Delirium Detection Using EEG
What and How to Measure

Arendina W. van der Kooi, PhD; Irene J. Zaal, MD; Francina A. Klijn, MD; Huiberdina L. Koek, MD, PhD; Ronald C. Meijer, MD; Frans S. Leijten, MD, PhD; and Arjen J. Slooter, MD, PhD

BACKGROUND: Despite its frequency and impact, delirium is poorly recognized in postoperative and critically ill patients. EEG is highly sensitive to delirium but, as currently used, it is not diagnostic. To develop an EEG-based tool for delirium detection with a limited number of electrodes, we determined the optimal electrode derivation and EEG characteristic to discriminate delirium from nondelirium.

METHODS: Standard EEGs were recorded in 28 patients with delirium and 28 age- and sex-matched patients who had undergone cardiothoracic surgery and were not delirious, as classified by experts using *Diagnostic and Statistical Manual of Mental Disorders, 4th edition*, criteria. The first minute of artifact-free EEG data with eyes closed as well as with eyes open was selected. For each derivation, six EEG parameters were evaluated. Using Mann-Whitney U tests, all combinations of derivations and parameters were compared between patients with delirium and those without. Corresponding P values, corrected for multiple testing, were ranked.

RESULTS: The largest difference between patients with and without delirium and highest area under the receiver operating curve (0.99; 95% CI, 0.97-1.00) was found during the eyes-closed periods of the EEG, using electrode derivation F8-Pz (frontal-parietal) and relative δ power (median [interquartile range (IQR)] for delirium, 0.59 [IQR, 0.47-0.71] and for nondelirium, 0.20 [IQR, 0.17-0.26]; P = .0000000000018). With a cutoff value of 0.37, it resulted in a sensitivity of 100% (95% CI, 100%-100%) and specificity of 96% (95% CI, 88%-100%).

CONCLUSIONS: In a homogenous population of nonsedated patients who had undergone cardiothoracic surgery, we observed that relative δ power from an eyes-closed EEG recording with only two electrodes in a frontal-parietal derivation can distinguish among patients who have delirium and those who do not.
Delirium is an acute disturbance of attention and cognition.\(^1\) It is a common disorder in postoperative and critically ill patients and associated with higher mortality and long-term cognitive impairment.\(^2,3\) Despite its frequency and impact, recognition of delirium by ICU physicians appeared to be poor (sensitivity, 29%).\(^4\) Therefore, several delirium assessment tools have been developed. The delirium assessment tool with the highest sensitivity in postoperative patients was the Nursing Delirium Symptom Checklist (sensitivity, 29%-95%),\(^5,6\) whereas the Confusion Assessment Method for the ICU (CAM-ICU) had the highest sensitivity in patients in the ICU (64%-100%).\(^7,9\)

It should be noted, however, that in these studies, delirium assessments were performed by a limited number of dedicated research nurses, and it is questionable whether these findings can be generalized to routine, daily practice, where numerous bedside nurses use these tools. The sensitivity of the Nursing Delirium Symptom Checklist in a real-life, postoperative setting has never been investigated,\(^3\) whereas the sensitivity of the CAM-ICU in routine ICU care appeared to be much lower than in a research setting.\(^\text{10}\) In a multicenter study in which teams of three delirium experts (psychiatrists, geriatricians, and neurologists) acted as the reference standard, the sensitivity of the CAM-ICU for delirium in routine, daily practice was low (<47%), whereas the specificity was high (98%).\(^\text{10}\) However, a study from the same institution where the CAM-ICU was developed showed a higher sensitivity (81%; specificity, 81%), but this was a single-center study in which research nurses acted as the reference standard.\(^\text{11}\) Furthermore, the CAM-ICU cannot be used to quantify the severity of delirium. As a consequence, delirium recognition is impaired and treatment is delayed, which may impair outcome.\(^\text{12}\)

A new approach to detect delirium, which may fit better in the culture of the postanesthesia care unit (PACU) and ICU, is to monitor physiologic alterations. Delirium is a manifestation of encephalopathy with altered function of neural networks.\(^\text{13}\) It has been known for decades that during delirium, EEG shows slowing of background activity.\(^\text{14}\) To use EEG for daily delirium screening is, however, time-consuming and unpractical, as it can only be performed and interpreted by trained personnel. EEG monitoring with automatic processing has become technically feasible.\(^\text{15}\) EEG-based detection with a limited number of electrodes and automatic processing is more practical and could possibly increase recognition of delirium. However, it is unclear which combination of EEG characteristic and electrode derivation would be the best in differentiating delirium from nondelirium.\(^\text{16}\) The objective of this study was to determine the electrode derivation and EEG characteristic that have the best capability to distinguish patients with delirium from patients without delirium. As a first step, we focused on a homogeneous population of patients who underwent cardiothoracic surgery and were admitted postoperatively to the ICU.

### Materials and Methods

**Study Design and Patients**

In this single-center observational study, EEGs were recorded in patients who had undergone cardiothoracic surgery and had or did not have delirium and who were matched, on group level, for age and sex. The included patients were admitted postoperatively to the ICU of the University Medical Center Utrecht. The institutional review board approved the study protocol (number 11-073), and written informed consent was obtained at the preoperative outpatient clinic or at hospital admission prior to surgery. Patients aged ≥50 years were eligible for this study if they were to undergo cardiothoracic surgery. We excluded patients with a history of any neurologic or psychiatric disease that may confound the diagnosis of delirium or the EEG. Patients with a previous cerebrovascular event were not excluded unless the event resulted in focal EEG alterations. In that case, the patient was replaced by another patient after evaluation of the EEG recording.

**Delirium Diagnosis and Data Collection**

Daily mental status assessment, including delirium screening, was conducted by research nurses and physicians using the Richmond Agitation and Sedation Scale (RASS) and CAM-ICU during the first five postoperative days.\(^\text{7,17}\) When surgery was complicated, daily mental status assessment was performed on the first 5 days that the patient was not in a comatose state, defined by a RASS score < −3 or a Glasgow coma score < 9.\(^\text{17,18}\) Patients positive for delirium by CAM-ICU received a neuropsychiatric evaluation conducted by a geriatrician, neurologist, or psychiatrist based on revised Diagnostic and Statistical Manual of Mental Disorders, 4th edition, Text Revision, criteria for delirium.\(^\text{1}\) This neuropsychiatric evaluation included assessment of the level of consciousness, attention, orientation, memory, language, and disorganized thinking. When the delirium expert classified the patient as “definite delirium,” the patient received an EEG recording. When the delirium expert had a different classification (delirium or nondelirium) based on the Diagnostic and Statistical Manual of Mental Disorders, 4th edition, Text Revision, criteria than the research nurse or physician based on the CAM-ICU, the patient was excluded. After matching on group level for age and sex, patients who were negative by the CAM-ICU also received a similar neuropsychiatric evaluation. When one of these patients was classified as “definite no delirium” by the delirium expert, an EEG recording was conducted. Several patient characteristics were registered: age, sex, APACHE (Acute Physiology and Chronic Health Evaluation) IV score,\(^\text{19}\) Charlson comorbidity index (CCI),\(^\text{20}\) European System for Cardiac Operative Risk Evaluation,\(^\text{21}\) surgery type, bypass time, and medication use during and in the 24 h prior to the EEG recording.

EEGs lasting 30 min were recorded in which patients were asked to keep their eyes open for 15 min and close them for the last 15 min of the recording. The ability to open and close eyes on command was not required for study inclusion. To avoid sleep during the EEG recording, patients were asked to conduct tasks like squeezing their hands at several time points during the recording. In addition to these tasks, patients...
Results

Fifty-eight patients were initially selected, but two equivocal cases were excluded after neuropsychiatric evaluation (Fig 1). In total, 56 patients remained, of whom 28 were delirious and 28 were not delirious. Of the patients with delirium, 14 had a hypoactive subtype (negative RASS scores), seven had a hyperactive subtype (positive RASS scores), and seven had a mixed subtype (both positive RASS and negative RASS scores) during the EEG registration. Five patients with delirium did not show an artifact-free epoch of 1 min with eyes open and were excluded for eyes-open analysis only. One patient with delirium did not have a 1-min epoch of artifact-free data with eyes closed and was excluded for eyes-closed analysis.

Patient characteristics are shown in Table 1. Patients with delirium differed from those without in APACHE IV scores, CCI scores, and haloperidol use. Two patients without delirium used haloperidol in the 24 h previous to the EEG, because of delirium 2 days before the EEG recording and they were tapering off the haloperidol.

Table 2 lists the 10 combinations of EEG derivations and EEG characteristics with the lowest P values for the eyes-closed condition. All combinations were statistically significantly different between delirium and non-delirium. The derivation F8-Pz for relative δ power showed the lowest P value (.0000000000018) and the largest area under the receiver operating curve (0.99). Also, neighboring electrodes of both F8 (eg, Fp2) and Pz (eg, P3 or O1) in combination with relative δ power were in the top five of smallest P values (Fig 2).

Table 3 shows the 10 combinations with the lowest P values for the eyes-open condition. Again, all combinations were statistically significantly different between...
Figure 1 – Flowchart of patient inclusion.

Table 1 shows the characteristics of the study population. The electrode derivation with the lowest P value was P7-P4 in combination with relative α power (P = .00000020), which was associated with an area under the receiver operating curve of 0.90.

When we restricted our analyses to patients with delirium and compared patients taking haloperidol to patients without haloperidol, there was no difference in the best combination for eyes closed (median relative δ power at F8-Pz: haloperidol, 0.58 [IQR, 0.40-0.70] vs no haloperidol, 0.64 [IQR, 0.53-0.71]; P = .39). Furthermore, for the best combination with eyes open (P7-P4, relative α), there was no difference between patients with delirium with or without haloperidol (median relative α at P7-P4: haloperidol, 0.11 [IQR, 0.08-0.14] vs no haloperidol, 0.15 [IQR, 0.10-0.15]; P = .37).

Discussion

In summary, we found that with only two electrodes and 1 min of EEG recording, large differences can be found between patients with and without delirium after cardiothoracic surgery. The largest difference between delirium and nondelirium (ie, the lowest P value) was observed in EEG epochs with eyes closed. In this condition, the optimal EEG characteristic and electrode derivation was the relative δ power in F8 (frontal lateral) to Pz (midline, parietal).

This study represents an innovative approach to detecting delirium. Using EEG with a limited number of electrodes and automatic processing may offer an objective tool to detect the encephalopathy that underlies...
delirium. Patients in the PACU and ICU are monitored for various physiologic alterations. Consequently, EEG-based detection of delirium may fit better in the local culture than cognitive testing. To our knowledge, our study is the first to systematically investigate what the best electrode derivation and EEG characteristic are to detect delirium. We showed that with only two electrodes and 1 min of recording, large differences can be found between patients with and without delirium. Two previous studies found that delirium can be detected with two electrodes, using T5-O1 (parietal to occipital) or C3-A1 (central to left ear) derivations in combination with EEG frequency analyses. However, the observed differences in these studies were much smaller than those found in our investigation.

Previously, we determined systematically which combination of bipolar electrode derivation and EEG characteristic showed the largest differences between patients with and without delirium. Unlike previous quantitative EEG studies in delirium, we did not observe that relative power was an important characteristic to distinguish patients with delirium from those without. Previous studies often stratified the frequency band into an upper and a lower part, with only the lower part proving to be significantly different between delirium and nondelirium. We did not use further division of the band, as this would have further increased the number of possible combinations to study.

Delirium is a common disorder with a prevalence up to 45% in the PACU and 32% when both the PACU and in-patient surgical ward are combined. In patients in the ICU, the prevalence may be even higher and ranges between 60% and 80% for patients on mechanical ventilation and 20% to 50% for those not mechanically ventilated. Although prevalence is high, multiple studies reported that in clinical practice, 50% of patients with delirium were missed using delirium screening tools such as the CAM-ICU. Specifity was high in these studies. A new diagnostic tool, therefore, should increase sensitivity to 50% to improve our current recognition of delirium.

**TABLE 2**

The 10 Combinations of EEG Derivation and Characteristic That Showed the Lowest P Value in Discriminating Delirium From Nondelirium With Patients’ Eyes Closed

<table>
<thead>
<tr>
<th>Rank</th>
<th>P Value</th>
<th>Derivation</th>
<th>Characteristic</th>
<th>Delirium, Median (IQR)</th>
<th>Nondelirium, Median (IQR)</th>
<th>AUC (95% CI)</th>
<th>Sensitivity, % (95% CI)</th>
<th>Specificity, % (95% CI)</th>
<th>Cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0000000000000018</td>
<td>F8-Pz</td>
<td>Rel. θ</td>
<td>0.59 (0.47-0.71)</td>
<td>0.20 (0.17-0.26)</td>
<td>0.99 (0.97-1.00)</td>
<td>100 (100-100)</td>
<td>96 (88-100)</td>
<td>0.37</td>
</tr>
<tr>
<td>2</td>
<td>0.0000000000000037</td>
<td>F8-P3</td>
<td>Rel. θ</td>
<td>0.59 (0.46-0.69)</td>
<td>0.19 (0.15-0.26)</td>
<td>0.99 (0.96-1.00)</td>
<td>96 (87-100)</td>
<td>96 (88-100)</td>
<td>0.35</td>
</tr>
<tr>
<td>3</td>
<td>0.0000000000000011</td>
<td>F8-O2</td>
<td>Rel. θ</td>
<td>0.60 (0.49-0.73)</td>
<td>0.23 (0.18-0.30)</td>
<td>0.99 (0.95-1.00)</td>
<td>96 (87-100)</td>
<td>96 (87-100)</td>
<td>0.42</td>
</tr>
<tr>
<td>4</td>
<td>0.0000000000000015</td>
<td>Fp2-O1</td>
<td>Rel. θ</td>
<td>0.66 (0.60-0.75)</td>
<td>0.27 (0.23-0.36)</td>
<td>0.99 (0.97-1.00)</td>
<td>96 (87-100)</td>
<td>95 (85-100)</td>
<td>0.42</td>
</tr>
<tr>
<td>5</td>
<td>0.00000000000000017</td>
<td>F8-F4</td>
<td>Rel. θ</td>
<td>0.60 (0.43-0.70)</td>
<td>0.20 (0.17-0.26)</td>
<td>0.98 (0.94-1.00)</td>
<td>96 (87-100)</td>
<td>92 (81-100)</td>
<td>0.35</td>
</tr>
<tr>
<td>6</td>
<td>0.00000000000000022</td>
<td>F8-O1</td>
<td>Rel. θ</td>
<td>0.62 (0.48-0.72)</td>
<td>0.22 (0.17-0.26)</td>
<td>0.99 (0.95-1.00)</td>
<td>96 (87-100)</td>
<td>95 (87-100)</td>
<td>0.37</td>
</tr>
<tr>
<td>7</td>
<td>0.00000000000000024</td>
<td>F8-Cz</td>
<td>Rel. θ</td>
<td>0.57 (0.46-0.67)</td>
<td>0.26 (0.20-0.33)</td>
<td>0.98 (0.94-1.00)</td>
<td>91 (80-100)</td>
<td>96 (88-100)</td>
<td>0.41</td>
</tr>
<tr>
<td>8</td>
<td>0.00000000000000024</td>
<td>F8-C3</td>
<td>Rel. θ</td>
<td>0.57 (0.49-0.67)</td>
<td>0.21 (0.17-0.30)</td>
<td>0.98 (0.94-1.00)</td>
<td>91 (80-100)</td>
<td>92 (81-100)</td>
<td>0.35</td>
</tr>
<tr>
<td>9</td>
<td>0.00000000000000029</td>
<td>Fp2-Pz</td>
<td>Rel. θ</td>
<td>0.64 (0.53-0.73)</td>
<td>0.28 (0.22-0.36)</td>
<td>0.99 (0.95-1.00)</td>
<td>100 (100-100)</td>
<td>95 (87-100)</td>
<td>0.42</td>
</tr>
<tr>
<td>10</td>
<td>0.00000000000000030</td>
<td>Cz-O1</td>
<td>Rel. θ</td>
<td>0.50 (0.37-0.57)</td>
<td>0.17 (0.10-0.25)</td>
<td>0.96 (0.91-1.00)</td>
<td>88 (76-100)</td>
<td>96 (89-100)</td>
<td>0.36</td>
</tr>
</tbody>
</table>

AUC = area under the curve of the receiver operating curve; Rel. θ = relative power in the θ frequency band. See Table 1 legend for expansion of other abbreviations.

All P values were <.00004. Therefore, all combinations in this table showed a statistically significantly difference between delirium and nondelirium.
posttest probability to a threshold below which it is acceptable to miss delirium.

Some limitations of the present study should be acknowledged. This study describes only the first step toward an objective delirium-detection tool. Due to the study design, in which only definitive cases were included, we cannot derive the incidence of delirium in this population. Therefore, parameters such as the area under the receiver operating curve, sensitivity, and specificity may be too optimistic and should be interpreted with caution. In our study, artifact-free data were manually selected for analysis and not automatically chosen. Automatic artifact-detection programs are available and already implemented in single-channel, sleep EEG analysis programs. A comparable automatic detection algorithm needs to be implemented in a future objective delirium-detection tool. The homogenous study population of patients who had undergone cardiothoracic surgery did not have residual sedation, which may be an issue in delirium detection in a general population of patients in the ICU. Besides, the etiology of delirium after cardiothoracic surgery differs from the etiology of ICU delirium. It is, therefore, unclear whether our findings can be extrapolated to a general ICU population. Another possible limitation of this study may be that differences in haloperidol use between the delirium and the nondelirium groups could have affected the results. It is, however, unlikely that this explains our findings, as we found no differences between users and nonusers of haloperidol within the group of patients with delirium. As delirium is caused by an underlying disease, unsurprisingly, there was a difference in APACHE IV score and CCI between patients with and without delirium. Furthermore, we cannot exclude that the results might be affected due to the problem that patients with delirium needed more reminders to keep their eyes closed or open than those without delirium. As both groups needed reminders to open or close their eyes, and only 1 min of EEG recording was analyzed, we expect that a possible effect of this problem will only be minimal. In this explorative study, patients with neurologic or psychiatric disease were excluded as well as doubtful cases, because these are the patients for whom making a diagnosis of delirium is the most difficult. In future validation studies, it should be explored whether...
In conclusion, we showed that with two electrodes and 1 min of EEG data, delirium can be discriminated from non-delirium under certain circumstances. This opens the prospect of EEG-based detection of delirium.

A serious limitation of this study is that the proposed method for discriminating delirium from non-delirium was not tested in an independent study population, which will be necessary to prove that this technique can be used as a diagnostic test for delirium. Therefore, future studies should validate the results of this study and investigate whether our findings can be generalized to an unselected population of patients in the ICU and to a general population of postoperative patients. To generalize the findings future studies should be conducted in unselected populations with blinding of the physician making the delirium diagnoses to the results of the EEG-based delirium monitor. The effects of residual sedation and automatic data selection should also be considered in these studies. The technique proposed in this study can also be applied to these subgroups.

**TABLE 3** The 10 Combinations of EEG Derivation and Characteristic That Showed the Lowest P Value in Discriminating Delirium From Nondelirium With Patients' Eyes Open

<table>
<thead>
<tr>
<th>Rank</th>
<th>P Value</th>
<th>Derivation</th>
<th>Characteristic</th>
<th>Delirium, Median (IQR)</th>
<th>Nondelirium, Median (IQR)</th>
<th>AUC (95% CI)</th>
<th>Sensitivity, % (95% CI)</th>
<th>Specificity, % (95% CI)</th>
<th>Cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00000020</td>
<td>P7-P4</td>
<td>Rel. α</td>
<td>0.12 (0.09-0.15)</td>
<td>0.33 (0.19-0.39)</td>
<td>0.90 (0.81-1.00)</td>
<td>85 (70-100)</td>
<td>91 (80-100)</td>
<td>0.16</td>
</tr>
<tr>
<td>2</td>
<td>0.0000042</td>
<td>P3-P4</td>
<td>Rel. α</td>
<td>0.14 (0.11-0.17)</td>
<td>0.34 (0.23-0.43)</td>
<td>0.89 (0.79-0.99)</td>
<td>81 (65-98)</td>
<td>91 (80-100)</td>
<td>0.19</td>
</tr>
<tr>
<td>3</td>
<td>0.000016</td>
<td>P7-O1</td>
<td>Rel. δ</td>
<td>0.44 (0.36-0.54)</td>
<td>0.24 (0.17-0.33)</td>
<td>0.88 (0.78-0.99)</td>
<td>86 (71-100)</td>
<td>85 (71-98)</td>
<td>0.35</td>
</tr>
<tr>
<td>4</td>
<td>0.000032</td>
<td>P7-O1</td>
<td>Rel. α</td>
<td>0.10 (0.08-0.14)</td>
<td>0.26 (0.19-0.33)</td>
<td>0.87 (0.77-0.98)</td>
<td>81 (64-98)</td>
<td>90 (79-100)</td>
<td>0.18</td>
</tr>
<tr>
<td>5</td>
<td>0.000035</td>
<td>P3-P4</td>
<td>S/F ratio</td>
<td>4.0 (2.5-5.2)</td>
<td>1.0 (0.6-1.7)</td>
<td>0.87 (0.75-0.97)</td>
<td>77 (60-95)</td>
<td>89 (77-100)</td>
<td>2.09</td>
</tr>
<tr>
<td>6</td>
<td>0.000040</td>
<td>P4-O1</td>
<td>Rel. α</td>
<td>0.13 (0.09-0.17)</td>
<td>0.29 (0.19-0.39)</td>
<td>0.87 (0.76-0.98)</td>
<td>78 (60-96)</td>
<td>90 (79-100)</td>
<td>0.18</td>
</tr>
<tr>
<td>7</td>
<td>0.000061</td>
<td>P7-O1</td>
<td>Rel. α</td>
<td>0.11 (0.09-0.16)</td>
<td>0.31 (0.19-0.39)</td>
<td>0.86 (0.75-0.98)</td>
<td>81 (64-98)</td>
<td>95 (87-100)</td>
<td>0.18</td>
</tr>
<tr>
<td>8</td>
<td>0.000079</td>
<td>P7-P4</td>
<td>S/F ratio</td>
<td>4.0 (2.9-5.6)</td>
<td>1.1 (0.7-2.2)</td>
<td>0.86 (0.74-0.97)</td>
<td>77 (60-95)</td>
<td>88 (76-100)</td>
<td>2.58</td>
</tr>
<tr>
<td>9</td>
<td>0.000094</td>
<td>P3-P8</td>
<td>Rel. α</td>
<td>0.13 (0.09-0.16)</td>
<td>0.32 (0.19-0.43)</td>
<td>0.86 (0.74-0.97)</td>
<td>78 (60-96)</td>
<td>90 (79-100)</td>
<td>0.18</td>
</tr>
<tr>
<td>10</td>
<td>0.00011</td>
<td>P7-O2</td>
<td>Rel. α</td>
<td>0.11 (0.09-0.15)</td>
<td>0.29 (0.19-0.37)</td>
<td>0.86 (0.75-0.97)</td>
<td>76 (58-94)</td>
<td>95 (87-100)</td>
<td>0.20</td>
</tr>
</tbody>
</table>

**Absence**

See Table 1 and 2 legends for expansion of other abbreviations.

a All P values were < .00056. Therefore, all combinations in this table showed a statistically significantly difference between delirium and non-delirium.
Acknowledgments


Financial/nonfinancial disclosures: The authors have reported to CHEST the following conflicts of interest: Drs van der Kooi and Slooter hold a patent (pending) through the University Medical Center of Utrecht entitled “Method and system for determining a parameter which is indicative of the accuracy of the data analysis in delirium in the elderly. Br J Anaesth. 2011;111(4):612-618. 6. Radtke FM, Franck M, Schneider M, et al. Comparison of three scores to screen for delirium in the recovery room. Br J Anaesth. 2008;101(3):338-343.


