Sleep Medicine 16 (2015) 173–180

Contents lists available at ScienceDirect

Sleep Medicine

journal homepage: www.elsevier.com/locate/sleep

Original Article

Polysomnography using abbreviated signal montages: impact on sleep and cortical arousal scoring

Warren R. Ruehland [a,b,*], Thomas J. Churchward [a], Linda M. Schachter [c], Tristia Lakey [d], Natalie Tarquinio [e], Fergal J. O'Donoghue [a], Maree Barnes [a], Peter D. Rochford [a]

[a] Institute for Breathing and Sleep, Austin Health, Heidelberg, Vic., Australia
[b] Department of Medicine (Austin Health/Northern Health), University of Melbourne, Heidelberg, Vic., Australia
[c] Sleep Services Australia, Melbourne, Vic., Australia
[d] Advanced Pulmonary and Sleep Diagnostics, Melbourne, Vic., Australia
[e] Bayside Sleep Analysis, Melbourne, Vic., Australia

ARTICLE INFO

Article history:
Received 28 April 2014
Received in revised form 7 October 2014
Accepted 13 November 2014
Available online 26 November 2014

Keywords:
Obstructive sleep apnoea
Diagnosis
Portable monitoring
Accuracy
Precision
Inter-scorer reliability

ABSTRACT

Objective: This study examined the impact of using two abbreviated signal montages on the accuracy, precision and inter-scorer reliability of polysomnography (PSG) sleep and arousal scoring, compared to a standard reference montage, in a cohort of patients investigated for obstructive sleep apnoea (OSA). One abbreviated montage incorporated two signals dedicated to sleep and arousal scoring, and the other incorporated a single signal.
Methods: Four scorers from two laboratories each scored 15 PSGS four times in random order: once using each abbreviated montage and twice using the reference montage.
Results: Use of the two-signal montage resulted in small changes in the distribution of sleep stages, a reduction in the arousal index and resultant reductions in sleep and arousal scoring agreement. For the one-signal montage, although similar magnitude sleep stage distribution changes were observed, there were larger reductions in the arousal index, and sleep and arousal scoring accuracy. Additionally, using the one-signal montage, there were statistically significant reductions in the precision of summary statistics including total sleep time (TST) and the amount of rapid eye movement (REM) sleep scored, and reductions in the inter-scorer reliability of REM sleep and arousal scoring.
Conclusions: These findings demonstrate that abbreviated signal montages may result in underestimation of the arousal index and, depending on the montage, poorer precision in TST and REM sleep scoring, with potential consequences for apnoea–hypopnoea index (AHI) measures and OSA diagnosis. The results highlight the importance of careful evaluation of PSG results when using portable devices that have restricted signals, and they offer guidance for future PSG and portable monitoring standards.

1. Introduction

It is standard clinical practice to confirm obstructive sleep apnoea (OSA) diagnosis using in-laboratory polysomnography (PSG); however, it is increasingly recognised that portable monitoring (PM) may be an acceptable alternative, with acknowledgement and understanding of accompanying limitations [1].

Although many PM devices have abbreviated signal recording capabilities compared to full PSG, there is an advantage in quantifying sleep and cortical arousals, requiring the recording of fast sampling rate signals such as electroencephalography (EEG), electrooculography (EOG), and electromyography (EMG). Recording of these signals allows: (i) assessment of the impact of any respiratory disturbance on sleep architecture; (ii) the use of total sleep time (TST) rather than total recording time (TRT) as the denominator in calculating indices of respiratory or sleep disturbance; (iii) verification of rapid eye movement (REM) sleep sampling, important due to the incidence of REM-related OSA, estimated to have a prevalence of approximately 35% in clinical OSA populations [2]; and (iv) the use of respiratory event scoring criteria requiring airflow reduction accompanied by cortical arousal.

Despite these theoretical advantages, the most recent clinical guidelines for use of PM to diagnose OSA [1] did not consider devices that were capable of measuring sleep. This was because there were no new data available comparing such devices to PSG since previous guidelines [3], stated that evidence was lacking to recommend their clinical use.

* Corresponding author. Institute for Breathing and Sleep, Austin Health, Ground Floor Bowen Centre, Studley Road, Heidelberg, Vic. 3084, Australia. Tel.: +61 3 9496 3528; fax: +61 3 9496 5128.
E-mail address: Warren.Ruehland@austin.org.au (W.R. Ruehland).

http://dx.doi.org/10.1016/j.sleep.2014.11.005
1389-9457/

PM devices with fast sampling capability may still be restricted to the number of signals that can be dedicated to the recording and scoring of sleep and cortical arousals. Current guidelines [4] recommend the use of six primary signals (three EEG, two EOG, and one EMG) for recording and scoring of sleep and cortical arousals in PSG. We have previously shown that the use of four primary signals, incorporating one EEG, results in only small changes in the distribution of sleep stages [5] and no statistically significant differences in sleep or cortical arousal scoring inter-scorer or intra-scorer reliability. However, there are no data to guide clinical practice on how further signal restrictions may impact the scoring of sleep and arousals. Such information is crucial for those using PM devices with limited signal recording capabilities.

This study aimed to examine the impact of using two abbreviated signal montages on the accuracy, precision, and inter-scorer reliability of PSG sleep and arousal scoring, compared to a standard reference montage, in a cohort of patients presenting for the investigation of OSA. One abbreviated montage incorporated two signals dedicated to recording and scoring of sleep and cortical arousals, whereas the other utilised a single signal.

2. Methods

2.1. Design

This study was a prospective, non-blinded, randomised comparison of sleep and arousal scoring using two abbreviated montages compared to a standard reference montage; it was approved by the institutional Human Research Ethics Committee.

2.2. Patient selection

The study utilised 15 single-night PSGs sourced during July and August 2006 from the Austin Health sleep laboratory in Melbourne, Australia, from consecutive patients investigated for OSA. PSGs were not considered if they were primarily conducted for non-OSA sleep disorders, research or treatment implementation.

2.3. PSG recordings

PSGs were recorded using Compumedics S-series or E-series monitoring equipment (Abbotsford, VIC, Australia). The recording configuration consisted of: one EEG signal (C4/A1), two EOG signals (left and right outer canthus (OC)/Fpz), one combined EEG/EOG signal (Fp1/LOC), submental EMG, electrocardiogram (ECG), nasal pressure, body position, thoracic and abdominal excursion (inductance plethysmography), oxygen saturation via finger pulse oximetry (Nellcor N-595; Nellcor Inc, Boulder, CO, USA), left and right leg movement and sound.

2.4. PSG scoring

Sleep and arousal scoring were performed manually, in a single pass, using Profusion PSG 2 software (Compumedics, Abbotsford, VIC, Australia), based on published standards available at the time of the study [6,7]. Apnoea–hypopnoea indices (AHIs) determined using "Chicago Criteria" [8] during the original clinical investigation characterised the patient sample.

During scoring, PSGs were configured to display one of three montages: (i) a reference montage (M_{Ref}) incorporating one EEG signal (C4/A1), two EOG signals and one EMG signal, selected as it was in the minimum configuration recommended for use in PSG [6]; (ii) an abbreviated two-signal montage (M_2) incorporating one EEG signal (C4/A1) and one EOG signal (LOC/Fpz); or (iii) an abbreviated one-signal montage (M_1) incorporating the single combined EEG/EOG signal (Fp1/ROC). Care was taken to ensure that the display size of all signals was identical regardless of the number of signals displayed.

During abbreviated montage scoring, the start and end of REM sleep were not defined by EMG changes, but they were defined by the presence/absence of REMs and stage 2 sleep features; arousals in REM did not require concurrent EMG elevation with EEG frequency shift.

2.5. Scorers

Four scorers from two separate Australian clinical sleep investigation services participated: two from Sleep Services Australia, Melbourne, and two from the Austin Hospital, Melbourne. All scorers participated in scoring concordance programmes and they were experienced in abbreviated montage scoring.

2.6. Protocol

For each scorer, all PSGs and versions were de-identified and presented in random order with the exception that no 2 versions of the same PSG were ever presented consecutively. A second copy of M_{Ref} (M_{Ref2}) was later scored to allow comparison of abbreviated montage accuracy and precision against intra-montage scoring repeatability. Thus, each scorer analysed all 15 PSGs four times each, twice using M_{Ref} and once each using M_1 and M_2.

2.7. Analysis

The analysis involved assessment of: (i) PSG summary statistic accuracy, (ii) PSG summary statistic precision, (iii) event-by-event/epoch-by-epoch accuracy, and (iv) event-by-event/epoch-by-epoch inter-scorer reliability. For all assessments of accuracy and precision, the mean value of all four scorers was used for statistical analysis. Distributions of the differences between numerous parameter pairs were skewed and so non-parametric Friedman tests were undertaken for all comparisons, with post hoc analysis conducted using Wilcoxon signed-rank tests.

2.7.1. Summary statistics accuracy

Statistical analysis compared repeated measure differences in PSG sleep and arousal summary statistics between M_1, M_2, M_{Ref}, and M_{Ref2}. The distribution of epoch-by-epoch sleep stage specific discordances was examined to elucidate the cause of any observed differences.

2.7.2. Summary statistics precision

Precision of PSG sleep and arousal summary statistics for M_1, M_2, and M_{Ref2} each were assessed using the median absolute deviation (MAD) about the median difference from M_{Ref}. Statistical analysis compared repeated measure differences in precision between M_1, M_2, and M_{Ref2}.

2.7.3. Epoch-by-epoch/event-by-event accuracy

Epoch-by-epoch accuracy of sleep and arousal scoring for M_1, M_2, and M_{Ref2} each versus M_{Ref} was assessed using Cohen's pair-wise kappa [9], modified for continuous measurements for arousal scoring [10]. Statistical analysis compared repeated measure differences in accuracy between M_1, M_2, and M_{Ref2}. Raw agreement, expressed as percentage agreement [11] for sleep and as proportion of specific agreement (PSA) for positive ratings [12] for arousals, was also presented for comparison.

2.7.4. Epoch-by-epoch/event-by-event inter-scorer reliability

Epoch-by-epoch inter-scorer reliability of sleep and arousal scoring for M_1, M_2, M_{Ref}, and M_{Ref2} each were assessed using Fleiss' multi-scorer kappa [11,12], modified for continuous measurements for