British Journal of Anaesthesia **106** (1): 101–11 (2011)
Advance Access publication 4 November 2010 · doi:10.1093/bja/aeq291

BJA

PAIN

Comparison of the Surgical Pleth Index™ with haemodynamic variables to assess nociception–anti-nociception balance during general anaesthesia

V. Bonhomme[1,2]*, K. Uutela[4], G. Hans[1], I. Maquoi[1], J. D. Born[3], J. F. Brichant[1], M. Lamy[1] and P. Hans[1,2]

[1] Department of Anaesthesia and Intensive Care Medicine, CHU Liege, Liege, Belgium
[2] University Department of Anaesthesia and Intensive Care Medicine, CHR Citadelle, Liege, Belgium
[3] Department of Neurosurgery, CHR Citadelle, Liege, Belgium
[4] GE Healthcare Finland, Oy, Helsinki, Finland

* Corresponding author. E-mail: vincent.bonhomme@chu.ulg.ac.be

Key points

- The Surgical Pleth Index offers a measure of the balance between noxious stimulation and anti-nociception during surgery.
- The SPI value could be affected by intravascular filling status of the patient and by conditions such as chronic hypertension.
- The current study compared heart rate (HR), mean arterial pressure (MAP), and the SPI for the measurement of the balance between nociception produced by a neurosurgical head holder and anti-nociception from a remifentanil infusion.
- The performance of the SPI response to head holder at indicating the anti-nociception level was comparable with that of MAP and HR.
- Low intravascular volume status and chronic treatment for high arterial pressure lowered the responses of those indexes to the stimulation.

Background. The Surgical Pleth Index (SPI) is proposed as a means to assess the balance between noxious stimulation and the anti-nociceptive effects of anaesthesia. In this study, we compared SPI, mean arterial pressure (MAP), and heart rate (HR) as a means of assessing this balance.

Methods. We studied a standard stimulus [head-holder insertion (HHI)] and varying remifentanil concentrations (CeREMI) in a group of patients undergoing neurosurgery. Patients receiving target-controlled infusions were randomly assigned to one of the three CeREMI (2, 4, or 6 ng ml^{-1}), whereas propofol target was fixed at 3 $\mu g\ ml^{-1}$. Steady state for both targets was achieved before HHI. Intravascular volume status (IVS) was evaluated using respiratory variations in arterial pressure. Prediction probability (Pk) and ordinal regression were used to assess SPI, MAP, and HR performance at indicating CeREMI, and the influence of IVS and chronic treatment for high arterial pressure, as possible confounding factors.

Results. The maximum SPI, MAP, or HR observed after HHI correctly indicated CeREMI in one of the two patients [accurate prediction rate (APR)=0.5]. When IVS and chronic treatment for high arterial pressure were taken into account, the APR was 0.6 for each individual variable and 0.8 when all of them predicted the same CeREMI. That increase in APR paralleled an increase in Pk from 0.63 to 0.89.

Conclusions. SPI, HR, and MAP are of comparable value at gauging noxious stimulation–CeREMI balance. Their interpretation is improved by taking account of IVS, treatment for chronic high arterial pressure, and concordance between their predictions.

Keywords: anaesthetic techniques, i.v.; monitoring, depth of anaesthesia; pain, acute

Accepted for publication: 13 August 2010

During the past few years, several tools have been developed to monitor the balance between the intensity of noxious stimulation and anti-nociception during surgery under general anaesthesia.[1] These tools aim at providing each patient with an appropriate, individually tailored anti-nociception regimen. Complex neural and neuroendocrine pathways are activated by noxious stimulation. This makes the identification of an ideal monitorable pharmacodynamic target difficult. In contrast to the monitoring of the hypnotic component of anaesthesia by the measurement of electroencephalogram-derived variables, assessment of the nociception–anti-nociception balance is indirect in nature, apart from evoked potentials.[2 3] Variables studied include the spinal withdrawal and H reflex,[4] the RIII reflex threshold,[5] the response spectral entropy of the electroencephalogram,[6 7] the skin vasomotor reflex,[8] the skin conductance,[9 10] and the pupil diameter.[11]

The recently renamed Surgical Pleth Index™ (SPI), previously known as the Surgical Stress Index™, has been demonstrated to be a function of the intensity of surgical

stimulation and the depth of the anti-noxious component of anaesthesia provided by an opiate infusion,[12] or regional anaesthesia.[13] In addition to its strong correlation to opiate concentration,[14] the increase in SPI in response to noxious stimulation has been demonstrated to be a predictor of the occurrence of patient movement.[15]

The calculation of SPI relies on a balanced sum of normalized heart beat intervals (HBIs) and plethysmographic pulse wave amplitudes,[12] both of which are controlled by the balance between sympathetic and parasympathetic tone. SPI has been shown to change with noxious stimulation even in patients receiving β-blocking agents, when compared with those receiving an appropriate dose of fentanyl.[16] However, other factors known to influence autonomic reactions independently of a noxious stimulus might interfere with the accuracy of this variable in evaluating the nociception–anti-nociception balance. Factors that may be relevant include intravascular volume status (IVS), diabetes, or chronic high arterial pressure and anti-hypertensive drugs.

This study was designed to compare the performance of SPI and haemodynamic variables, at gauging the nociception–anti-nociception balance of patients submitted to an intense and standardized noxious stimulation under general anaesthesia, and at identifying confounding factors that potentially impede their interpretation.[17] The standardized stimulus was the insertion of a pin head holder just before intracranial neurosurgery, in patients submitted to variable anti-nociception levels. Variation in anti-nociception was achieved by comparing patients receiving one of three different remifentanil concentrations as estimated by a pharmacokinetic model. IVS and treated chronic high arterial pressure were tested as potential confounding factors.

Methods

Patient recruitment, sample size, and randomization

After approval by our Institutional Review Board and informed consent, 33 patients undergoing intracranial neurosurgery were recruited for this prospective double-blind randomized study. Exclusion criteria included age below 18 and above 80, impaired cardiac function, and a past medical history of diabetes.

Sample size calculation was performed using G*Power© software (version 3.0.3, Franz Faul, Universitat Kiel, Germany)[18] and based on the intention to perform multiple regression analyses. Considering a squared multiple correlation coefficient (R^2) of 0.3 as being relevant, a set of three predictors (anti-nociception level indication, IVS, and history of chronic high arterial pressure), and an α-value of 0.05, a total sample size of 30 was required to achieve a power of 0.8.

Patients were randomly assigned to one of the three groups, according to the effect-site concentration of remifentanil (CeREMI) to be achieved during the study period, namely 2, 4, or 6 ng ml^{-1}. The randomization was obtained using an Excel™ (Microsoft™ Office Excel 2003, Microsoft Corporation, Luxembourg, Zaventem, Belgium) random number function-generated list that was available to the nurse in charge of preparing anaesthetic medications, but blinded to the anaesthesiologist in charge of the procedure. Blinded syringes were prepared with 25, 50, or 75 μg ml^{-1} remifentanil in normal saline. The patients were assigned a number in the randomization list according to the sequential order of their recruitment.

Anaesthesia protocol

All patients were planned for an early morning surgery and fasted for 6 h before induction of anaesthesia. Patients with a past medical history of chronic high arterial pressure received their usual anti-hypertensive medications on the morning of surgery at the time of premedication, except that angiotensin-convertase inhibitors and angiotensin II antagonists were withheld on that morning, but still given the day before. Premedication consisted of alprazolam 0.5 mg and atropine 0.5 mg given orally 1 h before surgery. Upon arrival in the operating theatre, two 18 G i.v. cannulae were sited and standard monitoring was applied including an Sp_{o_2} sensor. Careful attention was paid to the position of the Sp_{o_2} sensor on a thumb, and this was not changed throughout the study period. A crystalloid solution (Plasmalyte A®, Baxter International Inc., IL, USA) was infused through one venous cannula and a colloid solution (Voluven®, Fresenius Kabi, Bad Homburg, Germany) through the second one. The total infusion rate for both infusions combined was kept constant at 2 ml kg^{-1} h^{-1} throughout the study.

A target-controlled infusion (TCI) of remifentanil was started using an ASENA PK TCI® pump (Cardinal Health Alaris Products, Basingstoke, UK), and the pharmacokinetic model of Minto and colleagues.[19] The effect-site concentration target was set on the pump at 4 μg ml^{-1} in all patients, hence leading to a CeREMI of 2, 4, or 6 ng ml^{-1} once steady state had been obtained according to the remifentanil concentration in the blinded syringe (25, 50, or 75 μg ml^{-1}). Once CeREMI was achieved, a propofol TCI was started (Master TCI®, Fresenius Kabi) to obtain an effect-site concentration of 3 μg ml^{-1} (model of Marsh and colleagues).[20] The effect-site concentrations of propofol and remifentanil were kept constant throughout the study. Immediately after loss of consciousness, patients received a 0.2 mg kg^{-1} dose of cisatracurium to facilitate tracheal intubation. No further dose of neuromuscular blocking agent was administered thereafter. Once the airway was secured and mechanically ventilated with a 50% oxygen–air mixture (end-tidal carbon dioxide partial pressure maintained between 4.0 and 4.7 kPa), an arterial line was inserted into the radial artery at the wrist (20 G catheter), on the same side as the one for the Sp_{o_2} sensor, and contralateral to the non-invasive arterial pressure cuff. Normothermia was maintained throughout the study using a forced-air warming device.

Standard noxious stimulation

The standard noxious stimulation consisted of a Mayfield head-holder insertion (HHI) (Mayfield® Modified Skull Clamp, Integra™, Plainsboro, NJ, USA). This device allows

solid fixation and adequate placement of the head for most of intracranial surgical procedures. Its three pins apply a 60 N force on the skull through the skin. The head holder was inserted only when propofol and remifentanil target concentrations had both been achieved and kept constant for more than 15 min.

Evaluation of IVS

To assess the IVS of patients, the delta down (DD) was measured under steady-state anaesthetic conditions, immediately before HHI. DD is an indicator of fluid responsiveness, and hence of IVS in patients with normal cardiac function. It is defined as the difference between the systolic arterial pressure at the end of a respiratory pause, immediately before lung inflation, and its minimum value during the course of one respiratory cycle. It has been demonstrated to be as efficient as other surrogate measures at detecting fluid responsiveness during intracranial neurosurgery.[21] Values of DD higher than 5 mm Hg indicate that a patient will respond to fluid loading by an increase in arterial pressure. During DD acquisition, the tidal volume was set at 8 ml kg^{-1} and the frequency rate was reduced to 8 min^{-1} to mimic the conditions of a respiratory pause.

Data acquisition

The sequence of data acquisition is illustrated in Figure 1. Except for patient characteristics and DD measures which were manually recorded, all data were continuously acquired using the S/5® Collect software (version 4.0, GE Healthcare, Helsinki, Finland) and a laptop computer connected and synchronized to the Datex-Ohmeda AS3 monitor (GE Healthcare). The Sp_{O_2} waveform was continuously sampled at a rate of 300 Hz. The Sp_{O_2} waveform data were used for *post hoc* calculation of SPI. In addition, numerical values of heart rate (HR) and mean invasive arterial pressure (MAP) were continuously sampled at 10 s intervals.

Calculation of SPI

SPI calculation was performed off-line by a blinded investigator (K.U.). Photoplethysmographic pulse wave amplitude (PPGA) and HBI were extracted from plethysmographic signal of the Sp_{O_2} sensor and normalized to the range 0–100 using a histogram transformation. The SPI value was calculated by combining the normalized values as $SPI = 100 - (0.3 \times HBI_{norm} + 0.7 \times PPGA_{norm})$ and low-pass filtered with a 15 s moving median filter. SPI calculation has been described in detail by Huiku and colleagues.[12] This calculation produced one SPI value every second for each patient.

Data analysis

Normalization of distributions was checked when required. A two-tailed *P*-value of ≤ 0.05 was considered significant.

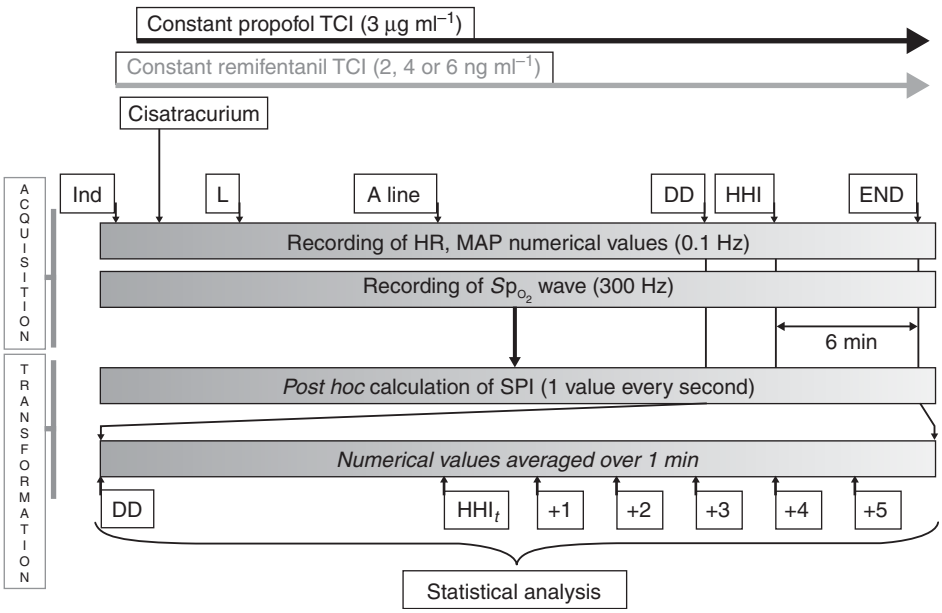

Fig 1 Sequence of data acquisition and transformation before statistical analysis. Data were continuously recorded at a 0.1 Hz rate for numerical values (HR and MAP) and at 300 Hz for the plethysmographic pulse wave (Sp_{O_2} wave). Ind, beginning of anaesthesia induction; L, laryngoscopy; A line, arterial line insertion; DD, delta down measurement; HHI_t, time of pin HHI; END, end of recording; TCI, target-controlled infusion; cisatracurium, moment of cisatracurium administration. The period of interest was starting at DD and finishing 6 min after HHI. After off-line SPI calculation, data of each individual patient were averaged over the minute immediately after points of interest: DD, HHI_t, and +1, +2, +3, +4, and +5 min after HHI_t.

For each patient, the SPI values (one per second), HR, and MAP values (one every 10 s) were averaged over the minutes immediately after each time point of interest (Fig. 1), precisely noted in the recording files: the time of DD measurement, the exact time of the pin HHI (HHI_t), and every minute after HHI_t until 5 min (+1, +2, +3, +4, and +5).

Within- and between-group comparisons of 1 min-averaged SPI, HR, and MAP were performed using two-way mixed-design ANOVA, and Tukey's HSD for *post hoc* comparisons (Datasim© software, Version 1.1, Drake R. Bradley, Department of Psychology, Bates College, Lewiston, ME, USA). Those data were reported as mean (SD).

The performance of 1 min-averaged SPI, HR, and MAP values at indicating the remifentanil concentration was assessed using the prediction probability (Pk) described by Smith and colleagues.[22 23] A Pk of 1 or 0 means a perfect prediction, whereas a Pk of 0.5 means that the predictive value of the parameter is not better than chance. The jack-knife method was used to compute the standard error of the estimate and 95% CI, so that the obtained estimates could be tested as whether they were significantly different from 0.5. Those Pk's were calculated for each 1 min-averaged variable recorded at the time of DD measurement, that is, under steady-state anaesthetic conditions in the absence of noxious stimulation, and for their maximal value recorded during the other time points of interest, that is, at HHI_t, +1, +2, +3, +4, or +5. Each above-defined 1 min-averaged variable of each patient was paired with the actual CeREMI received, leading to a maximum of 33 pairs of data for each Pk calculation.

The stepwise ordinal regression procedure of the SPSS© software (version 17.0.0, SPSS Inc., Chicago, IL, USA) was used to assess the relationship between the SPI, HR, or MAP value and the probability of having received 2, 4, or 6 ng ml^{-1} CeREMI. The maximum values of 1 min-averaged SPI, HR, and MAP recorded at the points of interest after HHI (namely HHI_t, +1, +2, +3, +4, or +5) were the independent variables (predictors), and the probability of having received 2, 4, or 6 ng ml^{-1} CeREMI was the dependent variable in each case (models for SPI, HR, or MAP). Hence, for a given value of an independent predictor, the corresponding model allowed calculation of the probability of having received each of the three different CeREMI. Out of those three probabilities, the highest determined the most probable CeREMI, or, in other words, the CeREMI predicted by the model. Details of the ordinal regression analysis are given in the Appendix.

The goodness-of-fit for each model was assessed using several tests (see Appendix for details). For each model with acceptable goodness-of-fit, the accurate prediction rate (APR) was calculated as the ratio between the number of subjects, whose predicted CeREMI (the concentration whose probability was highest according to the model) was the same as the true value, and the total number of subjects in the sample.

We then introduced supplementary independent variables in the models and checked for improved prediction accuracy. The first additional variable to be introduced was the DD value, followed by hypertension therapy coded as a binary variable. Each new model was evaluated using the same above-mentioned statistical tests and its APR.

To further check for an improvement in the performance of each model at indicating CeREMI adequately, Pk's were calculated using pairs of actual CeREMI and their corresponding concentration predicted by the models. Again, a maximum of 33 pairs of data were used for each Pk calculation.

Results

Patient characteristic data and type of surgery

From the initial 33 recruited patients, three had to be excluded from the study because of protocol violations (administration of vasoactive drugs to sustain arterial pressure before HHI). As indicated in Table 1, all three groups were comparable in terms of age, gender, ASA physical status, weight, height, and BMI. Fourteen patients had a past medical history of high arterial pressure and were treated using β-blocking agents, diuretics, calcium antagonists, α-blocking agents, angiotensin-convertase inhibitors, or angiotensin II inhibitors, alone or in association.

Surgery was planned for supratentorial tumours (n=18), supratentorial aneurysm (n=1), posterior fossa tumours (n=3), or posterior fossa functional pathologies (n=8).

Responses to HHI

The responses of SPI, HR, and MAP to HHI are illustrated in Figure 2. The significant differences within and between groups of patients according to the two-way mixed-design ANOVA and Tukey's HSD tests are detailed in the figure legend. They can be summarized as follows: HHI was associated with a significant increase in SPI, more marked for lower CeREMI. The same was observed for MAP. The increase in HR

Table 1 Characteristics for each group of patients (2, 4, or 6 ng ml^{-1} CeREMI) and the overall sample. *n*, number of patients; ASA, ASA physical status; PMH HAP, past medical history of high arterial pressure

	2 ng ml^{-1}	4 ng ml^{-1}	6 ng ml^{-1}	Overall
N	11	10	9	30
Age (yr) [mean (range)]	59 (35–76)	57 (20–80)	59 (44–74)	59 (20–80)
Gender (*n*) [male/female]	5/6	4/6	3/6	12/18
ASA (*n*) [I/II]	2/9	5/5	2/7	9/21
Weight (kg) [mean (SD)]	79 (14)	78 (19)	70 (9)	76 (15)
Height (m) [mean (SD)]	1.70 (0.05)	1.69 (0.12)	1.69 (0.10)	1.69 (0.09)
BMI (kg m^{-2}) [mean (SD)]	27 (5)	28 (4)	25 (3)	27 (5)
PMH HAP (*n*)	7	3	4	14

was less pronounced, as were differences between the groups for that variable.

Pk's of absolute indicator values

Under steady-state anaesthetic conditions and in the absence of noxious stimulation (time of DD measurement), only HR had a Pk for indicating the true CeREMI that was significantly different from chance alone (Table 2). In contrast, after HHI, the maximum observed value of all variables performed comparably in this regard and all had prediction probabilities which were significantly different from 0.5. The Pk's of maximum SPI, HR, or MAP ranged between 0.6 and 0.7. The power of this study at detecting a 0.1 Pk difference from 0.5, with a standard error of 0.07, an α threshold of 0.05, and a sample size of 30 was very close to unity.

Ordinal regressions

Since only maximum values of 1 min-averaged SPI, HR, and MAP observed after HHI had significant predictive values (Pk), ordinal regression models were constructed using those independent variables only. Models individually obtained for SPI, HR, and MAP (model #1) fitted the data at an acceptable level (data not shown) and exhibited reasonably strong relationships with the dependent variable, namely the probability of having received 2, 4, or 6 ng ml^{-1} CeREMI (Nagelkerke's pseudo-R^2 between 0.200 and 0.363, Table 3). Each model had an APR ranging between 0.50 and 0.57, which is better than chance in a three-option design (2, 4, or 6 ng ml^{-1} CeREMI) (Table 4). The Pk's obtained when comparing actual CeREMI and concentration predicted by the model were significantly different from 0.5, except for the SPI model, and were 0.63, 0.74, and 0.78 for the SPI, HR, and MAP model, respectively. Introducing additional independent variables such as DD (model #2) and the fact of having a past history of treated high arterial pressure (model #3) further improved each individual model, leading to APRs of 0.6 and Pk's of 0.76, 0.79, and 0.78, respectively. The APRs and Pk's were further improved when two out of the SPI, HR, and MAP models combining DD and the past medical history of high arterial pressure (model #3) were concordant in their prediction, that is, when two models predicted the same CeREMI, and the third model predicted a different one (Table 4). The best situation was observed when all models (SPI, HR, and MAP combining DD and the past medical history of high arterial pressure) were concordant (APR of 0.79 and Pk of 0.89). This concordance between models occurred in 14 of 30 patients, hence leading to an accurate prediction of remifentanil concentration in 11 of them.

The ordinal model obtained for SPI and combining that parameter with DD and treatment for hypertension (model #3) is illustrated in Figure 3. The graphs show the probability of having received 2, 4, or 6 ng ml^{-1} CeREMI as a function of maximum SPI observed early after HHI and as a function of DD in patients with no history of treated high arterial pressure (Fig. 3A) and in patients with such a history (Fig. 3B). For a given SPI value and a given DD value, the concentration predicted by the model is the one with the highest probability. Similar graphs can be drawn for HR and MAP. Figure 4 is an alternative graphical representation of the model for SPI. It shows the cut-off values of SPI above which or below which the probability of a given CeREMI is highest and may be seen as plane cuts of the surface–response curves presented in Figure 3, at the level of the intersections between surfaces. Similar graphs can also be drawn for the models of HR and MAP. A close look at Figures 3 and 4 shows that: (i) higher SPI values, and higher HR, or MAP values, are associated with higher probabilities of low CeREMI; (ii) an increase in DD (and hence a decrease in intravascular volume) reduces the SPI, HR, and MAP responses to noxious stimulation and hence reduces the thresholds above which or below which the probability of a given CeREMI is greatest; and (iii) those thresholds are even less in patients treated for hypertension.

Discussion

In this study, SPI, HR, and MAP observed after standardized noxious stimulation under propofol–remifentanil anaesthesia had comparable performances at predicting CeREMI, and those performances were better than chance. By reducing the responses of indicators to noxious stimulation, low intravascular volume and routine therapy for hypertension are factors that affect the interpretation of SPI, HR, and MAP. The best prediction accuracy is obtained when those factors are taken into account, and when SPI, HR, and MAP concord in their prediction. Finally, except for HR, none of the tested variables is able to predict CeREMI under stable anaesthetic conditions in the absence of noxious stimulation.

As far as our intention was to assess the ability of recorded variables to distinguish between three different CeREMI (ordinal three-point scale design), the best method to assess the relationship between the SPI, HR, or MAP value and the probability of having received 2, 4, or 6 ng ml^{-1} CeREMI was ordinal regression, and not a more simple ROC curve analysis, which only allows studying a two-point design.

The performance of maximum SPI, HR, or MAP alone at appropriately indicating CeREMI in our patients is comparable with the performance described by Struys and colleagues,[14] with the performance of SPI at detecting movers in response to noxious stimulation,[15] or with the performance of bispectral index, MAP, or HR at detecting the intensity of a calibrated electrical stimulus.[24] Although significantly different from chance alone in this three-option design, such a performance is far from perfect prediction, as further indicated by the APR of each variable, which is in the range of 0.5, and by the large inter-individual variability illustrated by the error bars (SD) in Figure 2. The reason for this poor performance is probably related to confounding factors, a number of which were identified in this study.

One may argue that the prediction of CeREMI could be improved by taking into account the difference between

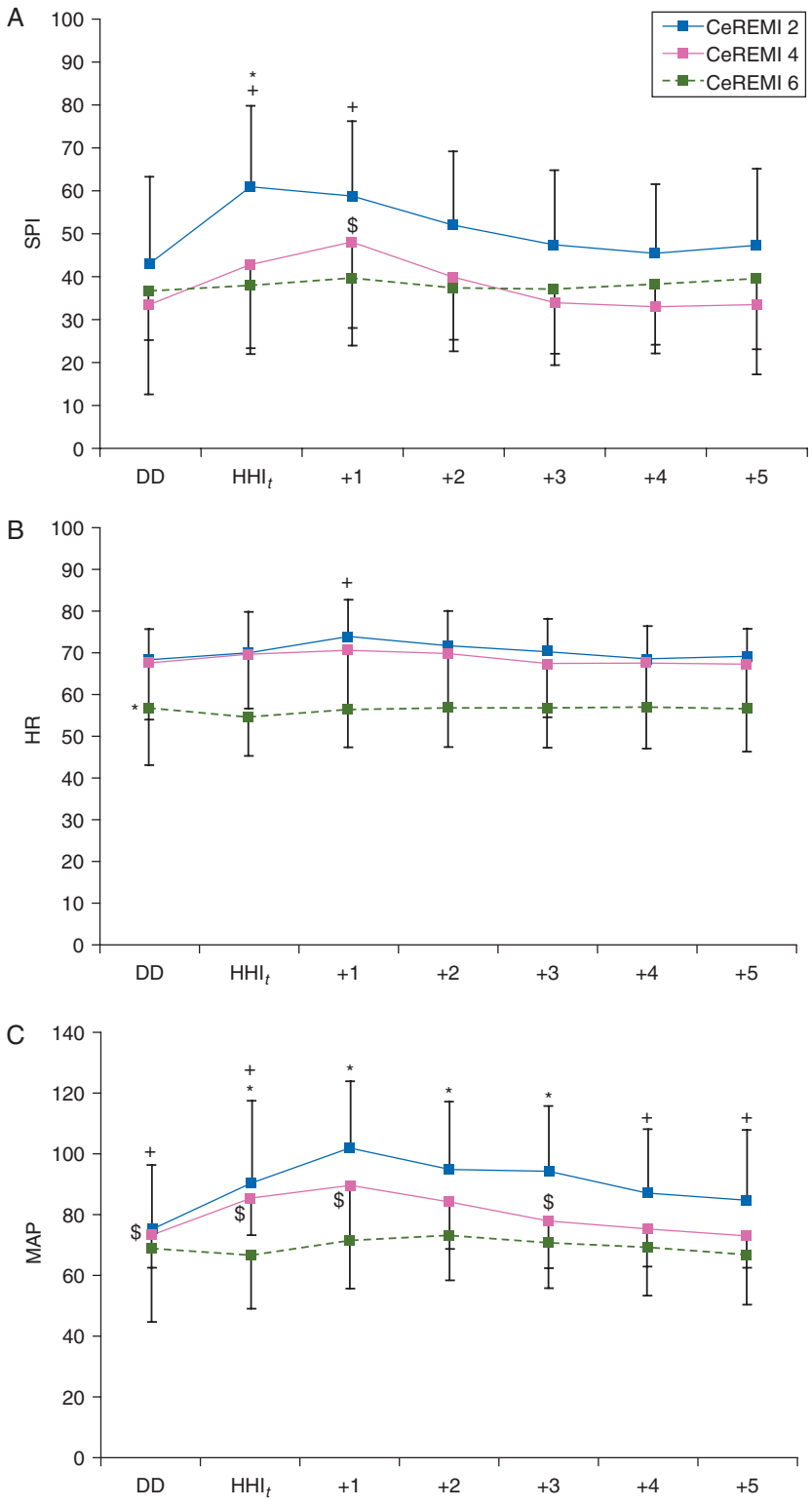

Fig 2 Evolution of 1 min-averaged SPI (A), HR (beat min^{-1}, B), and MAP (mm Hg, C) as a function of time events in each group of patients (see the Methods section for definition of time events). Data are within-group means and error bars correspond to standard deviations. Significant differences are the following. In (A): *at HHI$_t$, [CeREMI 2]>[CeREMI 6]; $^{+}$for CeREMI 2, [HHI$_t$ or +1]>[DD, +3, +4, or +5]; $^{\$}$for CeREMI 4, [+1]>[DD, +3, +4, or +5]. In (B): *globally, [CeREMI 6]< [CeREMI 2 or CeREMI 4]; $^{+}$globally, [+1]> [DD, +4, or +5]. In (C): *at HHI$_t$, +1, +2, or +3, [CeREMI 2]> [CeREMI 6]; $^{+}$for CeREMI 2, [DD]< [HHI$_t$, +1, +2, +3, +4, or +5], [HHI$_t$]< [+1], [+4]< [+1], and [+5]< [+1, +2, or +3]; $^{\$}$for CeREMI 4, [DD]<[HHI$_t$, +1, or +2], [HHI$_t$]>[+4, or +5], [+1]>[+3, +4, or +5], and [+3]>[+5].

Table 2 Pk's indicating the performance of 1 min-averaged SPI, HR, and MAP values at predicting remifentanil concentration. Those Pk's are given with standard error (SE), and 95% confidence interval (95% CI). They were calculated for values of each parameter recorded at the time of delta down measurement (DD) and for their maximum value recorded after HHI (maximum). *Pk value significantly different from 0.5

	Pk (SE)	95 % CI
SPI		
DD	0.55 (0.07)	0.41–0.68
Maximum	0.63* (0.06)	0.50–0.75
HR		
DD	0.63* (0.06)	0.51–0.75
Maximum	0.67* (0.05)	0.57–0.77
MAP		
DD	0.53 (0.07)	0.39–0.68
Maximum	0.70* (0.05)	0.59–0.80

Table 3 Nagelkerke's pseudo-R^2. Results of the test assessing the strength of the relationship between the dependent variable (probability of having received 2, 4, or 6 ng ml^{-1} remifentanil) and the independent variables [SPI, HR, MAP, IVS (DD), and past history of treated high arterial pressure (HAP)], according to the ordinal regression models (see Appendix for details). Note that strength improves with sophistication of the models

Independent variables	Model #	Additional independent variables	Nagelkerke's pseudo-R^2
SPI	1	None	0.200
	2	DD	0.311
	3	DD and HAP	0.394
HR	1	None	0.262
	2	DD	0.373
	3	DD and HAP	0.432
MAP	1	None	0.363
	2	DD	0.48
	3	DD and HAP	0.496

Table 4 Prediction accuracy of ordinal regression models. APR, accurate prediction rate; Pk, prediction probability calculated using actual remifentanil concentration against predicted concentration by the model with standard error (SE) and 95% confidence interval (95% CI)l. Abbreviations of independent variables are identical to those in Table 3. Those parameters are given for the models described in the Methods section and for the situation where models #3 for SPI, HR, and MAP are concordant in their prediction. *n* corresponds to the number of patients where concordance occurred. *Pk's that are significantly different from 0.5

Independent variables	Model #	Additional independent variables	APR	Pk (SE), 95% CI
SPI	1	None	0.50	0.63 (0.089), 0.44–0.81
	2	DD	0.57	0.70* (0.084), 0.53–0.87
	3	DD and HAP	0.60	0.76* (0.072), 0.61–0.90
HR	1	None	0.57	0.74* (0.073), 0.60–0.89
	2	DD	0.5	0.75* (0.061), 0.63–0.88
	3	DD and HAP	0.60	0.79* (0.059), 0.67–0.91
MAP	1	None	0.57	0.78* (0.065), 0.64–0.91
	2	DD	0.60	0.78* (0.067), 0.64–0.91
	3	DD and HAP	0.60	0.78* (0.067), 0.64–0.91
If models #3 are concordant in their prediction (*n*)				
SPI and HR (21)			0.71	0.87* (0.054), 0.76–0.98
SPI and MAP (17)			0.76	0.90* (0.049), 0.80–1.00
HR and MAP (20)			0.70	0.87* (0.045), 0.78–0.97
SPI, HR, and MAP (14)			0.79	0.89* (0.067), 0.75–1.00

the value of a given indicator before noxious stimulation under stable anaesthetic conditions and its maximum value immediately after the stimulation (delta indicator) rather than its maximum value alone. In this study, such changes did not perform better (results not shown), probably because baseline values were highly variable among subjects (Fig. 2). The literature gives conflicting information in this regard, since deltas may improve the detection of movers in response to noxious stimulation,[15] and maximum value performs better at predicting CeREMI.[14] Using maximum values may be closer to real clinical situations and eliminates the definition of a baseline value, and also defining reproducible baseline conditions.

This study highlights the importance of combining information to obtain the best prediction. In the domain of anti-nociception assessment, as in other monitoring domains, the practitioner must look at several variables, in this case SPI, HR, and MAP, and see whether their changes are concordant before drawing conclusions with the maximum accuracy. This combination of variables is not the only efficient combination.[25] In addition, other variables are of concern, since low intravascular volume and the fact of being treated for

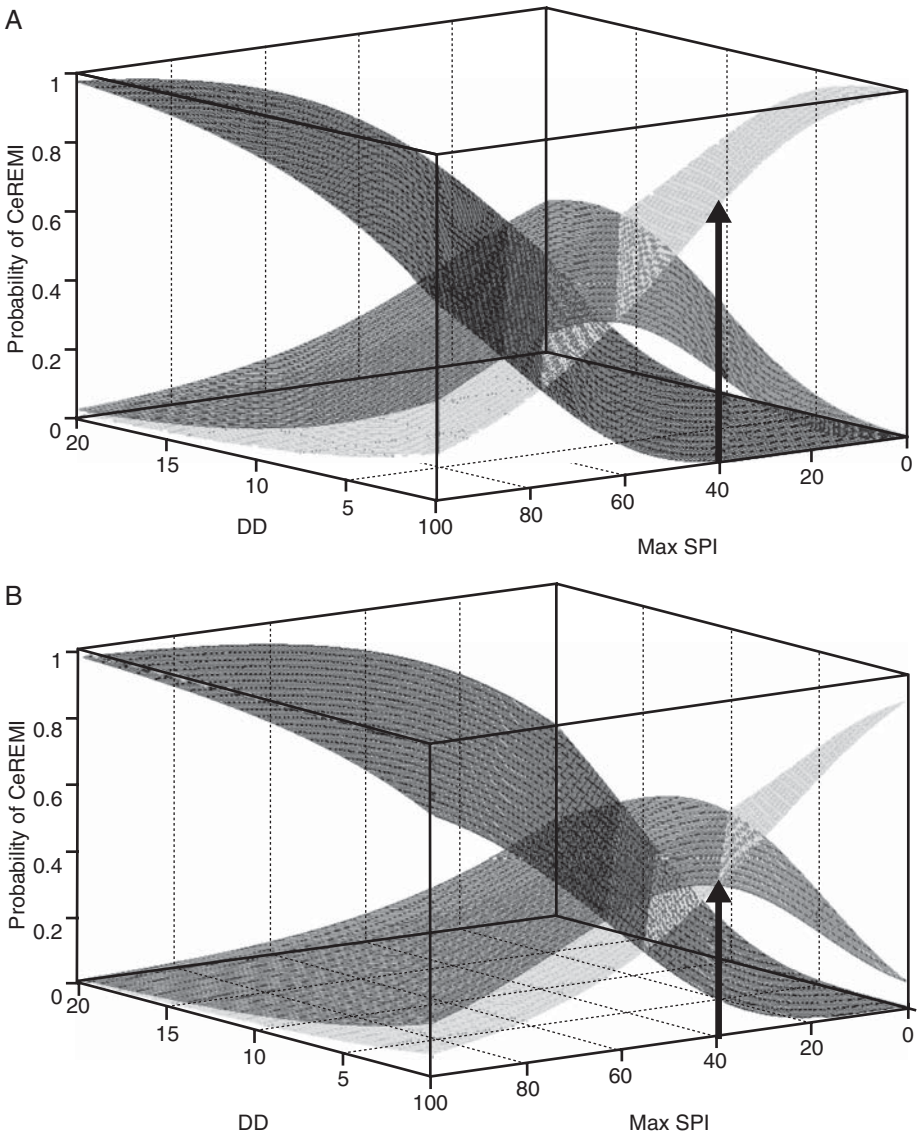

Fig 3 Surface–response curves of the probability of having received 2 ng ml^{-1} (black), 4 ng ml^{-1} (dark grey), or 6 ng ml^{-1} (light grey) CeREMI as a function of delta down (mm Hg, DD), and maximum 1 min-averaged SPI during the minutes following HHI, when patients have no past medical history of high arterial pressure (A), and when they do (B), according to the ordinal regression model #3. The arrows indicate the CeREMI whose probability is highest when maximum 1 min-averaged SPI is 40, and DD is 0, as predicted by the model. This CeREMI is 6 ng ml^{-1} in patients with no history of treated high arterial pressure and 4 ng ml^{-1} in treated hypertensive patients.

hypertension will be associated with lower responses of the concerned indicators to a noxious stimulation. For example, when maximum SPI is 40 and DD is 0 in a patient not on treatment for hypertension, the predicted concentration is 6 ng ml^{-1} (arrow in Fig. 3A). This predicted concentration becomes 4 ng ml^{-1} if the patient receives anti-hypertensive medications chronically (arrow in Fig. 3B).

Our results yield no evidence that SPI or MAP can predict CeREMI under stable anaesthetic conditions in the absence of noxious stimulation. In that situation, only HR had a significant predictive value. In contrast, as displayed in Figure 2, HR showed less variation after HHI than did the other parameters. This may be a sign that HR actually better reflects CeREMI than the nociception–CeREMI balance itself and is in accordance with the finding of Huiku and colleagues[12] that normalized HBI better correlates with CeREMI than with the intensity of noxious stimulation. One possible mechanism could be pharmacodynamic interactions between propofol and remifentanil that would differentially influence HR and the other parameters. The lack of SPI and MAP predictive values in the absence of noxious stimulation also highlights that the moment of the observation is important. Most procedures involve an intense and fairly reproducible noxious event at their beginning, such as

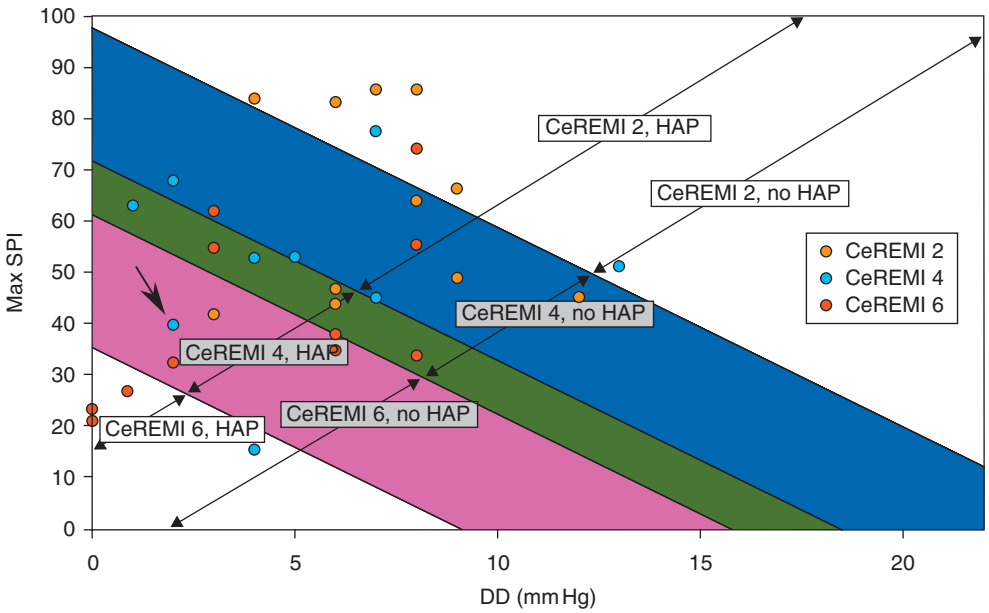

Fig 4 Thresholds of the maximum SPI value observed during the first 5 min after HHI and above which or below which the probability of having received 2, 4, or 6 ng ml^{-1} CeREMI is highest (plain lines), according to the ordinal regression model #3. Those thresholds are given for the situation of a past medical history of treated high arterial pressure (HAP), or not (no HAP). The arrows indicate the limits of the area where the probability of a 2, 4, or 6 ng ml^{-1} CeREMI is highest. Circles are data recorded in the sample of patients. The black arrow points out one observation: DD value was 2 and maximum SPI was 40. The most probable CeREMI received by that patient and predicted by the model would be 4 ng ml^{-1} if the patient has a past medical history of HAP and 6 ng ml^{-1} if not. That patient had such a history of high arterial pressure and actually received 4 ng ml^{-1}. In that case, the prediction by the model is accurate.

skin incision, or, in the case of intracranial neurosurgery, the HHI. This event generally occurs at the end of a period of stable anaesthetic conditions. Here, we show that the maximum response of the indices observed during the minutes after that particular event is of value for assessing the nociception–CeREMI balance. These observations may direct changes to CeREMI in the face of subsequent stimuli.

Maximum HR after noxious stimulation was lowered by low intravascular volume. This is surprising since one would expect higher HRs in hypovolaemic than in normovolaemic patients. However, the relationship between HR and intravascular volume is not continuous.[26] Indeed, moderate hypovolaemia is known to be associated with vagal reflex bradycardia, whereas tachycardia is obvious in the case of large hypovolaemia. Our observations are probably related to these pathophysiological mechanisms that may have been further influenced by the propofol–remifentanil combination. The smaller response in patients treated for hypertension is more easily understandable, as far as the cardiovascular depressant effect of anti-hypertensive medications keeps haemodynamic parameters at lower values.

Our results cannot directly be translated to other situations for the following reasons. First, the propofol–remifentanil combination has specific haemodynamic effects that may differ from those of other anaesthetic combinations. For example, a sevoflurane–remifentanil regimen is known to be much more frequently associated with episodes of bradycardia and hypotension than a propofol–remifentanil combination.[27] Secondly, specific categories of patients were excluded from our study, such as diabetic patients or patients with altered cardiac function, in order to limit the number of potential confounders and simplify analysis. In those patients, SPI, HR, or MAP could probably not be interpreted according to our models. Thirdly, the paradigm of nociception–anti-nociception balance studied here is specific. Although HHI can be considered a reproducible and standardized stimulus (60 N force applied on three skull points), it is not equivalent to other kinds of stimuli, such as skin incision, laryngoscopy, or tetanic electrical stimulation, in terms of intensity. Fourthly, the pharmacokinetic model of Minto and colleagues is not perfect and is associated with some degree of inter-individual variability in real CeREMI. In addition, a given CeREMI may not correspond to the same level of anti-nociception in all patients. Therefore, modifying CeREMI does not necessarily result in equivalent modifications of anti-nociception in all patients. Finally, there is a time delay in the SPI calculation by the commercially available SPI monitor. This probably does not influence the clinical accuracy of our results, as far as that delay is short, and as far as we looked at 1 min-averaged data during 5 min after noxious stimulation.

We conclude that SPI, MAP, and HR responses to a standardized noxious stimulation have comparable predictive values at indicating CeREMI. Interpretation of SPI during

general anaesthesia is improved by taking into account IVS and the fact of treatment for hypertension. Looking at maximum values obtained after noxious stimulation and combining information obtained from SPI, MAP, and HR offers the most accurate prediction. Our results may guide future research for developing new algorithms of anti-nociception assessment.

Acknowledgements

The authors would like to thank Kuider Zaoui, Isabelle Schaub, and Matthieu Clanet, and also the nursing staff of the neurosurgical operating theatre at the CHR Citadelle, Liege, Belgium, for their active participation in this study.

Conflict of interest

K.U. is an employee from GE Healthcare, Oy, Helsinki, Finland, the manufacturer of SPI.

Funding

This study was supported by an unrestricted educational grant from Astra-Zeneca and by the University Department of Anaesthesia and Intensive Care Medicine, CHU Liege, Liege, Belgium. GE Healthcare provided the S5® Collect software.

Appendix

Ordinal regression

Ordinal regression allows modelling the relationship between the probability of observing a particular score or less, in the present case a given remifentanil concentration or less, and several independent predictors. When the negative log–log function is used as the link function, the model has the following form:

$$-\ln[-\ln(\gamma_{i\text{less}})] = \beta_i + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

where ln is the natural logarithm, $\gamma_{i\text{less}}$ the probability of observing the score i or less, β_i a constant threshold for the score i, and β_{1-k} the weighting coefficients for each independent predictor x_{1-k}. In our three-category design, this model allows calculating $\gamma_{2\text{less}}$ that is the probability of observing a remifentanil concentration of 2 ng ml^{-1} (and not less because there is no lower remifentanil concentration) as a function of predictors such as the SPI value, the DD value, and so on. The same can be done for the probability of observing a concentration of 4 ng ml^{-1} or less, namely $\gamma_{4\text{less}}$. Hence, the probability of observing only a concentration of 4 corresponds to $\gamma_{4\text{less}} - \gamma_{2\text{less}}$. Since the probability of observing a concentration of 6 ng ml^{-1} (the highest remifentanil concentration) or less is 1 because there is no higher remifentanil concentration, the probability of observing only a concentration of 6 is $1 - \gamma_{4\text{less}}$.

Once the coefficients for the model have been obtained, it is necessary to check the accuracy of the model. This can be done through the use of several statistical tests. The χ^2 test on the change in -2 log-likelihood is a test of the null hypothesis that the coefficients for all the variables in the model are 0. When it has a significance level of <0.05, it means that the model with predictors is significantly better than the model without predictors. The Pearson goodness-of-fit statistic is an appreciation of the overall goodness of fit of the model. It tests the null hypothesis that the model fits the data, and hence, good models have large significance levels, at least higher than 0.05. The Nagelkerke pseudo-R^2 is a measure of the strength of the association between the dependent variable and the predictor variables. It is similar to the R^2 of a classical linear regression. A Nagelkerke pseudo-R^2 value ranging between 0.2 and 0.4 is considered satisfactory. Finally, the test of parallel lines checks the assumption that the relationships between the independent predictors and the link function of the probability for each category of the dependent variable are parallel lines. It takes also the form of a χ^2 test on a -2 log-likelihood difference. If this test is significant, meaning that the null hypothesis can be rejected, it is possible that the selected link function is incorrect for the data.

References

1 Guignard B. Monitoring analgesia. *Best Pract Res Clin Anaesthesiol* 2006; **20**: 161–80

2 Kochs E, Treede RD, Schulte am Esch J, Bromm B. Modulation of pain-related somatosensory evoked potentials by general anesthesia. *Anesth Analg* 1990; **71**: 225–30

3 Bonhomme V, Llabres V, Dewandre PY, Brichant JF, Hans P. Combined use of Bispectral Index and A-Line Autoregressive Index to assess anti-nociceptive component of balanced anaesthesia during lumbar arthrodesis. *Br J Anaesth* 2006; **96**: 353–60

4 Baars JH, Mager R, Dankert K, Hackbarth M, von DF, Rehberg B. Effects of sevoflurane and propofol on the nociceptive withdrawal reflex and on the H reflex. *Anesthesiology* 2009; **111**: 72–81

5 von DF, Hackbarth M, Mager R, Rehberg B, Baars JH. Monitoring of the responsiveness to noxious stimuli during anaesthesia with propofol and remifentanil by using RIII reflex threshold and bispectral index. *Br J Anaesth* 2010; **104**: 201–8

6 Viertio-Oja H, Maja V, Sarkela M, *et al.* Description of the Entropy algorithm as applied in the Datex-Ohmeda S/5 Entropy Module. *Acta Anaesthesiol Scand* 2004; **48**: 154–61

7 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**: 645–54

8 Shimoda O, Ikuta Y, Sakamoto M, Terasaki H. Skin vasomotor reflex predicts circulatory responses to laryngoscopy and intubation. *Anesthesiology* 1998; **88**: 297–304

9 Gjerstad AC, Storm H, Hagen R, Huiku M, Qvigstad E, Raeder J. Skin conductance or entropy for detection of non-noxious stimulation during different clinical levels of sedation. *Acta Anaesthesiol Scand* 2007; **51**: 1–7

10 Gjerstad AC, Storm H, Hagen R, Huiku M, Qvigstad E, Raeder J. Comparison of skin conductance with entropy during intubation, tetanic stimulation and emergence from general anaesthesia. *Acta Anaesthesiol Scand* 2007; **51**: 8–15

11 Barvais L, Engelman E, Eba JM, Coussaert E, Cantraine F, Kenny GN. Effect site concentrations of remifentanil and pupil response to noxious stimulation. *Br J Anaesth* 2003; **91**: 347–52

12 Huiku M, Uutela K, van Gils M, *et al.* Assessment of surgical stress during general anaesthesia. *Br J Anaesth* 2007; **98**: 447–55

13 Wennervirta J, Hynynen M, Koivusalo AM, Uutela K, Huiku M, Vakkuri A. Surgical stress index as a measure of nociception/anti-nociception balance during general anesthesia. *Acta Anaesthesiol Scand* 2008; **52**: 1038–45

14 Struys MM, Vanpeteghem C, Huiku M, Uutela K, Blyaert NB, Mortier EP. Changes in a surgical stress index in response to standardized pain stimuli during propofol-remifentanil infusion. *Br J Anaesth* 2007; **99**: 359–67

15 Gruenewald M, Meybohm P, Ilies C, *et al.* Influence of different remifentanil concentrations on the performance of the surgical stress index to detect a standardized painful stimulus during sevoflurane anaesthesia. *Br J Anaesth* 2009; **103**: 586–93

16 Ahonen J, Jokela R, Uutela K, Huiku M. Surgical stress index reflects surgical stress in gynaecological laparoscopic day-case surgery. *Br J Anaesth* 2007; **98**: 456–61

17 Hans P, Brichant JF, Dewandre PY, Born JD, Lamy M. Effects of two calculated plasma sufentanil concentrations on the hemodynamic and bispectral index responses to Mayfield head holder application. *J Neurosurg Anesthesiol* 1999; **11**: 81–5

18 Faul F, Erdfelder E. *GPOWER: A Priori, Post-hoc, and Compromise Power Analyses for MS-DOS (Computer Program)*. Bonn, FRG: Bonn University, Dep. of Psychology, 1992

19 Minto CF, Schnider TW, Shafer SL. Pharmacokinetics and pharmacodynamics of remifentanil. II. Model application. *Anesthesiology* 1997; **86**: 24–33

20 Marsh B, White M, Morton N, Kenny GN. Pharmacokinetic model driven infusion of propofol in children. *Br J Anaesth* 1991; **67**: 41–8

21 Deflandre E, Bonhomme V, Hans P. Delta down compared with delta pulse pressure as an indicator of volaemia during intracranial surgery. *Br J Anaesth* 2008; **100**: 245–50

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

23 Smith WD, Dutton RC, Smith NT. A measure of association for assessing prediction accuracy that is a generalization of non-parametric ROC area. *Stat Med* 1996; **15**: 1199–215

24 Takamatsu I, Ozaki M, Kazama T. Entropy indices *vs* the bispectral index for estimating nociception during sevoflurane anaesthesia. *Br J Anaesth* 2006; **96**: 620–6

25 Seitsonen ER, Korhonen IK, van Gils MJ, *et al.* EEG spectral entropy, heart rate, photoplethysmography and motor responses to skin incision during sevoflurane anaesthesia. *Acta Anaesthesiol Scand* 2005; **49**: 284–92

26 Secher NH, Jacobsen J, Friedman DB, Matzen S. Bradycardia during reversible hypovolaemic shock: associated neural reflex mechanisms and clinical implications. *Clin Exp Pharmacol Physiol* 1992; **19**: 733–43

27 Bonhomme V, Demoitie J, Schaub I, Hans P. Acid-base status and hemodynamic stability during propofol and sevoflurane-based anesthesia in patients undergoing uncomplicated intracranial surgery. *J Neurosurg Anesthesiol* 2009; **21**: 112–9