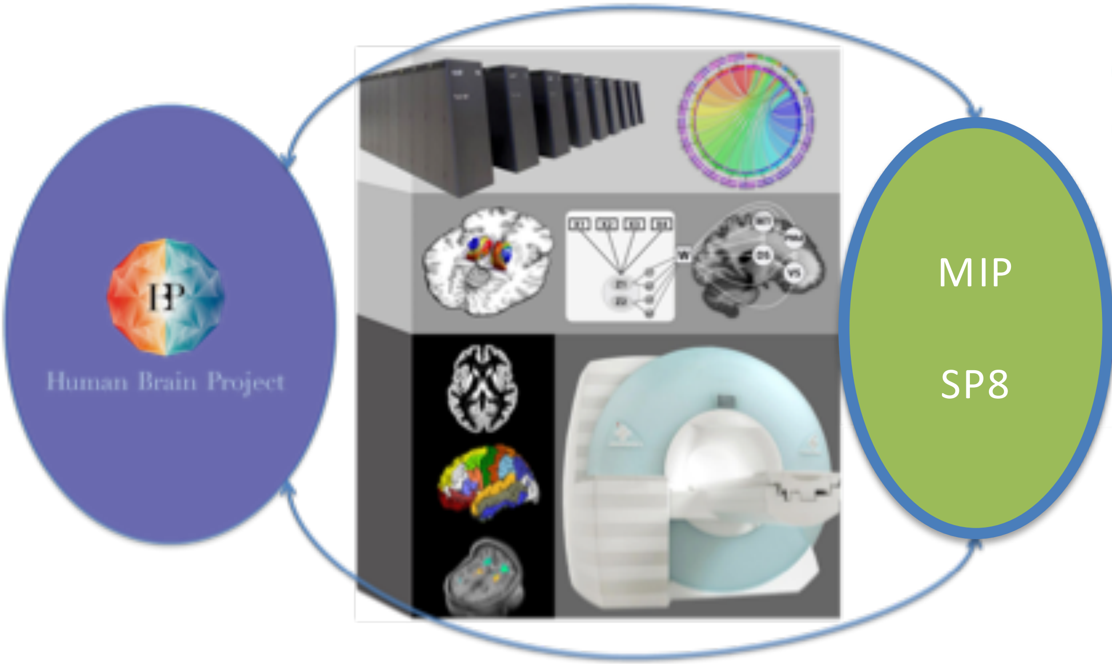
HUMAN BRAIN PROJECT

MEDICAL INFORMATICS PLATFORM



EU funded Collaborative project for understanding the human brain

25 Countries

400 Researchers **2013**

10 Years

Richard FRACKOWIAK

THE MEDICAL COMPONENT OF THE HBP

Alzheimer's disease: **20 per cent** beyond the age of 80; dependent within 3-5 years of onset.

Depression: the second most common condition in the world (WHO): **6 per cent** of the population in the Western world.

Cerebral vascular accidents: first cause of adult motor disability. **75 per cent** suffer residual disability.

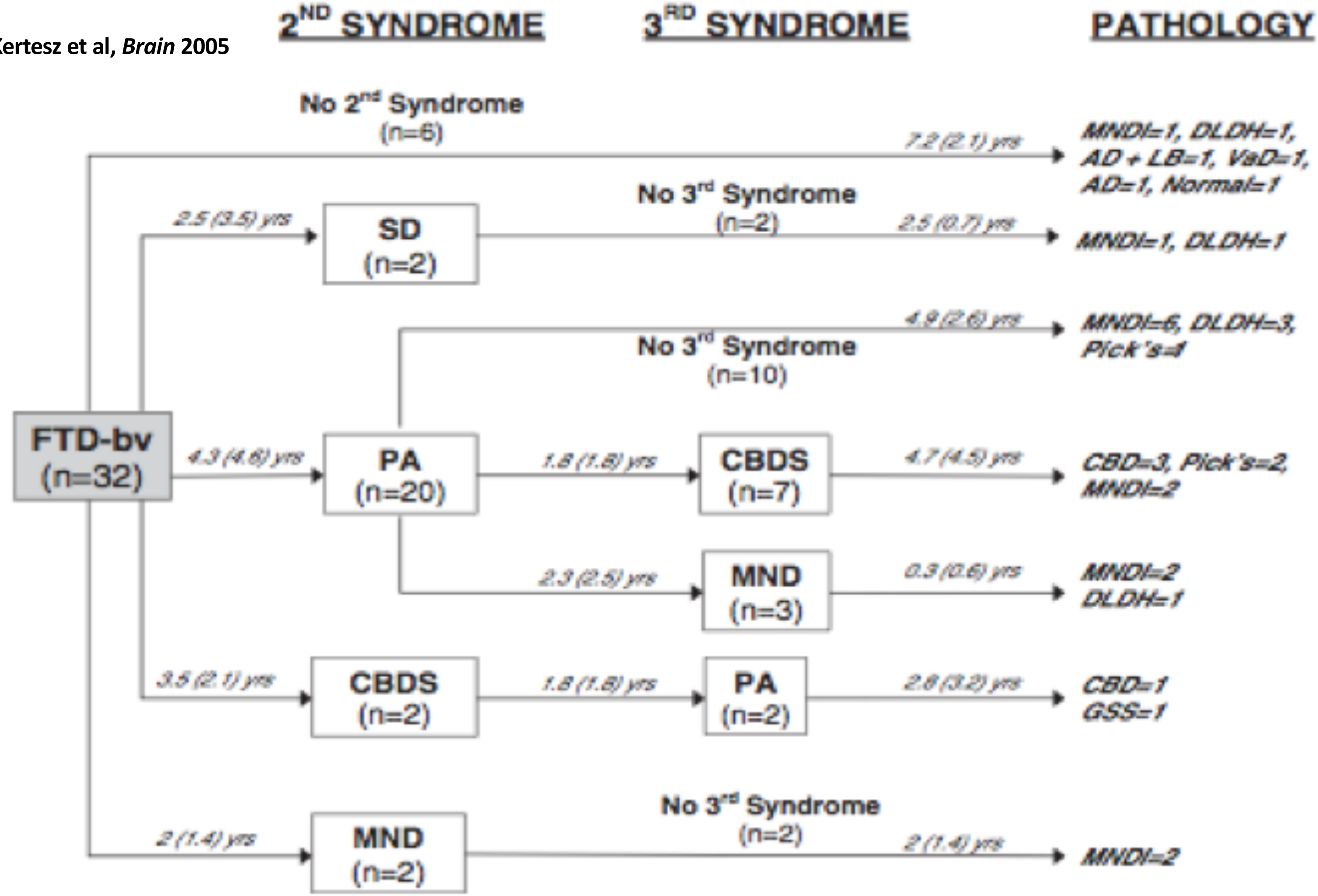
Parkinson's disease: second cause of motor disability. Affects **0.2 per cent** of the population.

Multiple sclerosis: mainly young people with dependency in **30 per cent**.

Epilepsy: 50 million people globally of which almost **50 per cent** are aged < 10 years**.** Social and familial repercussions are **lifelong.**

HAVE WE REACHED A DEAD END CLINICALLY?

Kertesz et al, *Brain* 2005



HYPOTHESIS 1

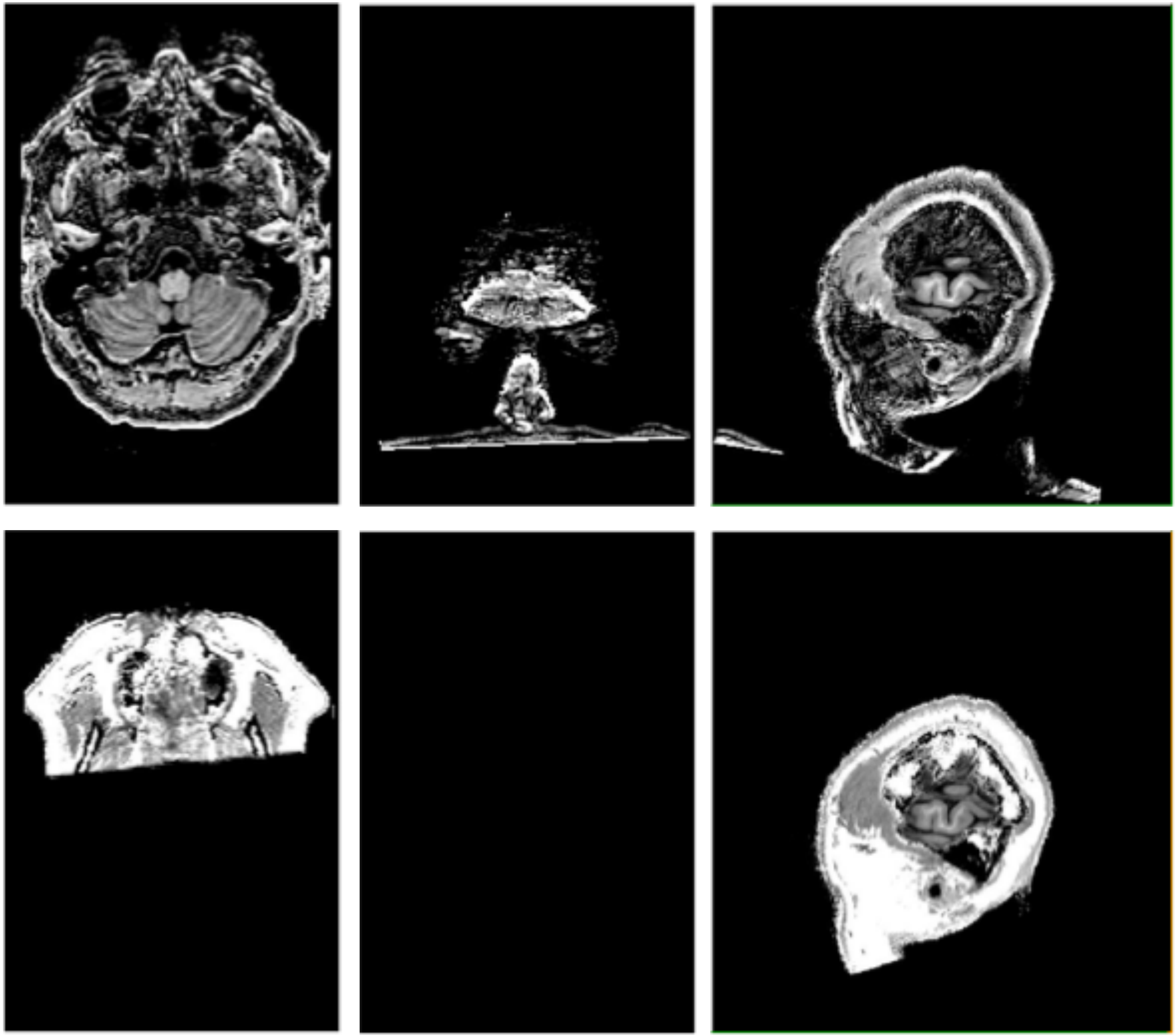
Phenomenology alone

is insufficiently discriminative

for diagnosis and prognosis

Genotyping does not

replace descriptive medicine



HYPOTHESIS 2

PATTERNS OF PIXEL ABNORMALITIES

ARE OF DIAGNOSTIC

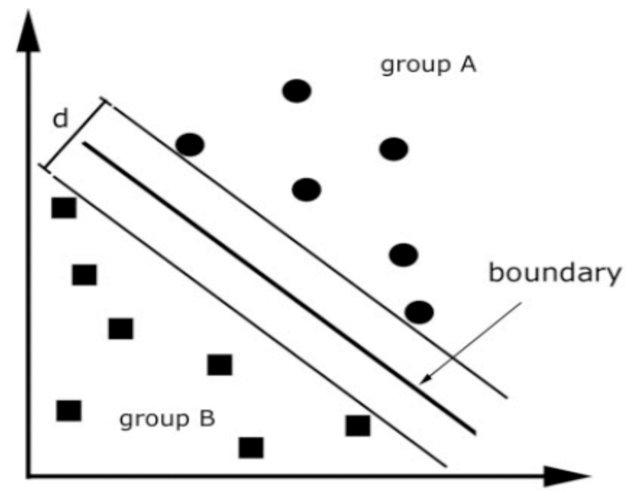
and/or

PROGNOSTIC SIGNIFICANCE

INFORMATICS CLASSIFY PATTERNS

	CORRECT	SENSITIVITY	SPECIFICITY
AD & CONTROLS **CLINICAL**	81%	61%	93%
AD 1 & CONTROLS **PATHOLOGY**	95%	95%	95%
AD 2 & CONTROLS **PATHOLOGY**	93%	100%	86%
AD 1 & CONTROLS vs AD 2 **PATHOLOGY**	96%	100%	93%

BINARY CLASSIFICATION BY SUPPORT VECTOR MACHINE



INFORMATICS REFINE DIAGNOSIS

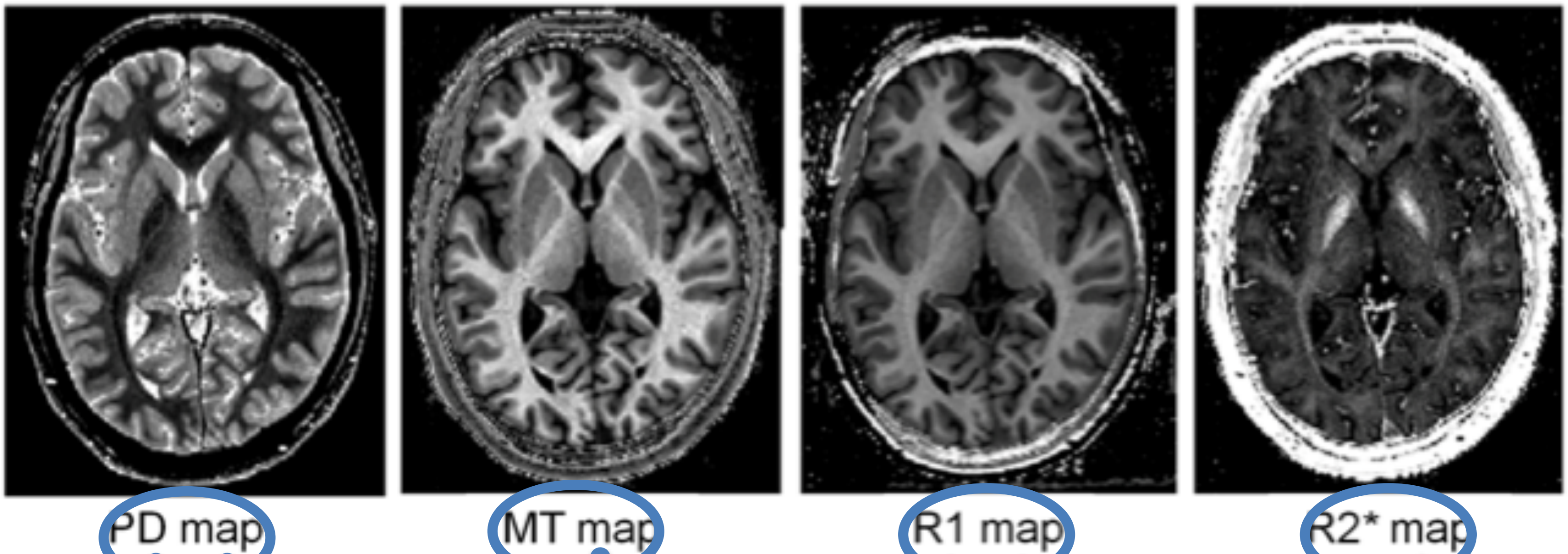
ADNI DATABASE

	CLINICAL AD	CLINICAL NC	
PATHOL AD+	15	3	**18**
PATHOL AD-	5	17	**22**
	20	**20**	

SVM ANALYSIS

QUANTITATIVE MULTI-PARAMETER MAPPING BASED ON BIOPHYSICAL MODELS

Lorio et al., 2016 *HBMapp*
Lutti et al., 2012 *PLOS One*
Draganski et al., 2011 *NeuroImage*



Proton density
Water content

Magnetization transfer saturation
Myelin content

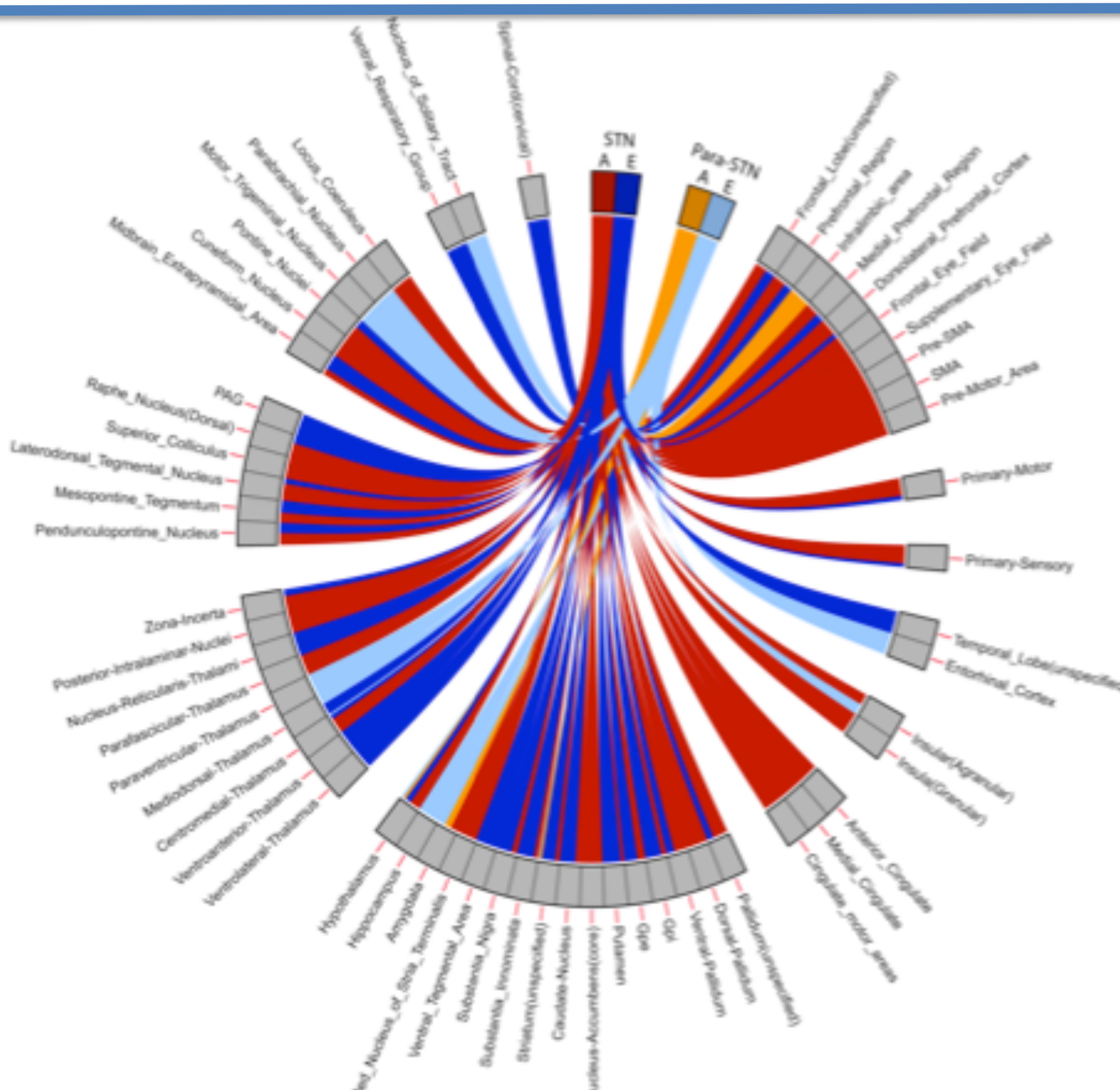
Longitutinal relaxation rate
Myelin content & water compart-mentalisation

Effective transverse relaxation rate
Iron content

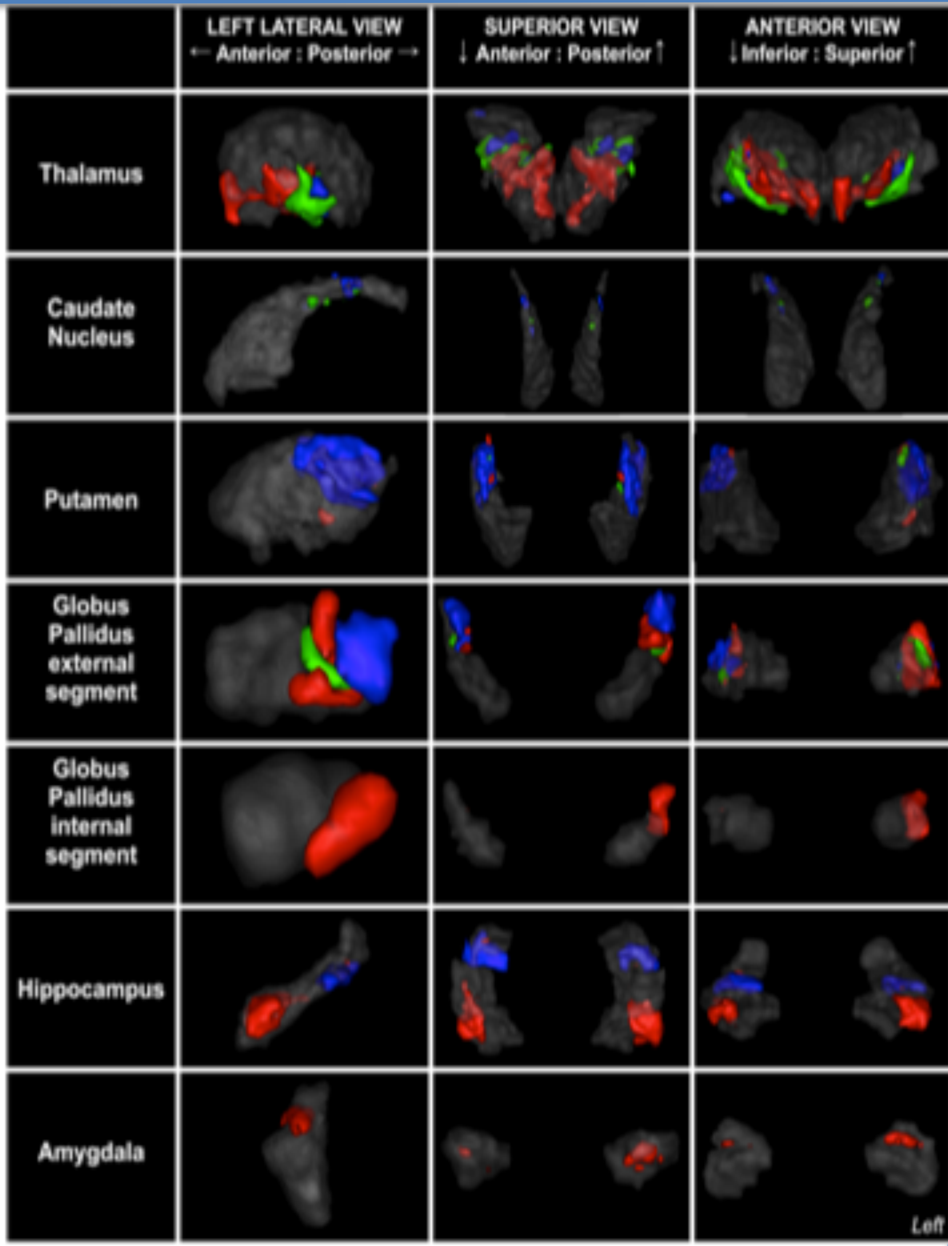
INFORMATICS INTEGRATE DATA

A review of the entire tract-tracing literature of the STN between 1947-2011 reveals connectivity between a broad array of cortical, sub-cortical and brainstem structures.

BLUE = EFFERENT
RED = AFFERENT



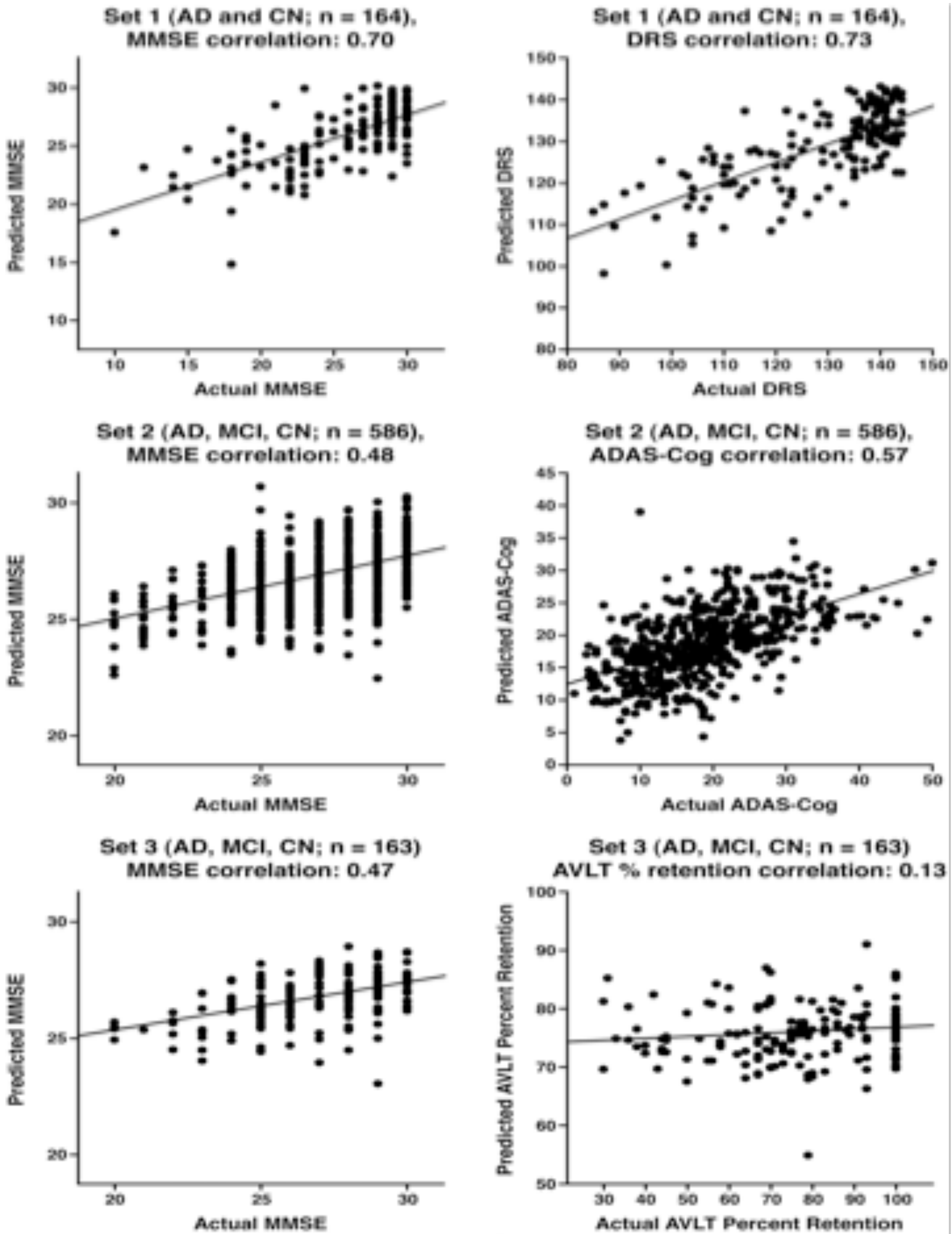
FUNCTIONAL CONNECTIONS



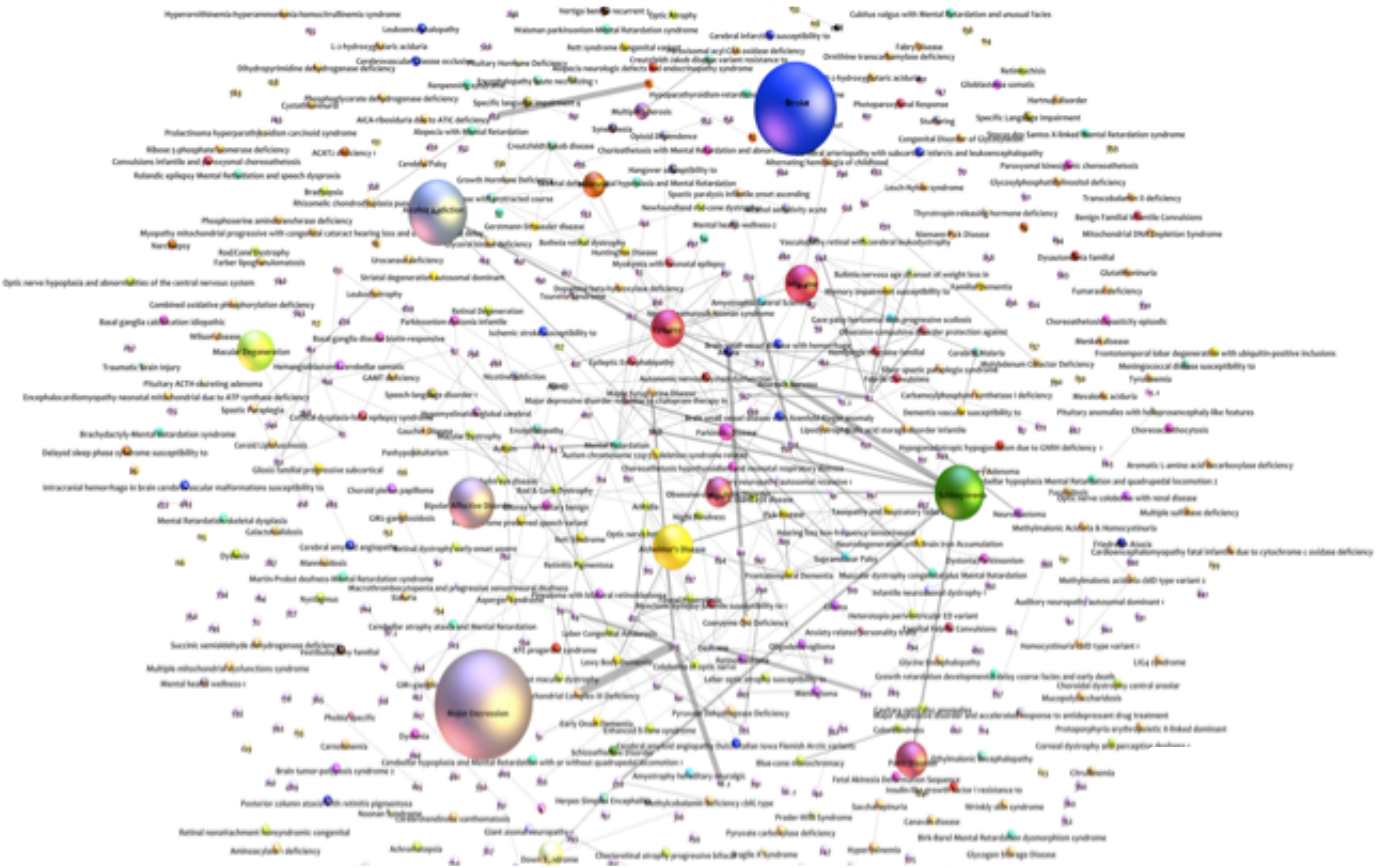
The posterior aspect of the STN projects to structures consistent with a motor structure:
Posterior putamen
Posterior GPe
Mid caudate nucleus
Ventro-lateral thalamic nuclei
Posterior Insula
Posterior hippocampus

The anterior aspect of the STN projects to structures consistent with a limbic structure:
Baso-lateral amygdala
Posterio-medial GPi
Inferio-mid putamen
Mid-GPe
Ventral-anterior and ventral-lateral thalamus
Anterior Insula
Anterior hippocampus

The middle "associative" STN projects to regions encompassing both the motor and limbic projections



PREDICTING CLINICAL SCORE COMPARING TO ASSESS RESERVE


Stroke
Macular Degeneration
Alzheimer's Disease
Schizophrenia
Major Depression

CONFRONTING PARADIGMS

CARTESIAN MODEL (TOP DOWN)

Mentally generated hypothesis
Mathematically expressed in a model
Confrontation with "relevant" data
Parameterisation and optimisation of model

Correlations (non-causal)

SIMULATION MODEL (BOTTOM UP)

Multimodal and multivariate data
Exhaustive mining to demonstrate coherent models
Exploration of these mathematical models as generated hypotheses
Investigation of hypotheses – clinical, mechanistic, prognostic, therapeutic

Knowledge (& causes)

FUNDAMENTAL ISSUE

CLINICAL AND NEUROSCIENCE DATA MUST BE INTEGRATED

SIMULATION is an analytical methodology & depends on high performance computing

SIMULATION is always bottom up – it is a reconstruction from real data

It generates complexity from simpler elements

It results in a PREDICTIVE MODEL that CONSTRAINS next level organisation

SIMULATION can start from any level but always bottom-up and data led

But data are useless unless they help reconstruction from one level to the next.

WE NEED GOOD PREDICTIVE SIMULATION MODELLING TO REDEFINE & REFINE NEURO-DIAGNOSES IN FUTURE DIAGNOSTIC MANUALS

BIG DATA & INFORMATICS

RUBBISH IN ≠ **RUBBISH OUT**

Signal additive
Noise suppressed
Avoids myth of perfect controls – whither RCTs?

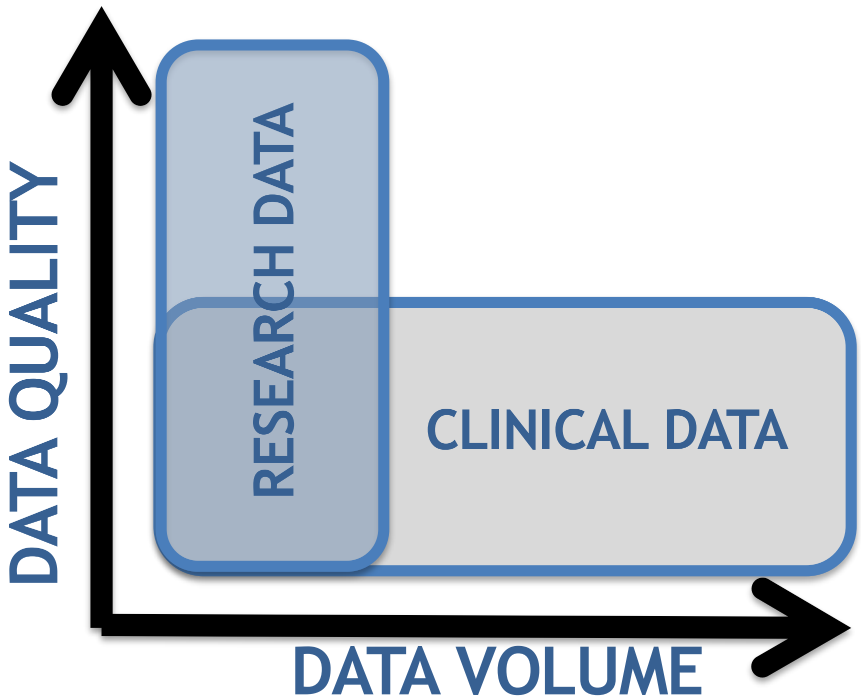
→ **GENERATES HYPOTHESES**

Biologically plausible
Explores multiple (all) models
Multi-scale & dimensional patterns
Additive over time
Built in reproducibility

→ **PREDICTIVE & CAUSAL**

How big data can help: Bradford Hill (1965)

DATA SOURCES AND CHALLENGES



HOSPITAL DATABASES

- NOT COMPLETE
- NOT STRUCTURED
- NOT STANDARDISED
- NOT CLEAN
- PROTECTED FOR PRIVACY
- PROTECTED AGAINST CORRUPTION

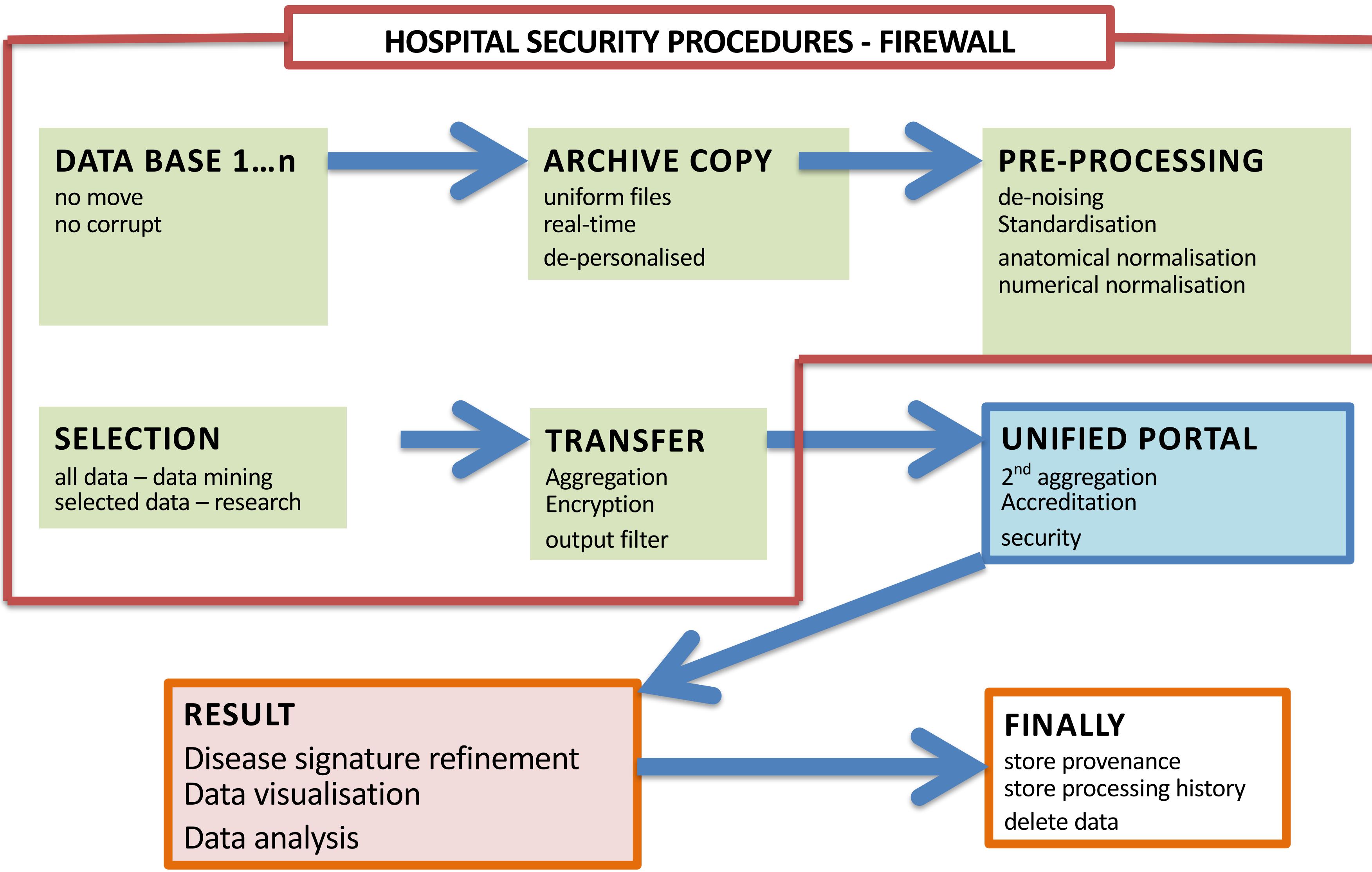
RESEARCH DATABASES

- PROTECTED CULTURALLY

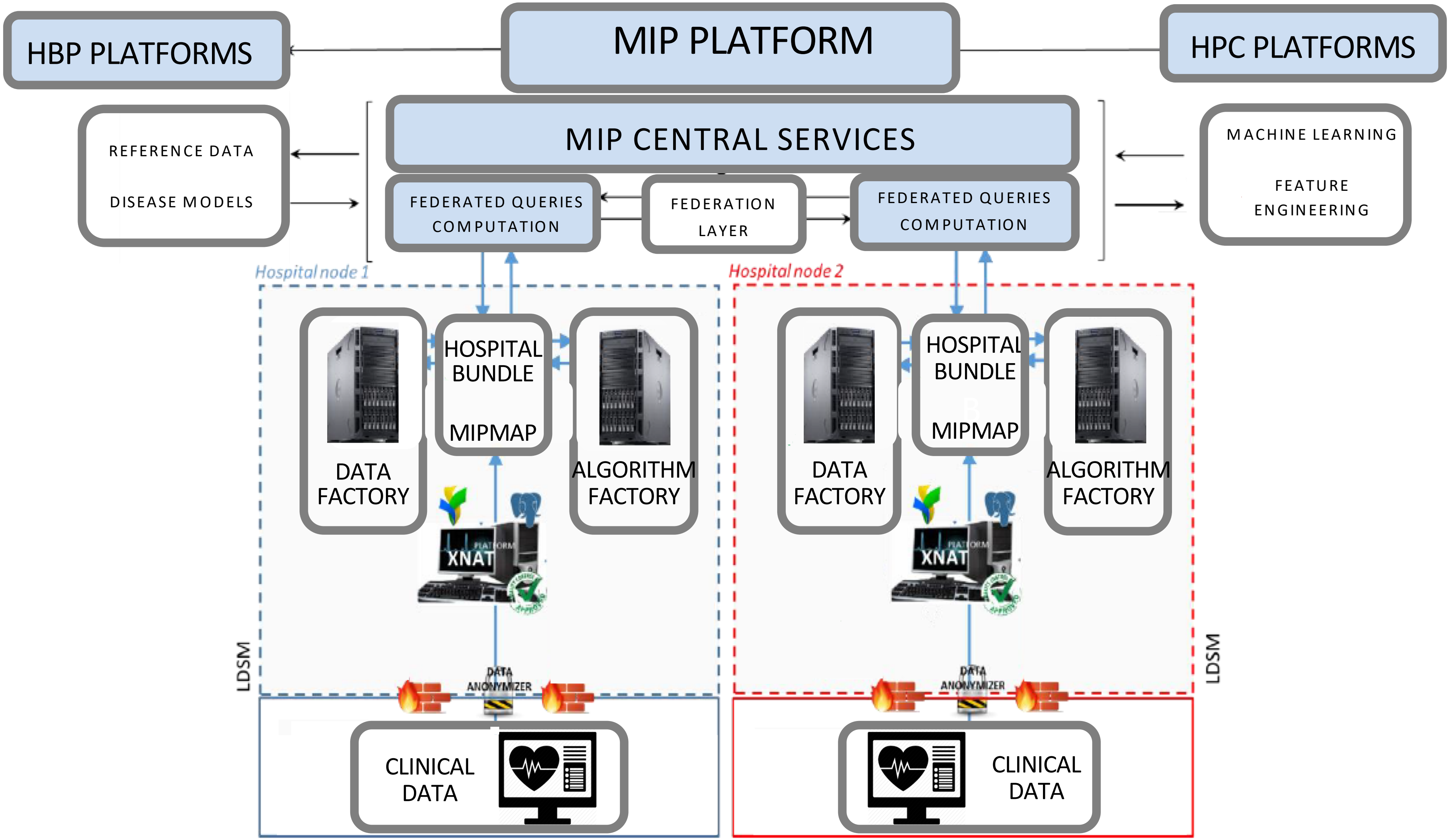
PHARMACEUTICAL DATABASES

- PROTECTED COMMERCIALLY

MEDICAL INFORMATICS PLATFORM
FEDERATING DATA



MEDICAL INFORMATICS PLATFORM



ETHICAL CHALLENGES

PRIVACY

- DE-PERSONALISATION
- ANONYMISATION

CONSENT

- BROAD CONSENT
- RETROSPECTIVE - PROSPECTIVE

MANAGEMENT OF ETHICS

- LOCAL ETHICS COMMITTEES
- VALUE AND CREDIBILITY OF SCIENCE

SUBJECTS & METHODS

We used 912 AD subjects – **ADNI DATABASE**

For a subsample of 508 we knew gender and age
For a subsample of 184 we knew the MMSE score

We used 5566 normal individuals – **THREE CITIES EPIDEMIOLOGICAL STUDY, FRANCE**

For a subsample of 2096 we knew gender and age
For a subsample of 2091 we knew the MMSE score

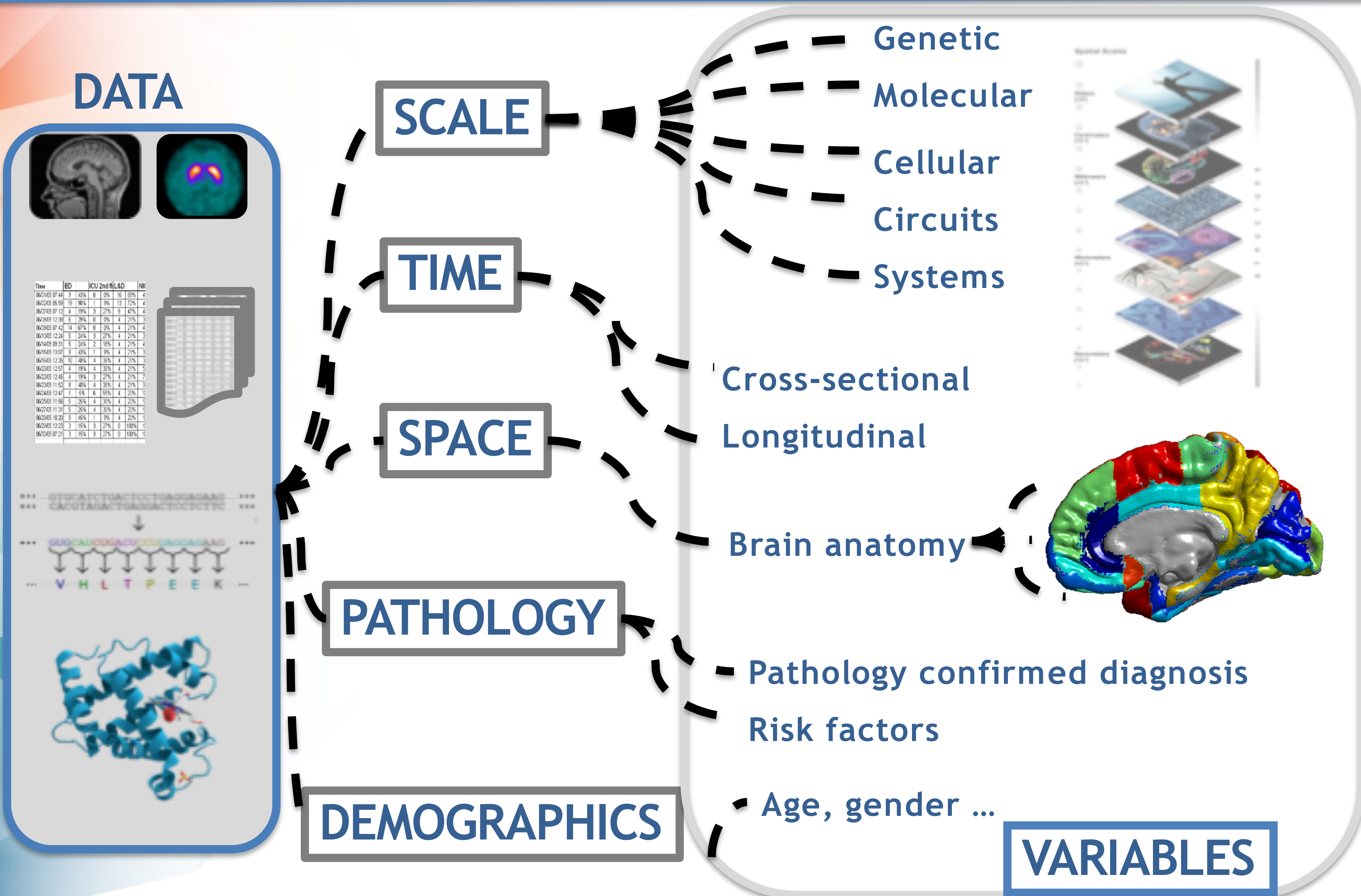
For learning we used half the dataset to create the classifier

The learning set = 3239 individuals (465 AD, 2774 controls)

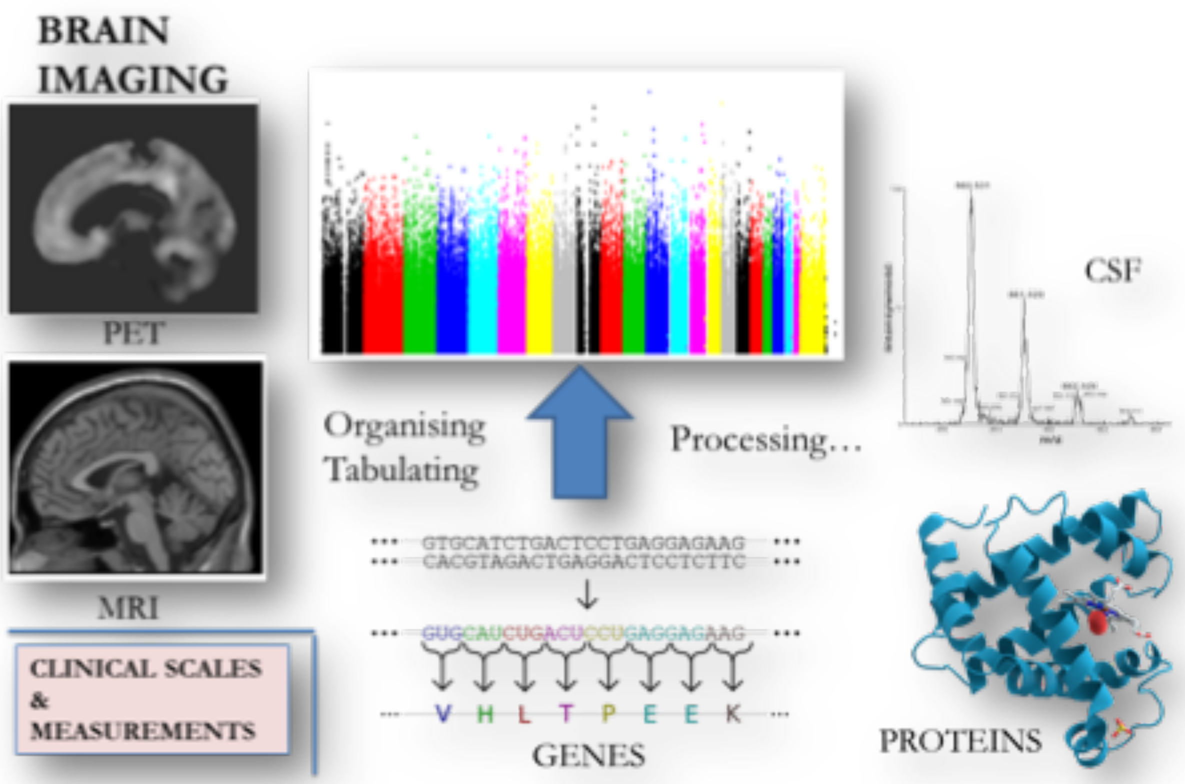
The other half was used to validate the classifier

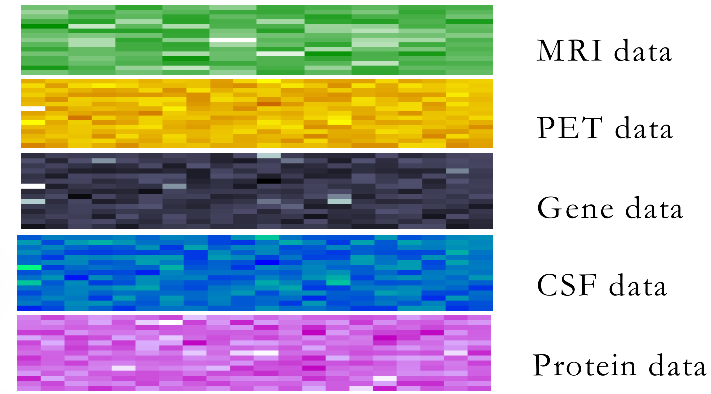
The testing set = 3239 individuals (447 AD, 2792 controls)

DATA EXTRACTION



DATA INTEGRATION

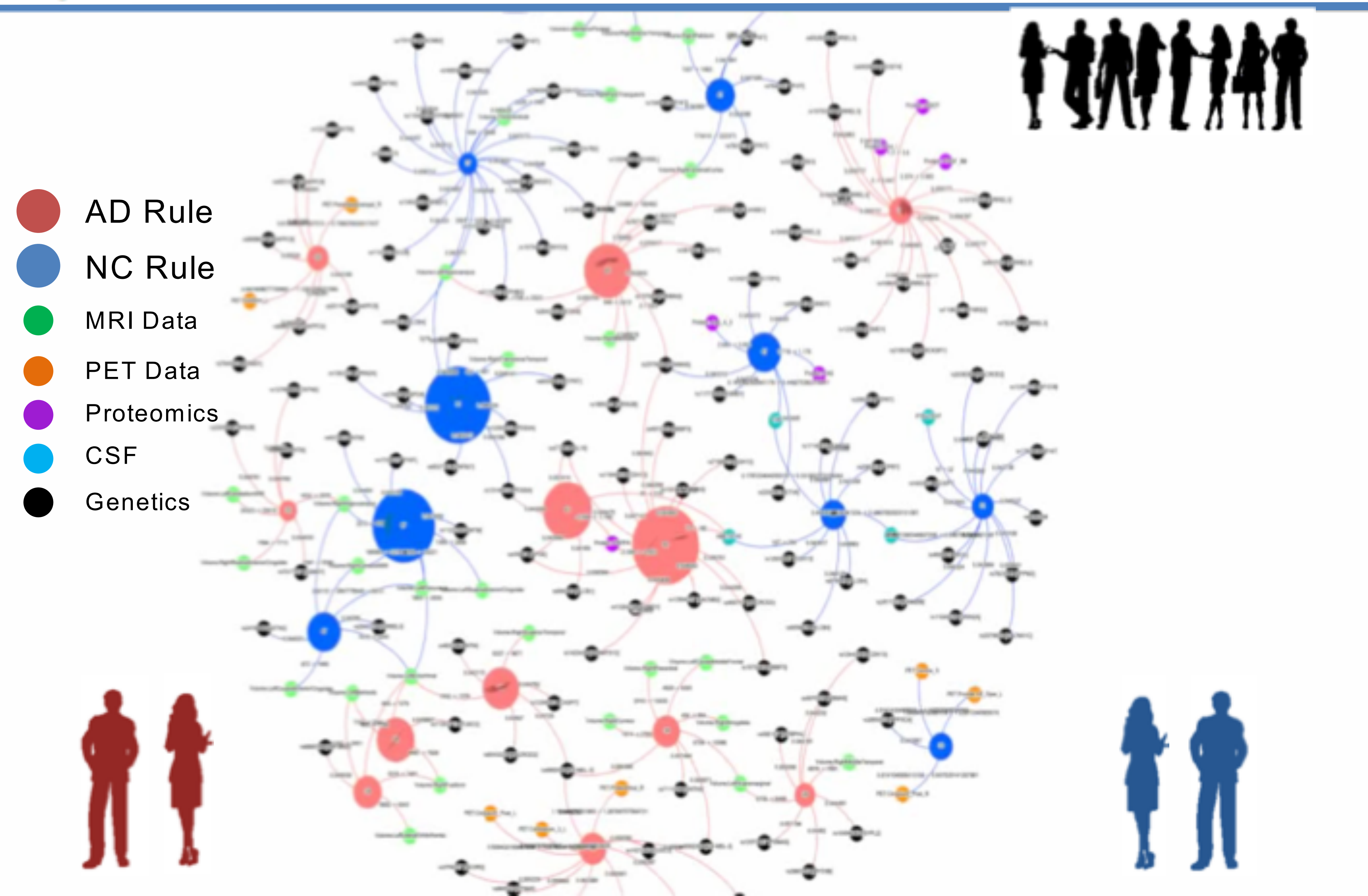




912 Alzheimer's patients
5566 Healthy controls

- Phenotype-led Semi-supervised clustering
- Biologically led classification
- High dimensional feature learning

LINKING CLINICS TO BIOLOGY

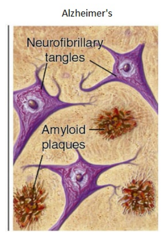


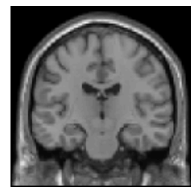
USING ANATOMY TO CONSTRAIN DIAGNOSIS

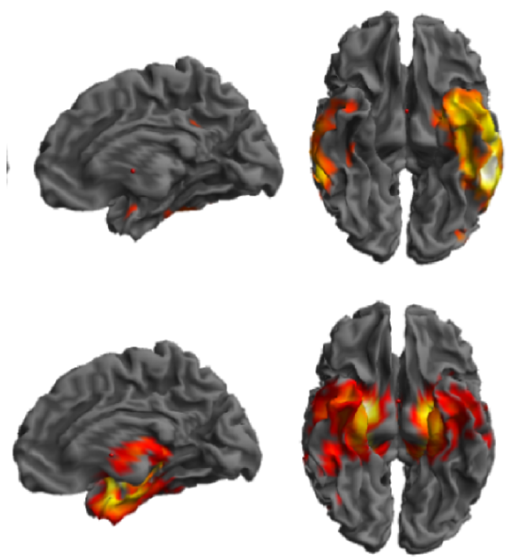
CLINICAL (SYNDROMIC) CLASSIFICATION

AD

HC







INTEGRATING PATHOLOGY INTO DIAGNOSIS

AD +
Symptoms +
Pathology +

AD -
Symptoms +
Pathology -

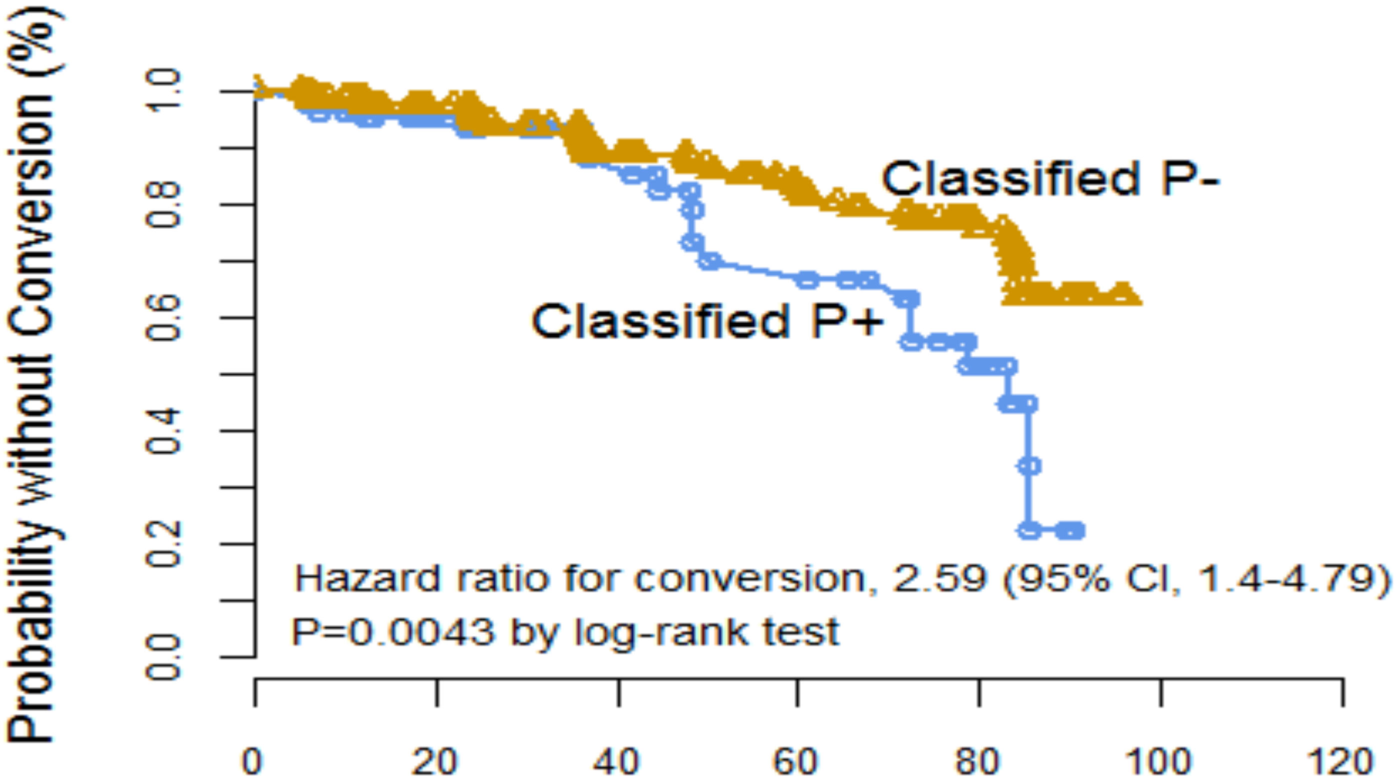
HC +
Symptoms -
Pathology +

HC -
Symptoms -
Pathology -

Brain atrophy pattern characteristic of pathological disease provides constraints on diagnosis

PREDICTING OUTCOME MORE ACCURATELY

NC subjects with two patterns of brain atrophy



ALL COGNITIVELY NORMAL ON RECRUITMENT AT BASELINE = 0

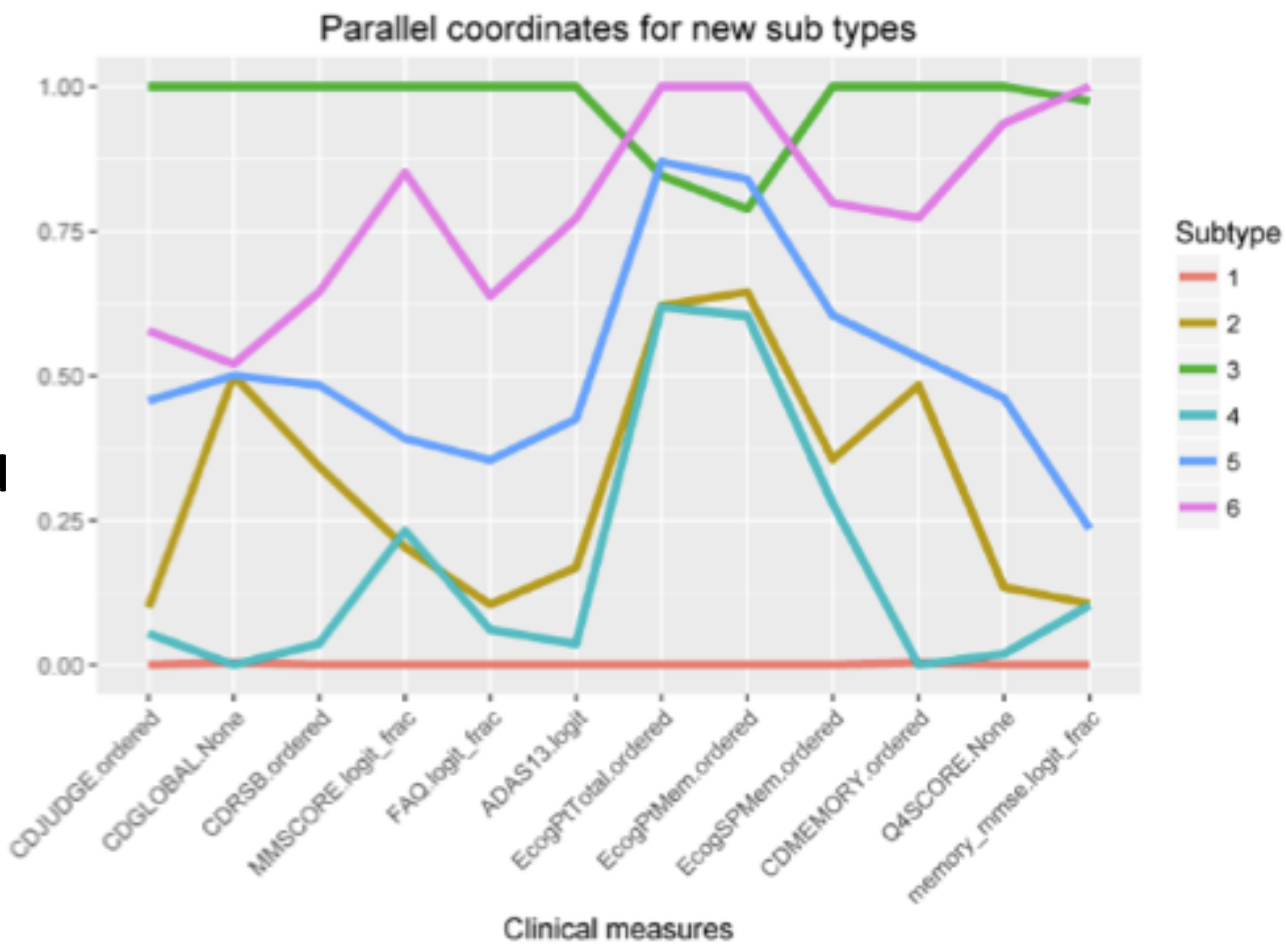
Result in terms of K-M survival

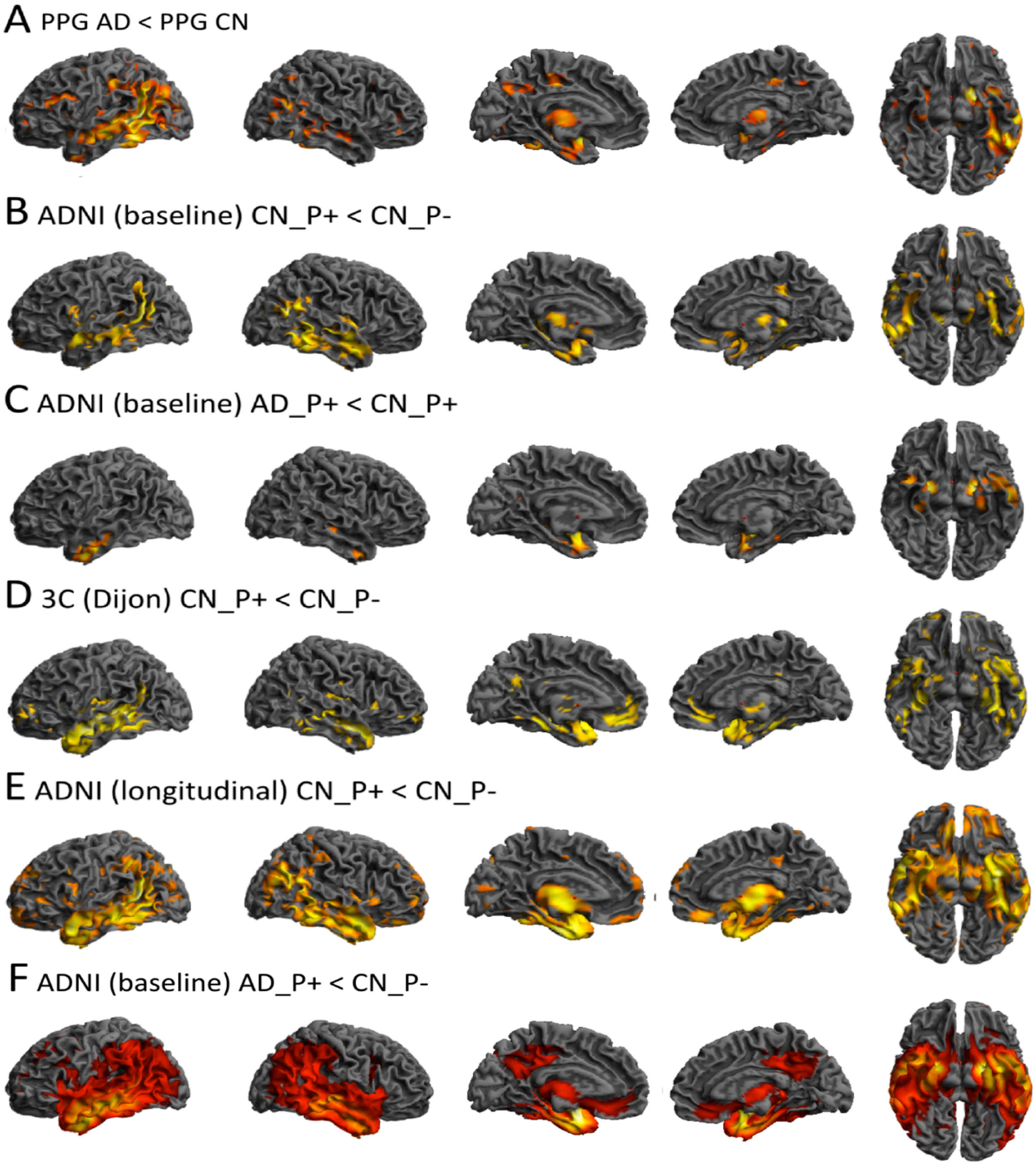
Jing Cui 2016 - in preparation

SUPERVISED SEARCH FOR AD SUBTYPES IN ADNI DATA

Six subtypes identified

1. Typical AD
2. Atypical AD
3. Hippocampus spared
4. Amnestic MCI
5. Depression suspected
6. Cognitively Normal



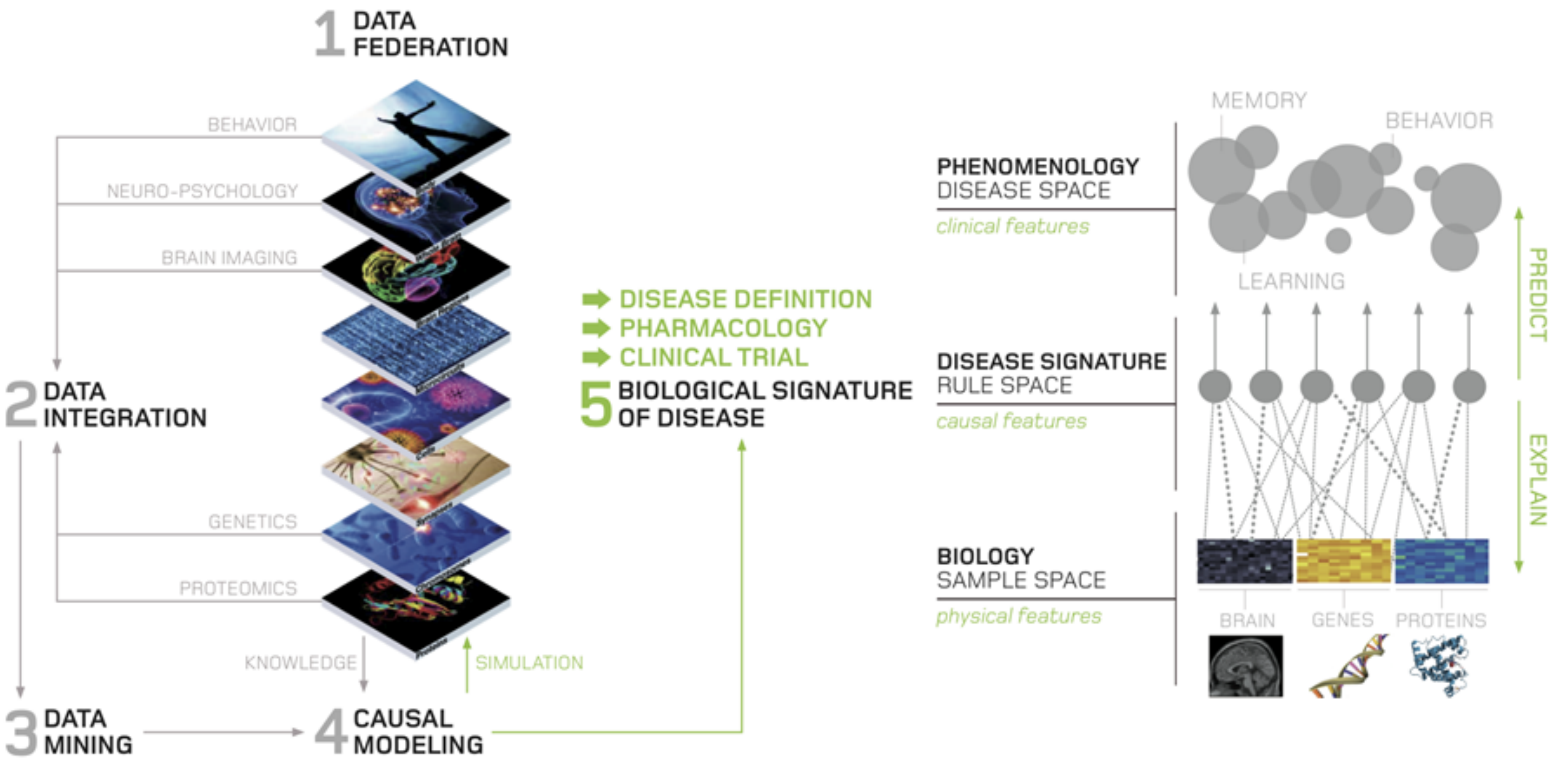


DISSECING CLINICAL GROUPS

A UNIVERSE OF PERSONALISED DIAGNOSES

MULTI-SCALE DISEASE SIGNATURES

The MIP provides methods to analyse federated data from hospitals, research centres and biobanks and aim to federate the different communities of users from these different locations.

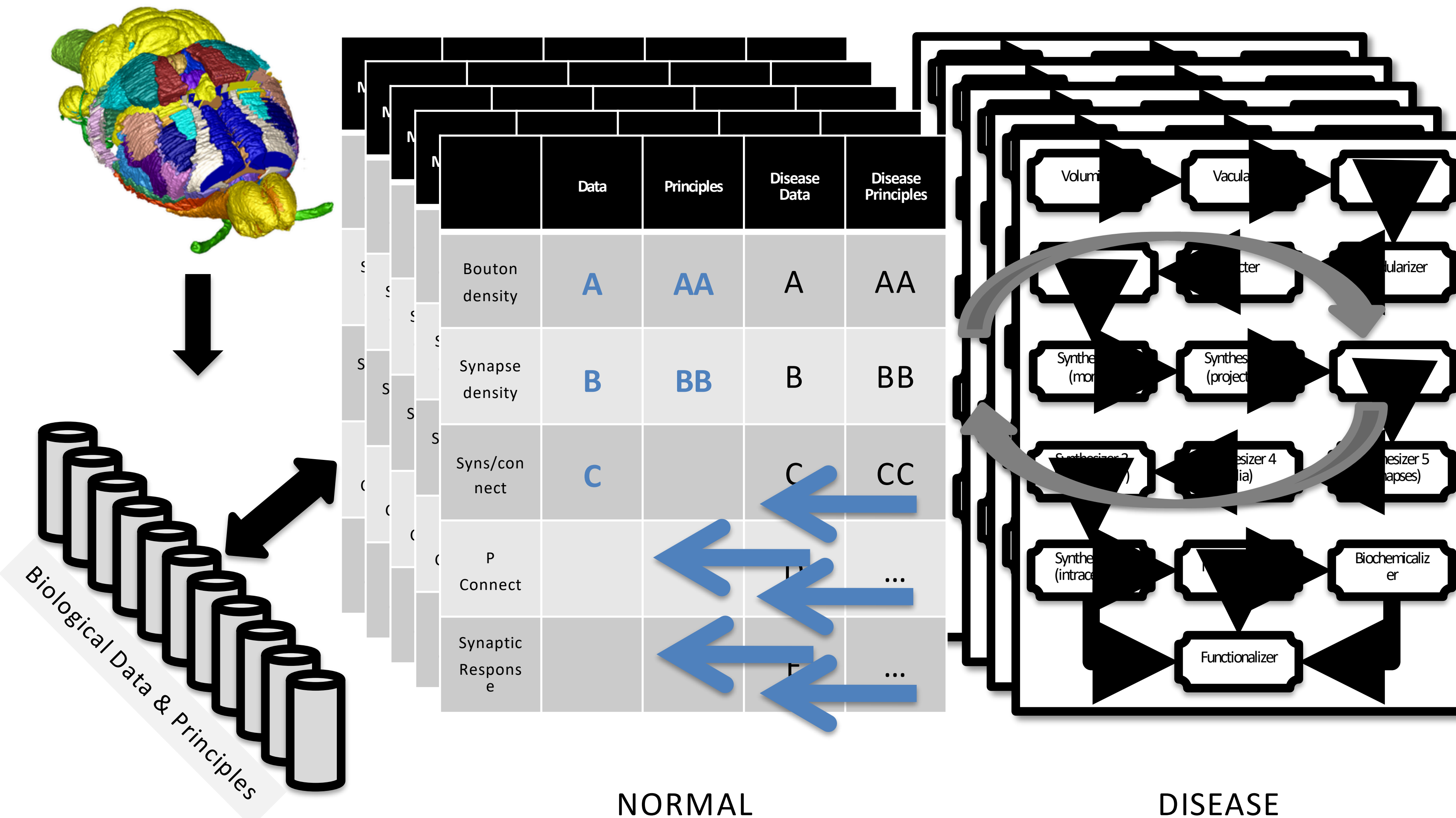


REVERSE ENGINEERING – TO VALIDATE

Brain Atlases
Data source

Biological Parameter Constraints & Biological Principles
Configurations

Multi-constraint Algorithms
Brain Reconstruction Workflows



DATA INTEGRATION

AND

CAUSAL ANALYSIS

DISEASE SIGNATURES

THEN

MEDICAL PHENOMENOLOGY

GENETIC & –OMIC CHARACTERISATION

DYSFUNCTIONAL SYSTEM IDENTIFICATION

(POLY)PHARMACOLOGY

ICONOCLASTIC CHALLENGES

A PIXEL is an INFORMATION ANCHOR with COORDINATES in STANDARD BRAIN ANATOMICAL SPACE THAT LINKS TO INFORMATION ABOUT IT IN THE LITERATURE

PATTERNS OF PIXEL ABNORMALITIES ARE OF DIAGNOSTIC and/or PROGNOSTIC SIGNIFICANCE

PATTERNS ARE THOUGHT TO BE MECHANISTICALLY SPECIFIC and SO MORE RELEVANT to THERAPEUTICS – TO BE PROVEN

THE FUTURE IS IN DATA-LED HYPOTHESIS GENERATION FOLLOWED by CLINICAL, BIOLOGICAL AND GENETIC CHARACTERISATION – DISEASE SIGNATURES

SUCH CHARACTERISATION SHOULD GENERATE THERAPEUTIC STRATEGIES AND TARGETS IN A MORE PRINCIPLED MANNER THAN AT PRESENT – TO BE PROVEN

RCTs MAY HAVE MORE FOCUSED USES – BIG DATA ANALYSES TO REPLACE DOUBLE BLIND CONTROLS BY REMOVING NOISE - MASSIVE DATA SETS DO NOT NEED TO BE COMPLETE, CLEAN OR UNIFORM – TO BE VALIDATED

THE ROLE OF DISRUPTIVE SCIENCE IN HBP-MIP

1. Move to a “**no database**” federated data analysis infrastructure

   - ✓ Security, privacy, research, ethics considerations
   - ✓ Advances in “virtualisation” “streaming” and “peer-to-peer” technologies
   - ✓ Use of products of EC funded research (eg Exareme)
   - ✓ Open source and cross-disciplinary specification
   - ✓ Unlocking hospital databases for research

2. Breaking **conservative** medical IT culture

   - ✓ Recruitments of university hospitals
   - ✓ Recruitment of structured research databases
   - ✓ Playing to “data sharing” revolution (NIH, EC, Wellcome initiatives)

3. Introduction of “**disease signatures**” concept

   - ✓ Cultural change from pure symptomatic & syndromic disease definitions
   - ✓ Preliminary classifications

HBP MEDICAL INFORMATICS PLATFORM

- Clinical neuroscientist
- Computer scientist
- Statistician neuroscientist
- Ethics

Bogdan Draganski
Ferath Kherif

Richard Frackowiak
Mira Marcus Kalish

Saso Dzezowski
Boudewijn Lelieveldt

Anastasia Ailamaki
Vasilis Vassalios

Yannis Ioannidis
Frank Schneider

Andrew Pocklington
John Ashburner

Alexis Brice
Kathinka Evers

Jean-Francois Dartigues

Giovanni Frisoni
Yoav Benjamini

Nada Lavrac
Thomas Heinis