J Anesth (2010) 24:518–525
DOI 10.1007/s00540-010-0952-z

ORIGINAL ARTICLE

Postoperative cognitive dysfunction: computerized and conventional tests showed only moderate inter-rater reliability

Finn M. Radtke · Martin Franck · Norbert Papkalla · Tim S. Herbig · Edith Weiss-Gerlach · Robin Kleinwaechter · Klaus D. Wernecke · Claudia D. Spies

Received: 9 December 2009 / Accepted: 7 April 2010 / Published online: 28 May 2010

Abstract
Purpose The incidence of postoperative cognitive dysfunction (POCD) depends on the test battery and calculation method used. The measurements may be performed with a paper and pencil test battery or with a computerized test battery. The objective of this study was to measure the incidence and congruence of POCD by comparing a computerized test battery with a paper and pencil test battery in the same patient population.
Methods In total, 67 patients were included: 30 consecutive in-patients undergoing surgery under general anesthesia and 37 non-surgical out-patients as control. Patients were tested with a paper and pencil test battery and a computerized test battery on inclusion in the study and 7 days later. Both test batteries covered the cognitive domains: visual attention, visual learning, memory, and speed of processing.
Results The computerized test battery classified 10.0% (95% CI 3.5–25.6%) of the patients as suffering from POCD whereas the paper and pencil test battery classified 30.0% (95% CI 16.7–47.9%) as suffering (95% CI for difference 3.9–36.5%, $P = 0.03$). The inter-rater reliability between both test batteries showed moderate agreement (Cohen's kappa of 0.41). All patients identified by use of the computerized test battery were also identified with the paper and pencil test battery. The paper and pencil test battery identified 6 additional cases.
Conclusion In our study we demonstrated that the incidence of POCD measured with computerized test battery and paper and pencil test battery showed moderate inter-rater reliability. Use of neuropsychological test batteries theoretically covering the same cognitive domains does not automatically lead to the same classification of POCD.

Keywords Cognitive deficit · Complication · Recovery · Neurological · Postoperative

F. M. Radtke and M. Franck contributed equally.

F. M. Radtke · M. Franck · N. Papkalla · T. S. Herbig · E. Weiss-Gerlach · R. Kleinwaechter · C. D. Spies (✉)
Department of Anesthesiology and Surgical Critical Care Medicine, Campus Charité Mitte and Campus Virchow-Klinikum, Charité, Universitätsmedizin, Charitéplatz 1, 10117 Berlin, Germany
e-mail: claudia.spies@charite.de

K. D. Wernecke
SOSTANA, Wildensteiner Straße 27, 10318 Berlin, Germany

Introduction

Postoperative cognitive dysfunction (POCD) is associated with higher mortality [1]. Symptoms may range from a mild cognitive deterioration that might not even be recognized by the patients themselves [2] to longer-lasting more apparent dysfunction of memory, information processing, and perception. Subjective self-reported cognitive symptoms do not substitute for objective cognitive testing, because they correlate poorly [3].

Large studies have reported an incidence of approximately 25% 1 week after surgery and approximately 10% after 3 months [2, 3]. It may occur in all age groups but elderly patients seem to be at increased risk [2, 4].

For measurement of POCD different aspects of cognition (advertence, concentration, verbal abilities, learning, and memory etc.) need to be assessed. Specific requirements for test selection have been recommended [5] in order to introduce uniformity into the assessment and definition of POCD.

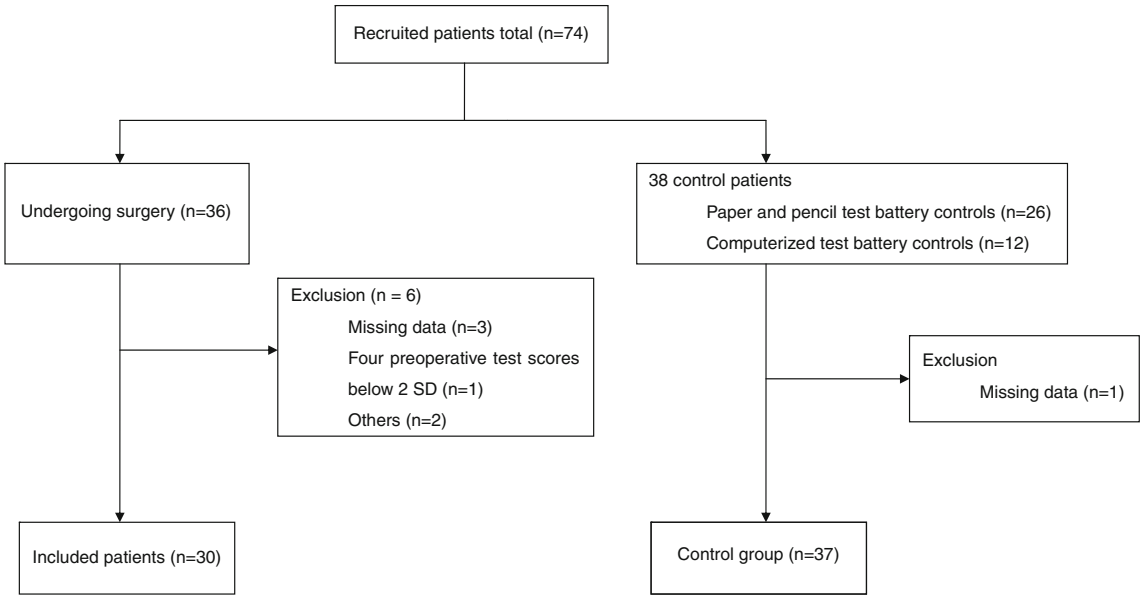

Fig. 1 Inclusion tree

Paper and pencil test batteries or computerized test batteries may be used for assessment of POCD. A computerized test battery enables more standardized administration, greater reliability, increased accuracy, improved data handling [6], and also enables random stimulus generation with nearly an unlimited number of alternative forms of test [7]. Additionally the test and retest sessions are less dependant on the test instructor. Therefore, test administration might be facilitated and large patient population might be tested at lower cost [8].

The objective of this study was to measure the incidences and congruence of POCD by comparing a computerized test battery with a paper and pencil test battery in the same patient population.

Methods

Patients

This ethically approved cohort study included 36 adult ASA PS III consecutive in-patients undergoing surgery and, as control group, 38 patients with severe systemic disease consistent with ASA PS III. Inclusion criteria were a mini mental state examination score of more than 23 points, general anesthesia and expected anesthesia duration of at least 180 min. Exclusion criteria were central nervous system disease, drug or alcohol abuse, severe visual or auditory handicap, or cardiac or neurosurgery. In addition patients with poor comprehension of the German language, illiteracy, or unwillingness to follow the study protocol were also excluded. 36 patients completed the paper and pencil test battery and the computerized test battery. Of these 36 patients 6 had to be excluded because of missing data, refusal to continue, or other reasons. Of the 38 control patients, one patient had to be excluded because of refusal to complete the second test (Fig. 1). Participants who performed two standard deviations (SDs) below the average in the preoperative test session in more than four tests were excluded because we assumed they were unable or unwilling to complete the given tasks with the necessary care.

Because comorbidities may trigger the incidence of POCD, particularly if they coincide with major surgery or long duration of surgery [3, 9], we chose ASA PS 3 patients for our study, to ensure a high incidence of POCD.

No restrictions were made regarding type or conduct of general anesthesia or postoperative care. Anesthesia was administered in accordance with the standard operating procedures of our department, published by Kox and colleagues [10]. General anesthesia was induced with thiopental, propofol, etomidate, or midazolam in combination with fentanyl or remifentanil, followed by neuromuscular block to facilitate endotracheal intubation. Anesthesia was maintained by total intravenous anesthesia using propofol or inhalation anesthetics using one of either desflurane, isoflurane, or sevoflurane. Patients received nitrous oxide at the discretion of the individual anesthesiologist in charge. The anesthesiologist was free to use opioid analgesics and muscle relaxants as needed.

Patients included were scheduled primarily for abdominal surgery with an expected postoperative length of stay of 7 days.

Table 1 Cognitive areas tested

Battery	Name of test	Ability tested	Measurements
Computerized	Detection Task	Speed of processing	Speed and accuracy
Computerized	Identification Task	Visual attention	Speed and accuracy
Computerized	One Card Learning	Visual learning, memory	Speed and accuracy
Paper and Pencil	Visual Verbal Learning	Visual learning, memory	Correct and false answers
Paper and Pencil	Stroop Color Word	Speed of processing, visual attention, cognitive flexibility	Correct answers and speed

Test selection

As suggested by Moller and colleagues [2] the test administration should not take more than 45 min per session as a longer test duration might result in a decrease of compliance and concentration. The computerized test battery and the paper pencil test battery required on average 20 min each. In order to measure the same aspects of cognitive function, each test battery should cover similar cognitive domains. Both test batteries contain cognitive tests covering the domains: visual attention, visual learning, memory, and speed of processing (Table 1).

Computerized test battery

The computerized test battery (developed by CogState, Melbourne, Australia) contained three cognitive tests. The tests administered were the Detection Task, the Identification Task, and the One Card Learning Task covering different areas of cognition (Table 1). Each test consisted of a card set with different tasks to complete. Each task was explained in detail to the participant. Preceding each test a trial period had to be completed so that each individual was made familiar with the upcoming task. Supervisors were allowed to correct participants during the trial period. The tests did not require more than two keys on the keyboard. The subject was asked to fulfill the tasks as rapidly and as accurately as possible. For each test average speed and accuracy, as the ratio of correct items to total items, were recorded.

Paper pencil test battery

The paper pencil test battery consisted of the Stroop Color Word Test [11] parts one to three and parallel versions of the Visual Verbal Learning Test [12] parts one to three including a delayed recall (part four). All the tests furnished excellent results in previous studies of POCD [2]. The participants were asked to complete the given task as rapidly and as accurately as possible. From the Visual Verbal Learning Test the number of correct items and the number of wrong items in the delayed recall, i.e. part four, were taken for POCD calculation. From the Stroop Color Word Test the time and the number of wrong items in the third part were taken for POCD calculation.

Test administration

The tests were carried out by two trained research assistants supervised by a psychologist. Written test instructions were used to guarantee standardized test execution. All patients were tested on the evening before surgery and 7 days after surgery or as soon as possible thereafter. To avoid disturbance the tests were administered in a separate room. First, parts I to III of the Visual Verbal Learning Test and the Stroop Color Word Test were administered. The patients were then asked to complete the computerized test battery before the delayed recall (part IV) of the Visual Verbal Learning Test.

POCD calculation

The reliable change index is one of the most reliable and valid methods for measuring POCD in individuals [13–15]. For POCD calculation the reliable change index with correction of the practice effect, as described by Rasmussen and colleagues [16], was used. For each test result a Z score (Eq. 1) was calculated. First the difference in test performance (Δx) between the preoperative and postoperative session was calculated for each individual patient. From this change the mean performance change of the control group ($\overline{\Delta x}_{\text{control}}$), representing the practice effect, was subtracted. The result is divided by the SD of the control group change ($\Delta x_{\text{control}}$).

$$Z=\frac{\Delta x-\overline{\Delta x}_{\text{control}}}{\text{SD}\ (\Delta x_{\text{control}})} \quad (1)$$

The composite Z score was calculated by dividing the sum of the Z scores of all tests of an individual by the SD of the sum of the Z scores of the control group.

All values were recorded so that test measurements indicating deterioration resulted in a negative Z score. Z scores for the paper and pencil version were derived from the Stroop Color Word Test (time and error) and from the

Table 2 Baseline characteristics of patients and controls

	Patient group ($n = 30$)	Control group ($n = 37$)	P value
Age (years)	66 (58–73)	71 (63–80)	0.02
Body mass index (kg/m^2)	27 (24–29)	28 (23–32)	0.48
Charlson comorbidity score	4.0 (3.0–8.0)	4.0 (2.0–5.0)	0.15
Gender (female fraction)	0.30	0.62	0.01

Data are median (25–75% percentile); the P value is from Fisher's exact test (for proportions) or from the Mann–Whitney U test (for continuous variables)

Table 3 Change in test performance for the paper and pencil test battery

	Patient group ($n = 30$)	Control group ($n = 26$)
Stroop Color Word duration (s)	−2.91 (−13.01–4.76)	3.24 (0.50–8.87)
Stroop Color Word errors	0.00 (0.00–0.00)	0.50 (0.00–2.00)
Visual Verbal Learning errors	0.00 (−1.00–0.00)	0.00 (0.00–0.00)
Visual Verbal Learning correct	0.00 (−1.25–2.25)	1.00 (−1.00–1.25)
Mini-mental state examination	30.00 (29.00–30.00)	29.00 (28.00–30.00)

Data are median (25–75% percentile)

Visual Verbal Learning Test (correct and false items). For the computerized test we used the Z scores derived from the Detection Task, the Identification Task, and the One Card Learning Task (speed and accuracy). For both test batteries we calculated composite Z scores.

Patients were regarded as suffering from POCD either when the composite Z score was lower than −1.96 or when at least two test endpoints had a Z score of less than −1.96. In that way individuals with general cognitive decline and patients with cognitive decline in single fields were detected.

The primary outcome of this study was POCD defined by the paper and pencil test battery and the computerized test battery after 7 days. We recorded the patient's pre and postoperative medical history and medication. All data were checked for consistency and plausibility.

All data are reported as medians with 25–75% range and proportions with 95%-confidence intervals. Proportions of independent samples were compared using Fisher's exact test. For continuous variables differences between independent groups were assessed using the Mann–Whitney U test. A P value <0.05 was considered as being significant. For statistical analysis we employed SPSS for Windows Version 15. Descriptive statistics were computed for all study variables.

Results

The median age for the 30 (9 female, 21 male) surgical patients was 66 years (25–75%-percentiles: 57.75–72.75). The average age for the 37 control group members (23 female, 14 male) was 71 years (25–75%-percentiles: 63.00–80.00). The control group was significantly older than the patient group ($P = 0.02$). Also gender differed significantly between these two groups ($P = 0.01$). Other baseline characteristics did not differ significantly between the control group and the patient group (Table 2).

In total, 7 patients (drop out quote 9.5%) had to be excluded from analysis (Fig. 1). Changes in test performance for the paper and pencil test battery and computerized test battery are shown in Tables 3 and 4. Patient characteristics for both test batteries are given in Tables 5 and 6.

The postoperative test was performed at a median of 7 days after the preoperative test (25–75%-percentiles 7.0–8.0 days).

9 of 30 patients [30.0% (95% CI 16.7–47.9%)] were scored positive for POCD according to the paper and pencil test battery and 3 out of 30 patients [10.0% (95% CI 3.5–25.6%)] according to the computerized test battery (95% CI for difference 3.9–36.5%, McNemar's test: $P = 0.03$) (Table 7).

Of the three patients defined as suffering from POCD by the computerized test battery two patients had two or more Z scores lower than −1.96. One patient had a composite Z score lower than −1.96. Of the nine patients defined as suffering from POCD by the paper and pencil test battery five patients had two or more Z scores lower than −1.96. Four patients also had a composite Z score lower than −1.96. The three POCD positive patients classified by the computerized test battery all had a Z score of less than −3

Table 4 Change in test performance for the computerized test battery

	Patient group ($n = 30$)	Control group ($n = 11$)
Speed Detection Task (ms)	−20.40 (−100.63–32.41)	−8.65 (−76.25–55.71)
Speed Identification Task (ms)	−35.44 (−66.95–21.20)	0.00 (−61.15–39.30)
Speed One-Card-Learning Task (ms)	103.28 (−33.59–469.15)	93.48 (−210.01–379.99)
Accuracy Detection Task	0.00 (−0.06–0.10)	0.05 (0.00–0.06)
Accuracy Identification Task	0.01 (−0.06–0.22)	0.06 (−0.06–0.10)
Accuracy One-Card-Learning Task	−0.08 (−0.22–0.03)	−0.01 (−0.17–0.04)
Mini-mental state examination	30.00 (29.00–30.00)	29.00 (27.00–30.00)

Data are median (25–75% percentile)

Table 5 Patient comparison classification by the computerized test battery

	NPOCD ($n = 27$)	POCD ($n = 3$)	P value
Age (years)	66.00 (58.00–72.00)	71.00 (47.00–76.00)	0.795
Body mass index (kg/m^2)	27.43 (23.03–29.00)	27.34 (24.44 35.26)	0.744
Mini-mental state examination	30.00 (29.00–30.00)	29.00 (25.00–30.00)	0.350
Test period (days)	8.00 (7.00–10.00)	8.00 (7.00–11.00)	0.795
Surgery duration (min)	223.00 (152.00–286.00)	296.00 (192.00–352.00)	0.315
Anesthesia duration (min)	320.00 (225.00–400.00)	435.00 (275.00–435.00)	0.226
Number of surgeries total	1.00 (1.00–2.00)	1.00 (1.00–1.00)	0.509
Length of stay (days)	17.00 (11.00–24.00)	20.00 (13.00–28.00)	0.554
Length of stay postoperative (days)	12.00 (9.00–17.00)	17.00 (10.00–27.00)	0.387
Intensive care unit (days)	1.00 (0.00–2.00)	1.00 (1.00–2.00)	0.845
Charlson comorbidity score	4.00 (3.00–8.00)	8.00 (5.00–9.00)	0.226

Data are median (25–75% percentile); the P value is from Fisher's exact test (for proportions) or from the Mann–Whitney U test (for continuous variables)

NPOCD non postoperative cognitive dysfunction, *POCD* postoperative cognitive dysfunction

in the paper and pencil test battery. Patients with POCD according to the paper and pencil test battery only, by failing in two single tests, were not positive in the computerized test battery.

Inter-rater agreement for the two test batteries was assessed with Cohen's kappa coefficient. The classifications showed the following accordance: of the 21 patients without POCD according to the paper and pencil test battery all were also negative according to the computerized test battery. 6 patients were classified differently by the test batteries. Of the 9 patients identified as suffering from cognitive decline by the paper and pencil test battery three were also identified by the computerized test battery. Cohen's kappa showed moderate agreement (0.41) according to the definition from Landis and Koch [17] (Table 7).

Median duration of surgery and anesthesia, and postoperative length of stay, were longer for patients classified as having POCD by the computerized test battery than for patients classified as having POCD by the paper and pencil test battery only (Table 8).

Discussion

The most important result of this study was that POCD classification by the computerized test battery and by the paper and pencil test battery showed moderate inter-rater agreement. The computerized test battery showed an incidence of 10% whereas the paper and pencil test battery showed an incidence of 30%. The incidences measured by use of the two test batteries are in accordance with previous publications ranging from 14.4% (6th day) in the study of Linstedt et al. [18] up to 40% in the study of Iohom et al. [19]. Abildstrom et al. [20] and Rasmussen et al. [21, 22] showed POCD incidences of about 25% in their studies. Linstedt et al. investigated 120 surgical patients whereas Iohom et al. investigated 42 inpatients undergoing

Table 6 Patient comparison classification by the paper and pencil test battery

	NPOCD ($n = 21$)	POCD ($n = 9$)	P value
Age (years)	66.00 (58.00–75.00)	67.00 (54.50–71.50)	0.824
Body mass index (kg/m^2)	27.33 (23.80–29.74)	27.42 (23.38–29.46)	0.965
Mini-mental state examination	30.00 (29.00–30.00)	29.00 (28.50–30.00)	0.137
Test period (days)	8.00 (7.00–10.00)	8.00 (7.00–10.50)	0.929
Surgery duration (min)	223.00 (152.00–310.50)	249.00 (185.50–285.50)	1.000
Anesthesia duration (min)	320.00 (240.00–392.50)	320.00 (237.50–432.50)	0.722
Number of surgeries total	1.00 (1.00–2.50)	1.00 (1.00–1.50)	0.476
Length of stay (days)	19.00 (12.00–25.00)	14.00 (10.50–22.50)	0.504
Length of stay postoperative (days)	13.00 (9.50–22.00)	12.00 (9.50–15.50)	0.533
Intensive care unit (days)	1.00 (0.00–2.00)	1.00 (0.50–2.50)	0.657
Charlson comorbidity score	3.00 (2.50–8.00)	8.00 (5.50–8.50)	0.086

Data are median (25–75% percentile); the P value is from Fisher's exact test (for proportions) or from the Mann–Whitney U test (for continuous variables)

NPOCD non postoperative cognitive dysfunction, *POCD* postoperative cognitive dysfunction

Table 7 Calculation of Cohen's kappa

	Computerized test battery		Total
	NPOCD[a]	POCD[b]	
Paper and pencil test battery			
NPOCD[a]	21	0	21
POCD[b]	6	3	9
Total	27	3	30

Cohen's kappa: 0.41 (McNemar's test: $P = 0.03$)

[a] Patient classified as not suffering from postoperative cognitive dysfunction

[b] Patient classified as suffering from postoperative cognitive dysfunction

laparoscopic cholecystectomy. Rasmussen and colleagues investigated patients under general anesthesia; in both studies they used the Visual Verbal Learning Test, the Concept Shifting Test, the Stroop Color Word Test, and the Letter-Digit-Coding Task for cognitive testing. Abildstrom et al. included patients scheduled for major surgery under general anesthesia using the same tests as used by Rasmussen et al. Our paper pencil test battery was similar to the tests used by Rasmussen and Abildstrom, with similar incidence.

The classification of POCD with the paper and pencil test battery differed significantly from the classification with the computerized test battery ($P = 0.03$). These findings reveal the difficulty of comparing results from two test batteries even if they cover the same cognitive domains and the same method of calculation is used. In a study by Silbert and colleagues [6], also comparing a computerized test battery with a paper and pencil test battery, higher incidence of POCD in patients undergoing cardiac surgery was shown (paper and pencil test battery 32%, computerized test battery 42%) with higher rates of agreement between the two batteries ($\kappa = 0.79$; $P = 0.001$). It is known that patients undergoing cardiac surgery are at increased risk of suffering from postoperative cognitive decline. In this study of Silbert and colleagues [6] all patients identified with POCD with the computerized test battery were also classified with the paper and pencil test battery, plus five additional cases. In our study the paper and pencil test battery identified six additional cases as having POCD. One explanation of these findings may be that the paper and pencil test battery defined patients with less cognitive decline as having POCD, whereas the computerized test battery may have only identified patients with a more severe form. These findings suggest that POCD is not a dichotomous variable but rather a continuous variable with degrees of severity. In two multi-center studies [2, 3] in which a paper and pencil test battery similar to our test battery was used, incidences of POCD in comparable patient populations differed significantly. The test center proved to be the most significant independent predictor of POCD. In the absence of a gold standard we are not able to determine which method is more accurate in defining POCD. However with regard to the methodology, more standardization may leave less room for human error. Computerized testing might possibly be advantageous with regard to recording and capturing of reaction times, and with regard to processing of data by minimizing human error.

Conclusion

Incidence of POCD measured with a computerized test battery and a paper and pencil test battery covering the same cognitive domains showed a moderate inter-rater

Table 8 Patient comparison between POCD by paper and pencil test battery only and POCD by computerized test battery

	$n = 6$[a]	$n = 3$[b]	P value
Age (years)	65.50 (56.25–70.50)	71.00 (47.00–76.00)	0.714
Body mass index (kg/m^2)	27.71 (21.87–28.77)	27.34 (24.44–35.26)	0.905
Mini-mental state examination	29.00 (28.75–30.00)	29.00 (25.00–30.00)	0.714
Test period (days)	7.50 (7.00–10.25)	8.00 (7.00–11.00)	0.714
Surgery duration (min)	220.00 (163.75–271.25)	296.00 (192.00–352.00)	0.167
Anesthesia duration (min)	287.50 (217.50–422.50)	435.00 (275.00–435.00)	0.167
Number of surgeries total	1.00 (0.75–2.25)	1.00 (1.00–1.00)	0.714
Length of stay (days)	12.50 (9.75–21.25)	20.00 (13.00–28.00)	0.262
Length of stay postoperative (days)	11.00 (8.50–12.50)	17.00 (10.00–27.00)	0.167
Intensive care unit (days)	1.50 (0.00–3.50)	1.00 (1.00–2.00)	0.905
Charlson comorbidity score	8.00 (4.50–8.75)	8.00 (5.00–9.00)	0.905

Data are median (25–75% percentile); the P value was from Fisher's exact test (for proportions) or from the Mann–Whitney U test (for continuous variables)

[a] Patients classified as having POCD by the paper and pencil test battery and not by the computerized test battery

[b] Patients classified as having POCD by computerized test battery

reliability. The use of neuropsychological test batteries theoretically covering the same cognitive domains do not automatically lead to the same classification of POCD. The computerized test battery measured a subset of the paper and pencil test battery positive patients, possibly including patients with more severe cognitive dysfunction.

Acknowledgments Financial supported by grants from the Charité-Universitätsmedizin Berlin, Germany.

References

1. 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.
2. 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.
3. 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.
4. Newman S, Stygall J, Hirani S, Shaefi S, Maze M. Postoperative cognitive dysfunction after noncardiac surgery—a systematic review. Anesthesiology. 2007;106:572–90.
5. Murkin JM, Newman SP, Stump DA, Blumenthal JA. Statement of consensus on assessment of neurobehavioral outcomes after cardiac surgery. Ann Thorac Surg. 1995;59:1289–95.
6. Silbert BS, Maruff P, Evered LA, Scott DA, Kalpokas M, Martin KJ, Lewis MS, Myles PS. Detection of cognitive decline after coronary surgery. A comparison of computerized and conventional tests. Br J Anaesth. 2004;92:814–20.
7. Collie A, Darby D, Maruff P. Computerised cognitive assessment of athletes with sports related head injury. Br J Sports Med. 2001;35:297–302.
8. Wild K, Howieson D, Webbe F, Seelyea A, Kayea J. Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement. 2008;4(6):428–37.
9. Müller SV, Krause N, Schmidt M, Münte TF. Cognitive dysfunction after abdominal surgery in elderly patients. Z Gerontol Geriatr. 2004;37(6):475–85.
10. Kox WJ, Spies C. Check-up Anästhesiologie Standards Anästhesie-Intensivmedizin-Schmerztherapie-Notfallmedizin. 2nd ed. New York: Springer; 2005.
11. Bohnen N, Twijnstra A, Jolles J. Performance in the Stroop color word test in relationship to the persistence of symptoms following mild head injury. Acta Neurol Scand. 1992;2:116–21.
12. Brand N, Jolles J. Learning and retrieval rate of words presented auditorily and visually. J Gen Psychol. 1985;112(2):201–10.
13. Collie A, Falleti MG, Silbert B, Maruff P. Determining the extent of cognitive change after coronary surgery: a review of statistical procedures. Ann Thorac Surg. 2002;73:2005–11.
14. Lewis MS, Maruff P, Silbert BS, Evered LA, Scott DA. The influence of different error estimates in the detection of post-operative cognitive dysfunction using reliable change indices with correction for practice effects. Arch Clin Neuropsychol. 2006;21:421–7.
15. Lewis MS, Maruff P, Silbert BS, Evered LA, Scott DA. The sensitivity and specificity of three common statistical rules for the classification of post-operative cognitive dysfunction following coronary artery bypass graft surgery. Acta Anaesthesiol Scand. 2006;50:50–7.
16. 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.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–74.
18. Linstedt U, Meyer O, Kropp P, Berkau A, Tapp E, Zenz M. Serum concentration of S-100 protein in assessment of cognitive dysfunction after general anesthesia in different types of surgery. Acta Anaesthesiol Scand. 2002;46:384–9.
19. Iohom 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;99:1245–52.
20. Abildstrom H, Rasmussen LS, Rentowl P, Hanning CD, Rasmussen H, Kristensen PA, Moller JT. Cognitive dysfunction 1–2 years after non-cardiac surgery in the elderly. Acta Anaesthesiol Scand. 2000;44:1246–51.
21. 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.
22. Rasmussen LS, O'Brien JT, Silverstein JH, Johnson TW, Siersma VD, Canet J, Jolles J, Hanning CD, Kuipers HM, Abildstrom H, Papaioannou A, Raeder J, Yli-Hankala A, Sneyd JR, Munoz L, Moller JT. Is peri-operative cortisol secretion related to post-operative cognitive dysfunction? Acta Anaesthesiol Scand. 2005;49:1225–31.