

Title

Use of Blood Pressure Measurements Extracted from the Electronic Health Record in Predicting Alzheimer's Disease: A Retrospective Cohort Study at Two Medical Centers

**Authors & Affiliations:** Donna Tjandra<sup>1</sup>, Raymond Q. Migrino<sup>2,3</sup> MD, Bruno Giordani, PhD<sup>4</sup>, Jenna Wiens, PhD<sup>1</sup>

<sup>1</sup>Computer Science and Engineering, University of Michigan, Ann Arbor, MI, USA;

<sup>2</sup>Phoenix Veterans Affairs Healthcare System, Phoenix, AZ 85012;

<sup>3</sup>University of Arizona, College of Medicine - Phoenix, Phoenix, AZ 85012;

<sup>4</sup>Neuropsychology Program, Department of Psychiatry, and Michigan Alzheimer's Disease Research Center, University of Michigan Ann Arbor, MI, USA

**Corresponding Author:** Jenna Wiens, PhD

2260 Hayward Street

University of Michigan,

Ann Arbor, MI 48109

email: [wiensj@umich.edu](mailto:wiensj@umich.edu)

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/alz.12676](https://doi.org/10.1002/alz.12676).

This article is protected by copyright. All rights reserved.

**DISCLOSURE STATEMENT:** The authors have nothing to disclose

**Key Words:** Alzheimer's disease, electronic health record, machine learning, risk prediction, blood pressure trajectory

**Word Count:** 1,500/1,500

## **Abstract**

**Introduction.** Studies investigating the relationship between blood pressure (BP) measurements from electronic health records (EHRs) and Alzheimer's disease (AD) rely on summary statistics, like BP variability, and have only been validated at a single institution. We hypothesize that leveraging BP trajectories can accurately estimate AD risk across different populations.

**Methods.** In a retrospective cohort study, EHR data from Veterans Affairs (VA) patients were used to train and internally validate a machine learning model to predict AD onset within five years. External validation was conducted on patients from Michigan Medicine (MM).

**Results.** The VA and MM cohorts included 5,488 and 1,372 patients, respectively. Model performance using BP trajectories was modest but comparable (AUROC=0.64 [95%CI=0.54-0.73] (VA) vs. AUROC=0.66 [95%CI=0.55-0.76] (MM)).

**Conclusion.** Approaches that directly leverage BP trajectories from EHR data could aid in AD risk stratification across institutions.

**Word count:** 150/150

## **Introduction**

Cardiovascular risk factors are associated with increased risk of Alzheimer's disease (AD)<sup>1</sup> and could be exploited to predict AD risk years before clinical diagnosis. Current work investigating relationships between AD risk and blood pressure (BP) primarily focuses on prospectively collected clinical trial data<sup>2-8</sup>. These trials are often limited in the amount of longitudinal data collected since sample sizes are frequently small (e.g., <1,000 individuals)<sup>2-3</sup>; the followup period is short (e.g., <three years)<sup>4-6</sup>; or the measurements are sparse (e.g., once every five years)<sup>7-8</sup>. To address these limitations, recent work used electronic health record (EHR) data, which contain decades of longitudinal data for thousands of patients, from a single institution to study the relationship between AD and BP<sup>9</sup>. However, assumptions were made regarding what aspects of the longitudinal measurements were important (e.g., BP variability). Moreover, these associations have only been validated on a single healthcare system. We expand on prior work

by i)using machine learning approaches to directly leverage longitudinal measurement trajectories without making assumptions and ii)validating on an external cohort. We hypothesize that i)using EHR-based BP trajectories can help predict AD onset and ii)performance will be on par with summary statistics.

## **Methods**

We describe the inclusion/exclusion criteria applied to two populations to obtain our study cohorts. This study was approved by the Institutional Review Boards at the University of Michigan and Phoenix Veterans Affairs.

### **Study Cohorts**

***Development and Internal Validation Cohort.*** We trained on patients from the five hospitals of the Veterans Affairs (VA) Veterans Integrated Service Network (VISN, formerly VISN18) region's Cerner EHR instance (Cerner Corporation, North Kansas city, MO)<sup>10</sup>. Patient timelines were aligned at the first available encounter between 68-72 years (i.e., we predicted AD onset for all patients at the first encounter between 68-72). We aligned to control for age and because AD incidence rises at 75 years<sup>11</sup>. We excluded patients labeled with AD at alignment, patients with <five years of followup without an AD label, and like previous work,<sup>9</sup> patients with <three measurements before alignment. Only patients with hypertension, identified by ICD (International Classification of Diseases) codes recorded  $\leq$ two years before alignment,<sup>12</sup> were included to control for the effect of hypertension on AD risk.

***External Validation Cohort.*** The external validation cohort included patients from Michigan Medicine's (MM) Epic EHR instance (Epic Systems Corporation, Verona, WI) aligned

between 68-72. To control for data availability, we only included patients with  $\geq 35$  BP measurements, since VA patients generally had more measurements than MM patients. 35 was chosen to match the average number of measurements over a patient's entire history before alignment. Cohort characteristics were compared between populations, using  $\chi^2$  tests for statistical significance.

**Outcome.** The model was trained to predict AD onset within five years of alignment. AD onset was labeled using a cohort discovery tool based on ICD codes for AD<sup>13</sup>.

### **Model Development & Evaluation**

**Data Preprocessing.** We focused on features that were easy to collect or recorded routinely, retrospectively extracting only those in **Table A3** for each cohort. Starting from alignment and going backward in time at six month intervals through five years of historical data, we recorded patient demographics (e.g., race), and the most recent vital sign measurements (e.g., the latest systolic BP measurement). For any missing measurement during the six-month interval, the previous value was carried forward, and a binary indicator denoting imputed values was set to 1. We also included the number of measurements taken within the six-month interval. Five years of historical lookback was chosen based on data availability.

**Model Training.** Our model, "BP Trajectories", was a long-short term memory (LSTM)<sup>14</sup> recurrent neural network trained with the development cohort, using features from "General Information" and "Trajectories" (**Table A3**). We also trained two baseline LSTMs. The first, "BP Stats", used all features from **Table A3** except BP trajectories. The second, "No BP", excluded both BP trajectories and summary statistics<sup>4,5,9,15,16</sup> (**Appendix A4**). Neural network parameters were optimized using Adam. We used early stopping and random search in the hyperparameter space for model selection (more in **Appendix A5**).

**Internal and External Validation.** On the held-out VA validation and external MM cohorts, we measured the AUROC and AUPR (area under the receiver operating characteristics and precision-recall curves, respectively), reporting empirical 95% confidence intervals from 1,000 bootstrapped samples. Statistical significance was tested using a resampling test<sup>17</sup>.

**Model Interpretation.** To visualize which trajectories “BP Trajectories” found important, we plotted the median and interquartile range of trajectories among the predicted high-risk ( $\geq 90^{\text{th}}$  risk percentile) and low-risk patients ( $\leq 10^{\text{th}}$  risk percentile) for the internal and external validation cohorts. Given the expected population differences in the internal and external cohorts (e.g., fraction female), we conducted a permutation importance analysis on MM to measure the extent of feature utilization by the model, reporting 95% confidence intervals (**Appendix A6**).

## Results

**Cohort Characteristics.** The development and internal validation cohorts included 5,488 and 1,372 patients, respectively. Across cohorts, 2.4% of patients experienced the outcome. The external MM cohort included 1,201 patients, 2.5% of which experienced the outcome (**Figure A1**). The internal and external validation cohorts had several differences (**Appendix A2**), including the proportion female (internal=2.6%; external=54.0%), proportion with dyslipidemia (internal=29%; external=70%), and median diastolic BP (internal=78mmHg [IQR=71-85]; external=72mmHg [IQR=67-78]). However, the median systolic BPs were similar (internal=135mmHg [IQR=123-148]; external=135mmHg [IQR=126-143]).

**Internal and External Validation.** On the VA validation cohort, “BP Trajectories” achieved AUROC=0.64, 95%CI=0.54-0.73 and AUPR=0.04, 95%CI=0.03-0.06. For MM, “BP Trajectories”

performed similarly (AUROC=0.66, 95%CI=0.55-0.76; AUPR=0.06, 95%CI=0.03-0.12) (**Figure 1**). Performance was comparable to “BP Stats” (**Figure 1**).

**Model Interpretation.** On the VA validation cohort, systolic and diastolic BP were consistently higher in the high-risk group than the low-risk group (**Figure 2**). MM had similar patterns for systolic BP. MM predictions were mostly affected by vitals, with 95%CI=0.043-0.188, 95%CI=-0.009-0.036 describing the drop in AUROC and AUPR, respectively, from permutation importance (**Table A4**).

## Discussion

We developed a model using EHR-based BP trajectories to predict AD onset. It was developed and internally validated using VA data and externally validated using MM data. The model had modest discriminative performance. Despite differences in health systems, EHR platforms, and patient populations, our patterns in discriminative performance (e.g., “BP Trajectories” was comparable to “BP Stats”) and observed high-/low-risk BP trajectory patterns were consistent, demonstrating the potential to generalize.

Our results highlight the potential for model interoperability across institutions.

Interoperability is a known challenge in healthcare due to differences in patient populations and medical/coding practices, and few studies have addressed it<sup>18</sup>. As hospitals collect more data,<sup>19</sup> addressing this will be crucial to improve healthcare practices.

Like previous work, high-risk patients generally had higher systolic and diastolic BP and greater variability<sup>3-9</sup>. Leveraging trajectories provides additional information by highlighting when these differences matter most. Summary statistics do not readily capture such differences.

While others used time-series trajectories to predict AD onset using datasets like ADNI,<sup>20</sup> we used EHR data. EHRs contain longitudinal data from routine clinical care (e.g., vitals). This

allows us to potentially develop screening tools for the general population to identify high-risk individuals in any healthcare system without requiring invasive tests. Such individuals could be recruited to clinical trials like ADNI for biological validation, providing more high-risk individuals for enrollment.

Although we focused on BP, this approach could potentially be used to identify meaningful longitudinal relationships among other features (e.g., image-based biomarkers). With BP, summary statistics important for predicting AD onset were established<sup>4,5,9,15,16</sup>. For new features, these statistics may not be established. Since using trajectories performed comparably to summary statistics for BP, one could potentially benefit from using trajectories when the longitudinal relationship between the feature and risk of AD onset is unknown, stimulating hypothesis generation.

Our study has several limitations. The cohort discovery tool used to identify AD patients had sensitivity=0.70<sup>13</sup>. We excluded patients with <35 BP measurements for MM, so the approach may not generalize to individuals with fewer measurements. Finally, the amount of BP data available was limited in terms of the lookback period and frequency of measurement, with missing rates between 10%-70%, (**Table A2**). While high, this reflects clinical practice, where patients may not have routinely collected measurements. We hypothesize that longer lookbacks and more routinely collected measurements could improve performance. However, we are encouraged that, despite high rates of missingness, the model could capture a predictive signal. We demonstrated the potential of using EHR-based BP trajectories to predict AD onset, and our results were consistent across two EHRs. Leveraging EHR trajectories could help uncover the relationship between BP and AD by discovering unrecognized temporal patterns. Such analyses could apply to other features/diseases without knowing which summary statistics are predictive.



## Acknowledgements

This research program is supported by the NIH/NIA funded Michigan Alzheimer's Disease Center ([5P30AG053760](#)) and the National Science Foundation ([IIS 2124127](#)). The views and conclusions in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the NSF, the NIH, the VA, or the United States Government. The authors have no conflicts of interest to report.

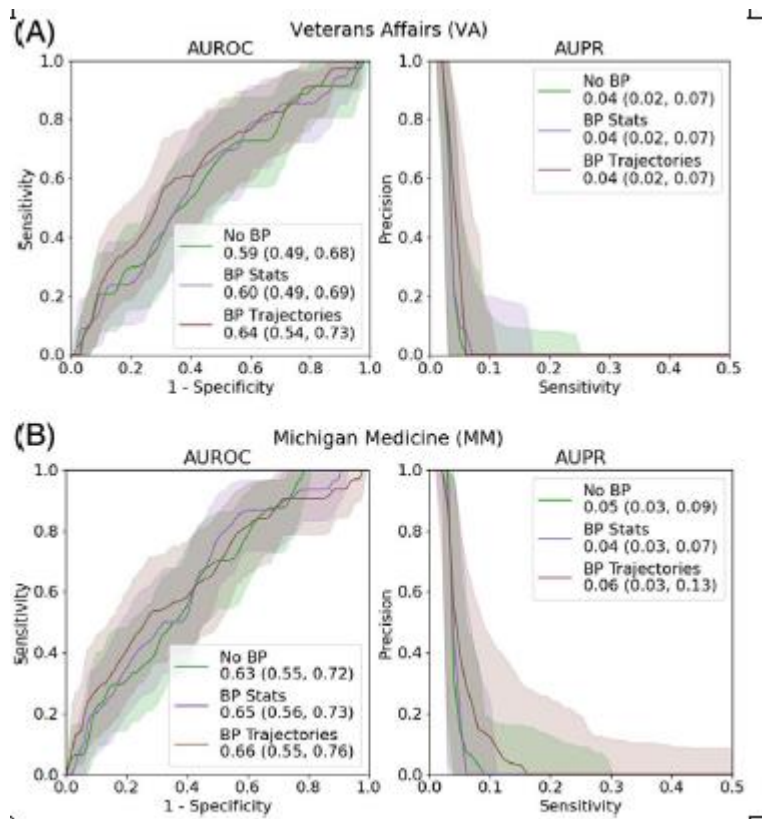
## References

1. Lennon MJ, Makkar SR, Crawford JD, Sachdev PS. Midlife hypertension and Alzheimer's disease: a systematic review and meta-analysis. *J Alzheimer's Dis.* 2019;71(1):307-16.
2. Hanon O, Latour F, Seux ML, et al. Evolution of blood pressure in patients with Alzheimer's disease: a one year survey of a French Cohort (REAL. FR). *J Nutr Health Aging.* 2005;9(2):106.
3. Ninomiya T, Ohara T, Hiraoka Y, et al. Midlife and late-life blood pressure and dementia in Japanese elderly: the Hisayama study. *Hypertension.* 2011;58(1):22-28.
4. Lattanzi S, Luzzi S, Provinciali L, Silvestrini M. Blood pressure variability predicts cognitive decline in Alzheimer's disease patients. *Neurobiol Aging.* 2014;35(10):2282-2287.
5. Wijsman LW, De Craen AJ, Muller M, et al. Blood pressure lowering medication, visit-to-visit blood pressure variability, and cognitive function in old age. *Am J Hypertens.* 2016;29(3):311-318.

6. de Heus RA, Olde Rikkert MG, Tully PJ, Lawlor BA, Claassen JA, NILVAD Study Group. Blood pressure variability and progression of clinical Alzheimer disease. *Hypertension*. 2019;74(5):1172-1180.
7. Kivipelto M, Helkala EL, Laakso MP, et al. Midlife vascular risk factors and Alzheimer's disease in later life: longitudinal, population based study. *BMJ*. 2001;322(7300):1447-1451.
8. Abell JG, Kivimäki M, Dugravot A, et al. Association between systolic blood pressure and dementia in the Whitehall II cohort study: role of age, duration, and threshold used to define hypertension. *Eur Heart J*. 2018;39(33):3119-3125.
9. Yoo JE, Shin DW, Han K, et al. Blood pressure variability and the risk of dementia: A Nationwide Cohort Study. *Hypertension*. 2020;75(4):982-990.
10. VA Informatics and Computing Infrastructure (VINCI), VA HSR RES 13-457, U.S. Department of Veterans Affairs. (2008). Retrieved July 15 2021, from <https://vaww.VINCI.med.va.gov>
11. Alzheimer's Association. 2021 Alzheimer's Disease Facts and Figures. *Alzheimers Dement*. 2021;17(3)
12. Pace R, Peters T, Rahme E, Dasgupta K. Validity of health administrative database definitions for hypertension: a systematic review. *Can J Cardiol*. 2017;33(8):1052-1059.
13. Tjandra D, Migrino RQ, Giordani B, Wiens J. Cohort discovery and risk stratification for Alzheimer's disease: an electronic health record-based approach. *Alzheimer's Dement.: Transl. Res. Clin. Interv*. 2020;6(1):e12035.

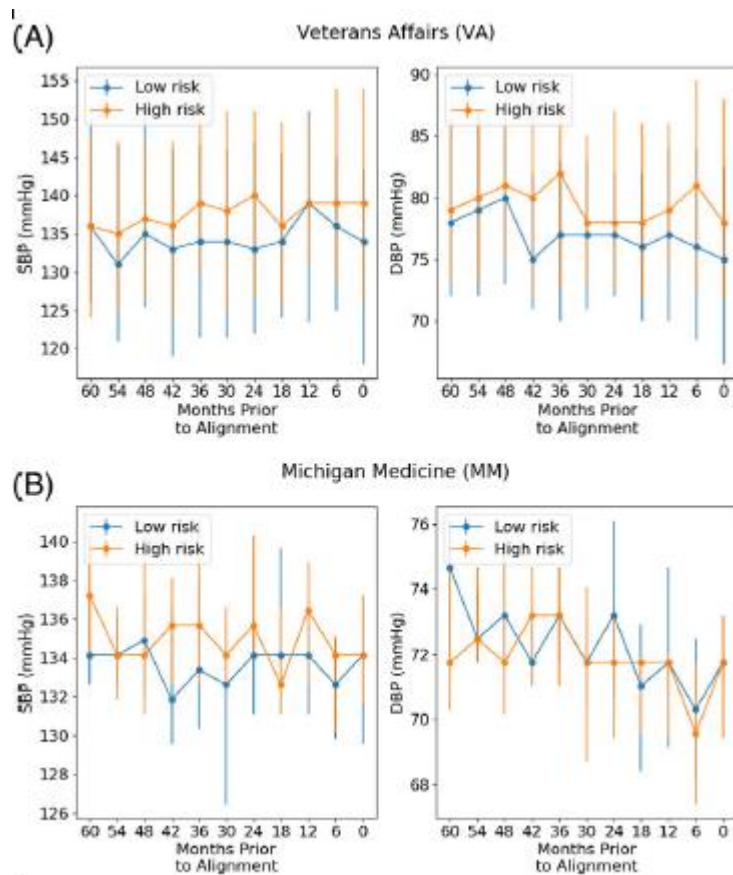
14. Hochreiter S, Schmidhuber J. LSTM can solve hard long time lag problems. *Conference Proceedings of Advances in neural information processing systems*. 1997;473-479.
15. Nwabuo CC, Yano Y, Moreira HT, et al. Association between visit-to-visit blood pressure variability in early adulthood and myocardial structure and function in later life. *JAMA Cardiol*. 2020;5(7):795–801. doi:10.1001/jamacardio.2020.0799
16. Yano Y. Visit-to-visit blood pressure variability—what is the current challenge?. *Am J Hypertens*. 2017;30(2):112-114.
17. Dixon PM. Bootstrap resampling. *Encyclopedia of Environmetrics*. Wiley. 2006:212-220.
18. Naveed A, Hu YF, Sigwele T, Mohi-Ud-Din G, Susanto M. Similarity Analyzer for Semantic Interoperability of Electronic Health Records Using Artificial Intelligence (AI). *J Sci Eng*. 2019;1(2):53-8.
19. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44-56.
20. Moore PJ, Lyons TJ, Gallacher J, Alzheimer's Disease Neuroimaging Initiative. Random forest prediction of Alzheimer's disease using pairwise selection from time series data. *PLoS One*. 2019 Feb 14;14(2):e0211558.

## Figure Captions



**Figure 1.** Overall performance. We show the AUROC and AUPR curves for all three models for each dataset. Error bars represent 95% confidence intervals. (a) VA dataset; (b) MM dataset

Author Ms



**Figure 2.** Median trajectories. We show the median trajectories for systolic and diastolic BP (abbreviated SBP and DBP, respectively) within different risk groups from the model trained on time series data. Risk groups were defined by the 10th and 90th percentiles of the model's predictions. Error bars represent interquartile ranges. (a) VA dataset; (b) MM dataset