

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

DISSERTATION

Validierung der Schweregradskala für die Confusion Assessment Method für Intensivstationen bei neurologischen und neurochirurgischen Intensivpatienten

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

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Dennis Ullrich
aus Berlin

Datum der Promotion: 17.09.2021

Teilergebnisse der vorliegenden Arbeit wurden zur Veröffentlichung eingereicht und angenommen bei:

Neurocritical Care, LCCN: 2002215176, ISSN: 1541-6933

Inhaltsverzeichnis

Abbildungs- und Tabellenverzeichnis.....	5
Abkürzungen	6
Zusammenfassung.....	7
Abstract	9
1. Einleitung	11
1.1 Definition des Delirs	12
1.2 Prävalenz des Delirs	13
1.3 Risikofaktoren für ein Delir	15
1.4 Relevanz des Delirs	17
1.5 Delirdiagnostik und -messinstrumente	19
1.6 Die CAM-ICU	25
1.7 Delirscreening bei neurologischen und neurochirurgischen Intensivpatienten.....	28
1.8 Zielsetzung dieser Arbeit	31
2. Material und Methoden	32
2.1 Patientenkollektiv, Ein- und Ausschlusskriterien, Erhebungszeiträume	32
2.1.1 Allgemeine Konzeption.....	33
2.1.2 Erhebung von Basis- und Outcomeparametern.....	34
2.2 Delirmessung mit der CAM-ICU (Screening) und nach DSM-IV TR (Goldstandard).....	35
2.2.1 CAM-ICU.....	35
2.2.2 DSM IV TR	37
2.2.3 Unterweisung und Schulung der CAM-ICU.....	38
2.2.4 Unterweisung und Schulung der neuropsychiatrischen Untersuchung und Evaluation nach DSM IV-TR.....	39
2.3 Besonderheiten der Datenerhebung	39
2.3.1 Referenzmessung und Wiederholungsmessungen.....	39
2.3.2 Berechnung der <i>severity scale for the CAM-ICU</i> (ssCAM-ICU)	40
2.4 Statistische Analyse.....	40
2.4.1 Datenauswertung	40
3. Ergebnisse	43
3.1 Patientenkollektiv.....	43
3.1.1 Charakteristika des Patientenkollektivs.....	44
3.2 Prävalenz des Delirs	45

3.3	Testgütekriterien für den ersten validen Messtag im Gesamtkollektiv	45
3.4	Testgütekriterien der repetitiven Messungen	46
3.5	Testgütekriterien für die repetitiven Messungen in den Subgruppen	47
3.6	ROC-Analyse und Youden-Index: ssCAM-ICU Gesamtkollektiv, erster valider Messtag	49
3.7	ROC-Analyse und Youden-Index: ssCAM-ICU Gesamtkollektiv, repetitive Messungen	50
3.8	ROC-Analyse und Youden-Index: ssCAM-ICU bei Schlaganfallpatienten, repetitive Messungen	51
3.9	ROC-Analyse und Youden-Index: ssCAM-ICU bei neurochirurgischen Patienten, repetitive Messungen	52
4.	Diskussion.....	53
4.1	Basischarakteristika des Patientenkollektivs und Delirprävalenz.....	54
4.2	Testgütekriterien der ssCAM-ICU im Gesamtkollektiv	57
4.3	Testgütekriterien der ssCAM-ICU in den Subgruppen.....	61
5.	Grenzen und Stärken der Arbeit	63
6.	Schlussfolgerungen und Ausblick.....	64
7.	Anhang.....	65
7.1	CAM-ICU Arbeitsblatt.....	65
7.2	ssCAM-ICU	66
7.3	GCS.....	67
7.4	BPS und BPS NI.....	67
7.5	SAPS II.....	69
7.6	APACHE II.....	70
8.	Literaturverzeichnis.....	71
9.	Eidesstattliche Versicherung	79
10.	Anteilsklärung	80
11.	Lebenslauf.....	81
12.	Publikationen im Rahmen der Dissertation.....	82
13.	Danksagung	84

Abbildungs- und Tabellenverzeichnis

Abbildungsverzeichnis

Abbildung 1.3:	Modell zur Delirentstehung.....	16
Abbildung 2.1.1:	Ablaufschema des Delirscreenings während der Datenerhebung.....	34
Abbildung 3.1:	Strobe Diagramm.....	43
Abbildung 3.6:	ROC-Kurve für die Punktwerte auf der ssCAM-ICU, erste valide Messung Gesamtkollektiv.....	49
Abbildung 3.7:	ROC-Kurve für die Punktwerte auf der ssCAM-ICU, repetitive Messungen Gesamtkollektiv.....	50
Abbildung 3.8:	ROC-Kurve für die Punktwerte auf der ssCAM-ICU, repetitive Messungen Schlaganfallpatienten.....	51
Abbildung 3.9:	ROC-Kurve für die Punktwerte auf der ssCAM-ICU, repetitive Messungen neurochirurgische Patienten.....	52
Abbildung 7.1:	CAM-ICU Arbeitsblatt.....	65
Abbildung 7.2:	ssCAM-ICU.....	66
Abbildung 7.5:	SAPS II.....	69
Abbildung 7.6:	APACHE II.....	70

Tabellenverzeichnis

Tabelle 1.1:	Merkmale des Delirs gemäß DSM V.....	12
Tabelle 2.1.2:	Erhobene Basisparameter.....	34
Tabelle 2.2.1:	Richmond Agitation Sedation Scale.....	37
Tabelle 2.4.1:	Formeln Testgütekriterien.....	41
Tabelle 2.4.2:	Youden-Index.....	42
Tabelle 3.1.1:	Basischarakteristika des Patientenkollektivs.....	44
Tabelle 3.3.1:	Kreuztabelle ssCAM-ICU vs. DSM IV TR, erster valider Messtag, Gesamtkollektiv.....	46
Tabelle 3.3.2:	Testgütekriterien erster valider Messtag, Gesamtkollektiv.....	46
Tabelle 3.4.1:	Kreuztabelle ssCAM-ICU vs. DSM IV TR, repetitive Messungen, Gesamtkollektiv.....	46
Tabelle 3.4.2:	Testgütekriterien repetitive Messungen, Gesamtkollektiv.....	46
Tabelle 3.5.1.1:	Kreuztabelle ssCAM-ICU vs. DSM IV TR, repetitive Messungen, Schlaganfallpatienten.....	47
Tabelle 3.5.1.2:	Testgütekriterien repetitive Messungen, Schlaganfallpatienten.....	47
Tabelle 3.5.2.1:	Kreuztabelle CAM-ICU vs. DSM IV TR, repetitive Messungen, neurochirurgische Patienten.....	47
Tabelle 3.5.2.2:	Testgütekriterien repetitive Messungen, neurochirurgische Patienten.....	48
Tabelle 7.3:	GCS.....	67
Tabelle 7.4.1:	BPS.....	67
Tabelle 7.4.2:	BPS-NI.....	68

Abkürzungen

APA	<i>American Psychological Association</i>
APACHE II	<i>Acute Physiology and Chronic Health Evaluation II</i>
ASE	<i>Attention Screening Examination</i>
AUC	<i>Area Under the Curve</i>
BPS	<i>Behavioral Pain Scale</i>
BPS-NI	<i>Behavioral Pain Scale Non-Intubated</i>
CAM-ICU	<i>Confusion Assessment Method for Intensive Care Unit</i>
CPR	Kardiopulmonale Reanimation
CVK	Charité Campus Virchow-Klinikum
DDS	<i>Delirium Detection Score</i>
DIVI	Deutsche interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin
DGAI	Deutsche Gesellschaft für Anästhesiologie und Intensivmedizin
DRS-R-98	<i>Delirium Rating Scale-Revised-98</i>
DSM III R	<i>Diagnostic and Statistical Manual of Mental Disorders 3. Edition Revision</i>
DSM IV TR	<i>Diagnostic and Statistical Manual of Mental Disorders 4. Edition, TextRevision</i>
DSM V	<i>Diagnostic and Statistical Manual of Mental Disorders 5. Edition</i>
ESCIM	<i>European Society of Intensive Care Medicine</i>
GCS	<i>Glasgow Coma Scale</i>
ICD-10	<i>International Statistical Classification of Diseases and Related Health Problems 10. Edition</i>
ICDSC	<i>Intensive Care Delirium Screening Checklist</i>
ICU	<i>Intensive Care Unit</i>
IQR	Interquartilsrange, Interquartilsabstand
KI	Konfidenzintervall
MMSE-Test	<i>Mini Mental State Examination</i>
NAS	Numerische Analogskala
NICE	<i>National Institute for Health and Care Excellence</i>
NPV	negative prädiktiver Wert
Nu-DESC	<i>Nursing Delirium Screenig Scale</i>
PPV	positive prädiktiver Wert
RASS	<i>Richmond Agitation Sedation Scale</i>
ROC	<i>Receiver-Operating-Characteristics</i>
SAPS II	<i>Simplified Acute Physiology Score II</i>
SOFA	<i>Sequential (Sepsis-related) Organ Failure Assessment</i>
SSD	Subsyndromales Delir
ssCAM-ICU	<i>severity scale for the Confusion Assessment Method for Intensive Care Unit</i>
sspCAM-ICU	<i>severity scale for the pediatric Confusion Assessment Method for Intensive Care Unit</i>
TIA	Transiente Ischämische Attacke
WHO	Weltgesundheitsorganisation

Zusammenfassung

Einleitung: Das Delir ist die häufigste akut-psychiatrische Erkrankung bei erwachsenen Intensivpatienten und mit einer erhöhten Morbidität für intensivmedizinische Komplikationen sowie einer dreifach erhöhten Mortalität nach sechs Monaten assoziiert. Es entzieht sich häufig einer Blickdiagnose, sodass ein routinemäßiges Delirscreening mit validierten Messinstrumenten in aktuellen Leitlinien empfohlen wird. Ordinale Messinstrumente können dabei Vorteile für die Diagnose eines subsyndromalen Delirs, die Verlaufsbeschreibung zur Abschätzung des Therapieerfolgs und eine Schweregrad-einschätzung bieten. Die CAM-ICU ist ein weit verbreitetes Messinstrument. Während die Validität ordinaler Delirmessinstrumente für internistische und chirurgische Intensivpatienten bereits gezeigt wurde, gibt es wenige Studien zu neurologischen und neurochirurgischen Intensivpatienten. Das Ziel dieser Arbeit besteht in der Validierung der Schweregradskala für die *Confusion Assessment Method für Intensivstationen* (ssCAM-ICU) bei Neurointensivpatienten.

Methodik: Im Rahmen einer prospektiven Observationsstudie wurden alle konsekutiv behandelten deutschsprachigen Patienten >18 Jahre, ohne anamnestisch bekannte Demenz oder psychiatrische Vorerkrankungen, einer universitären Neurointensivstation mittels ssCAM-ICU und *Diagnostic and Statistical Manual of Mental Disorders IV. Edition Text Revision* (DSM IV TR) als Referenzstandard täglich verblindet und standardisiert während der Dauer des Intensivaufenthalts untersucht. Es erfolgte eine Validierung der ssCAM-ICU, die Berechnung von *Receiver-Operating-Characteristics* (ROC)-Kurven und eine Cut-Off Optimierung mithilfe des Youden-Index für das Gesamtkollektiv sowie die Subgruppen der Schlaganfall- und neurochirurgischen Patienten. Für das Gesamtkollektiv wurde nach dem ersten validen Messtag und Wiederholungsmessungen unterschieden.

Ergebnisse: Es konnten 101 Patienten ausgewertet werden. Die Gesamtprävalenz des Delirs lag nach DSM IV TR bei 15,4% (n=30 Patienten). In den Wiederholungsmessungen (n=324) zeigte die ssCAM-ICU eine Sensitivität von 92% und eine Spezifität von 81%. In den Subgruppen wurden bei Schlaganfallpatienten vergleichbare (Sensitivität: 93,6; Spezifität 79,2%) bzw. bei neurochirurgischen Patienten schlechtere Testgütekriterien (Sensitivität: 87%; Spezifität 86,1%) festgestellt. Die *Area Under the Curve* der ROC-Kurven weist Werte eines akkuraten bis hochakkuraten Tests auf (AUC: 0,91; 0,9 bzw. 0,94). Im Rahmen der Cut-Off-Optimierung mittels des Youden-Index zeigt sich im Gesamtkollektiv und bei Schlaganfallpatienten ein zur vorhandenen Literatur vergleichbarer

optimaler Cut-Off, bei neurochirurgischen Patienten jedoch nach Adjustierung ein niedrigerer optimaler Cut-Off (Sensitivität: 95,7%; Spezifität 82,9%).

Schlussfolgerungen: Die ssCAM-ICU weist gute Testgütekriterien bei Neurointensivpatienten auf. Aufgrund der hohen Sensitivität bei guter Spezifität ist die ssCAM-ICU als Screeningtest geeignet. Unter Verwendung individualisierter Cut-Offs für Patienten mit Schlaganfall und neurochirurgische Patienten können in der jeweiligen Subgruppe bessere Testgütekriterien erreicht werden. Weitere Arbeiten müssen untersuchen, ob zusätzliche Faktoren Einfluss auf die Testgütekriterien haben können.

Abstract

Background and Objectives: Delirium is the most common acute psychiatric complication in adult ICU-patients. It is associated with adverse outcomes including increased mortality after six months and a higher burden of ICU-associated complications. Varying clinical manifestations and fluctuating course hamper diagnosis. Thus, clinical practice guidelines recommend routine use of validated delirium screening instruments for reliable diagnosis and early therapy initiation. Ordinal screening instruments may offer advantages for the diagnosis of subsyndromal delirium, the description of the clinical course, evaluation of therapy success and an assessment of its severity. CAM-ICU is a very common and frequently used delirium screening instrument. While high validity of ordinal delirium screening instruments has been shown for medical and surgical ICU patients, there is limited knowledge for neurocritical care. Aim of this work is the validation of the ordinal severity scale for the Confusion Assessment Method for the ICU (ssCAM-ICU) in neurocritical care patients.

Methods: Patients of our tertiary care neuro-ICU were enrolled in this prospective observational cohort study. Exclusion criteria were age <18 years, non-German speaking, known severe dementia or other psychiatric disease. Daily blinded assessment with ssCAM-ICU and DSM IV TR were performed. Validation of ssCAM-ICU was performed, Receiver-Operating-Characteristics (ROC-curves) were plotted and the optimal cut-off was determined by using Youden-Index. Furthermore, patients suffering ischaemic stroke or undergoing neurosurgery intervention were tested separately. Total population was evaluated for first assessment and consecutive measurements.

Main Results: All in all, 101 patients were enrolled and overall delirium prevalence was 15,4% (n=30 patients) according to reference standard. Test validity was high in 324 consecutive assessments (sensitivity: 92%, specificity: 81%). Subgroup of patients suffering ischaemic stroke showed similar results (sensitivity: 93,6%, specificity: 79,2%). In neurosurgery patients ssCAM-ICU demonstrated lower test validity (sensitivity: 87%, specificity: 86,1%). Area under the curve (AUC) indicates accurate to highly accurate test for all groups (AUC: 0,91; 0,9; 0,94). By using Youden-Index optimal cut-off was specified and indicates similar cut-off for entire population and stroke patients as given in literature. Although in neurosurgery patients lower cut-off demonstrated better validity (sensitivity: 95,7%, specificity: 82,9%).

Conclusions: The ssCAM-ICU demonstrated good test validity for a delirium screening instrument when used in this population. By using optimal cut-off for subgroups better test validity can be achieved. Further studies should mention the influence of other confounders on test validity.

1. Einleitung

Das Delir ist mit bis zu 87% betroffenen Patienten (vgl. Ely et al., 2001) die häufigste akut-psychiatrische Erkrankung auf der Intensivstation. Das Auftreten ist mit einer erhöhten Mortalität nach 6 Monaten (Hazard Ratio 3,2) (vgl. Ely et al., 2004) und mit zusätzlichen Kosten für die Gesundheitssysteme in Höhe von ca. 60.000\$ im ersten Jahr assoziiert (vgl. Leslie et al., 2008). Es ist durch einen plötzlichen Beginn und eine im Tagesverlauf fluktuierende Symptomatik gekennzeichnet. Die erkrankten Patienten zeigen Störungen des Bewusstseins, der Aufmerksamkeit und der Kognition. Es konnte in einer großen prospektiven Studie an 304 Patienten gezeigt werden, dass jeder zusätzliche Tag im Delir, trotz Berücksichtigung möglicher Einflussfaktoren, mit einer relativen Mortalitätserhöhung von 10% assoziiert ist (vgl. Pisani et al., 2009). Ferner steigt die Morbidität für Komplikationen (vgl. Ely et al., 2001a; Balas et al., 2009). So kommt es insbesondere bei bis zu 40% aller intensivmedizinisch behandelten Patienten zu einer deutlichen Einschränkung der kognitiven Fähigkeiten nach drei Monaten. Diese ist auch nach zwölf Monaten noch bei bis zu 34% der Patienten nachweisbar. Patienten, die ein Delir erlitten, zeigten nochmals schlechtere kognitive Fähigkeiten als die Vergleichsgruppe (vgl. Pandharipande et al., 2013). Obwohl es sich beim Delir des Intensivpatienten um eine häufig auftretende Erkrankung handelt, ist eine zuverlässige Diagnose, basierend auf einer klinischen Einschätzung, nur eingeschränkt möglich, weshalb es häufig übersehen wird (vgl. Devlin et al., 2007; van Eijk et al., 2009). Mittels validierter Messinstrumente kann ein Delir zuverlässiger erkannt werden (vgl. Devlin et al., 2007a; Devlin et al., 2008), sofern das verwendete Messinstrument für die spezifische Patientengruppe validiert ist und hohe Testgütekriterien nachgewiesen wurden (vgl. Luetz et al., 2010a). Zudem sollte eine gute Anwendbarkeit in der Routineversorgung gewährleistet sein. Die Forderung nach einem aktiven und standardisierten Delirscreening hat mittlerweile Einzug in die internationalen Leitlinien gefunden (vgl. Baron et al., 2015). Dies trägt auch der Tatsache Rechnung, dass durch multifaktorielle Interventionen sowohl die Dauer als auch die möglichen Folgen des Delirs positiv beeinflussbar sind (vgl. Inouye et al., 1999a; Schweickert et al., 2009).

Die Gruppe der neurologischen bzw. neurochirurgischen Patienten, die auf einer Intensivstation behandelt werden, ist allerdings bei den meisten Validierungsstudien für Delirmessinstrumente nicht untersucht worden (vgl. Ely et al, 2001; Luetz et al, 2010). Es ist somit bisher nicht sicher festzustellen, ob die für andere Patientenkollektive validierten Messinstrumente auf Neurointensivpatienten übertragbar sind.

1.1 Definition des Delirs

Die auf der Intensivstation am häufigsten akut auftretende psychiatrische Erkrankung mit zerebraler Dysfunktion ist das Delir (vgl. Ely et al., 2001). Historisch wurde häufig vom Durchgangssyndrom, einer ‚*ICU-Psychosis*‘ oder einem ‚*ICU-Syndrome*‘ berichtet. Alle diese Bezeichnungen lassen sich unter dem klinischen Syndrom des Delirs subsumieren (vgl. McGuire et al., 2000). International waren und sind weitere Bezeichnungen gebräuchlich, die eine genaue Abgrenzung in der Literatur deutlich erschweren (vgl. Morandi et al., 2008).

In der zehnten Auflage der *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) von 1992 definiert die Weltgesundheitsorganisation (WHO) das Delir unter der Schlüsselnummer F05.x als unspezifisches hirnorganisches Syndrom mit Bewusstseinsstörung. Zusätzlich kann eine Aufmerksamkeits-, Wahrnehmungs-, Denk- oder Gedächtnisstörung sowie eine Beeinträchtigung des Schlaf-Wach-Rhythmus, der Emotionalität oder der Psychomotorik mit schwankendem oder fluktuierendem Verlauf und akutem Beginn vorliegen (vgl. www.dimid.de, 2020). Neben der Bewusstseinsstörung müssen mindestens zwei der weiteren Symptome vorliegen, um ein Delir nach ICD-10 zu diagnostizieren.

Im Jahr 2013 wurde von der *American Psychological Association* (APA) ein mit Kommentierungen versehenes *Diagnostic and Statistical Manual of Mental Disorders V. Edition* (DSM V) herausgegeben. Hier wird das Delir wie in nachfolgender Tabelle definiert (vgl. APA, 2013):

Tabelle 1.1: Merkmale des Delirs gemäß DSM V, nach (APA, 2013)

Merkmal	Definition
Merkmal A	Aufmerksamkeitsstörung und qualitative oder quantitative Bewusstseinsstörung
Merkmal B	akuter Beginn oder fluktuierender Verlauf
Merkmal C	Störung der Kognition
Merkmal D	fehlender Hinweis auf Aggravation eines vorbestehenden neurokognitiven Defizits oder Koma
Merkmal E	Ätiologie aufgrund physiologischer oder anderer Gründe

Um nach dieser Definition ein Delir zu diagnostizieren, müssen alle fünf Merkmale erfüllt sein. Falls ein Patient die Definition des Delirs nach DSM V nicht vollständig erfüllt, jedoch eine deutliche klinische Einschränkung besteht, dann kann das abgeschwächte Delir diagnostiziert werden (*attenuated delirium syndrome*) (APA, 2013). Für die Diagnosestellung des abgeschwächten Delirs macht das DSM V keine expliziten Vorgaben.

Die Anwendung der Definition nach DSM V bzw. der vorherigen Versionen ist in der Forschungsliteratur verbreitet, da diese im Gegensatz zur Definition nach ICD-10 mit hilfreichen Handreichungen und Kommentierungen zur klinischen Anwendung versehen wurden. Des Weiteren muss die ICD-10 aufgrund ihrer internationalen Gültigkeit eine größere Bandbreite an interkulturellen Differenzen in den einzelnen Gesundheitssystemen abbilden. Dies gilt nicht für das DSM V, da es für die Anwendung im Gesundheitssystem der USA konzipiert wurde, weshalb die Definition spezifischer auf die Bedürfnisse einer homogenen Gruppe der Anwender fokussiert werden kann.

1.2 Prävalenz des Delirs

In der aktuell verfügbaren Literatur schwankt die Prävalenz des Delirs bei intensivmedizinisch behandelten Patienten zwischen 16% und 87% (vgl. Bergeron et al., 2001; Ely et al., 2001).

In einer prospektiven, monozentrischen Studie an 38 beatmeten und nicht beatmeten Patienten einer universitären, internistischen Intensivstation konnten Ely et al. eine Prävalenz von 87% zeigen (vgl. 2001). In einem Patientenkollektiv von 275 beatmeten, internistischen Patienten konnte durch dieselbe Arbeitsgruppe eine Prävalenz von 81,7% nachgewiesen werden (vgl. Ely et al., 2004). Dahingegen ergab sich in einer prospektiven, monozentrischen Studie mit 156 chirurgischen, beatmeten und nicht-beatmeten Intensivpatienten einer deutschen Universitätsklinik eine Prävalenz von 40,4% (vgl. Luetz et al., 2010). Dabei erfolgte der Einschluss unabhängig davon, ob eine Beatmung vorlag. Van Eijk et al. konnten an einem interdisziplinären Patientenkollektiv einer universitären Intensivstation mit 32 Behandlungsplätzen im Rahmen einer prospektiven, monozentrischen Arbeit eine Prävalenz von 34,1% beobachten (vgl. 2009). In einem gemischt chirurgisch/internistischen Kollektiv von 93 Patienten einer universitären Intensivstation wurde lediglich eine Prävalenz von 16% nachgewiesen (vgl. Bergeron et al., 2001). Bei dieser Arbeit wurden Patienten ausgeschlossen, die nicht binnen fünf Tagen mit den verwendeten Messinstrumenten untersuchbar waren. Alle Patienten, die ein Delir entwickelten, litten zudem an Erkrankungen, die ein Delir begünstigen oder imitieren könnten (vgl. Bergeron et al., 2001).

In den bisher genannten Untersuchungen wurde das Delirscreening stets im Rahmen eines streng kontrollierten Studienprotokolls unter Studienbedingungen und mit explizit geschultem Personal durchgeführt. Zusammenfassend ist die Prävalenz des Delirs in den einzelnen Arbeiten sehr heterogen

verteilt und zeigt beträchtliche Unterschiede, die auf das Patientenkollektiv (vgl. Luetz et al., 2010a), die Definition des Delirs (vgl. Meagher et al., 2014) und die verwendete Messmethodik (vgl. Luetz et al., 2010a) zurückzuführen sind.

Das in der vorliegenden Arbeit untersuchte Patientenkollektiv der Neurointensivpatienten ist bisher in sieben Validierungsstudien mit einem deutlichen Schwerpunkt auf neurologisch erkrankten Patienten untersucht worden, wobei sich die jeweiligen Studienprotokolle sehr heterogen darstellen.

So fanden Mitasova et al. im Rahmen einer monozentrischen, prospektiven Validierungsstudie in einer Kohorte von 129 Patienten, die auf einer universitären *Stroke Unit* behandelt wurden, eine Prävalenz von 42,6% (vgl. 2012). In einer Subgruppe des gemischten Kollektivs einer interdisziplinären Intensivstation einer niederländischen Universitätsklinik wurden 28 von 126 Patienten aufgrund einer neurologischen oder neurochirurgischen Diagnose behandelt. Davon entwickelten im Rahmen dieser monozentrischen, prospektiven Studie insgesamt 35,7% des Gesamtkollektivs ein Delir (vgl. van Eijk et al., 2009). Aus derselben Arbeitsgruppe sind Daten aus einer multizentrischen Anwendungsstudie auf Intensivstationen unterschiedlicher Versorgungsgrade bei routinemäßiger Anwendung eines Delirscores verfügbar. Dabei konnten 34 von 282 eingeschlossenen Patienten der Fachrichtung neurologisch bzw. neurochirurgisch zugeordnet werden. Nach Ausschluss der nicht untersuchbaren Patienten zeigten die verbliebenen 22 Patienten eine Delirprävalenz von 30,4% (vgl. van Eijk et al., 2011).

Im Bereich der neurochirurgischen Patienten konnte in einer auf 61 Intensivpatienten mit leichtem bis mittelschwerem Schädel-Hirn-Trauma beschränkten, prospektiven, multizentrischen Beobachtungsstudie zweier kanadischer Traumazentren eine Gesamtprävalenz von 45,9% nach den Kriterien der DSM IV TR festgestellt werden (vgl. Frenette et al., 2015). Daneben wiesen Wang et al. in einem Kollektiv von 128 neurochirurgischen Patienten, die aufgrund einer neurovaskulären Erkrankung behandelt wurden, mittels eines Delirmessinstruments eine Prävalenz von 42,2% nach (vgl. 2018). Darüber hinaus konnte in einer kürzlich veröffentlichten prospektiven, monozentrischen Studie für ein heterogeneres, neurochirurgisches Patientenkontingent auf einer universitären Intensivstation in Dänemark eine Prävalenz von 67% nach den Diagnosekriterien der ICD-10 ermittelt werden (vgl. Larsen et al., 2019). Das in der zitierten Arbeit untersuchte Patientenkollektiv bildet in der aktuell veröffentlichten Literatur den größten Querschnitt der neurochirurgischen Patienten ab. In einer weiteren Arbeit konnte an einer US-amerikanischen neurologischen Intensivstation mit *Stroke Unit* an 60

neurochirurgischen Patienten, die eine intracerebrale Blutung erlitten, nach den Kriterien des DSM V eine Prävalenz von 57% nachgewiesen werden (vgl. Reznik et al., 2020).

Abschließend ist festzustellen, dass die Prävalenz des Delirs bei Intensivpatienten in Abhängigkeit vom Patientenkollektiv, den verwendeten Messinstrumenten bzw. Definitionen des Delirs und der Methodik der einzelnen Studien beträchtlichen Schwankungen unterliegt.

1.3 Risikofaktoren für ein Delir

Im Rahmen epidemiologischer Studien konnten Risikofaktoren für die Entstehung eines Delirs identifiziert werden, wobei prädisponierende und präzipitierende Faktoren unterschieden werden. Als prädisponierende Faktoren werden die eine Erkrankung begünstigenden Faktoren bezeichnet. Diese sind in der Regel nicht oder nur eingeschränkt durch Interventionen während der Krankenhausbehandlung beeinflussbar. Häufig handelt es sich um bereits vorbestehende Erkrankungen oder kognitive Einschränkungen. Dahingegen handelt es sich bei präzipitierenden Faktoren um im Rahmen der Behandlung zusätzlich auftretende Faktoren, die in Verbindung mit den prädisponierenden Faktoren eine Erkrankung hervorrufen können und unter Umständen modifizierbar sind.

Bereits 1999 wurde ein multifaktorielles Modell zur Risikobewertung entwickelt, welches auf den zum damaligen Zeitpunkt verfügbaren epidemiologischen Daten für geriatrische Patienten basiert (vgl. Inouye, 1999). Nach diesem Modell erhöht sich das Risiko für das Entstehen eines Delirs durch das Zusammentreffen von prädisponierenden und präzipitierenden Faktoren. Hierbei wird die individuelle Verletzlichkeit durch prädisponierende Faktoren abgebildet, während die verschiedenen präzipitierenden Faktoren ein auslösendes Trauma darstellen. Nach diesem Modell bedarf ein Patient, der eine hohe Verletzlichkeit für ein Delir aufweist, nur eines verhältnismäßig geringen zusätzlichen Traumas, um die Entstehung eines Delirs zu begünstigen und umgekehrt.

Für internistisch/geriatrische Patienten konnten fortgeschrittenes Lebensalter, reduzierter kognitiver Status (z.B. Demenz), geringerer funktionaler Status, sensorische Einschränkung (auditiv/visuell), Mangelernährung oder Dehydratation, Substanzabusus und gravierende chronische oder akute Erkrankungen als prädisponierende Faktoren identifiziert werden (vgl. Inouye, 2006). Als präzipitierende Faktoren werden beispielsweise Medikamente, Substanzentzug, schwere Erkrankung, die Notwendigkeit eines operativen Eingriffs, Störungen des Tag-Nacht-Rhythmus, die individuelle

Erkrankungsschwere, Infektionen oder die Verwendung medizinischer Devices angegeben (vgl. Inouye, 2006). Für das Patientenkollektiv der Intensivpatienten konnten weitere Risikofaktoren identifiziert werden (vgl. Aldemir et al., 2001; Girard et al., 2008; Van Rompaey et al., 2009). Diese Risikofaktoren und eine Darstellung des Modells zeigt die nachfolgende Abbildung 1.3.

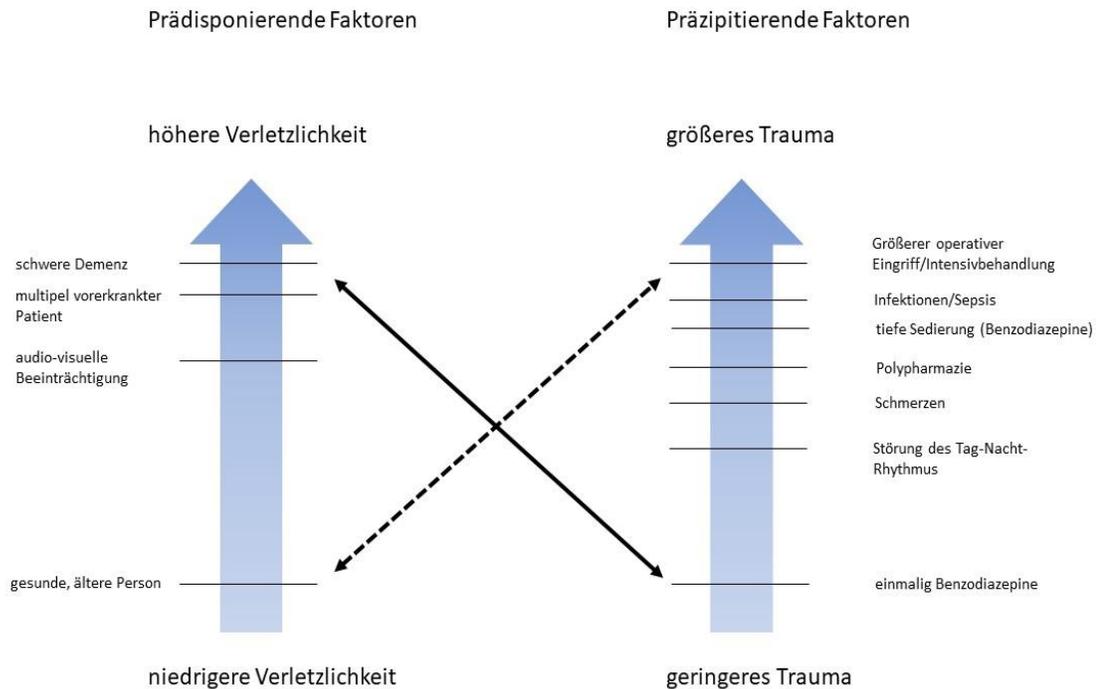


Abbildung 1.3: Modell zur Delirentstehung modifiziert nach (Inouye, 1999)

In einem Patientenkollektiv von 949 neurochirurgischen Patienten, die auf einer Normalstation behandelt wurden, konnten Zisper et al. nachweisen, dass ein Schlaganfall in der Anamnese als prädisponierender Faktor für die Delirentstehung anzusehen ist (vgl. 2019). Aus derselben Arbeitsgruppe wurden für 1487 neurologische Patienten, die auf einer Normalstation behandelt wurden, als prädisponierende Faktoren ein fortgeschrittenes Lebensalter, eine Unterbringung in Pflegeeinrichtungen vor dem aktuellen Krankenhausaufenthalt, eine Demenz und ein Substanzabusus festgestellt (vgl. Zisper et al., 2019a). Bei neurochirurgischen Patienten konnte ein Schädel-Hirn-Trauma und ein Hydrozephalus und bei neurologischen Patienten eine Epilepsie, eine intrakranielle Blutung oder ein ischämischer Schlaganfall als präzipitierender Faktor mit einem Delir assoziiert werden (vgl. Zisper et al., 2019; Zisper et al., 2019a).

1.4 Relevanz des Delirs

Die im Rahmen einer Intensivtherapie an einem Delir leidenden Patienten haben ein erhöhtes Risiko für Komplikationen. Wie Ely et al. bereits 2004 in einer prospektiven Arbeit an 275 internistischen Patienten nachwies, ist ein Delir nach Adjustierung für Einflussfaktoren mit einer längeren Krankenhausverweildauer von im Median zehn zusätzlichen Tagen und einer dreifach erhöhten Mortalität nach sechs Monaten assoziiert (vgl. Ely et al., 2004). Jeder zusätzliche Tag, an dem sich ein Patient im Delir befindet, erhöht die *Odds Ratio* für beispielsweise die individuelle Mortalität um 10% (vgl. Pisani et al., 2009), das weitere Verbleiben auf der Intensivstation um 10% sowie eine längere Behandlungsdauer im Krankenhaus um 20% (vgl. Ely et al., 2004). Bei beatmeten internistischen Patienten steigt nach Adjustierung für relevante Einflussfaktoren, wie Erkrankungsschwere, Vorerkrankungen oder Lebensalter, beim Vorliegen eines Delirs die *Odds Ratio* zu versterben auf 13. Zum Vergleich geben die Autoren die ebenso adjustierte *Odds Ratio* für ein Versterben im Rahmen eines schweren Schockgeschehens mit 12,91 an (vgl. Lin et al., 2004). Neben dem erhöhten Risiko zu versterben konnte ferner gezeigt werden, dass die Wahrscheinlichkeit deliranter Patienten, am Entlassungstag eine geistige Beeinträchtigung aufzuweisen, fast neunmal so hoch ist wie die nicht deliranter Patienten (Hazard Ratio: 9,1, $p=0,002$) (vgl. Ely et al., 2004). Bei Patienten, die ein fünftägiges Delir erlitten hatten, waren nach drei Monaten die Testergebnisse in mehreren neuropsychologischen Tests zur Abbildung der Alltagskompetenz um eine halbe Standardabweichung schlechter als bei vergleichbaren Patienten mit eintägigem Delir (vgl. Girard et al., 2010). Es wurden beispielsweise die Fähigkeit für Aufmerksamkeit und Konzentration, das Sprachverständnis und mnestiche Fähigkeiten untersucht sowie der Mini-Mental-State-Test durchgeführt (vgl. Folstein et al., 1975). Dieser Trend ist auch nach 12 Monaten weiterhin vorhanden (vgl. Girard et al., 2010). Die bisher größte multizentrische Studie zum kognitiven Outcome internistischer und chirurgischer Patienten konnte an 382 Patienten nach Intensivbehandlung dieses Ergebnis bestätigen. Es wurde gezeigt, dass 40% der Patienten, die intensivmedizinisch behandelt wurden, nach drei Monaten bei den verwendeten Tests im Median nahezu anderthalb Standardabweichungen schlechter waren als eine altersadjustierte Population (vgl. Pandharipande et al., 2013). Weitere 26% der Patienten schnitten sogar um zwei Standardabweichungen schlechter ab. Dieses Ergebnis entspricht dem Outcome nach einem mittelschweren Schädel-Hirn-Trauma bzw. einer beginnenden Demenz. Nach 12 Monaten bestätigten sich diese Befunde bei 34% bzw. 24% der untersuchten Patienten. Wenn die Patienten im Rahmen der Studie ein fünftägiges Delir entwickelten, dann schnitten sie im Median nochmals eine halbe bis eine Standardabweichung

schlechter ab als die sonstigen Patienten. Demzufolge stellt sich das Delir als unabhängiger Risikofaktor für ein schlechteres Abschneiden in den nach drei bzw. 12 Monaten verwendeten Tests dar (vgl. Pandharipande et al., 2013). Diese kognitiven Einschränkungen können im weit fortgeschrittenen Lebensalter nur eingeschränkt aufgeholt werden und erhöhen somit bei geriatrischen Patienten das Risiko für das Auftreten einer Demenz (vgl. Davis et al., 2012). So zeigten in einer Kohorte von 78 unfallchirurgischen Patienten mit hüftnaher Fraktur die Patienten, die ein perioperatives Delir erlitten, eine *Odds Ratio* von 5,66 bezüglich der Entwicklung einer Demenz innerhalb des fünfjährigen Follow-ups (vgl. Lundström et al., 2005). Unabhängig von den Einflussfaktoren erhöht ein Delir während einer intensivstationären Behandlung in einer Kohorte von 114 chirurgischen Patienten die *Odds Ratio* zur Hospitalisierung in Pflegeeinrichtungen nach intensivmedizinischer Behandlung statistisch signifikant auf 7,2 (vgl. Balas et al., 2009). Neben den schwerwiegenden individuellen Folgen entstehen auch für die Gesundheitssysteme erhebliche finanzielle Belastungen. Jeder zusätzliche Tag, den ein Patient im Delir auf einer Intensivstation gepflegt werden muss, geht mit zusätzlichen Kosten von ca. 440\$ für die Versorgung einher (vgl. Vasilevskis et al., 2018). Nach einer älteren Arbeit belaufen sich die durchschnittlichen zusätzlichen Kosten für Patienten, die ein Delir erlitten haben, im ersten Jahr unter Berücksichtigung der erhöhten Mortalität auf ca. 60.000\$ (vgl. Leslie et al., 2008).

Das im Rahmen der hier vorliegenden Arbeit untersuchte Patientenkollektiv der Neurointensivpatienten ist bisher nicht so umfassend untersucht worden wie internistische oder chirurgische Intensivpatienten. Wie Mitasova et al. nachweisen konnten, lässt sich in einem 129 Patienten umfassenden neurologischen Patientenkollektiv einer tschechischen *Stroke Unit* ebenfalls eine verlängerte Krankenhausbehandlung (18 vs. 12 Tage) und ein höherer mit dem Barthel-Index gemessener Hilfebedarf (15 vs. 37,5 Punkte) nachweisen (vgl. Mitasova et al., 2012). Nach einem sechsmonatigen Follow-up verbleibt ein weiterhin bestehender, jedoch nicht signifikanter Unterschied (27,5 vs. 75 Punkte). In den ersten sechs Monaten ist die Mortalität bei deliranten gegenüber den nicht deliranten Schlaganfallpatienten nicht signifikant erhöht (23,6% vs. 14,9%) (vgl. Mitasova et al., 2012). An einer Patientengruppe von 760 polnischen Schlaganfallpatienten einer universitären *Stroke Unit* konnten in einer retrospektiven Auswertung einer prospektiven Untersuchung diese Ergebnisse bestätigt werden (vgl. Kotfis et al., 2019). So wiesen delirante Patienten nicht nur einen längeren Krankenhausaufenthalt (12,1 vs. 10,3 Tage) und ein schlechteres funktionelles Outcome anhand des Rankin Scores (3,95 vs. 2,20) auf, sondern auch eine erhöhte Mortalität nach 7, 30 und 90 Tagen sowie einem Jahr. Zudem entwickelten sie häufiger kardiologische (38,84% vs. 13,62%), pulmonale (57,85% vs. 21,44%),

nephrologische (13,22% vs. 7,51%) und neurologische (9,09% vs. 3,60%) Komplikationen als Patienten ohne Delir (vgl. Kotfis et al., 2019).

Die mit dem Delir assoziierte erhöhte Mortalität, die erhöhte Morbidität für Komplikationen, eine geminderte kognitive Leistungsfähigkeit bis zur Begünstigung einer Demenz mit nachfolgender Unterbringung in Pflegeeinrichtungen und die hieraus folgenden erhöhten Kosten für die Sozialsysteme verdeutlichen die über den medizinischen Sektor hinausgehende Relevanz des Delirs.

1.5 Delirdiagnostik und -messinstrumente

Die Diagnose eines Delirs erfolgt in der Regel mittels einer strukturierten klinischen Anamnese und Untersuchung anhand der ICD-10- oder DSM V-Kriterien durch einen Facharzt für Psychiatrie (APA 2013).

Gemäß der aktuell vorliegenden Versorgungsleitlinie für die Intensivmedizin in Deutschland und internationaler Leitlinien, soll jeder Patient regelmäßig auf das Vorliegen eines Delirs untersucht werden. Die S3-Leitlinie *„Analgesie, Sedierung und Delirmanagement in der Intensivmedizin“* unter Federführung der Deutschen Gesellschaft für Anästhesiologie und Intensivmedizin (DGAI) und der Deutschen interdisziplinären Vereinigung für Intensiv- und Notfallmedizin (DIVI) fordert, diese Untersuchung auf das Vorliegen eines Delirs und die Dokumentation für jeden Patienten mindestens einmal pro Schicht durchzuführen. Eine derartige Untersuchung sollte in der Regel achtsündlich erfolgen und dokumentiert werden (vgl. Baron et al., 2015). Diese Empfehlung erfolgt in Übereinstimmung mit den amerikanischen Empfehlungen des *College of Critical Care Medicine* (vgl. Barr et al., 2013; Devlin et al., 2018), den britischen Guidelines des NICE-Instituts (NICE 2020) und den Empfehlungen der *European Society of Intensive Care Medicine* unter Verweis auf die vorgenannte S3-Leitlinie (ESCIM 2016). Diese routinemäßig vorgenommene Untersuchung auf das Vorliegen eines Delirs wird als Delirscreening bezeichnet.

Um jeden Patienten in den durch die Leitlinien geforderten Intervallen durch einen Facharzt für Psychiatrie nach dem Referenzstandard ICD-10 oder DSM V untersuchen zu können, wäre dieser im Schichtdienst auf der Intensivstation vorzuhalten. Dies ist allerdings nicht für alle intensivmedizinisch behandelten Patienten realisierbar, da hierfür eine große Menge an zusätzlichen personellen Ressourcen erforderlich wäre.

Um ein Delirscreening routinemäßig zu etablieren, ist die Entwicklung von Messinstrumenten für Intensivpatienten unter Berücksichtigung der entsprechenden Krankheitsbilder und des speziellen intensivmedizinischen Settings sinnvoll und notwendig (vgl. Ely et al., 2001a). Im Idealfall ist das Delirscreening durch das ohnehin im regulären Betrieb auf der Station vorhandene klinische Personal binnen einer oder weniger Minuten bettseitig durchführbar. Eine Notwendigkeit für apparative oder weitergehender Untersuchungen sollte nicht bestehen. Diese Forderungen tragen vor allem der hohen Arbeitsdichte in diesem Gebiet der Patientenversorgung Rechnung. Zudem sollte es möglich sein, intubierte und beatmete Patienten zu untersuchen (vgl. Ely et al., 2001). Diese Messinstrumente sollten in prospektiven Studien am jeweiligen Patientenkollektiv validiert worden sein und eine gute Reliabilität aufweisen (vgl. Pandharipande et al., 2005).

Zur Validierung verschiedener Messinstrumente wird eine richtlinienkonforme Übersetzung in die jeweilige Landessprache, eine kulturelle Adaptation gemäß den Kriterien der *International Society for Pharmacoeconomics and Outcome Research* (vgl. Wild et al., 2005) und eine verblindete Testung gegen den Referenzstandard gefordert (vgl. Luetz et al., 2010a).

Für eine sichere und valide Anwendung der Messinstrumente und die valide Durchführung eines entsprechenden Screenings im Routinebetrieb ist eine strukturierte Schulung im Rahmen des Implementierungsprozesses notwendig. Dies konnte bereits für Scores zur Beurteilung von Sedierungstiefe, Schmerz und Delir im Bereich der Intensivmedizin nachgewiesen werden. Hierzu stellten Radtke et al. ein modifiziertes Trainingskonzept vor (vgl. 2012). Während das vierteilige Standardtraining aus je einem abgehaltenen Vortrag sowie einem Beispielvideo je Score, Handouts zu den geschulten Scores und einer Einzelunterweisung am Bett bestand, wurde dieses im Interventionsarm zweimalig vierwöchentlich wiederholend und vertiefend vorgenommen. Es zeigte sich, dass die untersuchten Scores bei einer repetitiven Schulung häufiger auch 16 Wochen nach dem initialen Training durchgeführt wurden. Zusätzlich konnte nachgewiesen werden, dass eine Reduktion der Mortalität mit der häufigeren Durchführung der Scores zum Delirscreening und für die Beurteilung von Schmerzen assoziiert ist (vgl. Radtke et al., 2012). Van Eijk et al. konnten zeigen, dass selbst ein Delirmessinstrument, welches sonst gute Testgütekriterien zeigt, in einem multizentrischen Ansatz auf mehreren Intensivstationen unterschiedlicher Versorgungsstufen im Routinebetrieb schlechter abschneiden kann (vgl. 2011). Mehrere Autoren führen dies auf das fehlende Schulungskonzept und eine andere Durchführung des Screenings zurück (vgl. Patel and Kress, 2011; Vasilevskis et al., 2012).

Devlin et al. konnten nachweisen, dass sich bereits durch eine eintägige Schulung für 50 Intensivpflegekräfte die Verwendung (12% vs. 82%, $p < 0,005$) und die korrekte Anwendung eines Messinstruments (8% vs. 62%, $p < 0,005$) um den Faktor 7,5 steigern lässt. Die korrekte klinische Einschätzung zum Vorliegen eines Delirs konnte durch die hier durchgeführte Intervention um den Faktor 4,5 gesteigert werden (24% vs. 100%, $p < 0,005$) (vgl. Devlin et al., 2008). An 25 ärztlichen Intensivmedizinern einer internistischen Intensivstation wurde gezeigt, dass eine Schulung, bestehend aus einer folienbasierten Intervention (20 Folien), die Detektionsrate im Vergleich zu einem auf einem Messinstrument basierenden Screening durch einen erfahrenen Intensivmediziner für die Ärzte um 70% erhöhte (κ prä Schulung -0,14 vs. κ post Schulung 0,67) (vgl. Devlin et al., 2007).

Die verschiedenen Ausprägungen des Delirs sind von klinischer Bedeutung. So wird zwischen einem hyper- und einem hypoaktiven Delir unterschieden (vgl. Ely et al., 2001a). Je nach Patientenkollektiv werden in der Literatur Prävalenzen zwischen 43,5% und 67% für das hypoaktive und von 1% bis 1,6% für das hyperaktive Delir beschrieben (vgl. Peterson et al., 2006; Pandharipande et al., 2007; Robinson et al., 2011). Somit tritt das hypoaktive Delir deutlich häufiger auf als das hyperaktive. In einer kürzlich veröffentlichten Metaanalyse konnte gezeigt werden, dass insbesondere beatmete und schwer erkrankte Intensivpatienten eine erhöhte Prävalenz (35% bzw. 29%) für ein hypoaktives Delir aufweisen (vgl. Krewulak et al., 2018). Es gibt Hinweise darauf, dass das hypoaktive Delir mit einem schlechteren Outcome, wie beispielsweise einer Aspiration, Pneumonie, Immobilisation und nachfolgender Lungenembolie, assoziiert sein kann (vgl. O’Keeffe and Lavan, 1999; Ely et al., 2001a; Robinson et al., 2011). Ferner existiert ein Mischtyp der beiden Formen. Der Übergang zwischen den unterschiedlichen Typen ist fließend und jederzeit möglich (vgl. Pandharipande et al., 2007). Da es sich, wie bereits gezeigt wurde, beim Delir des Intensivpatienten keinesfalls um eine Blickdiagnose handelt (vgl. van Eijk et al., 2009) und der Patient im hypoaktiven Delir klinisch schwer zu erfassen ist, muss ein Delirmessinstrument dazu in der Lage sein, alle Subtypen des Delirs zuverlässig zu detektieren (vgl. Luetz et al., 2010a).

Mittlerweile sind zahlreiche validierte Delirmessinstrumente für das Screening veröffentlicht worden. Hierzu zählen die *Confusion Assessment Method for the Intensive Care Unit* (CAM-ICU) (vgl. Ely et al., 2001) und die *Intensive Care Delirium Screening Checklist* (ICDSC) (vgl. Bergeron et al., 2001), wobei darüber hinaus noch weitere Instrumente zur Detektion eines Delirs existieren. Die bereits zitierte S3-Leitlinie (vgl. Baron et al., 2015) nennt die *Nursing Delirium Screening Scale* (Nu-

DESC) (vgl. Gaudreau et al., 2005) und den *Delirium Detection Score* (DDS) (vgl. Otter et al., 2005). Im weiteren Verlauf soll jedoch vornehmlich auf die im Rahmen dieser Arbeit verwendete CAM-ICU eingegangen werden. Sowohl CAM-ICU als auch ICDSC sind zwischenzeitlich in einer kulturell adaptierten und gemäß den Kriterien der *International Society for Pharmacoeconomics and Outcome Research* (vgl. Wild et al., 2005) in die deutsche Sprache übersetzten Form verfügbar (vgl. Radtke et al., 2009; Guenther et al., 2010). Diese sind für Patientenkollektive aus unterschiedlichen medizinischen Disziplinen validiert und haben wiederholt gute bis sehr gute Testgütekriterien gezeigt (vgl. Ely et al., 2001; Luetz et al., 2010; Mitsova et al., 2012).

Neben dem beschriebenen Vollbild des Delirs mit seinen verschiedenen Subtypen gibt es in der Forschungsliteratur für geriatrische Patienten Hinweise auf das Vorliegen eines Symptomkomplexes, der sich fakultativ zu einem Delir entwickeln kann (vgl. Levkoff et al., 1996; Cole et al., 2003). Dieser wurde als subsyndromales Delir (SSD) bezeichnet und erstmalig in geriatrischen Patientenkollektiven nachgewiesen (vgl. Levkoff et al., 1996; Cole et al., 2003). Allerdings ist die Definition des SSD bei geriatrischen Patienten in der Literatur uneinheitlich (vgl. Levkoff et al., 1996; Cole et al., 2003). So zeigten diese Patienten nicht alle für die Diagnose des Delirs notwendigen Veränderungen, aber es existieren Anzeichen für eine Beeinträchtigung in mindestens einem der oben beschriebenen Merkmale bzw. in einem Merkmal des DSM III (vgl. Levkoff et al., 1996). Es konnte nachgewiesen werden, dass Patienten, die ein SSD erlitten, eine längere Krankenhausverweildauer, eine erhöhte Mortalität und eine höhere Unterbringungsrate in stationären Pflegeeinrichtungen hatten (vgl. Cole et al., 2003).

In der aktuellen Definition nach DSM V wird erstmalig das Delir als Erkrankung mit einem gewissen Spektrum der Erkrankungsschwere mit der Möglichkeit der Diagnose eines abgeschwächten Delirs beschrieben (vgl. APA, 2013). Eine Subsumption des SSD unter dieser Diagnose als abgeschwächtes Delir war jedoch nicht angestrebt (vgl. Serafim et al., 2017). Ebenso erreicht die im DSM V angebotene Definition des abgeschwächten Delirs nicht das Niveau der Definition des Delirs nach DSM V und ist in der internationalen Literatur nicht einheitlich akzeptiert (vgl. Serafim et al., 2017).

Im Jahr 2007 konnten Quimet et al. nachweisen, dass auch bei Intensivpatienten ein SSD auftreten kann (vgl. Ouimet et al., 2007a). Dabei wurde ein positives Ergebnis in mindestens einem der Merkmale der ICDSC als SSD bewertet. Die Patienten, die ein SSD erlitten, zeigten einen doppelt so langen Aufenthalt auf der Intensivstation (5,2 vs. 2,5 Tage; $p < 0,0001$) und eine nahezu fünffach erhöhte

Mortalität (10,6% vs. 2,4%; $p=0,0020$) im Vergleich zu Patienten, die weder ein Delir noch ein SSD erlitten. Im Vergleich zu den Patienten, die die Kriterien zur Diagnose eines Delirs nach ICDSC erfüllten, wurden die Patienten mit SSD halb so lange auf der Intensivstation behandelt (5,2 vs. 10,8 Tage; $p<0,0001$). Die Mortalität war um 5,3% geringer (10,6% vs. 15,9%; $p=0,1378$) (vgl. Ouimet et al., 2007a). Somit gibt es Hinweise darauf, dass intensivmedizinische Patienten, die ein SSD ausbilden, einen günstigeren klinischen Verlauf haben als Patienten, die ein Delir erleiden. Dennoch ist das Vorliegen eines SSD mit einem schlechteren Outcome assoziiert, wie beispielsweise einer Unterbringung in Pflegeeinrichtungen und einer längeren Behandlungsdauer auf der Intensivstation und im Krankenhaus (vgl. Ouimet et al., 2007a). Eine kürzlich veröffentlichte Metaanalyse bestätigt, unter Betrachtung von sechs relevanten Studien, einen längeren Krankenhausaufenthalt bei Patienten mit SSD, zeigt jedoch uneinheitliche Daten für eine Erhöhung der Mortalität (vgl. Serafim et al., 2017). In Bezug auf Intensivpatienten ist die Bewertung des SSD als Vorstufe zum Delir nicht einheitlich. Ouimet et al. geben an, dass es sich hierbei um eine dritte Kategorie zwischen ‚Delir‘ und ‚kein Delir‘ handele (vgl. 2007a). Auch die Definition des SSD bei Intensivpatienten ist uneinheitlich. So definierten Ouimet et al. dieses anhand eines positiven Merkmals auf der ICDSC, während Brummel et al. beispielsweise ein positives Merkmal der CAM-ICU als SSD bewerten (vgl. Ouimet et al., 2007a; Brummel et al., 2017).

Die Auswahl des Messinstruments sollte sich am Patientenkollektiv, der Validierung für das jeweilige Patientenkollektiv und der Intention für die Implementierung (z.B. hochsensitiver Suchtest oder Bestätigungstest, Art der Anwendung und Zielgruppe der Anwender) orientieren (vgl. Luetz et al., 2010a).

1.5.1 Skalenniveau der Delirmessinstrumente

In Abhängigkeit davon, welches Delirmessinstrument verwendet wird, kommen verschiedene Messniveaus zum Einsatz. Während bei der Verwendung eines dichotomen Instruments lediglich eine Diagnosefeststellung möglich ist, erlaubt ein ordinales Messinstrument zusätzlich eine Schweregradeinschätzung und eine Verlaufsbeschreibung (vgl. Brummel et al., 2017; Khan et al., 2017).

Klinisch stellt sich häufig die Frage, ob im Rahmen einer Therapie eine Verbesserung der Symptomatik erzielt wurde bzw. ob eine pharmakologische Therapie die erwartete Wirkung zeigt und ggf. beendet werden kann. Um dies festzustellen, kann ein ordinaler Score besser geeignet sein, da auch feinere Unterschiede oder Verläufe anhand numerischer Werte abgebildet werden können (vgl.

Inouye et al., 2014). Sofern ausschließlich ein dichotomes Messinstrument verwendet wird, ist bei einer Bewertung als ‚delirant‘ oder ‚nicht delirant‘ eine Erfassung feinerer Unterschiede erschwert, weshalb beispielsweise ein SSD in bis zu 10% der Fälle übersehen werden kann (vgl. Brummel et al., 2017).

Es gibt erste Hinweise darauf, dass bereits bei der Diagnose eines SSD eine Therapie durchgeführt werden kann und dies Einfluss auf das Outcome des Patienten haben kann. So verabreichten Hakim et al. in einer doppelblinden, placebokontrollierten, monozentrischen Studie an 101 Patienten nach elektiven herzchirurgischen Eingriffen unter Einsatz der Herz-Lungen-Maschine bereits bei Feststellung eines SSD entweder Risperidon oder Placebo (vgl. 2012). Das führte dazu, dass in der Interventionsgruppe weniger Patienten ein Delir entwickelten (13,7% vs. 34%, $p=0,031$). Eine Auswirkung auf die sekundäre Endpunkte Dauer und Schwere eines Delirs, die Behandlungsdauer auf der Intensivstation und im Krankenhaus sowie das Auftreten medikamentenspezifischer Nebenwirkungen konnte jedoch nicht nachgewiesen werden. Die Autoren schlussfolgerten daraus, dass durch die Verabreichung von Risperidon bereits bei Feststellung eines SSD das Auftretens eines Delirs verringert werden kann (vgl. Hakim et al., 2012). Gemäß einer kürzlich veröffentlichten Metaanalyse gibt es jedoch aufgrund geringer Studienzahlen und heterogenem Studiendesign in Verbindung mit hochselektierten Patientenkollektiven keine ausreichende Evidenz, um eine pharmakologische Delirprophylaxe bereits bei Vorliegen eines SSD zu empfehlen (vgl. Tao et al., 2018). Zudem befürworten die Autoren der zitierten Metaanalyse aufgrund der heterogenen Datenlage weitere Untersuchungen, die aufgrund einer größeren Datenbasis, eines homogeneren Studiendesigns und homogener Einschlusskriterien weitere Evidenz liefern könnten.

Mit der Verwendung eines ordinalen Messniveaus wird dem Vorliegen eines differenzierteren Bildes neurokognitiver Einschränkungen bei Intensivpatienten Rechnung getragen. Auf diese Weise wird neben der Diagnosestellung für ein Delir die Möglichkeit zur Feststellung eines SSD eröffnet. Zudem lassen sich die Folgen einer Intervention und ein klinischer Verlauf besser abschätzen. So bietet beispielsweise die ICDSC ein ordinales Messniveau und bildet auch das SSD ab (vgl. Ouimet et al., 2007a). Jedoch ist bisher keine Studie veröffentlicht worden, die untersucht, ob die ICDSC durch repetitive Messungen zur Einschätzung des klinischen Verlaufs und zur Abbildung therapeutischer Interventionen als Ordinalskala geeignet ist (vgl. Hakim et al., 2012).

Falls jedoch die Anwendung eines Delirmessinstrumentes auf einer Station bereits etabliert wurde, so stellt die Umstellung des Delirscreenings auf ein anderes Messinstrument aufgrund des zuvor beschriebenen Schulungsaufwands eine hohe logistische Hürde dar. Bei Verwendung der weit verbreiteten, in der Literatur umfangreich untersuchten und für verschiedene Patientenkollektive validierten CAM-ICU liegt bisher nur ein dichotom nominales Messniveau vor. Um die zuvor genannten Vorteile eines ordinalen Messinstrumentes zu nutzen und möglichst umfangreiche Ergebnisse aus der Anwendung nur eines Messinstrumentes zu erlangen, wird versucht, dieses Messniveau durch Modifikationen der CAM-ICU ebenfalls zu erreichen bzw. die Detektion eines SSD zu ermöglichen (vgl. Brummel et al., 2017; Khan et al., 2017).

1.6 Die CAM-ICU

Bereits 1990 publizierten Inouye et al. eine Methode, um internistische und geriatrische Patienten, die nicht auf einer Intensivstation behandelt werden mussten, routinemäßig auf das Vorliegen eines Delirs zu untersuchen. Die sogenannte *Confusion Assessment Method (CAM)* wurde gegen den damaligen Goldstandard, das DSM III R, validiert und zeigte gute Testgütekriterien (vgl. Inouye, 1990). Um auch bei beatmeten Patienten eine einfache, zeitsparende und verlässliche Möglichkeit zur Detektion des Delirs ohne Konsultation eines Psychiaters zur Verfügung zu stellen, entwickelten Ely et al. im Jahr 2001 die gegen den Goldstandard DSM IV TR validierte CAM-ICU (vgl. 2001). Hierbei handelt es sich um einen dichotomen Score, der eine Unterscheidung in ‚Delir‘ und ‚kein Delir‘ zulässt. Gemäß der originalen Arbeit von Ely et al. werden stets vier Merkmale in der immer gleichen Reihenfolge überprüft (vgl. 2001). Dies wurde als CAM-ICU-Arbeitsblatt veröffentlicht und ist online in verschiedenen Sprachen mit Schulungsmaterial und für die Durchführung notwendigem Material abrufbar (vgl. www.icudelirium.org, 2020). Zur Verkürzung der Durchführungszeit liegt eine leicht im Ablauf modifizierte Version der CAM-ICU mit dem CAM-ICU *Flowsheet* vor. Im Gegensatz zur originalen Publikation müssen bei der Verwendung des CAM-ICU *Flowsheets* nicht mehr bei jedem Untersuchungsgang alle vier Merkmale überprüft werden. Eine ausführliche Darstellung der einzelnen Merkmale und des Untersuchungsgangs der CAM-ICU folgt im Methodenteil dieser Arbeit.

Die CAM-ICU steht als richtlinienkonforme Übersetzung und als kulturell adaptierte Version in deutscher Sprache zur Verfügung (vgl. Guenther et al., 2010) und ist mittlerweile in verschiedene weitere Sprachen übersetzt worden. Zwischenzeitlich erfolgte die Validierung für unterschiedliche

Patientenkollektive und die Bestimmung der Testgütekriterien (vgl. Ely et al., 2001a; Guenther et al., 2010; Luetz et al., 2010; Mitasova et al., 2012). In Abhängigkeit vom Patientenkollektiv und dem verwendeten Referenzstandard konnten gute bis sehr gute Testgütekriterien nachgewiesen werden. So erreichten Ely et al. in der ursprünglichen Validierungsstudie an einem Patientenkollektiv von 38 internistischen Intensivpatienten eine Sensitivität von 95-100% und eine Spezifität von 89-93%. Eine hohe Übereinstimmung zwischen verschiedenen Untersuchern konnte mit einem κ von 0,79-0,95 erreicht werden (vgl. Ely et al., 2001). Für das CAM-ICU *Flowsheet* zeigten Guenther et al. in der deutschen Übersetzung an einem 54 Patienten umfassenden Patientenkollektiv einer chirurgischen Intensivstation einer Universitätsklinik eine Sensitivität von 88-92% und eine Spezifität von 100% bei einer *Interrater Reliability* κ von 0,96 (vgl. 2010).

Neben der hohen Interrater-Reliabilität der CAM-ICU besteht eine hohe Detektionsrate verschiedener Subtypen des Delirs und eine Durchführbarkeit innerhalb von 45-50 Sekunden bis zu wenigen Minuten (vgl. Pun et al., 2005; Guenther et al., 2010). Diese Zeit lässt sich durch die Verwendung des CAM-ICU *Flowsheet* nochmals senken.

Für die Anwendung im Routinebetrieb durch reguläre Intensivpflegekräfte ohne vorangegangene Schulung gibt es Hinweise auf eine Verschlechterung der Testgütekriterien. So konnte in einer multi-zentrischen Arbeit an zehn Institutionen unterschiedlicher Versorgungsstufen in der täglichen Routine eine Sensitivität von 47% und eine Spezifität von 95% nachgewiesen werden (vgl. van Eijk et al., 2011). Für nicht intensivpflichtige Patienten gibt es Hinweise darauf, dass die CAM-ICU schlechter geeignet sein könnte als andere Scores (vgl. Neufeld et al., 2011).

Um den bereits dargelegten Problemen bei der Verwendung eines dichotomen Scores entgegenzuwirken, gibt es Bemühungen, beispielsweise auch ein SSD anhand der CAM-ICU diagnostizieren zu können (vgl. Brummel et al., 2017). So haben Brummel et al. im Rahmen ihrer Studie jedes als positiv bewertete Merkmal der CAM-ICU als Hinweis auf ein SSD angesehen (vgl. 2017). Nach dieser Definition erlitten von 800 der untersuchten internistischen und chirurgischen Intensivpatienten insgesamt 86% während ihrer Behandlung auf der Intensivstation ein SSD mit einer medianen Dauer von 3 Tagen. Einer von vier Patienten entwickelte ausschließlich ein SSD, ohne jemals alle Kriterien für die Diagnose des Delirs mit der CAM-ICU zu erfüllen. Die Anzahl der Tage, die ein Patient an einem SSD litt, waren mit einer erhöhten Rate an Unterbringungen in Pflegeeinrichtungen nach Entlassung aus dem Krankenhaus assoziiert. Dieser Effekt blieb auch nach Adjustierung für relevante

Einflussfaktoren erhalten. Ein Patient, der niemals als delirant diagnostiziert wurde, aber 5 Tage als SSD-positiv klassifiziert wurde, hatte im Vergleich zu einem Patienten, der nur 1,5 Tage an einem SSD litt, eine erhöhte *Odds Ratio* von 4,2 (IQR 1,8-9,7), nach seiner Krankenhausbehandlung in einer Pflegeeinrichtung untergebracht zu werden. Daher schlussfolgern die Autoren, dass aktiv nach einem SSD gesucht werden sollte und der bloße Verdacht auf das Vorliegen eines SSD die Suche nach klinisch modifizierbaren Risikofaktoren für die Entwicklung eines Delirs rechtfertigt (vgl. Brummel et al., 2017). Dabei ist jedoch zu berücksichtigen, dass die hier verwendete Definition des SSD als sehr inklusiv angesehen werden muss. Bei der Verwendung einer restriktiveren Definition, die alle Patienten, die jemals ein Delir erlitten, in eine separate Gruppe ohne Berücksichtigung eines etwaig vorliegenden SSD einteilt, ergibt sich eine geringere Prävalenz für ein SSD von 26% (vgl. Brummel et al., 2017).

Um den zweiten Vorteil einer dezidierten Verlaufs- und Schwerebeschreibung des Delirs mittels der CAM-ICU zu realisieren, gibt es weitere Ansätze, um ein ordinales Messniveau zu erreichen. So stellten Khan et al. im Jahr 2017 die CAM-ICU-7 vor (vgl. 2017). Hierbei handelt es sich um eine aus den Daten der CAM-ICU berechnete Ordinalskala, die eine Abbildung der Schwere des Delirs ermöglichen soll. Zur Anwendung der CAM-ICU-7 muss eine Untersuchung gemäß des CAM-ICU-Arbeitsblattes mit der Erhebung aller vier Merkmale vorgenommen werden. Diesen Ergebnissen wird nun ein numerischer Wert zwischen null und zwei Punkten zugeordnet, wobei das Merkmal ‚fluktuierender Verlauf‘ oder ‚plötzlicher Beginn‘ eine Ausnahme bildet, da hier null bzw. ein Punkt vergeben werden. Wenn keine Fehler gemacht werden, erhält der Patient null Punkte, während sieben die maximal mögliche Punktzahl darstellt. Die Autoren schlagen eine Einteilung vor, wonach null bis zwei Punkte keinem Delir, drei bis fünf Punkte einem milden bzw. moderaten und sechs bzw. sieben Punkte einem schweren Delir entsprechen. Bezogen auf die als Referenzmodell zur Abschätzung der Delirschwere verwendeten *Delirium Rating Scale-Revised-98* (DRS-R-98) (vgl. Trzepacz et al., 2001), einem 16 Items umfassenden Messinstrument zur Detektion der Delirschwere, erreichte die CAM-ICU-7 ein Cronbachs alpha von $\alpha=0,85$. Dies entspricht einer guten bis sehr guten Korrelation (vgl. Khan et al., 2017). Da sich diese Studie auf die Untersuchung von deliranten Patienten beschränkte, ist keine Aussage zur Möglichkeit der Einschätzung eines unter Umständen vorliegenden SSD möglich.

Einen anderen Weg zur Beschreibung der Ausprägung der Symptome des Delirs eines Patienten publizierten Luetz et al. (vgl. 2016). So konnten sie in einem Kollektiv von 64 pädiatrischen Patienten, die auf einer universitären pädiatrischen Intensivstation behandelt werden mussten, in einer prospektiven Arbeit im Rahmen der Implementierung der für den pädiatrischen Bereich optimierten Version der CAM-ICU (pCAM-ICU) eine Schweregradskala entwickeln. Die hier vorgestellte Skala wurde als *severity scale for the pCAM-ICU* (sspCAM-ICU) benannt. Sie basiert auf einer vollständigen Erhebung der pCAM-ICU und errechnet aus den Ergebnissen des jeweiligen Merkmals einen Punktwert. Beim fehlerfreien Absolvieren der pCAM-ICU sind null Punkte zu erreichen, während ein Delir maximaler Ausprägung einem Punktwert von 19 Punkten entsprechen würde. Im Rahmen einer Evaluation des ersten Messtages zeigte die sspCAM-ICU im Vergleich zur pCAM-ICU eine bessere Sensitivität (84,6% vs. 76,9%) und eine vergleichbare Spezifität (98% vs. 98%). Bei den repetitiv vorgenommenen Messungen zeigte sie eine deutlich bessere Sensitivität (71,8% vs. 52,3%) bei vergleichbarer Spezifität (96,6% vs. 99,4%). Eine Validierung dieser ordinalen Skala für die CAM-ICU hat bei adulten Neurointensivpatienten nicht stattgefunden. Bisher ist lediglich eine Arbeit veröffentlicht worden, die dieses Messinstrument bei erwachsenen Patienten verwendet (vgl. Nacul et al., 2020).

1.7 Delirscreening bei neurologischen und neurochirurgischen Intensivpatienten

Ein wirksames und effektives Screening für das Vorliegen eines Delirs setzt die Existenz geeigneter und überprüfter Messinstrumente voraus. Jedoch sind Patienten mit neurologischer bzw. neurochirurgischer Erkrankung in den meisten Validierungsstudien *a priori* ausgeschlossen worden (vgl. Ely et al., 2001; Guenther et al., 2010; Luetz et al., 2010). So wurden in Bezug auf das spezielle Klientel der Neurointensivpatienten lediglich sieben Validierungsstudien veröffentlicht.

Bereits im Jahr 2009 untersuchten Van Eijk et al. im Rahmen einer Validierungsstudie an den Patienten einer interdisziplinären, universitären Intensivstation ein Patientenkollektiv von 126 Patienten, in das 28 neurologische bzw. neurochirurgische Patienten eingeschlossen waren. Die Messung erfolgte für die CAM-ICU, die ICDSC und den klinischen Eindruck des behandelnden Intensivmediziners gegen ein Expertenurteil nach den Kriterien des DSM IV TR. Für das Gesamtkollektiv wurde eine Sensitivität von 64%, 43% bzw. 29% und eine Spezifität von 88%, 95% bzw. 96% nachgewiesen (vgl. van Eijk et al., 2009). Eine nähere Betrachtung der Subgruppe erfolgte dabei jedoch nicht.

In einer prospektiven Validierungsstudie an 129 neurologischen Patienten einer tschechischen *Stroke Unit* verglichen Mitasova et al. die CAM-ICU mit einer Diagnose des Delirs nach dem DSM IV TR. Dabei konnten günstige Testgütekriterien mit einer Sensitivität von 76% und einer Spezifität von 98% nachgewiesen werden (vgl. Mitasova et al., 2012).

An einem kleinen Patientenkollektiv von 61 Patienten zweier Traumazentren, die ein leichtes bis mittelschweres Schädel-Hirn-Trauma erlitten hatten, konnte in einer prospektiven Untersuchung eine Validierung für die CAM-ICU und die ICDSC gegen die DSM IV erreicht werden. Mit einer Sensitivität von 62% bzw. 64% und einer Spezifität von 74% bzw. 79% unterschieden sich beide Messinstrumente nach Ansicht der Autoren nicht in den Testgütekriterien. Bei der Betrachtung der Subgruppen der Patienten mit einem mittelschweren Schädel-Hirn-Trauma oder der tiefer sedierten Patienten zeigten diese jedoch eine geringere Spezifität (vgl. Frenette et al., 2015).

Die bisher einzige Arbeit an einem breiten neurologischen bzw. neurochirurgischen Patientenkollektiv vergleicht monozentrisch und prospektiv an 49 Patienten die CAM-ICU und die ICDSC mit der Diagnose nach der ICD-10. Dabei konnte ein deutlicher Unterschied zwischen CAM-ICU und ICDSC nachgewiesen werden, wobei die Sensitivität mit 59% bzw. 85% und die Spezifität mit 56% bzw. 75% angegeben wurde (vgl. Larsen et al., 2019). Hierbei ist allerdings zu berücksichtigen, dass der Goldstandard nicht das DSM, sondern die Kriterien nach der ICD-10 waren.

Reznik et al. verglichen die CAM-ICU und die ICDSC prospektiv mit den Kriterien des DSM V an einem Patientenkollektiv von 60 Patienten, die eine intracerebrale Blutung erlitten hatten (vgl. 2020). Dabei zeigten die Messinstrumente gute bis akzeptable Testgütekriterien für die CAM-ICU und die ICDSC, sofern für die CAM-ICU die Untersuchung nach der Originalpublikation durchgeführt wurde. Die Sensitivität betrug 67,9% bzw. 77,1% und die Spezifität 88% bzw. 96,6%. Die Autoren testeten ebenfalls eine andere Interpretation der CAM-ICU. Während nach dem ursprünglichen Ansatz nur Patienten nicht untersucht werden können, die mindestens physische Stimuli benötigen, um eine Reaktion zu zeigen (RASS <-3), konnte in der Interpretation der Autoren auch eine neurologische Einschränkung aufgrund eines fokalneurologischen Defizits (z.B. Abulie oder Aphasie) eine weitere Anwendung des Messinstruments verhindern. Bei dieser Methode wird für die CAM-ICU eine Sensitivität von 40,7% und eine Spezifität von 88% angegeben (vgl. Reznik et al., 2020). Bei der Interpretation dieser Daten muss auch dem neueren Goldstandard nach DSM V Rechnung getragen werden.

Eine prospektive Validierung der ICDSC gegen das DSM V wurde an einer deutschen *Stroke Unit* an 346 Patienten unter besonderer Berücksichtigung des Anteils von Patienten mit einer Aphasie vorgenommen. In dem nicht aphasischen Patientenkollektiv von 115 Patienten konnte für das Delir eine Prävalenz von 27,8% ermittelt werden. Die ICDSC erreichte hier eine Sensitivität von 100% und eine Spezifität von 78% gegenüber dem Goldstandard. In der Subgruppe der 241 aphasischen Patienten zeigte sich eine Delirprävalenz von 36%. Für die ICDSC wurde in dieser Subgruppe die Sensitivität mit 98% und die Spezifität mit 55% angegeben. Durch Veränderung des Cut-Off Wertes von ≥ 4 Punkten zur Diagnose des Delirs nach ICDSC auf ≥ 5 Punkte konnte bei aphasischen Patienten eine Sensitivität von 90% und eine Spezifität von 75% erreicht werden (vgl. Boßelmann et al., 2019). Auch hier erfolgte die Testung gegen die neueren Kriterien des DSM V.

Von Hofen-Hohloch et al. verglichen die ICDSC und die CAM-ICU gegen die klinische Einschätzung des behandelnden Teams nach Schulung für ICD-10 und DSM V unter Berücksichtigung von neurologischen Komorbiditäten, wie Aphasie, Neglect/Hemineglect, veränderter Bewusstseinslage oder mangelndem Sprachverständnis (vgl. 2020). Die Untersuchung wurde an 123 Patienten einer universitären *Stroke Unit* und neurologischen Intensivstation durchgeführt. Insgesamt 58,5% der Patienten wurden wegen eines Schlaganfalls behandelt und 14,6% wurden beatmet. In Abhängigkeit vom verwendeten Messinstrument ergab sich im Vergleich zum Referenzstandard eine leicht erhöhte Delirprävalenz (18,7%, CAM-ICU: 23,6%, ICDSC 26,8%). Die CAM-ICU und die ICDSC zeigten vergleichbare Testgütekriterien mit einer Sensitivität von 66,9% vs. 69,9% und einer Spezifität von 93,3% vs. 93,9%. Unter Berücksichtigung der neurologischen Komorbiditäten waren die Merkmale ‚eingeschränktes Sprachverständnis‘ und ‚Hemineglect‘ signifikant häufiger mit einer positiven CAM-ICU assoziiert, was sich für die ICDSC nicht nachweisen ließ. Bei beiden Messinstrumenten war eine veränderte Bewusstseinslage mit einem RASS ≤ -1 und ≥ -3 und für die ICDSC ein RASS ≥ 1 häufiger mit einem positiven Ergebnis assoziiert. Die Autoren schlussfolgern daraus, dass beide Messinstrumente die Besonderheiten dieses neurologischen Patientenkollektivs nicht gut abbilden und empfehlen daher die Entwicklung spezieller Messinstrumente unter Berücksichtigung neurologischer Komorbiditäten (vgl. von Hofen-Hohloch et al., 2020).

Zusammenfassend ist festzustellen, dass keine breit angelegte Validierungsstudie an einem neurologischen bzw. neurochirurgischen Patientenkollektiv existiert, wobei die Schlaganfallpatienten das bisher am besten untersuchte Patientenkollektiv dieser Gruppe darstellen.

1.8 Zielsetzung dieser Arbeit

Beim Delir des Intensivpatienten handelt es sich um die häufigste syndromatisch-psychiatrische Manifestation einer sekundären Enzephalopathie auf der Intensivstation mit weitreichenden Folgen für den einzelnen Patienten und das gesamte Gesundheitssystem. Für das Delirscreening sind am jeweiligen Patientenkollektiv validierte Messinstrumente zu verwenden. Die CAM-ICU stellt ein weit verbreitetes und gut untersuchtes Messinstrument mit nominalem Skalenniveau dar. Bei der Verwendung von Messinstrumenten mit ordinalem Skalenniveau ist eine Erfassung des SSD, eine Schweregrad-einteilung und eine Verlaufsbeschreibung nach therapeutischen Interventionen möglich. Im Rahmen dieser Arbeit soll als Primärhypothese überprüft werden, inwieweit eine in der Literatur bereits vorgestellte Schweregradskala, die ssCAM-ICU, für das Patientenkollektiv der Neurointensivpatienten validiert werden kann. Im Rahmen der Sekundärhypothese sollen ferner die Testgütekriterien des neuen Messinstruments sowohl bei neurochirurgischen Patienten als auch bei Patienten mit Schlaganfall separat betrachtet werden. Dies erfolgt insbesondere, um eine Übertragbarkeit der Ergebnisse auf andere Bereiche, wie beispielsweise *Stroke Units*, zu ermöglichen.

2. Material und Methoden

Nachfolgend werden das Konzept und die genauen zeitlichen Abläufe während der Datenerhebung dargestellt. Im Anschluss erfolgt die Darstellung der verwandten Messinstrumente, eine Betrachtung angewandter Limitationen der Messinstrumente und die Darlegung der statistischen Auswertung.

2.1 Patientenkollektiv, Ein- und Ausschlusskriterien, Erhebungszeiträume

Bereits 2010 hatte die Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin, Campus Virchow-Klinikum und Charité Campus Mitte, auf Grundlage der deutschen Versorgungsleitlinie „Analgesie, Sedierung und Delirmanagement in der Intensivmedizin“ (Martin et al., 2010) ein mehrmals tägliches Delirmonitoring für jeden Intensivpatienten in die *Standard Operating Procedures* der Klinik übernommen. Auf der neurologischen und neurochirurgischen Intensivstation 1i am Campus Virchow-Klinikum wurde der Implementierungsprozess durch ein Qualitätsmanagementprojekt begleitet. Die Datenerfassung im Observationsdesign sowie die Publikation der Ergebnisse wurde durch die zuständige Ethikkommission (Ethikvotum: EA1/228/15) und den Datenschutz genehmigt. Eine Einwilligungserklärung der Patienten wurde nicht gefordert.

Als Ausschlusskriterien für die Implementierung wurden definiert:

- Alter unter 18 Jahren
- fehlende Kenntnisse der deutschen Sprache
- bekannte Demenz mit mehr als nur geringgradiger kognitiver Einschränkung
- andere psychiatrische Erkrankung in der Anamnese
- Ablehnung der einzelnen Untersuchung durch den Patienten.

Alle Patienten, die keines der genannten Ausschlusskriterien erfüllten, wurden täglich gemäß des Implementierungsalgorithmus untersucht. Die Untersuchung der Patienten endete mit der Verlegung von der Intensivstation oder dem Ende des Erhebungszeitraumes. Die Anzahl der Wiederholungsmessungen wurde durch die Erhebungszeiträume und die Intensivverweildauer der Patienten begrenzt.

Die Untersuchung wurde in zwei Untersuchungsblöcke geteilt. Der erste Erhebungszeitraum erstreckte sich vom 15. September 2011 bis zum 14. Oktober 2011 und der zweite Erhebungszeitraum vom 1. Februar 2012 bis zum 14. April 2012.

2.1.1 Allgemeine Konzeption

Während der oben genannten Blöcke erfolgten tägliche Visiten aller in die Untersuchung aufgenommenen Patienten zum Delirscreening. Dazu wurden die Patienten einmal täglich nach dem gleichen Ablaufschema untersucht. Um elf Uhr vormittags beginnend, erfolgte die Evaluation nach der ICDSC. Anschließend wurden die Patienten mit einem Versatz von 20 Minuten nach der CAM-ICU beurteilt. Weitere 20 Minuten später erfolgte die erste Beurteilung mit dem Referenzstandard nach DSM IV TR. Um der im Tagesverlauf schwankenden Symptomatik des zu untersuchenden Krankheitsbildes Rechnung zu tragen, fanden somit sämtliche Untersuchungen bei jedem einzelnen Patienten innerhalb von 90 Minuten statt. Während der gesamten Untersuchungsdauer führte jeder Untersucher stets die Einschätzung mit demselben Messinstrument durch. Die Einschätzung nach der CAM-ICU wurde von Dennis Ullrich, nach der ICDSC von Weronika Lisiuk und gemäß DSM IV TR durch Peter Braun vorgenommen.

Alle Untersuchungen wurden pseudonymisiert und bis zur Datenauswertung gegeneinander verblindet durchgeführt. Die Dokumentation erfolgte elektronisch. Für die Erfassung anamnestisch relevanter Sachverhalte, der Bewertung einer akut einsetzenden Symptomatik oder eines schwankenden Verlaufs war der Austausch mit den behandelnden pflegerischen oder ärztlichen Kollegen, ein Interview mit ggf. anwesenden Angehörigen oder das Aktenstudium zulässig. Die Entnahme bereits vorhandener Ergebnisse für das Delirscreening aus den elektronischen Patientenakten war nicht gestattet. Nach der Erhebung der einzelnen Ergebnisse erfolgte zeitnah die Dokumentation in einer Datenbank mittels LimeSurvey (LimeSurvey GmbH, Hamburg, Deutschland) auf internen Servern der Charité-Universitätsmedizin Berlin.

Im Anschluss daran wurden Schmerz- und Vigilanzscores für jeden Patienten täglich dokumentiert. Dies waren der Wert der aktuellen Vigilanz auf der *Glascow Coma Scale* (GCS) (vgl. Teasdale und Jennett, 1974) und eine Beurteilung der Analgesiequalität, die je nach Vigilanz des Patienten mittels der Numerischen Analog Skala (NAS), der *Behavioral Pain Scale* (BPS) (vgl. Payen et al., 2001) oder der *Behavioral Pain Scale non-intubated* (BPS-NI) (vgl. Chanques et al., 2009) erfolgte. Die verwendeten Messinstrumente sind im Anhang zu dieser Arbeit dargestellt. Ferner wurde täglich aus der

Patientenakte die Information übernommen, ob der Patient in den letzten 24h Medikamente mit sedierendem Wirkprofil erhalten hatte und deren Dosierung dokumentiert.

Die Durchführung der Datenerhebung ist in nachfolgender Abbildung 2.1.1 verdeutlicht.

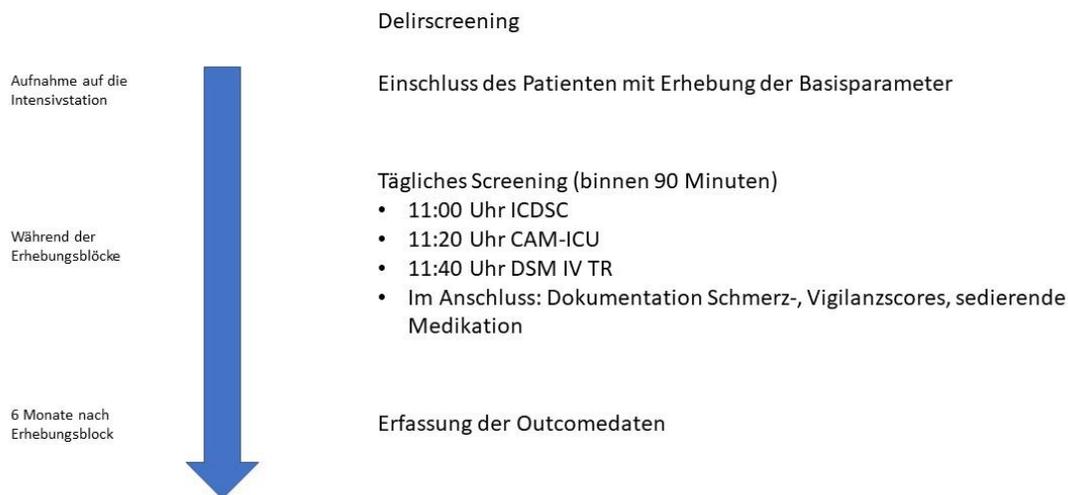


Abbildung 2.1.1: Ablaufschema des Delirscreenings während der Datenerhebung

2.1.2 Erhebung von Basis- und Outcomeparametern

Am Tag des Einschlusses wurden standardisierte biometrische und anamnestische Daten der Patienten erfasst und dokumentiert. Dabei handelte es sich um Alter, Gewicht, ideales Körpergewicht, Art und Grund der Aufnahme und die Information, ob vor dem Einschluss eine operative Therapie durchgeführt worden war. Die weiteren erhobenen Parameter sind in nachfolgender Tabelle 2.1.2 aufgelistet. Die Datenerhebung wurde pseudonymisiert durchgeführt.

Tabelle 2.1.2: Erhobene Basisparameter

Schlaganfall oder TIA in der Anamnese	stattgehabte Augenoperation
ggf. vorhandene Residuen	Alkohlabusus
bekannte Demenz	perioperative Stressprophylaxe
bekannte Enzephalopathie	perioperative Clonidingabe
kognitive Einschränkungen	jemals durchgeführte Reanimation

In zeitlichem Abstand von sechs Monaten nach Abschluss der Datenerhebung erfolgte jeweils der Export der Outcomedaten aus den Systemen des Medizincontrollings. Dabei wurden verschiedene Parameter entnommen, wie die fachärztliche Diagnose, die Länge des Krankenhaus- und

Intensivaufenthaltes, die Beatmungsdauer, das Ergebnis zweier Scores zur Erkrankungsschwere und die Entlassungsart. Die Erkrankungsschwere wurde standardmäßig mit der international weit verbreiteten *Acute Physiology and Chronic Health Evaluation II* (APACHE II) (vgl. Knaus et al., 1985) und dem *Simplified Acute Physiology Score II* (SAPS II) (vgl. Le Gall et al., 1993) objektiviert. Beide Scores sind im Anhang unter 7.5 und 7.6 in der jeweiligen deutschen Übersetzung zu finden. Sämtliche erhobene Daten wurden pseudonymisiert in LimeSurvey (LimeSurvey GmbH, Hamburg, Deutschland) auf internen Servern der Charité Universitätsmedizin Berlin gespeichert.

Alle Daten lagen am Ende der Datenerhebung im .xls Dateiformat vor (MS Excel, Microsoft Corporation, Redmond, Washington, USA). Es erfolgte eine Übertragung der Daten in SPSS 22 (IBM, Chicago, Illinois, USA) und eine Vorbereitung zur statistischen Analyse. Die endgültige statistische Analyse wurde mit Stata 16 (Stata Corp LLC, College Station, Texas, USA) durchgeführt.

2.2 Delirmessung mit der CAM-ICU (Screening) und nach DSM-IV TR (Goldstandard)

Nachfolgend werden die verwendeten Messinstrumente dargestellt und die durchgeführten Schulungsmaßnahmen erläutert.

2.2.1 CAM-ICU

Die CAM-ICU überprüft durch standardisierte Untersuchungen jedes Merkmal für das Vorliegen des Delirs nach DSM IV TR. Die hierfür vorgenommenen Untersuchungen werden nachfolgend dargestellt. Eine Abbildung des CAM-ICU Arbeitsblatts findet sich im Anhang zu dieser Arbeit unter 7.1.

Bevor die einzelnen Merkmale überprüft werden, muss eine dichotome Beurteilung des Patienten und seiner aktuellen Vigilanz mittels der *Richmond Agitation Sedation Scale* (RASS) durchgeführt werden. Diese wird ausführlich unter Merkmal 4 beschrieben.

Nach der Originalpublikation kann ein komatöser Patient nicht anhand der CAM-ICU beurteilt werden (vgl. Ely et al., 2001). Gemäß der kulturellen Adaptation und deutschsprachigen Übersetzung wird ein Patient mit einem Wert von RASS <-3 als komatös bezeichnet und ist daher nicht anhand der CAM-ICU zu untersuchen (vgl. Guenther et al., 2010).

Merkmal 1: akuter Beginn oder schwankender Verlauf

Dieses Merkmal wird als positiv beurteilt, wenn sich der geistige Zustand des Patienten innerhalb der letzten 24h akut verändert hat oder es im Verlauf der letzten 24h zu wiederholten Änderungen gekommen ist. Hierfür werden Informationen aus der Patientenakte, Beschreibungen des Pflegepersonals oder fremdanamnestic Angaben ausgewertet.

Merkmal 2: Störung der Aufmerksamkeit

Zur Beurteilung dieses Merkmals wird dem Patienten nach ausführlicher Erklärung das Wort A-N-A-N-A-S-B-A-U-M buchstabiert. Bei jedem genannten A muss der Patient die Hand des Untersuchers drücken. Tut er dies nicht oder bei einem anderen Buchstaben als A, so wird dies als Fehler gewertet. Alternativ kann eine Überprüfung dieses Merkmals auch mittels eines Testsets von gut zu differenzierenden Bildern erfolgen. Hierfür werden dem Patienten initial fünf Bilder präsentiert, die dieser dann in einem zweiten Schritt aus einem Set von zehn Bildern identifizieren muss. Bei drei oder mehr Fehlern wird diese Untersuchung als positiv angesehen.

Merkmal 3: Störung des formalen Denkens

Bei der Beurteilung dieses Merkmals wird überprüft, ob eine Störung des formalen Denkens bei dem zu untersuchenden Patienten nachweisbar ist. Hierfür werden ihm vier dichotome Fragen und eine einfache motorische Aufgabe gestellt. Die Fragen sind entweder verbal mit ‚Ja‘/ ‚Nein‘ oder alternativ durch Kopfnicken oder -schütteln zu beantworten, falls eine verbale Kommunikation nicht möglich sein sollte. Im Anschluss daran werden dem Patienten zwei erhobene Finger gezeigt und der Patient wird gebeten, diese Geste nachzuahmen. Nun soll der Patient mit der anderen Hand ebenso viele Finger oder – sofern eine Bewegungseinschränkung eines Armes besteht – einen Finger mehr zeigen, ohne dass der Untersucher weiterhin zwei Finger präsentiert. Jede falsch beantwortete Frage oder eine andere als die beschriebene Ausführung des motorischen Kommandos wird als Fehler gewertet. Merkmal 3 gilt als positiv, wenn der Patient mehr als einen Fehler macht.

Merkmal 4: Vigilanzänderung

Um die Vigilanz des Patienten zu überprüfen wird dieser mithilfe der *Richmond Agitation Sedation Scale* (vgl. Sessler et al., 2002) eingeschätzt. Diese ist in der nachfolgenden Tabelle 2.2.1 abgebildet. Für jedes Ergebnis, das vom Wert null auf der RASS abweicht, wird das Merkmal 4 als positiv bewertet.

Tabelle 2.2.1: Richmond Agitation Sedation Scale (Sessler et al., 2002)

Wert auf der RASS	Bezeichnung	Beschreibung
4	streitlustig	gewalttätiges Verhalten, unmittelbare Fremdgefährdung
3	sehr agitiert	aggressiv, entfernt Katheter und Drainagen
2	agitiert	häufige ungezielte Bewegungen, kämpft gegen Respirator
1	unruhig	ängstlich, kein aggressives Verhalten
0	aufmerksam, ruhig	Normalbefund
-1	schläfrig	jederzeit erweckbar, schläft intermittierend ein, Augenöffnen und Augenkontakt >10s
-2	leicht sediert	jederzeit erweckbar, schläft schneller wieder ein, Augenöffnen und Augenkontakt <10s
-3	moderat sediert	auf Ansprache Augenöffnen (ohne Fixieren) oder Bewegungen
-4	tief sediert	keine Reaktion auf Ansprache, auf Berührung Augenöffnen oder Bewegung
-5	nicht erweckbar	keine Reaktion auf Ansprache oder Berührung

Bei der RASS handelt es sich um eine validierte und bei intensivpflichtigen Patienten mehrfach als verlässlich nachgewiesene Skala zur Einschätzung der Sedierungstiefe. Sie gilt gemäß der aktuellen S3-Leitlinie im deutschsprachigen Raum als Goldstandard zur Einschätzung der Sedierungstiefe (vgl. Baron et al., 2015) und wird auch in den Empfehlungen des amerikanischen *College of Critical Care Medicine* als Goldstandard empfohlen (vgl. Barr et al., 2013).

Gemäß der CAM-ICU kann ein Patient als ‚delirpositiv‘ bewertet werden, wenn er positiv für die Merkmale 1 und 2 und zusätzlich positiv für Merkmal 3 oder 4 getestet wurde.

2.2.2 DSM IV TR

Während in der Einleitung die Definition des Delirs nach ICD-10 und DSM V dargelegt wurde, war zum Zeitpunkt der Datenerhebung dieser Arbeit die vorhergehende Version des *Diagnostic and Statistical Manual of Mental Disorders IV. Edition Text Revision* (DSM IV TR) maßgeblich. Für die Diagnose eines Delirs nach DSM IV TR müssen alle der vier folgenden Merkmale erfüllt sein (APA, 2002):

- quantitative Bewusstseinsstörung mit einer reduzierten Umgebungswahrnehmung und einer verminderten Fähigkeit, die Aufmerksamkeit zu richten, zu halten oder zu verlagern (Kriterium A)
- Denkstörung oder Wahrnehmungsstörung der Umwelt, die nicht auf eine beginnende, manifeste oder progrediente Demenz zurückzuführen ist (Kriterium B)
- kurzfristiger Beginn und im Tagesverlauf schwankende Symptomatik (Kriterium C)

- aus Anamnese, Klinik oder Laboruntersuchungen ergeben sich Hinweise, dass die Störung im Rahmen einer medizinischen Erkrankung aufgetreten ist (Kriterium D)

Die beiden Definitionen der APA unterscheiden sich insbesondere im jeweiligen Kriterium A. Während im DSM V eher eine Störung der Aufmerksamkeit bzw. der qualitativen Bewusstseinslage adressiert wird, fokussiert sich das DSM IV TR auf eine quantitative Bewusstseinsstörung. Zudem stellt ein Koma im DSM V nach Kriterium D ein Ausschlusskriterium für die Diagnose des Delirs dar, während das DSM IV TR ein Koma nicht explizit nicht ausschließt (vgl. EDA, 2014). Allerdings handelt es sich beim Koma um eine eigene Entität einer Erkrankung, die mit einem Delir nicht vereinbar ist. Die im Gegensatz zur ICD-10 internationale Verbreitung der DSM IV TR lässt sich damit erklären, dass die meisten Patienten, die eine Delirsymptomatik aufweisen, bei einer Bewertung nach der DSM IV TR eingeschlossen werden (vgl. Luetz et al., 2010a). Dies konnte auch durch eine europäische Umfrage an 200 Teilnehmern aus verschiedenen Fachrichtungen und Versorgungsstufen bestätigt werden (vgl. Morandi et al., 2013).

2.2.3 Unterweisung und Schulung der CAM-ICU

Die in der Untersuchung erhobenen Daten wurde von supervidiertem Fachpersonal (*Study Nurses/Doktoranden*) erhoben. Bei den Supervisoren handelte es sich um erfahrene Intensivmediziner. Das Personal wurde gemäß einem vierstufigen klinikinternen Konzept zur Schulung von Delirmessinstrumenten unterrichtet, das vorher publiziert wurde (vgl. Luetz et al., 2010).

Nach einer Vorstellung der verwendeten Scores durch einen erfahrenen Intensivmediziner und zwei *Study Nurses* wurden die Messinstrumente anhand schriftlicher Informationen und Vorträge erläutert. Diese beinhalteten das klinikinterne Trainingsmaterial und eine Auswahl an aktueller Literatur. In der zweiten Stufe erfolgte eine Einzelunterweisung, wobei die Anwendung der einzelnen Messinstrumente am Patientenbett demonstriert wurde. Im Rahmen einer einwöchigen Testphase untersuchte jeder Untersucher mehrere Patienten eines intensivmedizinischen Kollektivs ohne Neurointensivpatienten. Bei Fragen oder Unklarheiten konnte jederzeit ein erfahrener Intensivmediziner hinzugezogen werden. Dies stellt die dritte Stufe des beschriebenen Implementierungskonzepts dar. Zum Abschluss des Trainingskonzepts untersuchte jeder Untersucher fünf konsekutive Patienten gegeneinander verblindet. Diese Patienten waren vorher von zwei erfahrenen Intensivmedizinern mittels der Messinstrumente untersucht worden, was als Referenzstandard galt. Es zeigten sich sehr gute Übereinstimmungen mit den Referenzuntersuchern.

2.2.4 Unterweisung und Schulung der neuropsychiatrischen Untersuchung und Evaluation nach DSM IV-TR

Für die Schulung der Untersucher zur Erhebung des DSM IV TR wurde ebenfalls nach dem oben angegebenen Schulungskonzept vorgegangen. Dabei führte der klinisch erfahrenste Untersucher die strukturierte neuropsychologische Untersuchung und Bewertung nach DSM IV TR durch. Diese Untersuchung erfolgte mittels eines semistrukturierten Gesprächs, in dem wesentliche Elemente der *Mini Mental State Examination* (MMSE-Test) (vgl. Folstein et al., 1975) verwendet wurden. Angaben zur Anamnese des Patienten bezüglich eines vorbestehenden kognitiven Defizits, einer Demenz oder einer psychiatrischen Erkrankung zur Abgrenzung gegenüber dem Delir wurden durch Aktenstudium, Gespräche mit dem Pflegepersonal oder dem ärztlichen Dienst der Station und aus Gesprächen mit Angehörigen gewonnen.

2.3 Besonderheiten der Datenerhebung

Nachfolgend werden Besonderheiten während der Durchführung der Datenerhebung dargelegt.

2.3.1 Referenzmessung und Wiederholungsmessungen

Als primäre Referenzuntersuchung diente die Messung, bei der alle untersuchten Scores erstmals beim Patienten durchführbar waren. Im weiteren Verlauf wird der Begriff erster valider Messtag verwendet. Dabei kann es sich auch um einen Tag nach dem ersten Messtag handeln, sofern Ausschlusskriterien für die Durchführung der CAM-ICU vorlagen. Da die CAM-ICU, wie bereits dargestellt, eine strenge Vorgabe zur Nichtdurchführbarkeit beinhaltet, wurde der hier vorgegebene Cut-Off von RASS <-3 auf die anderen Scores übertragen. Jedoch wurde die RASS entwickelt, um die Tiefe einer Sedierung zu beurteilen (vgl. Sessler et al., 2002). Deshalb wurde bei Patienten, deren mangelnde Vigilanz nicht durch sedierende Medikamente zu erklären war, ein ebensolcher Cut-Off notwendig. Hierfür wurde die GCS aufgrund ihrer weiten Verbreitung in der Literatur und der einfachen Möglichkeit der Erhebung verwendet. Um eine hinreichende Näherung zu erreichen, wurde ein Wert von <9 gewählt. Diesem Vorgehen liegt die Überlegung zugrunde, dass ein Patient, um grundsätzlich mittels der CAM-ICU beurteilbar zu sein, mindestens dazu in der Lage sein muss, motorische Kommandos auszuführen und die Augen auf Berührung zu öffnen. Eine verbale Kommunikation ist nicht zwangsläufig erforderlich. Daher ergibt sich der Wert von 9 aus A2, M6, V1. Messtage, an denen die oben genannten Voraussetzungen nicht erfüllt waren, wurden von der statistischen Analyse ausgeschlossen.

2.3.2 Berechnung der severity scale for the CAM-ICU (ssCAM-ICU)

Gemäß der ursprünglich vorgestellten Version der CAM-ICU handelt es sich um eine dichotome, ordinale Skala, die lediglich die Einteilung in ‚Delir‘ oder ‚kein Delir‘ zulässt (vgl. Ely et al., 2001). Um jedoch unterschiedliche Schweregradeinteilungen vornehmen, ein eventuell vorliegendes SSD detektieren und Krankheitsverläufe genauer abbilden zu können, bietet sich eine Umrechnung des jeweiligen Testergebnisses auf ordinale Schweregradskala analog zum Vorgehen von Luetz et al. für pädiatrische Patienten an (vgl. 2016). Eine erste Adaptation dieses Messinstruments an erwachsene Intensivpatienten ist kürzlich veröffentlicht worden (vgl. Nacul et al., 2020).

Es erfolgt eine Umrechnung der jeweiligen Merkmale in einen Punktwert. So wird ein positives Merkmal 1 (akuter Beginn oder schwankender Verlauf) mit fünf Punkten bewertet. Falls Merkmal 1 negativ ist, werden null Punkte vergeben. Für jeden Fehler, den der Patient bei Merkmal 2 (Störung der Aufmerksamkeit/ASE-Bilder) macht, werden 0,5 Punkte addiert. Wenn dem Patienten bei der Untersuchung zu Merkmal 3 Fehler unterlaufen, dann wird für jeden Fehler ein Punkt addiert. Bei Merkmal 4 (Vigilanzänderung) wird bei einer veränderten RASS der jeweilige Betrag addiert. Exemplarisch ergibt ein RASS +/-2 einen Wert von 2 Punkten.

Die somit erhaltene Skala reicht von 0 Punkten (kein Delir) bis 19 Punkte (maximale Ausprägung des Delirs). In Analogie zu der pädiatrischen Version bietet sich die Benennung als *severity scale for the CAM-ICU* (ssCAM-ICU) bzw. Schweregradskala für die *Confusion Assessment Method* für Intensivstationen an. Bisher ist lediglich eine Arbeit zu diesem Messinstrument publiziert worden. Bei einem optimalen Cut-Off von $\geq 5,5$ Punkte auf der ssCAM-ICU wird eine Sensitivität von 91,7% und eine Spezifität von 82,5% für eine Kohorte internistischer bzw. chirurgischer Intensivpatienten angegeben (vgl. Nacul et al., 2020). Eine schematische Darstellung der ssCAM-ICU findet sich im Anhang zu dieser Arbeit (vgl. Abschnitt 7.2).

2.4 Statistische Analyse

Nachfolgend werden die Schritte der Datenaufbereitung, der statistischen Auswertung und der Präsentation dargelegt.

2.4.1 Datenauswertung

Die Beschreibung des Patientenkollektivs erfolgte bei ordinalen bzw. numerischen Daten jeweils durch Angabe des Medians mit den zugehörigen Grenzen des Interquartilsabstand (25. bis 75.

Perzentile). Bei nominalen Daten wurden die Häufigkeiten dargestellt. Die Unterschiede zwischen den Patienten, die bei keiner Messung nach DSM IV TR delirpositiv waren, und denen, die mindestens einmal delirpositiv waren, wurden bei ordinalen bzw. numerischen Variablen mittels des nichtparametrischen Wilcoxon-Mann-Whitney-Tests und bei nominalen Daten mittels des exakten Tests nach Fisher verglichen. Bei nominalen Daten mit mehr als zwei möglichen Zuständen wurde der Chi-Quadrat-Test angewandt. Die statistische Signifikanz wurde für einen p-Wert $\leq 0,05$ angenommen. Sämtliche Werte sind explorativ zu interpretieren. Es erfolgte keine Adjustierung für repetitive Messungen.

Für die statistische Auswertung wurde der erste valide Messtag von allen weiteren Messungen unterschieden. Die Testgütekriterien der ssCAM-ICU bezogen auf das DSM IV TR wurden gemäß des von Nacul et al. vorgeschlagenen Cut-Off $\geq 5,5$ Punkte auf der ssCAM-ICU sowohl für den ersten Messtag als auch für die nachfolgenden Messungen berechnet (vgl. 2020).

Mittels Kreuztabelle wurden folgende Testgütekriterien für die Beurteilung nach ssCAM-ICU, bezogen auf den Referenzstandard DSM IV TR, für den ersten validen Messtag berechnet, auf eine Nachkommastelle gerundet und als Wert mit dem 95% Konfidenzintervall angegeben:

Tabelle 2.4.1: Formeln Testgütekriterien

$\text{Sensitivität} = \frac{\{\text{richtig positive (RP)}\}}{\{\text{richtig positive (RP)} + \text{falsch negative (FN)}\}}$
$\text{Spezifität} = \frac{\{\text{richtig negative (RN)}\}}{\{\text{richtig negative (RN)} + \text{falsch positive (FP)}\}}$
$\text{positiv prädiktiver Wert (PPV)} = \frac{\{\text{richtig positive (RP)}\}}{\{\text{positive (RP + FP)}\}}$
$\text{negativ prädiktiver Wert (NPV)} = \frac{\{\text{richtig negative (RN)}\}}{\{\text{negative (RN + FN)}\}}$

Zur Bestimmung der Testgüte für die ssCAM-ICU wurde mittels *Receiver-Operating-Characteristics* (ROC)-Kurven grafisch dargestellt. Hierfür erfolgt eine Abbildung der Sensitivität auf der Ordinate und 1-Spezifität auf der Abszisse. Das optimale Ergebnis eines sehr sensitiven und sehr spezifischen Testes befindet sich in der linken oberen Ecke des Diagramms. Die diagnostische Aussagekraft des

Testes ist umso größer, je weiter sich das Ergebnis der linken oberen Ecke annähert. Als Qualitätsmaßstab für den verwendeten diagnostischen Test erfolgte die Bestimmung der *Area Under the Curve* (AUC) der ROC-Kurve und eine Einordnung nach dem Modell von Greiner et al. (vgl. 2000). Die ROC-Kurve für das Gesamtkollektiv wurde zunächst für den ersten validen Messtag und anschließend für die repetitiven Messungen erstellt. In den Subgruppen erfolgte die Auswertung der Wiederholungsmessungen. Für den ersten validen Messtag erfolgt die Angabe der AUC mit 95% Konfidenzintervall.

Im Sinne einer Grenzwertoptimierung wird rechnerisch das Wertepaar bestimmt, dessen Youden-Index (J) am größten ist (vgl. Youden, 1950). Der Youden-Index bewertet ausschließlich die Gesamtleistung des Tests und differenziert nicht zwischen Sensitivität und Spezifität. Definiert ist der Youden-Index als:

Tabelle 2.4.2: Youden-Index

$$J = \text{Sensitivität} + \text{Spezifität} - 1$$

Als Mindestmaß für die Sensitivität im Rahmen eines Screeningtests wurde zudem ein Wert von 80% als Kriterium festgelegt, in dessen Rahmen ein möglichst hoher Youden-Index (vgl. Youden, 1950) angestrebt wird, in Analogie zum Vorgehen von Nacul et al. (vgl. Nacul et al., 2020). Die Bestimmung des Youden-Index (vgl. Youden, 1950) erfolgte sowohl für den ersten Messtag als auch für die repetitiven Messungen im Gesamtkollektiv.

Zusätzlich wurden die Subgruppen der Schlaganfall- und neurochirurgischen Patienten betrachtet. Die Testgütekriterien für die repetitiven Messungen wurden mittels Kreuztabelle für diese Subgruppen bestimmt. Es werden für die repetitiven Messungen explizit keine Konfidenzintervalle angegeben, da es sich hierbei um Mehrfachmessungen handelt. Sofern man diesen Umstand nicht berücksichtigt, folgt hieraus eine künstliche Verkleinerung der Konfidenzintervalle und somit eine statistische Überschätzung der Wertigkeit des verwendeten Tests. Für die Untergruppen erfolgten eine ROC-Analyse und eine Optimierung der Cut-Off-Werte in Analogie zu dem oben beschriebenen Vorgehen für alle zur Verfügung stehenden Messungen. Sämtliche statistische Berechnungen und die Erstellung der verwendeten ROC-Kurven wurden mit STATA 16 (StataCorp LP, College Station, TX, USA) durchgeführt.

3. Ergebnisse

Nachfolgend werden die Ergebnisse des beschriebenen Studiendesigns wiedergegeben. Zunächst erfolgt eine Darstellung des Patientenkollektivs und anschließend die Auswertung der Delirmmessungen, wie in Abschnitt 2 bereits beschrieben.

3.1 Patientenkollektiv

Es wurden 181 Patienten, die auf der Intensivstation 1i des Charité Campus Virchow (CVK) behandelt wurden, bezüglich eines Einschlusses in die Untersuchung überprüft. Bei 37 Patienten lagen Ausschlusskriterien vor. Von den verbliebenen 144 Patienten wurden weitere 43 ausgeschlossen, da von diesen Patienten zu keinem Zeitpunkt ein komplettes Testset verfügbar war. Für die statistische Auswertung verblieben somit 101 Patienten mit 324 Wiederholungsmessungen (vgl. Abbildung 3.1).

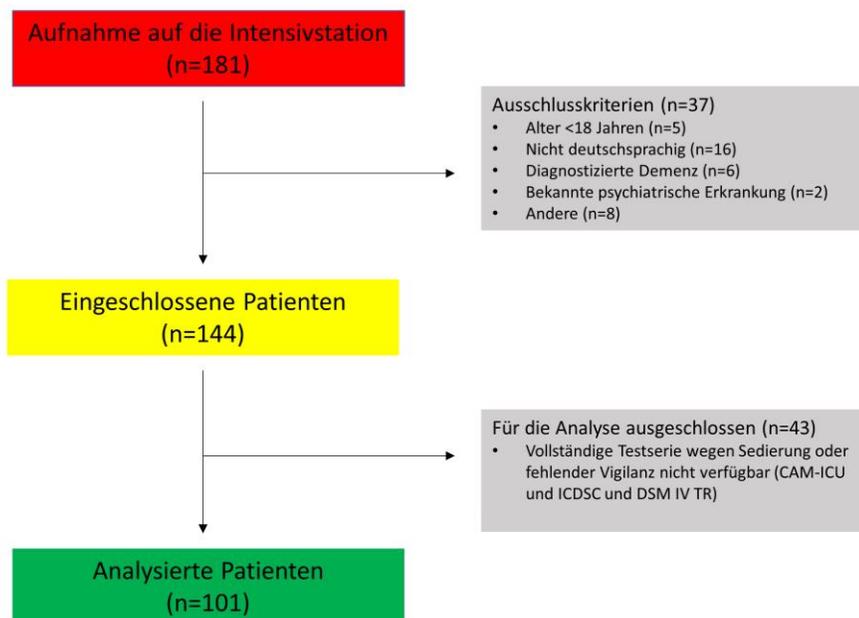


Abbildung 3.1: Strobe Diagramm, Andere: Patient nicht durch alle Untersucher zu untersuchen, Patient lehnt Teilnahme ab, Patient in moribunder Situation und Ableben binnen 24h wahrscheinlich

3.1.1 Charakteristika des Patientenkollektivs

Zur Auswertung der Basischarakteristika des Patientenkollektivs erfolgte eine Differenzierung nach delirpositiven und delirnegativen Patienten. Insgesamt 30 der 101 Patienten erhielten an mindestens einem Zeitpunkt eine positive Bewertung nach DSM IV TR. Die Basischarakteristika des Patientenkollektivs sind in Tabelle 3.1.1 dargestellt.

Tabelle 3.1.1: Basischarakteristika des Patientenkollektivs

	delirpositiv (n=30)*	delirnegativ (n=71)	p
Alter, <i>J</i> ^a	64,5 (58-75)	60 (50 - 76)	0,506 ^b
Geschlecht: männlich, <i>n</i>	20 (66,7%)	34 (47,9%)	0,126 ^c
Gewicht, <i>kg</i> ^a	78 (70 - 85)	71 (65-81)	0,160 ^b
APACHE II, <i>pkt.</i> ^a	20 (16 - 26)	13 (7 - 19), (n=67)	<0,001 ^b
SAPS II, <i>pkt.</i> ^a	39,5 (31 - 49)	25,5 (15 - 38), (n=70)	<0,001 ^b
Notfallaufnahme, <i>n</i>	25 (80,6%)	40 (56,3%)	0,012 ^c
CPR während der Behandlung, <i>n</i>	1 (3,3%)	1 (1,4%)	0,508 ^c
Schlaganfall/TIA in der Anamnese, <i>n</i>	8 (26,7%)	30 (36,2%)	0,179 ^c
maschinelle Beatmung, <i>n</i> ^e	22 (73,3%)	22 (31%)	<0,001 ^c
mittlere Dauer der maschinellen Beatmung, <i>h</i> ^a	105,5 (0 - 185)	0 (0 - 11)	<0,001 ^b
Dauer des Intensiv Aufenthaltes, <i>d</i> ^a	15,5 (9 - 22)	3 (1 - 6)	<0,001 ^b
Mortalität während des Aufenthaltes, <i>n</i>	5 (16,7%)	0	0,002 ^c
Aufnahmegrund			
neurochirurgische Intervention, <i>n</i>	17 (56,7%)	42 (59,2%)	0,530 ^d
kardiologisch, <i>n</i>	2 (6,7%)	1 (1,4%)	
neurologisch, <i>n</i>	10 (33,3%)	24 (33,8%)	
chirurgisch (ohne neurochirurgisch), <i>n</i>	1 (3,3%)	4 (5,6%)	
RASS			
erster Messtag ^a	-3 (-3 - -3), (n=1)	-1 (-1 - 0), (n=3)	0,157 ^b
Md aller Messungen ^a	-3 (-5 - -3), (n=13)	-3 (-5 - -1), (n=11)	0,177 ^b
GCS			
erster Messtag, <i>pkt.</i> ^a	14 (12 - 15), (n=42)	15 (14,5 - 15), (n=68)	<0,001 ^b
Md aller Messungen, <i>pkt.</i> ^a	14 (12 - 14), (n=29)	15 (15 - 15), (n=68)	<0,001 ^b
Schmerzassessment			
schmerzffreie Tage ^f	5,5 (3 - 10)	1 (1 - 3)	<0,001 ^b
Schmerztag ^f	0 (0 - 1)	0 (0 - 1)	0,578 ^b

*Delir positiv an mindestens einem Messtag nach DSM IV TR, Signifikanz wenn $p \leq 0,05$, ^aDaten dargestellt als Median (Interquartilsbereich), ^bWilcoxon-Mann-Whitney-Test, ^cExakter Test nach Fisher, ^dChi-Quadrat-Test, ^einvasive, maschinelle Ventilation auf der Intensivstation, ^fPatient schmerzfrei, wenn der Wert auf der NAS, BPS-NI, BPS jeweils <30% des Maximums, Schmerztag, wenn der Wert >30% des Maximums. *Abkürzungen: Md=Median; APACHE II=Acute Physiology and Chronic Health Examination II; SAPS II= Simplified Acute Physiology Score II; CPR=kardiopulmonale Reanimation; TIA=Transiente Ischämische Attacke; RASS=Richmond Agitation Sedation Scale; GCS=Glasgow Coma Scale; NAS=Numerische Analog Skala; BPS=Behavioral Pain Scale; BPS-NI=Behavioral Pain Scale – non intubated, (vgl. auch Weiss et al., 2021)*

Zwischen den delirpositiven und den delirnegativen Patienten gab es keine statistisch signifikanten Unterschiede bei der Betrachtung von Alter, Geschlecht, Körpergewicht, der Anzahl der Reanimationen, der Häufigkeit von Schlaganfall oder TIA in der Anamnese oder der Aufnahmegründe. Die

nicht-deliranten Patienten hatten niedrige APACHE II (13 vs. 20 Punkte; $p<0,001$) und SAPS II (25,5 vs. 39,5 Punkte; $p=0,012$) und wurden seltener notfallmäßig auf die Intensivstation aufgenommen (56,3% vs. 80,6%; $p=0,001$). Die Patienten mit einem Delir waren häufiger (73,3% vs. 31%, $p<0,001$) und länger (105,5 vs. 0 Stunden, $p<0,001$) invasiv beatmet und wurden länger auf der Intensivstation behandelt (15,5 vs. 3 Tage, $p<0,001$). In der Gruppe der deliranten Patienten war die Mortalität erhöht (16,7% vs. 0%, $p=0,002$). Bezüglich der Vigilanz zeigten die deliranten Patienten sowohl am ersten Messtag als auch in den repetitiven Messungen einen um einen Punkt geringeren Wert auf der GCS als die delirnegativen Patienten (14 vs. 15 Punkte; $p<0,001$).

3.2 Prävalenz des Delirs

Nach DSM IV TR konnte am ersten validen Messtag ein Delir bei 15 von 100 Messungen des Patientenkollektivs diagnostiziert werden. Dies entspricht einer Punktprävalenz von 15%. Bei den repetitiven Messungen wurden 50 von 324 Messungen als delirpositiv bewertet. Somit ergibt sich eine Gesamtprävalenz von 15,4%.

In der Subgruppe der 47 Schlaganfallpatienten wurde in den repetitiven Messungen mit dem Referenzstandard DSM IV TR die Diagnose Delir bei 31 von 180 Messungen gestellt, sodass die Prävalenz bei 17,2% liegt. In den Wiederholungsmessungen ließen sich in der Subgruppe der 58 neurochirurgischen Patienten 23 von 181 Messungen als delirpositiv bewerten, was einer Prävalenz von 12,7% entspricht.

3.3 Testgütekriterien für den ersten validen Messtag im Gesamtkollektiv

Am ersten validen Messtag konnten die Daten von 100 Patienten ausgewertet werden. Ein Patient musste ausgeschlossen werden, da die CAM-ICU am ersten validen Messtag wegen Ablehnung durch den Patienten nicht durchführbar war. Die Ergebnisse für das Delirscreening unter Berücksichtigung des Cut-Offs von $\geq 5,5$ Punkten auf der ssCAM-ICU analog zu Nacul et al. sind in der nachfolgenden Kreuztabelle (vgl. Tabelle 3.3.1) dargestellt (vgl. 2020).

Tabelle 3.3.1: Kreuztabelle ssCAM-ICU vs. DSM IV TR, erster valider Messtag, Gesamtkollektiv

Bewertung nach DSM IV TR	Bewertung nach ssCAM-ICU		Gesamt
	Positiv	Negativ	
Positiv	12	3	15
Negativ	13	72	85
Gesamt	25	75	100

Es ergibt sich eine Sensitivität von 82,4% und eine Spezifität von 84,7%. Der positiv prädiktive Wert beträgt 51,3% und der negativ prädiktive Wert 49,3% (vgl. Tabelle 3.3.2).

Tabelle 3.3.2: Testgütekriterien, erster valider Messtag, Gesamtkollektiv

	Sensitivität (95% KI) %	Spezifität (95% KI) %	positiv prädiktiver Wert (95 KI) %	negativ prädiktiver Wert (95% KI) %
ssCAM-ICU (n=100)	82,4 (74,9 – 89,9)	84,7 (77,6 – 91,8)	51,3 (41,5 – 61,1)	49,3 (39,5 – 59,1)

Aus Tabelle 3.3.1 ist ersichtlich, dass 13 Patienten falsch positiv und drei Patienten falsch negativ mit der ssCAM-ICU bewertet wurden.

3.4 Testgütekriterien der repetitiven Messungen

In den repetitiven Messungen des Gesamtkollektivs konnte insgesamt 324 Messungen ausgewertet werden. Die Ergebnisse für das Delirscreening unter Berücksichtigung des Cut-Offs von $\geq 5,5$ Punkte auf der ssCAM-ICU sind in der nachfolgenden Kreuztabelle (vgl. Tabelle 3.4.1) dargestellt.

Tabelle 3.4.1: Kreuztabelle ssCAM-ICU vs. DSM IV TR, repetitive Messungen, Gesamtkollektiv

Bewertung nach DSM IV TR	Bewertung nach ssCAM-ICU		Gesamt
	Positiv	Negativ	
Positiv	46	4	50
Negativ	52	222	274
Gesamt	98	226	324

Tabelle 3.4.2: Testgütekriterien, repetitive Messungen, Gesamtkollektiv

	Sensitivität, %	Spezifität, %	positiv prädiktiver Wert, %	negativ prädiktiver Wert, %
ssCAM-ICU (n=324)	92	81	55,2	45,4

Es ergibt sich eine Sensitivität von 92% und eine Spezifität von 81,0%. Der positiv prädiktive Wert kann mit 55,2% und der negativ prädiktive Wert mit 45,4% angegeben werden (vgl. Tabelle 3.4.2).

3.5 Testgütekriterien für die repetitiven Messungen in den Subgruppen

3.5.1 Schlaganfallpatienten

In der Gruppe der Schlaganfallpatienten konnten in den repetitiven Messungen 47 Patienten mit 180 Messungen untersucht werden. Es ergibt sich unter Berücksichtigung des Cut-Offs von $\geq 5,5$ Punkte auf der ssCAM-ICU folgende Kreuztabelle (vgl. Tabelle 3.5.1.1):

Tabelle 3.5.1.1: Kreuztabelle ssCAM-ICU vs. DSM IV TR, repetitive Messungen, Schlaganfallpatienten

Bewertung nach DSM IV	Bewertung nach ssCAM-ICU		Gesamt
	Positiv	Negativ	
Positiv	29	2	31
Negativ	31	118	149
Gesamt	60	120	180

Dies entspricht folgenden Testgütekriterien (vgl. Tabelle 3.5.1.2):

Tabelle 3.5.1.2: Testgütekriterien, repetitive Messungen, Schlaganfallpatienten

	Sensitivität, %	Spezifität, %	positiv prädiktiver Wert, %	negativ prädiktiver Wert, %
ssCAM-ICU (n=180)	93,6	79,2	56,0	44,6

Es ergibt sich eine Sensitivität von 93,6% und eine Spezifität von 79,2%. Der positiv prädiktive Wert kann mit 56,0% und der negativ prädiktive Wert mit 44,6% berechnet werden.

3.5.2 Neurochirurgische Patienten

In der Gruppe der neurochirurgischen Patienten konnten in den repetitiven Messungen 58 Patienten mit 181 Messungen untersucht werden. Unter Berücksichtigung des Cut-Offs von $\geq 5,5$ Punkte auf der ssCAM-ICU ergibt sich folgende Kreuztabelle (vgl. Tabelle 3.4.2.1):

Tabelle 3.5.2.1: Kreuztabelle ssCAM-ICU vs. DSM IV TR, repetitive Messungen, neurochirurgische Patienten

Bewertung nach DSM IV TR	Bewertung nach ssCAM-ICU		Gesamt
	Positiv	Negativ	
Positiv	3	20	23
Negativ	22	136	158
Gesamt	25	156	181

Aus Tabelle 3.5.2.1 lassen sich die in Tabelle 3.5.2.2 näher beschriebenen Testgütekriterien berechnen.

Tabelle 3.5.2.2: Testgütekriterien, erster valider Messtag, neurochirurgische Patienten

	Sensitivität, %	Spezifität, %	positiv prädiktiver Wert, %	negativ prädiktiver Wert, %
ssCAM-ICU (n=181)	87	86,1	52,6	48

Es ergibt sich eine Sensitivität von 87% und eine Spezifität von 86,1%. Der positiv prädiktive Wert beträgt 52,6% und der negativ prädiktive Wert 48% (vgl. Tabelle 3.4.2.2).

3.6 ROC-Analyse und Youden-Index: ssCAM-ICU Gesamtkollektiv, erster valider Messtag

Wenn, wie im Methodenteil dieser Arbeit beschrieben, die Sensitivität gegen 1-Spezifität aufgetragen wird, dann ergibt sich für die Ergebnisse der ssCAM-ICU am ersten validen Messtag eine ROC-Kurve mit einer AUC (95% KI) von 0,91 (0,85 – 0,97): (vgl. Abbildung 3.6).

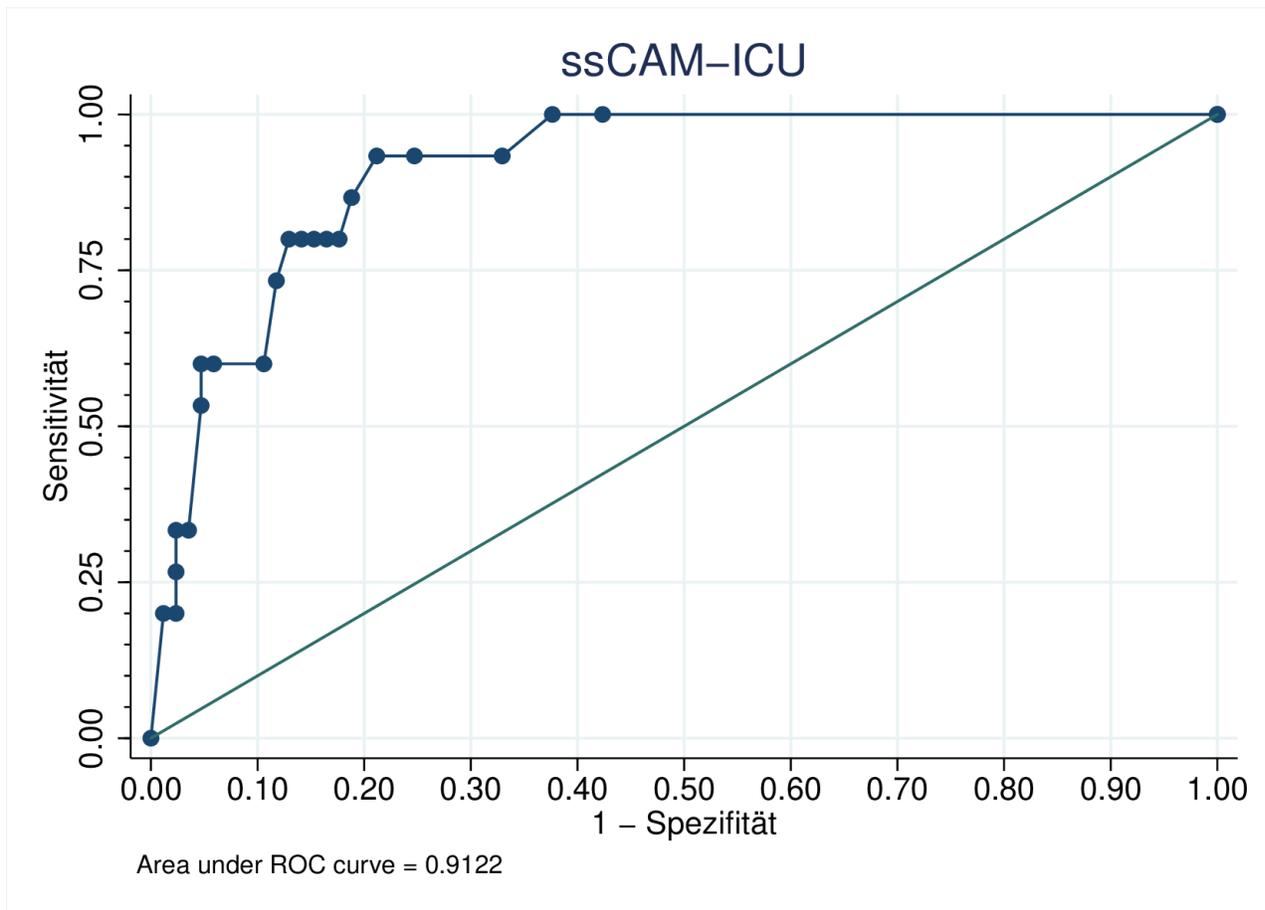


Abbildung 3.6: ROC-Kurve für die Punktwerte auf der ssCAM-ICU, erste valide Messung, Gesamtkollektiv, (vgl. auch Weiss et al., 2021)

Am ersten validen Messtag zeigt sich der optimale Cut-Off für die Diagnose eines Delirs bei $\geq 2,5$ Punkten auf der ssCAM-ICU. An diesem Punkt beträgt der Youden-Index 0,72 (vgl. Youden, 1950). Bei diesem Cut-Off liegt die Sensitivität bei 93,3% und die Spezifität bei 78,8%.

3.7 ROC-Analyse und Youden-Index: ssCAM-ICU Gesamtkollektiv, repetitive Messungen

Bei einer Wiederholung der ROC-Analyse und der Cut-Off Bestimmungen für die repetitiven Messungen ergibt sich folgende Grafik mit einer AUC von 0,91 (vgl. Abbildung 3.7).

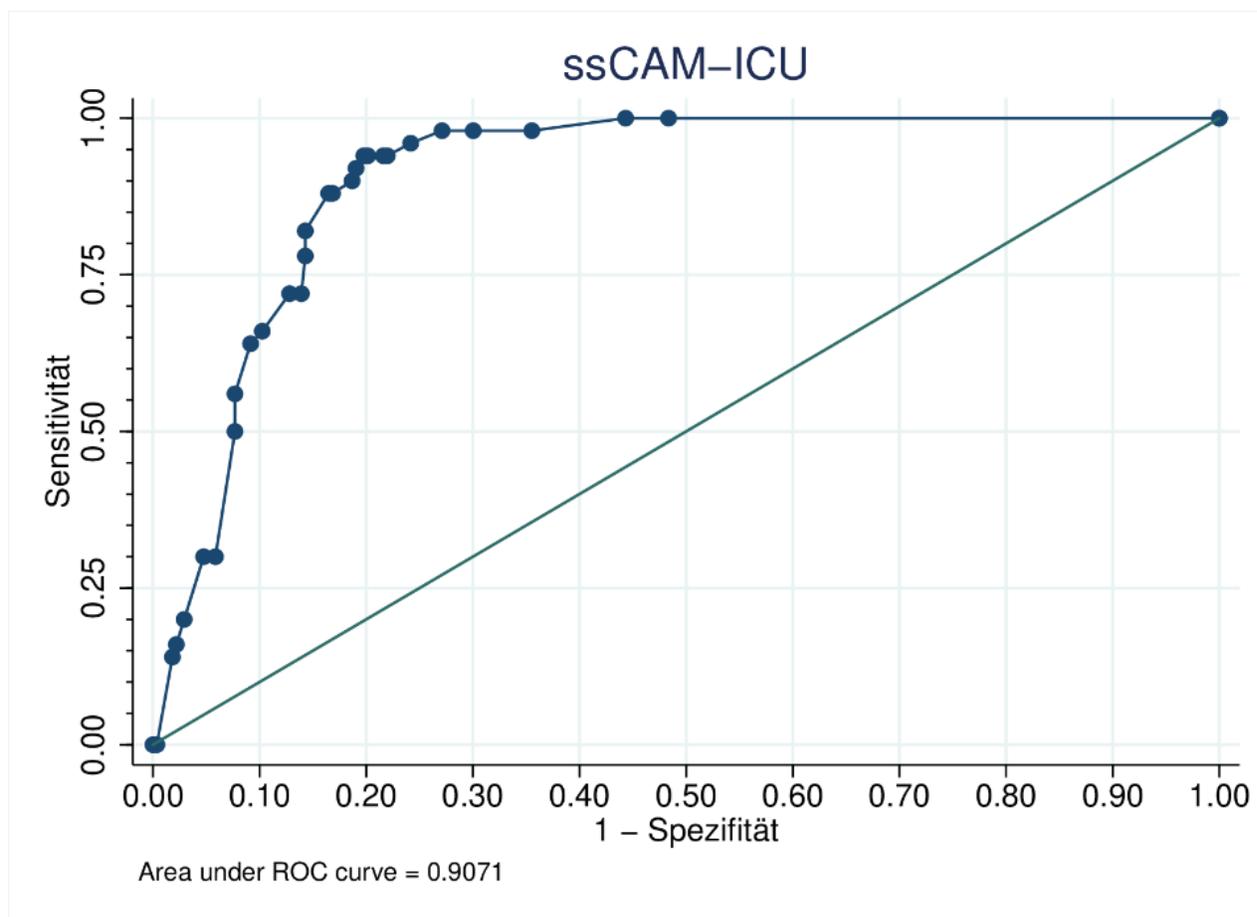


Abbildung 3.7: ROC-Kurve für die Punktwerte auf der ssCAM-ICU, repetitive Messungen, Gesamtkollektiv, (vgl. auch Weiss et al., 2021)

Der Cut-Off Wert mit der höchsten Sensitivität und der höchsten Spezifität beträgt ≥ 5 Punkte auf der ssCAM-ICU. Hier lassen sich eine Sensitivität von 94,0% und eine Spezifität von 80,2% nachweisen. Der Youden-Index (vgl. Youden, 1950) an diesem Punkt beträgt 0,74.

3.8 ROC-Analyse und Youden-Index: ssCAM-ICU bei Schlaganfallpatienten, repetitive Messungen

Im Rahmen einer Grenzwertoptimierungsanalyse für die Subgruppe der 47 Patienten, die wegen eines Schlaganfalls behandelt wurden, lässt sich folgende ROC-Kurve zeichnen (vgl. Abbildung 3.8). Für die Gruppe der Schlaganfallpatienten standen 180 Messungen zur Auswertung zur Verfügung, wovon 31 Messungen delirpositiv mit der DSM IV TR waren. Die AUC beträgt 0,9.

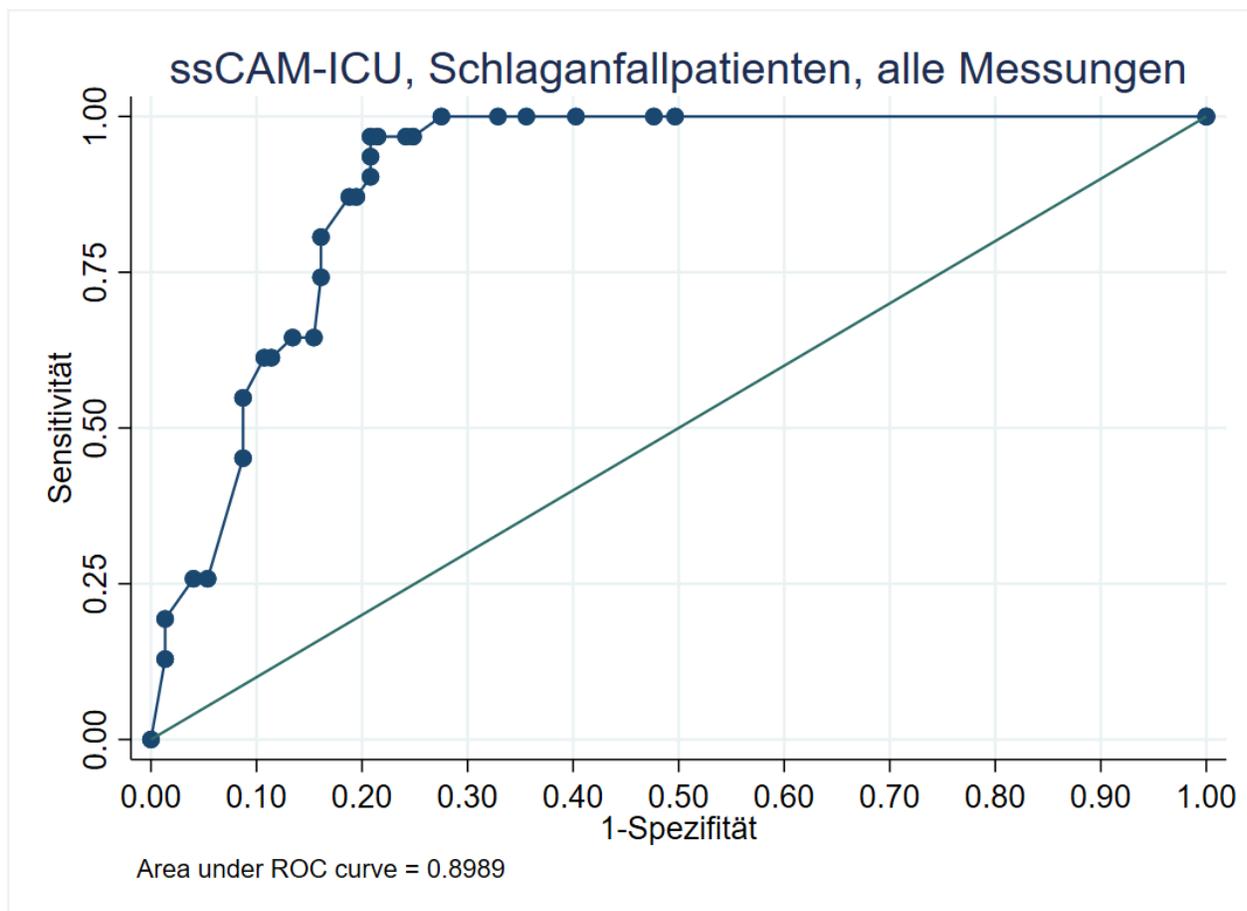


Abbildung 3.8: ROC-Kurve für die Punktwerte auf der ssCAM-ICU, repetitive Messungen, Schlaganfallpatienten

In der Subgruppe der Schlaganfallpatienten kann bei einem optimalem Cut-off von ≥ 5 und einem Youden-Index (vgl. Youden, 1950) von 0,76 eine Sensitivität von 96,8% und eine Spezifität von 79,2% angegeben werden.

3.9 ROC-Analyse und Youden-Index: ssCAM-ICU bei neurochirurgischen Patienten, repetitive Messungen

Im Rahmen einer Grenzwertoptimierungsanalyse für die Subgruppe der 58 Patienten, die wegen einer neurochirurgischen Intervention behandelt wurden, lässt sich folgende ROC-Kurve berechnen (vgl. Abbildung 3.9). Für die Gruppe der neurochirurgischen Patienten standen 181 Messungen mit 23 DSM IV TR positiven Messungen zur Verfügung. Die dargestellte ROC-Kurve hat eine AUC von 0,94.

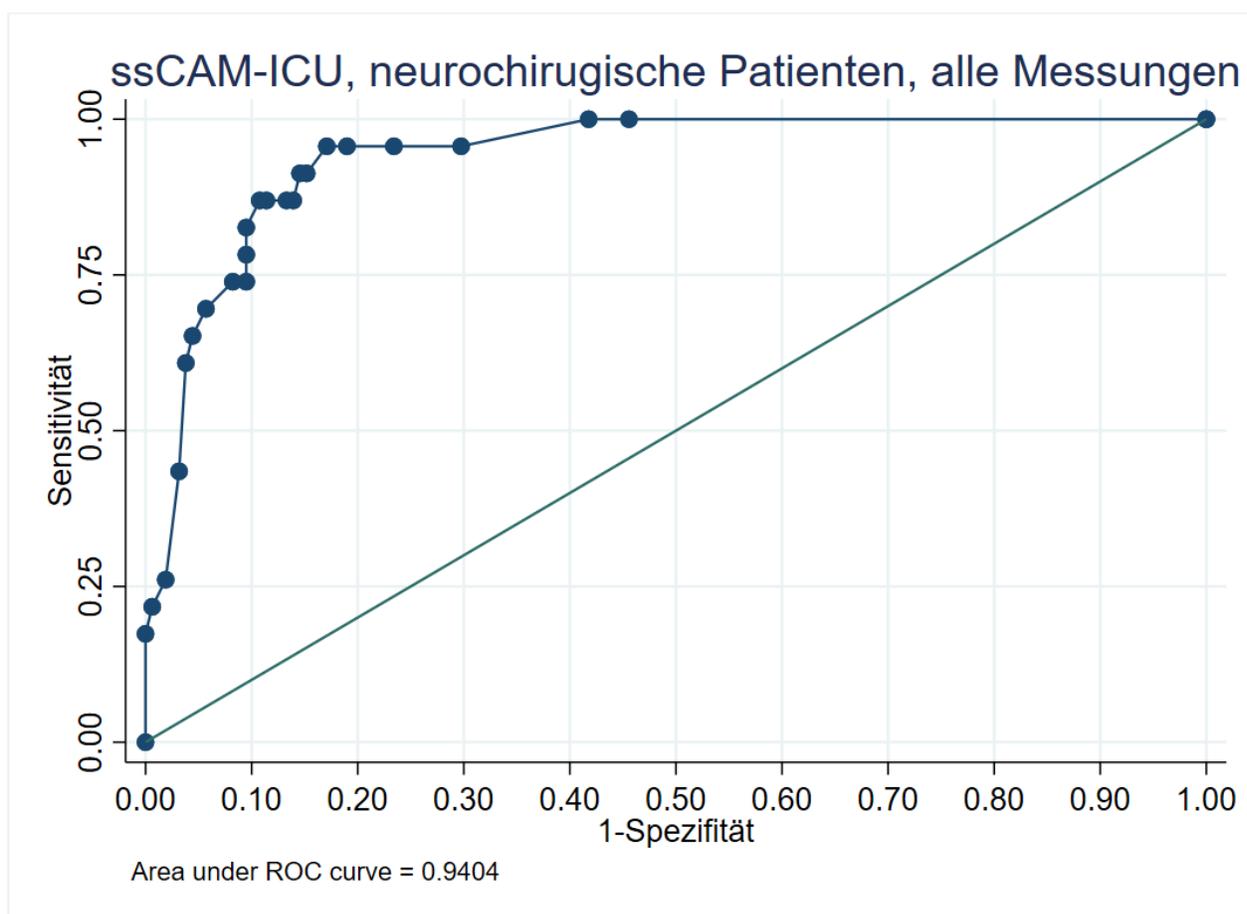


Abbildung 3.9: ROC-Kurve für die Punktwerte auf der ssCAM-ICU, repetitive Messungen, neurochirurgische Patienten

Für die neurochirurgischen Patienten ergibt sich ein optimaler Cut-Off von ≥ 3 mit einer Sensitivität von 95,7% und einer Spezifität von 82,9%. Der Youden-Index (vgl. Youden, 1950) beträgt 0,79.

4. Diskussion

Die vorliegende Arbeit untersucht die diagnostische Validität der ssCAM-ICU auf einer Intensivstation, die interdisziplinär sowohl neurologische als auch neurochirurgische Patienten behandelt. Als Cut-Off wurde der von Nacul et al. vorgeschlagene Wert $\geq 5,5$ Punkte auf der ssCAM-ICU verwendet (vgl. 2020).

Bei der hier vorliegenden Arbeit handelt es sich um eine Auswertung einer im Rahmen eines Qualitätsmanagementprojekts durchgeführten prospektiven Erhebung zur Validierung eines übersetzten Messinstrumentes gegenüber dem Referenzstandard des *Diagnostic and Statistical Manual of Mental Disorders* (DSM), das zum Erhebungszeitpunkt in der vierten textrevidierten Version vorlag (DSM IV TR). Dies stellt auch im internationalen Vergleich das Standardverfahren zur Validierung von Messinstrumenten dar (vgl. Wild et al., 2005; Luetz et al., 2010a).

Die Delirprävalenz in dem hier untersuchten Patientenkollektiv lag bei 15,4%. Am ersten validen Messtag ließen sich mit der ssCAM-ICU eine Sensitivität von 82,4% und eine Spezifität von 84,7% bei Verwendung eines Cut-Offs von $\geq 5,5$ Punkte auf der ssCAM-ICU nachweisen. Für die repetitiven Messungen ergab sich unter Berücksichtigung des gleichen Cut-Offs eine Sensitivität von 92% und eine Spezifität von 81%. In der hier vorliegenden Arbeit konnte somit nachgewiesen werden, dass die ssCAM-ICU am ersten Messtag eine gute und bei Wiederholungsmessungen eine sehr gute Sensitivität bei guter Spezifität unter Berücksichtigung des Cut-Offs von $\geq 5,5$ Punkte auf der ssCAM-ICU zur Detektion des Delirs zeigt.

In den Subgruppenanalysen bestätigte sich dieser Befund auch in den *a priori* festgelegten Gruppen Apoplex und neurochirurgische Interventionen unter Berücksichtigung des bereits genannten Cut-Offs (Schlaganfallpatienten: Sensitivität 93,6%, Spezifität 79,2%, neurochirurgische Patienten: Sensitivität 87%, Spezifität 86,1%).

Im Rahmen der für die Optimierung des Cut-Offs durchgeführten ROC-Analysen konnte für die ssCAM-ICU im Gesamtkollektiv exzellente Testgütekriterien für einen Screeningtest sowohl am ersten Messtag als auch in den repetitiven Messungen nachgewiesen werden (Sensitivität: 93,1% bzw. 94,1%, Spezifität 78,8% bzw. 80,2%). Der optimale Cut-Off lag am ersten validen Messtag bei $\geq 2,5$ Punkte auf der ssCAM-ICU. Bei den repetitiven Messungen ergab sich ein optimaler Cut-Off von ≥ 5 Punkte auf der ssCAM-ICU. Auch in den jeweiligen Subgruppen zeigte die ssCAM-ICU exzellente

Testgütekriterien für einen Screeningtest. In der Gruppe der Schlaganfallpatienten bzw. der neurochirurgischen Patienten wurde eine Sensitivität von 96,8% bzw. 95,7% und eine Spezifität von 79,2% bzw. 82,9% berechnet. Der Cut-Off variierte in den jeweiligen Subgruppen und wurde in der Gruppe der Schlaganfallpatienten mit ≥ 5 und in der Gruppe der neurochirurgischen Patienten mit ≥ 3 Punkte auf der ssCAM-ICU bestimmt.

4.1 Basischarakteristika des Patientenkollektivs und Delirprävalenz

4.1.1 *Patientenkollektiv*

Das hier abgebildete Patientenkollektiv repräsentiert einen breiten Querschnitt neurologischer bzw. neurochirurgischer Erkrankungen. Die Erhebung hat an einem universitären Krankenhaus der Maximalversorgung stattgefunden, weshalb die Krankheitsschwere des Patientenkollektivs im Vergleich zu anderen Arbeiten als vergleichbar hoch anzusehen ist (vgl. Ely et al., 2004; van Eijk et al., 2009; Luetz et al., 2010; van Eijk et al., 2011; Larsen et al., 2019; Nacul et al., 2020).

Bemerkenswert ist die in Bezug auf das im Rahmen dieser Arbeit untersuchte Gesamtkollektiv niedrige Prävalenz an beatmeten Patienten (43,6%). In der bisher international publizierten Literatur zum Delir ist dieser Wert vergleichsweise niedrig (vgl. Ely et al., 2001; Frenette et al., 2015; Larsen et al., 2019). Allerdings existieren insbesondere im Bereich der neurologischer Patienten Arbeiten, wie beispielsweise von Mitasova et al., die mit 7% beatmeten Patienten noch deutlich niedrigere Beatmungsfrequenzen aufweisen (vgl. 2012). Eine deutsche Arbeitsgruppe gibt einen Anteil von 14,6% an (vgl. von Hofen-Hohloch et al., 2020) und Reznik et al. berichten ebenfalls von einer geringeren Prävalenz von beatmeten Patienten (20%) (vgl. 2020).

Hierbei ist besonders zu beachten, dass Patienten mit persistierender Vigilanzstörung und Koma in der vorliegenden Arbeit ausgeschlossen werden mussten. Deshalb wurde ein Teil der beatmeten Neurointensivpatienten nicht erfasst, sodass diese Patienten in dem hier untersuchten Kollektiv unterrepräsentiert sind.

In Übereinstimmung mit der aktuellen Literatur zeigte sich, dass Patienten, die ein Delir erlitten, länger beatmet waren (vgl. Ely et al., 2004; Luetz et al., 2010; Wang et al., 2018) und einen fast fünfmal längere Intensivbehandlungsdauer hatten (vgl. Luetz et al., 2010; Ely et al., 2004; Wang et al., 2018). Delirante Patienten wiesen statistisch höhere Werte sowohl auf der APACHE II als auch im SAPS II und eine höhere Mortalität im Rahmen der Behandlung auf. Die Werte für die beiden Scores zur

Abschätzung der Erkrankungsschwere bzw. der Mortalität sind vergleichbar mit den Werten, die in anderen Untersuchungen an gemischten bzw. nicht neurologisch/neurochirurgischen Patientenkollektiven gezeigt werden konnten (vgl. Ely et al., 2001; Ely et al., 2004; van Eijk et al., 2009; Luetz et al., 2010; van Eijk et al., 2011). Sie entsprechen einer mittleren Erkrankungsschwere und somit einem mittleren Sterberisiko.

Innerhalb des hier untersuchten speziellen Patientenkollektivs waren jedoch größere Unterschiede erkennbar. So war der mittlere APACHE II in der deliranten Gruppe mit 20 Punkten fast doppelt so hoch wie in der nicht-deliranten Gruppe. Dies korreliert gut mit den Ergebnissen von Ouimet et al., die nachweisen konnten, dass bei Intensivpatienten ein Anstieg der APACHE II um einen Punkt mit einer Erhöhung des Risikos, ein Delir zu erleiden, um ca. 5% korreliert (vgl. 2007). Der SAPS II zeigte in der Gruppe der deliranten Patienten anderthalbfach so hohe Ergebnisse. Sowohl Frenette et al. als auch Wang et al. untersuchten neurochirurgische Patienten mit einem deutlich niedrigeren Wert auf der APACHE II (vgl. Frenette et al., 2015; Wang et al., 2018).

Nach der Bewertung mit der APACHE II war das Patientenkollektiv von Larsen et al. ähnlich schwer erkrankt (vgl. 2019). Außer den Arbeiten von Lütz et al. (vgl. Luetz et al., 2010) und Larsen et al. (vgl. Larsen et al., 2019) berücksichtigt keine weitere Veröffentlichung den SAPS zur Beschreibung der Krankheitsschwere. Die in beiden Studien angegebenen Werte korrelieren gut mit den in dieser Arbeit erhobenen Werten. Einige Studien verwendeten zur Beschreibung des Patientenkollektivs keinen intensivmedizinischen Score (vgl. Reznik et al., 2020; von Hofen-Hohloch et al., 2020) oder den SOFA (vgl. Mitasova et al., 2012). Der für den SOFA angegebenen Mittelwert von 1,1 und ein Maximum von 2,3 Punkten lässt auf ein Patientenkollektiv schließen, dessen Erkrankungsschwere als eher gering angesehen werden muss (vgl. Mitasova et al., 2012).

Die Mortalität während der aktuellen Krankenhausbehandlung war in der deliranten Gruppe erhöht (16,7% vs. 0%, $p=0,002$). Auch wenn die Mortalität während der Krankenhausbehandlung selten angegeben wird, so stellten Lütz et al. ebenfalls eine erhöhte Mortalität bei den deliranten Patienten fest, wobei die in der vorliegenden Arbeit nachgewiesene Mortalität noch übertroffen wurde (vgl. 2010). Da im Rahmen dieser Arbeit keine Nachverfolgung der Patienten geplant war, ist keine Aussage zur Mortalität nach einem längeren Zeitraum möglich.

4.1.2 Delirprävalenz

Die Delirprävalenz am ersten Messtag betrug 15% bzw. 15,4% unter Berücksichtigung der Gesamtbehandlungsdauer.

Van Eijk et al. zeigten in einer prospektiven, monozentrischen Studie an 126 Patienten einer interdisziplinären universitären Intensivstation eine Punktprävalenz von 34,1% am ersten Messtag (vgl. 2009). Für die 28 Patienten der neurologischen/neurochirurgischen Subgruppe ergibt sich eine Prävalenz von 35,7%. In einer größeren prospektiven Studie an mehreren Krankenhäusern unterschiedlicher Versorgungsstufen gelang am ersten Messtag der Nachweis einer Gesamtprävalenz von 41,4% (vgl. van Eijk et al., 2011). In diesem Patientenkollektiv von 282 Patienten betrug der Anteil neurologischer/neurochirurgischer Patienten 12,1%. Jedoch waren von diesen 34 neurologischen/neurochirurgischen Patienten insgesamt elf Patienten zu keinem Zeitpunkt zu evaluieren. Sieben der verbliebenen 23 Patienten zeigten ein Delir am ersten Messtag, was einer Prävalenz von 30,4% entspricht. In beiden Arbeiten erfolgte die Diagnosestellung des Delirs nach den Kriterien des DSM IV TR.

Bei der Betrachtung von Arbeiten mit einem Schwerpunkt auf neurologischen bzw. neurochirurgischen Intensivpatienten zeigen sich auf den ersten Blick deutlich höhere Werte in der Punktprävalenz. So diagnostizierten Mitasova et al. nach den Kriterien des DSM IV ein Delir am ersten Messtag bei 37 von 129 Patienten (28,7%). Bezogen auf die jemals delirpositiven Patienten ergibt sich eine Prävalenz von 67,3% am ersten Messtag (vgl. Mitasova et al., 2012). Die im Rahmen der hier vorliegenden Arbeit untersuchten Patienten waren im Mittel deutlich jünger als die in der prospektiven Studie von Mitasova et al. untersuchten Patienten. Ferner wurden durch Mitasova et al. 66 Patienten (51,2%) eingeschlossen, bei denen entweder eine Demenz vorlag oder ein erhöhtes Risiko für die Entwicklung einer Demenz vorlag. Diese Erkrankung stellt bei geriatrischen Patienten einen unabhängigen Risikofaktor für die Entwicklung eines Delirs dar (vgl. Inouye, 1999) und könnte ursächlich für die höhere Prävalenz sein. Des Weiteren stellt sich die Frage, wie konsequent Patienten mit Koma bzw. schwerer Bewusstseinsstörung ausgeschlossen wurden, da es hier zu mehr falsch positiven Testergebnissen kommen kann. Es ist jedoch anzumerken, dass die schwere Bewusstseinsstörung und das Koma eigene diagnostische Entitäten sind und nicht als Delir gelten. Insbesondere in der DSM-5 wird dieser Abgrenzung Rechnung getragen (vgl. EDA, 2014).

Weitere Arbeiten zeigten Gesamtprävalenzen zwischen 23,2% und 67% (vgl. Nacul et al., 2020; Wang et al., 2018; Frenette et al., 2015; Reznik et al., 2020; Larsen et al., 2019).

Die in dieser Arbeit nachgewiesene Punktprävalenz am ersten validen Messtag und den repetitiven Messungen ist im Vergleich zu den Angaben in der Literatur eher niedrig. Die wenigen verfügbaren Studien an ähnlichen Patientenkollektiven zeichnen ein heterogenes Bild. Je nach Größe des betrachteten Patientenkollektivs, der Art des Ein- bzw. Ausschlusses und der Messmethode in Verbindung mit den zugrundeliegenden Erkrankungen schwanken die Werte für die Punktprävalenz am ersten Messtag und den repetitiv durchgeführten Messungen erheblich. Mögliche Ursachen für die im Rahmen dieser Arbeit nachgewiesene geringe Prävalenz sind das geringe Durchschnittsalter des Patientenkollektivs, der Ausschluss von Patienten mit Anhalt für eine Demenz, der verhältnismäßig geringe Anteil an beatmeten Patienten und ggf. eine Unterrepräsentation der komplizierten Verläufe und schweren Erkrankungen, da diese Patienten unter Umständen niemals zu untersuchen waren.

4.2 Testgütekriterien der ssCAM-ICU im Gesamtkollektiv

Um die in der Einleitung dargelegten Schwächen eines dichotomen Messinstruments zu umgehen, ist eine ordinale Skala auf Basis der CAM-ICU entwickelt worden (vgl. Luetz et al., 2016; Nacul et al., 2020). Die ssCAM-ICU ist erstmals in einer pädiatrischen Version von Luetz et al. beschrieben worden (vgl. 2016) und wurde nun auch auf ein adultes Kollektiv übertragen und validiert (vgl. Nacul et al., 2020). Ziel der hier vorliegenden Arbeit ist die Validierung der ssCAM-ICU für das untersuchte Patientenkollektiv gegen den Referenzstandard DSM IV TR.

Da in der Literatur bisher nur eine Validierungsstudie zur ssCAM-ICU veröffentlicht wurde, wird zunächst eine Einordnung der Ergebnisse im Vergleich zur CAM-ICU vorgenommen. Die unter Berücksichtigung des Cut-Off von $\geq 5,5$ Punkte auf der ssCAM-ICU festgestellten Testgütekriterien für die ssCAM-ICU gegen den Referenzstandard DSM IV TR liegen unter denen von Ely et al. im Rahmen der Originalpublikation für die CAM-ICU veröffentlichten Testgütekriterien (Sensitivität 95-100%, Spezifität 89-93%) (vgl. 2001). Neben dieser Arbeit konnten lediglich Günther et al. eine ähnlich hohe Sensitivität erreichen (Sensitivität 88-93%, Spezifität 100%) (vgl. 2010). Im Vergleich zu den übrigen Publikationen, die die CAM-ICU gegen einen Referenzstandard untersuchen, zeigen sich den in der hier vorliegenden Arbeit nachgewiesenen Werten für die ssCAM-ICU ähnliche oder bessere Testgütekriterien.

Bei neurologischen bzw. neurochirurgischen Patienten stellt sich je nach verwendetem Referenzstandard und Testmethode ein heterogeneres Bild dar. Bei strenger Validierung der CAM-ICU gegen das

DSM IV in einer Kohorte von 129 Schlaganfallpatienten zeigen sich Testgütekriterien, die eine schlechtere Sensitivität und eine bessere Spezifität im Vergleich zur ssCAM-ICU aufweisen (Sensitivität 76% und Spezifität 98,1%) (vgl. Mitasova et al., 2012). In Abhängigkeit vom Ausmaß der Schädigung im Rahmen eines Schädel-Hirn-Traumas (mild: Sensitivität 51%, Spezifität 81% vs. moderat: 73% bzw. 43%) oder der Sedierung (RASS ≥ -1 : Sensitivität 61%, Spezifität 75%, vs. RASS -2/-3: 72% bzw. 57%) zeigte sich bei der CAM-ICU gegen den Referenzstandard DSM IV TR eine deutlich schlechtere diagnostische Validität (vgl. Frenette et al., 2015). Die in der hier vorliegenden Arbeit nachgewiesenen Werte für die ssCAM-ICU zeigen sowohl am ersten validen Messtag als auch in den repetitiven Messungen bessere Werte als bei Frenette et al. angegeben (vgl. 2015). Bezogen auf den Referenzstandard DSM V berichten Reznik et al. bei restriktiver Anwendung der CAM-ICU gemäß der ursprünglichen Validierungsstudie von Ely et al. (vgl. 2001) schlechtere Testgütekriterien mit einer Sensitivität von 67,9% und einer Spezifität von 88% (vgl. Reznik et al., 2020). Sofern neurologische Erkrankungen, die eine Einschränkung der Vigilanz bewirkten, oder eine Aphasie oder Abulie vorlagen, konnte bei liberalerer Anwendung nur eine Sensitivität von 40,7% und eine Spezifität von 88% nachgewiesen werden (vgl. Reznik et al., 2020).

Wenn als Referenzstandard die Definition des Delirs nach ICD-10 verwendet wird, dann ergeben sich mittelmäßige Testgütekriterien mit einer Sensitivität von 59% und einer Spezifität von 56% für die CAM-ICU (vgl. Larsen et al., 2019). Bei Betrachtung gegen eine klinische Einschätzung als Referenzstandard, der sowohl Kriterien der DSM V, der ICD-10 und einer klinischen Einschätzung des behandelnden Teams beinhaltet, zeigte die CAM-ICU eine Sensitivität von 66,9% und eine Spezifität von 93,3% (vgl. von Hofen-Hohloch et al., 2020). Die ssCAM-ICU weist im Gesamtkollektiv somit durchweg eine bessere Sensitivität und eine vergleichbare Spezifität bezogen auf die CAM-ICU bei Neurointensivpatienten auf. Selbstverständlich sollten bei der Interpretation der Daten der jeweilige Referenzstandard, die Selektion des Patientenkollektivs und die angewandte Methodik berücksichtigt werden.

Die hier untersuchte ssCAM-ICU ist bisher erst in einer Arbeit an erwachsenen Intensivpatienten validiert worden und es ergab sich dabei ein optimaler Cut-Off von $\geq 5,5$ Punkten auf der ssCAM-ICU. Die Sensitivität konnte mit 91,7% und die Spezifität mit 82,5% angegeben werden (vgl. Nacul et al., 2020). Eine Differenzierung zwischen der ersten validen Messung und den repetitiven Messungen

erfolgte nicht. Im Rahmen der zitierten Arbeit wurde keine Patienten untersucht, die neurologische oder neurochirurgische Erkrankungen erlitten.

Berücksichtigt man den vorgeschlagenen Cut-Off von $\geq 5,5$ Punkte auf der ssCAM-ICU (vgl. Nacul et al., 2020) zeigen sich bei den hier untersuchten Neurointensivpatienten am ersten validen Messtag schlechtere Testgütekriterien für die ssCAM-ICU. Unter Berücksichtigung der repetitiven Messungen lässt sich eine vergleichbare Sensitivität und Spezifität im hier untersuchten Patientenkollektiv nachweisen.

In der im Rahmen dieser Arbeit durchgeführten Grenzwertoptimierung mittels ROC-Analyse und Youden-Index zeigt sich jedoch ein anderer optimaler Cut-Off als jener, den Nacul et al. vorgeschlagen haben (vgl. 2020). Bei Betrachtung der ersten validen Messung des Gesamtkollektivs, ergibt sich ein Cut-Off $\geq 2,5$ Punkte auf der ssCAM-ICU. An diesem Punkt zeigt die ssCAM-ICU eine überlegene Sensitivität bei leicht schlechterer Spezifität. Nach dem Modell von Greiner et al. spricht die ermittelte AUC von 0,91 für einen hochakkuraten Test (vgl. 2000). Der Youden-Index an diesem Punkt ist mit 0,72 anzugeben (vgl. Youden, 1950).

Für die repetitiv durchgeführten Messungen zeigt sich ein ähnliches Bild. Der optimale Cut-Off im Rahmen der hier vorliegenden Arbeit kann mit ≥ 5 Punkten auf der ssCAM-ICU angegeben werden. Die Testgütekriterien zeigen hierbei vergleichbare Werte wie bei Nacul et al. (vgl. 2020). Der Youden-Index an diesem Punkt liegt bei 0,74 (vgl. Youden, 1950) und die AUC von 0,91 spricht für einen hochakkuraten Test (vgl. Greiner et al., 2000).

Der Unterschied bezüglich des Cut-Offs zwischen dem ersten validen Messtag und den repetitiven Messungen lässt sich aufgrund der geringeren Anzahl der in die Berechnung für den ersten validen Messtag einfließenden Messungen erklären. Es ist möglich, dass aufgrund der geringeren Anzahl an berücksichtigten Messungen bestimmte Punktwerte nicht überprüfbar sind, da sie am ersten validen Messtag nicht auftraten oder die Patienten ein zu diesem Zeitpunkt noch geringer ausgeprägtes Delir aufwiesen. Somit sollte eine entsprechende Cut-Off Bestimmung für den ersten validen Messtag an einem größeren Kollektiv von Neurointensivpatienten wiederholt werden. Ebenso muss berücksichtigt werden, dass Nacul et al. keine Differenzierung hinsichtlich der ersten validen Messung und den Wiederholungsmessungen für die ssCAM-ICU vornehmen (vgl. 2020). Betrachtet man weitere Arbeiten, zeigt sich ebenfalls, dass sich die diagnostische Validität der übrigen Delirmessinstrumente

bei Mehrfachmessungen in chirurgischen oder internistischen Intensivkohorten leicht verändern kann (vgl. Luetz et al, 2010; Luetz et al. 2016; Nacul et al., 2020). Sowohl eine Verbesserung als auch eine Verschlechterung der diagnostischen Validität sind beschrieben und sollten in weiteren Studien eingehender untersucht werden.

Bei Betrachtung des optimalen Cut-Offs für die repetitiven Messungen des Gesamtkollektivs zeigt sich ein, im Vergleich zu dem von Nacul et al. angegebenen, vergleichbarer Cut-Off (vgl. 2020). Auch die Testgütekriterien zeigen am optimalen Cut-Off ähnliche Werte. Da in der klinischen Routine Mehrfachmessungen der Realität entsprechen, sollte dies bei der Betrachtung des optimalen Cut-Offs berücksichtigt und der optimale Cut-Off anhand repetitiver Messungen ermittelt werden. Eine Verzerrung der Testgütekriterien aufgrund repetitiver Messungen erfolgt bei dem hier gewählten Vorgehen nicht, da es sich bei den einzelnen Werten der ROC-Kurve um Punktschätzwerte handelt, die voneinander nicht beeinflusst werden. Primär würde sich eine solche Verzerrung in einer künstlichen Verkleinerung der Konfidenzintervalle äußern.

Zusammenfassend ist festzustellen, dass die im Rahmen dieser Arbeit für die ssCAM-ICU erhobenen Werte für die Testgütekriterien den Angaben in der Literatur für die CAM-ICU weitgehend entsprechen oder sie sogar übertreffen. Problematisch sind neben der Verwendung eines anderen Messinstruments die Verwendung unterschiedlicher Referenzstandards, die unterschiedlichen Patientenkollektive, sowohl nach Größe des untersuchten Kollektivs, Erkrankungsart und -schwere, und die Heterogenität der Studiendesigns. Hierdurch wird die Vergleichbarkeit weiter eingeschränkt. Betrachtet man die einzige vorliegende Arbeit zur ssCAM-ICU zeigen sich vergleichbare Testgütekriterien und ein vergleichbarer Cut-Off. Es sollten jedoch weitere Arbeiten zur Überprüfung der ssCAM-ICU an unterschiedlichen Patientenkollektiven durchgeführt werden, um eine generelle Übertragbarkeit des ermittelten Cut-Offs zu überprüfen und einen dem jeweiligen Kollektiv angepassten Cut-Off festzulegen.

4.3 Testgütekriterien der ssCAM-ICU in den Subgruppen

Bei Schlaganfallpatienten zeigt sich eine sehr gute Sensitivität von 96,8% und eine gute Spezifität von 79,2% bei einem Cut-Off ≥ 5 Punkte auf der ssCAM-ICU. Der Youden-Index kann mit 0,76 angegeben werden (Youden, 1950) und die AUC von 0,9 spricht für einen akkuraten bis hoch akkuraten Test (vgl. Greiner et al., 2000). Berücksichtigt man den Cut-Off, den Nacul et al. vorgeschlagen haben (vgl. 2020), von $\geq 5,5$ Punkte auf der ssCAM-ICU, verschlechtert sich die Sensitivität auf 93,6% und die Spezifität bleibt mit 79,2% konstant. Die erreichten Testgütekriterien in dieser Subgruppe sind bei identischem Cut-Off mit den Ergebnissen im Gesamtkollektiv vergleichbar. Allerdings verdeutlicht die Grenzwertoptimierung mittels ROC-Analyse, dass der in einer Gruppe erhobene Cut-Off nicht ohne weiteres auf andere Patientenkollektive übertragen werden kann. Zwar sind im Rahmen der hier aufgezeigten Differenz von 0,5 Punkten auf der ssCAM-ICU die Testgütekriterien nur geringfügig niedriger, führen jedoch zu einer schlechteren Screeningleistung des Tests. Eine Verbesserung der Spezifität erfolgt indes nicht. Damit ist eine geringere Gesamtperformance des Tests assoziiert. Bei Anwendung in dieser Subgruppe übertrifft die ssCAM-ICU die von Mitasova et al. berichtete Sensitivität für die CAM-ICU deutlich (96,8% vs. 76,0%), bleibt jedoch weniger spezifisch (79,2% vs. 98,1%) (vgl. 2012). Zur Verwendung als Screeningtest ist die ssCAM-ICU aufgrund der hohen Sensitivität sehr gut geeignet.

In der Subgruppe der neurochirurgischen Patienten zeigt die ssCAM-ICU ebenfalls eine sehr gute Sensitivität (95,7%) bei guter Spezifität (82,9%). Die AUC von 0,94 weist auf einen hochakkuraten Test hin (vgl. Greiner et al., 2000). Der optimale Cut-Off liegt hier gemäß dem Youden-Index ($J=0,79$) (vgl. Youden, 1950) bei ≥ 3 Punkte auf der ssCAM-ICU. Aus der ROC-Kurve ist ersichtlich, dass bei einem Cut-Off ≥ 5 Punkte auf der ssCAM-ICU die Sensitivität auf 91,3% sinkt und die Spezifität auf 85,4% steigt. Bei diesem Cut-Off beträgt der Youden-Index 0,77 (vgl. Youden, 1950). Somit wäre eine direkte Übertragung des von Nacul et al. vorgeschlagenen Cut-Offs auf die Subgruppe der neurochirurgischen Patienten mit einer Verschlechterung der Testgütekriterien assoziiert (vgl. 2020). In Anbetracht der nur geringen Verschlechterung der Testgütekriterien und der geringen Differenz in den Youden-Indices lässt dieser Unterschied aber nur eine geringe Verschlechterung der Gesamtperformance des Tests erwarten. Hierbei ist zu berücksichtigen, dass der Youden-Index als Indikator lediglich die Gesamtperformance eines Tests erfasst. Es erfolgt keine Gewichtung zwischen Sensitivität und Spezifität. Als eine mögliche Erklärung für den errechneten geringeren optimalen Cut-Off im Kollektiv der neurochirurgischen Intensivpatienten gegenüber dem Gesamtkollektiv lässt

sich anführen, dass neurochirurgische Patienten mit schwereren Verläufen bzw. einem größeren Schädigungsmaß aufgrund der genannten Beschränkungen nicht in die Untersuchung eingeschlossen wurden bzw. unterrepräsentiert sind. Ferner ist die Prävalenz in dieser Subgruppe geringer als im untersuchten Gesamtkollektiv. Somit verwundert es nicht, dass 75,1% der Messungen unterhalb des Cut-Offs von ≥ 3 Punkte auf der ssCAM-ICU liegen.

Nichtsdestotrotz ist die ssCAM-ICU als Screeningtest aufgrund der hohen Sensitivität sehr gut geeignet und zeigt im Vergleich zu den Literaturwerten für die CAM-ICU durchweg bessere Testgütekriterien (vgl. Frenette et al., 2015; Larsen et al., 2019; Reznik et al., 2020; von Hofen-Hohloch et al., 2020). Allerdings könnte beispielsweise ein falsch negatives Ergebnis erhebliche Konsequenzen für die Patienten haben. Falls dies zu befürchten wäre, dann sollte ein Cut-Off mit einer höheren Sensitivität gewählt werden. Der hier bestimmte Cut-Off könnte entsprechend der Intention für die Verwendung des Scores verschoben werden. Wenn jedoch die ssCAM-ICU als Screeninginstrument verwendet werden soll, dann erscheint die höhere Sensitivität sinnvoll. Gegebenenfalls müsste dann in einem Grenzbereich ein Experte hinzugezogen werden, um das positive Screening zu verifizieren.

5. Grenzen und Stärken der Arbeit

Die Ergebnisse dieser Arbeit unterliegen gewissen Limitationen. Hierzu zählen der lange Zeitraum zwischen Datenerhebung und Veröffentlichung dieser Arbeit. Zudem können sich die getroffenen Aussagen lediglich auf das untersuchte Patientenkollektiv beziehen. Zwischenzeitlich ist die Intensivstation, auf der die Daten erhoben wurden, an einen anderen Standort verlegt worden, was mit einer Änderung des Patientenkollektivs einhergehen könnte (bspw. weniger Patienten mit Polytrauma/Schädel-Hirn-Trauma). Ebenso hat sich der Referenzstandard seit der Erhebung der Daten geändert, wobei in der Literatur bisher noch kein Konsens darüber besteht, welcher Referenzstandard als Optimum anzusehen ist (vgl. EDA, 2014). Somit kann die Berücksichtigung des DSM IV TR als Referenzstandard aufgrund der Verbreitung in der Literatur auch als Stärke dieser Arbeit angesehen werden.

Im Rahmen der Datenerhebung ist keine Erhebung neurologischer Defizite erfolgt. Dadurch könnte die Durchführbarkeit der jeweiligen Scores in Subgruppen von Patienten mit bestimmten neurologischen Defiziten gesondert untersucht und festgestellt werden, ob dies einen Einfluss auf die Durchführbarkeit und die Testgütekriterien des untersuchten Messinstruments hat.

Die ordinale ssCAM-ICU konnte in dem hier untersuchten Patientenkollektiv mit sehr guten Testgütekriterien validiert werden, weshalb sich diese als Screeninginstrument eignet. Jedoch muss ein für das jeweilige Kollektiv ermittelter optimaler Cut-Off berücksichtigt werden, um einer schlechteren Validität beispielsweise in Subgruppen vorzubeugen. Im Rahmen der Datenerhebung ist kein Score erhoben worden, mit dem eine Einschätzung der Delirschwere (bspw. DRS-R 98 (vgl. Trzepcz et al., 2001)) möglich wäre. Daher ist eine Validierung als ordinale Messinstrument im Rahmen dieser Arbeit zwar gelungen, jedoch lassen die hier vorgestellten Daten keinen Rückschluss auf die Fähigkeit zur Einschätzung der Schwere des Delirs zu.

Bisher ist keine Arbeit veröffentlicht worden, die ein so breites neurologisches/neurochirurgisches Patientenkollektiv auf einer Intensivstation und mit einem weit verbreiteten Referenzstandard untersucht hat. Neu aufgeworfene Fragestellungen, insbesondere zur Erfassung des Delirs mit einem ordinalen Messinstrument, wurden im Rahmen dieser Arbeit beantwortet.

6. Schlussfolgerungen und Ausblick

Das Delir auf der Intensivstation stellt sowohl für den einzelnen Patienten als auch das gesamte Gesundheitssystem eine hohe Belastung dar. Durch effektives Screening, rasche Detektion und eine Behandlung der Ursachen ist eine Verringerung der Mortalität und Morbidität möglich. Die aktuell vorhandenen Leitlinien fordern ein Screening mittels valider und getesteter Messinstrumente. Das Ziel der hier vorliegenden Arbeit bestand in der Validierung der ordinalen ssCAM-ICU und eine Cut-Off-Optimierung für neurologische und neurochirurgische Intensivpatienten sowie der Betrachtung in den Subgruppen der Patienten mit Schlaganfall und nach neurochirurgischer Intervention.

Im hier untersuchten Patientenkollektiv ist die Prävalenz niedrig. Die ssCAM-ICU zeigt mit den Literaturwerten vergleichbare oder bessere Testgütekriterien und ist in diesem Patientenkollektiv unter den dargelegten Limitationen anwendbar. Unter Berücksichtigung der verschiedenen Referenzstandards, der Patientenkollektive und der differierenden Ausschlusskriterien lassen sich die Unterschiede bei der Betrachtung einzelner Arbeiten im Vergleich zur CAM-ICU erklären.

Insbesondere zur Verwendung als Screeningtest ist die ordinale ssCAM-ICU sehr gut geeignet. Wie im Vergleich zur aktuell verfügbaren Literatur gezeigt werden konnte, muss jedoch der verwendete Cut-Off im jeweils zu untersuchenden Patientenkollektiv individuell festgelegt werden, da ansonsten ggf. eine Verschlechterung der Testgütekriterien auftreten kann. In den untersuchten Subgruppen der Schlaganfallpatienten und der neurochirurgischen Patienten zeigt die ssCAM-ICU im Vergleich zu den bisherigen Arbeiten überlegene Sensitivitäten bei leicht schlechteren Spezifitäten. Unter Berücksichtigung der Intention zur Implementierung muss der Cut-Off entsprechend gewählt werden.

Für künftige Arbeiten empfiehlt sich die Anwendung der ssCAM-ICU an anderen Patientenkollektiven, um ihre Anwendbarkeit in diesem Bereich zu überprüfen. Ebenso sollte eine Betrachtung der Testgütekriterien unter Berücksichtigung weiterer Einflussfaktoren und der neurologischen Komorbiditäten erfolgen. Um dem klinisch tätigen Intensivmediziner ein verlässliches Messinstrument zur Einschätzung der Delirschwere zur Verfügung zu stellen, ist eine Validierung gegen einen Score zur Bestimmung der Delirschwere notwendig. Zusammenfassend könnte die ssCAM-ICU die klinische Beurteilung des Delirs erheblich erleichtern.

7. Anhang

7.1 CAM-ICU Arbeitsblatt (www.icudelirium.org 2020)

Name, Vorname	Datum											
CAM-ICU Arbeitsblatt												
Merkmal 1: akuter Beginn oder schwankender Verlauf Positiv, wenn entweder in 1A oder 1B mit JA beantwortet	Positiv <input type="checkbox"/>	Negativ <input type="checkbox"/>										
<p>1 A: Ist der geistige Zustand des Pat. anders als vor der Erkrankung?</p> <p style="text-align: center;">ODER</p> <p>1 B: Zeigt der Pat. in den letzten 24 h Veränderungen in seinem Geisteszustand, z.B. anhand der Richmond-Skala (RASS), Glasgow Coma Scale (GCS) oder vorausgeganener Delir-Einstufung?</p>												
Merkmal 2: Aufmerksamkeitsstörung Positiv, wenn einer der beiden Scores (2A oder 2B) kleiner als 8 ist.	Positiv <input type="checkbox"/>	Negativ <input type="checkbox"/>										
<p>Zuerst die ASE-Buchstaben versuchen. Falls Pat. diesen Test durchführen kann und das Ergebnis eindeutig ist, Ergebnis dokumentieren und weiter zu Merkmal 3. Falls der Pat. den Test nicht schafft oder das Ergebnis nicht eindeutig ist, werden die ASE-Bilder angewendet. Falls beide Tests notwendig sind, werden die Ergebnisse der ASE-Bilder zur Einstufung verwendet.</p>												
<p>2 A: ASE-Buchstaben: Einstufung notieren (NE für nicht erfasst) <u>Anleitung:</u> Sagen Sie dem Patient: „Ich lese Ihnen jetzt hintereinander einige Buchstaben vor. Wenn Sie ein „A“ hören, drücken Sie meine Hand.“ Dann die folgenden Buchstaben in normaler Lautstärke vorlesen: A N A N A S B A U M (alternativ könnte z.B. A B R A K A D A B R verwendet werden) Einstufung: als Fehler wird gewertet, wenn Pat. die Hand bei einem „A“ nicht drückt und wenn Pat. die Hand bei einem anderen Buchstaben als dem „A“ drückt.</p>	Summe (von 10): _____											
<p>2 B: ASE-Bilder: Einstufung notieren (NE für nicht erfasst)</p>	Summe (von 10): _____											
Merkmal 3: unorganisiertes Denken Positiv, wenn die Summe aus Score 3A und 3B weniger als 4 ergibt	Positiv <input type="checkbox"/>	Negativ <input type="checkbox"/>										
<p>3A: Ja/Nein Fragen (entweder Set 1 oder Set 2 verwenden, falls notwendig tageweise abwechseln)</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: center;">Set 1</th> <th style="text-align: center;">Set 2</th> </tr> </thead> <tbody> <tr> <td>1. Schwimmt ein Stein auf dem Wasser?</td> <td>1. Schwimmt eine Ente auf dem Wasser?</td> </tr> <tr> <td>2. Gibt es Fische im Meer?</td> <td>2. Leben Elefanten im Meer?</td> </tr> <tr> <td>3. Wiegt ein Kilo mehr als 2 Kilo?</td> <td>3. Wiegen zwei Kilo mehr als ein Kilo?</td> </tr> <tr> <td>4. Kann man mit einem Hammer Nägel in die Wand schlagen?</td> <td>4. Kann man mit einem Hammer Holz sägen?</td> </tr> </tbody> </table>	Set 1	Set 2	1. Schwimmt ein Stein auf dem Wasser?	1. Schwimmt eine Ente auf dem Wasser?	2. Gibt es Fische im Meer?	2. Leben Elefanten im Meer?	3. Wiegt ein Kilo mehr als 2 Kilo?	3. Wiegen zwei Kilo mehr als ein Kilo?	4. Kann man mit einem Hammer Nägel in die Wand schlagen?	4. Kann man mit einem Hammer Holz sägen?	Summe (3A und 3B) _____ (max. 5)	
Set 1	Set 2											
1. Schwimmt ein Stein auf dem Wasser?	1. Schwimmt eine Ente auf dem Wasser?											
2. Gibt es Fische im Meer?	2. Leben Elefanten im Meer?											
3. Wiegt ein Kilo mehr als 2 Kilo?	3. Wiegen zwei Kilo mehr als ein Kilo?											
4. Kann man mit einem Hammer Nägel in die Wand schlagen?	4. Kann man mit einem Hammer Holz sägen?											
<p>Summe _____ (1 Punkt für jede richtige der 4 Antworten, max. also 4)</p>												
<p>3B: Aufforderung Sagen Sie dem Pat.: „Halten Sie so viele Finger hoch“, (Untersucher hält 2 Finger hoch). „jetzt machen Sie dasselbe mit der anderen Hand“ (ohne dass erneut die Anzahl der gewünschten Finger genannt wird). Falls Pat. nicht beide Arme bewegen kann, wird für den 2. Teil der Frage die Anleitung „fügen Sie einen Finger hinzu“ gegeben. Summe _____ (max. nur 1 Punkt, wenn Pat. alle Anleitungen vollständig ausführen kann)</p>												
Merkmal 4: Bewusstseinsstörung Positiv, wenn der aktuelle RASS von Null verschieden ist	Positiv <input type="checkbox"/>	Negativ <input type="checkbox"/>										
Gesamt CAM-ICU (Merkmale 1 und 2 UND entweder 3 oder 4 positiv)	Positiv <input type="checkbox"/>	Negativ <input type="checkbox"/>										

Copyright © 2002, E. Wesley Ely, MD, MPH und Vanderbilt University, all rights reserved

7.2 ssCAM-ICU (nach Ely 2001, Luetz 2016 und icudelirium.org 2020)

Schritte		Vorgehen	Punktwert
1.	Einschätzung mit der RASS	RASS ≤ 4: Screening nicht durchführen. Später erneut versuchen RASS ≥ 3: Screening durchführbar. Weiter mit Schritt 2	
		Merkmal 1: akuter Beginn oder schwankender Verlauf	
2.	Beurteilung mit der ssCAM-ICU	Gibt es eine akute Änderung im mentalen Status? (z.B. im Vergleich zu prä-OP?)	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
		Gibt es eine Fluktuation über den Tagesverlauf?	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
		Wenn mind. einmal „Ja“ angekreuzt wurde, addieren Sie 5 Punkte zum Punktwert.	
		Merkmal 2: Aufmerksamkeitsstörung (Attention-Screening-Examination ASE with Letters or Pictures)	
		Vigilanz-A-Test	ASE-Bilder
		<ul style="list-style-type: none"> Geben Sie dem Patienten die Hand. Überprüfen Sie, ob der Patient die Hand drücken kann. Geben Sie dem Patienten folgende Anweisung: „Ich lese Ihnen jetzt einige Buchstaben hintereinander vor. Drücken Sie jedes Mal meine Hand, wenn Sie den Buchstaben ‚A‘ hören.“ Buchstabieren folgendes Wort (ca. 1 Buchstabe/s): A N A N A S B A U M A B R A K A D A B R (alternativ) Fehler: jeder andere Buchstabe als ‚A‘ bei dem gedrückt wird und jedes ‚A‘, bei dem nicht gedrückt wird 	<ul style="list-style-type: none"> Schritt 1: <ul style="list-style-type: none"> Geben Sie folgende Anweisung: „Ich zeige Ihnen nun 5 Bilder, merken Sie sich diese bitte.“ Präsentieren Sie 5 Bilder für jeweils 3 Sekunden. Schritt 2: <ul style="list-style-type: none"> Geben Sie folgende Anweisung: „Ich zeige Ihnen noch mehr Bilder. Bitte zeigen Sie mir, ob sie das jeweilige Bild bereits gesehen haben.“ Präsentieren Sie nun 10 Bilder (5 neue, 5 alte) für jeweils 3 Sekunden. Jede Fehleinschätzung ist ein Fehler
		Für jeden Fehler addieren Sie 0,5 Punkte zum Punktwert.	
		Merkmal 3: unorganisiertes Denken	
		<ul style="list-style-type: none"> Geben sie folgende Anweisung: „Ich stelle Ihnen nun vier einfache Fragen, die Sie mit ‚Ja/Nein‘ oder Kopfnicken/Kopfschütteln beantworten können.“ Fragen Sie folgende Fragen (verwenden Sie abwechselnd beide Testsets): <ul style="list-style-type: none"> Schwimmt ein Stein auf dem Wasser? Gibt es Fische im Meer? Wiegt ein Kilo mehr als zwei Kilo? Kann man mit einem Hammer Nägel in die Wand schlagen? Geben Sie folgende Anweisung: „Halten Sie so viele Finger hoch.“ (Untersucher präsentiert 2 Finger) Geben Sie folgende Anweisung: „Jetzt machen Sie das dasselbe mit der anderen Hand/fügen Sie einen Finger hinzu.“ 	<ul style="list-style-type: none"> Schwimmt eine Ente auf dem Wasser? Leben Elefanten im Meer? Wiegen zwei Kilo mehr als ein Kilo? Kann man mit einem Hammer Holz sägen?
		Für jeden Fehler addieren Sie 1 Punkt zum Punktwert.	
		Merkmal 4: Bewusstseinsstörung	
		<ul style="list-style-type: none"> Bewerten Sie die aktuelle Vigilanz des Patienten mit der RASS. 	
		Für jeden RASS-Wert addieren Sie den Betrag zum Punktwert (z.B. RASS +4=4 Punkte, RASS -2=2 Punkte)	
		Addieren Sie die Punkte	
		Interpretation: <5 Punkte: kein Delir ≥5 - 19 Punkte: Delir	
3.	Punktwert berechnen		Summe

7.3 GCS (nach Teasdale und Jennett 1974)

Tabelle 7.3: GCS nach Teasdale und Jennett 1974

System	Beschreibung	Punkte
Augenöffnen	kein	1
	auf Schmerzreiz	2
	auf Ansprache	3
	spontan	4
beste verbale Reaktion	keine	1
	Stöhnen oder unzusammenhängende Laute	2
	unzusammenhängende Worte	3
	desorientiert	4
	konversationsfähig	5
beste motorische Reaktion	keine	1
	Strecken auf Schmerzreiz	2
	Beugen auf Schmerzreiz	3
	ungezielte Schmerzabwehr	4
	gezielte Schmerzabwehr	5
	befolgt Aufforderungen	6

7.4 BPS und BPS NI

Tabelle 7.4.1: BPS aus Baron et al. 2015, modifiziert nach Payen et al. 2001

Item	Beschreibung	Punkte	[Payen JF et al. 2001 (modifiziert)]
Gesichtsausdruck	Entspannt	1	
	Teilweise angespannt	2	
	Stark angespannt	3	
	Grimassieren	4	
Obere Extremität	Keine Bewegung	1	
	Teilweise Bewegung	2	
	Anziehen mit Bewegung der Finger	3	
	Ständiges Anziehen	4	
Adaptation an Beat- mungsgerät	Beatmung wird gut toleriert	1	
	Seltenes Husten	2	
	Kämpfen mit dem Beatmungsgerät	3	
	Kontrollierte Beatmung nicht möglich	4	

Tabelle 7.4.2: BPS-NI aus Baron 2015, modifiziert nach Chanques 2009

Item	Beschreibung	Punkte	[Chanques G et al. 2009 (modifiziert)]
Gesichtsausdruck	Entspannt	1	
	Teilweise angespannt	2	
	Stark angespannt	3	
	Grimassieren	4	
Obere Extremität	Keine Bewegung	1	
	Teilweise Bewegung	2	
	Anziehen mit Bewegung der Finger	3	
	Ständiges Anziehen	4	
Verbaler Ausdruck	Keine verbale Äußerung von Schmerzen	1	
	Seltenes Jammern/Stöhnen	2	
	Häufiges Jammern/Stöhnen	3	
	Weinen, inkl. verbaler Schmerzäußerungen/Luftanhalten	4	

7.5 SAPS II (www.divi.de 2020)

SAPS II

maßgeblich sind die schlechtesten Werte (höchste Punktzahl) in einem 24-Stunden-Zeitraum seit dem zur Intensiv-Aufnahme führenden Ereignis (Unfall, OP, Komplikation)

	Punkte bei abnormal niedrigen Werten				„normal“	Punkte bei abnormal hohen Werten					
	<40	40-69	70-119	120-159	>160	<40	40-59	60-69	70-74	75-79	≥80
Alter (Jahre)						<40	40-59 7	60-69 12	70-74 15	75-79 16	≥80 18
Herzfrequenz (pro Min.)		<40 11	40-69 2	70-119	120-159 4	≥160 7					
Blutdruck (syst., mmHg)		<70 13	70-99 5	100-199	≥200 2						
Temperatur (°C)				<39.0	≥39.0 3						
nur bei Beatmung o. Pulmonalis-Katheder PaO₂ (mmHg) / FiO₂	<100 11	100-199 9	≥200 6	-							
Urinausscheidung (L/Tag)		<0.5 11	0.5-0.99 4	≥1.0							
Harnstoff (mg/dl) oder Harnstoff-Stickstoff (mg/dl)				<60	60-179 6	≥180 10					
Leukozyten (10 ³ /mm ³)			<1.0 12	1.0-19.9	≥20 3						
Kalium (mmol/L)			<3 3	3.0-4.9	≥5.0 3						
Natrium (mmol/L)			<125 5	125-144	≥145 1						
Serum Bikarbonat (mEq/L)		<15 6	15-19 3	≥20							
Bilirubin (mg/dl)				<4.0	4.0-5.9 4	≥6.0 9					
Glasgow Coma Scale (vor Sedierung)	<6 26	6-8 13	9-10 7	11-13 5	14-15						
Vorerkrankungen				-	metastas. Karzinom 9	maligne hämatolog. Erkr. 10	AIDS 17				
Zuweisung auf ICU				elektiv chirurgisch	medizinisch (ohne OP) 6	ungeplant chirurgisch 8					

Le Gall et al. „A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study.“ JAMA (1993) 270: 2957-63

Abbildung 7.5: SAPS II von <https://www.divi.de/empfehlungen/publikationen/saps-ii-simplified-acute-physiology-score/viewdocument/135>, abgerufen am 18.08.2020

7.6 APACHE II (www.intensivecarunit.de 2020)

APACHE II Score Erhebungsbogen (Deutsch)

NEGATIVSTE WERTE	ERHOEBENE VARIABLEN	ABWEICHUNGEN NACH OBEN					ABWEICHUNGEN NACH UNTEN					Punkte		
		+4	+3	+2	+1	+0	+1	+2	+3	+4				
	Temperatur – rektal (°C)	≥41°C	0	39-40,9°C	0	38,5-39°C	0	36-38,4°C	0	34-35,9°C	0	30-31,9°C	0	≤29,9°C
	Mittlerer arterieller Druck (mmHg)	≥160	0	130-159	0	110-129	0	70-109	0	50-69	0	40-54	0	≤39
	Herzfrequenz HF /min	≥180	0	140-179	0	110-139	0	70-109	0	55-69	0	40-54	0	≤39
	Atemfrequenz AF /min	≥50	0	35-49	0	25-34	0	12-24	0	10-11	0	6-9	0	≤5
	Oxygenierung P _a O ₂ /mmHg, dann wird die Alveolo-arterielle Sauerstoffdifferenz (A-aDO ₂ berücksichtigt) ¹⁾ Bei einem P _a O ₂ < 5 wird der arterielle Sauerstoffdruck (paO ₂ mmHg) berücksichtigt	≥500	0	350-499	0	200-349	0	<200	0	61-70	0	55-60	0	≤55
	pH arteriell	≥7,7	0	7,6-7,69	0	7,5-7,59	0	7,35-7,49	0	>70	0	61-70	0	≤7,15
	Natrium mmol/l	≥180	0	160-179	0	155-159	0	130-149	0	120-129	0	111-119	0	≤110
	Kalium mmol/l	≥7	0	6-6,9	0	5,5-5,9	0	3,5-5,4	0	3-3,4	0	2,5-2,9	0	≤2,5
	Kreatinin mg/100ml Bei akuten Nierenversagen Punkte verdoppeln	≥3,5	0	2-3,4	0	1,5-1,9	0	0,6-1,4	0	<0,6	0	20-29,9	0	<20
	Hämokrit %	≥60	0	50-59,9	0	46-49,9	0	30-45,9	0	20-29,9	0	1-2,9	0	<1
	Leukozyten (x1000)	≥40	0	20-39,9	0	15-19,9	0	3-14,9	0	0	0	0	0	<1
	Glasgow-Coma-Skala (GCS)	Punkte = 15 – aktueller Glasgow-Coma-Score										Punkte gesamt:		

¹⁾Se berechnet sich aus aAO₂ (mmHg) = paO₂-p(aO) oder wird vom BGA-Gerät zur Verfügung gestellt

Age Points:	≤ 44 Jahre = 0 Punkte	45-54 = 2 Punkte	55-64 = 3 Punkte	65-74 = 5 Punkte	≥ 75 = 6 Punkte
Durch Biopsie gesicherte Zirrhose und festgestellter portaler Hochdruck					
Obere gastrointestinale Blutungen in der Vorgeschichte, ausgehend von einem portalen Hochdruck					
Vorhergehende Episoden mit hepatischer Enzephalopathie/ hepatischer Enzephalopathie/ hepatischem Koma					
New York Health Association (NYHA) Class IV					
Chronische restriktive, obstruktive oder gemischte Erkrankungen, die mit einer schweren Einschränkung bei leichten Aufgaben einhergehen (z.B. Unfähigkeit Haushalt zu führen)					
Bekannte chronische Myxome, Hypertrophie, sekundäre Polyzystie, schwere pulmonale Hypertension (>40mmHg) oder Abhängigkeit von Beatmung					
chronische Dialyse					
Der Patient hatte eine Therapie, welche die Abwehrkraft gegen Infektionen schwächt (z.B. Immunsuppression, Chemotherapie, Bestrahlung, langfristige oder hochdosierte Steroide)					
Erkrankungen, die mit einer immunschwäche einhergehen (z.B. Leukämie, Lymphom, AIDS)					
Nicht-operierter Patient					
Postoperativer Patient nach Notfall-OP					
Postoperativer Patient nach Elektiv-OP					

APACHE II Score:

	+		+		=
--	---	--	---	--	---

©2004 www.intensivecarunit.de/Andie Gerdes
- Ohne Gewähr -

Abbildung 7.6: APACHE II, <http://www.intensivecareunit.de/download/apache2.pdf>, abgerufen am 18.08.2020

8. Literaturverzeichnis

- APA, 2013. Diagnostic and Statistical Manual of Mental Disorders V. Edition. American Psychiatric Association. <https://doi.org/10.1176/appi.books.9780890425596>
- APA, 2002. DSM IV TR. American Psychiatric Publishing Inc. <https://doi.org/10.1176/appi.books.9781585622658>
- Aldemir, M., Ozen, S., Kara, I.H., Sir, A., Baç, B., 2001. Predisposing factors for delirium in the surgical intensive care unit. *Crit Care* 5, 265–70.
- Balas, M.C., Happ, M.B., Yang, W., Chelluri, L., Richmond, T., 2009. Outcomes Associated With Delirium in Older Patients in Surgical ICUs.. *Chest* 135, 18–25.
- Baron, R., Binder, A., Biniek, R., Braune, S., Buerkle, H., Dall, P., Demirakca, S., Eckardt, R., Eggers, V., Eichler, I., Fietze, I., Freys, S., Fründ, A., Garten, L., Gohrbandt, B., Harth, I., Hartl, W., Heppner, H.J., Horter, J., Huth, R., Janssens, U., Jungk, C., Kaeuper, K.M., Kessler, P., Kleinschmidt, S., Kochanek, M., Kumpf, M., Meiser, A., Mueller, A., Orth, M., Putensen, C., Roth, B., Schaefer, M., Schaefer, R., Schellongowski, P., Schindler, M., Schmitt, R., Scholz, J., Schroeder, S., Schwarzmann, G., Spies, C., Stinge, R., Tonner, P., Trieschmann, U., Tryba, M., Wappler, F., Waydhas, C., Weiss, B., Weisshaar, G., 2015. Evidence and consensus based guideline for the management of delirium, analgesia, and sedation in intensive care medicine. Revision 2015 (DAS-Guideline 2015) - short version. *Ger Med Sci* 13, Doc19.
- Barr, J., Fraser, G.L., Puntillo, K., Ely, E.W., Gélinas, C., Dasta, J.F., Davidson, J.E., Devlin, J.W., Kress, J.P., Joffe, A.M., Coursin, D.B., Herr, D.L., Tung, A., Robinson, B.R.H., Fontaine, D.K., Ramsay, M.A., Riker, R.R., Sessler, C.N., Pun, B., Skrobik, Y., Jaeschke, R., 2013. Clinical Practice Guidelines for the Management of Pain Agitation, and Delirium in Adult Patients in the Intensive Care Unit. *Critical Care Medicine* 41, 263–306. <https://doi.org/10.1097/ccm.0b013e3182783b72>
- Bergeron, N., Dubois, M.J., Dumont, M., Dial, S., Skrobik, Y., 2001. Intensive Care Delirium Screening Checklist: evaluation of a new screening tool. *Intensive Care Med* 27, 859–64.
- Boßelmann, C., Zurloh, J., Stefanou, M.I., Stadler, V., Weber, Y., Lerche, H., Poli, S., Ziemann, U., Mengel, A., 2019. Delirium Screening in Aphasic Patients With the Intensive Care Delirium Screening Checklist (ICDSC): A Prospective Cohort Study. *Front Neurol* 10, 1198.
- Brummel, N.E., Boehm, L.M., Girard, T.D., Pandharipande, P.P., Jackson, J.C., Hughes, C.G., Patel, M.B., Han, J.H., Vasilevskis, E.E., Thompson, J.L., Chandrasekhar, R., Bernard, G.R., Dittus, R.S., Ely, E.W., 2017. Subsyndromal Delirium and Institutionalization Among Patients With Critical Illness. *Am J Crit Care* 26, 447–455.
- Chanques, G., Payen, J.F., Mercier, G., de, L.S., Viel, E., Jung, B., Cissé, M., Lefrant, J.Y., Jaber, S., 2009. Assessing pain in non-intubated critically ill patients unable to self report: an adaptation of the Behavioral Pain Scale. *Intensive Care Med* 35, 2060–7.
- Cole, M., McCusker, J., Dendukuri, N., Han, L., 2003. The prognostic significance of subsyndromal delirium in elderly medical inpatients. *J Am Geriatr Soc* 51, 754–60.

Davis, D.H.J., Terrera, G.M., Keage, H., Rahkonen, T., Oinas, M., Matthews, F.E., Cunningham, C., Polvikoski, T., Sulkava, R., MacLulich, A.M.J., Brayne, C., 2012. Delirium is a strong risk factor for dementia in the oldest-old: a population-based cohort study. *Brain* 135, 2809–2816. <https://doi.org/10.1093/brain/aws190>

Devlin, J.W., Fong, J.J., Fraser, G.L., Riker, R.R., 2007. Delirium assessment in the critically ill. *Intensive Care Med* 33, 929–40.

Devlin, J.W., Fong, J.J., Schumaker, G., O'Connor, H., Ruthazer, R., Garpestad, E., 2007. Use of a validated delirium assessment tool improves the ability of physicians to identify delirium in medical intensive care unit patients. *Crit Care Med* 35, 2721–4; quiz 2725.

Devlin, J.W., Marquis, F., Riker, R.R., Robbins, T., Garpestad, E., Fong, J.J., Didomenico, D., Skrobik, Y., 2008. Combined didactic and scenario-based education improves the ability of intensive care unit staff to recognize delirium at the bedside. *Critical Care* 12, R19. <https://doi.org/10.1186/cc6793>

Devlin, J.W., Skrobik, Y., Gélinas, C., Needham, D.M., Slooter, A.J.C., Pandharipande, P.P., Watson, P.L., Weinhouse, G.L., Nunnally, M.E., Rochweg, B., Balas, M.C., van den Boogaard, M., Bosma, K.J., Brummel, N.E., Chanques, G., Denehy, L., Drouot, X., Fraser, G.L., Harris, J.E., Joffe, A.M., Kho, M.E., Kress, J.P., Lanphere, J.A., McKinley, S., Neufeld, K.J., Pisani, M.A., Payen, J.-F., Pun, B.T., Puntillo, K.A., Riker, R.R., Robinson, B.R.H., Shehabi, Y., Szumita, P.M., Winkelman, C., Centofanti, J.E., Price, C., Nikayin, S., Misak, C.J., Flood, P.D., Kiedrowski, K., Alhazzani, W., 2018. Clinical Practice Guidelines for the Prevention and Management of Pain Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU. *Critical Care Medicine* 46, e825–e873. <https://doi.org/10.1097/ccm.0000000000003299>

DIMDI 2020, <https://www.dimdi.de/static/de/klassifikationen/icd/icd-10-gm/kode-suche/htmlgm2020/block-f00-f09.htm>, abgerufen am 08.02.2020.

EDA, A.D.S., 2014. The DSM-5 criteria, level of arousal and delirium diagnosis: inclusiveness is safer. *BMC Med* 12, 141.

Ely, E.W., Margolin, R., Francis, J., May, L., Truman, B., Dittus, R., Speroff, T., Gautam, S., Bernard, G.R., Inouye, S.K., 2001. Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). *Crit Care Med* 29, 1370–9.

Ely, E.W., Shintani, A., Truman, B., Speroff, T., Gordon, S.M., Harrell, F.E.J., Inouye, S.K., Bernard, G.R., Dittus, R.S., 2004. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. *JAMA* 291, 1753–62.

Ely, E.W., Siegel, M.D., Inouye, S.K., 2001. Delirium in the intensive care unit: an under-recognized syndrome of organ dysfunction. *Semin Respir Crit Care Med* 22, 115–26.

ESCIM 2016, <https://www.esicm.org/article-review-new-guidelines-on-delirium-sedation-thiessen>, abgerufen am 11.02.2020.

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12, 189–98.

- Frenette, A.J., Bebawi, E.R., Deslauriers, L.C., Tessier, A.-A.L., Perreault, M.M., Delisle, M.-S., Bertrand, J.-C., Desjardins, M., Rico, P., Khwaja, K., Burry, L.D., Bernard, F., Williamson, D.R., 2015. Validation and comparison of CAM-ICU and ICDSC in mild and moderate traumatic brain injury patients. *Intensive Care Medicine* 42, 122–123. <https://doi.org/10.1007/s00134-015-3964-1>
- Gatsonis, C.A., 1995. Random-effects models for diagnostic accuracy data. *Acad Radiol* 2 Suppl 1, S14–21; discussion S57–67, S61–4 pa.
- Gaudreau, J.D., Gagnon, P., Harel, F., Tremblay, A., Roy, M.A., 2005. Fast, systematic, and continuous delirium assessment in hospitalized patients: the nursing delirium screening scale. *J Pain Symptom Manage* 29, 368–75.
- Genders, T.S., Spronk, S., Stijnen, T., Steyerberg, E.W., Lesaffre, E., Hunink, M.G., 2012. Methods for calculating sensitivity and specificity of clustered data: a tutorial. *Radiology* 265, 910–6.
- Girard, T.D., Jackson, J.C., Pandharipande, P.P., Pun, B.T., Thompson, J.L., Shintani, A.K., Gordon, S.M., Canonico, A.E., Dittus, R.S., Bernard, G.R., Ely, E.W., 2010. Delirium as a predictor of long-term cognitive impairment in survivors of critical illness. *Crit Care Med* 38, 1513–20.
- Girard, T.D., Pandharipande, P.P., Ely, E.W., 2008. Delirium in the intensive care unit. *Crit Care* 12 Suppl 3, S3.
- Greiner, M., Pfeiffer, D., Smith, R.D., 2000. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev Vet Med* 45, 23–41.
- Guenther, U., Popp, J., Koecher, L., Muders, T., Wrigge, H., Ely, E.W., Putensen, C., 2010. Validity and reliability of the CAM-ICU Flowsheet to diagnose delirium in surgical ICU patients. *J Crit Care* 25, 144–51.
- Hakim, S.M., Othman, A.I., Naoum, D.O., 2012. Early treatment with risperidone for subsyndromal delirium after on-pump cardiac surgery in the elderly: a randomized trial. *Anesthesiology* 116, 987–97.
- Inouye, S.K., 1999. Predisposing and precipitating factors for delirium in hospitalized older patients. *Dement Geriatr Cogn Disord* 10, 393–400.
- Inouye, S.K., 2006. Delirium in older persons. *N Engl J Med* 354, 1157–65.
- Inouye, S.K., 1990. Clarifying Confusion: The Confusion Assessment Method. *Annals of Internal Medicine* 113, 941. <https://doi.org/10.7326/0003-4819-113-12-941>
- Inouye, S.K., Bogardus, S.T.J., Charpentier, P.A., Leo-Summers, L., Acampora, D., Holford, T.R., Cooney, L.M.J., 1999. A multicomponent intervention to prevent delirium in hospitalized older patients. *N Engl J Med* 340, 669–76.
- Inouye, S.K., Kosar, C.M., Tommet, D., Schmitt, E.M., Puella, M.R., Saczynski, J.S., Marcantonio, E.R., Jones, R.N., 2014. The CAM-S: development and validation of a new scoring system for delirium severity in 2 cohorts. *Ann Intern Med* 160, 526–533.

- Khan, B.A., Perkins, A.J., Gao, S., Hui, S.L., Campbell, N.L., Farber, M.O., Chlan, L.L., Boustani, M.A., 2017. The Confusion Assessment Method for the ICU-7 Delirium Severity Scale: A Novel Delirium Severity Instrument for Use in the ICU. *Crit Care Med* 45, 851–857.
- Knaus, W.A., Draper, E.A., Wagner, D.P., Zimmerman, J.E., 1985. APACHE II: a severity of disease classification system. *Crit Care Med* 13, 818–29.
- Kotfis, K., Bott-Olejnik, M., Szylińska, A., Listewnik, M., Rotter, I., 2019. Characteristics, Risk Factors And Outcome Of Early-Onset Delirium In Elderly Patients With First Ever Acute Ischemic Stroke - A Prospective Observational Cohort Study. *Clin Interv Aging* 14, 1771–1782.
- Krewulak, K.D., Stelfox, H.T., Leigh, J.P., Ely, E.W., Fiest, K.M., 2018. Incidence and Prevalence of Delirium Subtypes in an Adult ICU: A Systematic Review and Meta-Analysis. *Crit Care Med* 46, 2029–2035.
- Larsen, L.K., Frøkjær, V.G., Nielsen, J.S., Skrobik, Y., Winkler, Y., Møller, K., Petersen, M., Egerod, I., 2019. Delirium assessment in neuro-critically ill patients: A validation study. *Acta Anaesthesiol Scand* 63, 352–359.
- Le Gall, J.-R., Lemeshow, S., Saulnier, F., 1993. A New Simplified Acute Physiology Score (SAPS II) Based on a European/North American Multicenter Study. *JAMA: The Journal of the American Medical Association* 270, 2957. <https://doi.org/10.1001/jama.1993.03510240069035>
- Leslie, D.L., Marcantonio, E.R., Zhang, Y., Leo-Summers, L., Inouye, S.K., 2008. One-Year Health Care Costs Associated With Delirium in the Elderly Population. *Archives of Internal Medicine* 168, 27. <https://doi.org/10.1001/archinternmed.2007.4>
- Levkoff, S.E., Liptzin, B., Cleary, P.D., Wetle, T., Evans, D.A., Rowe, J.W., Lipsitz, L.A., 1996. Subsyndromal Delirium. *Am J Geriatr Psychiatry* 4, 320–329.
- Lin, S.M., Liu, C.Y., Wang, C.H., Lin, H.C., Huang, C.D., Huang, P.Y., Fang, Y.F., Shieh, M.H., Kuo, H.P., 2004. The impact of delirium on the survival of mechanically ventilated patients. *Crit Care Med* 32, 2254–9.
- Luetz, A., Gensel, D., Müller, J., Weiss, B., Martiny, V., Heinz, A., Wernecke, K.D., Spies, C., 2016. Validity of Different Delirium Assessment Tools for Critically Ill Children: Covariates Matter. *Crit Care Med* 44, 2060–2069.
- Luetz, A., Heymann, A., Radtke, F.M., Chenitir, C., Neuhaus, U., Nachtigall, I., von, D.V., Marz, S., Eggers, V., Heinz, A., Wernecke, K.D., Spies, C.D., 2010. Different assessment tools for intensive care unit delirium: which score to use? *Crit Care Med* 38, 409–18.
- Luetz, A., Heymann, A., Radtke, F.M., Spies, C.D., 2010. [If delirium is not monitored it will often be not detected]. *Anesthesiol Intensivmed Notfallmed Schmerzther* 45, 106–11.
- Lundström, M., Edlund, A., Karlsson, S., Brännström, B., Bucht, G., Gustafson, Y., 2005. A multi-factorial intervention program reduces the duration of delirium, length of hospitalization, and mortality in delirious patients. *J Am Geriatr Soc* 53, 622–8.
- Martin, J., Heymann, A., Bäsell, K., Baron, R., Biniek, R., Bürkle, H., Dall, P., Dictus, C., Eggers, V., Eichler, I., Engelmann, L., Garten, L., Hartl, W., Haase, U., Huth, R., Kessler, P., Kleinschmidt,

S., Koppert, W., Kretz, F.J., Laubenthal, H., Marggraf, G., Meiser, A., Neugebauer, E., Neuhaus, U., Putensen, C., Quintel, M., Reske, A., Roth, B., Scholz, J., Schröder, S., Schreiter, D., Schüttler, J., Schwarzmann, G., Stingele, R., Tonner, P., Tränkle, P., Treede, R.D., Trupkovic, T., Tryba, M., Wappler, F., Waydhas, C., Spies, C., 2010. Evidence and consensus-based German guidelines for the management of analgesia, sedation and delirium in intensive care—short version. *Ger Med Sci* 8, Doc02.

McGuire, B.E., Basten, C.J., Ryan, C.J., Gallagher, J., 2000. Intensive care unit syndrome: a dangerous misnomer. *Arch Intern Med* 160, 906–9.

Meagher, D.J., Morandi, A., Inouye, S.K., Ely, W., Adamis, D., MacLulich, A.J., Rudolph, J.L., Neufeld, K., Leonard, M., Bellelli, G., Davis, D., Teodorczuk, A., Kreisel, S., Thomas, C., Hasemann, W., Timmons, S., O'Regan, N., Grover, S., Jabbar, F., Cullen, W., Dunne, C., Kamholz, B., Van, M.B.C., De, R.S.E., De, J.J., Trzepacz, P.T., 2014. Concordance between DSM-IV and DSM-5 criteria for delirium diagnosis in a pooled database of 768 prospectively evaluated patients using the delirium rating scale-revised-98. *BMC Med* 12, 164.

Mitasova, A., Kostalova, M., Bednarik, J., Michalcakova, R., Kasperek, T., Balabanova, P., Dusek, L., Vohanka, S., Ely, E.W., 2012. Poststroke delirium incidence and outcomes. *Critical Care Medicine* 40, 484–490. <https://doi.org/10.1097/ccm.0b013e318232da12>

Morandi, A., Davis, D., Taylor, J.K., Bellelli, G., Olofsson, B., Kreisel, S., Teodorczuk, A., Kamholz, B., Hasemann, W., Young, J., Agar, M., de, R.S.E., Meagher, D., Trabucchi, M., MacLulich, A.M., 2013. Consensus and variations in opinions on delirium care: a survey of European delirium specialists. *Int Psychogeriatr* 25, 2067–75.

Morandi, A., Pandharipande, P., Trabucchi, M., Rozzini, R., Mistraletti, G., Trompeo, A.C., Gregoretto, C., Gattinoni, L., Ranieri, M.V., Brochard, L., Annane, D., Putensen, C., Guenther, U., Fuentes, P., Tobar, E., Anzueto, A.R., Esteban, A., Skrobik, Y., Salluh, J.I., Soares, M., Granja, C., Stubhaug, A., de, R.S.E., Ely, E.W., 2008. Understanding international differences in terminology for delirium and other types of acute brain dysfunction in critically ill patients. *Intensive Care Med* 34, 1907–15.

Nacul, F.E., Paul, N., Spies, C.D., Sechting, H., Hecht, T., Dullinger, J.S., Piper, S.K., Luetz, A., Balzer, F.S., Wernecke, K.D., Sa, A.K., Barros, F.da C.C., Eymold, L., Chenitir, C., Weiss, B., 2020. Influence of Sedation Level and Ventilation Status on the Diagnostic Validity of Delirium Screening Tools in the ICU—An International, Prospective, Bi-Center Observational Study (IDeAS). *Medicina (Kaunas)* 56.

Neufeld, K.J., Hayat, M.J., Coughlin, J.M., Huberman, A.L., Leistikow, N.A., Krumm, S.K., Needham, D.M., 2011. Evaluation of Two Intensive Care Delirium Screening Tools for Non-Critically Ill Hospitalized Patients. *Psychosomatics* 52, 133–140. <https://doi.org/10.1016/j.psych.2010.12.018>

NICE 2020, <https://www.nice.org.uk/guidance/cg103>, abgerufen am 11.02.2020.

O'Keeffe, S.T., Lavan, J.N., 1999. Clinical significance of delirium subtypes in older people. *Age Ageing* 28, 115–9.

- Otter, H., Martin, J., Bäsell, K., von, H.C., Hein, O.V., Böllert, P., Jänsch, P., Behnisch, I., Wernecke, K.D., Konertz, W., Loening, S., Blohmer, J.U., Spies, C., 2005. Validity and reliability of the DDS for severity of delirium in the ICU. *Neurocrit Care* 2, 150–8.
- Ouimet, S., Kavanagh, B.P., Gottfried, S.B., Skrobik, Y., 2007. Incidence, risk factors and consequences of ICU delirium. *Intensive Care Med* 33, 66–73.
- Ouimet, S., Riker, R., Bergeon, N., Cossette, M., Kavanagh, B., Skrobik, Y., 2007. Subsyndromal delirium in the ICU: evidence for a disease spectrum. *Intensive Care Medicine* 33, 1007–1013. <https://doi.org/10.1007/s00134-007-0618-y>
- Pandharipande, P., Cotton, B.A., Shintani, A., Thompson, J., Costabile, S., Truman, P.B., Dittus, R., Ely, E.W., 2007. Motoric subtypes of delirium in mechanically ventilated surgical and trauma intensive care unit patients. *Intensive Care Med* 33, 1726–31.
- Pandharipande, P., Jackson, J., Ely, E.W., 2005. Delirium: acute cognitive dysfunction in the critically ill. *Curr Opin Crit Care* 11, 360–8.
- Pandharipande, P.P., Girard, T.D., Jackson, J.C., Morandi, A., Thompson, J.L., Pun, B.T., Brummel, N.E., Hughes, C.G., Vasilevskis, E.E., Shintani, A.K., Moons, K.G., Geervarghese, S.K., Canonico, A., Hopkins, R.O., Bernard, G.R., Dittus, R.S., Ely, E.W., 2013. Long-term cognitive impairment after critical illness. *N Engl J Med* 369, 1306–16.
- Patel, S.B., Kress, J.P., 2011. Accurate identification of delirium in the ICU: problems with translating the evidence in the real-life setting. *Am J Respir Crit Care Med* 184, 287–8.
- Payen, J.F., Bru, O., Bosson, J.L., Lagrasta, A., Novel, E., Deschaux, I., Lavagne, P., Jacquot, C., 2001. Assessing pain in critically ill sedated patients by using a behavioral pain scale. *Crit Care Med* 29, 2258–63.
- Peterson, J.F., Pun, B.T., Dittus, R.S., Thomason, J.W., Jackson, J.C., Shintani, A.K., Ely, E.W., 2006. Delirium and its motoric subtypes: a study of 614 critically ill patients. *J Am Geriatr Soc* 54, 479–84.
- Pisani, M.A., Kong, S.Y., Kasl, S.V., Murphy, T.E., Araujo, K.L., Van, N.P.H., 2009. Days of delirium are associated with 1-year mortality in an older intensive care unit population. *Am J Respir Crit Care Med* 180, 1092–7.
- Pun, B.T., Gordon, S.M., Peterson, J.F., Shintani, A.K., Jackson, J.C., Foss, J., Harding, S.D., Bernard, G.R., Dittus, R.S., Ely, E.W., 2005. Large-scale implementation of sedation and delirium monitoring in the intensive care unit: A report from two medical centers. *Critical Care Medicine* 33, 1199–1205. <https://doi.org/10.1097/01.ccm.0000166867.78320.ac>
- Radtke, F.M., Franck, M., Oppermann, S., Lütz, A., Seeling, M., Heymann, A., Kleinwächter, R., Kork, F., Skrobik, Y., Spies, C.D., 2009. [The Intensive Care Delirium Screening Checklist (ICDSC)–translation and validation of intensive care delirium checklist in accordance with guidelines]. *Anesthesiol Intensivmed Notfallmed Schmerzther* 44, 80–6.
- Radtke, F.M., Heymann, A., Franck, M., Maechler, F., Drews, T., Luetz, A., Nachtigall, I., Wernecke, K.D., Spies, C.D., 2012. How to implement monitoring tools for sedation, pain and delirium in the intensive care unit: an experimental cohort study. *Intensive Care Med* 38, 1974–81.

- Reznik, M.E., Drake, J., Margolis, S.A., Moody, S., Murray, K., Costa, S., Mahta, A., Wendell, L.C., Thompson, B.B., Rao, S.S., Barrett, A.M., Boukrina, O., Daiello, L.A., Asaad, W.F., Furie, K.L., Jones, R.N., 2020. Deconstructing Poststroke Delirium in a Prospective Cohort of Patients With Intracerebral Hemorrhage. *Crit Care Med* 48, 111–118.
- Robinson, T.N., Raeburn, C.D., Tran, Z.V., Brenner, L.A., Moss, M., 2011. Motor subtypes of postoperative delirium in older adults. *Arch Surg* 146, 295–300.
- Schweickert, W.D., Pohlman, M.C., Pohlman, A.S., Nigos, C., Pawlik, A.J., Esbrook, C.L., Spears, L., Miller, M., Franczyk, M., Deprizio, D., Schmidt, G.A., Bowman, A., Barr, R., McCallister, K.E., Hall, J.B., Kress, J.P., 2009. Early physical and occupational therapy in mechanically ventilated, critically ill patients: a randomised controlled trial. *Lancet* 373, 1874–82.
- Serafim, R.B., Soares, M., Bozza, F.A., Lapa, E.S.J.R., Dal-Pizzol, F., Paulino, M.C., Povoia, P., Saluh, J.I.F., 2017. Outcomes of subsyndromal delirium in ICU: a systematic review and meta-analysis. *Crit Care* 21, 179.
- Sessler, C.N., Gosnell, M.S., Grap, M.J., Brophy, G.M., O'Neal, P.V., Keane, K.A., Tesoro, E.P., Elswick, R.K., 2002. The Richmond Agitation-Sedation Scale. *American Journal of Respiratory and Critical Care Medicine* 166, 1338–1344. <https://doi.org/10.1164/rccm.2107138>
- Tao, R., Wang, X.W., Pang, L.J., Cheng, J., Wang, Y.M., Gao, G.Q., Liu, Y., Wang, C., 2018. Pharmacologic prevention of postoperative delirium after on-pump cardiac surgery: A meta-analysis of randomized trials. *Medicine (Baltimore)* 97, e12771.
- Teasdale, G., Jennett, B., 1974. Assessment of coma and impaired consciousness. A practical scale. *Lancet* 2, 81–4.
- Trzepacz, P.T., Mittal, D., Torres, R., Canary, K., Norton, J., Jimerson, N., 2001. Validation of the Delirium Rating Scale-revised-98: comparison with the delirium rating scale and the cognitive test for delirium. *J Neuropsychiatry Clin Neurosci* 13, 229–42.
- Van Rompaey, B., Elseviers, M.M., Schuurmans, M.J., Shortridge-Baggett, L.M., Truijen, S., Bossaert, L., 2009. Risk factors for delirium in intensive care patients: a prospective cohort study. *Crit Care* 13, R77.
- van Eijk, M.M., van Marum, R.J., Klijn, I.A., de, W.N., Kesecioglu, J., Slooter, A.J., 2009. Comparison of delirium assessment tools in a mixed intensive care unit. *Crit Care Med* 37, 1881–5.
- van Eijk, M.M., van, den B.M., van, M.R.J., Benner, P., Eikelenboom, P., Honing, M.L., van, der H.B., Horn, J., Izaks, G.J., Kalf, A., Karakus, A., Klijn, I.A., Kuiper, M.A., de, L.F.E., de, M.T., van, der M.R.C., Osse, R.J., de, R.S.E., Spronk, P.E., van, der V.P.H., van, G.W.A., Slooter, A.J., 2011. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. *Am J Respir Crit Care Med* 184, 340–4.
- Vasilevskis, E.E., Chandrasekhar, R., Holtze, C.H., Graves, J., Speroff, T., Girard, T.D., Patel, M.B., Hughes, C.G., Cao, A., Pandharipande, P.P., Ely, E.W., 2018. The Cost of ICU Delirium and Coma in the Intensive Care Unit Patient. *Medical Care* 56, 890–897. <https://doi.org/10.1097/mlr.0000000000000975>

Vasilevskis, E.E., Girard, T.D., Ely, E.W., 2012. The bedside diagnosis of ICU delirium: specificity is high, let's optimize sensitivity. *Am J Respir Crit Care Med* 185, 107–8; author reply 108.

Verbeke, G., 1997. Linear mixed models for longitudinal data. In: *Linear mixed models in practice*. Springer.

von Hofen-Hohloch, J., Awissus, C., Fischer, M.M., Michalski, D., Rumpf, J.J., Classen, J., 2020. Delirium Screening in Neurocritical Care and Stroke Unit Patients: A Pilot Study on the Influence of Neurological Deficits on CAM-ICU and ICDSC Outcome. *Neurocrit Care*.

Wang, J., Ji, Y., Wang, N., Chen, W., Bao, Y., Qin, Q., Xiao, Q., Li, S., 2018. Risk factors for the incidence of delirium in cerebrovascular patients in a Neurosurgery Intensive Care Unit: A prospective study. *J Clin Nurs* 27, 407–415.

Weiss, B.; Paul, N.; Spies, C.; Ullrich, D.; Ansorge, I.; Salih, F.; Wolf S.; Luetz, A., 2021. Influence of patient-specific covariates on test validity of two delirium screening instruments in neurocritical care patients (DEMON-ICU). Manuscript submitted for publication.

Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., Erikson, P., 2005. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in Health* 8, 94–104. <https://doi.org/10.1111/j.1524-4733.2005.04054.x>

www.divi.de, <https://www.divi.de/empfehlungen/publikationen/saps-ii-simplified-acute-physiology-score/viewdocument/135>, abgerufen am 18.08.2020.

www.icudelirium.org, <https://www.icudelirium.org/medical-professionals/downloads/resource-language-translations>, abgerufen am 22.07.2020.

www.intensivcareunit.de, <http://www.intensivcareunit.de/download/apache2.pdf>, abgerufen am 18.08.2020.

Youden, W.J., 1950. Index for rating diagnostic tests. *Cancer* 3, 32–5.

Zipser, C.M., Deuel, J., Ernst, J., Schubert, M., von, K.R., Böttger, S., 2019. The predisposing and precipitating risk factors for delirium in neurosurgery: a prospective cohort study of 949 patients. *Acta Neurochir (Wien)* 161, 1307–1315.

Zipser, C.M., Deuel, J., Ernst, J., Schubert, M., Weller, M., von, K.R., Boettger, S., 2019. Predisposing and precipitating factors for delirium in neurology: a prospective cohort study of 1487 patients. *J Neurol* 266, 3065–3075.

9. Eidesstattliche Versicherung

„Ich, Dennis Ullrich, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: ‚Validierung der Schweregradskala für die Confusion Assessment Method für Intensivstationen bei neurologischen und neurochirurgischen Intensivpatienten‘ (‚Validation of the severity scale for the Confusion Assessment Method for the Intensive Care Unit for neurological and neurosurgical patients‘) selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren/innen beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) werden von mir verantwortet.

Ich versichere ferner, dass ich die in Zusammenarbeit mit anderen Personen generierten Daten, Datenauswertungen und Schlussfolgerungen korrekt gekennzeichnet und meinen eigenen Beitrag sowie die Beiträge anderer Personen korrekt kenntlich gemacht habe (siehe Anteilserklärung). Texte oder Textteile, die gemeinsam mit anderen erstellt oder verwendet wurden, habe ich korrekt kenntlich gemacht.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Erstbetreuer/in, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; www.icmje.org) zur Autorenschaft eingehalten. Ich erkläre ferner, dass ich mich zur Einhaltung der Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis verpflichte.

Weiterhin versichere ich, dass ich diese Dissertation weder in gleicher noch in ähnlicher Form bereits an einer anderen Fakultät eingereicht habe.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§§ 156, 161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Berlin, den 30. März 2021

Unterschrift

10. Anteilserklärung

Dennis Ullrich hatte folgenden Anteil an den folgenden Publikationen:

Publikation 1:

Weiss, B.; Paul, N.; Spies, C.; Ullrich, D.; Ansorge, I.; Salih, F.; Wolf S.; Luetz, A., 2021. Influence of patient-specific covariates on test validity of two delirium screening instruments in neurocritical care patients (DEMON-ICU). *Eingereicht am 26.03.2021 bei Neurocritical Care*

Beitrag im Einzelnen (bitte detailliert ausführen):

- Koautor an der o.a. Veröffentlichung
- Mitwirkung an der Planung der Datenerhebung in Zusammenarbeit mit Prof. Dr. med. Alawi Lütz
- Maßgebliche Erhebung und Dokumentation der Daten
- Mitwirkung bei der statistischen Auswertung der erhobenen Daten
- Wesentliche Beiträge zur kritischen Würdigung der Resultate mit Identifikation der relevanten Aussagen der Studie einschließlich ihrer Limitationen
- Mitwirkung an dem zur Publikation führenden Manuskript gemeinsam mit Prof. Dr. med. Alawi Lütz

Unterschrift, Datum und Stempel des/der erstbetreuenden Hochschullehrers/in

Unterschrift des Doktoranden/der Doktorandin

11. Lebenslauf

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

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

12. Publikationen im Rahmen der Dissertation

Weiss, B.; Paul, N.; Spies, C.; Ullrich, D.; Ansorge, I.; Salih, F.; Wolf S.; Luetz, A., 2021. Influence of patient-specific covariates on test validity of two delirium screening instruments in neurocritical care patients (DEMON-ICU). Manuscript submitted for publication.

Eingereicht am 26. März 2021, Neurocritical Care, LCCN: 2002215176, ISSN: 1541-6933,

13. Danksagung

Prof. Dr. Alawi Lütz gilt mein ausdrücklicher Dank. Er hat die Untersuchung, in deren Rahmen diese Arbeit entstanden ist, maßgeblich geplant, die Doktoranden über die gesamte Laufzeit des Projekts betreut und unermüdlich konstruktive Kritik in jeder Phase der Arbeit geübt. Seine Expertise und kritischen Hinweise waren für mich bei der Erstellung der Arbeit von unschätzbarem Wert.

Mein Dank gilt Univ.-Prof. Dr. Claudia Spies für die Überlassung des Themas, der Möglichkeit, die untersuchte Fragestellung an ihrer Klinik zu bearbeiten, ihre Förderung und Ermutigung sowie Ihre konstruktive Kritik.

Für ihre Unterstützung im Rahmen der statistischen Auswertung, der kritischen Diskussion der vorliegenden Arbeit und dem gemeinsamen Erstellen des Manuskripts der Veröffentlichung danke ich Dr. Björn Weiß und Nicolas Paul. Für die Methodenkritik im Bereich der Statistik danke ich Dr. rer. nat. Sophie Piper.

Meinen Kollegen als Doktoranden Peter Braun und Weronika Lisiuk danke ich für die konstruktive Zusammenarbeit im Rahmen der Datenerhebung.

Ferner gilt mein Dank allen ärztlichen und pflegerischen Kollegen der Neurointensivstation 1i am Campus Virchow der Charité Universitätsmedizin Berlin, die durch ihre Unterstützung die Datenerhebung erst ermöglicht haben.

Für ihre unermüdliche Unterstützung während meiner gesamten bisherigen Ausbildung und ihren Glauben an mich danke ich meiner Familie und meinen Freunden.

Schließlich danke ich allen Patienten, die ich im Rahmen dieser Arbeit trotz ihrer schweren Erkrankung untersuchen durfte.