

Aus der Klinik für Anästhesiologie
mit Schwerpunkt operative Intensivmedizin
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Postoperative kognitive Dysfunktionen bei nicht kardiochirurgischen
Patienten – ein Vergleich zweier Testmethoden

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Norbert Papkalla
aus Ludwigshafen am Rhein

Gutachter: 1. Prof. Dr. med. C. Spies
2. Prof. Dr. med. E. Steinhagen-Thiessen
3. Prof. Dr. med. W. Wilhelm

Datum der Promotion: 05.06.2011

Abkürzungsverzeichnis

| | |
|--------|--|
| ASA | American Society of Anesthesiologists |
| CI | Confidence Interval (Konfidenzintervall) |
| CTB | Computerized Test Battery (computerbasierte Testbatterie) |
| DGAI | Deutsche Gesellschaft für Anästhesiologie und Intensivmedizin e.V. |
| DET | Detection Task |
| ICD | International Statistical Classification of Diseases and Related Health Problems |
| IDT | Identification Task |
| IQR | Interquartile Range (Interquartilsabstand) |
| IR | Impairment Index Rating |
| ISPOCD | International Study of Post-Operative Cognitive Dysfunction |
| KIS | Krankenhausinformationssystem |
| MMSE | Mini-mental State Examination |
| NPOCD | No postoperative cognitive dysfunction (keine postoperativen kognitiven Dysfunktionen) |
| OCLT | One Card Learning Task |
| POCD | Postoperative Cognitive Dysfunction (postoperative kognitive Dysfunktion) |
| PPTB | Paper Pencil Test Battery (Papier- und Bleistift-Testbatterie) |
| PS | Physical Status |
| RCI | Reliable Change Index |
| SD | Standard Deviation (Standardabweichung) |
| SOP | Standard Operating Procedures |
| SCWT | Stroop Colour Word Test |
| VVLT | Visual Verbal Learning Test |

Inhaltsverzeichnis

| | | |
|-----|--|----|
| 1 | Einleitung | 6 |
| 1.1 | Postoperative kognitive Dysfunktionen | 6 |
| 1.2 | Inzidenz und Bedeutung von POCD | 7 |
| 1.3 | Risikofaktoren für POCD | 8 |
| 1.4 | Detektion von POCD | 8 |
| 1.5 | Computerisierte versus Papier- und Bleistift-Testbatterien | 13 |
| 2 | Fragestellung | 15 |
| 3 | Methoden | 16 |
| 3.1 | Studiendesign | 16 |
| 3.2 | Patienten- und Kontrollgruppenkollektiv | 16 |
| 3.3 | Erhobene Daten | 17 |
| 3.4 | Testbatterien | 18 |
| 3.5 | POCD-Berechnung | 22 |
| 3.6 | Statistik | 24 |
| 4 | Ergebnisse | 25 |
| 4.1 | Ausgeschlossene Patienten | 26 |
| 4.2 | Patienten | 26 |
| 4.3 | Deskriptive Statistik | 26 |
| 4.4 | POCD-Klassifikationen und Interrater-Reliabilität | 29 |
| 4.5 | Basisdaten postoperativ nach Testbatterie | 31 |
| 5 | Diskussion | 35 |
| 5.1 | Inzidenzen | 35 |
| 5.2 | Diagnostische Tests | 36 |
| 5.3 | Limitationen dieser Studie | 38 |
| 5.4 | Interpretation und Ausblick | 42 |
| 6 | Zusammenfassung | 44 |
| 6.1 | Einleitung | 44 |
| 6.2 | Methodik | 44 |
| 6.3 | Ergebnisse | 44 |
| 6.4 | Diskussion | 44 |
| 7 | Literaturverzeichnis | 46 |
| 8 | Publikationen im Rahmen der Dissertation | 52 |
| 9 | Abbildungsverzeichnis | 53 |

| | |
|------------------------|----|
| Einleitung | 5 |
| 10 Tabellenverzeichnis | 54 |
| 11 Danksagung | 55 |
| 12 Lebenslauf | 56 |
| Erklärung | 57 |

1 Einleitung

1.1 Postoperative kognitive Dysfunktionen

Der Begriff postoperative kognitive Dysfunktion (Postoperative Cognitive Dysfunction, POCD) beschreibt die Abnahme der kognitiven Leistungsfähigkeit nach einer Operation^{1;2}. Die Symptome können von einem milden, nur wenige Tage anhaltenden Rückgang der kognitiven Leistungsfähigkeit, der vom Patienten häufig selbst nicht wahrgenommen wird, bis hin zu länger anhaltenden kognitiven Einschränkungen des Gedächtnis, der Informationsverarbeitung und der Auffassungsgabe reichen^{3;4}. Eine schwer ausgeprägte postoperative kognitive Dysfunktion kann die Selbstständigkeit des Patienten erheblich beeinträchtigen⁵. Diese Störung kann auch nach Eingriffen und Operationen, die ansonsten unkompliziert verlaufen sind, auftreten und wird zur perioperativen Morbidität gerechnet².

In der Literatur werden an Stelle von POCD häufig andere Termini wie Delir und Demenz verwendet⁶. Die International Statistical Classification of Diseases and Related Health Problems (ICD) Systematik in der 10. Version, German Modification (ICD-10-GM)⁷, hält für jeden der oben genannten Begriffe eine eigene Krankheitsgruppe bereit und ermöglicht so eine Abgrenzung der Termini. Nach dieser Klassifikation ist die leichte kognitive Störung (F06.7), in die sich die postoperative kognitive Dysfunktion einordnen lässt, charakterisiert durch „Gedächtnisstörungen, Lernschwierigkeiten und die verminderte Fähigkeit, sich längere Zeit auf eine Aufgabe zu konzentrieren. Oft besteht ein Gefühl geistiger Ermüdung bei dem Versuch, Aufgaben zu lösen. Objektiv erfolgreiches Lernen wird subjektiv als schwierig empfunden. Keines dieser Symptome ist so schwerwiegend, dass die Diagnose einer Demenz (F00-F03) oder eines Delirs (F05.-) gestellt werden kann“. POCD wird über eine Abnahme der Leistung in neuropsychologischen Tests zwischen einer präoperativen und einer postoperativen Sitzung definiert⁸⁻¹⁰.

Das Delir (F05.-) hingegen ist nach ICD-10-GM charakterisiert durch eine „gleichzeitig bestehende Störung des Bewusstseins und der Aufmerksamkeit, der Wahrnehmung, des Denkens, des Gedächtnisses, der Psychomotorik, der Emotionalität und des Schlaf-Wach-Rhythmus. Die Dauer ist sehr unterschiedlich und der Schweregrad reicht von leicht bis zu sehr schwer.“ Im Gegensatz zu POCD tritt ein Delir typischerweise in

den ersten Stunden bis Tagen nach einer Operation auf und wird im Allgemeinen als transient eingestuft¹¹.

Die Demenz (F00-F03) hingegen ist nach ICD-10-GM „ein Syndrom als Folge einer meist chronischen oder fortschreitenden Krankheit des Gehirns mit Störung vieler höherer kortikaler Funktionen, einschließlich Gedächtnis, Denken, Orientierung, Auffassung, Rechnen, Lernfähigkeit, Sprache und Urteilsvermögen.“

Klinisch können diese drei Krankheitsgruppen aber nicht immer eindeutig von einander getrennt werden⁶.

1.2 Inzidenz und Bedeutung von POCD

In der Fachliteratur wird zwischen früher POCD (Tage bis Wochen nach einem operativen Eingriff) und später POCD (Monate bis Jahre nach einem operativen Eingriff) unterschieden. Die Inzidenzen für frühe POCD werden in der Literatur zwischen 14,4% in der Studie von Linstedt et al.¹² und 40% in der Studie von Iohom et al.¹³ angegeben. Für späte POCD werden Inzidenzen von „nicht nachweisbar“ in der Studie von Hall et al.¹⁴ bis 53% in der Studie von Iohom et al.¹³ angegeben. Die zum Teil erheblichen Abweichungen der Inzidenzen sind auf Unterschiede in den Messmethoden und Patientenkollektiven zurückzuführen^{15;16}. Moller und Kollegen¹⁷ konnten in einer großen multizentrischen Untersuchung, der „International Study of Post-Operative Cognitive Dysfunction“ (ISPOCD), eine POCD-Inzidenz von 25,8% nach 7 Tagen und 9,8% nach 3 Monaten ermitteln. Abildstrom et al.¹⁸ konnten in einer Studie mit 336 Patienten zeigen, dass POCD bei ca. 1% der Patienten auch nach zwei Jahren nachweisbar ist.

Des Weiteren wird zwischen POCD nach nicht kardiochirurgischen Eingriffen und POCD nach kardiochirurgischen Eingriffen, die eine höhere Inzidenz von bis zu 80% aufweisen¹⁹, unterschieden. Die höhere POCD-Inzidenz nach kardiochirurgischen Eingriffen wird unter anderem auf die Phase des Herzstillstandes während der Operation sowie auf mikroembolische Ereignisse^{3;11;20}, die durch den Einsatz der Herz-Lungen-Maschine gefördert werden, zurückgeführt.

Bei einem Vergleich von Regionalverfahren mit Allgemeinanästhesien konnte durch Rasmussen et al.²¹, in einer randomisierten Studie mit 438 Teilnehmern über 60 Jahre, keine unterschiedliche Inzidenz bei späten POCD gezeigt werden. Allerdings konnte diese Studie in einer Analyse eine geringere Inzidenz von frühen POCD in der Gruppe mit Regionalverfahren zeigen. Rasmussen et al. deuten dies als einen Hinweis darauf, dass Narkotika oder die postoperative Schmerzbehandlung einen negativen Einfluss

auf die Entstehung von POCD haben könnten.

Die Observationsstudie von Steinmetz et al.²² mit 701 Patienten aus der ISPOCD 1 und ISPOCD 2 Studie (Nachverfolgung im Median 8,5 Jahre) konnte zeigen, dass POCD bei Entlassung aus dem Krankenhaus mit einem erhöhten Risiko für das vorzeitige Verlassen des Arbeitsmarktes sowie einer verlängerten Zahlung von Sozialleistungen assoziiert ist. Des Weiteren war in dieser Studie POCD nach 3 Monaten mit einer erhöhten Gesamtmortalität verbunden. Monk et al.²³ konnten zeigen, dass POCD bei Krankenhausentlassung mit einer erhöhten 3-Monats-Mortalität assoziiert ist. Bei Patienten, die sowohl nach einer Woche, als auch nach 3 Monaten POCD zeigten, konnte eine erhöhte 1-Jahres-Mortalität nachgewiesen werden. Außerdem scheinen Patienten mit POCD eine längere Krankenhausverweildauer und einen erhöhten Pflege- und Betreuungsaufwand zu haben²⁴.

Neben dieser direkten sozioökonomischen Relevanz ist POCD bei älteren Menschen mit einem Rückgang der Aktivitäten des täglichen Lebens verbunden, so dass ein erhöhter Bedarf an Unterstützung notwendig wird¹⁷.

1.3 Risikofaktoren für POCD

In einer Vielzahl von Studien wurden unterschiedliche Risikofaktoren für POCD identifiziert. Die bis jetzt größte durchgeführte multizentrische Kohortenstudie¹⁷ mit 1218 Teilnehmern konnte Alter, Bildungsstand, Anästhesiedauer, zweite Operation, postoperative Infektion und postoperative respiratorische Komplikationen als Risikofaktoren für frühe POCD nachweisen. Für späte POCD konnten nur Alter und präoperative Gabe von Benzodiazepinen als Risikofaktoren identifiziert werden. Monk und Kollegen²³ konnten in einer weiteren Studie zeigen, dass Alter, Schulbildung, ASA-Klassifikation (American Society of Anesthesiologists Physical Status (ASA-PS)), die Eingriffsart, die Länge des Krankenhausaufenthaltes und Delirium während des Krankenhausaufenthaltes Risikofaktoren für POCD sind.

1.4 Detektion von POCD

Da POCD sich oft subtil präsentiert und subjektive Einschätzungen des Patienten nur wenig mit objektiven Messungen korrelieren²⁵ ist es notwendig, vor und nach der Operation neuropsychologische Tests zur Detektion durchzuführen⁸⁻¹⁰. In der Literatur werden eine Vielzahl verschiedener Tests und Testkombinationen (sog. Testbatterien)

sowie Berechnungsmethoden zur Detektion von POCD vorgeschlagen. Um eine bessere Vergleichbarkeit der Ergebnisse zu ermöglichen, wurde 1995 eine Konsensusvereinbarung zur Detektion von POCD nach herzchirurgischen Eingriffen von Murkin et al.¹⁰ veröffentlicht. Diese beschreibt wesentliche Anforderungen an die Testauswahl, Durchführung und Berechnung und kann mit ihren Kernaussagen auch als Leitlinie für Studien mit nicht kardiochirurgischen Patienten dienen.

1.4.1 Testauswahl

In einer systematischen Übersichtsarbeit¹⁵ zum Thema postoperative kognitive Dysfunktionen nach nicht kardiochirurgischen Eingriffen wurden 46 Arbeiten analysiert. In diesen Arbeiten kamen 72 verschiedene kognitive Tests, unterteilt in die sieben Bereiche Gedächtnis und Lernen, Aufmerksamkeit und Konzentration, visuelle und räumliche Fähigkeiten, visuomotorische und manuelle Fähigkeiten, numerische Fähigkeiten, ausführende Funktionen und kombinierte Tests zur Anwendung. Die verwendeten Tests reichten von komplexen und aufwändigen Batterien wie der Wechsler Adult Intelligence Scale als traditionellem Intelligenzquotient-Test, bis zu einfachen, schnell durchzuführenden Screening-Tests wie der Mini-mental state examination (MMSE).

Postoperative kognitive Dysfunktionen werden über eine Abnahme der Leistungsfähigkeit in einer oder mehreren kognitiven Domänen definiert¹¹. Aus diesem Grund ist es nicht immer möglich, mit einem einzigen neuropsychologischen Test POCD nachzuweisen. Die Konsensusvereinbarung fordert deshalb bei der Zusammenstellung der Testbatterie auf eine ausgeglichene Balance der getesteten Domänen zu achten¹⁰. Ein weiterer wichtiger Punkt bei der Zusammenstellung der Testbatterie ist die Zeit, die zum Ausführen benötigt wird. Eine zu lange zeitliche Beanspruchung kann zu einer Abnahme der Testbereitschaft führen²⁶. Als akzeptables Maß wird in der Literatur eine maximale Testausführungszeit von 45 Minuten empfohlen¹⁷.

Die ausgewählten Tests sollten frei von kulturellen Einflüssen sein und keinem Boden- oder Deckeneffekt unterliegen¹⁰. Boden- und Deckeneffekte sind Messfehler die auftreten können, wenn der Test so schwer ist, dass kaum ein Proband eine minimale Punktzahl erreicht oder der Test so leicht ist, dass fast jeder Proband die volle Punktzahl erreicht. So werden Probanden mit niedriger Leistungsfähigkeit in sehr schweren Tests keine Steigerung durch mehrfache Testausführung erreichen, während Probanden mit hoher Leistungsfähigkeit zwischen den Testsitzungen eine Steigerung

der Testleistung aufweisen. Bei sehr leichten Tests gilt dies umgekehrt. Hier haben sich Tests, die die Zeit als Leistungsmerkmal haben, als vorteilhaft erwiesen, da hier Boden- und Deckeneffekte seltener auftreten¹⁹. Rasmussen und Kollegen empfehlen aber trotzdem, auf ein ausgeglichenes Verhältnis von zeit- und punktebasierten Tests in der Testbatterie zu achten¹⁹.

Die ausgewählten Tests sollten sowohl sensitiv sein als auch eine hohe Test-Retest-Reliabilität haben, um auch subtile Veränderungen der kognitiven Leistungsfähigkeit messen zu können^{10;19}. Ferner sollten sie für eine wiederholte Anwendung in einem chirurgischen Patientenkollektiv geeignet sein. So haben sich gerade klassische IQ-Tests bei der Detektion von POCD als zu unsensitiv erwiesen¹⁹.

1.4.2 Testbedingungen

Die prä- und postoperativen Testungen sollten unter den gleichen Bedingungen stattfinden¹⁰. Des Weiteren wird empfohlen jede Testung, um maximale Vergleichbarkeit zu gewährleisten, von einem geschulten Versuchsleiter nach einem standardisierten Protokoll durchführen zu lassen. In jeder Testsitzung sollten die gleichen Materialien zur Anwendung kommen¹⁹. Störungen während der Testsitzung können minimiert werden, indem ein separater Raum zur Testdurchführung benutzt wird¹⁹. Die Testungen sollten zur gleichen Uhrzeit stattfinden, um Tageszeit bedingte Schwankungen der kognitiven Leistungsfähigkeit möglichst auszuschließen.

1.4.3 Ausgangsmessung

Veränderungen der kognitiven Leistungsfähigkeiten nach Operationen lassen sich nur sinnvoll beurteilen, wenn ein Ausgangswert vorhanden ist.¹⁰ Die Ausgangsmessung repräsentiert die bestmögliche kognitive Leistung des Patienten¹⁹. Sie sollte nicht am Abend vor der Operation durchgeführt werden, da Angst und Müdigkeit einen negativen Einfluss auf das präoperative Ergebnis haben könnten. In der Literatur wird eine Zeitspanne von ein bis zwei Wochen vor der geplanten Operation vorgeschlagen¹⁹.

1.4.4 Lerneffekt

Lerneffekte in neuropsychologischen Tests können auftreten, wenn ein Proband einen Test mehrfach absolviert²⁷ und dadurch die Möglichkeit hat, seine Lösungsstrategien zu

optimieren. Als Folge können ab der zweiten Testsitzung bessere Testwerte erzielt werden, ohne dass es zu einer Veränderung der kognitiven Leistungsfähigkeit gekommen ist. Das Ausmaß des Lerneffektes hängt unter anderem von der Anzahl der Wiederholungen, dem Schwierigkeitsgrad des Tests, den Zeiträumen zwischen den Wiederholungen, dem Alter und den individuellen Fähigkeiten des Probanden ab^{19;28}.

Eine Nichtberücksichtigung des Lerneffekts in der Testauswertung kann dazu führen, dass ein kognitives Defizit bei Patienten, die zwar eine Steigerung ihres Testergebnisses (markiert mit ‡ in Abbildung 1) zwischen den beiden Testsitzungen erzielt haben, aber unter der durchschnittlichen Steigerung (markiert mit † in Abbildung 1) liegen, nicht erkannt wird²⁹. D. h. eine in Wahrheit stattfindende Abnahme der kognitiven Leistungsfähigkeit wird durch den Lerneffekt kompensiert und stellt sich unter Umständen als Steigerung dar. Um eine solche Verfälschung der Testergebnisse zu vermeiden, wird von mehreren Autoren gefordert den Lerneffekt zu berücksichtigen^{10;19}.

Der Lerneffekt kann im Wesentlichen auf zwei Arten minimiert werden¹⁹:

Erstens indem Tests verwendet werden, die per se einen minimalen Lerneffekt haben oder in mehreren parallelen Versionen^{11;29} (z.B. mit anderen Wörtern oder Zahlen) zur Verfügung stehen.

Zweitens indem eine Kontrollgruppe, bestehend aus Personen mit ähnlichen Eigenschaften bezüglich Alter und Grunderkrankung, die sich keinem chirurgischen Eingriff unterziehen müssen, im gleichen Test-Re-Test-Intervall getestet wird. Von den Ergebnissen der Patientengruppe wird dann die durchschnittliche Ergebnissteigerung der Kontrollgruppe abgezogen. Die so bereinigten Messwerte spiegeln die Veränderung der kognitiven Leistungsfähigkeit ohne Lerneffekt wieder. Einschränkend ist zu sagen, dass bei Tests, die auf dem Erlernen von Regeln basieren, der Lerneffekt nicht durch parallele Versionen minimiert werden kann, da erlernte Regeln nicht so schnell vergessen werden²⁸.

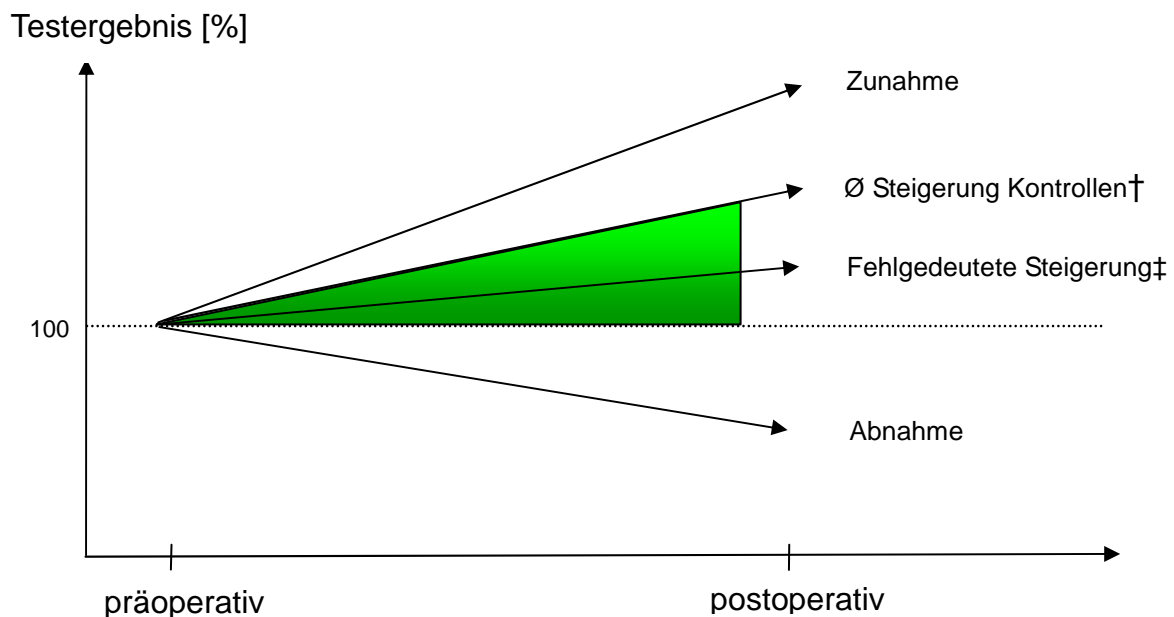


Abbildung 1 Lerneffekt bei wiederholter Testausführung modifiziert nach Newman et al.³⁰

1.4.5 Auswertung der Testergebnisse

In der Literatur werden verschiedene Methoden zur Auswertung der Testergebnisse und Ermittlung von POCD vorgeschlagen, die sich im Wesentlichen in zwei Gruppen einteilen lassen²⁶: Methoden, die auf einer Veränderung des Gruppenmittels basieren und Methoden, die eine individuelle Veränderung des Patienten berücksichtigen.

Der Vorteil der ersten Methode besteht darin, dass sich mit ihr zwei Gruppen sehr sensitiv miteinander vergleichen lassen. Hierzu wird meist ein t-Test verwendet, um einen statistischen Unterschied zwischen den Gruppen für verschiedene Testmerkmale nachzuweisen. Der Nachteil dieser Methode ist, dass der individuelle Verlauf eines einzelnen Patienten nicht berücksichtigt wird^{26;30}. Außerdem werden Subgruppen, wie z.B. ältere oder kränkere Probanden nicht erfasst¹⁹.

Bei der zweiten Methode wird anhand der Messergebnisse der neuropsychologischen Tests eine dichotome Einteilung in POCD ja/nein für jeden einzelnen Probanden vorgenommen. So kann für jeden Patienten individuell entschieden werden, ob eine Abnahme der kognitiven Leistungsfähigkeit zwischen den zwei Messzeitpunkten stattgefunden hat. In der Literatur werden hierzu im Wesentlichen zwei Definitionen ohne Kontrollgruppe beschrieben^{9;26}: Abnahme der kognitiven Leistungsfähigkeit um eine halbe oder eine Standardabweichung (SD) in mehr als 20% der Messungen oder eine Abnahme um mehr als 20% von der Ausgangsmessung in mehr als 20% der Tests

(20-20-Methode). Die Cut-Off-Werte dieser beiden Verfahren können beliebig definiert werden. Um eine Abnahme der kognitiven Leistungsfähigkeit nachweisen zu können, muss der Patient aber ein präoperatives Ergebnis erzielt haben, welches eine Abnahme von einer Standardabweichung überhaupt zulässt²⁶.

Eine weitere Möglichkeit zur Definition von POCD ist das Impairment Index Rating (IR)³¹. Mit dieser Methode wird anhand der durchschnittlichen Testleistung des Patientenkollektivs und der Standardabweichung eine Einteilung der Abnahme der kognitiven Leistungsfähigkeit in mehrere Gruppen vorgenommen. Dies ermöglicht eine Graduierung der kognitiven Störung.

Mahanna et al.³² verglichen, in einer Gruppe von 232 herzchirurgischen Patienten, die anhand dieser drei Methoden ermittelten POCD-Inzidenzen. Zum Zeitpunkt der Entlassung reichten die Inzidenzen von 66% für die 20-20-Methode bis zu 15,3% wenn das IR angewendet wurde. Mahanna und Kollegen folgerten, dass die verwendete Definition einen wesentlichen Einfluss auf die Inzidenz hat. Dies wurde in ähnlicher Form ebenfalls von Lewis et al.⁹ und Kneebone et al.³³ gezeigt.

Jacobson und Truax³⁴ entwickelten die Methode des Reliable Change Index (RCI), welche in der Folgezeit mehrfach modifiziert wurde, um den Lerneffekt einer Kontrollgruppe mit zu berücksichtigen³⁵. Der RCI beurteilt, ob die gemessene Veränderung der kognitiven Leistungsfähigkeit zwischen zwei Testsitzungen so groß ist, dass es unwahrscheinlich ist, dass die Änderung aufgrund von Messschwankungen zustande kam. Mit anderen Worten lässt sich mit dem RCI entscheiden, ob die Veränderung zwischen den Testsitzungen zufällig zustande kam oder nicht.

Nach allen Methoden, mit Ausnahme des IR, wird POCD anhand von stetigen oder diskreten Messwerten dichotom eingeteilt. Es findet keine weitere Graduierung von POCD statt. Eine die Methoden vergleichende Untersuchung⁹ kam zu dem Ergebnis, dass der RCI die am besten geeignete Methode zur Messung von POCD ist.

1.5 Computerisierte versus Papier- und Bleistift-Testbatterien

Neuropsychologische Tests werden traditionell als Papier- und Bleistift-Testbatterien (Paper Pencil Test Battery, PPTB) durchgeführt, es stehen aber auch immer mehr computerisierte Testbatterien (Computerized Test Battery, CTB) zur Verfügung.

Sowohl für CTBs als auch für PPTBs wird eine kurze Durchführungszeit, eine hohe Test-Re-Test-Reliabilität, eine hohe Sensitivität, ein minimaler Lerneffekt, eine große Bandbreite an möglichen Testwerten sowie das Fehlen von Boden- und Deckeneffekten

gefordert^{19;36}.

CTBs scheinen hier einige Vorteile gegenüber den konventionellen Testbatterien zu haben^{36;37}: So sollen computerisierte Testbatterien durch ihren hohen Grad an Standardisierung, sowohl bei der Testdurchführung als auch bei der Präsentation der Stimuli, eine höhere Test-Re-Test-Reliabilität erreichen. Des Weiteren können CTBs eine nahezu unendliche Anzahl unterschiedlicher Stimuli generieren und können so den Lerneffekt minimieren. Dadurch kann auf die Verwendung von parallelen Testversionen oft verzichtet werden. Computerisierte Testbatterien erlauben eine genauere Aufzeichnung der Daten, v. a. der Zeitmessungen, die durch manuelle Messmethoden nicht erreicht wird. Die Datenspeicherung und Datenauswertung geschieht automatisiert und minimiert so die Fehlerwahrscheinlichkeit. Weiterhin bieten CTBs oft die Möglichkeit der Online-Datenanalyse.

Ein weiterer Punkt ist die mögliche Kosteneinsparung durch die geringere Durchführungszeit³⁷. So können CTBs von ärztlichem Hilfspersonal durchgeführt werden und benötigen nicht mehr die ständige Anwesenheit eines neuropsychologisch ausgebildeten Versuchsleiters³⁷. Außerdem ist es dem Versuchsleiter bei der Vielzahl der präsentierten Stimuli oft nicht möglich sich ein Bild der Leistung des Patienten während der Testausführung zu machen. So können weder die weitere Testausführung noch die spätere Auswertung durch den Versuchsleiter verfälscht werden³⁸.

Mit der heute weiten Verbreitung von Computern bieten CTBs die Möglichkeit, eine Vielzahl von Patienten in kurzer Zeit auf kognitive Dysfunktionen zu testen.

Mögliche Nachteile einer computerisierten Testdurchführung werden von Wild et al.³⁷ in schlecht entworfenen User-Interfaces, der Möglichkeit der generellen Ablehnung von Computertechnik, v. a. durch ältere Probanden, sowie in der fehlenden Einschätzung eines erfahrenen Versuchsleiters gesehen. Auch darf nicht vorausgesetzt werden, dass neuropsychologische Tests nach Portierung in ein neues Medium (hier Computer) äquivalente Ergebnisse liefern. Unterschiede in der Testinstruktion, Stimuluspräsentation und Messart können zu signifikant anderen Ergebnissen führen³⁷. Silbert et al.³⁶ konnten zeigen, dass computerisierte Testbatterien prinzipiell geeignet sind, postoperative kognitive Dysfunktionen nach kardiochirurgischen Eingriffen nachzuweisen.

2 Fragestellung

Zur Detektion von POCD können sowohl Papier- und Bleistift-Testbatterien als auch computerisierte Testbatterien eingesetzt werden. In der Literatur wird beschrieben, dass computerisierte Testbatterien sich durch eine standardisierte Durchführung, eine höhere Reliabilität, eine erhöhte Genauigkeit bei der Messung sowie eine vereinfachte Handhabung der erhobenen Daten auszeichnen³⁶. Des Weiteren werden ihnen Vorteile bei der Testdurchführung, der Erzeugung von Teststimuli, der Bereitstellung alternativer Testformen und der Datenverarbeitung zugeschrieben^{36;37;39}.

Ziel der Studie war es, die Inzidenzen von POCD und die Übereinstimmung der Klassifikation, gemessen mit einer computerisierten und einer Papier- und Bleistift-Testbatterie, in demselben nicht kardiochirurgischen Patientenkollektiv zu untersuchen.

3 Methoden

3.1 Studiendesign

Die hier vorliegende prospektive Kohortenstudie wurde vom September 2006 bis April 2008 mit Zustimmung der lokalen Ethikkommission Charité - Universitätsmedizin Berlin durchgeführt. Alle Patienten haben eine Aufklärung über die Studieninhalte erhalten und gaben ihr schriftliches Einverständnis zur Teilnahme.

Alle Patienten wurden mit einer computerisierten Testbatterie und einer konventionellen Papier- und Bleistift-Testbatterie auf kognitive Dysfunktionen nach nicht kardiochirurgischen Eingriffen getestet. Zwei geschulte Mitarbeiter, supervidiert durch eine Psychologin, haben die Tests nach einem schriftlich fixierten Studienprotokoll durchgeführt. Die Daten wurden mittels eines Dokumentationsbogens erfasst und umgehend auf Plausibilität und Vollständigkeit geprüft.

Alle Eingriffe sind in Allgemeinanästhesie nach den aktuellen Standard Operating Procedures (SOP) der Klinik⁴⁰ durchgeführt worden. Das Studienprotokoll sah keine Einschränkungen über die Art der Einleitung oder die Aufrechterhaltung der Narkose vor.

3.2 Patienten- und Kontrollgruppenkollektiv

Patienten wurden in die Studie eingeschlossen, wenn alle unten genannten Einschlusskriterien erfüllt waren und kein Ausschlusskriterium vorlag. Ob die Deutschkenntnisse ausreichend waren, entschied der Versuchsleiter anhand des Aufklärungsgespräches und der Durchführung der Mini-mental state examination⁴¹ (MMSE).

3.2.1 Einschlusskriterien

Einschlusskriterien waren:

- ASA-PS III
- Mindestens 23 Punkte im MMSE-Test
- Durchführung einer Allgemeinanästhesie
- Geplante Anästhesiedauer von mindestens 180 Minuten

- Erwarteter stationärer Aufenthalt von mehr als 6 Tagen
- Schriftliches Einverständnis

3.2.2 Ausschlusskriterien

Ausschlusskriterien waren:

- Neurologische und/oder psychiatrische Grunderkrankungen
- Bekannter Medikamenten-, Alkohol- oder Drogenabusus
- Kardiochirurgische oder neurochirurgische Eingriffe
- Schwere Beeinträchtigung der Sehfähigkeit oder des Gehörs
- Mangelnde Deutschkenntnisse
- Unwillen oder Unfähigkeit des Patienten dem Studienprotokoll zu folgen
- Fehlende Daten in einem der neurokognitiven Tests
- Ablehnung eines einzelnen Tests oder einer Testsitzung
- Präoperative Messwerte von mehr als zwei Standardabweichungen unter dem Gruppenmittel in mehr als vier Tests

3.3 Erhobene Daten

Folgende Daten wurden den Anästhesiebögen bzw. dem Krankenhausinformationssystem (KIS) entnommen:

- Alter
- Geschlecht
- Größe
- Gewicht
- Charlson Comorbidity Score
- Anästhesiedauer
- Operationsdauer (Schnitt-Naht-Zeit)
- Anzahl der Voroperationen
- Gesamtliegedauer und postoperative Liegedauer
- Liegedauer auf Intensivstation

3.4 Testbatterien

Zur Detektion von POCD werden in der Literatur eine Vielzahl von neurokognitiven Tests vorgeschlagen¹⁵. Um bei der Durchführung der Tests eine optimale Compliance der Patienten zu gewährleisten, sollte die Gesamtausführungszeit der beiden Testbatterien nicht mehr als 45 Minuten betragen. Die Tests der PPTB wurden ausgewählt, da sie sich in früheren Studien als sensitiv zur Detektion postoperativer kognitiver Dysfunktionen erwiesen haben und eine gute bis sehr gute Test-Re-Test-Reliabilität haben^{42;43}. Beide Batterien decken die gleichen kognitiven Domänen ab und liefern eine vergleichbare Anzahl von Messwerten (siehe Tabelle 1).

3.4.1 Papier- und Bleistift-Testbatterie

Für die konventionelle Papier- und Bleistift-Testbatterie wurden der Stroop Colour Word Test⁴⁴ (SCWT) in der modifizierten Version nach Houx et al.⁴⁵ Teil 1 bis 3 sowie der Visual Verbal Learning Test⁴⁶ (VVL) modifiziert nach Brand und Jolles⁴⁷ Teil 1 bis 4 verwendet. Diese Tests haben sich zur Detektion von POCD¹⁷ in vorherigen Studien als sehr geeignet erwiesen. Alle Tests lagen in deutscher Sprache vor, für den VVL standen mehrere parallele Versionen zur Verfügung.

3.4.1.1 Stroop Colour Word Test

Der SCWT besteht aus drei aufeinander aufbauenden Teilen. Vor jedem Teil wird der Patient aufgefordert, die ihm gestellte Aufgabe so schnell wie möglich aber auch so sorgfältig wie möglich zu erledigen.

Im ersten Teil wird dem Probanden ein Blatt Papier, auf dem die Wörter „rot“, „grün“, „gelb“ und „blau“ in schwarzer Farbe in zufälliger Reihenfolge gedruckt sind, vorgelegt (siehe Abbildung 2). Der Proband wird aufgefordert nach dem Startsignal die Wörter laut vorzulesen.

Im zweiten Teil sind anstatt der Wörter Kästchen in den Farben „rot“, „grün“, „gelb“ und „blau“ abgebildet (siehe Abbildung 3). Nach dem Startsignal soll die Farbe der Kästchen laut benannt werden.

Im dritten Teil sind wieder die Wörter „rot“, „grün“, „gelb“ und „blau“ gedruckt, diesmal aber in einer der Farben „rot“, „grün“, „gelb“ und „blau“ die nicht der semantischen Bedeutung des Wortes entspricht (siehe Abbildung 4). Der Proband wird aufgefordert die Farbe, in der das Wort gedruckt ist, nicht das Wort laut zu benennen.

Für jeden Teil werden die Zeit und die Anzahl der Fehler gemessen.

grün rot gelb blau blau gelb rot blau gelb grün
 gelb gelb gelb rot blau gelb rot grün blau gelb

...

Abbildung 2 SCWT Teil 1



...

Abbildung 3 SCWT Teil 2

grün rot gelb blau blau gelb rot blau gelb grün
 gelb gelb gelb rot blau gelb rot grün blau gelb

...

Abbildung 4 SCWT Teil 3

3.4.1.2 Visual Verbal Learning Test

Der VVLT besteht aus drei Lernzyklen und einem Messzyklus, der 10 Minuten nach dem letzten Lernzyklus beginnt. In den ersten drei Zyklen wird dem Probanden eine Mappe mit 15 Wörtern, jeweils eines pro Seite, gezeigt. Auf jeder dieser Seiten ist ein einfaches deutsches Substantiv wie Baum, Haus etc. mittig gedruckt. Die Seiten werden vom Versuchsleiter im Abstand von ca. einer Sekunde umgeblättert. Der Proband wird aufgefordert sich möglichst alle Wörter zu merken. Direkt im Anschluss an das Durchblättern sollen die Wörter wiedergeben werden. Nach dem dritten Lernzyklus wird eine mindestens zehnmütige Pause eingelegt. Danach soll der Proband die Wörter benennen, an die er sich dann noch erinnern kann. Es werden sowohl die Anzahl der korrekten als auch die Anzahl der falschen Antworten gemessen.

Für die POCD-Berechnung wurde der dritte Teil des SCWT (Zeit und Fehler) sowie der

vierte Teil des VVLT (Anzahl korrekte und Anzahl falsche Antworten) verwendet. Tabelle 1 gibt eine Übersicht über die getesteten kognitiven Domänen sowie die gemessenen Variablen der beiden Testbatterien.

3.4.2 Computerisierte Testbatterie

Als computerisierte Testbatterie wurde die CogState ClinicalTrials Testbatterie, entwickelt von CogState Ltd, Melbourne, Australien, verwendet. Die Testbatterie beinhaltet den Detection Task (DET), den Identification Task (IDT) und den One Card Learning Task (OCLT). Die kognitiven Domänen der verwendeten Tests sind in Tabelle 1 dargestellt. Jede Aufgabe wurde dem Probanden vor der Durchführung ausführlich durch die Software erklärt. Die Probanden erhielten vor jedem Test die Möglichkeit sich mit der Aufgabe und dem Computer in einer kleinen Einführungsrunde vertraut zu machen. Während dieser Einführungsrunde wurden keine Daten aufgezeichnet und der Proband durfte Fragen stellen. Offensichtliche Missverständnisse inhaltlicher oder exekutiver Art wurden von dem Versuchsleiter korrigiert. Zur Ausführung der Tests wurden maximal zwei Tasten benötigt. Nach Beendigung der Einführungsrunde erfolgte die eigentliche Testung durch aktives Bestätigen der Eingabetaste. Die Probanden wurden aufgefordert, die ihnen gestellte Aufgabe so schnell und sorgfältig wie möglich zu erledigen. Zu jedem Test wurden die benötigte Zeit und die Genauigkeit ((Anzahl Antworten – falsche Antworten) / Anzahl Antworten) gemessen.

3.4.2.1 Detection Task

In der Mitte des Bildschirms erscheint ein Kartenstapel dessen oberste Karte verdeckt ist. Nach dem Starten des Tests dreht sich die oberste Karte nach ca. einer Sekunde automatisch um, so dass ihr Bild sichtbar wird (siehe Abbildung 5 und Abbildung 6). Die Aufgabe des Probanden besteht darin, die „Ja“-Taste zu drücken sobald dies geschieht. Nach Drücken der Taste verschwindet die aktuelle Karte und die nächste Karte wird verdeckt auf dem Stapel sichtbar und der Ablauf beginnt von neuem.

Drückt der Proband die „Ja“-Taste bevor das Bild der Karte sichtbar wird oder reagiert gar nicht auf das Umdrehen der Karte, so ertönt ein Fehlerlaut und die nächste Karte erscheint.

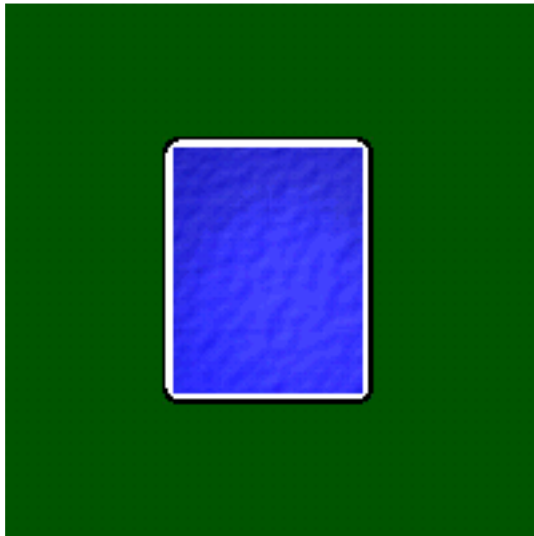


Abbildung 5 Detection Task I



Abbildung 6 Detection Task II

3.4.2.2 Identification Task

In der Mitte des Bildschirms erscheint erneut ein Kartenstapel, dessen oberste Karte verdeckt ist. Diesmal besteht die Aufgabe des Probanden darin, zügig zu entscheiden, ob die umgedrehte Karte rot ist oder nicht. Wenn die Karte ein rotes Bild zeigt, soll der Proband die „Ja“-Taste drücken, sonst die „Nein“-Taste (siehe Abbildung 7 und Abbildung 8).

Sobald eine Taste gedrückt wird oder die Präsentationszeit der Karte abgelaufen ist, verschwindet die oberste Karte und die nächste Karte wird mit dem Rücken nach oben sichtbar.



Abbildung 7 Identification Task I

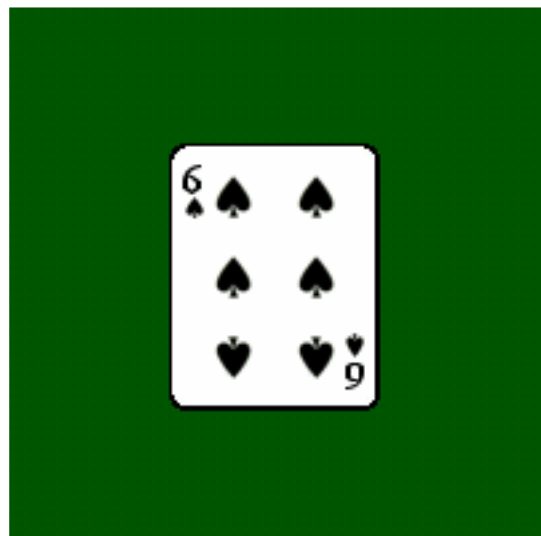


Abbildung 8 Identification Task II

3.4.2.3 One Card Learning Task

Beim OCLT finden ebenfalls die in Abbildung 7 und Abbildung 8 dargestellten Karten Verwendung. In der Mitte des Bildschirms erscheint ein Kartenstapel, dessen oberste Karte verdeckt ist. Jede Karte kann mehrfach in dem Kartenstapel vorkommen. Nach Starten des Tests dreht sich die oberste Karte für jeweils ca. eine Sekunde um und verschwindet danach. Die Aufgabe des Probanden besteht darin, für jede umgedrehte Karte zu entscheiden, ob er diese Karte schon einmal gesehen hat oder nicht. Wenn er glaubt die Karte schon einmal gesehen zu haben soll er die „Ja“-Taste drücken, sonst die „Nein“-Taste. Bei falscher oder keiner Antwort in der vorgegebenen Zeit ertönt ein Fehlerlaut und die nächste Karte erscheint.

3.4.3 Kognitive Domänen der Testbatterien

| | Test | Kognitive Domäne | Parameter |
|-------------|------------------------|---|---|
| CTB | Detection Task | Verarbeitungsgeschwindigkeit | Reaktionszeit & Genauigkeit |
| CTB | Identification Task | Visuelle Aufmerksamkeit | Reaktionszeit & Genauigkeit |
| CTB | One Card Learning Task | Visuelles Lernen, Gedächtnis | Reaktionszeit & Genauigkeit |
| PPTB | Visual Verbal Learning | Visuelles Lernen, Gedächtnis | Anzahl korrekter und falscher Antworten |
| PPTB | Stroop Colour Word | Verarbeitungsgeschwindigkeit, visuelle Aufmerksamkeit, kognitive Flexibilität | Zeit und Anzahl falscher Antworten |

Tabelle 1 Kognitive Domänen der Testbatterien

3.5 POCD-Berechnung

Zur POCD-Berechnung wurde der RCI in der von Rasmussen et al.¹⁹ beschriebenen und modifizierten Version verwendet. Bei diesem Verfahren wird für jeden einzelnen

Messwert ein so genannter Z-Wert berechnet. Dazu wird die Veränderung eines Testmerkmals wie Zeit oder Anzahl Fehler als Differenz ($\Delta x_{Patient}$) angegeben (eine negative Differenz zeigt eine Verschlechterung von der ersten zur zweiten Testsitzung an). Von dieser Differenz wird der Lerneffekt, der sich aus dem Mittelwert der Differenzen ($\Delta \bar{x}_{Kontrollen}$) zwischen erster und zweiter Testsitzung der Kontrollgruppe berechnet, abgezogen. Die erhaltene Differenz wird anschließend durch die Standardabweichung der Testdifferenz der Kontrollgruppe geteilt (siehe Formel 1). Die Differenzen der Messwerte zwischen den zwei Testzeitpunkten sind so berechnet, dass ein negatives Vorzeichen immer eine Verschlechterung der Testleistung bedeutet.

Die Z-Werte der einzelnen Tests entsprechen den Testleistungen der durch die jeweiligen neurokognitiven Tests abgedeckten Areale. Durch Summation und anschließende Division durch die Standardabweichung der Summe der Z-Werte des Kontrollkollektives kann ein kombinierter Z-Wert (Formel 2) berechnet werden. Dieser spiegelt die Gesamtleistung in der Testbatterie wieder. Ein Individuum wird als POCD positiv klassifiziert wenn:

- ein kombinierter Z-Wert $< -1,96$ vorliegt

und/oder

- mindestens zwei Einzeltests einen Z-Wert $< -1,96$ aufweisen

Bei einem Z-Wert kleiner $-1,96$ besteht nur bei 5% der Probanden die Wahrscheinlichkeit, dass dieser Wert zufällig zustande gekommen ist. Somit kann die Änderung der kognitiven Leistungsfähigkeit als signifikant angenommen werden²⁹. Bei einem Wert größer $-1,96$ wird angenommen, dass die Veränderung der kognitiven Leistungsfähigkeit in den messüblichen Schwankungen liegt und bei diesem Individuum kein POCD (NPOCD) vorliegt.

In der vorliegenden Studie wurde die POCD Klassifikation für die PPTB anhand der Testmerkmale Zeit und Fehler für den SCWT und Fehler und korrekte Antworten für den VVLT vorgenommen. Für die CTB wurden die Testmerkmale Zeit und Genauigkeit ((Anzahl Antworten – falsche Antworten) / Anzahl Antworten) für den DET, IDT und OCLT verwendet.

$$Z = \frac{\Delta x_{Patient} - \Delta \bar{x}_{Kontrollen}}{SD(\Delta x_{Kontrollen})} \quad (1)$$

$$Z_{kombiniert} = \frac{\sum Z_{Patienten}}{SD(\sum Z_{Kontrollen})} \quad (2)$$

3.6 Statistik

Alle kontinuierlichen Daten sind als Median mit 25-75% Quartile (Interquartile Range, IQR) angegeben. Kategorielle Daten werden als Prozentwerte mit dem 95%-Konfidenzintervall angegeben. Beim Vergleich zweier Gruppen wurde für Häufigkeiten der exakte Test nach Fisher verwendet, für kontinuierliche Daten der Mann-Whitney-U-Test.

Zum Vergleich der beiden Testbatterien wurde der McNemar-Test für die Häufigkeiten der Stichproben bzw. Cohens Kappa als Maß der Interrater-Reliabilität verwendet, und mit der modifizierten Version von Landis und Koch bewertet. Ein p-Wert kleiner 0,05 wurde als signifikant angenommen.

Für alle Studienvariablen wurden beschreibende statistische Werte dargestellt.

Alle statistischen Berechnungen wurden mit dem „Statistical Package of the Social Science Software“ (SPSS, Version 14.0) für Microsoft Windows durchgeführt.

4 Ergebnisse

In diese Studie wurden 36 ASA-PS III Patienten, die für einen elektiven nicht kardiochirurgischen Eingriff an der Charité - Universitätsmedizin Berlin, Campus Virchow und Campus Mitte stationär aufgenommen wurden, eingeschlossen. Als Kontrollgruppe wurden 38 Patienten, die sich keinem operativen Eingriff unterziehen mussten, rekrutiert. Die Probanden der Kontrollgruppe litten an schweren Allgemeinerkrankungen mit Leistungseinschränkungen, die einem ASA-PS III entsprechen.

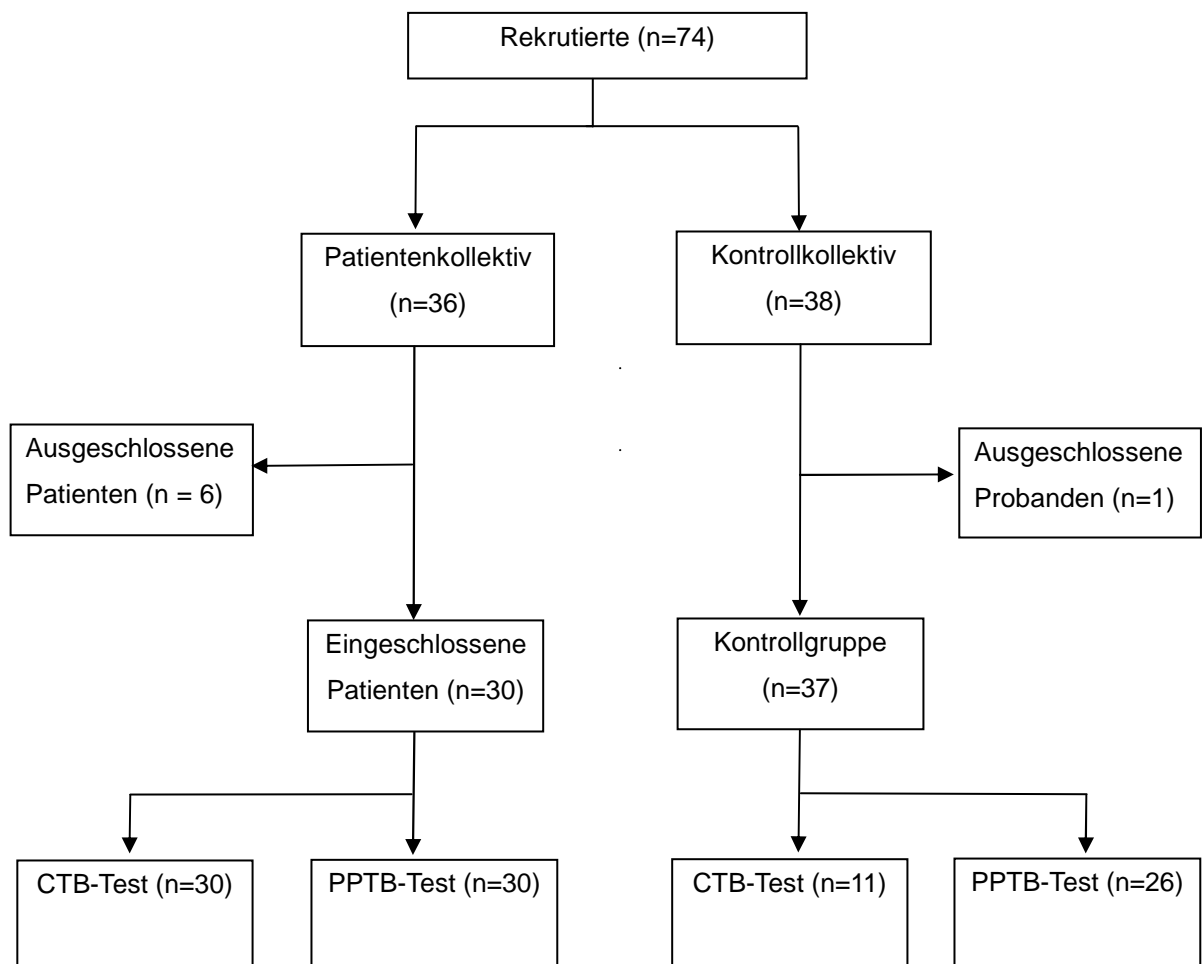


Abbildung 9 Flussdiagramm des Studienverlaufs

4.1 Ausgeschlossene Patienten

Von den 36 eingeschlossenen Patienten mussten insgesamt 6 Patienten ausgeschlossen werden. Davon drei Patienten wegen fehlender Daten, zwei Patienten wegen Ablehnung der zweiten Testung und ein Patient wegen präoperativen Messwerten, die in mindestens vier Tests zwei Standardabweichungen unter dem Gruppenmittel lagen (dies wurde als Unwillen/Unfähigkeit dem Studienprotokoll zu folgen bzw. die Tests mit Sorgfalt auszuführen interpretiert). Von den 38 rekrutierten nicht chirurgischen Kontrollpatienten musste einer wegen Ablehnung der zweiten Testsitzung ausgeschlossen werden.

Dies entspricht einer Gesamtausschlussquote von 9,5%. Ein vollständiges Flussdiagramm des Studienverlaufs ist in Abbildung 9 dargestellt.

4.2 Patienten

In die Analysen dieser Studie gingen Daten von insgesamt 67 Patienten (30 chirurgische Patienten und 37 ASA-PS III nicht chirurgische Kontrollpatienten) ein. Zu jedem dieser Patienten lag ein vollständiger Datensatz vor. Es wurden keine Werte extrapoliert.

4.3 Deskriptive Statistik

4.3.1 Basisdaten

Das Alter der 30 chirurgischen Patienten lag im Median bei 66 Jahren (IQR: 57,75 – 72,75). Das Kontrollkollektiv war im Median mit 71 Jahren (IQR: 63 – 80) statistisch signifikant älter als das Patientenkollektiv ($p=0,02$). Die Patientengruppe bestand zu 30% aus Frauen, wohingegen in die Kontrollgruppe 62% weibliche Probanden eingeschlossen wurden ($p=0,01$). Der BMI betrug in der Patientengruppe im Median 27,39 versus 27,90 in der Kontrollgruppe. Der Charlson Comorbidity Score lag im Median in beiden Gruppen bei 4,0. Für die beiden letzten Merkmale zeigte sich kein statistisch signifikanter Unterschied zwischen den Gruppen. Eine Aufstellung der Basisdaten gibt Tabelle 2 wieder.

| | Patientengruppe (n=30) | Kontrollgruppe (n=37) | p-Wert |
|--|-----------------------------------|----------------------------------|---------------|
| Alter [Jahre] | 66,00 (57,75 – 72,75) | 71,00 (63,00 – 80,00) | 0,02 |
| BMI [kg/m²] | 27,39 (24,09 – 29,37) | 27,90 (23,30 – 32,38) | 0,48 |
| Charlson Comorbidity Score | 4,00 (3,00 – 8,00) | 4,00 (2,00 – 5,00) | 0,15 |
| Geschlecht (Anteil weibliche Patienten) | 30% | 62% | 0,01 |

Tabelle 2 Basisdaten Patienten- und Kontrollkollektiv

Median (25-75% IQR)

4.3.2 Veränderung der Testergebnisse von prä- zu postoperativer Sitzung

4.3.2.1 PPTB

Die Kontrollgruppe für die PPTB zeigte in allen vier Tests im Median eine Verbesserung oder gleich bleibende Leistung von der präoperativen zur postoperativen Sitzung. Die Patientengruppe zeigte im Median eine Abnahme der Testleistung um 2,91 Sekunden für die Veränderung der Zeit im SCWT von der prä- zur postoperativen Testsitzung. Die anderen Testwerte blieben im Median unverändert.

Für den SCWT zeigte sich ein statistisch signifikanter Unterschied der Testleistungen zwischen der Kontrollgruppe und der Patientengruppe für die Veränderung der Zeit ($p=0,003$) und die Veränderung der Fehler ($p=0,016$). Für die Veränderung der Testergebnisse des VVLT zeigten sich keine statistisch signifikanten Unterschiede. Auch unterschieden sich die beiden Gruppen nicht hinsichtlich des präoperativen MMSE-Wertes.

Die Messergebnisse der Tests der PPTB sowie der MMSE-Wert für das Patientenkollektiv und das Kontrollkollektiv sind in Tabelle 3 dargestellt.

| | Patientengruppe (n=30) | PPTB Kontrollgruppe (n=26) | p-Werte |
|---|---------------------------|-------------------------------|---------|
| SCWT ΔZeit [sec] | -2,91 (-13,01 – 4,76) | 3,24 (0,50– 8,87) | 0,003 |
| SCWT ΔFehler | 0,00 (0,00 – 0,00) | 0,50 (0,00 – 2,00) | 0,016 |
| VVLT ΔFehler | 0,00 (-1,00 – 0,00) | 0,00 (0,00 – 0,00) | 0,511 |
| VVLT ΔRichtig | 0,00 (-1,25 – 2,25) | 1,00 (-1,00 – 1,25) | 0,817 |
| MMSE | 30,00 (29,00 – 30,00) | 29,00 (28,00 – 30,00) | 0,289 |

Tabelle 3 Messwerte PPTB Patienten- und Kontrollkollektiv

Median (25-75% IQR)

4.3.2.2 CTB

Die Kontrollgruppe der CTB zeigte im Median für die Veränderung der benötigten Zeit im DET sowie für die Veränderung der Genauigkeit im OCLT eine leichte Verschlechterung. Die Patientengruppe hingegen zeigte im Median eine Verschlechterung für die Veränderung der benötigten Zeit im DET und im IDT sowie eine Verschlechterung der Genauigkeit im OCLT.

Zwischen den Gruppen zeigte sich für die Veränderungen von prä- zu postoperativer Sitzung kein statistischer Unterschied. Der MMSE-Wert unterschied sich auch zwischen diesen beiden Gruppen nicht statistisch signifikant.

Die Messergebnisse der Tests der CTB sowie der MMSE-Wert für das Patientenkollektiv und das Kontrollkollektiv sind in Tabelle 4 dargestellt.

| | Patientengruppe (n=30) | CTB Kontrollgruppe (n=11) | p-Werte |
|---|---------------------------|------------------------------|---------|
| DET ΔZeit [ms] | -20,40 (-100,63 – 32,41) | -8,65 (-76,25 – 55,71) | 0,375 |
| IDT ΔZeit [ms] | -35,44 (-66,95 – 21,20) | 0,00 (-61,15 – 39,30) | 0,552 |

| | Patientengruppe (n=30) | CTB Kontrollgruppe (n=11) | p-Werte |
|--|---------------------------|------------------------------|---------|
| OCLT ΔZeit [ms] | 103,28 (-33,59 – 469,15) | 93,48 (-210,01 – 379,99) | 0,513 |
| DET ΔGenauigkeit | 0,00 (-0,06 – 0,10) | 0,05 (0,00 – 0,06) | 0,375 |
| IDT ΔGenauigkeit | 0,01 (-0,06 – 0,22) | 0,06 (-0,06 – 0,10) | 0,988 |
| OCLT ΔGenauigkeit | -0,08 (-0,22 – 0,03) | -0,01 (-0,17 – 0,04) | 0,391 |
| MMSE | 30,00 (29,00 – 30,00) | 29,00 (27,00 – 30,00) | 0,215 |

Tabelle 4 Messwerte CTB Patienten- und Kontrollkollektiv

Median (25-75% IQR)

4.4 POCD-Klassifikationen und Interrater-Reliabilität

Anhand der Veränderung der Testergebnisse zwischen den beiden Testsitzungen wurde, unabhängig von einander, für die CTB und die PPTB mittels des RCI eine dichotome Einteilung in POCD-positiv und POCD-negativ vorgenommen.

Die PPTB detektierte bei 9 von 30 Patienten POCD (30% [95%-CI: 16,7% - 47,9%]). Die CTB detektierte bei 3 der 30 Patienten POCD (10% [95%-CI: 3,5% - 25,6%]). Das 95%-Konfidenzintervall für den Unterschied zwischen den beiden Gruppen beträgt 3,9% - 36,5%, die Gruppen unterscheiden sich statistisch signifikant voneinander ($p = 0,03$) (Tabelle 5). Die drei mittels CTB detektierten Patienten wurden alle ebenfalls mit der PPTB detektiert, die PPTB klassifizierte sechs Patienten zusätzlich als POCD-positiv.

| | | CTB | | Gesamt | p-Wert |
|--------|-------|-------|------|--------|--------|
| | | NPOCD | POCD | | |
| PPTB | NPOCD | 21 | 0 | 21 | |
| | POCD | 6 | 3 | 9 | |
| Gesamt | | 27 | 3 | 30 | 0,03 |

Tabelle 5 Kreuztabelle Klassifikation nach CTB und PPTB

Die Interrater-Reliabilität ist ein Maß für die Übereinstimmung zweier unabhängiger Beobachter oder Messinstrumente (Rater) an einem Merkmal (hier die dichotome Einteilung in POCD und NPOCD mittels der beiden Testbatterien). Anders ausgedrückt beschreibt sie die Unabhängigkeit zweier Beobachter (hier Testbatterien) an einem Merkmal.

Zur Abschätzung der Übereinstimmung dient das Cohens Kappa (κ), welches bei vollkommener Übereinstimmung einen Wert von 1 annimmt und bei einer Übereinstimmung, die dem Zufall entspricht, den Wert 0. Landis und Koch⁴⁸ beschrieben folgende Interpretationen des Kappa-Wertes:

| | Interpretation (Englisch) | Interpretation (Deutsch) |
|-------------------------------------|---------------------------|-------------------------------------|
| $\kappa < 0$ | poor agreement | schlechte Übereinstimmung |
| $\kappa \geq 0$ & $\kappa \leq 0,2$ | slight agreement | geringfügige Übereinstimmung |
| $\kappa > 0,2$ & $\kappa \leq 0,4$ | fair agreement | anständige Übereinstimmung |
| $\kappa > 0,4$ & $\kappa \leq 0,6$ | moderate agreement | angemessene Übereinstimmung |
| $\kappa > 0,6$ & $\kappa \leq 0,8$ | substantial agreement | beachtliche Übereinstimmung |
| $\kappa > 0,8$ & $\kappa \leq 1$ | almost perfect agreement | beinahe vollkommene Übereinstimmung |

Tabelle 6 Interpretation des Cohens Kappa nach Landis und Koch

Für die beiden Testbatterien ergibt sich ein κ von 0,41, was einer angemessenen Übereinstimmung (moderate agreement) entspricht.

Wenn man annimmt, dass es sich hier bei der PPTB um den Goldstandard handelt, so ergeben sich für die diagnostischen Tests die in Tabelle 7 dargestellten Werte.

| | Wert |
|---|----------------------------|
| Sensitivität | 0,33 (95%-CI: 0,09 – 0,69) |
| Spezifität | 1,00 (95%-CI: 0,81 – 1,00) |
| Positiver prädiktiver Wert (PPW) | 1,00 (95%-CI: 0,31 – 1,00) |
| Negativer prädiktiver Wert (NPW) | 0,77 (95%-CI: 0,57 – 0,90) |

Tabelle 7 Diagnostische Tests

4.5 Basisdaten postoperativ nach Testbatterie

4.5.1 Vergleich POCD versus NPOCD entsprechend der CTB

In Tabelle 8 sind die Basisdaten der drei als POCD-positiv klassifizierten Patienten gegen die 27 als POCD-negativ klassifizierten Patienten, nach der Einteilung der CTB, dargestellt. Die Gruppe der Patienten mit POCD war im Median älter, hatte eine längere Schnitt-Naht-Zeit, eine längere Anästhesiedauer, eine längere Gesamtliegedauer und eine längere postoperative Liegedauer. Des Weiteren hatten sie einen höheren Charlson Comorbidity Score. Allerdings ließ sich für keines der Merkmale eine statistische Signifikanz nachweisen.

| | n=27 | n=3 | p-Wert |
|-------------------------------|-----------------------|-----------------------|--------|
| Alter [Jahre] | 66,00 (58,00 – 72,00) | 71,00 (47,00 – 76,00) | 0,795 |
| BMI [kg/m²] | 27,43 (23,03 – 29,00) | 27,34 (24,44 – 35,26) | 0,744 |
| MMSE | 30,00 (29,00 – 30,00) | 29,00 (25,00 – 30,00) | 0,350 |

| | n=27 | n=3 | p-Wert |
|---|--------------------------|--------------------------|--------|
| Postoperative Messung [Tage] | 8,00 (7,00 – 10,00) | 8,00 (7,00 – 11,00) | 0,795 |
| Schnitt-Naht-Dauer [min] | 223,00 (152,00 – 286,00) | 296,00 (192,00 – 352,00) | 0,315 |
| Anästhesiedauer [min] | 320,00 (225,00 – 400,00) | 435,00 (275,00 – 435,00) | 0,226 |
| Anzahl Vor-OPs | 1,00 (1,00 – 2,00) | 1,00 (1,00 – 1,00) | 0,509 |
| Liegezeit [Tage] | 17,00 (11,00 – 24,00) | 20,00 (13,00 – 28,00) | 0,554 |
| Postoperative Liegezeit [Tage] | 12,00 (9,00 – 17,00) | 17,00 (10,00 – 27,00) | 0,387 |
| Liegezeit Intensivstation [Tage] | 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 |

Tabelle 8 Basisdaten postoperativ - POCD-Einteilung nach CTB

Median (25-75% IQR)

4.5.2 Vergleich POCD versus NPOCD entsprechend der PPTB

In Tabelle 9 sind die Basisdaten der neun als POCD-positiv klassifizierten Patienten gegen die 21 als POCD-negativ klassifizierten Patienten, nach der Einteilung der PPTB, dargestellt. Für keines der betrachteten Merkmale konnte eine statistische Signifikanz nachgewiesen werden.

| | n=21 | n=9 | p-Wert |
|-------------------------------|-----------------------|-----------------------|--------|
| Alter [Jahre] | 66,00 (58,00 – 75,00) | 67,00 (54,50 – 71,50) | 0,824 |
| BMI [kg/m²] | 27,33 (23,80 – 29,74) | 27,42 (23,38 – 29,46) | 0,965 |

| | n=21 | n=9 | p-Wert |
|---|--------------------------|--------------------------|--------|
| MMSE | 30,00 (29,00 – 30,00) | 29,00 (28,50 – 30,00) | 0,137 |
| Postoperative Messung [Tage] | 8,00 (7,00 – 10,00) | 8,00 (7,00 – 10,50) | 0,929 |
| Schnitt-Naht-Dauer [min] | 223,00 (152,00 – 310,50) | 249,00 (185,50 – 285,50) | 1,000 |
| Anästhesiedauer [min] | 320,00 (240,00 – 392,50) | 320,00 (237,50 – 432,50) | 0,722 |
| Anzahl Vor-OPs | 1,00 (1,00 – 2,50) | 1,00 (1,00 – 1,50) | 0,476 |
| Liegezeit [Tage] | 19,00 (12,00 – 25,00) | 14,00 (10,50 – 22,50) | 0,504 |
| Postoperative Liegezeit [Tage] | 13,00 (9,50 – 22,00) | 12,00 (9,50 – 15,50) | 0,533 |
| Liegezeit Intensivstation [Tage] | 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 |

Tabelle 9 Basisdaten postoperativ - POCD-Einteilung nach PPTB

Median (25-75% IQR)

4.5.3 Vergleich POCD nach CTB versus POCD nach PPTB

Tabelle 10 vergleicht die postoperativen Basisdaten der drei von sowohl der CTB als auch der PPTB als POCD-positiv klassifizierten Patienten mit den Daten der sechs Patienten, die nur durch die PPTB als POCD-positiv klassifiziert wurden (siehe hierzu auch Tabelle 5). Die durch die CTB und die PPTB als POCD-positiv klassifizierten Patienten waren im Median älter, hatten eine längere OP- und Anästhesiedauer und einen längeren stationären und postoperativen Aufenthalt. Es zeigten sich jedoch keine statistisch signifikanten Unterschiede zwischen den Gruppen.

| | n=6 [†] | n=3 [‡] | p-Wert |
|--|--------------------------|--------------------------|--------|
| Alter [Jahre] | 65,50 (56,25 – 70,50) | 71,00 (47,00 – 76,00) | 0,714 |
| BMI [kg/m²] | 27,71 (21,87 – 28,77) | 27,34 (24,44 – 35,26) | 0,905 |
| MMSE | 29,00 (28,75 – 30,00) | 29,00 (25,00 – 30,00) | 0,714 |
| Postoperative Messung [Tage] | 7,50 (7,00 – 10,25) | 8,00 (7,00 – 11,00) | 0,714 |
| Schnitt-Naht-Dauer [min] | 220,00 (163,75 – 271,25) | 296,00 (192,00 – 352,00) | 0,167 |
| Anästhesiedauer [min] | 287,50 (217,50 – 422,50) | 435,00 (275,00 – 435,00) | 0,167 |
| Anzahl Vor-OPs | 1,00 (0,75 – 2,25) | 1,00 (1,00 – 1,00) | 0,714 |
| Liegezeit [Tage] | 12,50 (9,75 – 21,25) | 20,00 (13,00 – 28,00) | 0,262 |
| Postoperative Liegezeit [Tage] | 11,00 (8,50 – 12,50) | 17,00 (10,00 – 27,00) | 0,167 |
| Liegezeit Intensiv- station [Tagen] | 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 |

Tabelle 10 Basisdaten postoperativ - POCD positiv nach CTB versus POCD positiv nur nach PPTB

[†] POCD positiv nach PPTB aber nicht nach CTB

[‡] POCD positiv nach CTB und nach PPTB

Median (25-75% IQR)

5 Diskussion

5.1 Inzidenzen

Die Inzidenz postoperativer kognitiver Dysfunktionen, gemessen mit der konventionellen PPTB, lag in dieser Studie bei 30% (9 der 30 Patienten). Mit der CTB wurden in der gleichen Gruppe nur bei 10% (3 der 30 Patienten) POCD festgestellt. Die gemessenen Inzidenzen sind ähnlich zu den bereits publizierten Inzidenzen für POCD in nicht kardiochirurgischen Patientenkollektiven. Für frühe postoperative kognitive Dysfunktionen bei nicht kardiochirurgischen Patienten wird in der Literatur eine POCD-Inzidenz zwischen 14,4% in der Studie von Linstedt et al.¹² und 40% in der Studie von Iohom et al.¹³ angegeben. Rasmussen et al.^{21;49}, Abildstrom et al.¹⁸ und Moller et al.¹⁷ konnten für frühe POCD eine Inzidenz von ca. 20% bis 25% in ihren Studien zeigen. Diese Studien verwendeten eine konventionelle Papier- und Bleistift-Testbatterie, die unter anderem aus dem Visual Verbal Learning Test, dem Concept Shifting Test, dem Stroop Colour Word Test und dem Letter Digit Coding Test bestand und der in dieser Studie verwendeten Testbatterie ähnlich ist.

Silbert und Kollegen³⁶ verglichen in einer Arbeit die POCD-Inzidenzen gemessen mit einer PPTB und einer CTB. Sie beschrieben für die von ihnen verwendete PPTB eine POCD-Inzidenz von 32% und für die CTB von 42%. Die in der Arbeit von Silbert et al. benutzte CTB war der in dieser Studie verwendeten CTB sehr ähnlich (ebenfalls von Cogstate Ltd, Australien), die verwendete konventionelle Testbatterie unterscheidet sich aber deutlich von der in dieser Studie verwendeten. Die Unterschiede in den mittels CTB gemessenen POCD-Inzidenzen (42% bei Silbert et al. versus 10% in der vorliegenden Studie) könnten darin begründet sein, dass Silbert et al. ausschließlich kardiochirurgische Patienten, die sich einer Koronaren-Bypass-Operation unterziehen mussten, eingeschlossen haben, die vorliegende Studie aber ausschließlich nicht kardiochirurgische Patienten eingeschlossen hat. In der Literatur ist bekannt, dass kardiochirurgische Operationen mit einer erhöhten Inzidenz von POCD einhergehen^{3;50;51}. Für die Berechnung von POCD verwendete Silbert, wie die vorliegende Studie, einen RCI, legte aber einen Z-Wert von 1,65 als Grenze für die Einteilung POCD ja/nein fest. Dem Grenzwert kommt bei der Einteilung der kognitiven Veränderung in eine dichotome Variable eine entscheidende Bedeutung zu. Ein

niedrigerer Grenzwert hat zur Folge, dass Patienten mit einer milderer Veränderung der kognitiven Leistungsfähigkeit als POCD-positiv erkannt werden. Dies sowie die unterschiedlichen Studienpopulationen könnten zu den unterschiedlichen Ergebnissen geführt haben. Newman et al.¹⁵ kommen in ihrer Übersichtsarbeit zu dem Schluss, dass die große Anzahl der in der Literatur zur Detektion von POCD verwendeten neuropsychologischen Tests einen Vergleich der Studien mit unterschiedlichen Testbatterien äußerst schwer macht. Dies könnte die unterschiedlichen mittels PPTB gemessenen POCD-Inzidenzen erklären.

Einen weiteren Einfluss auf die Inzidenzen könnte die Anzahl der Patienten, die eine zweite Testung ablehnten, haben. Insgesamt mussten in dieser Studie 2 von 36 (5,5%) Patienten ausgeschlossen werden weil sie eine zweite Testung ablehnten. In der Literatur wird diskutiert, ob nicht gerade eine starke Beeinträchtigung durch POCD dazu führt, dass ein Patient nicht Willens oder in der Lage ist, die zweite Testung durchzuführen und sie deshalb ablehnt¹⁹. Die in der Literatur beschriebenen Ablehnungsraten liegen zwischen 3%⁵² und 28%⁵³.

Die Messung des präoperativen Ausgangswertes sollte unter optimalen Bedingungen stattfinden, da ansonsten eine Abnahme der kognitiven Leistungsfähigkeit nicht erkannt wird. Da Angst und Müdigkeit dazu führen, dass Patienten keine optimalen Testwerte erzielen empfehlen Rasmussen et al.¹⁹ die präoperative Testung ein bis zwei Wochen vor dem geplanten Operationstermin durchzuführen. Da die allermeisten Patienten aus organisatorischen Gründen einen Tag vor der Operation in den Anästhesieambulanzen gesehen werden, und dort erst der Einschluss erfolgte, wurden die präoperativen Testsitzungen in der vorliegenden Studie erst am Tag vor der Operation durchgeführt. Dies könnte ebenfalls die gemessenen Inzidenzen beeinflusst haben.

5.2 Diagnostische Tests

Die PPTB, als traditionell verwendete Testbatterie, stellt den bisher etablierten Goldstandard dar, so dass sich für die CTB eine Sensitivität von 33% ergibt. Die Sensitivität beschreibt den Anteil der als positiv klassifizierten Patienten zu den tatsächlich positiven Patienten. Dies bedeutet, dass zwei Drittel der Patienten mit POCD durch die CTB nicht erkannt wurden. Dagegen wäre die Spezifität, also das Verhältnis der als negativ klassifizierten Patienten zu den tatsächlich negativen Patienten 100%. In der Studie von Silbert et al.³⁶ lagen, mit der PPTB als Goldstandard, die Sensitivität bei 100% und die Spezifität bei 85%. Diese Unterschiede könnten zum

einen durch die unterschiedlichen Patientenkollektive, zum anderen durch die Verwendung unterschiedlicher konventioneller Testbatterien bedingt sein.

Wie Arrowsmith et al.²⁰ ausführten, kann es bei neurokognitiven Testungen keinen Goldstandard geben, gegen den verglichen werden kann, da sich POCD nur über die Abnahme der kognitiven Leistungsfähigkeit zwischen präoperativer und postoperativer Sitzung definiert. Es gibt kein organisches oder klinisches Korrelat, gegen das die Ergebnisse der kognitiven Tests verglichen werden könnten. Deshalb müssen diagnostische Tests beim Vergleich zweier kognitiver Testbatterien mit Vorsicht interpretiert werden.

Die ermittelten POCD-Inzidenzen (verbundene Stichproben) unterscheiden sich statistisch signifikant voneinander ($p=0,03$). Die Interrater-Reliabilität der beiden Testbatterien war mit einem Cohens Kappa von 0,41 niedrig. Im Allgemeinen werden Kappa-Werte zwischen 0,4 und 0,6 als noch akzeptabel angenommen. Dieses Ergebnis zeigt, dass zwei kognitive Testbatterien, selbst wenn sie die gleichen kognitiven Domänen (siehe Tabelle 1) abdecken und die gleiche Berechnungsmethode verwendet wurde, nicht automatisch zu derselben Klassifikation kommen. Da die äußeren Einflussfaktoren wie Alter, Bildung, Uhrzeit der Testung, räumliche Gegebenheiten und die Person, die den Test durchführte, für beide Testbatterien gleich waren, kommen sie als Gründe für die unterschiedlichen Inzidenzen nicht in Betracht. Ein Grund für die niedrige Übereinstimmung könnte der unterschiedliche Grad der Standardisierung der Testausführung gewesen sein. So ordnen Wild et al.³⁷ in ihrer Übersichtsarbeit die von Cogstate Ltd. entwickelte CTB als eine vom Patienten selbstständig ausführbare Batterie ein. Die Anweisungen werden vor jeder Aufgabe standardisiert am Bildschirm dargestellt. Die Testdurchführung der CTB verläuft somit normierter als dies für eine PPTB zu erreichen wäre. Zum Beispiel ist bei dem VVLT ein manuelles Umblättern der 15 Wörter notwendig. Dabei soll dem Probanden laut Studienprotokoll jedes Wort für ca. eine Sekunde gezeigt werden. Die Karten in den Tests der CTB hingegen werden den Probanden in immer den gleichen Zeitabständen präsentiert. Auch die von der CTB automatisiert erfassten Messwerte, im Gegensatz zu den von Hand abgestoppten Messwerten bei der PPTB, könnten zu valideren Ergebnissen beigetragen haben. In den multizentrischen Studien von Moller et al.¹⁷ und Johnson et al.²⁵, die beide eine konventionelle Testbatterie eingesetzt haben, erwies sich das durchführende Zentrum als starker Einflussfaktor für die gemessene Inzidenz von POCD. Dies könnte darauf hin deuten, dass Aspekte der Testaus- und Testdurchführung nicht unerheblich zu

unterschiedlichen Inzidenzen beitragen.

Ein weiterer Grund für die niedrige Interrater-Reliabilität könnte in der dichotomen Einteilung der Patienten in POCD ja/nein gelegen haben. Für beide Testbatterien wurde die gleiche Berechnungsmethode mit dem gleichen Grenzwert verwendet. Eine Einteilung stetiger Variablen in dichotome Werte geht aber mit einem Verlust an Informationen einher⁵⁴. Individuen, deren kognitive Veränderung nahe des Grenzwertes liegt, werden unter Umständen unterschiedlich klassifiziert anstatt einer ähnlichen Gruppe zugeordnet zu werden. Welche Veränderung der kognitiven Leistungsfähigkeit als neu entstandenes Defizit betrachtet wird ist Definitionssache³⁰. Die strikte Einteilung anhand eines vorher definierten Wertes könnte somit zu einer unvollständigen Übereinstimmung der Testbatterien beigetragen haben.

5.2.1 Risikofaktoren

Bekannte Risikofaktoren wie Alter^{17;23;55}, Schnitt-Naht-Zeit bzw. Anästhesiedauer^{17;25} und Gesamtliegedauer²³ konnten entgegen der Literatur weder für die PPTB noch für die CTB nachgewiesen werden. Ein Grund hierfür könnte die geringe Fallzahl, sowie bei der CTB die niedrige Inzidenz, die es schwierig macht einen statistisch signifikanten Unterschied aufzuzeigen, gewesen sein. Für die untersuchten Parameter postoperative Liegedauer und Liegezeit Intensivstation konnte ebenfalls für keine der beiden Testbatterien ein statistisch signifikanter Unterschied zwischen den als POCD-positiv und POCD-negativ klassifizierten Patienten gezeigt werden. Dies ist wahrscheinlich ebenfalls der geringen Fallzahl geschuldet.

5.3 Limitationen dieser Studie

5.3.1 Bias

5.3.1.1 Selektionsbias

In das Patientenkollektiv der Studie wurden 36 ASA-PS III Patienten, die sich einem elektiven nicht kardiochirurgischen Eingriff unterziehen mussten, eingeschlossen. Der Einschluss fand in der anästhesiologischen Ambulanz statt, in der Patienten aller großen chirurgischen Fachrichtungen gesehen werden. Es wurde weder zu speziellen

Zeitpunkten eingeschlossen, noch wurden Selektionen durch die einschließenden Personen vorgenommen. Die eingeschlossene Population entspricht somit im Wesentlichen einem gemischten Patientenkollektiv an einer Universitätsklinik. Die Einschränkung auf ASA-PS III Patienten wurde vorgenommen, da es Hinweise darauf gibt, dass ein hoher ASA-PS mit einer erhöhten Inzidenz von POCD einhergeht^{23;25} und dies einen Vergleich der beiden Testbatterien vereinfachen würde.

In der Kontrollgruppe fanden sich signifikant mehr Frauen ($p=0,02$). Dieser Selektionsbias könnte zu einer Verzerrung der gemessenen Inzidenzen geführt haben, wenn man annimmt, dass Frauen einen anderen Lerneffekt aufweisen wie Männer. In der Literatur findet sich hierfür aber kein Hinweis. Auch zeigte sich in den großen Studien kein Unterschied in der POCD-Inzidenz zwischen den Geschlechtern¹⁷.

Die Kontrollpatienten waren signifikant älter als das Patientenkollektiv (im Median 71 Jahre versus 66 Jahre $p=0,01$). Dieser Selektionsbias könnte ebenfalls die Ergebnisse der Patientengruppe verfälscht haben. Unter der Annahme, dass in einer Gruppe mit höherem Alter die Lernkurve zwischen zwei Sitzung niedriger ausfällt als in einer jüngeren Gruppe, würde dies bedeuten, dass durch den zu niedrig angesetzten Lerneffekt für die Patientengruppe die kognitive Abnahme nach der Operation unterschätzt wurde. Da sich die ermittelten Inzidenzen allerdings mit denen der Literatur^{12;13;17;21;49} decken, erscheint dies als eher unwahrscheinlich.

Das Studienprotokoll sah bei Einschluss keine Erhebung des Bildungsstandes vor, es ist aber bekannt, dass niedrige Schulbildung ein Risikofaktor für die Entstehung von POCD ist¹⁷, dies könnte die Ergebnisse ebenfalls verzerrt haben.

5.3.1.2 Informationsbias

Ein Erwartungseffekt kann dann auftreten, wenn der Untersuchungsleiter durch seine Erwartungshaltung oder durch Teilergebnisse den Test beeinflusst. So könnte z. B. ein Versuchsleiter geneigt gewesen sein, bei Patienten, die sehr schlecht im SCWT abgeschnitten haben, im VVLT die Seiten etwas schneller umzublättern, da er mit einem schlechten Gesamtergebnis gerechnet hat. Dies könnte ein Vorteil der computerisierten Testbatterie sein, da hier jeder Patient immer die gleichen Ausgangschancen hat und es bei der Vielzahl der Stimuli für einen Versuchsleiter extrem schwer sein dürfte, sich während der Laufzeit ein Bild von der Testleistung zu machen.

5.3.2 Störfaktoren

Da es sich in der vorliegenden Studie nicht um eine randomisierte Studie handelt, könnten Störfaktoren dafür verantwortlich gewesen sein, bekannte Risikofaktoren wie Alter und OP-Dauer nicht nachgewiesen zu haben. Durch Ausschluss von Studienteilnehmern mit bekannten psychiatrischen Erkrankungen, Operationen am zentralen Nervensystem oder am Herzen, bekannten Medikamenten-, Alkohol- oder Drogenabusus, mangelnden Deutschkenntnissen sowie Unwillen oder Unfähigkeit dem Studienprotokoll zu folgen, wurde versucht die Störfaktoren möglichst zu minimieren.

Die Definition von Ausschlusskriterien kann aber zur Folge haben, dass das eingeschlossene Kollektiv nicht mehr repräsentativ für alle an einer Klinik behandelten Patienten ist.²⁶ Dies ist aber keine Limitation des Studiendesigns dieser Studie sondern gilt allgemein.

Der MMSE-Test wurde in dieser Studie, wie in der Literatur empfohlen³, dazu verwendet Patienten mit einer vorbestehenden kognitiven Einschränkung auszuschließen, da bei diesen Patienten eine spätere Interpretation der Ergebnisse erschwert ist. Dennoch kann nicht ausgeschlossen werden, dass Patienten mit einer leichten, nicht durch den MMSE-Test erkannten, kognitiven Einschränkung eingeschlossen wurden.

5.3.3 Test-Auswahl

In der vorliegenden Studie wurden die beiden Testbatterien vorwiegend unter den Gesichtspunkten der abgedeckten kognitiven Domänen und der Gesamtausführungsdauer ausgesucht. Hier wurde eine gute Übereinstimmung erreicht.

Zur Detektion von POCD wurden in der Literatur für nicht kardiochirurgische Patientenkollektive über 70 verschiedene Tests¹⁵, in kardiochirurgischen Patientenkollektiven sogar über 350 verschiedene Tests⁵⁶ verwendet. Auch sind die Diagnosekriterien weder an spezielle kognitive Domänen gekoppelt noch ist eine Mindestabnahme der kognitiven Leistungsfähigkeit gefordert. Die gemessene POCD-Inzidenz wird wesentlich durch die ausgewählten Tests und die Zusammenstellung der Testbatterie beeinflusst¹. Es ist nicht auszuschließen, dass eine andere Zusammenstellung der Testbatterie zu anderen Resultaten geführt hätte. Dies ist aber ebenfalls keine Limitation dieser Studie sondern ein allgemeines Problem der POCD-Forschung. So konnten Lewis et al.⁵⁷ zeigen, dass in kardiochirurgischen Patientenkollektiven die Anzahl der Tests in einer Testbatterie Einfluss auf die Inzidenz

von POCD hat: Mit zwei Tests in der Testbatterie betrug die POCD-Inzidenz 13,3% mit sieben Tests 49,4%.

Dem Problem kann etwa dadurch begegnet werden, dass nicht nur jeder einzelne Test für sich betrachtet wird, sondern, wie auch in der vorliegenden Studie geschehen, eine Betrachtung in Form eines kombinierten Wertes für die Testleistung insgesamt durchgeführt wird um einen Gesamteindruck der Veränderung der kognitiven Leistungsfähigkeit zu erhalten.

5.3.4 Fallzahl

Die Anzahl der in diese Studie eingeschlossenen Patienten ist mit 36 relativ gering. Die Studie war als Pilotprojekt mit 38 Patienten geplant. Allerdings wurden nur 30 Patienten sowohl mit der CTB als auch mit der PPTB getestet. Wegen fehlender Daten, Ablehnung der zweiten Testung oder Unwillen dem Studienprotokoll zu folgen, mussten 6 Patienten ausgeschlossen werden. Mit einer Gesamtausschlussquote von 9,5% und einer Studienabbrucher-Quote von 5,5% im Patientenkollektiv liegt diese Studie im unteren Bereich der in anderen Studien berichteten Ausschlussquoten¹⁵. Wie von Rasmussen et al.¹⁹ empfohlen, wurden fehlende Messwerte nicht extrapoliert (mit Mittelwerten, schlechtestem gemessenem Testwert o. ä.). Dies führte zu einer Verringerung der Fallzahl, trägt aber zu einer höheren Datenqualität bei.

Nichtsdestoweniger konnte mit dieser Fallzahl ein statistisch signifikanter Unterschied zwischen den beiden Testbatterien gezeigt werden. Die Power dieser Studie liegt mit einem $\alpha=5\%$ (einseitig) bei 53%.

Die geringe Anzahl an POCD-positiven Patienten in der CTB- als auch in der PPTB-Messung macht es schwierig, innerhalb der beiden Testbatterien für die untersuchten Merkmale einen Unterschied aufzuzeigen und schränkt die statistische Aussagekraft ein. Mit einer größeren Fallzahl wäre es eventuell möglich gewesen, mit beiden Testbatterien, die aus der Literatur bekannten Risikofaktoren statistisch nachzuweisen.

5.3.5 Berechnungsmethode

Es gibt keine anerkannte Methode zur Berechnung der Inzidenz und Schwere von POCD^{11;58}. Methoden, die Mittelwerte analysieren, sind zur Ermittlung von Inzidenzen ungeeignet und lassen keine Aussagen über das Vorhandensein von POCD in einzelnen Individuen zu. RCIs hingegen lassen Aussagen zu einzelnen Patienten zu

und erlauben die Berechnung eines kombinierten Testwertes, der eine Aussage über die gesamte gemessene kognitive Leistungsfähigkeit macht. RCIs wurden in verschiedenen Studien^{9;17;25;35} zur Detektion von POCD angewendet bzw. untersucht und haben sich dort als geeignetes Instrument erwiesen, um zu entscheiden, ob es zwischen den beiden Testzeitpunkten zu einer signifikanten Änderung der kognitiven Leistungsfähigkeit gekommen ist²⁹. Bis zu einem gewissen Grad kann jedoch jede Berechnungsmethode in Frage gestellt werden³².

5.4 Interpretation und Ausblick

In dieser Arbeit konnte gezeigt werden, dass sowohl computerisierte Testbatterien als auch konventionelle Papier- und Bleistift-Testbatterien kognitive Veränderungen nach einer Operation nachweisen können. Die gemessenen Inzidenzen in der gegebenen Studienpopulation sind mit denen in der Literatur vergleichbar. Obwohl bei der Auswahl der beiden Testbatterien darauf geachtet wurde, dass beide die gleichen kognitiven Domänen abdecken, zeigte sich eine niedrige Interrater-Reliabilität. Dies könnte auf Unterschiede in den Methoden der beiden Testbatterien, aber auch auf die zu kleine Fallzahl zurückzuführen sein. Denkbar ist auch, dass die computerisierte Testbatterie durch ihre Vorteile in der Standardisierung und Testauswertung weniger Spielraum für Interpretationen der Versuchsleiter ließ, dadurch die Schwankungen in den Testsitzungen reduzierte und somit zu einer anderen Klassifikation kam als die konventionelle Testbatterie.

Das Auftreten postoperativer kognitiver Dysfunktionen nach nicht kardiochirurgischen Eingriffen bei einem nicht unerheblichen Teil der älteren Patienten gilt in der Literatur als gesichert¹⁵. Bis heute konnten aber keine Strategien entwickelt werden, um die Inzidenz zu senken. Auch werden Patienten in der Routineversorgung nicht auf postoperative kognitive Dysfunktionen hin untersucht, obwohl es mittlerweile gute Hinweise darauf gibt, dass POCD sowohl eine sozioökonomische Relevanz²² hat als auch mit einer erhöhten Mortalität²³ einhergeht.

Gerade für die Versorgung in der klinischen Routine könnten sich computerisierte Testbatterien als vorteilhaft erweisen, da sie eine große Anzahl kognitiver Fähigkeiten testen können, und Vorteile in der Testdurchführung wie Standardisierung, niedrige Lerneffekte sowie kaum vorhandene Boden- und Deckeneffekte haben. Des Weiteren bieten sie die Möglichkeit eine große Anzahl von Probanden in geringerer Zeit zu testen, benötigen nicht die ständige Anwesenheit eines neuropsychologisch geschulten

Arztes, bieten Unterstützung bei der elektronischen Auswertung und vereinfachen die Verwaltung der Testsitzungen^{36;37;39}.

Die von uns verwendete Cogstate Testbatterie musste auf einem Laptop installiert werden. Es wäre aber auch denkbar, Testbatterien zu entwickeln, die über das Internet ausgeführt werden können. So könnten Patienten zu einem späteren Zeitpunkt von zu Hause aus überprüfen, ob noch ein kognitives Defizit besteht. Wobei aber aus ethischen Gesichtspunkten bedacht werden muss, dass dann kein geschultes Personal zur Verfügung steht, welches dem Patienten hilft seine Testergebnisse richtig zu interpretieren.

Als Nachteile von computerisierten Testbatterien werden in der Literatur schlecht konzipierte Benutzeroberflächen sowie gerade für ältere Patienten der ungewohnte Umgang mit einem Computer angeführt³⁷. Auch wenn die Cogstate Testbatterie als „self-administrated“, also selbstdurchführbar, beschrieben wird, so lässt sich aus eigener Erfahrung sagen, dass ältere Menschen verbale Hilfestellung vom Testinstruktor benötigen, um die Aufgaben durchführen zu können. Diese Erfahrung haben auch de Jager et al.⁵⁹ in ihrer Arbeit beschrieben.

Ein generelles Problem in der POCD-Forschung ist das Fehlen einer einheitlichen Definition von POCD sowie die große Vielfalt an Test- und Berechnungsmethoden. Dies macht es schwierig Studien miteinander zu vergleichen^{11;15}. Aufgrund der oben aufgeführten Vorteile könnte computerisierten Testbatterien hierbei eine entscheidende Rolle zukommen.

6 Zusammenfassung

6.1 Einleitung

Die Inzidenz von postoperativen kognitiven Dysfunktionen (POCD) ist abhängig von der verwendeten Testbatterie und der angewendeten Berechnungsmethode. Die Messungen können mit einer Papier- und Bleistift-Testbatterie oder mit einer computerisierten Testbatterie durchgeführt werden. Ziel der Studie war es, die Inzidenzen und die Übereinstimmung von POCD, gemessen mit einer computerisierten Testbatterie und einer Papier- und Bleistift-Testbatterie, in demselben Patientenkollektiv zu bestimmen.

6.2 Methodik

Insgesamt wurden 74 Patienten eingeschlossen: 30 Patienten unterzogen sich einem operativen Eingriff unter Allgemeinanästhesie, 37 nicht chirurgische Patienten wurden als Kontrollgruppe eingeschlossen, 6 Patienten und ein Proband der Kontrollgruppe mussten ausgeschlossen werden. Die Patienten wurden im Abstand von 7 Tagen mit einer computerisierten Testbatterie und einer Papier- und Bleistift-Testbatterie getestet. Beide Testbatterien deckten dieselben kognitiven Domänen ab: Visuelles Lernen, Gedächtnis, Verarbeitungsgeschwindigkeit, visuelle Aufmerksamkeit.

6.3 Ergebnisse

Die computerisierte Testbatterie detektierte 10% [95%-CI: 3,5% - 25,6%] POCD wohingegen die Papier- und Bleistift-Testbatterie 30% [95%-CI: 16,7% - 47,9%] POCD detektierte ($p=0,03$). Die Interrater Reliabilität zwischen den beiden Testbatterien zeigte eine angemessene (moderate) Übereinstimmung (Cohens Kappa = 0,41). Alle Patienten, die durch die computerisierte Testbatterie erkannt wurden, wurden ebenfalls durch die Papier- und Bleistift-Testbatterie erkannt.

6.4 Diskussion

In der vorliegenden Studie konnte gezeigt werden, dass die POCD-Inzidenzen,

gemessen mit einer computerisierten Testbatterie und einer Papier- und Bleistift-Testbatterie, eine angemessene Übereinstimmung zeigen. Die Verwendung zweier neuropsychologischer Testbatterien, die theoretisch dieselben kognitiven Domänen abdecken, führt aber nicht automatisch zu denselben Klassifikationen.

7 Literaturverzeichnis

1. Rasmussen LS: Postoperative cognitive dysfunction: incidence and prevention. *Best Pract Res Clin Anaesthesiol* 2006; 20: 315-30
2. Müller SV, Krause N, Schmidt M, Münte TF: Cognitive dysfunction after abdominal surgery in elderly patients. *Z Gerontol Geriatr* 2004; 37 6: 475-85
3. Hanning CD: Postoperative cognitive dysfunction. *British Journal of Anaesthesia* 2005; 1: 82-7
4. Rasmussen LS: Defining postoperative cognitive dysfunction. *Eur J Anaesthesiol* 1998; 15: 761-4
5. Lewis MC, Nevo I, Paniagua MA, Ben Ari A, Pretto E, Eisdorfer S, Davidson E, Matot I, Eisdorfer C: Uncomplicated general anesthesia in the elderly results in cognitive decline: Does cognitive decline predict morbidity and mortality? *Medical Hypotheses* 2007; 68: 484-92
6. Engelhard K, Werner C: Postoperatives kognitives Defizit. *Anaesthesist* 2005; 54: 588-94
7. DIMDI Deutsches Institut für Medizinische Dokumentation und Information. ICD-10-GM Version 2009.
<http://www.dimdi.de/static/de/klassi/diagnosen/icd10/htmlgm2009/block-f00-f09.htm> . Zugriff: 15-10-2009.
8. Lewis MS, Maruff PT, Silbert BS: Examination of the use of cognitive domains in postoperative cognitive dysfunction after coronary artery bypass graft surgery. *Ann Thorac Surg* 2005; 80: 910-6
9. 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 Anaesthesiologica Scandinavica* 2006; 50: 50-7
10. Murkin JM, Newman SP, Stump DA, Blumenthal JA: Statement of Consensus on Assessment of Neurobehavioral Outcomes After Cardiac Surgery. *The Annals of Thoracic Surgery* 1995; 59: 1289-95
11. Caza N, Taha R, Qi Y&BG: The effects of surgery and anesthesia on memory and cognition. *Prog Brain Res* 2008; 169: 409-22
12. 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 Anaesthesiologica Scandinavica* 2002; 46: 384-9

13. 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. *Anesthesia and Analgesia* 2004; 99: 1245-52
14. Hall TA, McGwin G, Jr., Owsley C: Effect of cataract surgery on cognitive function in older adults. *J Am Geriatr Soc* 2005; 53: 2140-4
15. Newman S, Stygall J, Hirani S, Shaefi S, Maze M: Postoperative cognitive dysfunction after noncardiac surgery - A systematic review. *Anesthesiology* 2007; 106: 572-90
16. Fong HK, Sands LP, Leung JM: The role of postoperative analgesia in delirium and cognitive decline in elderly patients: a systematic review. *Anesth Analg* 2006; 102: 1255-66
17. 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
18. 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 Anaesthesiologica Scandinavica* 2000; 44: 1246-51
19. Rasmussen LS, Larsen K, Houx P, Skovgaard LT, Hanning CD, Moller JT: The assessment of postoperative cognitive function. *Acta Anaesthesiologica Scandinavica* 2001; 45: 275-89
20. Arrowsmith JE, Grocott HP, Reves JG, Newman MF: Central nervous system complications of cardiac surgery. *Br J Anaesth* 2000; 84: 378-93
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 Anaesthesiologica Scandinavica* 2003; 47: 260-6
22. Steinmetz J, Christensen KB, Lund T, Lohse N, Rasmussen LS: Long-term consequences of postoperative cognitive dysfunction. *Anesthesiology* 2009; 110: 548-55
23. 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

24. Biedler A, Juckenhofel S, Larsen R, Radtke F, Stotz A, Warmann J, Braune E, Dyttkowitz A, Henning F, Strickmann B, Lauven PM: Postoperative cognitive dysfunction in elderly patients. Based on the publication of the ISPOCD 1 study group in *Lancet* (1998) 351: 857-861. *Anaesthetist* 1999; 48: 884-95
25. 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
26. Rubens FD, Boodhwani M, Nathan H: Interpreting studies of cognitive function following cardiac surgery: a guide for surgical teams. *Perfusion-Uk* 2007; 22: 185-92
27. Mitrushina M, Satz P: Effect of repeated administration of a neuropsychological battery in the elderly. *Journal of Clinical Psychology* 1991; 790-801
28. Falletti MG, Maruff P, Collie A, Darby D: Practice effects associated with the repeated assessment of cognitive function using the CogState battery at 10-minute, one week and one month test-retest intervals. *J Clin Exp Neuropsychol* 2006; 28: 1095-112
29. Collie A, Maruff P, Makdissi M, McStephen M, Darby DG: Statistical procedures for determining the extent of cognitive change following concussion. *Br J Sports Med* 2004; 38: 273-8
30. Newman SP: Analysis and interpretation of neuropsychologic tests in cardiac surgery. *Ann Thorac Surg* 1995; 59: 1351-5
31. Heaton R, Paulsen JS, McAdams LA, Kuck J, Zisook S, Braff D, Harris J, Jeste DV: Neuropsychological deficits in schizophrenics. Relationship to age, chronicity, and dementia. *Arch Gen Psychiatry* 1994; 51: 469-76
32. Mahanna EP, Blumenthal JA, White WD, Croughwell ND, Clancy CP, Smith LR, Newman MF: Defining neuropsychological dysfunction after coronary artery bypass grafting. *Ann Thorac Surg* 1996; 61: 1342-7
33. Kneebone AC, Andrew MJ, Baker RA, Knight JL: Neuropsychologic changes after coronary artery bypass grafting: use of reliable change indices. *Ann Thorac Surg* 1998; 65: 1320-5
34. Jacobson NS, Truax P: Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991; 59: 12-9
35. 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. *Archives of Clinical*

- Neuropsychology 2006; 21: 421-7
36. 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. *British Journal of Anaesthesia* 2004; 92: 814-20
 37. Wild K, Howieson D, Webbe F, Seelye A, Kaye J: Status of computerized cognitive testing in aging: a systematic review. *Alzheimers Dement* 2008; 4(6): 428-37
 38. Maruff P, Thomas E, Cysique L, Brew B, Collie A, Snyder P, Pietrzak RH: Validity of the CogState brief battery: relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. *Arch Clin Neuropsychol* 2009; 24: 165-78
 39. Collie A, Darby D, Maruff P: Computerised cognitive assessment of athletes with sports related head injury. *Br J Sports Med* 2001; 35: 297-302
 40. Kox WJ, Spies C: *Check-up Anästhesiologie Standards Anästhesie - Intensivmedizin - Schmerztherapie - Notfallmedizin*, 2. edition. Springer-Verlag, 2005,
 41. Folstein MF, Folstein SE, McHugh PR: Mini-Mental State: a practical method for grading the state of patients for the clinician. *Journal of Psychiatric Research* 1975; 12: 189-98
 42. Rasmussen LS, Siersma VD: Postoperative cognitive dysfunction: true deterioration versus random variation. *Acta Anaesthesiologica Scandinavica* 2004; 48: 1137-43
 43. Lowe C, Rabbitt P: Test/re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: theoretical and practical issues. *Cambridge Neuropsychological Test Automated Battery. International Study of Post-Operative Cognitive Dysfunction. Neuropsychologia* 1998; 36: 915-23
 44. 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
 45. Houx PJ, Jolles J, Vreeling FW: Stroop interference: aging effects assessed with the Stroop Color-Word Test. *Exp Aging Res* 1993; 19: 209-24
 46. Rey A.: L'examen psychologique dans les cas d'encéphalopathie traumatique. *Archives de Psychologie* 1941; 28: 21
 47. Brand N, Jolles J: Learning and retrieval rate of words presented auditorily and visually. *J Gen Psychol* 1985; 201-10

48. Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-74
49. 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 Anaesthesiologica Scandinavica* 2005; 49: 1225-31
50. Bodolea C, Hagau N, Coman I, Pinteia S, Cristea IA, Cristea T, Negrutiu S: Postoperative cognitive dysfunction in elderly patients. An integrated psychological and medical approach. *Journal of Cognitive and Behavioral Psychotherapies* 2008; 8: 117-32
51. Funder KS, Steinmetz J, Rasmussen LS: Cognitive dysfunction after cardiovascular surgery. *Minerva Anestesiologica* 2009; 75: 329-32
52. Patel RL, Turtle MR, Chambers DJ, James DN, Newman S, Venn GE: Alpha-stat acid-base regulation during cardiopulmonary bypass improves neuropsychologic outcome in patients undergoing coronary artery bypass grafting. *J Thorac Cardiovasc Surg* 1996; 111: 1267-79
53. Forsman M, Olsnes BT, Semb G, Steen PA: Effects of nimodipine on cerebral blood flow and neuropsychological outcome after cardiac surgery. *Br J Anaesth* 1990; 65: 514-20
54. Altman D.G., Royston P.: The cost of dichotomising continuous variables. *British Medical Journal* 2006; 322: 1080
55. 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, Corsini LM, Silverstein JH, Nielsen IK, Moller JT: Cognitive dysfunction after minor surgery in the elderly. *Acta Anaesthesiologica Scandinavica* 2003; 47: 1204-10
56. Borowicz LM, Goldsborough MA, Selnes OA, McKhann GM: Neuropsychologic change after cardiac surgery: a critical review. *J Cardiothorac Vasc. Anesth* 1996; 10: 105-11
57. Lewis MS, Maruff P, Silbert BS, Evered LA, Scott DA: Detection of postoperative cognitive decline after coronary artery bypass graft surgery is affected by the number of neuropsychological tests in the assessment battery. *Ann Thorac Surg* 2006; 81: 2097-104
58. Blumenthal JA, Mahanna EP, Madden DJ, White WD, Croughwell ND, Newman MF: Methodological issues in the assessment of neuropsychologic function after cardiac surgery. *Ann Thorac Surg* 1995; 59: 1345-50

-
59. de Jager CA, Schrijnemaekers AC, Honey TEM: Detection of MCI in the clinic: evaluation of the sensitivity and specificity of a computerised test battery, the Hopkins Verbal Learning Test and the MMSE. *Age Ageing* 2009; 38: 455-60

8 Publikationen im Rahmen der Dissertation

Radtke FM, Franck M, Papkalla N, Herbig TS, Weiss-Gerlach E, Kleinwaechter R, Wernecke KD, Spies CD: Postoperative cognitive dysfunction - computerized and conventional tests showed only moderate interrater reliability. J Anesth. 2010; 24(4): 518-25

N. Papkalla, F. M. Radtke, M. Franck, T. Herbig, C. Spies. Postoperative kognitive Dysfunktionen – ein Vergleich zweier Testbatterien. Poster präsentiert auf: DAC 2009. Deutscher Anästhesiecongress der deutschen Gesellschaft für Anästhesiologie und Intensivmedizin (DGAI); 9.-12.05.2009; Leipzig

N. Papkalla, F. M. Radtke, M. Franck, T. Herbig, C. Spies. Postoperative Cognitive Dysfunction in high risk patients. Poster präsentiert auf: HAI 2007 - Der Hauptstadtkongress der DGAI für Anästhesiologie und Intensivtherapie; 13.-15.09.2007; Berlin

9 Abbildungsverzeichnis

| | |
|---|----|
| Abbildung 1 Lerneffekt bei wiederholter Testausführung modifiziert nach Newman et al. ³⁰ | 12 |
| Abbildung 2 SCWT Teil 1 | 19 |
| Abbildung 3 SCWT Teil 2 | 19 |
| Abbildung 4 SCWT Teil 3 | 19 |
| Abbildung 5 Detection Task I | 21 |
| Abbildung 6 Detection Task II | 21 |
| Abbildung 7 Identification Task I | 21 |
| Abbildung 8 Identification Task II | 21 |
| Abbildung 9 Flussdiagramm des Studienverlaufs | 25 |

10 Tabellenverzeichnis

| | |
|--|----|
| Tabelle 1 Kognitive Domänen der Testbatterien | 22 |
| Tabelle 2 Basisdaten Patienten- und Kontrollkollektiv | 27 |
| Tabelle 3 Messwerte PPTB Patienten- und Kontrollkollektiv..... | 28 |
| Tabelle 4 Messwerte CTB Patienten- und Kontrollkollektiv..... | 29 |
| Tabelle 5 Kreuztabelle Klassifikation nach CTB und PPTB | 29 |
| Tabelle 6 Interpretation des Cohens Kappa nach Landis und Koch | 30 |
| Tabelle 7 Diagnostische Tests | 31 |
| Tabelle 8 Basisdaten postoperativ - POCD-Einteilung nach CTB..... | 32 |
| Tabelle 9 Basisdaten postoperativ - POCD-Einteilung nach PPTB..... | 33 |
| Tabelle 10 Basisdaten postoperativ - POCD positiv nach CTB versus POCD positiv nur nach PPTB | 34 |

11 Danksagung

Zunächst möchte ich mich bei allen Patienten, die sich bereit erklärt haben an dieser Studie teilzunehmen, bedanken.

Des Weiteren danke ich Frau Prof. Dr. med. Claudia Spies für die Überlassung des Themas, das Einreichen der Genehmigungen für die Arbeit, die stetig gute Betreuung sowie der Korrektur meiner Arbeit.

Für die hervorragende Unterstützung bei der Durchführung sowie dem Schreiben der Arbeit danke ich Herrn Dr. med. Finn Radtke und Herrn Dr. med. Martin Franck.

Meinem Mitdoktoranden Tim Herbig danke ich für die allzeit gute Zusammenarbeit und die gegenseitige Unterstützung.

Meiner Familie und Freunden, insbesondere Cornelia Beierle, danke ich für die vielseitige Unterstützung bei der Erstellung dieser Arbeit.

Herrn Prof. Dr. rer. nat. habil. Wernecke (Geschäftsführer der SOSTANA GmbH und ehem. Direktor des Instituts für Medizinische Biometrie des Universitätsklinikum Charité) danke ich für die Beratung und Unterstützung bei der statistischen Auswertung der erhobenen Daten.

12 Lebenslauf

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Erklärung

„Ich, Norbert Papkalla, erkläre, dass ich die vorgelegte Dissertation mit dem Thema: „Postoperative kognitive Dysfunktionen bei nicht kardiochirurgischen Patienten – ein Vergleich zweier Testmethoden“ selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, ohne die (unzulässige) Hilfe Dritter verfasst und auch in Teilen keine Kopien anderer Arbeiten dargestellt habe.“

28.03.2011

Norbert Papkalla