CrossMark

Disease Progression Modeling: Key Concepts and Recent Developments

Sarah F. Cook[1,2] · Robert R. Bies[1,2,3]

Abstract Disease modeling involves the use of mathematical functions to describe quantitatively the time course of disease progression. In order to characterize the natural progression of disease, these models generally incorporate longitudinal data for some biomarker(s) of disease severity or can incorporate more direct measures of disease severity. Disease models are also often linked to pharmacokinetic–pharmacodynamic models so that the influence of drug treatment on disease progression can be quantified and evaluated. Regulatory agencies have embraced disease progression models as powerful tools that can be used to improve drug development productivity. This article provides a brief overview of key concepts in disease progression modeling followed by illustrative examples from models for Alzheimer's disease. Finally, recent novel applications in which disease progression models have been linked to cost-effectiveness analysis and genomic analysis are described.

Keywords Disease progression · Modeling and simulation · Drug development · Clinical pharmacology · Pharmacometrics

✉ Robert R. Bies
robertbi@buffalo.edu

1 Department of Pharmaceutical Sciences, University at Buffalo, State University of New York, 445 Kapoor Hall, Buffalo, NY 14214, USA

2 Department of Pharmacology and Therapeutics, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA

3 Centre for Addiction and Mental Health, University of Toronto, 1001 Queen Street West, Toronto, ON M6J 1H4, Canada

Introduction

Disease progression modeling involves the utilization of mathematical functions as quantitative descriptors of the time course of disease status. Disease models may incorporate biomarkers of disease severity and/or clinical outcomes to characterize the natural progression of disease. In clinical pharmacology, disease models are often integrated with pharmacokinetic–pharmacodynamic models in order to quantify the influence of drug treatment on disease progression. Seminal work in disease progression modeling was published in the early 1990s [1–5], but the field has truly flourished in more recent years, since the US Food and Drug Administration (FDA) embraced such quantitative models as powerful tools that can be used to improve drug development productivity. In 2004, the FDA launched its Critical Path Initiative, which aims to modernize drug development, and disease modeling was promoted as a key opportunity for advancement [6, 7]. Since that time, disease modeling has been widely adopted by the pharmaceutical industry [8]. In addition to encouraging the application of disease progression models to pharmaceutical research and development [6, 7, 9], the FDA has developed and made publicly available models for diseases of public health importance, including Alzheimer's disease, Parkinson's disease, bipolar disorder, obesity, and nonsmall cell lung cancer [10]. This article provides a brief review of the types and components of disease progression models followed by illustrative case studies in the area of Alzheimer's disease. Finally, recent novel applications of disease progression models to cost-effectiveness analysis and genome-wide analysis are described.

Disease Progression Modeling

Disease progression models can be broadly categorized into three classes: empirical, semi-mechanistic, and systems biology. Empirical models are purely data driven and do not describe underlying biological processes; they serve as mathematical frameworks for interpolation between observed data. At the other extreme, systems biology models of disease progression are physiologically based and incorporate mathematical representations of biological, pathophysiological, and pharmacological processes in as much molecular detail as possible. Semi-mechanistic models fall in between the poles of empirical and systems biology models. Importantly, the appropriate model type depends highly upon the problem at hand. Whereas empirical models are well suited to answering questions that are relatively narrow in scope (e.g., dose selection or clinical trial design and interpretation), the more mechanistic nature of semi-mechanistic and systems models enables application to broader scientific questions (e.g., prediction of effects of drugs with differing mechanisms of action, novel target identification, or risk projection based on biomarker data) [9, 11]. This review will first cover some considerations related to the types of data that can be utilized in disease progression models before discussing these three classes of models in greater detail.

Data

Various types of data can be incorporated into disease progression models [12]. Assessment of disease severity in many areas, especially neurological and mental health disorders, relies upon relatively subjective scoring systems based on patient-reported responses or clinician evaluation. Utilization of subjective measures in disease models poses particular challenges because such data tend to exhibit high variability, and it may therefore be necessary to include many subjects in order to obtain an adequate description of disease progression. Some examples of this type of data that have been used in disease progression models include the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog) [2, 3, 5, 13], the Hamilton Rating Scale for Depression (HAM-D) [14–16], and the Revised Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS-R) [17, 18]. Although these types of observations are often technically ordinal data, they are generally treated as continuous variables for the purposes of disease modeling because many of these scales comprise relatively wide ranges of possible scores (e.g., 0–70 for ADAS-Cog). Continuous biomarker data are also widely used in disease models. For instance, in a semi-mechanistic bone cycle model, C-terminal telopeptide of type I collagen (CTX) and osteocalcin were incorporated as serum biomarkers of bone resorption and formation, respectively, and bone mineral density was included as a measure of bone mass [19].

Additionally, continuous disease progression models can be linked to discrete clinical endpoints through survival or hazard functions. Approaches for linking disease progression to clinical outcome have been reviewed by Mould [20], and the author provided detailed examples of models linking the time course of chemotherapy-induced myelosuppression to the probability of developing febrile neutropenia [21] and tumor size to survival [22, 23].

Disease models are often developed to characterize chronic diseases that progress slowly, so these models commonly require incorporation of disease severity data that has been collected over extended time periods. Such data is rarely available from a single study, and data from several clinical studies must often be pooled in order to provide sufficient observations for modeling of disease progression. Consequently, model-based meta-analysis is an effective approach for disease modeling. Model-based meta-analysis involves model development using the combined results of multiple previous studies. Unlike traditional model development, which generally utilizes patient-level data from a limited number of clinical studies, model-based meta-analysis may incorporate both summary-level information obtained from literature references and patient-level information obtained from databases and/or available clinical trial data. Importantly, nonlinear mixed-effects modeling, which is widely utilized for empirical and semi-mechanistic disease modeling, can account for differences in study design and subject characteristics to reduce the potential for bias introduced by pooling data from multiple studies. Ahn and French have demonstrated meta-analytic approaches for fitting longitudinal summary-level data with NONMEM software, which is commonly used for nonlinear mixed-effects models of disease progression [24]. They provided illustrative NONMEM code for incorporation of study-level and/or treatment arm-level random effects, both with and without residual correlation, and they used simulated data to evaluate the importance of including these different model components. Mould has also reviewed the application of meta-analysis to disease progression modeling [20] and provided detailed discussion of meta-models for progression of Alzheimer's disease [25] and Crohn's disease [26]. Although model-based meta-analysis poses certain challenges, results from such analyses are often statistically stronger and more generalizable than those from single studies [27].

Empirical Models

The empirical approach has commonly been employed in areas where disease severity is routinely assessed via subjective scoring systems. For instance, empirical models have been utilized to model progression of Alzheimer's disease [2, 3, 5], Parkinson's disease [28, 29], depression [14, 15], schizophrenia [30, 31], amyotrophic lateral sclerosis [17], and Huntington's disease [32]. The simplest disease

progression model is linear, assuming a constant rate of change for a measure of disease status over time (Eq. (1)):

$$S(t) = S_0 + \alpha \times t \tag{1}$$

where $S(t)$ is disease status at time t, S_0 is the baseline disease status (y-intercept), and α is the rate of change of disease status (slope). A hypothetical linear disease model is depicted in Fig. 1. Although this function is quite simple, linear models have been successfully implemented to describe progression of several diseases, including Alzheimer's disease [2, 3, 5], schizophrenia [30], and Huntington's disease [32]. For many diseases, however, more complex equations must be used to describe the time course of progression adequately. Details about more complex models, such as asymptotic progress, physiological turnover, and growth and decay models, have been reported in other publications [12, 33]. Mould has also provided example code for implementing such models in NONMEM software [12].

It should be noted that baseline disease status, S_0, can be a complex parameter in disease progression models. It may be defined as some measure of disease status either at the time a disease started, at the time a disease was diagnosed, or at the time a subject enrolled in a clinical trial. Alternatively, if the disease status, $S(t)$, is defined as the *change* in a disease measure relative to a baseline value, rather than as the actual value of that measure, then S_0 may be fixed at zero. The precise definition of S_0 can have important implications for how the disease progresses from the baseline time. For instance, in
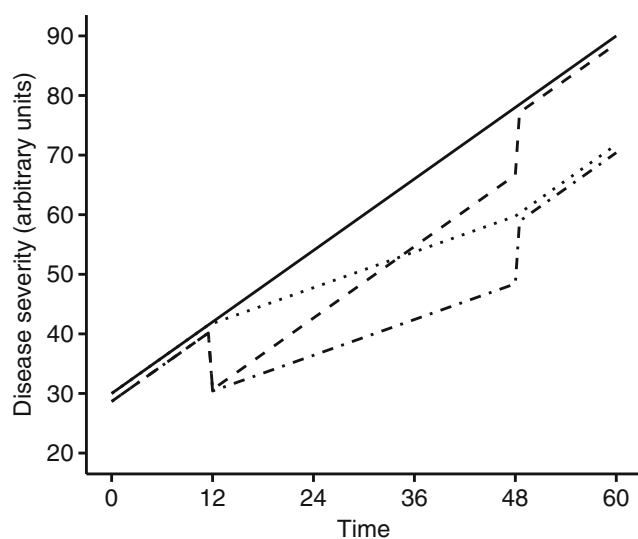

Fig. 1 A hypothetical empirical disease model with linear progression. The *solid line* represents disease progression with no treatment, and the *other lines* represent disease progression with treatment effects that start at time 12 and end at time 48. The *dashed line* illustrates the effect of a drug that provides symptomatic relief, the *dotted line* illustrates the effect of a drug that modifies the rate of disease progression, and the *dotted-and-dashed line* illustrates the combined effects of a disease-modifying drug that also provides symptomatic relief. Note that the lines have been offset slightly to prevent overlap

linear models of Alzheimer's disease progression, baseline disease status is generally defined near the time of trial enrollment, and baseline status has been observed to interact strongly with disease progression rate, which increased with increasing baseline severity [25, 34, 35]. This potential for interaction between baseline disease status and disease progression rate makes baseline severity a crucial factor in clinical trial design and in evaluation of clinical trial and disease model results.

Drug effects can be incorporated into disease models in different ways, and the manner in which a drug effect is included in a disease progression equation has implications for how the treatment influences the disease. For instance, a drug may provide only symptomatic relief. For the linear disease progression example depicted above (Eq. (1)), a symptomatic drug effect could be incorporated as follows (Eq. (2)):

$$S(t) = S_0 + \alpha \times t + E(t) \tag{2}$$

where $S(t)$ is disease status at time t, S_0 is the baseline disease status (y-intercept), α is the rate of change of disease status (slope), and $E(t)$ is a function incorporating the drug effect. In this case, the drug does not influence the rate of disease progression but merely provides a transient benefit by offsetting the disease severity; when treatment is discontinued, the disease will return to a state that is the same as if the treatment had not occurred (Fig. 1). Alternatively, a drug may alter the progression of a disease. For the linear disease progression example depicted above (Eq. (1)), such a disease-modifying effect could be incorporated as follows (Eq. (3)):

$$S(t) = S_0 + [\alpha + E(t)] \times t \tag{3}$$

where $S(t)$ is disease status at time t, S_0 is the baseline disease status (y-intercept), α is the rate of change of disease status (slope) in the absence of drug treatment, and $E(t)$ is a function incorporating the drug effect. Here, the drug alters the rate of disease progression; when treatment is discontinued, the rate of disease progression returns to that in the absence of treatment (α), but the disease severity will be offset from what it would have been if treatment had not occurred (Fig. 1). Methods for incorporating different types of drug effects into more complex disease progression models have been reported in other publications [12, 33] along with example NONMEM code [12].

One critical consideration when comparing drug effects to standard-of-care treatment is the fact that standard of care evolves over time. Improvements in standard of care are likely to be accompanied by changes in the typical time course of disease progression, so disease models and databases must be

updated regularly in order to consider the response of disease to novel treatments in an appropriate, current context.

Systems Biology Models

Systems biology models of disease progression incorporate as much molecular detail as possible. Consequently, systems models generally require extensive scientific collaboration and integration of in vitro, ex vivo, in vivo, nonclinical, and clinical data [11]. The complexity of systems disease modeling is evident in work that has been performed in the area of diabetes research. Diabetes is a chronic, complex disease, and a comprehensive model of diabetes progression would incorporate mathematical representations of the molecular pathways involved in glucose homeostasis, diabetes pathophysiology, and the diverse complications associated with hyperglycemia-induced tissue damage. After decades of modeling efforts, researchers are still working toward such a systems model [36]. Ajmera et al. recently reviewed developments in diabetes modeling from the past 50 years and identified challenges that must be addressed in order to achieve a complete description of the relevant regulatory network [37]. Additionally, as a collaborative resource for fellow researchers, Ajmera et al. encoded several key models from the literature in a standard model description language (SBML, systems biology mark-up language [38]) and shared these models via BioModels Database [39], which is an online repository for quantitative models of biological processes.

One example of a systems-based disease model is that for calcium homeostasis and bone remodeling published by Peterson and Riggs [40, 41]. The model incorporated relevant molecular pathways in physiological spaces representing the gut, vasculature, intracellular phosphate, kidneys, parathyroid gland, bone, and intracellular osteoblast components [40] (Fig. 2a). Although the model contained some empirical elements that could be refined through future experiments, the systems approach enabled adequate description of markers of calcium homeostasis, bone turnover, and bone mineral density under diverse disease states and therapeutic conditions, including hypoparathyroidism, hyperparathyroidism, renal insufficiency, antiresorptive therapy (denosumab), and anabolic therapy (teriparatide) [40, 41]. In 2014, in what has been described as "a watershed moment" for the discipline of quantitative systems pharmacology [42], clinical pharmacologists at the FDA used an open-source version of this model to evaluate the appropriateness of a proposed dosage regimen for a new biologic treatment for hypoparathyroidism [43].

Semi-mechanistic Models

Like empirical models, semi-mechanistic models are largely data driven, and like systems models, semi-mechanistic models incorporate knowledge about underlying biological,

pathophysiological, and pharmacological processes; however, in contrast to systems models, semi-mechanistic models are not intended to provide comprehensive depictions of molecular pathways. Instead, parsimony is preferred, and mathematical representations of biological processes generally incorporate just enough complexity to describe the relevant data accurately. For instance, van Schaick et al. recently reported a semi-mechanistic model of bone turnover and bone mineral density for evaluating the effects of osteoporosis treatments [19]. The semi-mechanistic approach was implemented because earlier mathematical models of bone remodeling were too complex to characterize adequately the time course of clinical biomarkers of bone turnover and the associated variability observed in clinical studies. Rather than including detailed representation of all molecular and cellular processes involved in bone remodeling, the semi-mechanistic model incorporated assumptions and simplified functions that were thought to capture the most critical aspects of the process. Specifically, the full model consisted of a closed form cyclical model with four compartments representing bone resorption, formation, primary mineralization, and secondary mineralization, a disease progression model representing bone loss in osteoporosis, a vitamin D and calcium supplementation (placebo) model, and a drug model for antiresorptive treatments (Fig. 2b). The full model accurately described the short-term (<1 year) and long-term (1–10 years) time course of bone turnover biomarkers and bone mineral density in postmenopausal women receiving treatment with either placebo or antiresorptive drugs (bisphosphonate or denosumab). Given the long duration of late-phase osteoporosis trials, incorporation of disease progression was a critical component of the full model. Potential future applications of this model include optimization of dosage regimens, prediction of effects of novel osteoporosis therapeutics, prediction of long-term bone mineral density values in patients based on bone turnover biomarker values observed in short-term phase I trials, and linkage of the model to a discrete clinical outcome, such as the probability of bone fracture.

Case Studies in Alzheimer's Disease

Alzheimer's disease is a chronic, neurodegenerative disease that afflicts an estimated 5.3 million Americans and is the sixth leading cause of death in the USA [44]. The chronic, degenerative nature of the disease combined with its high prevalence and societal burden has made this area ripe for application of disease progression modeling.

In a pair of companion papers published in 1992, Holford and Peace described the development and evaluation of an empirical disease model for progression of cognitive effects in Alzheimer's patients treated with the acetylcholinesterase inhibitor tacrine [2, 3]. This work, which proved seminal in
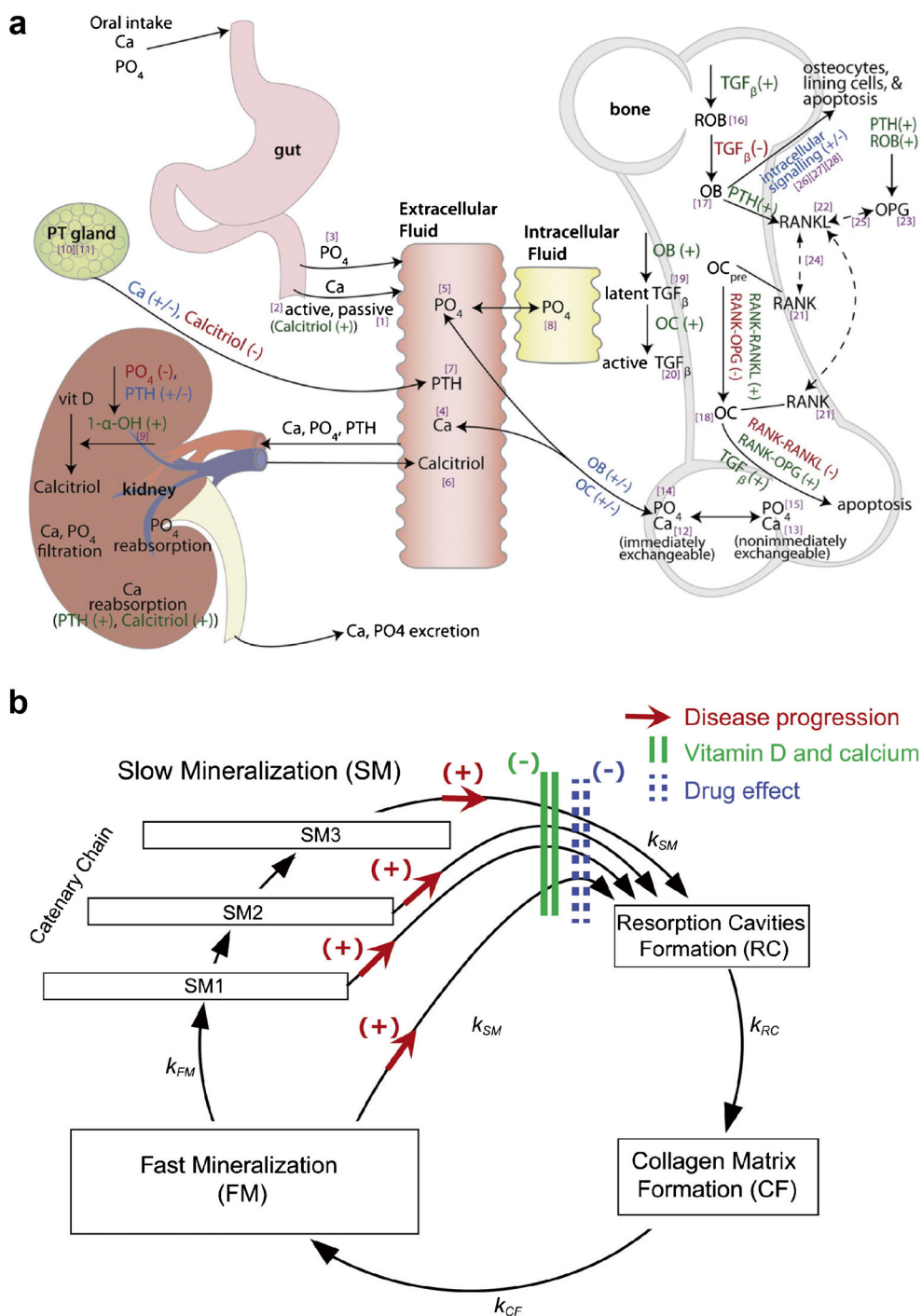
Fig. 2 **a** Schematic representation of the systems-based model describing calcium homeostasis and bone remodeling. *Solid arrows* indicate fluxes, *dashed arrows* indicate binding effects, *dark green plus signs* indicate stimulatory effects, *red minus signs* indicate inhibitory effects, *blue plus/minus signs* indicate bidirectional effects, and *bracketed purple numbers* refer to differential equation numbers in the source publication (reprinted with permission of Elsevier from [40], copyright 2009). **b** Schematic representation of the semi-mechanistic bone cycle model. The *red arrows* between the mineralization and resorption compartments indicate an increased transit rate due to disease progression, the *green bars* indicate a decreased transit rate due to vitamin D and calcium supplementation, and the *dashed blue bars* indicate a decreased transit rate due to a drug effect (reprinted with permission of Springer from [19], copyright Springer Science+Business Media New York 2015). *1-α-OH* 1-alpha-hydroxylase, *Ca* calcium, *OB* osteoblast, *OC* osteoclast, *OC*$_{pre}$ osteoclast precursor, *OPG* osteoprotegerin, *PO*$_4$ phosphate, *PT gland* parathyroid gland, *PTH* parathyroid hormone, *RANK* receptor activator of nuclear factor kappa B, *RANKL* RANK ligand, *ROB* responding osteoblast, *TGF*$_\beta$ transforming growth factor beta, *vit D* vitamin D

the application of nonlinear mixed-effects analysis to disease progression modeling, incorporated data from two clinical

trials of the same study design in subjects with probable Alzheimer's disease. One trial was conducted in the USA

and the other in France. The design involved an initial, randomized tacrine dose titration phase, a placebo baseline washout phase, a double-blind phase comparing placebo with the best-tolerated tacrine dose in patients who achieved a predefined response, and a blinded, sustained active phase with the best-tolerated tacrine dose in all patients. Patients were blinded to treatment throughout the trials. ADAS-Cog was the primary measure of disease status [13], and the final model successfully described the progression of ADAS-Cog, ADAS-Cog(*t*), with a linear model that incorporated symptomatic effects for both tacrine and placebo (Eq. (4)):

$$\text{ADAS-Cog}(t) = \text{ADAS-Cog}_0 + \alpha \times t + E_{\text{tacrine}}(t) + E_{\text{placebo}}(t) \tag{4}$$

where ADAS-Cog$_0$ is the baseline ADAS-Cog score; α is the rate of disease progression, which was estimated as 6.17 ADAS-Cog units/year; $E_{\text{tacrine}}(t)$ is a function incorporating the drug effect; and $E_{\text{placebo}}(t)$ is a function incorporating the placebo effect. The effect of tacrine shifted the disease progress curve by −2.99 ADAS-Cog units, and there were significant positive correlations between baseline ADAS-Cog score, disease progression rate, and the magnitude of the tacrine effect. Because subjects were blinded to treatment but aware of when changes in treatment occurred, the model included a placebo response at times of treatment change. The placebo model, $E_{\text{placebo}}(t)$, featured a Bateman-type function, which enabled estimation of the onset, offset, and magnitude of the placebo effect. Interestingly, the magnitude of the placebo effect was 76 % greater in the French trial compared to the US trial, and the offset of the placebo effect occurred more slowly in the French trial, perhaps reflecting cultural differences in expectations about treatment.

Over the past 25 years, many others have expanded upon the work of Holford and Peace. These efforts have been spurred on by prominent failures in Alzheimer's disease drug development, which have been partially attributed to inadequate clinical trial design [45]. In several recent publications, generalized logistic functions have been employed to characterize a sigmoidal progression of ADAS-Cog scores in Alzheimer's patients [46, 47, 48••]. For instance, Conrado et al. [48••] modeled data from the Coalition Against Major Diseases (CAMD) database for control-arm patients (both placebo and stable background Alzheimer's medication) from randomized, double-blind clinical trials in patients with mild-to-moderate Alzheimer's disease; the data were adequately described by the Richards function (a generalized logistic function) (Eq. (5)):

$$\text{ADAS-Cog}(t)$$
$$= \frac{\text{ADAS-Cog}_0 \times 70}{\left[\text{ADAS-Cog}_0^{\beta} + \left(70^{\beta} - \text{ADAS-Cog}_0^{\beta} \right) \times e^{-\beta r t} \right]^{1/\beta}} \tag{5}$$

where ADAS-Cog$_0$ is the baseline ADAS-Cog score, 70 is the maximum possible observed ADAS-Cog score, r is the intrinsic rate of disease progression, t is time, and β is the shape factor allowing for a noncentral inflection point of disease progression. This function describes an asymmetric, S-shaped curve with a nonlinear rate of disease progression that increases when ADAS-Cog is below the inflection point and decreases when ADAS-Cog is above the inflection point, which was estimated to occur at a score of 52. Thus, ADAS-Cog scores reach a plateau as they approach the maximum possible value. Compared to earlier linear models, the Richards function more appropriately constrains ADAS-Cog scores to the range of possible values (0–70) and also better characterizes the nonlinear, saturable progression of Alzheimer's disease, wherein disease scores increase exponentially during the early stages of disease and then the progression rate slows and approaches zero as ADAS-Cog scores approach 70. This plateau in the Richards function is consistent with the ceiling effect that has been observed in clinical practice when patients with severe Alzheimer's disease experience loss of speech and can no longer perform the tasks required to capture cognitive deterioration via ADAS-Cog. Additionally, in order to completely constrain model predictions to the possible range of ADAS-Cog scores, recent models have incorporated beta-distributed residual variability, which allows for a heteroscedastic error distribution with a variance that approaches zero as ADAS-Cog scores approach the boundaries [46, 48••]. Recent modeling efforts have also focused on identification of patient characteristics (i.e., covariates) that appear to influence the rate of disease progression. For instance, age and apolipoprotein E (*APOE*) genotype have consistently been identified as covariates affecting disease progression rate, with younger age and the presence of the *APOEε4* allele associated with faster progression [46, 47, 48••].

A milestone in Alzheimer's disease modeling was achieved in the summer of 2013 when the CAMD received regulatory endorsement from both the FDA and the European Medicines Agency (EMA) for its clinical trial simulation tool for mild-to-moderate Alzheimer's disease [49••]. At the core of this simulation tool is a disease progression model that incorporates many of the features described above. Disease progression is described by the longitudinal change in ADAS-Cog score over time, and scores advance along a sigmoidal curve with bounds at 0 and 70 [46]. The placebo effect is incorporated with a Bateman-type function [35], and symptomatic drug effects are described by an overall function offset [49••], as illustrated generically in Fig. 1 and Eq. (2). Although disease-modifying drugs for Alzheimer's disease are not currently available, the FDA and EMA agreed that claims of a disease-modifying drug effect could be substantiated by evidence of a drug-induced decrease in the rate of disease progression, provided that such a change in progression rate is

accompanied by a biomarker change that supports improvement in the underlying pathophysiology [49••]. For the purposes of clinical trial simulation, the tool also incorporates covariates, such as age and *APOEε4* allele status, which may influence disease progression. The FDA and EMA deemed this tool scientifically sound and fit for purpose in aiding design of future clinical trials in patients with mild-to-moderate Alzheimer's disease [49••], and intended applications include calculation of sample size, determination of optimal trial durations and treatment effect measurement times, identification of optimal data analysis methods, and comparison of the sensitivity of competing trial designs to assumptions about treatment effects [50]. Importantly, such a tool could not have been developed without construction and utilization of a large, collaborative, standardized database [51] that enabled integration of patient-level data with summary-level data from the literature. Similar efforts are currently underway in disease modeling of mild cognitive impairment and pre-symptomatic Alzheimer's disease [50, 52••].

Recent Developments

Integration of Disease Progression Modeling with Cost-Effectiveness Analysis

Widespread concern about rising healthcare costs has made cost-effectiveness analysis an increasingly important component of health technology assessment. At present, cost-effectiveness analyses for pharmaceutical agents primarily rely upon empirical models of disease progression, such as Markov models, to describe probabilistic transitions in the health states of patients over time [53]. Unfortunately, this approach relies on data from late-stage clinical trials and does not allow for model-informed prediction of cost-effectiveness during early stages of clinical development. Nevertheless, given the goal of providing effective, safe, and affordable treatment options in a timely fashion, there is increased interest among regulatory agencies, industry representatives, and coverage bodies in early integration of cost-effectiveness analysis with regulatory review [54].

Van Hasselt et al. recently described an integrated simulation framework for eribulin treatment in castration-resistant prostate cancer that included submodels for disease progression/clinical outcome, eribulin-induced toxicity, dose adaptation, and cost-effectiveness [55••]. The disease progression/clinical outcome model described the dynamics of prostate-specific antigen and its relationship to overall survival [56]. The models for eribulin-induced toxicity included dose-limiting neutropenia and other graded toxicities (nausea, fatigue, peripheral neuropathy, paresthesia, diarrhea, asthenia, and anemia), and the dose adaptation model incorporated definitions for when and how dose adjustments should be made

following toxicities. The hypothetical cost model included direct drug costs and adverse event-related costs. The cost model was only intended to demonstrate proof of concept and was, therefore, relatively simple; however, additional cost considerations, such as those associated with follow-up hospital visits or background mortality rates, could be readily incorporated into such a model. The integrated simulation was conducted in R using the deSolve [57] and MASS packages, and a detailed description of the implementation of cost-effectiveness analysis in R has been previously reported [58]. Using this integrated framework, the authors were able to simulate scenarios with different treatment protocols and patient characteristics and then evaluate the impact of these changes on expected efficacy and cost-effectiveness. Thus, the integrated framework successfully linked mechanistic clinical and pharmacological underpinnings to predictions about efficacy and cost. In addition to cancer, diabetes and heart disease are particularly costly therapeutic areas that are currently ripe for application of such integrated disease modeling strategies for cost-effectiveness analysis.

Integration of Disease Progression Modeling with Genomic Data Analysis

To date, the search for genetic variants that influence complex traits and diseases has largely relied upon phenotype data collected from individuals at single time points (i.e., cross-sectional measurements); however, interest has risen in integrating genotype data with longitudinal phenotype data that accounts for the progression of chronic disease [59, 60]. In a recent report, Wu et al. used an empirical nonlinear mixed-effects model to characterize lung function in asthma patients over time and then performed a genome-wide association study to identify genetic contributors to longitudinal phenotypes that were determined from the disease progression model [61••]. The authors modeled the time course of forced expiratory volume in 1 s (FEV_1) in children with mild-to-moderate asthma using longitudinal data (FEV_1 recorded every 2–4 months over a 4-year period) collected from 1041 children as part of the Childhood Asthma Management Program (CAMP) study [62]. FEV_1 was best predicted with an exponential function (Eq. (6)):

$$FEV_1 = e^{(\theta_1 \times age + \theta_2 \times height - \theta_3)} + \theta_{drug\ effect} \quad (6)$$

where θ_1 and θ_2 are the rates of change in FEV_1 associated with age and height, respectively, θ_3 is the baseline FEV_1, and $\theta_{drug\ effect}$ is the additive (symptomatic) improvement provided by the corticosteroid budesonide. In subsequent genome-wide analysis of a subset of study subjects (581 white patients), the individual-level parameter estimates for baseline FEV_1 (θ_3 in Eq. (6))

and for rate of change in FEV_1 with age (θ_1 in Eq. (6)) were used as longitudinal phenotypes, which were tested for association with 473,680 genotyped single nucleotide polymorphisms (SNPs). For comparison, genome-wide association analysis was also performed with a single-time-point phenotype (FEV_1 at 48 months). The genome-wide association study identified seven novel SNPs that were nominally associated with longitudinal lung function phenotypes (six SNPs for baseline FEV_1 and one SNP for rate of change in FEV_1 with age). Additionally, five SNPs that were previously associated with pulmonary function were suggestively associated ($p < 0.05$) with the longitudinal phenotypes, but the single-time-point phenotype was not associated ($p > 0.05$) with any of the previously reported SNPs that were considered. Thus, despite sample size limitations, genome-wide association analysis with the longitudinal phenotype appeared to enhance discovery of relevant genetic variants compared to analysis with the single-time-point phenotype [61••].

Conclusions

Disease models can be used to synthesize and quantitatively summarize knowledge about disease progression and the influence of drug treatment on disease severity. Models for chronic diseases that progress slowly often require incorporation of data that has been collected from multiple sources over extended time periods. Thus, collaborative research, data sharing, and database standardization are likely to be critical to the continued success of disease modeling endeavors. Recent years have seen notable effective collaborations between industry and regulatory agencies in the area of Alzheimer's disease modeling and simulation. Novel integration of disease progression modeling with other analyses, such as cost-effectiveness analysis and genomic data analysis, is likely to further expand the utility of this flexible, powerful modeling technique.

References

Papers of particular interest, published recently, have been highlighted as:
•• Of major importance

1. Griggs RC, Moxley 3rd RT, Mendell JR, Fenichel GM, Brooke MH, Pestronk A, et al. Prednisone in Duchenne dystrophy. A randomized, controlled trial defining the time course and dose response. Clinical Investigation of Duchenne Dystrophy Group. Arch Neurol. 1991;48(4):383–8.

2. Holford NH, Peace KE. Methodologic aspects of a population pharmacodynamic model for cognitive effects in Alzheimer patients treated with tacrine. Proc Natl Acad Sci U S A. 1992;89(23): 11466–70.

3. Holford NH, Peace KE. Results and validation of a population pharmacodynamic model for cognitive effects in Alzheimer patients treated with tacrine. Proc Natl Acad Sci U S A. 1992;89(23):11471–5.

4. Sale M, Sheiner LB, Volberding P, Blaschke TF. Zidovudine response relationships in early human immunodeficiency virus infection. Clin Pharmacol Ther. 1993;54(5):556–66.

5. Holford NH, Peace K. The effect of tacrine and lecithin in Alzheimer's disease. A population pharmacodynamic analysis of five clinical trials. Eur J Clin Pharmacol. 1994;47(1):17–23.

6. US Food and Drug Administration. Innovation or stagnation: challenge and opportunity on the critical path to new medical products. 2004. http://www.fda.gov/ScienceResearch/SpecialTopics /CriticalPathInitiative/CriticalPathOpportunitiesReports /ucm077262.htm. Accessed February 19, 2016.

7. Woodcock J, Woosley R. The FDA critical path initiative and its influence on new drug development. Annu Rev Med. 2008;59:1–12.

8. Lalonde RL, Kowalski KG, Hutmacher MM, Ewy W, Nichols DJ, Milligan PA, et al. Model-based drug development. Clin Pharmacol Ther. 2007;82(1):21–32.

9. Gobburu JV, Lesko LJ. Quantitative disease, drug, and trial models. Annu Rev Pharmacol Toxicol. 2009;49:291–301.

10. US Food and Drug Administration. Disease Specific Model Library. http://www.fda.gov/AboutFDA/CentersOffices /OfficeofMedicalProductsandTobacco/CDER/ucm180485.htm. Accessed February 19, 2016.

11. Visser SA, de Alwis DP, Kerbusch T, Stone JA, Allerheiligen SR. Implementation of quantitative and systems pharmacology in large pharma. CPT Pharmacometrics Syst Pharmacol. 2014;3(10):e142.

12. Mould DR. Developing models of disease progression. Pharmacometrics. John Wiley & Sons, Inc.; 2007. p. 547–81.

13. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry. 1984;141(11):1356–64.

14. Gomeni R, Merlo-Pich E. Bayesian modelling and ROC analysis to predict placebo responders using clinical score measured in the initial weeks of treatment in depression trials. Br J Clin Pharmacol. 2007;63(5):595–613.

15. Shang EY, Gibbs MA, Landen JW, Krams M, Russell T, Denman NG, et al. Evaluation of structural models to describe the effect of placebo upon the time course of major depressive disorder. J Pharmacokinet Pharmacodyn. 2009;36(1):63–80.

16. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol. 1967;6(4):278–96.

17. Gomeni R, Fava M. Amyotrophic lateral sclerosis disease progression model. Amyotroph Lateral Scler Frontotemporal Degener. 2014;15(1–2):119–29.

18. Cedarbaum JM, Stambler N, Malta E, Fuller C, Hilt D, Thurmond B, et al. The ALSFRS-R: a revised ALS functional rating scale that

incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). J Neurol Sci. 1999;169(1–2):13–21.

19. van Schaick E, Zheng J, Perez Ruixo JJ, Gieschke R, Jacqmin P. A semi-mechanistic model of bone mineral density and bone turnover based on a circular model of bone remodeling. J Pharmacokinet Pharmacodyn. 2015;42(4):315–32.

20. Mould DR. Models for disease progression: new approaches and uses. Clin Pharmacol Ther. 2012;92(1):125–31.

21. Hansson EK, Friberg LE. The shape of the myelosuppression time profile is related to the probability of developing neutropenic fever in patients with docetaxel-induced grade IV neutropenia. Cancer Chemother Pharmacol. 2012;69(4):881–90.

22. Wang Y, Sung C, Dartois C, Ramchandani R, Booth BP, Rock E, et al. Elucidation of relationship between tumor size and survival in non-small-cell lung cancer patients can aid early decision making in clinical drug development. Clin Pharmacol Ther. 2009;86(2):167–74.

23. Claret L, Girard P, Hoff PM, Van Cutsem E, Zuideveld KP, Jorga K, et al. Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics. J Clin Oncol. 2009;27(25):4103–8.

24. Ahn JE, French JL. Longitudinal aggregate data model-based meta-analysis with NONMEM: approaches to handling within treatment arm correlation. J Pharmacokinet Pharmacodyn. 2010;37(2):179–201.

25. Ito K, Ahadieh S, Corrigan B, French J, Fullerton T, Tensfeldt T. Disease progression meta-analysis model in Alzheimer's disease. Alzheimers Dement. 2010;6(1):39–53.

26. Pan W-J, Gibbs M, Frame B, Mould DR. Model based meta analyses of disease metrics in patients with Crohn's disease. Annual Meeting of the American College of Clinical Pharmacology; Chicago, IL, 2011.

27. Mandema JW, Gibbs M, Boyd RA, Wada DR, Pfister M. Model-based meta-analysis for comparative efficacy and safety: application in drug development and beyond. Clin Pharmacol Ther. 2011;90(6):766–9.

28. Holford NH, Chan PL, Nutt JG, Kieburtz K, Shoulson I. Disease progression and pharmacodynamics in Parkinson disease—evidence for functional protection with levodopa and other treatments. J Pharmacokinet Pharmacodyn. 2006;33(3):281–311.

29. Vu TC, Nutt JG, Holford NH. Progression of motor and nonmotor features of Parkinson's disease and their response to treatment. Br J Clin Pharmacol. 2012;74(2):267–83.

30. Kimko HC, Reele SS, Holford NH, Peck CC. Prediction of the outcome of a phase 3 clinical trial of an antischizophrenic agent (quetiapine fumarate) by simulation with a population pharmacokinetic and pharmacodynamic model. Clin Pharmacol Ther. 2000;68(5):568–77.

31. Potkin S, Agid O, Siu C, Watsky E, Vanderburg D, Remington G. Placebo response trajectories in short-term and long-term antipsychotic trials in schizophrenia. Schizophr Res. 2011;132(2–3):108–13.

32. Kuan WL, Kasis A, Yuan Y, Mason SL, Lazar AS, Barker RA, et al. Modelling the natural history of Huntington's disease progression. J Neurol Neurosurg Psychiatry. 2015;86(10):1143–9.

33. Holford NHG, Mould DR, Peck CC . Disease progress models. Principles of Clinical Pharmacology (Third Edition). Academic Press; 2012. p. 369–79.

34. Ito K, Corrigan B, Zhao Q, French J, Miller R, Soares H, et al. Disease progression model for cognitive deterioration from Alzheimer's Disease Neuroimaging Initiative database. Alzheimers Dement. 2011;7(2):151–60.

35. Ito K, Corrigan B, Romero K, Anziano R, Neville J, Stephenson D, et al. Understanding placebo responses in Alzheimer's disease clinical trials from the literature meta-data and CAMD database. J Alzheimers Dis. 2013;37(1):173–83.

36. Landersdorfer CB, Jusko WJ. Pharmacokinetic/pharmacodynamic modelling in diabetes mellitus. Clin Pharmacokinet. 2008;47(7):417–48.

37. Ajmera I, Swat M, Laibe C, Le Novere N, Chelliah V. The impact of mathematical modeling on the understanding of diabetes and related complications. CPT Pharmacometrics Syst Pharmacol. 2013;2(7):e54.

38. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics. 2003;19(4):524–31.

39. Li C, Donizelli M, Rodriguez N, Dharuri H, Endler L, Chelliah V, et al. BioModels Database: an enhanced, curated and annotated resource for published quantitative kinetic models. BMC Syst Biol. 2010;4:92.

40. Peterson MC, Riggs MM. A physiologically based mathematical model of integrated calcium homeostasis and bone remodeling. Bone. 2010;46(1):49–63.

41. Peterson MC, Riggs MM. Predicting nonlinear changes in bone mineral density over time using a multiscale systems pharmacology model. CPT Pharmacometrics Syst Pharmacol. 2012;1:e14.

42. 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(3):e00020.

43. US Food and Drug Administration. Transcript for the September 12, 2014 meeting of the Endocrinologic and Metabolic Drugs Advisory Committee (EMDAC). 2014. http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/EndocrinologicandMetabolicDrugsAdvisoryCommittee/ucm386727.htm. Accessed July 15, 2016.

44. Alzheimer's Association. 2015 Alzheimer's disease facts and figures. Alzheimers Dement. 2015;11(3):332–84.

45. Schneider LS, Mangialasche F, Andreasen N, Feldman H, Giacobini E, Jones R, et al. Clinical trials and late-stage drug development for Alzheimer's disease: an appraisal from 1984 to 2014. J Intern Med. 2014;275(3):251–83.

46. 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(5):479–98.

47. Samtani MN, Farnum M, Lobanov V, Yang E, Raghavan N, Dibernardo A, et al. An improved model for disease progression in patients from the Alzheimer's disease neuroimaging initiative. J Clin Pharmacol. 2012;52(5):629–44.

48••. Conrado DJ, Denney WS, Chen D, Ito K. An updated Alzheimer's disease progression model: incorporating non-linearity, beta regression, and a third-level random effect in NONMEM. J Pharmacokinet Pharmacodyn. 2014;41(6):581–98 .**One of the most recently published disease progression models for Alzheimer's disease. Includes NONMEM code for the final model illustrating implementation of beta regression and utilization of the Richards function for the structural disease progression model.**

49••. Romero K, Ito K, Rogers JA, Polhamus D, Qiu R, Stephenson D, et al. The future is now: model-based clinical trial design for Alzheimer's disease. Clin Pharmacol Ther. 2015;97(3):210–4 .**Describes the efforts of the Coalition Against Major Diseases that successfully advanced an Alzheimer's disease clinical trial simulation tool through regulatory review in the US and Europe. Provides an overview of the stages of database construction and standardization, model development and evaluation, and regulatory review.**

50. Coalition Against Major Diseases. http://c-path.org/programs/camd/. Accessed February 19, 2016.

51. 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(10):1212–21.

52••. Samtani MN, Raghavan N, Novak G, Nandy P, Narayan VA. Disease progression model for Clinical Dementia Rating-Sum of Boxes in mild cognitive impairment and Alzheimer's subjects from the Alzheimer's Disease Neuroimaging Initiative. Neuropsychiatr Dis Treat. 2014;10:929–52 .**Extends upon previous work in mild-to-moderate Alzheimer's disease to model disease progression in patients with late mild cognitive impairment and mild Alzheimer's disease. Incorporates screening of a wide range of potential covariates, including demographic information, APOEε4 allele status, baseline cognitive test scores, disease state, comedication use, hippocampal volume, and serum and cerebrospinal fluid biomarkers.**

53. Regnier ED, Shechter SM. State-space size considerations for disease-progression models. Stat Med. 2013;32(22):3862–80.

54. Fronsdal K, Pichler F, Mardhani-Bayne L, Henshall C, Rottingen JA, Morland B, et al. Interaction initiatives between regulatory, health technology assessment and coverage bodies, and industry. Int J Technol Assess Health Care. 2012;28(4):374–81.

55••. van Hasselt JG, Gupta A, Hussein Z, Beijnen JH, Schellens JH, Huitema AD. Integrated simulation framework for toxicity, dose intensity, disease progression, and cost effectiveness for castration-resistant prostate cancer treatment with eribulin. CPT Pharmacometrics Syst Pharmacol. 2015;4(7):374–85 .**Proof-of-concept model for integrated evaluation of the influence of clinical and pharmacological factors on treatment efficacy and cost. Provides a framework that can be utilized at early drug development stages to synthesize knowledge about disease progression, clinical outcome, adverse events, quality of life, and cost effectiveness.**

56. van Hasselt JG, Gupta A, Hussein Z, Beijnen JH, Schellens JH, Huitema AD. Disease progression/clinical outcome model for castration-resistant prostate cancer in patients treated with eribulin. CPT Pharmacometrics Syst Pharmacol. 2015;4(7):386–95.

57. Soetaert K, Petzoldt T, Setzer RW. Solving differential equations in R: package deSolve. J Stat Softw. 2010;33(9):1–25.

58. Frederix GW, van Hasselt JG, Severens JL, Hovels AM, Huitema AD, Raaijmakers JA, et al. Development of a framework for cohort simulation in cost-effectiveness analyses using a multistep ordinary differential equation solver algorithm in R. Med Decis Mak. 2013;33(6):780–92.

59. Beyene J, Hamid JS. Longitudinal data analysis in genome-wide association studies. Genet Epidemiol. 2014;38(Suppl 1):S68–73.

60. Wu Z, Hu Y, Melton PE. Longitudinal data analysis for genetic studies in the whole-genome sequencing era. Genet Epidemiol. 2014;38(Suppl 1):S74–80.

61••. Wu K, Gamazon ER, Im HK, Geeleher P, White SR, Solway J, et al. Genome-wide interrogation of longitudinal FEV1 in children with asthma. Am J Respir Crit Care Med. 2014;190(6):619–27 .**Utilizes a disease progression model to facilitate a genome-wide association study. A nonlinear mixed-effects model was employed to characterize lung function over time, and a genome-wide association study identified genetic contributors to longitudinal phenotypes that were determined from the disease progression model.**

62. The Childhood Asthma Management Program Research Group. Long-term effects of budesonide or nedocromil in children with asthma. The Childhood Asthma Management Program Research Group. N Engl J Med. 2000;343(15):1054–63.