

ALZHEIMER'S DISEASE COMMON RESEARCH ONTOLOGY AND PORTFOLIO ANALYSIS PROJECT

Coding Guidelines



TABLE OF CONTENTS

	Page
1. Introduction	3
– Project History	
– Purpose and Goals	
2. The Coding Process	4
– Projects Suitable for Coding	
– Structure of the Ontology	
– Steps for Assigning Codes to Funded Projects	
3. Process for Submitting Coded Grants	7
– Preparing Your Data for Submission	
– Required and Optional Data Fields for Submission	
Appendix 1 – CADRO Working Group Members and Contact Information	10
Appendix 2 – Common Alzheimer’s Disease Research Ontology	11
Appendix 3 – Further Clarification of Certain Topics	18

Introduction

Project History

The creation of the first iteration of the Common Alzheimer Disease Research Ontology (CADRO) is a collaborative effort between the National Institute on Aging (NIA) and the Alzheimer's Association (AA). The project began in May 2010 for the purpose of conducting a comparative analysis of the NIA and AA Alzheimer's disease (AD) research portfolios.

Purpose of the Ontology

The CADRO is designed to serve as a unified classification system that will enable the analysis of research portfolios, strategic planning, and coordination by federal and non-federal agencies supporting AD research in the United States and internationally. It is meant to be a dynamic portfolio analysis tool that can be used to:

- Track changes across the AD research landscape over time across various agencies and organizations.
- Identify research gaps and areas of overlap within and across funders of AD research.
- Identify collaborative opportunities aimed at advancing AD research and alleviating the socioeconomic burden of this devastating disease.

The ultimate goal of this project is to expand the use of the CADRO across federal and non-federal agencies that support AD research both in the US and internationally. To spur discussion and support decision-making, project data and related information would then be made available to the wider funding and research community through the development of a public online database.

The Coding Process

The purpose of this document is to provide guidance for organizations that support AD research regarding the classification of their portfolios and its inclusion in a central database (currently under development).

Please include funded projects:

- ☑ With 'Alzheimer's Disease' as the primary focus.
- ☑ Grants, projects and contracts active during FY2008 to the present.
[Note: Projects funded prior to FY2008 can also be included as long as they were active in FY2008.]

Please do not include grants, projects, and contracts awarded for the sole purpose of indirect costs (e.g., overhead, travel awards, loan repayment and buildings costs).

Structure of the Ontology

The CADRO is a *three-tier classification system*. To create the ontology, the NIA and the AA used projects funded from FY2004 through FY2010. Five broad research categories and two research resources-related categories were identified. A brief description of each category is presented below. Each of the categories is divided into research "topics" and some of these topics are further divided into "themes". The full CADRO listing the categories, topics, and themes can be found in Appendix 2.

Category A. Molecular Pathogenesis and Physiology of Alzheimer's Disease

This category includes research focused on the molecular and physiological processes underlying Alzheimer's disease pathogenesis and the genetic and epigenetic determinants of AD.

Category B. Diagnosis, Assessment and Disease Monitoring

This category includes research focused on the development, testing and validation of tools and methods for diagnosing and monitoring patients with AD from the preclinical phase of the disease through advanced dementia. These methods and tools include all types of novel and established biomarkers.

Category C. Translational Research and Clinical Interventions

This category aims to capture projects focused on the identification and development of therapies (small molecule, natural products and biologics) for AD from early therapeutic discovery through late stage preclinical development and all stages of clinical testing. Also included are projects focused on repurposing pharmacological agents already in use for other conditions as well as non-pharmacological interventions.

Category D. Epidemiology

This category includes all types of epidemiological studies (cross-sectional, prospective, and longitudinal) aimed at examining how a variety of genetic, lifestyle, and environmental factors influence the incidence, prevalence and clinical course of AD.

Category E. Care, Support and Health Economics of Alzheimer's disease

The research in this category includes projects aimed at improving the quality of care and quality of life for AD patients in a variety of care-giving settings (i.e. in the home, nursing home facilities, hospice programs) and across diverse populations. This category also includes research focused on alleviating the physical and emotional burden associated with caregiving as well as projects focused on assessing the socioeconomic burden of AD.

Category F. Research Resources

This category includes a variety of resources used to conduct, translate and disseminate high quality AD research such as research centers, infrastructure resources (i.e. various cores), data and tissue repositories and projects focused on generating disease models. Training and career development programs are also included in this category. Projects included in this category may also be reflected in Categories A-E based on scientific relevance.

Category G. Consortia and Public Private Partnerships

This category includes partnership enterprises created to enable major national and international efforts in basic and translational AD research. Projects categorized in this category may also be reflected in Categories A-F based on scientific relevance.

Steps for Assigning Codes to Funded Projects

1. Identify your relevant AD grants, projects, and contracts (for NIH - both extramural and intramural) based on the award title, abstract, and specific aims of the awarded application.
2. Identify and assign the appropriate research category, topic, and theme (if applicable).
 - Usually the project abstract together with the specific aims provide sufficient information to identify the most appropriate category, topic and theme. Other information about the investigator such as past research and publications should not be used in classification decisions for a given award.
 - Coding decisions should not be based on the future directions of the funded projects. For example a statement such as *“Understanding the role and structure of this molecule may lead to the identification of novel therapies for AD.”* cannot be used to classify this project as translational. Only the current aim of *“Understanding the role and structure of this molecule”* is relevant as to how it should be coded.
 - Most projects should be assigned to a single category, topic, and theme based on the best fit. In the case of grants with multiple components, notably grant mechanisms such as program projects grants (e.g. P01s) and some cooperative agreements, only the individual subprojects should be coded and included. This is meant to capture the multi-disciplinary nature associated with these types of grants.
3. To ensure quality control of this process, we recommend that each grant, project, and contract that has been identified for potential inclusion be independently coded by one or two additional reviewers with expertise in the relevant research category. This provides validation and cross-validation of the initial assignment. Differences among the reviewers should be reconciled before submitting.

Process for Submitting Funded Grants

Preparing Your Data for Inclusion in the Database

Once appropriate grants, projects, and contracts have been identified, a minimum set of ancillary information will be needed to enable users of such information to analyze and display data and build a comprehensive picture of Alzheimer's disease research funding.

To ensure ease of both submission and use, please use Microsoft Excel. The following fields must accompany each grant, project, and contract.

Field	Field Description
1. Name of Funding Organization	
2. Country of Funding Organization	
3. Project Terms	<i>Optional:</i> Key words associated with each project
4. Project or Grant Number	Project or identification code associated with specific project
5. Unique Identifier	A fully unique identifier for each grant. For many funders, this may be the same as the project or grant number. For NIH ICs, please use the Application ID and make sure that this is indicated at the subproject level when these are to be coded (rather than parent grants).
6. Project Title	
7. Administering IC*	<i>NIH Only:</i> Institute/Center overseeing the administering of grant, project, or contract; may also be co-sponsored by another IC or agency
8. Award Type*	<i>NIH Only:</i> '1-New awards', '2-Renewal', '5-non-competing continuation', etc.

9. Award Activity*	NIH Only: Type of project or grant mechanism used in awarding grant – e.g., ZIA for intramural, R01 for investigator initiated, U01 for Cooperative Agreement
--------------------	---

10. Support Year*	
-------------------	--

11. Subproject Number*	<i>NIH Only:</i> for P01 and U01 grants coded at the subproject level
------------------------	---

12. Project Official Information	Name of the Health Science Administrator or other program official who oversees the administration of the grant, project, or award
----------------------------------	--

13. Project Start Date	
------------------------	--

14. Project End Date	
----------------------	--

15. Award Solicitation	The notice of a funding opportunity in which the award, grant, or contract was in response to (leave blank if not applicable) <i>For NIH users:</i> The award solicitation includes the RFA/PA announcement number
------------------------	---

16. PI Last Name	
------------------	--

17. PI First Name	
-------------------	--

18. PI Middle Initial	
-----------------------	--

19. Grantee Name	Name of organization in which the grant, project or contract is awarded
------------------	---

20. Grantee City	
------------------	--

21. Grantee State	
-------------------	--

22. Grantee Zip	
-----------------	--

23. Grantee Country	
---------------------	--

24. Congressional District	<i>Optional:</i> Congressional District (if applicable) where the grant was awarded
25. ARRA Indicator*	<i>Federal Agencies Only:</i> grants and projects awarded in response to the funding provided through the American Recovery and Revitalization Act
26. Funding Mechanism Description	Descriptive code of the funding activity, including 'Individual Research', 'Center', 'New investigator Award', 'Small Business Grant'
27. FY	Fiscal year in which the grant, project, or contract receives funding or support
28. FY Total Cost	Funding allocated for associated with each FY in which the funded award was active
29. FY Total Cost* (subprojects)	<i>NIH Only:</i> funding allocated to subprojects associated with P01s and U01s (parent grant should not be submitted – only subprojects)
30. Category Name	Ontology - Research category associated with the grant, project, or subprojects
31. Topic Name	Ontology – Research topic (within each category) associated with each grant, project, or contract
32. Theme Name	Ontology – Research theme (within each category and topic) associated with each grant, project, or contract
33. *Category G – Consortia and Public-Private Partnerships (Y/N)	Please indicate whether this project, grant, or contract is part of a consortia or public-private partnership.
34. Grant, Project, or Contract URL	Website URL that may provide additional information or link users to the abstract of grant, project or contract. For NIH ICs, this will include the link to the project's respective NIH Reporter profile for grant, project, or contract

Appendix 1. CADRO Coordinating Group Members and Contact Information

National Institute on Aging

Charlene Liggins
Suzana Petanceska
Lorenzo Refolo
Laurie Ryan
Nina Silverberg

Alzheimer's Association

Maria Carrillo
Heather Snyder

NIA Contacts:

Charlene Liggins
ligginsc@mail.nih.gov

Lorenzo Refolo
refolol@mail.nih.gov

Alzheimer's Association Contact:

Heather Snyder
Heather.Snyder@alz.org

Appendix 2 - Common Alzheimer's Disease Research Ontology (CADRO)

The following is a listing of the CADRO Categories, Topics, and Themes as of November 2012.

Category A. Molecular Pathogenesis and Physiology of Alzheimer's Disease

This category includes research focused on the molecular and physiological processes underlying Alzheimer's disease pathogenesis and the genetic and epigenetic determinants of AD.

- 1. Amyloid**
 - a. APP Structure and Function
 - b. APP Processing
 - c. APP Signaling
 - d. Secretases
 - e. Amyloid beta Clearance
 - f. Amyloid beta Structure, Assembly, and Aggregation
 - g. Amyloid beta-Mediated Pathogenesis
- 2. Tau**
 - a. Normal Functions of Tau
 - b. Tau Phosphorylation, Metabolism, and Assembly
 - c. Tau-Mediated Pathogenesis (AD)
 - d. Tau-Mediated Pathogenesis (Tauopathies)
- 3. Presenilins**
 - a. Apoptosis
 - b. Calcium Signaling
 - c. Autophagy
 - d. Structure-Function Analysis
- 4. ApoE and Lipids**
 - a. ApoE in Abeta-Mediated AD Pathogenesis
 - b. ApoE in AD Pathogenesis Independent of Abeta
 - c. Brain Cholesterol Metabolism
 - d. Lipid-Mediated Signaling
 - e. Lipoprotein Receptors
 - f. Nuclear Receptors
 - g. Myelin
- 5. Brain Circuits and Synapses**
 - a. Synaptic Plasticity and Synaptic Dysfunction
 - b. Selective Vulnerability
 - c. Neurotransmitter Receptors Structure and Function
 - d. Network Function and Failure
- 6. Neurogenesis Cell Death**
 - a. Apoptosis
 - b. Oxidative Stress
 - c. Autophagy-Mediated Cell Death
 - d. Calcium-Mediated Cell Death
- 7. Immunity and Inflammation**
 - a. Astrocytes
 - b. Microglia
 - c. Innate Immunity
 - d. Immunotherapy – Mechanisms of Action
 - e. Inflammatory Mediators
- 8. Bioenergetics**
 - a. Mitochondria
 - b. CNS Glucose Metabolism and Ketogenesis
- 9. Vascular/Metabolic Factors**
 - a. Cerebrovascular Disease
 - b. BBB and Neurovascular Unit
 - c. Insulin Resistance and Type II Diabetes
 - d. Hypertension
 - e. Dyslipidemia
 - f. Atherosclerosis
 - g. Obesity
 - h. Metabolic Syndrome
- e. Cell Cycle Re-Entry**
- f. Ubiquitin Protease System**

10. Hormones

- a. Sex Hormones
- b. Growth Hormones
- c. Stress Hormones

11. Genetics

- a. Candidate Gene Approach
- b. Genome-Wide Approach
- c. Epigenetic and Epigenomic Approaches
- d. Genetic Architecture
- e. Gene-Gene and Gene Environment Interactions
- f. Expression Profiling
- g. Disease Pathways Identification
- h. Next Generation Sequencing
- i. Genetic Data Use and Analysis

12. Other

Category B. Diagnosis, Assessment, and Disease Monitoring

This category includes research focused on the development, testing and validation of tools and methods for diagnosing and monitoring patients with AD from the preclinical phase of the disease through advanced dementia. These methods and tools include all types of novel and established biomarkers.

- 1. Fluid Biomarkers**
 - a. CSF Biomarkers
 - b. Blood Biomarkers
- 2. Imaging Biomarkers**
 - a. PET Amyloid Imaging
 - b. PET Non-Amyloid Imaging
 - c. Functional MRI
 - d. Structural MRI
 - e. Other Brain Imaging Tools
- 3. Cognitive, Behavioral and Functional Assessment**
 - a. Cognitive
 - b. Behavioral
 - c. Functional
- 4. Multimodal Biomarkers**
- 5. Novel Biomarkers**
- 6. Novel Methodologies and Techniques**
- 7. Other**

Category C. Translational Research and Clinical Interventions

This category aims to capture projects focused on the identification and development of therapies (small molecule, natural products, and biologics) for AD from early therapeutic discovery through late stage preclinical development and all stages of clinical testing. Also included are projects focused on repurposing pharmacological agents already in use for other conditions as well as non-pharmacological interventions.

1. Drug Discovery (small molecules and biologics)

- a. Amyloid
- b. Tau
- c. ApoE, Lipids and Lipoprotein Receptors
- d. Neurotransmitter Receptors
- e. Neurogenesis
- f. Inflammation
- g. Oxidative Stress
- h. Cell death
- i. Proteostasis
- j. Metabolism and Bioenergetics
- k. Vasculature
- l. Growth Factors and Hormones
- m. Epigenetic Regulators
- n. Multi-target
- o. Unknown target
- p. Other

2. Preclinical Drug Development (small molecules and biologics)

- a. Amyloid
- b. Tau
- c. ApoE, Lipids and Lipoprotein Receptors
- d. Neurotransmitter Receptors
- e. Neurogenesis
- f. Inflammation
- g. Oxidative Stress
- h. Cell death
- i. Proteostasis

- j. Metabolism and Bioenergetics
- k. Vasculature
- l. Growth Factors and Hormones
- m. Epigenetic Regulators
- n. Multi-target
- o. Unknown target
- p. Other

3. Preclinical Proof of Concept for Non-Pharmacological Interventions

- a. Exercise
- b. Diet
- c. Enrichment
- d. Combination therapy
- e. Other

4. Clinical Trial Design

5. Early-stage Clinical Drug Development (Phase I and Phase II Clinical Trials)

- a. Amyloid
- b. Tau
- c. ApoE, Lipids and Lipoprotein Receptors
- d. Neurotransmitter Receptors
- e. Neurogenesis
- f. Inflammation
- g. Oxidative Stress
- h. Cell death
- i. Proteostasis
- j. Metabolism and Bioenergetics

- k. Vasculature
- l. Growth Factors and Hormones
- m. Epigenetic Regulators
- n. Multi-target
- o. Unknown target
- p. Other

6. Late-stage Clinical Drug Development (Phase III Clinical Trials)

- a. Amyloid
- b. Tau
- c. ApoE, Lipids Lipoprotein Receptors
- d. Neurotransmitter Receptors
- e. Neurogenesis
- f. Inflammation
- g. Oxidative Stress
- h. Cell death
- i. Proteostasis
- j. Metabolism and Bioenergetics
- k. Vasculature
- l. Growth Factors and Hormones
- m. Epigenetic Regulators
- n. Multi-target
- o. Unknown target
- p. Other

- 7. Non-pharmacological Interventions**
 - a. Exercise
 - b. Diet
 - c. Cognitive Training
 - d. Combination therapy
 - e. Other

- 8. Clinical Therapy Development for the Neuropsychiatric Symptoms of AD**
 - a. Pharmacological
 - b. Non-pharmacological

- 9. Clinical Ethics**

- 10. Other**

Category D. Epidemiology

This category includes all types of epidemiological studies (cross-sectional, prospective, and longitudinal) aimed to examine how a variety of genetic, lifestyle, and environmental factors influence the incidence, prevalence, and clinical course of AD.

- 1. Genetic/Epigenetic Risk**
- 2. Cardiovascular and Metabolic Factors**
- 3. Nutrition and Other Environmental Factors**
- 4. Multimodal Risk Factors**
 - a. Hispanics
 - b. African-Americans
 - c. Japanese-Americans
 - d. Multi-Racial/Cross-Cultural
 - e. International (Israel, Sweden, China, India)
 - f. Women
 - g. Oldest Old
 - h. Other
- 5. Other**

Category E. Care, Support and Health Economics of Alzheimer’s Disease

The research in this category includes projects aimed at improving the quality of care and quality of life for AD patients in a variety of care-giving settings (e.g., in the home, nursing home facilities, hospice programs) and across diverse populations. This category also includes research focused on alleviating the physical and emotional burden associated with caregiving as well as projects focused on assessing the socioeconomic burden of AD.

1. Care Interventions and Quality of Life

- a. Cognitive Training Interventions
- b. Health and Wellness
- c. Behavioral Interventions
- d. Hospice and End-of-Life Care
- e. Staff Training and Professional Development
- f. Assessment and Metrics
- g. Neuropsychological Interventions
- h. Other Interventions

2. Technology Assisted Care

- a. Personal Device Assisted Care
- b. Computer Assisted Care
- c. Environmental Modifications
- d. TV/Video Assisted Care
- e. Other Technology Assisted Care

3. Caregiver Support

- a. Caregiver Training
- b. Home-Based Support
- c. Behavioral Interventions
- d. Relationship Interventions
- e. Assessment and Metrics

4. Cultural Values and Beliefs

- a. Chinese
- b. American Indian
- c. African Americans
- d. Latino/Hispanic
- e. Japanese
- f. Assessment and Metrics

5. Economic Burden of Alzheimer’s Disease

6. Other

Category F. Research Resources

This category includes a variety of resources used to conduct, translate, and disseminate high quality AD research such as research centers, infrastructure (e.g., various cores), data and tissue repositories and projects focused on generating disease models. Training and career development programs are also included in this category. Projects in this category may also be reflected in Categories A-E based on scientific relevance.

1. Alzheimer’s Disease Centers

- a. Administrative Core
- b. Clinical Core
- c. Data Management and Statistics Core
- d. Education and Information Core
- e. Neuropathology Core
- f. Imaging Core
- g. Optional Cores

2. Other Types of Cores (e.g., program projects)**3. Professional and Career Development**

- a. Faculty Recruitment
- b. Clinical Scientist Career Development
- c. Training Grants
- d. Conferences/Workshops/Symposia

4. Repositories and Bioinformatics Tools and Resources

- a. Biobanks
- b. Data Repositories
- c. Bioinformatics

5. Infrastructure (including equipment, construction, technology, etc.)**6. Disease Models**

- a. Invertebrates
- b. Vertebrates
- c. Rodents
- d. Higher Mammals
- e. iPS Cells

7. Other**Category G. Consortia and Public Private Partnerships**

This category includes partnership enterprises created to enable major national and international efforts in basic and translational AD research. Projects categorized in this category may also be reflected in Categories A-F based on scientific relevance.

1. Consortia**2. Public Private Partnership**

Appendix 3 – Further Clarification of Certain Topics

Below are associated key words to help provide clarification on Category A. Topic 11. Genetics:

- a. Candidate gene approach: risk factors, gene identification, follow up study, gene confirmation, replication analysis of GWAS, deep sequencing, massively parallel sequencing, genetic polymorphism, phenotype, haplotype, case-control study, family based study, genetic pedigree, mutation, proband, linkage disequilibrium, population stratification, sibling, sib-pairs, endophenotype, minority population
- b. Genome-wide approach: association study (GWAS): genetic risk factors, genetic polymorphism, phenotype, phenotype-genotype interaction, age at onset, cognitive function, biomarkers, cerebrospinal fluid biomarkers (imaging, amyloid beta, tau, Lewy bodies, measures of mild cognitive function), haplotype, population substructure, case-control study, family based study, mutation, proband, population stratification, endophenotype, expression quantitative trait locus (eQTL), minority population
- c. Epigenetic and epigenomic approaches: gene expression, gene function, structure-function relationship, genetic heterogeneity, heritability, RNA-Seq, expression profiling, microarray, DNA methylation, RNA methylation, histone deacetylation, silenced gene, RNA epigenetics, expression quantitative trait locus (eQTL), Chip-Chip sequencing
- d. Genomic architecture: haplotype, gene desert, distant (trans) promotor or/and repressor / insulator (protector), gene function, copy number variation, DNA binding proteins, insertion/deletion (indel), copy number variation, haplotypes, unique genomic function, population stratification, heterozygote, homozygote, DNA damage, DNA repair
- e. Gene-gene and gene environment interactions: epistasis, allelic interactions, environmental risk factors, co-morbidities, haplotype, allelic interaction
- f. Expression profiling: gene expression profile, gene expression regulation, altered DNA – RNA interactions, altered DNA- protein products, phenotype, biomarkers, age at onset, linkage disequilibrium, mRNA, regulatory genes, genetic transcription, gene function, differential expression, microarray
- g. Disease pathways identification: identification of therapeutic targets, molecular diagnosis, cellular networks, AD genes in the same functional pathway, AD gene pathway interactions (AD genes across pathway analysis), prevention, mRNA, RNA sequencing, transcription, micro RNA sequencing, non-coding RNA, si RNA, differential expression, brainome, connectome

- a. Next generation sequencing: deep sequencing, massively parallel sequencing, high throughput, exome, unintended consequences of gene discovery*, family based approach, Mendelian gene, genetic pedigree, expression quantitative trait locus (eQTL), RNA-Seq, Chip-Chip sequencing, heritability

- i. Genetic Data use and analysis: new bioinformatic and statistical tools - Data acquisition, data collection, data integration, data management, data mining, data storage, data quality, data set, database, bioinformatics, firewall, encryption, sample size, high throughput analysis, covariates, multivariate linear regression, multi- dimensional variables, multi-dimensional covariates, algorithms, statistical modeling, theoretical properties, simulation modeling, statistical approach