Mechanism-based Stratification of Patients in Neurology: Lessons learned from the IMI-AETIONOMY project

Martin Hofmann-Apitius

Academic Co-Coordinator of AETIONOMY

Talk at Alzheimer Europe Conference, Den Hague,

October 24, 2019

innovative medicines initiative

Mission

To increase knowledge of the causes of Alzheimer´s and Parkinson´s Disease by generating a mechanism-based taxonomy; to validate the taxonomy in a prospective clinical study that demonstrates its suitability for identifying patient subgroups (based on discrete disease mechanisms); to support future drug development and lay the foundation for improved identification and treatment of patient subgroups currently classified as having AD or PD.

Inspired by patients. Driven by science.

Fraunhofer SCAI

efpia

AETIONOMY

The Concept of Mechanism-Based Taxonomies

In 2011, Kola and Bell published a remarkable paper in Nature Reviews Drug Discovery. With their **"Call to reform the taxonomy of human disease"** they proposed a new, **mechanism-based classification of human disease.**

A call to reform the taxonomy of human disease

Ismail Kola and John Bell

A coordinated effort to incorporate advances in the understanding of the molecular and genomic variations in common diseases, such as hypertension, into their diagnosis and treatment could transform drug development and medicine.

Many common human diseases are still diagnosed as if they were homogenous entities, using criteria that have hardly changed for more than a century. For example, a person with a systolic blood pressure of 140 mm Hg or greater and a diastolic blood pressure of 90 mm Hg or greater is diagnosed with hypertension, irrespective of the heterogeneous underlying molecular mechanisms in different individuals. Furthermore, the treatment approach for diseases that are diagnosed in this way is generic, with empiricism as its cornerstone. Continuing with the example of hypertension, the standard initial based on the presence of a shared mutation and/or a deregulated pathway, rather than on tumour location, has not yet been initiated to our knowledge, but is an approach that regulatory agencies may be comfortable with in the future.

The lack of recognition of disease heterogeneity in clinical development and medical practice has a number of well-known consequences. First, it will probably reduce the likelihood of success of clinical trials, perhaps more so for targeted therapies that have been pursued in recent years. Indeed, if the pathway that is being targeted

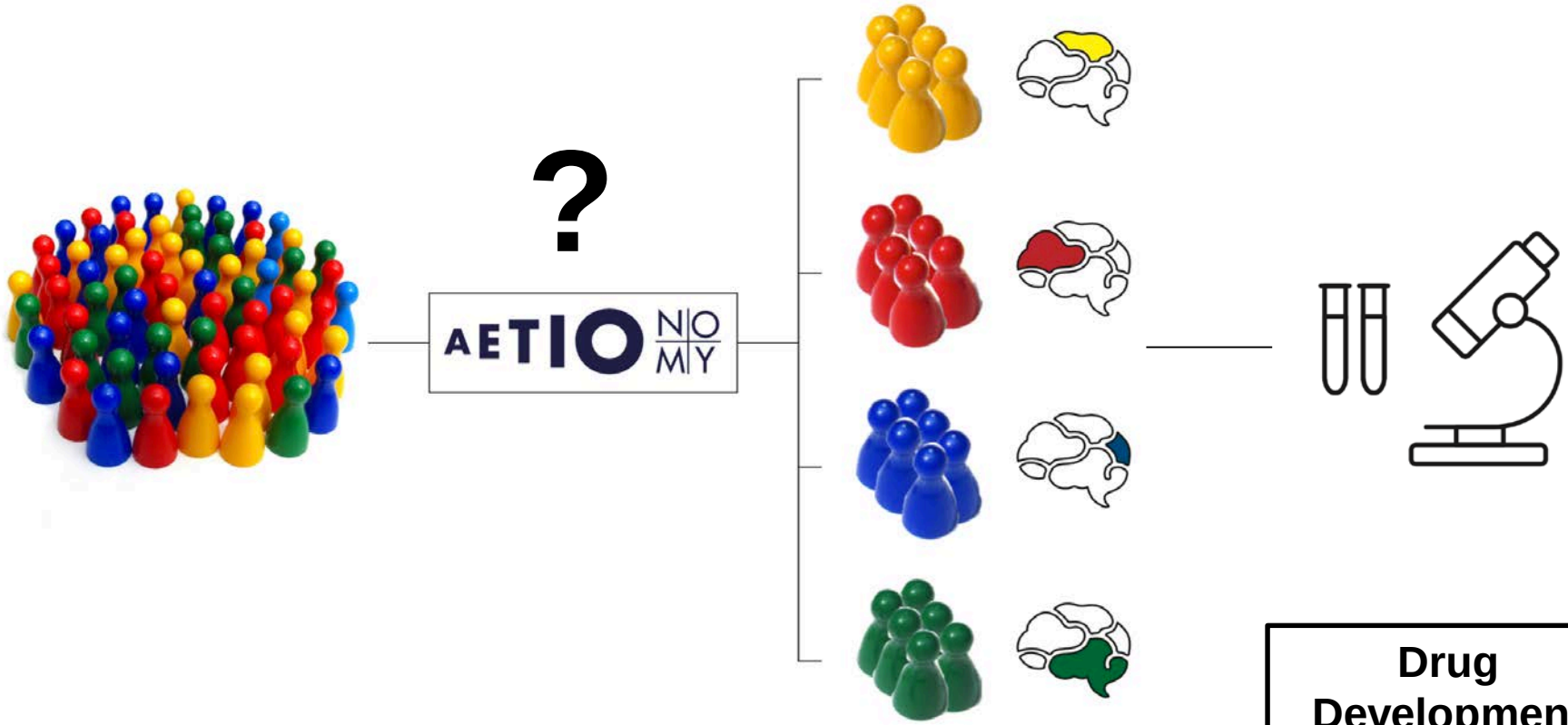
Kola, I., & Bell, J. (2011). A call to reform the taxonomy of human disease. *Nature Reviews Drug Discovery*, *10*(9), 641-642.

The Vision:

Stratifying Alzheimerism and Parkinsonism patients according to their individual (combinations of) pathophysiology mechanisms

AETIONOMY

Organising Knowledge about Neurodegenerative Disease Mechanisms for the Improvement of Drug Development and Therapy



Mixed Population

Stratified Population

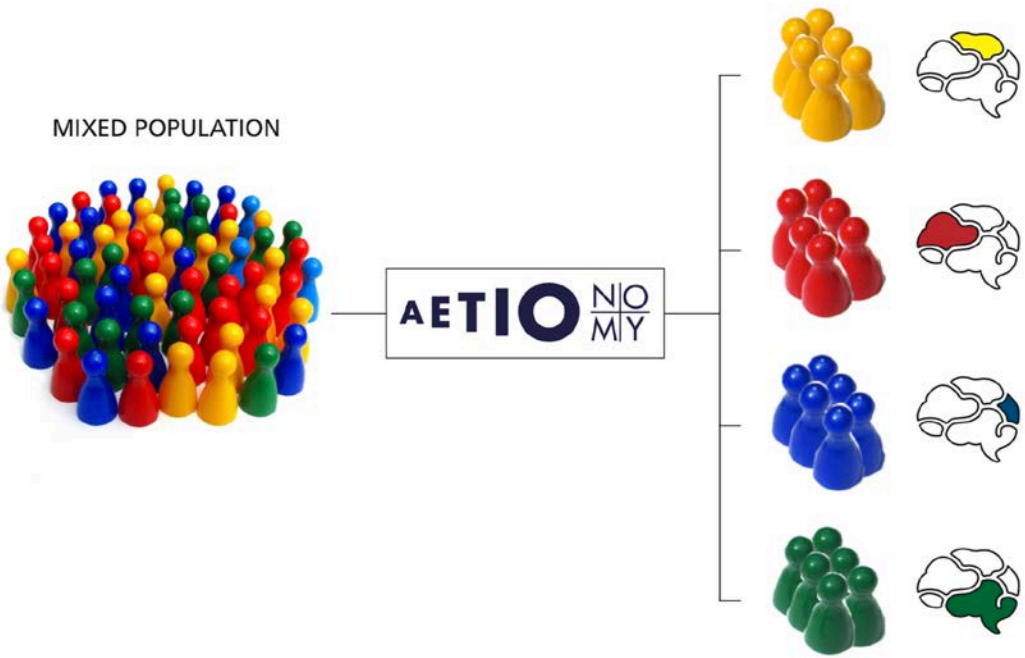
Drug Development and Therapy

ucb Fraunhofer SCAI efpia imi innovative medicines initiative AETIONOMY

The Reality:

Data and Knowledge about Pathophysiology Mechanisms are scattered, biased, heterogeneous and sometimes simply wrong.

Pathophysiology Mechanisms are Multimodal



- Molecular biomarkers
- Genetics
- Epigenetics
- Gene expression
- Proteomics
- “Pathway” dysregulation
- Cognition testing
- Imaging readouts
- Environment
- Sport
- Stress
- Published knowledge
- Expert knowledge

The Real Work:

What does it take to generate a “mechanism-based taxonomy of neurodegenerative diseases“?

The Original Concept

We thought:

- Let us do two different things:
  - Generate an **overview of mechanism hypotheses** (“pathophysiology graph”; “pathways”; “biomarkers”) and develop methods that test – on patient level data – whether patient subgroups can be associated with these mechanisms
  - Perform **unsupervised clustering and find patterns** in patient-level data that can be associated with “pathways” or “mechanisms”

- … that should provide us with a clue on patient subgroups ….

Fundamental Considerations

We need:

- A collection (an “inventory”) of **multimodal pathophysiology mechanisms** that can be tested (“challenged”) and validated by molecular and clinical study data.
- A comprehensive collection of available **patient-level data sets**, ideally longitudinal, rich in multimodal variables / readouts / features
- **Ways and methods to associate** pathophysiology mechanisms with the variables in clinical studies. (This may turn out to be non-trivial).
- Well-powered **data sets for validation**. If we can associate a multimodal pathophysiology mechanism with a subgroup of patients in a clinical study, we need to test the association in an independent clinical study.

The Problem-Solving Approach

Organising data and knowledge in the indication area and apply modelling and mining to gain new insights about disease mechanisms.

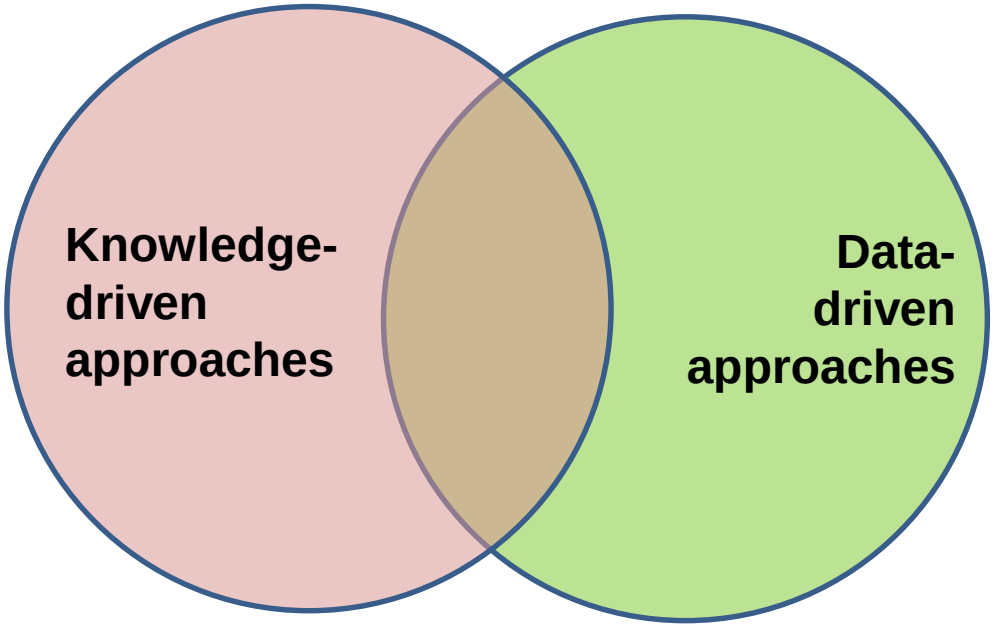
No large-scale new data generation, but rather:

work with what is out there.

AETIONOMY

Organising Knowledge about Neurodegenerative Disease Mechanisms for the Improvement of Drug Development and Therapy

- AETIONOMY KB
- **NeuroMMSig**
- Data Catalogue
- …



- Clustering
- Bayesian network models
- Longitudinal modeling
- tranSMART…

Hofmann-Apitius, M., et al. *International journal of molecular sciences* 16.12 (2015): 29179-29206.

Strategy .... and Implementation

The Challenge:

- A collection (an “inventory”) of **multimodal pathophysiology mechanisms** that can be tested (“challenged”) and validated by molecular and clinical study data.

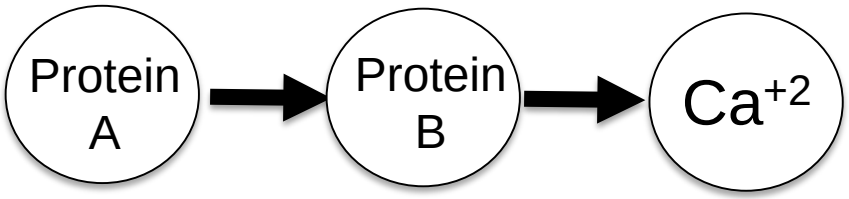
The Problem-Solving Approach:

- Systematic modeling of pathophysiology mechanisms using a dedicated graph-based modeling language. This resulted in **NeuroMMSig**, the “mechanism-enrichment server” for neurodegenerative diseases*.

NeuroMMSig*

*Domingo-Fernández, D., *et al*. "Multimodal Mechanistic Signatures for Neurodegenerative Diseases (NeuroMMSig): a web server for mechanism enrichment." *Bioinformatics* (2017).

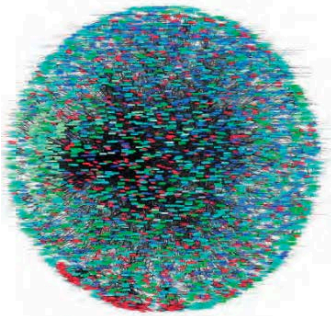
Capturing the disease-specific knowledge from literature

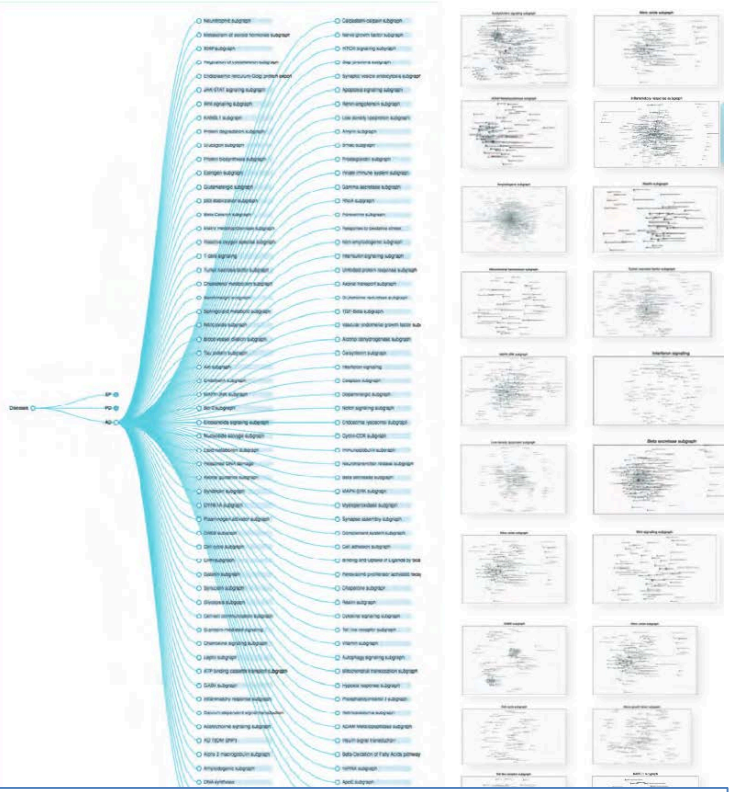


"[Protein A] increases [Protein B] leading to an increase in [Ca^{+2}]"

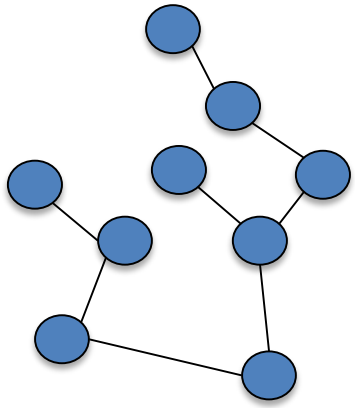
NeuroMMSig

Capturing the disease-specific knowledge from literature





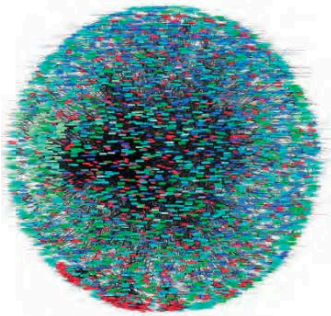
Classifying each relation in the network into the mechanism(s) they participate

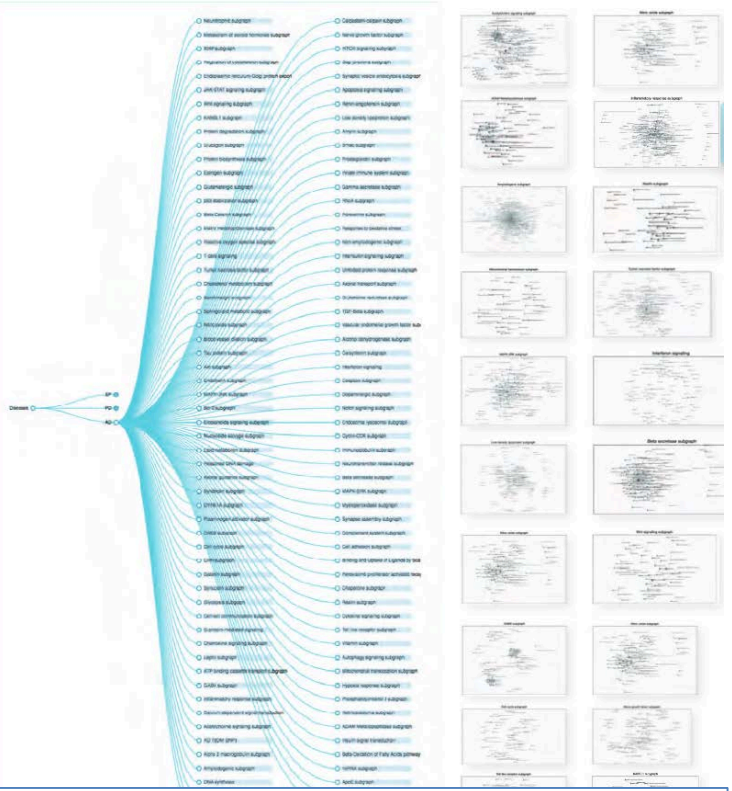


“[Protein A] leads to the generation of amyloid plaques”

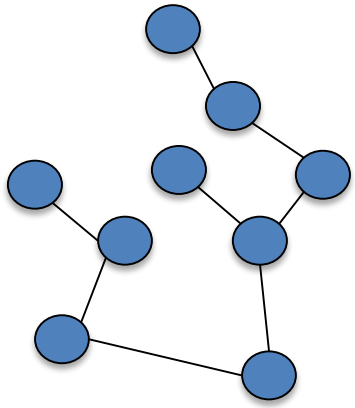
NeuroMMSig

Capturing the disease-specific knowledge from literature





Classifying each relation in the network into the mechanism(s) they participate

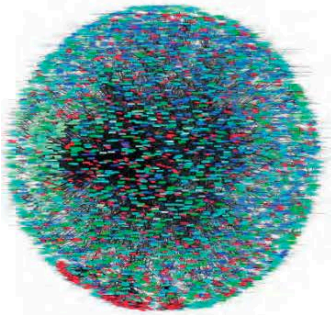


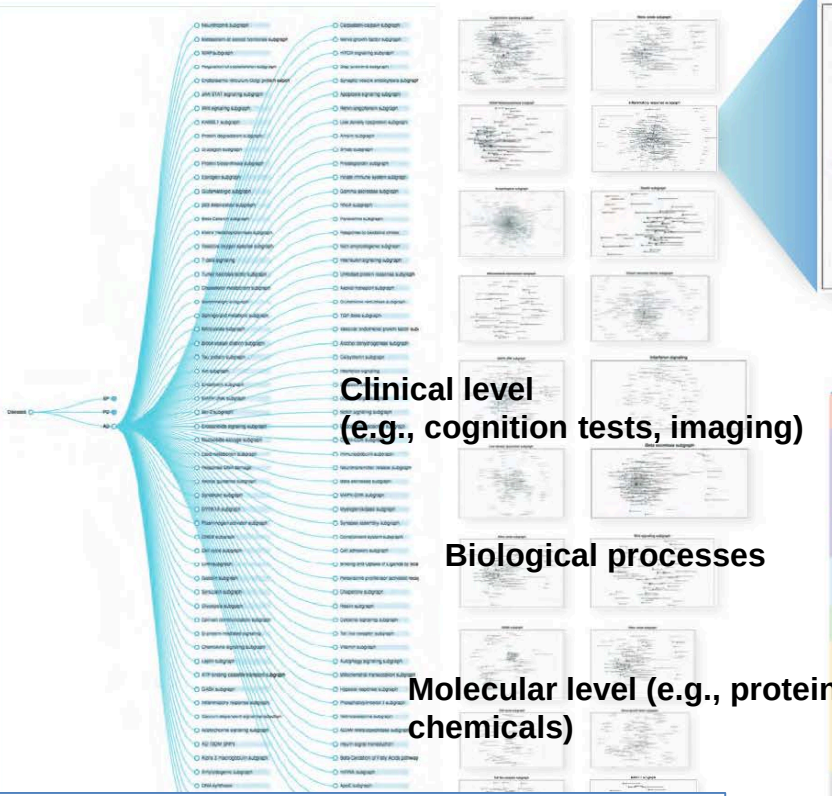
Amyloid cascade network

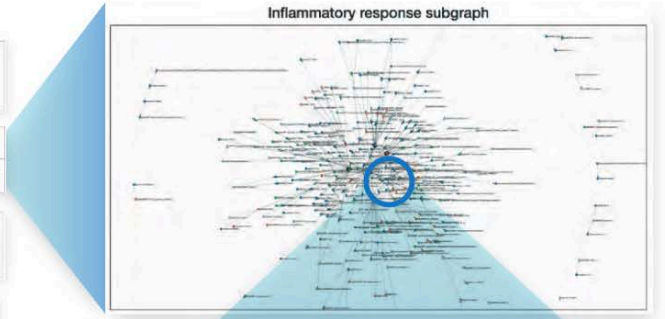
NeuroMMSig

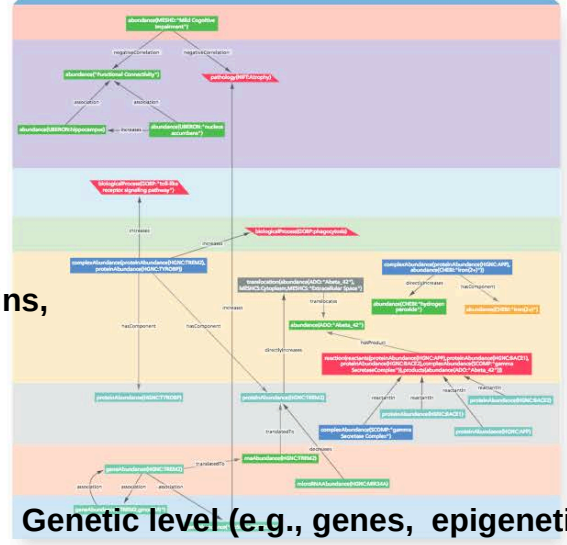
Neuroinflammation mechanistic subgraph representing multiple biological levels

Capturing the disease-specific knowledge from literature









Clinical level (e.g., cognition tests, imaging)

Biological processes

Molecular level (e.g., proteins, chemicals)

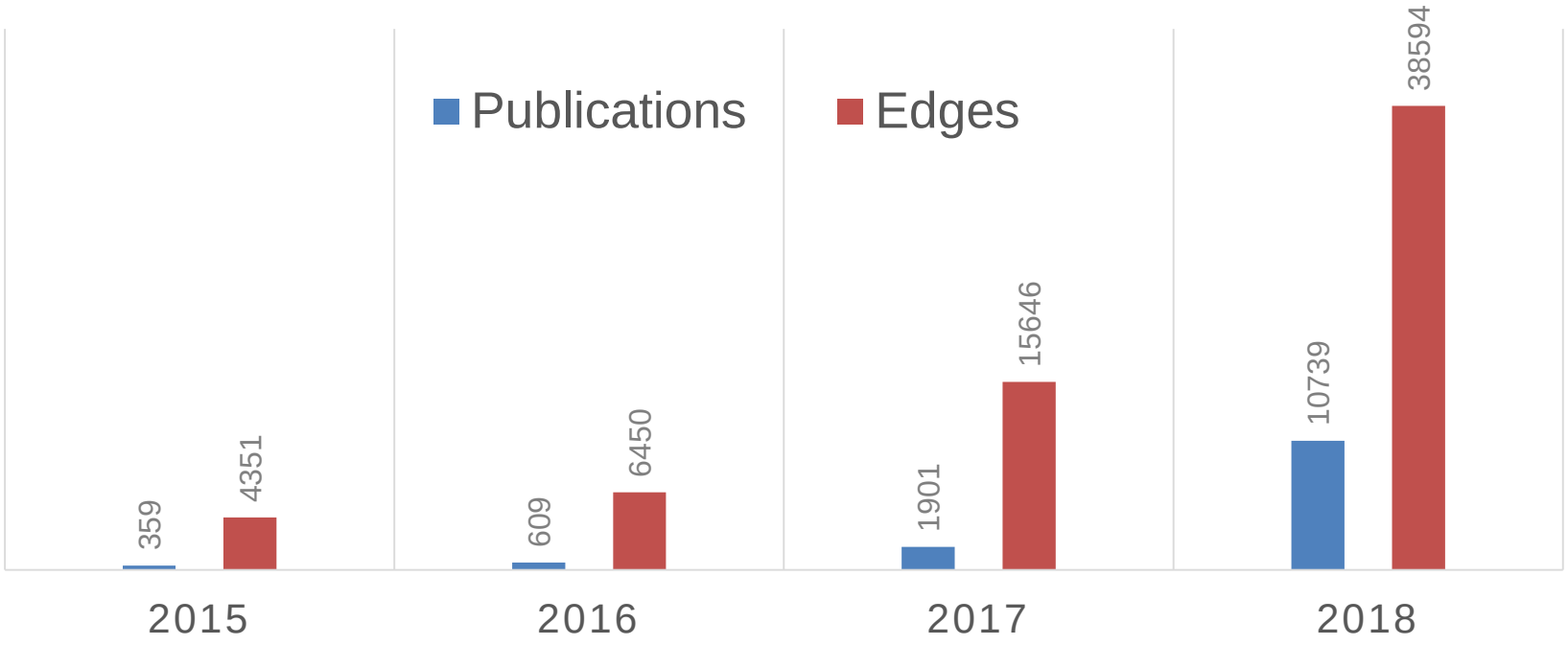
Genetic level (e.g., genes, epigenetics, variations)

Classifying each relation in the network into the mechanism(s) they participate

ucb Fraunhofer SCAI efpia imi innovative medicines initiative AETIONOMY

Growth during the project

NeuroMMSig AD



Online version

Production version

NeuroMMSig – Overview

- Comprises a candidate **mechanism collection** from three of the major neurological disorders
  - Alzheimer‘s (126), Parkinson‘s (76), and epilepsy (31)
  - (PTSD/TBI graphs with Cohen Veterans Bioscience)

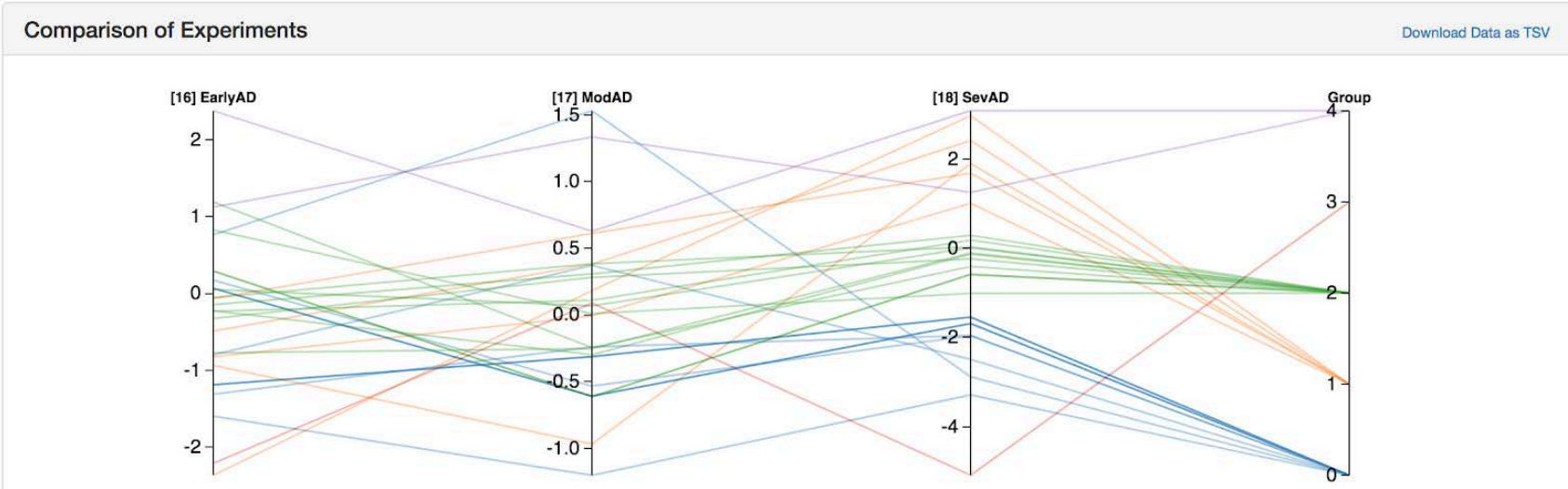
- **High resolution, disease-specific** pathophysiology graphs
  - Opposite to generalistic pathway databases such as KEGG and Reactome

- Candidate mechanisms are **computable networks**
  - Data can be used to contextualize hypotheses
  - Algorithmic and query functionalities built-in

NeuroMMSig - Analytics

1. Novel visualization
2. Data storage
3. Novel algorithmic implementations for patient stratification

*Hoyt, C. T., Domingo-Fernández, D, and Hofmann-Apitius, M. (2018). BEL Commons: an environment for exploration and analysis of networks encoded in Biological Expression Language. *Database.* In press (available online)



Which mechanism are up/down-regulated in different AD stages?

Clustering of mechanism trajectories on longitudinal data (e.g., MCI, AD)

Short Summary of Part I

- AETIONOMY has

  - generated **the largest inventory of disease mechanisms** for neurodegenerative diseases worldwide
  - these disease mechanisms are **represented in computable models** comprising cause-and-effect relationships
  - disease mechanism representations are inherently **multiscale and multimodal** and may integrate genetic variation information and imaging features in one graph
  - we made the **computable disease mechanism graphs freely accessible** though NeuroMMSig, the mechanism-enrichment server
  - the server is currently extended by dedicated algorithms and methods that support the **interpretation of patient-level data**

Dependencies on the Work of others ....

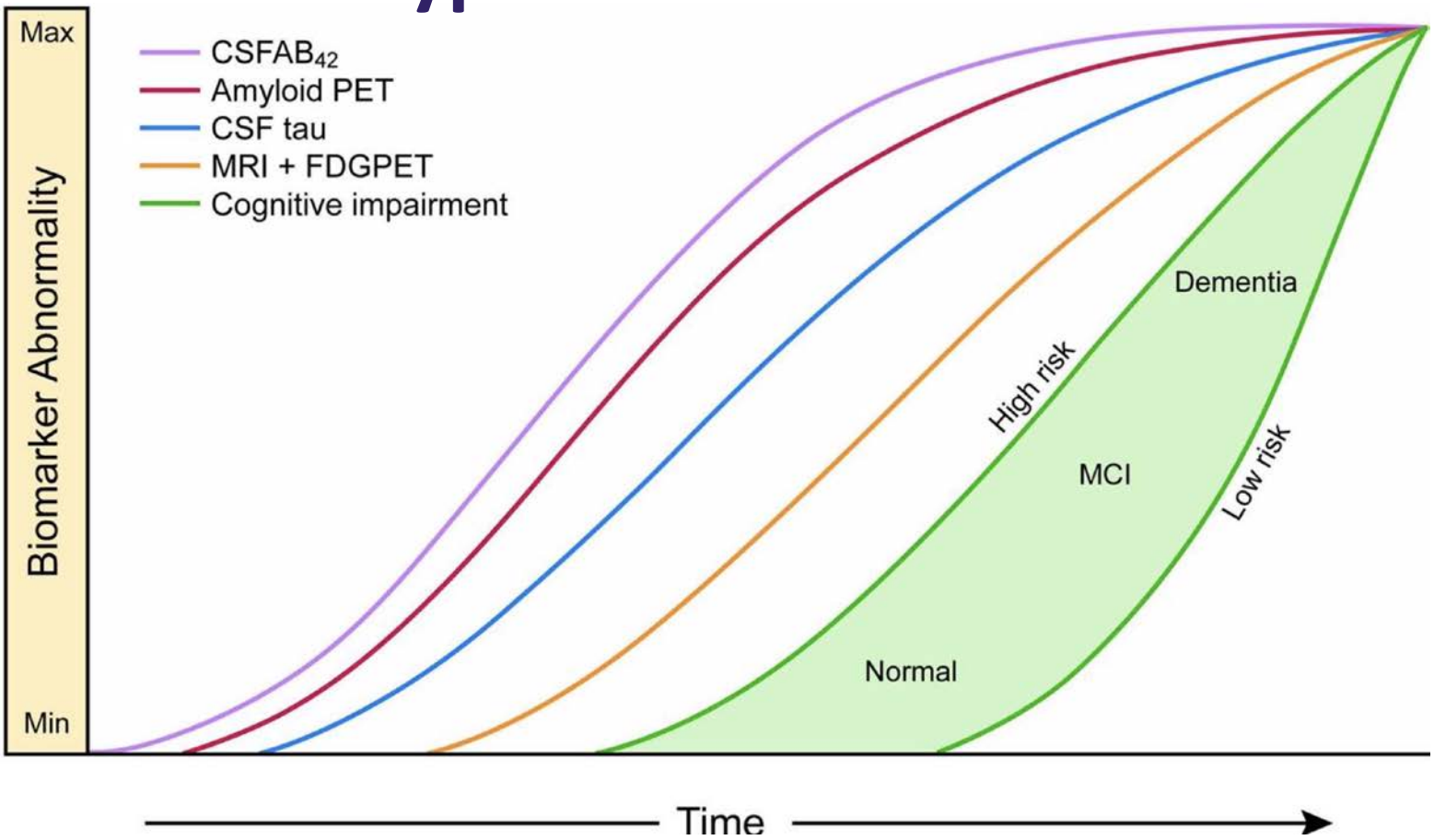
The Challenge:

- A comprehensive collection of available **patient-level data sets**, ideally longitudinal, so that we know, what “signature” of biomarkers is associated with disease progression (or disease risk)

The Problem-Solving Approach:

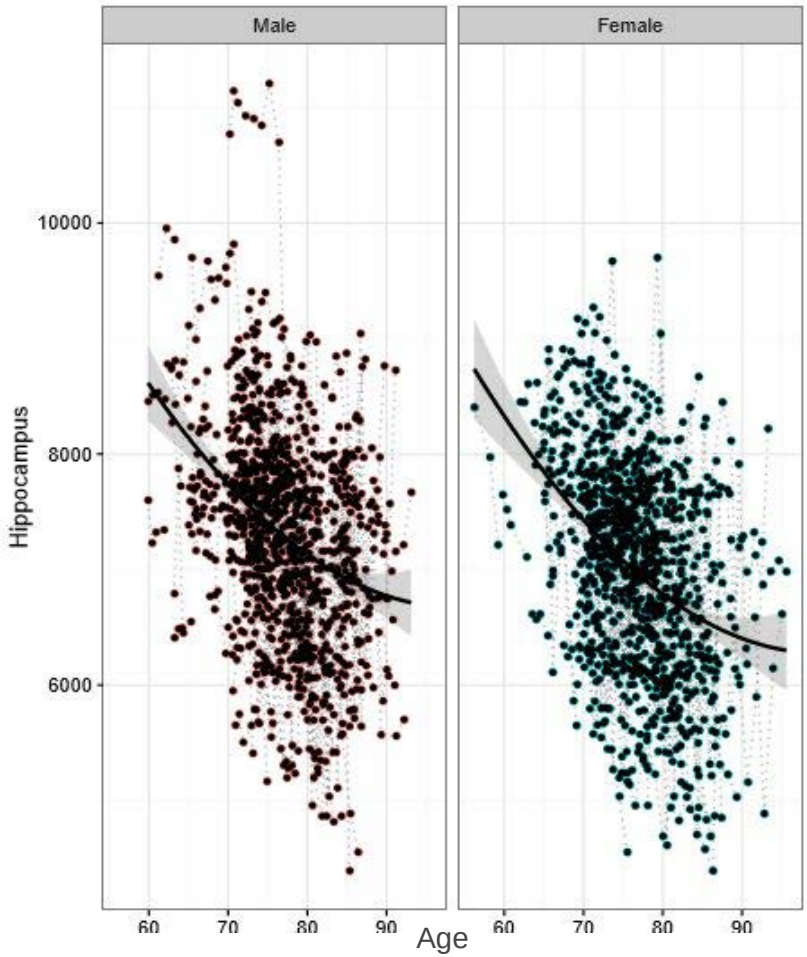
- Systematic harvesting, curation, pre-processing and comparative analysis of public patient-level data in AD and PD (ADNI, AddNeuroMed, AIBL, PPMI; others in preparation)
- Recruitment of the AETIONOMY PD cohort
- Alignment with other projects of the IMI AD platform (EMIF-AD and EPAD)

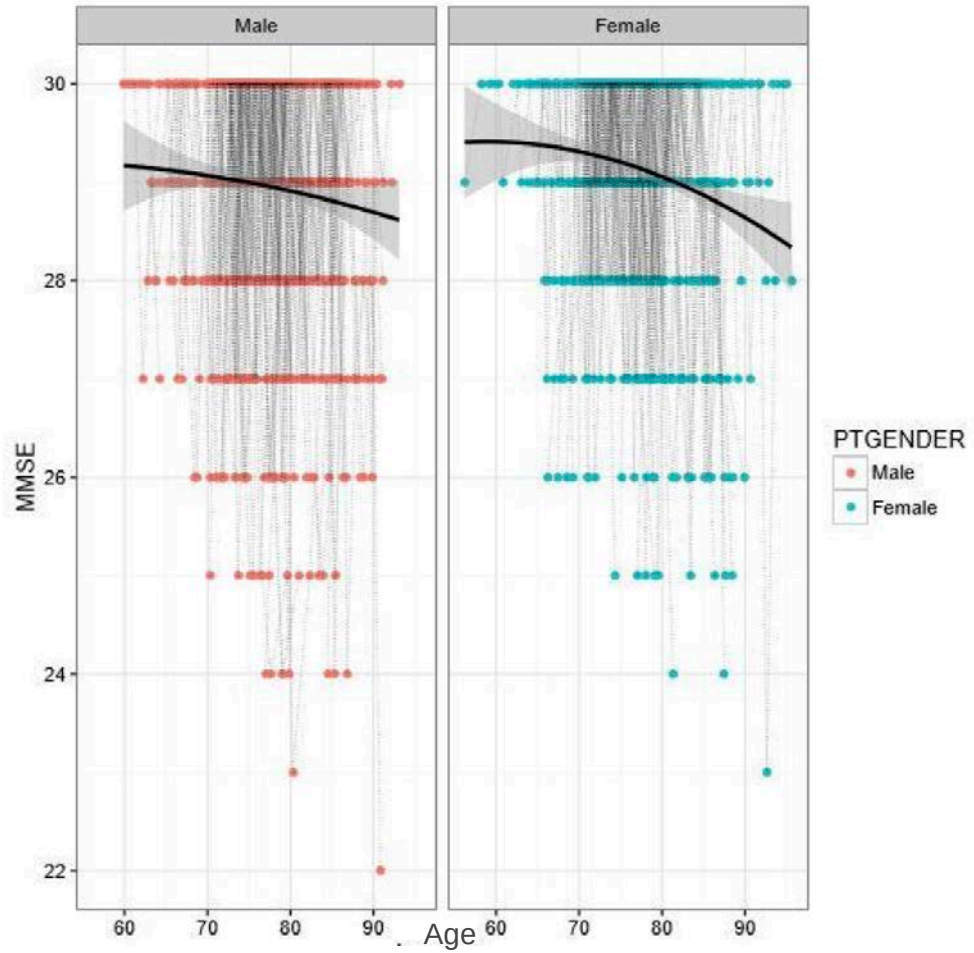
Hypothetical Model



"Update on hypothetical model of Alzheimer's disease biomarkers," Lancet neurology, vol. 12, no. 2, p. 207, 2013.

Reality Check



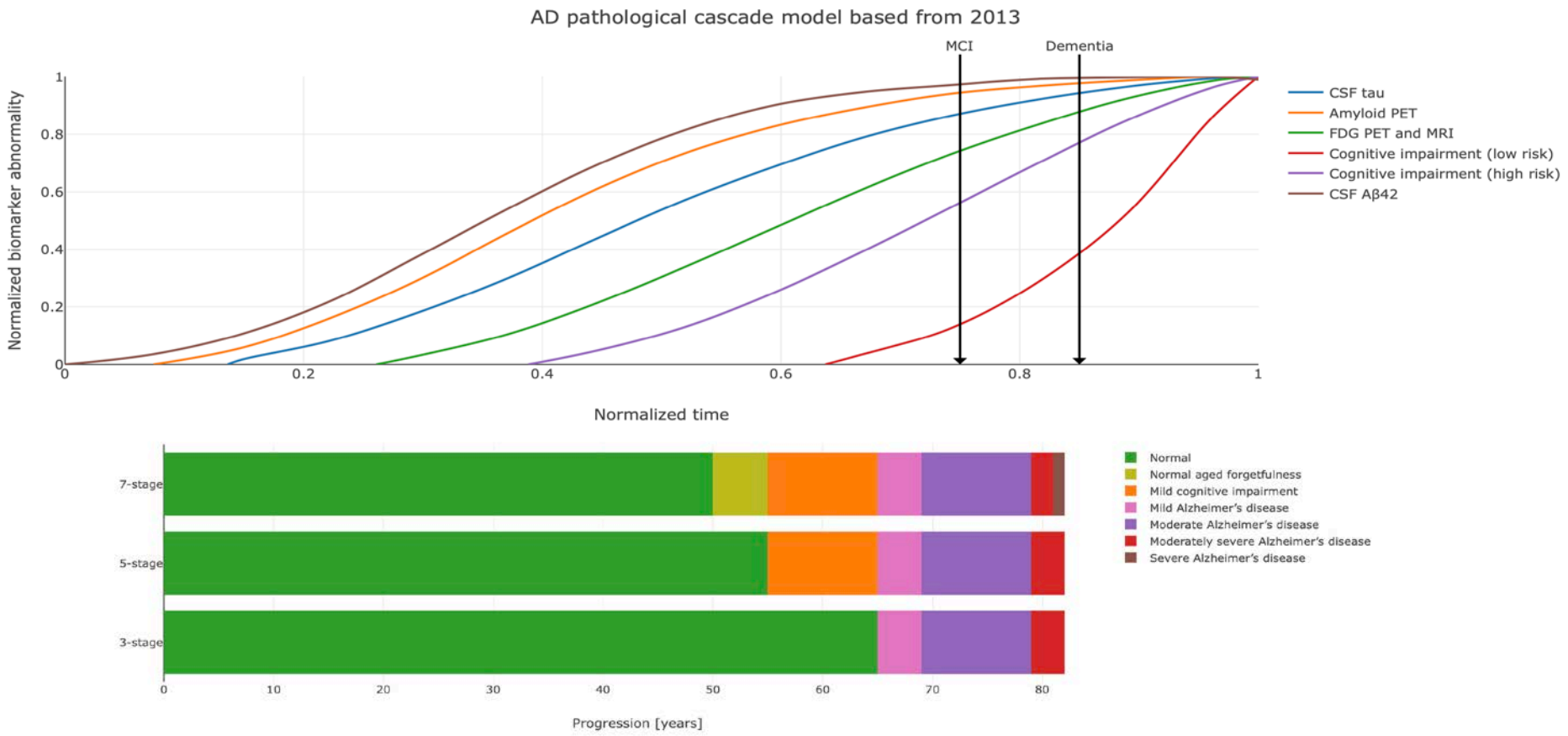


Hypothetical Model vs. Reality

Fundamental Question:

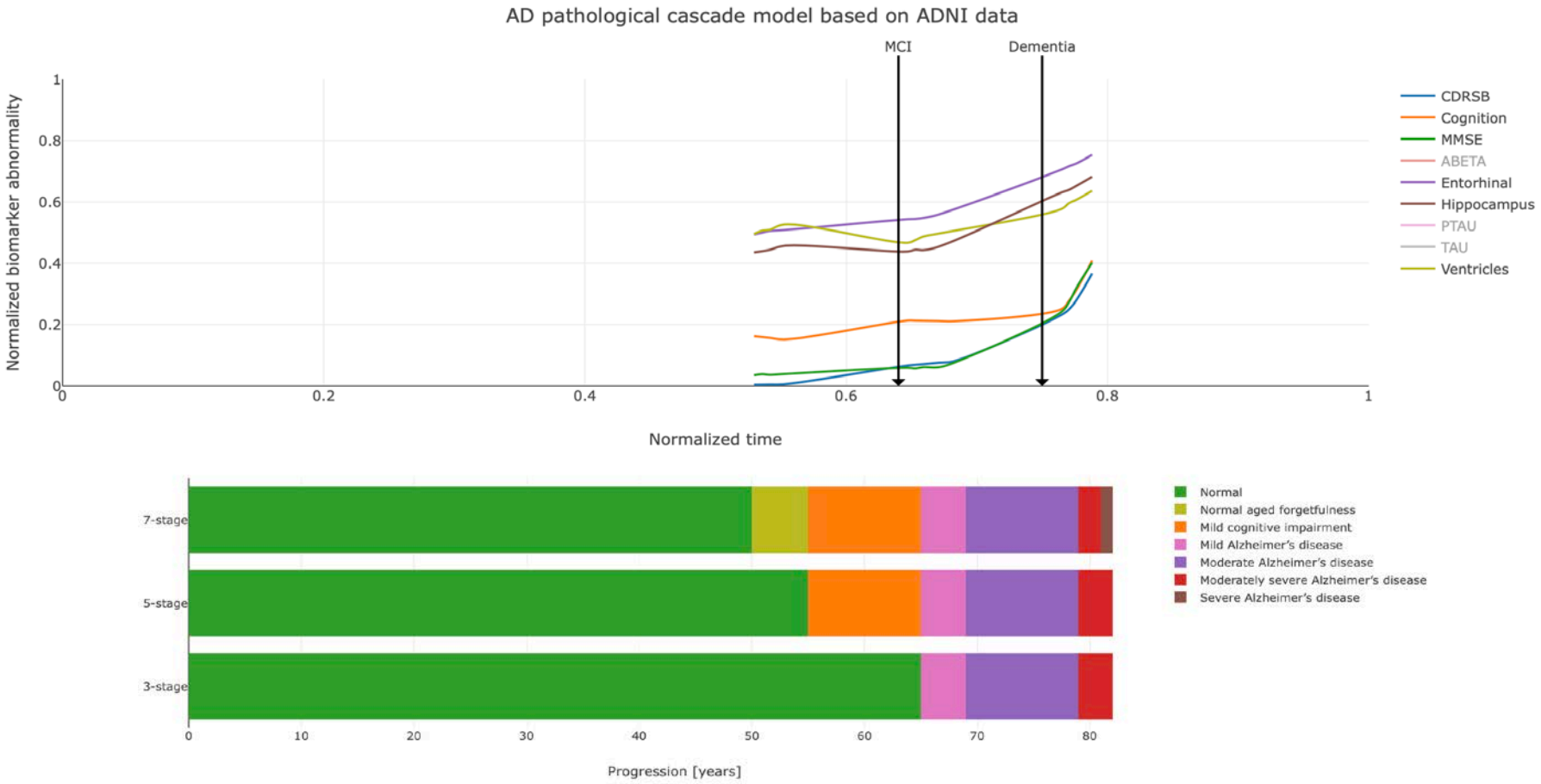
- Do ADNI biomarkers show the same trajectories like the hypothetical model published for AD?

Hypothetical Model: Computable



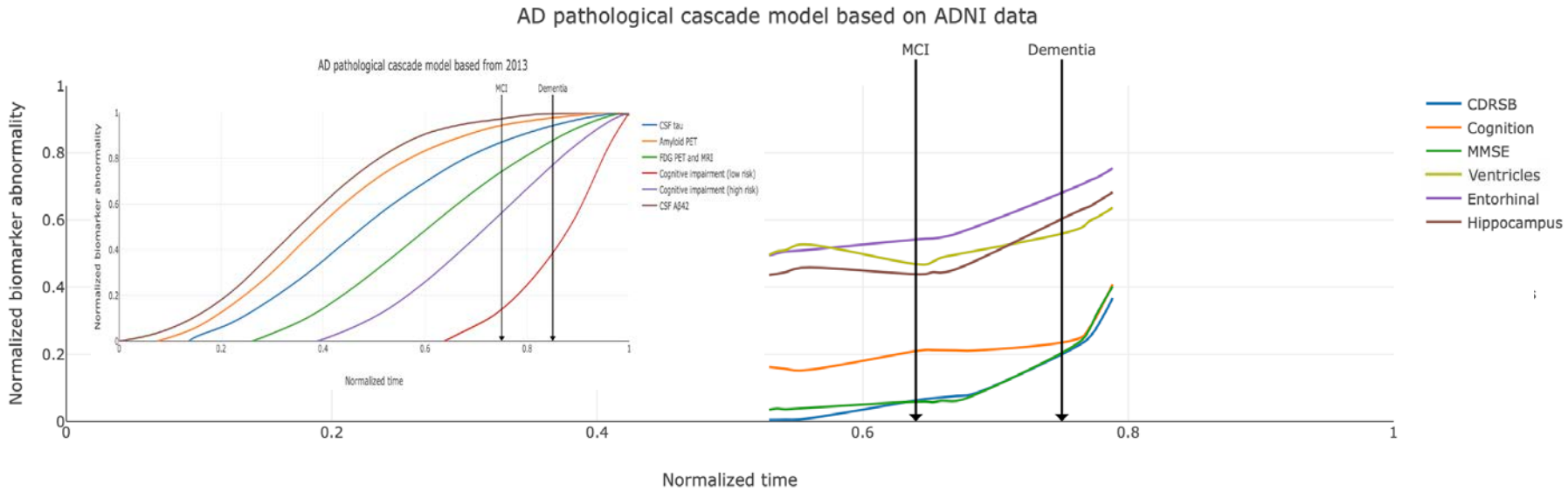
http://epad.scai.fraunhofer.de/longitudinal-adni

Computation of Trajectories



http://epad.scai.fraunhofer.de/longitudinal-adni

Computation of Trajectories



http://epad.scai.fraunhofer.de/longitudinal-adni

Short Summary of Part II

- AETIONOMY has
  - made use of the link to **EPAD**, the Alzheimer prevention trial. We have started modelling disease progression in the EPAD context and try to come up with mechanistic interpretation of biomarker trajectories in the course of AETIONOMY
  - to get started, we made the famous hypothetical model published by Jack et al (2013) “computable”. We have developed unified metrics that allow to compare biomarker trajectories sketched in the hypothetical model with biomarker trajectories extracted from ADNI
  - The reality check between the famous Jack et al., - model and our analysis of ADNI biomarker trajectories is sobering
  - Next step: we will plot biomarker trajectories of AddNeuroMed into the same coordinate system

Making Clinical Data Interoperable

The Challenge:

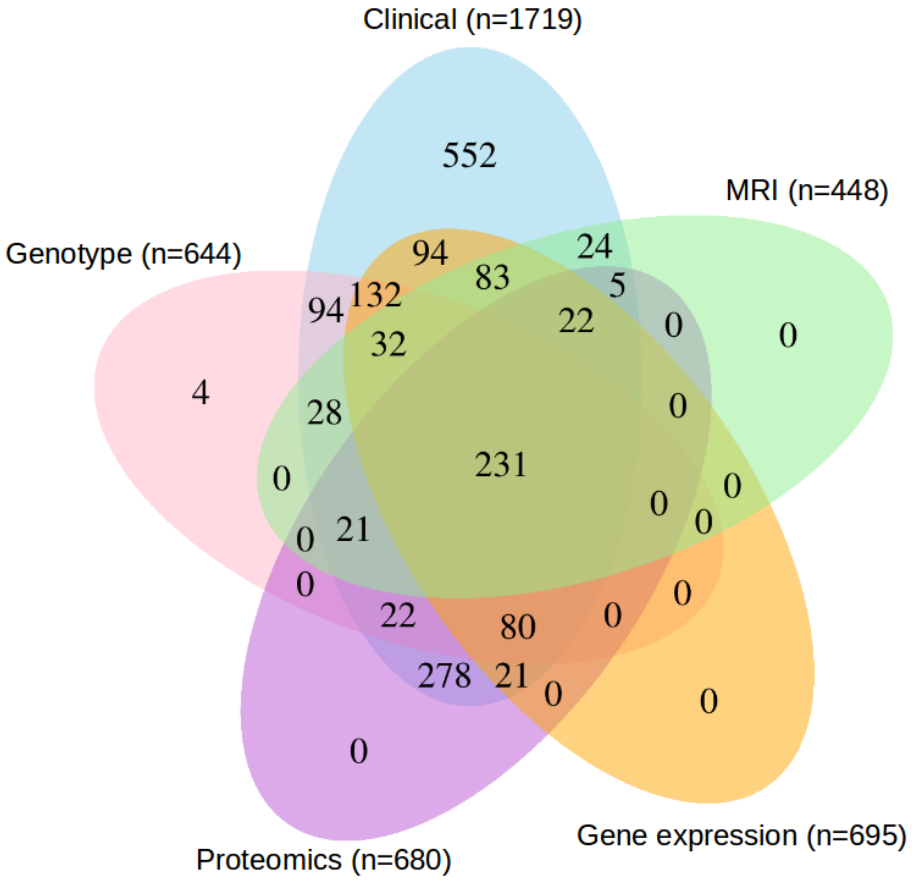
- Well-powered **data sets for validation**. If we can associate a multimodal pathophysiology mechanism with a subgroup of patients in a clinical study, we need to test the association in an independent clinical study.

The Problem-Solving Approach:

- Generation of AddNeuroMed – MERGE (a pre-processed, curated version of AddNeuroMed)
- Systematic comparative modeling of ADNI, AddNeuroMed, AIBL (and EMIF-1000, EPAD and ROSMAP)

Birkenbihl, Colin, *et al.,* manuscript in preparation
Balabin, Helena, *et al.,* manuscript in preparation (and already awarded with a prize)

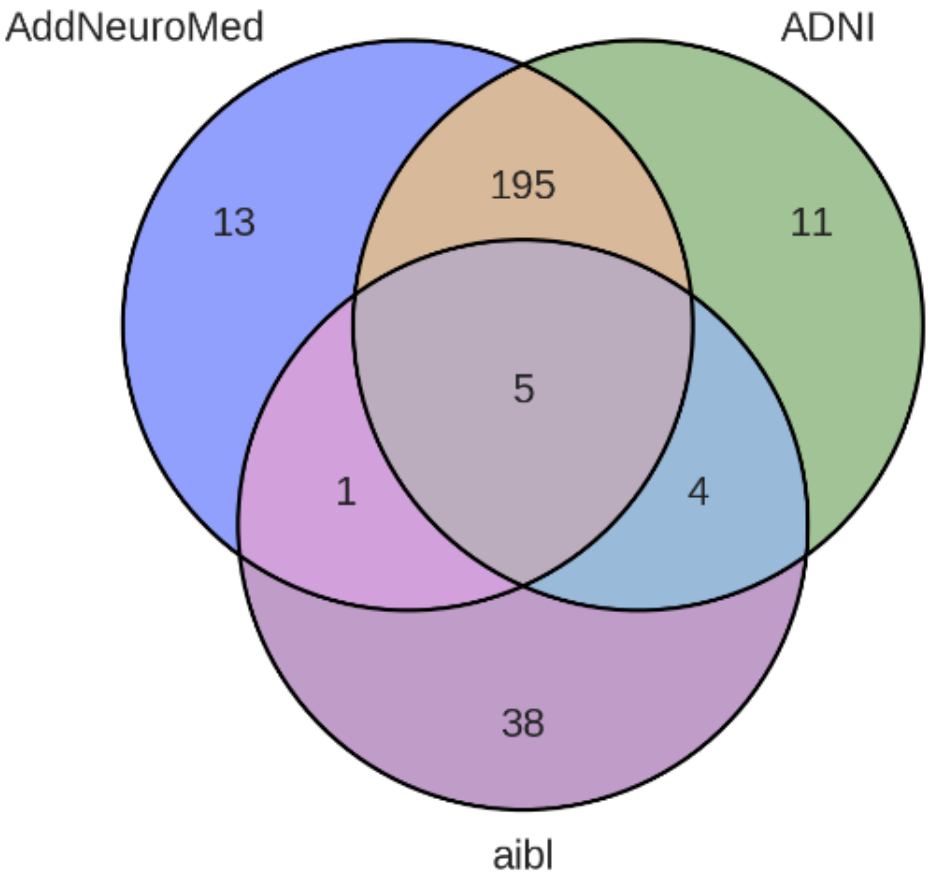
Patient Modality Overlap



- All modalities were made interoperable with each other

Feature Overlap

- Only features included into ADNIMERGE AddNeuroMed Merge and (to be) publicly released aibl data considered
- MRI, PET, Neurocognitive tests and CSF considered
- Redundant features excluded
- If all features present in the raw ADNI data would be included → higher overlap; but including is not trivial due to data documentation



New Algorithms ...

The Challenge:

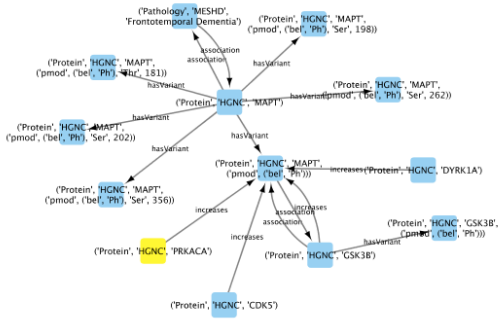
- **Ways and methods to associate** pathophysiology mechanisms with the variables in clinical studies. This may turn out to be non-trivial

The Problem-Solving Approach:

- Develop machine learning methods that allow us to establish links between candidate mechanisms and patient-level data
- Representation of patient-level data as probabilistic graph models (conditional dependency graphs; Bayesian networks) has been proven to work*

Khanna, Shashank, *et al.* "Using Multi-Scale Genetic, Neuroimaging and Clinical Data for Predicting Alzheimer's Disease and Reconstruction of Relevant Biological Mechanisms." Scientific reports 8.1 (2018): 11173.

An Approach for Mechanisms Based Patient Stratification



Shared AD / PD mechanisms (OpenBEL)

Mapping of SNPs + autoencoding

Unsupervised bi-clustering (sparse NMF)

Validation with external studies

- ROSMAP AD
- IDIBAPS AD
- AETIONOMY PD
- ICEBERG PD
- DIGPD

Clinical & biological understanding

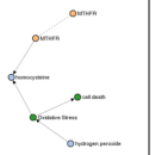
Biological contextualization

- Gene expression (ROSMAP)
- DNA methylation
- Proteomics
- Inflammation markers
- Histone marks (ROSMAP)

ongoing

Unsupervised joint clustering of Alzheimer's + Parkinson's patients

Most Discriminating Mechanisms in Detail

Subgraph Number	Genes in subgraph	Subgraph	Neighbourhood of the subgraph
Subgraph 5	MTHFR		
Subgraph 10	IL18, NLRP3		
Subgraph 12	AKT1		
Subgraph 13	MAPK9		

Cluster 3

- Follate metabolism (AD)
- Vitamin metabolism (AD)
- Epigenetic modification (PD)

Cluster 2

- IL signaling (AD, PD)
- Caspase signaling (AD, PD)

Cluster 1

- AKT/mTOR signaling (AD, PD)
- GBR10 signaling (PD)
- Nerve growth factor (AD)
- Matrix metalloproteinase (AD)

Cluster 4

- MAPK signaling (AD, PD)

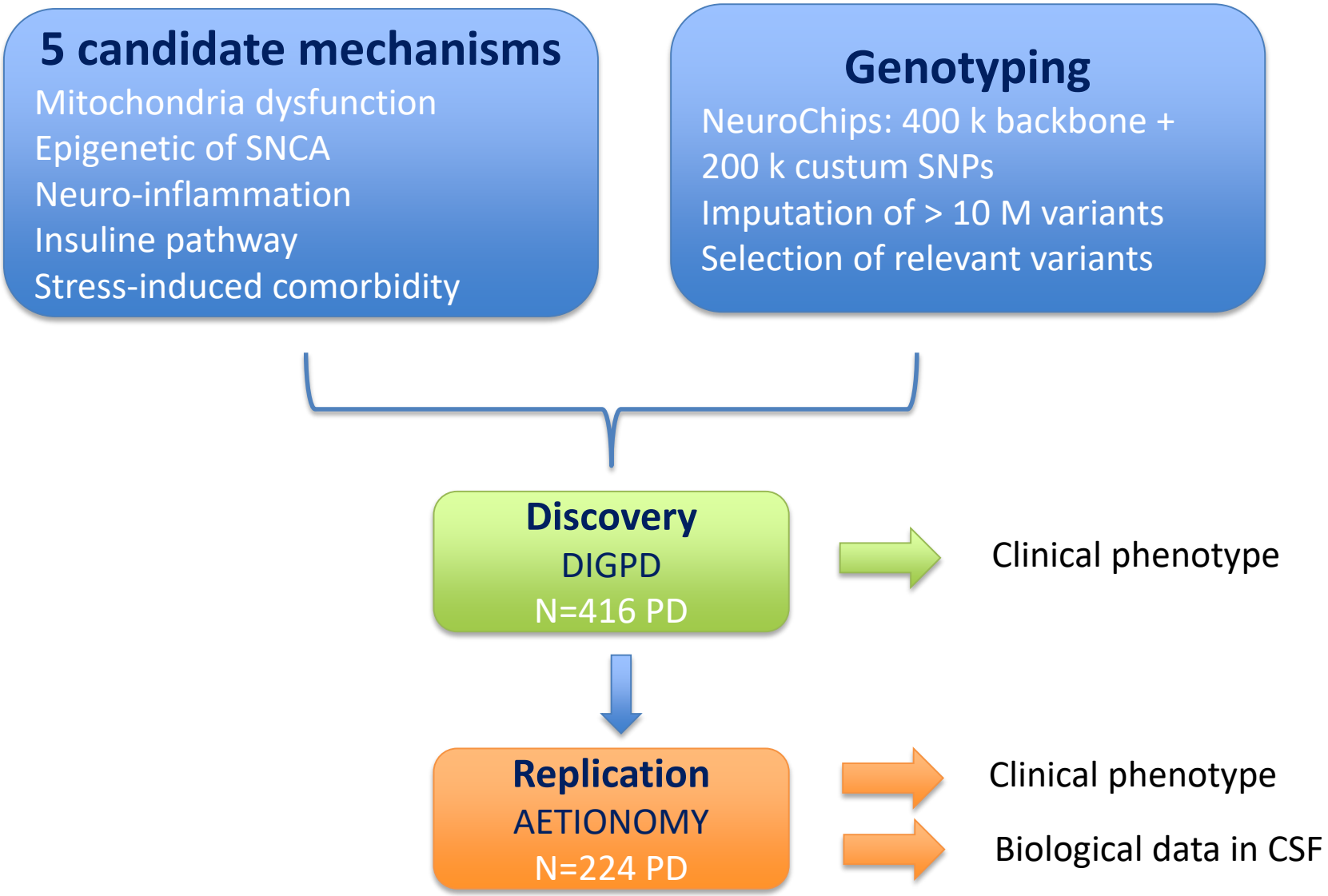
NeuroMMSig mechanisms

Joint sparse NMF based (bi-)clustering of ADNI + PPMI genotypes

Organising data and knowledge in the indication area and apply modelling and mining to gain new insights about disease mechanisms.
That is easily written on powerpoint. It needed a lot of organisation, synchronization and management.

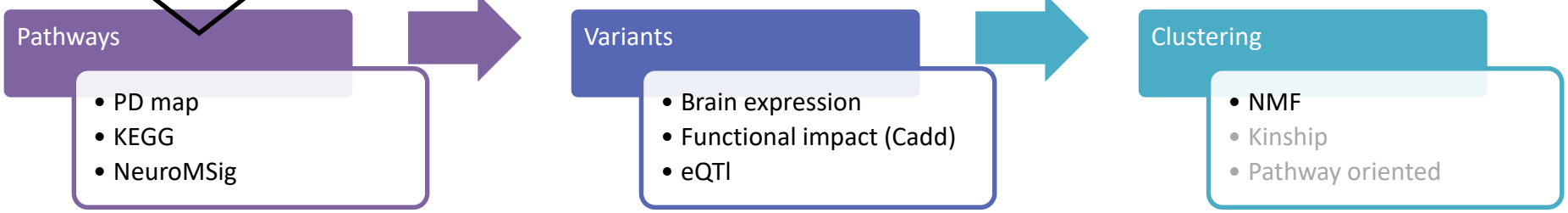
The Clinical Validation

Analysis pipeline



Genetic variant selection and clustering

5 candidate MECHANISMS



	N	Astroglial Inflammation	Insulin Signal Transduction	Mitochondrial Dysfunction	SNCA Methylation	Stress Induced Comorbidity
Total number of variants	956	354	221	285	237	76
Not shared		303	142	168	113	22
Common in 2 mechanisms		27	0	22	19	28
Common in 3 mechanisms		10	79	81	91	12
Common in 4 mechanisms		14	0	14	14	14

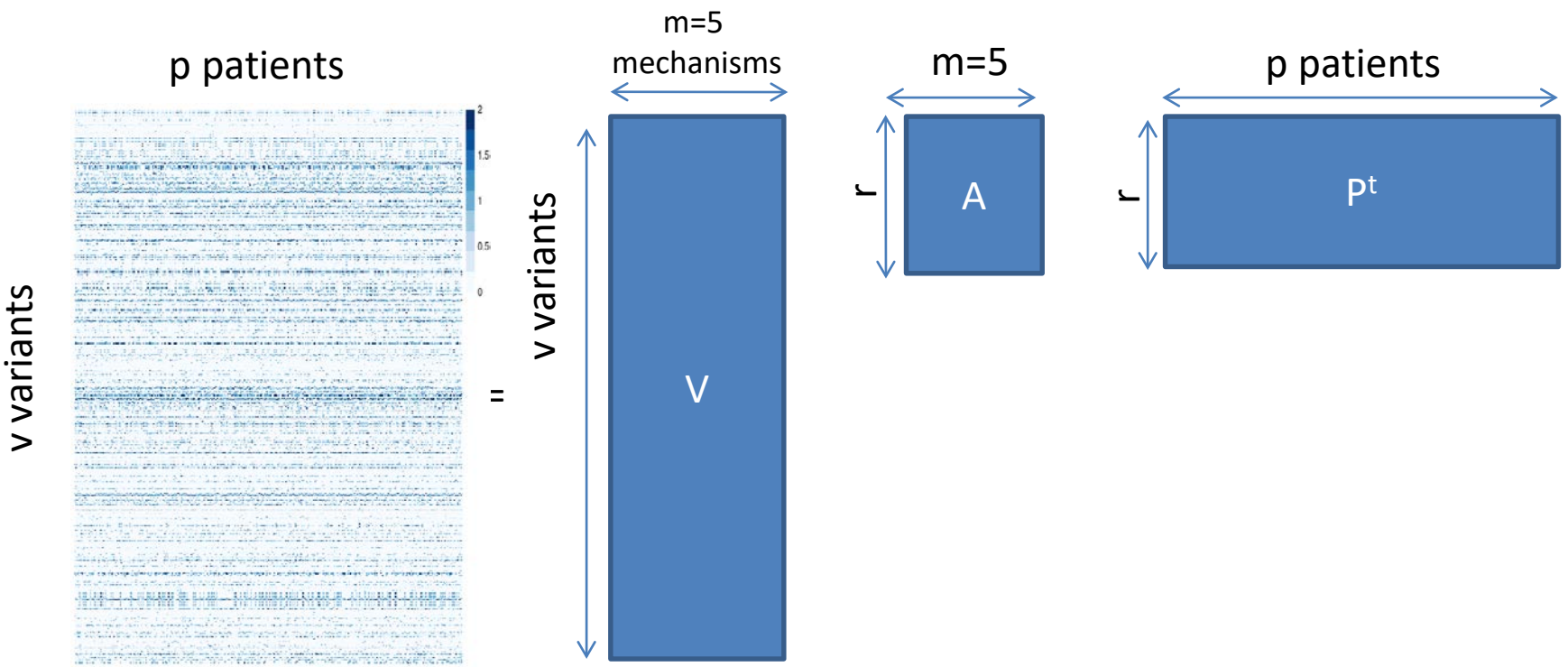
NMF: application to AETIONOMY

ICM, Paris
François-Xavier Lejeune
Fabrice Danjou
Boris Labrador

UCB
Holger Froehlich

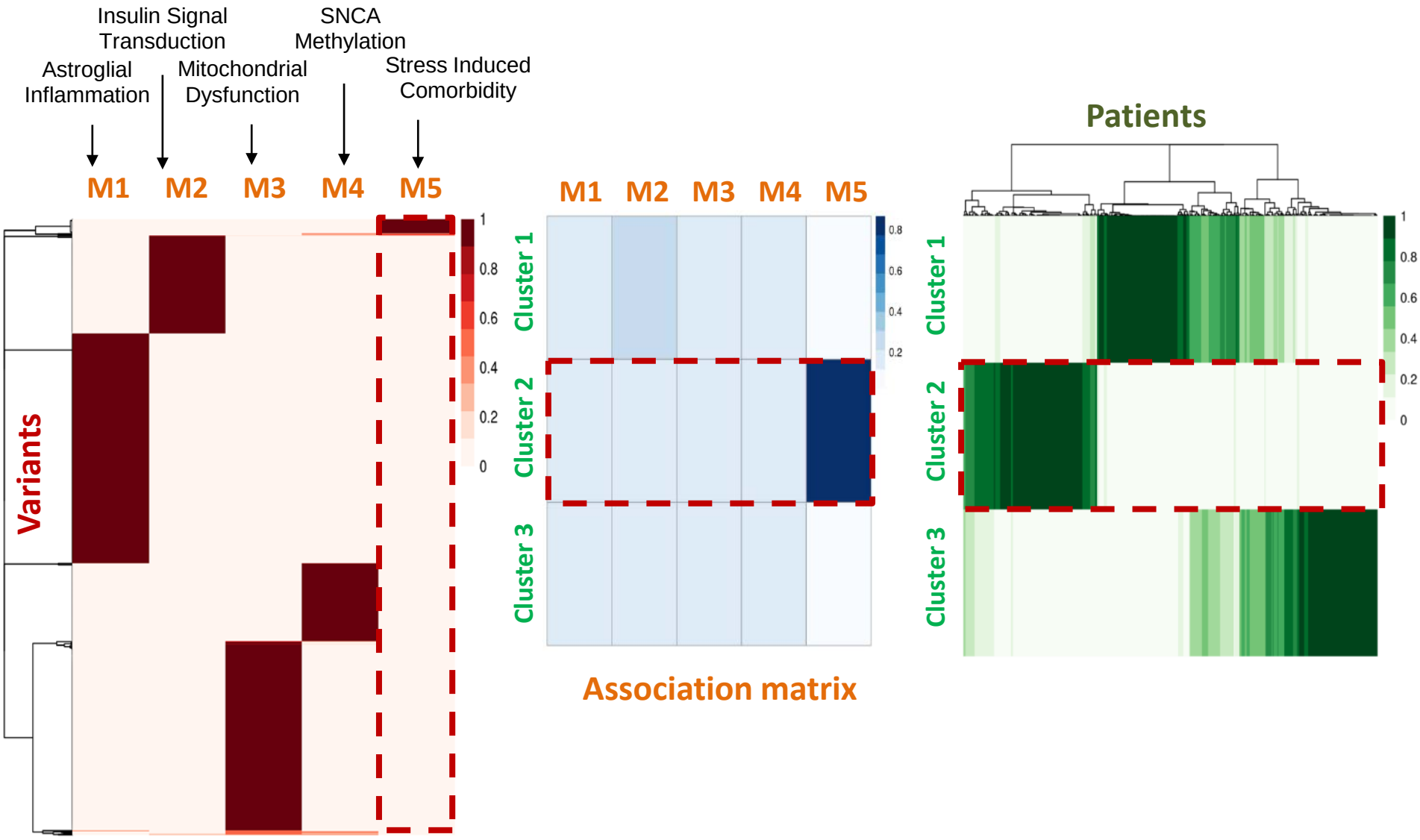
$$X \approx V \times A \times P^t$$

$$(v,p) = (v,m) \times (m,r) \times (r,p)$$

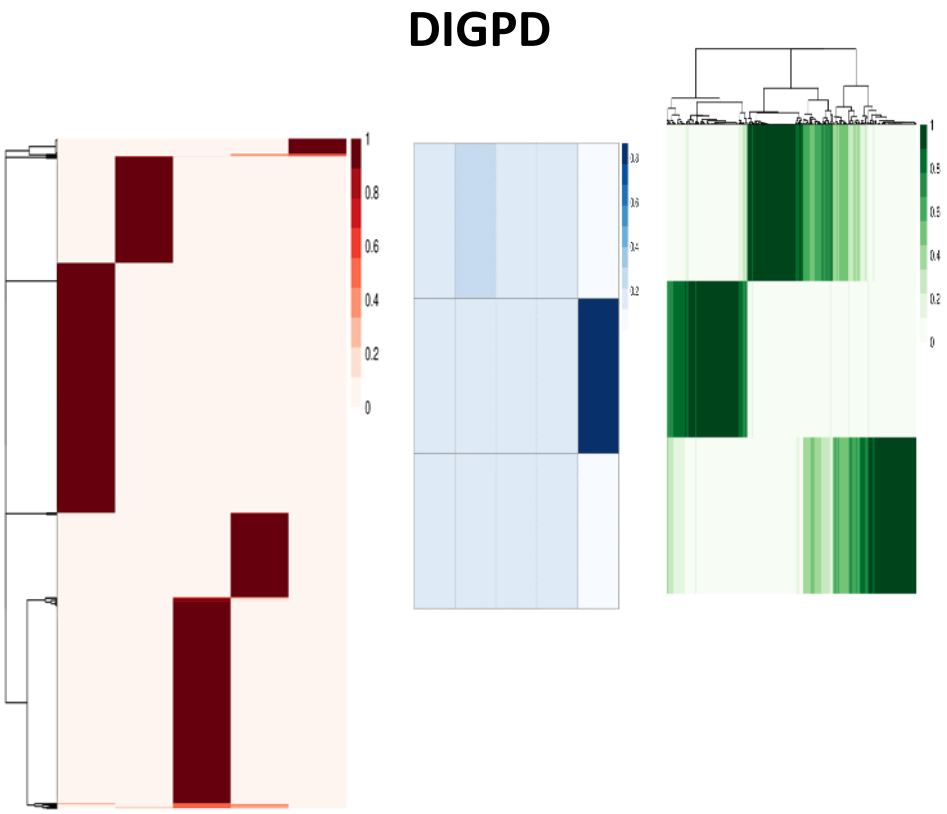


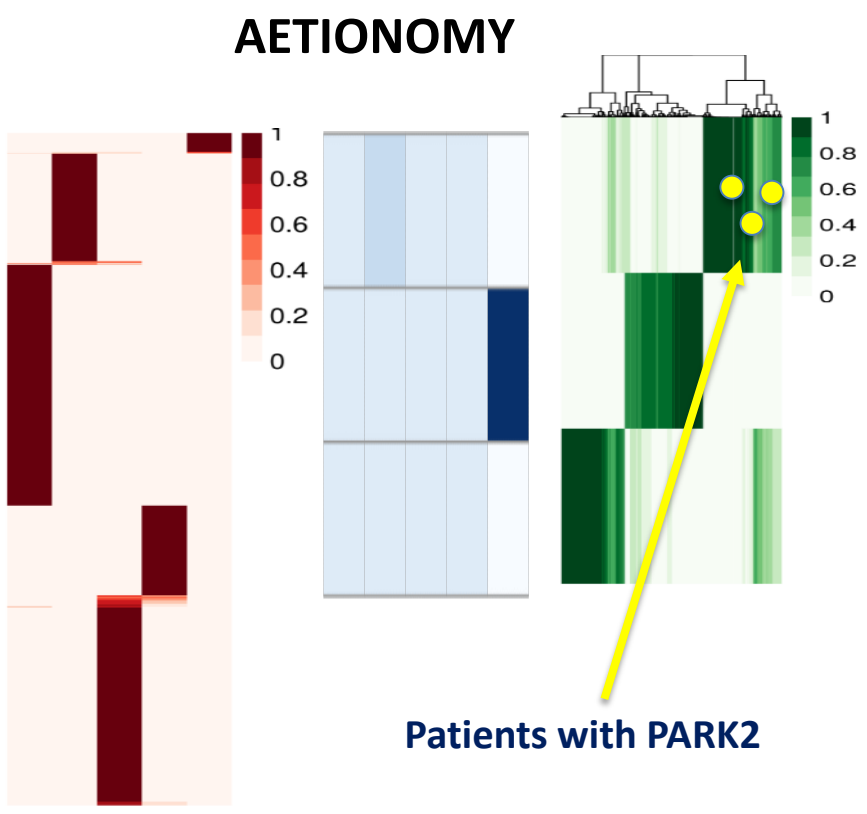
$$\min_{V,A,P} F(V, A, P) = \min_{V,A,P} ||X - V A P^t||^2 + Pen(V)$$

NMF in PD patients from the DIGPD cohort



Replication in the AETIONOMY cohort





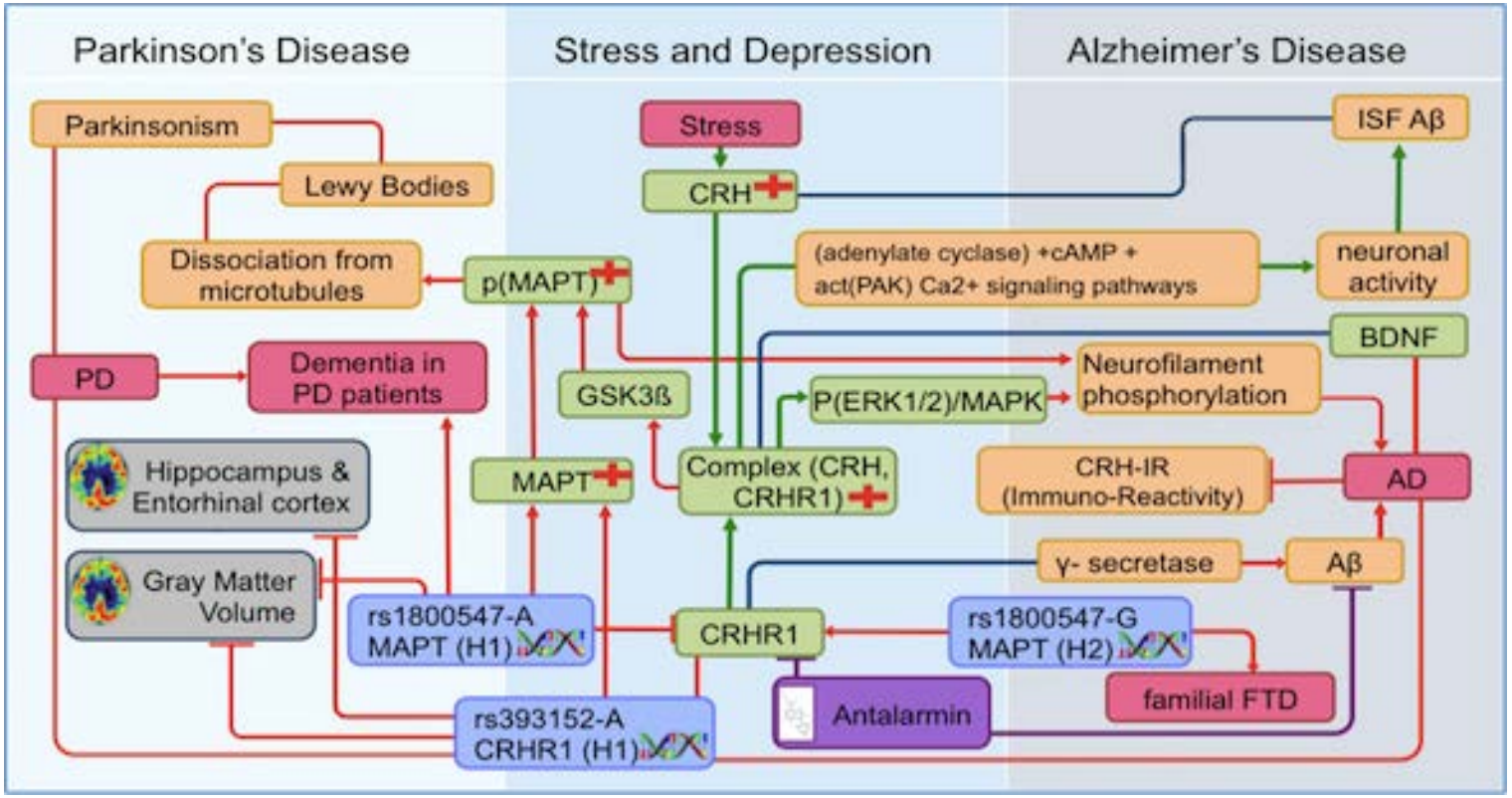
Similar variant map profile

Similar number of patients in clusters

Similar relationship with mechanisms

The Most Significant Mechanism in all Clinical Validation Exp.

Stress-depression/CRH release linking AD and PD



AETIONOMY – The Vision and the Reality

Take – home messages:

- AETIONOMY has generated the first version of a mechanism-based taxonomy for Alzheimerism and Parkinsonism
- With NeuroMMSig, the project has generated the largest inventory of computable disease mechanisms underlying neurodegeneration worldwide.
- With the Virtual Dementia Cohort concept, we break out of clinical data silos
- AETIONOMY has successfully developed strategies and new algorithms to associate mechanisms with biomarkers (and progression) in patient-level data.
- AETIONOMY has demonstrated in its clinical validation work package that stratification of patients according to disease mechanisms is possible.

AETIONOMY – Time to say THANK YOU!

The Coordinators would like to thank:

- The **funding body IMI**, the entire IMI team and in particular Elisabetta Vaudano for staying on our side all the time
- The **project office and project managers**; in particular Jacqueline Marovac, Stephan Springstubbe and Tobias Rechmann.
- All **Work Package leaders** for their tireless work and effort
- All **academic and all EFPIA partners** in the AETIONOMY project for their valuable contributions and the constructive collaboration
- All **partner projects** in IMI for fruitful collaboration
- **Simon Lovestone** and his team at the University of Oxford for sharing of data, sharing of thoughts and helping wherever they could
- **All patients** who consented to take part in the AETIONOMY cohort studies