Comparative evaluation of the bispectral index (BIS) and BISpro during propofol anaesthesia

Zhita Li1,*, Jiachun Cai2,*, Jiexiong Li1, Xianghui Xu3 and Limin Zheng1

Abstract
Objective: To compare the accuracy, correlation and agreement between the bispectral index (BIS) and BISpro during propofol anaesthesia.

Methods: The BIS, BISpro, heart rate, target-concentration of propofol and Observer’s Assessment of Alertness and Sedation (OAA/S) score were recorded every 30 s in female patients scheduled for hysteroscopic surgery. Propofol anaesthesia was induced by an initial target-controlled concentration (1.0 μg/ml) followed by a stepwise increase (0.5 μg/ml) until the patient was unresponsive. Spearman’s correlation coefficient and prediction probability were calculated for the association between sedation levels and the above parameters. The ability of investigated parameters to distinguish between OAA/S scores was analysed. Bland–Altman analysis was used to compare the agreement between BIS and BISpro. The BIS and BISpro cut-off values for lost response were also determined.

Results: Out of 30 patients in total, a high correlation was found between BIS and BISpro, and both correlated well with OAA/S score. Only BIS was able to distinguish all investigated OAA/S states accurately, but the ability to predict OAA/S score 5 to loss of response was comparable between BIS and BISpro. The calculated cut-off values were 68 for BIS and 70 for BISpro.

Conclusion: BISpro and BIS are reliable monitors of general anaesthesia during sedation.

Trial registration number: Chinese Clinical Trial Registry (URL: www.chictr.org.cn): ChiCTR1900024037 (retrospectively registered).

Keywords
Bispectral index, BIS, BISpro, propofol, sedation, anaesthesia, monitor

Date received: 15 October 2020; accepted: 18 February 2021

*These authors contributed equally to this work

Corresponding author:
Limin Zheng, Department of Anaesthesiology, Peking University Shenzhen Hospital, 1120 Lianhua Road, Shenzhen 518036, China.
Email: 1617463316@qq.com
**Introduction**

The measurement of anaesthesia depth remains an ongoing challenge in clinical practice, because it includes three components: sedation, analgesia and muscle relaxation. Several indices for measuring depth of anaesthesia have been commercialized and applied to clinical settings. The bispectral index (BIS) was the first value approved by the United States Food and Drug Administration for monitoring the depth of anaesthesia and is the most widely used clinically.

The BISpro monitor, developed by Maygreen Co. Ltd (Shenzhen, China), measures two indices of anaesthesia depth: BISpro and nociception index (NOX). The BISpro value is similar to the BIS value, which is based on the correlation of the phase between different frequency components of the electroencephalogram (EEG) with bispectral analysis. The BIS ranges from 0 (silent) to 100 (fully awake) and was developed using a database of EEG values and behavioural scales. Instead of using this previous database, the BISpro is based on raw physical principles (www.maygreen.com.cn/). Similar to BIS, the manufacturer of the BISpro monitor suggests that BISpro values ranging from 40–60 may prevent the adverse incidence of awareness during general anaesthesia.

The first aim of the present study was to compare BIS and BISpro values using the Observer’s Assessment of Alertness and Sedation (OAA/S) score, by analysing the accuracy of each variable in distinguishing between different clinically defined states during propofol anaesthesia. Specifically, the investigation focussed on whether the BISpro value is as reliable and accurate as the BIS value in reflecting changes during increasing propofol target concentration and sedation levels. A further aim was to determine an appropriate BISpro cut-off value that provided the optimal combination of sensitivity and specificity for predicting an OAA/S score of ≤2 (loss of response), using receiver operating characteristic (ROC) curve analysis.

**Patients and methods**

**Study population**

This prospective study was conducted at Peking University Shenzhen Hospital, Shenzhen, China between August 2019 and June 2020, and included sequentially enrolled female patients, aged between 18–55 years and with an American Society of Anaesthesiologists physical status classification of I or II, who were scheduled for hysteroscopic surgery. Exclusion criteria were: a history of neurological deficits; congestive heart failure; arrhythmias; bundle-branch block; severe liver and/or kidney dysfunction; pregnancy; antidepressant treatment within 1 year of enrolment; or dementia. The study was approved by Peking University Shenzhen Hospital Medical Ethics Committee (08/06/2019) and retrospectively registered at ChiCTR (ID: ChiCTR1900024037). Written informed consent was obtained from all participants. All procedures performed in this human study were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

**Study procedures and data collection**

No sedative or analgesic was administered preoperatively. On arrival in the operating room, an intravenous cannula was inserted in the dorsum of a hand, and Ringer’s lactate infusion was started. Electrocardiograms, SpO2, non-invasive blood pressure, BIS values (BIS VISTA™ Monitoring System, version 3.20; Aspect Medical Systems, Newton, MA, USA), and BISpro values
The EEG was registered for BIS using a BIS® sensor XP electrode (Aspect Medical Systems) and for BISpro using the relevant Maygreen electrode over the left or right area of the forehead, according to the manufacturers’ instructions. The side for temporal placement of BIS and BISpro electrodes was randomly selected by flipping a coin. The BIS was monitored using a two-channel referential montage, and calculated with a smoothing rate of 30 s. Both electrode impedances were kept below 10 kΩ. Supplemental oxygen was administered via a venturi mask at a flow rate of 3 l/min throughout the surgery.

After preoperative preparation, propofol anaesthesia (serial No. 16MM2728, Fresenius Kabi Austria GmbH, Austria) was induced by an initial target-controlled effect-site concentration (Ces) of 1.0 μg/ml, using the population pharmacokinetic sets by Schnider et al.6 and the Orchestra Base Primea (Fresenius Vial, Brézins, France), followed by a stepwise increase of 0.5 μg/ml until the patient was unresponsive. Based on the time to peak effect of propofol (1.6 min), every propofol concentration level was maintained for >2 min.7 The OAA/S scores were assessed on a scale from 1 (no response to mild prodding or shaking) to 5 (readily responds to name spoken in normal tone) (Table 1). To ensure that each patient received the same intensity of tone and stimulus, OAA/S scores were assessed by a specified investigator (ZL) for all patients every 30 s from the start of infusion. BIS, BISpro, heart rate (HR) and Ces of propofol was also recorded every 30 s. The loss of response was defined as an OAA/S value ≤ 2.

**Sample size calculation**

Based on previously published research, a difference of < 0.05 for prediction probability (PK) would not be of clinical significance, and with a t quotient calculated between the chosen difference in clinical importance and the standard error, the present study was calculated to require at least 30 patients.8

**Statistical analyses**

All data are presented as mean ± SD, median (range) or n patients. Changes in BIS, BISpro, HR, and Ces at different OAA/S scores were compared using the non-parametric Kruskal–Wallis H-test.

### Table 1. Observer’s Assessment of Alertness/Sedation (OAA/S) scale.

| Score | Responsiveness                                | Speech                          | Facial expression        | Eyes                           |
|-------|-----------------------------------------------|---------------------------------|--------------------------|--------------------------------|
| 1     | Does not respond to mild prodding or shaking  | Few recognized words           | Slurring or prominent slowing | Glazed and marked ptosis       |
| 2     | Responds only after mild prodding or shaking |                    | Marked relaxation (slack jaw) |                                |
| 3     | Responds only after the name is spoken loudly and/or repeatedly | |                          |                                |
| 4     | Lethargic response to name spoken in a normal tone | Mild slowing or thickening     | Mild relaxation           | Glazed or mild ptosis          |
| 5     | Responds readily to name spoken in normal tone | Normal                          | Normal                    | Clear, no ptosis               |

The final score is the sum of the Responsiveness, Speech, Facial expression, and Eyes component scores. Thus, a ‘wide awake’ score = 5 and a deeply sedated score = 1.
Two types of significant error for sedative assessment were defined and the number of times they occurred was calculated. A type I error was defined as BIS < 60 and BISpro > 80 or BIS < 40 and BISpro > 60, and a type II error was defined as BIS > 80 and BISpro < 60 or BIS > 60 and BISpro < 40.

Associations between OAA/S scores and BIS, BISpro, HR, and Ces were analysed using Spearman’s correlation coefficient. Moreover, the correlation between BIS and BISpro for all data pairs was evaluated by the nonparametric Spearman’s rank correlation coefficient. Bland–Altman analysis was used to compare the consistency between BIS and BISpro. Associations between OAA/S scores and BIS, BISpro, and HR were also assessed by determining PK. The accuracy of BIS, BISpro, and HR in distinguishing between different depths of sedation, such as an OAA/S score of 5 versus 4, 3, and loss of response, was analysed. PK was calculated and compared using the PKMACRO and PKDMACRO spreadsheets, as described by Smith et al. The jackknife method was used to compute the standard error of the estimate. A PK value < 0.5 indicates that discordance is more likely than concordance. To enable a comparison of PK, 1 –PK was used when the PK value was < 0.5. To evaluate the accuracy in distinguishing all investigated conditions, the following formula was used: 

\[ PK_{all} = PK_{OAA/S 5 versus 4} \times PK_{OAA/S 5 versus 3} \times PK_{OAA/S 5 versus loss of response} \]

All steps were weighted equally.

To compare the abilities of BIS and BISpro to identify different conscious states, and to obtain the BIS and BISpro cut-off values that estimate a hypnotic status, ROC curve analysis was used to determine the optimal combination of sensitivity and specificity for predicting an OAA/S score of ≤2. The closer the ROC curve is to the upper left corner, the higher the accuracy of the test.

Statistical analyses were performed using SPSS software, version 22.0 (IBM, Armonk, NY, USA), MedCalc statistical software, version 19.0.5 (MedCalc Software Ltd, Ostend, Belgium), GraphPad Prism, version 5.01 (GraphPad Software, San Diego, CA, USA), PKMACRO, and PKDMACRO. A P-value < 0.05 was considered statistically significant.

### Results

Thirty-one patients were enrolled, and one patient was excluded for a poor signal quality index bar at 10 min after induction. Thus, a total of 30 patients were included in the final analysis (mean age, 36.2 ± 8.4 years; height, 158.7 ± 6.8 cm; and weight, 53.8 ± 8.8 kg). The Ces for loss of response was 3.1 ± 0.3 μg/ml.

The two EEG parameters significantly decreased as OAA/S scores decreased, in that BIS and BISpro values were statistically lower in patients with OAA/S scores of 1 and 2 versus scores of 3–5 (P < 0.05; Table 2). One pair of values exhibited a type I error for sedative assessment and there were no type II errors (Table 3). BIS and BISpro also correlated well with OAA/S scores (r = 0.898, PK = 0.943 for BIS; and r = 0.806, PK = 0.879 for BISpro), but the PK of BIS was significantly better than the PK of BISpro (P < 0.01; Table 4). The correlation coefficient (r) for Ces (r = 0.761) was similar to BISpro, and no statistically significant differences in PK values were found between them (Table 4). Correlation coefficients and PK values for sedation levels were higher for BIS and BISpro than for HR (Table 4). A high correlation (r = 0.839, P < 0.001) was found between BIS and BISpro (Figure 1).

The results of Bland–Altman analysis are shown in Figure 2. The mean difference between BIS and BISpro was 1.7 with a standard error of the estimate of 7.91. The upper limit of agreement was 17.2 and the lower limit of agreement was
–13.8. Only 0.04% of points were outside of the 95% limit of agreement, but BIS and BISpro differed by more than 10 units, suggesting a poor agreement between them.

The probability of BIS, BISpro, HR, and Ces to predict change in OAA/S scores was assessed by PK (Table 5). Only BIS was able to distinguish all investigated states accurately (PK > 0.90 for each state change), and analysis revealed a PK_all of 0.933 for BIS and a PK_all of 0.840 for BISpro (Table 5). A lower PK for the
state between an OAA/S score of 5 and 4 was shown for BISpro (0.868) compared with BIS (0.937), however, the ability to predict the state change from OAA/S of 5 to loss of response was equivalent between BIS and BISpro.

The ROC curve analysis identified BIS and BISpro values with the best level of discrimination to obtain an OAA/S score of ≤2 during propofol sedation (Figure 3). The calculated BIS cut-off value was 68 (sensitivity of 92.4%, specificity of 86.3%, and area under the curve [AUC] of 0.940, \( P < 0.01 \)); and the BISpro cut-off value was 70 (sensitivity of 86.4%, specificity of 73.1%, and AUC of 0.879, \( P < 0.001 \)). The BIS value was closer to the upper left corner.

**Discussion**

The BIS is derived from recordings of large-volume EEGs and associated levels of consciousness and sedation of patients receiving different sedative infusions, combined into a database. The mixed information of the power and frequency of the EEG is then fitted with the bispectral analysis to obtain an optimal value, ranging from 0 to 100. When the number is reduced, cortical inhibition is deepened. In other words, BIS compares the EEG signal with a previous database to produce a dimensionless number that reflects the level of hypnosis.

BISpro is similar to BIS in that it also originates from EEG data, and is based on four basic waves: \( \delta \), \( \theta \), \( \alpha \), and \( \beta \). However, BISpro uses a different algorithm from BIS to quantify the depth of anaesthesia. Symbolic dynamics are used to divide the acquired EEG signals into linear and non-linear modules. Each part is marked by a symbol and converted into time series. EEG energy is calculated by a special frequency band, and the ratio of \( \alpha \) and \( \beta \) is determined by calculating the difference in energy between high and low frequency during anaesthesia. At the same time, the total number of suppression episodes during the 30 s of EEG is calculated in real-time, and the total amount of anaesthesia depth in the stationary phase and the quiescent period is quantified. The logical classification of the four-parameter values of brain waves is
Table 5. Prediction probability (PK) for the different stages of propofol anaesthesia in 30 patients.

| Parameter | OAA/S 5 versus 4 | OAA/S 5 versus 3 | OAA/S 5 versus loss of response |
|-----------|------------------|-----------------|-------------------------------|
| BIS       | PK 0.937, SE 0.017 | PK 0.996, SE 0.002 | PK 1.000, SE 0.000, PKall 0.933 |
| BISpro    | PK 0.858, SE 0.026\(^a\) | PK 0.980, SE 0.007 | PK 0.999, SE 0.002, PKall 0.840 |
| HR        | PK 0.614, SE 0.038\(^a,b\) | PK 0.701, SE 0.033\(^a,b\) | PK 0.719, SE 0.045\(^a,b\), PKall 0.617 |
| Ces       | PK 0.835, SE 0.026\(^a\) | PK 0.936, SE 0.013\(^b\) | PK 0.993, SE 0.003\(^c\), PKall 0.776 |

BIS, bispectral index; HR, heart rate; Ces, effect-site concentration of propofol; OAA/S, Observer’s Assessment of Alertness/Sedation; SE, standard error.

\(^a\)Versus BIS (P < 0.01); \(^b\)versus BISpro (P < 0.01); \(^c\)versus BIS (P < 0.05).

Good agreement is indicated by values within a range of 10 units above and below the mean difference.\(^11,12\) An unsatisfactory disagreement, shown by more than 10 units between BIS and BISpro, was found in the present study, with upper and lower agreement limits of 17.2 and −13.8, respectively.

A highly effective index of sedation should accurately discriminate between the investigated states of anaesthesia. A PK value of 0.5 indicates that the variable predicts the states no better than a chance of 50%, whereas a PK value of 1.0 means that the variable correctly predicts the states 100% of the time, with no overlap. Furthermore, a parameter that perfectly distinguishes all conditions with a PK of 1.0 results in a PK\(_{\text{all}}\) of 1.0 \times 1.0 \times 1.0 = 1.0; while a variable that predicts all conditions with a PK of 0.5 results in a PK\(_{\text{all}}\) of 0.5 \times 0.5 \times 0.5 = 0.125. Only BIS was able to accurately distinguish all investigated states with a PK >0.90. Based on the present data, BISpro was found to be limited by a worse PK for OAA/S score of 5 versus 4. This may be caused by an overlap of clinical signs between OAA/S scores of 5 and 4 that may have affected the researcher’s ability to differentiate between a score of 4 or 5. Another explanation for these discrepancies may be the differences between algorithms used in BIS and BISpro monitors. The
algorithms differ between BIS and BISpro values, and different artifact rejection methods might play a role, in addition to intrapatient variability. Findings similar to the present results were reported in a previous study, in which entropy index and BIS displayed no linear correlation outside a BIS range of 40–60. BIS is mainly calculated by the Beta Ratio during awake or light sedation, SyncFastSlow (bispectral component) predominates during the surgical level of hypnosis (BIS 40–60), and the Burst Suppression Ratio/Quazi algorithm detects a deep level of anaesthesia (BIS < 40). However, due to the lack of a specific algorithm for BISpro, we cannot comment on this phenomenon with any degree of certainty. Although the BIS and BISpro were not identical, the PK during transition from awake to lost response was equivalent (1.000 for BIS, and 0.999 for BISpro), suggesting that both were able to detect awareness during the operation.

The closer the ROC curve is to the upper left corner, the higher the accuracy of the test. The point closest to the ROC curve in the upper left corner is the best threshold with the least error and the fewest false positives and false negatives. BIS was more effective than BISpro in estimating the state of loss of response in the present findings, as its curve was closer to the upper left corner. BIS values at an OAA/S score of ≤ 2 were similar to those reported in previous research, and approximately 2 points less than BISpro.

The present study results may be limited by the fact that the investigated states were divided into five scales only by OAA/S assessment, which may have resulted in some overlap between adjacent scales, particularly between OAA/S score 5 and 4. This might have weakened the power of the monitor to distinguish between the different hypnotic levels. Previous studies have defined six anaesthetic states, comprising awake versus loss of response, awake versus loss of eyelash reflex, awake versus steady-state anaesthesia, steady-state anaesthesia versus the first reaction, and steady-state anaesthesia versus extubation, which might contribute to a more definite conclusion and more closely represent clinical settings. Another limitation was that the effect of electromyography was not considered, which might have played an important role in impairing the accuracy of BIS and BISpro.

In summary, BISpro showed a good correlation with BIS and OAA/S during propofol anaesthesia, but a poor agreement with BIS. BISpro provided the lowest accuracy in distinguishing the states between entirely awake and light sedation, but was close in power to BIS in distinguishing between awake and loss of response. Therefore, BISpro was shown to be a reliable monitor for general anaesthesia, but not an accurate index for light sedation.

Acknowledgements
We gratefully acknowledge the kind cooperation of Dr Tao Luo, Department of Anaesthesiology, Peking University Shenzhen Hospital, in the preparation and modification of this manuscript.

Author contributions
Zhitao Li helped design and conduct the study, analyse the data, and write the manuscript; Jiachun Cai helped design the study, and write the manuscript; Jiexiong Li helped design and conduct the study; Xianghui Xu helped design the study, and write the manuscript; and Limin Zheng helped design the study, and review and edit the manuscript.

Declaration of conflicting interest
The authors declare that there is no conflict of interest.

Funding
This research was supported by the Clinical Research Project of Peking University Shenzhen Hospital, Shenzhen, China. No. LCYJ202004.
References

1. Kreuer S, Biedler A, Larsen R, et al. Narcotrend monitoring allows faster emergence and a reduction of drug consumption in propofol–remifentanil anesthesia. *Anesthesiology* 2003; 99: 34–41.

2. Freye E. Cerebral monitoring in the operating room and the intensive care unit – an introductory for the clinician and a guide for the novice wanting to open a window to the brain. *J Clin Monit Comput* 2005; 19: 1–76.

3. Jameson LC and Sloan TB. Using EEG to monitor anesthesia drug effects during surgery. *J Clin Monit Comput* 2006; 20: 445–472.

4. Rampil IJ. A primer for EEG signal processing in anesthesia. *Anesthesiology* 1998; 89: 980–1002.

5. Kasuya Y, Govinda R, Rauch S, et al. The correlation between bispectral index and observational sedation scale in volunteers sedated with dexmedetomidine and propofol. *Anesth Analg* 2009; 109: 1811–1815.

6. Schnider TW, Minto CF, Shafer SL, et al. The influence of age on propofol pharmacodynamics. *Anesthesiology* 1999; 90: 1502–1516.

7. Liu N, Chazot T, Hamada S, et al. Closed-loop coadministration of propofol and remifentanil guided by bispectral index: a randomized multicenter study. *Anesth Analg* 2011; 112: 546–557.

8. Schmidt GN, Bischoff P, Standl T, et al. Comparative evaluation of the Datex-Ohmeda S/5 Entropy Module and the Bispectral Index monitor during propofol–remifentanil anesthesia. *Anesthesiology* 2004; 101: 1283–1290.

9. Myles PS and Cui J. Using the Bland–Altman method to measure agreement with repeated measures. *Br J Anaesth* 2007; 3: 309–311.

10. Smith WD, Dutton RC and Smith NT. Measuring the performance of anesthetic depth indicators. *Anesthesiology* 1996; 84: 38–51.

11. Musialowicz T, Lahtinen P, Pitknen O, et al. Comparison of spectral entropy and BIS VISTA™ monitor during general anesthesia for cardiac surgery. *J Clin Monit Comput* 2011; 25: 95–103.

12. Bonhomme V, Deflandre E and Hans P. Correlation and agreement between bispectral index and state entropy of the electroencephalogram during propofol anaesthesia. *Br J Anaesth* 2006; 97: 340–346.

13. Niedhart DJ, Kaiser HA, Jacobsohn E, et al. Intrapatient reproducibility of the BISxp monitor. *Anesthesiology* 2006; 104: 242–248.

14. Vakkuri A, Yli-Hankala A, Talja P, et al. Time-frequency balanced spectral entropy as a measure of anesthetic drug effect in central nervous system during sevoflurane, propofol, and thiopental anesthesia. *Acta Anaesthesiol Scand* 2004; 48: 145–153.

15. Kodaka M, Johansen JW and Sebel PS. The influence of gender on loss of consciousness with sevoflurane or propofol. *Anesth Analg* 2005; 101: 377–381.

16. LeBlanc JM, Dasta JF, Pruchnicki MC, et al. Bispectral index values, sedation-agitation scores, and plasma lorazepam concentrations in critically ill surgical patients. *Am J Crit Care* 2012; 21: 99–105.

17. Xu B, Li Z, Zhou D, et al. The influence of age on sensitivity to dexmedetomidine sedation during spinal anesthesia in lower limb orthopedic surgery. *Anesth Analg* 2017; 125: 1907–1910.