

Institut für Medizinische Psychologie und Verhaltensneurobiologie
Universität Tübingen

Funktionelle Bildgebungsverfahren in der Differenzierung
zwischen dem Vegetativen Zustand und dem Minimalen
Bewusstseinszustand:
Eine systematische Analyse und Metaanalyse

Inaugural Dissertation
zur Erlangung des Doktorgrades
der Zahnheilkunde

Der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von

Spiegelsberger, Benedikt Georg

2020

Dekan: Professor Dr. B. Pichler

1. Berichterstatter: Professor Dr. B. Kotchoubey

2. Berichterstatter: Professor Dr. M. Staudt

Tag der Disputation: 18.06.2020

Meiner Schwester

I. Inhaltsverzeichnis

I.	Inhaltsverzeichnis	IV
II.	Abbildungsverzeichnis	VII
III.	Tabellenverzeichnis	VIII
IV.	Abkürzungsverzeichnis	IX
1	Einleitung.....	1
1.1	Allgemeines zu Disorders of Consciousness	1
1.2	Unterscheidung Koma, Vegetative State und Minimally Conscious State ..	2
1.3	Problematik der Diagnosestellung	5
1.3.1	Allgemeines	5
1.3.2	Klinische Diagnostik nach der CRS	6
1.3.3	Erweiterte Diagnostik durch Bildgebung und elektrophysiologische Methoden	8
1.4	Fragestellung	16
2	Material und Methoden	17
2.1	Allgemeines zu Systematic Review und Metaanalyse	17
2.2	Reviewfrage	19
2.3	Literaturrecherche	20
2.3.1	Suchkriterien	20
2.3.2	Artikelauswahl	21
2.4	Analyse der Studien und Datenextraktion	28
2.4.1	Qualitätsbeurteilung	28
2.4.2	Kalkulation der Effektstärken	34

2.4.3	Korrelation und Metaanalyse.....	37
2.4.3.1	Korrelation.....	37
2.4.3.2	Metaanalyse.....	41
3	Ergebnisse.....	46
3.1	Ergebnisse der Qualitätsbeurteilung und der Korrelation.....	46
3.1.1	Patientenanzahl.....	50
3.1.2	Journal Impact Factor.....	51
3.1.3	Erscheinungsjahr.....	53
3.1.4	Mögliche Verzerrungseffekte.....	55
3.2	Ergebnisse des systematischen Reviews.....	60
3.3	Ergebnisse der Metaanalyse.....	63
3.3.1	Aktive Aufgaben.....	63
3.3.2	Passive Stimulationen.....	65
3.3.3	Resting-State-Methode.....	66
4	Diskussion.....	67
4.1	Limitationen.....	68
4.2	Metaanalyse.....	73
4.2.1	Aktive Aufgaben.....	73
4.2.2	Passive Stimulation.....	74
4.2.3	Resting-State-Methode.....	75
4.3	Schlussfolgerung.....	76
5	Zusammenfassung.....	78
6	Literaturverzeichnis.....	80
7	Anhang.....	87
8	Erklärung zum Eigenanteil der Dissertationsschrift.....	95

9	Danksagung	96
----------	-------------------------	-----------

II. Abbildungsverzeichnis

Abbildung 1: JFK Coma Recovery Scale-Revised (Zasler et al., 2013)	7
Abbildung 2: Gegenüberstellung FDG-PET gesunder Proband (linke Seite) versus VS-Patient (rechte Seite) (Di Perri et al., 2014).....	11
Abbildung 3: Beispielhafte Darstellung aktiver Hirnareale im fMRT (Bildgebung) .	12
Abbildung 4: Typische Darstellung einer EEG-Aufzeichnung(Hulick, 2016).....	14
Abbildung 5: Flow Chart zur Artikelauswahl	26
Abbildung 6: Varianten der möglichen Korrelationen (Schwarz, 2016)	38
Abbildung 7: Korrelation der Patientenzahl mit dem Standardfehler der Effektstärke	50
Abbildung 8: Korrelation Impact Faktor mit Patientenzahl.....	52
Abbildung 9: Korrelation Impact Faktor mit dem Standardfehler der Effektstärke .	52
Abbildung 10: Korrelation Erscheinungsjahr mit Patientenzahl	54
Abbildung 11: Korrelation Erscheinungsjahr mit Effektstärke	54
Abbildung 12: Korrelation Patientenzahl mit Effektstärke.....	55
Abbildung 13: Korrelation Standardfehler mit Effektstärke	56
Abbildung 14: Korrelation Bias mit Impact Faktor.....	58
Abbildung 15: Korrelation Bias mit Patientenzahl.....	59
Abbildung 16 Korrelation Bias mit Erscheinungsjahr	59
Abbildung 17: Forest-Plot der Aktive-State-Untersuchungsmethoden	64
Abbildung 18: Forest-Plot der Passive-State-Untersuchungsmethoden.....	65
Abbildung 19: Forest-Plot der Resting-State-Untersuchungsmethoden	66

III. Tabellenverzeichnis

Tabelle 1: Suchalgorithmen.....	22
Tabelle 2: Studienanzahl zu den entsprechenden Suchkriterien.....	24
Tabelle 3: Interpretation der Effektstärke d nach Cohen (Cohen, 1988)	36
Tabelle 4: Daten der relevanten 25 Artikel	47
Tabelle 5: Effektstärken und Standardfehler der jeweiligen Datensätze	49
Tabelle 6: Darstellung der Analyseergebnisse des Bias nach dem QUADAS-2- Prinzip	57
Tabelle 7: Übersicht aller verwendeten Studien	94

IV. Abkürzungsverzeichnis

BCI	Brain Computer Interface
BOLD-Signal	blood oxygenated level dependent signal
CRS -R	Coma Recovery Scale-Revised
DoC	Disorders of Consciousness
EEG	Electroencephalografie
EMCS	emerged from MCS
ERP/EP	Event Related Potential
FDG	Fluordesoxyglucose
fMRT (fMRI)	funktionelle Magnetresonanztomografie (engl. functional magnetic resonance imaging)
HDI	Hub Disruption index
ITCI	Tresholded Connectome Intactness
JIF	Journal Impact Factor
MCS	Minimally Conscious State
MRT	Magnetresonanztomografie
PET	Positronenemissionstomografie
PVS	Persistent Vegetative State
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
rs-fMRI	resting state- functional magnetic resonance imaging
sMRI	structural magnetic resonance imaging
UWS	Unresponsive Wakefulness Syndrom
VS	Vegetative State

1 Einleitung

1.1 Allgemeines zu Disorders of Consciousness

Die Fortschritte der modernen Medizin und insbesondere die der Intensivmedizin haben die Überlebenschancen von Patienten mit schweren Hirnschädigungen deutlich erhöht. Auslöser solcher drastischer Schädigungen können vielfältig sein. Hierzu können beispielhaft Traumata, Schlaganfälle, Herzinfarkte, bakterielle und virale Infektionen oder Stoffwechselstörungen zählen. Entweder direkt durch die traumatische oder entzündliche Schädigung oder als Folge der Sauerstoffmangelversorgung des Gehirnparenchyms kommt es zum partiellen oder totalen Funktionsverlust des Gehirns. Wird eine solche initiale Schädigung überlebt, befindet sich der Patient oftmals in einem Zustand, der als Koma bezeichnet wird. Allerdings kann es vorkommen, dass einige der Patienten durchaus ein vollständig wiederhergestelltes Bewusstsein oder sogar eine komplette Rehabilitation erlangen. Der Patientenanteil, dem dies nicht gelingt, verbleibt in verschiedenen Stadien einer chronischen Bewusstseinsstörung, den sogenannten Disorders of Consciousness (DoC). Unter diesem Begriff versteht man eine schwere Einschränkung oder auch den kompletten Verlust des Bewusstseins und somit der Fähigkeit, mit der Umwelt in Kontakt zu treten. Dieses klinische Erscheinungsbild wurde zum ersten Mal 1940 von W. Kretschmer beschrieben, der es als apallisch bezeichnete, da er - vom heutigen Standpunkt aus irrtümlicherweise - annahm, dass dabei alle kortikalen Funktionen vollständig ausfallen (a-pallisch = ohne Hirnrinde). Mit der Einführung der modernen Intensiv- und Unfallmedizin ab den 1950er - Jahren traten solche DoC vermehrt auf, da Patienten mit einer wie oben beschriebenen Gehirnschädigung dank dem Einsatz fortschrittlicher Techniken und Methoden deutlich höhere Überlebenschancen zeigten (Laureys, 2005). Ab den 1970ern kam hierzu erstmals die Möglichkeit, durch die Entwicklung erster bildgebender Verfahren und deren rasche Weiterentwicklung im Vergleich zur reinen

klinischen Untersuchung deutlich tiefere und aufschlussreichere Einblicke in die Funktion solch geschädigter Gehirne zu erhalten (Gantner et al., 2012).

1.2 Unterscheidung Koma, Vegetative State und Minimally Conscious State

In dem anfänglichen komatösen Zustand unmittelbar nach einer entsprechenden Schädigung ist keinerlei Reaktion, wie eine Augenöffnung oder Ähnliches, auf starke äußerer Reize, wie zum Beispiel wiederholte Schmerzstimuli, zu erwarten (Hannawi et al., 2015). Auch Medikamente wie Analgetika oder sedierende Arzneimittel, die das Bewusstsein beeinflussen, zeigen in dieser Phase des Komats keine Wirkung (Bender et al., 2015). So lässt sich in diesem Zustand keinerlei Kontakt der Patienten mit ihrer Umgebung erkennen. Normalerweise dauert diese initiale Phase nicht länger als 4 Wochen an. Danach erleiden die Patienten entweder einen Hirntod, gewinnen ihr volles Bewusstsein wieder zurück oder gehen in eine DoC über (Di Perri et al., 2014).

Der wichtigste und am häufigsten zu erwartende DoC ist als Vegetativer Zustand (Vegetative State, VS) bekannt. Synonym hierfür können Begriffe wie Wachkoma, apallisches Syndrom oder seit 2010 auch Unresponsive Wakefulness Syndrom (UWS) verwendet werden (Laureys et al., 2010). Im Folgenden wird der Begriff VS verwendet, der in den Studien zur funktionellen Bildgebung am häufigsten auftritt.

Im VS ist der Patient im Gegensatz zum Koma in der Lage seine Augen zu öffnen und ist durchaus im Stande auf einfache Reize wie zum Beispiel Schmerzreize zu reagieren. Ansonsten zeigt der Patient keine ersichtlichen Anzeichen von Bewusstsein (Bender et al., 2015). Die Funktionen von Zwischenhirn, Hirnstamm und Rückenmark bleiben nahezu erhalten, wodurch im Gegensatz zum Koma ein Zustand von Wachheit ohne Bewusstsein resultiert. Durch den weitgehenden Erhalt der Hirnstammfunktionen ist der Schlaf-Wach-Rhythmus annähernd regulär oder kaum beeinträchtigt (Andrews, 1997). Es wurde vermutet, dass es sich hierbei nur um den rein verhaltensmäßig ersichtlichen Schlaf handelt, sprich die Augen sind

geschlossen und die Muskulatur ist entspannt (Landsness et al., 2011). So zeigte eine italienische Studie von 2014, dass elektrophysiologische Schlafmuster, wie sie ein gesunder Erwachsener zeigt, nur bei 7,5 % der VS-Patienten zu erkennen sind (de Biase et al., 2014). Allerdings werden bestimmte Schlafphänomene, wie zum Beispiel Schlafspindeln, K-Komplexe oder REM-Schlaf, bei sehr vielen VS Patienten gefunden (Bedini et al., 2016) (Rossi Sebastiano et al., 2015) (Kang et al., 2014), wenngleich nur bei wenigen Patienten diese Phänomene das gesamte gesunde Schlafmuster ergeben.

Sehr gut beschreiben die durch die Multi-Society-Task-Force on PVS 1994 festgelegten diagnostischen Kriterien, den Vegetativen Status anhand der folgenden Befunde (1994):

- vollständiger Verlust des Bewusstseins bei weitgehend erhaltenem Hirnstamm
- spinale, hypothalamische und autonome Reflexe
- fehlende Fähigkeit, willkürliche und reproduzierbare Verhaltensänderungen infolge externer visueller, auditorischer, taktiler oder schmerzhafter Reize durchzuführen
- verlust der Sprache und des Sprachverständnisses
- Blasen- und Darmkontinenz

Befindet sich ein Patient länger als ein bis drei Monate in dem VS, so wurde dieser ehemals als Persistent Vegetative State (PVS) bezeichnet (von Wild et al., 2012). Für solche Patienten gab die Multi-Society-Task-Force on PVS von 1994 eine Lebenserwartung ab Krankheitseintritt von zwei bis fünf Jahren an. Dabei kann eine Überlebenschance von mehr als zehn Jahren als äußerst unwahrscheinlich angesehen werden. Inzwischen mehren sich aufgrund klinischer Erfahrungen allerdings Zweifel an diesen Zeitangaben und der Ausdruck PVS (in welchem der Buchstabe „P“ irrtümlicherweise als „permanent“ statt „persistent“ verstanden wird) sollte im derzeitigen Sprachgebrauch vermieden werden.

Bei einigen VS-Patienten können nach einer gewissen Zeit Anzeichen von wie auch immer gearteten reproduzierbaren Reaktionen auf die Umgebung festgestellt werden. In diesem Fall darf der Patient nicht weiter als „vegetative“ bezeichnet werden, vielmehr befindet er sich nun in einem Grenzzustand, dem sogenannten Minimally Conscious State (MCS) (Giacino et al., 2002). Charakteristisch dabei sind grundlegende, nicht reflexartige Verhaltensmuster wie die Blickfixierung und die Blickfolgebewegung. Zusätzlich kann die Fähigkeit, einfachen Befehlen und Aufforderungen nachzukommen, ausgeprägt sein (Bender et al., 2015). In dem Zustand von Koma und VS fehlen solche Verhaltensmuster gänzlich. Können solche nicht immer auf den ersten Blick offensichtlichen Verhaltensmuster detektiert werden, ist die Abgrenzung des MCS gegenüber dem Zustand von Koma und VS möglich. Dies ist dahingehend entscheidend, als sich selbstverständlich die Therapieoptionen und Heilungsaussichten in den jeweiligen Stadien der DoC wesentlich unterscheiden.

Ein solcher Zustand des MCS kann allerdings nicht nur über die Zwischenstufe des VS entwickelt werden, sondern kann auch direkt aus einem Koma hervorgehen. In sehr seltenen Fällen ist es auch möglich, dieses Syndrom direkt durch eine degenerative oder kongenitale, also angeborene, Schädigung des Nervensystems zu erlangen. Der MCS ist meist nur vorübergehend, kann sich allerdings in einigen Fällen chronifizieren und hin zu einer permanenten Situation entwickeln (Giacino et al., 2002).

Im Koma sind weder Wachheit noch eine Form von Bewusstsein nachweisbar, wobei im Gegensatz dazu der VS eine Form von Wachheit ohne Bewusstsein beschreibt. Der Vegetative Status ist also zusammenfassend durch ein Fehlen von bewussten und reproduzierbaren Reaktionen gekennzeichnet. Sobald jedoch einfache reproduzierbare Reaktionen erkennbar werden, wird angenommen, dass der MCS erreicht wurde. Eine funktionierende Kommunikation, entweder verbal oder nonverbal, zeigt die Rückgewinnung und damit Genesung des Bewusstseins an (Bender et al., 2015). Auch eine zuverlässige funktionelle Nutzung einfacher Gegenstände (zum Beispiel konsistente Versuche, sich mit einem Kamm zu

kämmen) weisen darauf hin, dass der Patient den MCS verlassen hat und zu einem höheren Bewusstseinszustand (oft als Exit-MCS oder EMCS bezeichnet) übergegangen ist.

1.3 Problematik der Diagnosestellung

1.3.1 Allgemeines

Die Schwierigkeit, diese doch sehr nahe beieinanderliegenden Diagnosen alleine durch rein klinische Beobachtung und Untersuchung des Patientenverhaltens korrekt zu stellen, ist offensichtlich. Die Herausforderung liegt dabei vor allem im Erkennen des vorhandenen oder eben nicht vorhandenen Bewusstseins trotz eingeschränkter beziehungsweise gänzlich fehlender Motorik und Aufmerksamkeit des Patienten. Diese Problematik der klinischen Unterscheidung, vor allem zwischen dem VS und dem MCS, zeigen vergangene Studien. Hier wurde aufgeführt, dass 37 – 43 % der mit VS diagnostizierten Patienten dennoch Anzeichen für ein Restbewusstsein aufweisen. Der am häufigsten berichtete Fehler stellt dabei die falsch negative Diagnose dar. Dabei wurden Patienten, die sich korrekterweise im MCS (manchmal sogar in einem höheren Bewusstseinszustand) befanden, als VS-Patienten fehldiagnostiziert (Schnakers et al., 2009) (Andrews et al., 1996). Diese häufige Fehldiagnose hat natürlich drastische ethische und praktische Auswirkungen für den Patienten und dessen Angehörige, aber selbstverständlich auch für das Pflegepersonal. Denn auch MCS Patienten leiden durchaus emotional unter fehlender Aufmerksamkeit, die aus einer fehlerhaften Einstufung ihrer Situation resultieren kann. Zudem kann eine für den Patienten falsche Therapie und Pflege eine mögliche Genesung beeinflussen. Auch die Entscheidung bezüglich der Fortführung lebenserhaltender Maßnahmen könnte durch die falsche Einschätzung des Krankheitsverlaufs negativ beeinflusst werden (Bender et al., 2015).

1.3.2 Klinische Diagnostik nach der CRS

Die international bestbewährte klinische Skala zur Unterscheidung von VS und MCS ist die revidierte Coma Recovery Scale (CRS-R), die eine Unterscheidung anhand der Summe von Punktwerten im Bereich von 0 (tiefst mögliches Koma) bis 23 (wach und voll orientiert) ermöglicht. Die Untersuchungskriterien können der Abbildung 1 entnommen werden. Dabei ist es wichtig, zu berücksichtigen, dass für die Diagnose des VS und MCS nicht die Summe aller Werte entscheidend ist, sondern lediglich einzelne Unterpunkte. So sind beispielsweise für die Diagnose des MCS lediglich die in Abbildung 1 mit * gekennzeichneten Rubriken von Bedeutung.

Coma Recovery Scale – Revised ^{©2004}

Record Sheet

This form should only be used in conjunction with the CRS-R Administration and Scoring Manual which defines guidelines for standardized application of the scale

Patient:					Diagnosis:					Etiology:									
Date of onset:					Date of Examination:														
Date																			
Week	Admission	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16			
AUDITORY FUNCTIONS																			
4	Consistent Movement to Command *																		
3	Reproducible Movement to Command *																		
2	Localization to Sound																		
1	Auditory Startle																		
0	None																		
VISUAL FUNCTIONS																			
5	Object Recognition*																		
4	Object Localization: Reaching*																		
3	Visual Pursuit *																		
2	Fixation*																		
1	Visual Startle																		
0	None																		
MOTOR FUNCTIONS																			
6	Functional Object Use**																		
5	Automatic Motor Response*																		
4	Object Manipulation*																		
3	Localization to Noxious Stimulation*																		
2	Flexion Withdrawal																		
1	Abnormal Posturing																		
0	None/Flaccid																		
OROMOTOR/ VERBAL FUNCTIONS																			
3	Intelligible Verbalization*																		
2	Vocalization / Oral Movement																		
1	Oral Reflexive Movement																		
0	None																		
COMMUNICATION SCALE																			
2	Functional: Accurate**																		
1	Non-functional: Intentional*																		
0	None																		
AROUSAL SCALE																			
3	Attention																		
2	Eye opening without stimulation																		
1	Eye opening with stimulation																		
0	no arousal response																		
TOTAL SCORE																			

Denotes emergence from MCS**
Denotes MCS*

Abbildung 1: JFK Coma Recovery Scale-Revised (Zasler et al., 2013)

Die Rate an Fehldiagnosen lässt sich trotz dieser standardisierten Methode und speziell geschulten Personals mit jahrelanger Erfahrung folgendermaßen erklären. Alle diese Untersuchungsmethoden hängen von einer visuell eindeutig ersichtlichen Reaktion des Patienten auf die gestellte Aufgabe ab, zum Beispiel einen gezeigten Finger mit den Augen zu fixieren und zu verfolgen. Bleiben die geforderten Reaktionen aus, so wird als Schlussfolgerung angenommen, der Patient sei nicht bei Bewusstsein und dadurch nicht in der Lage die geforderte Aufgabe korrekt auszuführen. Mit dieser Annahme erfolgt zwangsläufig die diagnostische Einstufung des Patienten im VS. Dabei wird die Tatsache, dass der Patient womöglich aufgrund verschiedener Dysfunktionen nicht in der Lage ist, die gestellte Aufgabe zu bewerkstelligen, nur schwer oder sogar gar nicht berücksichtigt. So kann zum Beispiel die Motorik so weit eingeschränkt sein, dass eine visuelle Fixierung unmöglich ist, obwohl die Aufgabe verstanden wurde. Eine weitere Fehlerquelle kann die mangelnde Motivation von Patienten zur Mitarbeit sein, die die Mitarbeit schlichtweg verweigern. Das bei diesen Patientengruppen durchaus vorhandene Bewusstsein wird mit all seinen negativen Folgen nicht erkannt und bleibt aus rein klinischer Betrachtung verborgen.

1.3.3 Erweiterte Diagnostik durch Bildgebung und elektrophysiologische Methoden

Die unter anderem aus den oben genannten Gründen resultierende hohe Rate an Fehldiagnosen veranlasst die Suche nach neuen, nicht rein klinischen, vom Patienten weitgehend unabhängigen Untersuchungsmethoden, um die Hirnfunktionen der DoC-Patienten besser verstehen und beurteilen zu können. Zudem sollen die Grenzen der klinischen Untersuchung von Verhaltensmustern zur Aufdeckung von Wahrnehmung und Bewusstsein überwunden werden. Sie können als Ergänzung zu etablierten klinischen Untersuchungsmethoden solche Patienten mit vorhandenen kognitiven Funktionen erfassen, die nicht in der Lage sind oder nicht die Motivation besitzen aktiv Zeichen von Bewusstsein zu zeigen. Im

Wesentlichen werden hierzu, wie nachfolgend erläutert, verschiedene testermittelte EEG-, MRT- und PET-Daten gewonnen und entsprechend ausgewertet. Dabei haben sich drei Gruppen von Paradigmen zur Erkennung von verborgenem Bewusstsein als effektiv erwiesen (Kondziella et al., 2016):

1. Aktive Denkmuster: Der Patient wird dabei dazu angeregt, selbstständig eine einfache Denkaufgabe zu lösen.
2. Passive Reizverarbeitung: Dem Patienten werden externe Reize dargeboten, aber keine aktiven Aufgaben.
3. Ruhezustand: Hierbei wird eine Hypothese über den Bewusstseinszustand des Patienten durch die Beobachtung spontaner Hirnaktivitäten einschließlich des Default Mode Networks aufgestellt. Das Default Mode Network oder auch Ruhezustandsnetzwerk beschreibt die Gehirnareale, die aktiv sind, wenn keinerlei Aktivitäten oder Aufgaben ausgeführt werden. Es umfasst unter anderem den präfrontalen Cortex, ein Teil des Gyrus cinguli, den Praecuneus und den Hippocampus (Lauren. et al., 2017).

Auch wenn die aktiven Denkmuster eine deutlich sicherere und zuverlässigere Methode zur Erkennung von verborgenem Bewusstsein darstellen, haben die passiven und Ruhezustandsuntersuchungen durchaus ihre Berechtigung (Kondziella et al., 2016). Die aktive Untersuchungsmethode kann den Untersucher und den Patienten vor dieselben Probleme wie die rein klinische Diagnostik stellen, nämlich dass das durchaus vorhandene Bewusstsein aus oben genannten Gründen nicht gezeigt und damit auch nicht erkannt werden kann. Jedoch erlauben die passiven und Ruhezustandsuntersuchungen, Anzeichen von Bewusstsein derjenigen Patienten zu erkennen, die nicht in der Lage sind, aktive Aufgaben zu bewerkstelligen oder diese nicht bewerkstelligen wollen. Das Unvermögen solche Aufgaben zu lösen kann wie bei der rein klinischen Untersuchung verschiedene Ursachen haben, zum Beispiel Aphasie (Störung des Sprachverständnisses und Sprachproduktion), schwere motorische Lähmungen, besonders in der Kombination mit Sehstörungen (Andrews et al., 1996), Depression und somit fehlende Kooperation oder schlichtweg Taubheit (Kondziella et al., 2016).

Das kann man in den diagnostisch-statistischen Begriffen ausdrücken: Die aktiven Methoden haben eine 100 %-Spezifität, weil es keine falsch positiven Befunde gibt. Der Preis dafür kann aber eine niedrige Sensitivität sein (viele falsch negative Befunde) (Lohninger, 11.11.2012).

Im Weiteren soll nun kurz die Funktionsweise der einzelnen und am häufigsten verwendeten modernen Untersuchungstechniken erläutert werden.

Positronenemissionstomografie (PET)

Die Positronenemissionstomografie ist ein nuklearmedizinisches Verfahren, bei dem die Emission von Positronen aus verschiedenen radioaktiven Substanzen (tracer) erfasst wird. Für eine PET-Untersuchung bekommen die Patienten meist eine veränderte Form der Glucose, die Fluordesoxyglucose (FDG). Da dieser Tracer vom Körper wie Glucose behandelt wird, hängt seine Aufnahme ins Gewebe maßgeblich von der jeweiligen Stoffwechselaktivität des zu untersuchenden Gewebes ab. Eine im Vergleich zum umgebenden Gewebe erhöhte Konzentration an FDG zeigt also eine gesteigerte Stoffwechselrate und damit zum Beispiel eine erhöhte neuronale Aktivität an. Eine verringerte Einlagerung der FDG im Vergleich zu anderem Gewebe derselben Art lässt auf eine reduzierte Stoffwechselrate und somit eine Schädigung dieses Gewebes schließen. Die vom radioaktiv markierten Gewebe ausgesendete Strahlung kann durch einen PET-Scanner erfasst und mittels Computer zum Bild verarbeitet werden. Zur Darstellung der Aktivität bei aktiven und passiven Denkmustern können wegen der im Vergleich zur FDG, kurzen Halbwertszeit dieser Moleküle anstatt der FDG Wassermoleküle mit radioaktiv markiertem Sauerstoff verwendet werden (Boly et al., 2008). Abbildung 2 zeigt das Ergebnis einer FDG-PET-Untersuchung bei einem gesunden Probanden auf der linken Seite und einem VS-Patienten auf der rechten Seite. Regionen mit hoher Glucoseaufnahme und damit hoher Aktivität werden in Rot dargestellt, Regionen mit verringerter Aufnahme in Blau. Der oben beschriebene Unterschied zwischen einem gesunden Probanden

und einem erkrankten wird hierbei durch die Wahl einer Abbildung mit besonders drastischem Kontrast zwischen gesund und erkrankt ersichtlich. Allerdings ist auch in durchschnittlichen Fällen eine Unterscheidung zwischen VS-Patienten und Gesunden zuverlässig möglich.

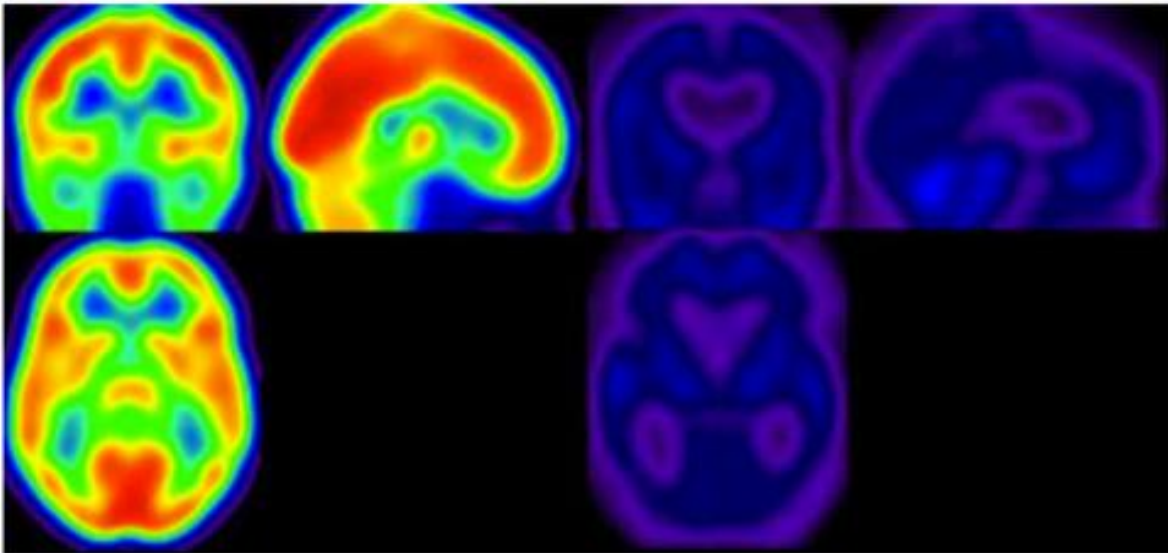


Abbildung 2: Gegenüberstellung FDG-PET gesunder Proband (linke Seite) versus VS-Patient (rechte Seite) (Di Perri et al., 2014)

Um eine exakte Lokalisation des untersuchten Gewebes zu ermöglichen, kann die PET-Untersuchung mit der konventionellen Computertomografie (CT) verknüpft werden. In kombinierten PET/CT-Geräten können beide Untersuchungen direkt hintereinander durchgeführt werden, wodurch sie sich sehr gut zur genauen Lokalisation eines PET-Befundes eignen.

Funktionelle Magnetresonanztomografie (fMRT)

Unter der funktionellen Magnetresonanztomografie versteht man eine Untersuchungsmethode, die auf den natürlichen negativen magnetischen Eigenschaften von oxygeniertem Hämoglobin und der positiven magnetischen Polarität von desoxygeniertem Hämoglobin beruht. Dadurch kann die Blutoxygenierung und damit die Änderung der Gewebsdurchblutung, die durch den Energiebedarf aktivierter Nervenzellen hervorgerufen wird, erfasst werden. Diese unterschiedlichen magnetischen Eigenschaften von Oxyhämoglobin und Desoxyhämoglobin werden in dem so genannten BOLD-Signal (Blood Oxygenated Level Dependent Signal) ermittelt und zur Generierung der Schnittbildserien genutzt (Di Perri et al., 2014). Der gesteigerte Stoffwechsel in bestimmten Hirnarealen, hervorgerufen durch deren gezielte Stimulation und damit Aktivierung, führt zwangsläufig auch zu einer gesteigerten regionalen zerebralen Durchblutung. Durch das somit veränderte Verhältnis von oxygeniertem zu desoxygeniertem Blut ist der Computer in der Lage, das stimulierte Hirnareal zu lokalisieren und visuell darzustellen. Beispielhaft ist die Darstellung aktivierter Hirnareale in Abbildung 3 zu erkennen.

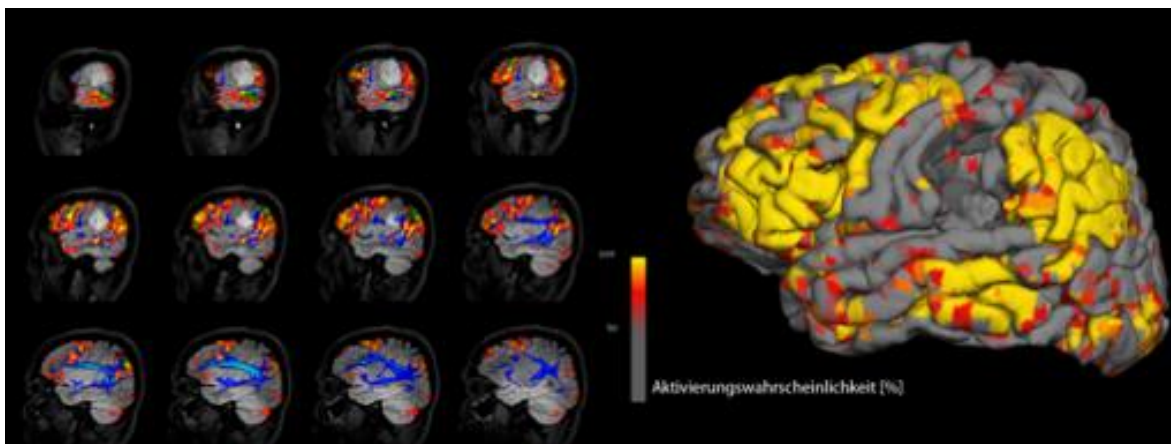


Abbildung 3: Beispielhafte Darstellung aktiver Hirnareale im fMRT (Bildgebung)

Das fMRT eignet sich hervorragend dazu, DoC-Patienten, die aus den oben beschriebenen Gründen nicht dazu in der Lage sind, aktive Aufgaben zu erfüllen, zu untersuchen. Mit Hilfe der fMRT ist es möglich, die Aktivierung von Hirnarealen auch in einem vollständigen Ruhezustand des Patienten nachzuweisen. Der Behandler kann somit Auskunft über verbliebene Gehirnfunktionen ohne Anstrengung oder Anzeichen einer Reaktion von Seiten des Patienten erhalten. Dieses im Ruhezustand aufgezeichnete fMRT kann zeitlich spontan auftretende Veränderungen im BOLD-Signal erfassen und so synchronisierte neuronale Aktivierungen zwischen verschiedenen Hirnarealen nachweisen (Biswal et al., 1997). Zudem kann es Patienten, die motorisch nicht in der Lage sind zu kommunizieren, mit der Technik des fMRT ermöglicht werden, mit ihrer Umgebung in Kontakt zu treten. Vergangene Studien zeigten, dass einige Patienten dadurch einfache Ja/Nein Fragen beantworten können. Dazu wurde die Vorstellung, Tennis zu spielen beziehungsweise sich durch das eigene Haus zu bewegen, genutzt. Wollte der Patient eine alternative Frage mit Ja beantworten, so sollte er zum Beispiel an Tennisspielen denken, bei der Antwort Nein an die Bewegung im eigenen Haus. Da diese beiden Vorstellungen zwei sehr deutlich unterschiedliche Aktivierungsmuster erzeugen (Boly et al., 2007), wird eine reproduzierbare, durch das fMRT registrierte Hirnaktivierung als zielgerichtete Antwort interpretiert (Monti et al., 2010) (Vogel et al., 2013). Der Vorteil dieser Bildgebung gegenüber der PET ist der vollständige Verzicht auf Bestrahlung und radioaktiver Kontrastmittel. Allerdings ist die Untersuchung mittels eines fMRT nicht für alle DoC-Patienten geeignet. Aus rein praktischen Gründen ist es oft aufgrund körperlicher Einschränkungen nicht möglich, jeden DoC-Patienten im MRT zu platzieren und zu untersuchen. Spontane Bewegungen der Patienten im MRT-Scanner verfälschen die Ergebnisse, sodass nur die Patienten untersucht werden können, die sich nicht viel bewegen. Auch stellen nicht-MRT-taugliche medizinische Gerätschaften, wie Schrittmacher oder Metallimplantate eine absolute Kontraindikation dar.

Elektroencephalografie (EEG)

Bei der Elektroencephalografie werden Potenzialänderungen des Gehirns an der Kopfhaut über Elektroden abgeleitet und aufgezeichnet. Die grafische Aufzeichnung der Stromkurven gibt die elektrische Aktivität der Hirnzellen wieder. Die Beurteilung der aufgezeichneten Hirnstromwellen erfolgt durch EEG-Bänder, also die Frequenzbänder, in die das EEG die Aufzeichnung unterteilt. Abbildung 4 zeigt die typische Darstellung einer solchen Aufzeichnung.

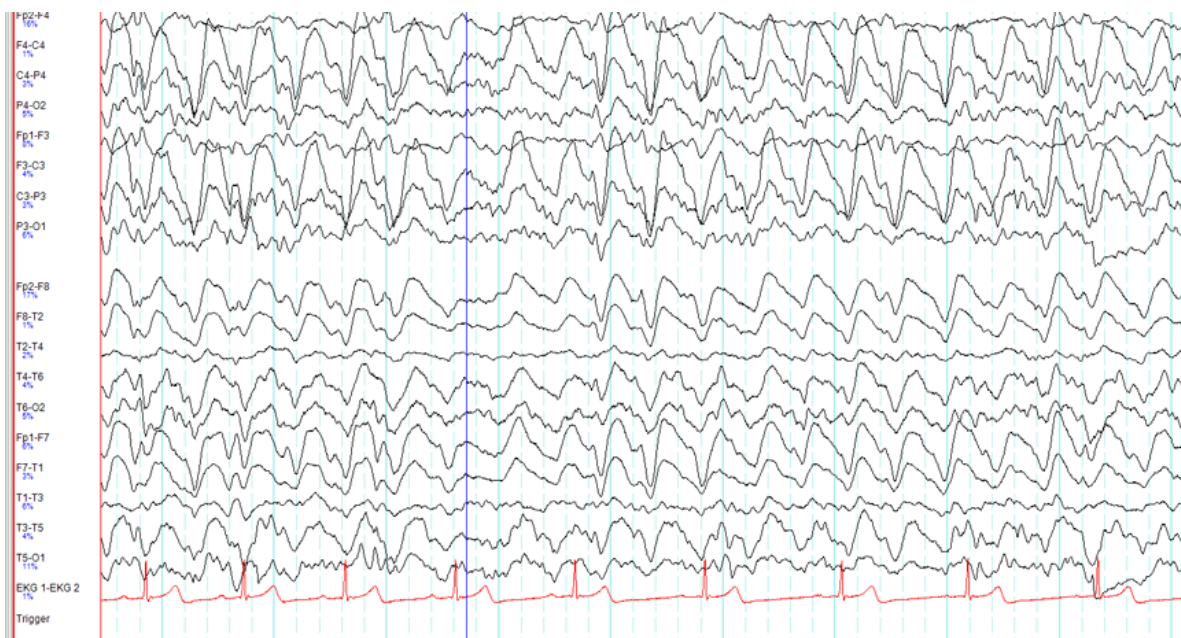


Abbildung 4: Typische Darstellung einer EEG-Aufzeichnung(Hulick, 2016)

Folgende Frequenzbereiche müssen unterschieden werden:

1. Alpha-Wellen: 8 – 13Hz
2. Beta-Wellen: 14 – 30Hz
3. Theta-Wellen: 4 – 7Hz
4. Delta-Wellen: 1 – 3 Hz

Im Allgemeinen gilt der Grundsatz, je niedriger die Frequenz der EEG-Welle, desto geringer die Aktivität der entsprechenden Hirnregion (Bertsch, 2007).

Wie auch das fMRT kann die EEG-Aufzeichnung dazu verwendet werden, verborgene Anteile der Erkenntnisfunktion und des Bewusstseins zu entdecken und dem Patienten die Möglichkeit zu bieten, mit der Umgebung in Kontakt zu treten (Di Perri et al., 2014). Das EEG wird auch genutzt, um Abnormitäten der Hirnaktivität aufzuspüren und aufzuzeichnen. Ein deutlich sichtbarer Effekt ist die Absenkung der Frequenz proportional zur Hirnschädigung. So ist der beim wachen und gesunden Erwachsenen vorherrschende Alpha-Rhythmus in Abhängigkeit von der Vorschädigung, weitgehend durch einen diffusen Theta- oder Delta-Rhythmus ersetzt. Der Delta-Rhythmus entspricht dem physiologischen Muster des Slow Wave Sleep eines gesunden erwachsenen Probanden. Können am Patienten Alpha- oder Beta-Wellen registriert werden, so zeigen diese nicht die für einen gesunden Erwachsenen üblichen charakteristischen Eigenschaften (Lehembre et al., 2012).

Ein äußerst nützliches Hilfsmittel sind die sogenannten Event Related Potentials (ERP), die durch verschiedene Stimuli ausgelöst werden können. Zudem kann damit auch die Zeit, die zur Verarbeitung der Information notwendig ist, ermittelt werden. Dies eröffnet beispielweise die Möglichkeit, die Erregungsleitung und Verarbeitung von visuellen, auditorischen oder emotionalen Stimuli zu überprüfen. Neueste Untersuchungen zeigten, dass sich durch aktive Denkmuster die Amplitude der entsprechenden Modalität registrierbar ändert, sobald sich der Patient auf die entsprechende Aufgabe konzentriert, beziehungsweise solange die Stimulation anhält (Lehembre et al., 2012). Eine solche signifikante Veränderung des EEG zwischen Ruhe- und Aktiver Phase kann ein deutlicher Hinweis auf ein verstecktes aber durchaus vorhandenes Bewusstsein sein. Wird ein solches EEG mit einem Computer verknüpft, der in der Lage ist die Unterschiede grafisch und vereinfacht darzustellen (Brain Computer Interface BCI), stellt dies für den Patienten ein wirkungsvolles Kommunikationsmittel dar. Als Nachteil des EEG gegenüber der Bildgebung mittels PET und fMRI, kann allerdings angesehen werden, dass hierbei

die Aktivität tiefer Hirnareale nicht abgebildet werden kann. Diese Aktivität ist jedoch zur Detektion von Bewusstsein im Bereich der DoC äußerst wichtig.

1.4 Fragestellung

In neuerer Zeit erscheinen einzelne empirische Studien einerseits und die klassischen narrativen Reviews andererseits als nicht mehr ausreichend, um einen genauen Überblick über die aktuelle Studien- und Forschungslage zu einem bestimmten Themengebiet zu geben.

Für einzelne empirische Untersuchungen müssen zahlreiche Daten mittels Experimenten, Untersuchungen und Beobachtungen durchgeführt werden. Anschließend erfolgt eine wissenschaftliche Aufarbeitung auf Grundlage statistischer Verfahren und Interpretation der gewonnenen Daten. Dies ist meist mit einem erheblichen Aufwand und hohen Kosten verbunden. Dennoch ist es vielfach nicht möglich, verlässliche Ergebnisse zu erzielen, da mit zu kleinen Stichproben gearbeitet wird und somit nur vermeintlich signifikante Ergebnisse erzielt werden (Hansueli Stamm, 1995).

Zudem ist es für den Einzelnen aufgrund der Fülle von Einzelstudien nur schwer möglich, alle für ein klinisches Problem relevanten Arbeiten zu ermitteln und zu bewerten. Aus diesem Grund werden meist Übersichtsarbeiten, so genannte Reviews, bevorzugt. Dabei ist zwischen narrative Reviews (unsystematischen) und systematischen Reviews zu unterscheiden, die zu einem Thema die Ergebnisse nach vorgegeben Kriterien zusammenfassen. Die systematischen Reviews fassen, im Gegensatz zu den narrative Reviews die Ergebnisse nicht nur zusammen, sondern arbeiten diese auch statistisch auf. Durch diese Zusammenfassung erhofft man sich ein genaueres Ergebnis, als es die Einzelstudien liefern können, deren Resultate oftmals weit streuen und breite Konfidenzintervalle aufweisen, sodass man sich nicht sicher sein kann, wo genau das eigentliche Ergebnis liegt. Die vielfach an ein systematisches Review angeschlossene Metaanalyse geht darüber

hinaus, fasst die Einzelergebnisse zu einem gepoolten Ergebnis zusammen und erlaubt dem Leser damit einen raschen Überblick über die Studienlage (Bown and Sutton, 2010). Dadurch haben systematische Reviews mit angeschlossener Metaanalyse stetig an Bedeutung gewonnen und darüber hinaus wird ihnen die höchste Stufe an Evidenz zugewiesen, sodass derartige Arbeiten oftmals als Grundlage für Leitlinien herangezogen werden (Sutton and Higgins, 2008, Sterne et al., 2001).

Ziel der vorliegenden Arbeit ist es zu erläutern inwieweit die diagnostischen Hilfsmittel im Bereich der bildgebenden Verfahren nun tatsächlich dazu dienen, die beschriebene äußerst problematische Unterscheidung zwischen MCS und VS zu erleichtern sowie effektiver zu gestalten. Dies soll an Hand des vorliegenden systematischen Reviews und der angeschlossenen Metaanalyse, soweit mit den vorhandenen Daten möglich, statistisch quantitativ geklärt werden. Damit kann sowohl eine Aussage über aktuell bereits angewendete als auch zukünftig mögliche Untersuchungsmethoden im Bereich der Bildgebung sowie deren diagnostische Effektivität und Wirksamkeit in Hinblick auf die Unterscheidung zwischen MCS und VS gegeben werden. Hierzu wurden zahlreiche Studien, die sich mit den diagnostischen Unterschieden zwischen MCS und VS beschäftigen, analysiert und mithilfe einer Metaanalyse auf ihre Aussagekraft untersucht.

2 Material und Methoden

2.1 Allgemeines zu Systematic Review und Metaanalyse

Aufgrund der stetigen Zunahme an wissenschaftlichen Publikationen gestaltet es sich äußerst schwierig, fortwährend auf dem neuesten Stand der Forschung zu bleiben. Aus diesem Grund ist es für einen ausgewogenen Überblick über die aktuelle Literatur und Studienlage zu einem speziellen Thema wichtig, diese in einer

Arbeit zusammenzufassen und gemeinsam zu bewerten. Genau dies ermöglichen die systematischen Übersichtsarbeiten, im englischen Sprachgebrauch auch als systematic reviews bezeichnet. Im Vergleich zu der traditionellen Form einer nicht systematischen Übersichtsarbeit ist es mithilfe der systematischen Übersichtsarbeiten möglich aus der Fülle an wissenschaftlichen Arbeiten durch eine im Folgenden beschriebene entsprechende systematische Auswahl und Auswertung einen optimalen Überblick über die aktuelle Forschungslage zu einem bestimmten Thema zu ermöglichen. In der traditionellen Form einer „erzählenden“ Übersichtsarbeit entfallen diese scharfen Auswahl- und Analyse Kriterien, was zu subjektiv bedingten Verzerrungen führen kann. So kann der Autor einer „erzählenden“ Übersicht bestimmte Studien, die seine Ansicht nicht bestätigen, weglassen (oftmals unbewusst) oder ihre Bedeutung herunterspielen, da die Kriterien, nach welchen die Studien eingeschlossen oder ihre Bedeutung eingestuft wird, nicht festgelegt werden. Deshalb können aufgrund rein beschreibender Übersichtsstudien keine zuverlässigen Aussagen und Empfehlungen getroffen werden.

Zu Beginn werden zu einer expliziten und klar formulierten Fragestellung sämtliche wissenschaftliche Forschungsergebnisse zusammengefasst und analysiert (hier: Diagnostische Unterscheidungsmöglichkeiten von Patienten in einem MCS von denjenigen in einem VS). Daher wurde auch für diese Arbeit, um eine gleichmäßige Literaturübersicht und statistische Beurteilung zu ermöglichen, das Verfahren der systematischen Übersichtsarbeit mit angeschlossener Metaanalyse gewählt. Allgemein werden von der Cochrane Collaboration folgende Charakteristika als wesentlich für ein Review definiert (Higgins JPT, 2011):

- Festlegung auf ein definiertes Thema
- Angabe reproduzierbarer Suchstrategien für Studien
- definierte Ein- und Ausschlusskriterien
- Auflistung aller identifizierter Studien
- Darstellung der einzelnen Studiencharakteristika mit Qualitätsanalyse

- Analyse der ausgewählten Studien mittels Metaanalyse
- Darstellung der Sensitivität der Studien
- strukturierte Zusammenfassung der Ziele, Methoden, Materialien und Ergebnisse

Die so erstellte Literaturübersicht stellt im nächsten Schritt die Grundlage der Metaanalyse dar, die der Erfassung und Auswertung der gewonnenen Daten und Evidenzen dient. Hierzu werden die Informationen der einzelnen Artikel zum selben Themengebiet extrahiert und mathematisch unter Berücksichtigung statistischer Signifikanzen und Effektstärken zu einem Gesamtergebnis gebündelt. Dieses Vorgehen erweist sich besonders im Hinblick auf die Thematik der Unterscheidung verschiedener DoC als äußerst sinnvoll. Die Aussagekraft und der Effekt einzelner Studien kann aufgrund kleiner Studiengrößen durchaus eingeschränkt sein. Durch eine integrierte und quantitative Bündelung der aus den einzelnen Studien gewonnenen Studienergebnissen ist eine deutliche Steigerung der Aussagekraft möglich.

2.2 Reviewfrage

Zur Entwicklung einer zielführenden Suchstrategie ist es von entscheidender Bedeutung, eine strukturierte Fragestellung zu formulieren. Hierzu wird die Reviewfrage in vier Komponenten untergliedert (Khan et al., 2003), die Population, die Intervention und Darstellung bezüglich der Fragestellung, das Ergebnis, sowie das Studiendesign, das zur Klärung der Fragestellung hilfreich erscheint.

Für diese Arbeit wurde folgende Frage formuliert:

Möglichkeiten der Differenzierung zwischen dem VS und dem MCS mittels funktioneller Bildgebungsverfahren.

Dabei erfolgt die Untergliederung in die oben beschriebenen vier Komponenten wie folgt:

- Population: Patienten mit den syndromalen Diagnosen Minimally Conscious State oder Vegetative State
- Intervention: Vergleich der beiden Patientengruppen anhand funktioneller bildgebender Diagnostik
- Ergebnis: Unterscheidungsmöglichkeiten der beiden Patientengruppen durch funktionelle bildgebende Verfahren
- Studiendesign: quantitative, statistische Studien

2.3 Literaturrecherche

2.3.1 Suchkriterien

Für die Auswahl der relevanten Artikel wurden folgende Auswahl- bzw. Ausschlusskriterien festgelegt:

- Diagnostisch neurophysiologisch vergleichende Beschreibung von Patienten mit der Diagnose MCS als auch VS (bzw. Synonyme wie UWS oder Wachkoma) anhand funktioneller bildgebender Verfahren
- Ausschluss der entsprechenden Artikel, wenn nur eine Patientengruppe (also VS oder MCS) untersucht wird
- Ausschluss, wenn VS/MCS Patienten nur erwähnt bzw. diskutiert werden, aber in der Studie nicht real untersucht wurden
- Ein bildgebendes Verfahren wie PET oder fMRT als Anwendungsmethode
- Ausschluss, wenn eine solche neurophysiologische Untersuchungsmethode auf Grundlage der Bildgebung nur erwähnt wird, aber nicht angewandt wurde

- Artikel, die sich mit der Verbesserung und Regeneration des Bewusstseins ausgehend vom Status des MCS und VS beschäftigen (auch hier gilt, dass beide Patientengruppen untersucht werden müssen).

Als weitere Kriterien gelten: Ausgeschlossen werden Studien, in denen morphologische (aber keine funktionellen) Bildgebungsverfahren angewendet wurden, sowie Übersichtsartikel, weil sie die Patientenstichproben weiter analysieren, die bereits in der primären Literatur analysiert worden sind.

2.3.2 Artikelauswahl

Anhand der beschriebenen Suchkriterien wurde mithilfe der Onlinedatenbank Pubmed (National Center for Biotechnology Information) die Literaturrecherche im Zeitraum vom 01.10.2016 bis zum 15.06.2018 durchgeführt. Um die Suchergebnisse optimal auf die Reviewfrage anzupassen, wurden die in folgender Tabelle 1 dargestellten Suchalgorithmen angewandt.

Suchanfrage	Suchalgorithmen
1	Disorders of conscious AND minimally conscious state
2	Disorders of conscious AND vegetative state
3	Minimally conscious state AND differences
4	Vegetative state AND differences
5	Minimally conscious state AND stimulation
6	Vegetative state AND stimulation
7	Vegetative state AND fMRI
8	Minimally conscious state AND fMRI
9	Vegetative state AND MRI
10	Minimally conscious state AND MRI
11	Vegetative State AND minimally conscious state
12	Minimally conscious state AND vegetative state AND prognosis
13	Vegetative State AND PET
14	Minimally Conscious State AND PET

Tabelle 1: Suchalgorithmen

Die so durchgeführte Suche ergab nach Addition der in Tabelle 1 dargestellten Einzelsuchergebnisse eine initiale Gesamtsumme von 7336 möglicherweise relevanten Artikeln. Im folgenden Schritt wurden all diese Artikel unter Anwendung der oben angeführten Ein- und Ausschlusskriterien einer ersten Relevanzprüfung anhand der Titel und Abstracts unterzogen. Nach der Entfernung aller Artikel, deren Abstracts für die vorliegende Arbeit irrelevant waren und nicht zu den Einschlusskriterien passten, verblieb eine Summe von 518 relevant erscheinender Artikel. Alle jene Artikel, die durch die Abstracts nicht eindeutig den Einschlusskriterien zugeordnet werden konnten, wurden in diesem Schritt zunächst als potentiell relevant angesehen.

Allerdings ist bei diesen 518 Artikeln noch nicht von der endgültigen zu untersuchenden Anzahl an Studien auszugehen, da die verschiedenen

Suchalgorithmen teilweise identische Ergebnisse liefern. Daher war es notwendig, die zu diesem Zeitpunkt als relevant erscheinenden Artikel, entsprechend der einzelnen Suchmasken gruppiert auf eventuell vorliegende Duplikate zu prüfen. Diese Prüfung erfolgte mit dem Programm Novus Scan (Labs, 2016). Hiernach verblieb eine Gesamtzahl von 124 Artikel, die nun durch genaue Betrachtung der Volltexte auf ihre Eignung für die vorliegende Arbeit untersucht wurden. Die Beschaffung der Volltexte zu den jeweiligen Studien erfolgte bei Bedarf über die Online-Plattform Subito (e.V.), sofern dies mittels pubmed.com nicht möglich war. Nach Sichtung und Bewertung der Volltexte, entsprechend der genannten Ein- und Ausschlusskriterien, verblieben 39 für diese Arbeit relevanten Studien. Eine detaillierte Auflistung der einzelnen Studienanzahlen zu den entsprechenden Suchkriterien ist Tabelle 2 zu entnehmen.

Suchanfrage	Suchalgorithmus	Erstes Suchergebnis	Erste Sichtung	Duplikate entfernt
1	Disorders of conscious AND minimally conscious state	137	30	5
2	Disorders of conscious AND vegetative state	153	31	9
3	Minimally conscious state AND differences	134	25	1
4	Vegetative state AND differences	340	23	8
5	Minimally conscious state AND stimulation	286	27	0
6	Vegetative state and AND stimulation	336	27	27
7	Vegetative state AND fMRI	393	30	24
8	Minimally conscious state AND fMRI	328	43	4
9	Vegetative state AND MRI	412	50	21
10	Minimally conscious state AND MRI	334	43	0
11	Vegetative State AND minimally conscious state	3413	134	0
12	Minimally conscious state AND vegetative state AND prognosis	884	30	30
13	Vegetative State AND PET	89	14	0
14	Minimally conscious state AND PET	89	11	0
Summe		7328	518	124

Tabelle 2: Studienanzahl zu den entsprechenden Suchkriterien

Die endgültig als relevant verbliebenen 39 Artikel wurden zur weiteren Auswertung in verschiedene Gruppen klassifiziert. Hierbei erfolgte die Unterteilung in qualitative Übersichtsarbeiten, statistische Arbeiten mit quantitativen Angaben und solche ohne quantitativen Angaben. Dabei sind die statistischen Arbeiten mit den entsprechenden quantitativen Angaben die Wichtigsten, da diese für die weitere Auswertung und Metaanalyse von Bedeutung sind. Damit ergibt sich für die Gruppe der quantitativen Studien mit den entsprechend notwendigen statistischen Angaben eine Gesamtzahl von 25 mit insgesamt 35 Datensätzen (eine Studie kann mehrere Datensätze beinhalten). 5 Artikel waren Übersichtsarbeiten und die restlichen 9 waren empirische Studien, die jedoch die für diese Arbeit essenziellen statistischen Angaben nicht aufführen. Für die angeschlossene Metaanalyse ergab sich, entsprechend der unten aufgeführten Kriterien, eine Anzahl von 14 Artikel mit in Summe 27 Datensätzen.

Das Vorgehen zur Selektion der relevanten Artikel kann in nachfolgendem Flow Chart (Abbildung 5) detailliert nachvollzogen werden.

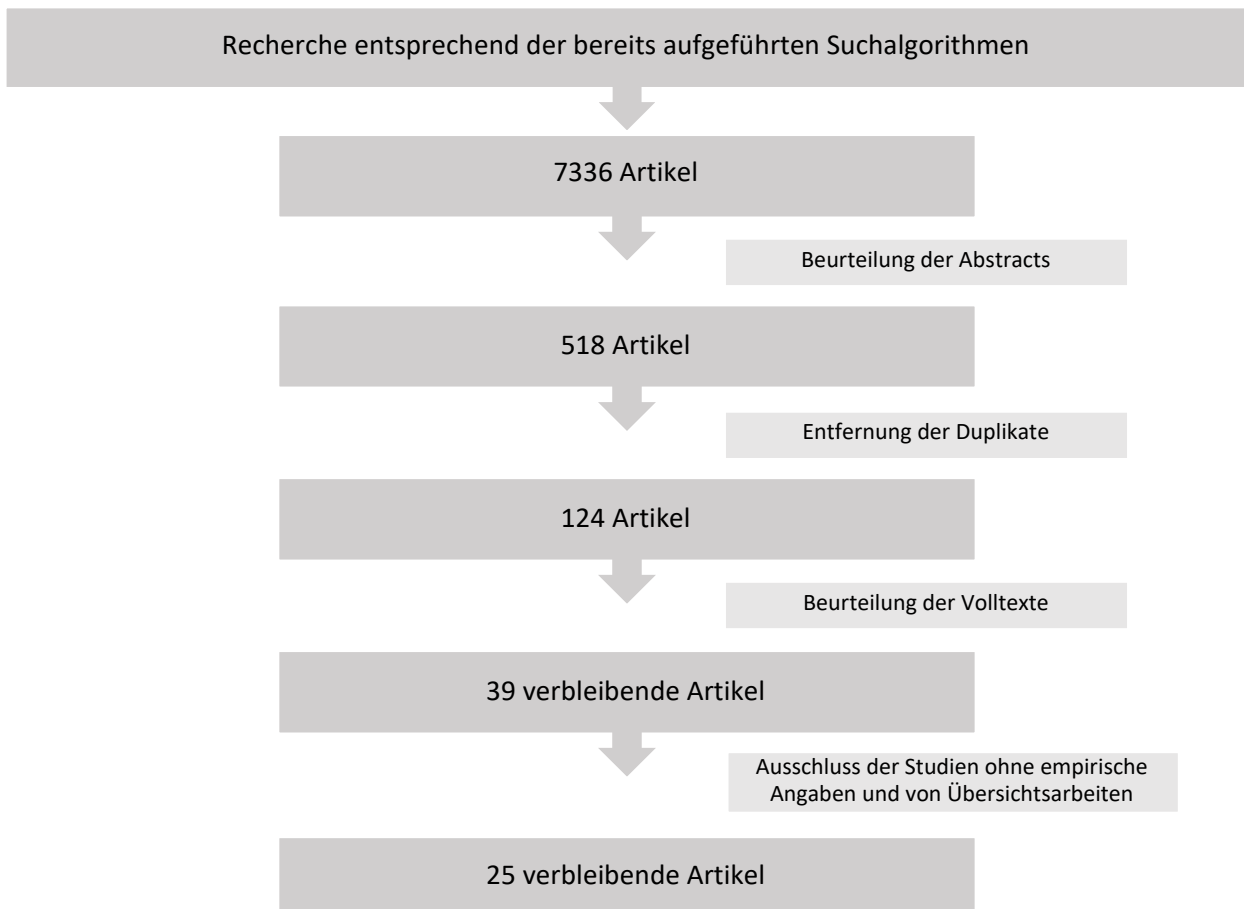


Abbildung 5: Flow Chart zur Artikelauswahl

Bei den unten aufgeführten Artikeln handelt es sich um die für die vorliegende Arbeit relevanten 25 Studien. Einen genauen Überblick dieser Artikel mitsamt der im nachfolgenden erläuterten Daten bietet Tabelle 7 im Anhang.

1. Implementing novel imaging methods for improved diagnosis of disorder of consciousness patients. (Bick et al., 2013)
2. Thalamo-frontal connectivity mediates top-down cognitive functions in disorders of consciousness. (Monti et al., 2015)
3. Willful modulation of brain activity in disorders of consciousness. (Monti et al., 2010)

4. Quantitative rates of brain glucose metabolism distinguish minimally conscious from vegetative state patients. (Stender et al., 2015)
5. Reductions of thalamic volume and regional shape changes in the vegetative and the minimally conscious states. (Fernandez-Espejo et al., 2010)
6. Multimodal Study of Default-mode network Integrity in disorders of consciousness. (Rosazza et al., 2016)
7. Central olfactory processing in patients with disorders of consciousness. (Nigri et al., 2016)
8. Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: a clinical validation study. (Stender et al., 2014)
9. Global functional connectivity reveals highly significant differences between the vegetative and minimally conscious state. (Kotchoubey et al., 2013)
10. fMRI activation with an "affective speech" paradigm in vegetative and minimally conscious states: applicability and prognostic value. (Piperno et al., 2012)
11. Deactivation of the default mode network as a marker of impaired consciousness: an fMRI study. (Crone et al., 2011)
12. A network approach to assessing cognition in disorders of consciousness. (Rodriguez Moreno et al., 2010)
13. Diffusion weighted imaging distinguishes the vegetative state from the minimally conscious state. (Fernandez-Espejo et al., 2011)
14. Cerebral response to speech in vegetative and minimally conscious states after traumatic brain injury. (Fernandez-Espejo et al., 2008)
15. Cerebral response to patient's own name in the vegetative and minimally conscious states. (Di et al., 2007)
16. A role for the Default Mode Network in the Base of Disorders of Consciousness. (Fernandez-Espejo et al., 2012)
17. The neural correlates of lexical processing in disorders of consciousness. (Nigri et al., 2017)

18. True or False? Activations of Language-related Areas in Patients with Disorders of Consciousness. (Kotchoubey et al., 2014)
19. Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. (Coleman et al., 2009)
20. Can mental imagery functional magnetic resonance imaging predict recovery in patients with disorders of consciousness? (Vogel et al., 2013)
21. Brain response to thermal stimulation predicts outcome of patients with chronic disorders of consciousness. (Li et al., 2015)
22. Degrees of functional connectome abnormality in disorders of consciousness. (Sinitsyn et al., 2018)
23. Functional MRI motor Imaginary Task to detect command following in traumatic. (Bodien et al., 2017)
24. Early detection of consciousness in patients with acute severe traumatic brain injury. (Edlow et al., 2017)
25. Functional Evaluation of Awareness in Vegetative and Minimally Conscious State. (Marino et al., 2017)

2.4 Analyse der Studien und Datenextraktion

2.4.1 Qualitätsbeurteilung

Als Nachteil des Vorgehens zur Erstellung eines Reviews ist die individuell stark divergente wissenschaftliche Qualität der dabei verwendeten Artikel anzusehen. Um die individuellen Qualitäten der jeweiligen Studien abschätzen und vergleichen zu können, wurden in dieser Arbeit drei Kriterien angewendet. Diese drei Kriterien zur Qualitätsbeurteilung beinhalten zum Ersten die Patientenzahl, zum Zweiten den sogenannten Journal Impact Factor und zuletzt die Ermittlung des Bias und optional der Applicability nach der Referenz auf die QUADAS-Methode (Quality Assessment

of Diagnostic Accuracy Studies). Im Folgenden werden die genannten Kriterien genauer betrachtet.

- Patientenanzahl

Der Umfang der analysierten Studien, also die Anzahl der jeweils untersuchten Patientengruppen, diente als einfaches und schnell ersichtliches Qualitätsmerkmal. Dabei wurde davon ausgegangen, dass die Größe der zu untersuchenden Patientengruppen auch mit der Zuverlässigkeit des Ergebnisses der jeweiligen Studie zusammenhängt. Je größer sich dabei die untersuchte Patientengruppe darstellt, desto aussagekräftiger sollte das Ergebnis sein.

- Journal Impact Factor

Der sogenannte Journal Impact Factor (JIF) oder schlicht Impact Factor (IF) der wissenschaftlichen Zeitschriften, welche den jeweiligen Artikel veröffentlichten, wurde als weiteres Qualitätsmerkmal herangezogen.

Obwohl der Impact Factor kein direktes Qualitätsmaß eines in der Zeitschrift veröffentlichten Artikels ist, kann er als indirektes Maß neben anderen verwendet werden, da gezeigt wurde, dass der Impact Factor direkt mit dem informellen Ruf und dem wissenschaftlichen Einfluss der entsprechenden Zeitschrift korreliert (Baethge, 2012). Um den Einfluss und die Wirkung einer wissenschaftlichen Zeitschrift zu bemessen, wird ermittelt, wie oft die zu evaluierende Zeitschrift in einer anderen Veröffentlichung der gleichen Fachrichtung zitiert wird. Allerdings ist anhand des Journal Impact Factors ein zuverlässiger Vergleich wissenschaftlicher Zeitschriften lediglich innerhalb eines Fachgebietes möglich, da sich die Größe des Fachgebietes und das je nach Fachgebiet variierende Zitierverhalten direkt auf den Journal Impact Factor auswirken. Bei genauer Betrachtung der Berechnung des Impact Factors wird dieser Zusammenhang klar ersichtlich. Die Berechnung erfolgt dabei auf Grundlage der in den zurückliegenden zwei Jahren veröffentlichten Artikel. Wenn also beispielhaft der Journal Impact Factor für 2017 bestimmt werden soll, erfolgt dies auf Basis der 2016 und 2015 veröffentlichten Artikeln. Zur Berechnung

werden nun die Zahl der Zitationen durch die Gesamtsumme der Quellartikel dividiert. Dafür ergibt sich somit folgende Formel (Lewandowski, 2006):

$$JIF = \frac{c}{s1 + s2}$$

Dabei steht JIF für den Journal Impact Factor der verwendeten wissenschaftlichen Zeitschriften, c für die Gesamtsumme der Zitationen der vorangegangenen beiden Jahre und $s1$ sowie $s2$ für die Zahl der Quellartikel der jeweiligen vergangenen Jahre. Allerdings steht der JIF auch durchaus in der Kritik, da es sich hierbei nur um ein sehr indirektes Maß handelt. Durch seine weite Verbreitung ist er aber dennoch durchaus von Interesse.

- Bias und Applicability

Bei diesem Analyseschritt wurde eine Qualitätsbeurteilung der einzelnen Artikel nach dem Prinzip von QUADAS-2 durchgeführt (Whiting et al., 2011). Im Jahr 2003 wurde durch eine Expertenkommission aus dem Bereich der diagnostischen Forschung QUADAS eingeführt. Ziel dabei war es, vor allem den Bias (Verzerrung der Studienergebnisse durch beispielsweise fehlerhafte Untersuchungsmethoden) und die Applicability (Anwendbarkeit der Ergebnisse auf den zu untersuchenden Fall) differenziert beurteilen zu können. Die Weiterentwicklung von QUADAS hin zu dem in dieser Arbeit verwendeten Prinzip des QUADAS-2 erfolgte, um die Qualität und damit auch die Richtigkeit der für ein Review verwendeten primären diagnostischen Studien zu bewerten. Zur Beurteilung des Bias wurden die folgenden Richtlinien auf der Basis von Penny F. Whiting et al., QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies angepasst (Whiting et al., 2011). Die Beurteilung erfolgte mithilfe eines Punktesystems von 0 bis 2. Da dieses Verfahren durchaus einer gewissen subjektiven Beeinflussung unterliegt, erfolgte die Beurteilung durch den Autor und eine weitere unabhängige Person.

Die Autoren unterstreichen (Whiting et al., 2011), dass die QUADAS-2-Prinzipien nur als Richtlinien angesehen werden sollen, die in jedem konkreten Fall entsprechend dem jeweiligen Wissensstand angepasst („tailored“) werden sollen. Als Ergebnis dieser Anpassung wurden hier für den Bias folgende Kriterien verwendet:

1. Patientenauswahl

Punktwert 0: Es liegen Anhaltspunkte dahingegen vor, dass die Patienten auf einer unklaren oder mangelhaft berichteten Grundlage ausgewählt worden sind.

Punktwert 1: Eine Vorauswahl oder entsprechender Ausschluss von Patienten für die Studie kann nicht nachvollzogen werden. Des Weiteren sind Studien, deren Patientengruppen entweder der VS oder der MCS weniger als zehn Patienten umfasst, niemals mit einem höheren Punktwert als 1 zu bewerten, da bei kleineren Gruppen die Wahrscheinlichkeit eines Bias grundsätzlich erhöht wird.

Punktwert 2: Entweder alle Patienten, die den Anforderungen entsprechen, oder eine streng per Zufall ausgewählte Patientengruppe nehmen an der Studie teil.

2. Index-Test

Hier wird der neurophysiologische Index (beispielsweise fMRT oder PET), welcher in den Studien zur Anwendung kommt, bewertet.

Die Fragen gestalten sich dabei folgendermaßen:

1. Wurde der Schwellenwert für eine Aktivierung bestimmter Hirnareale bereits vor der Analyse festgelegt?
2. Wurden tatsächlich alle erhobenen neurophysiologischen Werte veröffentlicht?
3. Sind alle gemessenen neurophysiologischen Kennzeichen mit den entsprechenden statistischen Kenngrößen, wie Durchschnittswerte, Standardabweichung oder Standardfehler, angegeben?

4. Hatten die Personen, welche die Auswertung der erhobenen neurophysiologischen Daten durchführten, im Voraus keine Kenntnis über die Diagnose der untersuchten Patienten (waren also blind bezüglich der Diagnose)?

Punktwert 0: Mindestens zwei der Fragen müssen mit NEIN beantwortet werden.

Punktwert 1: Eine der Fragen muss mit NEIN beantwortet werden.

Punktwert 2: Alle Fragen können mit JA beantwortet werden.

3. Referenz-Standard

Im Unterschied zu Punkt 1 „Patientenauswahl“, bei dem die Möglichkeit erwogen wird, dass MCS- und VS-Patienten mit einer vollkommen anderen Diagnose wie beispielsweise dem Locked-in-Syndrom verwechselt werden, wird hier die Möglichkeit einer Falschdiagnose zwischen den Befunden MCS und VS abgewogen.

Punktwert 0: Es handelt sich lediglich um eine klinische Untersuchung ohne eine quantitative Skala wie die CRS-R

Punktwert 1: Es wurde eine quantitative Skala verwendet, die zur Diagnose eingesetzten Personen hatten aber möglicherweise bereits vor der Bewertung, Kenntnis über die Diagnose.

Punktwert 2: Die neurophysiologische Diagnose wurde mithilfe einer quantitativen Skala und blind in Bezug auf vorhandene vorangegangene neurophysiologische Diagnosen erhoben.

4. Flow und Timing

In dieser Kategorie wird beurteilt, ob ein Flow-Diagramm zur Patientenauswahl vorhanden ist und wie viel Zeit zwischen der Erhebung des Referenz-Standard und dem Index-Test (d.h. zwischen der klinischen und der bildgebenden Untersuchung) liegt. Idealerweise werden diese beiden Untersuchungen an einem Patienten zur selben Zeit erhoben. Tritt, aus welchem Grund auch immer, eine Verzögerung

zwischen diesen beiden Untersuchungen auf, ist durch eine Verbesserung oder Verschlechterung des Zustandes des Patienten innerhalb dieser Zeit durchaus eine erhebliche Fehleinschätzung möglich.

Folgende Fragen dienen der Evaluierung:

1. Ist der Vorgang der Patientenauswahl mittels eines Flow-Diagramms dargestellt?
Aus diesem sollte hervorgehen, zu welchem Zeitpunkt wie viele Patienten, aus welchem Grund ausgeschlossen wurden.
2. War, falls ein zeitliches Intervall zwischen dem Referenz-Standard und der neurophysiologischen Untersuchung (zum Beispiel fMRI oder CT) vorlag, dieses für die beiden diagnostischen Gruppen im Durchschnitt identisch?
3. Für den Fall, dass das Intervall zwischen den Untersuchungen ausgesprochen lang war, ist zu klären, ob alle Patienten, die an der neurophysiologischen Untersuchung teilnahmen, auch nach dem Referenz-Standard beurteilt wurden.

Punktwert 0: Entweder Frage 2 oder Frage 3 kann mit NEIN beantwortet werden.

Punktwert 1: Frage 1 kann mit NEIN, aber Fragen 2 und 3 können mit JA beantwortet werden.

Punktwert 2: Die Antwort auf alle drei Fragen lautet JA.

Der Gesamtwert zur Beurteilung des Bias ergibt sich aus der Summe der Einzelkriterien. Dabei ist ein maximaler Punktwert von 8 bis hin zu einem minimalen Wert von 0 möglich. Somit bedeutet ein hoher Punktwert einen äußerst geringen Bias, also eine geringe Verzerrung des Studienergebnisses. Je höher damit der Punktwert im Bereich des Bias liegt, desto höher ist die Studienqualität.

Um eine in diesem Verfahren durchaus mögliche subjektive Verzerrung der Ergebnisse zu vermeiden, wurde die Analyse nach identischen Vorgaben zweimal unabhängig voneinander durchgeführt. Zu diesem Zweck wurde eine über die gesamte Arbeit und die erste Auswertung nicht informierte Mitarbeiterin des Institutes für medizinische Psychologie und Verhaltensneurobiologie der Eberhard

Karls Universität in Tübingen zur erneuten Qualitätsbeurteilung nach dem QUADAS-2-Prinzip hinzugezogen.

Betrug die Punktwertdifferenz beider Bewertungen in der Summe nicht mehr als 1, so wurde der Mittelwert aus beiden Bewertungen gebildet. Damit erklären sich auch die ungeraden Werte trotz der eigentlich nur geraden zu vergebenden Punktwerte pro Bewertungseinheit. Betrug die Divergenz mehr als 2 Punkte, so wurde durch Herrn Prof. Dr. Boris Kotchoubey eine der folgenden Optionen gewählt: Es wurde entweder der Bewertung des Autors oder der Bewertung der unabhängigen zweiten Bewerterin gefolgt oder wiederum der Mittelwert aus beiden gebildet. Herr Prof. Dr. Boris Kotchoubey durfte dabei keine dritte (eigene) Bewertung abgeben.

Auf eine gesonderte Betrachtung der Applicability wurde in dieser Arbeit verzichtet. Zum einen manifestierte sich durch die Sichtung der Bildgebungsstudien der Eindruck, dass die Ergebnisse durchaus auf die Zielpopulation anwendbar sind. Zum anderen kamen auch Konziella et al. (2016) in ihrem systematischen Review „Preserved conscious in vegetative and minimal conscious states: systematic review and metaanalysis“ zu dem Ergebnis, dass im Gegensatz zum Bias die Applicability in der Regel unproblematisch ist.

2.4.2 Kalkulation der Effektstärken

Für diese Arbeit wurde nun auf Basis der in den 25 hierfür untersuchten Studien experimentell ermittelten Patientendaten, meist mittels des Chi-Quadrat-Unabhängigkeitstests oder der Area Under the Curve, die sogenannte Effektstärke errechnet. Diese Effektstärke zu ermitteln ist von immenser Wichtigkeit, da nur mithilfe der Effektstärke die Bedeutsamkeit des Ergebnisses einer wissenschaftlichen Studie oder Arbeit eingeschätzt werden kann. Als Maß für die Effektstärke wurde in der vorliegenden Arbeit das bekannteste Effektstärkenmaß d nach Cohen angewandt (Cohen, 1988). Die Berechnung der Effektstärke d nach Cohen wurde anhand der Webseiten psychometrica.de (Lenhard W, 2016) und

vilespc01.wiwi.uni-oldenburg.de (Sklorz) durchgeführt. Problematisch gestaltete sich die Berechnung immer dann, wenn in einer 2 x 2-Berechnungstabelle der Wert einer gesamten Spalte null annahm. Dies ergab sich beispielsweise dann, wenn alle Patienten einer Patientengruppe auf ein Signal nicht reagierten. Wird die Kalkulation in dieser Konstellation dennoch durchgeführt, ergibt sich für d immer ein Wert, der im Unendlichen liegt. Abhilfe schafft hier die Möglichkeit, die Patienten so zu gruppieren, dass sich keine Gruppen mit der Anzahl 0 ergeben. Dies ist allerdings nicht immer ohne weiteres möglich, sodass wie von Nakagawa und Cuthill (Nakagawa and Cuthill, 2007) empfohlen vorgegangen werden muss. Nach dieser Methode wird in jeder Spalte der 2 x 2-Berechnungstabelle der Wert 0,5 angefügt. Für die vorliegende Arbeit erhalten die so ermittelten Effektstärken, nach eigenen Richtlinien, ein positives Vorzeichen für den Fall, dass die vergleichenden neurophysiologischen Untersuchungen der MCS-Patienten im medizinischen Sinne ein besseres Ergebnis erzielen als die der VS-Patienten. Andererseits werden die Effektstärken mit einem negativen Vorzeichen versehen, sollten die Ergebnisse der VS-Patienten besser als die der MCS-Patienten ausfallen. Tabelle 3 zeigt im Überblick die Interpretation der Effektstärken d nach Cohen (1988) (Lenhard W, 2016).

d	Interpretation nach Cohen (Cohen, 1988)
≤ 0	Negativer Effekt
0,0	Kein Effekt
0,1	
0,2	Kleiner Effekt
0,3	
0,4	
0,5	Mittlerer Effekt
0,6	
0,7	
0,8	Großer Effekt
0,9	
$\geq 1,0$	

Tabelle 3: Interpretation der Effektstärke d nach Cohen (Cohen, 1988)

Neben den Effektstärken wurden zudem im selben Schritt automatisch durch die jeweiligen oben angeführten Programme das 95 %-Konfidenzintervall (CI) und der entsprechende Standardfehler der Effektstärken ermittelt, da diese Parameter für die spätere Metaanalyse unerlässlich sind.

Darüber hinaus wurden im Rahmen der Studienanalyse relevante Variablen wie das Studienjahr, die veröffentlichende Zeitschrift, die Patientenzahl, sowie die Untersuchungsmethode erfasst.

2.4.3 Korrelation und Metaanalyse

2.4.3.1 Korrelation

Eine Korrelation beschreibt allgemein formuliert eine Wechselbeziehung zwischen zwei Variablen, beziehungsweise vergleicht die Merkmale zweier Variablen. Dabei wird ermittelt, inwieweit diese in einer Beziehung stehen und sich in einer Messzahl, dem sogenannten Korrelationskoeffizienten, ausdrücken lassen. Grundsätzlich ist es so möglich, jede Variable oder auch Merkmale einer Variablen mit einander zu vergleichen. Sinnvoll ist dies allerdings nur, wenn im Voraus ein gewisser Zusammenhang ersichtlich ist, sodass entweder das eine Merkmal durch das andere verursacht oder zumindest mitverursacht wird oder beide auf eine identische Ursache zurückzuführen sind. Dabei ist die Korrelation allerdings streng von der Kausalität abzugrenzen, bei der im Gegensatz zur Korrelation ein bewiesener Ursache-Wirkungszusammenhang bestehen muss. Allerdings gibt die Korrelation durchaus einen Hinweis auf eine mögliche Kausalität. Darüber hinaus ist zu berücksichtigen, dass Korrelationen stets ungerichtet sind, das bedeutet, sie geben keinen Aufschluss darüber, welche Variable die andere bedingt. Lediglich die Stärke des Zusammenhangs kann über den bereits erwähnten Korrelationskoeffizienten, der Werte zwischen -1 und +1 annehmen kann, angegeben werden (Schwarz, 2016).

Der Wert 0 bedeutet reiner Zufall, die Variablen stehen in keinerlei Zusammenhang. Nimmt der Koeffizient Werte zwischen 0 und +1 an, bedeutet der Wert +1 eine perfekte positive Korrelation. Das bedeutet somit in genau dem gleichen Grad, in dem sich das eine Merkmal verändert, verändert sich auch das andere. Dahingegen zeigen die Werte im Bereich von 0 bis -1 eine negative Korrelation an. Hohe Werte der einen Variablen, gehen mit niedrigen Werten der anderen Variablen einher (Brückler, 2017). Diese Form des Korrelationskoeffizienten wird nach ihrem Entdecker als Pearson-Korrelation bezeichnet. Eine solche Korrelation ist allerdings

nur bei linearen Zusammenhängen anwendbar. Die unterschiedlichen Zusammenhänge werden in Abbildung 6 ersichtlich.

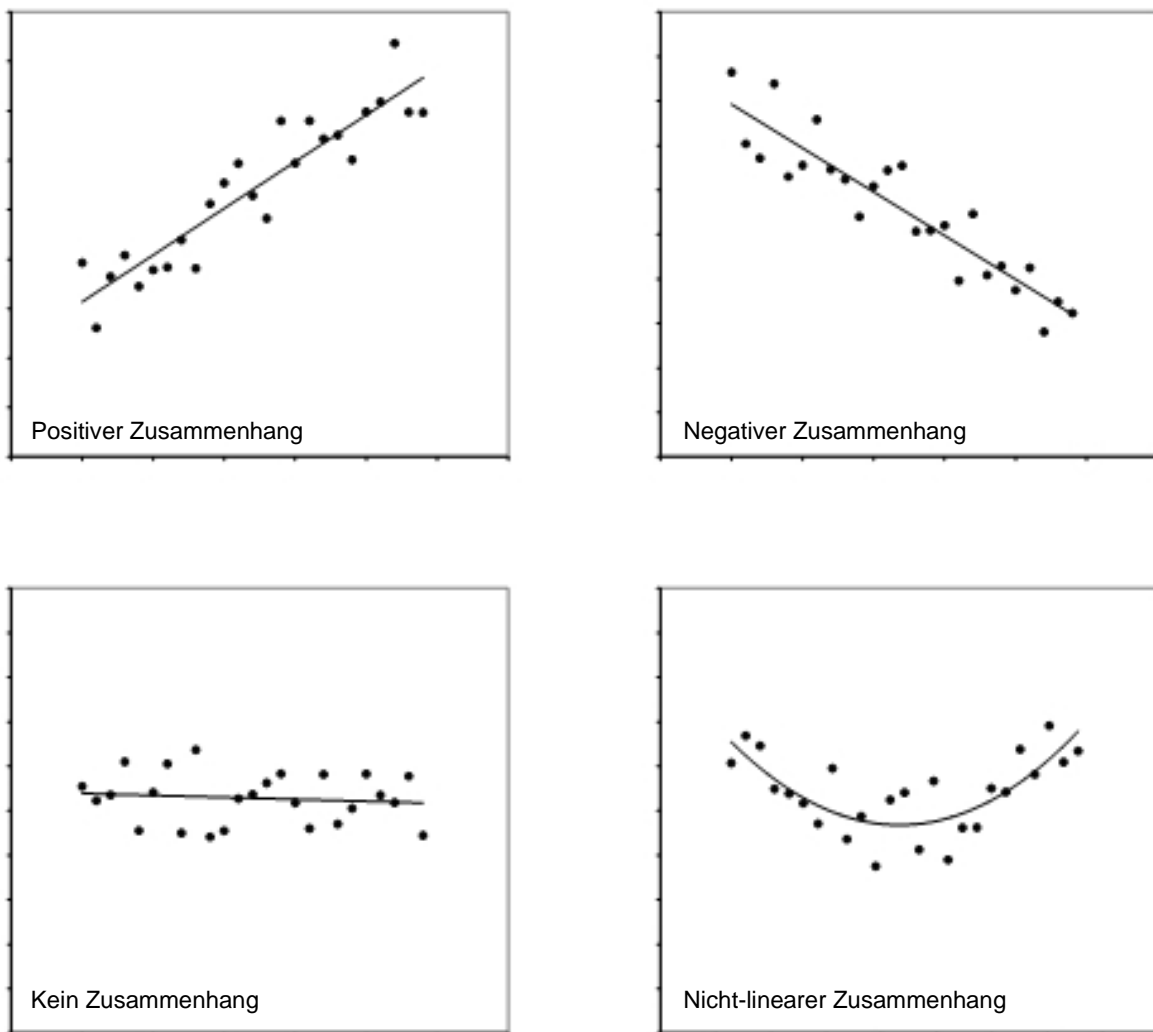


Abbildung 6: Varianten der möglichen Korrelationen (Schwarz, 2016)

Neben der Pearson-Korrelation, die hier der Vollständigkeit halber und für ein besseres Verständnis nicht außer Acht gelassen werden darf, besteht eine weitere Möglichkeit, das Korrelationsmaß anzugeben. Hierbei handelt es sich um ein parameterfreies Maß, den sogenannten Rangkorrelationskoeffizienten, bekannt

auch unter dem Namen der Spearman-Korrelation. Dieses Vorgehen wurde aus im folgenden aufgeführten Gründen für die vorliegende Arbeit gewählt.

Im Rahmen dieses nicht-parametrischen Verfahrens werden im Vergleich zu der Pearson-Korrelation deutlich geringere Anforderungen an die Verteilung der Messwerte gestellt. Die Daten müssen dabei nicht normalverteilt sein und auch kleine Stichproben und Ausreißer können mit dieser Methode erfasst werden (Schwarz, 2016). Vor allem aufgrund letzterer Vorteile wurde für die vorliegende Arbeit die Korrelation nach Spearman gewählt. Die Interpretation der Korrelationskoeffizienten erfolgt analog der Korrelation nach Pearson.

Hierbei ist zu beachten, dass zwar solche Daten wie das Publikationsjahr, der IF, der Bias und die Patientenzahl für jede Publikation gleich sind, sich für die Effektstärke und den Standardfehler der Effektstärke aber durchaus verschiedene unterschiedliche Werte pro Publikation ergaben. Somit wurden einerseits die Korrelationen jedes einzelnen Datensatzes errechnet und in den entsprechenden Grafiken dargestellt, sowie zusätzlich für die Effektstärke und die Standardfehler die Korrelationen mit deren gemittelten Werten bestimmt.

Für die vorliegende Arbeit wurden aus den ermittelten Datensätzen so folgende Korrelationen der zu untersuchenden 25 Artikel erstellt und bewertet:

Zum einen die Korrelation der Effektstärken mit der Patientenzahl und dem Standardfehler der Effektstärke, sowie zum anderen die Korrelation der Effektstärken mit dem Publikationsjahr und dem Journal Impact Factor, zusätzlich die Korrelation des Veröffentlichungsjahres mit der Patientenzahl. Zur Bewertung und Interpretation wurden folgende eigens gewählte Hypothesen aufgestellt:

1. Eine signifikante negative Korrelation der Effektgröße mit der Patientenzahl oder mit dem Standardfehler der Effektgröße würde auf einen Publikationsbias hinweisen.
2. Die Effektstärke korreliert mit dem Veröffentlichungsjahr negativ, was darauf hindeuten würde, dass in früheren Jahren größere Effekte erzielt wurden.

3. Wenn Journals mit höherem Ruf größere und damit zuverlässigere Studien publizieren als „durchschnittliche“ Zeitschriften, sollte die Stichprobengröße, also die Patientenanzahl positiv mit dem Journal Impact Factor korrelieren. Zudem kann angenommen werden, dass bei einer deutlichen negativen Korrelation des Standardfehlers mit der Patientenanzahl auch mit einer deutlich negativen Korrelation des Standardfehlers mit dem JIF zu rechnen ist.
4. Man kann annehmen, dass Ergebnisse, die zuerst in kleineren Stichproben ermittelt wurden, danach in größere Stichproben repliziert wurden. In diesem Fall sollte die Patientenanzahl positiv mit dem Veröffentlichungsjahr der entsprechenden Publikation korrelieren.

Die Berechnung der Korrelationen erfolgte mit dem Programme IBM SPSS Statistics 25. Hierbei können über den Menüpunkt „Analysieren“ die Optionen „Korrelationen“ und schließlich „Bivariant“ ausgewählt und die entsprechenden Korrelationskoeffizienten so ermittelt werden. Da die meisten o.g. Variablen, vor allem Stichprobengröße und Impact Factor, sehr schiefe Verteilungen haben, wurden Korrelationskoeffizienten nach Spearman berechnet. Trat der Fall auf, dass zu einer Studie mehrere Effektstärken und deren Standardfehler vorliegen, so wurde folgendermaßen vorgegangen: Es wurden die jeweilig einzelnen Korrelationen jeder Datensätze für sich gebildet und zusätzliche eine gemittelte Korrelation. Für den mittleren Standardfehler wird zunächst die Streuung entsprechend folgender Formel ermittelt.

$$\text{Streuung} = \text{SE}^2 * n$$

Aus der Streuung aller Datensätze einer Studie wird nun die mittlere Streuung gebildet. Der mittlere Standardfehler ergibt sich wie folgt:

$$\text{Mittlerer SE} = \sqrt{\frac{\text{mittlere Streuung}}{n}}$$

Die Richtung der Korrelation war im Voraus nicht bekannt, so musste die Signifikanz der Korrelation als „zweiseitig“ bewertet werden.

2.4.3.2 *Metaanalyse*

Die Metaanalyse ist eine Analysemethode, anhand derer quantitative Ergebnisse aus empirischen Untersuchungen zusammengefasst und untersucht werden können. Dabei handelt es sich um eine sogenannte Sekundäranalysemethode, die eine Integration quantitativer Resultate aus bereits existierenden Primäruntersuchungen ermöglicht (Hansueli Stamm, 1995).

Bei einer Primäruntersuchung handelt es sich um einen bereits vorhandenen, von anderen Untersuchungen unabhängigen Originalbeitrag (Fricke and Treinies, 1985). Wird ein solcher Originalbeitrag für sich alleine betrachtet, kann hier oftmals die Tatsache problematisch sein, dass die Stichprobengrößen zu klein gewählt wurden und dadurch lediglich vermeintlich signifikante Ergebnisse entstehen.

Im Gegensatz dazu sollen im Rahmen einer Sekundäruntersuchung die aus den Primäruntersuchungen bereits vorliegenden Forschungsergebnisse, zu einer oftmals neuen Fragestellung untersucht werden. Dabei wird ein Überblick über die bereits vorhandenen Resultate erstellt und zu einem allgemeingültigen Ergebnis zusammengeführt (Glass, 1976). Wie bei Hakim (1987) erläutert, sollen diese Sekundäranalysen in drei Subtypen untergliedert werden. Auf die beiden ersten Typen „Desk Research“ und „Research Review“ detailliert einzugehen, überschreitet den Rahmen der vorliegenden Arbeit. Bei der dritten Gruppe handelt es sich um die „Integrative Reviews“. Die Metaanalyse zählt zu dieser Klasse der Sekundäranalysen (Hakim, 2000) und soll deswegen an dieser Stelle kurz erläutert werden. Unter Berücksichtigung statistischer Signifikanzen und Effektstärken wird anhand der Integrative Reviews versucht Generalisierungen zu einem bestimmten Thema aus möglichst vielen empirischen Untersuchungen abzuleiten. Die Methode kann dann zur Anwendung kommen, wenn bereits zahlreiche

Primäruntersuchungen vorhanden sind und Zusammenhänge vermutet werden, aber noch nicht spezifisch sind (Hansueli Stamm, 1995).

Zusammenfassend umschreibt die in „Metaanalyse. Eine Einführung“ auf Seite 8 aufgeführte Definition nach Drinkmann sehr gut die Methodenfamilie der Metaanalyse:

„Metaanalyse soll sein: eine an den Kriterien empirischer Forschung orientierte Methode zur quantitativen Integration der Ergebnisse empirischer Untersuchungen sowie zur Analyse der Variabilität dieser Ergebnisse“ (Drinkmann, 1990).

Die Schritte zur Erstellung einer Metaanalyse decken sich im Grunde mit denen des systematischen Reviews, sodass dieses durchaus als Grundlage für eine Metaanalyse angesehen werden kann. Die Schritte zur Durchführung gestalten sich wie folgt (Hedges and Olkin, 1985):

1. Definition einer Fragestellung
2. Literaturrecherche
3. Statistische Auswertung
4. Interpretation

Die genaue Vorgehensweise ist mit der des systematischen Reviews identisch und wurde in den vorangehenden Abschnitten bereits beleuchtet.

Im Unterschied zum systematischen Review, kommen für die Metaanalyse allerdings nur solche Studien in Frage, welche den Untersuchungsgegenstand numerisch, also durch quantitative Messungen, erfassen.

Um eine Vergleichbarkeit der verschiedenen Studien zu ermöglichen, müssen wie bereits erläutert, die Effektstärken der einzelnen Studien errechnet werden und auf ein gemeinsames Effektstärkenmaß gebracht werden. Oftmals werden in den unterschiedlichen Primärstudien mehrere Effektstärken an der gleichen Stichprobe berechnet. Dabei ist es von enormer Bedeutung, erfolglose Resultate, soweit aufgeführt, nicht zu vernachlässigen und diese nicht-signifikanten Ergebnisse zu erhalten. Nur so ist es möglich, ein unverzerrtes Gesamtergebnis zu erzielen.

Allerdings wurden die einzelnen Effektstärken zu einer identischen Patientenzahl nicht gemittelt, sondern als selbstständige Ergebnisse betrachtet. Zudem verlangen die meisten Programme zur Berechnung der Metaanalyse zusätzlich den Standardfehler der Effektstärke. Dieser befindet sich in einem eindeutigen Verhältnis zum 95 % Konfidenzintervall:

$$(CI_h - CI_l) = 3,92 \times SE$$

wobei CI_h die höhere Grenze des Konfidenzintervalls, CI_l seine niedrigere Grenze, SE den Standardfehler darstellt.

Am Ende dieser Vorbereitungen werden in verschiedenen Metaanalysen, wie zum Beispiel in jener von Kotchoubey/Pavlov (Kotchoubey and Pavlov, 2018), Qualitätsschwellen eingeführt und die Studien, die diese Mindestanforderungen nicht erfüllen, werden aus der Metaanalyse ausgeschlossen. In dieser Arbeit wurden dafür zwei Mindestkriterien verwendet: Der Biaswert musste gleich oder größer als 3 sein (wie oben angeführt, je höher der Wert, umso kleiner der Bias) und die Gesamtzahl der Patienten musste mindestens 20 betragen. Bedenken wir, dass hier Studien analysiert wurden, die zwei diagnostische Gruppen (VS und MCS) vergleichen und dass diese Gruppen nur selten genau gleich sind, bedeutet $N = 20$, dass eine der zwei Gruppen oft nur 6 oder 7 Patienten einschloss. Das Kriterium $N = 20$ erscheint deshalb relativ liberal. Der Einschluss noch kleinerer Studien in die Metaanalyse wäre wohl nicht sinnvoll.

Die nach diesen Kriterien in Frage kommenden Studien wurden für eine noch differenziertere Betrachtung und Analyse in drei Gruppen unterteilt. Hierzu wurde nach den in den Primärstudien angewandten Untersuchungsmethoden unterschieden. Darin kamen entweder Resting-State, Passive-State oder Aktive-State-Methoden zur Anwendung.

Nach dieser Vorauswahl wurden insgesamt folgende 14 Studien unterteilt nach den Kategorien „Resting State“, „Passive State“ und „Aktive State“ für die Metaanalyse ermittelt:

Resting State:

1. Quantitative rates of brain glucose metabolism distinguish minimally conscious from vegetative state patients. (Stender et al., 2015)
2. Multimodal Study of Default-mode network Integrity in disorders of consciousness. (Rosazza et al., 2016)
3. Diffusion weighted imaging distinguishes the vegetative state from the minimally conscious state. (Fernandez-Espejo et al., 2011)
4. A role for the Default Mode Network in the Base of Disorders of Consciousness. (Fernandez-Espejo et al., 2012)
5. Degree of functional connectome abnormality in disorders of consciousness. (Sinitsyn et al., 2018)

Passive State:

1. Central olfactory processing in patients with disorders of consciousness. (Nigri et al., 2016)
2. fMRI activation with an "affective speech" paradigm in vegetative and minimally conscious States: applicability and prognostic value. (Piperno et al., 2012)
3. Deactivation of the default mode network as a marker of impaired consciousness: an fMRI study. (Crone et al., 2011)
4. True or False? Activations of Language-related Areas in Patients with Disorders of Consciousness. (Kotchoubey et al., 2014)
5. Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. (Coleman et al., 2009)
6. Brain response to thermal stimulation predicts outcome of patients with chronic disorders of consciousness. (Li et al., 2015)

Aktive State:

1. Willful modulation of brain activity in disorders of consciousness. (Monti et al., 2010)
2. Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: a clinical validation study. (Stender et al., 2014)
3. Can mental imagery functional magnetic resonance imaging predict recovery in patients with disorders of consciousness? (Vogel et al., 2013)

Zur Erstellung der Metaanalyse wurde das frei erhältliche Analyseprogramm „JASP“ genutzt (Team, 2018). Hier kann unter Angabe der jeweiligen Effektstärke und deren Standardfehler eine Metaanalyse kalkuliert werden. Zusätzliche optionale Angaben zur Patientenanzahl und der Artikeltitel können zu Beginn der Analyse eingefügt werden. Allerdings ist eine direkte Eingabe des Datensatzes in JASP nicht möglich. Dieser muss bereits als Excel- oder SPSS-Datei vorliegen und kann erst dann in JASP eingefügt und analysiert werden.

Unter dem Menüpunkt „Statistic“ ist es möglich, zusätzlich das 95 %-Konfidenzintervall auszuwählen. Wie für eine Metaanalyse üblich werden auch hier sowohl die Ergebnisse der Einzelstudien als auch der zusammengefasste Effekt in einem sogenannten Forest-Plot dargestellt. Da innerhalb der untersuchten Studiengruppen eine gewisse systematisch bedingte Heterogenität zu erwarten war, wurde hier mit einem Random-Effects-Modell gearbeitet. Hierbei wird die Variabilität zwischen den Primärstudien durch eine Verbreiterung des Konfidenzintervalls berücksichtigt. Dies führt somit in der Regel zu einer vorsichtigen und zurückhaltenden Einschätzung. Eine genaue Untersuchung der Heterogenität kann rein visuell an Hand des Forest Plot oder auch detaillierter mithilfe statistischer Methoden erfolgen. Das am häufigsten und somit auch hier verwendete in Prozentwerten von 0 % – 100 % angegebene Heterogenitätsmaß ist das von Cochran's Q abgeleitete I^2 . Von einer signifikanten Heterogenität der Primärstudien

kann ab einem Wert von $I^2 > 50\%$ ausgegangen werden (Higgins et al., 2003). Als ein weiterer Vorteil von I^2 können die bereits erläuterte einfache Interpretierbarkeit und die Unabhängigkeit von der Anzahl an in die Metaanalyse eingeschlossener Studien angeführt werden (Higgins and Thompson, 2002).

3 Ergebnisse

3.1 Ergebnisse der Qualitätsbeurteilung und der Korrelation

Die folgenden dargestellten Ergebnisse und Grafiken beruhen auf den in Tabelle 4 und Tabelle 5 aufgezeigten Daten. Aufgrund teilweise sehr ähnlicher Werte kann es in den Grafiken dazu kommen, dass sich einzelne Datenpunkte überlagern und so der Anschein entsteht, nicht alle Werte seien wiedergegeben worden.

Zudem sei darauf hingewiesen, dass die Korrelationsgeraden und-werte auf solchen Schaubildern verzerrt sein können, weil sie durch Ausreißer und nicht-normale Verteilungen beeinflusst werden können. Dagegen sind die im Text angegebenen Werte (Spearman's ρ) unverzerrt und von den Ausreißern nicht beeinflusst.

<i>Artikel</i>	<i>Patientenanzahl</i>	<i>Impact Faktor</i>	<i>Jahr</i>	<i>Bias</i>
Bick et al. 2013	11	2,26	2013	0,00
Monti et al, 2015	24	8,18	2015	2,00
Monti et al. 2010	46	53,48	2010	3,00
Stender et al 2015	35	5,40	2015	5,50
Fernandez-Espejo et al. 2010	9	4,29	2010	3,00
Rosazza et al. 2016	108	9,89	2016	4,50
Nigri et al. 2016	42	3,95	2016	3,00
Stender et al. 2014	122	45,22	2014	4,50
Kotchoubey et al. 2013	12	3,58	2013	2,00
Piperino et al. 2012	24	2,70	2012	5,00
Crone et al. 2011	25	4,09	2011	4,00
Moreno et al. 2010	8	8,01	2010	2,00
Fernandez-Espejo et al 2011	25	5,93	2011	4,00
Fernandez-Espejo et al. 2008	6	1,11	2008	2,00
Di et al. 2007	11	5,69	2007	2,00
Fernandez-Espejo et al. 2012	46	11,19	2012	3,00
Nigir et al. 2017	11	3,98	2017	3,50
Kotchoubey et al. 2014	55	3,45	2014	5,00
Coleman et al. 2009	41	9,49	2009	4,00
Vogel et al. 2013	22	2,44	2013	4,00
Li et al. 2015	22	3,09	2015	4,00
Sinitsyn et al. 2018	22	3,14	2018	3,50
Bodien et al. 2018	9	3,50	2017	2,50
Edlow et al. 2017	10	10,84	2017	3,50
Marino et al. 2017	50	5,42	2017	2,00

Tabelle 4: Daten der relevanten 25 Artikel

<i>Artikel</i>	<i>Effektstärke</i>	<i>Standardfehler</i>
Bick et al. 2013	0,12	0,86
Bick et al. 2013	-0,38	0,77
Bick et al. 2013	0,22	0,67
Bick et al. 2013	0,38	0,77
Bick et al. 2013	0,61	0,75
Monti et al, 2015	0,22	0,59
Monti et al. 2010	-1,01	0,64
Monti et al. 2010	-0,82	0,60
Stender et al 2015	1,01	0,36
Stender et al 2015	1,14	0,36
Stender et al 2015	1,38	0,37
Stender et al 2015	0,35	0,34
Stender et al 2015	0,24	0,33
Stender et al 2015	0,98	0,50
Fernandez-Espejo et al. 2010	1,38	0,37
Fernandez-Espejo et al. 2010	0,98	0,70
Rosazza et al. 2016	0,82	0,19
Rosazza et al. 2016	0,56	0,21
Rosazza et al. 2016	0,95	0,23
Nigri et al. 2016	0,73	0,32
Nigri et al. 2016	0,44	0,31
Stender et al. 2014	1,52	0,22
Stender et al. 2014	0,56	0,18
Kotchoubey et al. 2013	1,37	0,68
Kotchoubey et al. 2013	1,48	0,76
Kotchoubey et al. 2013	1,41	0,70
Kotchoubey et al. 2013	1,37	0,68
Kotchoubey et al. 2013	1,43	0,71
Kotchoubey et al. 2013	1,33	0,68
Kotchoubey et al. 2013	1,36	0,68

Piperino et al. 2012	0,41	0,41
Crone et al. 2011	0,27	0,53
Moreno et al. 2010	0,15	0,84
Fernandez-Espejo et al 2011	0,39	0,40
Fernandez-Espejo et al 2011	0,42	0,40
Fernandez-Espejo et al 2011	0,57	0,57
Fernandez-Espejo et al. 2008	0,00	0,95
Di et al. 2007	0,77	0,64
Fernandez-Espejo et al. 2012	0,69	0,36
Nigir et al. 2017	0,98	0,81
Kotchoubey et al. 2014	-0,38	0,33
Coleman et al. 2009	0,98	0,34
Vogel et al. 2013	0,43	0,43
Li et al. 2015	-0,43	0,43
Sinitsyn et a.l 2018	1,66	0,06
Sinitsyn et a.l 2018	1,04	0,12
Bodien et al. 2017	0,38	0,81
Bodien et al. 2017	0,15	0,44
Edlow et al. 2017	0,64	1,17
Edlow et al. 2017	0,42	1,16
Edlow et al. 2017	0,20	1,15
Marino et al. 2017	0,15	0,31
Marino et al. 2017	0,46	0,41

Tabelle 5: Effektstärken und Standardfehler der jeweiligen Datensätze

3.1.1 Patientenanzahl

Die durchschnittliche untersuchte Patientenzahl pro Studie beläuft sich auf einen Mittelwert von 31 ± 29 und einer Mediane von 22 mit einer Bandbreite von 6 bis 122 untersuchten Patienten. Daraus ergibt sich für die vorliegende Arbeit eine Gesamtzahl von 796 untersuchten Probanden. Dabei schließen 4 Studien weniger als 10 Patienten ein, 6 Studien untersuchen zwischen 10 und 20 Patienten und eine deutliche Mehrheit von 15 Studien beinhaltet eine Patientensumme von mehr als 20. Wie in den allgemeinen Annahmen gefordert korreliert der Standardfehler signifikant negativ mit der Stichprobengröße, also der Patientenzahl (Spearman's $\rho(\text{einzel}) = -0,79$ / Spearman's $\rho(\text{gemittelt}) = -0,80$).

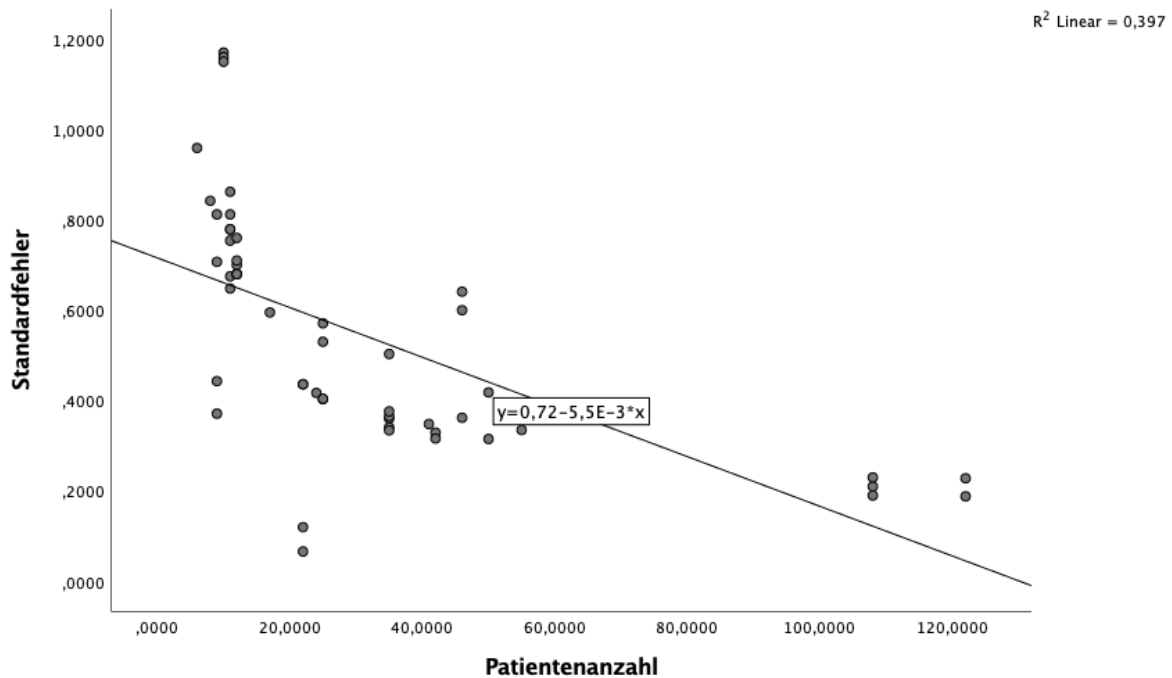


Abbildung 7: Korrelation der Patientenzahl mit dem Standardfehler der Effektstärke

3.1.2 Journal Impact Factor

Mithilfe der Online Datenbank InCite Journal Citation Report (Analytics, 2018) konnte der Impact Factor der jeweiligen Zeitschriften ermittelt werden.

Der mittlere Impact Factor beläuft sich auf einen Wert von $8,81 \pm 12,5$ mit einer Spannweite von 1 bis 53,5. Die Mediane liegt bei 4,29. Von den betrachteten Artikeln befinden sich 13 in einem Bereich, in dem der Impact Factor Werte zwischen 1 und 5 annimmt, 8 Artikel erhalten Werte zwischen 5 und 10 und 4 Artikel wiederum erzielen Punktwerte größer 10.

Die Hypothese, dass Studien mit einer vergleichbar höheren Anzahl an Teilnehmern in renommierten Journals, also solchen mit höherem Impact Factor, publiziert werden, konnte bestätigt werden. Der Impact Factor korreliert positiv mit der Patientenzahl (Spearman's $\rho = 0,43$) und zeigt hier auf dem 0,05-Niveau eine deutliche Signifikanz. Als weniger deutlich, aber dennoch tendenziell erkennbar erweist sich die Korrelation des Standardfehlers mit dem JIF (Spearman's $\rho(\text{einzel}) = 0,10$ / Spearman's $\rho(\text{gemittelt}) = 0,11$)

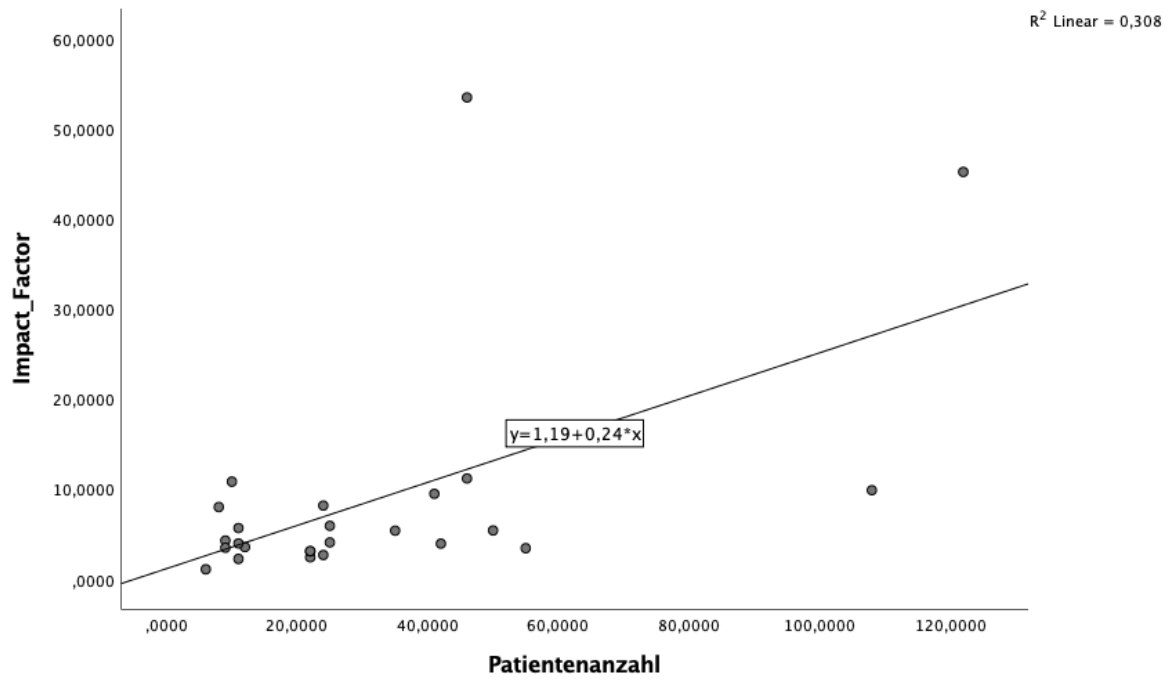


Abbildung 8: Korrelation Impact Faktor mit Patientenzahl

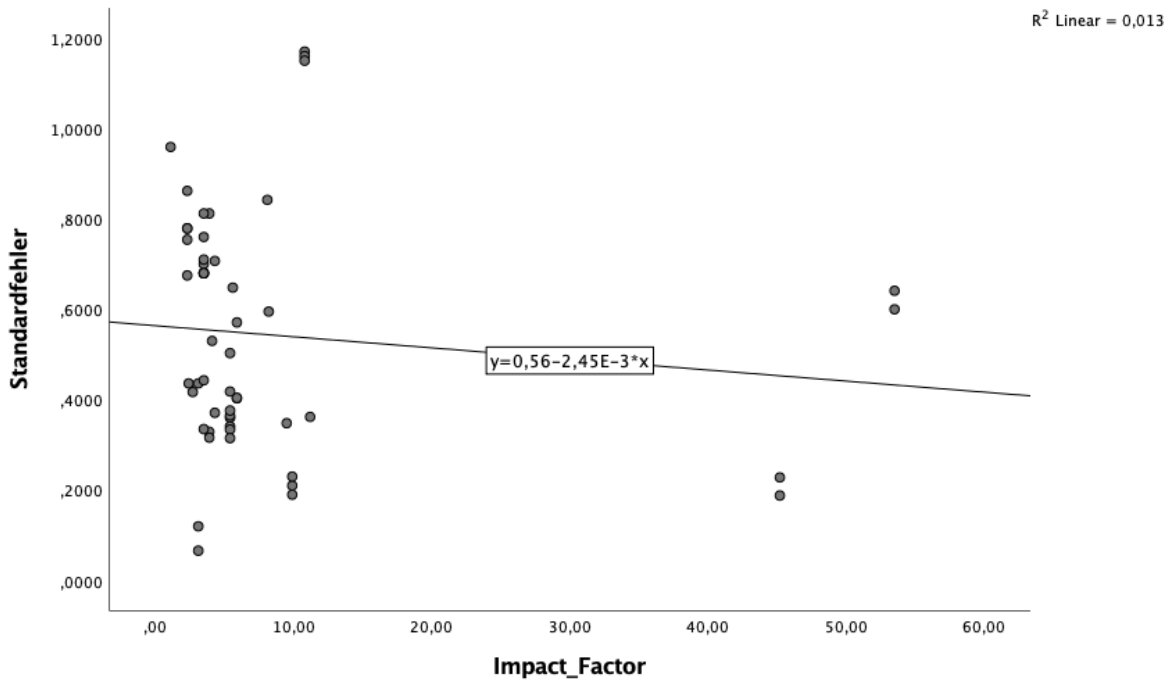


Abbildung 9: Korrelation Impact Faktor mit dem Standardfehler der Effektstärke

3.1.3 Erscheinungsjahr

Die untersuchten Artikel bewegen sich in einem Veröffentlichungszeitraum zwischen 2006 und 2018 (mittleres Veröffentlichungsjahr $2012 \pm 3,0$ Jahre/Mediane 2013). Wie in den selbst gewählten Hypothesen antizipiert, gibt es Grund zu folgender Annahme: Je aktueller die veröffentlichte Publikation ist, desto größer ist die untersuchte Probandengruppe oder desto kleiner ist die ermittelte Effektstärke (Brembs et al., 2013). Der Zusammenhang zwischen dem Veröffentlichungsjahr und der Patientenzahl ist zwar erkennbar, aber sehr schwach und nicht signifikant (Spearman's $\rho = 0,16$). Die Korrelation mit der Effektstärke liegt ebenfalls sehr nahe bei 0 (Spearman's $\rho(\text{einzel}) = 0,02$ / Spearman's $\rho(\text{gemittelt}) = 0,14$). Somit lässt sich für das Kriterium Erscheinungsjahr feststellen, dass dieses augenscheinlich nicht mit den erzielten Effekten und auch nicht mit der Größe der untersuchten Patientengruppe im Zusammenhang steht.

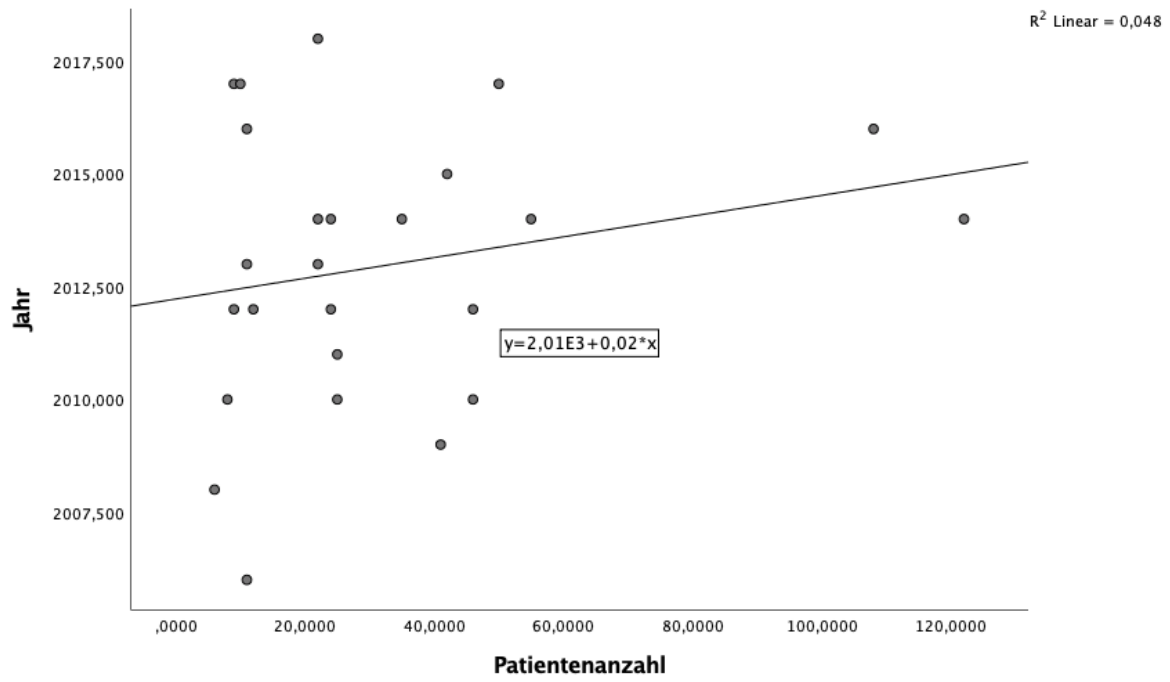


Abbildung 10: Korrelation Erscheinungsjahr mit Patientenanzahl

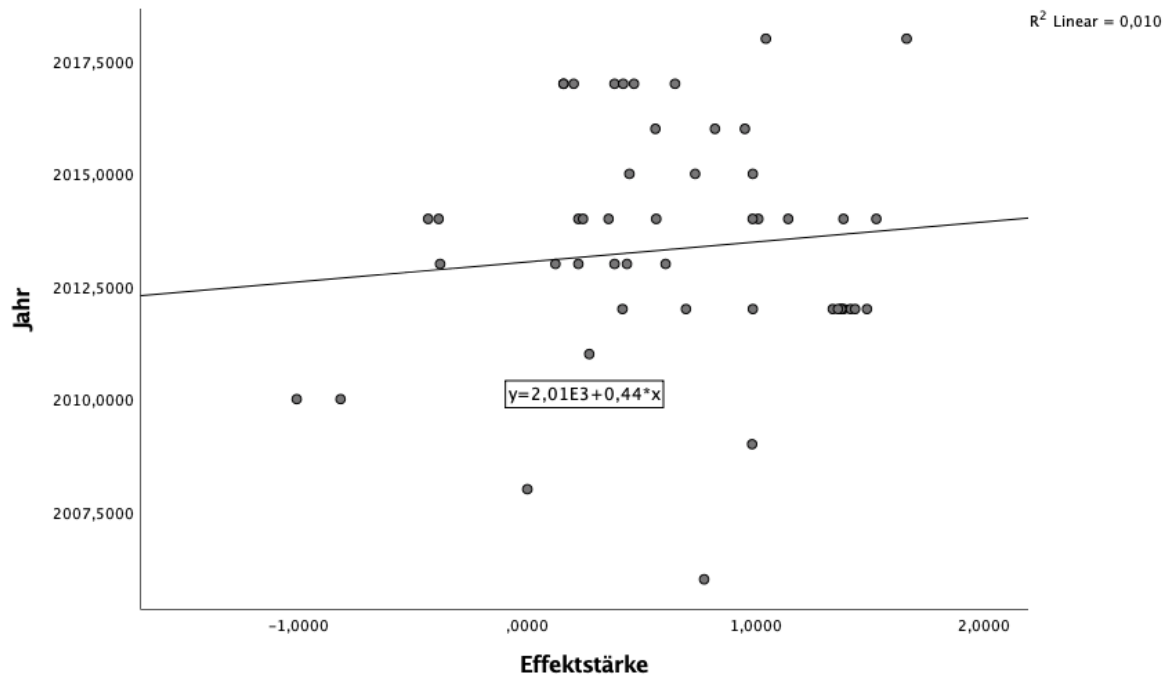


Abbildung 11: Korrelation Erscheinungsjahr mit Effektstärke

3.1.4 Mögliche Verzerrungseffekte

Die Korrelation zwischen der Effektgröße und der Patientenzahl liegt dicht bei 0 (Spearman's $\rho(\text{einzel}) = 0,05$ / Spearman's $\rho(\text{gemittelt}) = 0,07$). Die negative Korrelation zwischen der Effektstärke und dem Standardfehler der Effektstärke geht zwar in die aufgrund der Biashypothese erwartete Richtung, erreicht aber nicht das 5 %-Signifikanzniveau (Spearman's $\rho(\text{einzel}) = -0,15$ / Spearman's $\rho(\text{gemittelt}) = -0,28$). Somit gibt es keinen überzeugenden Hinweis auf das Vorhandensein einer Publikationsverzerrung. Deutlich wird dies auch in den folgenden Punktdiagrammen.

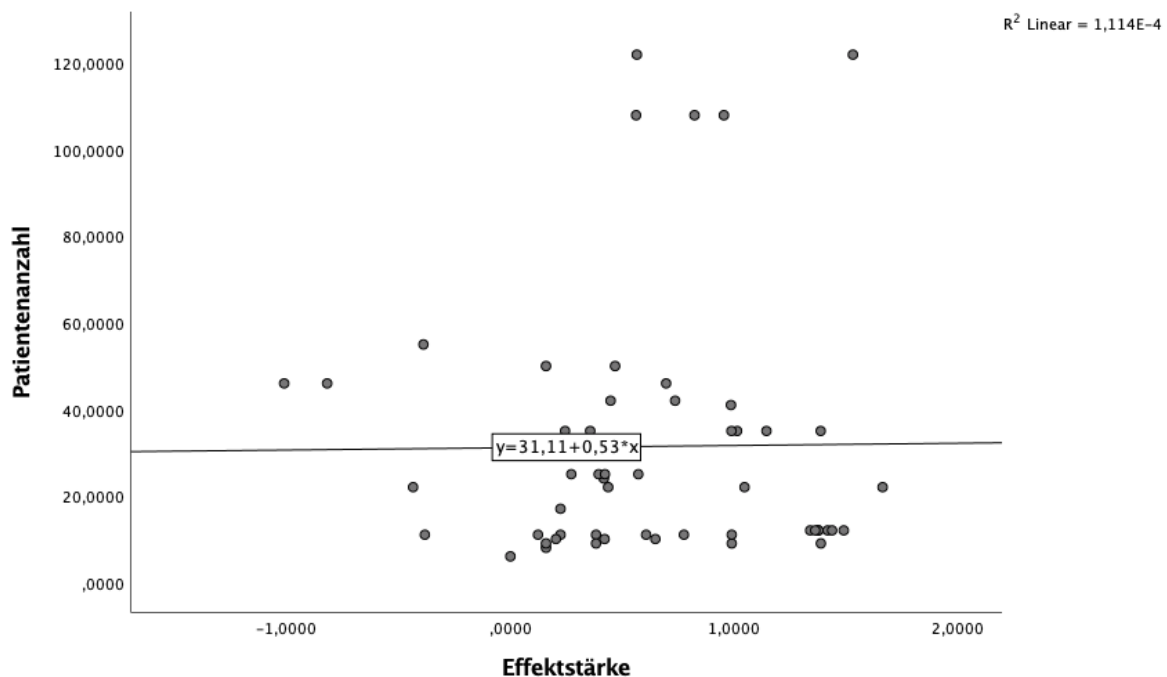


Abbildung 12: Korrelation Patientenzahl mit Effektstärke

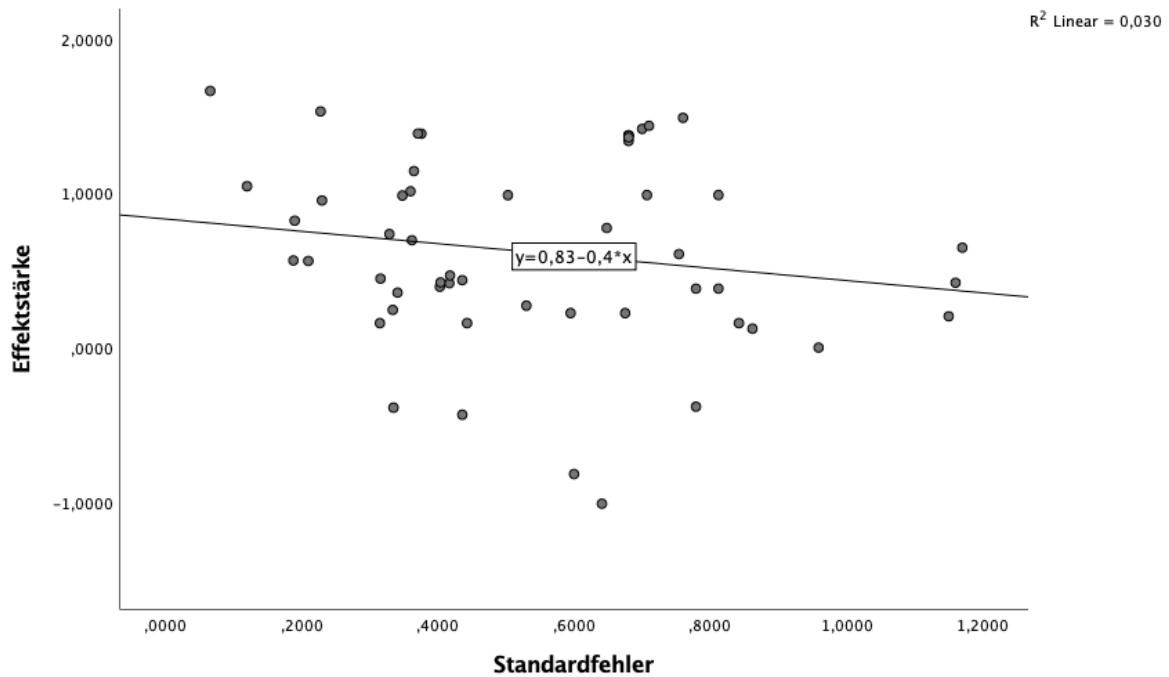


Abbildung 13: Korrelation Standardfehler mit Effektstärke

Eine detailliertere Beurteilung des Bias ermöglicht das Prinzip nach QUADAS-2. Für die einzelnen oben erläuterten Kriterien, „Patientenauswahl“, „Index“, „Reference Standard“ und „Flow und Timing“ ergaben sich nach genauer Analyse die in Tabelle 6 dargestellten Mittelwerte und Mediane. Die Bewertungen des Autors und die der unabhängigen Zweitbewerterin korrelieren mit einem guten Wert von $\rho = 0,71$ deutlich positiv.

		Patientenauswahl	Indextest	Reference Standard	Flow und Timing	Summe
N	Gültig	25	25	25	25	25
	Fehlend	0	0	0	0	0
Mittelwert		1,36	,18	,98	,74	3,26
Median		1,50	,00	1,00	1,00	3,50
Std.-Abweichung		,700	,35	,36	,66	1,25
Minimum		,00	,00	,00	,00	,00
Maximum		2,00	1,00	2,00	2,00	5,50

Tabelle 6: Darstellung der Analyseergebnisse des Bias nach dem QUADAS-2-Prinzip

Wie den Werten in Tabelle 6 zu entnehmen ist, kann durchaus von einem gewissen Bias ausgegangen werden. Obwohl ein Maximalwert von 5,5 für einen geringen Bias und folglich für eine hohe Studienqualität spricht, steht dem ein Mittelwert von lediglich 3,26 entgegen.

Zudem werden die Bias-Daten bezüglich folgender selbst gesetzter Hypothesen überprüft:

1. Je später der Veröffentlichungszeitpunkt, desto höher ist die Qualität der Studie anzusehen und somit um so geringer der Bias
2. Je höher der Impact Faktor, desto geringer ist der Bias
3. Je größer die Anzahl an Probanden, desto geringer ist der Bias

Die beiden Korrelationen zu Punkt 1 und Punkt 2 gingen zwar in die erwartete Richtung, waren aber nicht signifikant (Bias mit Veröffentlichungsjahr: $\rho = 0,16$; Bias mit Impact Factor: $\rho = 0,07$). Somit muss basierend auf dem hier betrachteten Zeitraum von einer weitgehend konstanten Qualität der Arbeiten ausgegangen werden. Auch der Unterschied zwischen den Zeitschriften mit höherem versus mittlerem Ruf zeigt keine Wirkung.

Deutlicher fällt der Vergleich zwischen den Biasdaten und der Patientenzahl aus. Die Korrelation zwischen diesen beiden Datensätzen beträgt $\rho = 0,51$. Je mehr Patienten untersucht wurden, desto höher sind die Werte für den Bias (d.h. desto geringer die Verzerrung).

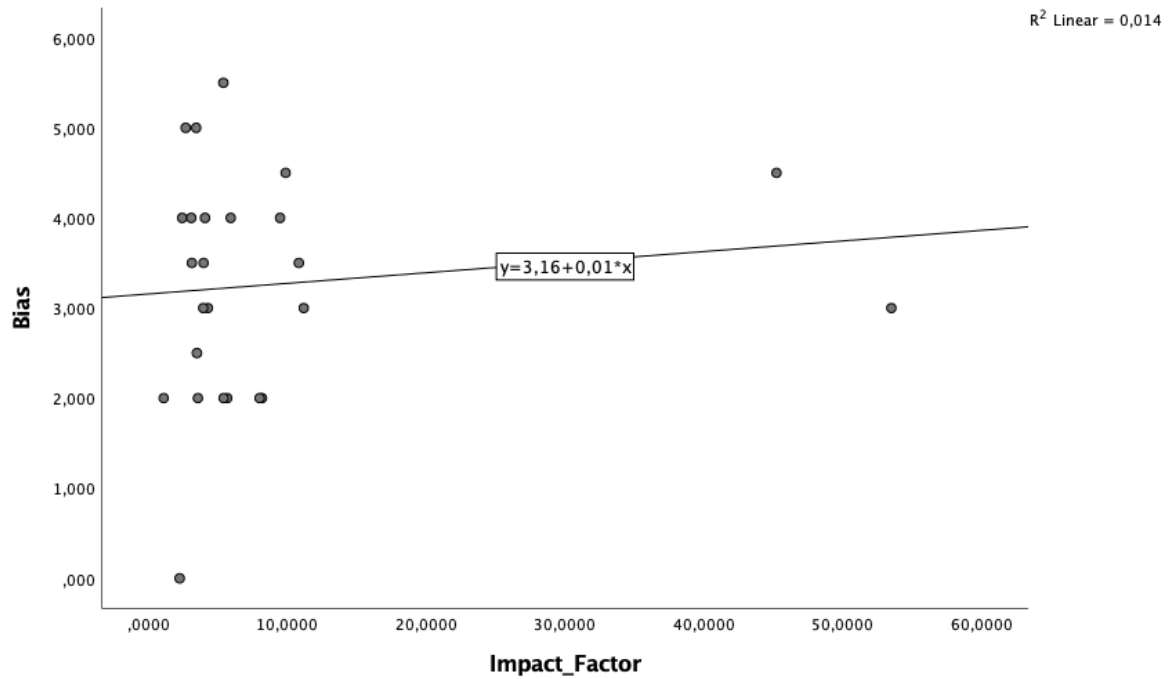


Abbildung 14: Korrelation Bias mit Impact Faktor

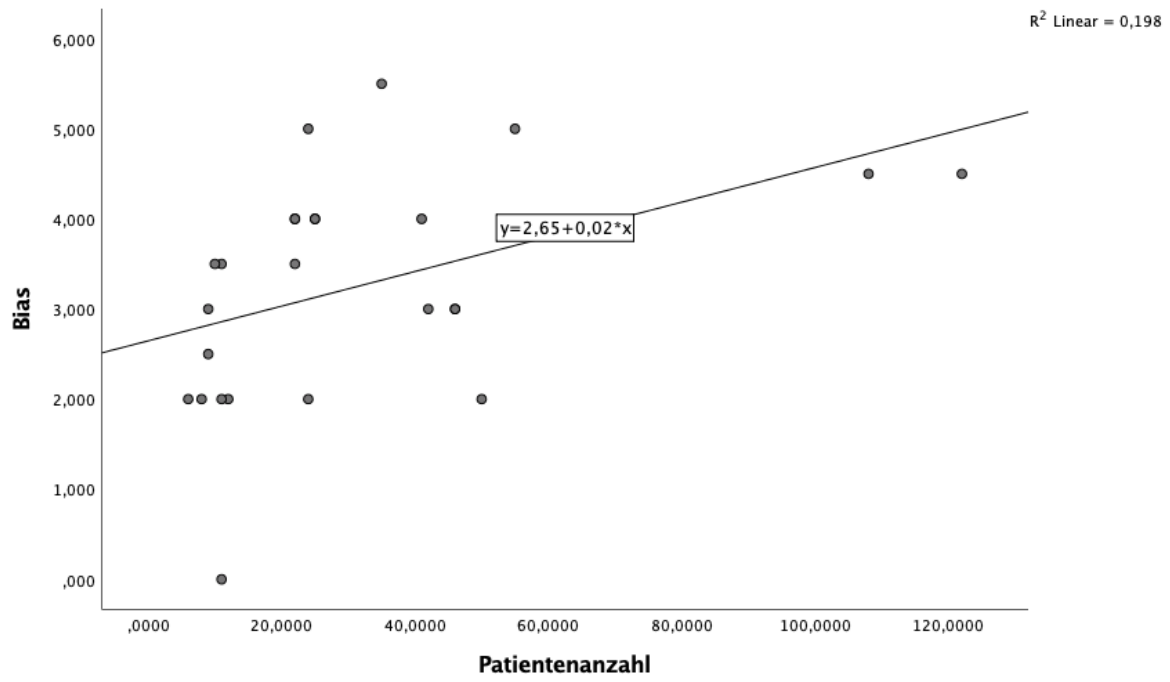


Abbildung 15: Korrelation Bias mit Patientenanzahl

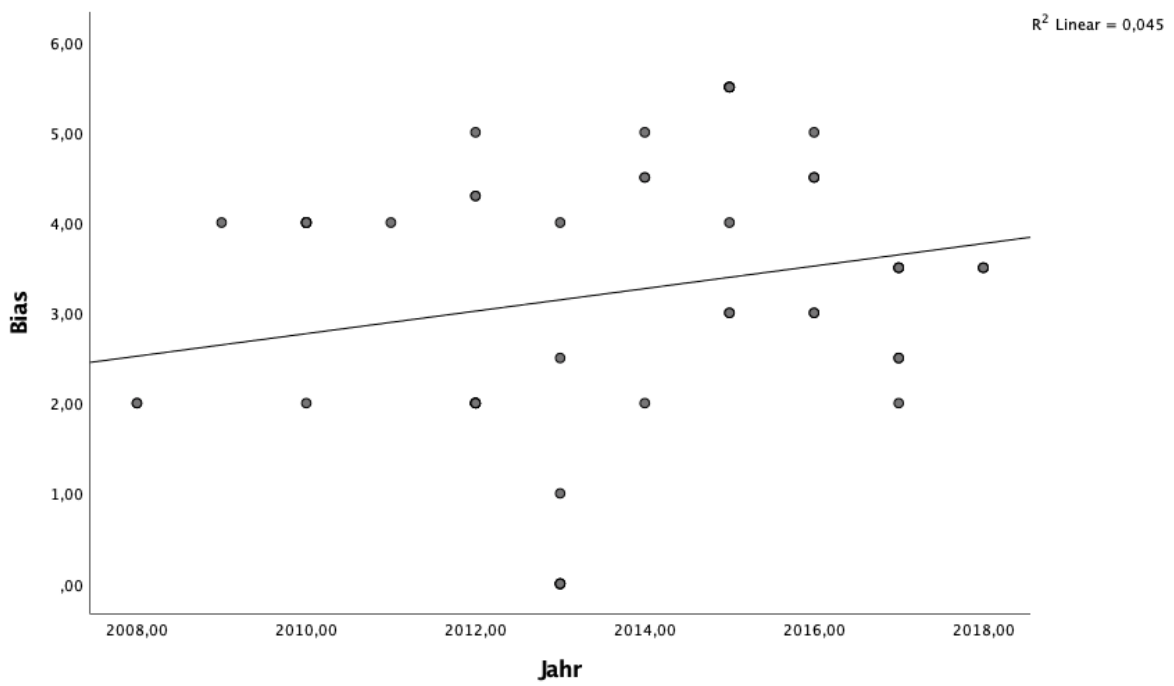


Abbildung 16 Korrelation Bias mit Erscheinungsjahr

3.2 Ergebnisse des systematischen Reviews

Hauptziel der vorliegenden Arbeit ist es, eine zuverlässige Unterscheidungsmethode zwischen dem MCS und VS mithilfe funktioneller bildgebender Verfahren zu ermitteln. Hierfür wurden, wie im Abschnitt zur Literaturrecherche ausführlich erläutert, aus einer Gesamtzahl von 124 relevant erscheinenden Artikeln, 39 als tatsächlich für diese Arbeit bedeutend identifiziert, wovon 25 wiederum mit den entsprechenden empirischen Angaben für die weitere Analyse von Bedeutung waren. Wie bereits aus Punkt 3.1.1 hervorgeht, reicht die Spannweite der untersuchten Patientenzahlen von lediglich 6 Fällen bis hin zu großen multizentrischen Studien mit bis zu 122 Fällen. Sinnvoll erscheint es hierbei, Untersuchungen, die mehr als 10 Patientenfälle betrachten, einzubeziehen, um eine verlässliche Aussage über die Unterscheidungsfähigkeit der angewandten Untersuchungsmethode bezüglich MCS und VS zu erhalten. Wird dieses Raster auf die 25 für diese Arbeit relevanten Artikeln angewandt, wird deutlich, dass lediglich folgende 4 Studien diesem Kriterium nicht entsprechen und eine Patientengruppe von weniger als 10 einschließen:

1. Reductions of thalamic volume and regional shape changes in the vegetative and the minimally conscious states (Fernandez-Espejo et al., 2010)
2. A network approach to assessing cognition in disorders of consciousness (Rodriguez Moreno et al., 2010)
3. Cerebral response to speech in vegetative and minimally conscious states after traumatic brain injury (Fernandez-Espejo et al., 2008)
4. Functional MRI motor Imaginary Task to detect command following in traumatic (Bodien et al., 2017)

In 64 % (25 Artikeln) der insgesamt 39 betrachteten relevanten Artikeln wurde mit den in den Studien angegebenen Werten eine Effektstärke errechnet. Im Falle der übrigen 14 Artikel konnte entweder aufgrund fehlender Angaben keine Effektstärke

errechnet werden oder es handelte sich bereits um Übersichtsarbeiten. Allerdings kommen in den meisten der betrachteten 25 Artikel mehrere Untersuchungsmethoden zur Anwendung. Dadurch übersteigt die Gesamtzahl an zu beurteilenden Datensätzen und Effektstärken die der vorhandenen Publikationen und es ergibt sich eine Summe von 53 Datensätzen aus den bewerteten 25 Artikeln. Aus diesen 53 Datensätzen geht eine mittlere Effektstärke nach Cohen von $d = 0,61$ und ein gewichteter Mittelwert von $d = 0,52$ hervor. Entsprechend der bereits erläuterten Interpretation von Cohen kann somit von einem mittleren Effekt ausgegangen werden. Werden lediglich die Studien mit einer Patientenanzahl von mehr als 10 betrachtet, ändert sich die mittlere Effektstärke kaum und es ergibt sich ein Wert von $d = 0,77$.

Zudem muss hinzugefügt werden, dass nicht, wie zu erwarten wäre, in jedem Fall Patienten in einem MCS bessere Untersuchungsergebnisse aufweisen als die in einem VS. In 10 % der untersuchten Datensätze (nicht Studien) schneiden Patienten in einem VS besser ab als die in einem MCS. Diese 10 % aller untersuchten Datensätze liegen ausschließlich innerhalb der Menge der Daten bezogen auf die fMRT-Untersuchungen. Keine der PET-Analysen kam zu dem Ergebnis, dass ein Patient im VS bessere Resultate erzielt als ein vergleichbarer im MCS.

Selbstverständlich müssen die jeweils angewandten Untersuchungsmethoden nicht nur in ihrer Gesamtheit, sondern zudem auch entsprechend ihrer diagnostischen Aussagekraft getrennt voneinander bewertet werden. Werden alle gesammelten Effektstärken gemeinsam betrachtet, ergibt sich, wie bereits erwähnt, ein mittlerer Effekt von $0,61 \pm 0,60$ und ein entsprechender Median von 0,56. Die untersuchten Methoden zur Diagnostik der verschiedenen DoC können also als durchaus zuverlässig angesehen werden. Von den insgesamt 53 untersuchten Datensätzen beruhen lediglich 8 auf der Methode der PET-Untersuchung, die übrigen auf der fMRT. Diese 8 Datensätze zur Diagnostik mit der PET ergeben einen mittleren Effekt von $0,95 \pm 0,45$ und einen Median von 0,99. Damit kann entsprechend der Cohen Interpretation von einem großen Effekt ausgegangen werden. Betrachtet man die

gewonnenen Daten in Bezug auf die fMRT-Untersuchungen alleine, so erhält man eine mittlere Effektstärke von $0,69 \pm 0,44$ und einen Median von 0,56.

Neben der beschriebenen Betrachtung der Gesamtdatenzahl ist zusätzlich eine differenziertere Betrachtung der Einzeldatensätze notwendig. Dabei ergaben sich für 24 (45 %) der 53 Datensätze ein großer Effekt für die Unterscheidungsmöglichkeit zwischen dem MCS und VS, für 8 (15 %) ein mittlerer Effekt und für 21 Datensätze (39 %) ein kleiner Effekt. Innerhalb der 4 Studien, die die selbst gewählte Mindestanzahl von 10 Patienten unterschreiten, zeigt nur die Studie von Fernandez-Espejo et al., 2010 einen großen Effekt. Die übrigen drei Studien lassen mit Effektstärken von teilweise deutlich unterhalb der Marke $d = 0,4$ keine Signifikanz erkennen.

Die gefundenen 24 Datensätze mit einem nach Cohen hohen Effekt verteilen sich auf insgesamt 9 Studien. Zu bemerken ist, dass diese Studien eine mittlere Patientenzahl von 44 und eine Gesamtpatientenzahl von 406 aufweisen. Damit zeigen von den untersuchten 25 Studien (nicht Datensätzen) 36 % ein hohes Effektmaß. Auf diesen Anteil der Studien entfallen 51 % der in Summe betrachteten Patienten. Im Vergleich hierzu steht eine mittlere Patientenzahl von 20,8, bei solchen Studien mit einer geringen Aussagekraft, die also ein Effektstärkenmaß unterhalb der 0,4 Grenze aufweisen. Dies bestätigt entgegen der bereits erläuterten fehlenden Korrelation zwischen der Effektgröße und der Patientenzahl die Annahme, dass signifikante Unterschiede eher in Arbeiten mit größeren Probandengruppen zu finden sind, denn in solchen mit geringer Patientenzahl. Ein Zusammenhang zwischen der Effektgröße und den durchführenden Forschergruppen konnte nicht hergestellt werden. Der Grund hierfür ist darin zu sehen, dass stets verschiedene Forschergruppen wenige Publikationen zu vorliegendem Thema publizierten. Zu nennen sind hier die Autorengruppen um Monti, Stender, Peri, Kotchoubey, Fernandez, Nigri und Dimitry mit mehr als zwei Veröffentlichungen, zudem die für diese Arbeit herangezogenen Artikel.

3.3 Ergebnisse der Metaanalyse

Entsprechend der oben bereits erläuterten Daten und der Vorgehensweise wurde die Metaanalyse getrennt für Aktive-, Passive- und Resting-State-Untersuchungsmethoden durchgeführt. Dabei wurden innerhalb der 14 für die Metaanalyse geeignete Studien 27 Datensätze untersucht. Wird die Summe der Patienten eines jeden Datensatzes ermittelt, ergibt dies eine betrachtete Gesamtpatientenzahl von 1649. Dabei entfallen 258 Patienten und 5 Datensätze auf die aktiven Untersuchungsmethoden, 251 Patienten und 7 Datensätze auf die passiven Untersuchungsmethoden und 699 Patienten mit 15 Datensätzen auf die Resting-State-Methode. Werden bereits vorab an dieser Stelle die Ergebnisse der Homogenitätsprüfung betrachtet, wird die Notwendigkeit der Verwendung der Random-Effects Methode deutlich. Für die Aktive-State-Methode ergibt sich ein I^2 von 89 %. Für die passiven Untersuchungsmethoden ergibt sich ein Wert für I^2 mit 58 %, für die Studien zur Resting-State-Untersuchungsmethode ein I^2 von 73 %.

3.3.1 Aktive Aufgaben

Der Abbildung 17 kann der Forest Plot des Random Effect-Modells zur Analyse der aktiven Untersuchungsmethoden entnommen werden. Den Symbolen der Studien zu Monti et al. und Vogel et al. kann hier bereits auf den ersten Blick die vergleichsweise geringe Studiengröße und das damit einhergehende große Konfidenzintervall entnommen werden. Dieses wird als durch die Symbole verlaufende horizontale Linien dargestellt. Je weiter (länger) sie sind, desto unsicherer ist, in welchem Bereich das eigentliche Ergebnis zu erwarten ist.

Die vertikale Linie auf dem Plot entspricht einer Effektstärke nach Cohen von $d = 0$, also keinem Effekt. Die statistische Signifikanz eines Effekts wird damit visualisiert, dass sich die Konfidenzintervalle vollständig auf der einen Seite der „Null-Effekt-Linie“ befinden, sie also nicht überschreiten. Man sieht, dass dies nur für die Studie

von Stender et al. mit ihrer hohen Patientenzahl der Fall ist. Zudem befinden sich auch nicht alle Effekte auf einer Seite der „Null-Effekt-Linie“, was bedeutet, dass in der Studie von Monti et al. MCS-Patienten besser abgeschnitten haben als VS-Patienten.

Allerdings kann erst durch Betrachtung des Gesamtergebnisses, also dem sogenannten gepoolten Ergebnis, ein abschließendes Urteil über die Aussage der Analyse getroffen werden. Das gepoolte Gesamtergebnis der Metaanalyse wird am Ende der Grafik als Raute dargestellt, wobei die Breite der Raute das Konfidenzintervall widerspiegelt. Deutlich ist in diesem Fall die Überschneidung des sehr weiten Konfidenzintervalls mit der „Kein-Effekt-Linie“ zu erkennen.

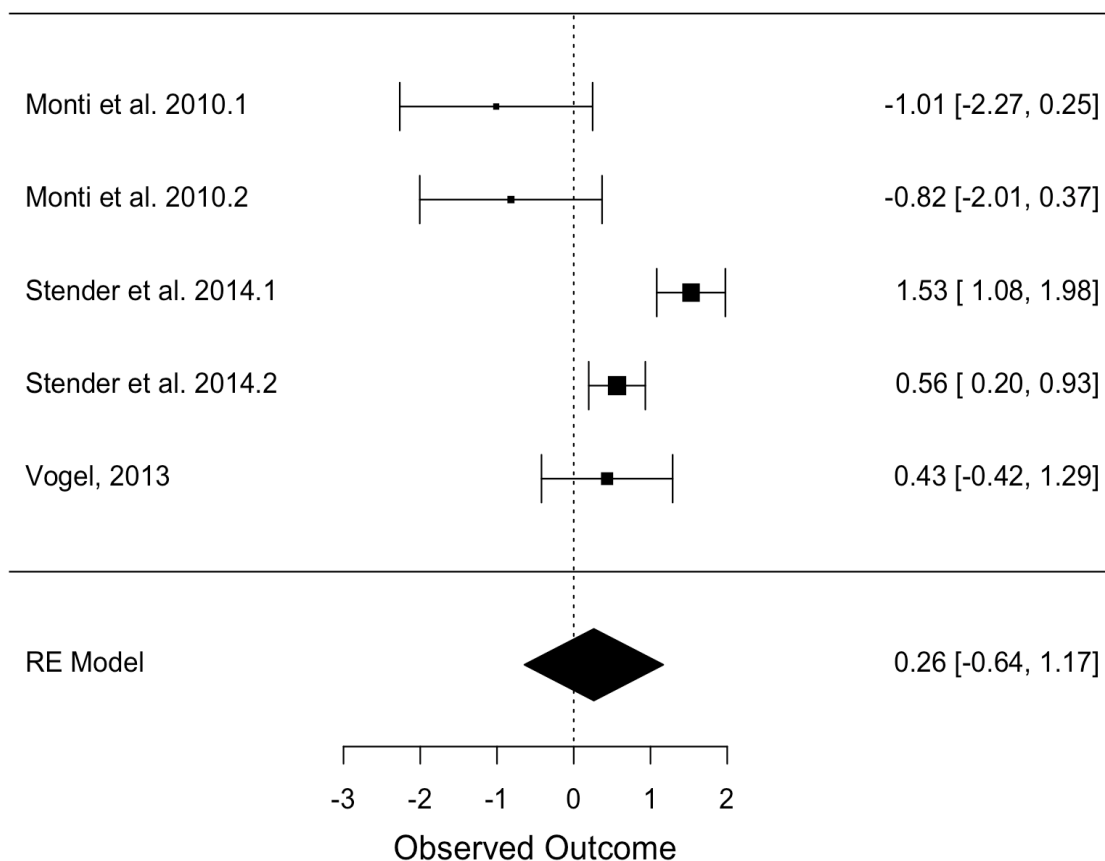


Abbildung 17: Forest-Plot der Aktive-State-Untersuchungsmethoden

3.3.2 Passive Stimulationen

Die Abbildung 18 stellt den Forest-Plot der Studien mit passiver Untersuchungsmethode dar.

Das Konfidenzintervall der gepoolten Gesamteffektstärke tangiert nur in einem äußerst geringen Umfang die vertikale „Kein-Effekt-Linie“, wobei eine gepoolte Gesamteffektstärke von $d = 0,36$ nach Cohen als dennoch kleiner Effekt anzusehen ist.

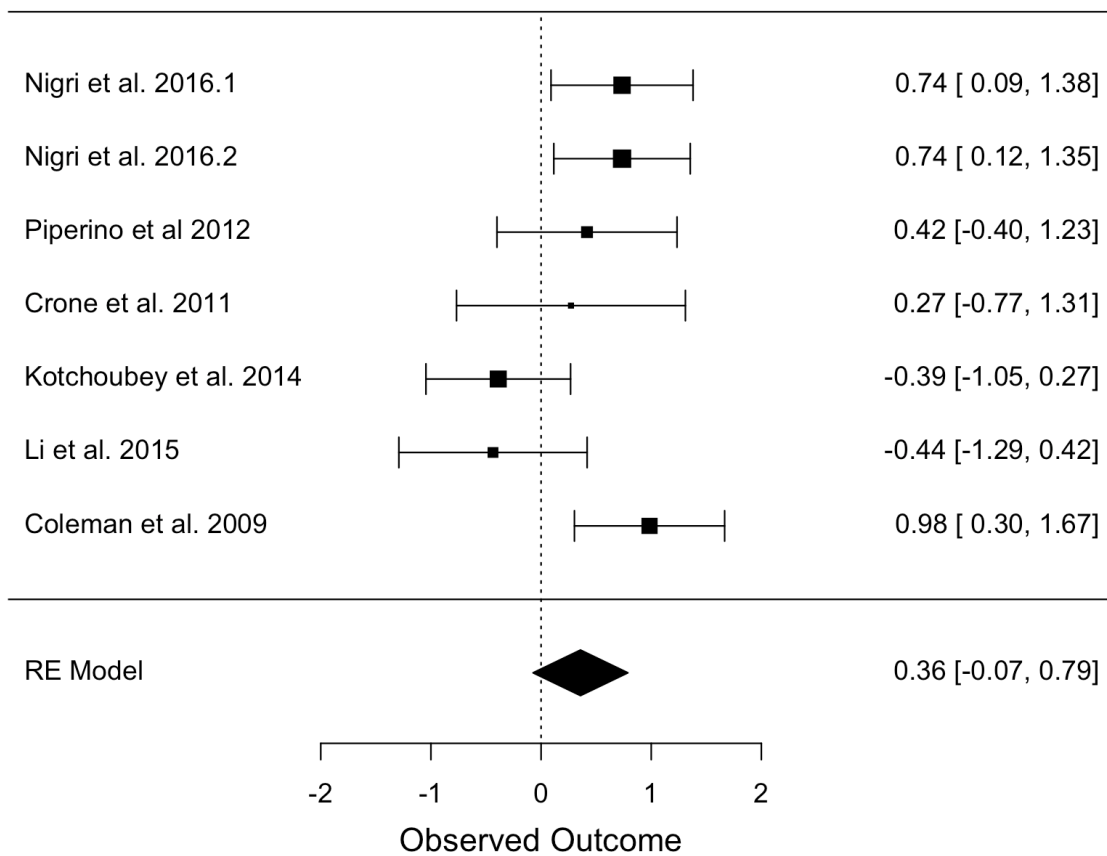


Abbildung 18: Forest-Plot der Passive-State-Untersuchungsmethoden

3.3.3 Resting-State-Methode

Dem Forest-Plot der Resting-State-Methode kann in Abbildung 19 auf den ersten Blick die Signifikanz dieser Untersuchungsmethode im Vergleich zu den beiden vorangegangenen Methoden entnommen werden. Das gepoolte Gesamtergebnis von $d = 0,87$ entspricht nach Cohen einem großen Effekt. Zudem befindet sich das erkennbar kleine Konfidenzintervall deutlich rechts der „Kein - Effekt - Linie“.

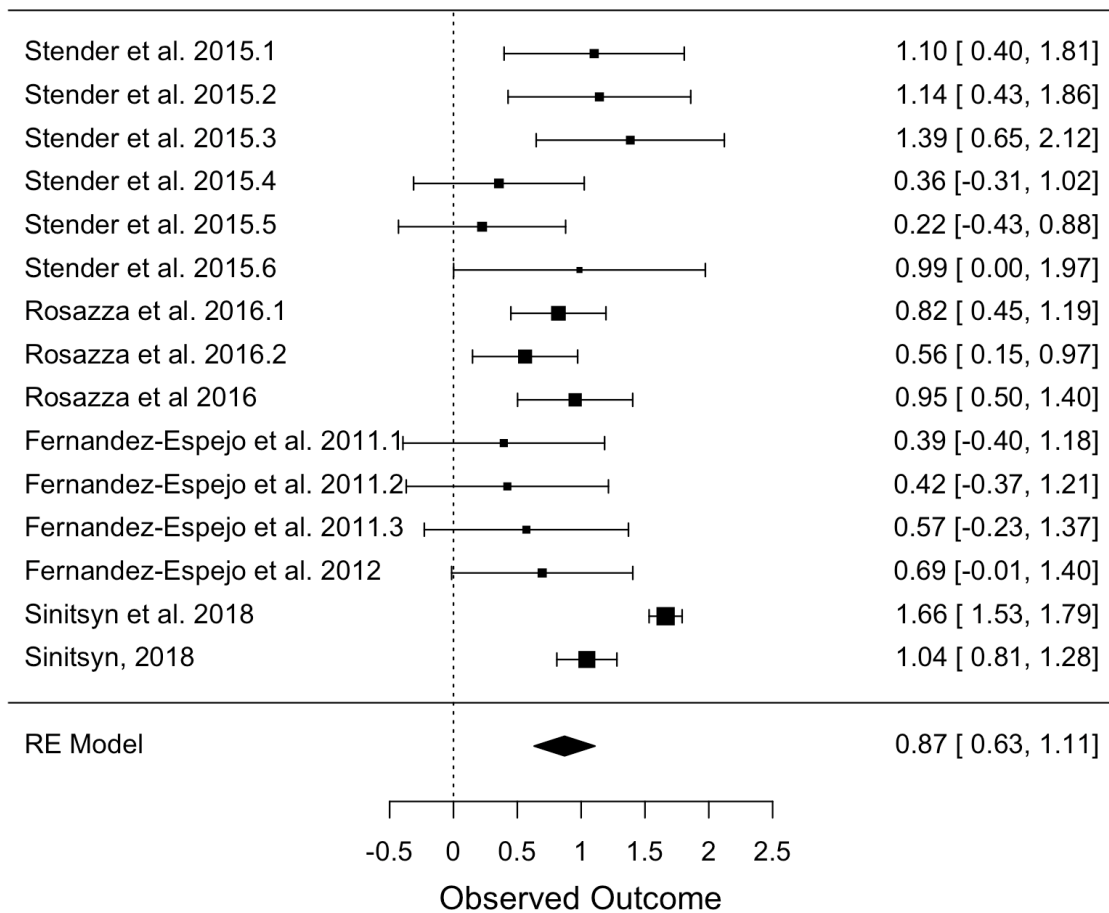


Abbildung 19: Forest-Plot der Resting-State-Untersuchungsmethoden

4 Diskussion

Das Hauptziel der Arbeit war eine verlässliche Darstellung der Sachlage auf dem Gebiet der Anwendung funktioneller bildgebender Verfahren zur Differenzierung der Bewusstseinszustände von Patienten, die nicht mehr in der Lage sind in einer wie auch immer gearteten Weise mit ihrer Umgebung in Kontakt zu treten. Seit Ende der 1990er Jahre (Giacino, 1997) wird diese Patientengruppe vor allem in zwei Kategorien unterteilt: Vegetativer Zustand (VS; Synonym Wachkoma, Unresponsive Wakefulness Syndrome) und Minimaler Bewusstseinszustand (MCS). Dennoch bleibt diese Differenzierung in höchstem Maße fehleranfällig und zwei in den letzten Jahren erschienene systematische Übersichtsarbeiten über neurophysiologische Befunde bei VS und MCS haben zu erheblichen Zweifeln geführt, inwieweit diese klinische Differenzierung eine neurophysiologische Basis hat. Es muss allerdings gesagt werden, dass eine dieser beiden systematischen Übersichten nur die Arbeiten mit ereigniskorrelierten Hirnpotentialen, aber keine fMRI- oder PET-Studien analysierte (Kotchoubey, 2017) und die andere berücksichtigte fMRI- und PET-Studien nur bis zum Erscheinungsjahr 2013 (Liberati et al., 2014), während in den Jahren 2015 – 2018 mehrere große Studien zu dieser Sachlage erschienen.

In der vorliegenden Studie sollten die Ergebnisse funktioneller Bildgebungsstudien zusammengefasst werden, welche in der Lage sein könnten, Hirnkorrelate des Bewusstseins von Patienten unabhängig von deren motorischer Funktion, im Wesentlichen auch ohne entsprechende Compliance des Patienten mit hoher räumlicher und zeitlicher Auflösung zu erfassen. Für diese Metaanalyse wurden hierzu Studien aus dem Bereich der fMRI- und PET-Untersuchung ausgewählt, die sich vergleichend mit der Diagnose VS und MCS beschäftigen.

4.1 Limitationen

Eine wichtige Limitation betrifft das eigene Vorgehen. Die hier untersuchten Artikel stammen ausschließlich aus der Pubmed-Datenbank. Keine weiteren Datenbanken wurden verwendet. Auch wurde davon abgesehen, mit Autoren der Primärstudien direkt in Kontakt zu treten, um möglicherweise unveröffentlichte Daten oder Ergebnisse zu erhalten.

Andere wichtige Einschränkungen stellen die Folgen der Limitationen in den analysierten Primärstudien dar. Zum einen werden möglicherweise nicht-signifikante oder negative Studienergebnisse teilweise nicht veröffentlicht oder nicht zur Veröffentlichung eingereicht. Sie könnten auch zum Beispiel in wenig bekannten Sammelbänden veröffentlicht und daher nur sehr schwer zugänglich sein. Dadurch kann es zu einer Verzerrung des Gesamtergebnisses zugunsten der positiven Resultate kommen. Aufschluss über einen solchen Publikationsbias, der dadurch zustande kam, dass negative Studienergebnisse abgelehnt oder vom Autor erst gar nicht zur Publikation angeboten wurden, kann die Korrelation der Effektstärke mit der Patientenzahl und dem Standardfehler der Effektstärke geben. Nur solche Verzerrungen können durch eine entsprechende negative Korrelation erfasst werden. Die aufgezeigten Ergebnisse einer annähernd bei 0 befindlichen Korrelation zwischen der Effektstärke und dem Standardfehler der Effektstärke und der Patientenzahl sprechen gegen einen solchen beschriebenen Publikationsbias. Damit kann davon ausgegangen werden, dass sowohl negative als auch positive Ergebnisse durch die Autoren der analysierten Primärstudien erwähnt wurden und dadurch in die vorliegende Metaanalyse eingeflossen sind. Im Gegensatz hierzu stießen Kotchoubey et al. (Kotchoubey and Pavlov, 2018) in ihrer Arbeit auf eine Korrelation des Standardfehlers mit der Effektstärke und einem damit einhergehenden Publikationsbias zugunsten positiver Ergebnisse.

Andererseits können Verzerrungen durch selektive Berichte über nur einen erfolgreichen Anteil einer Studie durch eine negative Korrelation nicht erfasst werden. Die Patientenzahl und in der Regel auch der Standardfehler sind für die

berichteten und nicht-berichteten Variablen identisch oder sehr ähnlich, da es sich um ein und dieselbe Studie handelt.

Eine weitere Möglichkeit der Verzerrung der veröffentlichten Ergebnisse ist in der Auswahl der Studienprobanden zu sehen. So kann es sein, dass zum Beispiel besonders schwer erkrankte VS-Patienten und entsprechend leicht erkrankte MCS-Patienten ausgewählt wurden. Dadurch würde in der gewählten Stichprobe ein deutlich größerer Unterschied der beiden Patientengruppen auftreten, als er sich in der Generalpopulation tatsächlich darstellt. Ein solches Vorgehen, d.h. die Auswahl von „klaren Fällen“ statt „durchschnittlicher Patienten“ ist in der Medizin, insbesondere in den ersten Stadien der Forschung, durchaus legitim, sollte aber als solches explizit dargestellt werden.

Des Weiteren kann eine fehlende Verblindung der Untersucher zu einer Verzerrung führen. So kann es sein, dass entweder die Neurologen, die die CRS-R-Diagnose stellen, über die Ergebnisse der Bildgebung informiert sind oder die Neurophysiologen, welche die Bildgebungsdaten analysieren, die klinische Diagnose kennen.

Lediglich in dem zu Beginn aufgezeigten Fall ist es möglich, einen Bias durch die negative Korrelation der Effektstärke mit der Patientenzahl und dem Standardfehler der Effektstärke festzustellen. In den weiteren Fällen gibt es keine direkten Methoden der Erfassung, sondern nur indirekte Maße wie QUADAS-2. In der vorliegenden Arbeit ergab die Qualitätsbeurteilung nach den Richtlinien von QUADAS-2 einen eher geringen Mittelwert von nur 3,2. Dieser geringe Wert ist hauptsächlich durch die Patientenauswahl, die Methode der Studiendarstellung und vor allem auch die fehlende Blindheit der Untersucher begründet. In den meisten Studien wurde die Patientenauswahl ohne eine klare Angabe über die Ausschluss- und Einschlussgründe der teilnehmenden Probanden getroffen. Die Problematik im Bereich des Index-Testes liegt darin begründet, dass in der überwiegenden Anzahl an Studien weder bereits vor der Untersuchung ein Schwellenwert bezüglich entsprechender Diagnosen genannt wurde, noch die Untersucher blind bezüglich einer bestehenden Diagnose waren. Zusätzlich kann die Kategorie des „Flow und

Timing“ die Beurteilung des Bias negativ beeinflusst haben. So sind in nur wenigen Fällen klare Angaben über den zeitlichen Ablauf zwischen Reference Standard und Index-Test vorhanden. Diese für eine hohe Studienqualität geforderten Ansprüche lassen sich in den meisten der für diese Arbeit untersuchten Studien nicht finden. Zu einem ähnlichen Schluss kommen Kotchoubey et al. in der Analyse der Publikation, die sich mit der Prognose von VS- und MCS-Patienten (nicht mit der diagnostischen Differenzierung zwischen VS und MCS, wie in vorliegender Studie) befasst haben. Sie berichteten, dass in mehr als 80 % aller analysierten Studien Untersucher bereits Kenntnis über die Ergebnisse einer entsprechenden Voruntersuchung hatten und dadurch möglicherweise in ihrer späteren Diagnostik beeinflusst wurden (Kotchoubey and Pavlov, 2018). Für das gesamte beleuchtete Forschungsgebiet zeigt sich ein gewisses Defizit im Bereich der Datenaufbereitung und Datenwiedergabe. Inwieweit die geforderten Qualitätsansprüche in der praktischen Umsetzung tatsächlich nicht erfüllt oder nur nicht protokolliert und dem Leser mitgeteilt wurden, ist im Nachhinein nicht oder nur sehr schwer nachzuvollziehen. Deswegen muss im Zweifelsfall im Rahmen der Beurteilung der Ergebnisse einzelner Primärstudien immer eher davon ausgegangen werden, dass die geforderten Qualitätsanforderungen bereits in der praktischen Durchführung nicht eingehalten worden sind. Dahingegen lässt der Reference Standard keine Bedenken zu, da in beinahe allen Studien die CRS-R zur Anwendung kam, die bekanntermaßen zu einer der verlässlichsten klinischen Evaluationsmethoden im Bereich der DoC zählt (Gerrard et al., 2014).

Des Weiteren können sogenannte nicht systematische Fehler, auch Zufallsfehler genannt, auftreten. Solche Zufallsfaktoren können das Ergebnis in jede Richtung verzerren und wenn sie mit einer systematischen Verzerrung zusammentreffen, diese verstärken. Vor allem kleine Stichprobengrößen und Heterogenität der Studien stellen solche Zufallsverzerrungen dar.

Im Rahmen der vorliegenden Arbeit wurden Studien mit einem Bias kleiner als 3 für die Metaanalyse ausgeschlossen, um eine Übertragung methodisch schlechter Einzelstudien nach dem „Garbage-in-Garbage-out-Problem“ (Bown and Sutton,

2010) auf das metaanalytische Gesamtergebnis zu vermeiden. Das Risiko, durch den Ausschluss einzelner qualitativ schlechter Studien einen gewissen Selektionsbias hervorzurufen, wurde dagegen als vernachlässigbar angesehen und in Kauf genommen.

Eine weitere generelle Problematik der Metaanalyse stellt die Heterogenität dar. Inwieweit eine Zusammenfassung zu einem gepoolten Ergebnis für eine Metaanalyse trotz Heterogenität sinnvoll ist, ist umstritten. Der Metaanalyse wird deshalb immer wieder vorgeworfen nicht vergleichbare Untersuchungen oder höchstens ähnliche Studien zu integrieren (Hansueli Stamm, 1995). Typische Faktoren, die diese Heterogenität verursachen könnten, wären in dem hier vorliegenden Fall die Zusammensetzung der Studiengruppen (Komorbiditäten, Spektrum des Schweregrads des jeweiligen DoC), die sich in den jeweiligen Subgruppen (Aktive-, Passive-, Resting-State) immer unterscheidende Untersuchungsmethode und auch die von den Untersuchern stark abhängige Diagnostik.

Allerdings gibt es unter den Statistikern auch Autoren, die der Heterogenität keine allzu große Bedeutung zumessen, da es im Grunde genommen nur auf die jeweilige Fragestellung ankommt. Stimmen die Fragestellungen der integrierten Primärarbeiten miteinander überein, so können diese trotz einer signifikanten Heterogenität metaanalytisch untersucht werden (Glass et al., 1981). Schließlich ist es zwangsläufig notwendig, alle Untersuchungsmethoden zu einer Thematik, auch wenn sich diese unterscheiden, zu berücksichtigen. Diese immer wieder auftretende Problematik wird auch als das „Äpfel-und-Birnen-Problem“ bezeichnet (Fricke and Treinies, 1985). Als weiteres Gegenargument kann aus Fricke & Treinies „Einführung in die Metaanalyse“ die Aussage von Glass angeführt werden, nachdem es schlichtweg notwendig ist, Äpfel und Birnen zu vergleichen, wenn Untersuchungen zum Thema Früchte angestellt werden sollen (Fricke and Treinies, 1985).

Von daher wurde die vorliegende Metaanalyse trotz eindeutiger Heterogenität durchgeführt. Dennoch sollten die zusammenfassenden Ergebnisse dieser Analyse

angesichts der Heterogenität und der beschriebenen methodischen Probleme mit Vorsicht betrachtet werden.

Interessant ist abschließend ein kurzer Vergleich mit der Qualität der für die Studie „Evoked and event - related potentials in disorders of consciousness: A quantitative review“ (Kotchoubey, 2017) verwendeten Studien. Hierbei kommt man zu dem Schluss, dass die Arbeiten über bildgebende Verfahren eine geringfügig bessere Qualität als solche über EP und ERP Daten aufweisen. Erstaunlicherweise zeigt sich zwar die mittlere untersuchte Patientenzahl in beiden Fällen mit einem Wert von 31 untersuchten Patienten identisch, allerdings wurde im Falle der EP und ERP Daten eine deutliche negative Korrelation zwischen der Patientenzahl und Effektstärke nachgewiesen. Ein solch deutlicher Hinweis auf eine Publikationsverzerrung konnte, wie oben erläutert, in dieser Arbeit nicht erbracht werden. Einen weiteren Unterschied stellt die hier gefundene positive Korrelation zwischen der Stichprobengröße und dem Impact Factor dar, die in der Studie von Kotchoubey (Kotchoubey, 2017) fehlte. Dies weist darauf hin, dass bei den bildgebenden Studien – im Gegensatz zu den elektrophysiologischen Studien – sich der Leser zumindest teilweise auf die Reputation der Zeitschrift verlassen kann, denn hier tendieren Zeitschriften mit höherem Ruf dazu, zuverlässigere Studien zu publizieren. Die deutlich negative Korrelation des Standardfehlers mit der Stichprobengröße und des Standardfehlers mit dem Journal Impact Factor, auch wenn dieser in einem nicht so deutlichen Umfang wie die Patientenzahl mit dem Standardfehler korreliert, unterstützt diese These.

4.2 Metaanalyse

4.2.1 Aktive Aufgaben

Werden die Ergebnisse zu den aktiven Untersuchungsmethoden betrachtet, wird vor allem mit Blick auf das gepoolte Gesamtergebnis schnell deutlich, dass diese Art der Diagnostik statistisch keine verlässliche Unterscheidung zwischen dem VS und MCS ermöglicht. Die Gesamteffektstärke von $d = 0,26$ kann als nur kleiner Effekt angesehen werden, zudem überschreitet das Konfidenzintervall die „Kein-Effekt-Linie“ deutlich. Damit muss davon ausgegangen werden, dass innerhalb des zu erwartenden Ergebnisses nicht nur kein Effekt auftritt, sondern sogar ein negativer Effekt erwartet werden kann. Dies bedeutet aber selbstverständlich nicht automatisch, dass dieses Mittel zur Unterscheidung klinisch in einzelnen Fällen nicht doch relevant sein kann. So kam die Studie von Stender et al. 2014 zu dem Schluss, den VS von einem MCS mit Hilfe einer aktiven Untersuchungsmethode verlässlich unterscheiden zu können. Die Werte für die FDG-PET mit $d = 1,53$ (CI: 1,08 - 1,98) und die fMRI-Diagnostik mit $d = 0,56$ (CI: 0,20 – 0,93) zeigten einen deutlichen Effekt. Dem entgegen stehen die negativen Effektstärken der Datensätze von Monti et al. 2010. Sie deuten auf ein besseres Abschneiden in den Untersuchungen der VS-Patienten im Vergleich mit den MCS-Patienten hin und stehen somit im Gegensatz zu dem erwarteten Ergebnis. In dieser Studie zeigten insgesamt nur fünf Patienten aus einer Stichprobe von insgesamt 45 Patienten eine signifikante Aktivierung, wovon sich vier in einem VS-Zustand befanden und nur ein Patient im MCS. Bei drei der vier VS-Patienten konnte durch einen nachträglich durchgeführten klinischen Test ein gewisser Grad an Wachheit nachgewiesen werden. Ein Patient war im Rahmen eines fMRI demnach sogar in der Lage auf simple „Ja/Nein“-Fragen zu antworten. Eine zuverlässige Unterscheidung zwischen dem VS und MCS war hiernach dennoch nicht möglich (Monti et al., 2010).

Eine Evaluation des Bewusstseins bei Patienten mit einer DoC ist also abschließend gesagt mit Hilfe aktiver Untersuchungsmethoden statistisch nicht definitiv möglich.

4.2.2 Passive Stimulation

Gut zu erkennen ist bei der Betrachtung der Studien von Nigri et al. 2016 und Coleman et al. 2009, dass diese das verlässlichste Ergebnis liefern. Deutlich wird dies an den geringen Konfidenzintervallen, die die vertikale Nulllinie nicht überschreiten. Zwar kommen Nigri et al. 2016 auch zu dem Schluss, dass eine Unterscheidung innerhalb der Patientengruppen nur schwer möglich ist, dennoch wird eine Aktivierung einzelner Hirnareale durch olfaktorische Stimulation erreicht. So zeigen 85 % der VS- und 100 % der MCS-Probanden eine Aktivierung des Cortex Piriformis und des primären olfaktorischen Cortex (Nigri et al., 2016). Wie wichtig eine Überprüfung der klinischen Evaluation des Bewusstseinszustandes sein kann, zeigt deutlich die Studie von Coleman et al. 2009. Zwei von insgesamt 22 als VS diagnostizierte Patienten zeigten auf einen Tonstimulus die gleichen Aktivierungen wie bei gesunde Probanden (Coleman et al., 2009).

Zwar weist auch die Arbeit von Kotchoubey et al. 2014 eine vergleichsweise große Studienpopulation und damit einhergehend ein geringes Konfidenzintervall auf, jedoch zeigt dieses Überschneidungen mit der „Kein-Effekt-Linie“. Von daher muss zumindest davon ausgegangen werden, dass die Möglichkeit besteht, dass es sich bei dem in diesem Fall gezeigten Effekt lediglich um eine zufällige Abweichung handelt und in Wahrheit kein Unterschied zwischen den Patientengruppen besteht. Dennoch wird auf den ersten Blick deutlich, dass die Arbeit von Kotchoubey et al. 2014 sowie die von Li et al. 2015 jeweils Resultate entgegen der Erwartung liefern, entsprechend der die MCS-Patienten grundsätzlich bessere Untersuchungsergebnisse erzielen sollten als die Patienten in einem VS. Die Ergebnisse der Arbeiten von Li et al. 2015, Piperino et al. 2012 und im Besonderen der von Crone et al. 2011 sind hinsichtlich der vorliegenden Auswertung mit Vorsicht zu bewerten. So zeigen sich hier doch klar die vergleichsweise kleinen untersuchten Patientengruppen und die deutlichen Überschneidungen mit der „Kein-Effekt-Linie“. Allerdings können die hier betrachteten passiven Methoden zur Unterscheidung des VS von einem MCS als durchaus verlässlicher als die der Aktive-State-Methodik

angesehen werden. Zum einen beläuft sich bei den Aktive-State-Untersuchungen die gepoolte Gesamteffektstärke nur auf einen Wert von 0,26 und ergibt so nach Cohen nur knapp einen kleinen Effekt. Im Vergleich hierzu wird mit den Passive-State-Methoden eine gepoolte Gesamteffektstärke von $d = 0,36$ erreicht, der allerdings auch noch immer als „kleiner Effekt“ einzuordnen ist. Zum anderen überschreitet das Konfidenzintervall der Gesamteffektstärke der Aktive-State-Methode, wie bereits aufgezeigt, deutlich die „Kein-Effekt-Linie“. Dies ist im Rahmen der Passive-State-Methode nicht in einem solch deutlichen Umfang wie im Bereich der Aktive-State-Methode der Fall. Zu einer ähnlichen Schlussfolgerung kommen Kondziella et al. 2016. Des Weiteren schreiben Kondziella et al. 2016, dass MCS-Patienten öfter auf gestellte Aufgaben reagieren als solche in einem VS. Vor allem scheinen MCS-Patienten in der Lage zu sein auf passive und aktive Paradigmen zu reagieren, wohingegen VS-Patienten nur in 15 % der von Kondziella et al. 2016 untersuchten Fälle in der Lage waren, auf aktive Aufgaben zu reagieren.

4.2.3 Resting-State-Methode

Bei Betrachtung der Ergebnisse in Bezug auf die Resting-State-Methode wird klar, dass es sich hierbei um die verlässlichste Methode zur Unterscheidung zwischen dem VS und dem MCS handelt. Der gepoolte Gesamteffekt von $d = 0,87$ (CI: 0,63 - 1,11) steht eindeutig für einen großen Effekt. Somit kann davon ausgegangen werden, dass durch die hier angewandten Methoden eine zuverlässige und signifikante Unterscheidungsmöglichkeit zwischen dem VS und MCS möglich ist.

Besonders hervorzuheben sind hierbei die Resultate von Sinitsyn et al. 2018 und Rosazza et al. 2016, die mit in Summe 130 untersuchten Patienten eine sehr große Population beleuchtet haben und damit verbunden auch nur ein kleines Konfidenzintervall zeigen. Rosazza et al. 2016 konnten mittels eines rs-fMRI (resting state fMRI) zwei klinisch als VS diagnostizierte Patienten dem MCS zuordnen. Allerdings weisen die Autoren darauf hin, dass für eine entsprechend notwendige

hohe Präzision bei der Bildgebung keinerlei Bewegungen die Aufnahme stören dürfen. Ist dies gegeben, eignet sich ein rs-fMRI sehr gut, um in unklaren Fällen zusätzliche Sicherheit in der Diagnostik zu erhalten (Rosazza et al., 2016). Sinitsyn et al. 2018 fanden eine Methode, um VS von MCS zu unterscheiden. Sowohl der von ihnen untersuchte ITCI (Index of Thresholded Connectome Intactnes) als auch der HDI (Hub Disruption Index) zeigten einen signifikanten Unterschied zwischen dem VS und MCS (Sinitsyn et al., 2018).

Interessant ist auch die Arbeit von Stender et al. 2015, die den Glucosestoffwechsel des Gehirns mittels PET-CT untersuchten. Hierbei zeigten nur die Messreihen zu Thalamus und Gehirnstamm kein signifikantes Ergebnis.

Abschließend kann somit gesagt werden, dass die Methode der aktiven Aufgaben die am wenigsten aussagekräftige ist. Im Gegensatz dazu unterscheidet die Resting-State-Methode entsprechend den hier gezeigten Ergebnissen am besten zwischen VS und MCS. Das abschließende Ergebnis der Methode der passiven Stimulation bleibt weniger klar; der Effekt liegt knapp an der Signifikanzgrenze, sodass empfohlen werden kann, weitere Studien hierzu abzuwarten.

Jedoch bedeutet eine statistische Aussagekraft nicht immer, dass dieses Ergebnis klinisch auch relevant ist. Genauso bedeutet umgekehrt ein statistisch nicht aussagekräftiges Ergebnis nicht, dass dieses im klinischen Alltag keine Bedeutung haben kann.

4.3 Schlussfolgerung

Obwohl die Qualität der gefundenen Studien in dem hier untersuchten Forschungsgebiet (v.a. PET- und fMRI-Verfahren) höher ist als in ähnlichen Studien mit der Anwendung ereigniskorrelierter Hirnpotentiale (Kotchoubey, 2017), erreichen auch diese Studien nicht den erhofften Standard, der in anderen Bereichen bereits üblich ist. Hier sollte und muss in Zukunft nachgebessert werden, um die Verlässlichkeit aufgezeigter Resultate zu steigern. Zudem würde das

Einhalten der geforderten Richtlinien keinen allzu großen zusätzlichen Mehraufwand in der praktischen Durchführung bedeuten und dennoch zu einer deutlichen Verbesserung der Studienqualität führen.

Wie auch in der Arbeit von Bender et al. 2015 wurde hier gezeigt, dass neben der bereits standardisierten klinischen Untersuchung an Hand der CRS-R zur Sicherung der Diagnose in zweifelhaften Situationen weitere diagnostische Mittel notwendig werden (Bender et al., 2015). Als ergänzendes Mittel zu der bereits durchgeführten klinischen Untersuchung sollte zukünftig generell eine wie oben aufgezeigte funktionelle Bildgebung durchgeführt werden. Zwar ist es hiermit alleine nur schwer möglich, den VS von einem MCS zu unterscheiden, allerdings kann vor dem Hintergrund der klinischen Untersuchung diese bestätigt beziehungsweise korrigiert werden. So gelang es des Öfteren, vor allem bei Patienten mit der klinischen Diagnose des VS, einen gewissen Grad an Bewusstsein nachzuweisen. Wirkungsvoll erscheinen hierbei vor allem die Untersuchungsmethoden aus dem Bereich der passiven Untersuchungsmethoden und besonders der Resting-State-Verfahren. Die Ergänzung der bereits erprobten klinischen Verfahren um solch zukunftsweisende Methoden kann und wird die Behandlung und den Umgang mit Patienten in einem Zustand der DoC und auch deren Angehörigen erheblich erleichtern.

5 Zusammenfassung

Eine verlässliche Methode zur Unterscheidung und Diagnostik von Patienten in einem minimalen (MCS) oder vegetativen Zustand (VS; Synonyme: Wachkoma; Unresponsive Wakefulness Syndrom) steht bis heute nicht zur Verfügung. Das Hauptziel dieser Arbeit war es eine quantitative Darstellung der Sachlage auf dem Gebiet der Anwendung funktioneller bildgebender Verfahren zur Differenzierung der Bewusstseinszustände von Patienten, die nicht mehr in der Lage sind, in einer wie auch immer gearteten Weise mit ihrer Umgebung in Kontakt zu treten, aufzuzeigen. Hierzu wurden aus der Pubmed-Datenbank im Zeitraum von 01.10.2016 bis zum 15.6.2018 eine Gesamtanzahl von 25 relevanten Artikeln, welche die entsprechenden statistischen Daten zur Unterscheidung der VS- von MCS-Patienten anhand funktioneller bildgebender Verfahren aufwiesen, ermittelt. Die betrachteten Studien untersuchten im Mittel eine Anzahl von 31 Patienten und wiesen in Summe eine Gesamtanzahl an 796 Patienten auf. Für die an das systematische Review angeschlossene Metaanalyse ergab sich aus den 25 Studien eine gesamte Anzahl von 53 zu analysierenden Datensätzen. Vor der Kalkulation der Metaanalyse wurde eine ausführliche Qualitätsbeurteilung, nach den Richtlinien von QUADAS-2 erhoben. Hierbei zeigte sich allgemein eine Schwäche der Studienqualität in dem gesamten hier betrachteten Forschungsgebiet, sodass sich im Durchschnitt für die Biaswerte lediglich ein Wert von 3,2 ergab (der Wert 0 entspricht den stärksten Bias, der Wert 8 bedeutet eine biasfreie Studie). Die Gründe hierfür liegen hauptsächlich in der fehlenden Verblindung der Untersucher, in der Methode der Patientenauswahl und in der Darstellung der untersuchten abhängigen Variablen.

In die angeschlossene Metaanalyse wurden nach der Einteilung der Studien entsprechend der Untersuchungsmethoden nach Aktiv-, Passiv- und Resting-State-Methode nur solche Studien mit einem Biaswert größer als oder gleich 3 aufgenommen. Hierbei konnte nachgewiesen werden, dass sich die Resting-State-Methode mit einer gepoolten Gesamteffektstärke von $d = 0,78$ (Konfidenzintervall

von 0,54 bis 1,02) als die verlässlichste bei der Unterscheidung der Bewusstseinszustände des VS und MCS herausstellte. Im Gegenteil dazu ist die Aktive-State-Methode die am wenigsten aussagekräftige. Das abschließende Ergebnis der Methode der passiven Stimulation bleibt mit einem nur knapp an der Signifikanzgrenze liegenden Effekt unklar, sodass empfohlen werden kann weitere Studien auf diesem Gebiet abzuwarten.

Die Ergänzung der bereits erprobten, standardisierten klinischen Untersuchungsverfahren um solche zukunftsweisenden bildgebenden Verfahren wird die Diagnostik im Bereich der Disorders of Consciousness (DoC) deutlich erleichtern und verbessern. Vor allem zur Sicherung der Diagnose in zweifelhaften Situationen sollte, wie die vorliegende Arbeit nahelegt, zukünftig stets eine zusätzliche funktionelle Bildgebung etabliert werden.

6 Literaturverzeichnis

- ANALYTICS, C. 2018. *InCites Journal Citation Reports* [Online]. Available: <https://jcr.incites.thomsonreuters.com/JCRLandingPageAction.action> [Accessed 2018].
- ANDREWS, K. 1997. Vegetative state--background and ethics. *J R Soc Med*, 90, 593-6.
- ANDREWS, K., MURPHY, L., MUNDAY, R. & LITTLEWOOD, C. 1996. Misdiagnosis of the vegetative state: retrospective study in a rehabilitation unit. *Bmj*, 313, 13-6.
- BAETHGE, C. 2012. Nichtperfekter Impact-Faktor. *Dtsch Arztebl International*, 109, 267-9.
- BEDINI, G., BERSANO, A., SEBASTIANO, D. R., SATTIN, D., CIARAFFA, F., TOSETTI, V., BRENNAN, G., FRANCESCHETTI, S., CIUSANI, E., LEONARDI, M., VELA-GOMEZ, J., BONCORAGLIO, G. B. & PARATI, E. A. 2016. Is Period3 Genotype Associated With Sleep and Recovery in Patients With Disorders of Consciousness? *Neurorehabil Neural Repair*, 30, 461-9.
- BENDER, A., JOX, R. J., GRILL, E., STRAUBE, A. & LULE, D. 2015. Persistent vegetative state and minimally conscious state: a systematic review and meta-analysis of diagnostic procedures. *Dtsch Arztebl Int*, 112, 235-42.
- BERTSCH, K. 2007. *Das EEG: Spontan-EEG und EKP* [Online]. Trier: Uni-Trier. Available: https://www.uni-trier.de/fileadmin/fb1/prof/PSB/TKP/Unterlagen_Lehre_2007_2008/Biopsychologie_seminar/07_Das_EEG.pdf [Accessed 19.06.2019 2019].
- BICK, A. S., LEKER, R. R., BEN-HUR, T. & LEVIN, N. 2013. Implementing novel imaging methods for improved diagnosis of disorder of consciousness patients. *J Neurol Sci*, 334, 130-8.
- BILDGEBUNG, A. N. U. N. Heidelberg: Neuroradiologie UK Heidelberg. Available: <https://www.klinikum.uni-heidelberg.de/neurologische-klinik/neuroradiologie/forschung/ag-neuroimaging-funktionelle-mrt/> [Accessed 19.06.2019 2019].
- BISWAL, B. B., VAN KYLEN, J. & HYDE, J. S. 1997. Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps. *NMR Biomed*, 10, 165-70.
- BODIEN, Y. G., GIACINO, J. T. & EDLOW, B. L. 2017. Functional MRI Motor Imagery Tasks to Detect Command Following in Traumatic Disorders of Consciousness. *Front Neurol*, 8, 688.
- BOLY, M., COLEMAN, M. R., DAVIS, M. H., HAMPSHIRE, A., BOR, D., MOONEN, G., MAQUET, P. A., PICKARD, J. D., LAUREYS, S. & OWEN, A. M. 2007. When thoughts become action: an fMRI paradigm to study

- volitional brain activity in non-communicative brain injured patients. *Neuroimage*, 36, 979-92.
- BOLY, M., FAYMONVILLE, M. E., SCHNAKERS, C., PEIGNEUX, P., LAMBERMONT, B., PHILLIPS, C., LANCELLOTTI, P., LUXEN, A., LAMY, M., MOONEN, G., MAQUET, P. & LAUREYS, S. 2008. Perception of pain in the minimally conscious state with PET activation: an observational study. *Lancet Neurol*, 7, 1013-20.
- BOWN, M. J. & SUTTON, A. J. 2010. Quality control in systematic reviews and meta-analyses. *Eur J Vasc Endovasc Surg*, 40, 669-77.
- BREMBBS, B., BUTTON, K. & MUNAFO, M. 2013. Deep impact: unintended consequences of journal rank. *Front Hum Neurosci*, 7, 291.
- BRÜCKLER, F. M. 2017. *Geschichte der Mathematik kompakt. Das Wichtigste aus Analysis, Wahrscheinlichkeitsrechnung, angewandter Mathematik, Topologie und Mengenlehre*, Springer Verlag
- COHEN, J. 1988. Statistical power analysis for the behavioral science (2.Auflage).
- COLEMAN, M. R., DAVIS, M. H., RODD, J. M., ROBSON, T., ALI, A., OWEN, A. M. & PICKARD, J. D. 2009. Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. *Brain*, 132, 2541-52.
- CRONE, J. S., LADURNER, G., HOLLER, Y., GOLASZEWSKI, S., TRINKA, E. & KRONBICHLER, M. 2011. Deactivation of the default mode network as a marker of impaired consciousness: an fMRI study. *PLoS One*, 6, e26373.
- DE BIASE, S., GIGLI, G. L., LORENZUT, S., BIANCONI, C., SFREDDO, P., ROSSATO, G., BASALDELLA, F., FUCCARO, M., CORICA, A., TONON, D., BARBONE, F. & VALENTE, M. 2014. The importance of polysomnography in the evaluation of prolonged disorders of consciousness: sleep recordings more adequately correlate than stimulus-related evoked potentials with patients' clinical status. *Sleep Med*, 15, 393-400.
- DI, H. B., YU, S. M., WENG, X. C., LAUREYS, S., YU, D., LI, J. Q., QIN, P. M., ZHU, Y. H., ZHANG, S. Z. & CHEN, Y. Z. 2007. Cerebral response to patient's own name in the vegetative and minimally conscious states. *Neurology*, 68, 895-9.
- DI PERRI, C., HEINE, L., AMICO, E., SODDU, A., LAUREYS, S. & DEMERTZI, A. 2014. Technology-based assessment in patients with disorders of consciousness. *Ann Ist Super Sanita*, 50, 209-20.
- DRINKMANN, A. 1990. *Methodenkritische Untersuchungen zur Metaanalyse*, Deutscher Studien Verlag.
- E.V., S. D. A. B. Available: <https://www.subito-doc.de> [Accessed].
- EDLOW, B. L., CHATELLE, C., SPENCER, C. A., CHU, C. J., BODIEN, Y. G., O'CONNOR, K. L., HIRSCHBERG, R. E., HOCHBERG, L. R., GIACINO, J. T., ROSENTHAL, E. S. & WU, O. 2017. Early detection of consciousness in patients with acute severe traumatic brain injury. *Brain*, 140, 2399-2414.
- FERNANDEZ-ESPEJO, D., BEKINSCHTEIN, T., MONTI, M. M., PICKARD, J. D., JUNQUE, C., COLEMAN, M. R. & OWEN, A. M. 2011. Diffusion weighted

- imaging distinguishes the vegetative state from the minimally conscious state. *Neuroimage*, 54, 103-12.
- FERNANDEZ-ESPEJO, D., JUNQUE, C., BERNABEU, M., ROIG-ROVIRA, T., VENDRELL, P. & MERCADER, J. M. 2010. Reductions of thalamic volume and regional shape changes in the vegetative and the minimally conscious states. *J Neurotrauma*, 27, 1187-93.
- FERNANDEZ-ESPEJO, D., JUNQUE, C., VENDRELL, P., BERNABEU, M., ROIG, T., BARGALLO, N. & MERCADER, J. M. 2008. Cerebral response to speech in vegetative and minimally conscious states after traumatic brain injury. *Brain Inj*, 22, 882-90.
- FERNANDEZ-ESPEJO, D., SODDU, A., CRUSE, D., PALACIOS, E. M., JUNQUE, C., VANHAUDENHUYSE, A., RIVAS, E., NEWCOMBE, V., MENON, D. K., PICKARD, J. D., LAUREYS, S. & OWEN, A. M. 2012. A role for the default mode network in the bases of disorders of consciousness. *Ann Neurol*, 72, 335-43.
- FRICKE, R. & TREINIES, G. 1985. *Einführung in die Metaanalyse*.
- GANTNER, I. S., BODART, O., LAUREYS, S. & DEMERTZI, A. 2012. Our rapidly changing understanding of acute and chronic disorders of consciousness: challenges for neurologists. *Future Neurology*, 8, 43-54.
- GERRARD, P., ZAFONTE, R. & GIACINO, J. T. 2014. Coma Recovery Scale-Revised: evidentiary support for hierarchical grading of level of consciousness. *Arch Phys Med Rehabil*, 95, 2335-41.
- GIACINO, J. T. 1997. Disorders of consciousness: differential diagnosis and neuropathologic features. *Semin Neurol*, 17, 105-11.
- GIACINO, J. T., ASHWAL, S., CHILDS, N., CRANFORD, R., JENNETT, B., KATZ, D. I., KELLY, J. P., ROSENBERG, J. H., WHYTE, J., ZAFONTE, R. D. & ZASLER, N. D. 2002. The minimally conscious state: definition and diagnostic criteria. *Neurology*, 58, 349-53.
- GLASS, G. V. 1976. Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher*, 5, 3-8.
- GLASS, G. V., MCGAW, B. & SMITH, M. L. 1981. *Meta-analysis in social research*, Sage Publications.
- HAKIM, C. 2000. *Research Design: Successful Designs for Social and Economic Research*, Routledge.
- HANNAWI, Y., LINDQUIST, M. A., CAFFO, B. S., SAIR, H. I. & STEVENS, R. D. 2015. Resting brain activity in disorders of consciousness: a systematic review and meta-analysis. *Neurology*, 84, 1272-80.
- HANSUELI STAMM, T. M. S. 1995. Metaanalyse. Eine Einführung. *Rainer Hampp Verlag*, 9.
- HEDGES, L. & OLKIN, I. 1985. *Statistical Methods in Meta-Analysis*.
- HIGGINS, J. P. & THOMPSON, S. G. 2002. Quantifying heterogeneity in a meta-analysis. *Stat Med*, 21, 1539-58.
- HIGGINS, J. P., THOMPSON, S. G., DEEKS, J. J. & ALTMAN, D. G. 2003. Measuring inconsistency in meta-analyses. *Bmj*, 327, 557-60.

- HIGGINS JPT, G. S. 2011. Cochrane Handbook for Systematic Reviews of Interventions.: TheCochraneCollaboration.
- HULICK, K. 2016. *Explainer: How to read brain activity* [Online]. Washington: ScienceNews for Students. Available: <https://www.sciencenewsforstudents.org/article/explainer-how-read-brain-activity> [Accessed 19.06.2019 2019].
- KANG, X. G., LI, L., WEI, D., XU, X. X., ZHAO, R., JING, Y. Y., SU, Y. Y., XIONG, L. Z., ZHAO, G. & JIANG, W. 2014. Development of a simple score to predict outcome for unresponsive wakefulness syndrome. *Crit Care*, 18, R37.
- KHAN, K. S., KUNZ, R., KLEIJNEN, J. & ANTES, G. 2003. Five steps to conducting a systematic review. *J R Soc Med*, 96, 118-21.
- KONDZIELLA, D., FRIBERG, C. K., FROKJAER, V. G., FABRICIUS, M. & MOLLER, K. 2016. Preserved consciousness in vegetative and minimal conscious states: systematic review and meta-analysis. *J Neurol Neurosurg Psychiatry*, 87, 485-92.
- KOTCHOUBEY, B. 2017. Evoked and event-related potentials in disorders of consciousness: A quantitative review. *Conscious Cogn*, 54, 155-167.
- KOTCHOUBEY, B., MERZ, S., LANG, S., MARKL, A., MULLER, F., YU, T. & SCHWARZBAUER, C. 2013. Global functional connectivity reveals highly significant differences between the vegetative and the minimally conscious state. *J Neurol*, 260, 975-83.
- KOTCHOUBEY, B. & PAVLOV, Y. G. 2018. A Systematic Review and Meta-Analysis of the Relationship Between Brain Data and the Outcome in Disorders of Consciousness. *Front Neurol*, 9, 315.
- KOTCHOUBEY, B., YU, T., MUELLER, F., VOGEL, D., VESER, S. & LANG, S. 2014. True or false? Activations of language-related areas in patients with disorders of consciousness. *Curr Pharm Des*, 20, 4239-47.
- LABS, T. N. 2016.
- LANDSNESS, E., BRUNO, M. A., NOIRHOMME, Q., RIEDNER, B., GOSSERIES, O., SCHNAKERS, C., MASSIMINI, M., LAUREYS, S., TONONI, G. & BOLY, M. 2011. Electrophysiological correlates of behavioural changes in vigilance in vegetative state and minimally conscious state. *Brain*, 134, 2222-32.
- LAUREN., M., LUCIANO, M., GLENDA, M., GEOFFREY, H., H., K. S. & ROUMEN, M. 2017. The Default Mode Network in Healthy Individuals: A Systematic Review and Meta-Analysis. *Brain Connectivity*, 7, 25-33.
- LAUREYS, S. 2005. Science and society: death, unconsciousness and the brain. *Nat Rev Neurosci*, 6, 899-909.
- LAUREYS, S., CELESIA, G. G., COHADON, F., LAVRIJSEN, J., LEON-CARRION, J., SANNITA, W. G., SAZBON, L., SCHMUTZHARD, E., VON WILD, K. R., ZEMAN, A. & DOLCE, G. 2010. Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. *BMC Med*, 8, 68.
- LEHEMBRE, R., GOSSERIES, O., LUGO, Z., JEDIDI, Z., CHATELLE, C., SADZOT, B., LAUREYS, S. & NOIRHOMME, Q. 2012. Electrophysiological

- investigations of brain function in coma, vegetative and minimally conscious patients. *Arch Ital Biol*, 150, 122-39.
- LENHARD W, L. A. 2016. *Calculation of Effect Sizes*. [Online]. Dettelbach. Available: https://www.psychometrica.de/effect_size.html [Accessed].
- LEWANDOWSKI, D. 2006. *Journal Impact Factor* [Online]. Available: http://www.forschungsinfo.de/iq/iq_inhalt.asp?agora/Journal_Impact_Factor/journal_impact_factor_inc.htmlXXXJournal%20Impact%20Factor [Accessed].
- LI, L., KANG, X. G., QI, S., XU, X. X., XIONG, L. Z., ZHAO, G., YIN, H. & JIANG, W. 2015. Brain response to thermal stimulation predicts outcome of patients with chronic disorders of consciousness. *Clin Neurophysiol*, 126, 1539-47.
- LIBERATI, G., HUNEFELDT, T. & OLIVETTI BELARDINELLI, M. 2014. Questioning the dichotomy between vegetative state and minimally conscious state: a review of the statistical evidence. *Front Hum Neurosci*, 8, 865.
- LOHNINGER, H. 11.11.2012. *Gütemaße für Klassifikatoren* [Online]. Available: http://www.statistics4u.com/fundstat_germ/ee_classifier_performance_metrics.html [Accessed 13.05.2019 2019].
- MARINO, S., BONANNO, L., CIURLEO, R., BAGLIERI, A., MORABITO, R., GUERRERA, S., RIFICI, C., GIORGIO, A., BRAMANTI, P. & DE STEFANO, N. 2017. Functional Evaluation of Awareness in Vegetative and Minimally Conscious State. *Open Neuroimag J*, 11, 17-25.
- MONTI, M. M., ROSENBERG, M., FINOIA, P., KAMAU, E., PICKARD, J. D. & OWEN, A. M. 2015. Thalamo-frontal connectivity mediates top-down cognitive functions in disorders of consciousness. *Neurology*, 84, 167-73.
- MONTI, M. M., VANHAUDENHUYSE, A., COLEMAN, M. R., BOLY, M., PICKARD, J. D., TSHIBANDA, L., OWEN, A. M. & LAUREYS, S. 2010. Willful modulation of brain activity in disorders of consciousness. *N Engl J Med*, 362, 579-89.
- NAKAGAWA, S. & CUTHILL, I. C. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev Camb Philos Soc*, 82, 591-605.
- NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION, U. S. N. L. O. M. Available: <https://www.ncbi.nlm.nih.gov/pubmed/> [Accessed].
- NIGRI, A., CATRICALA, E., FERRARO, S., BRUZZONE, M. G., D'INCERTI, L., SATTIN, D., SEBASTIANO, D. R., FRANCESCHETTI, S., MAROTTA, G., BENTI, R., LEONARDI, M. & CAPPA, S. F. 2017. The neural correlates of lexical processing in disorders of consciousness. *Brain Imaging Behav*, 11, 1526-1537.
- NIGRI, A., FERRARO, S., BRUZZONE, M. G., NAVA, S., D'INCERTI, L., BERTOLINO, N., SATTIN, D., LEONARDI, M. & LUNDSTROM, J. N. 2016. Central olfactory processing in patients with disorders of consciousness. *Eur J Neurol*, 23, 605-12.

- PIPERNO, R., BATTISTINI, A., CEVOLANI, D., MAFFEI, M., LEONARDI, M. & AGATI, R. 2012. fMRI activation with an "affective speech" paradigm in vegetative and minimally conscious States: applicability and prognostic value. *Neuroradiol J*, 25, 289-99.
- RODRIGUEZ MORENO, D., SCHIFF, N. D., GIACINO, J., KALMAR, K. & HIRSCH, J. 2010. A network approach to assessing cognition in disorders of consciousness. *Neurology*, 75, 1871-8.
- ROSAZZA, C., ANDRONACHE, A., SATTIN, D., BRUZZONE, M. G., MAROTTA, G., NIGRI, A., FERRARO, S., ROSSI SEBASTIANO, D., PORCU, L., BERSANO, A., BENTI, R., LEONARDI, M., D'INCERTI, L. & MINATI, L. 2016. Multimodal study of default-mode network integrity in disorders of consciousness. *Ann Neurol*, 79, 841-853.
- ROSSI SEBASTIANO, D., PANZICA, F., VISANI, E., ROTONDI, F., SCAIOLI, V., LEONARDI, M., SATTIN, D., D'INCERTI, L., PARATI, E., FERINI STRAMBI, L. & FRANCESCHETTI, S. 2015. Significance of multiple neurophysiological measures in patients with chronic disorders of consciousness. *Clin Neurophysiol*, 126, 558-64.
- SCHNAKERS, C., VANHAUDENHUYSE, A., GIACINO, J., VENTURA, M., BOLY, M., MAJERUS, S., MOONEN, G. & LAUREYS, S. 2009. Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. *BMC Neurol*, 9, 35.
- SCHWARZ, J. 2016. Universität Zürich. Available: https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/zusammenhange/rangkorrelation.html [Accessed 2018].
- SINITSYN, D. O., LEGOSTAEVA, L. A., KREMNEVA, E. I., MOROZOVA, S. N., POYDASHEVA, A. G., MOCHALOVA, E. G., CHERVYAKOVA, O. G., RYABINKINA, J. V., SUPONEVA, N. A. & PIRADOV, M. A. 2018. Degrees of functional connectome abnormality in disorders of consciousness. *Hum Brain Mapp*, 39, 2929-2940.
- SKLORZ, H. Oldenburg. Available: http://viles2.chi.make_chi_table.pl [Accessed 2017].
- STENDER, J., GOSSERIES, O., BRUNO, M. A., CHARLAND-VERVILLE, V., VANHAUDENHUYSE, A., DEMERTZI, A., CHATELLE, C., THONNARD, M., THIBAUT, A., HEINE, L., SODDU, A., BOLY, M., SCHNAKERS, C., GJEDDE, A. & LAUREYS, S. 2014. Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: a clinical validation study. *Lancet*, 384, 514-22.
- STENDER, J., KUPERS, R., RODELL, A., THIBAUT, A., CHATELLE, C., BRUNO, M. A., GEJL, M., BERNARD, C., HUSTINX, R., LAUREYS, S. & GJEDDE, A. 2015. Quantitative rates of brain glucose metabolism distinguish minimally conscious from vegetative state patients. *J Cereb Blood Flow Metab*, 35, 58-65.
- STERNE, J. A., EGGER, M. & SMITH, G. D. 2001. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *Bmj*, 323, 101-5.

- SUTTON, A. J. & HIGGINS, J. P. 2008. Recent developments in meta-analysis. *Stat Med*, 27, 625-50.
- TEAM, J. 2018. JASP. 0.9 ed. Netherlands.
- VOGEL, D., MARKL, A., YU, T., KOTCHOUBEY, B., LANG, S. & MULLER, F. 2013. Can mental imagery functional magnetic resonance imaging predict recovery in patients with disorders of consciousness? *Arch Phys Med Rehabil*, 94, 1891-8.
- VON WILD, K., LAUREYS, S. T., GERSTENBRAND, F., DOLCE, G. & ONOSE, G. 2012. The vegetative state--a syndrome in search of a name. *J Med Life*, 5, 3-15.
- WHITING, P. F., RUTJES, A. W., WESTWOOD, M. E., MALLETT, S., DEEKS, J. J., REITSMA, J. B., LEEFLANG, M. M., STERNE, J. A. & BOSSUYT, P. M. 2011. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*, 155, 529-36.
- ZASLER, N., H. ROSENBERG, J., GIACINO, J., KATZ, D., FILLEY, C. & KELLY, J. 2013. Hirschberg & Giacino VS MCS Dx Px Tx Neurol Clinics 2011.

7 Anhang

Im Anhang befindet sich eine Übersichtstabelle der verwendeten Artikel und der entsprechend notwendigen Daten.

Titel	Autoren und Veröffentlichungsjahr	Bias	Patientenanzahl	Impakt Faktor	Methode	Effektstärke
Implementing novel imaging methods for improved diagnosis of disorder of consciousness patients.	(Bick et al., 2013)	0	11 Patienten 6VS 5MCS	2,26 (2013)	fMRI	<p>Auditory: d = + 0,12 CI: -1,54 – 1,81 SE: 0,86</p> <p>Language: d = - 0,38 CI: -1,91 – 1,14 SE: 0,77</p> <p>Name: d = + 0,22 CI: -1,54 – 1,09 SE: 0,675</p> <p>Familiar voice: d = + 0,38 CI: -1,15 – 1,91 SE: 0,77</p> <p>Imaginary: d = + 0,60 CI: -1,96 – 0,75 SE: 0,69</p>
Thalamo-frontal connectivity mediates top-down cognitive functions in disorders of consciousness	(Monti et al., 2015)	2	24 Patienten 8 VS 16 MCS	8,18 (2014)	fMRI	<p>d = 0,22 CI: -0,94 – 1,39 SE: 0,59</p>
Willful modulation of brain activity in disorders of consciousness	(Monti et al., 2010)	3	46 Patienten 20 VS 26 MCS	53,48 (2010)	fMRI	<p>Motor Imaginary Task d = - 1,01 CI: -0,24 – 2,26</p>

						SE: 0,64 Spatial Imaginary Task: d = - 0,81 CI: -0,47 – 2,11 SE: 0,66
Quantitative rates of brain glucose metabolism distinguish minimally conscious from vegetative state patients.	(Stender et al., 2015)	5,5	35 Patienten 14 VS 21 MCS	5,40 (2014)	PET-CT	Cortex: d = + 1,01 CI: 0,30 – 1,71 SE: 0,36 Precuneus: d = + 1,14 CI: 0,42 – 1,85 SE: 0,36 Maximum Difference Map: d = + 1,38 CI: 0,64 – 2,12 SE: 0,37 Thalamus: d = + 0,35 CI: -0,31 – 1,02 SE: 0,34 Hirnstamm: d = + 0,24 CI: -0,421 – 0,91 SE: 0,33 Synaptische Verbindunge n: d = + 0,98 CI: 0,28 – 1,68 SE: 0,50

Reductions of thalamic volume and regional shape changes in the vegetative and the minimally conscious states.	(Fernandez-Espejo et al., 2010)	3	9 Patienten 4VS 5MCS	4,29 (2012)	MRT	Linke Hirnhälfte: d = +1,38 CI: -0,07 – 1,38 SE: 0,37 Rechte Hirnhälfte: d = + 0,98 CI: -0,38 – 2,38 SE: 0,70
Multimodal Study of Default-mode network Integrity in disorders of consciousness.	(Rosazza et al., 2016)	4,5	108 Patienten 72 VS 36 MCS	9,89 (2016)	fMRI, FDG- PET,	SMRI: d = + 0,82 CI: 0,45-1,20 SE: 0,19 rs-fMRI: d = + 0,56 CI: CI: 0,14 – 0,98 SE: 0,21 FDG - PET: d = + 0,95 CI: 0,51-1,40 SE: 0,23
Central olfactory processing in patients with disorders of consciousness.	(Nigri et al., 2016)	3	42 Patienten 32 VS 10 MCS	3,95 (2015)	fMRI	PC/AMY: Chi: 5,02 d = + 0,73 CI: 0,09 – 1,37 SE: 0,32 Höhere Ebene (IOFC): Chi: 2,63 d = + 0,44 CI: -0,17 – 1,06 SE: 0,31
Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: a clinical validation study.	(Stender et al., 2014)	4,5	122 Patienten 41 VS 81 MCS	45,21 (2014)	FDG - PET, fMRI	F-FDG PET: Chi-Quadrat: 45,9 d = + 1,51 CI: 1,12 – 1,93 SE: 0,22

						f-MRI: Chi- Quadrat: 9,26 d = + 0,51 CI: 0,19 – 0,93 SE: 0,18
Global functional connectivity reveals highly significant differences between the vegetative and the minimally conscious state.	(Kotchoubey et al., 2013)	2	12 Patienten 6VS 6MCS	3,57 (2012)	fMRI	cluster1: d = + 1,37 SE: 0,68 cluster2: d = + 1,48 SE: 0,76 cluster3. d = + 1,41 SE: 0,70 cluster4: d = + 1,37 SE: 0,68 cluster5: d = + 1,43 SE: 0,71 cluster6. d = + 1,33 SE: 0,68 cluster7: d = + 1,36 SE: 0,68
fMRI activation with an “affective speech” paradigm in vegetative and minimally conscious States: applicability and prognostic value.	(Piperno et al., 2012)	5	24 Patienten 16 VS 5 MCS	2,70 (2012)	fMRI	d = +0,41 CI: -0,40 – 1,23 SE. 0,41
Deactivation of the default mode network as a marker of impaired consciousness: an fMRI study.	(Crone et al., 2011)	4	25 Patienten 17 VS 8 MCS	4,09 (2011)	fMRI	d = + 0,27 CI: - 0,76 – 1,31 SE: 0,53

A network approach to assessing cognition in disorders of consciousness	(Rodriguez Moreno et al., 2010)	2	8 Patienten 3VS 5MCS	8,02 (2010)	fMRI	chi quadrat: 0,03 d = + 0,15 CI: -1,42 – 1,81 SE: 0,84
Diffusion weighted imaging distinguishes the vegetative state from the minimally conscious state.	(Fernandez-Espejo et al., 2011)	4	25 Patienten 10VS 15MCS	5,93 (2010)	MRT	Weißer Substanz: d = 0,39 CI: -0,39 – 1,18 SE: 0,40 Hirnstamm: d = 0,42 CI: -1,21 – 0,37 SE: 0,40 Thalamus: d = 0,57 CI: -1,37 – 0,22 SE: 0,41
Cerebral response to speech in vegetative und minimally conscious states after traumatic brain injury	(Fernandez-Espejo et al., 2008)	2	6 Patienten 3 VS 3 MCS	1,12 (2007)	fMRI	d = 0,32 CI = -1,87 – 1,87 SE: 0,95
Cerebral response to patient's own name in the vegetative and minimally conscious states.	(Di et al., 2007)	3	11 Patienten 7VS 4MCS	5,69 (2006)	fMRI	d = 0,77 CI: -2,04 – 0,49 SE: 0,64
A role for the Default Mode Network in the Base of Disorders of Consciousness	(Fernandez-Espejo et al., 2012)	3	46 Patienten 19 VS 27 MCS	11,19 (2012)	MRT	d = 0,69 CI: -1,29 - 0,17 SE: 0,36
The neural correlates of lexical processing in disorders of consciousness	(Nigri et al., 2017)	3	11 Patienten 4VS 7MCS	3,98 (2016)	fMRT/B OLD	d = 0,98 CI: -0,60 – 2,57 SE: 0,81

True or False? Activations of Language-related Areas in Patients with Disorders of Consciousness	(Kotchoubey et al., 2014)	5	55 Pat 29 VS 26 MCS	3,45 (2014)	fMRI Blutsauerstofflevel	Chi quadrat: 2,32 d = - 0,38 CI: -0,14 – 0,92 SE: 0,335
Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness	(Coleman et al., 2009)	4	41 Pat 22 VS 19 MCS	9,49 (2009)	fMRI	Chi quadrat: 8,05 d = + 0,98 CI: 0,30 – 1,66 SE: 0,34
Can mental imagery functional magnetic resonance imaging predict recovery in patients with disorders of consciousness?	(Vogel et al., 2013)	4	22 Pat 10 VS 12 MCS	2,44 (2013)	fMRI	Chi quadrat: 1,47 d = + 0,43 CI: -0,41 – 1,29 SE: 0,43
Brain response to thermal stimulation predicts outcome of patients with chronic disorders of consciousness	(Li et al., 2015)	4	22 Pat 12 VS 10 MCS	3,09 (2014)	fMRI	Chi quadrat: 1,56 d = - 0,43 CI: - 0,41 – 1,29 SE: 0,43
Degrees of functional connectome abnormality in disorders of consciousness	(Sinitsyn et al., 2018)	3,5	22 Pat 13 VS 9 MCS	3,14 (2017)	fMRI	Connectome intactness (ITCI) d = + 1,66 CI: 0,74 – 1 SE: 0,06 Hub disruption index d = + 1,04 CI: 0,53 – 1,00 SE: 0,12
Functional MRI motor Imaginary Task to detect command following in traumatic	(Bodien et al., 2017)	2,5	9 Pat VS 4 MCS 5	3,50 (2017)	fMRI	Handschütteln: d = + 0,38 CI: -1,97 – 1,20 SE: 0,812

disorders of consciousness						Tennis d = 0,15 CI: -1,57 – 1,89 SE: 0,44
Early detection of consciousness in patients with acute severe traumatic brain injury	(Edlow et al., 2017)	3,5	10 Pat 7 MCS 3 VS	10,84 (2017)	fMRI	Sprache: d = 0,64 CI: -0,34 – 1,23 SE 1,17 fMRI auf Musik: d = 0,42 CI = -1,87 – 2,3 SE = 1,16 Aktive Reaktion d = 0,20 CI = -0,50 – 1,86 SE = 1,15
Functional Evaluation of Awareness in Vegetative and Minimally Conscious State	(Marino et al., 2017)	2	50 Pat 23 VS 27 MCS	5,42 (2017)	fMRI	Diffuse signal: d = 0,15 CI = -0,94 – 1,39 SE = 0,31 Left frontal: d = 0,46 CI = -1,87 – 0,65 SE = 0,41

Tabelle 7: Übersicht aller verwendeten Studien

8 Erklärung zum Eigenanteil der Dissertationsschrift

Die Arbeit wurde am Institut für Medizinische Psychologie und Verhaltensneurobiologie unter Betreuung von Herrn Prof. Dr. Boris Kotchoubey durchgeführt.

Die Konzeption der Studie erfolgte in Zusammenarbeit mit Herrn Prof. Dr. Boris Kotchoubey.

Sämtliche Arbeitsschritte im Rahmen der Datenerhebung wurden eigenständig durch mich durchgeführt. Die statistische Auswertung erfolgte durch mich.

Ich versichere, das Manuskript selbstständig nach Anleitung durch Herrn Prof. Dr. Boris Kotchoubey verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.

Tübingen, den 17.09.2019

Benedikt Spiegelsberger

9 Danksagung

Ich danke ganz besonders Herrn Professor Dr. Boris Kotchoubey, meinem Doktorvater, für seine stete Geduld, Motivation und unbegrenzte Hilfe bei der Betreuung dieser Arbeit.

Ein besonderer Dank gilt meiner Familie und all Denen die zum Gelingen dieser Arbeit beigetragen haben.