# Seminar 1

Start Assignment

- Due 14 Sep by 23:59
- Points 0
- Submitting a file upload
- File types pdf, r, and rmd
- Available after 1 Sep at 11:30

# Seminar 1: Alzheimer's Disease

## General Instructions

Written solutions to all the tasks must be submitted before the **deadline: Sunday September 14 23:59** (as pdf files in Canvas, preferably including the most important parts of your code).

You are also expected to prepare an oral presentation of your solution for each of the tasks. The presentation should aim at taking 15 minutes, to leave room for questions and discussion. At the seminar, a member of your group (randomly chosen by the teacher) will be asked to present the solution.

We expect each of you to spend around 15h on these tasks; please plan your time and meetings keeping this in mind.

Please note that the course book (and the lectures) are not your only sources of information. There are lots of information available in the research/educational literature and on the Internet about statistics as well as R and its possibilities. You are very much encouraged to try and find methods not mentioned in the book or the lectures.

## Background: Alzheimer's Disease and the Amyloid-Beta Hypothesis

Alzheimer's disease (AD) is the most common form of dementia. AD is characterised by memory loss, language problems, behavioural issues and more. The disease progresses gradually, ultimately leading to death.

For many years, the amyloid hypothesis has been the dominating explanatory theory for Alzheimer's disease mechanism. The theoretical model states that AD-associated neurodegeneration is caused by the accumulation of Beta-amyloid plaques in the brain. This has informed the development of preclinical models of disease, diagnostic biomarkers (See Figure 1) and more recently, the development of pharmaceutical therapies that target amyloid beta plaque accumulation.

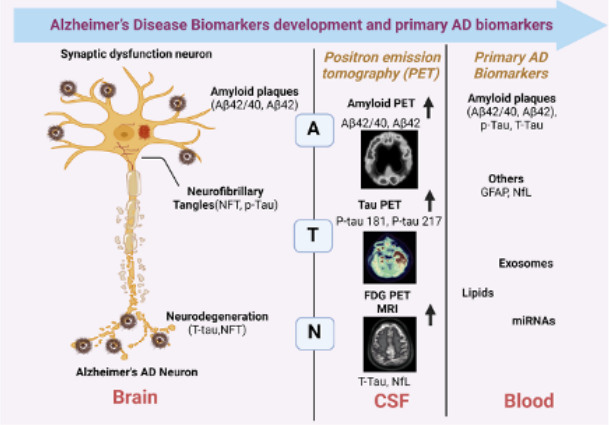


**Figure 1.** Primary Alzheimer's Disease Biomarkers (Predeepkiran *et al*. 2025; URL: https://www.aginganddisease.org/EN/10.14336/AD.2024.0286)

MECAS (Mamo-Eva-Chris-Adam-Smedby) Pharmaceutics Ltd is a pharmaceutical company carrying out discovery, research and development in the field of AD. The company was formed around pivotal research on the mechanism of Alzheimer's Disease (AD) and potential drug targets.

Together you form the data modelling team at MECAS Pharma.

## Task 1: The association between Alzheimer's Disease and Biomarkers NfL and B-tau181

MECAS Pharmaceutics has collaborated with an academic institution on a study of the association between diagnosed AD and the expression of the biomarkers Neurofilament light chain (NfL) and phosphorylated tau protein at threonine 181 (P-tau181).

The study data is given in the file, **Data_T1.csv (https://canvas.kth.se/courses/55919/files/9513181?wrap=1)** ↓ **(https://canvas.kth.se/courses/55919/files/9513181/download?download_frd=1)** . The study included 20 participants that suffered from AD (GROUP 1), and 20 matched controls that did not suffer from AD or any other form of diagnosed dementia (GROUP 0). Concentrations of NfL and P-tau181 were measured in both groups (unit: pg/mL).

1. Explore the dataset visually and by carrying out appropriate statistical analyses.
2. What conclusions can you draw?
3. How would you describe the statistical approach you are taking to solve this task? Discuss strengths and weaknesses of your approach.

# Task 2: Clinical evaluation of the candidate drug MECAS-123

MECAS Pharma has developed a promising molecule, named MECAS-123. Preclinical animal experiments have shown that MECAS-123 lowers the levels of NfL. This could potentially act to limit AD disease progression. MECAS-123 also has a favourable safety profile.

A parallel study will be carried out in matched AD patients. NfL and P-tau181 will be measured at three months following the experimental treatment and control treatment.

1. Based on the NfL data in AD patients in Task 1 - Perform power calculations and give a suggestion on the number of study participants given that we would like to detect a 30% reduction in the geometric mean of NfL.
2. Develop a statistical data analysis plan.
3. Retrieve the clinical data at your suggested sample size (N = study participants per study arm) using the CM2018 R package. You can do this using the following script:

```
# install and load necessary libraries
install.package(" devtools")
library(devtools)
install_github("adamdarwichkth/CM2018rpackage")
library(CM2018rpackage)
# run the data retrieval function to load data as dataframe,
# specifying the number of study participants per study arm (=N)
my_dataframe <- ad_trial_data(n_per_arm = N)
```

4. Carry out the statistical analysis based on your plan. What conclusions can be drawn?

# Task 3: The MECAS-123 dose finding study

The promising results of MECAS-123 and its favourable safety profile has led to the execution of a rudimentary dose finding study. A follow-up study was carried out to measure NfL in AD patients treated across four study arms: control treatment, experimental treatment (with MECAS-123) at a low, medium and high dose. A total of 80 patients were recruited (20 patients per study arm).

1. Propose a statistical analysis for multiple comparisons.
2. Carry out the analysis using the attached data (**Data_T3.csv (https://canvas.kth.se/courses/55919/files/9513182?wrap=1) (https://canvas.kth.se/courses/55919/files/9513182/download?download_frd=1)** ).
3. Plot the data, interpret and comment on the results.

# Task 4: Clinical outcome scales of cognitive impairment

The NfL biomarker data in Task 3 looked promising, to assess the clinical outcome MECAS Pharma needs a clinical measure of cognitive function.

The study in Task 3 was complemented with the Short Portable Mental Status Questionnaire (SPMSQ) as one of several potential clinical outcome scales. The SPMSQ is a short screening tool for testing the degree of cognitive impairment (Column Response with categories of Normal, Mild, Moderate and Severe cognitive impairment). The data is presented in **Data_T4A.csv (https://canvas.kth.se/courses/55919/files/9513183?wrap=1)** **(https://canvas.kth.se/courses/55919/files/9513183/download?download_frd=1)** .
Column Treatment, = 0 - control, 1 = low dose, 2 = medium dose, 3 = high dose.

1. Explore and present ways of visualising the questionnaire data across the four study arms.
2. Develop visualisations of SPMSQ given the biomarker data in Task 3.
3. SPMSQ data was collected for 25 individuals at two occasions to better understand inter-occasion variability (also called intra-individual variability). The data is available in **Data_T4B.csv (https://canvas.kth.se/courses/55919/files/9513184?wrap=1)** **(https://canvas.kth.se/courses/55919/files/9513184/download?download_frd=1)** . How would you visualise the SPMSQ for the same individuals at two occasions?