Effects of signal artefacts on electroencephalography spectral power during sleep: quantifying the effectiveness of automated artefact-rejection algorithms

JIANBO LIU[1], SRIDHAR RAMAKRISHNAN[1], SRINIVAS LAXMINARAYAN[1], MAXWELL NEAL[1], DAVID J. CASHMERE[2], ANNE GERMAIN[3] and JAQUES REIFMAN[1]

[1]Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, US Army Medical Research and Materiel Command, Fort Detrick, MD, USA; [2]University of Pittsburgh Medical Center, Pittsburgh, PA, USA; [3]Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

Keywords
EEG spectral power analysis, Brunner's algorithm, Doman's algorithm

Correspondence
Jaques Reifman PhD, Senior Research Scientist and Director, Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, US Army Medical Research and Materiel Command, ATTN: MCMR-TT, 504 Scott Street, Fort Detrick, MD 21702.
Tel.: +301-619-7915;
fax: +301-619-1983;
e-mail: jaques.reifman.civ@mail.mil

Accepted in revised form 21 May 2017; received 17 February 2017

DOI: 10.1111/jsr.12576

SUMMARY

Electroencephalography (EEG) recordings during sleep are often contaminated by muscle and ocular artefacts, which can affect the results of spectral power analyses significantly. However, the extent to which these artefacts affect EEG spectral power across different sleep states has not been quantified explicitly. Consequently, the effectiveness of automated artefact-rejection algorithms in minimizing these effects has not been characterized fully. To address these issues, we analysed standard 10-channel EEG recordings from 20 subjects during one night of sleep. We compared their spectral power when the recordings were contaminated by artefacts and after we removed them by visual inspection or by using automated artefact-rejection algorithms. During both rapid eye movement (REM) and non-REM (NREM) sleep, muscle artefacts contaminated no more than 5% of the EEG data across all channels. However, they corrupted delta, beta and gamma power levels substantially by up to 126, 171 and 938%, respectively, relative to the power level computed from artefact-free data. Although ocular artefacts were infrequent during NREM sleep, they affected up to 16% of the frontal and temporal EEG channels during REM sleep, primarily corrupting delta power by up to 33%. For both REM and NREM sleep, the automated artefact-rejection algorithms matched power levels to within ~10% of the artefact-free power level for each EEG channel and frequency band. In summary, although muscle and ocular artefacts affect only a small fraction of EEG data, they affect EEG spectral power significantly. This suggests the importance of using artefact-rejection algorithms before analysing EEG data.

INTRODUCTION

Muscle and ocular artefacts corrupt electroencephalography (EEG) signals recorded during sleep and alter their spectral power, potentially affecting the interpretation of such signals and their linkage to psychological disorders (Buysse *et al.*, 2001; Cohen *et al.*, 2013; Woodward *et al.*, 2000). Hence, to process increasingly large streams of EEG recordings, some research groups use well-established, automated artefact-rejection algorithms (Brunner *et al.*, 1996; Cohen *et al.*, 2013; Doman *et al.*, 1995), which provide for objective and consistent screening. However, the use of such algorithms has not been adopted universally, because the extent to which muscle and ocular artefacts affect spectral power values has not been quantified explicitly for whole-night rapid eye movement (REM) and non-REM (NREM) sleep. In addition, the extent to which these algorithms help match spectral power to their artefact-free levels has not been assessed fully. Providing a quantitative characterization of the effects of artefacts on spectral power and the effectiveness of artefact-rejection algorithms

will help to increase awareness for the need to properly post-process EEG signals before they are analysed.

The objective of this report is twofold: (1) to quantify the effects of muscle and ocular artefacts on the average EEG spectral power across different sleep states and EEG channels and (2) to assess the effectiveness of two previously developed automated algorithms—one for muscle artefacts (Brunner *et al*., 1996) and the other adapted for rejecting potential ocular artefacts (Cohen *et al*., 2013; Doman *et al*., 1995)—in minimizing the differences in the average spectral power of whole-night EEG recordings with respect to their artefact-free levels (annotated by visual detection and rejection).

METHODS

We analysed visually curated polysomnography (PSG) recordings of 20 subjects during 1 night of in-laboratory sleep from a previously conducted study (Cohen *et al*., 2013) at the University of Pittsburgh School of Medicine. The PSG recording montage consisted of bilateral frontal (F3 and F4), central (C3 and C4), temporal (T3 and T4), parietal (P3 and P4) and occipital (O1 and O2) EEG channels; right and left electro-oculogram (EOG) channels; and a bipolar submentalis electromyogram channel. We referenced the EEG and EOG channels to linked mastoids, filtered the channels to include frequencies from 0.3 to 100 Hz and sampled the data at 256 Hz. Both the University of Pittsburgh Institutional Review Board and the US Army Medical Research and Materiel Command Human Research Protection Office approved re-analyses of the data.

We analysed 74 NREM–REM cycles in total, consisting of 122 recording hours, scored previously in 20-s epoch segments according to the criteria of Rechtschaffen and Kales (1968). We partitioned each 20-s epoch into five non-overlapping 4-s epochs and used the 4-s epoch (channel by channel) as the basic unit for all subsequent analyses.

We inspected each EEG channel visually and annotated each 4-s epoch (one channel at a time) for artefacts, including muscle (possibly co-occurring with body/head movement), ocular, cardiac, sweat, respiration and technical (such as electrode pop, lead movement and environmental interference) artefacts (Anderer *et al*., 1999). One rater scored the EEG signals visually for artefacts and a second rater reviewed the scored artefacts. Any disagreements were resolved by consensus between the two raters. We used 4-s epochs without any artefact in each EEG channel to compute the (clean) spectral power P_C, which we used as the 'ground truth' for comparison.

To minimize the impact of muscle artefacts, we rejected transient high-frequency activities at 4-s epochs from each EEG channel (one EEG channel at a time, whenever the power between 26.25 Hz and 32.00 Hz of the 4-s epoch exceeded the moving median in a 3-min window centred around the 4-s epoch by a factor of 4), using a previously validated algorithm (Brunner *et al*., 1996). To minimize the impact of ocular artefacts during REM sleep, we first applied a previously developed algorithm for identifying REM events by detecting sharp opposite-phase deflections in the two EOG channels (Doman *et al*., 1995). We then rejected the 4-s epochs equally from all EEG channels during which at least one REM event was detected (because a REM event in EOG channels indicates a potentially contaminated 4-s epoch in all EEG channels). We note that Cohen *et al*. (2013) adopted the same procedure to reject potential ocular artefacts.

First, to assess the effects of muscle and ocular artefacts on EEG spectral power, for each NREM–REM sleep cycle (from each subject) and each sleep state (REM and NREM), we computed the power from the power spectral density estimated using standard periodograms (Campbell, 2009) by averaging the power spectra over all 4-s epochs devoid of other contaminations but muscle artefacts (P_{C+M}) and all 4-s epochs devoid of other contaminations but ocular artefacts (P_{C+O}). We used visual inspection to label the 4-s epochs. Subsequently, we compared them with the clean spectral power P_C by computing the percentage error $(P - P_C)/P_C \times 100\%$, where P denotes either P_{C+M} or P_{C+O}. Secondly, to assess the effectiveness of the previously developed algorithms in minimizing the effects of muscle and ocular artefacts for each NREM–REM sleep cycle (from each subject) and each sleep state (REM and NREM), we computed the percentage errors between the spectral power following exclusion of all 4-s epochs flagged by the automated algorithms (P_{BD}) and the clean spectral power P_C. We performed the above analysis for each of the six frequency bands [delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), sigma (12–16 Hz), beta (16–32 Hz) and gamma (32–50 Hz)] and 10 EEG channels. We reported the average percentage errors across the NREM–REM cycles obtained from the 20 subjects.

RESULTS

Effects of muscle and ocular artefacts on EEG spectral power

During REM sleep, based on visual inspection, muscle artefacts contaminated fewer than 5% of the 4-s epochs; they equally contaminated each EEG channel (Supporting information, Table S1). As expected, muscle artefacts resulted in disproportionally large increases in delta (49–126%), beta (66–171%) and gamma power (362–938%) (Fig. 1a, left panel; Supporting information, Table S1) and moderate increases in sigma (23–47%) and alpha power (10–19%). The spectral power of temporal EEG channels appeared to be more susceptible to muscle artefacts than that of other EEG channels (e.g., on average, the gamma power and beta power of temporal EEG channels increased by 870 and 154%, respectively, whereas the corresponding power of other channels increased by 459 and 86%). Ocular artefacts, in contrast, contaminated mainly the frontal (16% of the 4-s epochs), temporal (12%) and central EEG channels (12%),

and affected primarily delta power (25–33%) derived from temporal and frontal channels (Fig. 1a, right panel; Supporting information, Table S1). During NREM sleep, muscle artefacts were the main source of contamination, affecting approximately 3% of the 4-s epochs and causing increases of 32–71% in beta power and 118–262% in gamma power across different EEG channels.

Effectiveness of automated artefact-rejection algorithms

During REM sleep, applying both algorithms together effectively matched the spectral power to within ~10% of the P_C values for each EEG channel and frequency band (Fig. 1b, right panel; and Supporting information, Table S2). Applying both algorithms rejected between 25.5 and 26.4% of the 4-s epochs, most of which were rejected by Doman's algorithm. Fewer than 1% of the 4-s epochs were rejected by both algorithms (Supporting information, Table S3). Importantly, neither algorithm alone could match the spectral power to within 10% of the P_C values. Applying Brunner's algorithm alone left the delta power of frontal and temporal EEG channels largely uncorrected (27–38% error relative to P_C), whereas applying Doman's algorithm alone failed to match the beta power (35–88%) and gamma power (177–543%) across all EEG channels. During NREM sleep, Brunner's algorithm rejected fewer 4-s epochs (< 2%) than did visual inspection. Despite this lower rejection percentage, Brunner's algorithm effectively matched the NREM spectral power to within ~10% of the P_C values for each EEG channel and frequency band.

DISCUSSION

Although muscle and ocular artefacts affect only a small fraction of EEG data (3–16%), they inflate the spectral power of EEG signals significantly. The power spectra of

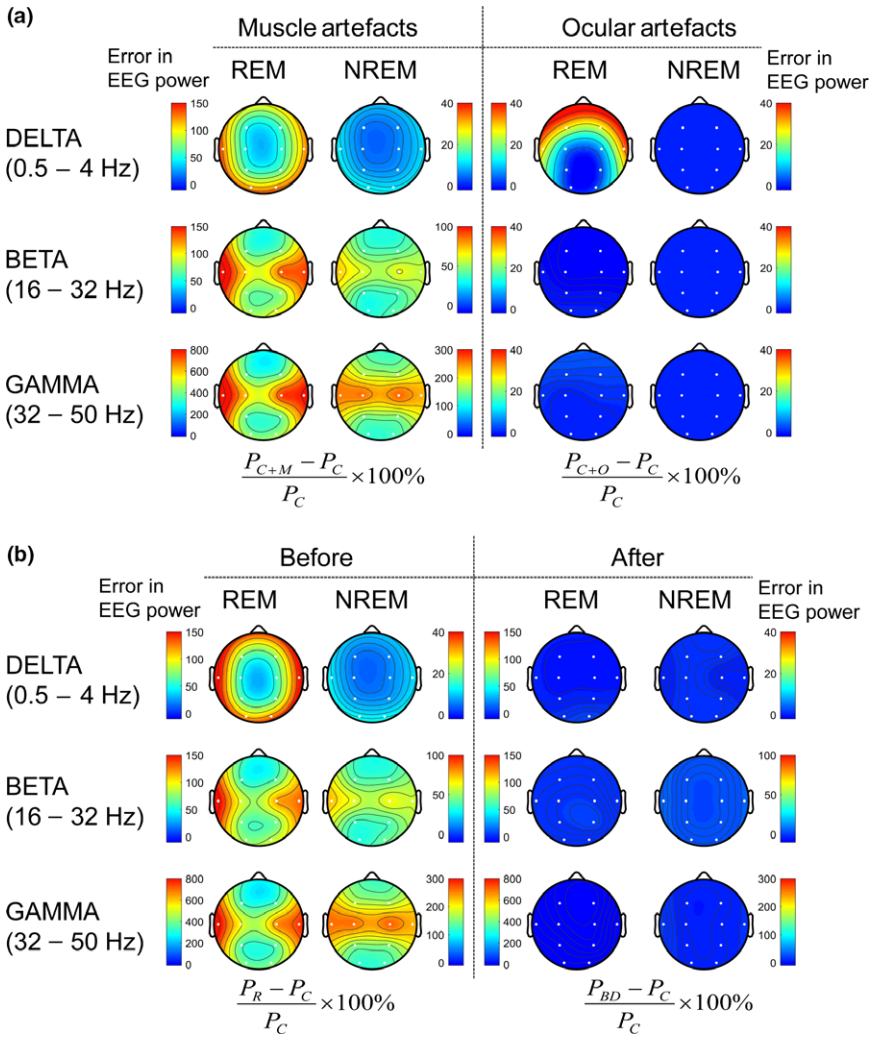

Figure 1. Topographical patterns of the effects of muscle and ocular artefacts on electroencephalography (EEG) spectral power (a), and the effectiveness of automated artefact-rejection algorithms (b) during rapid eye movement (REM) and non-REM (NREM) sleep. P_{BD}: power after automated artefact rejection; P_C: clean power; P_{C+M}: power contaminated by muscle artefact; P_{C+O}: power contaminated by ocular artefact; P_R: raw power.

delta, beta and gamma frequency bands were particularly susceptible to muscle and ocular artefacts, especially during REM sleep, with increases of up to 126, 171 and 938%, respectively. In contrast, in terms of percentage difference, typical group effects between subjects with and without post-traumatic stress disorder ranged from 20 to 35% for features derived from EEG spectral analysis (Cohen *et al*., 2013; Cowdin *et al*., 2014; Mellman *et al*., 2007; Richards *et al*., 2013). This strongly suggests the need to remove these artefacts before interpreting EEG power spectral data.

The artefact-rejection algorithms of Brunner and Doman effectively minimized the effects of muscle and ocular artefacts, respectively. Notably, applying these algorithms to our data set reduced the errors in EEG spectral power to within ~10% for each EEG channel and frequency band, without targeting other types of artefact specifically. This could be considered as adequate when compared with the natural variability in EEG spectral power across different subjects and sleep cycles [on average, the coefficient of variation was 48% (range: 27–94%) for REM sleep and 60% (range: 37–95%) for NREM sleep]. Further examination of the intrasubject variability in EEG spectral power between 2 nights revealed that the median absolute percentage difference in the spectral power was ~20%, which was well above the percentage errors we achieved after applying the automated algorithms. Therefore, we strongly recommend screening PSG data, prior to spectral analysis, for both muscle and ocular artefacts with these or similarly effective automated artefact-rejection routines. This is especially important when investigating potential biomarkers of risk or resilience for physical or psychiatric conditions, or reliable predictors of treatment response or treatment resistance to guide efforts in precision medicine. We acknowledge that the prevalence of different artefacts may vary across studies. If those other than muscle and ocular artefacts have considerable impact on the spectral power because their prevalence is high, they should also be considered in performing spectral analysis. The proposed methodology could be adapted for this purpose, because it is equally applicable to quantifying the effectiveness of other automated artefact-rejection algorithms.

Although we focused upon evaluating the effectiveness of artefact-rejection algorithms, artefact-correction techniques such as principal component analysis (PCA) and independent component analysis (ICA) (Jung *et al*., 2000; Vigario, 1997) attempt to correct contaminated signals instead of completely discarding them, thus avoiding the unnecessary rejection of brain EEG signals that may correlate with artefacts (such as phasic EEG activities during REM sleep). However, most PCA- or ICA-based techniques require manual identification of the underlying independent sources of the signal before their mixture can reproduce the original EEG signal (Vigario, 1997), and their effectiveness may be compromised when only a few EEG and EOG channels are available (which is not uncommon in many sleep biomarker studies). Moreover, an established method to evaluate the effectiveness of these artefact-correction techniques does not yet exist, given the difficulty of obtaining the underlying uncontaminated signals (Croft and Barry, 2000). Finally, we need to exercise caution when applying PCA or ICA to whole-night EEG recordings, because sleep EEGs are heterogeneous across sleep states and it is unclear whether a single PCA (ICA) mixture model would be adequate for all sleep states.

ACKNOWLEDGEMENTS

This was not an industry-supported study. This work was sponsored by US Defense Health Program (JPC-5; grant No. W81XWH-14-2-0145). The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the US Army or of the US Department of Defense. This paper has been approved for public release with unlimited distribution.

AUTHOR CONTRIBUTIONS

AG and JR conceived the research. JL, SR, SL, MN and DJC implemented the algorithms and performed the computations. AG and DJC provided data for analysis. JL, SR, SL and JR wrote the paper. All authors revised the paper.

CONFLICT OF INTEREST

The authors have indicated no financial conflicts of interest.

REFERENCES

Anderer, P., Roberts, S., Schlogl, A., *et al.* Artifact processing in computerized analysis of sleep EEG—a review. *Neuropsychobiology*, 1999, 40: 150–157.

Brunner, D. P., Vasko, R. C., Detka, C. S., Monahan, J. P., Reynolds, C. F. III and Kupfer, D. J. Muscle artifacts in the sleep EEG: automated detection and effect on all-night EEG power spectra. *J. Sleep Res*., 1996, 5: 155–164.

Buysse, D. J., Hall, M., Begley, A., *et al.* Sleep and treatment response in depression: new findings using power spectral analysis. *Psychiatry Res*., 2001, 103: 51–67.

Campbell, I. G. EEG recording and analysis for sleep research. *Curr. Protoc. Neurosci*., 2009, 49: 10.2.1–10.2.19.

Cohen, D. J., Begley, A., Alman, J. J., *et al.* Quantitative electroencephalography during rapid eye movement (REM) and non-REM sleep in combat-exposed veterans with and without post-traumatic stress disorder. *J. Sleep Res*., 2013, 22: 76–82.

Cowdin, N., Kobayashi, I. and Mellman, T. A. Theta frequency activity during rapid eye movement (REM) sleep is greater in people with resilience versus PTSD. *Exp. Brain Res*., 2014, 232: 1479–1485.

Croft, R. J. and Barry, R. J. Removal of ocular artifact from the EEG: a review. *Neurophysiol. Clin*., 2000, 30: 5–19.

Doman, J., Detka, C., Hoffman, T., *et al.* Automating the sleep laboratory: implementation and validation of digital recording and analysis. *Int. J. Biomed. Comput*., 1995, 38: 277–290.

Jung, T. P., Makeig, S., Humphries, C., *et al.* Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 2000, 37: 163–178.

Mellman, T. A., Pigeon, W. R., Nowell, P. D. and Nolan, B. Relationships between REM sleep findings and PTSD symptoms during the early aftermath of trauma. *J. Trauma. Stress*, 2007, 20: 893–901.

Rechtschaffen, A. and Kales, A. *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. US Government Printing Office, US Public Health Service, Washington, DC, 1968.

Richards, A., Metzler, T. J., Ruoff, L. M., *et al.* Sex differences in objective measures of sleep in post-traumatic stress disorder and healthy control subjects. *J. Sleep Res.*, 2013, 22: 679–687.

Vigario, R. N. Extraction of ocular artefacts from EEG using independent component analysis. *Electroencephalogr. Clin. Neurophysiol.*, 1997, 103: 395–404.

Woodward, S. H., Murburg, M. M. and Bliwise, D. L. PTSD-related hyperarousal assessed during sleep. *Physiol. Behav.*, 2000, 70: 197–203.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. Effects of muscle and ocular artefacts on electroencephalography (EEG) spectral power during rapid eye movement (REM) and non-REM (NREM) sleep.

Table S2. Effectiveness of automated artefact-rejection algorithms in matching electroencephalography (EEG) spectral power during rapid eye movement (REM) and non-REM (NREM) sleep.

Table S3. Prevalence of visually scored artefacts and fraction of 4-s epochs rejected by automated algorithms during rapid eye movement (REM) and non-REM (NREM) sleep.