Tools to screen and measure cognitive impairment after surgery and anesthesia

Sérgio Vide 1,2,3, Pedro L. Gambús 1,4,5

Summary

Cognition is essential to all aspects of our everyday life. Although we take our cognitive function for granted, the perioperative period is prone to several aggressions that might impair it. Postoperative cognitive dysfunction, has been the aim of many studies recently, and was shown to be very common with an incidence that can reach 40%, yielding not only impairment in cognition, but also longer hospital stays, higher costs and greater mortality. While several studies have revealed some of the mechanisms contributing to postoperative cognitive dysfunction, the search for the perfect instrument to screen and measure cognitive (dys)function has proven more elusive. The present paper aims to review several cognitive evaluation methods, discussing their advantages and disadvantages as well as their potential clinical applications in evaluating the dynamics of the recovery of cognitive function after anesthesia and surgery. The current availability of easy to use computerized tests might provide the tools necessary to identify patients at risk, and promptly provide them with the adequate course of action.

Postoperative cognitive dysfunction

Cognition requires several cortical and subcortical functions, such as perception, memory and information processing, which allows the individual to acquire knowledge, solve problems, and plan for the future [1]. These functions are essential for everyday living, and cognitive dysfunction appears when there is some form of impairment in these functional capacities.

1. Hospital CLINIC de Barcelona, Anesthesiology Department, Systems Pharmacology Effect Control & Modeling (SPEC-M) Research Group, Barcelona, Spain
2. Hospital Pedro Hispano, Department of Anesthesia, Matosinhos, Portugal
3. Centro Hospitalar Universitário do Porto, Department of Anesthesiology, Center for Clinical Research in Anesthesia, Porto, Portugal
4. University of California San Francisco (UCSF), Department of Anesthesia and Perioperative Care, San Francisco, California, USA
5. Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Neuroimmunology Research Group, Barcelona, Spain

Correspondence:
Pedro Gambús, Department of Anesthesiology, Hospital CLINIC de Barcelona, Villarroel 170, 08036 Barcelona, Spain.
p.lgambus@clinic.cat
One aspect of cognition in anesthesia is the strong effects of anesthetic drugs on cognitive functions which might affect normal recovery after general anesthesia or sedation. The complexity of evaluation of cognitive function has made it that it has not been until recently included as a relevant criteria to consider that a patient was fit to discharge. Given the increasing amount of procedures being performed under outpatient conditions, being able to evaluate the patient from a cognitive perspective a patient is fit for discharge acquires a relevant importance, at least as much importance as other criteria like absence of uncontrollable pain, hemodynamic stability, absence of nausea or tolerance to diet.

From a different, longer term view, changes in cognition after surgery have been reported since the beginning of anesthesia [2], and are currently known as postoperative cognitive dysfunction (POCD). Early POCD is common perioperatively (30–40% incidence) yielding longer hospital stays, higher costs and greater mortality [3]. This entity does not come as a surprise, as the perioperative period is prone to several aggressions (table 1) (stress, inflammation, pain, medications, anesthesia) [3], being some brains uniquely sensitive to these (table 1).

Although an unifying theory has proven elusive, from the several mechanisms proposed to contribute to POCD [4], both pain and the inflammatory process are accumulating consistent evidence [5]. The aggressions in the perioperative period can lead to an increase in pro-inflammatory cytokines. When there is no adequate balance through mechanisms of resolution of inflammation it is possible to subsequently generate a neuro-inflammatory process that facilitates the onset of cognitive dysfunction. The risk of POCD is greater in cases of increased surgical severity and with the occurrence of complications, as both of these factors are related to a higher magnitude of the inflammatory response [4,6]. Also, some authors have related the presence of Metabolic Syndrome with a higher risk of POCD [7] which could be related to the low-grade systemic inflammation characteristic of the Metabolic Syndrome [8,9]. Older animals and those with pathology of the CNS are more likely to develop inflammatory related alterations that go beyond the adaptive response of the sickness behaviour [6,10]. These observations add credence to the concept of diminished “cognitive reserve” in which elderly patients and those with existing neurodegeneration are more likely to suffer an episode of delirium [11,12]. Furthermore, both pain and opioid use can cause postoperative delirium [13], which can by itself lead to cognitive dysfunction [14]. Consequently, it is very important to adequately measure the nociception/antinociception balance, in order to provide the optimum level of analgesia.

Unfortunately, even though some of the mechanisms leading to these insults to the brain start to being understood, anesthesiologists are not totally aware of the dimension of this problem partly because an evaluation of brain function is absent neither in the formal preoperative assessment and during the course of the post-surgical recovery process [15]. The solution to this problem would be to routinely assess the cognitive function throughout the perioperative period, the same way other organ functions are [3]. The problem is that standard neuropsychologic test batteries, and many standard bedside screening tests are too cumbersome and time consuming to be practical in a perioperative evaluation [15]. The lack of a practical tool enabling the clinician to objectively assess and track the cognitive function trajectory throughout the perioperative period has hindered the knowledge on the "normal" pathway of the dynamics of cognitive function following the surgical process. So, it is still not well understood how different factors contribute to the normal recovery after a surgical intervention. Some proposals of this trajectory have already been made [16,17], but to really overcome this problem, it is necessary to identify a cognitive assessment instrument that is fast, objective, easily administered and scored, and has high interrater reliability [15].

Although the perfect instrument has yet to come, to label cognitive changes postoperatively as POCD, several instruments can be used. They can easily reveal some form of cognitive deterioration, but in order to classify them as POCD, it requires at least 2 measurements across the perioperative period. The timing of administration of these tests varies widely but the first test, the baseline, should be completed before surgery [18]. The present paper aims to review several cognitive evaluation methods, discussing their advantages and disadvantages as well as their potential clinical applications in evaluating the dynamics of the recovery of cognitive function after anesthesia and surgery.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Risk and precipitating factors for early/intermediate POCD. Adapted from [5] and [23]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk factors</td>
<td>Precipitating factors</td>
</tr>
<tr>
<td>Increasing age</td>
<td>Pain</td>
</tr>
<tr>
<td>Low education level</td>
<td>Postoperative infection</td>
</tr>
<tr>
<td>Burden of illness</td>
<td>Respiratory complications</td>
</tr>
<tr>
<td>Preoperative depression</td>
<td>Second operation within one week</td>
</tr>
<tr>
<td>Prior cognitive impairment</td>
<td>General Anesthesia vs regional anesthesia</td>
</tr>
<tr>
<td>Preoperative habits and drug use</td>
<td>Longer duration of anesthesia</td>
</tr>
<tr>
<td>Apolipoprotein E4</td>
<td></td>
</tr>
<tr>
<td>Cardiac surgery</td>
<td></td>
</tr>
</tbody>
</table>
Neuropsychological measures

Neuropsychological testing is essential to objectively measure cognitive function, and this can be achieved by assessing various aspects of intellectual function and personality [19]. The neuropsychological tests can be objective or subjective, be administered by an observer, or can be undertaken by the patients themselves with or without computer support [19]. They can be presented either as screen based, test batteries or can individually evaluate specific cognitive domains.

Different aspects of brain function can be assessed by using tests of problem solving, speed of information processing, flexibility and short-term memory. A consensus on POCD defined some core tests (table II) to assess cognitive function [20,21]. Additionally, the Paper and Pencil Memory Test (a test of sensorimotor speed and of the speed of recall) and the Four-Field Test (a test of psychomotor reaction time) can be used.

When combined, these tests allow an assessment of the cognitive function at the moment they are performed. However, as we discussed previously, to follow each patient cognitive course during time, different measurements should be taken repeatedly [51] (figure 1). This is applicable to all the neuropsychological measures we are going to discuss, and assumes a particular importance during the perioperative period, where significant changes can occur in a short period of time. Regarding the amount of changes required to make a determination of POCD, it depends on the degree of change that is considered statistically or clinically significant [18]; significance is one of the most controversial aspects. There are studies using 1 or more SD change on 2 tests, factor analysis, individual test analysis, and Z-score methodology in which a statistical measurement of a score’s relationship is compared with the mean in a group of scores [1,18]. This lack of agreement in the criteria for the assessment of POCD was evident in a study, where depending on the statistical definition, the number of patients in a single data set defined as POCD could vary from 70% to 16% [18,22]. A value of ≥ 2 in 2 or more tests or a composite score ≥ 2 from all tests was the cutoff level used in the International studies of postoperative cognitive dysfunction (iPOCD) to define POCD [1]. Exactly what would be considered clinically significant depends, to some extent, on the baseline status of the individual [18], so to a certain point, it is also important the personalization of this results, as different people in their daily life rely most on different cognitive functions.

Regardless the criteria used to classify cognitive changes as POCD, there are some issues that should be considered and addressed carefully, such as baseline performance and practice effects, issues that are pertinent to both the tests performed on computers or those with pen and paper.

Baseline performance

Baseline assessments allows the determination of whether an actual change in cognitive functioning occurred subsequent to the surgical event [23]. One advantage of using these tests in the perioperative period is that baseline performance may be determined. It is however essential that this initial measurement represents the optimum achievable by the patient. Otherwise, deterioration may not be detected. This may not be the case if the patient is studied on the day of admission or surgery when performance may be affected by tiredness or anxiety. Pre-existing disease can affect the performance but compensation for this factor is not possible. The optimal time for baseline assessment may be 1-2 weeks prior to surgery usually when the patient comes to the clinic for routine preoperative visit, preoperative tests and preparation [19]. Nearly half of studies of POCD have been conducted in adults undergoing cardiac surgery, a population at risk for cognitive changes because of underlying heart or vessel disease [23].

Practice effects

Practice effects refer to improvement in performance because of familiarity with test procedures [23] and have been noted for most neuropsychological tests both in volunteers and surgical patients. The degree of improvement with practice depends upon the number of times the test has been performed and the interval between tests. Generally, it is held that the practice effect diminishes with increasing familiarity and a “plateau” may be achieved, although this has not been well documented [19]. A common way of measuring practice effects is to measure the average improvement in performance for a group of matched controls [23].

Mean changes in control group scores from baseline to 1 week and baseline to 3 months provide measures of average learning effects, which are subtracted from changes in individual patient’s scores from scores from baseline to postoperative test. Each result is divided by the standard deviation of the change in control group scores to give a Z-score for individual patients in each subtest. Composite scores are averaged from the individual test scores of the battery [1]. The practice effect can be minimised with the use of tests, which have a proven low practice effect, and with the use of parallel versions instead of identical versions of the tests. This, however, tends to increase the variability and makes it more difficult to detect a difference between groups [19].

Paper and Pencil tests

Paper and pencil tests can be undertaken by patients themselves, or with the help of a professional. Different aspects of brain function can be assessed using these tests, but a comprehensive evaluation usually requires a battery of them. Although batteries containing larger numbers of neuropsychological tests are considered to have greater sensitivity to POCD, one unwanted consequence of increasing the number of tests in a battery is that the probability of falsely classifying cognitive change (i.e., type I error) also increases [24]. Thus, studies that
use large neuropsychological test batteries are likely to classify some individuals with POCD when they were normal. A comprehensive clinical neuropsychological examination takes about two and a half hours [20]. This is too time consuming and has limited the widespread of its use. However, shorter alternatives have been developed, allowing the screening for cognitive dysfunction in less time. Several of the available paper and pencil tests will be reviewed next, and each of them has its own advantages. However, their disadvantages are fairly similar: they are time consuming, they need an expert to review them, there can be inter-observer variability, and represent an increase in costs.

**Mini-Mental State Examination (MMSE)**

MMSE was designed as a screening tool for the clinical examination of patients with dementia. It was one of the first tests being used to detect cognitive dysfunction, and a good example of how these tests can be misused in this situation [1] as there is a marked learning effect [20]. This tool comprises a short battery of 20 individual tests covering 11 domains and totalling 30 points. Typical completion time is 8 min in cognitively unimpaired individuals and can take as much as 15 min in those with some form of cognitive impairment [25]. The cognitive skills measured are (1) time and place orientation, (2) registration (i.e., immediate auditory recall), (3) attention and calculation, (4) recall, and (5) language (receptive and expressive) [26]. From the total of 30 points, a cutoff levels of 23 can be used to differentiate dementia or cognitive impairment from normal functioning [26]. For most unimpaired adults, it is easy to score maximum. In addition, it has no parallel versions and consequently the same questions are repeatedly administered, resulting in learning effects [1]. It is therefore a suboptimal test to use in the setting of POCD.

**Montreal Cognitive Assessment**

The Montreal Cognitive Assessment (MoCA — http://www.mocatest.org) is a cognitive screening instrument developed to assess cognitive impairment and has been used in the setting of postoperative cognitive dysfunction in several studies [27-29]. It takes approximately 10 minutes to administer and has a total of 30 points. The cutoff for the MOCA is typically 26 out of 30 points, with scores below 26 considered as cognitive impaired [26]. It assesses multiple cognitive domains such as memory, language, executive functions, visuospatial skills,
calculation, abstraction, attention, concentration, and orientation [25]. One of the main reasons for the development of the MOCA was to create a measure that could detect MCI more effectively than the MMSE. It puts more emphasis on attention and executive functions than the MMSE, and contains a more difficult memory task, requiring the recitation and recall of five instead of three words [26]. MOCA is therefore more sensitive in detecting Mild Cognitive Impairment than MMSE, with a sensitivity of 90% vs. 18%.

Recently, MOCA was used to screen preoperatively patients aged 65 years and over, in whom no previous diagnosis of MCI had been made, yielding a potential prevalence of mild cognitive impairment of 56% [29].

Cognistat
The Neurobehavioral Cognitive Status Examination (Cognistat — http://www.cognistat.com) is a screening battery test that is more extensive in scope than the MMSE or MOCA, and is used as preliminary assessment of known or suspected neurologic disorders [26]. It assesses three general areas and five individual cognitive components as reflected in Table III. It takes longer than the MMSE and MOCA screens (10 to 20 minutes to administer). Although less implemented in the perioperative period, it has been used to assess postoperative cognitive dysfunction [30].

Post-operative Quality Recovery Scale (PQRS)
Postoperative Quality Recovery Scale (PQRS) was developed to assess how was the whole recovery process of surgical patients. It was developed by anesthesiologists and evaluates six domains of recovery (physiologic, nociceptive, emotive, activities of daily living (ADL), cognitive, and overall patient perspective) [31]. Although it was not designed to specifically assess cognition, in the cognitive domain tasks receive a performance score and can eventually be used for this purpose. In an attempt to improve the usability of this scale for detecting POCD, a modified version of the scoring in the cognitive domain of PQRS was studied [32]. It included a tolerance factor to account for performance variability, for changing the number of patients assessed as recovered, and also performing better to detect POCD.

Mini-Cog
The Mini-Cog involves a 3-item recall test for memory and a clock drawing test that serves as a distractor. It tests visual-spatial representation, recall, and executive function and has a sensitivity and specificity for the detection of dementia of 0.91 and 0.86, respectively [15]. It takes just 2 to 4 minutes to complete.

Mini-Cog can be graded on a 0 to 5 scale and investigators can be trained to grade the tests by reviewing information easily accessed via the Internet (http://www.alz.org/) [15,33,34].

Provable cognitive impairment can be defined as a score ≤ 2 [15]. It serves mainly as a screening tool, and as it is easy to implement, is recommended for all elderly patients preoperatively. If the screening score suggests impairment, a more complete assessment would be required [18].

Clock-in-the-Box (CIB)
The Clock-in-the-Box (CIB — http://www.heartbrain.com/cib) is a rapid (2-minute) cognitive screening tool, which assesses executive function components [15]. It involves a drawing test but includes no formal item recall, testing only working memory and planning/organization through providing the instructions in written form to the examinee and the response being provided in a predetermined location on the response form [15,34]. The written instructions, besides requiring the participant to read a page containing the instructions, and returning prior to completing the task on the response page, also reduces verbal administration cues comparing to the Mini-Cog [34]. It is a widely used cognitive screening tool that is simple and quick to administer and has been well accepted by both clinicians and patients [25]. Its application to the assessment of cognitive function perioperatively is fairly recent [15], but has been applied in other neurological and psychiatric disorders including: Alzheimer’s disease, Parkinson’s disease, Huntington’s disease, vascular disease, schizophrenia, stroke, and traumatic brain injury [25].

Other paper and pencil tests
There are several other paper and pencil tests that can be used to assess the cognitive function, but they have been less used in anesthesia, and its discussion is outside the scope of this review. Some of these are:
- The Quick Mild Cognitive Impairment Screen (Qmci)
- Six-Item Cognitive Impairment Test (6GIT);
- The General Practitioner Assessment of Cognition (GPCOG);
- TYM (Test Your Memory) Testing;
- DemTect;

<table>
<thead>
<tr>
<th>Table III</th>
<th>Areas and components assess through Cognistat</th>
</tr>
</thead>
<tbody>
<tr>
<td>General areas</td>
<td>Cognitive components</td>
</tr>
<tr>
<td>Level of consciousness</td>
<td>Language (comprehension, repetition, naming)</td>
</tr>
<tr>
<td>Orientation</td>
<td>Constructional ability (drawing from memory, arranging tiles)</td>
</tr>
<tr>
<td>Attention</td>
<td>Calculation skills</td>
</tr>
<tr>
<td>Memory</td>
<td>Executive skills (reasoning, judgment)</td>
</tr>
</tbody>
</table>
• Addenbrooke’s Cognitive Examinations;
• Erzikkeit’s Short Cognitive Performance Test (SKT).

**Computerized tests**

The use of computerized neurocognitive testing (CNT) goes back to several decades ago. The earliest were performance assessment batteries (PABs), used mostly in military and aerospace medicine, followed by World Health Organization which implemented PABs during the 1980s to evaluate neurocognitive impairment in industrial workers [35,36]. In the perioperative period, CNT have also demonstrated to be very useful and well suited to measure neurocognitive impairment in an environment where speed, efficiency and low cost are of great importance [37].

Computerized batteries offer a number of advantages over paper-and-pencil type tests: they are precise, accurate and can be timed to the nearest millisecond. In addition, they are easy to administer and score, have greater standardization and parallel versions may also be available which reduce practice effects [38]. Other advantages include the possibility of applying them at bedside using a tablet device, with instant results. The tests can be applied by any person, eliminating examiner effects and providing increased reliability [38].

The disadvantages include the possibility of unsupervised testing creating a suboptimal testing environment and the generation of vast amounts of data, whose clinical relevance may be hard to interpret properly [35]. Limitations in the ability to understand and manipulate information technology can also cause a negative effect as elderly tend not to be familiar with it [38].

**Cognitive Drug Research Computerized Assessment System (COGDRAS)**

Cognitive Drug Research Computerized Assessment System (COGDRAS, United BioSource Corp., Bethesda, MD), was developed for use in neuropharmacological research. It is based on more than 25 years of development and pharmaceutical industry experience, and it incorporates several computer-based measures to comprehensively assess cognition [39,40]. It was designed to be used repeatedly in any clinical population without showing training effects. It assesses the core aspects of cognitive function, such as attention, concentration, verbal and visuospatial recall and recognition, verbal and visual-spatial working memory, psychomotor speed and information processing speed, which globally underpin the ability to conduct the activities of daily living [38].

This battery requires approximately 20 to 25 minutes for completion and is available in more than 50 languages with at least 30 matched parallel. Although sensitivity and specificity data were not provided, the authors did provide data indicating that the COGDRAS is capable of identifying patients with cognitive dysfunction [39].

Benefits of the COGDRAS include the potential for Internet-based testing, ease of use, and long history of acceptance with the Food and Drug Administration and other regulatory agencies [39]. COGDRAS has been consistently used in surgery and anesthesia [41], and has helped to clarify the importance of intra-operative monitoring of anaesthetic depth and cerebral oxygenation in reducing post-operative cognitive impairment [42].

**The Computer Self Test (CST)**

The Computer Self Test (CST; University of Tennessee Medical Center, Knoxville, TN) was recently adapted from a paper-and-pencil version of the same battery [39].

CST is internet based and takes approximately 15 minutes to complete. It provides a user friendly interface with both written and oral instructions [39]. This battery demonstrates a high degree of sensitivity and specificity and is capable of accurately identifying cognitive impairment in patients with variable degrees of cognitive abnormality. In a case-control study [43], it accurately classified 96% of the cognitively impaired individuals as compared to controls, while the Mini-Mental Status Examination (MMSE) accurately classified 71% and the Mini-Cog 69% in the same respect.

It has been used in anesthesia [42], in RTC, to assess the potential benefits of intra-operative monitoring of anaesthetic depth and cerebral oxygenation as a pragmatic intervention to reduce post-operative cognitive impairment.

**CogState**

The CogState battery (CogState, Ltd., Melbourne, Victoria, Australia) was designed with an emphasis on creating cross-culturally valid tasks requiring minimal verbal instructions and responses. The use of a computer interface and minimal human assistance needed for administration make this battery generally accepted by participant populations of varied clinical domains, cultures, and ages [39]. Responses are made via a computer keyboard represented graphically on the screen, with responses using the “k” key for yes and “d” key for no [38]. Given that the CogState Brief Battery (CBB) is computerized, the administration, scoring and reporting is automated and highly standardized [44]. CBB also offers Internet-based administration, allowing for remote testing, and as far as we know it is the only measure with the current capability for real-time quality assurance and data reporting. This allows active monitoring by data safety monitoring boards or remote monitoring of cognitive changes in the home setting [39].

This cognitive test battery requires approximately 10 minutes for administration and consists of four cognitive tasks that measure psychomotor function, attention, working memory and memory [44,45]. Some of these tasks are: detection task (simple reaction time), identification task (choice reaction time), one-back task (working memory), one card learning task (visual learning), Groton Maze Learning task (executive...
function), continuous paired associate learning task (visual associate learning), and the international shopping list task (verbal learning and memory) [39]. Recent data from studies using the CBB suggests that composite scores, which are constructed from aggregating performance on the Detection and Identification tasks (i.e., an attention/psychomotor composite) and the learning and working memory tasks (i.e., a learning/working memory composite) may have greater sensitivity to cognitive impairment and decline when compared to scores from the individual CBB tasks [44]. While the CBB is not intended to replace formal neuropsychological assessment, the results of these recent studies do converge to suggest that it may be useful as a screening test for cognitive impairment in clinical settings. Establishing the reliability and stability of these composite scores could facilitate the use of composite cognitive measures to monitor changes in cognitive function during the perioperative period [44,46]. The self-administered Cogstate battery showed significant associations with several risk factors known to be associated with cognitive function. Future studies of cognitive aging may benefit from the numerous advantages of self-administered computerized testing [47].

Cogstate has also been used in anesthesia to assess cognitive impairment after sedation [46,48] and to compare the postoperative cognitive impairment with the measurements of biomarkers [49].

**Other computerized tests**
- NIH Toolbox.
- Mindestreams (Neurotrax).
- Cogtrack.

**Conclusions**

Recovery of cognitive function after anesthesia and surgery is a highly relevant topic that has not been adequately covered because of the lack of easy, fast tests and equipment to assess it at the bedside. It is of special relevance in the case of outpatients undergoing surgical procedures requiring general anesthesia or sedation. A description of the dynamics of recovery of cognitive function requires an objective measurement that could be assessed at different times. To identify cognitive impairment in patients before they come to the operating room, it requires that anesthesiologists would look for it. A paradigm shift is therefore necessary such that cognitive assessment becomes a routine part of the preoperative screening of elderly patients, not just a research tool. The task is an important and necessary one, and framing the approach toward identifying MCI (and even mild dementia, which is frequently not detected clinically) has tangible and theoretic benefits [50]. Some of these tests and tasks are useful to detect cognitive impairment preoperatively, but they may lack the sensitivity to adequately follow up these patients during the perioperative period. It is therefore important to choose the tests, which addresses the changes we want to measure.

Summarizing, scientific and technological advances have allowed the introduction of different computerized systems permitting the evaluation of several different aspects of cognition with relevant accuracy. Using these tools will change anesthesia practice both to exactly assess recovery from anesthetic effects and to define a patient fit for discharge as well as will allow a more accurate definition of cognitive function at different time points to define the spectrum of dynamic changes of cognitive function in the setting of Postoperative Cognitive Decline.

**Funding:** the authors declare no funding.

**Disclosure of interest:** the authors declare that they have no competing interest.

**References**

[8] Holvoet P. Relations between metabolic syndrome, oxidative stress and inflammation and...


