REVIEW ARTICLE

Manual Versus Automated Sleep Scoring for Diagnosis of Obstructive Sleep Apnoea

Dipti Gothi and Sonam Spalgais
Department of pulmonary Medicine ESI-PGIMSR Basaidarapur New Delhi 110015

DOI No:10.5958/0974-0155.2016.00002.4 *Indian J Sleep Med 2016; 11.1, 5-7*

Abstract

There are two types of sleep scoring techniques for evaluation of obstructive sleep apnea (OSA) on polysomnography, they are manual and automated scoring. Manual scoring with expert technician is considered as a gold standard scoring technique. It evaluates total sleep time, stage of sleep, and apnoea or hypopnea index (AHI) better than automated scoring. However, this technique needs more manpower, money, and infrastructure. Automated scoring technique is simple, cost effective, and less time consuming. Both techniques can be performed with home sleep testing or in laboratory polysomnography. Though, automated scoring technique is less accurate in diagnosis of mild form of OSA, it is a viable option for moderate and severe OSA especially where the patient load is high and facilities are limited.

Introduction

Obstructive sleep apnea (OSA) has become a public health problem due to its high frequency rate. The estimated prevalence of symptomatic OSA is between 3 and 8% in men and 1 and 5% in women[1]. Identification of true patients and accurate interpretation of disease severity directly affects the success of the treatment. Polysomnography (PSG) is used for recording neurophysiological, cardiorespiratory, and other physiological parameters during sleep at night[2]. It is the gold standard diagnostic method for OSA. The scoring of PSG parameters which determine OSA can be done by manual and automated (computational) techniques. Both the scoring techniques have their own advantages and disadvantages. Manual scoring is considered as a gold standard technique. However, automated scoring can also be useful if used judicially and appropriately.

Address for correspondence:
Dr Dipti Gothi
Professor, Pulmonary Medicine
ESI-PGIMSR, Delhi
E-mail: diptigothi@gmail.com

Manual scoring

Manual scoring technique requires a trained sleep technician to score initiation of sleep, stages of sleep, apnea or hypopnea index, sleep onset, and arousals and to follow each epoch on monitor. The length of one epoch is 30 seconds. So, 6 hour of sleep recording requires about 720 epochs to be evaluated separately. This process takes nearly 80-180 minutes, even when performed by the most experienced technician. Thus, it requires expertise and is time-consuming[3]. Since, manual scoring is done manually which has considerable interscorer and intrascorer variability makes its reliability and reproducibility questionable[4]. In India, manual scoring is usually performed for in-laboratory PSG. The cost of in-laboratory PSG is 3-5 times higher than that of home sleep testing. Manual scoring is rarely practiced for home sleep testing. It is essential for a practicing sleep physician to know the scoring system, so that in case of doubt even home sleep test can be crosschecked.

Automated scoring

In automated scoring, the installed software reports apnoeas, hypopnoeas, snoring, sleep onset, stages of

sleep, and the apnoea or hypopnea index[5]. Automated scoring can be performed in a shorter time without the need for expert technician support. There may be errors, particularly in recognizing the process of passing from the awake state to stage I and rapid eye movement (REM) sleep and in distinguishing arousal, epileptic activity, and parasomnia[6]. Automatic systems are not sufficiently validated and lack precision in discriminating sleep stages or detecting respiratory episodes in clinical practice[7]. However, it is a cost effective method as compared to manual scoring. In India, this scoring technique is usually done for home sleep testing. If automated scoring of laboratory performed PSG correlates with clinical finding then one may not review each and every study.

Studies on Manual verses automated scoring

Some studies have shown that automated scoring is not as good as manual scoring. Ozturket al.[8] compared 30 patients diagnosed with OSA and found the rate of consistency to be 58% between manual and automatic scoring. Asik Met al[3] in study of direct comparison of manual verses automated scoring showed that the percentage of rapid eye movement (REM) stage was significantly lower for all patients and all OSA subgroups in automated scoring. The non rapid eye movement (NREM) AHI was significantly higher in the automated scoring (p=0.002). The rate of specificity was only 90.0 and 86.6% for mild and moderate OSA, respectively. Barreiro B et al[7] in comparison between automatic and manual analysis showed that automated scoring underestimates the duration of the stages of REM sleep (P<0.007) and deep sleep (P<0.3). Aurora RN et al[9] compared 2 types of devices, Apnea Link Plus monitor and Emletta in 200 subjects. The difference between manual and automated AHI was 6.1 (95% CI, 4.9-7.3) and 4.6 (95% CI, 3.5-5.6) events/hour, respectively. Thus, agreement between automated and manual scoring of home sleep tests varies depending on the function of the portable device and definition of disordered breathing used. As per these studies, automated scoring underestimates the AHI compared with manual scoring.

There are a few other studies which favour automated scoring. Ernst G et al.[5] showed that there was no significant difference between automated and manual apnea or hypopnea indexes. The agreement between manual and automatic AHI for AHI ≥30 was 94%, with a Kappa coefficient of 0.83 (p< 0.001). The AUC-ROC, sensitivity, and specificity were 0.99, 86% and 97%, respectively. Asik M et al.[3] in a study on 120 patients on manual verses automated scoring showed that there were no statistically significant differences in the total AHI and REM. AHI between two scoring techniques (p=0.053 and p=0.319, respectively). Sensitivity, specificity, positive predictive value, and negative predictive value of automated scoring were 98.88, 93.33, 97.80, and 95.55%, respectively. Masa JF et al.[10] in study of 366 randomized patients had a specificity of 93% for automated and 94% for manual scorings in patients of home sleep testing. In an another study average scores of the 10 technologists for the 70 pairs of automated scoring verses manual scoring showed that the automatic system yielded results that were similar to those obtained by experienced technologists[4].

Thus, overall the rate of consistency has been reported to range from 60% to >90% in the literature[3,4,11,12]. Possible causes of this inconsistency between the two techniques is postulated to be due to variation in devices in different studies, the restricted number of samples in some studies, or inhomogeneous cases of the sampling in terms of the severity of OSA and the expertise of technician. The calculated NREM stage N1, N2, N3 REM sleep time is likely to be better with an experienced technologists. Arousals, apneas, and hypopnea are also likely to be scored better with an experienced technician.

Indian scenario

There is lack of proper studies on manual scoring versus automated scoring in India. Based on the available data, the question arises whether manual scoring should be done to achieve accurate AHI at the cost of huge manpower, money and infrastructure. There is a lack of proper population prevalence study of OSA in India due to paucity of health care facilities with multi-channel polysomnography equipment. The various studies in last decade reported that the prevalence of OSA was nearly 3–4% in India[13]. With a population of approximately 1.2 billion it is estimated that nearly 40–50 million suffer from OSA. There are only about 500 PSG laboratories across India. With nearly 600 sensitized technician across the country, it is difficult to know how many are well-trained technician. It is not technically possible to evaluate all 40–50 million sleep studies by 600 technicians.

Besides, there is night-to-night variability in AHI scoring and it also varies with different technician in manual technique. It should also be remembered that there is insufficient evidence to draw conclusions regarding the efficacy and/or effectiveness of CPAP treatment for mild OSA[14,15].

Overall, proper sleep time detected by manual scoring makes a significant difference in the diagnosis of mild OSA. The automated scoring is possibly more cost-effective for moderate to severe OSA. If automated scoring generates mild OSA or there is discrepancy between history, examination, and automatically generated report, then review with manual scoring is required even if it is a home sleep testing.

References

1. **Pien GW**, Rosen IM and Fields BG. Sleep Apnea Syndromes: Central and Obstructive. In:Grippi MA, editor. Fishman's Textbook of Pulmonary Disease and Disorders New York: *McGraw Hill Medicine;* 2015.p.3042-43.
2. **Kokturk O**. Diagnostic methods and polysomnography for respiratory disorders during sleep. In: Respiratory system and diseases. Ozlu T, Metinta M, Karadað M, Kaya A (Eds). Ýstanbul: *Ýstanbul Týp Kitapevi;* 2010.p.2109-25.
3. **Asýk M**, Bostanc A and Turhan M. Comparison of Manual and Automated Scoring Techniques in Polysomnography.*Turk Arch Otolaryngol* 2014;52:17-21.
4. **Malhotra A,** Younes M, Kuna ST,Benca R, Kushida CA, Walsh J, et al. Performance of an Automated Polysomnography Scoring System Versus Computer-Assisted Manual Scoring.*Sleep*. 2013;36:573-82
5. **Ernst G**, Bosio M, Salvado A, Nogueira F, Nigro C and Borsini E. Comparative Study between Sequential Automaticn and Manual Home Respiratory Polygraphy Scoring Using a Three-Channel Device: Impact of the Manual Editing of Events to Identify Severe Obstructive Sleep Apnea.*Sleep Disorders*2015.p.1-5.Article ID 314534,http://dx.doi.org/10.1155/2015/314534
6. **Kokturk O**. Scoring sleep recordings. *Solunum* 2013;15:14-29.
7. **Barreiro B**, Badosa G, Quintana S, Esteban L, and Heredia JL. Comparison Between Automatic and Manual Analysis in the Diagnosis of Obstructive Sleep Apnea-Hypopnea Syndrome. *Arch Bronconeumol* 2003;39:544-8
8. **Ozturk O**, Mutlu LC, Saðcan G, Deniz Y, Cuhadaroðlu C. The concordance of manuel (visual) scoring and automatic analysis in sleep staging. *Tuberk Toraks* 2009;57:306-13.
9. **Aurora RN,** Swartz R, and Punjabi N M. Misclassifi cation of OSA Severity With Automated Scoring of Home Sleep Recordings. *Chest* 2015;147(3):719–727.
10. **Masa JF**, Corral J, Pereira R, Duran-Cantolla J, Cabello M and Hernandez-Blasco L. et al. Effectiveness of sequential automatic manual home respiratory polygraphy scoring. *Eur Respir J* 2013;41:879–887.
11. **Schaltenbrand N**, Lengelle R, Toussaint M, Luthringer R, Carelli G, Jacqmin A, et al. Sleep stage scoring using the neural network model: comparison between visual and automatic analysis in normal subjects and patients. *Sleep* 1996;19:26-35.
12. **Redline S,** Budhiraja R, Kapur V, Marcus CL, Mateika JH, Mehra R, et al. The scoring of respiratory events in sleep: reliability and validity. *J Clin Sleep Med* 2007;3:169-200.
13. **Sharma SK** and Ahluwalia. Epidemiology of adult obstructive sleep apnoea syndrome in India. *Indian J Med Res* 2010;131:171-175
14. **Rosenthal L,** Gerhardstein R, Lumley A, Guido P, Day R, Syron ML. et al. CPAP therapy in patients with mild OSA: implementation and treatment outcome. *Sleep Med*.2000;1:215-220.
15. **Engleman HM,** Kingshott RN, Wraith PK, Mackay TW, Deary IJ and Douglas NJ. Randomized placebo-controlled crossover trial of continuous positive airway pressure for mild sleep Apnea/Hypopnea syndrome. *Am J RespirCrit Care Med* 1999;159(2):461-7.