Alzheimer's & Dementia 12 (2016) 1014-1021

Review Article

Big data to smart data in Alzheimer's disease: The brain health modeling initiative to foster actionable knowledge

Hugo Geerts[a,*], Penny A. Dacks[b], Viswanath Devanarayan[c], Magali Haas[d], Zaven S. Khachaturian[e], Mark Forrest Gordon[f], Stuart Maudsley[g], Klaus Romero[h], Diane Stephenson[h], on behalf of the Brain Health Modeling Initiative (BHMI)

[a]*In Silico Biosciences, Inc., Berwyn, PA, USA*
[b]*Alzheimer's Drug Discovery Foundation, New York, NY, USA*
[c]*Exploratory Statistics, AbbVie, Souderton, PA, USA*
[d]*Orion Bionetworks, Inc., Cambridge MA, USA*
[e]*PAD2020, Potomac, MD, USA*
[f]*Boehringer Ingelheim Pharmaceuticals, Ridgefield, CT, USA*
[g]*VIB Department of Molecular Genetics, Institute Born-Bunge, University of Antwerp, Antwerp, Belgium*
[h]*C-Path Institute, Tucson, AZ, USA*

Abstract Massive investment and technological advances in the collection of extensive and longitudinal information on thousands of Alzheimer patients results in large amounts of data. These "big-data" databases can potentially advance CNS research and drug development. However, although necessary, they are not sufficient, and we posit that they must be matched with analytical methods that go beyond retrospective data-driven associations with various clinical phenotypes. Although these empirically derived associations can generate novel and useful hypotheses, they need to be organically integrated in a quantitative understanding of the pathology that can be actionable for drug discovery and development. We argue that mechanism-based modeling and simulation approaches, where existing domain knowledge is formally integrated using complexity science and quantitative systems pharmacology can be combined with data-driven analytics to generate predictive actionable knowledge for drug discovery programs, target validation, and optimization of clinical development.

Keywords: Brain disorders; Alzheimer's dementia; Complexity theory; Systems pharmacology; Systems biology; Drug discovery and development

1. Introduction

Among all the enduring disabling conditions of our increasingly aged society, chronic brain disorders, such as dementia/Alzheimer's disease (AD), are the leading contributors to spiraling health care costs as well as individual and caregiver burden. Leading health organizations across the world have estimated that brain disorders (neurological, psychiatric, brain injury, or pain) will affect one in five individuals in their lifetime with an associated cost of more than 2 trillion US dollars annually in the United States and Europe alone [1]. This exceeds the annual combined burden of cardiovascular disease, cancer, and diabetes, and it is expected to rise with increasing life expectancy. Among these brain disorders, dementia represents one of the largest burdens to our aging societies [2], afflicting more than 35 million people worldwide [3]. Today, there are no effective therapies for these conditions, despite enormous financial and research investments. This reality has galvanized a global effort launched by the G8 Summit on Dementia in 2014; however, we suggest that the large

*Corresponding author. Tel.: +1-267-679-8090; Fax: +1-484-802-3478.
E-mail address: Hugo-Geerts@In-Silico-Biosciences.com

http://dx.doi.org/10.1016/j.jalz.2016.04.008

investment in research and development can substantially benefit from integrative predictive modeling.

From 2002 to 2012, 99.6% of the clinical trials of disease-modifying treatments for AD have failed [4]. In stark contrast, between 1995 and 2010, approximately 300 interventions were reported to reduce pathology and/or improve behavior in transgenic AD mouse models [5]. Although changes in preclinical animal research design might somewhat improve the predictive value of these animal models [6], rodents will continue to be fundamentally different from humans [7].

AD is inherently a multifactorial syndrome, and individual patients present with a wide variety of pathologies, as a consequence of comorbidities, life history, and genotypes (Fig 1). In fact, neuropathologic evidence suggests that different pathologies converge in the brains of elderly people with dementia [8,9]. This suggests that the biological processes driving the clinical phenotype can differ markedly from patient to patient. In addition, up to one-third of nondemented, high-functioning seniors may harbor underlying pathology to an extent that would be expected to cause dementia. So far reductionist molecular biological approaches have failed to explain this phenomenon [9].

The complexity of clinical trials for AD has also contributed to the therapeutic failure rate. The clinical outcome metrics related to cognition and function are highly variable, not only due to the inherent variability in the pathological processes (see above), but also the impact of co-medications and genotypes both within and across patient groups, necessitating large sample size and treatment duration to detect remediation. New modeling efforts such as the precompetitive consortium, the Coalition Against Major Diseases, can help develop tools to optimize the efficiency of clinical trial design [10]. Biomarkers can quantify neuropathology and its progression but the use of single molecular biomarkers in isolation has unfortunately not successfully predicted the functional and cognitive outcomes relevant to patients.

2. From reductionism to integration

The prevailing paradigm for scientific inquiry in the neuroscience field has used classical reductionism, an approach wherein explanation of entire systems is predicated in terms of their individual, constituent parts, and their interactions. This molecular biology approach, often based on data-driven correlation analysis, is basically a bottom-up strategy, where the resulting outcome is defined usually as a consequence of a single set of linear assumptions. This often negates the many nonlinear interactions between subsystems and the appearance of emergent properties that cannot be reduced to a single target.

The case of beta-amyloid modulation as a therapeutic approach for AD illustrates the problems associated with a statistical approach that correlates a clinical phenotype with genetic information. The most optimistic perspective on the failure of this approach is that these trials have been conducted too late in the course of the disease, a failure in the trial design rather than the targets, and that the solution is to conduct trials in prodromal conditions [11]. However, the assumption that "reducing beta-amyloid load" leads to cognitive improvement is probably a major oversimplification of the complex biology of beta-amyloid in the human AD brain that we are gradually starting to understand. Recent studies indeed document different aggregation dynamics [12], different formation and clearance in the human brain [13], different neuroprotective versus neurotoxic properties of the shorter versus longer amyloid peptides [14], and the complex nonlinear interaction of co-medications and genotypes on clinical cognitive readouts [15]. In other words, even if beta-amyloid is the correct therapeutic target, successful drug development will likely require a more sophisticated understanding of its complex dynamics. In addition, nonamyloid processes such as tau pathology, neuroinflammation, and oxidative stress interact with beta-amyloid physiology resulting most likely in an idiosyncratic cognitive trajectory for each AD patient.

With the development of systems biology, the concept of circuit and network insights was combined with multivariate analyses, resulting in an integrative approach that starts with the patient. It "is about putting together rather than taking apart, integration rather than reduction. It requires that we develop ways of thinking about integration that are as rigorous as our reductionist programs, but different…It means changing our philosophy, in the full sense of the term" [16]. Quantitative systems pharmacology (QSP) goes one step further by adding formalized domain expertise about the biological nature such as enzyme kinetics and interaction between drugs and targets of the different parts of key circuits and pathways. In this way, causation is explicitly integrated in the modeling. In the case of central nervous system (CNS) disorders, QSP also integrates this information into biophysically realistic neuronal networks, the firing properties of which can be associated with a clinical phenotype [17].

3. From big data to smart data

As part of the new approaches to reduce clinical trial failure rates, global efforts are now shifting toward a focus on gathering "big data" [18]. The integration of large clinical data sets is viewed as a potentially powerful approach to expedite medical discovery, and there is justifiable enthusiasm based on results of global studies of disease progression and large-scale genomics efforts [19]. Advanced deep analytical approaches have been developed and are covered by other publications in the field of bioinformatics [20,21] and pharmacology [22]; however, specific case studies for brain disease are limited and these publications are typically written for a very narrow specialty audience.

These large-scale data collection efforts, such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) consortium, and the European initiatives (http://www.alzheimers.org.uk/site/scripts/download_info.php? fileID=2273) will yield