Algorithm for Quantifying Frontal EMG Responsiveness for Sedation Monitoring

*Timo Petteri Lapinlampi, Hanna Elina Viertiö-Oja, Matti Helin,
Kimmo Henrik Uutela, Mika Olli Kristian Särkelä, Anne Vakkuri,
Gordon Bryan Young, Timothy Simon Walsh*

ABSTRACT: *Introduction:* To study stimulation-related facial electromyographic (FEMG) activity in intensive care unit (ICU) patients, develop an algorithm for quantifying the FEMG activity, and to optimize the algorithm for monitoring the sedation state of ICU patients. *Methods:* First, the characteristics of FEMG response patterns related to vocal stimulation of 17 ICU patients were studied. Second, we collected continuous FEMG data from 30 ICU patients. Based on these data, we developed the Responsiveness Index (RI) algorithm that quantifies FEMG responses. Third, we compared the RI values with clinical sedation level assessments and adjusted algorithm parameters for best performance. *Results:* In patients who produced a clinically observed response to the vocal stimulus, the poststimulus FEMG power was 0.33 μV higher than the prestimulus power. In nonresponding patients, there was no difference. The sensitivity and specificity of the developed RI for detecting deep sedation in the subgroup with low probability of encephalopathy were 0.90 and 0.79, respectively. *Conclusion:* Consistent FEMG patterns were found related to standard stimulation of ICU patients. A simple and robust algorithm was developed and good correlation with clinical sedation scores achieved in the development data.

RÉSUMÉ: Surveillance de la sédation au moyen d'un algorithme quantifiant la réactivité frontale à l'EMG. *Contexte:* Nous avons développé un algorithme pour quantifier l'activité en lien à la stimulation lors de l'électromyographie faciale (EMGF) chez des patients hospitalisés à l'unité de soins intensifs (USI) et nous avons optimisé l'algorithme pour surveiller l'état de sédation de ces patients. *Méthode:* Nous avons d'abord étudié les caractéristiques à l'EMGF de la réponse à la stimulation vocale chez 17 patients. Nous avons ensuite recueilli des données EMGF en continu chez 30 patients de l'USI. À l'aide de ces données, nous avons développé l'algorithme de l'Indice de réactivité (IR) qui quantifie les réponses à l'EMGF. Puis nous avons comparé les valeurs de l'IR aux évaluations du niveau clinique de sédation et nous avons ajusté les paramètres de l'algorithme afin d'optimiser sa performance. *Résultats:* Chez les patients qui avaient une réponse à la stimulation vocale observable cliniquement, la puissance à l'EMGF poststimulus était de 0,33 μV plus élevée que la puissance préstimulus, alors qu'il n'y avait pas de différence chez les patients qui ne répondaient pas au stimulus. La sensibilité et la spécificité de l'IR que nous avons développé pour détecter une sédation profonde dans le sous-groupe chez qui la probabilité d'une encéphalopathie était faible étaient de 0,90 et 0,79 respectivement. *Conclusion:* Nous avons observé des profils constants en lien à la stimulation standard chez des patients hospitalisés à l'USI. Nous avons développé un algorithme simple et robuste, et nous avons démontré une bonne corrélation aux scores de sédation clinique obtenus lors du développement de l'algorithme.

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] Oversedation during critical illness is associated with adverse outcomes, and clinical strategies designed to avoid this can decrease ventilation times and intensive care unit (ICU) length of stay.[2,3] These strategies can also improve clinical outcomes, decrease complications (e.g., 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 costs of sedative drugs and is considered best practice in current 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] Clinical sedation scales are, however, prone to interrater variability and provide only intermittent information of the patient's state. The most

commonly used assessment scale is the Ramsay scale,[9] which controversially has not been validated in an ICU environment.

Electroencephalogram (EEG)-based monitoring devices such as the Bispectral Index (BIS) (Aspect Medical Systems, Needham MA) and the Entropy measurement (GE Healthcare, Helsinki Finland) are used in monitoring the depth of anesthesia in the operating room. The BIS monitor is also intended for use in the ICU, but although correlation between BIS and Ramsay has been established,[10] studies have shown the confounding effect of frontal electromyographic (FEMG) activity on both the Entropy[11] and BIS[12] in the ICU setting.

In conscious, unparalyzed, and lightly anesthetized patients, painful and/or stressful stimuli are associated with increased FEMG amplitude.[13] FEMG power may also increase in association with certain sounds in deeply anesthetized or comatose patients who are clinically unresponsive to surgical or electrical stimulation.[13] Because FEMG is a major confounder for EEG-based algorithms, we hypothesized that measuring FEMG activity would provide a surrogate measure of sedation depth. Because deepening sedation tends to suppress naturally occurring arousals as well as responses to external stimuli,[14] measuring the number and magnitude of these responses should give a surrogate measure of the depth of sedation.

For quantifying FEMG activity in ICU patients, we developed a novel algorithm that analyzes the response patterns in the spontaneous FEMG signal measured from the forehead and outputs a single digit with arbitrary units, the Responsiveness Index (RI), ranging from 0, corresponding to no variability in FEMG activity, to 100, corresponding to strong and variable FEMG activity. The characteristics of the response patterns in the FEMG signal to predefined stimuli were studied as part of the algorithm development. The measurement itself does not use active stimulation because the patient is continuously subjected to several sources of external and internal stimuli in the ICU. External sources of stimulation can include lights, noise, and caregiving procedures, whereas internal stimuli can include pain, stress, and anxiety.

We have previously evaluated the RI performance in critically ill patients recovering from cardiac surgery and shown a better correlation to clinical sedation scores compared with state and response entropy.[15] The current article first describes a study of EMG responses and the physiological background that led to the approach used in RI and then describes the development of the algorithm details using the development data.

METHODS

Physiological Background

There is abundant neuroanatomical and neurophysiological support for an interaction of the components of the reticular formation (RF) in the brainstem that are responsible for arousal and pain response and of the motor nuclei in the brainstem that are responsible for activation of skeletal muscles of the forehead.[16] The facial muscles are innervated by the seventh cranial nerve (facial nerve) that emerges from the brainstem between the pons and the medulla. The facial nerve nucleus is situated within the RF in the lower pons. The dorsal aspect of the facial nucleus contributes to innervation of the forehead muscles, which are monitored with surface electrodes used for RI calculation. The third cranial (oculomotor) nucleus, which is intimately associated

with the RF in the dorsal midbrain, sends fibers that accompany the seventh nerve in its course. These supply the anterior part of the occipitalis-frontalis, the orbicularis palpebrarum, and the corrugator supercilii.

The rostral pontine RF is also known to regulate somatic motor tone,[17-19] and the progression from wakefulness to rapid eye movement sleep is marked by a decrease of this tone.[20] The activity of the area is modulated by a number of neurotransmitters, notably gamma aminobutyric acid (GABA).[21-25] The evolution of wakefulness from isoflurane anesthesia has been characterized experimentally by increasing GABA levels in the RF, which increases EMG and decreases EEG activity.[26]

Even anesthetized patients may show limb movements as a response to painful stimulation. These reflexes can be spinal, and it is suggested that immobility caused by anesthesia is primarily produced by effect on the spinal cord.[27,28] In contrast, the frontal muscles are controlled by the brainstem and require brain activation for even simple reflexes.

With increasing level of consciousness, cortical processing and the facial muscles expressing the emotional state of the patient are likely to generate additional responses. Startle-type reflexes can be induced by many types of stimulation.[29] Finally, when the patient is fully conscious, FEMG activity resulting from speech and other voluntarily generated muscle activations occurs. Our hypothesis is that different levels of sedation are reflected at the threshold of generating frontal muscle movements and therefore the FEMG response amounts will be correlated with the levels of sedation.

METHODS

The development of the RI algorithm was divided into three stages: characterizing ICU patients' FEMG activity related to standardized stimuli, developing the algorithm by an iterative process, and evaluating the algorithm with clinical data. The purpose of the first stage was to test the hypothesis that a standardized vocal stimulus causes a consistent response pattern in the frontal EMG activity of ICU patients. In the second stage, the observed FEMG response characteristics would be used to develop an algorithm that quantifies FEMG activity. The third stage would be an initial test on the clinical usability of the algorithm.

FEMG Response to Vocal Stimulus

To study the characteristics of the frontal FEMG response to stimuli, we used a previously collected data set of 19 ICU patients (11 females, eight males) with a mean age (standard deviation [SD]) of 46.5 (17.1) years. Patients were admitted to the Helsinki Surgical Hospital (Helsinki, Finland) ICU because of hepatic encephalopathy. After receiving written informed consent from the patient or immediate family member, each subject was monitored for 2-7 days. Approval for the study was obtained from the local ethics committee. Monitoring was divided into active and passive periods. During active monitoring periods (~8 hours per day on average) the patients were stimulated once every 3 hours with a prerecorded vocal stimulus. The command "open your eyes" (in Finnish) was used and applied via headphones (Sony MDR-XD200) at an 85-dB intensity. A positive response was recorded if the patient opened his or her eyes within 10 seconds from the start of the stimulus. The stimuli's time stamps were
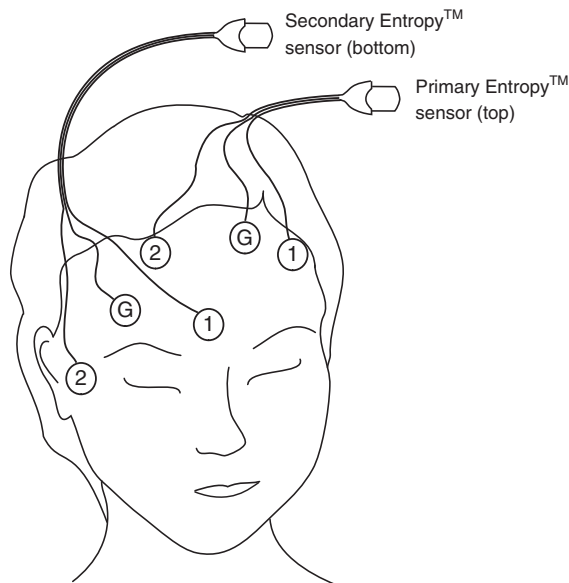
Figure 1: Placement of primary and secondary Entropy sensors in the recordings. Numbers 1 and 2 denote the active electrodes; G represents the ground electrode.

recorded automatically and the research nurse recorded the response (positive or negative) manually. During passive monitoring periods, the stimuli were not applied.

The FEMG signals were recorded with the GE Datex-Ohmeda Entropy module connected to a GE Datex-Ohmeda S/5 monitor. The primary Entropy sensor was placed on the patient's forehead symmetrically across the hemispheres approximately 4 cm above the orbital bridge. A secondary Entropy sensor was placed below the first one, with one electrode on the temple and the other above the contralateral eye (Figure 1). Two sensors were used to study the effect of placing the electrodes on different facial muscles. The primary sensor covered the frontalis muscle on both hemispheres. The negative pole of the secondary electrode was placed at the temporalis muscle and the positive pole on the frontalis and the corrugator supercilii. The sampling frequency of the measurement was 400 Hz and the frequency band was 0.5-150 Hz.

Epochs starting 2 minutes before each stimulus and ending 2 minutes after the stimulus were extracted from the FEMG waveform. A visual inspection was performed to remove epochs that contained artifacts, including poor electrode contacts and impedance tests performed by either one of the sensors. All epochs containing artifacts or more than 20 seconds of missing data were removed before analysis. From a total of 415 assessments, the research nurse classified 244 as positive response and 171 as no response.

An estimate of the FEMG power signal was calculated by dividing each epoch to 0.5 s long windows. Root mean square power from the 65-95 Hz frequency band was calculated using the fast-Fourier transform. The frequency band was chosen conservatively to easily avoid the mains frequencies (and multiples) and to contain dominantly FEMG activity.

The median and 25th and 75th quartile FEMG powers were analyzed separately for patients who produced a clinically observable response (opened eyes within 10 seconds from the stimulus) and for those who did not respond to the stimulus.

By observing individual FEMG power curves, we discovered that the latency from the start of the stimulus to the FEMG response onset differed within and between patients. For comparing the FEMG responses, we defined a feature-aligned FEMG response by matching the individual responses by response onset rather than the stimulus onset. Response onset time was defined as the first time point after the stimulus when FEMG power increased to a value greater than the 20-second mean FEMG power plus 3 SDs before the stimulus. Only the cases in which the response onset occurred within 30 seconds or less after the stimulus were included. Feature-aligned median and 25th percentile and 75th percentile powers were calculated to describe the form of a typical FEMG response. Of the 244 assessments classified as positive response, 53 were rejected, leaving 191 assessments for the feature-aligned analysis. Wilcoxon signed-rank test was performed to a median power before and after stimulus to analyze the significance of differences in the power levels.

Responsiveness Algorithm Description

In the second stage, the RI algorithm was developed following an iterative process using real ICU patient data (described in the following section). The starting point was to detect patterns resembling the FEMG response patterns related to the vocal stimulus and quantify the amount and magnitude of those patterns over time. The performance of the algorithm was measured by comparing the output with the clinical sedation scores (the modified Ramsay score) in the development data set. Based on the performance metrics in the development data, the algorithm was tuned by iterating the algorithm parameters.

Development Data Set

The development data for the RI algorithm have been published in the team's previous article (Walsh et al., 2008).[11] The inclusion criteria for the patients were (1) admission to the ICU because of a critical illness, (2) requirement for mechanical ventilation, and (3) written consent from the patient or relatives. The exclusion criteria were (1) drug overdose, (2) need for neuromuscular paralysis at the time of enrollment, (3) age < 16 years, (4) pregnancy, and (5) status epilepticus.

After local ethics committee approval and written informed consent, each subject was monitored for a maximum of 72 hours. Monitoring was stopped earlier if mechanical ventilation was discontinued, the patient or a relative requested discontinuation, the patient died, or technical reasons made continuation impossible. Two Entropy sensors were placed on the patient's forehead as described in Figure 2. As a reference measure for the depth of sedation, we used a self-developed modified Ramsay scoring system that used standardized vocal, loud vocal, and tetanic stimuli. The vocal stimulus was a verbal command "open your eyes" spoken to the patient. The loud vocal stimulus was a similar verbal command with louder intensity. The tetanic stimulus was a repeated electrical stimulus (50 Hz, 40 mA, 250 ms) applied to the ulnar nerve. The purpose of this modification was (1) to diminish the variation in the assessment by using machine-generated stimuli and (2) to enable discrimination of deep sedation levels in which the patient is unconscious by applying tetanic stimulus. The assessment started with the lightest stimulus and proceeded step by step to stronger stimuli. If the patient at any step showed an observable response, no further stimuli were applied.
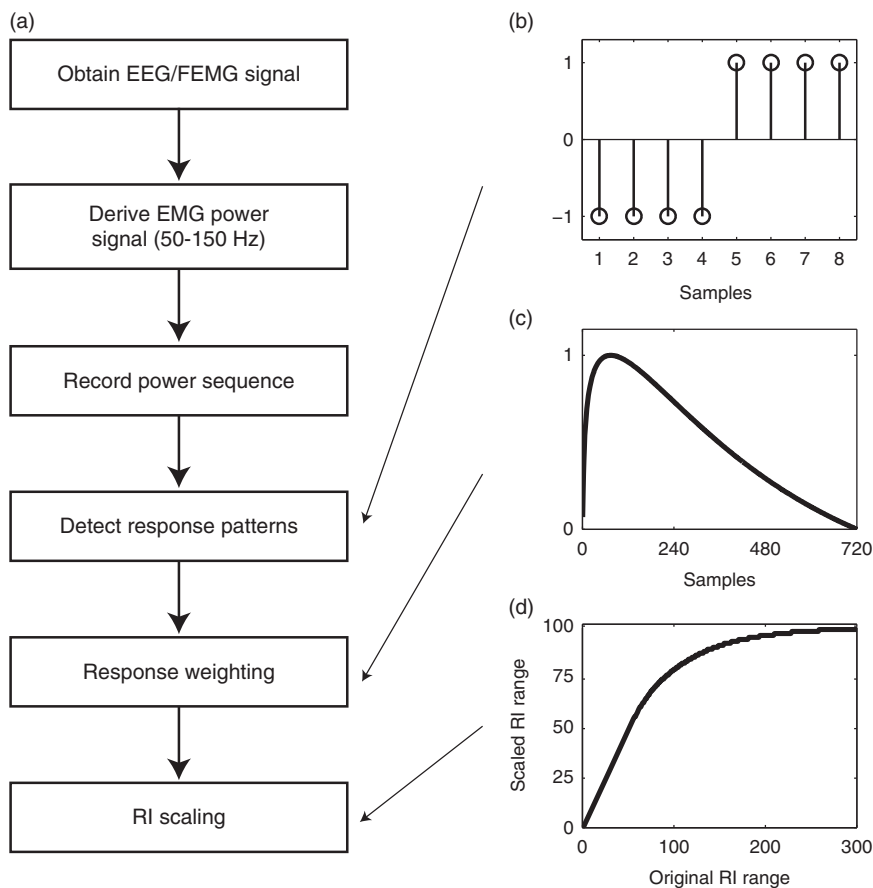
Figure 2: *Flow chart of the responsiveness index algorithm. (a, b) The impulse response of the filter used to detect the response patterns from the FEMG data. (c) The weighting function g and (d) the scaling function S used in the RI algorithm.*

The correspondence of the standard stimuli, the modified Ramsay scores, and the expected responses is shown in Table 1. A single trained research nurse performed all the modified Ramsay assessments, minimizing variation between the assessments. The interval between the assessments was at least 30 minutes and depended on the clinical duties of the research nurse. The RI values were not shown to the clinical staff during data collection.

Based on the clinician's review of the patient record and patient state after regaining consciousness, a retrospective classification into two subgroups was made: group 1 had a low probability of encephalopathy (17 patients) and group 2 (13 patients) had a high probability of encephalopathy. This exploratory analysis attempted to identify patients with significant encephalopathy and/or delirium caused by critical illness, which would be a confounder to normal responses to vocal and possibly physical stimuli.

Algorithm Description

The RI algorithm (Figure 2) calculates the root mean square FEMG power in 5-second epochs using the frequency band

Table 1: Definition of the modified Ramsay scale

Modified Ramsay score	Definition	Response definition
1	Spontaneous eye opening and/or agitation	N.A.
2	Spontaneous eye opening	N.A.
3	Response to vocal stimulus	Patient opens eyes within 10 seconds of the stimulus
4	Response to loud auditory stimulus	A gross, purposeful movement of the head or extremities within 10 seconds of the stimulus
5	Response to tetanic stimulus	A gross, purposeful movement of the head or extremities within 10 seconds of the stimulus
6	No response to tetanic stimulus	N.A.

N.A. = not applicable.

between 50 and 150 Hz. The power calculation is performed in a time domain after applying a high-pass finite impulse response filter and a 10-Hz comb filter to the signal. The comb filter is used to attenuate the mains frequencies (50 and 60 Hz and multiples). The power value time series is then subjected to a filter that extracts the steep rises related to patient responses from the signal. The filter is implemented as a finite impulse response filter with a step-shaped impulse response depicted in Figure 2b. The shape of the filter was chosen based on the results of the median feature-aligned FEMG power response analysis. For constant FEMG activity, the filter output is zero regardless of the baseline level. The length of the filter is eight samples corresponding to 40 seconds in the FEMG power time series. If the filtered power value at time point t is denoted by $P_F(t)$, the RI value at time t is written as

$$RI(t) = S\left(\sum_{n=t-\Delta}^{t} g(n) log\left(\frac{P_F(n)}{1\mu V} + 1 \right) \right)$$

where S is a scaling function, g is a weighting function, and Δ is the number of P_F values included in the summation. Hence, the RI is a weighted average of the logarithm of the P_F values adding 1 µV in the given time window. The addition of 1 µV was introduced to set the minimum value of the logarithm to zero. The purpose of the weighting function g (Figure 2c) is to operate as a low-pass filter and to give more weight to recent FEMG changes compared with those occurring in the past. The value of Δ was set at 720 samples (corresponding to 1 hour). The weighting function g gives a 50% contribution to the latest approximate 17 minutes of the 1-hour time window in the RI calculation.

For enhanced usability, the RI value is scaled as a 0-100 range by a scaling function S shown in Figure 2d. The introduction of the scaling was driven by the clinical need to maximize discrimination within deeper sedation states where clinical assessments based on intermittent assessments lack discrimination. The scaling is linear in the range 0-50, which is the relevant range for detecting deeper sedation states, whereas values >50 are scaled nonlinearly. The maximum RI value is set at 100.

The RI algorithm is tolerant to typical EEG artifacts such as patient eye movements, blinks, and movements because these phenomena are present in the low end of the frequency spectrum and are discarded in the filtering steps of the algorithm. Short-lasting (<1 second) EMG bursts in this context are regarded as artifacts and also filtered out. The frequency band used in the calculation ranges from 50 to 150 Hz with removal of the main frequencies. FEMG, when measured with surface electrodes with relatively large area, has a wide frequency spectrum, and the choice of the frequency band is not critical as long as a wide enough band is used.

## ROC Analysis and Cutoff Values

RI values were calculated from all the patients in the development data set. The RI values preceding each modified Ramsay level assessment by 1 minute were extracted from the data. The 1-minute advance was used to avoid the effect of the Ramsay assessment procedure on the RI value. The sensitivity and specificity and the related RI threshold value for detecting deeper sedation states (defined as modified Ramsay levels 5 and 6) were determined by receiver operating characteristic (ROC) analysis. Sensitivities and specificities were calculated by altering the threshold RI value. The cost of false negatives and false positives

Table 2: Number of observations across modified Ramsay scores in the development data

	Sensor	Ramsay scores	Number of rejected RI values				Number of RI values included in the analysis						
			No RI available	Poor contact	Too little data	Total	Ramsay 1	Ramsay 2	Ramsay 3	Ramsay 4	Ramsay 5	Ramsay 6	Total
All (n = 30)	Primary	487	36	12	78	126	4	80	137	20	103	17	361
	Secondary	478	11	10	64	85	8	87	150	19	119	19	402
Subgroup 1 (n=17)	Primary	211	1	6	34	41	4	47	65	12	39	3	170
	Secondary	211	6	5	23	34	4	49	68	12	40	4	177
Subgroup 2 (n=13)	Primary	276	35	6	44	85	0	33	72	8	64	14	191
	Secondary	276	8	5	39	52	4	37	82	7	79	15	224

RI = Responsiveness Index.
Subgroup 1/2 = low/high probability of encephalopathy.
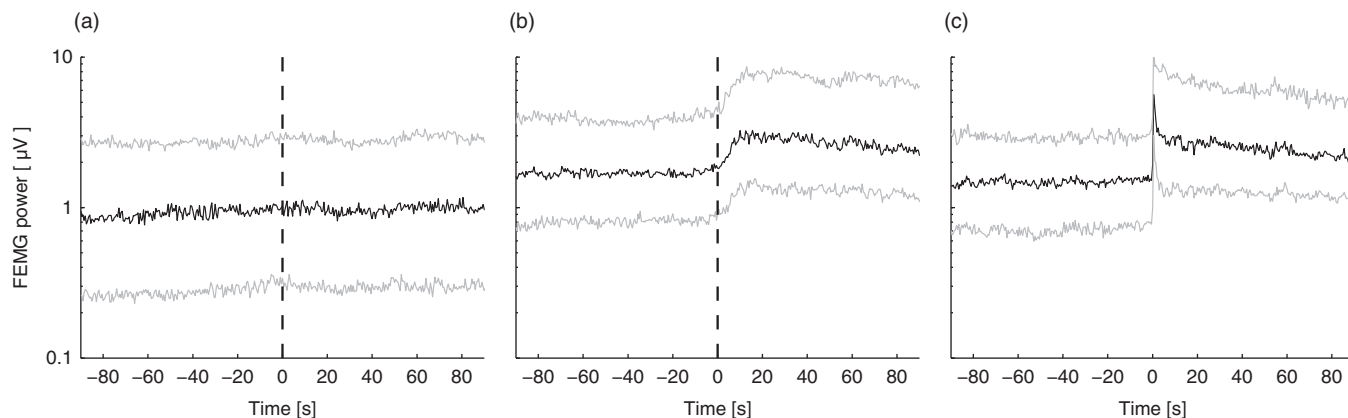
Figure 3: *Median FEMG power (25th and 75th quartiles) related to (a) the patients not responding to the vocal stimulus and (b) the patients responding. (c) Feature-aligned power. Time t = 0 is the time of the stimulus onset in (a, b). In (c), the time t = 0 indicates the response onset.*

was presumed to be equal, so the optimal threshold was selected by maximizing the sum of sensitivity and specificity. This corresponds to choosing a point on the ROC curve maximizing the distance from the diagonal.

The total number of Ramsay scores, number of rejected data points (with reasoning), and distribution of data points to different Ramsay levels are shown in Table 2.

RESULTS

FEMG Response to Vocal Stimulus

The median and 25th and 75th quartile FEMG powers from the primary Entropy sensor are shown in Figure 3 for patients not responding to the vocal stimulus (a) and for the patients who did respond (b) to the stimulus. In the responding group, the poststimulus FEMG power was significantly higher than the prestimulus power, with a median difference of 0.33 µV (quartiles: −0.00, 1.38, $p < 0.001$) between 30 seconds median power before and after the stimulus. In the nonresponding group, the pre- and poststimulus powers did not differ (median difference 0.00 µV, quartiles: −0.04, 0.05, $p = 0.73$). For the secondary Entropy sensor, the results were similar. In the responding group, poststimulus FEMG power was significantly higher than the prestimulus power with a median difference of 0.23 µV (quartiles: −0.07, 0.86,

$p < 0.001$). In the nonresponding group, the pre- and poststimulus powers did not differ (median difference 0.00 µV, quartiles: −0.05, 0.05, $p = 0.96$).

As shown in Figure 3c, the feature-aligned FEMG power shows a very sharp rise followed by an elevated baseline level after the response. The initial peak in the aligned response curve may be enhanced by the way the alignment was performed based on the power data itself and by transient muscle activity related to the eye-opening command. However, a fast rise followed by a sustained power increase was a typical finding also in the unaveraged EMG power; therefore, the stepwise increase pattern was used for the RI algorithm.

RI Algorithm Performance

The data set included 30 critically ill patients' general ICU data from Edinburgh Royal Infirmary (22 males, 8 females) with a mean age (SD) of 59.0 (17.5) years. The algorithm performance was measured within the development data set with the purpose of guiding the iterative algorithm development. The reported results are calculated with the final algorithm settings described previously in the Algorithm Description section.

The RI value distributions across the modified Ramsay levels in all patients and subgroups 1 and 2 are shown in Figure 4.
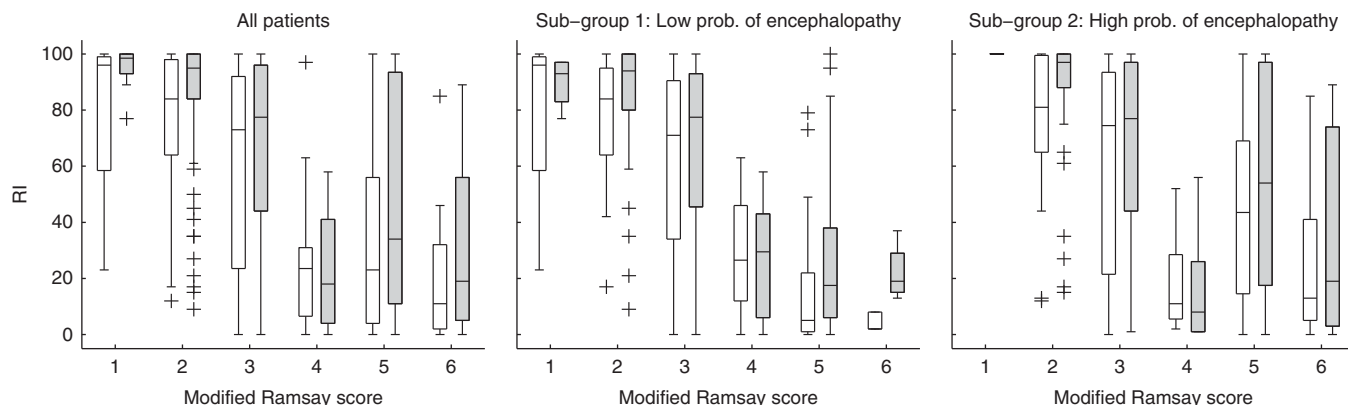
Figure 4: *Boxes and whiskers of RI values across different modified Ramsay levels. The white boxes indicate the primary Entropy sensor and the gray boxes the secondary Entropy sensor. The horizontal lines within the boxes denote median RI values and the box edges the 25th and 75th quartiles. Statistical outliers denoted by a + sign are defined as data points whose distance from the 25th or 75th quartile lines was greater than 1.5 times the interquartile range.*

616

Table 3: The optimal Responsiveness Index threshold values and related sensitivities and specificities related to detecting deep sedation (Richmond Agitation-Sedation Scale > 5) with the RI algorithm

Cohort	Sensor	Optimal threshold	Sensitivity	Specificity
All	Primary	57	0.80	0.68
	Secondary	57	0.67	0.70
Subgroup 1	Primary	35	0.90	0.79
	Secondary	59	0.91	0.71
Subgroup 2	Primary	57	0.72	0.65
	Secondary	57	0.56	0.67

Subgroup 1/2 = low/high probability of encephalopathy.

The optimal threshold values for detecting deep sedation in all patients and subgroups 1 and 2 are presented in Table 3. The corresponding ROC curves are shown in Figure 5. Figure 6 shows an example of a patient case in the development data set (subgroup 1) with EEG, FEMG power, filtered FEMG power, RI, and the modified Ramsay score.

DISCUSSION

The RI algorithm has the potential to be a robust and objective method for quantifying the balance between the sedation state and the stimulation level of an ICU patient. It is not a direct measure of the depth of sedation and should not be treated as such. For example, a patient receiving little or no sedation and sleeping calmly may respond strongly to stimulation-related care procedures or noise. The same patient will show low RI values if he or she is sleeping calmly and no stimulation is applied.

There are some limitations related to this study. First, the patient group used in the first phase (characterization of ICU patients' FEMG activity) was limited to a specific condition and does not represent general ICU patients. Furthermore, hepatic encephalopathy may cause neural alterations that in turn might affect patients' responses to stimuli. Second, the development data set used in the RI algorithm development was quite small, containing only 30 patients. However, the algorithm is relatively

simple and robust with few adjustable parameters. Good results were previously reported in another patient population,[15] which provides confidence in the method and warrants further investigations in larger patient populations.

Some of the potentially encephalopathic patients showed increased levels of EMG responses without purposeful movements as reaction to the auditory stimulus (Figure 3, modified Ramsay scores 5 and 6). Our hypothesis is that encephalopathy may alter the patients' reactions to stimuli. In particular, responses to stimuli expecting a cognitive reaction (such as a command to open one's eyes) might be compromised. This would make traditional sedation assessments (including the modified Ramsay score used in this study) unreliable in these patients because responses might be suppressed from the encephalopathy instead of the sedation state.

In the analysis, we defined Ramsay levels 5 and 6 as deeper sedation. The definition of deep sedation, however, depends on the clinical requirements of individual patients. Recent studies recommend a target Ramsay score of 3[30] or from 3 to 4[4] in general. Guidelines, however, recommend that the need for sedation should be assessed according to patient and situation.[31] The main purpose of the RI algorithm is to alert the hospital staff to deep sedation. The discrimination at lighter sedation levels was considered less important in the development phase because light sedation is usually easily observable from patient behavior. The time scale of the RI algorithm is rather long (60 minutes on average) and therefore it is not applicable in cases where the time scale of events is short, such as procedural sedation.

Measuring the performance of the RI algorithm by comparing it with clinical sedation scales such as Ramsay and the Richmond Agitation-Sedation Scale may understate the actual performance of the algorithm. These clinical scales are the current gold standard, but they measure a slightly different phenomenon. Our hypothesis is that the RI algorithm measures the balance between the sedative drug effect and the level of stimulation (i.e., if the sedation level remains constant, the RI value may change when the stimulation level changes). The clinical scales, on the other hand, measure patients' response to certain stimuli as perceived by a caring nurse.

Using neuromuscular blocking agents (NMBA) is believed to prevent the use of a developed monitoring method because it is based on changes in FEMG activity. However, upper facial
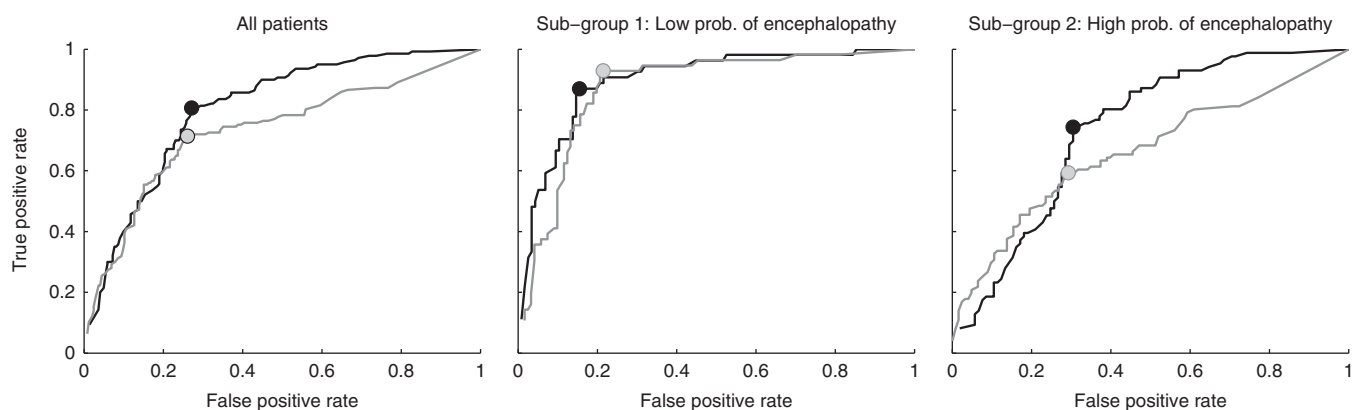
Figure 5: *The ROC curves for the primary (black curves) and secondary (gray curves). Entropy sensors in all patients and the subgroups with a low and high probability of encephalopathy. The optimal RI threshold values are shown as filled circles.*
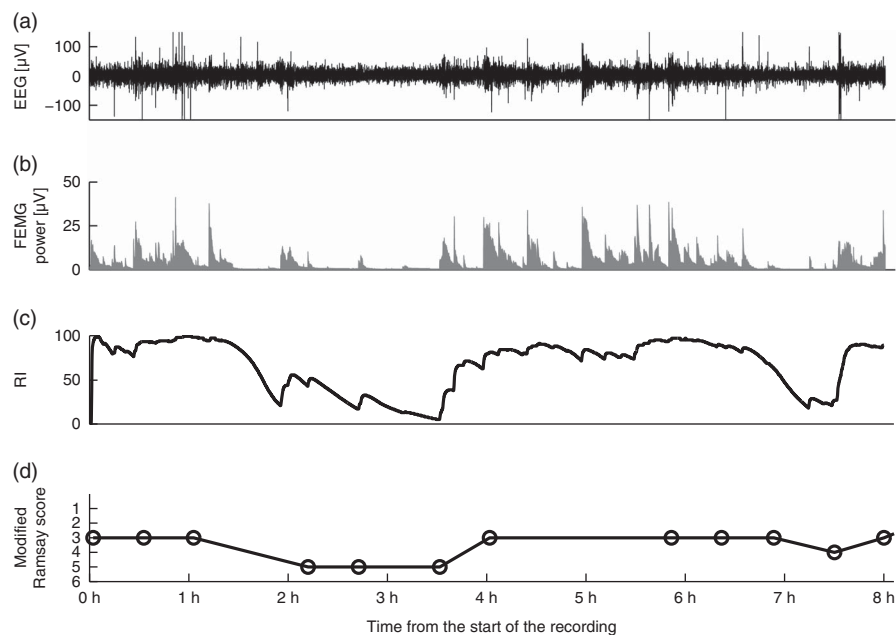
Figure 6: *Algorithm example from actual patient case in the development data set in sub-group 1. From top to bottom are shown (a) the EEG waveform, (b) FEMG power, (c) the Responsiveness Index, and (d) the modified Ramsay score.*

muscles are more resistant to the effect of NMBAs than the hand muscles, where the effect of the neuromuscular block is usually measured.[32] Today the use of NMBAs in the ICU is restricted to limited patient groups and are not recommended in the current guidelines.[5]

This study did not address the patients' related outcomes, which could be looked at in a bigger trial.

STATEMENT OF AUTHORSHIP

TL carried out the statistical analysis for RI algorithm performance, participated in the RI algorithm development, and was the main writer of the manuscript. HV participated in the design of the study and RI algorithm development. MH carried out the statistical analysis for FEMG responses to vocal stimuli. KU participated in the design of the study and RI algorithm development. MS participated in the RI algorithm development. AV participated in the design of the study and data collection from the hepatic encephalopathy patients. BY wrote the Physiological Background section. TW participated in the design of the study and data collection from the general ICU patients. All authors read and approved the final manuscript.

REFERENCES

1. Park G, Coursin D, Ely EW, et al. Balancing sedation and analgesia in the critically ill. Crit Care Clin. 2001;17:1015-27.
2. Gehlbach BK, Kress JP. Sedation in the intensive care unit. Curr Opin Crit Care. 2002;8:290-8.
3. Tonner PH, Weiler N, Paris A, Scholz J. Sedation and analgesia in the intensive care unit. Curr Opin Anaesthesiol. 2003;16:113-21.
4. Kress JP, Pohlman AS, O'Connor MF, Hall JB. Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation. N Engl J Med. 2000;342:1471-7.
5. Nasraway SA, Jacobi J, Murray MJ, Lumb PD. Task Force of the American College of Critical Care Medicine of the Society of Critical Care Medicine and the American Society of Health-System Pharmacists, American College of Chest Physicians. Sedation, analgesia, and neuromuscular blockade of the critically ill adult: revised clinical practice guidelines for 2002. Crit Care Med. 2002;30:117-8.
6. Riker RR, Picard JT, Fraser GL. Prospective evaluation of the Sedation-Agitation Scale for adult critically ill patients. Crit Care Med. 1999;27:1325-9.
7. De Jonghe B, Cook D, Appere-De-Vecchi C, Guyatt G, Meade M, Outin H. Using and understanding sedation scoring systems: a systematic review. Intensive Care Med. 2000;26:275-85.
8. Ely EW, Truman B, Shintani A, et al. Monitoring sedation status over time in ICU patients: reliability and validity of the Richmond Agitation-Sedation Scale (RASS). JAMA. 2003;289:2983-91.
9. Jackson DL, Proudfoot CW, Cann KF, Walsh TS. The incidence of sub-optimal sedation in the ICU: a systematic review. Crit Care. 2009;13:R204.

10. Mondello E, Siliotti R, Noto G, et al. Bispectral Index in ICU: correlation with Ramsay score on assessment of sedation level. J Clin Monit Comput. 2002;17:271-7.

11. Walsh TS, Ramsay P, Lapinlampi TP, Särkelä MOK, Viertiö-Oja HE, Meriläinen PT. An assessment of the validity of spectral entropy as a measure of sedation state in mechanically ventilated critically ill patients. Intensive Care Med. 2008;34:308-15.

12. Vivien B, Di Maria S, Ouattara A, Langeron O, Coriat P, Riou B. Overestimation of Bispectral Index in sedated intensive care patient revealed by administration of muscle relaxant. Anesthesiology. 2003;99:9-17.

13. Edmonds HL Jr, Couture LJ, Stolzy SL, Paloheimo M. Quantitative surface electromyography in anesthesia and critical care. Int J Clin Monit Comput. 1986;3:135-45.

14. Dutton RC, Smith WD, Bennett HL, Archer S, Smith NT. Craniofacial electromyogram activation response: another indicator of anesthetic depth. J Clin Monit. 1997;14:5-17.

15. Walsh TS, Lapinlampi TP, Ramsay P, Särkelä MOK, Uutela KH, Viertiö-Oja HE. Responsiveness of the frontal EMG for monitoring the sedation state of critically ill patients. Br J Anaesth. 2011; 107:710-8.

16. Kandel ER, Schwartz JH, Jessell TM. Principles of neural science, 4th ed. New York. McGraw-Hill; 2000.

17. Jouvet M. Research on the neural structures and responsible mechanisms in different phases of physiological sleep. Arch Ital Biol. 1962;100:125-206.

18. Henley K, Morrison AR. A re-evaluation of the effects of lesions of the pontine tegmentum and locus coeruleus on phenomena of paradoxical sleep in the cat. Acta Neurobiol Exp (Wars). 1974;34:215-32.

19. Hendricks JC, Morrison AR, Mann GL. Different behaviors during paradoxical sleep without atonia depend on pontine lesion site. Brain Res. 1982;239:81-105.

20. Wills N, Chase MH. Brain stem control of masseteric reflex activity during sleep and wakefulness: mesencephalon and pons. Exp Neurol. 1979;64:98-117.

21. Morales FR, Engelhardt JK, Soja PJ, Pereda AE, Chase MH. Motoneuron properties during motor inhibition produced by microinjection of carbachol into the pontine reticular formation of the decerebrate cat. J Neurophysiol. 1987;57:1118-29.

22. Camacho-Arroyo I, Alvarado R, Manjarrez J, Tapia R. Microinjections of muscimol and bicuculline into the pontine reticular formation modify the sleep-waking cycle in the rat. Neurosci Lett. 1991;129:95-7.

23. Lai YY, Siegel JM. Pontomedullary glutamate receptors mediating locomotion and muscle tone suppression. J Neurosci. 1991;11:2931-7.

24. Xi MC, Morales FR, Chase MH. Evidence that wakefulness and REM sleep are controlled by a GABAergic pontine mechanism. J Neurophysiol. 1999;82:2015-9.

25. Xi MC, Morales FR, Chase MH. The motor inhibitory system operating during active sleep is tonically suppressed by GABAergic mechanisms during other states. J Neurophysiol. 2001;86:1908-15.

26. Vanini G, Watson CJ, Lydic R, Baghdoyan HA. GABAergic neurotransmission in the pontine reticular formation modulates hypnosis, immobility, and breathing during isoflurane anesthesia. Anesthesiology. 2008;109:978-88.

27. Rampil IJ, Mason P, Singh H. Anesthetic potency (MAC) is independent of forebrain structures in the rat. Anesthesiology. 1993; 78:707-12.

28. Rampil IJ. Anesthetic potency is not altered after hypothermic spinal cord transection in rats. Anesthesiology. 1994;80:606-10.

29. Brown P, Rothwell JC, Thompson PD, Britton TC, Day BL, Marsden CD. New observations on the normal auditory startle reflex in man. Brain. 1991;114:1891-902.

30. Brook AD, Ahrens TS, Schaiff R, et al. Effect of a nursing implemented sedation protocol on the duration of mechanical ventilation. Crit Care Med. 1999;27:2609-15.

31. Jacobi J, Fraser GL, Coursin DB, et al; Task Force of the American College of Critical Care Medicine (ACCM) of the Society of Critical Care Medicine (SCCM), American Society of Health-System Pharmacists (ASHP), American College of Chest Physicians. Clinical practice guidelines for the sustained use of sedatives and analgesics in the critically ill adult. Crit Care Med. 2002;30:119-41.

32. Paloheimo MPJ, Wilson RC, Edmonds HL, Lucas LF, Triantafillou AN. Comparison of neuromuscular blockage in upper facial and hypothenar muscles. J Clin Monit. 1988;44:256-60.