Alzheimer's & Dementia 12 (2016) 1022-1030

Review Article

Big data to smart data in Alzheimer's disease: Real-world examples of advanced modeling and simulation

Magali Haas[a,*], Diane Stephenson[b], Klaus Romero[b], Mark Forrest Gordon[c], Neta Zach[d], Hugo Geerts[e], on behalf of the Brain Health Modeling Initiative (BHMI)

[a]*Orion Bionetworks, Inc., Cambridge, MA, USA*
[b]*C-Path Institute, Tucson, AZ, USA*
[c]*Boehringer Ingelheim Pharmaceuticals, Ridgefield, CT, USA*
[d]*Prize4Life, Tel Aviv, Israel*
[e]*In Silico Biosciences, Berwyn, PA, USA*

Abstract Many disease-modifying clinical development programs in Alzheimer's disease (AD) have failed to date, and development of new and advanced preclinical models that generate actionable knowledge is desperately needed. This review reports on computer-based modeling and simulation approach as a powerful tool in AD research. Statistical data-analysis techniques can identify associations between certain data and phenotypes, such as diagnosis or disease progression. Other approaches integrate domain expertise in a formalized mathematical way to understand how specific components of pathology integrate into complex brain networks. Private-public partnerships focused on data sharing, causal inference and pathway-based analysis, crowdsourcing, and mechanism-based quantitative systems modeling represent successful real-world modeling examples with substantial impact on CNS diseases. Similar to other disease indications, successful real-world examples of advanced simulation can generate actionable support of drug discovery and development in AD, illustrating the value that can be generated for different stakeholders.

Keywords: Brain disorders; Alzheimer's dementia; Crowdsourcing; Private-public partnership; Quantitative systems pharmacology; Regulatory

1. Introduction

A major problem with Alzheimer's disease (AD) drug development is the lack of reliable models to predict clinical efficacy during drug development. In stark contrast to the 99.6% failure rate of clinical trials, roughly 300 interventions were reported to reduce pathology and/or improve behavior in transgenic AD mouse models between 1995 and 2010 [1].

This illustrates the need for an alternative more humanized set of models that is better aligned with the clinical situation. Modeling and simulation might be a powerful alternative, especially when based on human clinical data and with appropriate domain expertise. Existing "big data analytics" cover a vast computational space ranging from bottom-up dynamical systems modeling to top-down probabilistic causal approaches. A variety of methodologic frameworks have been developed for modeling and analyzing complex multivariate data that can be adapted to neuroscience in general and AD in particular.

Despite extensive investment to identify biomarkers for central nervous system (CNS) disorders, a large gap remains between exploratory biomarkers and their validation and integration into routine clinical practice. This gap exists in part because conventional analysis methods focus on single biomarker analysis, whereas most disease phenotypes arise not from single genes and proteins but from a complex

*Corresponding author. Tel.: +1-609-218-2908.
E-mail address: magali.haas@cohenbio.org

http://dx.doi.org/10.1016/j.jalz.2016.05.005

network of molecular interactions. Therefore, an alternative approach for identifying disease markers adopts network biology and pathway databases to infer differences between populations of patients and the role of genes in disease in the context of networks and pathways. The activity of these entire networks and pathways can then be used as part of the statistical analysis procedure—rather than, for example, the expression levels of single genes—to characterize or stratify patients into similar groups. The identification of these pathways would also enable the identification of new targets for disease intervention (either preventatively or acutely).

Systems biology is an integrative approach that combines theoretical modeling and direct experimentation. Theoretical, or computational, models provide insights into experimental observations, and experiments can provide data needed for model creation or can confirm or refute model findings. Many of our scientific breakthroughs and technologic advances in other engineering industries rely on models simulated on high-performance computers.

This revolution is now beginning to touch the fields of neuroscience and brain research.

This article documents a few real-world examples of such novel frameworks that can support these efforts in various CNS indications. The report describes public-private partnerships (PPPs) with Clinical Path Institute (C-Path) as an example, The Orion Bionetwork for Multiple Sclerosis, a few examples of quantitative systems pharmacology (QSP) for CNS, a crowdsourcing example of modeling in amyotrophic lateral sclerosis (ALS), and examples where regulatory agencies used modeling and simulations. Although these approaches are certainly not perfect, they illustrate what can be achieved in practical and actionable terms for addressing well-defined questions and problems. Importantly, they also identify possible pathways toward concrete solutions for various stakeholders, including patient foundations, academics, regulatory agencies, and pharmaceutical companies.

2. Public-private partnerships

There is a growing expansion in the number of AD PPPs given the recognition that collaboration is essential [2].

The flagship PPPs, Alzheimer's Disease Neuroimaging Initiative (ADNI) and Parkinson's Progressive Markers Initiative (PPMI), are revolutionizing the understanding of brain diseases, such as AD and Parkinson's disease (PD), by demonstrating that pathologic hallmarks originate in the brain decades before the onset of symptoms. Despite these significant advances in understanding the longitudinal progression of disease, disappointing results in clinical trials that have evaluated biomarkers and potential disease modification continue to pose challenges. Two critical success factors underlie why the ADNI and PPMI collaborations are so impactful: (1) agreement on open sharing of patient-level data to the broad research community and (2) agreement to comply with consensus data standards for collection of all data including patient diagnostic criteria and biomarker standardized acquisition parameters. Although numerous publications exist on methodologies for biomarker assay standardization, there is no current mechanism in place for sponsors conducting clinical trials to agree and conform to consensus standards for data collection or open sharing of patient data.

3. The C-Path experience

C-Path was formed in 2004 to advance the goals set forth in the Critical Path Initiative of the U.S. Food and Drug Administration (FDA) [3]. One of C-Path's many precompetitive consortia, the Coalition Against Major Diseases (CAMD), was launched in September 2008 to develop new technologies and methods to accelerate drug development for AD and PD. CAMD convenes pharmaceutical industry, research and patient advocacy organizations, regulatory and other government agencies, and academia to achieve formal regulatory decisions endorsing outcome measures, modeling and simulation, and biomarkers by creating consensus data standards and enabling precompetitive data sharing platforms [4].

Through partnership with Clinical Data Interchange Standards Consortium (CDISC), C-Path has successfully developed data standards for AD, PD, polycystic kidney disease, tuberculosis, multiple sclerosis (MS) and most recently, schizophrenia and traumatic train injury. The AD CDISC standards represented the first such disease-specific standards and a version 2.0 of the AD CDISC standards, completed in early 2014, incorporates biomarkers and early stages of the AD spectrum, specifically targeted at the stage of mild cognitive impairment. These therapeutic area standards were developed with funding support from the FDA and represent the preferred format by regulatory agencies for new drug applications for expedited review. By 2017, submission of new drug applications in CDISC standards will likely be required by the FDA, suggesting that clinical trials initiating at the present time should adopt these standards. Using these standards, CAMD remapped control-arm patient-level data from AD clinical trials to populate the CAMD AD database. This standardization facilitates analyses of the data as a single-integrated source. Pooling data in this fashion allow analysts to query all trials or subsets of trials in the database without writing programming statements for each new study. The database is available to qualified researchers and currently consists of placebo clinical data from 24 trials of about 6500 subjects in total [5]. The CAMD AD database was fundamental to developing a drug-disease-trial model for mild-to-moderate AD, which became the first regulatory-endorsed quantitative clinical trial simulation tool. This model describes the longitudinal progression of the 11-item AD Assessment Scale Cognitive sub-scale (ADAS-Cog) in AD patients from both natural history and randomized clinical trials, placebo effect, patient dropouts, and different types of treatment effects [6]. With

endorsement from both the FDA and European Medicines Agency, this tool is being adopted to optimize clinical trial design for candidate therapeutics in AD [7].

4. Orion Bionetworks: A case study for computational modeling of multiple sclerosis

MS is a leading cause of disability in young adults, affecting >2.5 million individuals worldwide [8]. The pathogenic mechanisms that lead to the loss of immune homeostasis, myelin and axonal injury, and progressive neurological symptoms are incompletely understood [9,10]. The clinical course of MS varies greatly, reflecting the considerable complexity of pathogenesis and disease expression. The current generation of disease-modifying therapies generally treats the immune-modulatory components of disease and is proven effective only for patients with the relapse-remitting form of MS (RRMS). Diagnostic, prognostic, and response biomarkers would aid in clinical development of next-generation therapeutics as well as delivery of precision clinical care [11].

Orion Bionetworks, Inc., a Cambridge, MA–based nonprofit research organization, was founded to bridge the translational divide for brain disorders through the adoption of a systems biology and modeling approach enabled through a unique public-private alliance partnership model.

The Orion Bionetworks Flagship Program (launched in 2013) has successfully piloted the establishment of an MS Bionetwork comprised of academic (Neuroscience Institute of the Brigham and Women's Hospital), advocacy (Accelerated Cure Project for MS), computational (GNS Healthcare, MetaCell, Thomson Reuters), informatics (Converge by Deloitte, Rancho Biosciences, Exaptive), and online patient community partners (PatientsLikeMe [PLM]) with funding sponsorship provided by Janssen Pharmaceuticals and public philanthropic donations. Deidentified data from three databases were curated and loaded into a cloud-based data knowledge management system called TranSMART. The integrated repository includes >9000 subjects with MS and related conditions.

The alliance developed a roadmap for joint execution with specific scientific aims. Using a computational systems modeling approach and our integrated database, we sought to answer the following research questions:

- What is the natural course of MS disease progression based on patient-derived and clinically derived data sources?
- Does the current nosology of clinically isolated syndrome/RRMS/primary progressive MS/secondary progressive MS accurately reflect this pattern?
- Are there biomarkers of prognosis?
- Could we dissect the structure of the MS patient population based on molecular pathways or other variables?
- How many patient subsets are there?
- How does this structure relate to disease course?
- How does this structure relate to treatment response?
- What are the most common comorbidities associated with MS based on patient-reported experience, and could we use the clinically derived data to model pathogenic mechanisms?
- What constitutes a "relapse" event and what are the triggers?
- Could we build a systems model to understand the pathophysiology of neurodegeneration in MS?

To date, the alliance has generated three models (discussed below) based on the data repository using different algorithmic approaches.

4.1. Phenotypic prognostic model

To develop our phenotypic prognostic model, data from a database amassed by PLM on >35,000 patients with MS was used. The database comprised two major sets of patient self-reported outcomes: (1) Multiple Sclerosis Rating Scale (MSRS) and (2) General Symptoms Scale as well as key demographic variables on each patient (e.g., age, gender, diagnosis, location). The PLM online platform uses a structured data collection process that largely mirrors data collected from a clinical research environment [12]. Key variables collected in the MSRS included walking, upper limb function, vision, speech, swallowing, thinking/memory/cognition, and sensation/burning/pain. Key outcomes collected in the General Symptoms Scale included anxious mood, depressed mood, fatigue, insomnia, pain, bladder problems, bowel problems, brain fog, emotional lability, somnolence, mood swings, sexual dysfunction, and stiffness/spasticity. Summary statistics were generated for all key outcomes (e.g., symptom severity over time). We also generated a linear stepwise-regression model to project the MS course for an individual patient based on their history of symptoms. Complete results have been submitted for publication.

4.2. Molecular prognostic models

To develop our molecular prognostic models, data from a clinical study database called CLIMB [13] were used to build two separate models. GNS Healthcare developed a data-driven model using Bayesian causal inference modeling [14], whereas Thomson Reuters used a pathway-based analysis approach that incorporated "priors" from the literature [15]. It has been shown in a number of studies that approaches based on utilization of prior knowledge perform better (e.g., accuracy, reproducibility, and robustness of the models) than purely data-driven approaches for such applications as biomarker identification, patient stratification [15–20], and clustering analysis for subgroup identification [21]. A network-based analysis approach was tested for MS based on genome-wide association studies (GWAS) data [22].

These molecular models were generated from the CLIMB database, which includes >2000 RRMS patient's longitudinal (up to 5 years) data including demographic, clinical, magnetic resonance imaging (MRI), and genetic (GWAS and

transcriptome) data. Both algorithmic approaches were used to model disease progression (time to relapse or reduced functioning on Kurtzkee Functional Scale) as a function of clinical and/or molecular variables. Complete results have been submitted for publication elsewhere.

The alliance quickly recognized the need to develop new informatics and visualization tools to work with these large data sets and models.

4.3. Bioinformatics infrastructure

Orion Bionetworks has funded the development and deployment of several open-source and proprietary data management and modeling and visualization tools. Orion has an established data management system called TranSMART that houses all the deidentified clinical and phenotypic data provided to the alliance [23]. This cloud-based infrastructure is secure and password-protected for use by the alliance partners. TranSMART is a knowledge management platform that enables scientists to develop and refine research hypotheses by investigating correlations between genetic and phenotypic data, and assessing their analytical results in the context of published literature and other work [24]. Its Web interface, backed by a powerful R engine, provides a standardized process for data analysis and visualization. With our alliance partner, Exaptive Inc., we have developed plug-in visualization tools that support hypothesis testing and simulations.

In summary, by using data sets of convenience, we were able to build two powerful *in silico* models that can be used to run simulations of different conditions to help us interrogate complex systems pathways leading to different disease outcomes. This modeling exercise allowed us to identify critical gaps to build more robust and predictive models for MS and other brain disorders.

5. Prediction of clinical outcomes with quantitative systems pharmacology in schizophrenia and AD

QSP is an advanced mechanism-based simulation platform of biophysically realistic humanized neuronal circuits, consisting of different types of neurons and using Hodgkin–Huxley equations to calculate the actual time-dependent membrane potential changes [25]. Voltage-gated ion channels can be modulated by various neurotransmitters, such as dopamine, serotonin, norepinephrine, and acetylcholine, and human pathology can be introduced using clinical imaging and postmortem data. The platform's architecture is based on the human brain neuroanatomy and neurophysiology and includes a number of cortico-striatal-thalamo-cortical loops with the appropriate different cell types. Localization of membrane receptors and ion channels and their intracellular regulation is derived from literature and reflects the domain expertise collected in the neuroscience community. Key biological coupling parameters between receptor activation and subsequent change in voltage-gated ion channels are calibrated using a correlation between model outcome and actual historical clinical outcomes for a large number of pharmacological interventions, such as ADAS-Cog changes over time with different doses of acetylcholinesterases and 5-H^6 antagonist SB742457 [26].

This section documents a number of cases where the QSP platform has been used successfully in CNS disorders.

The computer model blindly predicted, purely based on the preclinical pharmacology, that PF-04995274, a selective 5-HT^4 partial agonist with beneficial improvement in traditional preclinical animal models, would actually worsen cognitive outcomes in a phase I scopolamine study with human volunteers. This unexpected clinical prediction was indeed confirmed subsequently [27], and the clinical development project was halted, despite the fact that back-up compounds with different pharmacology showed a greater effect on the predictive cognitive outcome and could have had a higher chance of success in clinical trials for AD. Conversely, the QSP platform identified a different AD patient population likely to respond successfully to the selected clinical candidate [27].

In this case, the translational gap between the animal and computer models in predicting the clinical response was largely driven by the differences in serotonin dynamics and interaction of the candidate drug with the human targets versus the rodent receptors. Because of the *in silico* nature and the calibration with human imaging studies, the humanized platform captured the serotonin dynamics of the human brain better than rodent models.

The predictive power of QSP in schizophrenia has been demonstrated with the phase II predictions of JNJ37822681, a low-affinity selective D_2R antagonist and ocaperidone, a high-affinity multi-target D_2 antagonist [28]. In this study, the modelers were kept blinded to the actual clinical outcome and the QSP model correctly predicted a high–motor side-effect liability that was not observed in preclinical animal models and that led to the demise of the clinical development project. In contrast, the model was unable to predict the relative high placebo effect in the ocaperidone trial, but correctly and quantitatively predicted the relative clinical improvement (i.e. the difference) of ocaperidone compared with placebo.

A third example refers to the new 5-HT_{2C} agonist vabicaserin as a stand-alone in schizophrenia [29]. This new target was introduced in the neuronal circuitry based on preclinical animal neurophysiology, and the dose-dependent outcome of an agonist at this receptor as a stand-alone therapy in schizophrenia patients was simulated. Comparison with the actual phase II outcomes showed that the model approximately predicted the effect size of the drug on the total Positive and Negative Syndrome Scale score, which was much lower than reference antipsychotics and therefore was insufficient to warrant a full clinical development program. Note that the preclinical animal models that usually give a binary response (yes/no) resulted in an equipotent response compared to reference antipsychotics, suggesting that this target would be a valuable development project.

As an example of personalized medicine, the QSP platform was used to identify the biological rationale for responders to the antipsychotic iloperidone [30] in schizophrenia patients. In this study, the drug pharmacology was kept constant, whereas the biological coupling parameters were allowed to fluctuate around their optimal values; optimization of iloperidone response in this model identified the coupling between D_4R and AMPA on pyramidal cells as the major driver, in line with one reported genotype in a more traditional pharmacogenomics analysis [31]. This approach allows in principle to identify the pathways and circuits and therefore the single-nucleotide polymorphism candidates that would drive clinical response to a specific therapeutic intervention.

In another example with implications for polypharmacy treatment in real-life, the nonlinear pharmacodynamic interactions between acetylcholinesterases, memantine, and smoking in combination with specific antipsychotics was simulated in cognitive impairment in schizophrenia [32]. Although the study identified a few cases of positive synergy, most of the results suggested that pharmacodynamic drug-drug interactions had a negative impact on cognition. Interestingly, this study generated hypotheses as to why reported clinical results with these compounds were so variable, ranging from a negative effect to a positive synergy.

These examples demonstrate that integrating QSP approaches at an early stage in drug discovery can help pharmaceutical companies identify the optimal clinical candidate and validate relevant targets. Alternatively, as demonstrated in the example of the 5-HT^4 agonist, the approach can identify stronger and better clinical candidates in discovery research. By using virtual patients, such an approach can provide guidance for personalized and rational therapeutic polypharmacy. Fig. 1 shows a possible strategy to implement this form of QSP in a clinical development program.

6. Using crowdsourcing to bridge between big data tools and neurodegenerative diseases

ALS is a neurodegenerative disease that specifically attacks motor neurons, leading to progressive paralysis and death, typically within 3–5 years. One in 1000 people will die of ALS [33], and it has no effective treatment. The heterogeneity of ALS manifestation—similar to all

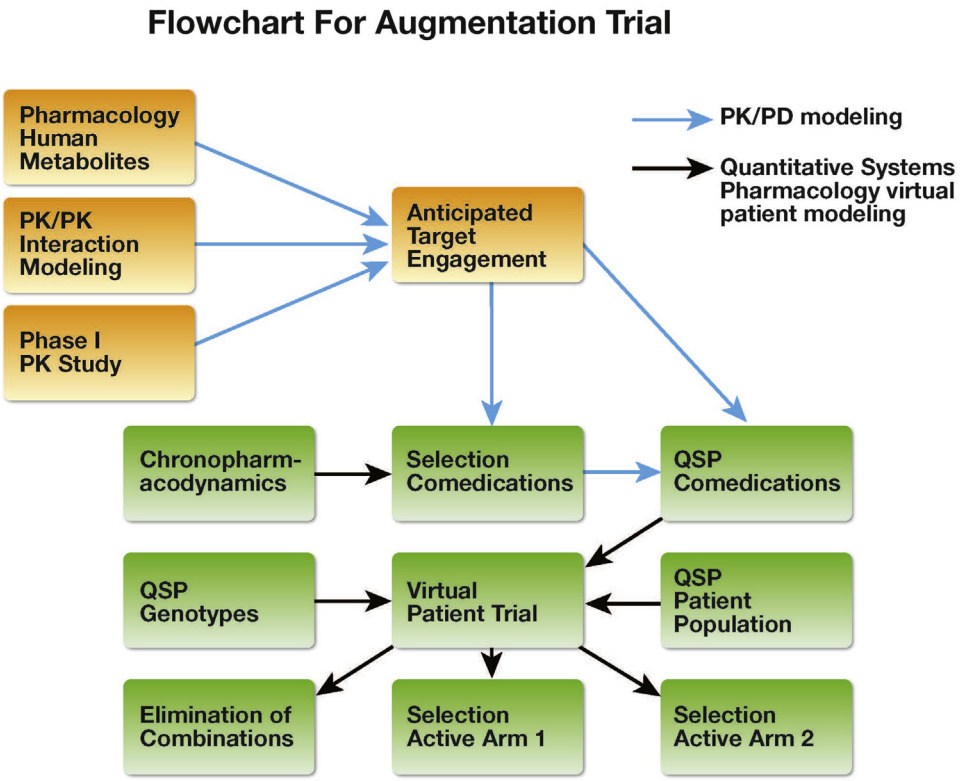

Fig. 1. Possible virtual patient strategy by implementing quantitative systems pharmacology (QSP) in a clinical development program for a novel drug target. Pharmacokinetic (PK) modeling (blue lines) based on phase I studies can be used to derive a model for the dose and time dependency of plasma levels. PK-PK interaction with co-medications can be simulated based on the known affinity of the drugs for metabolizing enzymes and transporters. QSP simulation (black lines) based on PET tracer displacement studies or noninvasive biomarkers such as BOLD fMRI or phEEG can be used to obtain an estimate of anticipated target engagement. The effect of co-medications and genotypes on a pharmacodynamic (PD) readout can be derived in the mechanism-based platform from the pharmacology of CNS-active co-medications and the implementation of common genotypes. Different disease states can be simulated, and available data on the circadian rhythm of key targets can be incorporated into the QSP platform to determine the optimal conditions of dose selection, patient population, inclusion/exclusion criteria, and genotype stratification. Finally, a virtual patient trial where individual "subjects" are sampled from probability distributions around genotypic, and biological variability can be combined with PK variability to better estimate the number of patients needed to identify a clinical robust signal.

neurodegenerative diseases—presents substantial barriers to the planning and interpretation of clinical trials for ALS treatments, leading to large, expensive clinical trials and even worse, the potential failure of an effective drug that will not reach patients solely due to the variability of patients in the trial. Indeed, recent failures in large clinical trials have left the drug industry and the patient population looking for new approaches and solutions.

Prize4Life, a nonprofit dedicated to the acceleration of treatments and a cure for ALS, was determined to address this need and remove the barriers for the development of ALS treatments. To effectively address this problem, two important tools are needed:

(1) A large data set of clinical, longitudinal, and patient information, with as diverse data sets as possible, to provide a comprehensive overview of the patient disease state,
(2) New computational approaches obtainable through crowdsourcing.

Large sample data sets are critical for identifying statistically significant and biologically relevant variables, particularly for diseases resulting from the complex interplay of genetic and environmental factors. To reach such large sample data sets, pooled clinical trial data sets have proven to be an invaluable resource for researchers seeking to unravel other complex neurodegenerative diseases [4,6,34,35]. With that in mind, Prize4Life collaborated with the Neurological Clinical Research Institute at Massachusetts General Hospital to create the Pooled Resource Open-Access ALS Clinical Trials (PRO-ACT, www.ALSdatabase.org) platform with funding from the ALS Therapy Alliance and in partnership with the Northeast ALS Consortium. The vision of the PRO-ACT project was to accelerate and enhance translational ALS research by designing and building a data set that would contain the merged data from as many completed ALS clinical trials as possible. Containing >8600 patients [36], PRO-ACT was launched as an open access platform for researchers in December 2012. Since then, it has been used by over 1000 researchers.

One effective way to foster the development of effective and diverse computational tools, and facilitate an unbiased assessment of their performance, is through crowdsourcing [37]. Crowdsourcing also allows for reaching out to new disciplines previously not engaged in ALS research.

To address the question of the variability in the progression of ALS, a subset of the PRO-ACT data was used before its public launch for an international crowdsourcing initiative, The DREAM-Phil Bowen ALS Prediction Prize4Life. The challenge asked solvers to predict the rate of progression of ALS for individual patients over the course of a year, using only 3 months of data. Algorithms were evaluated in a statistically rigorous blinded assessment. The prize ($50,000) was to be awarded for the most accurate methods to predict ALS progression. The challenge was administered through a collaboration between the Dialogue for Reverse Engineering Assessments and Methods (DREAM) initiative and Prize4Life using the InnoCentive Platform.

The challenge drew 1073 solvers from 63 countries and resulted in the submission of 37 unique algorithms from which two winning entries were identified, with one team coming from an analytic marketing company, Sentrana, Washington DC, and another team from Stanford University. In a post-challenge survey, 80% of the solvers who responded indicated they had minimal knowledge regarding ALS before participation.

The best-performing algorithms predicted disease progression better than both a baseline model and clinicians using the same data. Clinical trial modeling indicates that using the algorithms should enable a substantial reduction of >20% in the population size of a clinical trial required to demonstrate a drug effect, which translates to reduction of millions of dollars in the costs of all clinical trials using the data [38]. The algorithms are now in use or under consideration by several pharmaceutical companies and researchers.

Finally, the challenge was able to also offer new insights into factors predicting ALS progression. The prediction algorithms not only corroborated several previously identified predictive features, including uric acid, age, site of onset, and time from onset, the algorithms also uncovered several clinical measurements formerly unknown to be predictive of disease progression, including creatinine, creatine kinase, pulse, phosphorus, and blood pressure. These readily obtained measures hold promise to further our understanding and, one day, treatment of ALS.

Overall, this example demonstrates the potential of using pooled clinical trial data and crowdsourcing initiatives to develop new models and tools for better research and development in neurodegenerative diseases. More broadly, it suggests that improved access to better, larger, and more diverse data sets of patient data holds a great promise for improving clinical development and clinical practice beyond its current state.

7. Other examples of modeling influencing regulatory decisions

There are a few examples where modeling and simulation have helped influence regulatory agencies to approve a drug under slightly different conditions without the need for additional clinical trials. In post-herpetic neuralgia, regulatory approval for specific doses of gabapentin was granted based on sophisticated exposure-response modeling using five clinical trials with different doses [39,40] without the need to conduct another clinical trial. The approved label mentioned explicitly that "pharmacokinetic/pharmacodynamic (PK/PD) modeling provided confirmatory evidence of efficacy across all doses." The anti-epileptic oxcarbazepine (Trileptal) was approved as monotherapy in children, based on advanced PK/PD modeling [41] and subsequently confirmed in a clinical trial [42].

Recently, an advanced mechanism-based QSP model was used for simulating the impact of human recombinant parathyroid hormone [43] with regard to alternative dosing regimens. This suggests that regulatory agencies are becoming increasingly interested in advanced multiscale modeling and simulation beyond the more traditional PK/PD modeling.

8. Discussion

This position paper discusses some examples and real-world tests of models and simulations for generating actionable knowledge from many databases that are currently being collected and the domain expertise that has been generated in the neuroscience community. For AD, we interpret "actionable" as knowledge that can be used to actively support drug discovery programs for therapeutic interventions (in terms of target validation and identification of the optimal modulation strategy for the selected target), to identify possible responder populations for specific targets, and to validate clinical readouts that can demonstrate relevant changes in disease progression.

The examples in this report range from sharing data initiatives in precompetitive consortia such as C-Path and ADNI over crowdsourcing to better define the historical disease progression in ALS and identifying novel pathways in MS to mechanism-based predictive modeling of virtual human patients that could inform the selection of clinical candidates for drug development, improve clinical trial design, or guide rational polypharmacy in real-world situations.

The examples also illustrate the wide variety of relevant stakeholders, ranging from preclinical and clinical researchers, regulatory agencies, and pharmaceutical companies. Obviously in all these projects, patients remain the most important stakeholders and provide the most important information for the development and validation of the different analysis techniques. The engagement of regulators enables decisions that have implications across multiple targets independent of the sponsor and the precise mechanism of action of drug candidates. Ultimately, the purpose of predictive modeling in this disease area is to get the right drug to the right patients in the shortest amount of time.

It is also important to realize that different scientific questions need to be addressed with different levels of detail in various models. First and foremost, data sharing in a standardized format is an essential requirement for generating actionable knowledge. This applies both to data from real-life situations and to data from clinical trials. Initiatives such as C-Path and the Innovative Medicines Initiative are making progress on this issue.

For some qualitative relationships (e.g., the factors driving clinical progression), deep analytical approaches based on advanced statistical approaches are probably sufficient, as illustrated in the example of ALS and Orion BioNetworks. These approaches can also address more complex quantitative issues such as a better estimation of the progression of placebo trajectory.

For more predictive and quantitative questions that are relevant to actual drug discovery and development (i.e., in what direction and how much do we need to therapeutically affect one or more biological processes to have a substantial impact on the clinical phenotype in a specific patient population), formalized integration of extensive neurobiological domain expertise seems to be a solution to define better the causality between different biological processes as illustrated by the QSP.

This last example also illustrates the predictive nature, that is, the ability to estimate the clinical trial outcome of a new intervention, purely based on preclinical information. Going from purely exploratory analyses to approaches that are more predictive in nature is absolutely essential to support the development of new and better therapeutic interventions. By combining data and advanced modeling algorithms with formalized domain expertise of neurophysiologists, clinicians, imagers, and neuropharmacologists, this approach has the capability of being a game changer in the quest for new and effective drugs for AD.

In summary, this article shows a number of real-world examples of predictive modeling in CNS diseases in general and in AD in particular. Impact for this approach is evident for all phases of drug development. The different examples, using various analytical approaches, illustrate the impact and level of detail that can be achieved for the different stakeholders. These approaches could be a good starting point to expand the toolbox of the next-generation analysis techniques that will help bring the right drug to the right patient.

RESEARCH IN CONTEXT

1. Systematic review: We searched for various real-world examples for predictive modeling for CNS disorders that had shown actual real benefit for helping drug discovery and development.
2. Interpretation: We described private-public partnerships focused on data sharing, causal inference, and pathway-based analysis in multiple sclerosis, crowdsourcing in ALS and mechanism-based quantitative systems modeling in Alzheimer's disease with predictive value and potential impact on R&D. In at least one case, a quantitative and unexpected prediction was confirmed in a subsequent clinical trial.
3. Future directions: We intend to update the list on a regular basis to demonstrate that such in silico approaches can have a valuable impact on CNS research and development programs.

References

[1] Zahs KR, Ashe KH. 'Too much good news' - are Alzheimer mouse models trying to tell us how to prevent, not cure, Alzheimer's disease? Trends Neurosci 2010;33:381–9.

[2] Snyder HM, Bain LJ, Egge R, Carrillo MC. Alzheimer's disease public-private partnerships: A landscape of the global nonprofit community. Alzheimers Dement 2013;9:466–71.

[3] Brumfield M. The Critical Path Institute: transforming competitors into collaborators. Nat Rev Drug Discov 2014;13:785–6.

[4] Romero K, de Mars M, Frank D, Anthony M, Neville J, Kirby L, et al. The Coalition Against Major Diseases: developing tools for an integrated drug development process for Alzheimer's and Parkinson's diseases. Clin Pharmacol Ther 2009;86:365–7.

[5] Neville J, Kopko S, Broadbent S, Aviles E, Stafford R, Solinsky CM, et al. Development of a unified clinical trial database for Alzheimer's disease. Alzheimers Dement 2015;11:1212–21.

[6] Rogers JA, Polhamus D, Gillespie WR, Ito K, Romero K, Qiu R, et al. Combining patient-level and summary-level data for Alzheimer's disease modeling and simulation: a beta regression meta-analysis. J Pharmacokinet Pharmacodyn 2012;39:479–98.

[7] Romero K, Sinha V, Allerheiligen S, Danhof M, Pinheiro J, Kruhlak N, et al. Modeling and simulation for medical product development and evaluation: highlights from the FDA-C-Path-ISOP 2013 workshop. J Pharmacokinet Pharmacodyn 2014;41:545–52.

[8] Ramagopalan SV, Sadovnick AD. Epidemiology of multiple sclerosis. Neurol Clin 2011;29:207–17.

[9] Sospedra M, Martin R. Immunology of multiple sclerosis. Annu Rev Immunol 2005;23:683–747.

[10] Mahad DH, Trapp BD, Lassmann H. Pathological mechanisms in progressive multiple sclerosis. Lancet Neurol 2015;14:183–93.

[11] Comabella M, Montalban X. Body fluid biomarkers in multiple sclerosis. Lancet Neurol 2014;13:113–26.

[12] Bove R, Secor E, Healy BC, Musallam A, Vaughan T, Glanz BI, et al. Evaluation of an online platform for multiple sclerosis research: patient description, validation of severity scale, and exploration of BMI effects on disease course. PLoS One 2013;8:e59707.

[13] Gauthier SA, Glanz BI, Mandel M, Weiner HL. A model for the comprehensive investigation of a chronic autoimmune disease: the multiple sclerosis CLIMB study. Autoimmun Rev 2006;5:532–6.

[14] Xing H, McDonagh PD, Bienkowska J, Cashorali T, Runge K, Miller RE, et al. Causal modeling using network ensemble simulations of genetic and gene expression data predicts genes involved in rheumatoid arthritis. PLoS Comput Biol 2011;7:e1001105.

[15] Lee E, Chuang HY, Kim JW, Ideker T, Lee D. Inferring pathway activity toward precise disease classification. PLoS Comput Biol 2008; 4:e1000217.

[16] Su J, Yoon BJ, Dougherty ER. Accurate and reliable cancer classification based on probabilistic inference of pathway activity. PLoS One 2009;4:e8161.

[17] Johannes M, Brase JC, Frohlich H, Gade S, Gehrmann M, Falth M, et al. Integration of pathway knowledge into a reweighted recursive feature elimination approach for risk stratification of cancer patients. Bioinformatics 2010;26:2136–44.

[18] Chuang HY, Rassenti L, Salcedo M, Licon K, Kohlmann A, Haferlach T, et al. Subnetwork-based analysis of chronic lymphocytic leukemia identifies pathways that associate with disease progression. Blood 2012;120:2639–49.

[19] Kim S, Kon M, DeLisi C. Pathway-based classification of cancer subtypes. Biol Direct 2012;7:21.

[20] Zhang M, Zhang L, Zou J, Yao C, Xiao H, Liu Q, et al. Evaluating reproducibility of differential expression discoveries in microarray studies by considering correlated molecular changes. Bioinformatics 2009;25:1662–8.

[21] Hofree M, Shen JP, Carter H, Gross A, Ideker T. Network-based stratification of tumor mutations. Nat Methods 2013;10:1108–15.

[22] International Multiple Sclerosis Genetics Consortium. Network-based multiple sclerosis pathway analysis with GWAS data from 15,000 cases and 30,000 controls. Am J Hum Genet 2013; 92:854–65.

[23] Athey BD, Braxenthaler M, Haas M, Guo Y. tranSMART: An open source and community-driven informatics and data sharing platform for clinical and translational research. AMIA Jt Summits Transl Sci Proc 2013;2013:6–8.

[24] Szalma S, Koka V, Khasanova T, Perakslis ED. Effective knowledge management in translational medicine. J Transl Med 2010;8:68.

[25] Geerts H, Spiros A, Roberts P, Carr R. Quantitative systems pharmacology as an extension of PK/PD modeling in CNS research and development. J Pharmacokinet Pharmacodyn 2013;40:257–65.

[26] Roberts PD, Spiros A, Geerts H. Simulations of symptomatic treatments for Alzheimer's disease: computational analysis of pathology and mechanisms of drug action. Alzheimers Res Ther 2012;4:50.

[27] Nicholas T, Duvvuri S, Leurent C, Raunig D, Rapp T, Iredale R, et al. Systems pharmacology modeling in neuroscience: prediction and outcome of PF-04995274, a 5HT4 partial agonist, in a clinical scopolamine impairment trial. Adv Alzheimers Dis 2013;2:83–98.

[28] Geerts H, Spiros A, Roberts P, Twyman R, Alphs L, Grace AA. Blinded prospective evaluation of computer-based mechanistic schizophrenia disease model for predicting drug response. PLoS One 2012;7:e49732.

[29] Liu J, Ogden A, Comery TA, Spiros A, Roberts P, Geerts H. Prediction of efficacy of vabicaserin, a 5-ht2c agonist, for the treatment of schizophrenia using a quantitative systems pharmacology model. CPT Pharmacometrics Syst Pharmacol 2014;3:e111.

[30] Geerts H, Roberts P, Spiros A, Potkin S. Understanding responder neurobiology in schizophrenia using a quantitative systems pharmacology model: Application to iloperidone. J Psychopharmacol 2015; 29:372–82.

[31] Lavedan C, Licamele L, Volpi S, Hamilton J, Heaton C, Mack K, et al. Association of the NPAS3 gene and five other loci with response to the antipsychotic iloperidone identified in a whole genome association study. Mol Psychiatry 2009;14:804–19.

[32] Geerts H, Spiros A, Roberts P. Assessing the synergy between cholinomimetics and memantine as augmentation therapy in cognitive impairment in schizophrenia. A virtual human patient trial using quantitative systems pharmacology. Front Pharmacol 2015;6:198.

[33] Kiernan MC, Vucic S, Cheah BC, Turner MR, Eisen A, Hardiman O, et al. Amyotrophic lateral sclerosis. Lancet 2011;377:942–55.

[34] Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. Brain 1999;122:871–82.

[35] Ontaneda D, LaRocca N, Coetzee T, Rudick R. Revisiting the multiple sclerosis functional composite: proceedings from the National Multiple Sclerosis Society (NMSS) Task Force on Clinical Disability Measures. Mult Scler 2012;18:1074–80.

[36] Atassi N, Berry J, Shui A, Zach N, Sherman A, Sinani E, et al. The PRO-ACT database: design, initial analyses, and predictive features. Neurology 2014;83:1719–25.

[37] Marbach D, Costello JC, Kuffner R, Vega NM, Prill RJ, Camacho DM, et al. Wisdom of crowds for robust gene network inference. Nat Methods 2012;9:796–804.

[38] Kuffner R, Zach N, Norel R, Hawe J, Schoenfeld D, Wang L, et al. Crowdsourced analysis of clinical trial data to predict amyotrophic lateral sclerosis progression. Nat Biotechnol 2015;33:51–7.

[39] Miller R, Ewy W, Corrigan BW, Ouellet D, Hermann D, Kowalski KG, et al. How modeling and simulation have enhanced decision making in new drug development. J Pharmacokinet Pharmacodyn 2005; 32:185–97.

[40] Bockbrader HN, Wesche D, Miller R, Chapel S, Janiczek N, Burger P. A comparison of the pharmacokinetics and pharmacodynamics of pregabalin and gabapentin. Clin Pharmacokinet 2010;49:661–9.

[41] Ahmad A, Garnett WR. Simulated fluctuations in plasma drug concentrations for patients receiving oxcarbazepine or carbamazepine extended-release capsules. Clin Drug Investig 2005;25:669–73.

[42] Eun SH, Kim HD, Chung HJ, Kang HC, Lee JS, Kim JS, et al. A multicenter trial of oxcarbazepine oral suspension monotherapy in children newly diagnosed with partial seizures: a clinical and cognitive evaluation. Seizure 2012;21:679–84.

[43] Peterson MC, Riggs MM. FDA advisory meeting clinical pharmacology review utilizes a quantitative systems pharmacology (QSP) model: a watershed moment? CPT Pharmacometrics Syst Pharmacol 2015;4:e00020.