Facilitating CPAP Adherence with Personalized Recommendations Using Artificial Neural Networks

Matheus Araujo*, Tara Pereira*, Jaideep Srivastava*, Conrad Iber†
*Computer Science & Engineering
University of Minnesota, Minneapolis, USA
arauj021@umn.edu, perei087@umn.edu, srivasta@umn.edu
†M Health Fairview, Minneapolis, Minnesota
ciber1@fairview.org

***Abstract*—Sleep apnea is a common sleep disorder that, if left untreated, can have critical complications to the individual. The most common and effective treatment for sleep apnea is the Continuous Positive Airway Pressure (CPAP) therapy. But it has a long-term adherence rate as low as 60% due to discomfort and other factors. Although previous research has attempted to increase CPAP usage, there has been little to no change in its average adherence for the past two decades. This paper attempts to change this scenario using a large longitudinal dataset combined with a Recurrent Neural Network model to generate therapy use recommendations after one month of therapy. We performed a retrospective cohort analysis on 3380 patients during their first six months of therapy and compared our personalized recommendation system with the current generic recommendations made by sleep physicians. We show that recommendations generated by our artificial neural network model are easier to achieve since they are significantly closer to patients' therapy progress while being equally successful in maintaining therapy adherence.**

***Index Terms*—Recommendation Systems, Therapy Adherence, Recurrent Neural Network, Sleep Apnea, Sleep Medicine**

I. Introduction

About 22 million people in the United States suffer from sleep apnea, and that number is steadily increasing each year [2]. Sleep apnea symptoms can include loud snoring, breathing cessation episodes, abrupt awakenings, chronic fatigue, headaches, attention problems, and irritability. In severe cases, if left untreated, sleep apnea can have critical complications, such as high blood pressure, heart problems, type 2 diabetes, liver problems, and difficulties with medications. Eventually, these conditions could lead to an early death [14].

Continuous positive airway pressure (CPAP) is a mask connected to an electronic that delivers positive airflow when an apnea event is detected, helping patients sleep. It is the most common and effective treatment for sleep apnea. However, usage is low in patients due to several reasons, including usage habits, leakage caused by an ill-fitting mask, or unrecognized factors governing adaptation [1]. In general, long-term therapy adherence to CPAP can be defined as using the CPAP for $\geq$ 4 hours per night for at least 70% of the days in the 6th month of therapy [7]. Unfortunately, over the last twenty years, non-adherence to CPAP has remained as high as 30-40% and has not shown any meaningful improvement despite numerous suggested interventions [13]. Motivational enhancement, desensitization protocols, and in-clinic daytime studies(PAP NAP) are recommended to improve therapy adherence, but more randomized controlled trials are required for evaluation [10].

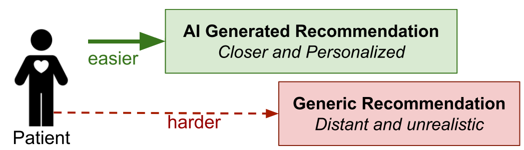

Fig. 1. AI Generated Recommendation should be personalized and closer to current patient state when compared to generic but harder to follow recommendations.

Recently, sleep physicians have gained the ability to leverage the potential of big data regarding behavior in CPAP therapy, enabling the possibility of advanced analytics to personalize CPAP usage recommendations and its management. For example, M Health Fairview Sleep Clinics use their large dataset for predicting future therapy non-adherence, applying an early intervention on those who are likely to not adhere to the therapy [3]. In addition to these predictions, we can combine big data with deep learning techniques to propose new clinical processes, including adaptively refining therapy recommendations.

In this work, we investigate the use of Deep Learning using a Recurrent Neural Network (RNN) model that is able to generate customized therapy use recommendations with improved adherence. The model is responsible for receiving the previous therapy signals collected from CPAP devices and generating future recommendations for each individual in order to enhance overall therapy adherence in the long-term.

We hypothesize that one of the factors that involve non-adherence is linked to generalized CPAP usage recommendations, which in general is standard for all patients, making it difficult for those who require more flexibility in their therapy goals. Our approach can be compared to the case where overweight individuals should not set their weight loss goals too high according to ideal cultural standards; otherwise, going to the gym can mean frustration due to the goal being too distant. There is no reason to force a patient with a generic CPAP therapy use recommendation if we can generate a more personalized recommendation as illustrated in Figure

2372-9198/21/$31.00 ©2021 IEEE
DOI 10.1109/CBMS52027.2021.00093

1. The therapy trajectory of each individual is unique, and each patient has their challenges. Thus, personalized recommendations can be much easier to follow, and in this work, we show they can be at least as effective as generic recommendations.

II. Related Work

People can fail to translate their intentions into healthy behavior. However, having accessible goals, where an individual's purpose is aligned with their expectation, has been associated with more significant prediction of behavior by intention [4]. This phenomenon is also observed in the example of weight loss, where unrealistic goals can prevent the success of weight loss attempts [16]. Overall, individuals are more likely to have a firmer belief in their capacity to execute their actions if their goals are personalized. Thus, a more personalized goal unlocks the expected behavior from the individual [5].

This work shows how to use a novel AI-based system for a personalized recommendation that aligns with the unprecedented advances in clinical data analysis. Recent computational approaches have radically changed how medicine is being made with more personalized processes [8] including in the context of the CPAP therapy [9].

III. Methodology

In this section, we show the steps to train a Recurrent Neural Network (RNN) that receives the collected therapy data from the first month of use of each patient as an input and generates a recommendation to be followed by each patient during the next 5 months of therapy.

A. Data

The data used in this study was shared by Fairview Sleep Clinics and collected from a group of 3380 patients diagnosed with sleep apnea and uses CPAP. The recordings' dates range from March 2009 to October 2017. Using a retrospective cohort study approach, we split the patients into a train (2000 patients 60%) and test set (1380 patients - 40%). All the patients have at least 6 months of data, and only the first 6 months were used in the scope of this project. Each patient's data has daily granularity.

The CPAP appliances provide the following daily signals about each day of therapy:

- the amount of time the machine was used in hours
- the pressure of the mask in cmH2O
- the air leakage from the mask in L/min
- the average number of apnea-hypopnea index (AHI) events per hour during the night of sleep.

Our proposed model generates a recommended value for each signal. Out of the 4 signals, only time used per night, air pressure, and air leakage are under the patient's control. Thus, AHI values are not included in the model's recommendation.

B. The Traditional Generic Recommendation

To define the generic recommendation, we used previous literature suggestions for what the average hours of use, pressure settings, and air leak recommended should be for an average patient during their CPAP therapy. We compare the proposed personalized recommendations against the generic recommendation in terms of ease of use and therapy adherence efficacy.

For average hours of use, we found that a minimum of 4.0 hours per night is recommended for an improvement on the Epworth Sleepiness Scale (sleep quality). However, 7.5 hours is required to improve disease-specific quality of life, and functional status [17]. The ideal pressure setting is 10.0 cmH2O, although some cases require a custom adjustment based on a patient's sleep apnea severity [12]. Unintentional air leakage can occur through the CPAP mask, but the ideal is no leakage, and it should be kept at an average threshold below 25.0 L/min [11].

Thus, for the scope of this paper, we specified that a generic recommendation is average therapy use of 7.5 hours per night, an average pressure setting of 10.0cmH20, and an average air leakage of 0.0 L/min.

C. Personalized Recommendation Using RNN

The personalized AI recommendations are generated from a deep neural network model that receives previous therapy information from each patient as input and generates an array of values corresponding to the expected mean of each therapy recorded signal as output. The output array has 3 values corresponding to the mean values for time on the face, air pressure, and air leakage. These 3 values define the recommendation that a sleep physician should make for the patient to follow after the first month of therapy.

One important caveat of our training process is that we only train the model on examples of patients that we know are long-term adherent to the therapy. Thus, the model learns how to map the first-month time-series data into the next 5 months' mean values from previously selected adherent patients. Consequently, out of the 2000 patients in the training set, only 1300 adherent patients were used. We also standardized the input to have a zero mean, as it is a good practice while training deep neural networks.

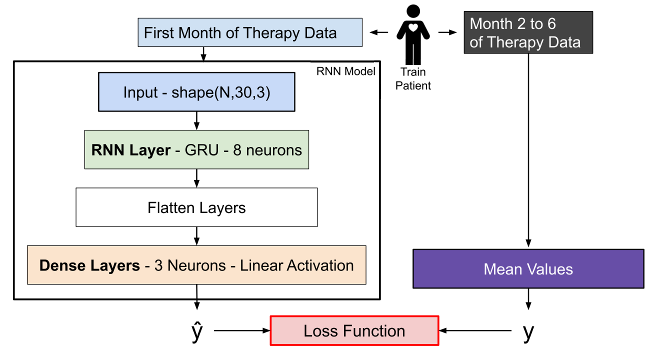

Fig. 2. Overview of how we split the data from patients to train our RNN Model. We also highlight the internal model architecture.

In Figure 2, we show an overview of the training process. The first month of therapy data is separated from the other 5 subsequent months. For the model, the goal while training is to predict the future mean of the 3 controllable CPAP signals

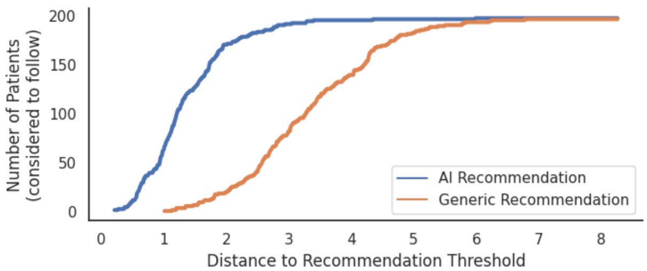

Fig. 3. Recommendation Strategies and Distance to Recommendation Threshold: The x-axis denotes the Manhattan distance threshold between a recommendation strategy and patients' therapy progress, and the y-axis is the number of patients following either AI (darker color) or the generic (lighter color) recommendation. The neural network approach recommends CPAP settings much closer to patients' therapy progress when compared to generic recommendations..

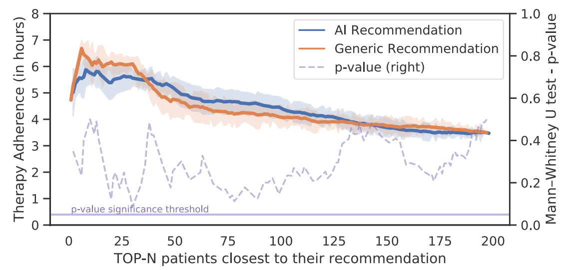

Fig. 4. On the y-axis, we present the average therapy usage in the 6th month for the TOP-N closest patient. On the x-axis, we vary N from 1 to 198. The dashed line shows the Mann–Whitney U test p-value between therapy usage groups. There is no scenario in which the p-value is less than 0.05, and the confidence interval overlaps across the plot.

given the first-month therapy. Thus, its output layer has 3 neurons with a linear activation. The model has an RNN layer with 8 Gated Recurrent Units (GRUs) and was trained over 500 epochs using Adam optimization as the stochastic gradient descent algorithm with the mean squared error as the loss function. We used the Tensorflow 2.2 python library as our deep learning framework.

After training, when a new patient is introduced to the RNN, the input is a tensor of shape 30x3, corresponding to their 3 signals from CPAP recordings in the first 30 days. It is expected that the model generates the closest, realistic recommendation for a new patient that incentivizes adherence in the next 5 months. Because we only used adherent patients as samples during the training supports this expectation. Moreover, we investigated this assumption in all generated recommendations for patients in the test set. We observed absolutely no AI-generated recommendation suggesting to patients use their devices less than they used on average during the first 30 days. All recommendations tried to push the patients for higher adherence.

D. Evaluating Therapy Recommendation

We evaluate the quality of the personalized AI-generated recommendations regarding how much easier they are to be followed and how effective they are compared to generic recommendations. All the results were made using patients from the test set, which the model never observed while training. Following a clinical procedure where intervention is only made for low-adherent patients, we only evaluated recommendations for those patients who have an average therapy use lower than 4 hours during their first month. Thus, out of the 1380 patients on the test set, we evaluate our recommendation on the trajectory of 198 non-adherent patients.

We assume that the closer a recommendation is to the patient therapy progress after the first month, the easier the recommendation is to be followed. Therefore, we compute how the patient behaved after the recommendation was generated by computing the actual mean of the 3 CPAP's controllable signals from the patient history data, forming an array of 3 values of the same shape as the AI-generated and the generic recommendations. We then computed the Manhattan distance between these two arrays to evaluate how close patients were to each recommendation. The results of this analysis are shown in Figure 3.

To evaluate the effectiveness of the personalized recommendations, we compare the average therapy use during the 6th month of therapy of the TOP-N patients closest to the generic recommendation and the TOP-N patients closest to the AI-generated recommendation. The comparison was made from N=1 to N=198. If the effectiveness of the AI-generated recommendation is as effective as a generic recommendation, these two curves should not be statistically different. We evaluate the statistical difference by comparing the confidence interval of both curves. Also, we perform for every N in the TOP-N closest patients, an Mann–Whitney U test comparison between therapy usage of patients, if p-value > 0.05 for every N, both recommendations types are not different. The result is shown in Figure 4.

IV. Results & Discussion

We plot in Figure 3 the number of patients that followed the recommendation for each threshold distance to the recommendation. Our purpose is to show how the recommendation generated by the RNN model is more accessible (e.g., closer) for a greater number of patients. Both curves start lower since only a small group of patients are very close to the recommendation. As we increase the threshold distance, more patients would be considered as following the recommendation. The accessibility of our recommendations can be easily observed by the left shift of the AI-generated recommendation curve (in blue). The first patient who followed our recommendation has a distance of 0.21 in contrast to 1.13 for the generic recommendation. When considering the TOP-50 patients, the AI recommendation has an average distance of 0.93, and the generic recommendation has an average distance of 2.6. We also found that for the same number of patients to be considered in the therapy, the generic recommendation is an average 2.6x (+/- 0.07) more distant than the personalized recommendation. This analysis demonstrates how far and, therefore, harder it is to achieve the generic recommendation when compared to the AI-generated recommendation.

In Figure 4, we evaluate if the AI-generated recommendation and the generic recommendation have the same effectiveness. The results consider only the 6th month of therapy use because it is how the long-term adherence is defined. We hypothesized that for a small N, the average therapy use of the TOP-N closest patients would be high for both types of recommendations. For a large N, the average therapy use of both curves would converge to the average therapy use of the population (since all the patients would be in the TOP-N). We observe that for N=6 to N=30, the average hours of use from patients who followed the generic recommendation is higher than our recommendations. This may happen for those few extremely motivated patients who could change behavior according to the generic recommendation from doctors. For $N>36$ to $N<143$, we observe that the AI recommendation had a high average use. By observing the overlapping confidence interval between both curves, we can state that they are not different. To assure this conclusion, we computed a Mann–Whitney U test between therapy uses of both AI and generic recommendations for each TOP-N bin of patients. We plotted with a dashed purple line the p-values from each test. At all points in the curve, the p-values were more than 0.05.

Both results showed that while our personalized AI-generated recommendations are closer to patients' therapy reality, they result in similar average hours of CPAP usage in their 6th month of treatment compared to a generic recommendation. The key point is that closer goals tend to be more manageable and achievable. Hence, the AI recommendation can embrace a larger number of patients while also producing the same effect as a generic recommendation.

In future work, we should investigate a clustering analysis approach, where similarity-based recommendations among cohorts can be generated for a new patient.

A. Clinical Application

Our model generates a recommendation on day 30 for patients to follow throughout their treatment. In a real scenario, if a patient reaches their suggested recommendation, they should be encouraged to keep increasing usage until the ideal usage. This could be done by training models for subsequent days of follow-up. Although we could have chosen any day to generate the therapy recommendation, we select the first month since most CPAP patients will have a follow-up appointment with their doctor a month after the therapy start [6].

One may question the ability of the patients to follow our personalized recommendation. As shown by [17], mask leakage can be controlled by sleeping position, regular cleanings, or fit. Also, pressure settings can be changed in the CPAP machine, and the number of hours can be controlled by Cognitive Behavioral Therapy (CBT) and sleep hygiene. Even if the patient cannot tune to the exact proposed recommendation, our personalized goal will turn the therapy into a much less frustrating experience.

Our approach introduces the possibility of creating a patient-driven application service that patients can check independently, encouraging active participation by using self-tracking and up-to-date, personalized guidance. This personalized type of care benefits both patients and hospital staff [15].

V. Conclusion

To address low adherence challenges in long-term CPAP therapy, we attempt to generate a more achievable recommendation of therapy use. We trained a Recurrent Neural Network (RNN) to generate recommendations using known adherent patients as training examples. In a retrospective cohort analysis, we observed that patients who followed the AI-generated recommendation had similar hours of usage as patients who followed a generic recommendation in their 6th month of treatment. However, we also showed that our personalized recommendation is, on average, 2.6x closer to the patients' therapy progress, making the goal more likely to be reached. As a consequence, patients will perceive success more often and keep motivated.

References

[1] Cpap therapy, 2017. https://www.sleepapnea.org/treat/cpap-therapy/.

[2] Sleep apnea information for clinicians, 2017. https://www.sleepapnea.org/learn/sleep-apnea-information-clinicians/.

[3] Araujo, M., Kazaglis, L., Bhojwani, R., Iber, C., Khadanga, S., and Srivastava, J. 1078 machine learning to predict pap adherence and compliance in tele-health management. *Sleep 41* (2018).

[4] Avishai, A., Conner, M., and Sheeran, P. Setting realistic health goals: Antecedents and consequences. *Annals of Behavioral Medicine 53*, 12 (2019), 1020–1031.

[5] Bandura, A., and Locke, E. A. Negative self-efficacy and goal effects revisited. *Journal of applied psychology 88*, 1 (2003), 87.

[6] Cervenka, T., and Iber, C. EHR integration of pap devices in sleep medicine implementation setting in the clinical. *Sleep Med Clin. 15(3)* (2020), 377–382.

[7] Choi, J.-A., Yoon, I.-Y., Han, E.-G., and Lee, S. Subjective and objective cpap compliance in patients with obstructive sleep apnea syndrome. *Sleep Medicine Research 2*, 2 (2011), 63–68.

[8] Cirillo, D., and Valencia, A. Big data analytics for personalized medicine. *Current opinion in biotechnology 58* (2019), 161–167.

[9] Joymangul, J. S., Sekhari, A., Moalla, N., and Grasset, O. Obstructive sleep apnea compliance: modeling home care patient profiles. In *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)* (2020), IEEE, pp. 397–402.

[10] Lance, C. Increasing adherence to pap therapy for patients with sleep apnea, Feb 2020.

[11] Lebret, M., Arnol, N., Martinot, J.-B., Lambert, L., Tamisier, R., Pepin, J.-L., and Borel, J. C. Determinants of unintentional leak during cpap treatment in obstructive sleep apnea syndrome. *Chest 153*.

[12] Repasky, D., 2020. https://www.cpap.com/blog/cpap-pressure-setting-cpap-apap-bipap/.

[13] Rotenberg, B., Murariu, D., and Pang, K. Trends in cpap adherence over twenty years of data collection: a flattened curve. *Journal of Otolaryngology - Head & Neck Surgery 45* (12 2016).

[14] Staff, M. C. Sleep apnea, 2018. https://www.mayoclinic.org/diseases-conditions/sleep-apnea/symptoms-causes/syc-20377631.

[15] Swan, M. Emerging patient-driven health care models: An examination of health social networks, consumer personalized medicine and quantified self-tracking. *International Journal of Environmental Research and Public Health 6* (2009), 492 – 525.

[16] Wamsteker, E. W., Geenen, R., Zelissen, P. M., van Furth, E. F., and Iestra, J. Unrealistic weight-loss goals among obese patients are associated with age and causal attributions. *Journal of the American Dietetic Association 109*, 11 (2009), 1903–1908.

[17] Weaver, T., Maislin, G., Dinges, D., Bloxham, T., Greenberg, H., Kader, G., Mahowald, M., Younger, J., and Pack, A. Relationship between hours of cpap use and achieving normal levels of sleepiness and treatment. *Sleep 30* (06 2007), 711–9.