Anesthesiology 2009; 110:1068–76

Postoperative Neurocognitive Dysfunction in Elderly Patients after Xenon versus *Propofol Anesthesia for Major Noncardiac Surgery*

A Double-blinded Randomized Controlled Pilot Study

Jan Höcker, M.D., Claudia Stapelfeldt, M.D.,* Jörn Leiendecker, M.D.,* Patrick Meybohm, M.D.,* Robert Hanss, M.D.,* Jens Scholz, M.D.,† Berthold Bein, M.D., D.E.A.A.‡*

Background: **Postoperative cognitive dysfunction (POCD) in elderly patients after noncardiac surgery is a common problem. The noble gas xenon has been demonstrated to exert substantial neuroprotective properties in animal studies. Therefore, this study was designed to assess POCD after xenon anesthesia in comparison to propofol in elderly patients undergoing major noncardiac surgery.**

Methods: **After approval of the local ethical committee was obtained, 101 patients (American Society of Anesthesiologists physical status I–III; age, 65–83 yr) undergoing elective abdominal or urologic surgery (duration, > 2 h) were enrolled into this randomized, double-blinded controlled pilot study. Patients received anesthesia with sufentanil and either propofol or xenon and were assessed before treatment and 1, 6, and 30 days after treatment using a neuropsychological test battery based on previous studies investigating POCD.**

Results: **There were no significant differences in terms of age, American Society of Anesthesiologists status, education, duration of surgery, administered analgetics, and preoperative neurocognitive status between study groups. POCD as classified was present in 22 patients (44%) of the xenon group *versus* 25 patients (50%) of the propofol group 1 day after treatment, in 6 xenon patients (12%) *versus* 9 propofol patients (18%) 6 days after treatment, and in 3 xenon patients (6%) *versus* 6 propofol patients (12%) 30 days after treatment. These differences were not statistically significant.**

Conclusion: **Postoperative impairment of neurocognitive function was observed in a substantial proportion of elderly patients even 30 days after treatment. Xenon-based anesthesia was not associated with decreased incidence of POCD in comparison to propofol.**

FOR years, postoperative cognitive dysfunction (POCD) was mainly regarded as a problem associated with cardiac surgery. In 1998, however, the International Study of Postoperative Cognitive Dysfunction (ISPOCD1) demonstrated POCD also after major noncardiac surgery in 26% of elderly patients (aged 60 yrs or older) 1 week after general anesthesia and in 10% even 3 months after general anesthesia.[1] In middle-aged patients (aged 40–60 yrs), the incidence of POCD was significantly lower after 1 week and also after 3 months, when POCD had almost disappeared.[2] Recently, several independent risk factors for POCD at 3 months after major surgery have been identified, including increased age (60 yrs or older), a previous cerebral vascular incident with no residues and POCD at discharge from hospital.[3] This study also revealed that POCD at hospital discharge is associated with an increased risk of long-term cognitive problems only in elder patients. Furthermore, patients with POCD are at an increased risk to die in the first year after surgery.[3] However, different anesthetic regimes have not been evaluated in these studies.

Among the currently available anesthetic agents, especially the noble gas xenon has been shown to induce neuroprotective effects in different animal models of cerebral ischemia or neurologic dysfunction.[4–6] In an animal model of cardiopulmonary bypass-induced neurologic and neurocognitive dysfunction, xenon administration resulted in an improved neurocognitive outcome.[4] Xenon's anesthetic action as well as its neuroprotective potential have been attributed, at least in part, to the ability of inhibiting glutamatergic *N*-methyl-D-aspartate receptors,[7] which play a major role in the dissemination of acute neuronal injury.[8]

Our study was designed to investigate the hypothesis that xenon anesthesia in comparison to propofol may decrease the rate of postoperative impairment and cognitive dysfunction after long-lasting major noncardiac surgery. We analyzed the incidence of early (1 and 6 days after surgery) neurocognitive impairment and intermediate (30 days after surgery) POCD in elderly patients (aged 65 yrs or older).

Materials and Methods

Study Design

The study was performed as a prospective, double-blind, randomized pilot trial. Sealed envelopes for each patient were used for randomization. The physician performing the neuropsychological tests and the patient were blinded to the anesthetic used.

* Staff Anesthesiologist, † Professor and Chair, ‡ Associate Professor, Department of Anaesthesiology and Intensive Care Medicine, University Hospital Schleswig-Holstein, Campus Kiel, Germany.

Received from the Department of Anaesthesiology and Intensive Care Medicine, University Hospital Schleswig-Holstein, Campus Kiel, Germany. Submitted for publication May 18, 2008. Accepted for publication December 10, 2008. Support was provided solely from institutional and/or departmental sources. Air Liquide, Duesseldorf, Germany, provided the xenon

Address correspondence to Dr. Höcker: Department of Anaesthesiology and Intensive Care Medicine, University Hospital Schleswig-Holstein, Campus Kiel, Schwanenweg 21, 24105 Kiel, Germany. hoecker@anaesthesie.uni-kiel.de.

Table 1. Number of Patients with Cognitive Decline or Improvement in Each Neuropsychological Test 1 Day, 6 Days, and 30 Days after Surgery Compared with Baseline Values

	Number of Patients with Classified Decline (Improvement)					
	After 1 Day		After 6 Days		After 30 Days	
	XE	PRO	XE	PRO	XE	PRO
Memory						
RAVLT 1–3	NI	NI	5 (1)	6 (0)	3 (4)	2 (5)
RAVLT LT	NI	NI	5 (0)	7 (2)	4 (1)	5 (7)
Attention						
STROOP No. 1	19 (3)	26 (2)	9 (6)	11 (3)	7 (5)	4 (6)
STROOP No. 2	2 (4)	6 (2)	2 (8)	1 (4)	0 (1)	1 (0)
TMT	7 (2)	8 (1)	3 (3)	2 (3)	1 (2)	4 (2)
Digit span	7 (6)	4 (6)	5 (5)	2 (8)	2 (2)	4 (8)
DSST	6 (3)	6 (0)	4 (3)	4 (1)	2 (2)	0 (3)
Motor skills						
PBT dominant	11 (0)	13 (0)	5 (1)	1 (5)	0 (6)	0 (5)
PBT nondominant	7 (0)	10 (3)	3 (1)	2 (5)	0 (1)	2 (3)
Executive function						
VFT semantic	13 (8)	12 (7)	7 (8)	3 (7)	1 (4)	6 (6)
VFT phonetic	8 (4)	13 (5)	7 (4)	6 (4)	5 (4)	2 (5)
Total number of patients (n)	22 (1)	25 (2)	6 (7)	9 (8)	3 (11)	6 (8)

Digit Span = digit span test; DSST = digit symbol substitution test; NI = not investigated; PBT dominant = Purdue pegboard test, performance with preferred hand; PBT nondominant = Purdue pegboard test, performance with nonpreferred hand/other hand; PRO = propofol; RAVLT 1–3 = Rey's auditorial verbal learning test, first to third presentation of words (short-term memory); RAVLT LT = Rey's auditorial verbal learning test, long-term memory; STROOP No. 1 = Stroop color word interference test, first run; STROOP No. 2 = Stroop color word interference test, second run; TMT = trail making test; VFT semantic = verbal fluency test, semantic categories; VFT phonetic = verbal fluency test, phonetic categories; XE = xenon (details of the test performance are described in Materials and Methods).

Subject Enrollment

After approval by the institutional review board (University Kiel, Schleswig-Holstein, Germany) and written informed consent, 114 patients (aged 65 yrs or older) undergoing elective major noncardiac surgery at the University Hospital Schleswig-Holstein, Campus Kiel, between September 2006 and January 2008 were enrolled in the study. Thirteen patients were lost to follow-up. Data of 101 patients were included in the final analysis. Additional inclusion criteria were abdominal or urologic surgery under general anesthesia of more than 2 h, fluency in German, ability to read, and American Society of Anesthesiologists physical status physical status I–III. Patients were excluded from the study in case of diseases of the central nervous system, including dementia (defined as a Mini-Mental score examination [MMSE] score below 24), a current or past psychiatric illnesss, current use of tranquilizers or antidepressants, drug dependence (including alcoholism), severe visual, auditory, or motor handicap or acute infections.

Preoperative Evaluation and Neurocognitive Assessment

All patients were visited and tested the day before surgery. Demographic data were analyzed, including education, medical history, and current medication. Signs and symptoms of depression were assessed preoperatively using the Beck depression inventory,[9] and signs of dementia were investigated with the MMSE.

Neuropsychological evaluation comprised functions of the cognitive domains memory, attention, executive function, and motor skills and was performed 1 day before (baseline) and 1 day, 6 days (± 1 day), and 30 days (± 2 days) after surgery. The test battery used was based on the neuropsychological tests used in the International Studies of POCD 1 and 2[1,2] supplemented by evaluation of motor skills and verbal fluency. If patients did not perform all four neurocognitive evaluations, they were excluded from the study.

One day after surgery, Rey's auditorial verbal learning test was not performed because of time and organizational restrictions (see tables 1–3). All patients were tested by one experienced physician. At 30 days (± 2 days) after surgery, patients were visited at home (except 5 patients who were still hospitalized) to perform the final neuropsychological evaluation.

Memory was evaluated with Rey's auditorial verbal learning test (RAVLT).[10] The number of words recalled after the first, the second, and the third presentation of a 10-word list were summed up to determine a patient's short-term memory (RAVLT 1-3). A free recall after 20 min was used to evaluate long-term memory (RAVLT long term). Parallel versions of the word list were randomly used to minimize practice effects.

Attention was tested with the trail making test (TMT) part A. The patient had to connect consecutive numbers as fast as possible. The required time was measured. Furthermore, patients performed the Stroop color word

Table 2. Results of Neuropsychological Assessment before and 1 Day, 6 Days, and 30 Days after Surgery

	Before Surgery		After 1 Day		After 6 Days		After 30 Days	
Neurocognitive Test	XE	PRO	XE	PRO	XE	PRO	XE	PRO
Mini-Mental Status Score	28 (2.1)	27 (2.4)	NI	NI	NI	NI	NI	NI
Depression Inventory Score	4.2 (2.6)	4.6 (3.1)	NI	NI	4.9 (3.0)	5.2 (3.2)	5.5 (4.5)	5.0 (4.8)
Memory								
RAVLT 1–3	18 (8–6)	18 (9–27)	NI	NI	17 (6–32)	15 (4–30)	19 (9–30)	19 (10–32)
RAVLT LT	6 (2–10)	5 (0–9)	NI	NI	5 (2–10)	5 (1–9)	6 (2–10)	5 (0–9)
Attention								
STROOP No. 1	14 (8–25)	13 (7–27)	18 (10–47)	18 (8–48)	16 (8–42)	15 (7–40)	16 (9–44)	12 (1–24)
STROOP No. 2	41 (16–112)	40 (16–95)	41 (14–87)	44 (16–98)	35 (11–94)	36 (13–93)	34 (12–107)	34 (11–96)
TMT	51 (22–88)	47 (28–76)	56 (29–98)	54 (24–78)	46 (23–102)	44 (26–76)	43 (29–86)	45 (20–96)
Digit span	12.5 (3.5)	12.8 (3.5)	12.0 (3.5)	12.1 (3.2)	12.9 (3.9)	13.3 (3.7)	12.8 (3.8)	13.7 (4.4)
DSST	35.3 (10.8)	37.1 (9.7)	32.8 (10.9)	33.8 (10.6)	38.3 (12.6)	36.5 (11.5)	37.4 (10.6)	40.5 (9.6)
Motor skills								
PBT dominant	96 (21)	104 (27)	115 (39)	129 (52)	102 (39)	104 (34)	92 (28)	90 (23)
PBT nondominant	104 (36)	116 (40)	129 (49)	132 (54)	109 (44)	109 (38)	103 (40)	111 (42)
Executive function								
VFT semantic	18.6 (7.4)	18.8 (7.6)	17.0 (6.3)	18.1 (7.3)	19.2 (7.0)	18.3 (6.8)	20.9 (7.0)	19.6 (7.9)
VFT phonetic	10.7 (4.3)	10.0 (4.4)	9.1 (4.0)	8.5 (3.9)	10.6 (4.4)	9.4 (4.4)	11.2 (5.5)	10.0 (4.0)

Data are presented as mean (SD) if normally distributed. Nonparametric data are presented as median (range). Digit Span = digit span test; DSST = digit symbol substitution test; NI = not investigated; PBT dominant = Purdue pegboard test, performance with preferred hand; PBT nondominant = Purdue pegboard test, performance with nonpreferred hand/other hand; PRO = propofol; RAVLT 1–3 = Rey's auditorial verbal learning test, first to third presentation of words (short-term memory); RAVLT LT = Rey's auditorial verbal learning test, long-term memory; STROOP No. 1 = Stroop color word interference test, first run; STROOP No. 2 = Stroop color word interference test, second run; TMT = trail making test; VFT semantic = verbal fluency test, semantic categories; VFT phonetic = verbal fluency test, phonetic categories; XE = xenon (details of the test performance are described in Materials And Methods).

interference test (STROOP test) by receiving a sheet of paper with the words green, blue, yellow, and red (in German) in incongruent ink colors. In the first run, the participants had to read out the words loudly, ignoring the name of the color (STROOP test No. 1). In the second run, the patients read the ink color ignoring the meaning of the word (STROOP test No. 2). We measured the reading time of both runs. Patients were also tested with the Digit symbol substitution test (DSST). Patients had to substitute defined symbols on a sheet of paper with numbers. The number of correctly substituted symbols in 90 s was counted and transformed into a score. Finally, we performed the Digit Span test. First, the patients listened to a list of 3–9 digits and had to repeat

Table 3. Standardized Change Scores of Neuropsychological Assessment 1 Day, 6 Days, and 30 Days after Surgery

	After 1 Day			After 6 Days			After 30 Days		
	XE	PRO	P	XE	PRO	P	XE	PRO	P
Memory									
RAVLT 1–3	NI	NI	NI	–0.24 (0.35)	–0.32 (0.47)	0.09	0.20 (1.08)	0.22 (0.76)	0.46
RAVLT LT	NI	NI	NI	–0.16 (0.43)	–0.12 (0.52)	0.24	–0.06 (0.77)	0.15 (1.19)	0.19
Attention									
STROOP No. 1	–1.37 (2.61)	–1.18 (1.54)	0.37	–0.10 (2.57)	–0.49 (1.38)	0.18	0.24 (1.14)	0.10 (0.84)	0.39
STROOP No. 2	–0.04 (0.74)	–0.21 (0.78)	0.14	0.23 (1.33)	0.25 (0.73)	0.47	0.29 (1.29)	0.36 (0.57)	0.39
TMT	–0.28 (1.39)	–0.38 (0.76)	0.33	0.23 (0.79)	0.14 (0.62)	0.27	0.40 (0.47)	0.08 (0.72)	0.12
Digit span	0.01 (1.08)	–0.17 (1.06)	0.20	0.17 (1.05)	0.09 (0.51)	0.30	0.03 (0.44)	0.16 (0.63)	0.15
DSST	–0.35 (0.79)	–0.45 (1.02)	0.29	0.16 (1.16)	–0.10 (0.74)	0.11	0.19 (0.89)	0.32 (0.53)	0.22
Motor skills									
PBT dominant	–0.4 (1.48)	–0.61 (0.84)	0.21	–0.24 (0.97)	0.04 (0.63)	0.05	0.28 (0.61)	0.36 (0.58)	0.13
PBT nondominant	–0.6 (0.73)	–0.54 (1.40)	0.41	–0.09 (0.63)	0.21 (1.02)	0.06	0.05 (0.40)	0.11 (0.70)	0.34
Executive function									
VFT semantic	–0.14 (1.31)	–0.08 (1.31)	0.41	0.15 (1.02)	–0.02 (1.10)	0.22	0.34 (1.21)	–0.02 (0.96)	0.08
VFT phonetic	–0.03 (1.02)	–0.11 (1.06)	0.36	–0.03 (1.02)	–0.11 (1.06)	0.36	–0.02 (1.09)	0.04 (0.75)	0.39
Summarized change score	–2.92 (5.43)	–3.56 (4.39)	0.26	0.49 (4.25)	–0.01 (3.14)	0.25	1.42 (3.50)	1.61 (2.78)	0.38

Negative scores indicate deterioration; positive scores indicate improvement of cognitive function. Data are presented as mean (SD). Digit Span = digit span test; DSST = digit symbol substitution test; NI = not investigated; PBT dominant = Purdue pegboard test, performance with preferred hand; PBT nondominant = Purdue pegboard test, performance with nonpreferred hand/other hand; PRO = propofol; RAVLT 1–3 = Rey's auditorial verbal learning test, first to third presentation of words (short-term memory); RAVLT LT = Rey's auditorial verbal learning test, long-term memory; STROOP No. 1 = Stroop color word interference test, first run; STROOP No. 2 = Stroop color word interference test, second run; TMT = trail making test; VFT semantic = verbal fluency test, semantic categories; VFT phonetic = verbal fluency test, phonetic categories; XE = xenon (details of the test performance are described in Materials and Methods).

them correctly. In the second run, the digits had to be repeated inverse to the listening order. The total number of correctly cited digits was measured.

Executive functions were quantified with the verbal fluency test (VFT), including semantic and phonetic categories.[10] These categories were pseudorandomized each time, choosing male and female first names, animals, countries, professions, hobbies, and plants for semantic categories and different alphabetical letters for phonetic fluency. The sum of all runs was scored.

Motor skills were evaluated with the Purdue pegboard test (PBT).[10] The patients had to place pegs into the appropriate slots with their preferred hand (dominant) in the first and the other hand (nondominant) in the second run. The required time was measured, and the sum of both runs was quantified.

Protocol

Patients were premedicated with 0.07–0.1 mg/kg midazolam 1 h before anesthesia. Standard monitoring for every patient included continuous registration of electrocardiogram, pulse oximetry, relaxometry, and invasive or noninvasive measurement of blood pressure (1- to 5-min intervals). Moreover, Bispectral Index (version 4.0) (BIS XP sensor electrodes, Aspect Medical Systems, Norwood, MA; and BIS module, Datex-Ohmeda, Instrumentarium Corp., Helsinki, Finland) as well as State and Response Entropy (M-entropy module, Datex-Ohmeda, Instrumentarium Corp.) were recorded. The measurement of the electroencephalogram-derived parameters was performed by the manufacturer's standard electrodes mounted side by side on the patient's temporal-frontal forehead after careful skin preparation. Their position was pseudorandomized by flipping a coin to avoid brain hemispheric effects. Smoothing time was set to a 15-s interval for the BIS. The moving frequency-related time windows for entropy calculation were 15–60 s and 2–15 s for State Entropy and Response Entropy. Sampling rate for the raw electroencephalogram was 256 Hz for BIS and 400 Hz for entropy. Electrode impedances were considered as acceptable if less than 10 kΩ for BIS and 7.5 kΩ for entropy. The signal quality index of the BIS sensor was checked every 10 min to ensure appropriate signal quality. BIS was used as additional monitoring and not primarily for guidance of anesthesia according to a defined target value. However, a patient safety algorithm monitoring was performed. In case of BIS levels of more than 60 during maintenance of anesthesia, a rescue medication (propofol bolus 0.5 mg/kg) would be administered to guarantee a sufficient depth of anesthesia. These patients had to be excluded from the study because the study design did not allow any supplementary anesthetic in both groups.

Anesthesia was induced in all patients with 0.2–0.4 mg/kg etomidate and 0.3–0.5 μg/kg sufentanil. Tracheal intubation was facilitated with 0.6 mg/kg rocuronium followed by at least 10 min of denitrogenization with 100% O_2. Anesthesia was maintained with 0.1–0.2 mg/kg bolus etomidate during denitrogenization. Afterwards, patients were randomly assigned to receive anesthesia with either xenon (60% in O_2) using a closed-circuit anesthesia machine (Physioflex; Draeger, Luebeck, Germany) or propofol (Propofol-Lipuro 1%; Braun, Melsungen, Germany) (as target controlled infusion; target plasma level, 2–4 μg/ml plasma) and sufentanil. An Alaris Asena PK target controlled infusion pump (software version 3.2.15; Alaris Medical Systems, Baesweiler, Germany) was used to infuse propofol. Target plasma concentrations were estimated by the target controlled infusion pump based on the Marsh model.[11] Xenon in medical quality was provided by Air Liquide (Duesseldorf, Germany). The inspiratory fraction of oxygen in both groups was 35 ± 2%. An increase in systolic blood pressure or heart rate by more than 20% from baseline (10 min after the end of the denitrogenization period) was treated with 10-μg bolus of sufentanil. If the surgical procedure required muscle relaxation, additional 10-mg boluses of rocuronium were given. At the end of surgery, patients received neostigmine to reverse muscle relaxation if train-of-four relaxometry revealed a train-of-four ratio less than 0.9. Vasoactive drugs could be administered during anesthesia if required at the decision of the attending anesthesiologist. Every patient received 1 g of metamizole IV 15 min before end of surgery. Administration of anesthetics was stopped when all surgical interventions were completed. Extubation was performed when the patient breathed spontaneously and opened eyes on command. If necessary, 3.75–7.5 mg of piritramide was administered. After extubation, the patient was transferred to the postanesthesia care unit. After patients were discharged from postanesthesia care unit, pain therapy was performed by a patient-controlled analgesia device for IV infusion of 1.5- to 2-mg bolus of piritramide. The patients were evaluated daily with a numeric rating scale to quantify pain intensity.

Statistical Analyses

Statistical analyses were performed using commercially available statistics software (SPSS for Windows, version 11.5, SPSS Inc., Chicago, IL; GraphPad Prism 5.0, GraphPad Software, San Diego, CA).

Values between groups were compared using unpaired Student *t* test or Mann–Whitney U-test. Proportions were compared with Fisher exact test or chi-square test, as appropriate. Confidence intervals (95% CI) for the event rates actual proportions of POCD after 30 days, 6 days, and 1 day were calculated. Hemodynamic data and BIS values were analyzed using two-way repeated measures analysis of variance with Bonferroni correction for multiple comparisons factoring for time and anesthetic agent. $P < 0.05$ was considered statistically significant. Gaussian distribution of each of the preprocedural

neuropsychological test results was examined with the Kolmogorov-Smirnov test. The results of the RAVLT, STROOP test, and trail making test were not normally distributed, and logarithmic transformation was performed to achieve normal distribution.[12] We applied the definition of POCD, which has been used by previous studies investigating cognitive changes after different surgical procedures.[13–15] A cognitive change was assumed if the preoperative to postoperative difference in 2 or more tasks assessing different cognitive domains exceeded more than one SD (table 3). To analyze how many patients of each group cognitively declined or potentially improved, we calculated the SD of each preoperative test on the basis of all patients. Due to a lacking control group, the influence of learning effects on neurocognitive testing could not be analyzed. Therefore, the evaluation of cognitive decline or improvement was limited to a between-group comparison.

Finally, we calculated a standardized change score (postoperative test result subtracted from preoperative [baseline] test result divided through the test specific preoperative SD) for each patient. This score indicates the individual change in performance. If appropriate, we changed the algebraic sign so that positive changes indicate improvement, whereas negative signs reflect deterioration. We calculated the sum of standardized change scores of all tests and compared the patients of the xenon group to the patients of the propofol group by a one-way analysis of variance.

The primary outcome was the mean probability of a reduction of POCD as defined in our study in the xenon group 30 days after surgery. Sample-size calculation was based on two previous studies investigating POCD after xenon *versus* propofol or desflurane anesthesia.[16,17] However, both did not reveal significant differences. Furthermore, sample-size calculation was based on *in vitro* and *in vivo* animal studies demonstrating rather large neuroprotective effect sizes in different models of neuronal injury.[4,5,18] Therefore, we made the reasonable assumption that the incidence of POCD could be reduced from 30% in the propofol group to 10% in the xenon group, and we tripled the number of patients enrolled compared to previous studies. For a power of 80% and an α-error of 5%, a total of 49 patients in each group was calculated. To compensate for dropouts, we enrolled 57 patients in each group. This power analysis, including the underlying assumption that was not based on ideal empirical basis, should be regarded as a starting point to recruit a sample for further investigation of this issue.

Results

There were no significant differences between groups with respect to operative and demographic data and to perioperatively and postoperatively administered analgesics or postoperative pain intensity (tables 4 and 5).

Table 4. Demographic Data, Duration and Type of Surgery, and Duration of Anesthesia in Patients in the Xenon Group and the Propofol Group

	Xenon (n = 50)	Propofol (n = 51)
Age, yr	72 (6)	72 (6)
Gender, F:M	11:39	11:40
Weight, kg	78 (11)	81 (12)
Body mass index, kg/m^2	25 (2.7)	26 (3.5)
ASA physical status, I:II:III	0:38:12	1:34:16
Duration of surgery, min	170 (75)	178 (64)
Type of surgery, abdominal:urologic	19:31	16:35
Duration of anesthesia, min	218 (73)	217 (55)

Data are presented as mean (SD) or absolute number. ASA = American Society of Anesthesiologists; IV = intravenously; M = male; PACU = postoperative care unit; PO = per os. No difference between groups.

A total of 13 of 114 patients were lost to follow-up (6 patients of the xenon group; 7 patients of the propofol group), so 101 patients were analyzed. The reasons for loss to follow-up neurocognitive assessment are presented in table 6. No significant differences of the preoperative test results (baseline) were detected between study groups. Furthermore, no significant differences of the duration of the preoperative and postoperative hospitalization of patients were detected between study groups (table 7).

Neurocognitive Changes 1 Day after Treatment Compared to Baseline Results

In two or more tests of different cognitive domains, 22 patients of the xenon group (44%) *versus* 25 patients of the propofol group (50%) showed cognitive deterioration (table 1, 2). However, one patient of the xenon

Table 5. Analgesics Administered Perioperatively and Postoperatively within First 2 Hours at PACU and after Discharge from PACU

Analgesic Drug	Xenon (n = 50)	Propofol (n = 51)
Perioperatively administered		
Sufentanil, μg	65 (25–105)	65 (25–115)
Postoperatively administered at PACU		
Piritramide, mg	7.5 (0–15)	11.4 (3.8–18.4)
Postoperatively administered after discharge from PACU (1–3 postoperative day)		
Piritramide IV, mg/d	16 (3–42)	18 (4–38)
Metamizole IV, g/d	4 (0–5)	4 (0–5)
Postoperatively administered after discharge from PACU (4–6 postoperative day)		
Piritramide IV, mg/d	8 (0–16)	6 (0–20)
Metamizole IV, g/d	4 (0–5)	3 (0–5)
Tramadole PO, mg/d	26 (0–100)	32 (0–100)

Data are presented as median (range). PACU = postoperative care unit. No difference between groups.

Table 6. Characteristics of Patients Lost to Follow-up Neurocognitive Assessment

	Xenon (n = 6)	Propofol (n = 7)
Refused further assessment	4	3
Further assessment not possible – other than refusal	2	3
Dead	0	1

group and three patients of the propofol group improved in their performance. These observations did not reach statistical significance (95% CI −0.09–0.19]. A comparison of the standardized change scores revealed no significant differences between groups (table 3). Neurocognitive decline was detected in 47 (47%) of 101 patients, whereas 4 patients improved. Most of the patients classified as cognitively declined failed to adequately perform the STROOP test, which assesses domain memory, and the Purdue Pegboard Test, which assesses motor skills (table 1).

Neurocognitive Changes 6 Days after Treatment Compared to Baseline Results

Six days after treatment, the complete comprehensive test battery was tested, including the RAVLT to investigate the cognitive domain memory. Six patients of the xenon group (12%) *versus* 9 patients of the propofol group (18%) showed neurocognitive impairment (95% CI −0.08–0.31), whereas 7 xenon patients and 8 propofol patients improved. Standardized change scores did not differ significantly.

Neurocognitive Changes 30 Days after Treatment Compared to Baseline Results

One month after treatment, POCD according to our definition was recognized in 3 patients of the xenon group (6%). On the other hand, 11 xenon-patients (22%) showed improved performance. In the propofol group, 6 patients (12%) deteriorated compared with baseline values, whereas 8 patients (16%) improved (95% CI −0.06–0.42). Again, neither these differences nor the summarized change scores of all performed tests were significant between groups.

Hemodynamic Data and BIS Values

Before induction of anesthesia, baseline values of heart rate, mean arterial blood pressure, BIS, entropy, and oxygen saturation were not different between the xenon and the propofol group.

Heart rate was significantly lower in the xenon group 15, 30, 45 min ($P < 0.001$), and 90 min ($P = 0.039$) after induction of anesthesia, whereas heart rate in the propofol group decreased 5 min after extubation compared to the xenon group ($P = 0.001$). Mean arterial blood pressure in the xenon group was higher than in the propofol group 15 and 30 min ($P = 0.001$) and 45 min ($P = 0.016$) after induction of anesthesia. These differences disappeared later. The amount of vasoactive drugs used was comparable in both groups. No differences between groups were found when analyzing BIS (fig. 1) and oxygen saturation values. The analysis of entropy values was not possible because of technical problems with loss of data.

Table 7. Duration of the Preoperative and Postoperative Hospitalization of Patients in Both Study Groups

	Xenon (n = 50)	Propofol (n = 51)
Preoperative hospitalization, days	1.9 (0.5–7)	1.8 (0.5–5)
Postoperative hospitalization, days	11 (3–32)	13 (3–28)

Data are presented as median (range). No difference between groups.

Discussion

In our study, we did not detect significant differences comparing the neurocognitive function after xenon *versus* propofol anesthesia in elderly patients. One day after treatment, 47% of the patients showed neurocognitive impairment as defined in our study. Six days after treatment, POCD was recognized in 15% of the patients enrolled, whereas this rate decreased to 9% 30 days after surgery. Hemodynamic data between groups were significantly different. During maintenance of anesthesia, mean arterial blood pressure was increased, whereas heart rate was decreased in the xenon group compared with the propofol group.

POCD is not only associated with cardiac surgery, but it is also a common problem of noncardiac surgery. However, definition and assessment of POCD are still inconsistent, and the reported incidence ranges between 7% and 71% 7–9 days after surgery[19,20] and between 6% and 56% 42–84 days after surgery.[2,20,21] Independent risk factors for POCD after major surgery have been recently detected, including increasing age or a history of a previous cerebral vascular accident.[3]

Several studies investigated the influence of different anesthetic techniques on the incidence of POCD, *e.g.*, intravenous *versus* inhalational anesthesia or general *versus* regional anesthesia.[22,23] Though Rasmussen *et al.* found no difference in the incidence of cognitive dysfunction after regional *versus* general anesthesia 3 months after surgery, there may be a decreased incidence of POCD early after surgery.[23] In agreement with these results, Papaioannou *et al.* demonstrated a more frequent cognitive impairment in patients who received general anesthesia *versus* regional anesthesia during the first 3 postoperative days.[24] However, there is still no evidence for an advantage of volatile or intravenous anesthetics, regarding intermediate or long-term postoperative cognitive impairment after noncardiac surgery.

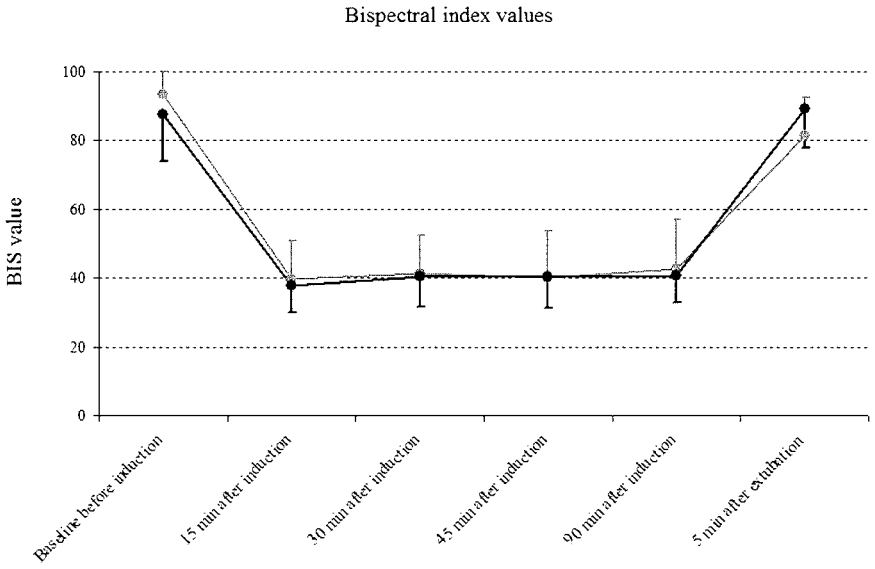

Fig. 1. Bispectral index (BIS) values in patients undergoing anesthesia with either propofol (gray; n = 51) or xenon (black; n = 50). Values are mean (SD).

The noble gas xenon is associated with rapid emergence and early recovery from anesthesia and therefore may be beneficial to prevent neuronal damage and to reduce neurocognitive impairment after surgery. Neuroprotective properties of xenon have been clearly demonstrated in various *in vitro* and *in vivo* studies of neuronal injury and cerebral ischemia. *In vitro*, xenon concentration-dependent reduced neuronal injury induced by *N*-methyl-D-aspartate, glutamate, or oxygen deprivation.[5] *In vivo* studies in rats revealed xenon to decrease *N*-methyl-D-aspartate induced neuronal injury.[25] Moreover, Schmidt *et al.* demonstrated that xenon attenuates cerebral damage after ischemia following cardiac arrest in pigs.[6]

Neuroprotective properties may ideally result in reduced postoperative neurocognitive impairment. Ma *et al.* studied the effect of xenon on cardiopulmonary bypass induced neurologic dysfunction in rats and demonstrated an attenuation of these impairments by xenon.[4] In humans, especially in elder patients with a higher incidence of POCD,[1] so far there is no evidence for an effect of xenon on postoperative neurocognitive impairment. Coburn *et al.* compared emergence and early cognitive function (< 72 h after surgery) in 38 patients aged 65 to 75 yrs after xenon or desflurane anesthesia. In agreement with our results, these authors found no differences between groups regarding POCD. Faster emergence was recognized in the xenon group, but POCD was not evaluated beyond 72 after surgery [17] (table 8).

In our study, we compared xenon with propofol instead of another volatile anesthetic. In different animal models of cerebral ischemia, neuroprotective properties of propofol comparable to xenon were described.[26–28] Furthermore, in a previous trial investigating patients' self-evaluation 4–12 weeks after xenon *versus* propofol anesthesia, no differences in early orientation and the satisfaction with either anesthetic regimen were found.[29] However, self-evaluation (*e.g.*, with questionnaires) has been criticized due to only poor correlation with the results of neuropsychological testing.[15]

Rasmussen *et al.* compared xenon and propofol for supplementary general anesthesia in 35 patients undergoing knee replacement in spinal anesthesia. No difference was detected in cognitive function between the third and fifth days and 3 months postoperatively (table 8). However, the authors admitted that this may be attributed to insufficient sample-size rather than the absence of a true difference.[16] Therefore, to our opinion, propofol was the most appropriate and interesting anesthetic agent for a comparison with xenon.

Table 8. Clinical Studies Evaluating Postoperative Cognitive Function after Xenon Anesthesia—Anesthetic Regime, Demographics, and Assessment of Cognitive Function

Author	Year of Publication	Anesthetic Regime/ Groups	Age of Patients	Patients Scheduled/Patients Finally Included	Intervals for Evaluation	Assessment of Cognitive Function	Significant Difference in Cognitive Function
Coburn et al.[17]	2007	Xenon *vs.* Desflurane/2	65–75 yr	38/38	6–12 h; 66–72 h	TAP	No
Rasmussen et al.[16]	2006	Xenon *vs*. Propofol as Supplement for SPA/2	> 60 yr	39/35	3–5 d; 10–14 weeks	VVL, CST, STROOP, LDC	No

CST = concept shifting task; LDC = letter digit coding task, SPA = spinal anesthesia; STROOP = Stroop colour word interference test; TAP = test for attentional performance; VVL = visual verbal learning.

One problem of studies investigating POCD is the inconsistent definition of this term and its methods of assessment. Therefore, the results of different studies cannot be compared easily. We adopted a definition similar to those previously used in the ISPOCD 1 and 2 studies.[1,2] In these studies, individual changes of each patient compared to baseline pretreatment values were analyzed. Furthermore, a z-score evaluation as a composite score for a comparison with an age-matched control population was added. We were mainly interested in comparing two anesthetic regimes, so we did not include an additional control population in our analysis. However, we calculated standardized change scores, indicating individual changes of patients. In contrast to other studies, our definition of POCD required deterioration of more than 1 SD in 2 or more tests assessing different cognitive domains after a recommended restrictive definition of cognitive dysfunction.[15]

In comparison to previous studies investigating POCD in elder patients, the incidence of POCD in our study 6 days (15%) and 30 days after surgery (9%) was relatively low. Several studies in which a standardized change score (z-score) was applied for patient assessment revealed cognitive decline in 6.8% to 32.7% of patients 7 days after surgery depending on type of surgery, patients' age, and neurocognitive testing.[1,2,19,30] Ancelin *et al.* analyzed POCD in elderly patients 42 days after surgery. Cognitive deterioration was realized in 56% of patients if deterioration in one of 21 scores was regarded and in only 11% of patients if deterioration in four or more scores was regarded.[20] Once more, these differences emphasize the importance of consistent interpretation of data and standardized criteria for cognitive decline.

Several reasons for the relatively low incidence of POCD and the lack of improved neurocognitive performance in the xenon group in our study have to be discussed. First, a low sensitivity of neurocognitive testing may be hypothesized. On the other hand, our assessment revealed a large number of patients who improved neurocognitive performance, indicating that the sensitivity of the test battery was sufficient.

Interestingly, more patients showed improved rather than deteriorated neurocognitive function independent from the anesthetic regimen 30 days after surgery. However, this result is limited by the fact that no additional control group was analyzed to control for learning or time-dependent effects. Nevertheless, it may indicate neurocognitive and possibly also psychological impairment before major surgery. Possibly, neuropsychological baseline testing should have been performed earlier than 1 day before surgery, which was required for administrative and organizational reasons in our study.

Second, our definition of cognitive decline was relatively strict. Furthermore, both xenon and propofol are modern anesthetics that enable rapid emergence and recovery from anesthesia. This may explain a reduced incidence of POCD compared to previous studies evaluating POCD after anesthesia with traditional anesthetics. Finally, a low absolute incidence of POCD, as found in our study, decreases the probability to detect differences between anesthetic groups.

Furthermore, 13 of 114 enrolled patients (11%) were lost for final analysis because of different reasons (table 6). Although these patients were equally allocated to both treatment groups, the missing data had probably influenced our results by an undetermined effect on the observed event rates.

Some limitations of our study must be noted. Motivation of the patients 1 day after surgery was limited, so we decided to perform an incomplete neurocognitive test battery at that time. Therefore, POCD assessing four different cognitive domains was only evaluated 6 days and 30 days after surgery, whereas only three cognitive domains were tested 1 day after surgery.

The design and statistical power of our study as a pilot study did not allow for definitive conclusions regarding a reduction of cognitive dysfunction by xenon. This is emphasized by the increasing 95% CIs for the event rates of POCD from the first postoperative testing 1 day after surgery (95% CI −0.09–0.19) to the second testing after 6 days (95% CI −0.08–0.31) and the final testing after 30 days (95% CI −0.06–0.42), indicating a noticeable uncertainty in the present results and challenging their clinical significance. Nevertheless, the increasing asymmetry of the 95% CIs denotes a trend towards an advantage in the xenon group. On the basis of our data, however, the number of patients needed to treat with xenon to prevent one patient from POCD is quite high (approximately 17). Therefore, further studies with more statistical power are needed to convincingly demonstrate a decreased rate of POCD after xenon anesthesia.

Regarding our results, the assumption of a 20% reduction of POCD in the xenon group was too optimistic. On the other hand, the high costs of a xenon anesthesia in comparison to propofol may be justified only by a clinically relevant reduction of the incidence of POCD. Moreover, previous *in vitro* and *in vivo* studies investigating xenon effects in different models of neuronal injury revealed substantial neuroprotective properties, suggesting potential neuroprotection or neurocognitive benefit in clinical settings, too.[4,5,18]

Finally, comparable BIS values were observed in both groups. However, this did not automatically guarantee a comparable depth of anesthesia and a similar dosing of anesthetics. The BIS algorithm was developed empirically on electroencephalogram data of patients receiving common anesthetics such as propofol or isoflurane; therefore, its validity for xenon is debatable.[31] Investigating how xenon affects BIS in comparison to isoflurane, Goto *et al.* found a lower reliability of BIS as an indicator of depth of anesthesia in xenon patients.[32]

Also, Laitio *et al.* found that BIS and entropy performed well during xenon anesthesia after steady state was achieved, but there were delays in the detection of the actual clinical state during induction and emergence. Interestingly, xenon-induced changes in electroencephalogram closely resemble those induced by propofol.[33]

Nevertheless, in our study BIS was only used as a monitoring parameter and not as a target parameter to titrate anesthetics according to a defined value or range. Therefore, we abstained from permanent adjustments of the inspiratory xenon concentration to avoid a strong increase of xenon consumption associated with higher costs, an unrealistic perspective for future applications. On the other hand, these adjustments were simply not necessary to perform a stabile and consistent xenon anesthesia.

In conclusion, in elderly patients undergoing major noncardiac surgery, we could not demonstrate significant differences in postoperative neurocognitive function 1 day, 6 days, and 30 days after surgery between a xenon-based and a total intravenous anesthetic regimen. Previously demonstrated neuroprotective properties of xenon did not result in a reduced incidence of POCD in comparison to propofol. A trial with more statistical power is needed to definitively answer the question regarding neuroprotective properties of xenon in daily clinical practice.

The authors thank Britta Raitschew, M.S., Medical Student, Benita Barnewitz, M.S., Medical Student, Birgit Zastrow, R.N., and Christian Casadio, IT-Technician (Department of Anaesthesiology and Intensive Care Medicine, University Hospital Schleswig-Holstein, Campus Kiel, Germany) for their enthusiastic support.

References

1. Moller JT, Cluitmans P, Rasmussen LS, Houx P, Rasmussen H, Canet J, Rabbitt P, Jolles J, Larsen K, Hanning CD, Langeron O, Johnson T, Lauven PM, Kristensen PA, Biedler A, van Beem H, Fraidakis O, Silverstein JH, Beneken JE, Gravenstein JS: Long-term postoperative cognitive dysfunction in the elderly ISPOCD1 study. ISPOCD investigators. International Study of Post-Operative Cognitive Dysfunction. Lancet 1998; 351:857–61

2. Johnson T, Monk T, Rasmussen LS, Abildstrom H, Houx P, Korttila K, Kuipers HM, Hanning CD, Siersma VD, Kristensen D, Canet J, Ibanaz MT, Moller JT: Postoperative cognitive dysfunction in middle-aged patients. ANESTHESIOLOGY 2002; 96:1351–7

3. Monk TG, Weldon BC, Garvan CW, Dede DE, van der Aa MT, Heilman KM, Gravenstein JS: Predictors of cognitive dysfunction after major noncardiac surgery. ANESTHESIOLOGY 2008; 108:18–30

4. Ma D, Yang H, Lynch J, Franks NP, Maze M, Grocott HP: Xenon attenuates cardiopulmonary bypass-induced neurologic and neurocognitive dysfunction in the rat. ANESTHESIOLOGY 2003; 98:690–8

5. Wilhelm S, Ma D, Maze M, Franks NP: Effects of xenon on *in vitro* and *in vivo* models of neuronal injury. ANESTHESIOLOGY 2002; 96:1485–91

6. Schmidt M, Marx T, Gloggl E, Reinelt H, Schirmer U: Xenon attenuates cerebral damage after ischemia in pigs. ANESTHESIOLOGY 2005; 102:929–36

7. Franks NP, Dickinson R, de Sousa SL, Hall AC, Lieb WR: How does xenon produce anaesthesia? Nature 1998; 396:324

8. Choi DW, Koh JY, Peters S: Pharmacology of glutamate neurotoxicity in cortical cell culture: Attenuation by NMDA antagonists. J Neurosci 1988; 8:185–96

9. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J: An inventory for measuring depression. Arch Gen Psychiatry 1961; 4:561–71

10. Lezak MD: Neuropsychological assessment. New York, Oxford Press, 1995

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

12. Bland JM, Altman DG: The use of transformation when comparing two means. BMJ 1996; 312:1153

13. Smith PL, Treasure T, Newman SP, Joseph P, Ell PJ, Schneidau A, Harrison MJ: Cerebral consequences of cardiopulmonary bypass. Lancet 1986; 1:823–5

14. Crawley F, Stygall J, Lunn S, Harrison M, Brown MM, Newman S: Comparison of microembolism detected by transcranial Doppler and neuropsychological sequelae of carotid surgery and percutaneous transluminal angioplasty. Stroke 2000; 31:1329–34

15. Rasmussen LS, Larsen K, Houx P, Skovgaard LT, Hanning CD, Moller JT: The assessment of postoperative cognitive function. Acta Anaesthesiol Scand 2001; 45:275–89

16. Rasmussen LS, Schmehl W, Jakobsson J: Comparison of xenon with propofol for supplementary general anaesthesia for knee replacement: A randomized study. Br J Anaesth 2006; 97:154–9

17. Coburn M, Baumert JH, Roertgen D, Tiedel V, Fries M, Hein M, Kunitz O, Fimm B, Rossaint R: Emergence and early cognitive function in the elderly after xenon or desflurane anaesthesia: A double-blinded randomized controlled trial. Br J Anaesth 2007; 98:756–62

18. Homi HM, Yokoo N, Ma D, Warner DS, Franks NP, Maze M, Grocott HP: The neuroprotective effect of xenon administration during transient middle cerebral artery occlusion in mice. ANESTHESIOLOGY 2003; 99:876–81

19. Canet J, Raeder J, Rasmussen LS, Enlund M, Kuipers HM, Hanning CD, Jolles J, Korttila K, Siersma VD, Dodds C, Abildstrom H, Sneyd JR, Vila P, Johnson T, Munoz Corsini L, Silverstein JH, Nielsen IK, Moller JT: Cognitive dysfunction after minor surgery in the elderly. Acta Anaesthesiol Scand 2003; 47:1204–10

20. Ancelin ML, de Roquefeuil G, Ledesert B, Bonnel F, Cheminal JC, Ritchie K: Exposure to anaesthetic agents, cognitive functioning and depressive symptomatology in the elderly. Br J Psychiatry 2001; 178:360–6

21. Johom G, Szarvas S, Larney V, O'Brien J, Buckley E, Butler M, Shorten G: Perioperative plasma concentrations of stable nitric oxide products are predictive of cognitive dysfunction after laparoscopic cholecystectomy. Anesth Analg 2004; 99:1251–52

22. Enlund M, Mentell O, Flenninger A, Horneman G, Ronquist G: Evidence of cerebral dysfunction associated with isoflurane- or propofol-based anaesthesia for orthognathic surgery, as assessed by biochemical and neuropsychological methods. Ups J Med Sci 1998; 103:43–59

23. Rasmussen LS, Johnson T, Kuipers HM, Kristensen D, Siersma VD, Vila P, Jolles J, Papaioannou A, Abildstrom H, Silverstein JH, Bonal JA, Raeder J, Nielsen IK, Korttila K, Munoz L, Dodds C, Hanning CD, Moller JT: Does anaesthesia cause postoperative cognitive dysfunction? A randomised study of regional *versus* general anaesthesia in 438 elderly patients. Acta Anaesthesiol Scand 2003; 47: 260–6

24. Papaioannou A, Fraidakis O, Michaloudis D, Balalis C, Askitopoulou H: The impact of the type of anaesthesia on cognitive status and delirium during the first postoperative days in elderly patients. Eur J Anaesthesiol 2005; 22:492–9

25. Ma D, Wilhelm S, Maze M, Franks NP: Neuroprotective and neurotoxic properties of the 'inert' gas, xenon. Br J Anaesth 2002; 89:739–46

26. Gelb AW, Bayona NA, Wilson JX, Cechetto DF: Propofol anesthesia compared to awake reduces infarct size in rats. ANESTHESIOLOGY 2002; 96:1183–90

27. Bayona NA, Gelb AW, Jiang Z, Wilson JX, Urquhart BL, Cechetto DF: Propofol neuroprotection in cerebral ischemia and its effects on low-molecular-weight antioxidants and skilled motor tasks. ANESTHESIOLOGY 2004; 100:1151–9

28. Engelhard K, Werner C, Eberspacher E, Pape M, Stegemann U, Kellermann K, Hollweck R, Hutzler P, Kochs E: Influence of propofol on neuronal damage and apoptotic factors after incomplete cerebral ischemia and reperfusion in rats: A long-term observation. ANESTHESIOLOGY 2004; 101:912–7

29. Coburn M, Kunitz O, Baumert JH, Hecker K, Rossaint R: Patients' self-evaluation after 4–12 weeks following xenon or propofol anaesthesia: A comparison. Eur J Anaesthesiol 2005; 22:870–4

30. Rasmussen LS, Moller JT: Central nervous system dysfunction after anesthesia in the geriatric patient. Anesthesiol Clin North America 2000; 18:59–70

31. Rampil IJ: A primer for EEG signal processing in anesthesia. ANESTHESIOLOGY 1998; 89:980–1002

32. Goto T, Nakata Y, Saito H, Ishiguro Y, Niimi Y, Suwa K, Morita S: Bispectral analysis of the electroencephalogram does not predict responsiveness to verbal command in patients emerging from xenon anaesthesia. Br J Anaesth 2000; 85:359–63

33. Laitio RM, Kaskinoro K, Sarkela MO, Kaisti KK, Salmi E, Maksimow A, Langsjo JW, Aantaa R, Kangas K, Jaaskelainen S, Scheinin H: Bispectral index, entropy, and quantitative electroencephalogram during single-agent xenon anesthesia. ANESTHESIOLOGY 2008; 108:63–70