

# Use of Language to Identify Stages of Dementia



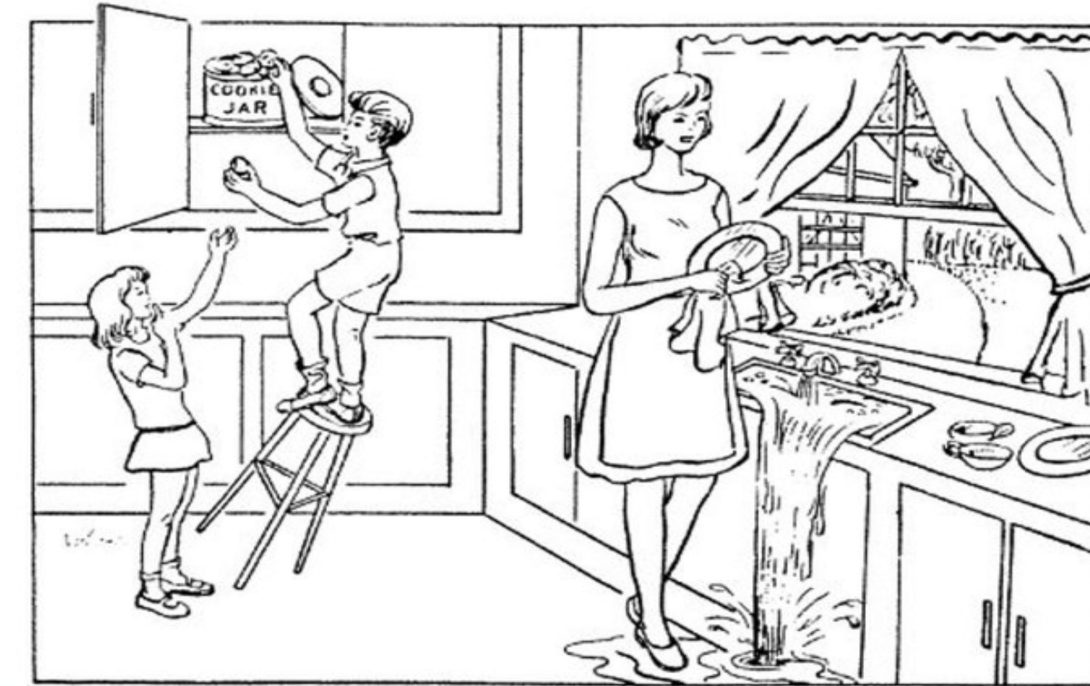
By: Xander Brick, Sophia Hill & Gbenuola Olaiya

Advisor: Dr. Joel Greenhouse

Client: Dr. Davida Fromm

## Background and Introduction

Dementia is a general term for loss of memory and other mental abilities severe enough to interfere with everyday life. Dementia describes a wide range of symptoms, such as difficulty with memory and language.



Cookie Theft Picture (CTP)

The Cookie Theft Picture (CTP) is a commonly used assessment tool to investigate markers of dementia.

**Project Goal: Evaluate language in dementia by analyzing core lexicon usage in the CTP task.**

**Task: Subjects were asked to describe what they see in the picture**

## Data

### Groups

- Control
- Mild Cognitive Impairment (MCI)
- Probable Alzheimer's Disease (Probable AD)

Groups	Number of Patients
Control	243
Probable AD	234
MCI	41

### Data Generation

- Generated by performing Computerized Language Analysis (CLAN) on the audio files of subject descriptions of the Cookie Theft Picture above.

### Data Description

- 518 Pittsburgh Individuals
- Key Variables
  - Time
    - Total time it took each individual to describe the picture
    - Time until each core lexicon word was said
  - Demographic Information: Age, Gender
  - Frequency of Core Lexicon Usage

## Methods

### Kaplan Meier Survival Curve

- Visualizes the distribution of time until word is first said.

### Log-Rank Test

- Test to compare the distribution of time across groups

## Results

We first wished to analyze how long the three main groups, those in the Control, MCI, and Probable AD groups, took to say certain words when describing the picture. (Figures 1,2,3)

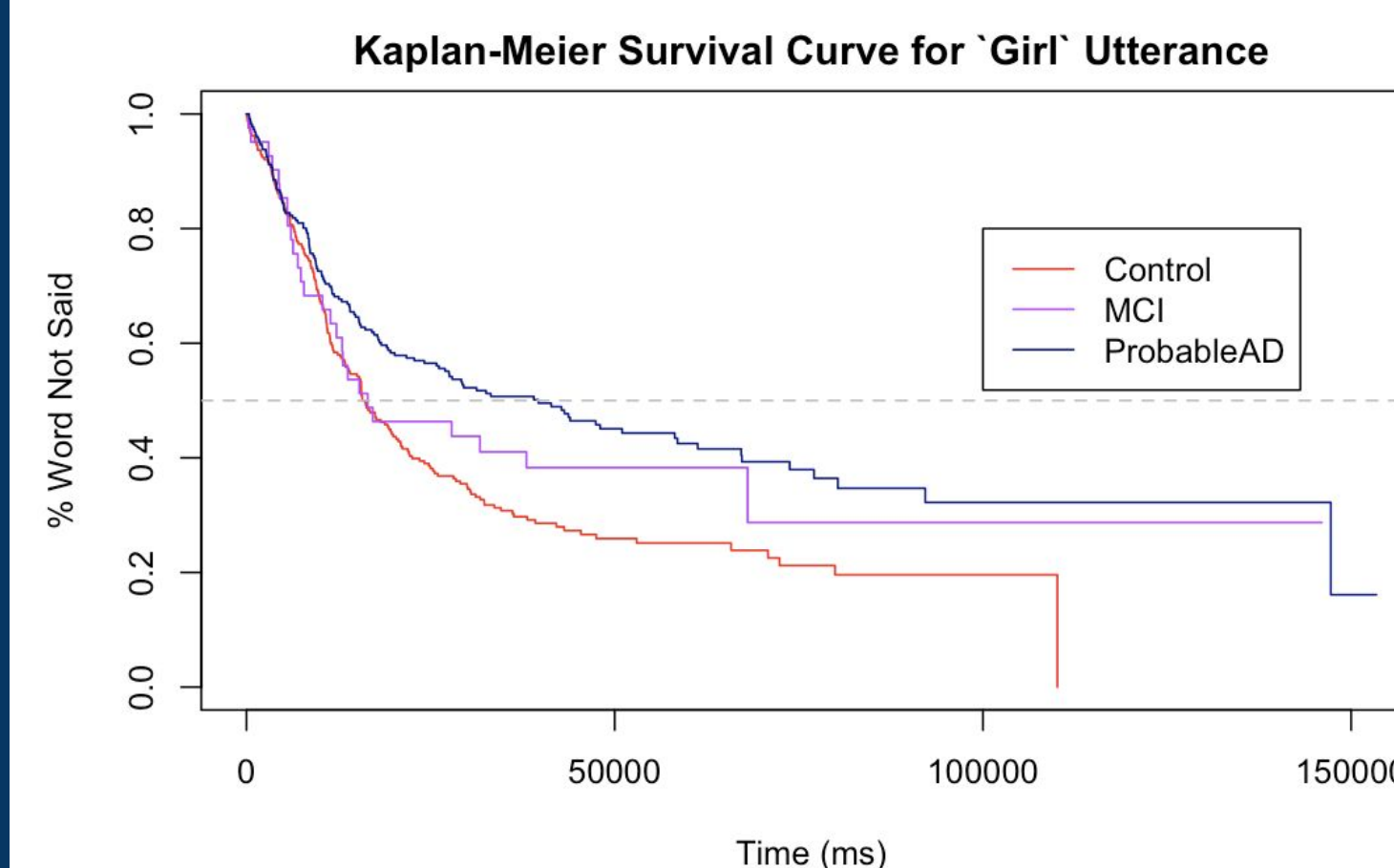


Figure 1

Curve showing the time until the word "girl" was said through the three groups. We note that the Probable AD group starts falling behind the others, then later the MCI group also slows down compared to the control. Log-Rank Test:  $p \leq 5 \times 10^{-4}$

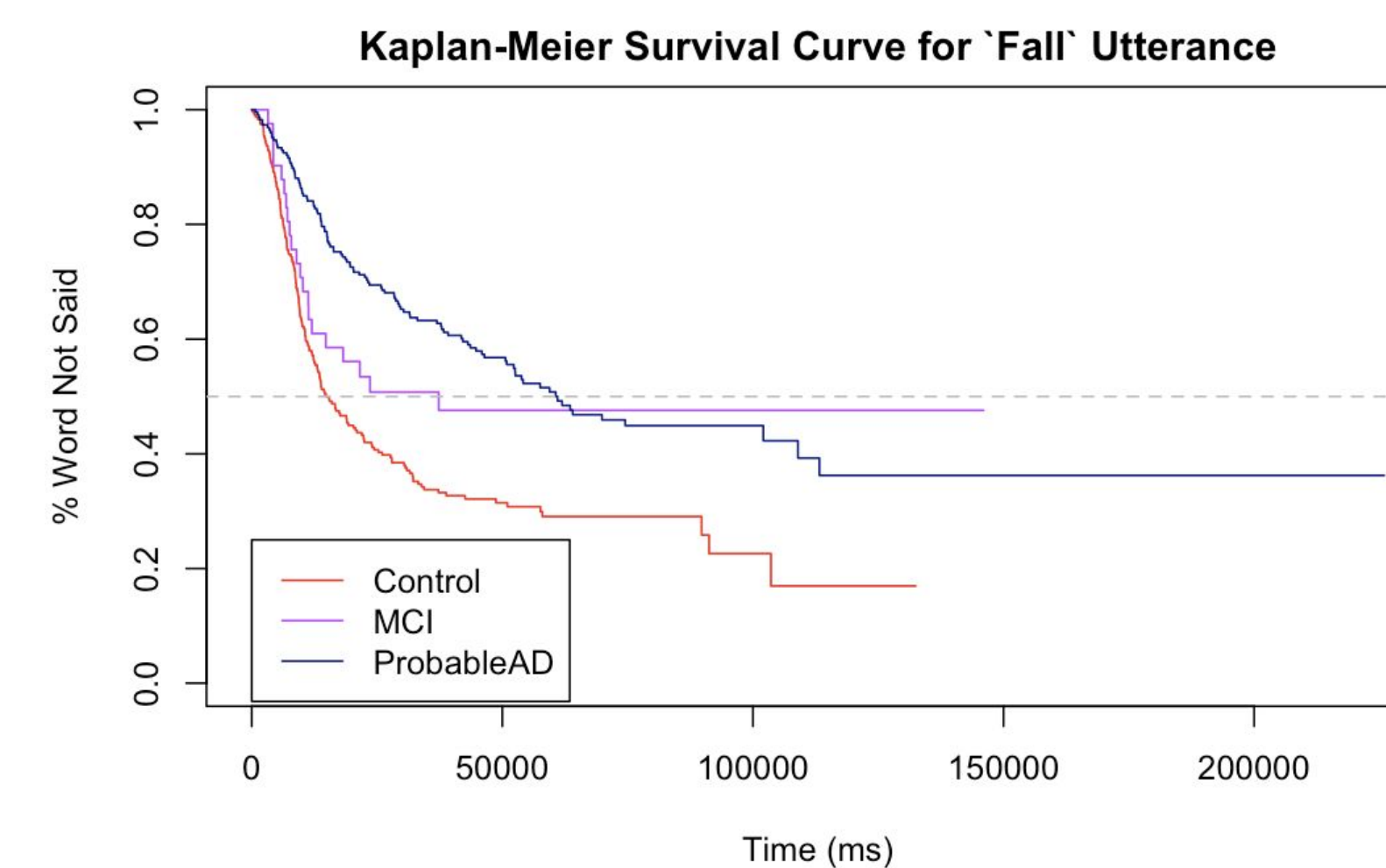


Figure 2

Here we once again see the same pattern for the word "fall", where there is an almost immediate difference between the Control and ProbableAD groups, followed by MCI which stays with control until eventually falling behind. Log-Rank Test:  $p \leq 2 \times 10^{-8}$

Next we wanted to see if there was a difference in time until a word was said within regions of the image between the Control and Probable AD Groups, (Figure 4)

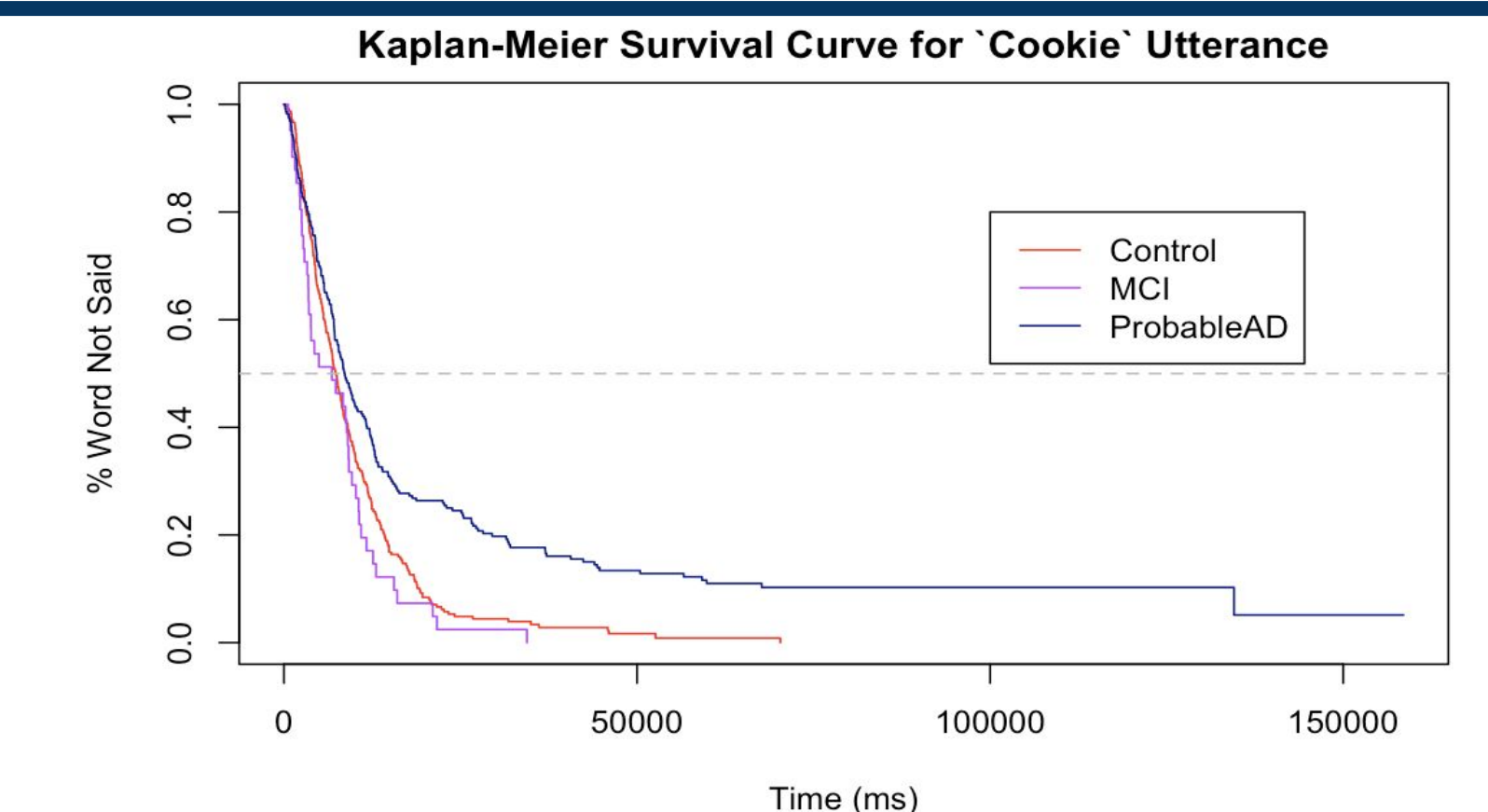


Figure 3

For the first half of the people in each group, there doesn't appear to be that substantial difference in the time until ~30% of people in the probable AD group stopped responding, where they slowed down substantially compared to other groups. Log-Rank Test:  $p \leq 8 \times 10^{-7}$

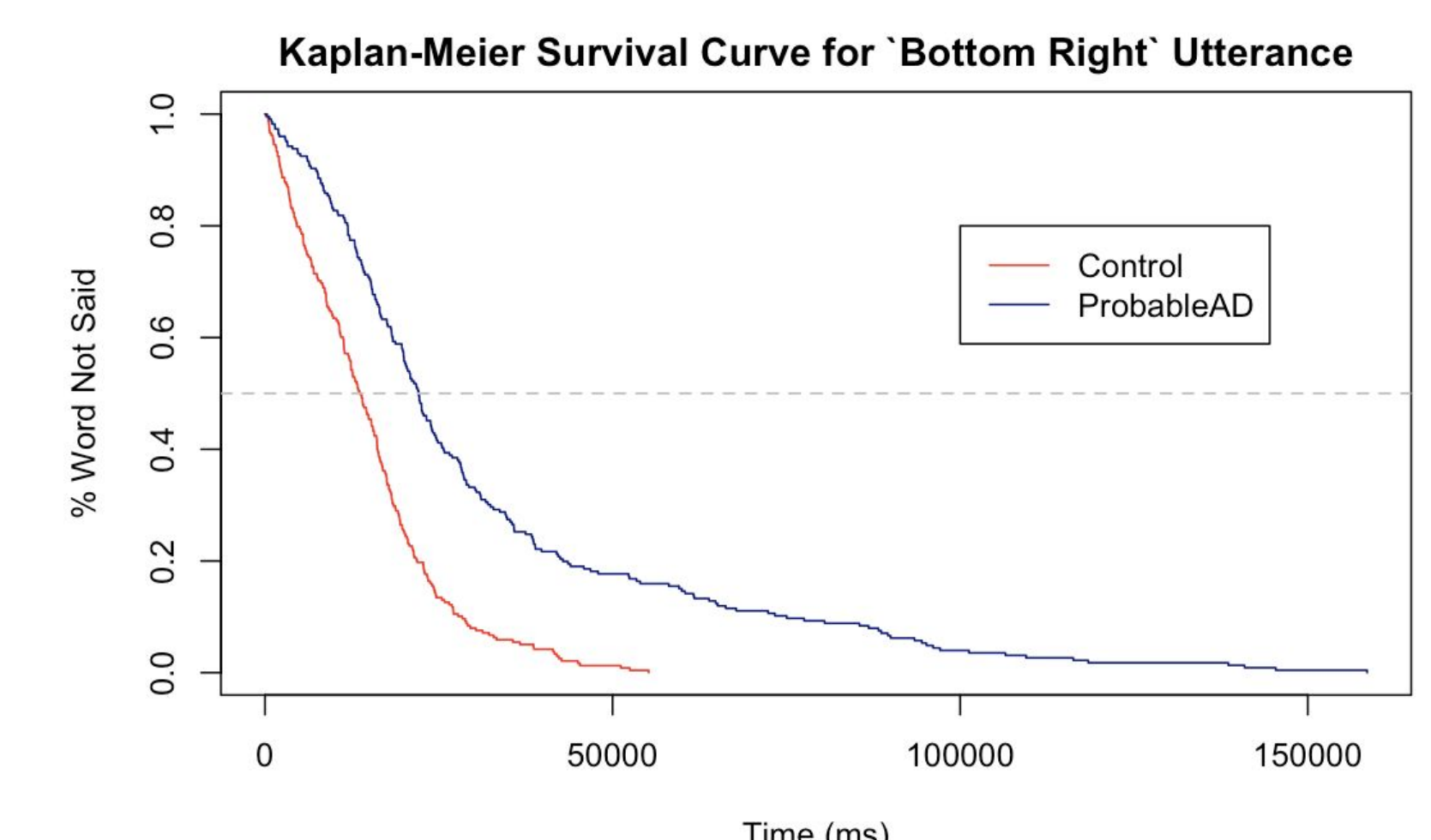


Figure 4

We see a substantial difference between the two groups in terms of time until saying one of the words in the bottom right (Run, Sink, Water, Flow). Log-Rank Test:  $p \leq 2 \times 10^{-16}$

## Conclusions

### Main findings

- Statistically significant differences in time-until words between groups
- Difference in time until different quadrants of the image

### Implications

- Use of language may help identify and track development of dementia

### Further Research

- Replicating results with other datasets

## References

- TalkBank Repository. Retrieved from <https://talkbank.org/>
- Dalton SG, Stark BC, Fromm D, Apple K, MacWhinney B, Rensch A, Rowedder M. **Validation of an Automated Procedure for Calculating Core Lexicon From Transcripts.**