Epilepsy & Behavior 34 (2014) 81–85

Contents lists available at ScienceDirect

Epilepsy & Behavior

journal homepage: www.elsevier.com/locate/yebeh

Diagnostic accuracy of microEEG: A miniature, wireless EEG device

Arthur C. Grant [a,b,*], Samah G. Abdel-Baki [c], Ahmet Omurtag [c], Richard Sinert [d], Geetha Chari [a], Schweta Malhotra [d], Jeremy Weedon [e], Andre A. Fenton [b,f], Shahriar Zehtabchi [d]

[a] Department of Neurology, SUNY Downstate Medical Center, Brooklyn, NY 11203, USA
[b] Department of Physiology and Pharmacology, SUNY Downstate Medical Center, Brooklyn, NY 11203, USA
[c] Bio-Signal Group Corporation, Brooklyn, NY 11226, USA
[d] Department of Emergency Medicine, SUNY Downstate Medical Center, Brooklyn, NY 11203, USA
[e] The Scientific Computing Center, SUNY Downstate Medical Center, Brooklyn, NY 11203, USA
[f] Center for Neural Science, New York University, New York, NY 10003, USA

ARTICLE INFO

Article history:
Received 31 October 2013
Revised 27 January 2014
Accepted 17 March 2014
Available online xxxx

Keywords:
Sensitivity and specificity
EEG
Altered mental status
Diagnostic accuracy

ABSTRACT

Measuring the diagnostic accuracy (DA) of an EEG device is unconventional and complicated by imperfect interrater reliability. We sought to compare the DA of a miniature, wireless, battery-powered EEG device ("microEEG") to a reference EEG machine in emergency department (ED) patients with altered mental status (AMS). Two hundred twenty-five ED patients with AMS underwent 3 EEGs. Two EEGs, EEG1 (Nicolet Monitor, "reference") and EEG2 (microEEG) were recorded simultaneously with EEG cup electrodes using a signal splitter. The remaining study, EEG3, was recorded with microEEG using an electrode cap immediately before or after EEG1/EEG2. The official EEG1 interpretation was considered the gold standard (EEG1-GS). EEG1, 2, and 3 were de-identified and blindly interpreted by two independent readers. A generalized mixed linear model was used to estimate the sensitivity and specificity of these interpretations relative to EEG1-GS and to compute a diagnostic odds ratio (DOR). Seventy-nine percent of EEG1-GS were abnormal. Neither the DOR nor the κ_f representing interrater reliabilities differed significantly between EEG1, EEG2, and EEG3. The mean setup time was 27 min for EEG1/EEG2 and 12 min for EEG3. The mean electrode impedance of EEG3 recordings was 12.6 kΩ (SD: 31.9 kΩ). The diagnostic accuracy of microEEG was comparable to that of the reference system and was not reduced when the EEG electrodes had high and unbalanced impedances. A common practice with many scientific instruments, measurement of EEG device DA provides an independent and quantitative assessment of device performance.

1. Introduction

Electroencephalography can provide useful clinical data. In the emergency department (ED), EEG can narrow the differential diagnosis of patients with altered mental status (AMS) and potentially help avoid unnecessary tests, procedures, and admissions. An EEG is necessary to diagnose nonconvulsive seizures and nonconvulsive status epilepticus, diagnoses that typically trigger emergent and specific treatment protocols. We previously found that the EEG is abnormal in 78% of ED patients with AMS, including 5% with electrographic seizures [1].

Although EEG is a useful, noninvasive, and relatively inexpensive diagnostic test, its routine use in the ED faces several obstacles. Hospital EEG laboratories are rarely open around the clock. Very few EDs are equipped with EEG machines or staffed with a technologist who can properly apply EEG electrodes and record a technically adequate study in an environment that is often electrically hostile. The long electrode wires and EEG machine may impede patient movement and limit access of medical personnel to the patient.

MicroEEG is a miniature, wireless, and battery-powered EEG device that can potentially overcome these limitations to routine use of EEG in the ED. Beyond engineering and safety testing required by the FDA, the device has undergone extensive analyses that were driven, in part, by skepticism among experienced emergency medicine physicians and neurologists that such a device would perform as intended in the ED environment. For instance, we previously reported technical and clinical analyses of microEEG recordings obtained from healthy volunteers and compared them to recordings made with a common commercially available system [2]. Those experiments demonstrated high concordance in the time and frequency domains between microEEG and the standard system. In a parallel recording made simultaneously with microEEG and the standard system, imperfect correlation between the signals was due primarily to relatively greater 60-Hz noise in the standard system [2].

This transvalue of microEEG continues with the present clinical trial designed to measure the diagnostic accuracy of microEEG in identifying abnormal EEG activity in ED patients with AMS. Measuring and

* Corresponding author at: SUNY Downstate Medical Center, Comprehensive Epilepsy Center, 450 Clarkson Ave., Box 1275, Brooklyn, NY 11203, USA. Tel.: +1 718 270 2959; fax: +1 718 270 4711.
E-mail address: arthur.grant@downstate.edu (A.C. Grant).

http://dx.doi.org/10.1016/j.yebeh.2014.03.015
1525-5050/

reporting the diagnostic accuracy of new scientific and medical instruments are common practice. However, it is rarely performed, if ever, for new EEG devices, possibly because the process is complicated by three apodictic facts. First, there is no gold standard for EEG interpretations. Second, interrater agreement on interpretation of complete EEG recordings is imperfect and poorly studied. Third, estimation of the new device's true accuracy must account for the accuracy of both the reference device and the EEG interpreters. These challenges led to two additional studies, the results of which are utilized in the present analyses [3,4].

2. Methods

2.1. Subjects

The study was conducted at Kings County Hospital Center (KCHC) and Downstate Medical Center (DMC), which together have >200,000 ED visits annually. All ED patients ≥13 years old with AMS were screened for study participation until the enrollment target was achieved. Altered mental status was defined as any alteration in consciousness, responsiveness, or alertness. Exclusion criteria included an immediately correctable cause of AMS (finger stick or serum glucose less than 60 mg/dl, body temperature below 35.0 °C, hyperthermia, heat exhaustion or heat stroke, and opioid overdose responding to naloxone), hemodynamic instability (SBP < 90 mm Hg), and inability to undergo an EEG recording (e.g., severe scalp injury and combativeness). All data were collected prospectively. The study was approved by the joint KCHC and DMC institutional review board, which waived the requirement for informed consent. Surrogate consent was obtained from a legally authorized representative when available. The study was registered on ClinicalTrials.gov (# NCT01355211).

2.2. EEG devices

MicroEEG measures 9.4 × 4.4 × 3.8 cm and weighs 88 g (Fig. 1). It digitizes the EEG signals close to the electrodes, transmits the digital data wirelessly to a personal computer located within 10 m, and stores the data on an onboard memory card. Custom software running on the PC controls the device, measures electrode impedances, displays the signals, allows entry of annotations, and writes data to the hard disk. The manufacturers' specifications for microEEG and the Nicolet Monitor (Natus Medical Inc., San Carlos, CA), which was used as the reference system, are shown in Table 1.

2.3. EEG studies

The EEG studies were obtained 24/7 by a pool of five technologists (each with at least 5 years of experience) and one physician with research and clinical EEG experience (SAB). We intended to examine microEEG performance both with standard EEG cup electrodes applied individually and with an electrode cap that can be quickly applied by minimally trained assistants in the ED. This goal led to the recording of three EEGs from each subject. Study EEG1 was recorded with the reference system, and EEG2 was recorded simultaneously with microEEG using a signal splitter [2]. Both EEG1 and EEG2 were obtained with 9-mm gold-plated cup electrodes placed according to the international 10–20 system, with all electrode impedances below 5 kΩ. Study EEG3 was recorded immediately before or after EEG1/EEG2 with microEEG using an electrode cap composed of an elastic spandex-type fabric with recessed pure tin 10–20 system electrodes (ECI, Inc., Eaton, OH). A small amount of electrogel was injected through a hole in the center of each electrode to minimize impedances, but there were no impedance restrictions for cap electrodes. The order of EEG1/EEG2 and EEG3 was randomized. All EEG recordings were 30 min in duration.

Fig. 1. MicroEEG attached to a headband.

Table 1
Technical specifications for microEEG and the Nicolet Monitor.

Product	MicroEEG	Nicolet Monitor
Company	Bio-Signal Group Corp.	CareFusion Corp.
A/D converter resolution	16	16
Voltage resolution (μV)	0.15	0.15
Maximum input range	10 mV p. to p.	10 mV p. to p.
Sampling rate	1000 Hz	Up to 2000 Hz
Bandwidth	0.15–500 Hz	0.16–1000 Hz
Input impedance (MΩ)	>100	>100
Separately test GND/REF impedance	Yes	No
Number of channels	32	64
Channel crosstalk	<−126 dB	<−40 dB
Noise	<2.4 μV p. to p. (0.1–100 Hz)	<1.5 μV p. to p. (0.1–70 Hz)

2.4. EEG interpreters and interpretations

The absence of a true gold standard and the presence of imperfect interrater agreement on EEG interpretations led to the following design and designations for the EEG interpretation process. Study EEG1 was the official recording that became part of each patient's medical record. Its final interpretation was considered the gold standard for this study and is designated EEG1-GS. Gold standard EEG interpretations were "unblinded" in that the interpreter was aware of patient age, medications, presenting symptoms, clinical history, and technologist comments within the EEG recording. The fact that the accuracy of EEG1-GS may be less than perfect is taken into account in our assessment of microEEG (see Section 2.7).

The three EEG studies were de-identified, including removal of technologist comments, and subject to blinded interpretation by a group of six epileptologists board-certified in clinical neurophysiology. The three de-identified EEGs from each subject were interpreted independently by two readers, unaware that they were interpreting all three studies from the same subject. Interpreters knew only that the EEGs were recorded from ED patients ≥13 years old with AMS. Interpreters assigned each study to one of six diagnostic categories: status epilepticus (SE), seizure (Sz), interictal epileptiform discharges (Ep, with or without independent slowing), slowing only (Sl), normal (Nl), and uninterpretable. Burst suppression and triphasic waves were categorized as "slowing." EEG1-GS interpretations were converted to the six category format by one of the authors (ACG).

2.5. Estimation of sensitivity and specificity of interpretation of de-identified EEG1, 2, and 3 relative to EEG1-GS

The diagnostic accuracy of microEEG was estimated in two conditions. In the first condition, the six EEG categories were condensed to two: abnormal and normal. Data from patients whose EEG1-GS assessment was "uninterpretable" were excluded from this analysis.

A generalized mixed linear model was constructed as follows: Let X_i be an indicator variable corresponding to the event that EEG1-GS of the ith patient is rated abnormal, and let Y_{ijk} indicate the event that the EEG of the ith patient in the jth "measurement mode" (i.e., de-identified EEG1, 2, or 3) is rated abnormal by the kth rater ($k = 1 \ldots 6$). Y_{ijk} was modeled as a Bernoulli random variable with $E(Y_{ijk}) = \pi$, such that $\ln[\pi / (1 - \pi)] = \beta_0 + \beta_1 X_i + a_i + b_k$, where a_i and b_k are random patient and rater effects, respectively, independently and normally distributed with mean $\mu_{a,b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and covariance $\sum_{a,b} = \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{bmatrix}$. This model allows for intercorrelation of outcomes among multiple observations of the same patient and also among multiple observations supplied by the same blinded rater. Kenward–Roger adjustments of standard errors and denominator degrees of freedom were applied. Asymptotic confidence intervals (CIs) were constructed for the quantities $\beta_0 + \beta_1$ and $-\beta_0$; inverse-link transformations were applied to the limits of these CIs to generate measurement mode-specific CIs for estimates of sensitivity Prob ($Y_{ij} = 1 \mid X_i = 1$) and specificity Prob ($Y_{ij} = 0 \mid X_i = 0$), i.e., $\exp(\beta_0 + \beta_1) / [1 + \exp(\beta_0 + \beta_1)]$ and $\exp(-\beta_0) / [1 + \exp(-\beta_0)]$, respectively.

A similar method was used to compute a CI for the diagnostic odds ratio (DOR), i.e., $\exp(\beta_1)$. The DOR is an expression of the utility of a test in a single parameter rather than a (sensitivity and specificity) pairing. It is interpretable as the odds ratio for an abnormal blinded result associated with the gold standard result's being abnormal vs. normal. For instance, if DOR = 2 for EEG1, then the odds of an abnormal EEG1 result from a randomly-selected blinded rater are twice as high if the unblinded gold standard rating is abnormal than if it is not.

In an initial analysis, β and $\sum$ parameters were estimated separately for each measurement mode. In a subsequent pooled analysis, a measurement mode indicator Z_j was introduced as a second fixed factor; a significance test for this term was used to detect the presence of differential bias (i.e., variability among the measurement modes in their propensity to categorize patients as abnormal); also, the existence of an $X_i Z_j$ interaction was tested to detect heterogeneity of EEG discriminability (i.e., DOR) across measurement modes.

In the ED setting, diagnosing status epilepticus and ongoing seizures is paramount, and, ideally, we would also determine the diagnostic accuracy of microEEG with EEG1-GS interpretations condensed to 1) SE + Sz and 2) all other. However, since the number of SE and Sz studies was very small, the resulting diagnostic parameters would be meaningless because of extremely large confidence intervals. Therefore, the second condition in which the diagnostic accuracy of microEEG was estimated was with the six EEG categories condensed to: 1) SE + Sz + Ep and 2) Sl + Nl (again excluding "uninterpretable" studies from the analysis). This grouping separated the studies into those that are relatively more significant in the ED setting and those that are relatively benign. The analyses were performed as described above, substituting SE + Sz + Ep for abnormal.

To estimate the probability of an uninterpretable EEG in each measurement mode, a single generalized mixed linear model was constructed similar to the ones used to estimate sensitivity and specificity. In this case, the event of interest was an uninterpretable result rather than an abnormal result, and the only fixed effect was measurement mode.

Interrater reliability was assessed separately for each measurement mode using all six EEG diagnostic categories including uninterpretable. The Fleiss kappa statistic is reported. A χ^2 test was conducted of heterogeneity of kappa values across measurement modes. Statistical Analysis System (SAS Institute, Cary, NC, USA) Release 9.2 statistical software was used for all analyses.

2.6. A priori sample size analysis

Using projections that 30% of EEG recordings would yield abnormal results and that the sensitivity of EEG2 and EEG3 would both be 95%, 243 patients would be required to achieve 95% CIs around sensitivity estimates having a width of 5% [5]. Allowing for a 5–10% rate of failure to collect adequate data yields approximately 260 subjects to be enrolled.

2.7. Procedure for correcting measured sensitivity and specificity

Ideally, the accuracy of microEEG would be determined by directly comparing its outcome with the patient disease state. With a sufficiently large patient population, this would have allowed its true sensitivity and specificity to be determined from a 2 × 2 frequency table. The typical, and, under our circumstances, best available method is to compare the interpretation of microEEG recordings with the interpretation of reference device recordings. This method yields the measured accuracy of microEEG and is confounded by two distinct effects. First, the standard system may not be perfectly accurate — meaning the degree to which it can generate a recording that would lead to a correct classification when interpreted by a perfectly accurate interpreter. Second, EEG interpretation accuracy is less than perfect. Omurtag and Fenton showed that imperfections in interpretation and in the reference system exert a downward bias on the measured sensitivity and specificity and that the set of pairs of interpreter sensitivity and specificity which correspond to a single value of kappa generates a range of corrected sensitivity and specificity values [4]. They derive a formula for systematically correcting measured sensitivity and specificity as a function of standard system accuracy, the prevalence of abnormal gold standard interpretations, and interrater agreement [4]. This correction method is employed in the Results section.

3. Results

3.1. Subjects

From May, 2011 through February, 2012, 302 patients were screened, and 261 were enrolled. Because of operator error, 36 EEG2 or EEG3 studies were not properly saved, leaving 225 subjects with all 3 EEGs available for review. Subject clinical data and their six category gold standard (EEG1-GS) findings are shown in Tables 2 and 3, respectively. There were no adverse events.

3.2. Diagnostic parameters for each measurement mode

Sensitivity, specificity, and diagnostic odds ratio estimated separately for each measurement mode are shown in Table 4. Point estimates of sensitivity and specificity in Table 4A resemble closely those obtained from simple 2 × 2 frequency tables. There were no significant differences among the raters in their propensity to identify cases as abnormal (EEG1: $p = 0.061$, EEG2: $p = 0.169$, and EEG3: $p = 1.000$).

Table 2
Subject characteristics.

Variable	n	%
Age	225	60[a]
Female	119	53
Presenting symptom		
Confusion	90	40
Lethargy	55	24
Coma/unresponsive	56	25
Agitation	29	13
Delirium	7	3
Unexpected psychosis	6	3
Seizure in the field	73	32
Seizure in the ED	53	24
AC medication in the field	17	8
AC medication in the ED	146	65

Abbreviations: AC = anticonvulsive.
[a] Median.

Table 3
Gold standard EEG findings.

Category	Frequency	Percent
Slowing	136	61
Normal	43	19
Epileptiform discharges	31	14
Status epilepticus	7	3
Uninterpretable	5	2
Seizure	3	1

In the pooled analysis, the test of homogeneity of the DOR across measurement modes yielded $p = 0.416$; we conclude that the three DOR estimates do not differ significantly from one another. After removing the $X_i Z_j$ interaction term from the model, there was no evidence of differential bias ($p = 0.449$); we conclude that the three modes do not differ significantly in their propensity to identify cases as abnormal. Model-based tests of differing sensitivity and specificity between EEG1 and EEG2 yielded $p = 0.305$ and 0.487, respectively; we conclude that these 2 modes do not differ significantly in either sensitivity or specificity.

The analyses described above for the data in Table 4A were repeated for the data in Table 4B, i.e., for the diagnostic parameters derived with EEGs grouped into SE + Sz + Ep and Sl + Nl, with identical qualitative results.

Fleiss kappas representing interrater reliabilities (with 95% CIs) for EEG1, 2, and 3 were 0.44 (0.35, 0.54), 0.47 (0.37, 0.56), and 0.47 (0.37, 0.56), respectively. A test of differences among these coefficients yielded $p = 0.842$; we conclude that interrater agreement does not differ significantly across measurement modes.

3.3. Corrected sensitivity and specificity

Corrected sensitivity and specificity for each measurement mode were also determined for the standard condition with EEG1-GS interpretations condensed to abnormal and normal (Table 4A). Briefly, since the accuracy of the standard EEG device is unknown, we arbitrarily assumed that its sensitivity (s0) and specificity (p0) were equal (i.e., s0 = p0). We then derived the range of corrected sensitivities and specificities for two arbitrary but illustrative values of s0 and p0, namely, 1.0 (i.e., the standard system was perfect or maximally accurate) and 0.9 (i.e., the standard system was "very good" but not perfect). These values are shown in Table 5.

3.4. EEG setup time, electrode impedances, and uninterpretable studies

The mean EEG setup time, i.e., time to initialize the EEG equipment, apply electrodes or the electrocap, and begin the EEG recording, was 27 min (SD: 2 min) for EEG1/EEG2 and 12 min (SD: 2 min) for EEG3. The mean impedance of all EEG electrodes from a random sample of 25 EEG3 recordings was 12.6 kΩ (SD: 31.9 kΩ) with a minimum of 1 kΩ and a maximum of 214 kΩ.

Since each de-identified EEG underwent two interpretations, there were 450 interpretations for each measurement mode. Of those 450 interpretations, 40 were "uninterpretable" for EEG1 studies (29 unique subjects), 30 were "uninterpretable" for EEG2 studies (22 unique subjects), and 51 were "uninterpretable" for EEG3 studies (32 unique subjects). There was a significant measurement mode effect ($p = 0.032$). Post hoc pairwise comparisons with bootstrap-adjusted p values showed that the estimated probability of an uninterpretable result differed significantly only between EEG2 and EEG3 ($p = 0.023$).

Table 4
Diagnostic parameters estimated separately for each measurement mode.

Measurement mode	Sensitivity (95% CI)	Specificity (95% CI)	DOR (95% CI)
A. EEGs grouped into abnormal and normal			
EEG1	0.93 (0.86, 0.96)	0.46 (0.29, 0.64)	10.7 (5.1, 22.2)
EEG2	0.90 (0.84, 0.94)	0.51 (0.35, 0.67)	9.8 (5.0, 19.2)
EEG3	0.91 (0.87, 0.94)	0.41 (0.28, 0.56)	6.7 (3.3, 13.8)
B. EEGs grouped into SE + Sz + Ep and Sl + Nl			
EEG1	0.65 (0.49, 0.79)	0.84 (0.74, 0.90)	9.7 (5.2, 18.0)
EEG2	0.62 (0.45, 0.77)	0.89 (0.81, 0.94)	9.8 (5.0, 19.2)
EEG3	0.55 (0.36, 0.72)	0.84 (0.73, 0.91)	6.5 (3.3, 12.5)

Abbreviations: DOR = diagnostic odds ratio, CI = confidence interval.

Table 5
Corrected measured sensitivity and specificity.

Measurement mode	Sensitivity			Specificity		
	Value	Min	Max	Value	Min	Max
A. Assuming a perfect reference system with S0 = p0 = 1.0						
EEG1	0.94	0.93	1	0.64	0.63	1
EEG2	0.91	0.90	0.96	0.69	0.52	0.69
EEG3	0.92	0.91	0.98	0.55	0.41	0.55
B. Assuming an imperfect reference system with s0 and p0 arbitrarily set to 0.9						
EEG1	0.97	0.95	1	0.85	0.84	1
EEG2	0.94	0.92	1	0.91	0.76	0.91
EEG3	0.94	0.93	1	0.72	0.61	0.73

s0: sensitivity of reference system, p0: specificity of reference system.
Min and max refer to the minimum and maximum values obtained by varying interpreter sensitivity and specificity among all possible values for the measured Fleiss kappas of 0.44, 0.47, and 0.47 for EEG1, 2, and 3, respectively.

3.5. Sample size reconsidered

Although the study is slightly underpowered relative to initial projections, the prevalence of abnormality was considerably higher than projected. Thus, confidence intervals for sensitivity estimates are at least as narrow as originally desired.

4. Discussion

We performed a prospective clinical trial of microEEG to measure its sensitivity and specificity compared to a "reference" EEG machine, the Nicolet Monitor. The EEG recordings were acquired from ED patients with AMS, a population where the advantages of a miniature, wireless, and battery-powered EEG device are particularly valuable. Using the unblinded interpretation (abnormal or normal) of EEGs recorded from the reference system as a gold standard, the measured sensitivity and specificity of EEG1 and EEG2 did not differ significantly in either measurement condition, i.e., with EEGs divided into abnormal and normal and with EEGs divided into SE + Sz + Ep and Sl + Nl. Thus, the diagnostic accuracy of microEEG is comparable to that of the reference system in this clinical setting.

Assuming a perfect reference system and accounting for the Fleiss kappa values for interrater agreement, the corrected specificities of each measurement mode were substantially higher than the measured values but remained significantly lower than the corrected sensitivities (Table 5). These differences in sensitivity and specificity are expected for at least two reasons. First, the prevalence of abnormal EEG1-GS studies was much higher than the prevalence of normal studies. Second, equivocal EEG findings such as mild diffuse slowing are more likely to be interpreted as abnormal when the interpreter is blind to patient age, state, history, and medications.

The EEG3 mean setup time of 12 min was considerably lower than the 27 min for EEG1/EEG2. These time savings may be especially valuable in the ED, where there is high demand for access to the patient and the expectation that studies will be completed expeditiously. As anticipated, electrode cap impedances and interelectrode impedance differences were substantially higher than longstanding professional society recommendations [6,7]. Both of these variables contribute independently to increased common mode noise, which can result in an uninterpretable study [7]. Nonetheless, the probability that a randomly selected interpretation would be "uninterpretable" (i.e., due to artifact)

did not differ significantly between EEG3 and EEG1. The most likely reason that more EEG3 studies were not obscured by 60-Hz noise is the relatively short electrode wires when microEEG is used with an electrode cap [2]. In that configuration, the maximum electrode wire length was 45 cm compared to the 140 cm length of standard electrode wires.

Because EEG3 was recorded immediately before or after EEG1/EEG2, it was a priori a different study. Most EEG3 findings should be identical to those of the corresponding EEG1/EEG2, but different findings could occur if the patient were undergoing simultaneous treatment for seizures or status epilepticus, seizing intermittently, or had other intermittent findings such as interictal epileptiform discharges. The fact that EEG3 was a different study could account for its lower DOR compared to EEG1 and EEG2 (Table 4), although the differences were not statistically significant.

An inevitable limitation of comparing devices whose output is subject to human interpretation is the imperfect nature of the interpretive process. Aware of this limitation, we performed a separate study of EEG intra- and interrater reliability [3]. Briefly, a pool of six epileptologists interpreted 300 EEGs in such a way as to generate both intra- and interrater reliability data, as well as other variables of interest. The Cohen kappas for 15 reader pairs ranged from 0.29 to 0.62, with an aggregated Fleiss kappa of 0.44. This value is very similar to the kappas obtained in this study and to results from other studies [8–11].

That the diagnostic accuracy of microEEG is not compromised by the use of an electrode cap suggests that EEGs adequate for the ED setting can be recorded by personnel other than trained EEG technologists. That hypothesis was tested in a separate study in which ED patients with AMS were randomized to receive standard care or standard care plus a microEEG recorded with an electrode cap by personnel without formal EEG technologist training. Outcome variables included the impact of microEEG data on the differential diagnosis, patient management, and patient outcome [12].

5. Conclusions

The diagnostic accuracy of a new EEG device compared to a standard or reference system can be measured in a clinical setting. For this measurement to be meaningful, the methodology must account for the absence of definitive “gold standard” EEG interpretations, imperfect interrater agreement on EEG interpretations, and the theoretical possibility that the reference device is imperfect. Using such methodology, we showed that the diagnostic accuracy of microEEG is comparable to that of a Nicolet Monitor in ED patients ≥13 years old with AMS, a population with a very high prevalence of EEG abnormalities. Setup time for EEG recordings was shortened by 56% when they were obtained with an electrode cap rather than individual electrodes without compromising diagnostic accuracy.

Acknowledgments

This study was supported by NIH grant 1RC3NS070658 to Bio-Signal Group (BSG) Corp. with a subcontract to SUNY Downstate Medical Center. The authors thank Dr. Ewa Koziorynska, Dr. Douglas Maus, Dr. Tresa McSween, Dr. Katherine Mortati, and Dr. Alexandra Reznikov for interpreting the EEG studies. The authors also acknowledge the assistance of Saroj Kunnakkat, Vanessa Arnedo, and Madeleine Coleman with manuscript preparation.

Conflict of interest statement

Dr. Grant serves on the Bio-Signal Group (BSG), Corp. advisory board. All income derived from this position is donated directly from BSG to the Downstate College of Medicine Foundation.

Dr. Abdel Baki is an employee of BSG, owns stock options in BSG, and is a coinventor on US patents pending 61/554,743 and 13/284,886.

Dr. Omurtag was previously an employee of BSG and is a coinventor on US patents pending 61/554,743 and 13/284,886.

Dr. Fenton is a shareholding founder and president of BSG; is the inventor on US patent 7767195 and a coinventor on US patents pending 13/61,554,743, 13/284,886, 11/694,855, 11/694,816, 10/314,890, and 10/425,023; is an Associate Editor at the Journal of Neuroscience; and has had travel paid for by BSG.

The remaining authors have no conflicts of interest.

References

[1] Zehtabchi S, Abdel-Baki S, Omurtag A, Sinert R, Chari G, Malhotra S, et al. Prevalence of non-convulsive seizure and other electroencephalographic abnormalities in emergency department patients with altered mental status. Am J Emerg Med 2013;31:1578–82.
[2] Omurtag A, Abdel-Baki SG, Chari G, Cracco R, Zehtabchi S, Fenton A, et al. Technical and clinical analysis of microEEG: a miniature wireless EEG device designed to record high-quality EEG in the emergency department. Int J Emerg Med 2012;5:35.
[3] Grant AC, Abdel-Baki SG, Weedon J, Arnedo V, Chari G, Koziorynska E, et al. EEG interpretation reliability and interpreter confidence: A large single-center study. Epilepsy Behav 2014;32:102–7.
[4] Omurtag A, Fenton AA. Assessing diagnostic tests: how to correct for the combined effects of interpretation and reference standard. PLoS One 2012;7.
[5] Privitera M, Hoffman M, Moore JL, Jester D. EEG detection of nontonic–clonic status epilepticus in patients with altered consciousness. Epilepsy Res 1994;18:155–66.
[6] Epstein C, Bej M, Foldvary-Schaefer N, Langerlund T, Helmers S, Husain A. Guideline 1: minimum technical requirements for performing clinical electroencephalography. J Clin Neurophysiol 2006;23:86–91.
[7] Legatt AD. Impairment of common mode rejection by mismatched electrode impedances: quantitative analysis. Am J EEG Technol 1995;35:296–302.
[8] Azuma H, Hori S, Nakanishi M, Fujimoto S, Ichikawa N, Furukawa TA. An intervention to improve the interrater reliability of clinical EEG interpretations. Psychiatry Clin Neurosci 2003;57:485–9.
[9] Gerber PA, Chapman KE, Chung SS, Drees C, Maganti RK, Ng Y-T, et al. Interobserver agreement in the interpretation of EEG patterns in critically ill adults. J Clin Neurophysiol 2008;25:241–9.
[10] Piccinelli P, Viri M, Zucca C, Borgatti R, Romeo A, Giordano L, et al. Inter-rater reliability of the EEG reading in patients with childhood idiopathic epilepsy. Epilepsy Res 2005;66:195–8.
[11] Stroink H, Schimsheimer RJ, de Weerd AW, Geerts AT, Arts WF, Peeters EA, et al. Interobserver reliability of visual interpretation of electroencephalograms in children with newly diagnosed seizures. Dev Med Child Neurol 2006;48:374–7.
[12] Zehtabchi S, Abdel Baki SG, Omurtag A, Sinert R, Chari G, Roodsari GS, et al. Effect of microEEG on clinical management and outcomes of emergency department patients with altered mental status: a randomized controlled trial. Acad Emerg Med 2014;21:283–91.