Visual and automatic classification of the cyclic alternating pattern in electroencephalography during sleep

R. Largo, M.C. Lopes, K. Spruyt, C. Guilleminault, Y.P. Wang and A.C. Rosa

1LaSEEB – Evolutionary Systems and Biomedical Engineering Laboratory, Institute for Systems and Robotics (ISR-Lisboa), Instituto Superior Técnico (IST), University of Lisbon, Lisbon, Portugal
2Instituto de Psiquiatria (PRATA), Hospital das Clínicas (HCFMUSP), Faculdade de Medicina, Universidade de São Paulo, São Paulo, SP, Brasil
3Lyon Neuroscience Research Center, INSERM U1028-CNRS UMR 5292 Waking Team, School of Medicine, University Claude Bernard, Lyon, France
4Sleep Disorders Clinic, Stanford University Medical Center, Stanford, CA, USA
5Instituto de Psiquiatria (LIM-23), Hospital das Clínicas (HCFMUSP), Faculdade de Medicina, Universidade de São Paulo, São Paulo, SP, Brasil
6Escola Superior de Tecnologia de Setúbal, Instituto Politécnico de Setúbal, Setúbal, Portugal

Abstract

Cyclic alternating pattern (CAP) is a neurophysiological pattern that can be visually scored by international criteria. The aim of this study was to verify the feasibility of visual CAP scoring using only one channel of sleep electroencephalogram (EEG) to evaluate the inter-scorer agreement in a variety of recordings, and to compare agreement between visual scoring and automatic scoring systems. Sixteen hours of single-channel European data format recordings from four different sleep laboratories with either C4-A1 or C3-A2 channels and with different sampling frequencies were used in this study. Seven independent scorers applied visual scoring according to international criteria. Two automatic blind scorings were also evaluated. Event-based inter-scorer agreement analysis was performed. The pairwise inter-scorer agreement (PWISA) was between 55.5 and 84.3%. The average PWISA was above 60% for all scorers and the global average was 69.9%. Automatic scoring systems showed similar results to those of visual scoring. The study showed that CAP could be scored using only one EEG channel. Therefore, CAP scoring might also be integrated in sleep scoring features and automatic scoring systems having similar performances to visual sleep scoring systems.

Key words: Cyclic alternating pattern; Sleep; Visual scoring; Automatic scoring

Introduction

Cyclic alternating pattern (CAP) in non-rapid eye movement (NREM) sleep is a neurophysiological pattern that can be visually scored by international criteria, and has been described in several conditions over a period of 20 years (1). Visual CAP scoring is a new way to measure sleep instability, and it could be a feasible tool to detect changes in brain plasticity. Some authors have been describing the essential roles of spontaneous brain activity in basic brain functioning and in the understanding of the working mechanisms of the brain (2,3). In fact, the sleep slow waves are homeostatically regulated and linked to learning and plasticity processes (4). The CAP pattern may therefore be a feasible marker of sleep instability showing the mechanisms involved, which allows a better understanding of neural plasticity during sleep (5). Moreover, this scoring system allows recognition of sleep arousals and disturbances that are not scored with the international criteria for scoring the sleep macrostructure.

CAP analysis has been performed by detection of electrocortical events with regular intervals in a range of seconds during NREM sleep (6). These events may be clearly distinguishable from the background electroencephalogram (EEG) rhythm as abrupt frequency shifts or amplitude changes. Two phases (A and B) are present as part of a CAP cycle and recur within 2 to 60 s (7). When neither of the phases (A and B) is identifiable, sleep has reached a new stable non-CAP (NCAP) state. The identification of CAP during sleep allows a different approach to
investigate NREM sleep in many sleep disorders. It allows the recognition of disturbances of NREM sleep that are not visually identified using either the international manual of Rechtschaffen and Kales (8) or the American Sleep Disorders Association scoring criteria (9) for the identification of short (3 s or longer) EEG arousals (10). CAP has also been useful in the investigation of normal sleep (11,12), sleep changes by acoustic perturbations (13), cardiorespiratory perturbations (14–16), and neurological conditions (17,18). It advances our understanding of the physiologic components of sleep (19–23) and normal changes occurring during nocturnal sleep (24–26).

One limitation for the application of CAP analysis in sleep medicine has been the extra time needed to apply such a scoring system in clinical practice. Namely, the classification of short duration events in sleep EEG like the A phases of CAP is a laborious and error-prone task. Nonetheless, the clinical meaning has been linked to the detection of sleep instability in mild sleep disorders (5). The effects of age, patient group, and recording characteristics on the scoring agreement, however, are not completely known (27–31). For instance, studies investigating specific medical syndromes call upon age-matched controls run in the same laboratory. Therefore, there is great interest in developing an automatic scoring system that could be useful in clinical practice.

Another limitation is the concern about the feasibility, accuracy, and reliability of CAP classification independent of sleep stages using a single central EEG lead (C4-A1 or C3-A2). According to Rosa et al. (32), CAP scoring can be done with only one channel of EEG, and the evaluation of the inter-scorer agreement in a variety of recordings was useful in comparison with an automatic scoring system.

The main objective of this work was to assess the viability and performance of visual and automatic scoring on a single EEG channel, as suggested by the published CAP scoring rules in Terzano et al. (1), using recordings from different laboratories of subjects with different age, gender, health condition, and sampling frequency.

Material and Methods

Sixteen hours of single-channel European data format (EDF) in eight polysomnographic recordings from four different sleep laboratories with either C4-A1 or C3-A2 channel and with different sampling frequencies (100, 128, and 256 Hz) were used. In order to speed up the visual classification process, 2 h of NREM sleep were extracted from each recording, 1 h epochs of NREM from the first half and the other 1 h epochs from the second half of the night, totaling 16 files.

Methods

Seven scorers (G, N, P, Q, X, Z, and Y) performed visual scoring according to CAP scoring criteria described in Terzano et al. (1), and two automatic scoring systems (V and A) were used in the study. All scorers were blind to any information on each selected segment. Five visual and one automatic CAP scoring was performed using Somnologica Science 3 software version 3.3.2 (Flaga Hf, Iceland). The automatic scoring “A” is described in Rosa et al. (33) and “V” is fully described in Largo et al. (34). All the phase A scorings were converted into International Standards Organization (ISO) time format and converted to the same sampling frequency (when appropriate) for agreement analysis (35).

Analysis

Two types of analysis were performed: the first was the time-based agreement of the scorings, where the scorings were compared for every sampling point and the second was event-based, where an agreement was scored whenever 2 events have an intersection longer than 0.5 s. In this report, only the event-based analysis is presented.

Figure 1 shows the different possibilities for the event-based analysis: when 2 events intersect for more than a specific period of time, this is considered a “hit” (T: True), otherwise events are marked either as a “miss” (Fn: false negative) or as a “false” (Fp: false positive). As an example, in Figure 1, if we take the top trace (visual) as the reference, we have one T, one Fn, and 2 Fp.

Following the above example, a specific agreement analysis can be proposed as shown in Table 1. The “mutual agreement” (MA) is the measure of the agreement between any two scorers without taking either of them as the reference (36). We can only calculate the sensitivity of scorer 2 in relation to reference scorer 1 (SS) and the positive predictive value (PP). The number of disagreements (false negatives and false positives) per minute can be an alternative index.

Results

The total number of events scored by the 2 automatic scoring systems and the 7 visual scorers are presented in Table 2. Their values were between 939 and 2036. All pairwise scorers (automatic or visual) were compared file-by-file and the average results of MA are reported in

![Figure 1. Event-based scoring agreement analysis. Whenever two events (A) intersect for more than a specific period, it is considered a hit (T). Other events are marked either as a miss (Fn: False negative) or a false (Fp: False positive).](image-url)
Table 2. Inter-scorers mutual agreement.

| Scorer | Average MA% | # Events |
|--------|--------------|----------|
| a      | 71.1         | 1882     |
| g      | 73.9         | 1864     |
| n      | 59.7         | 1693     |
| p      | 73.1         | 1689     |
| q      | 80.8         | 1430     |
| x      | 65.3         | 2036     |
| z      | 72.2         | 1625     |
| y      | 66.9         | 1126     |

Scorer by scorer mutual agreement (MA%) table is symmetrical; therefore, only the upper matrix without diagonal is presented. The last two columns are average MA% and total number of events for each scorer.

Table 2. The table is symmetrical, therefore only the upper triangular matrix is presented. The inter-scorer MA values were between 55.5% (AP) and 84.3% (VA). The average MA for each scorer compared with all the others is presented in Table 2, and shows a minimum of 60.6% and a maximum of 73.5%. The global average was 69.9% with standard deviation of 6.9%.

We calculated the MA between laboratories (Figure 2). The sensitivity for all pairwise file-by-file comparisons showed high density values around 80%. The limit curve of this graph suggested a "Pareto front" for the scoring process, indicating also that the best averaged MA sensitivity was 90%. The sensitivity was compared to PP value for all pairwise file-by-file comparisons (Figure 3). Figure 4 shows the MA scatter plot versus PP value for all pairwise file-by-file comparisons and is very similar to Figure 3. This was expected because the comparison was not made with a reference but between two scorers. Figure 5 shows a scatter plot of sensitivity versus PP value for all pairwise file-by-file comparisons. The PP value for all pairwise file-by-file comparisons was generated. This was expected because the comparison was not made with a reference but between two scorers. The inter-scorer MA for all files was above 80%; as an example, the comparison between two of the most experienced scorers was 87% MA.

All file averages are reported in Table 3. The upper triangle shows the sensitivity values that were between 43.3 and 91.4%. The lower triangle shows the positive predictive values that were between 42.6 and 94.2%. Average values for each scorer compared with all the others are reported in the last column and last line. The global average and the standard deviation were 78.1 and 11.5%, respectively, for sensitivity, and 70.0 and 12.7%, respectively, for PP values.

![Figure 2: Inter-scorer MA between laboratories](image)

**Discussion**

This is the first study to show that CAP can be scored using only a single EEG channel. Although the CAP scoring experience of different scorers varied considerably.

---

**Table 1. Truth table for events scoring.**

| Phase A (1=Ref) | T Fp Positive Predictive value=\(T/(T+Fp)\) | \((Fp+Fn)/2\) | Mutual Agreement MA=\(T/(T+(Fn+Fp)/2)\) |
|----------------|---------------------------------------------|---------------|------------------------------------------|
| Phase A (2)    | T                                           | Fp            |                           |
| Phase B (2)    | Fn                                          |               |                           |
| Sensitivity=\(T/(T+Fn)\) |                           |               |                           |

Ref: reference; 1: Scorer 1; 2: Scorer 2; T: true positive; Fp: false positive; Fn: false negative; MA: mutual agreement.
(from newcomer to very experienced), the results were in ranges seen for inter-scorer agreement in similar studies (32). An interesting finding in our study was that the automatic scoring system (scorer V) had the highest overall MA (73.4%) and sensitivity (84.9%). The examples reported in Figures 7 and 8 show some of the difficulties

Figure 2. Scatter plot of all mutual agreement (%) values in different files.

Figure 3. Mutual agreement scatter plot versus sensitivity for all pairwise file-by-file comparisons. This graph shows high density values in 80% neighborhood.
encountered in visual scoring. There were inconsistencies in the CAP patterns selected by any given subject when such visual difficulty occurred, but the automatic system was consistent in its choices, based on amplitude and frequency criteria.

Our results demonstrated that the automatic analysis of CAP has a clinical application. Because the number of events to be visually scored was usually large (several hundred up to thousands), it was necessary to mark the events boundary accurately, otherwise the classification of

**Figure 4.** Mutual agreement scatter plot versus positive predictive value for all pairwise file-by-file comparisons.

**Figure 5.** Sensitivity scatter plot versus positive predictive value for all pairwise file-by-file comparisons.
the events into different A subtypes and the precision of
the phase durations, cycles, and alternating conditions
could be compromised. There is already some published
data on visual CAP scoring agreement with automatic
scoring systems (37,38) for events such as arousals (39)
or CAP A phases. However, besides the overall values of
correctness, specificity, and sensitivity, these studies
only report detailed information on the detection method.
The analysis of the performance of the automatic scoring
system was done on a limited number of tracings and
in a limited part of the night. The reference scoring,
furthermore, has been described by one or two experts.
The inter-scorer reliability in scoring CAP parameters of
normal sleep has only been evaluated in two papers
(32,40) and was based on only a few scorers, while intra-
lab and automatic scoring is evaluated in Rosa et al. (32).
In Ferri et al. (40), 4 human scorers participated in the
study and reported values of the Kendall W coefficient of

Table 3. Inter-scorers average sensitivity (SS%) and positive predictive values (PP%).

| SS% | PP% | v  | a  | g  | N  | p  | q  | x  | z  | y  | Average SS% |
|-----|-----|----|----|----|----|----|----|----|----|----|-------------|
| v   | .   | 84.9| 80.8| 81.2| 90.7| 85.6| 80.9| 85.9| 87.7| 84.7|             |
| a   | 83.9| .   | 78.2| 77.5| 88.4| 82.2| 76.4| 83.9| 85.6| 82.1|             |
| g   | 68.2| 65.0| .   | 77.6| 83.9| 78.4| 67.1| 75.7| 82.4| 78.0|             |
| n   | 70.6| 65.8| 79.8| .   | 88.9| 84.2| 71.4| 75.9| 87.4| 80.5|             |
| p   | 47.0| 42.6| 54.4| 49.1| .   | 54.4| 43.3| 48.0| 58.8| 69.6|             |
| q   | 66.4| 62.2| 72.2| 73.5| 86.4| .   | 66.2| 72.4| 83.8| 75.9|             |
| x   | 82.1| 77.5| 82.9| 84.1| 94.2| 88.8| .   | 84.7| 91.4| 72.7|             |
| z   | 73.7| 71.2| 78.1| 75.4| 84.5| 80.7| 69.4| .   | 86.1| 76.6|             |
| y   | 56.5| 52.1| 61.1| 62.5| 77.1| 67.3| 53.5| 60.1| .   | 82.9|             |
| Average PP% | 68.6 | 65.0 | 70.2 | 70.1 | 66.9 | 74.7 | 79.1 | 74.1 | 61.3 |             |

All files average sensitivity (upper triangle) and positive predictive values (lower triangle). Last column and last line represent inter-
scorers average sensitivity and positive predictive values.

Figure 6. Box plot with all files inter-scorer mutual agreement with scorer Z. The boxes are between 1st and 3rd quartile. The
interquartile range is between 20 and 50%. The line inside is the median. Scorers: vz, az, gz, nz, pz, qz, xz, and yz.

Figure 7. Sleep stage N3, one-minute epoch scorings. We found agreement between scorers.
concordance above 0.9 for total CAP time, CAP time in sleep stage 2, and percentage of A phases in sequence; the CAP rate also showed a high value.

Moreover, ambulatory monitoring is increasingly performed in clinical practice and often one EEG lead is the only available EEG information. Our results showed that just one EEG lead was well correlated between human and automatic scoring. Development of automatic scoring will allow greater usage of CAP scoring in clinical practice, particularly when using ambulatory monitoring with few recording leads. Considering the potential large clinical usage of such automatic scoring systems and our encouraging results, there is no doubt that CAP scoring could be easily integrated into conventional sleep scoring. Our findings further suggest that usage of a single EEG channel for automatic scoring could be more refined by using visual analysis of several EEG leads.

Some limitations of the CAP analyses, however, have to be addressed, particularly the clinical meaning of the CAP rate. For now, it has only been proposed as a NREM sleep tool to evaluate sleep instability. Probably due to the amplitude criteria, it is not applied to REM sleep. However, the automatic scoring of CAP following the default decisions made by specialists in CAP analyses may assist the users in improving their recognition of NREM sleep. In conclusion, once several NREM sleep disturbances can be assessed by CAP analyses, the automatic scoring system or CAP might be applied for research purposes to better understand sleep phenomena.

Acknowledgments

We are very grateful to Gabriela Alves, Magneide Brito, Rafaële Ferri, Mario Terzano, Robert Thomas, and Marco Zucconi for their important contribution to visual CAP scoring in the sleep EEG recordings. We thank Medcare-Somnologica for providing the investigators with the Science 3.3 program and its automatic analysis software and the Stanford University meeting for establishment of default criteria for the usage of a single EEG channel to CAP analyses. The review of visual scoring between experts and naïve scorers was supported by an educational grant from Sanofi-Aventis International. This work was supported by LARSyS (PEst-OE/EEI/LA0009/2013).

References

1. Terzano MG, Parrino L, Smerieri A, Chervin R, Chokroverty S, Guilleminault C, et al. Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep. Sleep Med 2002; 3: 187–199, doi: 10.1016/S1389-9457(02)00003-5.
2. Liu X, Zhu XH, Zhang Y, Chen W. Neural origin of spontaneous hemodynamic fluctuations in rats under burst-suppression anesthesia condition. Cereb Cortex 2011; 21: 374–384, doi: 10.1093/cercor/bhq105.
3. Esser SK, Hill SL, Tononi G. Sleep homeostasis and cortical synchronisation: I. Modeling the effects of synaptic strength on sleep slow waves. Sleep 2007; 30: 1617–1630, doi: 10.1093/sleep/30.12.1617.
4. Massimini M, Tononi G, Huber R. Slow waves, synaptic plasticity and information processing: insights from transcranial magnetic stimulation and high-density EEG experiments. Eur J Neurosci 2009; 29: 1761–1770, doi: 10.1111/j.1460-9568.2009.06720.x.
5. Parrino L, Ferri R, Bruni O, Terzano MG. Cyclic alternating pattern (CAP): the marker of sleep instability. *Sleep Med Rev* 2012; 16: 27–45, doi: 10.1016/j.smrv.2011.02.003.

6. Terzano MG, Parrino L, Rosa A, Palomba V, Smerieri A. CAP and arousals in the structural development of sleep: an integrative perspective. *Sleep Med* 2002; 3: 221–229, doi: 10.1016/S1389-9457(02)00009-6.

7. Ferri R, Bruni O, Miano S, Piazzì G, Spruyt K, Gozal D, et al. The time structure of the cyclic alternating pattern during sleep. *Sleep* 2006; 29: 693–699, doi: 10.1093/sleep/29.5.693.

8. Rechtschaffen A, Kales A. Manual of standardized terminology: techniques and scoring system for sleep stages of human subjects. Los Angeles: UCLA Brain Information Service/Brain Research Institute; 1968.

9. American Sleep Disorders Association. EEG arousals: scoring rules and examples. A preliminary report. *Sleep* 1992; 15: 173–184, doi: 10.1093/sleep/15.2.173.

10. Smirne S, Ferini-Strambi L. Clinical applications of cyclic alternating pattern. *Electroencephalogr Clin Neurophysiol Suppl* 1999; 50: 109–112.

11. Terzano MG, Mancia D, Salati MR, Costani G, Decembrino A, Parrino L. The cyclic alternating pattern as a physiologic component of normal NREM sleep. *Sleep* 1985; 8: 137–145, doi: 10.1093/sleep/8.2.137.

12. Terzano MG, Parrino L, Fioriti G, Spaggiari MC, Pirolli A. Morphologic and functional features of cyclic alternating pattern (CAP) sequences in normal NREM sleep. *Func Neuro* 1986; 1: 29–41.

13. Terzano MG, Parrino L, Fioriti G, Orofiamma B, Depoortere H. Modifications of sleep structure induced by increasing levels of acoustic perturbation in normal subjects. *Electroencephalogr Clin Neurophysiol* 1990; 76: 29–38, doi: 10.1016/0013-4694(90)90055-O.

14. Ferini-Strambi L, Bianchi A, Zucconi M, Oldani A, Castro-novo V, Smirne S. The impact of cyclic alternating pattern on heart rate variability during normal sleep. *Sleep* 2001; 24: 881–885, doi: 10.1093/sleep/24.8.881.

15. Terzano MG, Parrino L, Smerieri A, Carli F, Nobili L, Donadio S, et al. CAP and arousals are involved in the homeostatic and ultradian sleep processes. *J Sleep Res* 2005; 14: 359–368, doi: 10.1111/j.1365-2869.2005.00479.x.

16. Terzano MG, Parrino L, Boselli M, Smerieri A, Spaggiari MC. CAP components and EEG synchronization in the first 3 sleep cycles. *Clin Neurophysiol* 2000; 111: 283–290, doi: 10.1016/S1388-2457(99)00245-X.

17. Ferri R, Guilleminault C, Lopes MC. Cyclic alternating pattern as a sign of brain instability during sleep. *Neurologia* 2006; 21: 304–311.

18. Ferri R, Rundo F, Bruni O, Terzano MG, Stam CJ. Regional scalp EEG slow-wave synchronization during sleep cyclic alternating pattern A1 subtypes. *Neurosci Lett* 2006; 404: 352–357, doi: 10.1016/j.neulet.2006.06.008.

19. Smerieri A, Parrino L, Agosti M, Ferri R, Terzano MG. Cyclic alternating pattern sequences and non-cyclic alternating pattern periods in human sleep. *Clin Neurophysiol* 2007; 118: 2305–2313, doi: 10.1016/j.clinph.2007.07.001.

20. Parrino L, Boselli M, Spaggiari MC, Smerieri A, Terzano MG. Cyclic alternating pattern (CAP) in normal sleep: polysomnographic parameters in different age groups. *Electroencephalogr Clin Neurophysiol* 1998; 107: 439–450, doi: 10.1016/S0013-4694(98)00108-4.

21. Bruni O, Ferri R, Miano S, Verrillo E, Vittori E, Della Marca G, et al. Sleep cyclic alternating pattern in normal school-age children. *Clin Neurophysiol* 2002; 113: 1806–1814, doi: 10.1016/S1388-2457(02)00265-1.

22. Bruni O, Ferri R, Miano S, Verrillo E, Vittori E, Farina B, et al. Sleep cyclic alternating pattern in normal preschool-aged children. *Sleep* 2005; 28: 220–230, doi: 10.1093/sleep/28.2.220.

23. Bruni O, Ferri R, Miano S, Piazzì G, Guilleminault C, Tu S, et al. Cyclic alternating pattern in peripubertal children. *Sleep* 2005; 28: 215–219, doi: 10.1093/sleep/28.2.215.

24. Ferri R, Bruni O, Miano S, Verrillo E, Vittori E, Farina B, et al. Visual and automatic cyclic alternating pattern (CAP) scoring: inter-rater reliability study. *Arq Neuropsiquiatr* 2006; 64: 578–581, doi: 10.1590/S0004-282X2006000400008.

25. da Rosa AC, Paiva T. Automatic detection of K-complexes: validation in normals and dysthyemic patients. *Sleep* 1993; 16: 239–248.

26. Lago R, Rosa A. Wavelets based detection of a phases in sleep EEG. *Proc II Int Conf Computacional Bioengineering* 2005; 2: 1105–1115.

27. Mariani S, Manfredini E, Rosso V, Grassi A, Mendez MO, Alba A, et al. Efficient automatic classifiers for the detection of A phases of the cyclic alternating pattern in sleep. *Med
Biol Eng Comput 2012; 50: 359–372, doi: 10.1007/s11517-012-0881-0.

36. Klemens B. Mutual information as a measure of intercoder agreement. J Official Statistics 2012; 28: 395–412.

37. Navona C, Barcaro U, Bonanni E, Di MF, Maestri M, Murri L. An automatic method for the recognition and classification of the Aphasases of the cyclic alternating pattern. Clin Neurophysiol 2002; 113: 1826–1831, doi: 10.1016/S1388-2457(02)00284-5.

38. Rosa AC, Parrino L, Terzano MG. Automatic detection of cyclic alternating pattern (CAP) sequences in sleep: preliminary results. Clin Neurophysiol 1999; 110: 585–592, doi: 10.1016/S1388-2457(98)00030-3.

39. De Carli F, Nobili L, Gelcich P, Ferrillo F. A method for the automatic detection of arousals during sleep. Sleep 1999; 22: 561–572, doi: 10.1093/sleep/22.5.561.

40. Ferri R, Bruni O, Miano S, Smerieri A, Spruyt K, Terzano MG. Inter-rater reliability of sleep cyclic alternating pattern (CAP) scoring and validation of a new computer-assisted CAP scoring method. Clin Neurophysiol 2005; 116: 696–707, doi: 10.1016/j.clinph.2004.09.021.