Semiautomatic quantification of spiking in patients with continuous spikes and waves in sleep: Sensitivity to settings and correspondence to visual assessment

M.E. Peltola a,c,*, K. Palmu b,d, E. Liukkonen c, E. Gailly c, S. Vanhatalo b,e

Department of Clinical Neurophysiology, Turku University Hospital, Finland
Department of Children’s Clinical Neurophysiology, Helsinki University Central Hospital, Finland
Epilepsy Unit, Department of Pediatric Neurology, Helsinki University Central Hospital, Finland
Department of Biomedical Engineering and Computational Science, School of Science, Aalto University, Espoo, Finland
Department of Neurological Sciences, University of Helsinki, Finland

Article history:
Accepted 2 December 2011
Available online xxxx

Keywords:
CSWS
Electrical status epilepticus during sleep ESES
Spike index
Automated EEG analysis

Abstract

Objective: To define the optimal analysis protocol for semiautomatic quantification of spike index (SI) in continuous spikes and waves in sleep (CSWS).

Methods: Ten overnight EEGs (nine patients) with abundant spiking were used to quantify SI with a previously published semiautomatic quantification based on spike detection with BESA software. We studied (i) dependency of SI on maximal interspike interval (maxISI) defining the continuous discharge, (ii) sensitivity of SI to variations in the spike search protocol, and (iii) stability of SI over time. Finally, the semiautomatic method was compared with the quantification based on visual scoring by two neurophysiologists.

Results: MaxISI of 3 s appeared to yield the best combination of sensitivity and stability in SI quantification. The SI of the first hour of sleep did not differ significantly from the SI of the whole night. Mean error of the semiautomatic method compared to visual scoring was only seven percentage units.

Conclusions: Semiautomatic quantification of SI functions well with maxISI of 3 s, and the first hour of sleep represents the whole night SI with a clinically relevant accuracy.

Significance: This method opens a possibility for objective quantification of near-continuous epileptiform spiking during sleep, and it supports the use of shorter epochs for quantitative assessment of CSWS.

Highlights

• An objective paradigm for quantification of continuous spikes and waves in sleep (CSWS) is needed for both scientific and clinical use.
• Semiautomatic quantification of spike index (SI) with appropriate parameter settings is a robust and a promising tool.
• SI of the first hour of sleep is representative of the whole night SI.

1. Introduction

Continuous spikes and waves in sleep (CSWS), or electrical status epilepticus during sleep (ESES) is an EEG diagnosis, originally defined as having at least 85% of NREM sleep covered by continuous spikes and waves (Patry et al., 1971; Tassinari et al., 2000, 2009). CSWS is usually associated with epileptic encephalopathy manifesting as variable cognitive and behavioral impairments which are permanent in most patients (Roulet Perez et al., 1993; Veggiotti et al., 1999; Tassinari et al., 2000; Scholtes et al., 2005; Liukkonen et al., 2010). The quantified (percentage) amount of spiking is considered important for both assessing the impact of spiking activity on cognition (Billard et al., 1990; Beaumanoir, 1995; Guzzetta et al., 2005; Van Hirtum-Das et al., 2006; Scheltens-de Boer, 2009), and for evaluation of the success of drug treatment (Aeby et al., 2005; Inutsuka et al., 2006). In this context, it is striking that there is no unambiguous definition of how to calculate the percentage of spiking. An objective paradigm for quantification of CSWS would be necessary for both scientific and clinical use.

CSWS analysis is based on visually estimating the amount of spike and wave discharges, and it yields a dichotomic classification of CSWS (present or absent). More advanced and laborious...
approach is to calculate a spike index (SI), which indicates the percentage of NREM sleep covered by spike and waves (Morikawa et al., 1985; Galanopoulou et al., 2000; Aeby et al., 2005). The methods previously used to determine SI are, however, too diverse to enable comparability between studies (Patry et al., 1971; Tassinari et al., 2000; Aeby et al., 2005; Inutsuka et al., 2006; Galanopoulou et al., 2000; Liukkonen et al., 2010). The core problems here are the striking lack of agreement about (i) how to define a ‘continuous’ discharge, and (ii) what would be the minimum duration of sleep needed for a reliable SI calculation.

Manual quantitation of SI is too time-consuming to be used in the clinical setting, or even in most scientific research. To overcome this, a semiautomated method for calculation of SI was recently published by Larsson et al. (2009, 2010). This method is based on a spike search using spatio-temporal pattern match in the commercial BESA Research software (Scherg et al., 2002; Bast et al., 2004), followed by a further analysis of the spike detection logs with a purpose-built MATLAB script (hereafter called PS). Another recent study has used the semiautomated spike detection of ESES in Landau–Kleffner syndrome with a different focus and algorithm (Martín Miguel et al., 2011).

Table 1

Electroclinical details of the patients.

<table>
<thead>
<tr>
<th>EEG</th>
<th>Pat.</th>
<th>Age at epilepsy</th>
<th>Age at CSWS</th>
<th>Age at EEG studied</th>
<th>Etiology</th>
<th>MRI</th>
<th>Seizure type at EEG studied</th>
<th>EEG spike focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2.5</td>
<td>6.0</td>
<td>6.0</td>
<td>Unknown</td>
<td>Normal</td>
<td>Psychomotor → right clonic</td>
<td>FL, PR</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4.2</td>
<td>9.2</td>
<td>9.7</td>
<td>Unknown</td>
<td>Normal</td>
<td>Psychomotor</td>
<td>FCR</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4.2</td>
<td>6.1</td>
<td>7.3</td>
<td>Symptomatic perinatal vascular</td>
<td>MCA infarction</td>
<td>Atypical absence</td>
<td>FL, TPL, FR</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2.8</td>
<td>5.7</td>
<td>7.6</td>
<td>Symptomatic perinatal vascular</td>
<td>PVL, HC</td>
<td>Atypical absence</td>
<td>FL, PL, FR</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5.6</td>
<td>5.6</td>
<td>5.6</td>
<td>Idiopathic (familial)</td>
<td>Normal</td>
<td>Atypical absence</td>
<td>FL, PL</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>4.0</td>
<td>6.4</td>
<td>6.4</td>
<td>Idiopathic (familial)</td>
<td>Normal</td>
<td>Facial clonic</td>
<td>CI, TL, CR</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>2.5</td>
<td>5.4</td>
<td>5.9</td>
<td>Symptomatic perinatal vascular</td>
<td>PVL, right thalamic lesion</td>
<td>Facial clonic</td>
<td>FL, FR, TR, PR</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>3.8</td>
<td>6.2</td>
<td>8.0</td>
<td>Symptomatic perinatal vascular</td>
<td>PVL, HC</td>
<td>Psychomotor → right clonic</td>
<td>FL</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>2.3</td>
<td>2.8</td>
<td>3.1</td>
<td>Idiopathic</td>
<td>Normal</td>
<td>Absence</td>
<td>FCR</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>4.2</td>
<td>9.2</td>
<td>12.2</td>
<td>Unknown</td>
<td>Normal</td>
<td>Facial clonic</td>
<td>Left arm clonic</td>
</tr>
</tbody>
</table>

* Eye blinking and short arrest (absence unverified), F = frontal, T = temporal, C = central, P = parietal, TP = tempo-parietal, FC = fronto-central, L = left, R = right, MCA = middle cerebral artery, PVL = periventricular leucomalacia, HC = hydrocephalus. EEG focus with most abundant spiking in CSWS is marked in bold.
2. Methods

The flowchart of study design is presented in Fig. 1.

2.1. EEG data

Ten overnight EEGs (nine subjects) recorded between March 2006 and June 2010 due to suspected or previously diagnosed CSWS were randomly selected from the database of the Epilepsy Unit of the Helsinki University Central Hospital. Electroclinical details of the patients are shown in Table 1. The recordings were made as a part of clinical epilepsy evaluation with a Telefactor video-EEG equipment by using 26–35 scalp electrodes placed according to the international 10-20/10-10 system. Data was recorded at 200 Hz.

This study was approved by the Ethics Committee of the Hospital for Children and Adolescents, Helsinki University Central Hospital.

2.2. Standard spike search

Spikes were searched from the whole EEG recording with the semiautomated in-built spatio-temporal search function of BESA Research® software (v.5.3, MEGIS GmbH, Gräfelfing, Germany) (Scherg et al. 2002; Bast et al. 2004). EEG was filtered with a combination of 1.6 Hz (6 dB/oct.; forward) low cutoff and 35 Hz (24 dB/oct.; zerophase) high cutoff filters. Spike search was created in two steps by using an initial and a final template. The initial template was created by selecting one representative prototype spike manually. Visual inspections were done from traditional and source montages. Source montage transforms EEG into signals of regional source activity, and functions like a spatial filter to aid in localization. Then, an automated spike search was performed with a correlation threshold of 75%, using the whole length of the prototype spike waveform from one EEG channel as initial template. An average spike waveform (n ~ 100) was created from a search from the first hour of sleep. This average spike served as the final template for further search of spikes over all channels from the whole night recording with a correlation threshold of 50%. To assess the accuracy of spike search, the results of each two-step run of automated search were inspected visually from five randomly selected pages during the first 2 min epochs. The above search procedure was visualized from five randomly selected pages during the first two-step run of automated search were inspected visually from five randomly selected pages during the first 2 min epochs.

2.3. Quantification of spike index (SI)

The SI was calculated using the PS script (v.2, courtesy of Dr. Larsson) (Larsson et al. 2009, 2010). The SI was analyzed from consequent 2 min epochs (Larsson et al. 2009), and it describes the proportion of time with continuous discharges in the given epoch. A critical parameter here is the maximal interspike interval (maxISI), which defines how long pauses between successive spikes are allowed in what is considered a continuous discharge (Supplementary Fig. S1). Prior studies by Larsson et al. (2009, 2010) have only reported results with maxISI of 3 s, but we decided to assess SI results with seven maxISI values ranging from 1 to 7 s (maxISI1–maxISI7, respectively). The output of SI in 2 min epochs was imported to Excel® for further calculations. The SI was calculated both from the first hour of sleep and the whole-night sleep. Following the original papers (Larsson et al. 2009, 2010), the whole-night SI was calculated by including only SIs that were above the mean SI level, which was considered mainly to represent NREM periods of sleep. The SI of the first hour of sleep was calculated from the continuous one hour of EEG signal beginning from the sleep onset.

2.4. Natural distribution of ISIs

The distribution of ISIs was evaluated from the first hour of sleep using the event log file of spike detections from BESA®. We motivated this approach by assuming that spike intervals within continuous discharges may be distinct from those that are observed during pauses, i.e. between the continuous discharges. Hence, the natural distribution of ISIs could indirectly provide the range of suitable maxISI values to be used in the SI calculation. Spike intervals of 0.2 s or less were excluded as technical artefacts because they are within the length of individual spike and wave, and a considerable proportion of them came from multiple detections of the same spike by different templates.

2.5. Temporal variation of SI

The overnight variation of SI during sleep was assessed from normalized SI (2 min epochs) averaged in 20 or 60 min epochs. Normalization was done by dividing each SI by the overnight mean SI. Finally, the temporal stability of SI in shorter time scales was estimated from varying the epoch length between 2 and 60 min.

2.6. Variations of the spike search protocol

The sensitivity of SI to variations in BESA spike search was assessed by using additional spike search procedures and subsequent quantitation of SI from the first hour of sleep in eight EEG recordings.

First, we assessed how stringency in spike search would affect SI output. We searched spikes using three different combinations of correlation thresholds in our two-step search procedure that always started from the same initial template (visually selected spike): for the least stringent search (“low correlation”), we used a correlation threshold of 65% for the search over one EEG channel with the initial template, and a threshold of 40% for the search over all channels with the final template. These thresholds were then elevated to 75% and 50%, or 85% and 60%, to obtain spikes with “moderate correlation” (the standard method in Larsson et al. (2009)) and “high correlation”, respectively. This was applied to the four EEGs (Nos. 2, 6, 9 and 10) with only one spike focus.

Second, we aimed to assess how critical the quality of spike search is for SI quantitation. To attain this, we altered the number of prototype spikes and thus, the number of initial spike templates while keeping correlation percentages constant (75% and 50%) in the spike searches. This was applied to four EEGs (Nos. 3–5 and 8) where multiple spike templates were needed to tag at least 80% of spikes. The goal was to obtain one satisfactory, one good and one excellent spike match. The quality of spike search was assessed visually from five randomly selected pages during the first...
hour of sleep. The reference, rated as “good”, was a standard spike match involving at least 80% of spikes. An “excellent” match was created by adding a template and a subsequent search for those spikes that were missed after a good search. Likewise, a satisfactory match was obtained by excluding tags created by one of the templates used for the good spike match. The result of spike search after this kind of exclusion was disproportionately poor in one EEG, and therefore one template was replaced by another instead of exclusion.

2.7. Visual scoring

Two experienced clinical neurophysiologists (M.P. and S.V.) analyzed the first hour of sleep independently in all 10 EEGs, using similar filter settings, montages, as well as time (15 s/window) and amplitude (400 nAm–2.5 μAm) scaling. The readers tagged manually all ISIs of approximately 3 s or more. Tags (i.e. time points) were imported to Excel® and ISIs of more than 3 s were regarded as epochs without discharge. Akin to the semiautomated SI analysis, SI was then calculated as the proportion of continuous discharge. The mean SI of the two readers was taken as the gold standard. Finally, semiautomated SI obtained by the spike search with the moderate correlation percentages and max ISI3 was compared with the gold standard.

2.8. Statistics

Normality of the data was assessed by Shapiro–Wilk test (SPSS®). Differences between two groups were evaluated with the Mann–Whitney U test, and considered significant if $p < 0.05$ (two-tailed). The pair-wise correlation between scorings was analyzed with the Spearman rank correlation test.

3. Results

3.1. SI and maxSI

Expectedly, the SI increased with the increase in duration of max ISI (Fig. 2A) until a ceiling effect could be observed at around 3 s (2–4 s) in each recording. The maxSI of less than 3 s resulted in a clearly increasing intersubject variation and lower SI estimates. Comparison of maxSI2 and maxSI3 in one patient before and after drug treatment (Fig. 2D) showed, that the relative change due to treatment is markedly similar with both maxSI values.

3.2. Natural characteristics of ISI and SI

3.2.1. Distribution of ISIs

ISIs during spiking was examined to find a more physiological argument for the choice of maxSI. A representative histogram of ISI durations during the first hour of sleep in one EEG is shown in Fig. 2B (for all subjects, see Supplementary Fig. S2). While 64–100% (mean 91%) of all ISIs were less than 2.0 s, almost all (77–100%; mean 96%) ISIs were shorter than 3.0 s (Fig. 2C). These observations together suggest that using maxSI of 3 s (maxSI3) would cover nearly all spiking.

3.2.2. Temporal variation of SI

The stability of SI over a whole night sleep was assessed using maxSI3. The normalized SI averaged over epochs of 20 or 60 min showed an apparently random overnight fluctuation, with a slight tendency toward higher SI values at the beginning of sleep (Fig. 3A, Supplementary Fig. S3). Notably, there was no significant difference between the SI estimate of the first sleep hour and the whole night sleep ($p = 0.097$) (Fig. 3B).

Fig. 2. Relation between maxSI and SI. (A) The effect of maxSI (1–7 s) on SI quantified from 10 overnight sleep EEGs. (B) The distribution of ISIs in one EEG. ISIs of 0.2 s or less (gray bars) were discarded as technical artefacts, typically double detections of the same spike. (C) The cumulative percentage of spike intervals (0.3–7.0 s) in 10 EEGs. (D) The treatment effect on SI calculated with maxSI2 and maxSI3 in one patient.

Fig. 3. Temporal variability of SI. (A) The normalized SI by 20 min epochs for the whole night (all recordings). Missing values are SIs below the mean SI of the whole night, and they mostly represent SI of REM-sleep or awake. (B) The difference between the SI of the first hour of sleep and the SI of the whole night. (C) The SI of the first hour of sleep calculated by epoch lengths of 2–60 min from one EEG.

Fig. 4. Effect of spike search on SI. Spike search was done with three different (low, moderate and high) correlation percentages from four EEGs. Graph (A) shows the effect of correlation percentage on SI when SI is quantified with maxISI2 and graph (B) with maxISI3. The effect of the quality (excellent, good and satisfactory) of spike search on SI is seen in graph (C) with maxISI2 and in graph (D) with maxISI3.
The effect of epoch length (2–60 min) on the SI within an individual is shown in Fig. 3C (for all subjects, see Supplementary Fig. S4). A reasonably stable SI was obtained from already a 30 min long epoch in eight of 10 EEGs (EEGs No. 2–9, Supplementary Fig. S4). The difference between the SI of 30 and 60 min epoch was 0.3–4.0 (mean 2.0) SI percentage units in those eight patients, and 0.3–13.5 (mean 3.9) SI percentage units in all 10 patients.

### 3.3. Stability of SI for variations in the spike search

The SI of the first hour of sleep was calculated from the BESA \textsuperscript{a} template matching with a high, a moderate, and a low correlation percentage from the same spike template (Fig. 4A and B). The SI calculated with the high and the low correlation percentage differed from the SI calculated with the moderate correlation percentage by 1.5–14.3 (mean 6.1) SI percentage units with maxISI2, and this difference was substantially reduced (0.06–6.4; mean 3.3) when maxISI was increased to 3 s. Corresponding relative differences were 3.0–9.9% (mean 9.2%) for the maxISI2 and 0.07–9.7% (mean 4.2%) for the maxISI3. These imply that maxISI3 is considerably less sensitive to variation in the stringency of spike template matching.

Another type of variation in the quality of BESA search was produced by varying the amount of initial spike templates and keeping the correlation threshold constant in the spike search. This resulted in three different outputs: satisfactory, good and excellent spike search. The SI calculated with both maxISI2 and maxISI3 was comparable between excellent and good searches. The insufficient number of spike templates did clearly affect both SI estimates, ISI2 being more labile (Fig. 4C and D).

### 3.4. Comparison of visual and semiautomatic scoring

The SIs calculated from the visual scorings of the two independent EEG readers were strongly correlated (Spearman $r = 0.94$, $p < 0.01$) (Fig. 5). The mean difference between the semiautomatic SI and the gold standard generated by expert markings was only 7.0 (range 0.4–19.6) SI percentage units ($p = 0.05$). The relative error between the semiautomatic SI and the gold standard was 0.5–23.2% (mean 7.6%). The semiautomatic method gave typically (in eight out of 10 patients) slightly lower SI than the visual analysis. Notably, both visual scoring and semiautomatic SI calculation resulted in SI values that exceeded 80% in all but one patient (Fig. 5).

### 4. Discussion

Our observations suggest that semiautomatic SI quantification is reasonably robust, stable, as well as intuitively clear, hence holding promise for both clinical and scientific use. While our findings are fully compatible with the earlier work (Larsson et al., 2009, 2010), we extend the existing literature by demonstrating the effects of user-defined settings in SI quantification, as well as the natural characteristics of spiking in patients with CSWS. Also, we show how the method compares with visual scoring. As a result, this study yielded recommendations for the semiautomatic quantification of SI.

### 4.1. MaxISI

Our observations confirm that SI and the definition of continuous discharge are strongly influenced by the choice of maxISI. It would be ideal to have a solid physiological basis, such as a potential refractory mechanism, as an argument for a particular maxISI value. To the best of our knowledge, however, such evidence is not available. In the original papers by Larsson and co-workers, the maxISI (3 s) was argued with the presumption that each spike would affect the brain for 3 s (Mirskey and Vanburen, 1965; Rugland, 1990; Binnie and Marston, 1992). While we feel that their cited physiological evidence is more permissive than instructive, Larsson and co-workers still come to a practical conclusion that is congruent with our study based on actual EEG spiking behavior in our CSWS patients. Our approach to settle this question was to search through our real EEG datasets in order to see the relation between SI and maxISI, as well as to see the natural distribution of spike intervals in the EEG signal. We hoped that these findings would lead to a choice of maxISI that could (i) catch the natural spiking behavior in the EEG, (ii) be sufficiently stable across patients and over time, and yet (iii) be sensitive to changes associated with treatment and recovery.

Our findings suggest that the optimal duration for a maxISI is 3 s. We showed that maxISI shorter than 3 s will rapidly decrease the SI estimate, whereas maxISI longer than 4 s (up to 7 s) only modestly increases the SI estimate. This is in line with the observed distribution of inter-spike intervals concentrating at around 1–2 s with a long tail to the right. The distribution we observed predicts that a short maxISI will yield a potentially labile SI estimate due to its sensitivity to mere ‘random variation’ in spiking behavior, such as what we showed to occur at time scales from minutes to hours (see Fig. 3A). An unnecessarily long maxISI above ceiling level may cause opposite problems. It may lead to a situation where SI loses its sensitivity to reflect true SI changes such as treatment effects, or SI may fail to correctly recognize patients that are near the elusive diagnostic SI threshold. Finally, our results also show that reducing the maxISI from 3 to 2 s does render the paradigm more labile if there is any uncertainty in the spike search.

### 4.2. Validation

Any automated detection or quantitation method is ideally assessed against a gold standard. Since there are no standardized routines for assessing spiking activity, we decided to compare our semiautomatic SI quantification to manual markings of human EEG experts that applied maxISI criteria similar to the automated PS script (taken as gold standard). The mean error of semiautomated SI compared with gold standard was only seven SI percentage units. We feel that this is a clinically acceptable accuracy, and importantly, it also falls well within the intrindividual SI fluctuation that we observed during a whole night.

### 4.3. Overnight sleep vs. short naps

CSWS diagnosis is conventionally based on a whole night sleep (Morikawa et al., 1985; Tassinari et al., 2000; Guzzetta et al., 2005; Inutsuka et al., 2006; Liukkonen et al., 2010). Since both the
recording and analyses of a whole night EEG are time consuming and laborious, a short nap EEG would be an attractive alternative for routine screening or even diagnosis of CSWS. Larsson et al. (2010) found that the SI calculated from a morning nap after sleep deprivation was comparable with the whole-night SI, albeit the nap SI tended to be higher than the mean SI of the whole night. Earlier studies have demonstrated that spike density peaks during nap SI tended to be higher than the mean SI of the whole night. Deprivation was comparable with the whole-night SI, albeit the search and analysis of epileptic spikes (BESA Epilepsy) opens an attractive avenue to fully automated SI quantitation, which needs to be developed and tested in a separate, prospective study.

4.4. Limitations and future directions

For a comprehensive optimization of the semiautomatic SI quantification method, we would have needed more patients with varying levels of spiking activity. All our results are from EEGs with SI over 50%, and all our patients did indeed have a CSWS diagnosis or CSWS was suspected from a previous nap EEG. Also, our gold standard was created by using visual rules (maxSI level) close to those implemented in the SI quantification; hence they were not fully independent.

The main challenge in the present protocol relates to the way how different spike types are recognized and how template spikes are selected so that at least 80% of spikes are found. Recent implementation of the hyperclustering algorithm into the automated search and analysis of epileptic spikes (BESA Epilepsy®) opens an attractive avenue to fully automated SI quantitation, which needs to be developed and tested in a separate, prospective study.

Acknowledgments

This study was supported by Grants from Finnish Epilepsy Research Foundation (MEP), National Graduate School of Clinical Investigation (MEP), Foundation for Paediatric Research (KP), Finnish Academy (SV), and from the European Community’s Seventh Framework Programme European Community FP7-PEOPLE-2009-IOF, Grant agreement No. 254235 (SV). The authors have no conflict of interest to disclose.

Appendix A. Supplementary data


References


