Original Article

Clinical EEG and Neuroscience
1–10

DOI: 10.1177/1550059418767584
journals.sagepub.com/home/eeg
SAGE

Diagnostic Performance and Utility of Quantitative EEG Analyses in Delirium: Confirmatory Results From a Large Retrospective Case-Control Study

Robert Fleischmann[1,2*], Steffi Tränkner[1*], Rouven Bathe-Peters[3], Maria Rönnefarth[1], Sein Schmidt[1], Stephan J. Schreiber[4], and Stephan A. Brandt[1]

Abstract

Background. The lack of objective disease markers is a major cause of misdiagnosis and nonstandardized approaches in delirium. Recent studies conducted in well-selected patients and confined study environments suggest that quantitative electroencephalography (qEEG) can provide such markers. We hypothesize that qEEG helps remedy diagnostic uncertainty not only in well-defined study cohorts but also in a heterogeneous hospital population. *Methods*. In this retrospective case-control study, EEG power spectra of delirious patients and age-/gender-matched controls (n = 31 and n = 345, respectively) were fitted in a linear model to test their performance as binary classifiers. We subsequently evaluated the diagnostic performance of the best classifiers in control samples with normal EEGs (n = 534) and real-world samples including pathologic findings (n = 4294). Test reliability was estimated through split-half analyses. *Results*. We found that the combination of spectral power at F3-P4 at 2 Hz (area under the curve [AUC] = .994) and C3-O1 at 19 Hz (AUC = .993) provided a sensitivity of 100% and a specificity of 99% to identify delirious patients among normal controls. These classifiers also yielded a false positive rate as low as 5% and increased the pretest probability of being delirious by 57% in an unselected real-world sample. Split-half reliabilities were .98 and .99, respectively. *Conclusion*. This retrospective study yielded preliminary evidence that qEEG provides excellent diagnostic performance to identify delirious patients even outside confined study environments. It furthermore revealed reduced beta power as a novel specific finding in delirium and that a normal EEG excludes delirium. Prospective studies including parameters of pretest probability and delirium severity are required to elaborate on these promising findings.

Keywords

delirium, quantitative electroencephalography, diagnostic performance, validation, reliability, adults

Received May 28, 2017; revised December 13, 2017; accepted February 26, 2018.

Introduction

Delirium describes an acute confusional state that is associated with cognitive impairment in one or more domains in the context of an underlying medical condition, medication or intoxication.[1] It is estimated to affect up to 70% of hospitalized patients, with a relative increase in correlation with patient age and disease severity.[2,3] Substantial health care costs and increased morbidity, including in-hospital complications and sequelae such as cognitive impairment, underline the associated burden.[4-6] Despite the fact that delirium is generally considered a reversible condition, diagnosing and treating patients so as to avoid short- and long-term complications remains a challenge. Importantly, the diagnostic process is closely tied to the clinical experience of the treating physician.[7-9] In this context, the precision medicine initiative has highlighted that objective disease marker, that is, biomarkers, are key to guide individualized approaches to patients in order to improve individual outcome.[10] Such biomarkers are as yet unavailable in delirium, which poses an explanation for the current lack of standardized approaches.[11,12] Brain imaging modalities such as computed tomography (CT) and magnetic resonance imaging

[1]Vision and Motor System Research Group, Department of Neurology, Charité–Universitätsmedizin Berlin, Berlin, Germany
[2]Neurointensive Care Unit, Department of Neurology, Charité–Universitätsmedizin Berlin, Berlin, Germany
[3]St. Hedwig Hospital, Department of Psychiatry, Charité–Universitätsmedizin Berlin, Berlin, Germany
[4]Department of Neurology, Asklepios Fachklinikum Brandenburg, Brandenburg, Germany

[*]Equal contribution, arbitrary order

Corresponding Author:
Robert Fleischmann, MD, Department of Neurology, Charite Universitatsmedizin Berlin Klinik fur Neurologie mit Experimenteller Neurologie, Charité Campus Mitte, Charitéplatz 1, Berlin, 10117, Germany.
Email: robert.fleischmann@uni-greifswald.de

(MRI) have thus far not been proven to be beneficial in the diagnostic process.[13-15]

Electroencephalography (EEG) is, in contrast, one of the most promising tools for providing diagnostic biomarkers that could help improve diagnostic accuracy in delirium.[16] In addition to being a noninvasive, easy to use, and amply available tool, it is already well established that EEG should be used in every patient suspected for delirium to check for presence of nonconvulsive epileptic activity.[17] In line with this notion, the mainstay of standard EEG (sEEG) analysis remains its qualitative interpretation by an experienced EEG reader, which can already yield decisive information for the diagnosis of delirium, for example, when patients are not amenable to clinical bedside tests or when bedside tests are inconclusive.[18] sEEG can furthermore indicate potential causes of delirium with classic delirium manifesting as diffuse background slowing while delirium tremens presents with a normal EEG or even faster rhythms.[19,20] In contrast, neuroleptic malignant syndrome often causes only minor slowing while metabolic encephalopathies typically include bi- and triphasic waves.[19] Extending on this notion that sEEG can be of help in the diagnostic process, 1 case-control study investigating 28 patients undergoing cardiothoracic surgery reported that quantitative EEG (qEEG) can in fact be used not only to assist the diagnostic process but also to diagnose patients with delirium providing a sensitivity of 100% and specificity of 96%.[16] While these results from a selected patient group are encouraging, they were obtained in a confined study environment and thus require confirmation. This interpretation is supported by recent studies that report substantially reduced sensitivity and specificity of gold standard clinical tools for delirium in the field as compared with study contexts.[21-23] For example, meta-analyses revealed a pooled sensitivity of 80% of the Confusion Assessment Method (CAM) in clinical routine settings and as low as 47% in the intensive care setting (CAM-ICU) as opposed to a sensitivity of 94% to 100% in study environments.[21,23] The diagnostic yield can hence be severely limited in clinical routine settings but it remains unclear if similar trends apply for more objective neurophysiological methods.

The objective of this study was hence to estimate and potentially validate the diagnostic performance of qEEG analyses in a retrospective case-control study, including EEG data from several thousand in-hospital patients. This larger data set allows for reliability estimates that are currently unavailable. For this purpose, we first used EEG data of a well-defined group of delirious patients to establish binary classifiers suitable to separate delirium from age- and gender-matched controls using linear regression methods. Extending on previous studies, we included delirious patients with multiple etiologies of delirium and 2 further control groups to test the diagnostic performance of identified classifiers. The latter control groups included either subjects with normal EEG readings, which were not previously used to define classifiers, or all EEG data in the database. Control groups thus included patients with all diseases prevalent in a university hospital setting including neurological or psychiatric disease that may substantially confound the diagnostic performance of qEEG.

Our hypothesis was that binary classifiers would be well suited to separate delirious patients form normal controls, that is, that a normal EEG excludes delirium, but that diagnostic accuracy would drop substantially when pathologic EEGs were included. Classifiers should also specifically include delta or theta band activity since both are consistently shown to be increased in delirious patients.[24-26]

Materials and Methods

This study complied with local ethics and data protection regulations for retrospective studies and adhered to the Helsinki declaration in its latest revision. Formal consent of the data protection board was obtained since substantial amounts of pseudonymous clinical routine data had to be handled.

Study Design and Patient Selection

We used EEG data recorded between the years 2004 and 2014 at a German university hospital (Charité–Universitätsmedizin Berlin, Germany). As part of clinical routine, all EEG data were being stored in a local database along with a brief medical history and a preliminary diagnosis at the time of EEG acquisition.

This study was conducted in a retrospective case-control study design and EEG data were for our purposes divided into 4 groups for successive analytic steps: DELIRIUM, CONTROL1, CONTROL2, and CONTROL3. A graphical summary of the study design is given in Figure 1, a summary of patient characteristics is given in Table 1.

The first analysis aimed to determine EEG characteristics, that is, binary classifiers, that best discriminate patients with confirmed delirium (DELIRIUM; n = 31, age 75.5 ± 10.9 years, 12 females) from age- and gender-matched controls with reported normal EEG patterns (CONTROL1; n = 345, age 75.3 ± 13.5 years, 134 females). Delirious patients were defined by a positive CAM-ICU rating documented within less than 24 hours before the EEG time stamp. Exclusion criteria for the DELIRIUM group included: structural brain disease and delirium due to alcohol withdrawal since EEG patterns are known to differ in these patients.[27]

The diagnostic performance of identified classifiers was then determined in the context of 2 further control groups. First and in order to test the hypothesis that a normal EEG excludes delirium, we evaluated the classifiers' capability to distinguish delirious patients from an unmatched sample of patients with normal EEGs that were not previously used (CONTROL2; n = 534, age 55.2 ± 17.6 years, 262 females). A normal EEG was defined according to the guidelines of the International Federation of Clinical Neurophysiology.[28] We finally evaluated how the performance of classifiers would change in a real-world sample without including any a priori information about patient conditions (CONTROL3; n = 4294, age 52.9 ± 18.9 years, 2171 females).

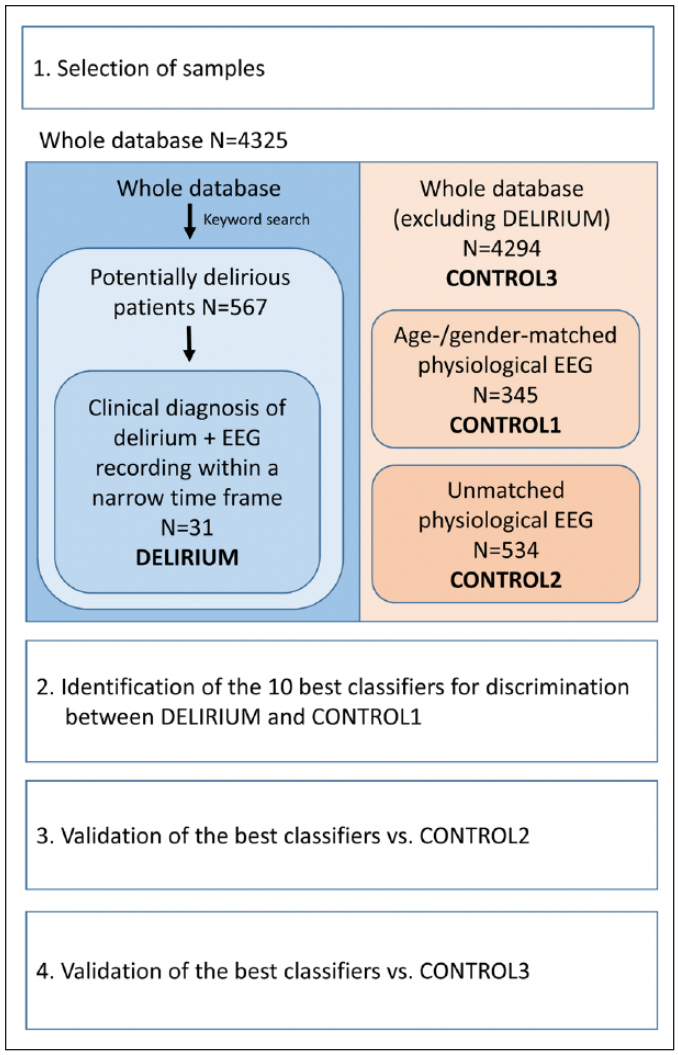

Figure 1. Flowchart of the study design. Step 1: Overview of the selection process for the DELIRIUM, CONTROL1, CONTROL2, and CONTROL3 groups. The whole database consisted of 4325 EEG recordings. From these recordings, 31 could be assigned to patients that were delirious at the time of EEG acquisition. From the remainder EEG data, we drew 3 samples that consisted either of age-/gender-matched physiological controls (CONTROL1), of unmatched physiological controls (CONTROL2) or real-world data without preselection with respect to patient or EEG characteristics (CONTROL3). A step-wise approach was chosen to estimate optimal EEG characteristics for the identification of delirium (step 2) and to test the diagnostic performance versus physiological (step 3) and real-world (step 4) data.

EEG Recording and Processing

EEG were digitally recorded through a commercially available clinical EEG system (Galileo.NET, BE Light system, EB Neuro S.p.A., Firenze, Italy). The montage consisted of standard electrode numbers and positions of the international 10-20 system. The sampling rate was 256 Hz. Routine recordings typically lasted 20 minutes with eyes closed shortly interrupted by provocation maneuvers. We included only epochs that were free of provocation maneuvers.

EEG processing was done with Matlab 2016a (MathWorks, Natick, MA, USA) and the additional Fieldtrip toolbox.[29] Preprocessing included low-pass filtering at 70 Hz, notch filtering at 50 Hz and detrending. Artifacts were dealt with in a stepwise approach: First, data were split into 10-second trials and rejected if excessive variance indicated artifacts and outlier data. Remainder trials were visually inspected and rejected if confounding data were present. Subsequently, all EEG data were standardized by their mean power to allow for valid comparisons between subjects. Power was calculated through a multitaper method fast Fourier transform using Hanning tapers without spectral smoothing.

Analysis

Power spectra of all possible bipolar EEG sensor combinations and frequencies were compared between delirious patients (DELIRIUM) and age-/gender-matched controls (CONTROL1) in order to determine EEG characteristics that best differentiate the 2 conditions. This was done by fitting data in a generalized linear model and calculating a receiver operating characteristic (ROC) for each comparison. Specifically, for each comparison, identified by integer m from 1 to 14 700 (210 channel combinations * 70 frequencies), we calculated a logistic regression model according to the model equation

$$y_{m,i} = p_{m,i}$$

where y is the binary response variable (delirium: yes or no), p is the normalized spectral power for a given channel combination at a given frequency, and i denotes the individual patient from 1 to n. We then obtained the optimal operating point of the resulting ROC and its area under the curve (AUC). The magnitude of the AUC ranges from 0 to 1 with a perfect classifier having a value of 1 and is a direct indicator for the performance of a classifier. The diagnostic performance of the ten best classifiers was determined in the context of unmatched normal EEGs (DELIRIUM vs CONTROL2) and all available EEG data (DELIRUM vs CONTROL3). We added likelihood ratios (LRs) to the analysis since these are a suitable means to determine how test results change the pretest probability that a condition exists in a patient which better reflects the clinical situation[30]:

$$\mathrm{LR+} = \mathrm{sensitivity} / (1 - \mathrm{specificity})$$

The change in pretest probability by a positive test is in linear approximation given by

$$\mathrm{probability}(\%) = \log(\mathrm{LR}) \times .19$$

We then calculated positive and negative predictive values to fully describe the test performance. Subsequently, we tested whether a combination of classifiers from different frequency bands increased the diagnostic yield.

Table 1. Patient Characteristics: Demographic Data and Description of Delirium Subtype in the Study Population.[a]

	DELIRIUM	CONTROL1	CONTROL2	CONTROL3
Sample size (male/female)	31 (19/12)	345 (211/134)	534 (272/262)	4294 (2123/2171)
Mean age (years)	75.5	75.3	55.2	52.9
Standard deviation (years)	10.9	13.5	17.6	18.9
Additional information	Presumed etiology of delirium: Postoperative: 10 Postoperative + sepsis: 7 Sepsis: 5 CNS infection: 4 Metabolic: 1 Unknown: 4	Patients with a normal EEG. Most common reasons for EEG recording in clinical routine were the following suspected or proofed disorders and symptoms: 1. Epilepsy 2. Disorders of sensitivity or motor function 3. Syncopations 4. Dementia 5. Headaches 6. Vertigo 7. Inflammatory events 8. Falls 9. Stroke 10. Memory disorders		Patients with either normal or pathologic EEG.
Motor subtype	Hypoactive: 11 Hyperactive: 5 Mixed: 8 Unknown: 7			
CNS active medication[b] in the past 24hours	28 yes 3 no			

Abbreviation: CNS, central nervous system.
[a]The DELIRIUM and CONTROL1 group were age- and gender-matched and statistically tested for the hypothesis that both samples are from the same population ($P = .95$, z test). For details on how the study population was selected, refer to the section Study Design and Patient Selection.
[b]For details, see the appendix.

Test reliability was estimated through split-half analyses. For this purpose, all data used to calculate the diagnostic performance of identified EEG classifiers (ie, DELIRIUM, CONTROL2, and CONTROL3) were divided into 2 groups in an odd-even manner. The Spearman-Brown formula was then applied to calculate reliability coefficients for the identification of delirious patients among controls with normal EEG patterns or unselected real-world data.

Results

Classifiers Separating Delirium From Matched Physiologic Controls

The 10 best classifiers that distinguish patients with delirium from patients without delirium were confined to the delta and beta frequency range and predominantly included frontal and central sensor locations (Figure 2). Classifiers with the best ROC characteristics were in the delta band at F3-C4 (AUC = .994), F3-P4 (AUC = .994), and O2-F3 (AUC = .993).

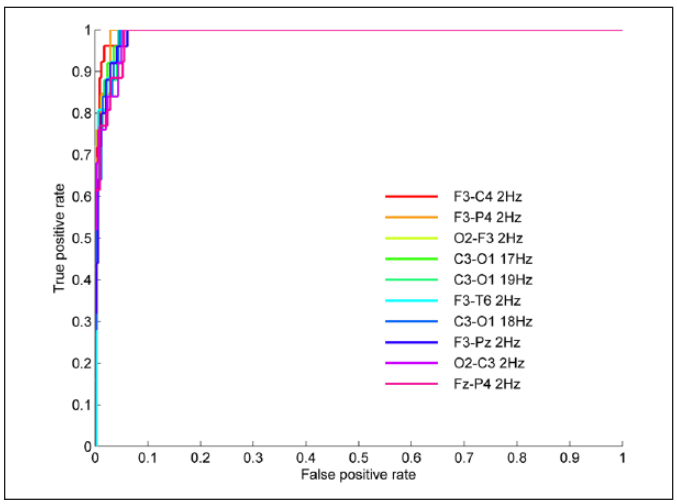

Figure 2. Receiver operating characteristics of the 10 best bipolar montages discriminating between delirious patients (DELIRIUM) and age- and gender matched controls (CONTROL1).

Diagnostic Performance of Classifiers in an Unmatched Sample of Physiologic Controls

In line with results from the previous analyses the best classifiers were in the delta frequency range (Figure 3, Table 2). The classifiers with the best ROCs were again F3-P4 (AUC = .964) and O2-F3 (AUC = .969) at 2 Hz. All classifiers provided a sensitivity ranging from 92% to 100% while specificity was more variable and ranged from 72.66% to 97%. A combination of the best frequency specific classifiers, that is, F3-P4 at 2Hz and C3-O1 at 19 Hz, yielded a further improvement of test performance and provided a sensitivity of 100% and specificity of 99.4% (AUC: = .992). The likelihood ratio for this combination was 172.41 indicating an approximated change in pretest probability of 98%. This combination of classifiers identified all patients of the DELIRIUM group and 3 false positives, but no false negatives. The Spearman-Brown coefficient of split-half reliability was .98. Among the 3 false positive assignments, there were 2 patients who had a history of epileptic seizures and 1 received an EEG because of concentration deficits following viral encephalitis.

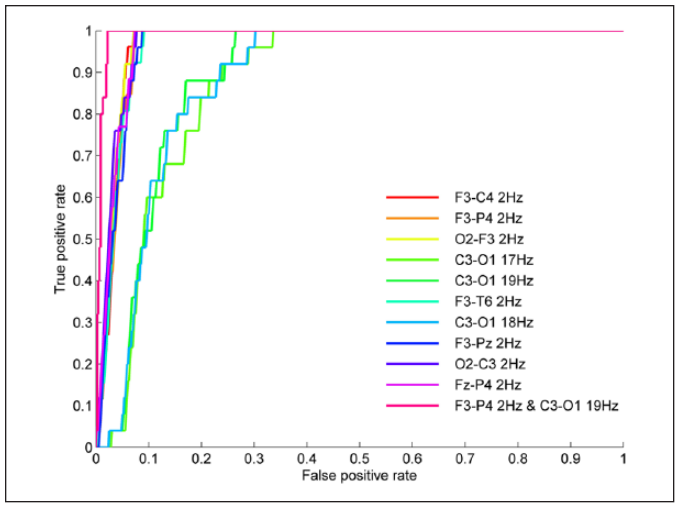

Figure 3. Receiver operating characteristics of the 10 best bipolar montages and the corresponding frequency range discriminating between delirious patients (DELIRIUM) and patients with a physiological EEG (CONTROL2).

Diagnostic Performance of Classifiers in an Unselected Real-World Population

The best classifiers in this scenario were F3-C4 (AUC = .969) and F3-P4 (AUC = .97), and provided a sensitivity of 100% and specificity of 90.98% and 92.55%, respectively (Figures 4 and 5, Table 3). Again, the combination of the two best single classifiers F3-P4 at 2Hz and C3-O1 at 19Hz increased the diagnostic performance providing a sensitivity of 100% and specificity of 94.91% (AUC = .974). The false positive rate was about 5% due to 175 of 4325 patients that were identified as being delirious according to EEG criteria although they were classified as non-delirious based on the brief patient history included in the EEG report. We screened available remainder documentation (eg, charts, reports) of respective patients for information that was potentially omitted in the EEG history and indeed, 10 of these patients did in fact have a clear diagnosis of delirium (clinical rating and received appropriate medication). For the remainder, the level of consciousness and whether or not the patient suffered from delirium was not clearly documented. A total of 69 patients had electrophysiological characteristics of epileptic brain activity, 29 were admitted for stroke, 21 for dementia, and 8 for hypoxic brain injury. There were also individual cases of central nervous system infection, Parkinson's disease, posterior reversible encephalopathy syndrome (PRES), migraine, opioid withdrawal and condition after neurosurgery. The likelihood ratio of 19.65 indicates an increase of the pretest probability by 57% and thus a significant diagnostic yield if test results are positive. Also, the false negative rate remained zero. The Spearman-Brown coefficient of split-half reliability was .99.

Discussion

This is the first study to show that a combination of few EEG sensors can be used to accurately identify and diagnose patients with delirium even in a mixed cohort of several thousand patients. The diagnostic performance of quantitative EEG outpaced that of classic clinical tests despite the lack of a priori information about the patients' condition. Our results furthermore provide novel evidence that decreased beta power is a specific finding in delirium and that a normal EEG rules out delirium.

Identification of the Best EEG Classifiers for the Diagnosis of Delirium

Since Romano and Engel[31] were able to link a slowing of EEG background activity to delirium, numerous studies have validated their findings. Most prominent findings include an increase in relative delta and theta power, and the decrease of posterior alpha power.[20,25,32-35] These hallmarks are implemented in clinical routine interpretations of sEEG since they can easily be qualitatively identified and assist in the diagnostic process.[18] Our results extend on this notion and not only provide further evidence that increased delta power is a typical finding in patients suffering from delirium but that it is also specific when present to a certain extent and at specific sensor locations in qEEG. This is in contrast to common concepts of qualitative EEG interpretations, which state that background slowing was an unspecific finding. Findings from this study also extend classic concepts of EEG changes in delirium by demonstrating that decreased beta power is also a typical finding. These changes, in contradistinction to those in the theta and alpha band, seem to be also highly specific. A thorough discussion of the diagnostic performance of reported classifiers can be found in the next paragraph. At this point, one may wonder why decreased beta power has not previously been reported to be a characteristic of delirium. An intriguing explanation is that studies in the field mainly employed qualitative or semiquantitative analyses of EEG data and may thus have overlooked the effect of beta power. There are as yet few qEEG studies that tried to establish criteria for an objective and reliable diagnosis of delirium and thus need to be discussed in more detail. Two recent studies of van der Kooi et al. focused on the most promising EEG characteristics for the identification of delirium. They found that the relative delta power at F8-Pz provided the best diagnostic performance with a sensitivity of 100% and a specificity of 96%.[16,36] Also ranked in the top 4 were delta power changes at F8-P3, F8-O2, and Fp2-O1. In line with our results, their sensor combinations spanned frontal and pario-occipital regions and performed best in the delta band. To the best of our knowledge, there are only 2 further studies from 1989 and 1997 that investigated quantitative EEG findings in delirium. Koponen et al[37] reported that EEG changes in the alpha and delta range located at T6-O2 and T5-O1 were the most prominent findings in 51 delirious patients and that these changes were correlated with the severity of cognitive deterioration. However, they only evaluated temporo-occipital sensor combinations and may thus have overlooked effects at more frontal and central locations. Matsushima et al[33] found that theta to alpha power ratios were increased at C3 and O1 locations in a sample of 10 delirious patients. This result is well in line with our findings of

Table 2. Results for DELIRIUM Versus CONTROL2.[a]

Reference Channel	Channel	Frequency (Hz)	Cutoff (μV^2/Hz)	LR	Sensitivity (%)	Specificity (%)	False Positive Rate (%)	False Negative Rate (%)
F3	C4	2	>7.05	36.68	96.15	97.38	2.62	3.85
F3	P4	2	>10.02	41.55	100	97.59	2.41	0
O2	F3	2	>11.08	34.18	96.00	97.19	2.81	4.00
C3	O1	17	<0.29	4.35	92.00	78.84	21.16	8.00
C3	O1	19	<0.27	4.33	100	76.89	23.11	0
F3	T6	2	>11.44	22.48	96.15	95.72	4.28	3.85
C3	O1	18	<0.32	3.66	100	72.66	27.34	0
F3	Pz	2	>8.97	20.54	100	95.13	4.87	0
O2	C3	2	>6.04	28.48	96.00	96.63	3.37	4.00
Fz	P4	2	>8.55	33.38	100	97.00	3.00	0
F3	P4	2	>10.02	172.41	100	99.42	0.58	0
C3	O1	19	<0.27					

Abbreviation: LR, likelihood ratio.
[a]Diagnostic performance of the 10 best single bipolar montages (according to their area under the curve (AUC) and best combination of montages (last row) for the discrimination of patients with delirium (DELIRIUM) from patients with a physiological EEG (CONTROL2).

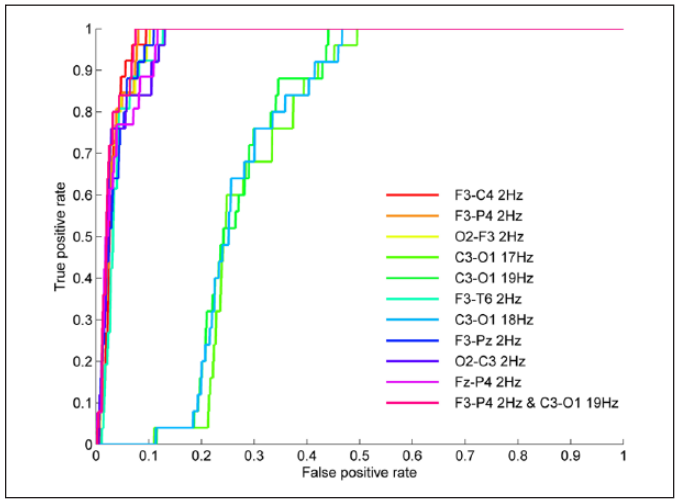

Figure 4. Receiver operating characteristics of the 10 best bipolar montages discriminating between delirious patients (DELIRIUM) and the whole database, including patients with physiological and pathological EEGs (CONTROL3).

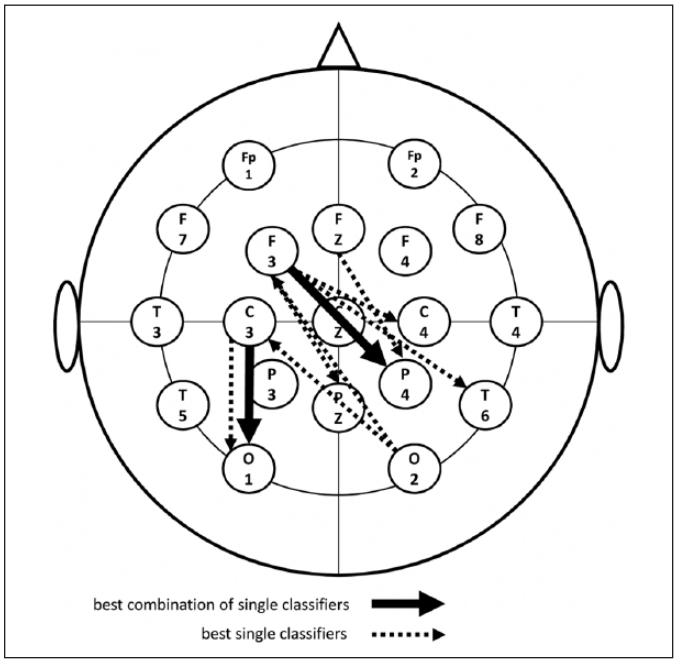

Figure 5. Graphical summary of the best single classifiers and the best combination for the discrimination of delirium patients within the whole database consisting of physiological and pathological EEG recordings other than delirium mainly deriving out of the delta and beta frequency band and mainly including frontal to parieto-occipital derivations. For detailed information on the performance of the classifiers, see Table 3.

C3-O1 power spectra to be significantly changed in the delta (increase) and beta (decrease) band which also indicates an increase of the slow to fast EEG activity ratio in delirium. In summary, our results provide a synthesis of findings from the few available quantitative EEG studies in delirium. We furthermore establish reduced beta activity as a specific finding. It remains elusive and to be investigated in future studies if qEEG can indicate potential causes of delirium as commonly suggested for particular configurations of sEEG.[19]

Diagnostic Performance of EEG Classifiers

The retrospective study design enabled us to examine a mixed population with multiple etiologies for delirium. This distinguishes our study from most former studies that focused on one specific type of delirium (eg, postoperative delirium), a fact that is important to consider when interpreting reported findings since prevalence and causes of delirium are known to differ between clinical settings.[38-40]

Clinical tools such as CAM/CAM-ICU are currently the gold standard and most widely used to diagnose patients with delirium. The broad acceptance and recommendation to use

Table 3. Results for DELIRIUM Versus CONTROL3.[a]

Reference Channel	Channel	Frequency (Hz)	Cutoff (μV^2/Hz)	LR	Sensitivity (%)	Specificity (%)	False Positive Rate (%)	False Negative Rate (%)
F3	C4	2	>7.02	11.09	100	90.98	0.02	0
F3	P4	2	>10.02	13.42	100	92.55	7.45	0
O2	F3	2	>11.03	10.30	100	90.29	9.71	0
C3	O1	17	<0.33	2.11	96	54.46	45.54	4.00
C3	O1	19	<0.27	2.25	100	55.62	44.38	0
F3	T6	2	>11.4	8.18	100	87.78	12.22	0
C3	O1	18	<0.32	2.12	100	53.34	46.66	0
F3	Pz	2	>9.41	9.51	100	89.48	10.52	0
O2	C3	2	>6.02	7.92	100	87.38	12.62	0
Fz	P4	2	>8.60	8.95	100	88.83	11.17	0
F3	P4	2	>10.02	19.65	100	94.91	5.09	0
C3	O1	19	<0.27					

Abbreviation: LR = likelihood ratio.
[a]Diagnostic performance of the 10 best single bipolar montages (according to their area under the curve) and best combination of montages (last row) for the discrimination of patients with delirium (DELIRIUM) from remainder patients of the whole database (CONTROL3).

these tools is challenged by recent meta-analyses reporting a pooled sensitivity of 80% and specificity of 95.9% that may even be lower in clinical routine settings.[21,23,41] A systematic review of the CAM revealed a false positive rate of up to 10% and that the interrater reliability could be as low as .81.[5] Despite these results, clinical tools remained the gold standard for the diagnosis of delirium over the past decades not least since alternative tools were not available. In this context, we provide evidence that the diagnostic performance of simple quantitative analyses of a limited number of EEG sensor combinations may outpace classic clinical tests. Even with an admittedly confounded and heterogeneous population due to the nature of our retrospective study design, we still reached a surprisingly high diagnostic accuracy. This leads us to believe that the diagnostic performance should be at least comparable in prospective controlled studies or even higher. Consequently, our results provide the intriguing perspective to use EEG for the standardized diagnosis and monitoring of delirium, particularly in situations when thorough repeated cognitive evaluations are impractical. Exemplary situations could be the presence of neurological disease such as aphasia and insufficient patient cooperation, level of consciousness or staffing. Particularly in patients with difficult to assess mental status, abbreviated EEGs including only most relevant electrode locations and minimum recording durations were suggested and could be particularly helpful since preparation of a complete EEG set may impair patient cooperativity and preclude meaningful conclusions.[42]

Patients Falsely Classified as Delirious From a Real-World Sample

We were able to identify all patients that were classified with delirium. However, it is of particular interest to elaborate on cases that were classified false positive in order to increase the specificity of our suggested approach. It is important to bear in mind that we included only patients for the classifier definition that were clearly delirious during the period of EEG acquisition. We can therefore not rule out that the remaining several thousand EEGs included data from other delirious patients. We thus elaborated on patients that were classified false positive since these might in fact have been delirious but not have fulfilled our inclusion criteria for the DELIRIUM group. Indeed, a thorough review of the medical charts revealed that 10 patients clearly had a clinical diagnosis of delirium. Most of the remaining false positive cases had a history of epileptic seizures, which is known to be a possible etiology of delirious episodes.[17] Also, others suffered from potential etiologies of delirium or from diseases that are closely linked to delirium (eg, sepsis, stroke, and dementia).[4,25,43] Unfortunately, we were retrospectively unable to tell if delirium was present in these patients or not. However, the percentage of patients classified as being delirious, including false positive cases, roughly concurs with the least expected number of delirious patients in an in-hospital setting, that is, about 5% (206 out of 4325 patients in the database).[44,45] While it is important not to overinterpret results, it is possible that a considerable amount of the false positive cases may in fact have been delirious. It is hence possible that the real-world performance of the classifier combination is even higher than stated.

Incorporation of qEEG in Clinical Routine

This retrospective study neither evaluated the implementation of qEEG in clinical routine nor did it compare the diagnostic accuracy between an expert EEG reader and qEEG or the combination of both. It is yet important to address how qEEG can be incorporated into clinical diagnostic pathways. First of all, there is currently no software available that can replace the interpretation of sEEG data by an experienced EEG reader as is implemented in current guidelines.[46] The current focus of qEEG research is therefore to assist clinicians with the interpretation

and to provide biomarkers for clinical and research purposes.[46,47] A prevailing issue is that qEEG requires cleaning data from artifacts, which cannot be done unsupervised. Semi-automatic rejection methods such as identification of outlier data by its variance and presentation of potentially confounded periods to the EEG reader for final judgement might be a viable compromise between time efficiency and validity.[29] Yet, even with perfect artifact removal, it remains to be examined in future studies to what extent diagnostic, therapeutic or prognostic conclusions can be drawn for various entities. This study revealed that at least in the context of delirium, qEEG may be of substantial help in inconclusive cases.

Limitations

The most obvious limitation of this study is its retrospective character. While this design allowed for the inclusion of a significant number of control subjects, it is also associated with typical shortcomings. These include lack of standardized recordkeeping and application of diagnostic criteria. It is hence possible that we underestimated the prevalence of delirium since we had to rely on certain keywords to search the database for delirious patients. Another limitation is that we cannot exclude confounding effects of psychotropic drugs on EEG activity. However, we tested for medication effects and did not find a significant difference in the EEG activity; neither did any other quantitative EEG study in delirium. A meta-analysis of psychotropic drugs' effects on EEG activity concluded that only 5% to 8% of cases were affected and that beta band activity would be increased, that is, affected in the opposite direction to changes reported in our study.[48] It is also possible that the fluctuating course of delirium severity may have caused routine EEG recordings to not be obtained when delirium was most severe.

Another important limitation of qEEG criteria is that their development critically depends on the underlying normative database.[49] This is a general issue in electrophysiological practice and current guidelines therefore suggest that laboratories create their own reference values.[50] While this might be viable for simple nerve conduction studies, this procedure is time consuming and methodologically challenging in qEEG since it must take sample adequacy and validity into account.[51] This also challenges the selection of our control, yet all EEG data were recorded by trained assistants that adhered to published guidelines for EEG recordings. We also visually screened every EEG for artifacts and removed confounding data as well as provocation maneuvers (eg, hyperventilation, photic stimulation). Hence, no relevant proportions of confounding data should be present in our normative data set.

Conclusions

This retrospective study provides preliminary evidence that EEG classifiers may outpace the diagnostic performance of clinical tests in delirium even outside confined study environments. Results furthermore indicate that a normal EEG excludes delirium and that suggested EEG criteria could substantially affect the pretest probability of being delirious. Prospective studies including parameters of pretest probability, delirium severity and cognitive outcome are required to elaborate on these promising findings.

Appendix

EEG records contained information about its clinical indication and the EEG reader's description and interpretation. Original (German) search terms for the identification of patients with a diagnosis of delirium within the hospital EEG database are enlisted below. Resulting patients were carefully hand-screened for eligibility for the DELIRIUM group.

Delir
verwirr*
bewusst* bewußt*
Encephalopathie Enzephalopathie
Encephalitis Enzephalitis
org. affektive Störung
org. Psychose
Wesensveränderung
Orientierungsstörung
kognitive Einbußen unklarer Ätiologie
Hirnfunktionsstörung
Desorientiertheit
Vigilanzminderung

Acknowledgments

We thank Juliette Wallace for English proofreading.

Author Contributions

RF and ST had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of data analysis. SS, RBP, and MR contributed to data analysis and interpretation. RF, SJS, and SAB were involved in clinical interpretation of data. RF and ST contributed equally to drafting of the manuscript. RBP, MR, SS, SJS, and SAB substantially contributed to revision of the manuscript. All authors approved the manuscript prior to submission.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Robert Fleischmann https://orcid.org/0000-0001-8159-2658

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, (DSM-5®)*. Washington, DC: American Psychiatric Association; 2013.

2. Elie M, Rousseau F, Cole M, Primeau F, McCusker J, Bellavance F. Prevalence and detection of delirium in elderly emergency department patients. *CMAJ*. 2000;163:977-981.
3. McNicoll L, Pisani MA, Zhang Y, Ely EW, Siegel MD, Inouye SK. Delirium in the intensive care unit: occurrence and clinical course in older patients. *J Am Geriatr Soc*. 2003;51:591-598.
4. Fong TG, Tulebaev SR, Inouye SK. Delirium in elderly adults: diagnosis, prevention and treatment. *Nat Rev Neurol*. 2009;5:210-220. doi:10.1038/nrneurol.2009.24.
5. Leslie DL, Inouye SK. The importance of delirium: economic and societal costs. *J Am Geriatr Soc*. 2011;59(suppl 2):S241-S243. doi:10.1111/j.1532-5415.2011.03671.x.
6. Leslie DL, Marcantonio ER, Zhang Y, Leo-Summers L, Inouye SK. One-year health care costs associated with delirium in the elderly population. *Arch Intern Med*. 2008;168:27-32. doi:10.1001/archinternmed.2007.4.
7. De J, Wand AP. Delirium screening: a systematic review of delirium sreening tools in hospitalized patients. *Gerontologist*. 2015;55:1079-1099. doi:10.1093/geront/gnv100.
8. Levkoff SE, Evans DA, Liptzin B, et al. Delirium. The occurrence and persistence of symptoms among elderly hospitalized patients. *Arch Intern Med*. 1992;152:334-340.
9. van Eijk MM, van Marum RJ, Klijn IA, de Wit N, Kesecioglu J, Slooter AJ. Comparison of delirium assessment tools in a mixed intensive care unit. *Crit Care Med*. 2009;37:1881-1885. doi:10.1097/CCM.0b013e3181a00118.
10. Ashley EA. The precision medicine initiative: a new national effort. *JAMA*. 2015;313:2119-2120. doi:10.1001/jama.2015.3595.
11. Luetz A, Balzer F, Radtke FM, et al. Delirium, sedation and analgesia in the intensive care unit: a multinational, two-part survey among intensivists. *PLoS One*. 2014;9:e110935. doi:10.1371/journal.pone.0110935.
12. Pandharipande P, Shintani A, Peterson J, et al. Lorazepam is an independent risk factor for transitioning to delirium in intensive care unit patients. *Anesthesiology*. 2006;104:21-26.
13. Alsop DC, Fearing MA, Johnson K, Sperling R, Fong TG, Inouye SK. The role of neuroimaging in elucidating delirium pathophysiology. *J Gerontol A Biol Sci Med Sci*. 2006;61:1287-1293.
14. Soiza RL, Sharma V, Ferguson K, Shenkin SD, Seymour DG, Maclullich AM. Neuroimaging studies of delirium: a systematic review. *J Psychosom Res*. 2008;65:239-248. doi:10.1016/j.jpsychores.2008.05.021.
15. Vijayakrishnan R, Ramasubramanian A, Dhand S. Utility of head CT scan for acute inpatient delirium. *Hosp Top*. 2015;93:9-12. doi:10.1080/00185868.2015.1012928.
16. van der Kooi AW, Zaal IJ, Klijn FA, et al. Delirium detection using EEG: what and how to measure. *Chest*. 2015;147:94-101. doi:10.1378/chest.13-3050.
17. Naeije G, Depondt C, Meeus C, Korpak K, Pepersack T, Legros B. EEG patterns compatible with nonconvulsive status epilepticus are common in elderly patients with delirium: a prospective study with continuous EEG monitoring. *Epilepsy Behav*. 2014;36:18-21. doi:10.1016/j.yebeh.2014.04.012.
18. Sidhu KS, Balon R, Ajluni V, Boutros NN. Standard EEG and the difficult-to-assess mental status. *Ann Clin Psychiatry*. 2009;21:103-108.
19. Hughes JR. A review of the usefulness of the standard EEG in psychiatry. *Clin Electroencephalogr*. 1996;27:35-39.
20. Jacobson SA, Leuchter AF, Walter DO. Conventional and quantitative EEG in the diagnosis of delirium among the elderly. *J Neurol Neurosurg Psychiatry*. 1993;56:153-158.
21. Gusmao-Flores D, Salluh JI, Chalhub RÁ, Quarantini LC. The confusion assessment method for the intensive care unit (CAM-ICU) and intensive care delirium screening checklist (ICDSC) for the diagnosis of delirium: a systematic review and meta-analysis of clinical studies. *Crit Care*. 2012;16:R115. doi:10.1186/cc11407.
22. Reade MC, Finfer S. Sedation and delirium in the intensive care unit. *New Engl J Med*. 2014;370:444-454. doi:10.1056/NEJMra1208705.
23. van Eijk MM, van den Boogaard M, van Marum RJ, et al. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. *Am J Respir Crit Care Med*. 2011;184:340-344. doi:10.1164/rccm.201101-0065OC.
24. Brenner RP. Utility of EEG in delirium: past views and current practice. *Int Psychogeriatr*. 1991;3:211-229.
25. Thomas C, Hestermann U, Kopitz J, et al. Serum anticholinergic activity and cerebral cholinergic dysfunction: an EEG study in frail elderly with and without delirium. *BMC Neurosci*. 2008;9:86. doi:10.1186/1471-2202-9-86.
26. van der Kooi AW, Leijten FS, van der Wekken RJ, Slooter AJ. What are the opportunities for EEG-based monitoring of delirium in the ICU? *J Neuropsychiatr Clin Neurosci*. 2012;24:472-477. doi:10.1176/appi.neuropsych.11110347.
27. Williams ST. Pathophysiology of encephalopathy and delirium. *J Clin Neurophysiol*. 2013;30:435-437. doi:10.1097/WNP.0b013e3182a73e04.
28. Noachtar S, Binnie C, Ebersole J, Mauguière F, Sakamoto A, Westmoreland B. A glossary of terms most commonly used by clinical electroencephalographers and proposal for the report form for the EEG findings. The International Federation of Clinical Neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl*. 1999;52:21-41.
29. Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*. 2011;2011:156869. doi:10.1155/2011/156869.
30. McGee S. Simplifying likelihood ratios. *J Gen Intern Med*. 2002;17:646-649.
31. Romano J, Engel GL. Delirium: I. Electroencephalographic data. *Arch Neurol Psychiatry*. 1944;51:356-377. doi:10.1001/archneurpsyc.1944.02290280054003.
32. Katz IR, Mossey J, Sussman N, et al. Bedside clinical and electrophysiological assessment: assessment of change in vulnerable patients. *Int Psychogeriatr*. 1991;3:289-300.
33. Matsushima E, Nakajima K, Moriya H, Matsuura M, Motomiya T, Kojima T. A psychophysiological study of the development of delirium in coronary care units. *Biol Psychiatry*. 1997;41:1211-1217.
34. Plaschke K, Hill H, Engelhardt R, et al. EEG changes and serum anticholinergic activity measured in patients with delirium in the intensive care unit. *Anaesthesia*. 2007;62:1217-1223. doi:10.1111/j.1365-2044.2007.05255.x.
35. Reischies FM, Neuhaus AH, Hansen ML, Mientus S, Mulert C, Gallinat J. Electrophysiological and neuropsychological analysis of a delirious state: the role of the anterior cingulate gyrus. *Psychiatry Res*. 2005;138:171-181. doi:10.1016/j.pscychresns.2004.06.008.
36. van der Kooi AW, Slooter AJ, van Het Klooster MA, Leijten FS. EEG in delirium: increased spectral variability and decreased complexity. *Clin Neurophysiol*. 2014;125:2137-2139. doi:10.1016/j.clinph.2014.02.010.

37. Koponen H, Partanen J, Pääkkönen A, Mattila E, Riekkinen PJ. EEG spectral analysis in delirium. *J Neurol Neurosurg Psychiatry*. 1989;52:980-985.
38. Chang YL, Tsai YF, Lin PJ, Chen MC, Liu CY. Prevalence and risk factors for postoperative delirium in a cardiovascular intensive care unit. *Am J Crit Care*. 2008;17:567-575.
39. de Lange E, Verhaak PF, van der Meer K. Prevalence, presentation and prognosis of delirium in older people in the population, at home and in long term care: a review. *Int J Geriatr Psychiatry*. 2013;28:127-134. doi:10.1002/gps.3814.
40. Giraud K, Vuylsteke A. Point-prevalence of delirium in intensive care units. *Anaesthesia*. 2014;69:394. doi:10.1111/anae.12649.
41. Reade MC, Eastwood GM, Peck L, Bellomo R, Baldwin I. Routine use of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) by bedside nurses may underdiagnose delirium. *Crit Care Resusc*. 2011;13:217-224.
42. Javanbakht A, Amirsadri A, Arfken C, Dewald O, Boutros NN. Standard EEG study of acute psychiatric patients with difficult to assess mental status. *Am Assoc Emerg Psychiatry*. 2012;10:1-5.
43. Polito A, Eischwald F, Maho AL, et al. Pattern of brain injury in the acute setting of human septic shock. *Crit Care*. 2013;17:R204. doi:10.1186/cc12899.
44. Francis J. Delirium in older patients. *J Am Geriatr Soc*. 1992;40:829-838.
45. Inouye SK, Rushing JT, Foreman MD, Palmer RM, Pompei P. Does delirium contribute to poor hospital outcomes? A three-site epidemiologic study. *J Gen Intern Med*. 1998;13:234-242.
46. Nuwer MR, Lehmann D, da Silva FL, Matsuoka S, Sutherling W, Vibert JF. IFCN guidelines for topographic and frequency analysis of EEGs and EPs.The International Federation of Clinical Neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl*. 1999;52:15-20.
47. Stephan KE, Iglesias S, Heinzle J, Diaconescu AO. Translational perspectives for computational neuroimaging. *Neuron*. 2015;87:716-732. doi:10.1016/j.neuron.2015.07.008.
48. Aiyer R, Novakovic V, Barkin RL. A systematic review on the impact of psychotropic drugs on electroencephalogram waveforms in psychiatry. *Postgrad Med*. 2016;128:656-664. doi:10.1080/00325481.2016.1218261.
49. Johnstone J, Gunkelman J. Use of databases in QEEG evaluation. *J Neurotherapy*. 2003;7:31-52.
50. Fuglsang-Frederiksen A, Pugdahl K. Current status on electrodiagnostic standards and guidelines in neuromuscular disorders. *Clin Neurophysiol*. 2011;122:440-455. doi:10.1016/j.clinph.2010.06.025.
51. Prichep LS. Use of normative databases and statistical methods in demonstrating clinical utility of QEEG: importance and cautions. *Clin EEG Neurosci*. 2005;36:82-87. doi:10.1177/155005940503600207.