BIS and state entropy of the EEG - comparing apples and oranges

S. Pilge* and G. Schneider

Department of Anaesthesiology, HELIOS Clinic Wuppertal, Witten/Herdecke University, Heusnerstr. 40, Wuppertal 42283, Germany

*Corresponding author. E-mail: stefanie.pilge@uni-wh.de

Monitoring of anaesthetic effects on the main target organ of anaesthesia, the brain, has evolved from a research tool to a clinical method and may help to adjust anaesthetic drugs to specific individual requirements (i.e., to avoid inadequate anaesthetic depth with the risk of awareness or excessively deep hypnosis).

The term ‘anaesthetic depth’ is poorly defined and dates back to a period when general anaesthesia was induced by a single substance. A combination of effects and side effects were used to quantify ‘depth of anaesthesia’. Today, a combination of drugs is used to induce general anaesthesia, and several aspects of anaesthetic effects are induced by specific drugs. Consequently, present definitions try to follow this combinatorial approach. Different ‘components’ of anaesthesia seem to follow the different types of drugs administered to induce general anaesthesia. This results in the core problem: the difficulty to find an all-embracing definition of anaesthetic depth—which would be the basis to develop suitable assessment methods. Thus, the present definition must remain very vague: according to the ASA task force, depth of anaesthesia or depth of hypnosis refers to a continuum of progressive central nervous system depression and decreased responsiveness to stimulation. The assessment of clinical signs as responses to stimulation has been established in several clinical sedation scores. This definition may not entirely follow the present concept of general anaesthesia. Responsiveness to stimulation may reflect the ‘hypnotic component’, but may also be influenced by the ‘antinociceptive component’, because analgesia and antinociception may also reduce responses to painful stimuli. It remains unclear whether a separation in these different components, as driven by classes of drugs is clearly possible. In addition, many drug effects seem to follow an ‘all-or-none’ principle, or a very steep drug response curve. Consequently, the concept of ‘depth’ has been questioned, because a human may not be a submarine during anaesthesia. Still, an increase of anaesthetic drugs induces graded changes of the clinical status of a patient.

The use of clinical signs may not be reliable in measuring the hypnotic component of anaesthesia. The key features of EEG changes in response to different doses of anaesthetic agents and clinical state responses were established and suggest the clinical feasibility of depth of anaesthesia (DOA) monitors: these are devices that record or process brain electrical activity, usually from frontal EEG, as a surrogate for the drug effects on the entire central nervous system (CNS). Several parameters are calculated from the EEG, and these parameters are combined and mathematically converted into a continuous measure typically scaled from 0 to 100.

Clinical measures of consciousness represent an important component in the development process of an EEG based index. As far as it is known, the majority of commercially available EEG based monitors were validated with different clinical scales during their development process (e.g., the manufacturers of bispectral index (BIS), cerebral state index and patient state index monitor referred to the modified OAA/S score, to define which EEG based subparameters, or a combination of them correlated best with the clinical state). In contrast to other indices, entropy based EEG monitoring calculates entropy of the EEG from awake to deep anaesthesia (i.e., the same parameter is calculated over a wide range, only the time window of the EEG increases with increasing drug effects). The time-frequency balanced spectral entropy is calculated by the Datex-Ohmeda Entropy Module. With this approach, clinical signs are not necessarily required for the combination of sub-parameters, but for the scaling of the index values.

Aside from the underlying algorithm—each index was developed on the basis of a separate training data set with different patients or volunteers and different anaesthetic regimens.

As a consequence, the performance of a monitor can only be as good as the corresponding database. Furthermore, the ability of an index value to recognize the hypnotic state can only be as reliable as the underlying method to classify the clinical state during development process.

Today, different EEG-based monitors and indices are on the market. Mostly, the underlying algorithms remain proprietary. There is neither a possibility to thoroughly comprehend the performance of one DOA index nor to explain outliers—both would be substantial for a valid interpretation of EEG analysis and for the reasonable transfer of these findings to clinical context.

Furthermore, comparability of different DOA monitors is limited. There is no reason why different DOA index values should show similar trajectories across time, because they are based on different processing times, different artifact rejection algorithms and different calculation algorithms for generating an index.

But how to compare apples and oranges?

One possible approach for the comparison of different DOA monitors may be the analysis of the unprocessed EEG sequences that underlie index computation. Information about the underlying raw (time domain) EEG signals may allow for the evaluation of the potential influence of distinct EEG findings/artifacts on index calculation: index calculation may be influenced because
of physiological EEG pattern (e.g. beta or delta arousal), pathophysiological features (e.g. epileptiform pattern, electrocogragram, electromyogram (EMG)), or technical artifacts (e.g. from weak electrode contact or cautery). Such an approach may contribute to the elucidation of the mechanistic causes of discrepancies/outliers between different index values: if EEG is influenced by ‘pathology’, indices may be misleading. If two indices disagree, one of them or both can be ‘wrong’ because of these reasons.

So far, only one comparative study has used the information of unprocessed EEG for such a comparison. In an offline reanalysis of EEG data from two anaesthetic regimens (propofol and sevoflurane anaesthesia), EEG was inspected for potential causes of erroneous index calculation of BIS and state entropy (SE) during analysis of identical EEG signals. With information about the level of consciousness, EEG sequences were identified which produced index values that were in conflict with clinical examination. These sequences were analysed with respect to frequency characteristics (frequency bands/SEF95), and visually inspected for physiological EEG pattern and pathophysiological features. High frequency signals and eye blinks were present in the majority of all EEG sequences and may in particular account for index values that falsely indicate consciousness. Compared with BIS, SE more often showed false classifications of the clinical state at transitions between consciousness and unconsciousness. There was a trend toward higher incidence of high frequency signals and eye blinks during sevoflurane compared with propofol anaesthesia. Unfortunately, it remains impossible to exclusively evaluate the impact of artifacts on the EEG signal or to ascertain if these are causal for erroneous index calculation: the exact algorithms of index calculation have not been revealed in detail by the manufacturers. Further, it is unknown how monitor configuration settings of BIS and SE (e.g. filter settings) influence signal processing. In addition, most artifacts are not characterized by a stereotype response. As a result of these findings, a uniform explanation for failure of DOA monitors seems unlikely.

In the present issue of the British Journal of Anaesthesia, Aho and colleagues present a prospective study in 65 patients during general anaesthesia with sevoflurane. The authors analysed the raw EEG signals underlying conflicting BIS and SE values during surgical anaesthesia in order to identify EEG characteristics that cause discrepancies between both index values. The study endpoint - disagreement between BIS and SE values - was defined as a difference ≥10 index points for ≥60 s. The underlying raw EEG signal was visually inspected and analysed in time and frequency domain. The endpoint was observed in 11% of all BIS and SE index pairs. The authors concluded that the following EEG findings can cause index discrepancies: during burst suppression and alpha or beta activity, SE values were higher than BIS. A delta-theta dominance resulted in a higher BIS values. Furthermore, EMG and electro-cautery accounted for differences between index values. In the present analysis, only EEG sequences during surgical anaesthesia were included. These may not be as artifact-contaminated as EEG sequences during lighter levels of anaesthesia, in particular at the transitions between unconsciousness and unconsciousness.

The study of Aho and colleagues offers interesting insights in EEG analysis, but important study limitations have to be discussed. Despite the opportunity of a prospective study design, their present approach misses the target: their study design lacks of a golden standard for the comparison of EEG based indices and the used endpoint to define index differences may be considered as clinically irrelevant.

As a result of index computation, the recommended ranges of processed EEG based index values for the respective levels of hypnosis reflect probability functions rather than absolute, clinical or pharmacological classifications. The authors defined discrepancies between BIS and SE whenever the index values differed ≥10 index points. There is no evidence that index values of different EEG based monitors should agree. They rather may run in a continuous, but not in a linear manner-as a result of manifold and partially unknown reasons (e.g. different calculation algorithms or underlying training data sets). Therefore, one must expect differences between index values of different monitors, because an index value does not quantify a physiological parameter, but a constructed probability function based on a clinical assessment of a poorly defined endpoint.

Furthermore, there is little reason to believe that different index values obtained from different persons during ‘comparable levels of consciousness’ would be clinically meaningful. As an example: is an index value of 100 before induction of general anaesthesia different to an index value of 90?

Index differences ≥10 units do not necessarily imply that both monitors display different levels of anaesthesia and may have no clinical relevance.

Thus, the endpoint may not be appropriate for a comparison of different depth of anaesthesia monitors.

The study design of Aho and colleagues lacks a golden standard for the level of hypnosis (i.e. an assessment of whether an index value indicates the correct level of anaesthesia). Thus, the study misses the opportunity to answer the key questions of mechanistic causes of index discrepancies. These questions may be: what was different about the raw EEG of the patients that were unresponsive (responsive) but had a high (low) BIS/SE values indicating consciousness (unconsciousness)? What was different about the EEG of the patients that had a high BIS but a low SE or vice versa?

As long as anaesthesia indices are based on probabilistic approaches, the golden standard remains the clinical assessment of the level of consciousness. During surgical anaesthesia, the determination of a clinical endpoint is difficult. The classification of EEG stages may represent an alternative solution.

The investigation of causal sources (e.g. specific EEG pattern or artifacts) that may mislead one or both monitors to indicate a level of anaesthesia, that does not agree with the golden standard (clinical or EEG assessment), requires the analysis of all EEG sequences. Further, it should be described how often the respective EEG pattern/artifact is identified in EEG sequences underlying index values, that agree/disagree with the golden standard.

It is important that current anaesthesia indices reflect a probability function of a clinical state rather than a physiological parameter. As these indices of the hypnotic component of anaesthesia may not follow a common definition of this hypnotic component, and may not be scaled to the same dose-response curve, index values are not comparable simply by a comparison of numbers. Therefore, not everything that calculates a dimensionless number from 0 to 100 is comparable with each other. The EEG is more than just a number, and blind trust in a calculated number should be avoided, because it may result in a comparison of apples and oranges.

References

1. American Society of Anesthesiologists Task Force on Intraoperative Awareness. Practice advisory for intraoperative
doi:10.1093/bja/aev231

Anaesthesia for preterm Caesarean delivery: is it different from term deliveries?

C. A. Wong

Department of Anaesthesiology, Northwestern University Feinberg School of Medicine, 251 East Huron Street, F5-704, Chicago, IL, USA
E-mail: c-wong2@northwestern.edu

Worldwide, more than 15 million babies are born annually preterm (fewer than 37 completed weeks’ gestational age), approximately one in 10 births.1 In England and Wales, 7% of births are preterm,2 and in the USA, more than 11%.3 Prematurity is the leading cause of newborn death, and survivors may face lifelong disability.1 Preterm birth may result from the preterm onset of spontaneous labour or from obstetric intervention for maternal or fetal indications, when the risk of preterm birth is deemed to be lower than the risk to the mother or fetus of continuing the pregnancy.

The optimal mode of delivery for preterm fetuses, especially very early preterm, is a matter of much discussion in the obstetric community. Except for specific indications (e.g. breech presentation), evidence is lacking that Caesarean delivery results in better neonatal outcomes than vaginal delivery. However, a larger proportion of preterm infants are born by Caesarean than term infants. In the USA in 2013, 46.6% of preterm fetuses were delivered by Caesarean.4 The mothers of these infants require surgical anaesthesia for delivery, either general or neuraxial, and by extension, the preterm fetus will be exposed to the adverse (or beneficial) effects of these techniques.

There is little evidence that mode of anaesthesia impacts outcomes for term infants.5 It is conceivable, however, that the mode of delivery anaesthesia might influence outcomes for these tiniest of humans. It is generally believed that the preterm fetus is more susceptible to the depressant effects of anaesthetic drugs than the term fetus;6 thus, one might hypothesize that neuraxial anaesthesia is preferable for preterm deliveries. However, delivery is often urgent in these patients, and the induction of general anaesthesia is faster than neuraxial anaesthesia. An understanding of the patterns of use of anaesthesia in this population might be a helpful first step in determining whether the mode of anaesthesia plays a role in outcome.

In this issue of the British Journal of Anaesthesia, Butwick and colleagues7 report a secondary analysis of data from the United States National Institutes of Child Health and Human Development Maternal–Fetal Medicine Units Network Cesarean Registry. Data were collected prospectively between 1999 and 2002 from 19 US academic medical centres. All women with a Caesarean delivery were included in the first 2 yr of the Registry, and only women with previous Caesarean delivery were included in the last 2 yr.

The aim of the study by Butwick and colleagues7 was to identify risk factors for general anaesthesia in a population of women undergoing preterm Caesarean delivery. Potential risk factors included maternal age, predelivery body mass index, race or ethnicity, gestational age, singleton or multiple gestation, primary or repeat Caesarean delivery, fetal presentation at delivery, and the presence of hypertensive disorders of pregnancy, labour, premature preterm rupture of membranes, and emergency indication for delivery. Overall, 82.4% of women received neuraxial anaesthesia and 17.6% received general anaesthesia. This distribution is significantly different from the distribution reported for the entire Registry (93% received neuraxial anaesthesia).7 Thus, it appears that preterm infants were at higher risk for exposure to general anaesthesia than term infants.

In the final multivariate model of the study, race or ethnicity, haemolysis, and elevated liver enzymes and low platelet (HELLP) syndrome or eclampsia (combined) were independently