



Continuity with caveats in anesthesia: state and response entropy of the EEG

Max Ebensperger¹ · Matthias Kreuzer¹ · Stephan Kratzer² · Gerhard Schneider¹ · Stefan Schwerin¹

Received: 21 November 2023 / Accepted: 22 January 2024
© The Author(s) 2024

Abstract

The growing use of neuromonitoring in general anesthesia provides detailed insights into the effects of anesthetics on the brain. Our study focuses on the processed EEG indices State Entropy (SE), Response Entropy (RE), and Burst Suppression Ratio (BSR) of the GE EntropyTM Module, which serve as surrogate measures for estimating the level of anesthesia. While retrospectively analyzing SE and RE index values from patient records, we encountered a technical anomaly with a conspicuous distribution of index values. In this single-center, retrospective study, we analyzed processed intraoperative electroencephalographic (EEG) data from 15,608 patients who underwent general anesthesia. We employed various data visualization techniques, including histograms and heat maps, and fitted custom non-Gaussian curves. Individual patients' anesthetic periods were evaluated in detail. To compare distributions, we utilized the Kolmogorov–Smirnov test and Kullback–Leibler divergence. The analysis also included the influence of the BSR on the distribution of SE and RE values. We identified distinct *pillar* indices for both SE and RE, i.e., index values with a higher probability of occurrence than others. These *pillar* index values were not age-dependent and followed a non-equidistant distribution pattern. This phenomenon occurs independently of the BSR distribution. SE and RE index values do not adhere to a continuous distribution, instead displaying prominent *pillar* indices with a consistent pattern of occurrence across all age groups. The specific features of the underlying algorithm responsible for this pattern remain elusive.

Keywords Anesthesia · Electroencephalogram · Burst suppression · State entropy · Response entropy

1 Introduction

1.1 Background

To monitor the hypnotic component of anesthesia, the electroencephalograph (EEG) is a non-invasive measurement that directly monitors anesthetic-induced changes in the target organ, the brain [1, 2]. Commercial monitoring systems use a reduced EEG montage to record from the patient's forehead. The systems translate the raw EEG into an index, which provides information regarding the patient's anesthetic level, i.e., the hypnotic component of anesthesia [3–6]. Most of these monitoring systems use algorithms to detect changes in the EEG spectrum by tracking the transition from a low-amplitude, high-frequency EEG present in the awake state to a high-amplitude, low-frequency EEG during anesthetic-induced unconsciousness [1]. The monitoring systems translate these changes into dimensionless indices that inversely correlate with the hypnotic component of anesthesia. These indices are mostly proprietary,

✉ Matthias Kreuzer
m.kreuzer@tum.de

Max Ebensperger
max.ebensperger@campus.lmu.de

Stephan Kratzer
stephan.kratzer@hessing-stiftung.de

Gerhard Schneider
gerhard.schneider@tum.de

Stefan Schwerin
stefan.schwerin@tum.de

¹ Department of Anesthesiology and Intensive Care, School of Medicine and Health, Technical University of Munich, Ismaningerstr. 22, 81675 Munich, Germany

² Abteilung für Anästhesiologie, Intensiv- und Schmerzmedizin, Hessing Stiftung, Hessingstraße 17, 86199 Augsburg, Germany

and we only have partial insight into how these indices are derived [3–7]. Results from reverse-engineering one of the most commonly used systems, the Bispectral Index™ (BIS), indicate that even the partially provided information may be misleading [8]. Other monitoring systems, like the GE Entropy™ Module with the State Entropy (SE) and the Response Entropy (RE), appear more transparent, as parts of the algorithms were explained [4, 9]. The SE and RE seem to be calculated from the spectral entropy [4, 10], which is the Shannon Entropy [11] applied to the power spectrum of the EEG. Another article explains the steps for calculating the Burst Suppression Ratio (BSR) [9]. The BSR information is straightforward to interpret, it indicates the percentage of suppressed EEG duration within a defined period. The *depth-of-anesthesia* indices are more complex because the output of the respective mathematical algorithm is matched on a scale ranging from 0 to 100 for RE or – as in the case of SE – to 91 [4]. Nonetheless, the exact mechanism by which the algorithm executes this matching process remains obscure.

1.2 Objectives

In our retrospective data analysis, our primary aim was to investigate the presence of age-dependent distribution patterns in SE and RE index values. While examining the large dataset comprising SE, RE, and BSR data, we discovered an unanticipated clustering of particular *pillar* index values, a finding which we elaborate on further in this article. These findings challenge the prevailing claim that SE and RE index values originate from a continuous matching of spectral entropy values onto a *depth-of-anesthesia* scale [4].

2 Methods

2.1 Electronic patient records

For this retrospective, single-center study, we performed an analysis of processed intraoperative EEG data from a cohort of 15,608 patients who underwent general anesthesia. Our hospital's electronic records contained index values of SE, RE, and BSR from the GE Entropy™ module. Additionally, these records included demographic information such as patient age, sex (male or female), Body Mass Index, ASA Physical Status Classification and whether inhalation anesthesia with volatile anesthetic gases or total intravenous anesthesia (TIVA) was used for anesthesia maintenance. Furthermore, the records featured multiple perioperative procedural timestamps and time periods, outlining the total perioperative time. This includes: start of anesthesia, anesthesiological case clearance, surgical incision, and end of anesthesia, in accordance with the

German Perioperative Procedural Time Glossary [12]. We included all patients aged 18 and older in our analysis. The trend data (RE, SE, and BSR) were recorded at 10-second intervals. All analyses and plots were computed using MATLAB R2023a [13].

2.2 Definition of time intervals for analysis

For certain aspects of our analysis, we selected specific time intervals to isolate periods during which patients can be expected to reach a stable state of general anesthesia. To define this interval, we adjusted two of the procedural timestamps to better align with our research objectives. Specifically, we used the timestamp *surgical incision* as a reference point and subtracted 5 min to delineate the beginning of our target period from anesthesia induction with more dynamic pharmacological interventions and concomitant EEG changes. To exclude patients who were awakening from anesthesia, we subtracted 10 min from the *end of anesthesia* timestamp. This approximated differentiation allowed us to evaluate the indices for (i) the entire procedural period (including states of patient wakefulness and awakening in the operating room) and (ii) the surgical core period with likely stable anesthesia-induced unconsciousness.

2.3 Analysis of SE and RE index values

As previously mentioned, both SE and RE index values are derived from spectral entropy analysis. The frequency ranges used for calculating these indices are 0.8–32 Hz for SE and 0.8–47 Hz for RE, as specified in Viertio-Oja et al.'s work [4]. Notably, RE encompasses information from higher-frequency components, capturing a greater extent of recorded muscle activity. In this study, we present the results for both SE and RE separately, as well as their respective difference for each index value pair, Δ (RE–SE). This difference contains valuable information, as a Δ (RE–SE) > 10 might be associated with impending arousal [4, 14]. Both RE and SE are initially computed as spectral entropy values. However, in order to ease interpretation by clinicians, these values are converted into whole integer indices, ranging from 0 to 91 (SE) or 100 (RE). This transformation is purportedly accomplished via application of a monotonic spline function, which is chosen to ensure that the resulting curve is perfectly smooth, devoid of any abrupt changes or irregularities, while still preserving the essential information. It's important to note that this transformation does not evenly space the values from the original scale. This deliberate choice is made to optimize data resolution within the SE original range of 0.5 to 1.0 [4].

2.4 Statistical analysis

For parts of our analyses, we focused on the surgical core period only, because including the values from awake or awakening patients during the entire procedural period with high SE and RE index values would have constituted a biasing factor for our distribution fits. Throughout this study, we explicitly indicate whether the data presented are derived from the entire procedural or the surgical period. We visualize our findings with histograms, heat maps, and trends. In order to calculate probabilities of index value distributions, the data is initially portioned into discrete intervals known as bins. Subsequently, the data points within each of these bins are enumerated. This is followed by a normalization step, where the counts within each bin are divided by the total count of data points. This ensures that the sum of probabilities equals 1. For the indices' distribution during the surgical period, we created an expected, continuous probability distribution by fitting a custom, interpolated, non-Gaussian curve to the histogram of observed probabilities and then subtracted the *pillar* index values above the fit. For this purpose, we applied the MATLAB linear interpolation function *interp* to estimate values for the vacant positions of excluded *pillar* index values. After the fitting, the custom curve was transferred back to the original histogram. This adjustment was made for SE values (2 to 76; manually chosen for visual fit optimization) to provide a more accurate representation of the *pillar* index values. The expected probability distribution was then renormalized. Observed and expected distributions of probabilities were compared using Kullback–Leibler divergence as an information-based measure of disparity (MATLAB function *klDiv*), [15] and the Kolmogorov–Smirnov test as a nonparametric goodness-of-fit test for distributions, using a significance threshold of $p < 0.05$; (MATLAB function *kstest2*), [16]. We further fitted mono-exponential functions [17]. Due to the non-parametric nature of distributions, all values are presented as medians with 25th (Q1) and 75th (Q3) percentiles or as percentages (%) in case of categorical variables.

2.5 Bias

The process of visually recognizing *pillar* index values could potentially introduce unintended biases. Nevertheless, we have taken deliberate steps to mitigate this by explicitly specifying and illustrating the SE and RE index values that we have categorized as *pillar* values and we list them here for the surgical period (3, 5, 6, 7, 9, 11, 12, 14, 15, 16, 18, 20, 22, 24, 25, 26, 27, 28, 29, 30, 31, 33, 35, 37, 41, 45, 51, 59, 66, 72, 77). This transparency extends to our use of expected distribution fits, which includes non-Gaussian modified fits. These measures are aimed at providing a

clearer demarcation between the observed distribution and the expected distribution.

2.6 Missing data

We aimed to clean our large dataset by removing any fully or partially incomplete or invalid index values pertaining to SE, RE, and BSR. This was achieved by implementing exclusion criteria, as demonstrated and detailed in Fig. S1. Additionally, all *NaN* values, which in MATLAB represent non-real or complex numbers, were automatically disregarded in all arithmetic functions as well as in all statistical analyses.

3 Results

3.1 Patient characteristics

The median age of the included patients was 59 [43–72] years. 54% were male and 46% were female. The median surgery duration was 120 [79–179] minutes. Due to the high percentage of incomplete data sets, we excluded any surgery exceeding a duration of 8 hours (273 out of 15,550 patients). A flowchart showing all exclusion criteria is presented in Fig. S1. The corresponding histograms with surgery durations and age distributions are presented as supplemental Fig. S2.

3.2 Distribution of SE and RE index values

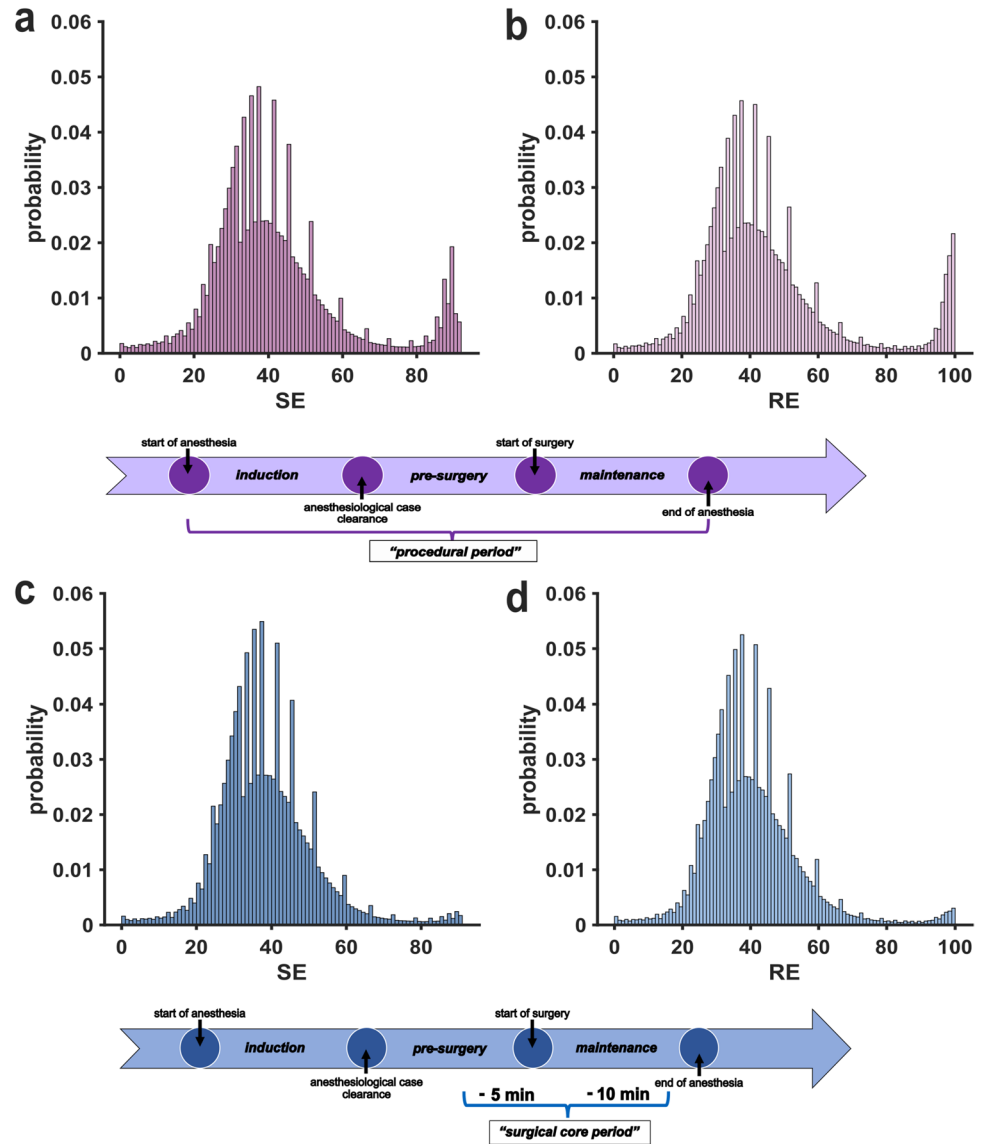
3.2.1 Overall distribution of SE and RE

The first step comprised the visualization of the available index data. Figure 1 contains the histograms that show a bimodal probability distribution for the SE and RE data over the entire procedural period and all age groups combined with index accumulation in the range to be expected under general anesthesia (<60) and in the awake range (>90) (a, b). When only considering the surgical period, the number of high index values (>80) strongly decreases (c, d), as expected. Within these distributions, distinct values for both SE and RE with a sharply increased probability in comparison to directly adjacent values can be visually identified. These peak values, which we term *pillar* indices, were not distributed isometrically. This observation suggests that the algorithm outputs are not equally matched onto the 0–91 (SE) or 0–100 (RE) scale.

3.2.2 Distribution of SE and RE index values stratified by age

The heat maps in Fig. 2 show the index densities as a function of age for the entire procedural period (a, b) and only

Fig. 1 Histograms of state entropy (SE) and response entropy (RE) probability distributions for all age groups. Multiple index values show a distinctively higher probability compared to adjacent values and contrast the expected continuous index value distribution. Below the histogram, respective timelines indicate the relevant time points and periods. **a** Histogram of probabilities for SE index values during the procedural period. **b** Histogram of probabilities for RE index values during the procedural period. **c** Histogram of probabilities for SE index values during the surgical period. **d** Histogram of probabilities for RE index values during the surgical period



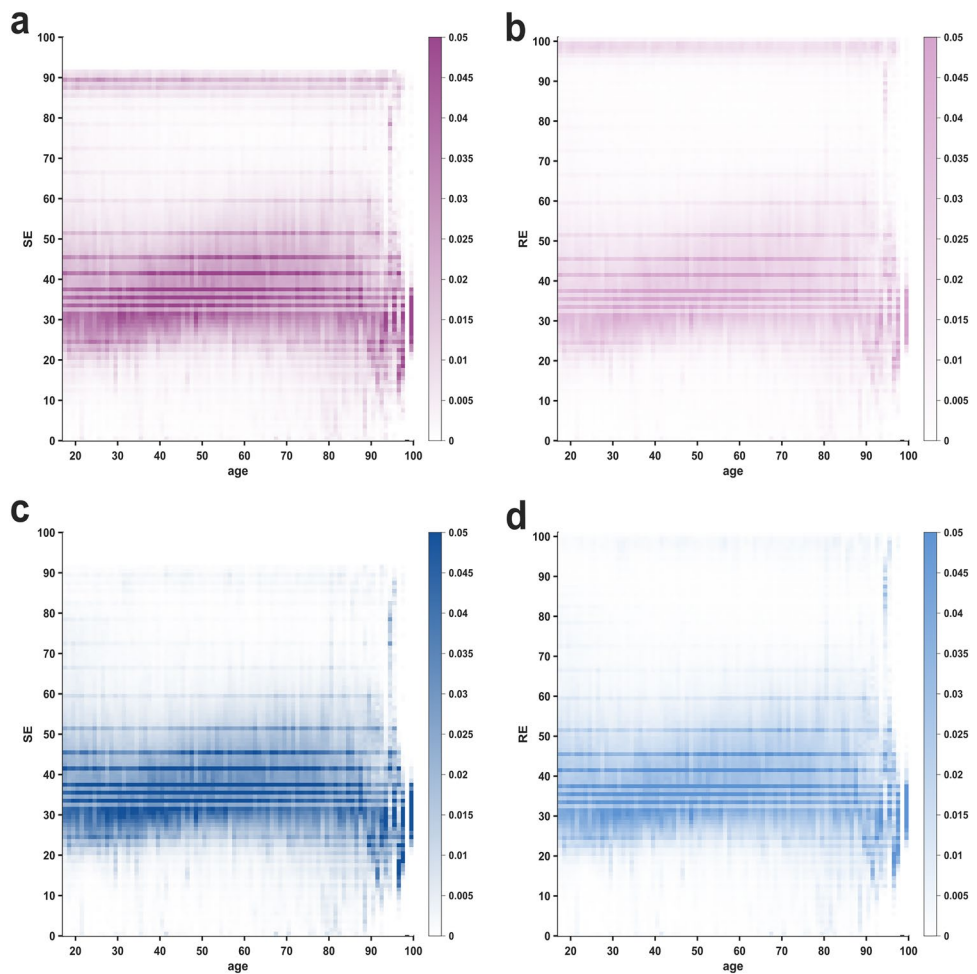
the surgical period (c, d). We observed perfectly horizontal, high-density lines in the heat maps, denoting a significant concentration of particular index values. These values are identical to the *pillar* indices observed in the histograms. The distinct lines show a consistently higher occurrence probability for SE and RE values across all age groups. Their occurrence therefore appears to be independent of patient age, i.e. to be of technical nature. The marked change in distribution patterns beyond the age of 90 years can be attributed to the very limited number of patients included in these age groups, which leads to stronger data dispersion or missing values altogether. As in the histograms, the horizontal lines follow the same, non-isometrical distribution pattern over the index range. Especially in older patients (c, d), we found higher probabilities for very low index values (<30) and high, *awake* values (>80). We also present the heat maps of the differences between RE and SE in supplemental Fig.

S3. Notably, the Δ (RE–SE) does not display distinct *pillar* index values. Instead, the variation appears to be spread in a more continuous manner. This observation suggests that there is a perfect overlap in the *pillar* values between SE and RE.

3.2.3 Analyzing individual trends of SE and RE index values

The histograms and heat maps revealed that *pillar* indices occur much more frequently compared to adjacent index values. To elucidate this phenomenon and to discern the underlying mechanisms, we further analyzed the SE and RE frequency distributions. How SE and RE occurrence behaves on an individual level is exemplified in Fig. 3. It shows the SE and RE trend data of different patients during their respective surgical interventions with two different time scales. The trends covering the entire surgical period show a

Fig. 2 Heat maps of state entropy (SE) and response entropy (RE) probability distributions as a function of age. The horizontal, darkened lines indicate a distinctively higher probability occurrence for the respective SE and RE index values as well as a high consistency over the entire age range. Greater data dispersion in the age groups over 90 years is attributed to the limited sample size of patients above 90 years of age. **a** Heat map of SE index values during the procedural period. **b** Heat map of RE index values during the procedural period. **c** Heat map of SE index values during the surgical period. **d** Heat map of RE index values during the surgical period



continuous course of the indices, which seemingly fluctuate freely in the range recommended for anesthesia (40–60) and there is no indication of the described *pillar* index values, see Fig. 3a. However, upon examination of the trends with increased temporal resolution, we noticed that the indices appear to linger at certain values for a prolonged duration. Furthermore, the indices exhibit peak points with subsequent trend reversion, which appear to coincide with *pillar* index values and may account for their more frequent occurrence (see Fig. 3b).

3.3 Time distribution across specific SE and RE index values

3.3.1 Maximum durations and consecutive runs for SE and RE index values

We further evaluated the cumulative time spent in individual index values and the maximum duration, during which an index value did not change. The histograms in supplemental Fig. S4 show that the cumulative time spent in the previously identified *pillar* indices was, as one might expect,

much higher. The maximum duration during which the index remained unchanged was also higher in the case of *pillar* index values. In Fig. 4, we display the maximum consecutive runs (10-second epochs per run) of individual SE and RE index values, both for the entire procedural and the surgical period, respectively. Because it may be argued that the maximum run duration was influenced by artifacts, we also included the 99th and 95th percentiles.

3.3.2 Consecutive occurrence of *pillar* indices and adjacent index values

Having established that *pillar* SE and RE index values occur with greater frequency and longer duration than other index values, we extended our analysis. We compared these *pillar* indices to directly adjacent index values and plotted the number of observations against consecutive occurrences. Both *pillar* index values and the respective adjacent index values are well described by negative exponential functions and show excellent fits ($R^2 > 0.99$). However, *pillar* index values demonstrate differences in both the rate of decay as well as the vertical shift. Across different segment lengths,

Fig. 3 Trend data of state entropy (SE) and response entropy (RE) values recorded during the surgical period from four exemplary patients of different age groups. SE and RE data are presented in dark and light blue. Procedural timestamps are marked out vertically and *pillar* index values are indicated as grey horizontal lines. **a** Trend data for four patients throughout the entire surgical duration showed no discernible *pillar* indices and fell within the anticipated value range. **b** Sections with increased temporal resolution from the same four patients. Visible are plateauing values as well as peak values, which appear to constitute turning points and coincide with *pillar* index values

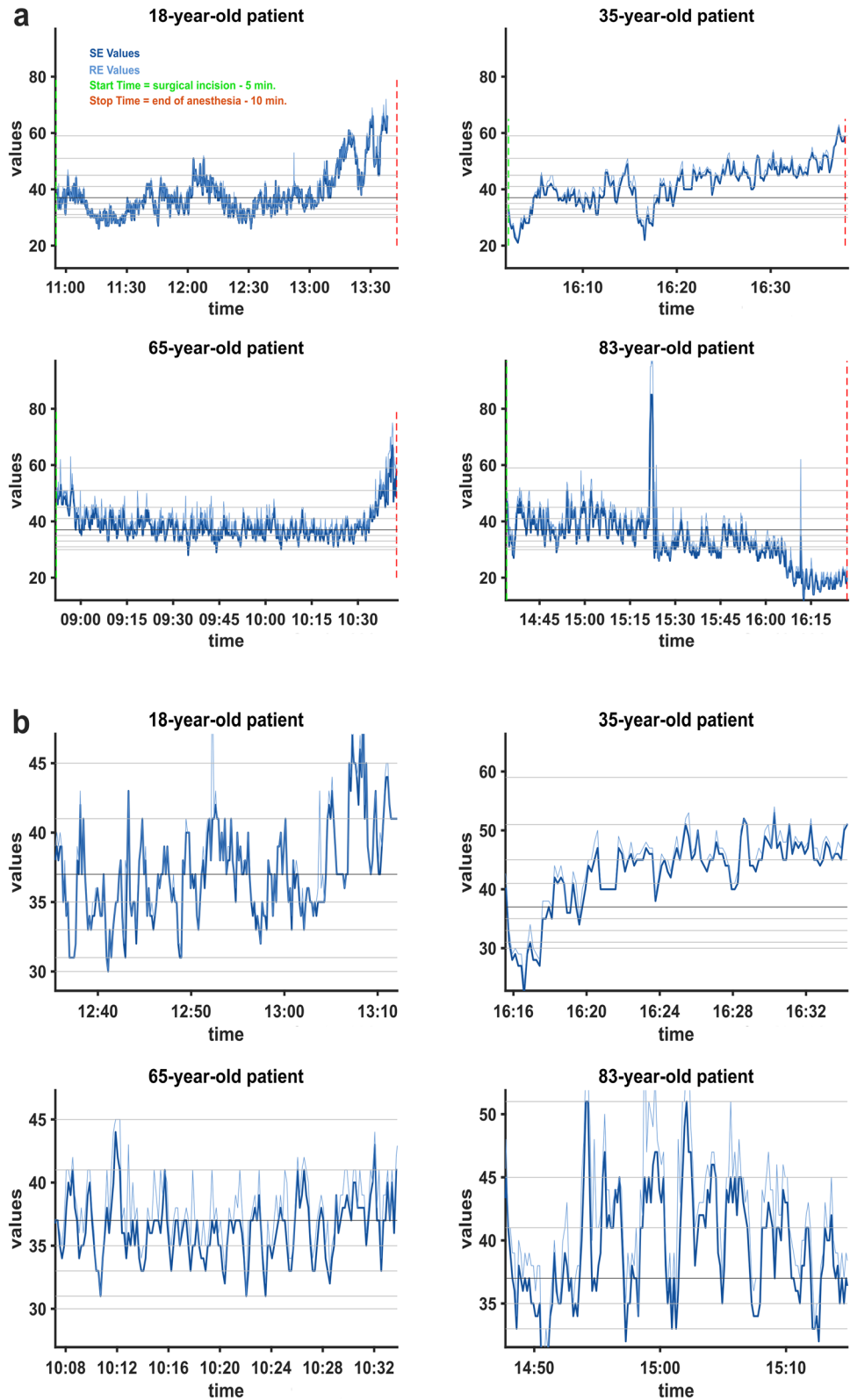
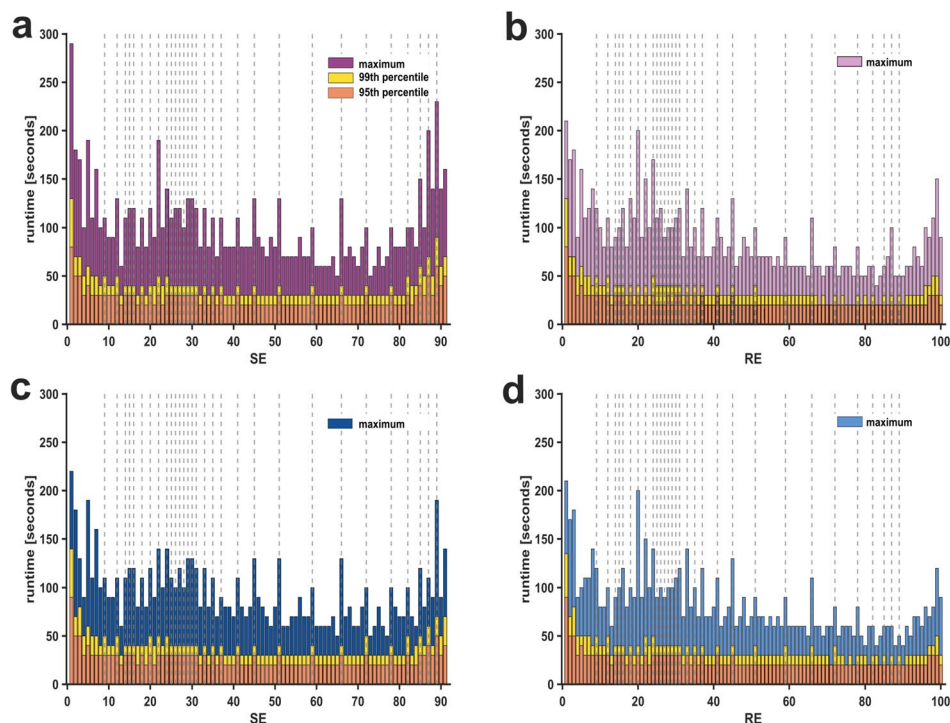


Fig. 4 Maximum consecutive runtime spent in state entropy (SE) and response entropy (RE) values for the entire procedural period and the surgical period. The time spent in specific SE and RE index values indicated as consecutive runs with 10-second epochs per run. Maximum time spent, 99th, and 95th percentile are displayed separately. The previously identified *pillar* index values are highlighted with vertical grey dotted lines. **a** Consecutive runs of SE indices in the procedural period. **b** Consecutive runs of RE indices in the procedural period. **c** Consecutive runs of SE indices in the surgical period. **d** Consecutive runs of RE indices in the surgical period



pillar indices were over-represented compared to adjacent indices. Further details can be found in the supplemental Fig. S5 and Table S1, which contains the parameters of the exponential fits.

3.4 Interpolation for SE index values during the surgical period

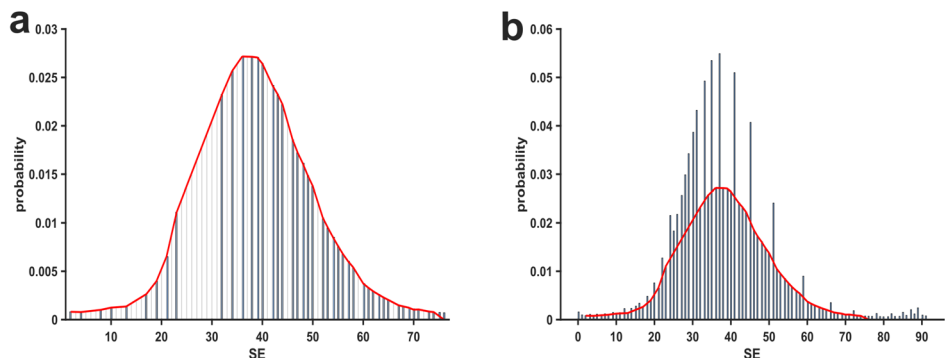
To evaluate the percentage of *pillar* index values above an expected distribution (presupposing a pattern of continuously increasing probabilities towards a peak value before continuously decreasing), we manually modeled a fit based on the (visually identified) non-peak values, as shown in Fig. 5a. The fit reveals a nearly symmetrical distribution around a maximum point of SE equal to 36. The fraction of

excess probability above the fit as depicted in Fig. 5b was 25.46% for SE index values during the surgical period.

3.5 Analyzing distributional discrepancies: Kullback–Leibler divergence and Kolmogorov–Smirnov tests

In order to evaluate the differences between observed and expected distributions, we first calculated the Kullback–Leibler divergence (KL), which was 0.048. To contextualize this metric more meaningfully, we also tested the interpolated values against a linear, *ideal* value distribution, resulting in a KL divergence of 0.465. Additionally, we compared the observed probability distribution against the same linear, ideal probability distribution, obtaining a KL divergence of 0.553. These results indicate that the interpolated probability

Fig. 5 Fit modeling and peak identification of State Entropy (SE) index values. **a** Histogram of the SE index values that were included to create a custom, non-gaussian fit, superimposed as a red line. **b** Histogram of SE index values including *pillar* index values and the modeled fit superimposed in red, which was used to quantify the excess probability of *pillar* index values



distribution only deviates by approximately one-tenth from the observed distribution in comparison to linear values. Furthermore, we performed the Kolmogorov–Smirnov test twice. First, we used the *Ksstat* function to compare the measured probability distribution against a normal distribution (CDF), which produced a KS statistic (D) of 0.176 ($p = 0.018$). This significant difference between the two distributions suggests that the measured probabilities differ significantly from a normal distribution. Secondly, we used the *Kstest2* to compare the measured probabilities against our interpolated probabilities. In this analysis, the KS statistic (D) was found to be 0.095 ($p = 0.879$), indicating that the probability distribution of the observed data is consistent with the one of interpolated, *expected* probability distribution.

3.6 Evaluation of burst suppression ratio (BSR)

When evaluating the BSR index with the same histogram and heat map approach as SE and RE for the entire procedural period, we found that the BSR did not show *pillar* index values, but a rather continuous decline in occurrence probability, which could be well described with an exponential decay function ($R^2 = 0.948$), as depicted in Fig. 6a. The heat map visualizing the probability distribution of BSR values across the age range showed a continuous behavior without horizontal lines, in contrast to RE and SE. The BSR heat map is presented in Fig. 6b. To prevent scaling distortions in the plots, only BSR values > 0 are included. This is because the most frequently observed BSR value was zero, representing absence of any burst suppression as detected by the algorithm, and accounted for 74.24% of the whole BSR dataset. When plotting the BSR indices as a function of SE indices, the *pillar* indices are readily discernible as horizontal lines, while there are no vertical lines for BSR,

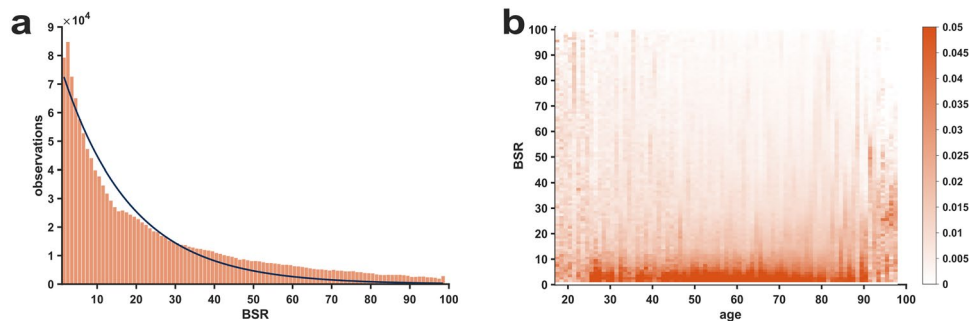


Fig. 6 Distribution of Burst Suppression Ratio (BSR) indices > 0 for the entire procedural period. For clarity, BSR of 0 was omitted as it accounted for approximately three-fourths of all BSR values. **a** Histogram showing the count distribution of BSR values > 0 . The decay appears continuous without discrete *pillar* indices and can be

further suggesting a continuous distribution of BSR values, see Fig. S6.

3.7 Exploration of potential confounders

To identify potential confounders, we further investigated the distribution of SE/RE index values depending on Body Mass Index, ASA Physical Status Classification, surgery department, and the choice between the use of volatile anesthetic gases or total intravenous anesthesia (TIVA) for the surgical period. In our analysis, the *pillar* index values persisted for all conditions, and no confounding factors could be identified. The histograms showcasing the distribution of SE/RE values grouped according to volatile anesthetic gases versus the use of TIVA are presented in Fig. S7.

4 Discussion

Every manufacturer of EEG-based monitoring systems emphasizes the primacy of processed index values for navigating the level of anesthesia. The utilization of these indices, like BIS and other systems, was discussed as a means to prevent intraoperative awareness [18], although some publications show contradicting results [19, 20]. Other publications have demonstrated that these indices show a (mostly unknown) temporal lag [21, 22], which may depend on artifact content and in part on the electromyographic component [23, 24], or that they overlook or misinterpret burst suppression [25–27]. While the algorithms generating these indices are proprietary, other freely accessible parameters have demonstrated effective – and in some cases even superior – performance in monitoring the anesthetic level [28, 29]. Therefore, relying on an index whose precise design is not fully disclosed may not be the most advisable approach to establish trust both in clinical practice and research settings,

well described by an exponential decay function, which is superimposed in the figure. **b** Heat map of BSR > 0 -probability distribution, stratified by age, that shows a continuous fading of probability with increasing BSR index values

particularly in light of these limitations. There are publications that have uncovered substantial components of the BIS algorithm via reverse engineering, showing its reliance on the low-gamma waveband and contradicting earlier beliefs about its dependence on a bispectral index [8, 30]. Additionally, there have been instances highlighting contradictory information provided by simultaneous BSR and SE outputs [31].

Our results point towards discrepancies between the prevailing explanation of how the SE/RE indices are calculated and what we can observe clinically. As stated in the manuscript describing the algorithm [4], the spectral entropy is matched onto the 0–91 (SE) or 0–100 (RE) scale per spline interpolation. Subsequently, the spline function is characterized as smooth without “kinks”. This leads to the assumption that the index distribution was also intended to be smooth and not contain *pillar* index values with overly high occurrence probability, which we found in our dataset and describe in this study. We could also observe that in the RE and SE range of approximately 30 to 90, the *pillar* index values were further apart from each other than in the lower index range. For the higher range with values between 50 and 90, the presented spline function in the description paper was also steepest [4].

While the clinical implications of our findings have yet to be explored in further detail, they could indeed impact the use of the indices for research purposes. Different monitoring systems were shown to provide discordant recommendations based on identical EEG recordings with patterns of emergence from anesthesia [32]. Much like other clinical parameters that require monitoring during general anesthesia — such as electrocardiography — the interpretation and classification guiding clinical decisions should not be contingent upon the specific monitoring device employed. Further studies have described discrepancies between processed EEG monitoring devices, for example, simultaneously recorded BIS and SE index values during general anesthesia with sevoflurane [33]. The underlying uncertainties and deviations from the expected behavior of processed EEG indices might be particularly problematic when studies compare or aim for the identical target value range with different monitoring systems [34]. Comparability is further impeded by the observation that processed EEG indices, specifically SE/RE and BIS demonstrate an extensive inter-individual variability with regard to defined clinical endpoints as well as anesthetic-specific differences [35]. Another study describes that SE and RE values are not well matched by two different BIS target corridors [36], which might be partially explained by discontinuous *pillar* indices. When comparing anesthesiologists’ assessments of BIS or SE, the SE “errors” were clustered while BIS “errors” were more continuously distributed, maybe reflecting the non-continuous SE index value distribution. In another study, anesthesiologists tried to

titrate anesthesia to an SE index corridor between 40 and 60, but had difficulties to not go below these limits or to get back into the limits from lower SE index values [37]. Ultimately, only 45% of the recorded values fell within the predetermined range. The skewed distribution leaning towards low index values is particularly evident in Supplemental Fig. 2 of the study, suggesting that achieving the desired depth of anesthesia as indicated by the Entropy module poses a challenge even when it is the primary target intervention of a prospective study. In a separate study, 76 patients in the intervention group received anesthesia based on the Surgical Pleth Index™ (GE Healthcare) and SE, targeting an SE range of 40 to 60. Interestingly, the mean SE values for both the control and intervention groups leaned towards the lower end of the range. Specifically, during the period from after intubation to during surgery, the intervention group exhibited mean SE values ranging between 40–45, while the control group showed values between 39–43. No significant differences were observed between the groups [38]. While this evidence may be relatively modest, it is noteworthy that when aiming explicitly for a SE range of 40 to 60 on what is presumed to be a continuous scale, one might anticipate mean values closer to 50. Another group also noticed a strong tendency towards SE index values below 40 during anesthesia maintenance. The mean values reported for their measurement points correlate in parts with the *pillar* indices found in this study, i.e., a SE of 30, 37, or 41 [39]. These findings could suggest that the Entropy algorithm itself may generate skewed distributions, displaying a tendency toward lower index values. Schmidt and colleagues presented figures, where, similarly to our findings, index values with a higher probability of occurrence can be identified [40] and Vanluchene and colleagues described a smoother behavior for the BIS index values rather than SE [41]. However, even the BIS algorithm seems to generate peak values as well as sudden changes of index value occurrence at distinct boundary values [42].

4.1 Key results

Researchers and clinical anesthesiologists should be aware that the reduction of the available EEG signal into an index leads to a strong loss of information. If, as is the case with the discussed commercial indices, the algorithms are not fully disclosed, the optimal assessment of a patient’s anesthetic level is hampered. Overall, these findings advocate for a broader perspective in monitoring practices, moving beyond an index-centered approach. The information provided by EEG-based monitoring devices can offer valuable insights for navigating anesthesia, yet the overall management should integrate all information available. Barnard and colleagues stated that the combination of using the index and being capable of interpreting the raw EEG may be the best

strategy to safely navigate a patient under general anesthesia [43]. When relying on a single number, parameters that are better understandable – like the spectral edge frequency or the individually calculated spectral entropy – may be the preferential choice. EEG, in its own right, serves as a potent tool for evaluating the anesthesia state of patients, and in the future, shifting the focus away from proprietary indices in favor of assessing the raw EEG as well as more refined EEG-based monitoring tools, could be a more promising direction.

4.2 Limitations

Due to its retrospective nature, this study has several limitations. We did not have access to raw EEG data or EEG band-power, which would have been valuable for a more in-depth investigation of *pillar* index values. Anesthesia regimes were not explicitly administered according to a predefined target range for SE and RE values (however, the manufacturer recommended range of 40 to 60 is included in the monitor output). This may introduce a potential confounding factor and could partially contribute to a higher probability of low index values. We further did not include an in-depth analysis of the relationship between SE/RE/BSR-index values and demographic parameters, drug concentrations, or intraoperative hemodynamic parameters. This decision was based on the fact that we are addressing a technical phenomenon within a universally applicable (*one-size-fits-all*) system, which, by default, does not consider these factors [4].

5 Conclusion

Given the discontinuous behavior of SE and RE, it appears that there is not a consistent alignment of the anesthetic level on the index scale. To enhance comprehension of the processed indices, it is imperative that the algorithms employed are accessible to all. EEG is an essential information source for effective patient monitoring. Anesthesiologists should also possess the skills to interpret raw EEG data and its spectral representation to address any deficiencies in the processed indices.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10877-024-01130-9>.

Author contributions All authors contributed to the study conception and design. Data collection and analysis were performed by Max Ebensperger, Matthias Keuzer, and Stefan Schwerin. The first draft of the manuscript was written by Max Ebensperger and Stefan Schwerin. Matthias Kreuzer, Stephan Kratzer, and Gerhard Schneider critically discussed the results and revised the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Declarations

Competing interests Matthias Kreuzer is named as an inventor for a patent dealing with spectral EEG features and age (U.S. Provisional Patent Application No. 62/914,183). Gerhard Schneider and Matthias Kreuzer are named as inventors for a patent filed on a novel method for intraoperative EEG monitoring (U.S. Patent Application Serial No. 62/960,947). Gerhard Schneider, Matthias Kreuzer, and Stephan Kratzer are also named as inventors for a patent dealing with the EEG features during anesthesia emergence (U.S. Provisional Patent Application No. 63/459,294). Stefan Schwerin and Max Ebensperger declare no conflict of interest.

Ethical approval The institutional ethics committee of Technical University of Munich/rechts der Isar approved the study on 12 December 2018 (No. 572/18 s; Chairperson: Prof. Schmidt) and waived the requirement for written informed consent of patients or the next of kin. Individual patients cannot be identified based on the data included in this study. Electronic records from the years 2016 to 2017 were included for analysis. This manuscript adheres to the applicable STROBE guidelines [44].

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Brown EN, Lydic R, Schiff ND. General anesthesia, sleep, and coma. *N Engl J Med*. 2010;363(27):2638–50. <https://doi.org/10.1056/NEJMr0808281>.
2. Teplan M. Fundamentals of EEG measurement. *Meas Sci Rev*. 2002;2(2):1–11.
3. Rampil IJ. A primer for EEG signal processing in anesthesia. *Anesthesiology*. 1998;89:980–1002. <https://doi.org/10.1097/0000542-199810000-00023>.
4. Viertio-Oja H, Maja V, Sarkela M, Talja P, Tenkanen N, Tolvanen-Laakso H, Paloheimo M, Vakkuri A, Yli-Hankala A, Merilainen P. Description of the entropy algorithm as applied in the Datex-Ohmeda S/5 entropy module. *Acta Anaesthesiol Scand*. 2004;48(2):154–61. <https://doi.org/10.1111/j.0001-5172.2004.00322.x>.
5. Drover D, Ortega HR. Patient state index. *Best Pract Res Clin Anaesthesiol*. 2006;20(1):121–8. <https://doi.org/10.1016/j.bpa.2005.07.008>.
6. Jensen E, Valencia J, López A, Anglada T, Agustí M, Ramos Y, Serra R, Jospin M, Pineda P, Gambus P. Monitoring hypnotic effect and nociception with two EEG-derived indices, qCON and

- qNOX, during general anaesthesia. *Acta Anaesthesiol Scand.* 2014;58(8):933–41. <https://doi.org/10.1111/aas.12359>.
7. Kreuer S, Biedler A, Larsen R, Altmann S, Wilhelm W. Narcotrend monitoring allows faster emergence and a reduction of drug consumption in propofol-remifentanyl anesthesia. *Anesthesiology.* 2003;99(1):34–41. <https://doi.org/10.1097/00000542-200307000-00009>.
 8. Connor CW. Open reimplementations of the BIS algorithms for depth of anesthesia. *Anesth Analg.* 2022;135(4):855–64. <https://doi.org/10.1213/ANE.0000000000006119>.
 9. Särkelä M, Mustola S, Seppänen T, Koskinen M, Lepola P, Suominen K, Juvonen T, Tolvanen-Laakso H, Jääntti V. Automatic analysis and monitoring of burst suppression in anesthesia. *J Clin Monit Comput.* 2002;17(2):125. <https://doi.org/10.1023/a:1016393904439>.
 10. Steyn-Ross DA, Steyn-Ross ML, Wilcocks LC, Sleight JW. Toward a theory of the general-anesthetic-induced phase transition of the cerebral cortex. II. numerical simulations, spectral entropy, and correlation times. *Phys Rev E.* 2001;64(1): 011918. <https://doi.org/10.1103/PhysRevE.64.011917>.
 11. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J.* 1948;27(3):379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
 12. Bauer M., Auhuber T., Kraus R., Rüggeberg J., Wardemann K., Mueller P., Taube C., Diemer M., Schuster M., (2020) The german perioperative procedural time glossary (2020 edition*) a joint recommendation by the bda, bdc, vopm, vopmo, ogari and sfopm. *Anästhesiologie und Intensivmedizin* 61, 516–531 <https://doi.org/10.19224/ai2020.516>
 13. Inc. T.M.: MATLAB Version: 9.14.0.2337262 (R2023a). <https://www.mathworks.com>
 14. Wheeler P, Hoffman WE, Baughman VL, Koenig H. Response entropy increases during painful stimulation. *J Neurosurg Anesthesiol.* 2005;17(2):86–90. <https://doi.org/10.1097/01.ana.0000151408.62650.b5>.
 15. Joyce J.M. Kullback-leibler divergence. In: *International Encyclopedia of Statistical Science*, pp. 720–722. Springer, Berlin, (2011). https://doi.org/10.1007/978-3-642-04898-2_327
 16. Dodge Y. *The Concise Encyclopedia of Statistics*. Springer, New York (2008). https://doi.org/10.1007/978-0-387-32833-1_214
 17. Clapham C, Nicholson J, Nicholson JR. *The Concise Oxford Dictionary of Mathematics*. Oxford University Press, Oxford; 2014. <https://doi.org/10.1093/acref/9780199235940.001.0001>
 18. Myles PS, Leslie K, McNeil J, Forbes A, Chan MTV. Bispectral index monitoring to prevent awareness during anaesthesia: the B-Aware randomised controlled trial. *Lancet.* 2004;363(9423):1757–63. [https://doi.org/10.1016/S0140-6736\(04\)16300-9](https://doi.org/10.1016/S0140-6736(04)16300-9).
 19. Russell IF. The Narcotrend ‘depth of anaesthesia’ monitor cannot reliably detect consciousness during general anaesthesia: an investigation using the isolated forearm technique. *Br J Anaesth.* 2006;96(3):346–52. <https://doi.org/10.1093/bja/ael017>.
 20. Avidan MS, Zhang L, Burnside BA, Finkel KJ, Searleman AC, Selvidge JA, Saager L, Turner MS, Rao S, Bottros M, et al. Anesthesia awareness and the bispectral index. *N Engl J Med.* 2008;358(11):1097–108. <https://doi.org/10.1056/NEJMoa0707361>.
 21. Pilge S, Zanner R, Schneider G, Blum J, Kreuzer M, Kochs E. Time delay of index calculation: analysis of cerebral state, bispectral, and narcotrend indices. *Anesthesiology.* 2006;104(3):488–94. <https://doi.org/10.1097/00000542-200603000-00016>.
 22. Kreuzer M, Zanner R, Pilge S, Paprotny S, Kochs EF, Schneider G. Time delay of monitors of the hypnotic component of anesthesia: analysis of state entropy and index of consciousness. *Anesth Analg.* 2012;115(2):315–9. <https://doi.org/10.1213/ANE.0b013e31825801ea>.
 23. Messner M, Beese U, Romstöck J, Dinkel M, Tschaikowsky K. The bispectral index declines during neuromuscular block in fully awake persons. *Anesth Analg.* 2003;97(2):488–91. <https://doi.org/10.1213/01.ANE.0000072741.78244.C0>.
 24. Schuller PJ, Newell S, Strickland PA, Barry JJ. Response of bispectral index to neuromuscular block in awake volunteers. *Br J Anaesth.* 2015;115(Suppl 1):95–103. <https://doi.org/10.1093/bja/aev072>.
 25. Hart S, Buchannan C, Sleight J. A failure of m-entropy™ to correctly detect burst suppression leading to sevoflurane overdosage. *Anaesth Intensive Care.* 2009;37(6):1002–4. <https://doi.org/10.1177/0310057X0903700619>.
 26. Muhlhöfer W, Zak R, Kamal T, Rizvi B, Sands L, Yuan M, Zhang X, Leung J. Burst-suppression ratio underestimates absolute duration of electroencephalogram suppression compared with visual analysis of intraoperative electroencephalogram. *Br J Anaesth.* 2017;118(5):755–61. <https://doi.org/10.1093/bja/aex054>.
 27. Fleischmann A, Georgii M-T, Schuessler J, Schneider G, Pilge S, Kreuzer M. Always assess the raw electroencephalogram: why automated burst suppression detection may not detect all episodes. *Anesth Analg.* 2023;136(2):346–54. <https://doi.org/10.1213/ANE.0000000000006098>.
 28. Olofsen E, Sleight JW, Dahan A. Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect. *Br J Anaesth.* 2008;101(6):810–21. <https://doi.org/10.1093/bja/aen290>.
 29. Schneider G, Jordan D, Schwarz G, Bischoff P, Kalkman CJ, Kuppe H, Rundshagen I, Omerovic A, Kreuzer M, Stockmanns G, Kochs EF. Monitoring depth of anesthesia utilizing a combination of electroencephalographic and standard measures. *Anesthesiology.* 2014;120(4):819–28. <https://doi.org/10.1097/ALN.000000000000151>.
 30. Bruhn J, Bouillon TW, Shafer SL. Bispectral index (BIS) and burst suppression: revealing a part of the BIS algorithm. *J Clin Monit Comput.* 2000;16:593–6. <https://doi.org/10.1023/A:1012216600170>.
 31. Georgii M-T, Pilge S, Schneider G, Kreuzer M. State entropy and burst suppression ratio can show contradictory information: a retrospective study. *Eur J Anaesthesiol.* 2020;37(12):1084–92. <https://doi.org/10.1097/EJA.0000000000001312>.
 32. Hight D, Kreuzer M, Ugen G, Schuller P, Stüber F, Sleight J, Kaiser HA. Five commercial ‘depth of anaesthesia’ monitors provide discordant clinical recommendations in response to identical emergence-like EEG signals. *Br J Anaesth.* 2023;130(5):536–45. <https://doi.org/10.1016/j.bja.2022.12.026>.
 33. Aho A, Kamata K, Jääntti V, Kulkas A, Hagihira S, Huhtala H, Yli-Hankala A. Comparison of bispectral index and entropy values with electroencephalogram during surgical anaesthesia with sevoflurane. *Br J Anaesth.* 2015;115(2):258–66. <https://doi.org/10.1093/bja/aev206>.
 34. Aimé I, Verroust N, Masson-Lefoll C, Taylor G, Laloë P, Liu N, Fischler M. Does monitoring bispectral index or spectral entropy reduce sevoflurane use? *Anesth Analg.* 2006;103(6):1469–77. <https://doi.org/10.1213/01.ane.0000246838.93153.23>.
 35. Kaskinoro K, Maksimow A, Långsjö J, Aantaa R, Jääskeläinen S, Kaisti K, Särkelä M, Scheinin H. Wide inter-individual variability of bispectral index and spectral entropy at loss of consciousness during increasing concentrations of dexmedetomidine, propofol, and sevoflurane. *Br J Anaesth.* 2011;107(4):573–80. <https://doi.org/10.1093/bja/aer196>.
 36. Lehmann A, Schmidt M, Zeitler C, Kiessling A-H, Isagro F, Boldt J. Bispectral index and electroencephalographic entropy in patients undergoing aortocoronary bypass grafting. *Eur J Anaesthesiol.* 2007;24(9):751–60. <https://doi.org/10.1017/S0265021506002249>.

37. Gruenewald M, Harju J, Preckel B, Molnár Z, Yli-Hankala A, Roskopf F, Koers L, Orban A, Bein B. Comparison of adequacy of anaesthesia monitoring with standard clinical practice monitoring during routine general anaesthesia: An international, multicentre, single-blinded randomised controlled trial. *Eur J Anaesthesiol.* 2021;38(1):73. <https://doi.org/10.1097/EJA.0000000000001357>.
38. Bergmann I, Göhner A, Crozier TA, Hesjedal B, Wiese CH, Popov A, Bauer M, Hinz JM. Surgical pleth index-guided remifentanil administration reduces remifentanil and propofol consumption and shortens recovery times in outpatient anaesthesia. *Br J Anaesth.* 2013;110(4):622–8. <https://doi.org/10.1093/bja/aes426>.
39. Hasak L, Wujtewicz M, Owczuk R. Assessment of the depth of anaesthesia during inhalational and intravenous induction of general anaesthesia. *Anaesthesiol Intensive Ther.* 2014;46(4):274–9. <https://doi.org/10.5603/AIT.2014.0044>.
40. Schmidt GN, Bischoff P, Standl T, Hellstern A, Teuber O, Esch J. Comparative evaluation of the datex-ohmeda s/5 entropy module and the bispectral index® monitor during propofol-remifentanil anaesthesia. *J Am Soc Anesthesiol.* 2004;101(6):1283–90. <https://doi.org/10.1097/00000542-200412000-00007>.
41. Vanluchene AL, Struys MM, Heyse BE, Mortier EP. Spectral entropy measurement of patient responsiveness during propofol and remifentanil. A comparison with the bispectral index. *Br J Anaesth.* 2004;93(5):645–54. <https://doi.org/10.1093/bja/ae251>.
42. Lee H-C, Ryu H-G, Park Y, Yoon SB, Yang SM, Oh H-W, Jung C-W. Data driven investigation of bispectral index algorithm. *Sci Rep.* 2019;9(1):13769. <https://doi.org/10.1038/s41598-019-50391-x>.
43. Barnard J, Bennett C, Voss L, Sleigh J. Can anaesthetists be taught to interpret the effects of general anaesthesia on the electroencephalogram? comparison of performance with the BIS and spectral entropy. *Br J Anaesth.* 2007;99(4):532–7. <https://doi.org/10.1093/bja/aem198>.
44. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (strobe) statement: guidelines for reporting observational studies. *Lancet.* 2007;370(9596):1453–7. <https://doi.org/10.7326/0003-4819-147-8-200710160-00010>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.