Routine Blood Count and Chemistry can predict AD dementia risk up to ten years ahead

Amir Glik[1,2,3,4], Omry Arbiv[5,6], Keshet Pardo[2,4], Orit Raphaeli[5,3,6], Hayim Raclaw[3], Roy Kait[3], Ofri Kait[3], Anat Goldstein[5,3,6] and Chen Hajaj[5,3,6]

[1]Cognitive Neurology Clinic, Rabin Medical Center – Beilinson Hospital, Petach Tikva 4941492, Israel, [2]Department of Neurology, Rabin Medical Center – Beilinson Hospital, Petach Tikva 4941492, Israel, [3]ALZAI Health Corp, Toronto, Ontario M5X 1G5, Canada, [4]Faculty of Medicine, Tel Aviv University, Tel Aviv 6997801, Israel, [5]Industrial Engineering and Management, Ariel University, Ariel 40700, Israel, [6]Data Science and Artificial Intelligence Research Center, Ariel University, Ariel 40700, Israel.

Introduction

- The pre-clinical phase of AD is long, and contains any of the pathological changes that characterize AD disease.
- With the advent of novel and effective treatments in the early stages of Alzheimer's disease (AD), and the identification of modifiable risk factors for the condition, it has become paramount to develop predictive methods for identifying at-risk individuals and detecting AD patients as early as possible.
- A good screening tool should be able to effectively detect the disease in question, be safe, reasonably priced, and lead to improved healthcare outcomes.
- The current diagnostic tests are invasive, expensive, and inaccessible.
- The use of machine learning (ML) has become very popular in medical research across all fields, and that includes AD. Since ML models requires large data sets, many of these studies are based on data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), and relay heavily on cognitive tests and scores, demographic data, CSF biomarkers and imaging data.

Aim

- To utilize ML models on routine blood results (CBC. Routine chemistry) done for various other reasons, to predict AD dementia risk in cognitive healthy population.
- Enabling a new risk prediction tool, which can be deployed on mass populations instantly.

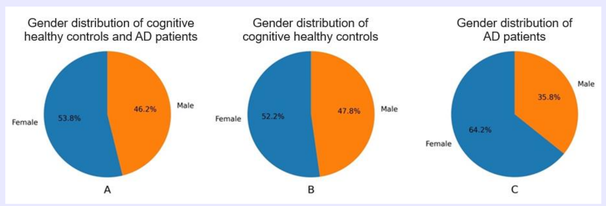

Figure 2. Gender distribution of cognitive healthy controls and AD patients in our cohort.

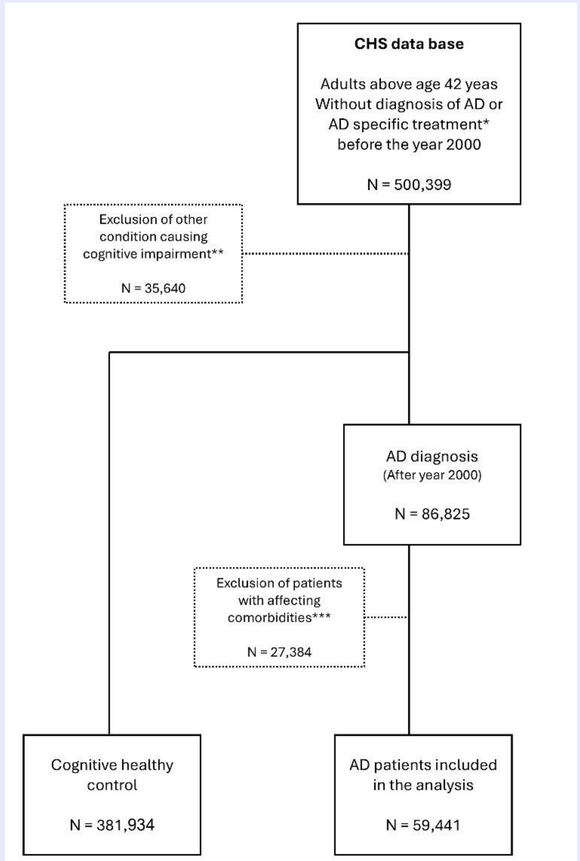

Figure 1. Study Flow Chart.

Abb: CHS, Clalit health services; AD Alzheimer's disease

*AD specific treatment: Donepezil, Galantamine, Rivastigmine, and Memantine

**Other conditions causing cognitive impairment: CNS malignancy, Creutzfeldt-Jacob Disease, Drug & Alcohol-Induced Dementia, other neurodegenerative disorders (Parkinson's disease, Parkinson's disease dementia, corticobasal degeneration, frontotemporal dementia, Lewy body dementia, Huntington), HIV and Multiple sclerosis

***Affecting comorbidities prior to AD diagnosis: Epilepsy, Ischemic stroke, Intracerebral Hemorrhage, Psychotic & affective syndromes.

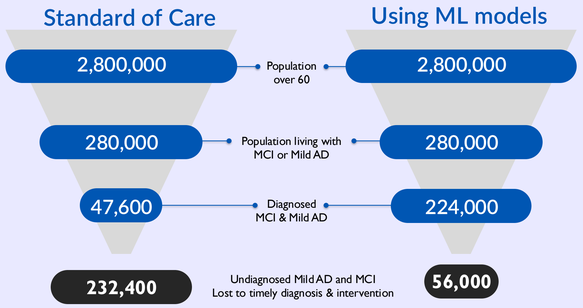

Figure 4. Using ML models may increase MCI/Mild AD diagnosis rates from ~17% to ~80%. (numeric values represent the number of subjects)

History (years)	Prediction (years)	Sensitivity	Specificity	PPV	NPV	AUC
1	1	0.72	0.74	0.28	0.95	0.80
	3	0.79	0.71	0.27	0.96	0.83
	8	0.68	0.75	0.23	0.95	0.78
3	1	0.76	0.76	0.30	0.96	0.83
	3	0.78	0.74	0.29	0.96	0.84
	8	0.76	0.72	0.22	0.97	0.80
8	1	0.74	0.73	0.27	0.96	0.81
	3	0.72	0.75	0.26	0.96	0.80
	8	0.82	0.81	0.28	0.98	0.89

Table 1. Different models results based on blood exam history and AD risk prediction horizons.

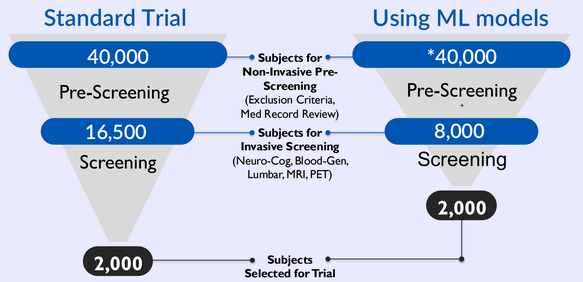

Figure 5. ML models may reduce screen failure rates in clinical trials and shorten recruitment periods. (numeric values represent the number of subjects)

* The utilization of ML models will diminish the necessity to screen 40,000 individuals to a sample size of 8,000 subjects.

Methods

- We interrogated a community cohort from the Clalit health care services above the age 42y during the years 2000 to 2022.
- Each subject had at least one routine blood count and basic chemistry and up to ten consecutive blood exams done on a yearly basis.
- All blood exams were taken while the subjects was cognitively healthy.
- AD Diagnosis was based on the 2011 NIA guidelines.
- Different predictive ML models were trained for different combinations of historical periods (1 to 10 years) and prediction horizons (1 to10 years).
- 80% of the dataset was used for model training, while the remaining 20% was reserved for testing.
- For each model, we calculated the Sensitivity, Specificity, PPV, NPV and AUC.

Results

- 500,399 adults above the age of 42 in the year 2000 were allocated in the database, we excluded patients with other conditions causing cognitive impairment (Fig. 1).
- 59,441 AD patients were included in the analysis.
- 381,934 subjects were cognitive healthy controls.
- The Gender distribution was consistent with numbers described in other cohorts (Fig 2).
- The Highest Sensitivity and Specificity were around 80%, the NPV was above 95% (Table 1).

Discussion

- Using routine blood samples data, done for various other reasons, may improve AD risk prediction, stratified by years.
- The model can predict who will develop the disease and when it will happen.
- The model can be deployed on databases without requiring patients to visit the clinic.
- Integrating this model into public health systems will augment the diagnosis rates of MCI and Mild AD dementia, thereby facilitating access to novel disease-modifying medications (Fig. 4) and enabling the postponement of disease progression.
- The model can be incorporated as a pre-screening tool in preventive or therapeutic clinical trials, thereby reducing the screen failure rate and saving time and money (Fig. 5).