A Precision Medicine Initiative for Alzheimer's disease: the road ahead to biomarker-guided integrative disease modeling.

Harald Hampel, Sid O'bryant, Stanley Durrleman, Erfan Younesi, Katrine Rojkova, Valentina Escott-Price, Jean-Christophe Corvol, Karl Broich, Bruno Dubois, Simone Lista

A Precision Medicine Initiative for Alzheimer's disease – The Road ahead to Biomarker-guided Integrative Disease Modeling*

Harald Hampel[a,b,*], Sid E. O'Bryant[c], Stanley Durrleman[d,e], Erfan Younesi[f], Katrine Rojkova[a,b], Valentina Escott-Price[g], Jean-Christophe Corvol[h], Karl Broich[i], Bruno Dubois[b], Simone Lista[a,b,*] for the Alzheimer Precision Medicine Initiative (APMI)

[a]*AXA Research Fund & UPMC Chair, Paris, France*

[b]*Sorbonne Universités, Université Pierre et Marie Curie (UPMC) Paris 06, Inserm, CNRS, Institut du cerveau et de la moelle (ICM), Département de Neurologie, Institut de la Mémoire et de la Maladie d'Alzheimer (IM2A), Hôpital Pitié-Salpêtrière, Boulevard de l'hôpital, F-75013, Paris, France*

[c]*Institute for Healthy Aging, University of North Texas Health Science Center, Fort Worth, TX USA*

[d]*ARAMIS Lab, Inria Paris, 75013, Paris, France*

[e]*Sorbonne Universités, UPMC Univ Paris 06, Inserm, CNRS, Institut du cerveau et la moelle (ICM) - Hôpital Pitié-Salpêtrière, Boulevard de l'hôpital, F-75013, Paris, France*

[f]*European Society for Translational Medicine, Schwindgasse 4/7, 1040, Vienna, Austria*

[g]*Medical Research Council (MRC) Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, Wales, UK*

[h]*Sorbonne Université, Université Pierre et Marie Curie (UPMC) Paris 06 UMR S 1127, Institut National de Santé et en Recherche Médicale (INSERM) U 1127 and CIC-1422, Centre National de Recherche Scientifique (CNRS) U 7225, Institut du Cerveau et de la Moelle Epinière (ICM), Assistance Publique Hôpitaux de Paris, Département de Neurologie, Hôpital Pitié-Salpêtrière, F-75013, Paris, France*

[i]*President, Federal Institute for Drugs and Medical Devices (BfArM), Bonn, Germany*

*** Corresponding Authors:**

Harald Hampel, MD, PhD, MA, MSc

Simone Lista, PhD

AXA Research Fund & UPMC Chair,

Sorbonne Universités, Université Pierre et Marie Curie (UPMC) Paris 06, Inserm, CNRS,

Institut du cerveau et de la moelle (ICM), Département de Neurologie,

Institut de la Mémoire et de la Maladie d'Alzheimer (IM2A), Hôpital Pitié-Salpêtrière,

47 Boulevard de l'hôpital, F-75013, Paris, France

Phone: +33 1 57 27 44 81

Fax: +33 1 42 16 75 16

E-mail: harald.hampel@med.uni-muenchen.de (H. Hampel)

E-Mail: slista@libero.it (S. Lista)

Manuscript Requirements

Title character count	122
Word count Abstract	220 words
Word count Manuscript	5,787 words
Number of References:	94
Number of Figures:	2
Number of Tables:	1

Short Title (no more than 20 letters): PM initiative in AD

***This work was partly presented by HH in a plenary lecture at the International Menopause Society's (IMS) 15th World Congress on Menopause, 30th September 2016, Prague, Czech Republic**

ABSTRACT

After intense scientific exploration and more than a decade of failed trials, Alzheimer's disease (AD) remains a fatal global epidemic. A traditional research and drug development paradigm continues to target heterogeneous late-stage clinically phenotyped patients with single "magic bullet" drugs.

Here, we propose that it is time for a paradigm shift towards the implementation of precision medicine (PM) for enhanced risk screening, detection, treatment, and prevention of AD. The overarching structure of how PM for AD can be achieved will be provided through the convergence of breakthrough technological advances, including big data science, systems biology, genomic sequencing, blood-based biomarkers, integrated disease modeling and P4 medicine. It is hypothesized that deconstructing AD into multiple genetic and biological subsets existing within this heterogeneous target population will provide an effective PM strategy for treating individual patients with the specific agent(s) that are likely to work best based on the specific individual biological makeup.

The Alzheimer's Precision Medicine Initiative (APMI) is an international collaboration of leading interdisciplinary clinicians and scientists devoted towards the implementation of PM in Neurology, Psychiatry and Neuroscience. It is hypothesized that successful realization of PM in AD and other neurodegenerative diseases will result in breakthrough therapies, such as in oncology, with optimized safety profiles, better responder rates and treatment responses, particularly through biomarker guided early preclinical disease stage clinical trials.

KEY WORDS

Alzheimer's disease; precision medicine; precision medicine initiative; Alzheimer precision medicine initiative; systems biology; systems medicine; biomarkers; integrative disease modeling.

INTRODUCTION

The technology industry has entered the field of medicine and boldly aims to eliminate disease itself. It may well succeed in a focused global interdisciplinary effort because of a convergence of exponentially advancing technologies, including big and deep data science, computing, artificial intelligence, sensors, and genomic sequencing. Sequencing of the human genome, first completed and published in Nature in 2001[1] took years for completion at a cost of about $3 billion. Today, this is possible in less than a day for about $1,000, with costs falling so fast that, by 2022, genome sequencing may be cheaper than a blood test. Now that it has been mapped into bits that computers can process, the genome is transforming into an information technology. With increasingly large sample sizes and tools such as IBM's artificial intelligence system – Watson – scientists are gaining a better understanding of how genes affect health and disease; how food, physical and cognitive exercise, and medicines we take affect the complex interplay between our genes and environment.

We are going to see more medical advances in the next decade than happened in the past century. Within the coming years, our genome, epigenome, transcriptome, proteome, metabolome, microbiome, interactome, brain network connectome, motor and sensory systems, cognition, behavior, lifestyle, and environment will all be mapped, stored, analyzed, integrated, and individualized. Precise and prescriptive-medicine systems supported by artificial intelligence will help us feel better, be healthier and live longer.

On January 20 2015, U.S. President Obama outlined the great opportunities for Medicine through Precision Medicine (PM) in his State of the Union Address and announced a national PM Initiative (PMI) on January 30 (https://www.whitehouse.gov/precision-medicine), flanked by the publication of the PMI Cohort Program (PMI-CP) by Francis Collins and Harold Varmus in the New England Journal of Medicine on that day[2]. Since then, the PMI-CP is establishing an infrastructure and organization to coordinate and enroll the large national research cohort of one million participants in 2016, supported by a $215 million in federal support. The PMI-CP targets oncology challenges as a first priority but is also expected to expand and transfertilize to other relevant disease areas.

After more than a decade of failed therapy trials and one of the lowest success rates in drug development in medicine, the time has come to launch an international Alzheimer PMI (APMI) and link it with the U.S. PMI and other related global initiatives. The establishment of a PM paradigm for Alzheimer's disease (AD), an exponentially growing complex polygenic brain disease, requires the incorporation of an array of converging breakthrough

technological developments and methods. Systems theory allows for the conceptualization of novel and original models to elucidate all systems levels (assessed by systems biology [SB] and systems neurophysiology) and different data types in space and time of the complex, non-linear, dynamic, and chronically progressive nature of the genetically, biologically, pathologically, and clinically heterogeneous construct of "AD"[3,4], a historical term defined by the clinical description of first patients and related first brain histopathological observations[5]. For more than 100 years, Alois Alzheimer's pioneering AD syndrome was the target of scientific exploration. A major step forward was the discovery of single gene mutations on chromosomes 21, 14, and 1 resulting in an overproduction of the amyloid beta (Aβ) peptide[6] and causing a progressive linear mechanistic neurodegenerative disease leading to familial early onset (30-60 years) AD dementia (EOAD) in very small subsets of individuals (<5%). From this mutation model, transgenic animals and the majority of AD anti-amyloid drug development programs were generated[7,8]. To date, a major barrier towards a next evolutionary step to the PM paradigm for AD is the prevailing "implicit" assumption that the cellular, molecular pathophysiological mechanisms and the biological endophenotype of EOAD can be perceived as the original linear mono-mechanistic amyloid cascade model[9] and extrapolated to the pathophysiology of the complex polygenic sporadic late-onset AD (LOAD), representing the vast majority of all affected AD patients[3]. To date, the AlzGene database demonstrates substantial genetic heterogeneity in LOAD patients, 695 genes with 2,975 polymorphisms, more than 20 genome-wide associated risk variants have been described (available at http://www.alzgene.org/). However, in a number of the known over 200 autosomal dominant mutations leading to EOAD different biomarker profiles and clinical phenotypes were demonstrated, therefore being far fromhomogeneous[10,11]. Moreover, even in EOAD, the accumulation of pathological Aβ peptide and tau protein never occurs without disruption of other major pathophysiological systems (e.g. inflammation, oxidative stress, metabolic alterations), thus again underlining the genetic and biological complexity of the AD construct in general[3,4]. In order to untangle this complexity through deconstruction of AD into multiple genetic and biological subsets, advancing biomedical research provides a variety of data from patients' complex and diverse pathophysiology through innovative, converging exploratory biomolecular tools and neuroimaging modalities. The resulting heterogeneous, multidimensional big and deep data are in the process of being standardized and integrated via computational and data science methods in the form of mechanistic disease models, according to the integrative disease modeling (IDM) conception (**Figure 1**)[12].

In this perspective, we outline key aspects and issues for transformation and implementation of an AD PM paradigm to advance both treatment and prevention strategies in AD. Notably, this results into an innovative scientific taxonomy, a differentiated working language (**Table 1**) for reality-based medicine, which identifies evidence from real-life scenarios.

This and the theoretical background presented in our recently published perspective *"PRECISION MEDICINE - The Golden Gate for Detection, Treatment and Prevention of Alzheimer's Disease"*[13,14] provides the theory and roadmap for a new scientific movement framed by the newly established *Alzheimer Precision Medicine Initiative (APMI)*. APMI and its planned cohort program (*APMI-CP*) is an international network of leading interdisciplinary clinicians, scientists and researchers devoted towards the transformation of Neurology and Psychiatry through the implementation of PM, a reintegration of neuroscience and its clinical specialties into the theoretical framework of medicine. It is hypothesized, like in parts of oncology already achieved, that successful and consequent implementation of PM for AD and other neurodegenerative diseases along the disease spectrum of brain proteinopathies will result in breakthrough single and combination therapies, with optimized safety profiles, better responder rates and more convincing treatment responses, particularly in early disease stage clinical trials, providing substantial benefits to individuals on the way to or patients suffering from this devastating disease.

A PRECISION MEDICINE PARADIGM FOR ALZHEIMER'S DISEASE

The adoption of a PM paradigm for generating innovative strategies to treat, prevent and cure complex diseases is not a novel concept. For decades, the oncology field was at loss as to how to treat patients inevitably dying from advanced late-stage cancer; however, today, mortality and cure rates, at least for some forms of cancer, are far beyond initial expectations.

On the other hand, after over 100 years of accumulating scientific knowledge, there is no therapeutic solution for prevention and cure in AD which remains 100% fatal. Available treatments, approved for late potentially irreversible clinical disease stages only, offer marginal clinical benefits. It is time for a paradigm shift, with the field of oncology providing a previously validated model for successful implementation of a PM model that the AD/neurodegeneration field can at least partially adopt[13].

The concept of PM aims at tailoring medical treatment to the individual genetic drivers, pathophysiological and clinical characteristics of the disease for each single patient[15]. In other

words, it aims at tailoring disease prevention and treatment to the individual's specific biological makeup (customized treatment), which is in sharp contrast to the ongoing "one-drug-fits-all" approach. Given the highly complex nature of AD, the likelihood of identifying a single drug to provide meaningful benefit to every patient is minimal, at best. This is the situation in other areas such as oncology and cardiology. A key methodological framework required for successfully implementing the PM is the incorporation of the exploratory, integrative, and interdisciplinary systems approach of SB, complemented by systems neurophysiology[3,4].

SB allows a system level approach to drug discovery − with special reference to drug target identification, validation, and screening assay development – that embraces the whole complexity of disease pathophysiology. Recent years have witnessed significant success in biomarker-guided therapeutic strategies in advanced translational research fields of biomedicine – including oncology and cardiovascular medicine. The traditional reductionistic categorical nosology of "neurodegenerative diseases" reflects advanced late stages of fragmented clinical phenotypes and syndromes with different or overlapping histopathological patterns. Although continuous working group efforts to refine categorical diagnostic criteria improved the diagnostic reliability and accuracy, particularly after integrating biomarkers as part of the criteria[16], the validity of current categorical nosological systems for neurodegenerative diseases remains limited. A step in the right direction is represented by the recent formulation of unbiased agnostic biomarker classification systems for AD and neurodegenerative diseases to identify and grade risk in normal elderly people. The goal is to identify the full spectrum of the specific biological alterations in elderly individuals at risk long before the appearance of first clinical symptoms[13].

We hypothesize that using PM in the fields of Neurology, Psychiatry, and Neuroscience will trigger a paradigm shift in the medical practice of brain diseases towards preclinical detection and effective early interventions. Prevention strategies can be employed before any substantial disease progression has occurred, with a strong focus on individualized care.

Among the objectives of PM are to introduce new paradigms for early detection, classification/differential diagnosis, treatment, and prevention of neurodegenerative diseases (better proteinopathies of the brain), based on individual biological differences, as reflected by multimodal biological indicators, biomarkers[4,17,18]. In this regard, evolving evidence of AD biomarkers has been obtained during the last 20 years from studies performed in neurogenetics/neuroepigenetics[19-21], neurochemistry[22-24] – the latter having been conducted

both on cerebrospinal (CSF)[25-27] and blood (plasma/serum)[28-32] – as well as in structural/functional/metabolic neuroimaging[33-35], and neurophysiology[36,37]. Following the oncology model, it is anticipated that innovative biomarker studies, combined with SB, will identify specific diagnostic, prognostic, and predictive biomarker signatures in order to tailor the therapy to individual patients[13]. Additionally, biomarker-guided PM removes today's "trial-and-error" strategy to pharmacological interventions, which has significant medical consequences for patients and healthcare system[38]. As stated by the Institute of Medicine (IOM) Committee Recommendations for Advancing Appropriate Use of Biomarker Tests (companion diagnostics) for Molecularly Targeted Therapies, the ultimately goal of PM is to improve both the quality of patient care and clinical outcomes[39].

In summary, the PM conception is in the process of being applied to AD and across a rapidly increasing number of other neurodegenerative diseases owing to: (I) the development of high-throughput "omic" tools designed for screening biomedical samples and (II) the setting-up of large-scale biological datasets. Consequent development, validation, and implementation of biomarker-guided interventions built on the SB conceptual framework will accelerate the path to PM for AD.

SYSTEMS THEORY AND SYSTEMS BIOLOGY PARADIGMS

In order to achieve PM, innovative theoretical concepts and strategies need to be embraced. The traditional reductionistic approach in AD research aims to characterize single pathophysiological pathways affecting specific components of particular systems in a linear, non-dynamic and over-simplified manner. This myopic view has resulted in a limited representation of complex pathophysiological processes and their interactions. For instance, the prevailing amyloid cascade hypothesis speculates that the Aβ peptide is the cause of AD and that, as a result, targeting Aβ should lead to substantial disease modification at advanced clinical stages in LOAD. This assumption has been challenged by numerous failures of phase III clinical trials aimed at modulating Aβ production[40] or increasing clearance from the brain[41,42]. Additionally, as stated previously, in no complex model of LOAD Aβ seems to occur in isolation of biological dysfunction related to other systems. An agnostic, hypothesis-free, unbiased systems theory approach seems better suited to explain the complex and heterogeneous origin and time course of the pathophysiological failure underlying different forms of AD[4]. For multifactorial diseases like AD, comprehensive holistic systems-level approaches are necessary; this is the case of the SB model, which aims at understanding the

genotype-phenotype relationships and the mechanisms at the level of genome/epigenome, transcriptome, microRNome, proteome/peptidome, metabolome/lipidome, microbiome, lifestyle, and environmental factors participating in complex cellular networks[3,4,43]. Correspondingly, SB is based on: (I) advanced molecular and high-throughput "omics" methods disclosing and characterizing biomarkers associated with disease mechanisms and (II) computational and integrative network biology tools for assimilating multimodal information to comprehensively understand the systems-level dysfunction[44]. Longitudinal investigations using the above mentioned SB-based methodologies can provide a full characterization of the complex molecular pathophysiology of both single gene and sporadic forms of AD. The working hypothesis is that most if not all AD subforms evolve through non-linear dynamic convergence of alterations and/or failures in several "systems", networks, signalling pathways, or pathophysiological processes[3]. As a result, the specific intervention needed for a particular individual would depend on the specific system-level alteration and/or dysfunction at a given time point, which may change as a function of time and progression with (i.e. the specifically effective treatment for a given patient may vary over time).

"OMICS"-BASED TECHNOLOGIES FOR BIOMARKER IDENTIFICATION

PM is biomarker-guided medicine. According to the Food and Drug Administration (FDA) & the National Institutes of Health (NIH) Biomarkers, Endpoints, and other Tools (BEST) Resource, biomarker categories can be divided into the following categories: (I) susceptibility/risk biomarker, (II) diagnostic biomarker, (III) monitoring biomarker, (IV) prognostic biomarker, (V) predictive biomarker, (VI) pharmacodynamic/response biomarker, and (VII) safety biomarker[45]. In the AD field, however, such fine grained separation between different types of biomarkers is largely absent. For example, it is assumed that amyloid positivity is both a diagnostic and predictive biomarker, which may or may not be the case for given patients. However, the fine-grained specifications of the exact function of each biomarker (or biomarker profile) is required to advance PM in AD[45]. When combining this fine-grained categorization of specific types of biomarkers with the evolution of the "omic" technologies currently available under the SB methods, there now exists the foundation for building the PM based paradigm for treating and preventing AD across the spectrum of disease progression[45]. In genomics, the development of less expensive and comprehensive genome-wide arrays paved the way to the genome-wide association studies (GWAS). However, although initial results were promising, numerous GWAS were disappointing due

to inadequate sample size, limitation of arrays for certain genetic variations (genetic markers), and/or heterogeneity in phenotype[46,47] as well as the focus on finding the gene(s) responsible for AD rather than looking for subsets of AD cases. Big collaborations, such as the International Genomics of Alzheimer's Project (IGAP), and advanced genomic imputation techniques (*in silico*) generated highly consistent GWAS results[48,49], which replicate and provide insights to underlying biological pathways. Notably, the introduction of next generation sequencing (NGS)-based methods led to significantly improve the genomic analyses. Particularly, unbiased whole-genome sequencing (WGS) and whole-exome sequencing (WES) support the identification of many genetic variants, including SNPs, single nucleotide variants, small insertions/deletions, and structural and genomic variants[50,51]. Besides genomics, high-throughput screening methods led to substantial AD-related discoveries in other "omic" areas, especially proteomics[52,53] and metabolomics/lipidomics[54-56] that may change over time by contrast to the genome.

The "omics"-based screening of disease states is supposed to result in improved personalized, mechanistically-based interventions (therapeutic and/or preventive) by revealing precise patterns of biomarkers and molecular signatures underlying the exact molecular pathophysiological mechanisms active in specific disease states and in individual patients[57]. Substantial attempts are ongoing to explicate key pathways functions, signalling network organization, and organism-level responses via high-throughput biological data (for instance, global gene expression, comprehensive proteomic data)[58].

Notably, applying SB to blood-based "omic" technologies to promote the PM paradigm for AD will enable two primary advances for improved patient outcomes[13,59]: (I) generation and validation of enhanced multi-stage neurodiagnostic processes and (II) identification of targeted therapeutic intervention strategies for specific patients or subgroups of patients[59]. As with the PM paradigm successfully implemented in oncology, a primary key to success is the generation of early detection biomarkers identifying patients before significant pathological accumulation. As with other frontline detection strategies, blood-based tools detecting patients within primary care settings in the earliest stages of disease progress will foster a multi-stage diagnostic process for appropriate referrals to CSF and positron emission tomography (PET) biomarker methods. Additionally, once such a multi-stage process is established, it would provide support to the global AD clinical trials community. The second advancement will be the identification of which specific patients are most likely to benefit from precise and definite interventions. Applying SB for the analysis of multi-level blood-

based “omic” data will facilitate the segregation of patient populations into biologically-based subgroups that can be further scrutinized for targeted interventions. Using SB methods to create diagnostic biomarkers of specific subsets of AD patients will have a tremendous impact on the advancement of the PM paradigm in AD[59].

THE PATH FROM “BIG DATA” TO “SMART DATA”

SB aims at exploring the enormous complexity of biological systems by (I) outlining the components of the system, (II) clarifying their interrelations, and (III) defining their spatio-temporal dynamics needed for executing their biological functions[60]. According to the Workshop “From Systems Biology to Systems Medicine” – organized by the European Commission, Directorate of Health – the application of SB-based strategies to medical research/practice is referred to as systems medicine. The objective of systems medicine is to integrate a variety of massive biomedical data at all levels of the cellular organization by employing global, integrative, and dynamic statistical and computational modeling to elucidate the pathophysiological mechanisms, diagnosis, prognosis, and therapy of the disease (https://ec.europa.eu/research/health/large-scale/pdf/systems-medicine-workshop-report-june-2010_en.pdf). The era of “omics” sciences – describing complex biological systems in an integrative, non-reductionistic (holistic) manner – led to the generation of large-scale and heterogeneous biomedical data and allowed entering the area of “big data” in Biology and Medicine[13]. Big data is a comprehensive expression referring to the complexity, challenges, and new opportunities presented by the combined analysis of data. These data sources include the heterogeneous, complex, disorganized, massive, and multidimensional data (from molecular/cellular data, to conventional clinical data, to enormous amounts of imaging, demographic, and environmental data) extensively produced by academic institutions, clinics, and mobile devices[60,61]. These datasets, due to their large sizes and complexities, cannot be analyzed using the traditional ways of processing the data. Big data usually display: (I) significantly enormous amount of data, (II) elevated speed of data production, and (III) heterogeneity of data generated by using different modalities. Such features are typically found in several large assortments of data[62]. In this regard, the Obama Administration announced the Big Data Research and Development Initiative (https://www.whitehouse.gov/blog/2012/03/29/big-data-big-deal) aimed at targeting personalized medicine through the Genomic Information System for Integrated Science (GenISIS) program to improve health care for Veterans. Of note, in 2012, the U.S. National

Institutes of Health (NIH) launched the Big Data to Knowledge (BD2K) Initiative (https://datascience.nih.gov/bd2k), to support the research of innovative methods to speed up the integration of big data into biomedical research[63]. In order be successful, the PMI must utilize innovative methods for collecting/managing big data. This is accomplished by remarkable progresses in information technology that provided significant reductions in terms of costs of data storage, and substantial increases in analytic capabilities, thus enabling the collection and examination of exceptionally large datasets in biomedicine. Particularly, the development and implementation of electronic health records (EHRs) allow gathering/preserving longitudinal health care records and clinical data at limited costs. EHRs represent a key source of clinical data to examine biological and environmental contributions to a large number of conditions and health outcomes. Additionally, there has been an exponential growth in terms of adoption of personal mobile technologies – including phones, apps, wearables, in-home devices – as innovative way to collect health information (mobile health or "mHealth") aimed at collecting clinical relevant information in a more ecological environment, and improving patient care and advancing research. Data generated from increasingly sophisticated software applications can enrich self-reported data on lifestyle and environment, thus providing researchers with a well-defined view into these factors previously difficult to capture.

The holistic paradigm of systems medicine utilizes all assortments of biological information – DNA/RNA, protein/peptides, metabolites/lipids, other small molecules, cells, tissues, organs, individuals, social networks, and external environmental signals – integrating them in such a way that predictive and "actionable" models for health and disease are generated[60]. Presently, unparalleled amounts of heterogeneous data are being gathered with content in AD, ranging from genetic/epigenetic and molecular "omic" disciplines to clinical phenotypes of patients. The production of such big data is expected to radically renovate the development of effective therapies for AD, under the condition that such data are converted into "actionable" knowledge[64]. In the AD domain knowledge is defined as "actionable" when it can be utilized to actively support drug discovery & development programs for therapeutic interventions, to define potential groups of responders to specific targets, and to validate clinical data that can indicate the presence of substantial changes during the advancement of the disease[65]. The integration of large clinical datasets is considered as a potentially powerful approach to accelerate medical discovery based on recent results of world-wide studies of disease progression and large-scale genomics efforts[66]. Innovative analytical methods have

been developed both in the field of bioinformatics[67,68] and in pharmacology[69]. According to Geerts and colleagues (2016), the attempts of gathering large-scale data, such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) consortium (http://adni.loni.usc.edu/) are assumed to yield maximal impact if they are combined with advanced predictive modeling approaches where large neurobiological domain expertise is formally integrated. Such an approach is referred to as converting big data into "smart data" with the aim of producing "actionable" knowledge. This strategy is envisioned to emphasize the link between the domain of data-information and that of "actionable" knowledge, i.e. the notion of "smart data"[64]. The creation of "actionable" knowledge is intended to support the development of novel paradigms that will be used for therapeutic purposes. In this regard, the Brain Health Modeling Initiative (BHMI) has been recently introduced[64] with the objective of accelerating the development and validation of biomarkers and therapeutic agents by highlighting the role of integrative analytic tools and mechanism-based computational methods. These will improve the mining of complex big datasets to obtain an increasingly accurate and "actionable" understanding of the disease. This initiative is expected to result in the development of more successful treatments or an improved efficient screening of patients with AD-specific pathophysiology or an enhanced match between biomarkers and therapeutic targets[64,70]. This will fully support the application of the PM paradigm[2] involving matching patients to a therapy focused on the pathophysiological features of their disease. In summary, the "big data" advancements provide a platform upon which the "omics" data can be understood from a SB standpoint for the generation of a PM paradigm in AD (**Figure 1**).

THE PATH TO INTEGRATIVE DISEASE MODELING AND P4 MEDICINE INCLUDING ONTOLOGY AND TAXONOMY ISSUES

Neurodegenerative diseases or protein misfolding disorders/proteinopathies of the brain leading to neurodegeneration, present a large continuous spectrum of phenotypical subtypes and substantial complexity in terms of genetics, pathophysiology and molecular and topographical progression patterns. This is partly due to the high complexity of the evolutionary driven human brain organization, lack of direct access to brain tissue in living patients, and large heterogeneity and overlapping clinical signs, symptoms and syndromes. As a result, biomedical research in the field of neurodegenerative diseases including AD is presently focused on indirectly obtaining data from the brains of patients using a variety of clinical, molecular neurophysiological and neuroimaging assessments and technologies.

However, a major challenge is that these data are vast, heterogeneous, and scattered. Thus, a multidimensional data space has been created over time that needs standardization, management, integration, and analysis. Computational approaches have been recently developed to facilitate these processes and present the integrated data in the form of mechanistic disease models, i.e. IDM. The IDM methodology aims at linking molecular, neurophysiological and neuroimaging data across multiple physiological levels to clinical readouts and phenotypic observations in such a way that further analytical actions, including mode-of-action simulation and outcome prediction, can be performed on integrative models.

The first and the most important step towards integrative modeling of the multidimensional data space of neurodegenerative diseases is represented by the standardization and representation of the knowledge domain of brain diseases.

Ontologies provide a secure and stable tool to achieve this goal. An ontology is a formal naming and definition of the types, properties, and interrelationships of the entities that really or fundamentally exist for a particular domain of discourse. It is thus a practical application of philosophical ontology, with a taxonomy. The ontology compartmentalizes the variables needed for some set of computations and establishes the relationships between them. The fields of artificial intelligence, the Semantic Web, systems engineering, software engineering, biomedical informatics, library science, enterprise bookmarking, and information architecture all create ontologies to limit complexity and to organize information. The ontology can then be applied to problem solving. The most well-known ontology in the biomedical community is Gene Ontology (GO) (http://geneontology.org/), which organizes expert knowledge about genes, their function, cellular location, and biological processes[71]. However, GO is devoid of disease concept, meaning that gene function and processes have been annotated for physiological conditions. Recently, several attempts have been undertaken to formalize the knowledge domain of neurodegenerative diseases for the use in integrative disease models. Notably, Malhotra and colleagues (2014) pioneered the development of the AD Ontology[72] and Younesi and colleagues (2015) published the first ontology draft for Parkinson's Disease[73]. Other ontologies have been consequently created[74,75]. Interestingly, Iyappan and colleagues (2016) defined a pathway terminology system representing a comprehensive set of signalling pathways and biological events to mine the knowledge domain of AD and visualize the perturbed pathways on top of their corresponding anatomical locations in human brain[76]. Using these domain-specific ontologies and vocabularies, it is possible to annotate and

harmonize both quantitative (e.g. expression metadata) and qualitative (e.g. textual) data. Such curated datasets provide the substrate for data integration.

After standardization and annotation of scattered datasets, these quality datasets should be properly managed and handled within specific computational platforms so that they can become amenable to modeling and further analysis. One such platform is the AETIONOMY knowledgebase, an IMI-funded resource aiming at gathering, organizing, and managing knowledge and data on neurodegenerative diseases with focus on AD and Parkinson's disease (available at http://www.aetionomy.eu/en/vision.html). The AETIONOMY platform is based on the tranSMART architecture, a knowledge management platform assisting researchers to generate data-driven hypotheses (available at http://transmartfoundation.org/).

The integration of curated datasets within data management platforms requires robust modeling methods. Although many algorithms have been introduced to integrate quantitative and qualitative data and visualize them in the form of correlational and causal networks, these networks are usually not multidimensional and model one or two types of relations between a few numbers of biological entities. For instance, the AlzPathway map provides a collection of signalling pathways and biological events taught to be involved in the AD pathogenesis[77]. However, these maps lack of other data dimensions such as SNPs, epigenetic regulators, and clinical outcomes; moreover, they are not amenable to computational reasoning and dynamic simulation. In this regard, Biological Expression Language (BEL) is a modeling language addressing these caveats to a large extent. The principles underlying this modeling approach include standard representation of triples and their relationships (i.e. subject-relation-object) with clear directions for the relations as cause and effect (available at http://openbel.org/).

Probably, the most important utilization of IDM is prediction of outcome. BEL models have been successfully applied to predicting AD pathophysiological mechanisms and biomarker identification tasks[78]. For instance, they can be subjected to reasoning algorithms such as Network Perturbation Analysis to predict the network behaviour under disease conditions using transcriptomic data[79].

Having said that, disease-specific integrative computational models play a crucial role in the IDM paradigm and constitute foundations for "actionable" P4 medicine measures in the area of AD and other neurodegenerative diseases (**Figure 1**). Accordingly, the integrative disease models are expected to support decision making for (I) early diagnosis of brain disease progression with mechanistic biomarkers (predictive), (II) stratification of individuals at high risk of developing neurodegenerative diseases based on mechanistic comorbidities

(preventive), (III) tailoring treatment to the right patient population (personalized), and (IV) optimizing "actionable" plans for the benefit of patients based on patient-centric information in EHRs and patients' feedback in social media[12].

DATA SCIENCE AND INTEGRATIVE DISEASE MODELING: FROM CONCEPTS TO METHODS TO CLINICAL APPLICATION

"Data science", a new scientific field at the interface between mathematics, statistics, and computer science, had already a deep impact on many aspects of human activities, thanks to the automatic prediction of individual behaviour from personal data. Back up by advances in artificial intelligence and machine learning, this field is likely to play a pivotal role in the emergence of PM.

Large multimodal observational studies are acquired within clinical research with the aim to better understand the pathophysiology of diseases and identify potential therapeutic targets. In most cases though, such datasets serve to validate a specific set of hypotheses, with the common belief that data have a short time lapse before their value expires. The potential value of retrospective use of these data with hypothesis-free approaches is currently underestimated. However, they represent a huge potential to design effective systems to automatically position each patient into a disease progression scale, predict symptoms onset, find pathological sub-types, monitor disease progression, and predict treatment efficacy for each patient.

Nevertheless, this potential can only be exploited by the development of specific methodologies. In this regard, neurodegenerative diseases, spanning dynamically over decades in the life of individuals, represent great theoretical and computational challenges. The idea is to design an artificial intelligence system that will synthetize the changes observed in the data into one or several integrative digital models of disease progression that, in turn, will be used to estimate the current and future state of a certain patient given its past observations. Designing such systems requires: (I) building long-term disease progression models from collecting short-term observation data, where each patient is examined a different number of times, at different time-points; (II) integrating in the model various categories of structured data: clinical measurements, biological (blood/CSF) markers, structural/functional/metabolic imaging data, and potentially molecular data; (III) temporally co-registering the disease trajectories of every patient, which may start at a different age with a different pace and pathophysiological pattern; (IV) accounting for inter-individual variability in terms of spatio-temporal patterns of disease progression, since each individual

shows different anatomical/physiological/functional characteristics and these features will change in each patient in a different way; (V) turning descriptive scenarios of disease progression into predictive systems.

The integration of spatio-temporal measurements into a digital model of disease progression is often based on the idea of regressing measurements against an estimated time to disease onset[80,81]. However, the goal of automatic diagnosis/staging/prognosis systems is estimating such a time to disease onset, which cannot be then a pre-requisite. This difficulty led to pragmatic solutions to temporally re-align disease trajectories, especially for set of unstructured measurements without spatial organization like whole images[82-85]. Another challenge is not only describing the scenario of events occurring during the disease course, but also the variability of such a scenario among different individuals. Mixed-effects models seem to be a piece of choice to account for both population and individual effects[86], and, therefore, pave the way to digital models that may be personalized to individual cases.

Such mixed-effect models may be used together with the combination of spatial normalization and temporal alignment, with re-synchronisation of the individual timeline using the concept of "time warps"[87,88] or permutation of discrete events[89]. These ideas led to predictive staging systems, where patients are given an estimated stage of disease progression[89,90].

The question of how to disentangle variations in measurements and in pace of changes from longitudinal observations still remains open. However, recent theoretical developments for estimating the statistical distribution of individual trajectories are promising[91,92]. A practical example of what can be done within this framework is illustrated in **Figure 2** representing Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-Cog) scores of 248 individuals pooled into 4 categories of cognitive functions (i.e. memory, concentration, language, and praxis).

CONCLUDING REMARKS

The field of medicine is at a tipping point, a transformational stage, with emerging paradigm shifts in how we conceptualize medical science and research. Clearly one, if not the major breakthrough objective is set out to accomplish PM, facilitated by converging advances in theory and technology, such as SB, genomic sequencing, exploratory high-throughput analyses, the emergence of blood-based biomarkers, data and computational science, integrated disease modeling, EHRs, smart technologies and P4 medicine. Towards this end,

the PMI-CP (https://www.whitehouse.gov/precision-medicine and https://www.nih.gov/precision-medicine-initiative-cohort-program) – launched in January 2015 by the U.S. President Obama – is noteworthy. Under this program, over a million U.S. citizens are expected to provide their genetic data, biological samples and behavioral data, which will be extensively characterized and subsequently linked to EHRs. High-density data that will be generated through the application of SB will be invaluable to dissect the molecular underpinnings of different common complex diseases, which will lead to the development of safe and effective individually-tailored biomarker-guided therapies.

The fields of Neurology, Psychiatry and Neuroscience have yet to follow these giant footsteps and reintegrate into Medicine embracing PM as primary target of concerted efforts. The field AD and other neurodegenerative diseases (or brain protein misfolding disorders/proteinopathies with neurodegenerative pathology) has a larger economic impact than cancer, yet surprisingly an equivalent and appropriate level of funding has by far never been granted to the global epidemic of AD[93]. Learning from the advances in oncology, there is no choice than to accelerate the acceptance and implementation of the PM paradigm for improved health outcomes, which will have a significant impact on worldwide economic outcomes. After more than a decade of clinical trial failures the acceptance of a PM paradigm for AD research and drug development is gaining *momentum* and this is why the time is right (in the worst of times with exponentially rising investments into R&D and equally decreasing rates of a success in therapy developments) for us to initiate and establish a global initiative, the APMI consortium (**Figure 1**). Consequently, PM in the field of AD and neurodegeneration should target genetic risk and the molecular stages of disease, meaning at the earliest preclinical asymptomatic stage[94], when the disease is potentially reversible, tailored to delay, stop – and possibly prevent – the progression to clinical signs and symptoms. Both citizens (active participants and no longer study "patients") and policy-makers need to become more actively engaged with caregivers, basic scientists, and clinical researchers in a common effort to internationalize, centralize and revolutionize the current approach to clinical and translational neurological and psychiatric research. Inevitably, this requires a radical theoretical and cultural shift from traditional concepts, based on the treatment of late stage diseases guided by heterogeneous clinical phenotypes treated by hypothesized "one-size fits all - magic bullet therapies", to the patient centered PM-based approach, focused on early screening for risk and detection of biology, with customized targeted and biomarker-guided therapies to achieve effective and safe prevention and therapy

grounded on the biological characteristics of the individual patient. The vision and objective of the APMI is to facilitate a paradigm shift and transformation of research towards PM through international interdisciplinary networking and collaboration.

SOURCE OF FUNDING

HH and **JCC** are supported by the AXA Research Fund, the Fondation Université Pierre et Marie Curie and the Fondation pour la Recherche sur Alzheimer, Paris, France. The research leading to these results has received funding from the program "Investissements d'avenir" ANR-10-IAIHU-06 (**HH**, **JCC**).

SEO is supported by the National Institute on Aging of the National Institutes of Health under Award Numbers R01AG051848 and R56AG054073. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

SD is funded by the European Research Council (ERC) under grant agreement No 678304, European Union's Horizon 2020 research and innovation programme under grant agreement No 666992, and the program "Investissements d'avenir" ANR-10-IAIHU-06.

CONFLICT OF INTEREST

HH declares no competing financial interests related to the present article. He serves as Senior Associate Editor for the journal Alzheimer's & Dementia®; he has been a scientific consultant and/ or speaker and/or attended scientific advisory boards of Axovant, Anavex, Eli Lilly and company, GE Healthcare, Cytox, Jung Diagnostics, Roche, Biogen Idec, Takeda-Zinfandel, Oryzon Genomics; and receives research support from the Association for Alzheimer Research (Paris), Pierre and Marie Curie University (Paris), Pfizer & Avid (paid to institution); and has patents as co-inventor, but received no royalties: A patent in vitro multiparameter determination method for the Diagnosis and early diagnosis of neurodegenerative disorders. Patent number: 8916388 Issued. A patent in vitro procedure for diagnosis and early diagnosis of neurodegenerative diseases. Patent number: 8298784 Issued. A patent Neurodegenerative Markers for Psychiatric Conditions. Publication number: 20120196300 Issued. A patent IN VITRO MULTIPARAMETER DETERMINATION METHOD FOR THE DIAGNOSIS AND EARLY DIAGNOSIS OF NEURODEGENERATIVE DISORDERS. Publication number: 20100062463 Issued. A patent IN VITRO METHOD FOR THE DIAGNOSIS AND EARLY DIAGNOSIS OF NEURODEGENERATIVE DISORDERS. Publication number: 20100035286 Issued. A patent In vitro Procedure for Diagnosis and Early Diagnosis of Neurodegenerative Diseases. Publication number: 20090263822 Issued. A patent in vitro method for the diagnosis of neurodegenerative diseases. Patent number: 7547553 Issued. A patent CSF Diagnostic in Vitro Method for Diagnosis of Dementias and Neuroinflammatory Diseases. Publication number: 20080206797 Issued. A patent in vitro Method For the Diagnosis of Neurodegenerative Diseases. Publication number: 20080199966 Issued. A patent Neurodegenerative Markers for Psychiatric Conditions. Publication number: 20080131921 Issued. **SEO** has the following patents pending related to precision medicine: PCT/US2011/036496 and PCT/US2014/067562 (additional patent filed). He has served on an advisory board for and received honoraria from Roche and has equity in Cx Precision Medicine, Inc. **VEP** reports personal fees from Cytox Ltd. **JCC** declares no competing financial interests related to the present article. He has been scientific consultant for BMS, Zambon, Pfizer, Abbvie, Ipsen, Clevexel, Amarantus, and received research grants from the

Michael J Fox Foundation, French Ministry of Health, French Ministry of Research. He owns stock options from B&A Therapeutics. **BD** reports personal fees from Eli Lilly. **SL** has received lecture honoraria from Roche. **EY**, **SD**, **KR**, and **KB** reports no conflicts of interest.

REFERENCES

1 Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860-921

2 Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015;372:793-5

3 Lista S, Khachaturian ZS, Rujescu D, Garaci F, Dubois B, Hampel H. Application of Systems Theory in Longitudinal Studies on the Origin and Progression of Alzheimer's Disease. *Methods Mol Biol* 2016;1303:49-67

4 Hampel H, Lista S, Khachaturian ZS. Development of biomarkers to chart all Alzheimer's disease stages: the royal road to cutting the therapeutic Gordian Knot. *Alzheimers Dement* 2012;8:312-36

5 Alzheimer A. Über einen eigenartigen schweren. Erkrankungsprozeß der Hirnrinde. *Neurologisches Centralblatt* 1906;23:1129-36

6 Kang J, Lemaire HG, Unterbeck A, et al. The precursor of Alzheimer’s disease amyloid A4 protein resembles a cell-surface receptor. *Nature* 1987;325:733-6

7 Games D, Adams D, Alessandrini R, et al. Alzheimer-type neuropathology in transgenic mice overexpressing V717F beta-amyloid precursor protein. *Nature* 1995;373:523-7

8 Schenk D, Barbour R, Dunn W, et al. Immunization with amyloid-beta attenuates Alzheimer-disease-like pathology in the PDAPP mouse. *Nature* 1999;400:173-7

9 Hardy JA, Higgins GA. Alzheimer's disease: the amyloid cascade hypothesis. *Science* 1992;256:184-5

10 Van der Flier WM. Clinical heterogeneity in familial Alzheimer's disease. *Lancet Neurol* 2016. doi: 10.1016/S1474-4422(16)30275-7

11 Portelius E, Andreasson U, Ringman JM, et al. Distinct cerebrospinal fluid amyloid beta peptide signatures in sporadic and PSEN1 A431E-associated familial Alzheimer's disease. *Mol Neurodegener* 2010;5:2

12 Younesi E, Hofmann-Apitius M. From integrative disease modeling to predictive, preventive, personalized and participatory (P4) medicine. *EPMA J* 2013;4:23

13 Hampel H, O’Bryant SE, Castrillo JI, et al. PRECISION MEDICINE - The Golden Gate for Detection, Treatment and Prevention of Alzheimer’s Disease. *J Prev Alz Dis* 2016. doi: 10.14283/jpad.2016.112

14 Khachaturian ZS, Khachaturian AS. The Paradox of Research on Dementia-Alzheimer's Disease. *J Prev Alz Dis* 2016. doi: 10.14283/jpad.2016.117

15 National Research Council (US) Committee on a Framework for Developing a New Taxonomy of Disease. *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease*. Washington (DC): National Academies Press (US), 2011

16 Dubois B, Feldman HH, Jacova C, et al. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol* 2014;13:614-29.

17 Hampel H, Lista S, Teipel SJ, et al. Perspective on future role of biological markers in clinical therapy trials of Alzheimer's disease: a long-range point of view beyond 2020. *Biochem Pharmacol* 2014;88:426-49

18 Hampel H, Lista S. Use of biomarkers and imaging to assess pathophysiology, mechanisms of action and target engagement. *J Nutr Health Aging* 2013;17:54-63

19 Hampel H, Lista S. Alzheimer disease: from inherited to sporadic AD-crossing the biomarker bridge. *Nat Rev Neurol* 2012;8:598-600

20 Zetzsche T, Rujescu D, Hardy J, Hampel H. Advances and perspectives from genetic research: development of biological markers in Alzheimer's disease. *Expert Rev Mol Diagn* 2010; 10:667-690

21 Lista S, Garaci FG, Toschi N, Hampel H. Imaging epigenetics in Alzheimer's disease. *Curr Pharm Des* 2013;19:6393-415

22 Lista S, O'Bryant SE, Blennow K, et al. Biomarkers in sporadic and familial Alzheimer's Disease. *J Alzheimers Dis* 2015;47:291-317.

23 Rosen C, Hansson O, Blennow K, Zetterberg H. Fluid biomarkers in Alzheimer's disease - current concepts. *Mol Neurodegener* 2013;8:20.

24 Blennow K, Zetterberg H, Fagan AM. Fluid biomarkers in Alzheimer disease. *Cold Spring Harb Perspect Med* 2012;2:a006221.

25 Blennow K, Hampel H, Weiner M, Zetterberg H. Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nat Rev Neurol* 2010;6:131-44.

26 Hampel H, Shen Y, Walsh DM, et al. Biological markers of amyloid beta-related mechanisms in Alzheimer's disease. *Exp Neurol* 2010;223:334-46.

27 Hampel H, Blennow K, Shaw LM, Hoessler YC, Zetterberg H, Trojanowski JQ. Total and phosphorylated tau protein as biological markers of Alzheimer's disease. *Exp Gerontol* 2010;45:30-40.

28 O'Bryant SE, Lista S, Rissman RA, et al. Comparing biological markers of Alzheimer's disease across blood fraction and platforms: Comparing apples to oranges. *Alzheimers Dement (Amst)* 2016;3:27-34.

29 O'Bryant SE, Gupta V, Henriksen K, et al. Guidelines for the standardization of preanalytic variables for blood-based biomarker studies in Alzheimer's disease research. *Alzheimers Dement* 2015;11:549-60.

30 Henriksen K, O'Bryant SE, Hampel H, et al. The future of blood-based biomarkers for Alzheimer's disease. *Alzheimers Dement* 2014;10:115-31.

31 Snyder HM, Carrillo MC, Grodstein F, et al. Developing novel blood-based biomarkers for Alzheimer's disease. *Alzheimers Dement* 2014;10:109-14.

32 Gupta VB, Sundaram R, Martins RN. Multiplex biomarkers in blood. *Alzheimers Res Ther* 2013;5:31.

33 Teipel SJ, Grothe M, Lista S, Toschi N, Garaci FG, Hampel H. Relevance of magnetic resonance imaging for early detection and diagnosis of Alzheimer disease. *Med Clin North Am* 2013;97:399-424.

34 Ewers M, Sperling RA, Klunk WE, Weiner MW, Hampel H. Neuroimaging markers for the prediction and early diagnosis of Alzheimer's disease dementia. *Trends Neurosci* 2011;34:430-42.

35 Lista S, Garaci FG, Ewers M, et al. CSF Abeta1-42 combined with neuroimaging biomarkers in the early detection, diagnosis and prediction of Alzheimer's disease. *Alzheimers Dement* 2014;10:381-92.

36 Poil SS, de Haan W, van der Flier WM, Mansvelder HD, Scheltens P, Linkenkaer-Hansen K. Integrative EEG biomarkers predict progression to Alzheimer's disease at the MCI stage. *Front Aging Neurosci* 2013;5:58.

37 Stam CJ. Use of magnetoencephalography (MEG) to study functional brain networks in neurodegenerative disorders. *J Neurol Sci* 2010;289:128-34.

38 Jørgensen JT. Companion diagnostics: the key to personalized medicine. Foreword. *Expert Rev Mol Diagn* 2015;15:153-6.

39 Lyman GH, Moses HL. Biomarker Tests for Molecularly Targeted Therapies--The Key to Unlocking Precision Medicine. *N Engl J Med* 2016;375:4-6.

40 Doody RS, Raman R, Farlow M, et al. A phase 3 trial of semagacestat for treatment of Alzheimer's disease. *N Engl J Med* 2013;369:341-50

41 Salloway S, Sperling R, Fox NC, et al. Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N Engl J Med* 2014;370:322-33

42 Doody RS, Thomas RG, Farlow M, et al. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. *N Engl J Med* 2014;370:311-21

43 Noorbakhsh F, Overall CM, Power C. Deciphering complex mechanisms in neurodegenerative diseases: the advent of systems biology. *Trends Neurosci* 2009;32:88-100

44 Castrillo JI, Oliver SG. Alzheimer's as a Systems-Level Disease Involving the Interplay of Multiple Cellular Networks. *Methods Mol Biol* 2016;1303:3-48

45 FDA-NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and other Tools) Resource*. Silver Spring (MD): Food and Drug Administration (US), Bethesda (MD), National Institutes of Health (US), 2016-

46 Daly AK. Genome-wide association studies in pharmacogenomics. *Nat Rev Genet* 2010;11:241-6

47 Zhang J, Chiodini R, Badr A, Zhang G. The impact of next-generation sequencing on genomics. *J Genet Genomics* 2011;38:95-109

48 Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* 2013;45:1452-8

49 Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014;511:421-7

50 Daly AK. Pharmacogenetics and human genetic polymorphisms. *Biochem J* 2010;429:435-49

51 Guerreiro R, Wojtas A, Bras J, et al. TREM2 variants in Alzheimer's disease. *N Engl J Med* 2013;368:117-27

52 Brinkmalm A, Portelius E, Öhrfelt A, et al. Explorative and targeted neuroproteomics in Alzheimer's disease. *Biochim Biophys Acta* 2015;854:769-78

53 Veenstra TD. Neuroproteomic tools for battling Alzheimer's disease. *Proteomics* 2016. doi: 10.1002/pmic.201600211.

54 Proitsi P, Kim M, Association of blood lipids with Alzheimer's disease: A comprehensive lipidomics analysis. *Alzheimers Dement* 2016. doi: 10.1016/j.jalz.2016.08.003

55 Trushina E, Mielke MM. Recent advances in the application of metabolomics to Alzheimer's Disease. *Biochim Biophys Acta* 2014;1842:1232-9.

56 Czech C, Berndt P, Busch K, et al. Metabolite profiling of Alzheimer's disease cerebrospinal fluid. *PLoS One* 2012;7:e31501.

57 Weston AD, Hood L. Systems biology, proteomics, and the future of health care: toward predictive, preventative, and personalized medicine. *J Proteome Res* 2004;3:179-96

58 Butcher EC, Berg EL, Kunkel EJ. Systems biology in drug discovery. *Nat Biotechnol* 2004;22:1253-9

59 O'Bryant SE, Mielke MM, Rissman RA, et al. Blood based biomarkers in Alzheimer's disease: Current state of the science and a novel collaborative paradigm for advancing discovery to clinic. *Alzheimers Dement* 2016. In press

60 Hood L, Flores M. A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. *N Biotechnol* 2012;29:613-24

61 Gligorijević V, Malod-Dognin N, Pržulj N. Integrative methods for analysing big data in precision medicine. *Proteomics* 2016;16:741-58

62 Beyer MA, Laney D. *The Importance of 'Big Data': A Definition*. Stamford, CT: Gartner, 2012

63 Margolis R, Derr L, Dunn M, et al. The National Institutes of Health's Big Data to Knowledge (BD2K) initiative: capitalizing on biomedical big data. *J Am Med Inform Assoc* 2014;21:957-8

64 Geerts H, Dacks PA, Devanarayan V, et al. Big data to smart data in Alzheimer's disease: The brain health modeling initiative to foster actionable knowledge. *Alzheimers Dement* 2016;12:1014-21

65 Haas M, Stephenson D, Romero K, Gordon MF, Zach N, Geerts H; Brain Health Modeling Initiative (BHMI). Big data to smart data in Alzheimer's disease: Real-world examples of advanced modeling and simulation. *Alzheimers Dement* 2016;12:1022-30

66 Saykin AJ, Shen L, Yao X, et al. Genetic studies of quantitative MCI and AD phenotypes in ADNI: Progress, opportunities, and plans. *Alzheimers Dement* 2015;11:792-814

67 Choi IY, Kim TM, Kim MS, Mun SK, Chung YJ. Perspectives on clinical informatics: integrating large-scale clinical, genomic, and health information for clinical care. *Genomics Inform* 2013;11:186-90

68 Bai JP, Abernethy DR. Systems pharmacology to predict drug toxicity: integration across levels of biological organization. *Annu Rev Pharmacol Toxicol* 2013;53:451-73

69 Maudsley S,Martin B, Janssens J, et al. Informatic deconvolution of biased GPCR signaling mechanisms from in vivo pharmacological experimentation. *Methods* 2016;92:51-63

70 Peck RW. The right dose for every patient: a key step for precision medicine. Nat Rev *Drug Discov* 2016;15:145-6

71 Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology, The Gene Ontology Consortium. *Nat Genet* 2000;25:25-9

72 Malhotra A, Younesi E, Gündel M, Müller B, Heneka MT, Hofmann-Apitius M. ADO: A disease ontology representing the domain knowledge specific to Alzheimer's disease. Alzheimers Dement. 2014;10:238-46

73 Younesi E, Malhotra A, Gündel M, et al. PDON: Parkinson's disease ontology for representation and modeling of the Parkinson's disease knowledge domain. *Theor Biol Med Model* 2015;12:20

74 Sahoo SS, Lhatoo SD, Gupta DK, et al. Epilepsy and seizure ontology: towards an epilepsy informatics infrastructure for clinical research and patient care. *J Am Med Inform Assoc* 2014;21:82-9

75 Malhotra A, Gündel M, Rajput AM, et al. Knowledge retrieval from PubMed abstracts and electronic medical records with the Multiple Sclerosis Ontology. *PLoS One* 2015;10:e0116718

76 Iyappan A, Gündel M, Shahid M, et al. Towards a Pathway Inventory of the Human Brain for Modeling Disease Mechanisms Underlying Neurodegeneration. *J Alzheimers Dis* 2016;52:1343-60

77 Mizuno S, Iijima R, Ogishima S, et al. AlzPathway: a comprehensive map of signaling pathways of Alzheimer's disease. *BMC Syst Biol* 2012;6:52

78 Kodamullil AT, Younesi E, Naz M, Bagewadi S, Hofmann-Apitius M. Computable cause-and-effect models of healthy and Alzheimer's disease states and their mechanistic differential analysis. *Alzheimers Dement* 2015;11:1329-39

79 Martin F, Sewer A, Talikka M, Xiang Y, Hoeng J, Peitsch MC. Quantification of biological network perturbations for mechanistic insight and diagnostics using two-layer causal models. *BMC Bioinformatics* 2014;15:238

80 Benzinger TL, Blazey T, Jack CR Jr, et al. Regional variability of imaging biomarkers in autosomal dominant Alzheimer's disease. *Proc Natl Acad Sci USA* 2013;110:E4502-9

81 Iturria-Medina Y, Sotero RC, Toussaint PJ, Mateos-Pérez JM, Evans AC; Alzheimer's Disease Neuroimaging Initiative. Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nat Commun* 2016;7:11934

82 Donohue MC, Jacqmin-Gadda H, Le Goff M, et al. Estimating long-term multivariate progression from short-term data. *Alzheimers Dement* 2014;10:S400-10

83 Samtani MN, Raghavan N, Shi Y, et al. Disease progression model in subjects with mild cognitive impairment from the Alzheimer's disease neuroimaging initiative: CSF biomarkers predict population subtypes. *Br J Clin Pharmacol* 2013;75:146-61

84 Delor I, Charoin JE, Gieschke R, Retout S, Jacqmin P. Modeling Alzheimer's Disease Progression Using Disease Onset Time and Disease Trajectory Concepts Applied to CDR-SOB Scores From ADNI. *CPT Pharmacometrics Syst Pharmacol* 2013;2:e78

85 Gaser C, Franke K, Klöppel S, Koutsouleris N, Sauer H; Alzheimer's Disease Neuroimaging Initiative. BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease. *PLoS One* 2013;8:e67346

86 Lavielle M. *Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools*. Series: Chapman & Hall/CRC Biostatistics Series, Boca Raton (FL): CRC Press, Taylor & Francis Group (US), 2015

87 Durrleman S, Pennec X, Trouvé A, Gerig G, Ayache N. Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets. *Med Image Comput Comput Assist Interv* 2009;12:297-304

88 Durrleman S, Pennec X, Trouvé A, Braga J, Gerig G, Ayache N. Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data. *Int J Comput Vis* 2013;103:22-59

89 Fonteijn HM, Modat M, Clarkson MJ, et al. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. *Neuroimage* 2012;60:1880-9

90 Jedynak BM, Lang A, Liu B, et al. A computational neurodegenerative disease progression score: method and results with the Alzheimer's disease Neuroimaging Initiative cohort. *Neuroimage* 2012;63:1478-86

91 Schiratti JB, Allassonniere S, Colliot O, Durrleman S. Learning spatiotemporal trajectories from manifold-valued longitudinal data. Presented at the *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, 2015: 2404-12

92 Schiratti JB, Allassonniere S, Routier A, Colliot O, Durrleman S. A mixed-effects model with time reparametrization for longitudinal univariate manifold-valued data. Presented at the 24th International Conference, *IPMI 2015 - Information Processing in Medical Imaging*, Sabhal Mòr Ostaig, Isle of Skye, UK, Jun 2015:564-75

93 Hurd MD, Martorell P, Delavande A, Mullen KJ, Langa KM. Monetary costs of dementia in the United States. *N Engl J Med* 2013;368:1326-34

94 Dubois B, Hampel H, Feldman HH, et al. Preclinical Alzheimer’s disease: Definition, natural history, and diagnostic criteria. *Alzheimers Dement* 2016;12:292-323

FIGURE CAPTIONS

Figure 1. Translational bench to bedside data flow within the conceptual framework of the Alzheimer Precision Medicine Initiative (APMI).

The IDM Data Sciences lifecycle takes advantage of both data- and knowledge-driven approaches so that quantitative (biomolecular, neuroimaging, neurophysiological and clinical data) and qualitative (literature data) are first represented in a harmonized, standardized format to be prepared for proper management within an integrative computational infrastructure. Once these heterogeneous datasets and scales are leveraged, the next step is to integrate them using modeling algorithms that allow for further analysis such as predictive operations (reasoning, simulation, and visualization). The output should be an "actionable" model that predicts the trajectory of individual patient-centric detection or treatment within the P4M implementation.

Abbreviations: APMI cohort or APMI-CP, Alzheimer Precision Medicine Initiative cohort or cohort program; IDM, Integrative Disease Modeling; P4M, P4 (Predictive, Preventive, Personalized, and Participatory) medicine.

Figure 2. Digital model of cognitive decline in AD, built from the ADAS-Cog scores (pooled into 4 categories of cognitive functions) observed repeatedly in 248 subjects with an average of 5 follow-up visits, who progressed from prodromal stage MCI to AD dementia during the observation period (data derived from the ADNI). The model consists of an average disease progression model showing the typical scenario of cognitive decline from prodromal symptomatic (MCI) to syndromal symptomatic (dementia) stages of AD (panel a), and a set of parameters showing the variability of this scenario in terms of age at disease onset

(panel b), pace of cognitive decline (panel c), temporal ordering and relative delay between declines of different cognitive functions (panels d e and e) within the studied population.

Abbreviations: AD, Alzheimer’s disease; ADAS-Cog, Alzheimer’s Disease Assessment Scale-cognitive subscale; ADNI, Alzheimer's Disease Neuroimaging Initiative; MCI, mild cognitive impairment.

Table 1: Evolving lexicon and terminology of Precision Medicine

Concept	Abbreviation	Definition
Big Data		A repository of large amounts of data sets generated by data mining tools. Big Data includes information obtained through systems theory- and, knowledge-based approaches and clinical records
Biomarkers	BMs	A defined characteristic that is measured as an indicator of normal biological processes, pathogenic process, or response to an exposure or intervention, including therapeutic interventions. Molecular, histologic, radiographic, or physiological characteristics are types of biomarkers. A biomarker is not an assessment of how an individual feels, functions or survives. Categories of biomarkers include: susceptibility/risk biomarker, diagnostic biomarker, monitoring biomarker, prognostic biomarker, predictive biomarker, pharmacodynamics/response biomarker and safety biomarker[45]
Data Science		Interdisciplinary field about processes and systems to extract knowledge from data in different forms, either structured or unstructured, which is a continuation of some of the data analysis fields including statistics, artificial intelligence, machine learning, data mining, and predictive analytics
Electronic Health Records	EHRs	Systematized gathering of population electronically-stored health information and clinical data in a digital format. These registries can be shared across different health care settings through network systems. EHRs eliminate the need to track down a patient's previous paper medical records and assist in ensuring that data are accurate and legible
Genomic Medicine		Discipline utilizing personal genomic information for diagnostic characterization and the development of therapeutic plans
Integrative Disease Modeling	IDM	Multidisciplinary approach to standardize, manage, integrate, and interpret multiple sources of structured and unstructured quantitative and qualitative data across biological scales using computational models that assist decision making for translation of patient-specific molecular mechanisms into tailored clinical applications
Mobile Health	mHealth	Smart personal mobile devices (phones, wearables, in-home devices and Apps) collecting health information aimed at improving patient care
Omic Disciplines		High-throughput screening tools aimed at fully collecting, characterizing and quantifying pools of biological molecules

		(DNA sequences, transcripts, proteins, metabolites/lipids) that translate into the structure, function, and dynamics of an organism and/or organisms
Ontology		Formal naming and designation of the types, properties, and interactions of the entities that really or fundamentally exist for a specific domain of discourse
P4 (Predictive, Preventive, Personalized, and Participatory) Medicine	P4M	Translational medicine component of the Precision Medicine paradigm. It is a clinical practice model aimed at applying knowledge, tools, and strategies of systems medicine. It involves generation, mining, and integration of enormous amounts of data on individual patients to produce predictive and "actionable" models of wellness and disease
Personalized Medicine		Component of P4 Medicine aiming at tailoring treatment for individual patients in contrast with "one-size fits-all" or traditional "magic bullet" drug approach
Precision Medicine	PM	Translational science paradigm related to both health and disease. PM is a biomarker-guided medicine on systems-levels taking into account methodological advancements and discoveries of the comprehensive pathophysiological profiles of complex polygenic, multi-factorial neurodegenerative diseases (proteinopathies of the brain). It aims at optimizing the effectiveness of disease prevention and therapy, by considering (customized) an individual's specific biological makeup (e.g. genetic, biochemical, phenotypic, lifestyle, and psychosocial characteristics) for targeted interventions through P4M implementation
Systems Biology	SB	Evolving hypothesis-free, exploratory, holistic (non-reductionistic), global, integrative, and interdisciplinary paradigm using advances in multimodal high-throughput technological platforms that enable the examination of networks of biological pathways where elevated amounts of structurally and functionally different molecules are simultaneously explored over time at a system level (i.e., at the level of cells, group of cells, tissues, organs, apparatuses, or even whole organisms)
Systems Medicine		Holistic paradigm applying systems biology-based strategies to medical research. It aims at integrating a variety of considerable biomedical data at all levels of the cellular organization (by employing global, integrative, and statistical/mathematical/computational modelling) to explicate the pathophysiological mechanisms, prognosis, diagnosis, and treatment of diseases
Systems Theory	ST	Translational research theory of the Precision Medicine paradigm. It is an interdisciplinary conceptual framework allowing for the conceptualization of novel/original models to

	extract and explicate all systems levels and different spatiotemporal data types of complex polygenic diseases
Taxonomy	Scientific classification into groups based on shared characteristics and natural relationships. Taxonomy adds a relation dimension between individual items and is defined as a way to group similar items together