Vision

Towards a mechanism-based taxonomy of neurodegenerative diseases

Martin Hofmann-Apitius

Academic Co-Coordinator of AETIONOMY

AETIONOMY FINAL SYMPOSIUM

November 29, 2018

Bonn, Germany

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

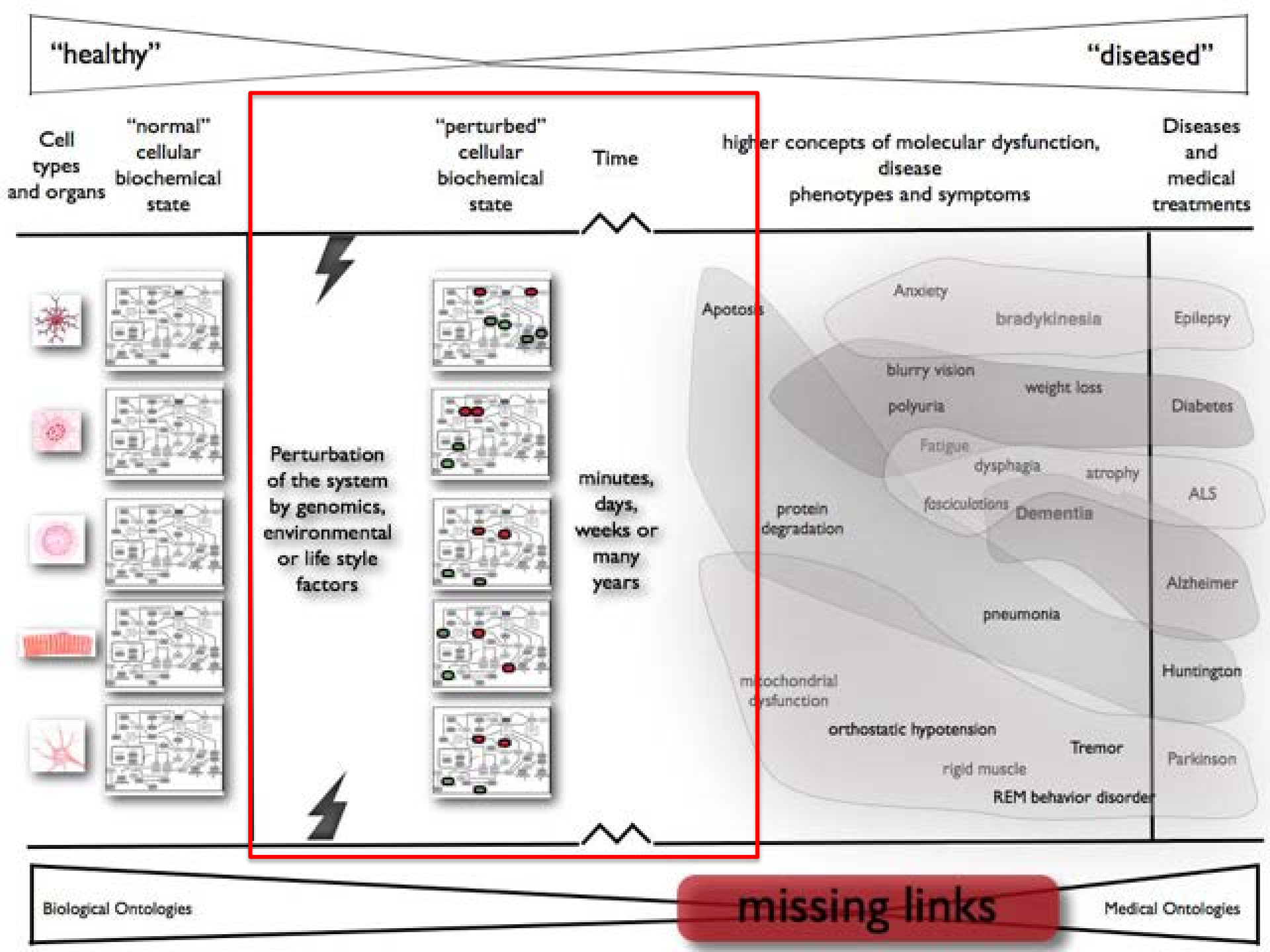
A coordinated effort to incorporate advances in the understanding of the molecular and genomic variations in common diseases, such as hypertension, into their diagnosis and treatment could transform drug development and medicine.

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.

A Non-Trivial Challenge: Development of a "mechanism-based taxonomy for neurodegenerative diseases"

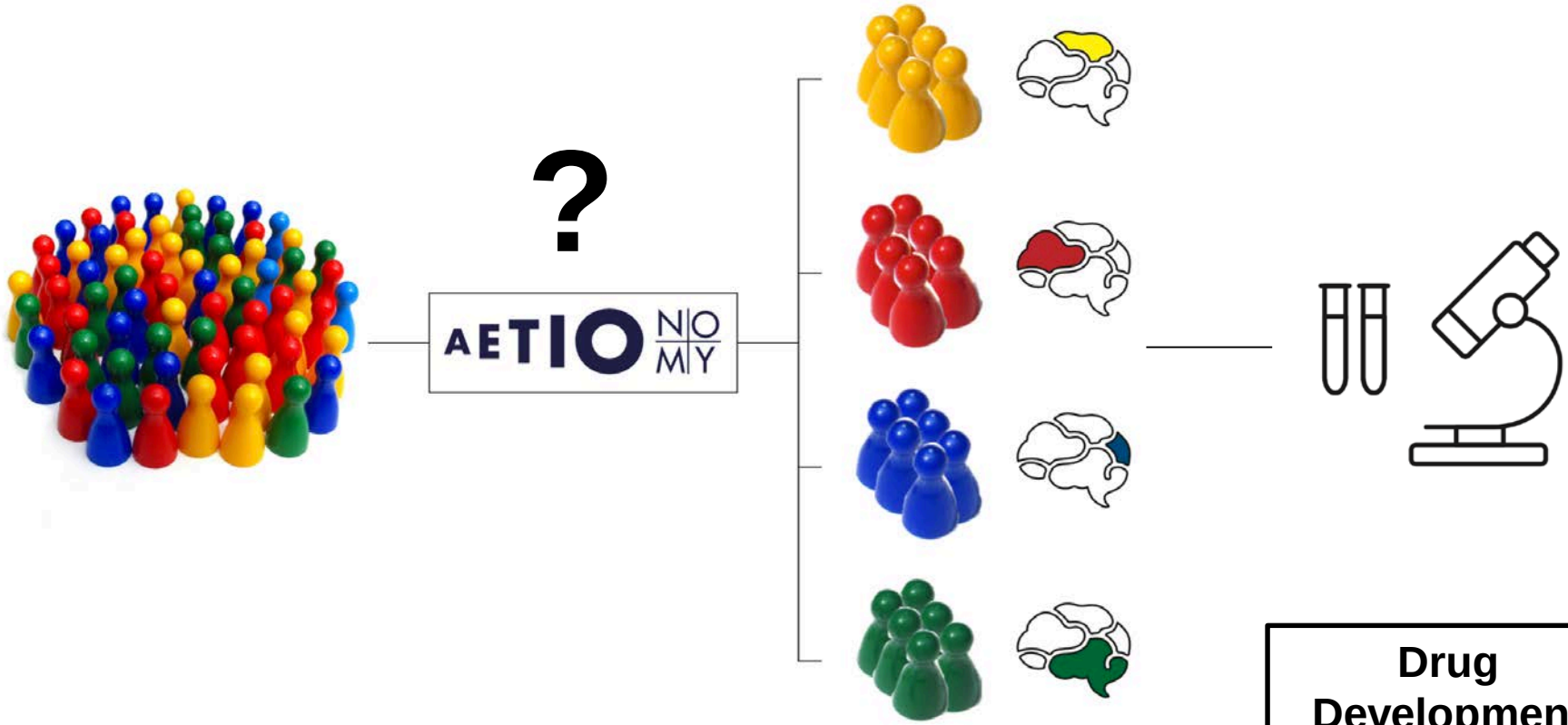

"healthy"
"diseased"
Cell types and organs
"normal" cellular biochemical state
"perturbed" cellular biochemical state
Time
higher concepts of molecular dysfunction, disease phenotypes and symptoms
Diseases and medical treatments
Perturbation of the system by genomics, environmental or life style factors
minutes, days, weeks or many years
Apotosis
Anxiety
bradykinesia
Epilepsy
blurry vision
weight loss
polyuria
Diabetes
Fatigue
dysphagia
atrophy
fasciculations
ALS
protein degradation
Dementia
Alzheimer
pneumonia
Huntington
mitochondrial dysfunction
orthostatic hypotension
Tremor
Parkinson
rigid muscle
REM behavior disorder
Biological Ontologies
missing links
Medical Ontologies

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

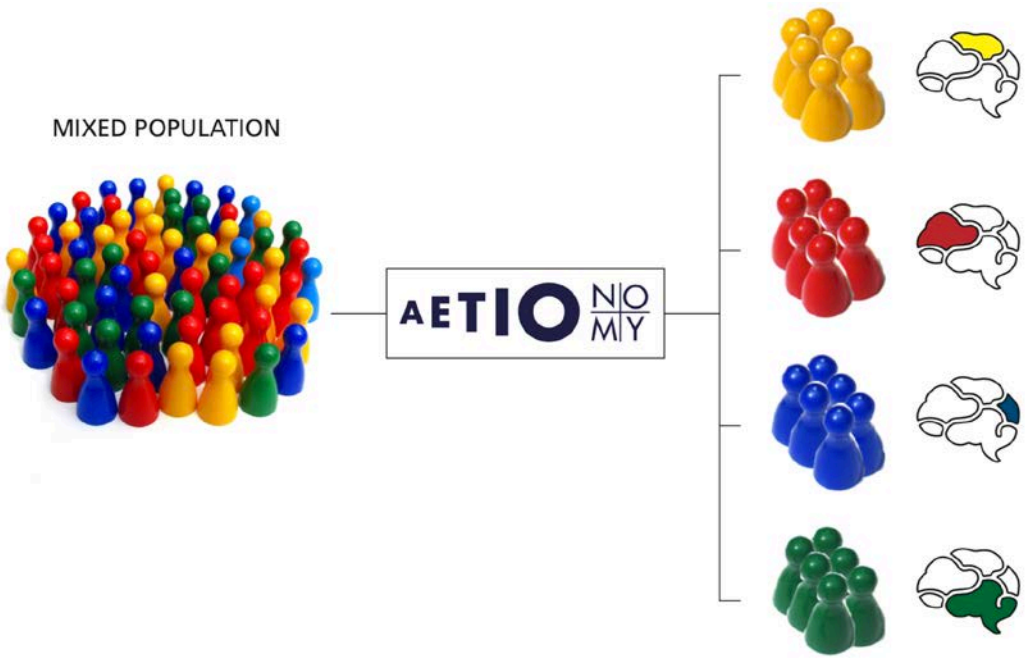
Drug Development and Therapy

ucb Fraunhofer SCAI efpia imi innovative medicines initiative AETIONOMY

The Reality:

Data and Knowledge about Pathophysiology Mechanisms are scattered, biased, heterogeneous and somtimes false.

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 Work:

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

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, so that we know, what “signature” of biomarkers is associated with disease progression (or disease risk).
- **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.

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*.

Domingo-Fernández, D., *et al.* (2017). Multimodal Mechanistic Signatures for Neurodegenerative Diseases (NeuroMMSig): a web server for mechanism enrichment. *Bioinformatics*, *33*(22), 3679-3681.

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 and pre-processing 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)

… Patient Involvement, Ethics and Legal

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:

Patient Involvement, Ethics and Legal

- Systematic harvesting, curation and pre-processing 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)

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.

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)

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.

Annotation, Curation, Quality Control, Interoperability of Data and Models;
Mining and new insights:
the Work Packages

WP2: The AETIONOMY Knowledge Base

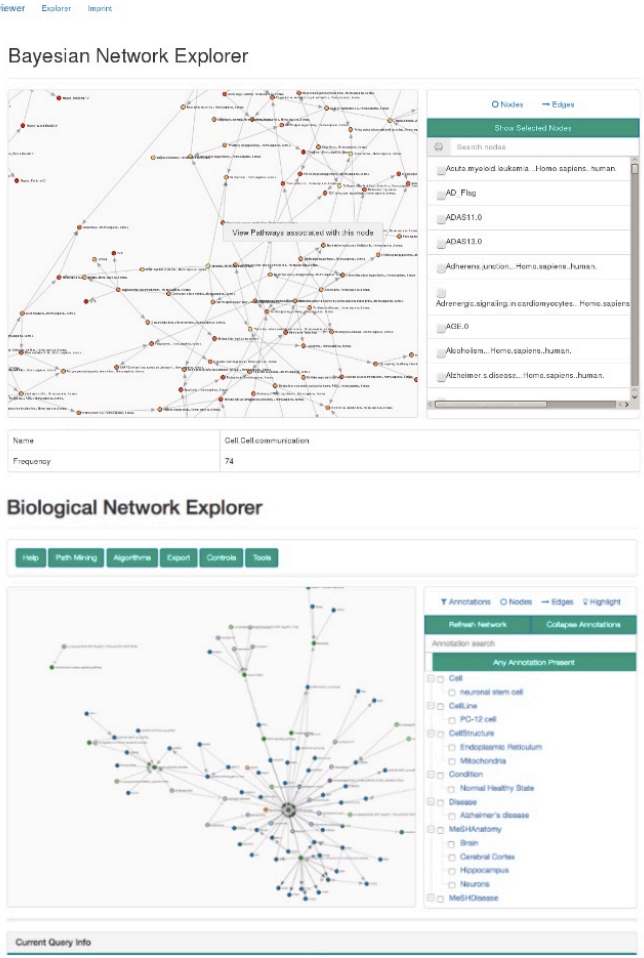
A Knowledge Base (KB) comprises curated data, models and methods to analyze data and models. The AETIONOMY work package that delivered the knowledge base had to

- Organise / curate multi-omics data and clinical data so that they are **FAIR**.
- The AETIONOMY KB will be maintained **for the next five years after the end of the funded period** of the project.
- Through the link to ELIXIR, we make the AETIONOMY KB a **sustainable resource** for translational neurodegeneration research.
- To overcome legal restrictions linked to patient-level data analysis, we have invented the concept of **“Virtual Dementia Cohorts”** (VDC). See also FRONTIERS in BIG DATA

WP3: The Models and the Mining

All disease modeling work and the data analytics / data science tasks in AETIONOMY are organized in WP3. This work package had to deliver the initial version of the mechanism-based taxonomy. In WP3, we organized the work that lead to:

- **NeuroMMSig,** the **mechanism-enrichment server** with its ”inventory” AD, PD (and EP) mechanisms
- **New algorithms** that allow to establish associations between mechanism-graphs and patient-level data
- **Progression models of disease**, based on biomarker trajectory mode (inspired by our link to EPAD) and longitudinal Bayesian models of disease progression
- **Data science approaches** that allow us to test for the “distance” between a real-world cohort and its derivative, the Virtual Dementia Cohort. This ensures, that our Virtual Patients are close to reality.



WP4: The Patients the Law and our Ethics

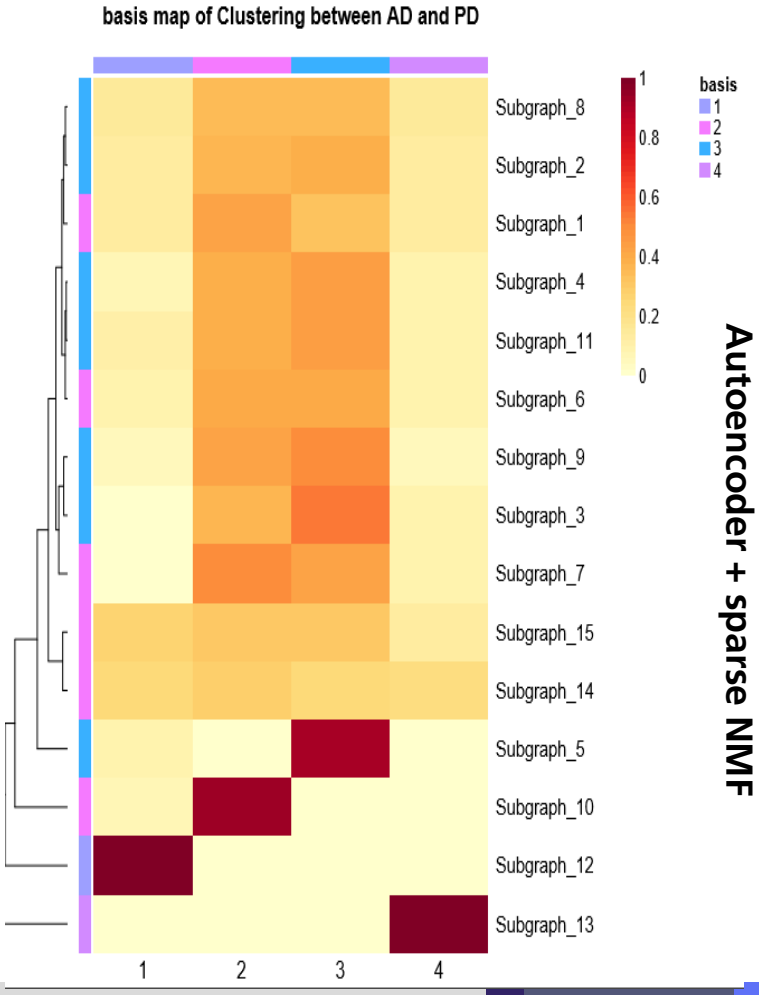
Discussing patient interests is particularly challenging when talking about major neurodegenerative diseases. In WP4, our colleagues from Alzheimer Europe and from the University of Hannover guided us

- In all tasks of the clinical work package (WP5)
- They have given us confidence in the Virtual Dementia Cohort concept through a clear assessment of the legal state of Virtual Patients
- They have helped us communicating to patient organizations and lay audiences
- They were always prepared to provide input on the ethics side of our work

WP5: The Clinical Research Work

All clinical and experimental work in AETIONOMY is organized in WP5. This work package provided patient-level data; suggested candidate mechanisms for the taxonomy and performed all initial validation work. Intensive collaboration between data scientists in WP3 and clinical researchers in WP5 has lead to:

- **Identification of 7 candidate mechanisms** that have – in part – been tested for their potential to identify patient subgroups
- Implementation of a **validation approach** for candidate mechanisms in the AETIONOMY PD cohort
- Implementation of a validation approach for candidate mechanisms based on **massive parallel proteomics profiling**



AETIONOMY – The Vision and the Reality

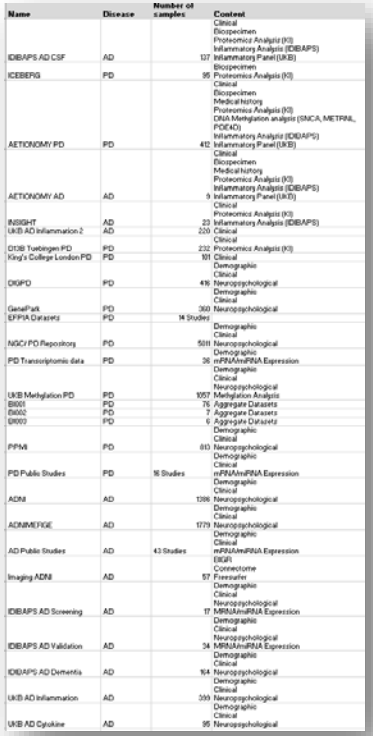
Take – home messages:

- AETIONOMY has generated the first version of a mechanism-based taxonomy for Alzheimerism and Parkinsonism
- AETIONOMY has generated a resource, the AETIONOMY Knowledge Base that contains high-quality, curated data, and computable models of disease
- 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.
- Validation of the first attempts at mechanism-based stratification of patients is under way, but will keep us busy beyond the end of the funded period of the project.

Looking ahead:

Legacy: What remains?

- Experienced partnerships
- New projects building directly on AETIONOMY such as **VirtualBrainCloud** (with Fraunhofer, Oxford and ICM) and **MENTA-COM** (with Luxembourg, ICM, Oxford and Fraunhofer)
- Crosstalk to other projects, such as PHAGO, HBP, EPAD, RADAR-AD, FAIRplus
- Future developments, like the Virtual Cohort approach have heavily resonated with the community; a special issue of FRONTIERS in Big Data will deal with that topic
- AETIONOMY Knowledge base maintained via ELIXIR-LU



Name	Disease	Number of samples	Content
			Clinical
			Biospecimen
			Proteomics Analysis (KI)
			Inflammatory Analysis (IDBAPS)
IDBAPS AD CSF	AD	137	Inflammatory Panel (UKB)
			Biospecimen
ICEBERG	PD	95	Proteomics Analysis (KI)
			Clinical
			Biospecimen
			Medical history
			Proteomics Analysis (KI)
			DNA Methylation analysis (SNCA, METRNL, PDE4D)
			Inflammatory Analysis (IDBAPS)
AETIONOMY PD	PD	412	Inflammatory Panel (UKB)
			Clinical
			Biospecimen
			Medical history
			Proteomics Analysis (KI)
			Inflammatory Analysis (IDBAPS)
AETIONOMY AD	AD	9	Inflammatory Panel (UKB)
			Clinical
			Proteomics Analysis (KI)
INSIGHT	AD	23	Inflammatory Analysis (IDBAPS)
UKB AD Inflammation 2	AD	220	Clinical
			Clinical
DIGI Tuebingen PD	PD	232	Proteomics Analysis (KI)
King's College London PD	PD	101	Clinical
			Demographic
			Clinical
DIGPD	PD	416	Neuropsychological
			Demographic
			Clinical
GenePark	PD	360	Neuropsychological
EPPIA Datasets	PD	14 Studies	
			Demographic
			Clinical
NGCI PD Repository	PD	5011	Neuropsychological
			Demographic
PD Transcriptomic data	PD	36	mRNA/miRNA Expression
			Demographic
			Clinical
			Neuropsychological
UKB Methylation PD	PD	1057	Methylation Analysis
EMIF	PD	76	Aggregate Datasets
EMIF	PD	7	Aggregate Datasets
EMIF	PD	6	Aggregate Datasets
			Demographic
			Clinical
PPMI	PD	813	Neuropsychological
			Demographic
			Clinical
PD Public Studies	PD	16 Studies	mRNA/miRNA Expression
			Demographic
			Clinical
ADNI	AD	1286	Neuropsychological
			Demographic
			Clinical
ADNIMERGE	AD	1779	Neuropsychological
			Demographic
			Clinical
AD Public Studies	AD	43 Studies	mRNA/miRNA Expression
			EEGFL
			Connectome
Imaging ADNI	AD	57	Freesurfer
			Demographic
			Clinical
			Neuropsychological
IDBAPS AD Screening	AD	17	MRNA/miRNA Expression
			Demographic
			Clinical
			Neuropsychological
IDBAPS AD Validation	AD	34	MRNA/miRNA Expression
			Demographic
			Clinical
IDIBAPS AD Dementia	AD	164	Neuropsychological
			Demographic
			Clinical
UKB AD Inflammation	AD	399	Neuropsychological
			Demographic
			Clinical
UKB AD Cytokine	AD	95	Neuropsychological

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