COALITION AGAINST
MAJOR DISEASES
(CAMD)

WORK SCOPE 1.1
JULY 2009
MEMBERS

- Abbott Laboratories
- Alliance for Aging Research
- Alzheimer’s Association
- Alzheimer’s Foundation of America
- AstraZeneca Pharmaceuticals LP
- Bristol-Myers Squibb Company
- Daiichi Sankyo
- Eli Lilly and Company
- F. Hoffmann La Roche, Ltd.
- Forest Research Institute
- Genentech, Inc.
- GlaxoSmithKline
- Johnson & Johnson
- National Health Council
- Novartis Pharmaceutical Corporation
- Parkinson’s Action Network
- Parkinson’s Disease Foundation
- Pfizer, Inc.
- sanofi-aventis, US, Inc.
- Schering-Plough
- Wyeth Pharmaceuticals, Inc.

NON-VOTING PARTICIPANTS

- U.S. Food & Drug Administration (FDA)
- European Medicine Agency (EMEA)
- National Institute on Aging (NIA)
- National Institute of Neurological Disorders (NINDS)

IN COLLABORATION WITH

- Engelberg Center for Health Care Reform at the Brookings Institution
# TABLE OF CONTENTS

**OVERVIEW** ...............................................................................................................................................3

**CHAPTER I: INTRODUCTION** .............................................................................................................4

CAMD Case Statement ..................................................................................................................................4
Collaboration ..................................................................................................................................................7
Background ..................................................................................................................................................7
Initial Workgroups .......................................................................................................................................10
Workflow Among Workgroups .....................................................................................................................11
Structure of Workgroups ...............................................................................................................................12

**CHAPTER II: PROPOSED WORKGROUPS FOR RESEARCH** .................................................................13

Workgroup 1 – Data ......................................................................................................................................13
  Subgroup A: Data Sources, Standards and Sharing ..................................................................................14
  Subgroup B: Data Infrastructure ..............................................................................................................19
Workgroup 2 – Disease-Progression Modeling ..........................................................................................26
  Process to Determine/Compile Data Needs .............................................................................................28
  Quantitative Neurodegenerative Disease Model Creation ...................................................................29
Workgroup 3 – Biomarker Evaluation .........................................................................................................32
  Process .....................................................................................................................................................32
  Deliverables ...........................................................................................................................................34
Workgroup 4 – Health Authorities Submissions .......................................................................................35

Alliances with Patient Communities ..........................................................................................................36

**CHAPTER III: SUPPORT AND COMMUNICATIONS TEAM** ................................................................37

Administration and Project Management ..................................................................................................37
Announcements and Ongoing External Communications ...........................................................................38

**APPENDICES** ......................................................................................................................................40

Appendix A: Glossary of Terms and Acronyms .......................................................................................40
Appendix B: Overview and Purpose of the CAMD Legal Agreement .......................................................42
Appendix C: CAMD Participant Roles ......................................................................................................44

**REFERENCES CITED** .............................................................................................................................45
OVERVIEW

The goal of the Coalition Against Major Diseases (CAMD) is to bring together major pharmaceutical companies, the U.S. Food and Drug Administration (FDA), the European Medicine Agency (EMEA), the National Institute on Aging (NIA), the National Institute of Neurological Disorders and Stroke (NINDS), and patient groups in a collaboration to develop new knowledge that will enhance the industry’s ability to develop innovative new therapies. CAMD will focus first on Alzheimer’s and Parkinson’s diseases and then expand to other diseases.

This work scope defines the role of the initial four workgroups that will generate new knowledge resulting in tools to improve the medical product development process. The major deliverables of the coalition are:

- To submit biomarkers to the FDA for qualification to accurately diagnose disease, stratify patient populations, and predict patient outcomes.
- To submit multiparameter models of disease progression to the FDA for qualification that can be used to project the effects of potential diagnostics and treatments, as well as influence the design of clinical trials.
- To develop an integrated database from completed trials in a common Clinical Data Interchange Standards Consortium (CDISC) standard format usable for research by coalition members and others.

A workgroup has been created to work on each major deliverable. A fourth workgroup will be formed to assist in the creation of the dossiers for submission to the FDA.

It is not the intent of the coalition to duplicate current efforts already under way in these areas, but instead to leverage existing data and knowledge, create consensus on methods to advance product development, and make the methods available for broad applications. Where appropriate, the resulting applied data and new information will be submitted for FDA review, with the goal to have them qualified, and in all cases, to have them widely available for use in new medical product development.

The coalition is being founded and supported by the Critical Path Institute (C-Path) in collaboration with the Engelberg Center for Health Care Reform at the Brookings Institution. CAMD is a self-governing entity advised by scientists from the FDA, EMEA, and the National Institutes of Health (NIH) and directed by its members, who are pharmaceutical and biotech companies and patient groups committed to advancing the care of patients with neurodegenerative diseases.
One of the greatest challenges facing biomedical sciences in the 21st century is the development of fundamentally better treatments for neurodegenerative diseases. The two most prevalent of these, Alzheimer’s disease and Parkinson’s disease, exert a heavy and rapidly growing human and economic burden on our society. Our lack of knowledge about the specific cause(s) of either disease is a major obstacle to the development of treatments that would have the potential to either cure or prevent.

In order to replicate the successful process that was used to develop HIV/AIDS drugs during the 1980s, the Critical Path Institute (C-Path), in collaboration with the Engelberg Center for Health Care Reform at the Brookings Institution, has formed the Coalition Against Major Diseases (CAMD), which includes advisors from the FDA, EMEA, NIH (NIA and NINDS), and academia. Members include patient groups and the medical products industry. A coordinating committee with representation from each member organization will direct, prioritize, coordinate, and oversee the work of the coalition. CAMD will focus first on Alzheimer’s and Parkinson’s diseases and then expand to other diseases with significant public health implications.

The initial, very ambitious goal of CAMD is to establish a common research support infrastructure for using pooled control or placebo patient data from clinical trials to create quantitative disease-progression models for both Alzheimer’s disease and Parkinson’s disease. These models will utilize already defined data standards when possible, and the coalition will develop new data standards when gaps exist. Coalition members also will work together to incorporate data on imaging and biochemical and molecular biomarkers that have the greatest potential to identify those patients with the maximum likelihood of deriving benefit from and/or the likelihood of harm from specific therapies. Importantly, where appropriate and useful, CAMD members will collaborate to assemble, evaluate, and submit the evidence supporting requests for the FDA to designate such tools as qualified for use in drug development. These newly qualified tools then will be made publicly available for all scientists and commercial developers to utilize.

The CAMD Coordinating Committee, with 21 members represented, met on September 22, 2008, in Washington, D.C., and authorized C-Path to enroll additional participants and to draft a detailed work scope of activities to be executed over the next few years. By January 2009 the draft work scope was completed and began with the creation of four major workgroups that will be supported by a team focused on providing the information technology infrastructure and the interorganizational communication necessary for effective collaboration.

It is not the intent of the coalition to duplicate current efforts already under way in any of these areas, but instead to gather, integrate, share and leverage existing data and knowledge by making them available for broad application and to have the resulting analytical tools and processes qualified by the FDA where appropriate, so they may be widely utilized in new medical product development.
Alzheimer’s and Parkinson’s diseases have been chosen initially because of their prevalence, impact on morbidity and mortality, extensive prior data available, and current investment by academia, government and industry for better disease understanding and interventions for prevention or better treatment. The overall roadmap for the CAMD activities is shown in the following diagram (Figure 1).

The execution phase of the coalition began in early 2009, with a meeting of the coordinating council to launch the initial workgroups. The evidence will be evaluated by the working groups and submitted to the FDA as voluntary exploratory data submissions (VXDS). Once an internal consensus is reached within CAMD that there is adequate evidence to support the qualification of a given biomarker, a Biomarker Qualification Review Team (BQRT) will be formed at the FDA to review the submission package for qualification. Acceptance by the health authority agencies will mean that the new methods are then qualified for use in drug development in a defined manner.

The new methods and tools from CAMD are expected to have a significant effect on the drug development process. Today, most drug candidates are identified because they have demonstrated some measurable change in a laboratory model that is considered relevant to a disease. This lab assay, if it can be linked to patients with the disease, often has the potential to be a useful biomarker during drug development and perhaps in clinical practice. Biomarkers can identify subsets of patients with a disease who have distinct patterns of progression or outcomes. In CAMD, data integration and sharing are planned to create a quantitative disease-progression model that includes biomarkers that potentially identify discrete patient subsets of the disease. Increasingly, a “disease model” and a “drug model” are integrated, and modeling and simulation are used to simulate and design clinical trials with a greater chance of success. The drug model includes information about its pharmacokinetics, pharmacodynamics, and the patient-specific factors that influence its actions (e.g., CYP isom form metabolism). Disease models enable greater accuracy in predicting the occurrence of clinically significant events and the influence of potentially confounding factors.
CAMD will address the need for a more reliable and predictable process of drug development for these diseases by creating scientific consensus for new tools and methods qualified by the FDA for use in drug development. As shown below in Figure 2, inputs for creating the disease model will come from many established and new sources. The disease model and biomarkers will be reviewed by the FDA and “qualified for use” where possible, and will become part of the drug development process. The combination of qualified biomarkers and quantitative disease models will provide important tools for the pharmaceutical industry to use to identify potential new therapies, to learn more about when to more quickly terminate candidate drugs that have a low probability of success, and to ensure that when a new therapy is found, it has a much higher probability to move through the development and regulatory system successfully and more rapidly.

CAMD also will provide a framework for continuous learning about the diseases because the database and the disease models will be enriched as new information becomes available. Furthermore, the experience with Alzheimer’s and Parkinson’s disease will be readily expanded to a systematic exploration for many other important diseases.

Figure 2. CAMD’s contribution to integrated drug development.
COLLABORATION

There is a great deal of research under way today in both Alzheimer’s and Parkinson’s diseases. The Alzheimer’s Study Group recently published a report on the challenges of effectively utilizing all of this new science to generate effective therapies for patients. The CAMD staff, as part of the development of this Work Scope, has maintained close communication with those organizations involved in active areas of research. The Biomarker Consortium, Alzheimer’s Disease Neuroimaging Project (ADNI), Parkinson’s Study Group (PSG), and Alzheimer’s Disease Cooperative Study (ADCS) have all been contacted and invited to collaborate with CAMD.

CAMD’s initial work involves the sharing of clinical trial information by our corporate members to generate the core information needed to build new disease models that represent populations. In addition, we will also seek to incorporate data from concurrent research by others to supplement these disease models, as well as to enrich these models with priority biomarker information from other research efforts. The active participation of the FDA and their commitment to review data for possible qualification make CAMD novel and especially important. We will seek qualification for these models and biomarkers and place them into the public record so everyone can use them to enhance research and drug development for these diseases.

BACKGROUND

ALZHEIMER’S DISEASE

It has been more than 100 years since Dr. Alois Alzheimer first described the case of Auguste D., a patient with rapidly failing memory, confusion, disorientation, trouble expressing her thoughts, and unfounded suspicions about her family and hospital staff. Today, Alzheimer’s disease (AD)—the most common form of dementia—affects 4.6 million new patients worldwide each year. There are 5.1 million cases in the U.S., most of them receiving care under Medicare. By 2030, the number of Americans 65 and older with AD will have grown by 50 percent to 7.7 million. AD is estimated to afflict about 10 percent of people over age 65 and 30 to 50 percent of those over age 85. As a cause of death, AD grew by 33 percent from 2000 to 2004, compared to declines in the percentages of deaths caused by heart disease, stroke, and breast cancer. The direct and indirect costs of AD and other dementias, including Medicare and Medicaid costs and the indirect cost to employers of caregivers, is more than $148 billion annually in the U.S. For the 10 million Americans caring for a person with AD or other dementia, the annual burden (in terms of reduced productivity and lower health status) has been estimated at $60 billion.

In the past 20 years, more than 300 drugs have entered testing for AD, yet only five have been approved in developed countries. While these five drugs have some impact on symptoms for patients who have already developed signs of disease, they do not fundamentally alter its course, and their cost-effectiveness has been questioned. Though effective AD drugs are badly needed, progress has been difficult.
There is currently a worldwide effort to identify and validate biomarkers of AD. The identification of biomolecules—proteins, genes and their pathways—combined with advanced imaging technologies, promises to inform on risk, prevention, identification of treatment targets and response, and disease progression. It is believed that AD is a heterogeneous disease, comprising several subtypes, each of which may have different etiopathogenetic mechanisms leading ultimately to a similar pathology. In addition, the two anatomical hallmarks of AD, amyloid plaques and neurofibrillary tangles, are now known, respectively, to be associated with aging in the absence of dementia and are found in other disease states as well as AD.

For these reasons, it is highly likely that panels of biomarkers, combined with imaging technologies, will be needed to speed progress in accurate early diagnosis and successful treatment strategies. The real push is toward development of antecedent markers to enable early detection or prevention of disease, based on the hypothesis that biochemical changes occur far in advance of cognitive decline and brain damage.

To date, four genes associated with an increased risk of AD have been identified. Mutations in presenilin 1 (PS1) and presenilin 2 (PS2) are thought to cause early-onset AD. Mutations in genes expressing amyloid precursor protein (APP) associated with β- and γ-secretase are another rare cause of early-onset familial AD. Mutations in PS1, PS2 and APP that are associated with AD all result in increased β-amyloid-42 (Aβ42) production. These mutations, however, are rare, and no genes to date have been found to be associated with the common sporadic form of AD. However, certain alleles of the apolipoprotein E (ApoE) gene have been associated with risk of disease; the E4 allele with increased risk, and the E2 allele with decreased risk.

Over the last decade, biomarkers derived from cerebrospinal fluid (CSF) were shown to correlate with pathogenic processes in AD and have high potential as diagnostic markers. The combinations of elevated total Tau (t-Tau), or phosphorylated Tau proteins (p-Tau), and low β-amyloid-42 (Aβ42) are currently the only CSF biomarkers with high sensitivity and specificity for differentiating early-stage AD from other dementias. Their stability in individual patients over time also makes them promising markers for monitoring treatment response in clinical trials with potential disease-modifying drugs. Other research suggests that levels of P-Tau and isoprostanes combined with imaging (MRI or FDG-PET) may be more sensitive in assessing drug interventions. On the horizon are additional proteins such as visinin-like protein 1 (VLP-1), a brain injury marker, and the enzyme BACE1. In a recent study, CSF VLP-1 concentrations were shown to correlate with tau proteins and with scores on the Mini-Mental State Examination, a standard clinical diagnostic tool. When combined with Aβ42 and tau proteins, VLP-1 improved diagnostic accuracy. Similarly, BACE1 correlates with beta amyloid and the ApoE4 allele, and when combined may be useful as a predictive panel of disease markers.

An important goal is to develop noninvasive early diagnostic markers. While confirmatory trials are needed, CD69 values in white blood cells allowed researchers to differentiate between patients with Alzheimer’s disease and those with Parkinson’s disease with greater than 90 percent accuracy. The measure also distinguished normal subjects from patients with AD with 88 percent sensitivity, and 82
percent of the time when the individuals had no cognitive deficits. Several additional blood-based biomarkers for predicting onset of disease and diagnosis, including Aβ42, APP, and BACE, are currently in development.

Neuroimaging (MRI and PET) technologies and markers also have received much attention and research effort. The ability to produce images of amyloid in the brains of living people represents a great advance in diagnosing Alzheimer’s disease. In addition, imaging can help confirm the value of surrogate markers and may serve to assess the efficacy of drug interventions. While there are some limitations due to the half-life of some radiotracers, requiring them to be made on-site, improved tracers that could be used in more community settings are being developed.

**PARKINSON’S DISEASE**

Parkinson’s disease (PD) is a chronic degenerative disorder of the central nervous system. British physician James Parkinson first described it as “the shaking palsy” in 1817. The U.S. prevalence of PD is estimated at one million cases, second only to AD among neurodegenerative diseases. The average age of onset is 60 years, though PD can strike adults at any age. The total cost to the nation is estimated to exceed $25 billion annually. The risk of PD increases with age, so analysts expect the financial and public health impact of this disease to increase as the population ages.

Parkinson’s disease is characterized by specific movement disorders, including tremor, rigidity, slow movement, and postural instability. These symptoms usually begin gradually and worsen with time. As they become more pronounced, patients may have difficulty walking, talking, or completing other simple tasks. Advanced PD patients also may suffer from dementia, a disconcerting, but perhaps informative, nexus between these two disabling diseases.

The specific cause of PD remains unknown. PD occurs when neurons in an area of the brain known as the substantia nigra die or become impaired, reducing production of dopamine—a neurotransmitter required to produce smooth, purposeful movement. Recent studies have shown that people with PD also have loss of the nerve endings that produce norepinephrine, which may explain the nonmotor features of PD, including fatigue and abnormalities of blood pressure regulation. Many brain cells of people with PD contain Lewy bodies—unusual deposits of proteins—but researchers do not yet know what role they play in development of the disease. Several genetic mutations have been associated with PD, and many more genes have been tentatively linked to the disorder. Although the importance of genetics in PD is increasingly recognized, most researchers believe environmental exposures also increase a person’s risk of developing the disease. Certain toxic chemicals, trauma, and some viruses are known environmental triggers for PD. At a cellular level, mitochondrial dysfunction, oxidative stress, inflammation, and many other processes may contribute to PD, but the actual cause of the dopamine cell death for most patients is still undetermined. Unmet clinical needs thus include the validation of biomarkers and imaging techniques, such as 18F-dopa PET, for their ability to better discriminate disease types, leading to a better understanding of etiology, and ultimately, treatment.
In the 1990s, there was optimism among patient advocates and researchers that emerging science would enable PD to be cured in as little as five to ten years. In 2000, the NIH issued an ambitious Parkinson’s Disease Research Agenda that would leverage new research capabilities, such as high-throughput drug screening, array technologies, and tissue repositories, to better understand PD and to develop new pharmacological, surgical, cell implantation, gene, and rehabilitation therapies. However, that optimism has not yet turned into the needed effective new therapies. Of the five new drugs approved for PD by the FDA since 2004 (rasagiline, rotigotine transdermal, ropinerole extended release, selegiline dissolving tablets, and carbidopa/levodopa rapidly dissolving tablets), only rasagiline is a novel compound—the other four are reformulations of existing products. Current treatments for PD, like those for AD, are to relieve symptoms, rather than reverse or prevent underlying causes, and they lose their effectiveness over time.

INITIAL WORKGROUPS

This work scope document describes the processes that will guide activity to begin in early 2009. In this document, a number of workgroups are identified to conduct the tasks related to various projects. They are listed below, and others may be created as necessary. Finally, a team will provide logistical and communications support for the coalition.

**Workgroup (WG) 1: Data**
- Compile data identified by Workgroups 2 and 3
- Establish data standards and data remapping
- Provide data management infrastructure

**Workgroup (WG) 2: Disease-Progression Modeling**
- Determine clinical trial data requirements
- Create quantitative disease models
- Determine data requirements for inclusion of electronic health records

**Workgroup (WG) 3: Biomarker Evaluation**
- Determine biomarker and imaging data requirements
- Select biomarkers to subset patients in the model

**Workgroup (WG) 4: Health Authorities Submissions**

**Support and Communications Team**
WORKFLOW AMONG WORKGROUPS

WG 2 (Disease Progression Modeling) and WG 3 (Biomarker Evaluation) will identify for WG 1 (Data) what data it will need and will assist with compiling these data. WG 1 will request the data from the member companies, NIH, FDA (if feasible and as permitted), etc.; standardize/remap the data; put the data into an infrastructure; and make the data available to WG s 2 and 3 as the data are ready. WG 2 will define the data needed for model building and work with the data as they become available to create
a disease model (empirical now - mechanistic at first, based on clinical trial data), while WG 3 works with the data to evaluate and identify the most promising biomarkers, including imaging. Candidate biomarkers identified by WG 3 will be referred back to WG 2 for the development of more sophisticated mechanistic disease models. Once biomarkers and disease models have been adequately developed with sufficient evidence, WG 4 will submit them to the FDA for consideration and possible qualification.

**STRUCTURE OF WORKGROUPS**

The current structure of the workgroups is shown above. Each of the three initial workgroups has a direct liaison with the coordinating committee. The workgroups are led by two co-chairs, one from C-Path and the other from an industry member. Workgroups 2 and 3 have formed separate Alzheimer’s and Parkinson’s disease teams in order to focus the knowledge and expertise in the distinct areas.
CHAPTER II: PROPOSED WORKGROUPS FOR RESEARCH

WORKGROUP 1: DATA

OVERVIEW

The Data Workgroup will focus on enabling the efforts of the Biomarker Evaluation and Disease-Progression Modeling workgroups. This will be accomplished by working with these groups to understand their needs and priorities for information and then working with the member companies to compile this information from the source data. This section addresses the technical requirements of gathering, assessing, standardizing, and pooling disparate sources of clinical and laboratory data into an integrated database. It describes the work needed to develop standards for remapping and integrating data from a variety of sources and formats into a common format based on open, consensus-based standards where feasible (e.g., CDISC and Health Level 7 [HL7]). It also addresses the architecture specifications: how all pieces of the coalition’s information technology (IT) interact to integrate, submit, store and access the data for analysis in either a centralized or federated system.

OBJECTIVES

The Data Workgroup will address two primary objectives, possibly via separate collaborative subgroups: 1) to convert all data domains and individual data elements requested by the Disease-Progression Modeling and Biomarker Evaluation workgroups to a standard usable format to populate the integrated database, and 2) to provide the infrastructure/architecture of the entire IT system for the coalition.

In support of these primary objectives, the Data Workgroup will make linkages to key personnel within the member companies, FDA, EMEA, NIH, NIH-funded efforts, academia, patient groups, and other consortia to collect the information and data that are relevant to the scope of CAMD. This workgroup will develop data usability criteria in consultation with the Disease-Progression Modeling Workgroup, and will use these criteria to assess fitness of candidate source data for inclusion in the project. The Data Workgroup will work closely with the other workgroups to understand their evolving data needs and to anticipate future needs so data conversion can occur in a timely manner. The Data Workgroup will need to address quality control throughout the process of data conversion and will coordinate, when applicable, submission of data to FDA for qualification.

In consideration of developing the infrastructure, the Data Workgroup will evaluate and recommend potential hardware and technology products and vendors for their ability to meet the security and functionality requirements of CAMD. Additionally, in consultation with the Disease-Progression Modeling Workgroup, the Data Workgroup will assist in decisions as to whether or not to acquire any available commercial off-the-shelf (COTS) analysis tools, or to develop custom analysis tools that will fit the requirements of the project.
SUBGROUP A: DATA SOURCES, STANDARDS AND SHARING

Overview
This subsection describes the first role of the Data Workgroup (WG 1): developing technical standards for remapping and integrating disparate sources of clinical data. An overview of the workflow is presented in Figure 3, and deliverables and resources required for meeting timeline goals are outlined below.

Introduction
When considering the various data needs of the coalition, it is important to distinguish between data requirements of the disease models, the standards for integrating the data, the requirements for data storage infrastructure, and the tools for data analysis. The primary objective of the data team described in this section is to compile and convert source data into a standardized format for semantic interoperability and integration in the disease model database. This team will also play a significant role in the infrastructure considerations addressed in the next section. The data requirements of the disease model refer to those individual elements or domains deemed necessary by the Disease-Progression Modeling and Biomarker Evaluation workgroups to draw meaningful scientific conclusions, and should therefore be considered deliverables of those workgroups. It will be important to establish best practices standards for both data mining/modeling and reporting models. These standards will likely evolve over time, but they are important considerations in setting the framework for how tasks will be accomplished.

The Data Workgroup will work with national standards organizations, such as the Clinical Data Interchange Standards Consortium (CDISC) and Health Level 7 (HL7), to ensure that its efforts are neither redundant nor contradictory to the standardization efforts currently being developed and implemented by the FDA and in the pharmaceutical and, if possible, healthcare industries. The workgroup will work with the FDA to ensure all data for CAMD are formatted in a manner that is accessible by the agency.

Specific Aspects
A broad overview of the workflow among the workgroups is presented in Figure 3. The Disease-Progression Modeling Workgroup’s process for constructing models is iterative; each model produced will build on the previous models as more conclusions are drawn, more hypotheses are generated, and more types of data are included. As the Disease-Progression Modeling and Biomarker Evaluation workgroups identify new or more complex candidate processes, tools and biomarkers, they will request the representative data required for modeling these factors from the Data Workgroup, which then will produce these data in a usable format. Considering that the usability of preexisting data will limit all future modeling activities, an essential early and ongoing task of the Data Workgroup will be to establish a process for assessing the usability of clinical study data and study acceptance in a prospective manner.

The Data Workgroup also will need to consider that while the first sources of data for the Disease-Progression Modeling Workgroup will be limited to clinical trial data from the member companies, additional sources of data such as the ADNI (Alzheimer’s Disease Neuroimaging Initiative) will be used to enrich the models as the work progresses.
As the complexity of the models increase, the scope of contributing data will also increase. Figure 5 below illustrates how the various sources of data build on the previous sources as data from NIH clinical studies and consortia (e.g., ADNI), available biomarkers data, and ultimately, personal health record and electronic health record data that are incorporated into the model. This last source of data is a later-phase goal but is worth pursuing since the data obtained may help corroborate the clinical observations and the conclusions drawn by the model, and may verify their relevance to the general population. Additionally, the electronic medical record and personal health record data also may help the CAMD effort serve as a valuable patient resource and a clinical trial recruitment tool.

The approach of converting only the requested data for a given exercise is the most efficient since it allows the minimum set of data elements or domains to undergo the significant task of conversion, reducing the required effort and resources. Conversion of entire bodies of data from the available studies is an enormous task that is likely unnecessary; many of the data collected in these studies will not be required by the models. Regardless of the chosen format, the workload is significant, so workload alone should not be a factor in terms of choosing the specific new standard. The first modeling goal of CAMD will be to create a relatively basic empirical model founded on relatively fewer clinical domains and clinical data elements than are represented in an entire clinical trial. The Disease-Progression Modeling Workgroup also will request basic demographic information and a variety of other supporting data. For much of this data, there are likely to be existing CDISC standards. The Data Workgroup will work with its CDISC experts, and, when necessary, CDISC leaders to determine how best to fit CAMD-requested data into Study Data Tabulation Model (SDTM).

CAMD will submit models and real data to the FDA for assessment and qualification, and by using CDISC and HL7 standards, the coalition can ensure that the FDA will be able to receive and work with the submissions and with the standard interactive analytical tools being implemented. Additionally, companies currently conducting studies will benefit from knowing that the proactive conversion of data will help them respond to any potential regulatory requests that may come years after the study is submitted. Considering that CAMD efforts will continue for many years, it is inevitable that the coalition will eventually receive data from its members in SDTM format. It would be unreasonable to expect that these future data should be retrofitted to some other outdated standard used by the coalition. Finally, remap-
ping existing data to CDISC standards is becoming a prescribed process. The chance of finding metadata-driven and fully auditable COTS analytical tools, in-house expertise, and/or qualified consultants to assist in the process is therefore increased.

Whether or not new CDISC domain standards for Parkinson’s disease and Alzheimer’s disease will need to be created, or whether the existing standards can accommodate the disease model data requirements remains to be seen as the complexity of the models grows. Most likely, as new candidate biochemical and imaging biomarkers emerge, new therapeutic-area standards to accommodate these elements will need to be created. CAMD will not delay modeling progress while awaiting final development of these new domain standards (assuming new domains are necessary). Therefore, it will be necessary to keep CDISC abreast of coalition activities as it moves forward to avoid contradictory or duplicated efforts in the future. It is also important to note that the defined CDISC process for developing new disease-specific domains involves assembling subject-matter experts, many of whom likely will have been involved in the CAMD effort, thereby reducing the chance of conflicting efforts.

Realistic detailed estimations of time and resource requirements to achieve the goals above are not feasible without first knowing the quantity and current state of the source data that will need to be converted and integrated. However, several generalizations can be made and extrapolated by applying the examples relayed to us by experts who have embarked on similar efforts. Generally speaking, the amount of effort required in converting the data depends on answers to the following questions:

- How many clinical studies need to be remapped?
- How many different sponsors are involved?
- How many domains need to be converted to the new standard?
- How many standard variables are needed?
- How many standard variables are needed?
- Will new domains need to be created?
- What degree of terminology mapping is required?
- How well documented are the available study data?
- How accessible are the people who are familiar with the data and the study protocol?
- What is the current state and degree of standardization of legacy datasets to be transformed?
- How useful are the analysis tools to detect errors in the data (date variables containing dates, plus other foreign data from shifted columns)?

The largest effort involved in these undertakings will be the preliminary metadata analysis. Therefore, the last two items above touch upon resources that are germane to the conversion project. The FDA and CDISC recently engaged in a similar effort—The Integrated Pilot Database Project—wherein it was attempted to integrate data from 29 clinical studies involving 8 clinical domains, 8 sponsors, and a total of 13 compounds. This work, the first of its kind, took 4 FTEs (full-time equivalent employees) and nearly 6 months to complete, and only 8 of the original 29 studies were ultimately integrated. Other estimates from CDISC “registered solution providers” indicate that a typical conversion of ~40 source data domains, resulting in 20 to 25 SDTM domains, takes approximately 4 weeks for a well-documented, more recent study, and up to 3 months for older studies. In their estimation, the task requires 4 “roles” (though
not necessarily 4 FTEs), which they define as job developer, project manager, statistician, and a clinical lead who is familiar with the data, as well as an expert medical reviewer to assess the soundness of the integrated data. The work of this reviewer will be facilitated by having interactive access to graphic displays of “Patient Profile Views” of individual patient data and graphic displays of aggregated data, including the “Napoleon’s March” graphic display.

Given the above considerations, it is realistic to expect that within the first 5 months, the Data Workgroup will assemble, agree upon a standard, and inventory/characterize the first sets of source data to meet the modelers’ needs for the empirical model based on clinical data elements. Implementation of an error-checking tool early in the process will minimize finding problems at the end of the analytical process. However, allowing time for team and workflow normalization, and the possibility of encountering some poor-quality data, it is reasonable to allow another 3 to 6 months for the conversion of the first set of requested domains and their integration into the disease-model database. In other words, within the first 8 to 11 months, the coalition can expect to have the necessary data to create its first empirical model, based on an initial dataset residing in a shared infrastructure. It is important to note that more aggressive timelines can be achieved by taking leaner approaches to documentation and validation, relying instead on COTS validation tools. The feasibility of this will be discussed with the FDA. In addition, the degree of commitment from participants can drastically affect the timeline. For optimum progress, a small team of people working nearly full time in the initial stages is ideal. CAMD will consult with the appropriate experts to ensure the coalition can meet its goals with quality and timeliness.

Within two years of launch, the coalition can expect to have gained economies of scale. A well-established process for categorizing and remapping source data should be in place, and the Data Workgroup can be expected to be proactive and responsive to the needs of the Disease-Progression Modeling Workgroup to deliver the data it requires. Specific optimized analytical tools may need to be created to do additional assessments. Also, by this time (if not sooner) a more robust infrastructure should be in place with suitable data-storage solutions, and GUI (graphical user interface) portals to the data and analysis tools that have been chosen by the Disease-Progression Modeling Workgroup will be integrated into the infrastructure. Within five years, any new CDISC domains that were necessary will have been long completed, the coalition will have progressed through the hierarchy of data sources to include EHR (electronic health record) data, and it will possibly be accepting PHR (personal health record) data. The database may be serving as a clinical trial recruitment tool for interested patients by this time.

**Deliverables, Considerations, and Resource Requirements**

I. Data deliverables
   A. Survey of currently available data, including a list of member companies that have generated or are generating study data, broken down by:
      1. Number of studies per disease (AD and PD)
      2. Types of studies by disease (phase II, III, failed studies, etc.)
      3. Number of investigators and number of patients per investigator
   B. All final, annotated blank CRFs (case report forms) from candidate studies
   C. All final approved protocols from candidate studies
D. Final raw clinical datasets that match the annotated CRFs (even though not all data will be converted)
E. All diagnostics and supplemental data (labs, ECG, ePRO, etc.)
F. Possibly useful: analysis data sets used for clinical study reports (CSRs)

II. Data-assessment tasks
A. Verify all datasets can be accessed and are not corrupted
B. Develop and implement criteria for assessing study data of value to the project
   1. Data team role: criteria for quality of data with regards to technical standards
      (flag suspicious data; do not attempt cleaning at this stage)
   2. Modeling team role: communicate criteria for scientific utility of data (small number, dropout rates, trial duration, etc., that may render data undesirable)
C. Evaluate and understand the protocols; metadata analysis

III. Data conversion
A. Personnel requirements (roles)
   1. Of the coalition
      a. Facilitator(s) to coordinate the activities among the various contributing groups
      b. Administrative support staff
   2. Of each contributing group
      a. Project manager - preferably experienced in data conversion
      b. Lead SDTM analyst
      c. SDTM programming leader (SAS programmer—works with lead analyst to communicate and resolve inconsistencies)
      e. Statistician (familiar with each study to be converted)
      f. Clinical lead (familiar with the study objectives)
      g. Any support staff required, to be determined by each group
B. Special considerations
   1. Other possible sources of collaborative data sharing
      a. ADNI/LONI (Laboratory of NeuroImaging)
      b. Other NIA databases
      c. PD-DOC (Parkinson’s Disease Data and Organizing Center)
      d. EHRs
      e. Others (ADCS, REGARDS Study, etc.)
   2. Overall workflow considerations
      a. Each member company converts its own data before submitting to CAMD, or
      b. A coordinated, shared resource work plan to convert all contributed data will be established.
SUBGROUP B: DATA INFRASTRUCTURE

The data life cycle needs to accommodate the varied activities to support transformation of data into information and knowledge, accounting for future data modalities and complex analysis.

The data model and database architecture must accommodate individual contributing sites and must include workflows that can integrate with the downstream analysis. This section describes the robust infrastructure meeting these criteria. It is important to note that less sophisticated systems can and should be used in earlier stages of the CAMD effort. See Figure 6 to the right for an outline of data infrastructure goals.

The data infrastructure must be designed to support:

- Efficient data infusion: receipt and ETL (extract, transform and load) operations
- Quality assurance steps before incorporating the data into the repository
- Partitioning and mapping data for integrity and quality benchmarking
- Data integration across disparate sources through the use of standards—with extensibility in mind as new standards emerge
- A mapping of data provenance and dating, incorporated with tools for data-quality rating and archiving features for old or infrequently used data sets
- Risk avoidance via proper backup and recovery procedures
- Security within the transmission, storage, and manipulation of data
- Efficient structure for data analyses and reporting
- Workflow, visualization, and collaboration tools for consumers of the data
- An ability to integrate with other external sources and repositories for greater knowledge aggregation and broader data analysis
- Information in multiple standard-compliant formats for data analytics and data-mining activities
- Sharing both directly derived and inferred content
- Monitoring for availability, performance, and security
Our proposed basic model for the receipt and infusion of data from disparate sources into a single, cohesive data warehouse is shown in Figure 7 below.

This model assumes source data are coming from multiple sources (e.g., pharmaceutical companies “A” through “Z” and others). The pertinent concepts of the data flow of this model are shown in the table on the following page.
Though it is extremely important to capture quality data in a consistent model, of equal importance is the ability for coalition members to access, view, and manipulate that data in a reliable, secure manner once it has been consolidated.

<table>
<thead>
<tr>
<th>Step No.</th>
<th>CAMD Data Consolidation Flow Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The model allows for the likelihood that multiple studies have been performed at multiple pharmaceutical companies and at various times over the years, so the associated data conform to no prearranged standard at this step.</td>
</tr>
<tr>
<td>2</td>
<td>The data are currently held in internal, proprietary formats.</td>
</tr>
<tr>
<td>3</td>
<td>The onus is placed upon the participating coalition members to transform the data or have them transformed into a commonly accepted format, or one of multiple allowable formats. The data format will be prescribed by CAMD and is likely to be CDISC (see section on Data Sources, Standards, and Pooling for greater detail).</td>
</tr>
<tr>
<td>4</td>
<td>Data will be transferred from each participating CAMD contributor across a wide-area network (WAN) in an encrypted manner. The WAN may be the Internet, or a set of point-to-point connections. The transfer mechanism has not yet been determined but will be set by CAMD standard and mutual agreement. The mechanism may be push or pull, and the data transfer standards may use secure file transfer protocol (FTPS), SOAP, and/or HTTPS.</td>
</tr>
<tr>
<td>5</td>
<td>A set of protection mechanisms is suggested to secure the perimeter of the CAMD/C-PATH infrastructure. These mechanisms will include firewall(s), intrusion detection/prevention systems, and forward and reverse proxies. These mechanisms will be configured to allow the transfer of data from authorized CAMD partners.</td>
</tr>
<tr>
<td>6</td>
<td>Incoming data will be received in a data store tuned for the rapid receipt and writing of data, likely an Online Transaction Processing (OLTP) structure.</td>
</tr>
<tr>
<td>7</td>
<td>Incoming data will be validated before being written to the data repository. The validation process will provide informative responses back to CAMD participant organizations when/if their data transmissions are rejected for validation errors.</td>
</tr>
<tr>
<td>8</td>
<td>Valid data may need to undergo additional ETL processes to be written into the CAMD repository.</td>
</tr>
<tr>
<td>9</td>
<td>Valid data will be written to the CAMD OLTP repository. Data may be associated with additional metadata-associated data provenance and data quality.</td>
</tr>
</tbody>
</table>
The basic model for the access and use of the consolidated data set is shown in the diagram below, and described in the following table. Note: The data-access model begins with “Step 9” (from the previous data consolidation flow model), meaning the data-access model assumes that the data have already been transferred, quality-checked, transformed as necessary, and are resident in the CAMD repository. See Figure 8 and the table below for the CAMD data access model and flow description.

Figure 8. Proposed data access model.

<table>
<thead>
<tr>
<th>Step No.</th>
<th>CAMD Data Consolidation Flow Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>The mode begins with the assumption that quality-controlled, standardized-format data now reside in the CAMD (OLTP) database.</td>
</tr>
<tr>
<td>10</td>
<td>Based on business rules, certain data may be periodically archived. This could be because the data have passed criteria for useful life, have failed to reach threshold for utility, or have fallen below rated quality standards. For one reason or another, these data are not longer considered valid for typical reporting, but the coalition does not want the data lost forever since they may have some value for specific future data analyses.</td>
</tr>
<tr>
<td>11</td>
<td>It is envisioned that repository data may need to be further manipulated or transformed (perhaps consolidated or merged) for reporting purposes. The reporting data structure likely will follow and OLAP structure.</td>
</tr>
<tr>
<td>12</td>
<td>Data access applications will make data calls and retrieve data on behalf of a number of applications.</td>
</tr>
<tr>
<td>Step No.</td>
<td>CAMD Data Consolidation Flow Description</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------</td>
</tr>
</tbody>
</table>
| 13      | Envisioned data access, view, and manipulation applications (collaboration software) will include, but are not limited to:  
          - A knowledge base/document repository with the ability to rate/grade the relevance of received information  
          - A Wiki encyclopedia  
          - Data exports to allow users to export portions of the data in a variety of useful formats  
          - Modeling: the ability to apply the aggregated data set(s) to disease models  
          - Forums for the exchange of information among CAMD participants across organizations  
          - Both ad-hoc and pre-canned reporting mechanisms  
          - Sophisticated search mechanisms  
          - Other applications yet to be determined |
| 14      | Operating system and application security layer: Applications will be protected by an application security layer, including but not limited to:  
          - Single sign-on functionality (SSO)  
          - Secured remote accessibility  
          - Strong authentication  
          - Access rights granted by explicit permissions  
          - Visible data provenance as a part of the metadata  
          - Security audit features  
          - Internal enclaves with minimal ports, protocols, and services allowed across enclave barriers  
          - 24 x 7 monitoring for faults, performance issues, and security concerns |
| 15      | Perimeter transmission security: In addition to the application and operating system security, the infrastructure perimeter will be protected by what is primarily transmission security (see also Step 5). These mechanisms will include firewall(s), intrusion detection/prevention systems, and forward and reverse proxies. These will be configured to allow the transfer of data from authorized CAMD partners. |
| 16      | Valid CAMD users will issue requests after authenticating and being authorized through strong authentication mechanisms, and requests will be issued across encrypted channels. (Based on CAMD participants’ mutual agreement, public users may be given access to a more limited set of data without as strong a set of security constraints.) |
To accomplish the Data Consolidation and Data Access flow models as described above, the following infrastructure concept is proposed. Figure 9 below is meant to be a conceptual model, not a representation of physical connectivity.

Incoming data transfers and requests for information will arrive, encrypted at the perimeter of the C-Path infrastructure. These service requests will be handled by redundant (no single point-of-failure) infrastructure throughout. The perimeter security (e.g., firewalls, intrusion detection, and proxies) will pass service requests to load-balanced Web services through a limited set of coalition-agreed-upon ports, protocols and services. The Web servers alone shall talk to application services, which in turn shall talk alone to database services. The services will be separated by enclave boundaries using a “defense-in-depth” approach, so that minimal services need to be opened between enclaves.

This structure will limit CAMD’s risk for:

- Unauthorized data access
- Data loss
- Data theft
- Data tampering
- Downtime/outages
- Other problems

Details of the infrastructure-specific configurations and settings will be determined by C-Path as CAMD participants form definitive opinions on answers to the following questions:
Availability
- What is the allowable downtime, i.e., mean time between failure (MTBF) and mean time to repair (MTTR)?
- Is redundancy across all infrastructure components warranted at the associated cost?
- What is the continuity of operations plan?
  - How often must configurations be backed up?
- Will daily incremental backups and weekly full backups suffice?
  - How quickly must restore procedures be executed?
  - Is offsite backup warranted?
  - Is a disaster-recovery site warranted?
- Other considerations

Processing
- Are open-source products (e.g., Linux, Apache, PostgreSQL) acceptable to the CAMD participants in order to minimize cost while still providing a scalable solution?
- What are the response requirements?
- How many simultaneous users are envisioned?
- What type of load is envisioned? CPU bound? Disk I/O intensive?
- Other considerations

Storage
- What volume of materials is envisioned?
- What performance metrics are envisioned?
- Under what conditions (e.g., age, quality, infrequency of use) would data be taken offline?
- Is data archiving necessary?
- In addition to data provenance issues, what additional metadata are required?
- Other considerations

Data Mining
- Will the volume and usage be such that a separate set of infrastructures for data mining is warranted?
- What quality-validation rules does the coalition desire to enforce?
- What is the nature of the data, and what sorts of reports and queries are of greatest interest?
- Other considerations

Security
- How stringent should the coalition be with regard to authentication and authorization?
- What specific ports, protocols, and services should be allowable at the perimeter and between enclaves?
- What level of automated monitoring and corrective action would CAMD like to see?
- What forum will evaluate security violations, and what are the punitive actions, if any, for violations?
- Other considerations

Further detail regarding the data infrastructure is left pending until agreement is reached regarding these basic concepts, namely the Data Consolidation Flow Model, the Data Access Flow Model, and the basic design of the supporting infrastructure. Once these basics have reached consensus, this workgroup will hold requirements meeting(s) with CAMD representatives to flesh out further details of the data infrastructure design.
WORKGROUP 2: DISEASE-PROGRESSION MODELING

OVERVIEW

This section describes the application of mathematical models to characterize disease progression, and how the modeling process is a continuum that depends on the existing knowledge about specific conditions, as well as the different data needed to develop such models. In the case of Alzheimer’s and Parkinson’s diseases, the existing knowledge provides information about clinically observable phenomena, which can be used to create empirical disease-progression models. A particular application for quantitative disease, drug, and trial models is to help make more efficient decisions in the drug development process. The two key strategic goals of this workgroup are to provide a library of trial designs, endpoints, and analysis options for 1) early, exploratory, or phase II clinical trials and 2) late, adequate and well-controlled, or phase III, clinical trials.

Previous models developed by FDA or member companies will act as the basis for further model development. The work will be conducted considering the differences and similarities of Parkinson’s and Alzheimer’s disease in the context of neurodegeneration, from the clinical to the biological spectrum of both conditions (see Figure 10 below).

Figure 10. AD and PD in the context of neurodegeneration
As further knowledge regarding biomarkers, mechanistically defined subpopulations, and imaging parameters become incorporated into the models and associated with outcomes that have been defined as clinically relevant, it will be possible to examine potential relationships and associations.

Once the overall models contain sufficient data about the complex interactions among parameters in the individual disease models, a systems dynamics approach can be applied to create whole system biology models. Criteria then can be established for consensus identification of candidate parameters (biomarkers, imaging markers, etc.) that can be submitted to the FDA with a request that they be deemed “qualified” for specific use(s) in drug development. The subsequent results from the parameters’ application in drug development can be considered confirmation of knowledge that is then used to enrich the model and improving its predictive accuracy.

OBJECTIVES

This workgroup will pool data to create robust (in terms of scope and predictive accuracy), quantitative disease-progression models for use in drug development for Alzheimer’s and Parkinson’s diseases. The data used to create such models initially will be pooled from control arms of clinical trials performed by CAMD member companies and full clinical trials that were considered failures due to either efficacy or safety concerns, as well as publicly available sources like the ANDI database.

Additional data on relevant biomarkers (laboratory tests, imaging parameters, etc.) will be progressively incorporated into the models as they are generated by the respective team of the coalition. Other data sources that could have future value relevant to the workgroup include electronic health records of less systematically identified, characterized, and controlled populations compared to individuals participating in industry or NIH-sponsored clinical trials.

Further Objectives

- Describe modeling techniques and their context of application
- Describe how and when those modeling techniques could be applied to generate robust quantitative disease-progression models for use in drug development for Alzheimer’s and Parkinson’s diseases
- Define the clinical trial endpoints that will be incorporated into the model
- Evaluate the level of correlation among multiple endpoints (ADAS-Cog vs. DAD vs. NTB, etc.)
- Define the level of correlation among the subscales within each scale
- Evaluate the longitudinal predictability of each scale and its applicability in drug development

Milestones Needed to Achieve the Workgroup’s Objectives

- Identify prospective or retrospective datasets that will inform the model development
- Design a predefined model scope, as well as a data and knowledge analysis plan for all drug response and disease scales of interest
- Design a predefined plan for model development, evaluation, calibration, and implementation
• Continuously incorporate candidate biomarkers of disease stratification, activity/severity, and progression, both for Parkinson’s and Alzheimer’s diseases, as identified and qualified by the Biomarker Evaluation Workgroup

**PROCESS TO DETERMINE/COMPILE DATA NEEDS**

As explained below, the process to develop robust quantitative disease models will follow a downwardly integrative approach, beginning with empirical models based on clinically observed phenomena (see Figure 11 below.)

![Figure 11. Disease-progression modeling.](image)

The process to determine which variables are relevant for the development of robust models will be based on their strengths of association to disease progression, which includes the following:

- Medical history, including age of first diagnosis
- Basic demographic information
- Standard therapy and dose tracking
- Non-pharmacological actions (food, rest, and exercise)
- Frequency of follow-up visits and procedures
- Clinical score evolution over time (medical and self-assessed)
- Additional laboratory and imaging tests performed
- Patient dropout rates

Additional pathophysiologic data, such as biomarkers of risk, disease stratification, activity, progression, and prognosis, will be incorporated at further stages in the evolution of the models. The process to determine the relevance of such data will co-evolve with the development of CAMD, as the Biomarker Evaluation Workgroup and other worldwide scientific groups generate new information.
QUANTITATIVE NEURODEGENERATIVE DISEASE MODELS

Robust neurodegenerative disease-progression models offer quantitative insights into disease behavior and how it can be modified. These models can increase efficiency and decrease risk of errors in drug-development decisions by overcoming the complexity and uncertainties of the disease-drug-treatment interaction.

Many factors that influence outcomes can be considered simultaneously with the use of simulations, and the weight of knowledge and data gaps may be more systematically ranked and prioritized. A model-based drug-development approach can increase efficiency by integrating all pertinent prior information into a predictive model, which in turn can be used to guide the design of clinical trials and drug development strategy.20-23

A systematic model-based framework to maximize knowledge and data interpretation from prior and ongoing clinical trial information, as well as relevant preclinical and laboratory research, is critically needed.24 Such an approach can be used to characterize and quantify natural disease progression, placebo and drug effects, informed dose selection, as well as trial execution variables (patient discontinuation rates, compliance, self medication, specific design, etc.) from multiple trials using patient-level data.25,26

A large spectrum of modeling methods and techniques will be considered and may be used, depending on different factors, such as the questions to be resolved; the stage of drug research and development at which these questions arise; data availability; length and time scales of the problem (molecular to patients, weeks to years); and manpower and time constraints. These methods include the following:

- Empirical models, which often may originate from a “top-down” approach (from patient population data into pathophysiological phenomena, and as driven by drug dose and pharmacokinetics.
- Semi-empirical or semi-mechanistic, with increasingly explicit representation of pathophysiological, cellular, molecular, marker, and/or drug pharmacokinetics mechanisms. These may include mathematical transforms to generate measurable outputs that can be related to measurable events.
- Mechanistic or “first principle” models, which often are built upon a “bottom-up” approach (from molecular pathways to pathophysiology and clinical endpoints, and as driven by drug dose and pharmacokinetics).
- Search and inference algorithms, which may be applied across length and time scales and are aimed at establishing links among variables of interest, for which measurements are available.
- A combination or integration thereof to generate systems biology models of high complexity.20,23,25,26

As an illustration, an empirical model may describe the shape or trajectory of disease markers (biochemical tests, imaging patterns, clinical scores, etc.) over time, and may try to relate changes in these markers to treatment interventions. These models may have little or no relevance to biological basis of disease progression. These trajectories and their variations can be useful for simulating clinical trials and detecting potential drug effects.27,28 At this level of observation, both linear and nonlinear approaches can be applied to provide an estimate of disease states that relate to progression.29 Examples of Alzheimer’s
disease models include the work of Holford and Peace,\textsuperscript{21,22} as well as Lockwood et al.\textsuperscript{30} Another more detailed approach uses mathematical transforms that provide an output without necessarily incorporating explicit interconnections of underlying biological mechanisms.\textsuperscript{20} These kinds of models are particularly useful in oscillating, periodic, or aperiodic contexts, wherein the most detailed level, such as the relationship between heart rate variability and cardiac events,\textsuperscript{30-32} cannot be adequately defined.

A mechanistically oriented model incorporates a representation of key molecular signaling, cellular responses, or pathophysiological processes (or a combination of these) involved in disease-state manifestation and progression. Such a representation often incorporates feedback and feed-forward regulatory loops that maintain homeostasis in nonlinear, adaptive systems, thus providing significant insights into disease progression.\textsuperscript{20} Such models may simulate alterations in one or more parameters over a defined time period of interest. In addition, the function of multiple parameters may be assessed simultaneously to provide significant insights into the dynamics of the system.\textsuperscript{33-35}

All approaches require data from well-designed and well-executed studies, whether in vitro, in vivo or clinical, for the model to be adequately informed and useful. Model-driven experimental design, with a targeted collection of time-series (longitudinal) data, can help uncover underlying mechanisms and their associated kinetics within biological and pathophysiological regulatory loops.\textsuperscript{20} A model-driven experimental design should consider important aspects such as the timing and pattern of pharmacokinetic and pharmacodynamic effects, response-hysteresis phenomena, e.g., the time scale for the biological system to shift from and return to a baseline state after an event or intervention, etc.\textsuperscript{20,29} All these components should be integrated continuously and progressively into a model-based framework.

In the case of quantitative neurodegenerative disease models, it is essential to take into account a detailed clinical score evolution (Unified Parkinson’s Disease Rating Scale, or UPDRS, and Alzheimer’s Disease Assessment Scale-Cognitive subscale, or ADAS-Cog), patient dropout rates, genotype and other biomarkers, as well as environmental history data as extracted from human and possibly in vivo/in vitro studies, both previous and ongoing. The model also should consider all relevant pharmacokinetic and pharmacodynamic pathways (e.g., Cytochrome P450 polymorphisms, etc.) that may affect treatment response and contain relevant covariates such as acute and chronic treatment effects.\textsuperscript{21,36,37}

The qualification of candidate efficacy and safety biomarkers is a fundamental component of an evolving quantitative disease-progression model. Since biomarker qualification is context-dependent, and the context is provided by the drug-disease model and the data justifying the model, biomarker qualification and drug-disease modeling are intimately linked. Information about how the model responds (or not) to a drug intervention represents a fundamental opportunity to enrich any evolving quantitative disease-progression model.

With overall features such as different disease states and progressions being adequately represented, an integrated drug-disease model should reproduce clinically observed trends or trajectories correctly, by accurately modeling control and subsequent treatment responses to various drug classes.\textsuperscript{20-23,25-29} A key aspect in the case of neurodegenerative diseases is the need for the model to be able to correctly represent differing response and progression rates to novel therapies in characterized patient subpopulations.
As briefly described above, different—and often complementary—modeling methods and techniques may be used to create quantitative disease models. In the specific case of Parkinson’s and Alzheimer’s diseases, given the limited understanding of the molecular and physiological causal chains that lead to the diseases and their progression, a downwardly integrative approach may be preferred initially. Satellite modeling efforts (e.g., models borrowing from in vitro and/or in vivo studies and scaling up to human, molecular marker kinetic models in various body compartments, model-based scaling from animal to patients, etc.) may be used to indirectly support the top-down approach.

The proposed approach would start by considering clinically observed phenomena, which are included in each disease’s clinical score metrics (UPDRS and ADAS-Cog), followed by an attempt to represent these top levels with key pathophysiological functions that feed into clinical scores or their subcomponents. In this downwardly integrative attempt, the role of certain biomarkers may be ascertained more fully, thereby contributing to providing a basis for their potential qualification.

Certainly, there will be some knowledge and evidence gaps, which will be noted and considered for further analysis, e.g., via hypothesis testing through simulations of the larger modeling framework, as well as targeted experimental work or evaluation via the satellite models described above. In some instances, physiologically reasonable assumptions can temporarily attempt to fill some of those gaps. Model calibration and robustness need to be tested at each stage of development and at each scale of interest against available output data.

Although Parkinson’s and Alzheimer’s diseases are, according to the best understanding available, two different medical conditions, there are areas of pathophysiological similarity and clinical overlap. For this reason, a constant interaction between the teams focusing on each disease is crucial for the success of CAMD. Specific teams can carry out work in each particular condition, but with a regular and periodical effort toward common standards and integration of advances. Such a continuous process of information-sharing can provide useful insights for scientists working on each disease.

Since FDA scientists have developed an initial quantitative disease model for PD while member companies have worked on other in-house models for AD, the proposed approach for this workgroup is to start by cross-testing and enriching the existing models, with a continuous feedback based on the experience gathered at each stage.

Ultimately, the application of quantitative mathematical modeling can provide useful and significant insights into the nature of neurodegenerative diseases and their potential response to pharmacological interventions. The level of detail required by the models will be dictated by the nature of the questions posed regarding both diseases and, by extension, the level of knowledge required in order to sufficiently address these.
WORKGROUP 3: BIOMARKER EVALUATION

OVERVIEW

The focus of the Biomarker Evaluation Workgroup will be to establish a process and execute the plan for compiling and evaluating the scientific merit of reported biomarkers (including imaging data) that are potentially useful in drug or diagnostic test development. Initially, through collaboration with the Data Workgroup, data format and requirements, as well as IT support for the database, will be established. An initial list of candidate biomarkers will be compiled from members and other available data sources, and evidence will be thoroughly and systematically reviewed. Gap analysis will be performed, and programs will be implemented to address gaps. When biomarkers are determined to be adequately supported by evidence, and if deemed appropriate, they will move into the biomarker qualification process described in Chapter 1. The ultimate objective is to identify and qualify biomarkers that have promise in the development of new medical products to improve the management and outcome of these diseases.

OBJECTIVES

• Establish a process for collating a comprehensive list of potential biomarkers using publicly available data and CAMD members’ proprietary data
• Establish an evidence-based process incorporating current methods to assess scientific strength of candidate biomarkers
• Determine which biomarkers should have priority for further development efforts, as well as specific use contexts
• Determine when biomarkers have sufficient evidence to be submitted to the biomarker qualification process with FDA and EMEA (European Medicines Agency)
• Integrate biomarker data into quantitative disease-progression models, and use evolving disease models to help with the evaluation of biomarkers

PROCESS

The Biomarker Evaluation Workgroup will evaluate and prioritize candidate biomarkers for further development, submission, and qualification based on analysis of established databases, incorporation into disease-progression models, and internal CAMD member unpublished data and published reports. The Biomarker Evaluation Workgroup will include representatives from the member companies, member patient organizations, NIH, academia and other research organizations, FDA, and C-Path. The mission of this workgroup will be to establish and execute a systematic process for review of candidate biomarkers and to evaluate the scientific merit/strength of evidence supporting biomarkers for disease detection, activity, progression, predisposition, prediction of response to treatment, negative response markers, and safety biomarkers.

The first phase of this biomarker review effort will focus on process definition, e.g., defining the data sources for candidate biomarkers by compiling a comprehensive list of existent imaging and biochemi-
Coalition against Major Diseases

Work Scope

cal biomarkers, defining the process for adding new biomarkers to the list, determining criteria for evaluation, and establishing the method of rating the degree of confidence in each biomarker according to established methods of review.\cite{43, 44} Data on exploratory biomarkers will be shared with the FDA in multiple Voluntary Exploratory Data Submissions (VXDS) meetings. Working together with the Data Workgroup, information management also will be addressed, including development of a standardized format for the biomarker evidence database, as well as the infrastructure and method for storage and sharing of biomarker information. Figure 12 below outlines this process.

![Biomarker evaluation process and estimated timeline.](image)

The second phase will be the ongoing evaluation of evidence and the selection of the most promising biomarkers for regulatory review, qualification (where appropriate), and acceptance and use by the scientific community for a specific purpose. The workgroup will evaluate external reports and member organization data through literature presentations and internal data presentations, with the participation of additional invited experts. Evidence will be thoroughly reviewed, including rigorous statistical analysis of data. Candidate biomarkers will be ranked, and evaluation of top-tier biomarkers will continue...
with formulation of the claim of specific purpose(s) for each biomarker, review of evidence supporting claims, and identification of gaps in supporting evidence. A clinical validation program, for either de novo and/or re-mining of retrospective data, will be designed to address gaps. It is desirable to develop clinical program templates that can be used as a tool to guide clinical program development in a standardized approach, but it is also anticipated that the specific program for qualification will be biomarker and drug development claim-specific. When claims are determined to be adequately supported by evidence, these will be considered for submission to the biomarker qualification process by the Health Authorities Submission Workgroup, as described in the section below on Workgroup 4.

This process as described is open and dynamic, with data supporting candidate biomarkers entering the evaluation process on an ongoing basis. As data become available the workgroup will re-evaluate them and reprioritize biomarkers based on cumulative evidence.

Note: The timeline above (Figure 12) is a best estimate for evaluation and qualification of top-tier biomarkers expected to be further along in development. Qualification of additional biomarkers, with less supporting evidence and therefore requiring more clinical programs, is expected to have an extended timeline.

**DELIVERABLES**

*Nine months*
- Standardized format for biomarker database
- Criteria and process for evaluation of scientific merit of evidence
- Infrastructure to support data and access
- Comprehensive list of existent candidate biomarkers with existing scientific evidence mined from public and CAMD member databases, including formulating claims of use
- Initial prioritization of biomarkers
- Process for updating database

*Nine to fifteen months*
- Data from additional sources on current biomarkers incorporated into the existing database
- Gap analysis of top-tier
- Definition and initiation of clinical plan to address gaps

*Fifteen to twenty-four months*
- Submission of biomarkers to Workgroup 4 for submission to FDA/EMEA for qualification for specific use through the BQRT at the FDA
WORKGROUP 4: HEALTH AUTHORITIES SUBMISSIONS

OVERVIEW

The Health Authorities Submissions Workgroup (HASWG) has as its mandate to be the interface between the other working groups and government health authorities (agencies) for the purpose of regulatory submissions. Thus this group does not develop or analyze biomarker data or disease models, but rather assures that the progress of these efforts is informed by the current thinking relevant to biomarker and model qualifications of health authorities. Importantly, when such biomarker data or disease models reach a level of maturity that warrants regulatory review, the HASWG will be responsible for seeing that the collation of such information is ready for submission to the health authorities. The HASWG will prepare the submission package, following current health authority submission practices, and will serve as the primary interface for communication among the CAMD and the health authorities during the review process.

FDA personnel will participate in the functional workgroups as advisors and, in some cases, it will have direct scientific involvement. On the other hand, there will be no FDA direct participation on the HASWG. Certainly, the HASWG may have key contacts within the health authorities, and these contacts may offer advice regarding data format and procedural issues, but to maintain objectivity in the review process, these contacts will not participate in the submission preparation; in a similar fashion, reviewers at the FDA must not have been involved in CAMD activity.

OBJECTIVES

The HASWG has as its primary objective the preparation and submission of data-set packages for review by health authorities, and the successful qualification of the submitted biomarkers and/or disease models for use in regulated drug development.

A series of regular “grand rounds” events at the FDA, examining the state of the field and engaging outside experts, may be pursued prior to formal data submission to the FDA if the grand rounds serve as a valuable opportunity to communicate thinking and to engage in discussion with a large number of FDA regulatory scientists and staff members.

Submission of disease-progression models and biomarkers developed out of CAMD work will follow the path developed for biomarker submission that the Predictive Safety Testing Consortium presented to both the FDA and EMEA. For both agencies, the process will begin with a brief two-page notification of the intent to seek qualification. The process will proceed in the manner planned by FDA and EMEA for their own functions.

Aside from the requirements of the work, a dedicated team from the members of regulatory affairs departments, experienced in both FDA and EMEA submissions, is expected to manage the preparation of submission materials. These individuals should serve as the primary points of contact on the submission for biomarker qualification.
ALLIANCES WITH PATIENT COMMUNITIES

OVERVIEW

The Critical Path Institute (C-Path) strives to promote collaboration among biopharmaceutical companies and includes federal agencies such as FDA and NIH. The first priority for the C-Path Coalition Against Major Diseases (CAMD) is to develop and qualify for use the most modern tools and methods for drug development for Alzheimer’s and Parkinson’s diseases. In recognition that the patient communities represent a valuable resource of knowledge, experience, service, advocacy, and support for research, CAMD includes the participation by and contributions from patient groups that have become coalition members.

The patient groups will actively participate in the CAMD technical workgroups to accelerate the development of new therapies for Alzheimer’s and Parkinson’s diseases. The patient groups consist of Alzheimer’s and Parkinson’s organizations, as well as other not-for-profit organizations that promote health.

OBJECTIVES

The patient groups will bring to CAMD their collective knowledge of the entire Alzheimer’s and Parkinson’s disease communities, including patients, researchers, scientists, and companies working on these diseases. Under most circumstances, the member patient groups will represent the voices of patients—people living with Alzheimer’s and Parkinson’s diseases—and their caregivers. The patient groups will ensure that patient issues are at the forefront of all discussions and convey a sense of urgency. Rather than being a separate workgroup, the CAMD patient groups will seek to be integrated with and contribute to the CAMD technical workgroups, which in turn will strive to improve the development of therapies for Alzheimer’s and Parkinson’s diseases. The patient groups also will serve in a leadership role in communicating between CAMD and the broader patient community when appropriate and allowable under the confidentiality restrictions of the legal agreement.

As part of this process, CAMD will invite scientists supported by the patient groups and the federal government to present their data to the technical workgroups in the hopes of increasing the transparency of research, promoting dialogue among multidisciplinary scientists, and speeding the transition from discovery to therapeutics.
CHAPTER III: SUPPORT AND COMMUNICATIONS TEAM

ADMINISTRATION AND PROJECT MANAGEMENT

OVERVIEW

The Support and Communications Team is responsible for supporting the overall CAMD project coordination and communication.

OBJECTIVES

• Establish a project management staff support and infrastructure that will maintain project plans for each workgroup, as well as track and report progress
• Facilitate collaboration and information-sharing among members of CAMD

SPECIFIC TASKS

1. Provide overall coordination of workgroup efforts. Tasks include:
   • Develop and maintain workgroup schedules and milestones
   • Maintain detailed work plans, schedule meetings, and document actions and the status of projects
   • Standardize the use of Microsoft Project; using a baseline for tracking, assign task owners for accountability, templates for filters, and a process for keeping information up to date
   • Standardize a general communication and collaboration tool
   • Manage expenses and other resources needed by the workgroups
   • Create templates for important documents
   • Maintain a documentation database with necessary security and check-out controls

2. Establish a web domain and password-protected site for members of the coalition. This “intranet” will have a basic structure for communication within the coalition, including a main page for announcements of general interest, and pages where each workgroup can post draft documents and share comments. Microsoft SharePoint will be the collaboration software.

   All members will be given an ID and password for access to the CAMD site. Workgroup leaders will approve additional users.

3. Establish teleconference, web conference, and videoconference capabilities as required to support CAMD workgroup activities.

4. To review progress, C-Path and Brookings will host formal annual meetings (perhaps more frequent) of the CAMD coordinating committee. To keep CAMD members informed of progress on an ongoing basis, more frequent communications will be conducted through e-mail.
ANNOUNCEMENTS AND ONGOING EXTERNAL COMMUNICATIONS

OVERVIEW

The Support and Communications Team is also responsible for announcements and ongoing external communications of CAMD benefits and progress to the industry, medical and business communities, and the media.

External communications will maximize the impact of CAMD announcements. This will include securing media coverage and industry/medical publication of significant accomplishments. CAMD will be highlighted as an example of effective government/private partnership. This external communication is intended to generate significant visibility for CAMD and to lead to enhanced opportunities to bring additional support to the project.

OBJECTIVES

- Gather broad exposure and press coverage from CAMD launch announcement and future progress, developing awareness in the industry, patient communities, medical community, academia, and the general public
- Show the unique and complementary value of the CAMD effort
- Develop a pre- and post-launch CAMD press plan
- Develop a pre- and post-launch CAMD communications plan to government officials, industry and medical publications, etc.

SPECIFIC TASKS

Specific tasks and resources required for each are outlined below.

CAMD launch event in Washington, D.C., in 2009:
- Prepare CAMD launch via broad press release across media platforms

Post-launch communications plan:

CAMD communications will involve a tiered approach to information, beginning with a macro message and becoming progressively more technical as needed. The macro message will allow CAMD to emerge as a major health-care initiative that is valued and respected by the public. Communications will be directed toward the following audiences/constituencies:

- Media
  - Establish a detailed plan for periodic press releases to show metrics, transparency, and accountability throughout the CAMD process
- Develop continuous and updated content from scientists and other members to stay active with media and provide specific messaging for sophisticated or targeted audiences
- Secure prominent placement and continued coverage in industry and medical publications

• Members
Members will be kept apprised of news, successes and developments through the CAMD project management team.

• Patient Community
Patient groups will be a key resource for communicating important findings to the patient community throughout the timeline of the coalition.

• General Public
The general public will be educated regarding the magnitude of the Alzheimer’s and Parkinson’s problems and the successes of CAMD.
APPENDICES

APPENDIX A: GLOSSARY OF TERMS AND ACRONYMS

**AD**: Alzheimer’s Disease

**ADaM**: Analysis Data Model (a CDISC standard)

**ADAS-Cog** (Alzheimer’s Disease Assessment Scale-Cognitive subscale): The most commonly used primary outcome instrument in clinical trials for treatments of dementia. It focuses on memory loss, so it is often complemented by the application of the Mini-Mental State Examination (MMSE).

**ADNI**: Alzheimer’s Disease Neuroimaging Initiative (part of LONI)

**BQRT**: Biomarker Qualification Review Team

**CAMD**: Coalition Against Major Diseases (part of C-Path)

**CDISC**: Clinical Data Interchange Standards Consortium

**CFIC**: Center for FDA and Industry Collaboration

**COTS**: Commercial Off-the-Shelf (software)

**C-Path**: Critical Path Institute

**CPI**: Critical Path Initiative (part of the FDA)

**CRF**: Case Report Form

**CSR**: Clinical Study Report

**CSF**: Cerebrospinal Fluid

**EHR**: Electronic Health Records

**EMEA**: European Medicines Agency

**FDA**: U.S. Food and Drug Administration

**FTE**: Full-time equivalent (employee)
GUI: Graphical User Interface

HL-7: Health Level 7

HHS: U.S. Department of Health and Human Services

IT: Information Technology

LONI: Laboratory of Neuro Imaging (UCLA)

MMSE: Mini-Mental State Evaluation. A screening tool frequently used by health-care providers to assess overall brain function. It is often used to evaluate patients with possible Alzheimer’s disease or another related dementia.

NIA: National Institute on Aging

NINDS: National Institute for Neurological Disorders and Stroke

NIH: National Institutes of Health (part of Health and Human Services)

PD: Parkinson’s Disease

PD-DOC: Parkinson’s Disease Data and Organizing Center

PK: Pharmacokinetics

PRO: Patient-Reported Outcomes

SDTM: Study Data Tabulation Model (A CDISC standard)

UPDRS (Unified Parkinson’s Disease Rating Scale): Rating score used to follow the prospective course of Parkinson’s disease, based on the patient’s experience with the disease and a motor examination. It is composed of the following sections: mentation, behavior and mood; experiences of daily living; motor examination; and complications of therapy.
APPENDIX B: OVERVIEW AND PURPOSE OF CAMD LEGAL AGREEMENT

The CAMD legal agreement provides the framework for the operation of the Coalition Against Major Diseases. Below is a brief summary of important provisions contained in the legal agreement.

Section 2: Statement of Purpose: Research Plan

2.1: Statement of Purpose
To form a coalition of members representing any of the following:
1. Patient organizations
2. Pharmaceutical companies
3. Biotechnology companies
4. Diagnostic and medical device companies
5. Food and Drug Administration
6. National Institute on Aging or other unit of NIH
7. Other governmental entities
8. Research institutions
9. Nonprofit organizations

To enable sharing of information about neurodegenerative diseases:
   Alzheimer’s disease
   Parkinson’s disease

To expedite the development of safe new treatments, cures and preventions by creating quantitative disease-progression models

To qualify biomarkers for use in drug development and to incorporate their use in drug development as clinical diagnostic tools

Each signatory of the agreement will appoint one voting representative to serve on a coordinating committee that will govern and guide the work of the consortium. Although not explicit in the agreement, it is expected that the coordinating committee will encourage each member to assign scientists to participate in each working group (see Section 2.2). Members will be expected to contribute data, expert guidance, advocacy for the mission, or other important contributions.

2.2: Research Program
The overall research program of the CAMD will be conducted through one or more research projects. The research plan and budget for each research project will be subject to approval by the CAMD’s Coordinating Committee, which will be incorporated into a project agreement that is mutually agreed to by C-Path (on behalf of the CAMD), the member(s) participating in the research project, and, if applicable, any approved participating third-party contractor(s).
Each research project will be overseen by a working group, composed of representatives of each of the parties to such project agreement.

2.3: Materials Transfer
In order to facilitate the research project, it is expected that biological samples, reagents or chemicals may be exchanged by the members from time to time—this section allows the members to share materials throughout the research project without the need for a separate materials transfer agreement for each item shared.

2.4: Policy regarding compliance with antitrust policy
By reference to Exhibit A, this informs all members and attests that the members will act in compliance with antitrust laws.

2.5: Disclaimer of partnership
This states the intent of the members to interact as independent contractors, and not to create any agency, joint venture, or similar arrangement.

Section 4. Management

4.1: Consortium Management
   Director – will be a C-Path employee
   Co-Director – will be a member’s representative elected by members
   Coordinating Committee – Each member has one voting representative on the committee.

4.2: Coordinating Committee – This section defines its responsibilities, e.g.:
   Develop research plan
   Oversee research groups
   Approve research projects and budgets
   Will include nonvoting FDA or EMEA representative(s)

Section 6. Confidentiality

6.1: Confidentiality – This section enables all members to share their intellectual property without loss of their rights. This is an essential element to ensure the development of new therapies.

6.2: Exceptions – standard language for confidentiality agreement

6.3: Authorized disclosure – standard language to disclose information

6.4: Publications – standard language that encourages prompt publication and respects confidentiality of information being shared and has standard language to allow time for filing patents.

6.5: Publicity – Permission required for mention of partners in publicity regarding research or technical publicity, but not required if only mentioning their name.
APPENDIX C: CAMD PARTICIPANT ROLES

Voting members of the Coordinating Council:

- Designated representatives of the member patient groups and industry. Responsible to be the conduit and champions for CAMD activities within their organizations. Together, they are the governing group for CAMD work plans and all matters pertaining to the coalition.
- Critical Path Institute - Director of CAMD

Nonvoting participants and advisors:

- FDA, EMEA and NIA representatives (technical advisors to the Coordinating Council and workgroups)
- Brookings Institution - cosponsor (with C-Path) of CAMD public meetings

Workgroup members:

- Scientists from relevant disciplines within the industry and patient group members to define and, in some cases, execute activity defined by that workgroup’s work scope
- Patient group members to provide guidance on the direction and priority of work and assist in public policy messaging
- Scientific and advisory personnel from the FDA, EMEA and NIH

Additional CAMD members:

- To be determined by the Coordinating Council and as defined by the CAMD Consortium Legal Agreement

Additional CAMD participants:

- Other organizations or individuals that are brought into workgroups to provide information, services, advice, etc., have a confidential disclosure agreement in place with C-Path, and are approved by the director of CAMD.
REFERENCES CITED


