognitive flexibility

The Utility of a computerized algorithm Based on a Multi-Domain Profile of Measures for the Diagnosis of attention deficit/hyperactivity Disorder. Crippa, Salvatore et al. 2017, *Frontiers in Psychiatry*.

ML models | Examples

Single domain

Measures	Accuracy (mean ± sd)	Sensitivity (mean ± sd)	Specificity (mean ± sd)
Neuropsychological	62 ± 17	70 ± 27	57 ± 24
Biological	66 ± 21	58 ± 40	73 ± 29
NIRS OXY	57 ± 27	48 ± 47	67 ± 33
NIRS DEOXY	78 ± 22	72 ± 34	82 ± 24
NIRS OXY + DEOXY	72 ± 32	73 ± 29	68 ± 43

Multi-domain

Measures	Accuracy (mean ± sd)	Sensitivity (mean ± sd)	Specificity (mean ± sd)
NPS + BIO + NIRS OXY			
NPS + BIO + NIRS DEOXY			
NPS + NIRS OXY + NIRS DEOXY			
BIO + NIRS OXY + NIRS DEOXY			
NPS + BIO + NIRS OXY + NIRS DEOXY			

The Utility of a computerized algorithm Based on a Multi-Domain Profile of Measures for the Diagnosis of attention deficit/hyperactivity Disorder. *Crippa, Salvatore et al. 2017, Frontiers in Psychiatry.*

ML models | Examples

Single domain

Measures	Accuracy (mean ± sd)	Sensitivity (mean ± sd)	Specificity (mean ± sd)
Neuropsychological	62 ± 17	70 ± 27	57 ± 24
Biological	66 ± 21	58 ± 40	73 ± 29
NIRS OXY	57 ± 27	48 ± 47	67 ± 33
NIRS DEOXY	78 ± 22	72 ± 34	82 ± 24
NIRS OXY + DEOXY	72 ± 32	73 ± 29	68 ± 43

Multi-domain

Measures	Accuracy (mean ± sd)	Sensitivity (mean ± sd)	Specificity (mean ± sd)
NPS + BIO + NIRS OXY	71 ± 10	70 ± 27	73 ± 24
NPS + BIO + NIRS DEOXY	81 ± 15	73 ± 24	87 ± 22
NPS + NIRS OXY + NIRS DEOXY	78 ± 18	70 ± 36	87 ± 22
BIO + NIRS OXY + NIRS DEOXY	77 ± 21	63 ± 31	90 ± 21
NPS + BIO + NIRS OXY + NIRS DEOXY	76 ± 16	83 ± 22	68 ± 23

The Utility of a computerized algorithm Based on a Multi-Domain Profile of Measures for the Diagnosis of attention deficit/hyperactivity Disorder. *Crippa, Salvatore et al. 2017, Frontiers in Psychiatry.*