ELSEVIER

Sleep Medicine xxx (2006) xxx–xxx

SLEEP MEDICINE

www.elsevier.com/locate/sleep

Original article

Predicting effective continuous positive airway pressure in sleep apnea using an artificial neural network

Ali A. El Solh *, Zaher Aldik, Moutaz Alnabhan, Brydon Grant

Western New York Respiratory Research Center, Department of Medicine, Division of Pulmonary, Critical Care, and Sleep Medicine, University at Buffalo, School of Medicine and Biomedical Sciences, 462 Grider Street, Buffalo, NY, USA
Veterans Affairs Medical Center, Buffalo, NY, USA

Received 7 July 2006; received in revised form 8 September 2006; accepted 11 September 2006

Abstract

Background: Mathematical formulas have been less than adequate in assessing the optimal continuous positive airway pressure (CPAP) level in patients with obstructive sleep apnea (OSA). The objectives of the study were (1) to develop an artificial neural network (ANN) using demographic and anthropometric information to predict optimal CPAP level based on an overnight titration study and (2) to compare the predicted pressures derived from the ANN to the pressures computed from a previously described regression equation.
Methods: A general regression neural network was used to develop the predictive model. The derivation cohort included 311 consecutive patients who underwent CPAP titration at a University-affiliated Sleep Center. The model was validated subsequently on 98 participants from a private sleep laboratory.
Results: The correlation coefficients between the optimal pressure determined by the titration study and the predicted pressure by the ANN were 0.86 (95% confidence interval [CI] 0.83–0.88; $p < 0.001$) for the derivation cohort and 0.85 (95% CI 0.78–0.9; $p < 0.001$) for the validation cohort, respectively. Whereas there was no significant difference between the optimal pressure obtained during overnight polysomnography and the predicted pressure estimated by the ANN ($p = 0.4$), the estimated pressure derived from the regression equation underestimated the optimal pressure in both the derivation and the validation group, respectively.
Conclusion: The optimal CPAP level predicted by the ANN provides a more accurate assessment of the pressure derived from the historic regression equation.

Keywords: Titration; Obstructive sleep apnea; Artificial neural network; Continuous positive airway pressure

1. Introduction

Obstructive sleep apnea (OSA) is a relatively common problem with potentially serious health consequences [1]. It has been linked to increased risk of mortality and morbidity due to cardiovascular and neurophysiologic disorders [2]. Nasal continuous positive airway pressure (CPAP) is considered a well established and effective therapy for this disorder [3]. Compliance with treatment leads invariably to enhanced vigilance, improved quality of life, and reduced traffic accidents [4].

In order to derive the most effective pressure, CPAP titration is performed in the sleep laboratory during which the pressure is gradually increased until apneas and hypopneas are abolished in all sleep stages and in all body positions. The technique is, however, time-consuming and labor-intensive. Furthermore, the duration of the study may not be sufficient to attain this goal because of patient's poor ability to sleep in this environment or due to difficulty attaining an appropriate

* Corresponding author. Tel.: +1 716 898 3375; fax +1 716 898 6139.
E-mail address: solh@buffalo.edu (A.A. El Solh).

1389-9457/$ - see front matter © 2006 Published by Elsevier B.V.
doi:10.1016/j.sleep.2006.09.005

pressure. A predictive algorithm based on demographic, anthropometric, and polysomnographic data was developed to facilitate the selection of a starting pressure during the overnight titration study [5], but the performance of this model was inconsistent when validated by other centers [6,7]. One of the potential reasons for the lack of reproducibility is the complex relationship of behavioral processes with nonlinear attributes. In areas of complex interactions, the artificial neural network (ANN) has been found to be a more appropriate alternative to linear, parametric statistical tools due to its inherent property of seeking information embedded in relationships among variables thought to be independent.

Neural networks are computation systems that process information in parallel, using large numbers of simple units, and that excel in tasks involving pattern recognition. These intrinsic properties of the neural networks have been translated into a higher performance accuracy in outcome prediction compared to expert opinion or conventional statistical methods [8,9]. Hence, we hypothesized that the ability to estimate the optimal pressure (Popt) can be improved by using computer analyses involving neural networks. To test this hypothesis, we first applied an ANN to the analysis of data from patients with documented OSA and validated it prospectively on a separate cohort. Second, we compared the predictive accuracy of ANN to the previously published predictive model of CPAP titration.

2. Materials and methods

2.1. Study population

The study protocol was approved by the Institutional Review Board of the University at Buffalo. The Ethics Committee agreed to waive the need for informed consent. The derivation cohort included consecutive patients who underwent CPAP titration for documented OSA by polysomnography between January 2005 and August 2005 at the University-affiliated Sleep Center. The validation cohort represented patients with OSA who underwent a titration study between September 2005 and November of 2005 at a private sleep laboratory (Sleep Disorders Center of Western New York). Demographic information (age, gender) and anthropomorphic measurements (neck circumference (NC), height, weight, and body mass index (BMI)) were obtained from the computerized data records which also included the initial apnea-hypopnea index (AHI) on the diagnostic study and the set of pressures used during CPAP titration. The estimated optimal pressure derived from the regression equation (RE) [5] ($Ppred(RE) = (0.16 \times NC) + (0.13 \times BMI) + (0.04 \times AHI) - 5.12$) was also calculated.

2.2. Sleep studies and CPAP titration

All participants underwent standard overnight polysomnography with recordings of electroencephalogram (EEG), electrooculogram (EOG), submental and bilateral leg electromyograms (EMGs), and electrocardiogram (ECG). Airflow was measured qualitatively by an oral-nasal thermistor and respiratory effort by thoracoabdominal piezoelectric belts (Piezo Crystals, EPM Systems, Midlothian, VA). Measurement of arterial oxyhemoglobin saturation was performed with a pulse oximeter (ASC: Nellcor N-200, Nellcor Puritan Bennett, St. Louis, MO). All signals were collected and digitized on a computerized polysomnography system (ASC: Rembrandt, Aerosep Corporation, Buffalo, NY at University affiliated Sleep Center and Alice 3 system, Healthdyne Technologies, Marietta, GA at the Sleep Disorders Center of Western New York). Sleep stages were recorded in 30-s epochs using the Rechtschaffen and Kales sleep scoring criteria [10]. Each epoch was analyzed for the number of apneas, hypopneas, arousals, and oxygen desaturations. Apnea was defined as the absence of airflow for more than 10 s. Hypopnea was defined as a visible reduction in airflow lasting at least 10 s associated with either a 4% decrease in arterial oxyhemoglobin saturation or an EEG arousal. An arousal was defined according to the criteria proposed by the Atlas Task Force [11].

CPAP titration was conducted on a subsequent night in the sleep laboratory. Patients were initiated at a pressure of 4 cm H_2O. The pressure was gradually increased by 1 cm H_2O every 20 min until the level at which apnea, hypopnea, snoring, and recurrent oxyhemoglobin desaturations, but not arousals, were eliminated. The optimal pressure was defined as the lowest pressure at which the patient had an AHI < 5. Patients who failed to achieve a Popt during CPAP titration were not included in the analysis. In both centers, the definitions and CPAP titration protocol were identical.

2.3. Design of the artificial neural network

A general regression neural network (GRNN) was used in the development of the predictive model [12] using commercially available software (Neuroshell 2, Ward Systems, Frederick, MD). The advantage of the GRNN lies in the fact that whereas conventional nonlinear regression techniques involve a priori specification of the structure of the regression equations to yield a best fit for the data presented, the GRNN circumvents these restrictions by adjusting the surface dimension in which the regression surface resides without constraining it to a specific form. Generalization is optimized by modifying the smoothing factor which determines how tightly the network matches its predictions to the data in the training patterns.

Please cite this article in press as: El Solh AA et al., Predicting effective continuous positive airway pressure in sleep ..., Sleep Med (2006), doi:10.1016/j.sleep.2006.09.005

A three-layer structure was used in the development of the neural network: an input layer, a hidden layer, and an output layer. The input variables selected for the ANN were based on similar parameters used in the regression equation published by Miljeteig and colleagues [5]. Intervening layers of processors, called hidden units, detect higher-order features in the input layer, analyze the signal, and relay the output to other neurons to make a correct response. The number of neurons in the hidden layer is determined by the number of patterns in the training set as GRNNs require one neuron per pattern processed. The output layer of the GRNN provides an estimate of the optimal pressure for the CPAP device to reduce or abolish apneic events.

A fivefold cross-validation approach was used for evaluation [13]. The entire data set of the derivation group was divided with a random number generator into five subsets. Four of the five subsets were pooled and used for training. The data from the fifth subset were used as an evaluation set during training. The entire process was repeated four additional times by rotating the subset that was used as the evaluation set during training. The mean square error was computed for each of the five neural networks on the entire derivation data set. The mean square errors were averaged, and the ANN that had a mean square error closest to the average was selected.

2.4. Statistical analysis

Data are summarized as mean ± standard deviation (SD) for normally distributed variables or median ± 95% CI otherwise. For continuous variables, difference in mean values was assessed using Student's *t*-test or the Mann–Whitney *U*-test. Categorical values were compared using the χ^2 or the Fisher exact test when appropriate. Comparisons between Popt, Ppred(ANN), and Ppred(RE) were made using one-way analysis of variance. A post-hoc test (Dunn's test) was used on all pairwise comparisons. A Spearman correlation was performed to assess the relationship between the actual optimal pressure and the predicted pressures (ANN and RE). Model comparisons were assessed based on the confidence intervals. Agreement between measurements was assessed also by the method of Bland and Altman [14]. Statistical significance was set at $p < 0.05$ (two-tail).

3. Results

A total of 343 patients were identified for inclusion in the derivation cohort. Twenty-nine patients were excluded because of failure to achieve a pressure setting where AHI ⩽ 5 events/h, and two did not complete the titration study. As for the validation group, an optimal pressure was not attained in seven of the 105 patients. Table 1 displays the characteristics of the study population. There were no significant differences in age, gender ratio, neck circumference, BMI, or total sleep time between the derivation and the validation cohort. The distribution of the severity of sleep apnea was also comparable between the two groups.

Five variables were selected to form the input layer: age, gender, BMI, neck circumference, and baseline AHI. The mean square error of the ANN selected was 3.8. The predicted optimal pressures by the neural network for the derivation and the validation cohort are presented in Table 2. Overall, there was no significant difference between the optimal pressure obtained during an overnight polysomnography and the predicted pressure estimated by the ANN ($p = 0.4$). However, the estimated pressure derived from the regression equation underestimated the optimal pressure in both the derivation and the validation group, respectively. The histograms of the differences between Popt and Ppred(ANN) and Popt and Ppred(RE) for the derivation group and validation group are shown in Figs. 1(A) and (B) and Figs. 2(A) and (B), respectively. The correlation coefficient between Popt and Ppred(ANN) in the derivation cohort was 0.86 (95% confidence interval [CI] 0.83–0.88; $p < 0.001$) compared to 0.62 (95% CI 0.54–0.68; $p < 0.001$) for the correlation coefficient between Popt and Ppred(RE). In the validation cohort, the correlation coefficients between Popt and Ppred(ANN) and between Popt and Ppred(RE) were 0.85 (95% CI 0.78–0.9; $p < 0.001$) and 0.6 (95% CI 0.53–

Table 1
Characteristics of the study population

	Derivation cohort ($n = 311$)	Validation cohort ($n = 98$)	*P*-value
Age (years)	49.6 ± 12.4	51.4 ± 12.3	0.2
Gender (M/F)	184/127	50/48	0.2
Neck circumference (cm)	16.7 ± 1.9	17.1 ± 1.9	0.1
BMI (kg/m^2)	35 (34–37)	37 (34–41)	0.1[a]
Epworth score	10 (9–10)	12 (10–14)	0.07[a]
Total sleep time (min)	352 ± 63	361 ± 67	0.3
AHI (h^{-1})	33 (28–38)	39 (29–47)	0.5[a]
AHI 5 to <15/h, n (%)	52 (17%)	17 (17%)	1.0
AHI 15 to <30/h, n (%)	91 (29%)	22 (22%)	
AHI ⩾ 30/h, n (%)	168 (54%)	59 (61%)	

[a] Mann–Whitney test.

Table 2
Comparison of the pressures obtained by polysomnography, neural network, and regression analysis (median, IQ)

	Popt	Ppred (ANN)	Ppred (RE)
Derivation cohort	8.0 (7.0–12.0)	8.3 (7.6–11.1)	3.8 (2.5–5.5)[*,†]
Validation cohort	8.0 (6.0–12.0)	8.9 (7.5–11.9)	4.5 (2.8–6.0)[*,†]

[*] $p < 0.01$ compared to Popt (Dunn's test).
[†] $p < 0.01$ compared to Ppred (ANN) (Dunn's test).

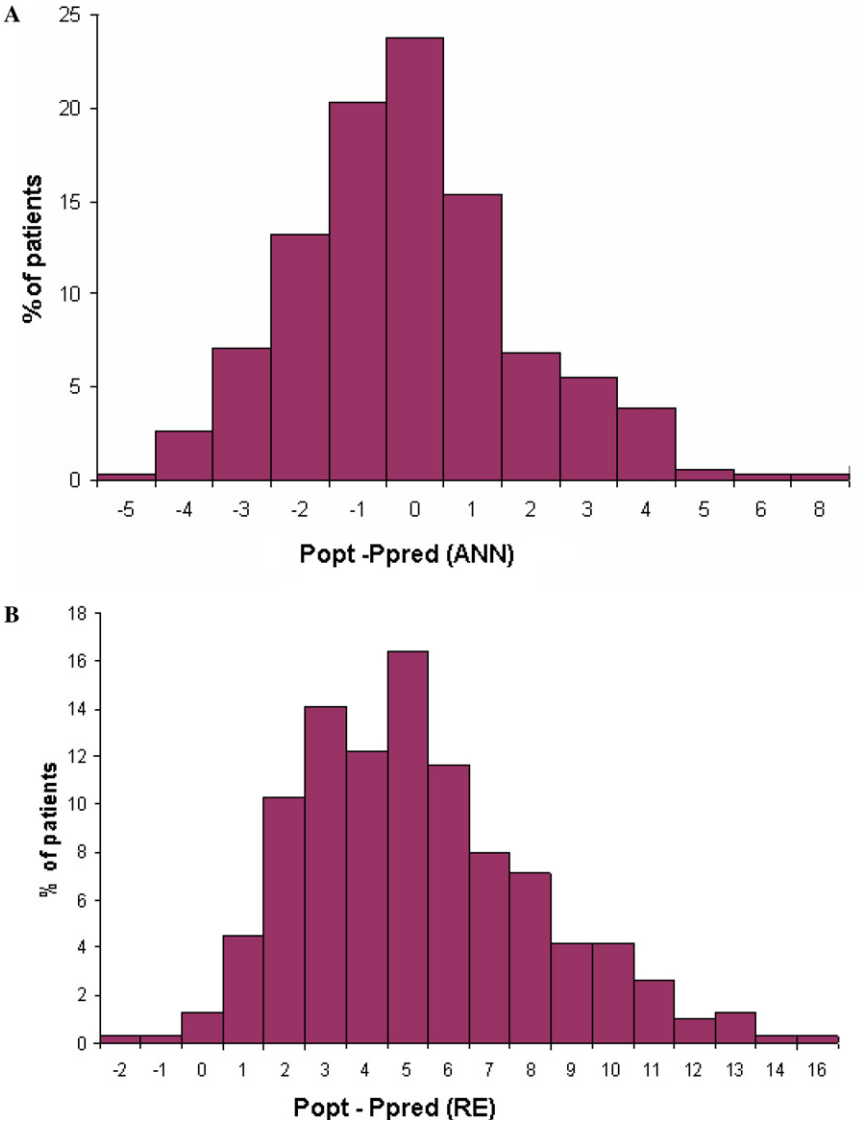

Fig. 1. Histograms of the differences between optimal pressures and predicted pressures by the artificial neural network (A) and by the regression equation (B) in the derivation cohort. ANN = artificial neural network; RE = regression equation.

0.76; $p < 0.001$), respectively. In both the derivation and the validation cohorts, the performance of ANN was superior to the logistic regression model ($p < 0.001$ and $p < 0.001$, respectively).

Fig. 3 shows the level of agreement between the Popt and Ppred(ANN) using the Bland and Altman analysis for the entire cohort. The plot reveals that the majority of the estimated pressures by the neural network fall within 95% confidence interval from the calculated paired pressure mean.

4. Discussion

The optimal prescription for CPAP therapy in patients with OSA is that which most effectively prevents the adverse consequences of OSA while causing the least discomfort and the lowest risk of complications. A central element of the CPAP prescription is the pressure level which is typically derived from a titration study. Various solutions have been proposed as alternatives to conventional titration: "partial-night" trials [15], automatic titration with auto-CPAP devices [16], and pressure prediction using mathematical formulas [5,17]. The present study is the first to present a validated ANN to predict effective CPAP that relies on a combination of anthropomorphic and clinical data, the majority of which have been found to be significantly correlated with optimal CPAP [5].

The findings of the study point to high performance accuracy of the ANN when compared with overnight polysomnography. When applied to CPAP titration, the pressure established by the neural network fell within 3 cm of H_2O above or below the optimal pressure set by polysomnography for 92% of the patients. In

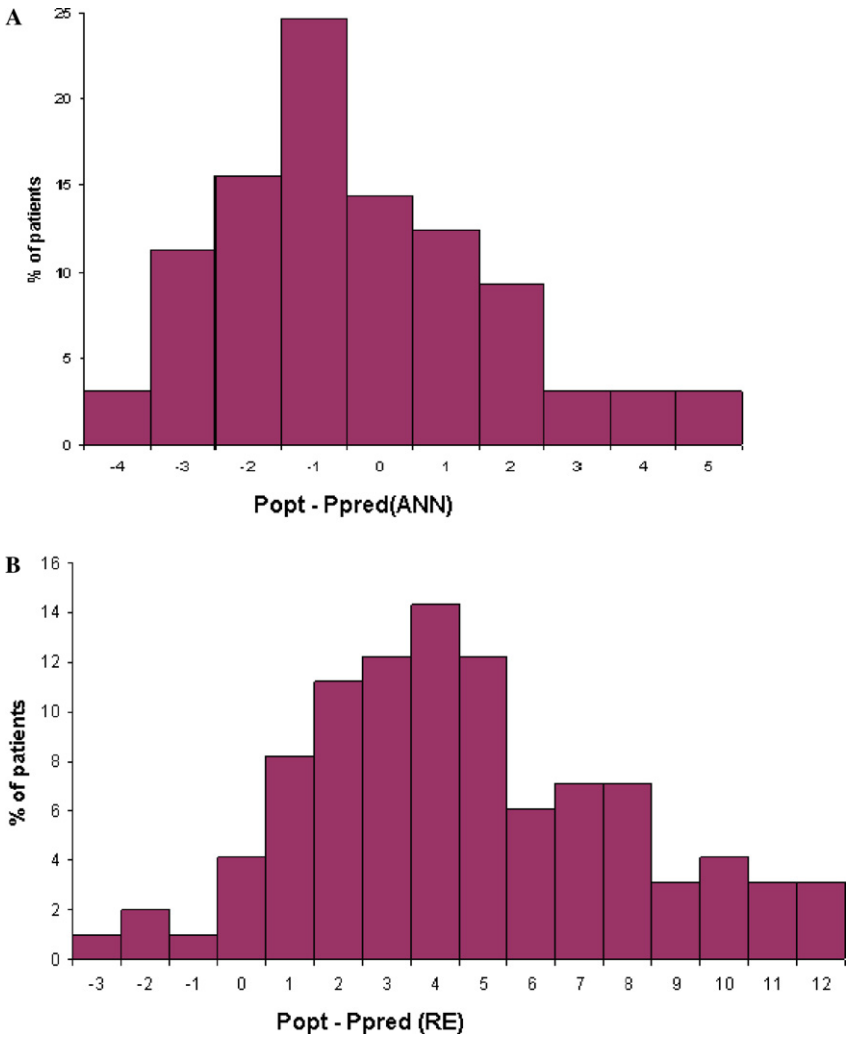

Fig. 2. Histograms of the differences between optimal pressures and predicted pressures by the artificial neural network (A) and by the regression equation (B) in the validation cohort. ANN = artificial neural network; RE = regression equation.

contrast, the overall accuracy of the regression equation was poor, as only 40% of patients had their estimated pressure fall within 3 cm of H_2O of the optimal pressure determined by overnight sleep study. These results coincide with the observations of previously published validation studies from other sleep centers [6,7]. Overall, the regression equation tended to underestimate the optimal pressure in both the derivation and the validation cohort. We attribute the deterioration in the predictive ability of the regression equation to the phenomenon of "model drift" [18]. The model drift could stem from either a modified definition of apneas/hypopneas, an improved sensitivity of diagnostic tools, or change in the disease pattern. It has been argued also that the discrepancy in optimum pressure prediction by the regression equation may be attributed to a difference in the population under study. Considering that the derivation of the regression equation was performed in a mostly male population, a preponderance of female participants might have skewed the CPAP prediction, as women tend to have a lower severity of sleep apnea and smaller neck circumference [19,20]. While the gender distribution of our population was equivalent in both cohorts, the neural network included a gender adjustment to account for inherent differences in sleep characteristics.

It is intrinsic to any predictive model that the ANN-predicted optimal pressure may overestimate the effective CPAP. In such an event, the CPAP level can then be decreased by 1 cm H_2O during the CPAP titration every 20 min until the level at which apnea, hypopnea, snoring, or recurrent oxyhemoglobin desaturations would recur. However, we have to acknowledge that the effectiveness of such a strategy can only be assessed during a prospective study.

The exclusion of patients with unsuccessful titration in our study might explain the higher level of accuracy of the neural network compared to the regression equation. Analysis of those patients who were excluded did not, however, reveal a common pattern or characteristic.

Please cite this article in press as: El Solh AA et al., Predicting effective continuous positive airway pressure in sleep ..., Sleep Med (2006), doi:10.1016/j.sleep.2006.09.005

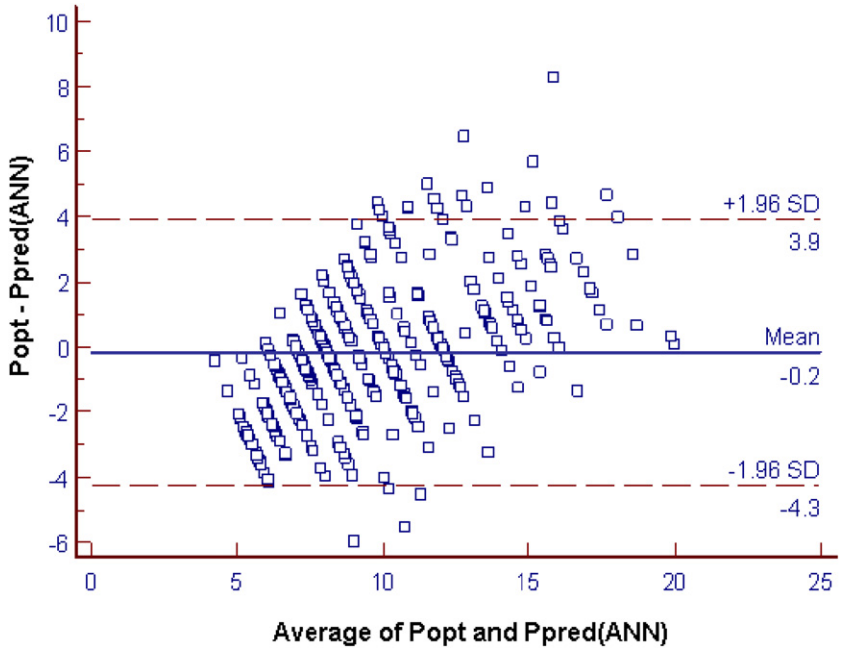

Fig. 3. Bland and Altman for optimal pressures vs. predicted pressures by the artificial neural network for the entire study population.

However, the high level of CPAP titration failure reported in other series of 16–40% [6,21] despite the use of the regression equation suggests the underperformance of the predictive model and underlines the need for a more accurate and cost-effective algorithm. A prospective implementation of the neural network will be required to assess the impact of this technology on the rate of CPAP titration failure.

A major departure from previous studies is the cutoff we have used to define titration success. We have selected an AHI < 5/h compared to AHI $\leqslant 10$/h used in other studies. Despite the fact that an AHI $\leqslant 10$/h was considered as the criterion for defining OSA in screening studies [22,23] or for successful titration when relying on mathematical formulas [5,21], we stipulated that the likelihood of achieving a clinically significant optimal pressure is increased when an AHI of <5/h was targeted, thus requiring a smaller number of pressure increments during a titration study.

There are several potential limitations to the study. First, neural networks have the ability to approximate predictive output to any desirable degree of accuracy when provided with enough running time. This could result in overfitting, particularly when there is an attempt to increase the processing power of the network by adding a large number of hidden neurons. In this case, the network will end up learning not only the training set but also the noise in the data, which leads to poor generalization. The accuracy of prediction observed in the validation set points, however, tends to argue against this possibility and reinforces the fact that the network architecture is based on robust features rather than memorizing the idiosyncrasies embedded in the data set. We should mention that the validation study was conducted on a set where the optimal pressure was determined by a regular titration study first. This step is important for two reasons: to assess the reproducibility of the model in a setting other than the one used to develop the predictive model, and to remove the potential for bias that could occur from being aware of the ANN output beforehand (blinding effect). Once this step is deemed successful, the ANN would then be used to examine its effect on CPAP titration. Second, the accuracy of the network is subject to advances in technology, improvement in sensitivity of diagnostic equipment (i.e., use of nasal pressure to detect airflow), and changes in disease definition. Similar to any prognostic model, periodic recalibration of the neural network is thus required to maintain accuracy. Third, a frequently cited limitation in the literature is the fact that little is known about the pathways used by the ANN to predict outcome [24]. These pathways are complex and do not convey an understanding of the structure of reasoning. Unlike the logistic regression equation, the relationships between variables are not explicit. The superior predictive ability of the ANN, however, would offset this limitation. With the wide availability of computers and modern software in medical practice, the neural network algorithm can be published on a website allowing easily accessibility for daily use.

In summary, the proposed ANN outperformed the traditional regression equation in predicting optimal CPAP. While it is not intended to be a substitute for overnight polysomnography, the high level of agreement between ANN and overnight polysomnography indi-

Please cite this article in press as: El Solh AA et al., Predicting effective continuous positive airway pressure in sleep ..., Sleep Med (2006), doi:10.1016/j.sleep.2006.09.005

cates that the ANN may be used to facilitate the titration study by providing baseline pressure from which to start CPAP titration.

References

[1] Young T, Palta M, Dempsey J, et al. The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 1993;328:1230–5.
[2] Wright J, Johns R, Melville A, et al. Health effects of obstructive sleep apnea and the effectiveness of continuous positive airway pressure: a systematic review of the research evidence. BMJ 1997;314:851–60.
[3] Jenkinson C, Davies RJ, Mullins R, et al. Comparison of therapeutic and subtherapeutic nasal continuous positive airway pressure for obstructive sleep apnoea: a randomised prospective parallel trial. Lancet 1999;353:2100–5.
[4] Montserrat J, Ferrer M, Hernandez L, et al. Effectiveness of CPAP treatment in daytime function in sleep apnea syndrome: a randomized controlled study with an optimized placebo. Am J Respir Crit Care Med 2001;164:608–13.
[5] Miljeteig H, Hoffstein V. Determinants of continuous positive airway pressure level for treatment of obstructive sleep apnea. Am Rev Respir Dis 1993;147:1526–30.
[6] Rowley J, Tarbichi A, Badr S. The use of a predicted CPAP equation improves CPAP titration success. Sleep Breath 2005;9:26–32.
[7] Gokcebay N, Iqbal S, Yang K, et al. Accuracy of CPAP predicted from anthropometric and polysomnographic indices. Sleep 1996;19:600–1.
[8] Baxt WG. Use of an artificial neural network for the diagnosis of myocardial infarction. Ann Intern Med 1991;115:843–8.
[9] Rutledge R, Fakhry S. The use of neural networking and polynomial regression to predict the outcome in the critically ill: a comparison with APACHE II [abstract]. Crit Care Med 1993;21(Supplement):S213.
[10] Rechtschaffen A, Kales A. A manual of standardized technology, techniques and scoring system for sleep stages of human subjects. Los Angeles, CA: Brain Information Service, Brain Information Institute, University of California; 1968.
[11] The Atlas Task Force. EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Task Force of the American Sleep Disorders Association. Sleep 1992;15:173–84.
[12] Specht D. A general regression neural network. IEEE Trans Neural Networks 1991;2(6):568–76.
[13] Weiss SM, Kulikowski CA. How to estimate the true performance of a learning system. In: Computer systems that learn. Palo Alto, CA: Morgan Kaufmann; 1991.
[14] Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;11:307–10.
[15] Sanders MH, Kern NB, Costantino JP, et al. Adequacy of prescribing positive airway pressure therapy by mask for sleep apnea on the basis of a partial-night trial. Am Rev Respir Dis 1993;147:1169–74.
[16] Meurice JC, Marc I, Series F. Efficacy of autoCPAP in the treatment of obstructive sleep apnea/hypopnea syndrome. Am J Resp Crit Care Med 1996;153:794–8.
[17] Panagou P, Mermingis H, Tsipra S, et al. Prediction of effective continuous positive airway pressure (CPAP) in obstructive sleep apnea syndrome (OSA). Eur Respir J 2005;26(Supplement):111S.
[18] Grant BJB, Choi Y, Sathananthan A, et al. Model drift in a questionnaire that predicts severity of obstructive sleep apnea. Sleep 2006;29:A210.
[19] O'Connor C, Thornley K, Hanley P. Gender differences in the polysomnographic features of obstructive sleep apnea. Am J Respir Crit Care Med 2000;161:1465–72.
[20] Rowley J, Aboussouan L, Badr M. The use of clinical prediction formulas in the evaluation of obstructive sleep apnea. Sleep 2000;23:929–38.
[21] Oliver Z, Hoffstein V. Predicting effective continuous positive airway pressure. Chest 2000;117:1061–4.
[22] El Solh A, Mador J, Ten-Brock E, et al. Validity of neural network in sleep apnea. Sleep 1999;22:105–11.
[23] Flemons W, Whitelaw W, Brant R, et al. Likelihood ratios for a sleep apnea clinical prediction rule. Am J Respir Crit Care Med 1994;150:1279–85.
[24] Hart A, Wyatt J. Evaluating black-boxes as medical decision aids: issues arising from a study of neural networks. Med Inform 1990;15:229–36.

Please cite this article in press as: El Solh AA et al., Predicting effective continuous positive airway pressure in sleep ..., Sleep Med (2006), doi:10.1016/j.sleep.2006.09.005