ORIGINAL

Timothy S. Walsh
Pamela Ramsay
T. Petteri Lapinlampi
Mika O. K. Särkelä
Hanna E. Viertiö-Oja
Pekka T. Meriläinen

An assessment of the validity of spectral entropy as a measure of sedation state in mechanically ventilated critically ill patients

T. S. Walsh (✉) · P. Ramsay
Edinburgh Royal Infirmary,
51 Little France Crescent, Old Dalkeith
Road, EH16 4SA Edinburgh, Scotland
e-mail: twalsh@staffmail.ed.ac.uk
Tel.: +44-131-2421186
Fax: +44-131-2423138

T. P. Lapinlampi · M. O. K. Särkelä ·
H. E. Viertiö-Oja · P. T. Meriläinen
GE Healthcare Finland Oy,
Kuortaneenkatu 2, 00510 Helsinki, Finland

P. T. Meriläinen
Helsinki University of Technology,
Otakaari 1, P.O. Box 1000, 02015 TKK
Espoo, Finland

Abstract *Objective:* To assess whether the Entropy Module (GE Healthcare, Helsinki, Finland), a device to measure hypnosis in anesthesia, is a valid measure of sedation state in critically ill patients by comparing clinically assessed sedation state with Spectral Entropy *Design:* Prospective observational study. *Setting:* Teaching hospital general ICU. *Patients and participants:* 30 intubated, mechanically ventilated patients without primary neurological diagnoses or drug overdose receiving continuous sedation. *Interventions:* Monitoring of EEG and fEMG activity via forehead electrodes for up to 72 h and assessments of conscious level using a modified Ramsay Sedation Scale. *Measurements and results:* 475 trained observer assessments were made and compared with concurrent Entropy numbers. Median State (SE) and Response (RE) Entropy values decreased as Ramsay score increased, but wide variation occurred, especially in Ramsay 4–6 categories. Discrimination between different sedation scores [mean (SEM) P_K value: RE 0.713 (0.019); SE 0.710 (0.019)] and between lighter (Ramsay 1–3) vs. deeper (Ramsay 4–6) sedation ranges was inadequate [P_K: RE 0.750 (0.025); SE 0.748 (0.025)]. fEMG power decreased with increasing Ramsay score but was often significant even at Ramsay 4–6 states. Frequent "on–off" effects occurred for both RE and SE, which were associated with fEMG activity. Values switched from low to high values even in deeply sedated patients. High Entropy values during deeper sedation were strongly associated with simultaneous high relative fEMG powers. *Conclusions:* Entropy of the frontal EEG does not discriminate sedation state adequately for clinical use in ICU patients. Facial EMG is a major confounder in clinical sedation ranges.

Keywords Sedation monitoring · Electroencephalogram · Critical illness · Spectral entropy · Sedatives

Introduction

Sedation is integral to the management of critically ill patients requiring mechanical ventilation. It is currently provided by either infusions or intermittent boluses of drugs with hypnotic and analgesic properties [1]. Over-sedation is associated with adverse outcomes during critical illness, and clinical strategies designed to avoid this can decrease ventilation times and intensive care unit (ICU) length of stay [2, 3]. These strategies may also improve clinical outcomes, decrease some complications, such as ventilator-associated pneumonia, and decrease illness cost [4].

At present, sedation is usually managed by clinical assessments, often in conjunction with a protocol for adjusting drug doses. This approach has been shown to decrease ventilation times and the costs of sedative drugs and is considered best practice in recent guidelines [5]. Several clinical scales have been proposed and validated

for assessing sedation level reliably; these are integral parts of most sedation protocols [6–8]. One problem with these scales is that they lack discriminative power in the over-sedated range. This means that clinical assessments lack the ability to distinguish between patients who are heavily sedated, but would regain consciousness rapidly after sedation withdrawal, from those who would have delayed recovery in consciousness because of drug accumulation or coma/encephalopathy. This is clinically relevant because the patient in whom deeper levels of sedation are frequently maintained in order to facilitate more invasive or uncomfortable treatments are often sicker. Under these circumstances sedative drug pharmacokinetics and dynamics are less predictable.

A monitoring system that could detect over-sedation in critically ill patients would be useful, particularly if it could discriminate different degrees of deep sedation and encourage staff to consider decreasing drug doses [9]. Several systems based on electroencephalogram (EEG) analysis via forehead electrodes have been developed as depth-of-anesthesia monitors. The Bispectral Index (BIS, Aspect Medical Systems, Newton, Mass.) is widely used for anesthesia. Although promoted as an intensive care monitor, published literature has contrasting conclusions regarding its value in this setting [10]. Spectral Entropy measurement (Entropy™ Module, GE Healthcare, Helsinki, Finland) was developed for anesthesia monitoring and has not been evaluated in intensive care [11]. The module can be used to record raw EEG signals as well as Entropy values, allowing off-line examination of data.

The primary aim of this study was to assess whether Entropy, a measure of hypnosis in anesthetized patients, could be used to measure sedation status in general intensive care patients during routine clinical management by comparing values with a clinical sedation scale.

Patients and methods

We studied 30 patients admitted to a single teaching hospital ICU over a treatment period of up to 72 h. Inclusion criteria were a requirement for mechanical ventilation, sedation using continuous infusions of either midazolam or propofol and concomitant analgesic drugs if clinically indicated, and informed consent from relatives. Exclusion criteria were: (a) a patient in whom brain injury, namely hypoxic brain injury, traumatic brain injury, or intracranial hemorrhage, were considered present at the time of enrollment to the study; (b) drug overdose as admission diagnosis; (c) a patient requiring neuromuscular paralysis at the time of screening for the study; or (d) status epilepticus. We also excluded patients known to be clinically deaf or who had chronic neuromuscular disorders or brain disease that might interfere with normal clinical sedation assessment. The ethics committee and local institution gave permission to carry out the study.

Study design and protocol

We performed a non-interventional prospective cohort study. A standard disposable GE Datex-Ohmeda Entropy sensor (GE Healthcare, Helsinki, Finland) was applied to the forehead symmetrically relative to the midline. Each sensor comprised a strip that included one electrode each for the left and right hemispheres, and one central ground electrode. Sensors were changed every 24 h. The monitor performed an automatic impedance test every 10 min to ensure electrical contact fidelity. Periods of poor electrode contact (impedance > 5 kOhm) were rejected from the analyses.

The sensors were connected to a GE Datex-Ohmeda S/5 Compact anesthesia monitor, provided with the Entropy Module (GE Healthcare, Helsinki, Finland). The Entropy Module calculates two different Entropy parameters. State Entropy (SE) is derived over the frequency range 0.8–32 Hz; it includes most of the EEG power and in an anesthetized patient primarily reflects activity of cortical neurons, whereas contribution from facial muscle activity is small. Response Entropy (RE) is derived over the frequency range 0.8–47 Hz; it includes in addition to the cortical component a significant contribution from muscle activity that dominates the high-frequency part of the spectrum. An Entropy number ranging from 0 to 100 is displayed for each parameter, with 0 indicating complete suppression of cortical activity and 100 the normal awake state. For anesthesia a value in the 40–60 range is associated with a low probability of consciousness. The data recording included raw EEG/EMG waveforms with a 400-Hz sampling frequency plus the SE and RE parameters. All data were recorded by a dedicated laptop computer equipped with S/5 Collect software (GE Healthcare, Helsinki, Finland). A detailed description of the entropy algorithm has been published previously by Viertiö-Oja and colleagues [11].

Once data recording had started, it continued until one of the following end points was reached: (a) patient regained consciousness and mechanical ventilation was discontinued; (b) 72 h had elapsed; (c) the patient or a relative requested discontinuation of the monitoring and/or withdrawal from the study; or (d) the patient had died.

Observations and management of patients during the protocol

Patients received routine clinical management throughout the study period determined by caring clinicians. Most patients were sedated with propofol as first-choice sedative. Some patients received midazolam. Sedatives were administered by continuous infusion with additional boluses as considered clinically appropriate. Analgesia was provided with alfentanil or morphine infusions. The choice of sedative and analgesic drugs, and the doses prescribed, were

determined by medical and nursing staff and were not controlled for the purpose of the study. A sedation score assessment was carried out up to every 30 min using a modified Ramsay scoring system (Table 1). The published scoring system was modified to standardize the stimuli, specifically by including a tetanic stimulus at deep sedation levels. Each Ramsay score was compared with the median Entropy numbers for the 1-min period preceding each assessment. This approach was used to avoid the potential confounding effect of stimulation during the Ramsay assessment. All Ramsay assessments included in the analysis were carried out by a single member of the research team (P.R.), who was blinded to the Entropy numbers. In addition, a laptop-based notation file was used to record all events during periods of observation. We aimed to achieve a mean of 15–20 single observer assessments per patient, which would generate a total of 450–600 paired data for analysis.

Analysis

Assessment of criterion validity

To assess the criterion validity of entropy we examined the correlation between SE, RE, and clinical sedation score as reference standard. We calculated the distribution of SE and RE values corresponding to each modified Ramsay level and presented the results graphically by box-and-whisker plots. The performance of the Entropy parameters in indicating the depth of sedation as given by the Ramsay score was tested with the prediction probability, P_K, which is a variant of Kim's measure of association [12]. All prediction probabilities and their standard errors were estimated with the jackknife method as described by Smith [12] et al. With this approach a P_K value of 0.5 indicates no predictive ability compared with the reference (in this case clinical sedation level), and a P_K value of 1 indicates perfect prediction. The calculations were performed with Excel software (Microsoft, Redmond, Wash.) using a custom spreadsheet macro, PKMACRO, developed by Smith and colleagues [12]. We calculated P_K values for the ability of SE and RE to

distinguish each Ramsay score category from the other categories.

Assessment of construct validity

We used two approaches to assess construct validity. Firstly, as we were primarily interested in the ability of Entropy to distinguish "lighter" sedation ranges from "deeper" sedation states we calculated the P_K value for SE and RE for discriminating patients in Ramsay range 1–3 from those in Ramsay range 4–6. Secondly, we plotted continuous entropy data and compared it with intermittent sedation scores and our annotation files. In this analysis we also examined the relationship between entropy values and facial EMG power, which was recalculated from the data files. This was done from the sum of the frequency components between 55 and 145 Hz, excluding the 100-Hz mains frequency multiple.

Comparisons between different Ramsay levels were carried out using a Kruskal–Wallis test for non-parametric data. If significant, a Mann–Whitney U test was used to test inter-group differences. The significance level was set at 5%.

Results

Characteristics of the 30 patients are shown in Table 2. A total of 1200 h of EEG/EMG monitoring were carried out. The median duration of monitoring per patent was 40 h (first quartile 26 h, third quartile 66 h; range 2–76 h). The single assessor made a total of 475 sedation assessments during the monitoring period; of these, 59 were rejected from the analysis due to poor Entropy data quality at the time of the assessment, leaving 416 assessments that were used in the analyses.

Table 1 The modified Ramsay Sedation Assessment Scale used in the study

Score	Clinical state or response to stimulation
1	Patient anxious and/or agitated and/or restless
2	Spontaneous eye openings
3	No spontaneous eye opening, response to vocal stimulus
4	No response to vocal stimulus, response to loud stimulus
5	No response to loud stimulus, response to tetanic (50 Hz, 40 mA, 0.25 s pulses, duration 4 s) stimulus
6	No response to tetanic stimulus

Table 2 Characteristics of the patients studied

Age (years) Mean (SD; range)	59 (17; 17–82)
Male, *n* (%)	22 (73)
Duration of ICU admission prior to study (hours). Median (range)	88 (2–300)
SOFA score at start of study Median (quartiles; range)	6 (5, 8; 3–11)
Diagnosis at ICU admission	
Pneumonia	7
Bowel perforation or infarction	6
Major vascular surgery	4
Septic shock	3
Trauma	3
Pancreatitis	2
Gastrointestinal bleeding	2
Liver transplant	1
Cardiac failure	1
Fulminant hepatic failure	1

Fig. 1 Box-and-whisker plots
show the State (*top panel*) and
Response (*bottom panel*)
Entropy (*SE, RE*) values that
were observed at different
Ramsay sedation scores using
data pooled from all patients
observed during the study.
Boxes indicate interquartile
range; *horizontal line within
box* indicates median value.
The number of observations
on which the plots are based
is shown in Table 2.
* *p* < 0.05 for adjacent groups
(Mann–Whitney U test)
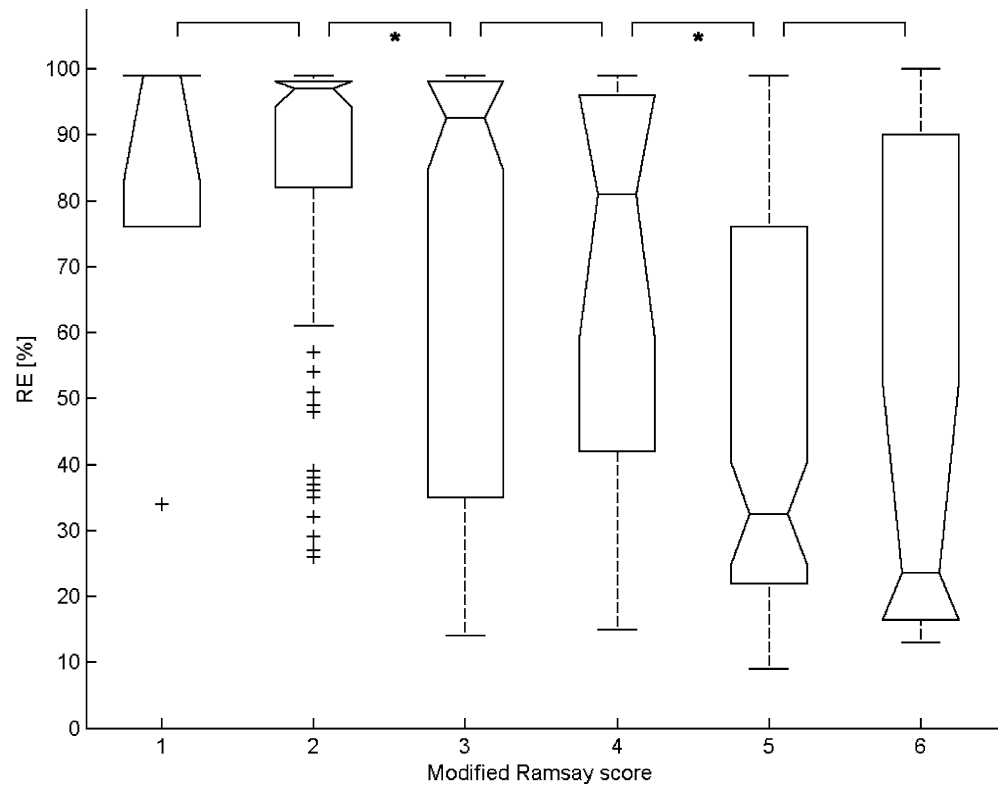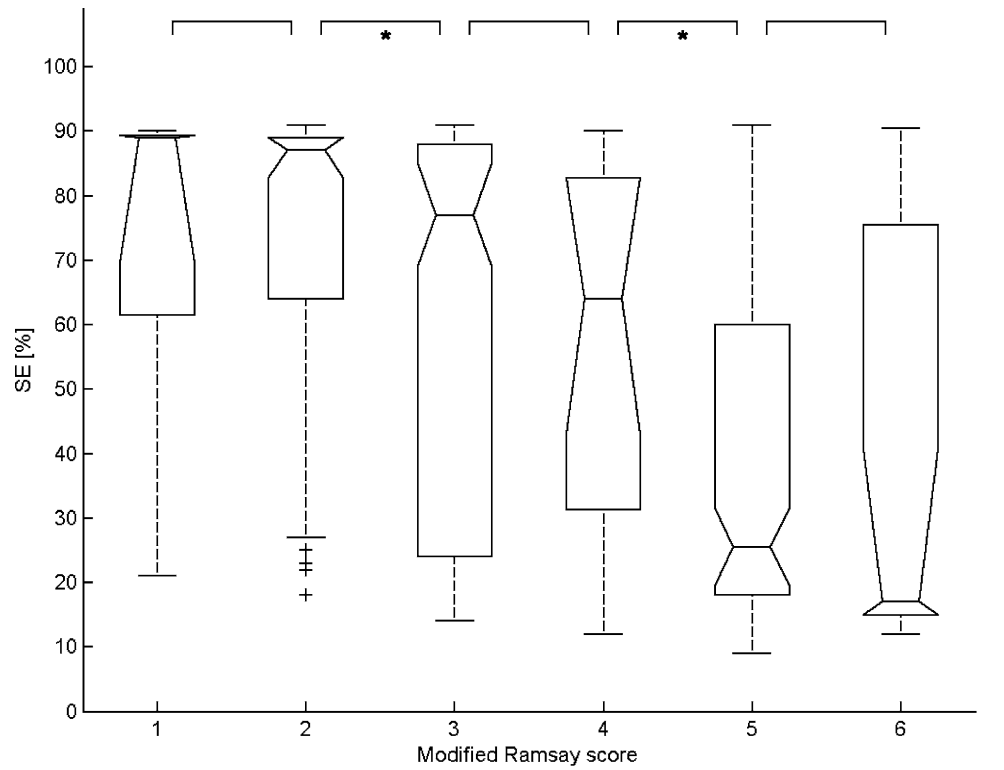
The median (first, third quartile; range) number of trained observer sedation assessments per patient was 14.5 (8, 21.5; 3–32). The number of observations made across Ramsay scores 1–6 were 5, 90, 168, 16, 119, and 18, respectively. This broadly categorized patients into "deeper" sedation state (Ramsay 4–6, $n = 153$ assessments) and "lighter" sedation state (Ramsay 1–3, $n = 263$ assessments).

Assessment of criterion validity

The values for SE and RE in relation to Ramsay score are shown in Fig. 1. Although median values did decrease as Ramsay scores progressed from 1 to 6 ($p < 0.05$ between levels 2–3 and 4–5), there was a wide range in values for each category, particularly for the Ramsay 3–6 range. These ranged from values suggesting deep anaesthesia (< 40) to values suggesting very light sedation or normal consciousness (> 80) even for the Ramsay 5–6 patients, who were non-responsive. The mean (SEM) P_K value of RE and SE for discriminating each sedation level from all other levels were 0.713 (0.019) and 0.710 (0.019), respectively. Although these values indicated some predictive power (value > 0.5), criterion validity was inadequate as a clinically useful test to distinguish each clinical sedation level from other levels.

Construct validity

The mean (SEM) P_K value of RE and SE for discriminating patients in "lighter" (Ramsay 1–3) from "deeper" (Ramsay 4–6) sedation states were 0.750 (0.025) and 0.748 (0.025), respectively. This suggested inadequate construct validity to distinguish lighter from deeper sedation.

We observed a clear pattern on visual inspection of the data. There were frequent "on–off" effects in entropy number where values changed rapidly from low to high levels, and vice versa. This was particularly noticeable during deeper sedation states (Fig. 2). During lighter sedation states entropy numbers tended to be consistently very high (Fig. 3). When we examined the fEMG power corresponding to the on–off effect, we found that fEMG activations corresponded to switches to high entropy numbers. The fEMG power varied within each Ramsay score category; high levels of fEMG power were observed during some periods even in over-sedated patients (Ramsay 5–6). The distribution of fEMG power calculated for each Ramsay sedation assessment is shown in Fig. 4. There was a correlation between fEMG power and Ramsay score. The P_K value for fEMG for distinguishing each sedation score was similar to the entropy numbers [0.711 (0.019)]. The P_K for distinguishing Ramsay 1–3 from 4–6 was higher than for the entropy numbers 0.764 (0.024).


Fig. 2 An individual patient plot illustrates the relation between the Entropy values and facial EMG power (*fEMG power*) when the clinical sedation score indicated that the patient had sluggish responses to stimuli. The Ramsay score was 5 throughout the recording (*bottom panel*). Despite this, clear periods of fEMG activation occurred manifest as acute increases in fEMG power (*middle panel*). Some of these responses were spontaneous and some in response to clinical stimuli (evident from the detailed annotation files). During any fEMG response both State and Response Entropy numbers increased to very high "awake" levels despite no change in Ramsay score (*top panel*). This created an "on–off" effect in Entropy that correlated with fEMG activations
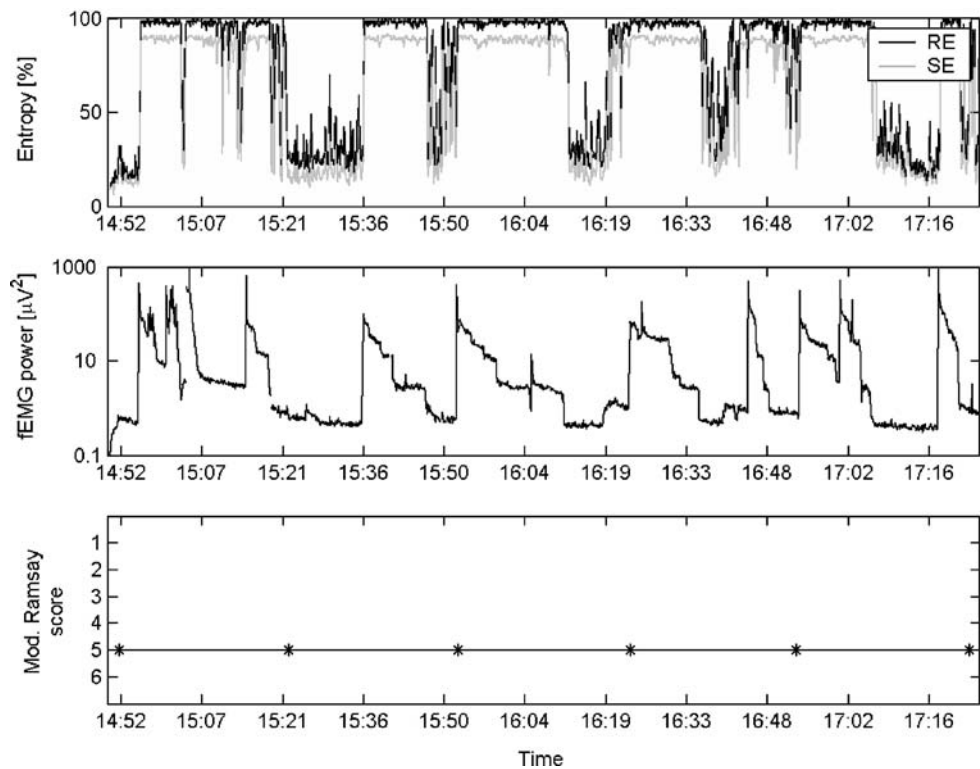
Fig. 3 An individual patient plot illustrates the relation between the Entropy values and facial EMG power (*fEMG power*) when a patient was emerging from sedation as drugs were reduced. The Ramsay score increased indicating transition from sluggish response to stimuli (level 5) to a cooperative, awake, and tranquil patient (level 2) (*bottom panel*). The fEMG power was high throughout the period of observation and increased as sedation level progressed from 5 through to 2 (*middle panel*). Both State and Response Entropy numbers (*SE, RE, top panel*) had a frequent "on–off" effect during deeper sedation, and subsequently a persistent very high value throughout emergence. The lack of discrimination for Entropy numbers between sedation levels 5 to 2 is clear
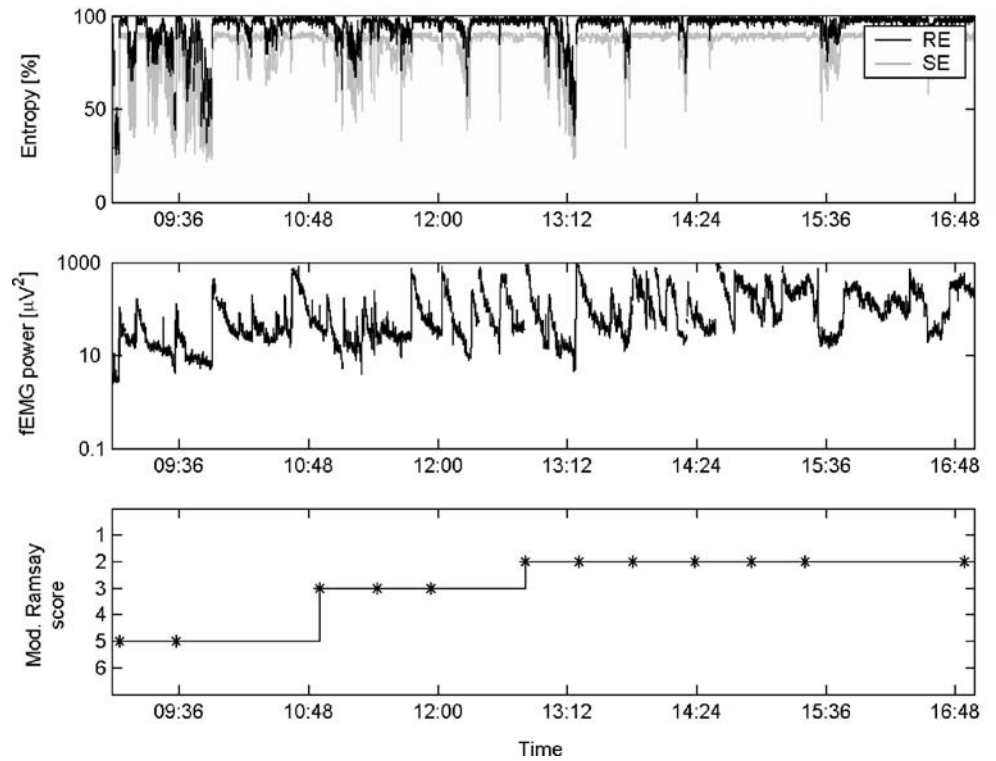
Fig. 4 Box-and-whisker plot shows the facial EMG power (*fEMG power*) observed at each Ramsay sedation assessment level. * $p < 0.05$ for adjacent groups (Mann–Whitney U test)

314
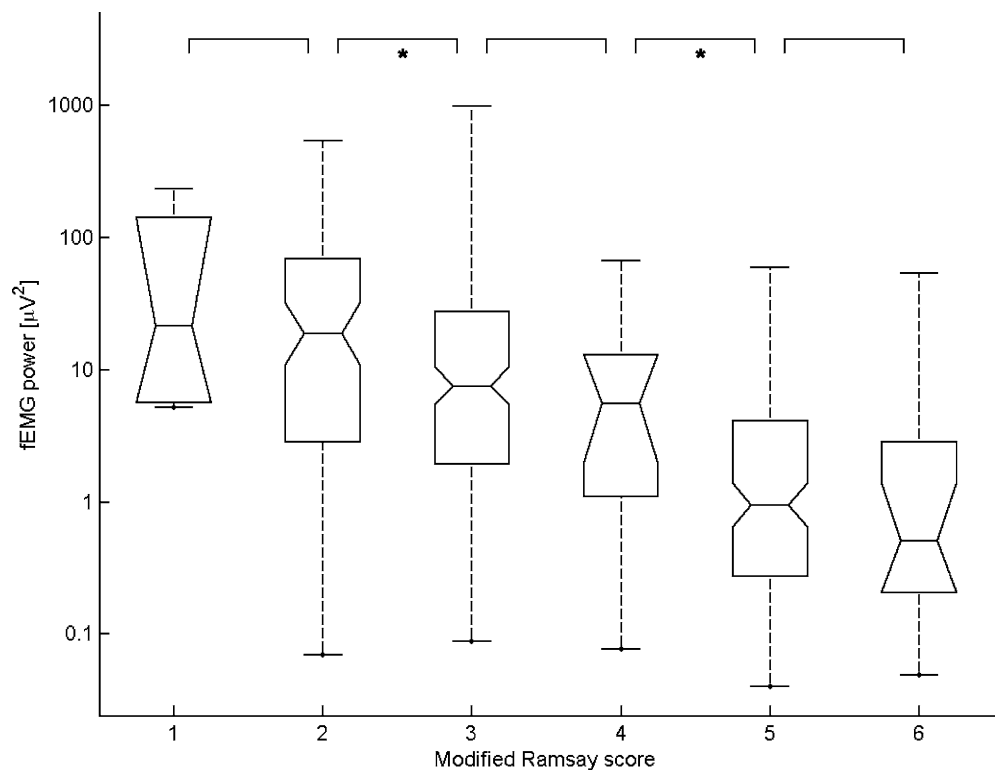
Discussion

We have shown that State and Response Entropy values do not discriminate well between different clinical sedation levels in non-paralyzed intensive care patients. Although median Entropy values decreased in association with increasing sedation level, the variation in observed values was high. These observations suggest that entropy has inadequate validity as a measure of sedation state under routine clinical conditions. Facial EMG activation is a plausible explanation for the poor validity; it was observed across all clinical sedation ranges and often created an "on–off" effect in Entropy numbers, particularly in more sedated patients.

We studied patients during routine clinical care at various stages of their illness and with a range of conditions. Our findings are likely to be relevant to most non-neurological ICU patients. We obtained a large number of data points based on observations by a single trained observer. Inter- and intra-rater variability were therefore minimized as a potential source of bias. We observed relatively few under-sedated patients (Ramsay score 1) but do not consider this important because monitors are not required in this situation. Most observations were made in the Ramsay 2, 3, and 5 levels, which include sedation levels considered optimal in most guidelines (Ramsay 2–3) and the heavy/over-sedated ranges (Ramsay 5–6). We believe a useful intensive care sedation monitor must reliably detect over-sedation and distinguish it from optimal sedation. Although no single state can be considered an optimal sedation level for all ICU patients, the impression that Entropy numbers do not adequately distinguish lighter from deeper sedation states was supported by the P_K values. If the monitor was intended to alert staff to excessive sedation by presenting a low Entropy number, it has limited value in non-paralyzed patients because high numbers were often generated in patients assessed as Ramsay 5 or 6. Paradoxically, this could lead to increases, rather than decreases, in sedation.

Entropy parameters were developed for use during anesthesia, when the main goal is avoidance of light anesthesia. During anesthesia State Entropy correlates well with surgical anesthetic level, and an increase in Response Entropy, which intentionally includes the higher fEMG frequencies, acts as a warning if the patient frowns in response to surgical stimulation [13]. The system was not developed to monitor patients in the ICU environment who receive frequent intense stimuli, such as tracheal suctioning, and in whom the goal is a responsive patient. High Entropy values observed in Ramsay levels 5 and 6 were associated with fEMG activations and higher fEMG power.

Our data suggest that, unlike during anesthesia, during ICU sedation much lower fEMG frequencies, including those below 32 Hz, are problematic confounders to the Entropy algorithm. This conjecture is supported by our previous observations that in the ICU, fEMG may dominate the forehead EEG signal down to frequencies as low as 22 Hz [14]. We showed that bursts of "high" fEMG were frequently observed during and in between formal sedation assessments, and were consistently associated with high entropies numbers.

This is the first study to evaluate the Entropy Module in sedated critically ill patients. Several studies have evaluated the Bispectral Index (BIS) as a sedation monitor in the ICU [10]. The BIS algorithm is different from Entropy, but it is potentially subject to similar confounders. Case reports and a volunteer study have shown that neuromuscular blockade without altering sedative drugs, or in non-sedated volunteers, decreases BIS numbers in association with abolition of fEMG [15, 16]. This indicates that fEMG is an important component of the BIS number as we have observed with Entropy. A recent controlled study found that BIS numbers decreased after administering a neuromuscular blocker to ICU patients with Ramsay score 4–5 but did not change when the Ramsay score was 6 [17]. These observations are consistent with an important confounding effect from fEMG at optimum sedation levels. Recent versions of BIS (BIS-XP) aim to adjust for fEMG activity and present fEMG power. In an observational study similar in design to ours Ely and colleagues found a trend for BIS-XP values to decrease with greater depth of clinical sedation, but observed a wide range in BIS-XP values, including those usually associated with awake individuals, even during deep sedation [18]. The authors also observed significant fEMG power during deep sedation and a strong correlation between fEMG power and BIS number, confirming the findings of others [19]. These data suggest that fEMG is likely to be an important confounder to all currently available consciousness monitors that analyze EEG from frontal electrodes.

Conclusion

In conclusion, we have shown that Entropy measured from frontal EEG has low validity to distinguish clinical sedation state in critically ill patients managed under routine clinical conditions due to strong interference from facial EMG activity.

References

1. Park G, Coursin D, Ely EW, England M, Fraser GL, Mantz J, McKinley S, Ramsay M, Scholz J, Singer M, Sladen R, Vender JS, Wild L (2001) Balancing sedation and analgesia in the critically ill. Crit Care Clin 17:1015–1027
2. Tonner PH, Weiler N, Paris A, Scholz J (2003) Sedation and analgesia in the intensive care unit. Curr Opin Anaesthesiol 16:113–121
3. Gehlbach BK, Kress JP (2002) Sedation in the intensive care unit. Curr Opin Crit Care 8:290–298
4. Kress JP, Pohlman AS, O'Connor MF, Hall JB (2000) Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation. N Engl J Med 342:1471–1477
5. Nasraway SA, Jacobi J, Murray MJ, Lumb PD (2002) Sedation, analgesia, and neuromuscular blockade of the critically ill adult: revised clinical practice guidelines for 2002. Crit Care Med 30:117–118.
6. Riker RR, Picard JT, Fraser GL (1999) Prospective evaluation of the Sedation–Agitation Scale for adult critically ill patients. Crit Care Med 27:1325–1329
7. Ely EW, Truman B, Shintani A,Thomason JWW, Wheeler AP, Gordon S, Francis J, Speroff T, Gautam S, Margolin R, Sessler CN, Dittus RS, Bernard GR (2003) Monitoring sedation status over time in ICU patients: reliability and validity of the Richmond Agitation–Sedation Scale (RASS). J Am Med Assoc 289:2983–2991
8. De Jonghe B, Cook D, Appere-De-Vecchi C, Guyatt G, Meade M, Outin H (2000) Using and understanding sedation scoring systems: a systematic review. Intensive Care Med 26:275–285
9. Walsh TS, Ramsay P, Kinnunen R (2004) Monitoring sedation in the intensive care unit: Can "black boxes" help us? Intensive Care Medicine 30:1511–1513
10. Nasraway SA (2005) The Bispectral Index: expanded performance for everyday use in the intensive care unit? Crit Care Med 33:685–687
11. Viertiö-Oja H, Maja V, Särkelä M, Talja P, Tenkanen N, Tolvanen-Laakso H, Paloheimo M, Vakkuri A, Yli-Hankala A, Merilainen P (2004) Description of the Entropy algorithm as applied in the Datex-Ohmeda S/5 Entropy Module. Acta Anaesthesiol Scand 48:154–161
12. Smith WD, Dutton RC, Smith NT (1996) Measuring the performance of anesthetic depth indicators. Anesthesiology 84:38–51
13. Ellerkmann RK, Liermann VM, Alves TM, Wenningmann I, Kreuer S, Wilhelm W, Roepcke H, Hoeft A, Bruhn J (2004) Spectral entropy and bispectral index as measures of the electroencephalographic effects of sevoflurane. Anesthesiology 101:1275–1282
14. Rautee R, Sampson T, Särkelä M, Melto S, Hovilehto S, van Gils M (2004) Application of spectral entropy to EEG and facial EMG frequency bands for the assessment of level of sedation in ICU. Proc 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, California, 1–5 September, pp 3481–3484
15. Vivien B, Maria S di, Ouattara A, Langeron O, Coriat P, Riou B (2003) Overestimation of bispectral index in sedated intensive care patients revealed by administration of muscle relaxants. Anesthesiology 99:9–17
16. Messner M, Beese U, Romstock J, Dinkel M, Tschaikowsky K (2003) The bispectral index declines during neuromuscular block in fully awake persons. Anesth Analg 97:488–491
17. Inoue S, Kawaguchi M, Sasaoka N, Hirai K, Furuya H (2006) Effects of neuromuscular block on systemic and cerebral hemodynamics and bispectral index during moderate or deep sedation in critically ill patients. Intensive Care Medicine 32:391–397
18. Ely EW, Truman B, Manzi DJ, Sigl JC, Shintani A, Bernard GR (2004) Consciousness monitoring in ventilated patients: Bispectral EEG monitors arousal not delirium. Intensive Care Med 30:1537–1543
19. Tonner PH, Wei C, Bein B, Weiler N, Paris A, Scholz J (2005) Comparison of two bispectral index algorithms in monitoring sedation in postoperative intensive care patients. Crit Care Med 33:580–584