

Exploring the Effect of Eligibility Criteria on AD Severity and Severe Adverse Event in Eligible Patients

Aokun Chen

Department of Health Outcome
and Biomedical Informatics,
College of Medicine
University of Florida
Gainesville, FL, USA
chenaokun1990@ufl.edu

Qian Li

Department of Health Outcome
and Biomedical Informatics,
College of Medicine
University of Florida
Gainesville, FL, USA
liqian@ufl.edu

Elizabeth Shenkman

Department of Health Outcome
and Biomedical Informatics,
College of Medicine
University of Florida
Gainesville, FL, USA
eshenkman@ufl.edu

Yonghui Wu

Department of Health Outcome
and Biomedical Informatics,
College of Medicine
University of Florida
Gainesville, FL, USA
yonghui.wu@ufl.edu

Yi Guo

Department of Health Outcome
and Biomedical Informatics,
College of Medicine
University of Florida
Gainesville, FL, USA
yiguo@ufl.edu

Jiang Bian

Department of Health Outcome
and Biomedical Informatics,
College of Medicine
University of Florida
Gainesville, FL, USA
bianjiang@ufl.edu

Abstract—Clinical trials were vital tools to prove the effectiveness and safety of medications. To maximize generalizability, the study sample should represent the sample population and the target population. However, the clinical trial design tends to favor the evaluation of drug safety and procedure (i.e., internal validity) without clear knowledge of its penalty on trial generalizability (i.e., external validity). Alzheimer's Disease (AD) trials are known to have generalizability issues. Thus, in this study, we explore the effect of eligibility criteria on the AD severity patients and the severe adverse event (SAE) among the eligible patients.

Keywords—clinical trial, eligibility criteria, machine learning, SHAP value.

I. INTRODUCTION

Clinical trials, especially randomized clinical trials, were vital tools to prove the effectiveness and safety of medications. [1] In clinical trials, three major populations are involved: target population, study population, and study sample. The target population refers to the patients who benefited from the trailed medication. The study population is the patients that fulfill the eligibility criteria (EC). The study sample is the patients enrolled in the study. To maximum generalizability, the study sample should represent the study population and the target population. However, the clinical trial design tends to adopt eligibility criteria in favor of the evaluation of drug safety and procedure (i.e., internal validity). [2–4] Alzheimer's Disease (AD) trials are known to have the issue of the mismatched study population and target population. [5–7] In this study, we seek to identify ECs that would lead to the mismatch in AD severity and the likelihood of experiencing severe adverse events (SAEs). In this process, we built an EC bank, extracted a cohort of AD patients

with AD severity information, and explored the effect of ECs on AD severity and SAE with machine learning (ML) methods and SHapley Additive exPlanation (SHAP) values. The outline of our system is shown in Fig. 2.

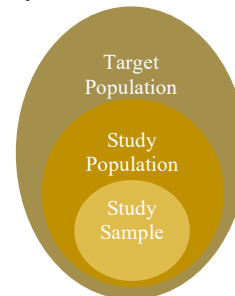


Fig. 1. Populations in the clinical trial.

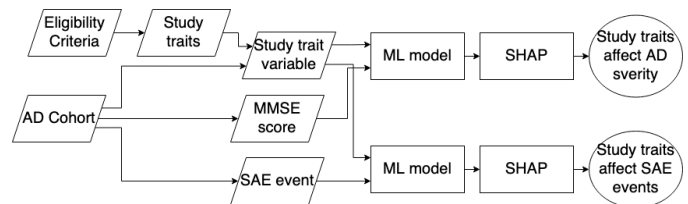


Fig. 2. System outline

II. METHOD

We extracted ECs from 13 phase III and IV AD trials on the safety and effectiveness of four FDA-approved AD medications. ECs were decomposed into study traits that each defines a single

demographic (e.g., age, race-ethnicity), diagnoses (e.g., history of diabetes), lab test results (e.g., fasting glucose), or other measurement used to define the study population.

From the University of Florida Integrated Data Repository (IDR), we extracted a cohort of AD patients from 2012 to 2021. We measured AD severity with the Mini-Mental Statement Examination (MMSE) score (0-30, cognitive impairment < 25) [8] extracted from the clinical narratives with an NLP tool [9].

The SAEs reported by the drug label of the corresponding medication were extracted within the SAE observation window between the first prescription of the AD medication and the 30 days after the last prescription (Fig. 3).

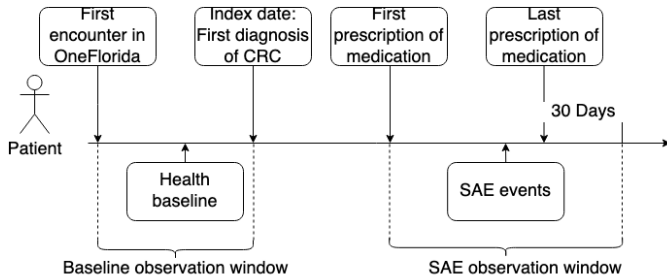


Fig. 3. SAE observation window.

For each study trait, we extracted each patient's eligibility based on their EHR and record them as study trait variables. The corresponding study trait variable will be coded 0 if the patient did not match the measurement described by the corresponding study trait and 1 otherwise.

In our analysis, we first modeled the association between study traits & AD severity, and study traits & number of SAE events with three ML methods: extreme gradient boost regressor (XGBoost) [10], Adaboost regressor (Adaboost) [11], and support vector machine regressor (SVM) [12] stratified by age, sex, race, and ethnicity. We performed the analysis on a trial basis with the patients on the trialed medication in our AD cohort. Then, we calculated the SHAP value with the best-performed model. SHAP value is an artificial intelligence (AI) explanation tool that measured the contribution of each variable to the change in the outcomes. [13] With the SHAP value, we could quantify the change that study traits imposed on AD severity (measured by MMSE score) and the number of SAEs. The study traits with major effects on AD severity and the number of SAEs were identified with the mean absolute SHAP value.

III. RESULTS

In total, we extracted 1,024 AD patients with MMSE scores. The average age of our AD cohort was 80.63 years with the majority being female patients (64.7%). Regarding race and ethnicity, most of the patients were white (76.7%) and non-Hispanic (94.7%) (TABLE I).

We excluded the trials that contains too few patients and analyzed the eligibility criteria from 4 trials: NCT00097916, NCT00235716, NCT00305903, and NCT00478205. The SHAP values were calculated based on the XGBoost models as it outperformed the rest in all 4 trials on mean absolute error.

TABLE I. THE DEMOGRAPHIC OF THE AD COHORT WITH MMSE

	AD patient with MMSE (n = 1,026)
Age (SD ^a)	80.63 (9.40)
Sex	
female (%)	664 (64.7%)
male (%)	362 (35.3%)
Race	
White (%)	787 (76.7%)
Black (%)	191 (18.6%)
Other (%)	39 (3.8%)
Unknown (%)	9 (0.9%)
Ethnicity	
Hispanic (%)	42 (4.1%)
Non-Hispanic (%)	972 (94.7%)
Unknown (%)	12 (1.2%)

^a SD: standard deviation.

From NCT00097916, we identified 5 study traits (shown in Fig. 4, i.e., mental disorder, disorder of endocrine system, kidney disease, Donepezil prescription, and liver disease) affecting the AD severity of the patients, and 7 study traits (shown in Fig. 5, i.e., disorder of coronary artery, peripheral nerve system, drug dependence, cardiac dysrhythmias, renal failure, transient ischemic attack, diabetes, and asthma) related to the number of SAEs (absolute SHAP value > 0.1).

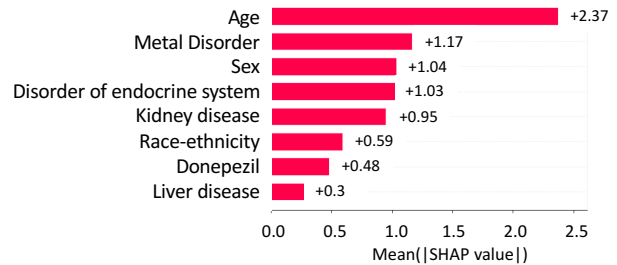


Fig. 4. Absolute SHAP value of variables on AD severity in NCT00097916

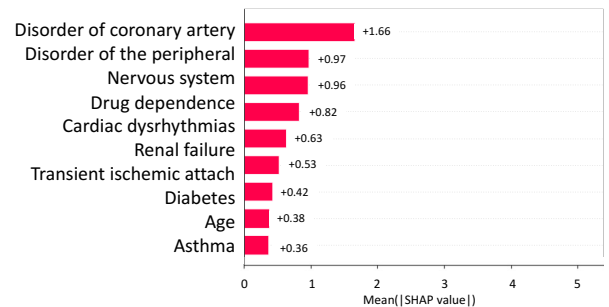


Fig. 5. Absolute SHAP value of variables on SAE in NCT00097916

From NCT00235716, we identified 3 study traits (shown in Fig. 6 and Fig. 7, i.e., delirium, vitamin B12 deficiency, and Donepezil prescription) affecting both the AD severity and the number of SAEs (absolute SHAP value > 0.1).

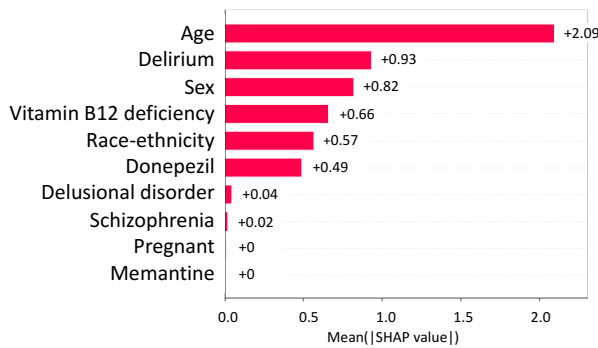


Fig. 6. Absolute SHAP value of variables on AD severity in NCT00235716

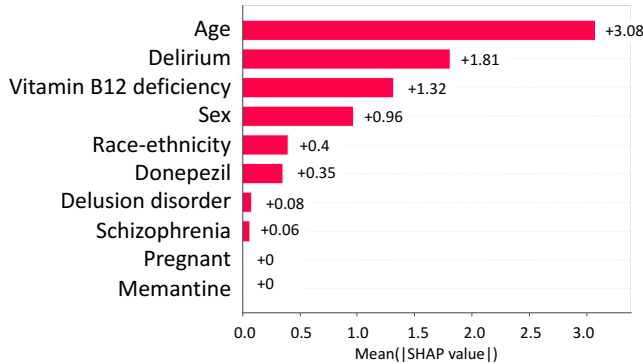


Fig. 7. Absolute SHAP value of variables on SAE in NCT00235716

From NCT00305903, we identified 5 study traits (shown in Fig. 8 and Fig. 9, i.e., kidney disease, disorder of endocrine system, myocardial infarction, disorder of respiratory system, and disorder of digest system) affecting both the AD severity and the number of SAEs (absolute SHAP value > 0.1).

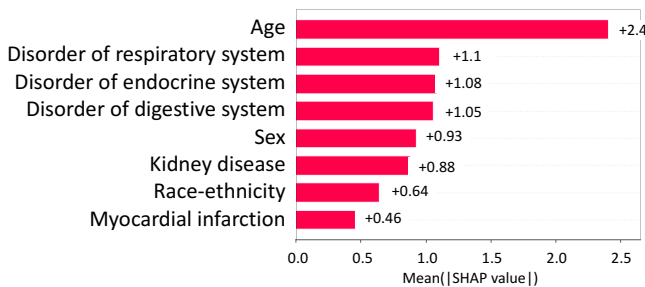


Fig. 8. Absolute SHAP value of variables on AD severity in NCT00305903

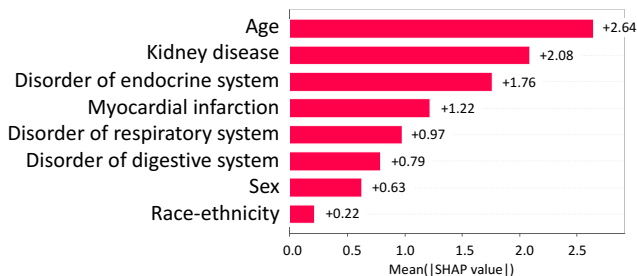


Fig. 9. Absolute SHAP value of variables on SAE in NCT00305903

From NCT00478205, we identified 8 study traits (shown in Fig. 10, i.e., kidney disease, Benzodiazepine prescription, Memantine prescription, disorder of endocrine system, disorder of respiratory system, sleep disorder, urinary incontinence) affecting the number of SAEs (absolute SHAP value > 0.1).

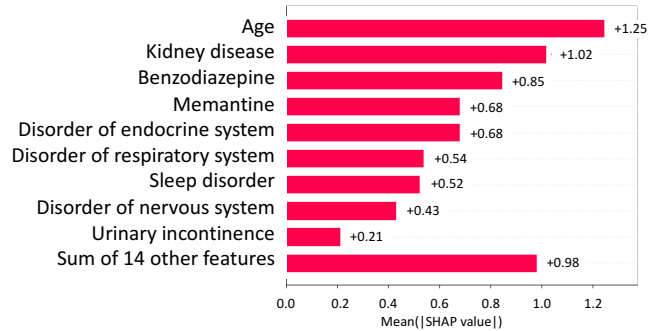


Fig. 10. Absolute SHAP value of variables on AD severity in NCT00478205

IV. CONCLUSION

In this study, we identified 14 unique study traits that affected the patients' AD severity and 20 unique study traits that affected the number of SAEs. This analysis served as the first step in the development of a future system to guide AD clinical trial design.

REFERENCES

- [1] P. M. Spieth, A. S. Kubasch, A. I. Penzlin, B. M.-W. Illigens, K. Barlinn, and T. Siepmann, "Randomized controlled trials - a matter of design," *Neuropsychiatr. Dis. Treat.*, vol. 12, pp. 1341–1349, Jun. 2016.
- [2] S. Karim, Y. Xu, S. Kong, O. Abdel-Rahman, M. L. Quan, and W. Y. Cheung, "Generalisability of common oncology clinical trial eligibility criteria in the real world," *Clin. Oncol. (R Coll. Radiol.)*, vol. 31, no. 9, pp. e160–e166, Sep. 2019.
- [3] J. He, D. R. Morales, and B. Guthrie, "Exclusion rates in randomized controlled trials of treatments for physical conditions: a systematic review," *Trials*, vol. 21, no. 1, p. 228, Feb. 2020.
- [4] Q. Li, Y. Guo, Z. He, H. Zhang, T. J. George Jr, and J. Bian, "Using real-world data to rationalize clinical trials eligibility criteria design: A case study of Alzheimer's disease trials," *AMIA Annu. Symp. Proc.*, vol. 2020, pp. 717–726, 2020.
- [5] R. Raman *et al.*, "Disparities by Race and Ethnicity Among Adults Recruited for a Preclinical Alzheimer Disease Trial," *JAMA Netw Open*, vol. 4, no. 7, p. e2114364, Jul. 2021.
- [6] J. Shin and P. M. Doraiswamy, "Underrepresentation of African-Americans in Alzheimer's trials: A call for affirmative action," *Front. Aging Neurosci.*, vol. 8, p. 123, Jun. 2016.
- [7] K. Fredriksen-Goldsen *et al.*, "Design and development of the first randomized controlled trial of an intervention (IDEA) for sexual and gender minority older adults living with dementia and care partners," *Contemp. Clin. Trials*, vol. 128, no. 107143, p. 107143, Mar. 2023.
- [8] M. F. Folstein, L. N. Robins, and J. E. Helzer, "The Mini-Mental State Examination," *Arch. Gen. Psychiatry*, vol. 40, no. 7, p. 812, Jul. 1983.
- [9] Z. Chen *et al.*, "Assess the documentation of cognitive tests and biomarkers in electronic health records via natural language processing for Alzheimer's disease and related dementias," *Int. J. Med. Inform.*, vol. 170, no. 104973, p. 104973, Feb. 2023.
- [10] T. Chen and C. Guestrin, "XGBoost," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA, 2016.
- [11] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.
- [12] M. Awad and R. Khanna, "Support Vector Regression," in *Efficient Learning Machines*, Berkeley, CA: Apress, 2015, pp. 67–80.

- [13] S. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” *arXiv [cs.AI]*, 22-May-2017.