Deep Learning to Predict PAP Adherence in Obstructive Sleep Apnea

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

KAISER PERMANENTE®

UW Medicine
DEPARTMENT OF NEUROLOGY

ensodata research

UCLA Health®

Beth Israel Deaconess Medical Center

Introduction

Machine Learning (ML) algorithms to predict Positive Airway Pressure (PAP) adherence may support personalized clinical management. Models were developed to predict adherence at various time-points after PAP initiation and in moving time windows.

Methodology

- Deep neural network (DNN) models were trained utilizing daily PAP data (Kaiser Permanente, Southern California). The DNN was evaluated with 10-fold cross-validation on N=21,397 patients.
- Algorithms developed included
  - (a) Models 1 and 2 utilizing early usage to predict adherence at 90-days and 1-year.
  - (b) Model 3 which utilized 14 and 30-day moving windows to predict subsequent usage.
  - Regression analyses compared ML and Naïve (i.e., future use equals previous use) predictions versus the Actual adherence values observed.

Results

- Model 1 predicted "% days without usage" for first 90-days based on first 7, 14, 21, 30-days of input and at 1-year (90-day window) based on the first 30, 60, 90, 180-days of input.
  - ML was superior to Naïve in predicting adherence [R2 for ML versus Naïve compared to Actuals for different input days (all p< 0.05):
  - At 90-days: 0.495-vs-0.193; 0.660-vs-0.465; 0.748-vs-0.607; 0.828-vs-0.735.
  - At 1-year: 0.362-vs-0.104; 0.463-vs-0.247; 0.513-vs-0.339; 0.680-vs-0.547.
- Model 2 predicted "hours/night" of use—ML did not outperform the Naïve prediction with similar R 2
  - When ML predicted < 3 hours/night, nearly all patients had "no significant usage" at 1-year.
  - The naïve model had no differentiating threshold to predict this outcome.

Results Continued

Model 1: 90-day CPAP Days Used >0 Hours Forecast

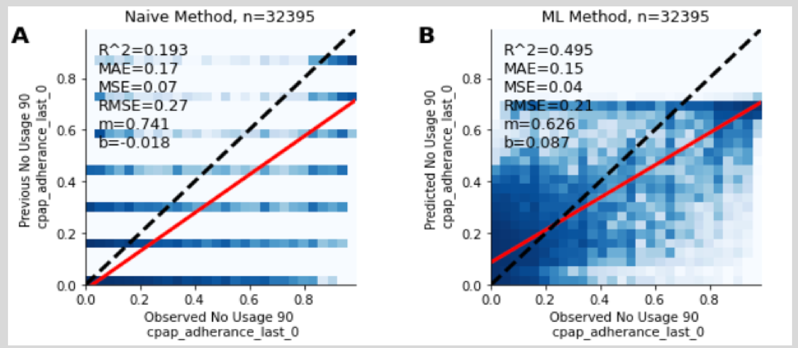

Figure 1. Naive vs. ML methods based on CPAP Usage at the 7 day mark.

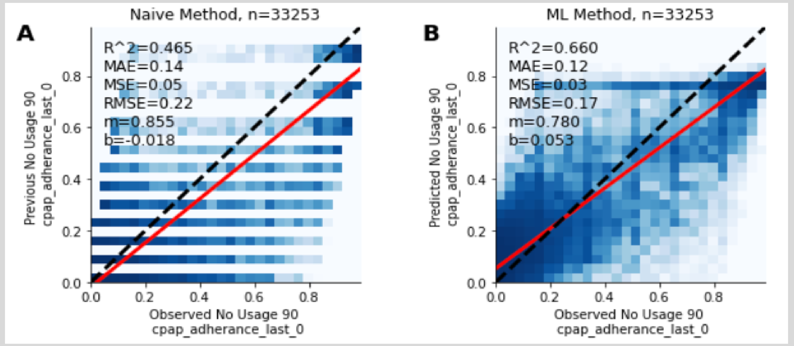

Figure 2. Naive vs. ML methods based on CPAP Usage at the 14 day mark.

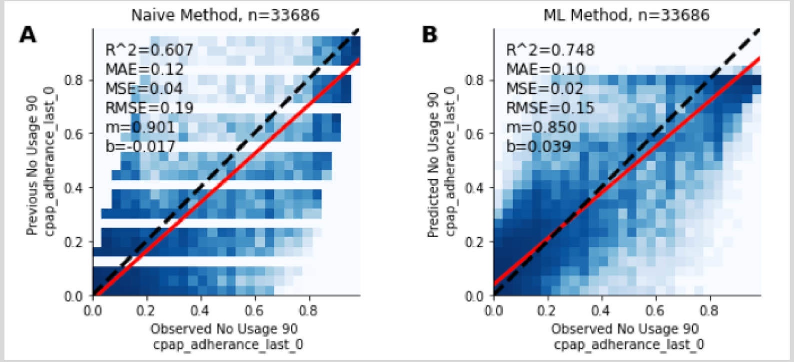

Figure 3. Naive vs. ML methods based on CPAP Usage at the 21 day mark.

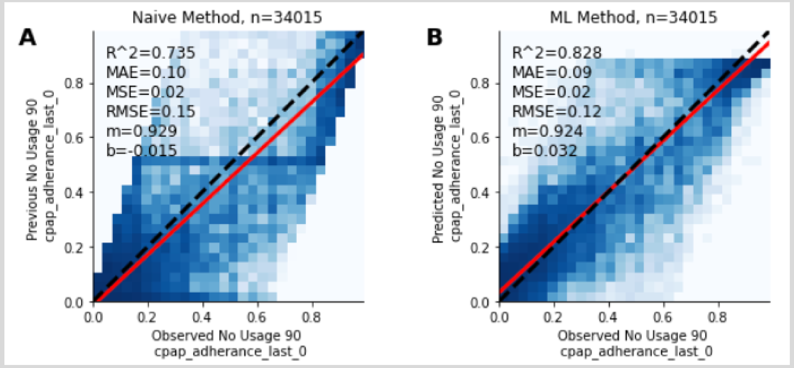

Figure 4. Naive vs. ML methods based on CPAP Usage at the 30 day mark.

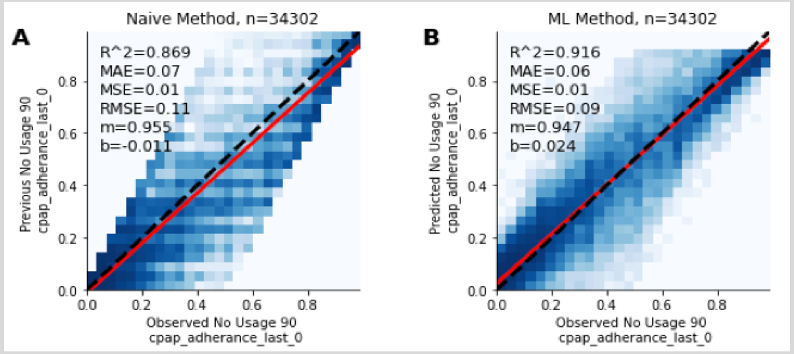

Figure 5. Naive vs. ML methods based on CPAP Usage at the 60 day mark.

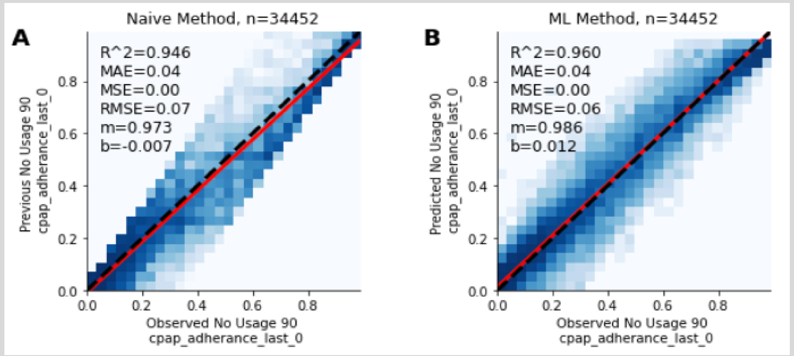

Figure 6. Naive vs. ML methods based on CPAP Usage at the 90 day mark.

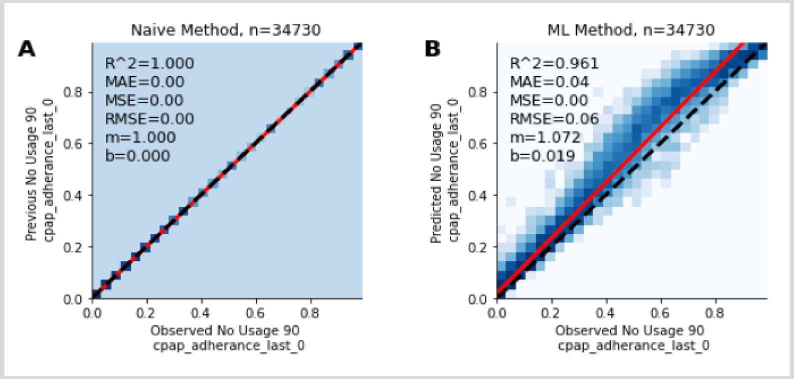

Figure 7. Naive vs. ML methods based on CPAP Usage at the 1 year mark. At one year, the "naive" model is comparing data against itself, creating a perfect line, (left).

Results Continued

Model 2: 90-day Hours/Night Usage Forecast

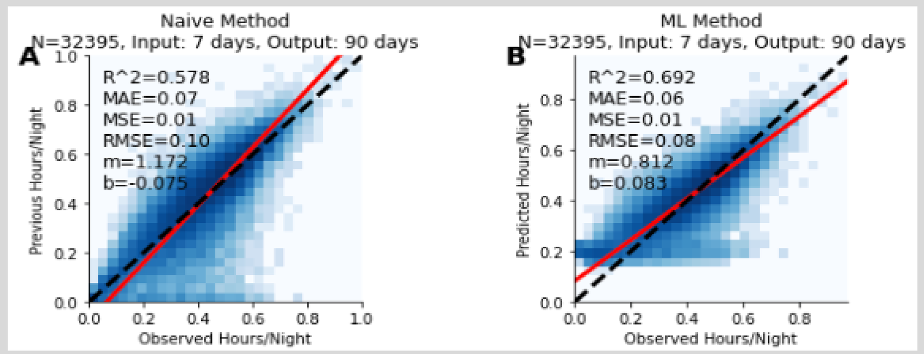

Figure 8. Usage at the 7 day mark

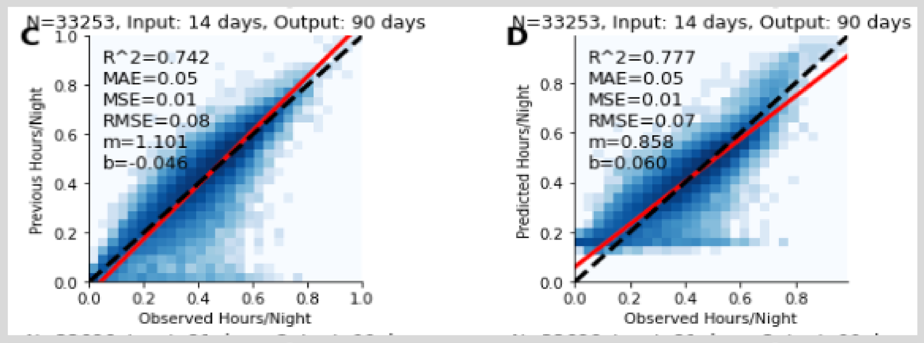

Figure 9. Usage at the 14 day mark

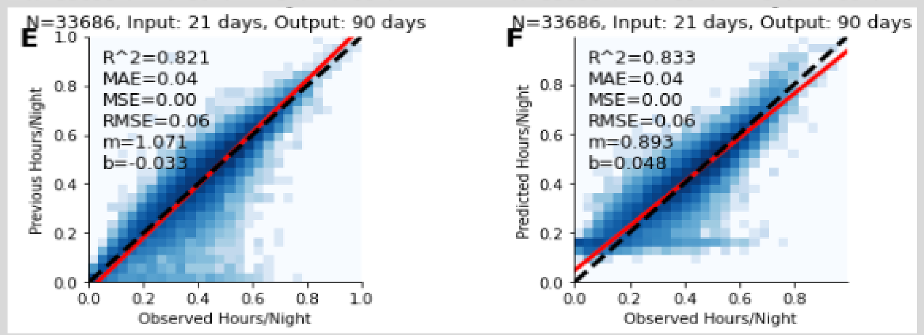

Figure 10. Usage at the 21 day mark

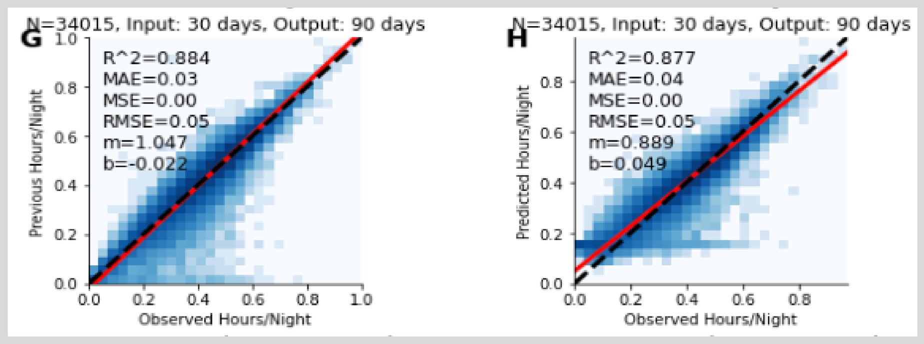

Figure 11. Usage at the 30 day mark

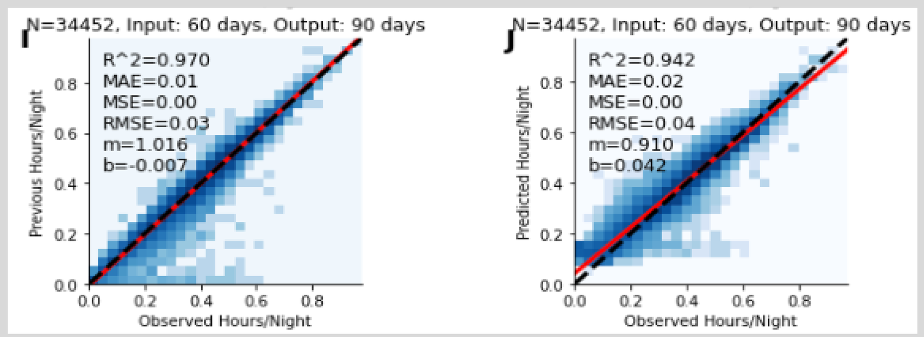

Figure 12. Usage at the 60 day mark

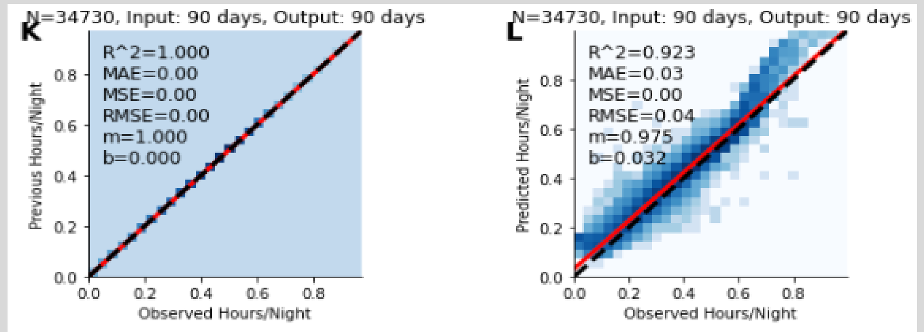

Figure 13. Usage at the 90 day mark.

Model 3 utilized different windows of PAP usage to predict subsequent usage.

- ML predictive accuracy was similar using 14 or 30-days of input.
- R2 for ML vs. Actuals in predicting 7, 14, and 30-day "% days used ≥4 hours" were (all p< 0.05).
  - 0.687, 0.701, 0.699 using 14-days of input.
  - 0.582, 0.702, 0.77 using 30-days of input.

Conclusions

- ML algorithms based on PAP usage can predict future adherence, potentially supporting personalized treatment decisions and preemptive interventions when upcoming non-adherence is predicted.
- The results show that different kinds of treatment usage behavior can be modeled.
  - The # Days Used >0 Hours represents behavior around nightly usage and potential factors that would cause an individual to forgo usage for all together.
  - The Hours/Night Usage represents behavior around usage if used during a night and potential factors that would cause an individual to use for either a portion or the entire night.
- The behavioral phenotypes we can forecast allow clinical staff the resources to create intervention strategies and understand at a detailed level where patients may struggle with adherence.

Future Work

- We can build upon this research by looking at different behavior metrics and the potential to forecast those behaviors.
- For example, more detailed metrics characterizing intermittent usage sessions intra-night would help differentiate behaviors that might cause a patient to start and stop treatment in the middle of night.
- Further, a metric characterizing usage patterns dependent on day of the week, seasonal patterns, and annual holidays would help characterize unique usage patterns that would help a clinical coach in their intervention strategies.
- Lastly, a metric characterizing usage patterns with respect to CPAP supplies and hardware would help characterize how the treatment device may be affecting adherence.
- These potential metrics in addition to others will help capture a more comprehensive picture of treatment adherence to aid in behavioral coaching and intervention.