July 2015

Newsletter Issue 2

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

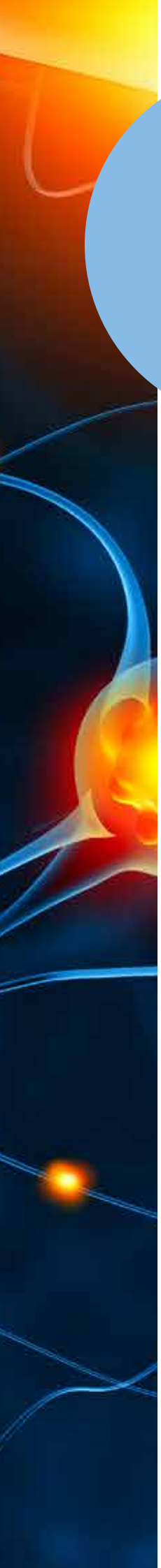
IN THIS ISSUE

Project Coordinators:

UCB Biopharma
Professor Duncan McHale

Fraunhofer SCAI
Professor Martin Hofmann-Apitius

www.aetionomy.org



EDITORIAL

Towards a mechanism-based taxonomy of Alzheimer´s and Parkinson´s disease

AETIONOMY is a consortium brought together under the European Innovative Medicines Initiative (IMI) to tackle the problem of the classification of neuro-degenerative diseases. Revising the taxonomy (classification) will take many years to complete and to be accepted by the biomedical community. AETIONOMY is the first step on this journey.

AETIONOMY is novel in terms of both its scientific approach and its scale. There is a large body of published literature on the potential causes of Alzheimer's and Parkinson's disease and a significant number of major collaborations already funded and working well. The majority of these are looking at individual hypotheses or approaches to the problem e.g. genetic association studies, imaging studies, non-motor Parkinson's disease or familial Alzheimer's disease. Rather than start another similar approach, AETIONOMY will identify all of the available data either from published literature, publically available datasets or datasets from our collaborators. A common framework will be developed which, once this data is curated, will allow the integration of this data within the AETIONOMY Knowledge Base (KB). Once this has been achieved, novel data mining and visualisation approaches will be used to identify the pathophysiological changes occurring in the disease process at a molecular level. The knowledge extracted from the datasets will be used to cluster individual patients into separate mechanism based sub-groups leading to a new taxonomy of Alzheimer's and Parkinson's disease. The final part of AETIONOMY is an observational clinical study, which will be used to collect biomarkers and prospectively validate 1 or more of the proposed new sub-groups.

Currently, the established disease classification systems such as ICD (International Classification of Disease) makes use of phenotypes measured clinically or using standard laboratory and imaging techniques to establish major types and subtypes of disease. In contrast to the established disease classification systems, the "mechanism-based taxonomy" will be based on the knowledge of the biological pathways involved in the aetiology of the disease leading to a classification of disease classes and subclasses.

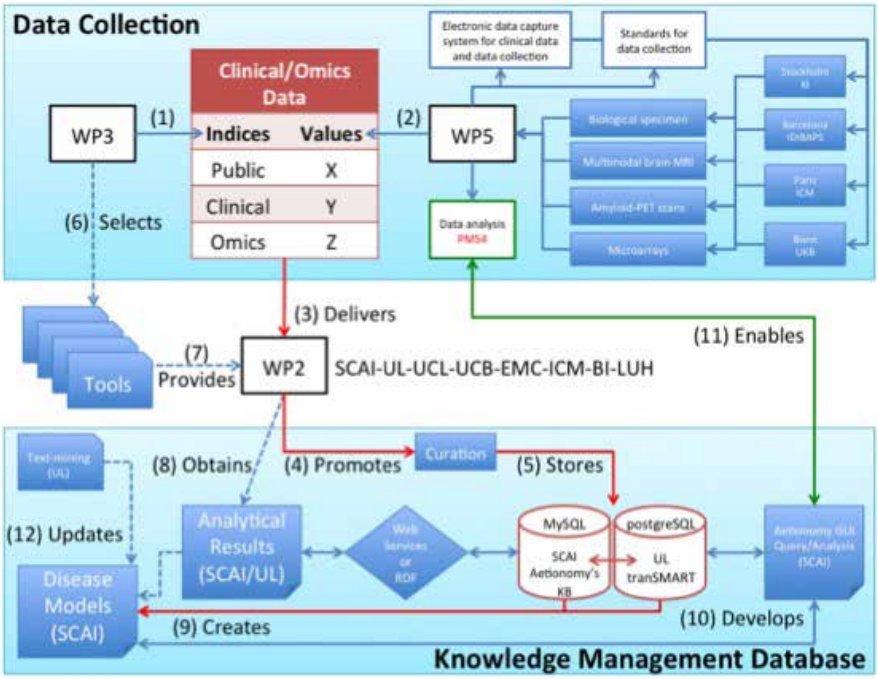
The Consortium is jointly led by Professor Martin Hofmann-Apitius from the Fraunhofer Institute for Algorithms and Scientific Computing (SCAI) in Germany and Professor Duncan McHale from the biopharmaceutical company UCB Biopharma SPRL in Belgium.

AETIONOMY

PROJECT DATA FLOW

The AETIONOMY Knowledge Base

AETIONOMY's Knowledge Base is the central information hub of the whole AETIONOMY Project data flow. The schema shown below depicts the main data flow from data sources to the final destination in the AETIONOMY Knowledge Base. The following workflow is envisioned: a) the data mining team in WP3 indicates specific public, clinical and OMICS indices relevant to AETIONOMY (1) b) the clinical partners in WP5 collect the different values for these indices at retrospective and prospective studies available at different sites (2) c) Data is being made available to WP2 (3) d) WP2 promotes curation of the data (4) to be stored adequately (5) e) WP3 selects specific analytical tools to analyze the available data (6) and provides these tool to WP2 in order to be linked to stored data (7). f) The obtained analytical results are integrated to KB using specific APIs (8). Both stored data and analytical results support the creation of disease models (9). g) Finally, a graphical user interface (GUI) is developed (10) by WP2 to enable other users to analyze the data (11). h) Text-mining systems are developed to update the disease models concepts (12). The components for data storage, analytical results management, data integration and analysis tools are named Knowledge Management Database (KMDB) in the scope of AETIONOMY.



The AETIONOMY modeling, mining and validation strategies

The AETIONOMY Knowledge Management Database (KMDB) is composed by two main inter-communicable layers: the Fraunhofer SCAI AETIONOMY Knowledge Base (KB) and the University of Luxembourg AETIONOMY tranSMART platform.

AETIONOMY KNOWLEDGE BASE

Collection of Information and Data for the AETIONOMY KB

Within the AETIONOMY KB, public datasets are represented by

- Clinical and Imaging data
- Genomic Wide Association Studies and comparative Toxicogenomics data (genomic level)
- Gene Expression data (transcriptome level)
- Protein Information, Domain and Interaction data (proteomic level)
- Pathways, Ontologies and Disease Models data (system level)

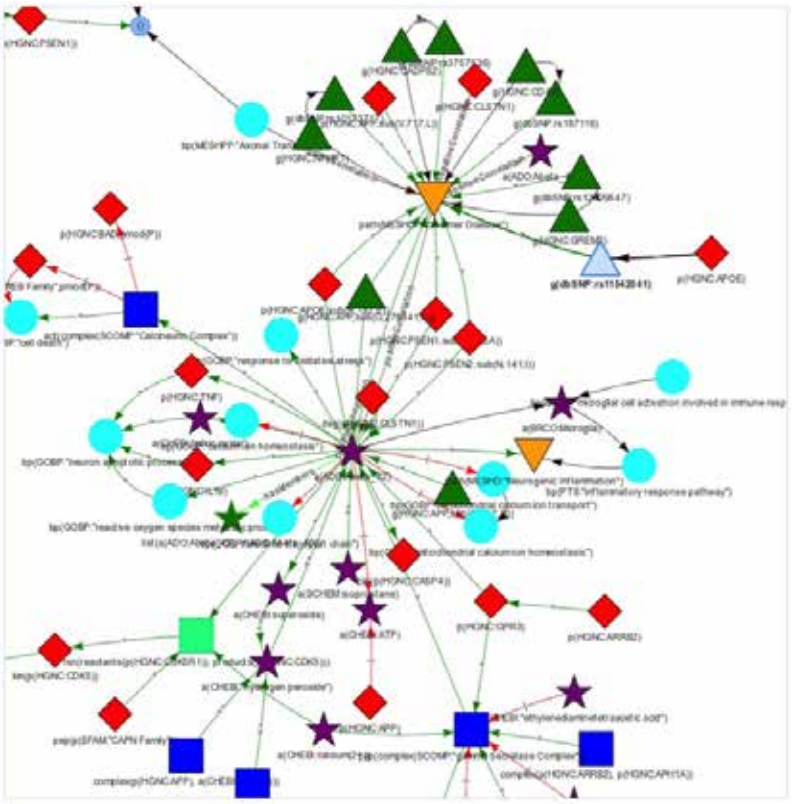
Amongst these, GWAS, Gene Expression, Protein Interactions and Disease Models have undergone an extensive curation and harmonisation process. An adaptation of the established eTRIKS data curation principles served as guidelines for that curation and re-annotation process. Curation and harmonisation of these datasets were supported by several ontologies and standard terminologies that have been created at Fraunhofer SCAI including disease-specific ontologies for Alzheimer's (ADO) and Parkinson's disease (PDON), as well as a brain and cell type terminology (BRCT), and clinical trial terminology (CTO). These datasets have been organized under the Knowledge domains tab in the above-mentioned interface. Access to these ontologies and terminologies is provided under the Ontologies tab.

The screenshot above shows the joint graphical user interface which combines services of the AETIONOMY Knowledge Base with its two different source installations, TranSMART (UL) and MySQL (Fraunhofer).

More information at aetionomy.scai.fraunhofer.de

Computer-processable disease models supporting hypothesis formation

In the real world, scientists often describe biological processes in the form of pathways and chains of interactions between molecules. Most of this information is hidden in the text portion of scientific publications and needs to be extracted and organized in models. Biological Expression Language (BEL) is a state-of-the-art modeling syntax to represent literature-derived cause-and-effect relationships among biological entities as networks of interactions. These models not only represent a comprehensive view on core pathophysiology pathways, but also cover a broad spectrum of events that are linked to clinical outcomes often seen in AD and PD patients. Thus, such disease-specific models in the AETIONOMY project lend support to clinical researchers to generate hypotheses for disease subtype stratification by providing an integrative and mechanistic framework for further data-driven analyses. For instance, gene expression values that have been derived from cohort analyses can be projected onto these context-specific models to validate mechanistic hypotheses. The AETIONOMY KB currently hosts an array of BEL models such as Alzheimer's disease, Parkinson's disease, Neuroinflammation, and SWEDD (Scans Without Evidence for Dopamine Deficiency) models. Currently, a proof-of-principle analysis has been undertaken to investigate mechanistic differences between PD and SWEDD conditions with the idea in mind to identify signatures that can guide biomarker discovery. These, and other models are currently being analysed for signatures that bear the potential to discriminate Parkinsonism subtypes as defined by the Parkinson's Disease Ontology (PDON).

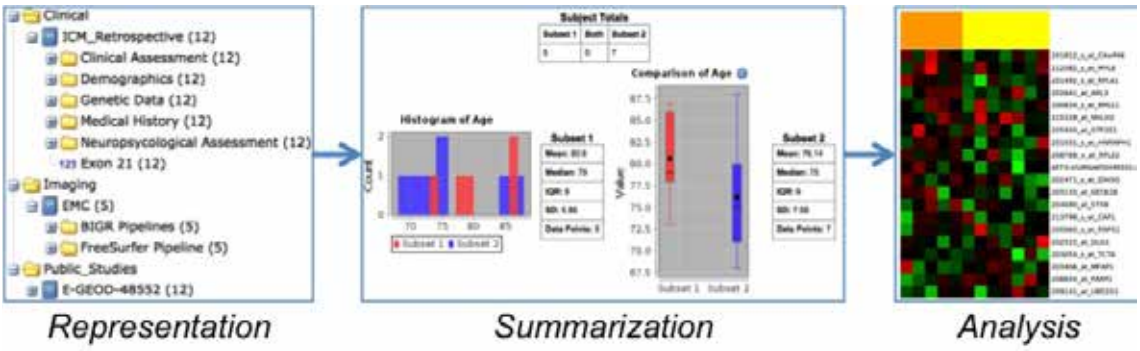


TRANSMART PLATFORM

AETIONOMY´s tranSMART Platform

AETIONOMY is a multi-centre project taking advantage of the different expertise and technology provided by various partners all over Europe. Likewise, the project must deal with the diversity of data provided by these partners. Meanwhile, the challenge is focused on the integration of different retrospective studies that can shed some light in the classification of neurodegenerative disease (NDD) subtypes, such as AD and PD. Different clinical, laboratorial, multi-omics and imaging data should, therefore, be integrated in the same platform, in order to permit a broader picture of the features observed within and across study's cohorts.

To cope with the enormous diversity of translational research projects, such as AETIONOMY, the European scientific community is gradually opting for tranSMART as the preferred data warehouse system for data integration. The WP2 partner **University of Luxembourg** (lead by Dr. Reinhard Schneider) has deployed a dedicated instance of tranSMART to attend AETIONOMY's demand on variable data storage (http://aetionomy.uni.lu/transmart).



AETIONOMY's tranSMART server hosted at the University of Luxembourg. Different capabilities of the system are depicted such as representation, summarization and analysis multiple data types.

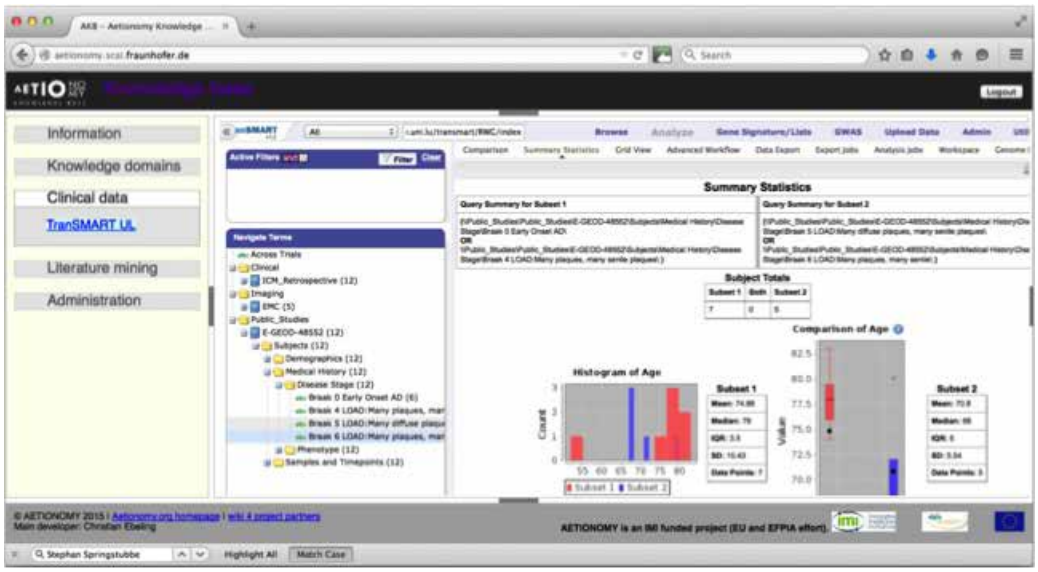
As showed on the figure above, initial data descriptors have already been represented on the server for Clinical and Imaging retrospective biomarkers from ICM and EMC partners, whereas a placeholder is prepared to receive selected public studies relevant to AETIONOMY.

tranSMART v1.2

tranSMART Architecture

tranSMART is rapidly developing into the de facto standard for data management in clinical omics research. The UL tranSMART web application is developed following an N-Tier architecture having JavaScript, GSP/JSP and Jason/XML as the main components of it's web-based user interface. The core components of UL tranSMART are modeled on a PostgreSQL server, this server inherits the incremental backup policy performed by the SIU (Informatics Service of the University) and using its infrastructure. tranSMART is deployed with all its web-components, AETIONOMY data will profit by using this resource, given that the data is loaded following native tranSMART ETL processes. Once loaded, access to specific datasets will be granted to partners following a well-defined data sharing policy negotiated between AETIONOMY partners, whereas for external studies, we will sign all needed agreements prior to release of third-party data on our platform. All this is facilitated by using the existing tranSMART users privilege management system.

Thank to its native RESTful API, all data deposited under the tranSMART database schema can be exported to external on-line applications, making it easy to integrate AETIONOMY KB data to external analytical pipelines.



UL tranSMART server represented as a major component of AETIONOMY's Knowledge Base.

Physically, the AETIONOMY tranSMART server currently operates at the HPC facilities of partner UL with a total of 369 computing nodes (4004 cores, 45.901 TFlops) and a shared storage capacity of 934.4 TB (+ 1064 TB for backup). WP2 uses a dedicated virtual machine within this infrastructure that fits the needs of AETIONOMY.

HPC Facilities at University of Luxembourg

PARTNER PROFILES

Joint Expertise for Research

The structure and size of the project consortium and the proposed project activities have been formulated to ensure a cost effective approach while realising the project objectives.

In this issue we introduce two Work Package 2 partners, who are involved in the generation of the AETIONOMY Knowledge Base: University of Luxembourg and Boehringer Ingelheim.

University of Luxembourg

The Luxembourg Centre for Systems Biomedicine (LCSB) is an interdisciplinary research centre at the University of Luxembourg. It is accelerating biomedical research by closing the link between systems biology and medical research. Collaboration between biologists, medical and computer scientists, physicists, engineers as well as mathematicians is offering new insights in complex systems like cells, organs and organisms. These findings are essential for understanding principal mechanisms of disease pathogenesis and for developing new tools in diagnostics and therapy. Neurodegenerative diseases like Parkinson's disease and description of diseases as networks are in the focus of research at LCSB. The Centre has established strategic partnerships with leading academic institutions and industry worldwide and has founded two spin-off companies to accelerate the translation of fundamental research results into (clinical) applications.

Bioinformatics Core @ LCSB

The Bioinformatics Core, led by **Dr. Reinhard Schneider**, is responsible for the efficient data flow between the experimental groups and the theoretical and medical oriented groups. The group focuses on the development of (clinical) data management and storage systems and set up cost- and time-efficient data analysis pipelines. To this end, the group also develops new algorithms in various fields like data mining and visualisation to help to understand and interpret the data. Together with the high-performance computing group of the University they are responsible to setup and run large computer and storage facilities with thousands of processors and an ever-increasing number of hard disks. Currently the LCSB hosts the biggest hardware installation in the Luxembourgian academic landscape.

More information at wwwen.uni.lu/lcsb

uni.lu
UNIVERSITÄT LUXEMBURG

Boehringer Ingelheim

The Boehringer Ingelheim group is one of the world's 20 leading pharmaceutical companies. Headquartered in Ingelheim, Germany, Boehringer Ingelheim operates globally with 146 affiliates and a total of more than 47,700 employees. The focus of the **family-owned company**, founded in 1885, is researching, developing, manufacturing and marketing new medications of high therapeutic value for human and veterinary medicine.

More than 8,000 highly qualified people of Boehringer Ingelheim's over 47,700 employees work at either one of our three major research & development sites (Biberach and Ingelheim, Germany, Ridgefield, USA, and Vienna, Austria) or two support centres (Kobe, Japan and Milan, Italy).

Research & development has been the foundation of Boehringer Ingelheim's success and continues to be the major driver of innovative, new medicines for the treatment of diseases with high unmet medical need. Furthermore, in the constant quest for pharmaceutical innovation Boehringer Ingelheim has successful ongoing collaborations and is actively seeking new collaborations with external partners, ranging from academic institutions to biopharmaceutical enterprises and start-up companies. Currently, more than 100 projects in the companies' research organization are ongoing.
In 2014 the R&D spending reached 2.7 billion Euro, which represents about 20 % of the net sales. Boehringer Ingelheim has a long-standing reputation in supporting basic science in academia. During the past decade, from 2005 to 2014, there were 1,414 studies conducted or sponsored with 115 compounds in 95 countries from all regions of the world, striving to expand the company's international network of academic, industry and not-for-profit partners who share the goal to meet the medical needs of the patients. "Currently, more than 50% of our early-mid stage pipeline is filled with products that are derived from external innovation with our partners. We think together we can improve the health of people around the world."

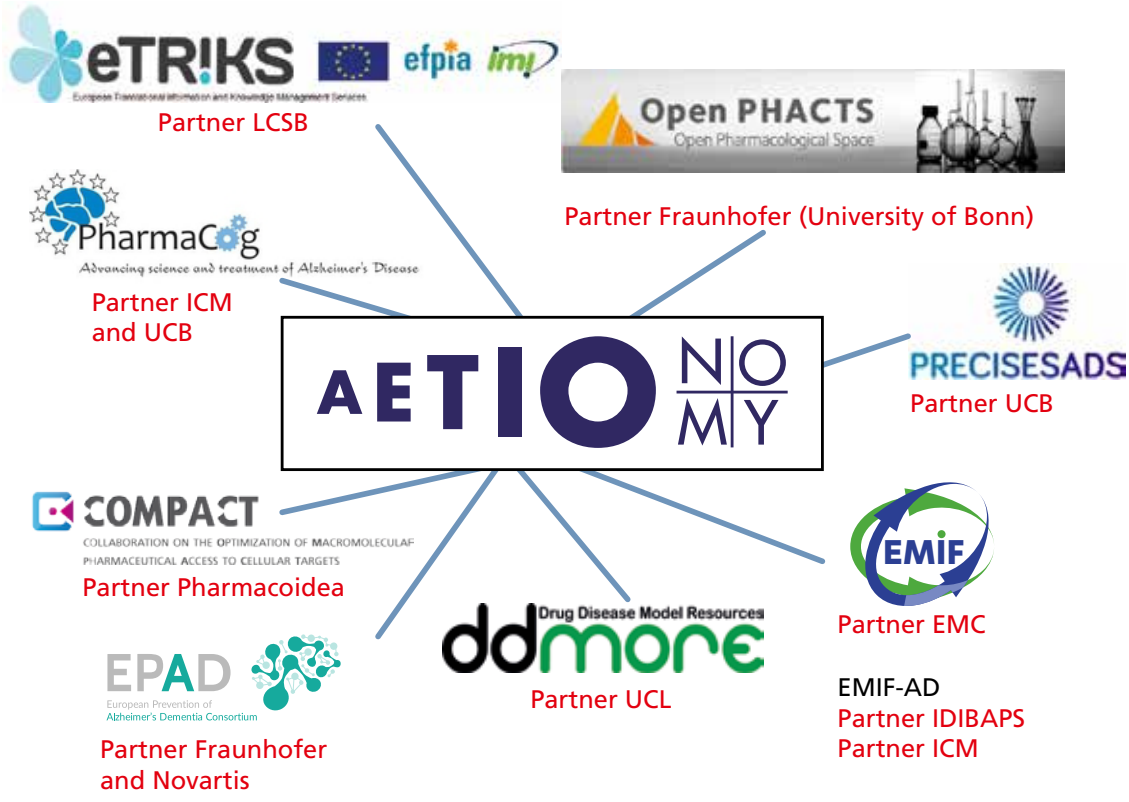
Boehringer Ingelheim

More information at www.boehringer-ingelheim.com

RELATED INITIATIVES

Exchange of Data, Knowledge and Experience

The majority of the EFPIA and academic partners in AETIONOMY are involved in other IMI projects and thus are able to mediate uptake of essential developments made in other IMI projects or feedback results of AETIONOMY to other IMI projects. Of utmost importance for AETIONOMY are the involvement of University of Luxembourg with **eTRIKS**; the involvement of Fraunhofer with **OpenPHACTS**, the coordination of **EMIF** through partner EMC and the involvement of partners ICM in **PharmaCog**, UCL in **DDMoRe** and Pharmacoidea in **COMPACT**. Moreover, two partners in AETIONOMY (IDIBAPS and ICM) are involved in **EMIF-AD**, the Alzheimer pillar of the EMIF project. Besides these, AETIONOMY partners are active in more than 40 IMI projects.



AETIONOMY and related research initiatives on Alzheimer's and Parkinson's diseases

In the following we would like to briefly introduce initiatives, which are collaborating intensively with AETIONOMY.

IMI AD Platform

Innovative Medicines Initiative Alzheimer's Disease Platform

- IMI Alzheimer's disease projects AETIONOMY, EMIF and EPAD have a combined budget of €138 million and jointly address many key challenges for medicines research and development.
- Collaboration via IMI Alzheimer's Disease Research Platform will enable faster progress.
- The platform will have a global reach through a Memorandum of Understanding between IMI and the Global Alzheimer's Platform (GAP).
- Announcement comes at 12th International Conference on Alzheimer's and Parkinson's Diseases and Related Neurological Disorders (AD/PD 2015) and in wake of major WHO conference on dementia.

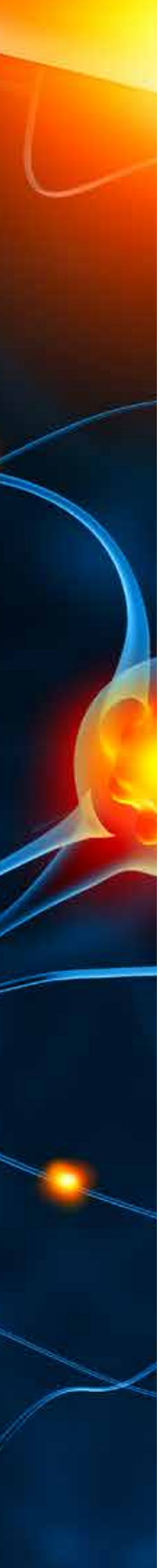
The Innovative Medicines Initiative (IMI) and its AETIONOMY, EMIF and EPAD projects announced the establishment of the IMI Alzheimer's Disease Research Platform. The platform will facilitate collaboration between the three projects, helping them to deliver results faster and more effectively. Additionally, IMI and the Global Alzheimer's Platform (GAP) are announcing their plans to sign a Memorandum of Understanding to accelerate Alzheimer's drug development by building a global, standing, trial-ready platform for Alzheimer's drug development.

The announcements came during a symposium held at the 12th International Conference on Alzheimer's and Parkinson's Diseases and Related Neurological Disorders (AD/PD 2015) in March 2015, and in the wake of a major World Health Organization (WHO) conference on dementia.

Dementia already affects over 35 million people globally, and as populations age, this figure is set to rise to over 115 million by 2050. The disease places a huge and growing burden on health and social care systems and on the families and caregivers of those affected. Yet despite decades of research, there is still neither treatment nor cure for the disease.

The challenge of developing new, effective treatments for dementia is simply too great for any organisation to tackle alone, and so IMI has launched a number of projects that bring together leading experts from the pharmaceutical industry, universities, small biotech companies, and patient organisations from across Europe and beyond. The three projects in the new IMI Alzheimer's Disease Research Platform have a combined budget of €138 million and address complementary areas of Alzheimer's disease research.

AETIONOMY EMIF EPAD
AD Platform



AETIONOMY is paving the way towards a new approach to the classification of neurodegenerative diseases, particularly Alzheimer's and Parkinson's diseases, thereby improving drug development and increasing patients' chances of receiving a treatment that works for them.

EMIF is developing a common information framework of patient-level data that will link up and facilitate access to diverse medical and research data sources, opening up new avenues of research, particularly in the fields of Alzheimer's disease and obesity.

EPAD is pioneering a new, more flexible approach to clinical trials of innovative Alzheimer's disease treatments designed for people who have the disease but have not yet developed dementia.

'The European Union has a long tradition of fostering research collaboration,' said Jean Georges, Executive Director of Alzheimer Europe, which is a partner in all three projects. 'The creation of the IMI Alzheimer's Disease Research Platform is another great example of European research projects working together to improve our understanding of dementia and to give hope to the 8.4 million Europeans affected by dementia of a cure of the condition in the future. Alzheimer Europe is delighted to support all three projects by representing the views of people with dementia and their carers in the research consortia and by making the research results available to the wider general public.'

The projects are keen to collaborate more closely with other Alzheimer's research projects around the world. The global reach of the platform will be aided by the signature of a Memorandum of Understanding between IMI and the Global Alzheimer's Platform (GAP). In addition, international collaboration is already built into the IMI projects. For example, EPAD, EMIF and the Medical Research Council's Dementias Platform UK are already linked and a number of EPAD partners are directly involved in other initiatives such as GAP.

'Irene Norstedt, IMI Acting Executive Director commented: 'Alzheimer's disease is a global challenge that requires a global solution, and it is in this spirit that the IMI Alzheimer's Disease Research Platform is reaching out to other initiatives on Alzheimer's disease around the world. Everyone working on Alzheimer's disease needs to pull together if we want to deliver results that will help us to end the suffering caused by this terrible disease.'

AETIONOMY EMIF EPAD
AD Platform

EMIF-AD

European Medical Information Framework

The EMIF project aims to develop a common information framework of patient-level data that will link up and facilitate access to diverse medical and research data sources, opening up new avenues of research for scientists. To provide a focus and guidance for the development of the framework, the project will focus initially on questions relating to obesity and Alzheimer's disease. In the subproject EMIF-AD the challenge will be to identify predictors of Alzheimer's Disease (AD) in the pre-clinical and prodromal phase, with the support of EMIF-Platform.

Recent years have seen an explosion in the number of databases containing medical and research data, from Electronic Health Records (EHRs), cohort studies (in which a group of individuals are followed for a number of years), disease-specific studies, and biobanks, to name a few. Because this data is scattered across diverse platforms, it cannot be fully exploited. The table shows data available through AD cohorts:

Number of subjects	Controls	SCI	MCI
All	31219	3173	11976
Plasma	3537	1764	3961
DNA	8763	3163	5731
RNA	1637	484	2128
CSF	846	1336	3757
Urine	2307	1116	2349
MRI	3449	1035	5586
FDG-PET	551	308	1942
Amyloid PET	335	143	492
EEG	301	773	1173

Either access to results from analysis or access to samples/scans

Linking up the data would allow scientists to significantly advance medical research and drug development. However, in practice, this is rather difficult. Not only is the data fragmented, differences in coding systems and languages, plus legal and ethical restraints, hamper efforts to combine these sources of data. Furthermore, there are often information gaps. There is therefore a need for a single system that allows researchers to link data on an immense scale, including patient health records, research data, survey and administrative data, imaging, social, environmental and economic data. Such a system should also be able to bring together data from different populations; this would increase sample sizes and facilitate the study of rare or highly specific subgroups.

EMIF will develop a common information framework that will not only facilitate access to existing data sources, but ease the creation of links between sources and, where needed, collect additional information. The work will require the team to address a number of issues, including data standards, semantic interoperability, ethics, data privacy, legal issues, and the development of an IT platform that allows access to multiple data sources. Obesity and dementia are two of the greatest healthcare challenges of our time; EMIF's work will pave the way for new diagnostic tools and treatments to help patients with these conditions.

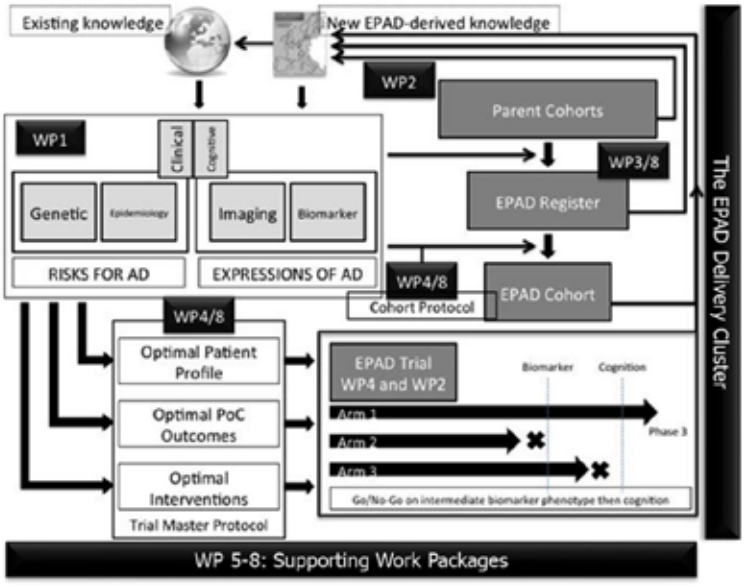
More Information at www.emif.eu

EPAD

European Prevention of Alzheimer's Dementia Consortium

The **EPAD** project aims to **develop an infrastructure that efficiently enables the undertaking of adaptive, multi-arm Proof of Concept studies for early and accurate decisions on the ongoing development of drug candidates or drug combinations**. This includes evaluating patients' reactions to a drug early in a clinical trial and modifying the trial according to these reactions. The EPAD project will initially run for five years. The platform will draw European participants, whose records are already part of existing national/regional cohort or register studies, into an EPAD register of approximately 24,000 people. From this group, 6,000 people will be asked to join a pan-European EPAD Cohort for consistent, longitudinal follow-up, and approximately 1,500 of them will eventually be invited to enter the standing EPAD Proof of Concept Trial. This approach aims to ensure EPAD has access to an at-risk population showing biomarker evidence of Alzheimer's disease prior to the development of dementia. All **data collected from the cohort and trial will become publically available** for analysis to improve disease models in the pre-dementia phase of Alzheimers' disease. This should lead to more accurate stratification for trial selection, improved measure-ments of outcomes and a greater understanding of Alzheimer's disease processes before dementia develops. This project has numerous advantages over current approaches including: excellent pre-trial characterisation of subjects to inform selection and reduce screen failure; establishment of the highest possible quality **Trial Delivery Centres (TDC's)** across Europe; rapid decision making on the likely success of a drug (or combination of drugs) in subsequent confirmatory trials and access to a shared placebo group.

The project is divided into **eight Work Packages, shown with their inter-processing in the following graphics:**



More information at www.ep-ad.org

EPAD
European Prevention of Alzheimer's Dementia Consortium

REFERENCES

Publications and Articles

We do emphasize the role and importance of joint publications, reflecting the close collaboration between industrial and academic partners in AETIONOMY, which will soon be intensified.

Hofmann-Apitius, M.; Alarcón-Riquelme, M. E.; Chamberlain, C.; McHale, D. (2015): *Towards the taxonomy of human disease.* In: Nature Reviews Drug Discovery 14, 75–76 (2015), doi:10.1038/nrd4537.

José Luis Molinuevo, Pablo Ripolles, Marta Simó, Albert Lladó, Jaume Olives, Mircea Balasa, Anna Antonell, Antoni Rodriguez-Fonells, Lorena Rami (2014): *White matter changes in preclinical Alzheimer ' s disease: a magnetic resonance imaging-diffusion tensor imaging study on cognitively normal older people with positive amyloid b protein 42 levels.* In: Neurobiology of Aging, Volume 35, Issue 12, December 2014, Pages 2671–2680.

Anandhi Iyappan, Shweta Bagewadi, Matthew Page, Martin Hofmann-Apitius , and Philipp Senger (2015): *NeuroRDF: Semantic Data Integration Strategies for Modeling Neurodegenerative Diseases. In: Proceedings of 6th International Symposium on Semantic Mining in Biomedicine (SMBM).*

Imprint

Prof. Duncan McHale
UCB Biopharma SPRL
Vicepresident, Head of Global Exploratory Development
Allée de la Recherche, 60
1070 Brussels, Belgium
duncan.mchale@ucb.com

Photo Acknowledgements:
Cover and Back: fotolia; All other photos: project partners

IMI Projects AETIONOMY, EMIF, EPAD:
The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertalking under grant agreements n° 115568, 115372, 115736 resources of which are composed of financial contribution from the European Union´s Seventh Framework Programme (FP7/2017-2013) and EFPIA companies´ in kind contribution.

More information:
www.aetionomy.org
www.imi.europa.eu