

# Silent Risks: Validating a Dementia-Prediction Tool on the MIMIC-IV Dataset

<sup>1</sup>Heran Russom, <sup>2</sup>Kuk Jin Jang, <sup>2</sup>Kevin B. Johnson

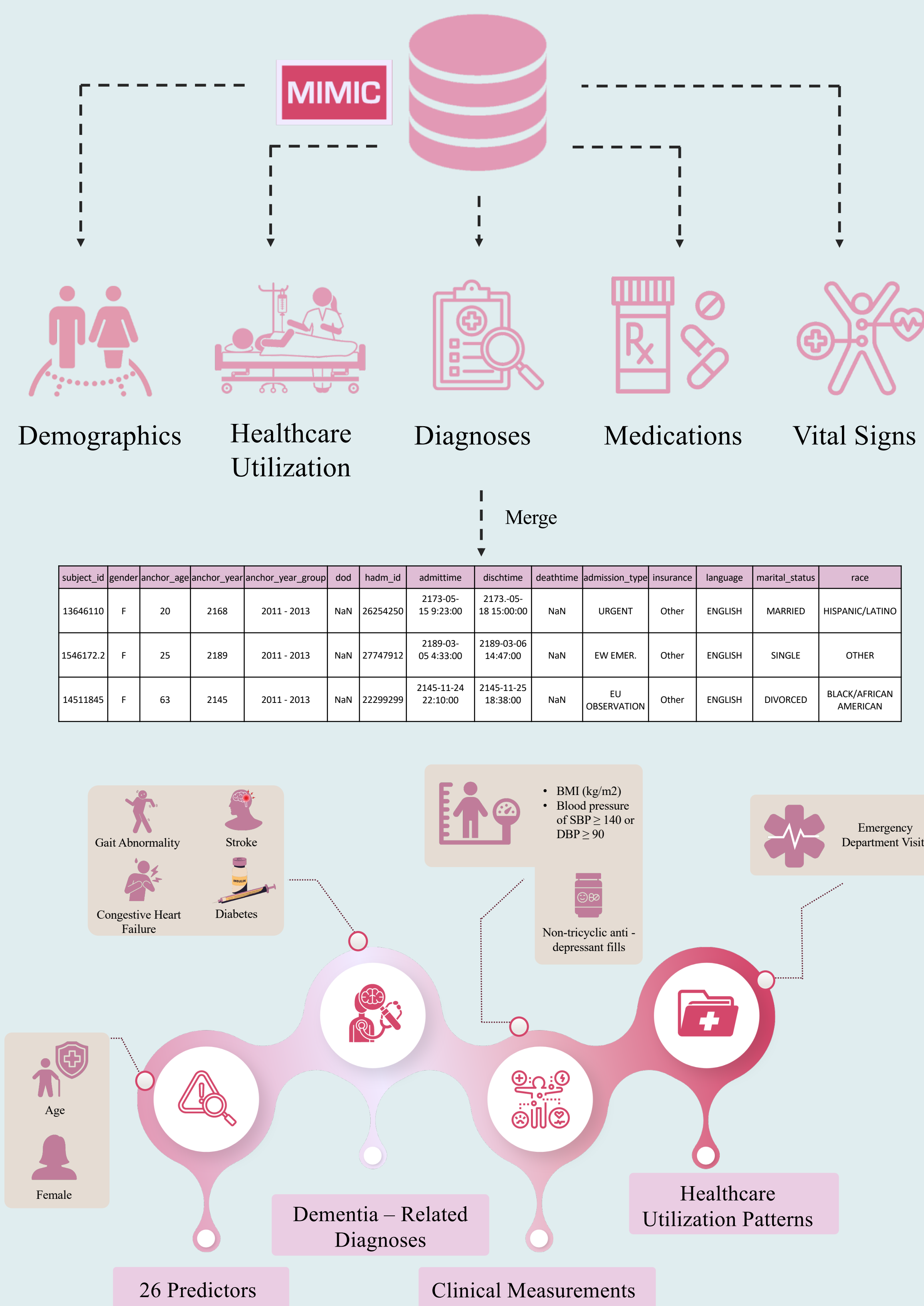
<sup>1</sup>Inclusion, Diversity, Equity, and Learner Research Program, University of Pennsylvania, <sup>2</sup>Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania



## 1 Introduction

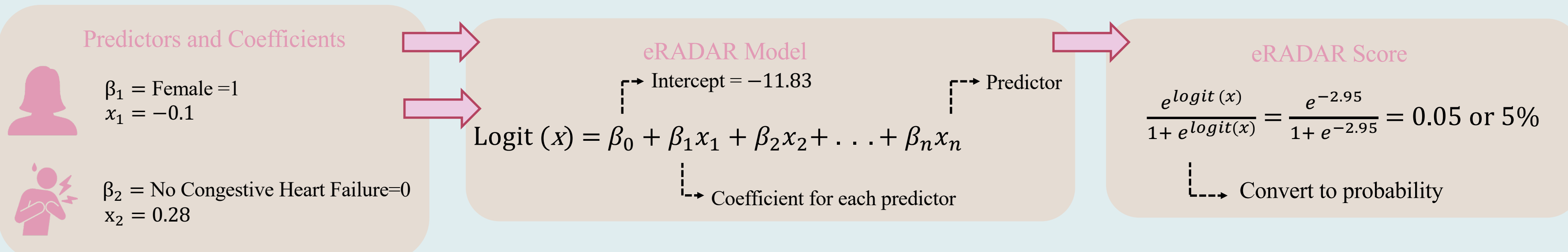
- Dementia is a chronic condition affecting over five million people in the U.S.
- Lack of diagnosis results in delayed treatment and intervention.
- The Electronic Health Records Risk of Alzheimer's and Dementia Assessment Rule (eRADAR) uses electronic health records (EHR) to find dementia risks in patients.
- The Medical Information Mart for Intensive Care (MIMIC-IV) dataset has EHR data from 2008 to 2019, providing patient diagnoses, demographics, vital signs, drugs, etc.
- We applied the logistic regression model of eRADAR to create a score for each patient in the MIMIC dataset and investigated the trends in the data.

## 2 MIMIC - IV



**Figure 1.** MIMIC – IV tables were queried and preprocessed for the relevant information necessary to apply the e-RADAR restricted model.

## 3 e-RADAR



**Figure 2.** The eRADAR restricted model was applied using a logistic regression equation to calculate each patient's score.

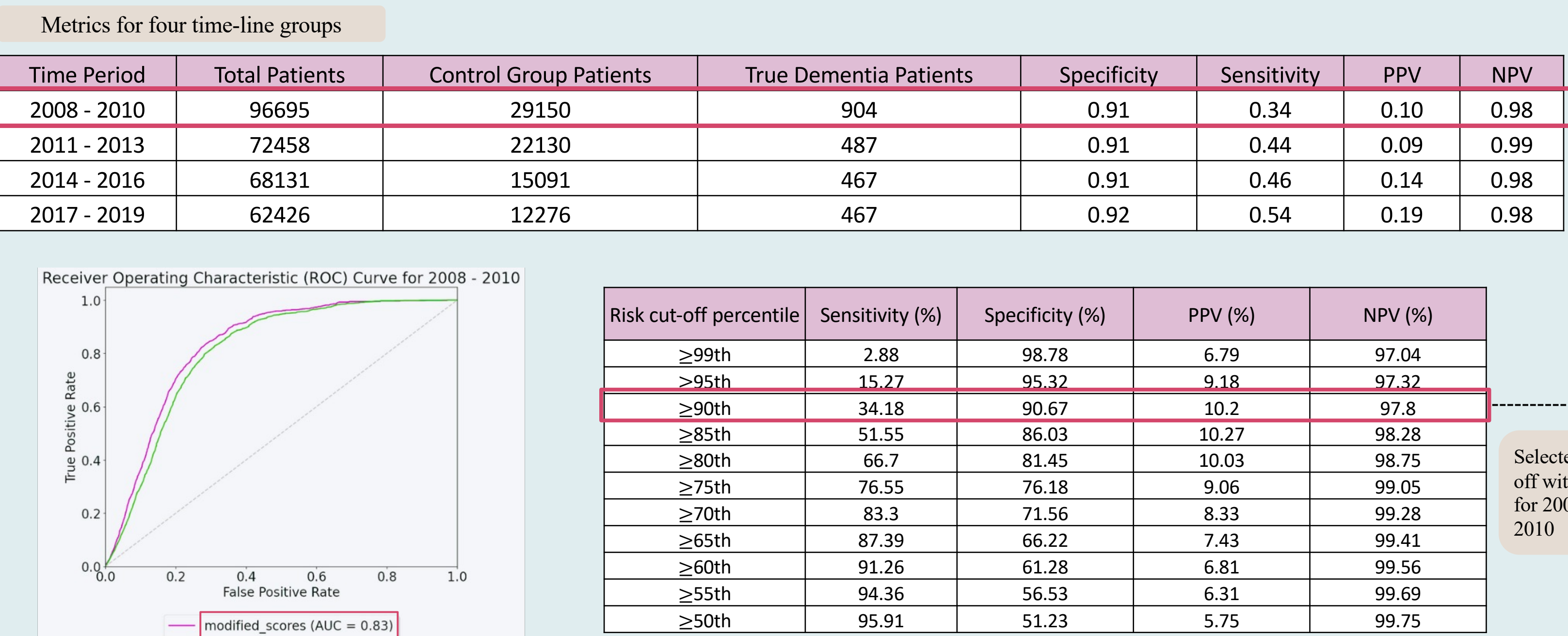
Subject_id	Actual Diagnosis	90th percentile	e-RADAR score	Result
19805351	True	0.06	0.22	True Positive
19796596	False	0.06	0.05	True Negative
19802563	False	0.06	0.08	False Positive
19800340	True	0.06	0.02	False Negative

**Table 1.** Example metrics for the 90<sup>th</sup> percentile threshold cut-off of high-risk patients.

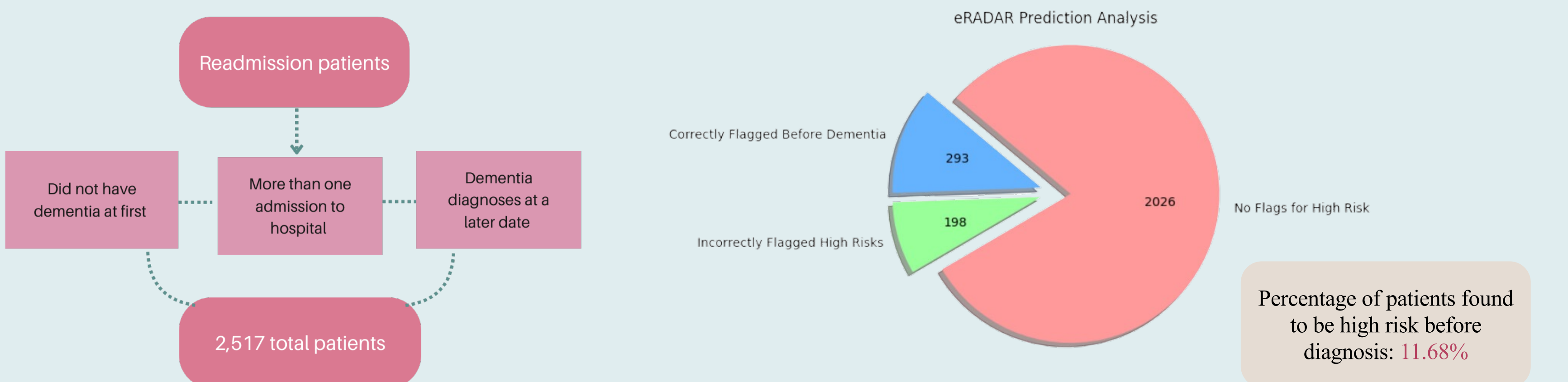
Group	Number of Patients used for Predictions
Total Patients	299,712
Control Patients	180,640
Dementia Patients	5,296

**Table 2.** Number of patients in each group for analyses.

## 4 e-RADAR Analyses



**Figure 3.** Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and various risk percentile thresholds for the year-group 2008 to 2010.



**Figure 4.** How well does eRADAR predict the risk of dementia in patients before they are actually diagnosed?

## 5 Discussion

- Metrics demonstrate the model has moderate predictive accuracy.
- Limitations: MIMIC mainly contains inpatient data, which influences the inclusion of all coefficients from the restricted model.
- Potential for identifying patients at risk before dementia diagnosis.
- Improved model can support early intervention, targeted monitoring, and timely treatment care.

## 6 Conclusion

- Computed eRADAR scores for the MIMIC-IV Dataset.
- Predictors like outpatient, physical therapy, and address visits were excluded because they referred to outpatient data.
- In the data frame with all patients, most had missing data, which impacted the model's evaluation.
- The 90th percentile cutoff was selected to assess the proportion of positive predictions and the number of patients who are not at risk for dementia.
- The MIMIC dataset contained patients with several readmissions, allowing for investigation of the model's accuracy in identifying high-risk patients.

## 7 Future Directions

- Examine methods to identify outpatient information in the MIMIC dataset.
- Further investigation: subgroup analyses, impact of resources, socioeconomic status, etc.
- Do scores increase closer to the time of dementia diagnosis?
- What is the average time it takes for eRADAR to detect a patient's risk of dementia before they are diagnosed?
- Applying score to clinical video visits.

## 8 Acknowledgments ; References

“The eRADAR™ algorithm<sup>1,2</sup> was developed by Dr. Deborah Barnes of the University of California, San Francisco (UCSF) and Dr. Sascha Dublin of Kaiser Permanente Washington (KPWA) with funding from Tideswell at UCSF and the National Institute on Aging (R24 AG045050, R56 AG056417). The views and opinions expressed do not necessarily reflect those of the investigators, institutions, or U.S. government.”

This project was supported by funding from URISE, MARC, IDEAL Research. Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania.

