

## Principal Investigator

**First Name:** Andreas  
**Last Name:** Monsch  
**Degree:** PhD  
**Primary Affiliation:** University of Basel  
**E-mail:** [alex.knight@holmusk.com](mailto:alex.knight@holmusk.com)  
**Phone number:** +4179 514 3161  
**Address:** Burgfelderstrasse 101

**City:** Basel  
**State or Province:** N/A  
**Zip or Postal Code:** 4055  
**Country:** Switzerland  
**SCOPUS ID:** 7003427149

## General Information

### Key Personnel (in addition to PI):

**First Name:** Seyed  
**Last name:** Valizadeh  
**Degree:** PhD.  
**Primary Affiliation:** University Hospital Zurich  
**SCOPUS ID:** 55655146200

**Are external grants or funds being used to support this research?:** External grants or funds are being used to support this research.

**Project Funding Source:** Synapsis Foundation Research Grant; plus support from Holmusk, Singapore

**How did you learn about the YODA Project?:** Other

## Conflict of Interest

[https://yoda.yale.edu/system/files/2021\\_5\\_13\\_yoda\\_project\\_coi\\_form\\_am\\_aum.pdf](https://yoda.yale.edu/system/files/2021_5_13_yoda_project_coi_form_am_aum.pdf)  
[https://yoda.yale.edu/system/files/yoda\\_project\\_coi\\_form\\_sv\\_sv.pdf](https://yoda.yale.edu/system/files/yoda_project_coi_form_sv_sv.pdf)

## Certification

**Certification:** All information is complete; I (PI) am responsible for the research; data will not be used to support litigious/commercial aims.

**Data Use Agreement Training:** As the Principal Investigator of this study, I certify that I have completed the YODA Project Data Use Agreement Training

1. [NCT00575055 - ELN115727-302 - A Phase 3, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Efficacy and Safety Trial of Bapineuzumab \(AAB-001, ELN115727\) In Patients With Mild to Moderate Alzheimer's Disease Who Are Apolipoprotein E4 Carriers](#)
2. [NCT00574132 - ELN115727-301 - A Phase 3, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Efficacy and Safety Trial of Bapineuzumab \(AAB-001, ELN115727\) In Patients With Mild to Moderate Alzheimer's Disease Who Are Apolipoprotein E4 Non- Carriers](#)

**What type of data are you looking for?:** Individual Participant-Level Data, which includes Full CSR and all supporting documentation

## Research Proposal

### Project Title

Getting new insight out of failed AD trials

### Narrative Summary:

Alzheimer's Disease (AD) is the number one cause of death without a single approved disease-modifying drug. Over 20 years, thousands of patients have been enrolled in 156 clinical trials, costing USD 250B-300B. 152 turned out negative, bringing only 4 new drugs to market. Why did so many fail? Across all of the failed trials, there were successes at the individual patient level; but little is known about why such successes emerged. We will apply our semi-mechanistic machine learning method to identify novel "imaging biomarkers" to track disease progression; predict individual patients' clinical decline in the placebo arm; and evaluate drug response at the patient level in the treatment arm.

### Scientific Abstract:

#### Background

Alzheimer's disease (AD) is a significant health burden in Western societies, with a prevalence of 50M patients and an estimated care cost expected to rise from \$300B to \$1T by 2050. No disease-modifying drugs have been approved in the last 20 years, despite 156 clinical trials being conducted, at a total cost of up to \$300B. It is not clear why they failed, but it is likely that in all of them there were successes for individuals or subgroups of patients. We propose to investigate this by applying advanced machine learning (ML) approaches to the data from failed trials, in this case those of bapineuzumab.

#### Objectives

- Identify novel imaging biomarkers that predict cognitive decline
- Predict trajectories of clinical decline and biomarker changes in the placebo arm
- Evaluate drug response at the patient level in the treatment arm

#### Study Design

Apply ML methodologies to quantitatively assess clinically relevant biomarkers driving AD progression. Model disease outcome using the placebo arms, which will then enable us to model the effect of the drug at an individual patient level.

#### Participants

Subjects who participated in the vMRI sub-studies in either trial (ApoE4 carriers/non-carriers, active/placebo)

#### Main Outcome Measure(s)

Cognitive and functional scores

#### Statistical Analysis

We will model the relationship between the vMRI-derived variables and cognitive scores using General Linear Models, Structural Equation Modelling and related methods. This model will predict outcomes and identify the specific brain-regions with high relevance to the disease trajectory prediction.

### Brief Project Background and Statement of Project Significance:

AD is the most frequent cause of dementia in Western societies, with an estimated world-wide prevalence of 50M patients and almost 10M new patients each year, which is – with ageing populations – expected to rise significantly. WHO calculates the cost of care for AD as \$300B today, expected to rise to \$1T by 2050. AD is the number one cause of mortality for which not a single disease-modifying drug is approved.

Over the last 20 years, thousands of patients in different stages of AD enrolled in 156 clinical trials. These patients

were cognitively healthy subjects with a biomarker burden of AD, and/or with mild cognitive impairment (MCI), and/or with early Alzheimer's dementia, who were given new compounds or placebo for up to several years. Although these drug trials were carried out diligently and the pharmaceutical industry spent about \$250B to \$300B, 152 of them turned out negative, bringing only four molecules for symptomatic treatment to the market (Ceyzériat 2020).

Questions raised by these failed trials are: Was the dose too low? Were the patients enrolled at an early enough stage of the disease? Were the primary end points appropriate? Was the duration of the trial sufficient? Did subgroups of patients respond? Across all of the failed trials, there were successes at the individual patient level, where improvements of cognitive benefit or significant slowing of progression were observed. Unfortunately, little is published, let alone understood, about why such successes emerged. This project is setting out to change this.

The data generated in Clinical Trials (CT) and Real-World Evidence (RWE) are invaluable data assets for continued exploration of AD progression. The CT data is particularly valuable for its meticulously collected biomarkers at consistent time intervals from a selected cohort of patients. In order to develop personalized medicine for AD, modern research is moving away from broad clinical diagnoses, and towards trying to understand disease onset and progression at a more granular level. To achieve this, well-defined early biomarkers and related clinical tests are required that are reliable and indicative of disease progression. This understanding is still very poor and cannot be addressed with traditional approaches. Technological advances in AI and a better understanding of brain function opens up unprecedented opportunities to better understand AD pathophysiology.

Machine learning has allowed us to rapidly explore large amounts of data to derive insights on AD diagnostics and prognosis. These models have demonstrated high accuracy on the MCI – AD conversion probability but fail to provide insights into the clinically observable changes in disease progression (Rathore et al. 2017). Here we propose to use a variety of ML approaches to model disease progression and the efficacy of the drug in patient subgroups.

### **Specific Aims of the Project:**

Despite global efforts in R&D, AD remains an incurable disease, and we still have major gaps in our understanding of its onset and progression, calling for innovative strategies. Here we propose to address these gaps by using machine learning methodologies that allow for the quantitative assessment of clinically relevant biomarkers that drive AD progression in the form of clinical decline.

With Holmusk's machine learning models, we aim to:

- 1) identify novel brain biomarkers from MRI imaging that are predictors of cognitive decline,
- 2) predict clinical decline and biomarker changes in the placebo arm, and
- 3) evaluate drug response at the patient level in the treatment arm.

Our hypothesis is that the clinical trial results contain information about individual patient outcomes that can be revealed by machine learning analyses, giving insights into the biology of AD that can be used to refine the search for new treatments.

### **What is the purpose of the analysis being proposed? Please select all that apply.**

New research question to examine treatment effectiveness on secondary endpoints and/or within subgroup populations  
Participant-level data meta-analysis  
Participant-level data meta-analysis pooling data from YODA Project with other additional data sources  
Develop or refine statistical methods

## **Research Methods**

### **Data Source and Inclusion/Exclusion Criteria to be used to define the patient sample for your study:**

Data source:

NCT00575055 and NCT00574132 trials, including cognitive, functional, and health outcome assessments, imaging results and biomarker data.

**Inclusion criteria:**

Volumetric MRI variables available (sub-study participants) at baseline and at week 78.  
Cognitive scores available through week 78.

**Exclusion criteria:**

No exclusion criteria planned.

**Main Outcome Measure and how it will be categorized/defined for your study:**

The cognitive scores (ADAS-Cog, DAD, and NTB) are the primary outcome measures of interest.

**Main Predictor/Independent Variable and how it will be categorized/defined for your study:**

Volumetric MRI data including Brain Boundary Shift Integral, Volume Boundary Shift Integral, Hippocampal Boundary Shift Integral, Whole Brain Volume, Ventricular Volume, Hippocampal Volume. Any additional variables extracted from MRI data, including volumes, area, and thickness of specific brain regions.

Holmusk has previously shown that cognitive scores can be predicted from the imaging data from the ADNI study, specifically MRI, using a 3D CNN model. With this data set, as the raw images are not available, we will aim to model the relationship between changes in the imaging-derived variables and the selected outcome measures.

**Other Variables of Interest that will be used in your analysis and how they will be categorized/defined for your study:**

We will also include the available biomarkers, genetic data (ApoE genotype) and any demographic and medication information in our analysis.

**Statistical Analysis Plan:**

We will perform a 3-stage analysis of the available imaging-derived variables (and other relevant data) and their relationship to the outcome measures (primarily ADAS-Cog13).

1. Check for possible differences within groups
2. Develop a General Linear Model to find likely trends and correlations between changes in particular brain compartments and cognitive scores.
3. Use Structural equation modelling to extract the latent relationship between brain compartment and behavioural test result.

In addition to general linear models, we will also explore the use of other ML techniques such as Random Forests, and Deep Neural Networks ("deep learning").

Longitudinal brain biomarkers are combined with tabular patient data, such as medical history, medication, blood and CSF biomarkers, and cognitive sub-scores. Note that all relevant neuropsychological assessments, including cognitive tests, will be used for this section. The data is then processed in a biologically relevant manner and can be used for statistical and machine learning approaches (Russell et al., 2014).

The placebo arm of randomized CTs allows us to further understand the heterogeneity of the population under normal conditions (with no treatment). Our strategy consists of analysing and training machine learning algorithms on the placebo arm to predict AD outcomes. The model will provide insights into two clinical trial endpoints: change in clinical test scores and/or biomarkers.

We will then apply the algorithm to the treatment arm to observe the changes that occurred in the patients that were given the drug. We will then compare the two arms to get insights into the true drug effects using statistical methods including propensity score-based covariate balancing and causal inference.

The whole process can be summarised as follows:

- 1) Analytics and Biomarker extraction
  - a. Tabular data processing
  - b. Holmusk algorithm brain-region specific biomarker extraction

- 2) Placebo arm disease modelling
  - a. Data Analytics and Statistical analysis
  - b. Clustering Algorithm for patient sub-stratification
  - c. Training disease progression model: Patient-level cognitive decline and biomarker changes

- 3) Investigation of true drug effects
  - a. Evaluating treatment arm effects by measuring the difference between the true patient cognitive decline versus the placebo arm decline

In addition, we will employ our understanding of physiology and our previous experience with modelling AD progression based on brain imaging data. Physiologically informed models can be semi-mechanistic and therefore more explainable and informative than pure “black box” ML models. For example, Holmusk has developed a Quantitative Systems Pharmacology (QSP) model of AD progression which represents an unbiased and comprehensive description of the pathological pathways of AD. Holmusk has also developed a 3-D Convolutional Neural Network that can predict cognitive state and identify related brain regions by analysing multimodal imaging data from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database. While these models are not directly applicable to the current data set, they will inform our model design and the interpretation of the results.

Software Used:

Python

#### **Project Timeline:**

We estimate the following project milestone dates, assuming a data release date of 1st November 2021. We understand the data release agreement would expire one year later, which we think would be sufficient although should the schedule below slip, we would request a further year extension to ensure the analysis can be completed.

01/11/2021 Project start (assuming data available)  
01/01/2022 Data engineering and exploration completed  
01/04/2022 Analytics & Biomarker Extraction completed  
01/07/2022 Placebo arm modelling completed  
01/10/2022 Prediction of true drug effects completed  
01/11/2022 Data agreement expires (approximately)  
01/12/2022 First Draft Manuscript complete  
01/02/2023 Manuscript Submitted for Publication; Results reported back to YODA project

#### **Dissemination Plan:**

It is planned to submit and publish the results of this project in a peer-reviewed scientific journal, such as “Alzheimer’s Research and Therapy”. In addition, it is planned to submit a poster and/or platform presentation at the peer-reviewed Alzheimer Association International Conference (AAIC).

#### **Bibliography:**

Shinohara RT, Sweeney EM, Goldsmith J, Shiee N, Mateen FJ, Calabresi PA, et al. (2014) Statistical normalization techniques for magnetic resonance imaging. *NeuroImage Clin.*; 6:9–19.

Rathore S, Habes M, Iftikhar MA, Shacklett A, Davatzikos C. (2017) A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer’s disease and its prodromal stages. *Neuroimage.* 155:530–48.

Sarkar J, Dwivedi G, Chen Q, Sheu IE, Paich M, Chelini CM, et al. (2018) A long-term mechanistic computational model of physiological factors driving the onset of type 2 diabetes in an individual. DiStefano J, editor. *PLoS One.* 13(2):e0192472.

Roy N, Sarkar J. Generalized Modeling Framework for Disease Progression in Chronic & Mental Health. [https://www.holmusk.com/assets/whitepapers/Holmusk\\_Galaxis\\_1.0\\_04.pdf](https://www.holmusk.com/assets/whitepapers/Holmusk_Galaxis_1.0_04.pdf)

Ceyzériat, K., Zilli, T., Millet, P., Frisoni, G. B., Garibotto, V., & Tournier, B. B. (2020). Learning from the Past: A Review of Clinical Trials Targeting Amyloid, Tau and Neuroinflammation in Alzheimer’s Disease. *Current Alzheimer Research*, 17(2), 112–125

