Deep Learning Classification of Future PAP Adherence based on CMS and other Adherence Criteria

Sam Rusk, BS1, Yoav N. Nygate, MS1, Chris R. Fernandez, MS1, Jiaxiao M. Shi, PhD2, Jessica Arguelles, BS2, Matthew T. Klimper, BS2, Nathaniel F. Watson, MD, MS3, Robert Stretch, MD4, Michelle R. Zeidler, MD5, MS, Anupamjeet Sekhon, MD2, Kendra Becker, MD2, Joseph Kim, MD2, Dennis Hwang, MD2

1 EnsoData Research, Ensodata, Madison, WI, USA
2 Kaiser Permanente, Southern California, USA
3 Department of Neurology, University of Washington School of Medicine, Seattle, WA
4 Department of Internal Medicine and Pulmonary Disease, Beth Israel Deaconess Medical Center, Boston, MA
5 Department of Pulmonology, UCLA Health, Los Angeles, CA

ensodata research

UW Medicine
DEPARTMENT OF NEUROLOGY

KAISER PERMANENTE®

UCLA Health®

Beth Israel Deaconess Medical Center

Introduction

- Optimizing obstructive sleep apnea outcomes is dependent on sufficient positive airway pressure (PAP) adherence.
- Personalizing treatment options and applying effective behavioral interventions early in the course of routine patient issues with therapy may result in improved treatment success.
- We apply and study the ability of predictive AI models to enable earlier interventions, improve adherence, or pivot to other therapy strategies.
- In this work, we demonstrate Deep Learning models for forecasting future PAP use to identify patients at risk for non-adherence.

Data

Overview

- A cohort of subjects with daily PAP usage recorded during 2015-2021 was assembled (N=34,015)

Input:

- Number of minutes the PAP machine was used during each day for the first 30-days after initiation of the treatment

Output

- **Simple Compliance (compliance_last):**
  - Positive if ≥4-hours of usage per night for ≥70% of nights during specific timeframe (3-month, 6-month or 1 year).
  - Negative otherwise.
- CMS Compliance (compliance_cms_last):
  - Positive if ≥4-hours of usage per night for ≥70% of nights at one of the rolling 30-day windows during specific timeframe (3-month, 6-month or 1 year).
  - Negative otherwise.

Methodology

- Machine learning approaches:
  - Deep Neural Network: Convolutional Neural Network and Long Short-Term Memory (CNN-LSTM) architecture. Evaluated with 10-fold cross-validation.
  - Naïve Method: Estimate adherence at each specific time point with the adherence during the first 30-days.
- Performance Metrics:
  - Sensitivity: sensitivity for non-adherers
  - Specificity: specificity for adherent patients
  - ROC-AUC: aggregate performance metric

Classification Performance: Future Adherence

Adherance Target	First 30-Day Naïve	First 30-Day ML
	ROC-AUC	
3-Month	0.96	0.97
6-Month	0.88	0.89
12-Month	0.83	0.84
	Sensitivity	
3-Month	0.89	0.91
6-Month	0.81	0.82
12-Month	0.76	0.76
	Specificity	
3-Month	0.88	0.91
6-Month	0.79	0.82
12-Month	0.76	0.77

Figure 1. A table of 3-month, 6-month, and 12-month time point ROC-AUC (top), Sensitivity (middle), and Specificity (bottom) performance estimates. In each experimental endpoint analyzed, the ML methods demonstrated greater performance than the simple (Naive) methodology. Note across all 3 statistical performance measures, 3-month performance tended to be higher than 6-month, which respectively was higher 12-month performance measures.

Receiver Operating Curve (ROC) Analysis for 3, 6, and 12 Month Adherence Ranges

3-Month Adherence Classification (CMS Criteria)

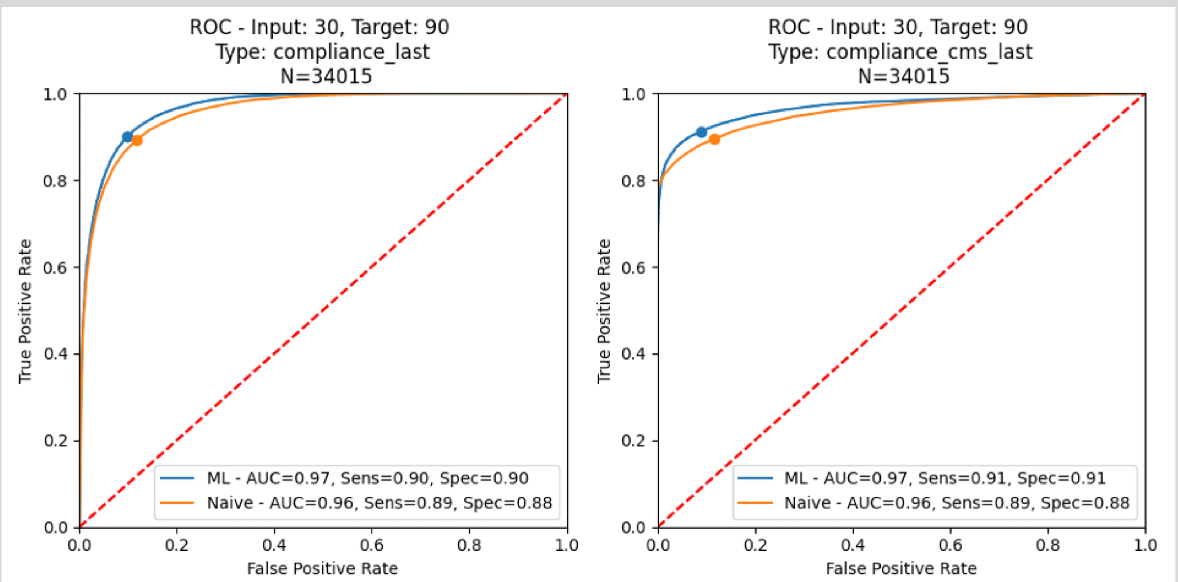

Figure 2. Receiver Operating Characteristic curves for ML and Naive models on the 3-month prediction task.

6-Month Adherence Classification (CMS Criteria)

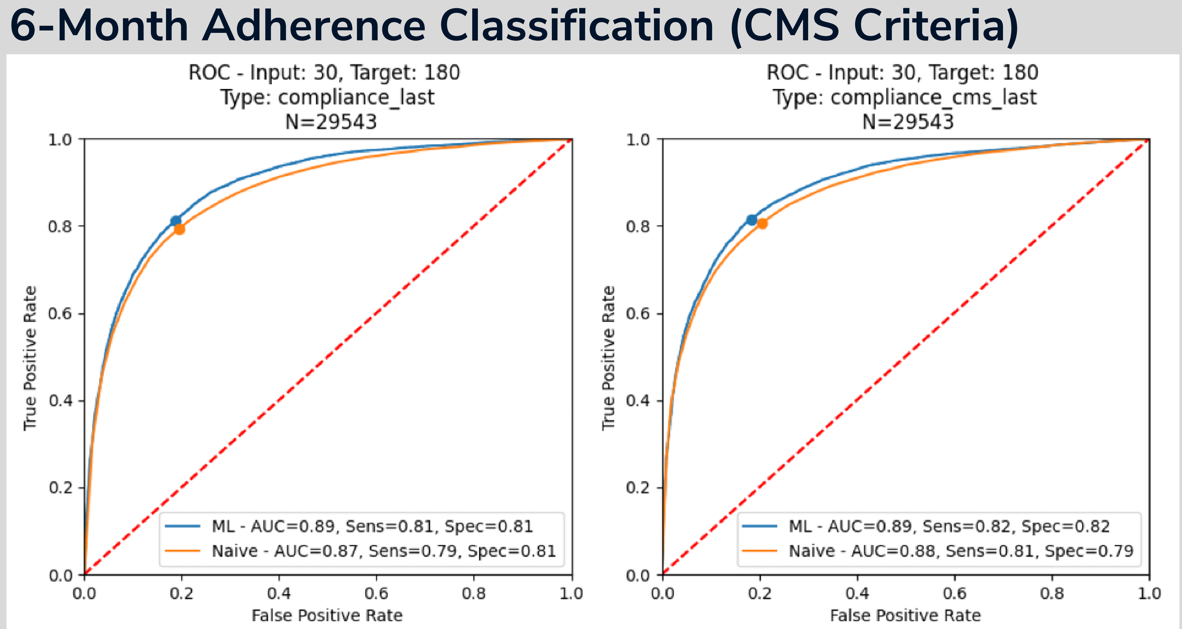

Figure 3. Receiver Operating Characteristic curves for ML and Naive models on the 6-month prediction task.

12-Month Adherence Classification (CMS Criteria)

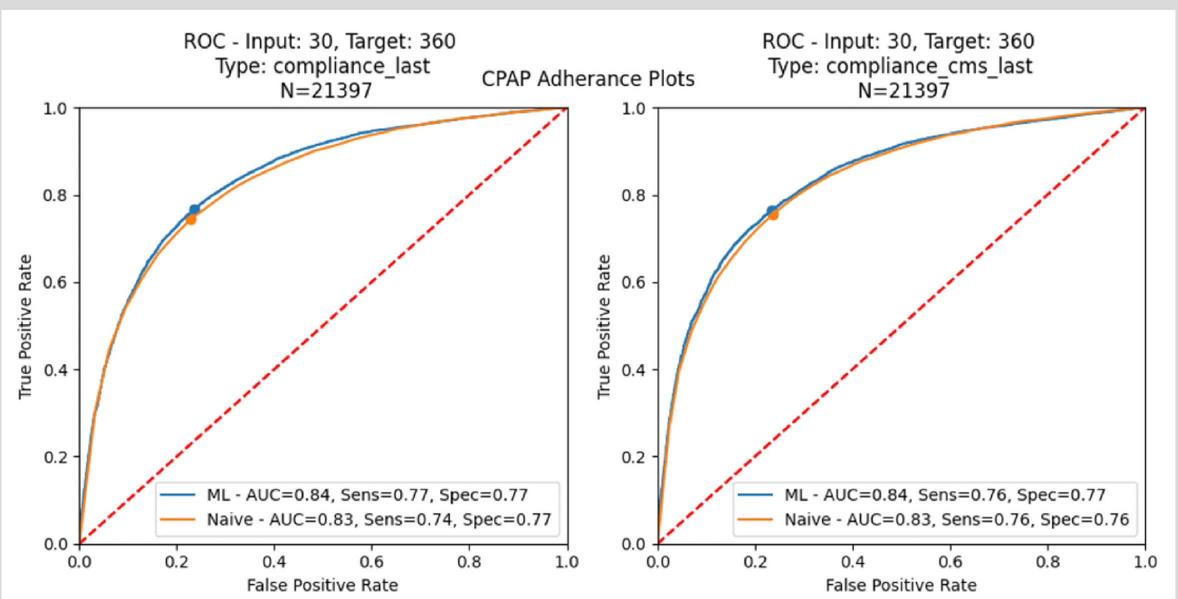

Figure 4. Receiver Operating Characteristic curves for ML and Naive models on the 3-month prediction task.

Results

- The DNN models predicted CMS adherence with a sensitivity of 91%, 82%, 76% and a specificity of 91%, 82%, 77%, for 3-month, 6-month, and 1-year endpoints, with ROC-AUC values of 0.97, 0.89, and 0.84 respectively.
- Both Standard and CMS definitions of compliance show similar classification performance.

Conclusions

- DNN models demonstrated strong predictive performance for PAP adherence, as defined by the CMS adherence criteria, measured by sensitivity, specificity, and overall ROC-AUC results at 90-day, 6-month, and 1-year timepoints.
- AI approaches show promise as early predictors of the likelihood to meet key therapy utilization thresholds within the first four weeks of therapy, enabling early PAP intervention or transition to alternative therapies.
- Performance for Naive and DL classification are similar suggesting that predicative inputs beyond simple CPAP usage may be required to improve predictive performance.

Future Work

- Treatment usage is a largely behaviorally influenced metric and other non CPAP based inputs may be critical to further refine adherence classification accuracy.
- Future work may include demographic variables, socio-economic variables, diagnostic sleep study variables, medication usage variables, and survey variables.
- A more comprehensive data representation of patients may help improve adherence classification performance and help understanding of which factors most influence CPAP adherence.