

Aus der Klinik für Anästhesiologie m.S. operative Intensivmedizin (CCM/CVK) der
Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Klinische und ökonomische Effekte aus der Anwendung von Qualitätsindikatoren
in der Intensivmedizin am Beispiel der Überwachung von Schmerz, Sedierung und
Delir

zur Erlangung des akademischen Grades
Doctor rerum medicinalium (Dr. rer. medic.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Marc Deffland
aus Wildeshausen

Datum der Promotion: 04. März 2022

Inhaltsverzeichnis

ABKÜRZUNGSVERZEICHNIS.....	IV
TABELLEN- UND ABBILDUNGSVERZEICHNIS.....	V
ABSTRACT (deutsch)	VI
ABSTRACT (englisch).....	VII
1. Einleitung	1
1.1 Relevanz und Kosten der Intensivmedizin.....	1
1.2 Qualitätsindikatoren in der Intensivmedizin	2
1.2.1 Hintergrund	2
1.2.2 Schmerz-, Sedierungs- und Delirmanagement	3
1.2.3 Qualitätsindikatoren in der Praxis	5
1.3 Vergütung der Intensivmedizin in Deutschland.....	5
1.3.3 Vergütungssystem von Krankenhäusern.....	5
1.3.4 Besonderheiten der Intensivmedizin.....	7
1.4 Zielstellung der Arbeit.....	9
2. Material und Methoden.....	10
2.1 Studiendesign	10
2.2 Variablen.....	11
2.2.1 Unabhängige Variablen.....	11
2.2.2 Abhängige Variablen	12
2.3 Analyse	13
2.3.1 Klassische statistische Ansätze.....	14
2.3.2 Machine-Learning-Algorithmus und Zusammenhangsanalyse	14
3. Ergebnisse	16

3.1	Patient*innen-Charakteristik	16
3.2	Univariate und multiple lineare Analyse	17
3.2.1	Univariate Analyse.....	17
3.2.2	Multiple lineare Analyse	20
3.3	Boruta und MIC.....	21
3.2.1	Boruta.....	21
3.2.2	MIC.....	21
4	Diskussion.....	23
4.1	Interpretation der Ergebnisse	23
4.2	Stärken und Limitationen der Studie	26
5.	Schlussfolgerung.....	28
	LITERATURVERZEICHNIS.....	VII
	EIDESSTÄTLICHE VERSICHERUNG UND ANTEILSERKLÄRUNG	XII
	AUSZUG JOURNAL SUMMARY LIST	XV
	PUBLIKATION.....	XVII
	LEBENS LAUF.....	XXXI
	PUBLIKATIONS LISTE	XXXII
	DANKSAGUNG.....	XXXIII

ABKÜRZUNGSVERZEICHNIS

APACHE	Acute Physiology and Chronic Health Evaluation
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
BPS	Behavior Pain Scale
CAM-ICU	Confusion Assessment Method for the Intensive Care Unit
DIVI	Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin
DRG	Diagnosis Related Groups
ESICM	European Society of Intensive Care Medicine
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
ITS	Intensivstation
InEK	Institut für das Entgeltsystem im Krankenhaus
MIC	Maximal Information Coefficient
OPS	Operations- und Prozedurenschlüssel
RASS	Richmond Agitation Sedation Scale
VAS	Visual Analog Scale

TABELLEN- UND ABBILDUNGSVERZEICHNIS

Abbildung 1: Beispielhafter Auszug aus einem IneK-Definitionsbandbuch (Quelle: IneK).....	6
Abbildung 2: Auszug aus dem Fallpauschalenkatalog 2020 (Quelle: IneK)	6
Abbildung 3: Darstellung der Berechnung (Quelle: Eigene Darstellung)	11
Abbildung 4: Beispiel für eine Visual Agitation Scale (Quelle: Eigene Darstellung)	12
Abbildung 5: Consort-Flussdiagramm (Quelle: Eigene Darstellung)	16
Tabelle 1: Patient*innen-Charakteristik.....	17
Tabelle 2: Überwachung und klinische Effekte	19
Tabelle 3: Überwachung und ökonomische Effekte	19
Tabelle 4: Multiple lineare Regression.....	20
Abbildung 6: Ergebnisse Boruta and MIC (Quelle: Boruta und eigene Darstellung)	22

ABSTRACT (deutsch)

Background Die Intensivmedizin hat durch die Behandlung von lebensbedrohlich erkrankten Patient*innen hohe medizinische Relevanz. Die apparativen, personellen und organisatorischen Standards bedingen eine zusätzliche ökonomische Relevanz. Eine Maßnahme zur Sicherstellung einer optimalen Behandlung von Patient*innen sind Qualitätsindikatoren. Diese verlangen die Überwachung von Schmerz, Sedierung und Delir in jeder Schicht. Es wurde die These aufgestellt, dass die durchgängige Überwachung von Schmerz, Sedierung und Delir das klinische Outcome verbessert und das ökonomische Outcome aufgrund der fehlenden finanziellen Vergütung verschlechtert. Die Bestätigung der These könnte somit ein Indikator für Fehlanreize im Vergütungssystem sein.

Methoden Es handelt sich um eine retrospektive Studie unter Nutzung von Routinedaten. Es wurden univariate und multiple lineare Analysen sowie Machine-Learning-Algorithmen und eine moderne Zusammenhangsanalyse zum Verhältnis der Überwachung von Schmerz, Sedierung und Delir sowie klinischem und ökonomischem Outcome durchgeführt.

Ergebnisse Es wurden 1.323 erwachsene Patient*innen einbezogen, welche mindestens einen Tag zwischen Aufnahme und Entlassung vom 1. Januar bis 31. Dezember 2016 auf einer Intensivstation behandelt wurden. Eine durchgängige Überwachung von Schmerz, Sedierung und Delir war mit einer geringeren Krankenhaus-Verweildauer (bspw. Schmerz-Überwachung 13 vs. 10 Tage; $p < 0.001$), ITS-Verweildauer sowie Beatmungsdauer verbunden. Neben diesen positiven klinischen Effekten wurden negative ökonomische Effekte anhand von geringerem Case Mix pro Tag sowie Gewinn pro Tag festgestellt. Eine multiple lineare Analyse konnte alle genannten Ergebnisse nicht bestätigen. Weiterhin wurde auf Basis eines Machine-Learning-Algorithmus festgestellt, dass die Überwachung von Schmerz, Sedierung und Delir relevanter für klinisches und ökonomisches Outcome ist als der Schweregrad.

Schlussfolgerung Die Einhaltung der Leitlinien zur Schmerz-, Sedierungs- und Delir-Überwachungen auf Intensivstationen sind relevant für das klinische und ökonomische Outcome. In der univariaten Analyse ist die durchgängige Überwachung verbunden mit besserem klinischen und schlechterem ökonomischen Outcome. Dies konnte jedoch mit der multiplen linearen Analyse nicht bestätigt werden. Es besteht die Notwendigkeit sich stärker mit (Fehl-)Anreizen im Vergütungssystem auseinander zu setzen.

ABSTRACT (englisch)

Background Intensive care medicine has a high relevance due to the treatment of threateningly ill patients. It is therefore also resource intensive. Because of this medical and economic background, it is necessary to adhere to quality indicators. One quality indicator requires the monitoring of all cases of pain, sedation and delirium in every shift. It has been hypothesized that continuous monitoring of pain, sedation and delirium improves clinical outcome and reduces economic values due to the lack of financial compensation. Confirmation of the thesis could thus be an indicator of false incentives in the compensation system.

Methods This is a retrospective study using routine data. Univariate and multiple linear analyses were used as well as a machine-learning algorithm and a modern correlation analysis to analyze the relationship between monitoring of pain, sedation and delirium and clinical and economic outcome.

Results 1,323 adult patients were included. These patients were treated in an intensive care unit for at least one day between admission and discharge from January 1 to December 31, 2016. Continuous monitoring of pain, sedation and delirium was found with a lower hospital length of stay (e.g. pain monitoring 13 vs. 10 days; $p < 0.001$), ICU length of stay and ventilation duration based on univariate analysis. In addition to the positive clinical effects, negative economic effects were determined based on a lower case mix per day and profit per day based on a univariate analysis. A multiple linear analysis could not confirm these results. Furthermore, based on a machine-learning algorithm, monitoring of pain, sedation and delirium is more relevant for clinical and economic outcome than severity.

Conclusion Adherence with guidelines for pain, sedation and delirium monitoring in intensive care units is relevant for clinical and economic outcome. In univariate analysis, continuous monitoring is associated with better clinical and worse economic outcome. However, this could not be confirmed by the multiple linear analysis. A further analysis of the incentives in reimbursement strategies is necessary.

1. Einleitung

1.1 Relevanz und Kosten der Intensivmedizin

Auf intensivmedizinischen Stationen eines Krankenhauses werden lebensbedrohlich oder kritisch erkrankte Patient*innen gepflegt, therapiert und überwacht. Ziele der Intensivmedizin sind daher eine niedrige Rate an (vermeidbaren) Todesfällen, der Erhalt von Organfunktionen, Lebensqualität und Alltagsfunktionalität bei gleichzeitig möglichst kurzer Beatmungsdauer.[1]

Die Intensivmedizin benötigt dafür einen umfangreichen apparativen Standard, hoch ausgebildetes Personal und aufeinander abgestimmte Abläufe.

Diese hohen Anforderungen an eine professionelle Organisation, spiegeln sich auch in einem hohen Ressourcenverbrauch durch Intensivstationen wider: Intensivstationen repräsentieren nach einer nordamerikanischen Studie 13,3% der Kosten eines Krankenhauses, 4,2% der nationalen Gesundheitsausgaben sowie 0,56% des Bruttoinlandsproduktes und verursachen damit die höchsten Kosten pro Patient*in und Behandlungstag.[2]

Ressourcen im Gesundheitswesen sind begrenzt und müssen effizient genutzt werden, um eine optimale Gesundheitsvorsorge zu gewährleisten. Zur Erreichung des Ziels einer effizienten Ressourcennutzung, sollten entsprechende Anreize im jeweiligen System integriert sein.

Die intensivmedizinische Behandlung hat zusammenfassend sowohl medizinisch als auch ökonomisch eine hohe Relevanz für jedes einzelne Krankenhaus sowie das gesamte Gesundheitssystem.

Eine Maßnahme zur medizinisch optimalen Behandlung von Patient*innen auf Intensivstationen können Qualitätsindikatoren sein.

1.2 Qualitätsindikatoren in der Intensivmedizin

1.2.1 Hintergrund

Unter Qualitätsindikatoren werden Kennzahlen verstanden, aus denen sich indirekt Qualität abbilden lässt. Qualitätsindikatoren sollen die Integration von evidenzbasierten Leitlinien in die klinischen Prozesse durch regelmäßige Evaluationen unterstützen.[3]

Bereits im Jahr 2001 wurde der Gemeinsame Bundesausschuss (G-BA) im Rahmen des Gesetzes zur Modernisierung der gesetzlichen Krankenversicherung (GKV-Modernisierungsgesetz) mit der Organisation und Steuerung einer externen Qualitätssicherung mithilfe von Qualitätsindikatoren in definierten Leistungsbereichen beauftragt. Die Teilnahme an diesem Qualitätssicherungsverfahren ist für Krankenhäuser verpflichtend. Zu den definierten Leistungsbereichen gehören beispielsweise die Geburtshilfe, Orthopädie, Hygiene sowie Kardiologie und Herzchirurgie. Im Rahmen der Geburtshilfe wird als Qualitätsindikator beispielsweise die Anwesenheit eines Pädiaters bei Frühgeborenen mit einem Gestationsalter von 24+0 bis unter 35+0 ins Verhältnis zu allen Geburten von Frühgeborenen mit einem Gestationsalter von 24+0 bis unter 35+0 gesetzt.

Die Intensivmedizin wird von diesen bundesweit verpflichtenden Qualitätsindikatoren nicht als Leistungsbereich erfasst. Das erste Set von Qualitätsindikatoren in der Intensivmedizin in Deutschland wurde 2010 vom Verein „Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin (DIVI)“ veröffentlicht. Bereits zuvor gab es Impulse der Johns Hopkins University in Baltimore. Bis heute wurde das Set von Qualitätsindikatoren in der Intensivmedizin bereits zweimal, zuletzt im Jahr 2017, überarbeitet. Die Nutzung dieser Qualitätsindikatoren ist gesetzlich nicht vorgeschrieben.

Das Set der Organisation „Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin (DIVI)“ enthält die folgenden zehn Qualitätsindikatoren:

- I. Tägliche multiprofessionelle und interdisziplinäre klinische Visite mit Dokumentation von Tageszielen
- II. Management von Sedierung, Analgesie und Delir
- III. Patientenadaptierte Beatmung
- IV. Frühzeitige Entwöhnung von einer invasiven Beatmung (Weaning)
- V. Überwachung der Maßnahmen zur Infektionsprävention
- VI. Maßnahmen zum Infektionsmanagement
- VII. Frühe enterale Ernährung
- VIII. Dokumentation einer strukturierten Patienten- und Angehörigenkommunikation
- IX. Frühmobilisation
- X. Leitung der Intensivstation

In dieser Studie wird der Qualitätsindikator „Management von Analgesie, Sedierung, Delir“ und dessen klinische und ökonomische Effekte analysiert.

Der Qualitätsindikator „Management von Analgesie, Sedierung, Delir“ basiert auf der aktuell publizierten S3-Leitlinie „Analgesie, Sedierung und Delirmanagement in der Intensivmedizin“ der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. (AWMF).

1.2.2 Schmerz-, Sedierungs- und Delirmanagement

Intensivmedizinisch behandelte Patient*innen leiden häufig unter Schmerzen, die ein wesentlicher Stressfaktor sind und zu Unruhe führen.[4, 5] Daher ist die Schmerztherapie ein erster Ansatzpunkt, um die Sedierung der Patient*innen zu minimieren. Dieser Ansatz findet breite Zustimmung, beispielweise im Rahmen der Konsensusempfehlungen der European Society of Intensive Care Medicine (ESICM), die ein frühzeitiges, protokollbasiertes Management von „Wachheit/minimaler Sedierung“ (sogenanntes eCASH concept) empfehlen.[6]

Eine zu tiefe Sedierung von Patient*innen auf der Intensivstation ist mit unerwünschten Ergebnissen verbunden. Dazu gehören eine längere Beatmungsdauer und ein erhöhtes Mortalitätsrisiko. [7, 8] Daher wird ein Ansatz empfohlen, der die Sedierung auf ein Minimum reduziert.

Das Delir ist eine häufig auftretende Hirnfunktionsstörung bei schwerkranken Patient*innen, welche mit unerwünschten Ergebnissen verbunden ist.[9, 10] Zu den unerwünschten Ereignissen gehören längere Verweildauern im Krankenhaus sowie auf der Intensivstation (ITS) und eine langfristige kognitive Beeinträchtigung der Patient*innen.[11] Die Entwicklung des Delirs bei Patient*innen auf der Intensivstation kann durch eine hohe Volatilität in der Sedierung sowie eine nicht adäquate Schmerzbehandlung begünstigt werden.[12, 13]

Die dargestellte Begünstigung des Delirs durch unzureichendes Schmerz- und Sedierungsmanagement sowie der aufgeführte Zusammenhang von Schmerztherapie und Sedierung deuten bereits darauf hin, dass die drei Werte nicht unabhängig voneinander betrachtet werden können. Vielmehr sind diese daher als Teil eines Gesamtkonzeptes zu verstehen.[14]

Der Qualitätsindikator „Management von Analgesie, Sedierung, Delir“ des DIVI erfasst daher alle Werte in einem Indikator. Der Indikator ist zweigeteilt. Zum einen wird geprüft, ob eine Verfahrensanweisung vorliegt, die Schmerz-, Sedierungs- und Delir-Management erfasst. Diese Ausprägung des Qualitätsindikators kann lediglich mit ja oder nein beantwortet werden. Zum anderen soll das Sedierungs- und Schmerzniveau sowie die Präsenz eines Delirs anhand validierter Skalen, alle acht Stunden oder wenn sich die klinische Situation verändert, überwacht werden. Somit kann die Ausprägung des Qualitätsindikators, insbesondere bei längeren ITS-Verweildauern, sehr unterschiedlich ausfallen.

1.2.3 Qualitätsindikatoren in der Praxis

Obwohl die Überwachung von Schmerz, Sedierung und Delir inhärenter Bestandteil der klinischen Praxis sein soll, zeigen Studien geringe Implementierungsquoten.[15–18] Diese empirischen Ergebnisse demonstrieren die Herausforderung der Umsetzung von evidenzbasierten Leitlinien in die klinische Praxis.[19] Sich mit den Implementierungsstrategien und Implementierungshindernissen beschäftigende Studien befassen sich daher mit Schulungen, gewohntem Verhalten und bestehenden Organisationsstrukturen.[20–22]

Ein weiterer Grund für geringe Implementierungsraten könnte das Fehlen eines finanziellen Anreizes für das Krankenhaus in dem zugrundeliegenden Vergütungssystem des jeweiligen Landes sein.

1.3 Vergütung der Intensivmedizin in Deutschland

1.3.3 Vergütungssystem von Krankenhäusern

Die Vergütung (der Betriebskosten) von Krankenhausleistungen erfolgt in Deutschland seit 2004 über ein leistungsorientiertes und pauschalisiertes Vergütungssystem auf Basis von Diagnosis Related Groups (DRG).

Zuvor wurden Krankenhäuser auf Basis von abteilungsspezifischen, kostendeckenden tagesgleichen Pflegesätzen vergütet. Ziel der Umstellung auf das DRG-System war die Herstellung von Transparenz über Leistungen und Kosten, die Einführung einer strikt leistungsorientierten Vergütung, höherer Wettbewerb, Reduzierung der Verweildauern, Ausschöpfung von Wirtschaftlichkeitsreserven sowie eine längerfristige Stabilisierung der Ausgaben der gesetzlichen Krankenkassen.[23]

In einem DRG-System werden Pauschalen pro Fall berechnet. Die Pauschalen entstehen durch die Multiplikation eines bundeslandesweit geltenden Basisfallwertes mit einem

Relativgewicht. Das Relativgewicht wird im Wesentlichen auf Grundlage der Hauptdiagnose ermittelt. Daneben werden auch weitere Variablen, wie beispielsweise Verweildauer, Nebendiagnose, Operations- und Prozedurenschlüssel (OPS), Geschlecht und Alter einbezogen. Die daraus resultierende Klassifikation wird durch einen DRG-Grupper vorgenommen, der jeden Fall (anhand der in mehreren tausendseitigen Definitionshandbücher beschriebenen Algorithmen) prüft. Diese Definitionshandbücher werden vom Institut für das Entgeltsystem im Krankenhaus (InEK) herausgegeben.

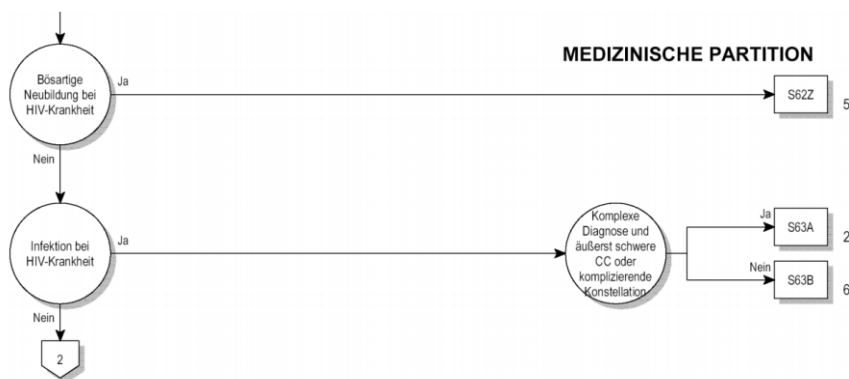


Abbildung 1: Beispielhafter Auszug aus einem IneK-Definitionshandbuch (Quelle: IneK)

Der in Abbildung 1 dargestellte Algorithmus kann in den rechts abgedruckten DRGs münden. Alle DRGs sind in der Regel nach dem Muster: Buchstabe – Zahl mit zwei Ziffern – Buchstabe aufgeteilt. Wie in Abbildung 1 zu sehen, beispielsweise „S63B“. Die Hauptdiagnosegruppe „S“ steht für ‚HIV‘, die Zahl 63 für die Behandlungsart bzw. Partition ‚M = medizinische Fallpauschale‘ und die Endung „A“ oder „B“ sagt etwas über den Schweregrad aus. Wird der Schweregrad nicht weiter unterteilt, wird der Buchstabe „Z“ genutzt (siehe Abbildung 1, oben)

Die Höhe des genauen Relativgewichtes einer DRG kann dem jeweiligen Fallpauschalenkatalog des Jahres entnommen werden (Abbildung 2).

DRG	Parti-tion	Bezeichnung ⁶⁾	Bewertungsrelation bei Hauptabteilung
1	2	3	4
S62Z	M	Bösartige Neubildung bei HIV-Krankheit	0,932
S63A	M	Infektion bei HIV-Krankheit mit komplexer Diagnose und äußerst schweren CC oder mit komplizierender Konstellation	3,070
S63B	M	Infektion bei HIV-Krankheit ohne komplexe Diagnose oder ohne äußerst schwere CC, ohne komplizierende Konstellation	1,275

Abbildung 2: Auszug aus dem Fallpauschalenkatalog 2020 (Quelle: IneK)

Eine Behandlung mit der DRG ‚S63A‘ hat somit ein Relativgewicht (Case Mix) in Höhe von 3,070. Durch die Multiplikation des Landesbasisfallwertes Berlin in Höhe von 3.659,00EUR müssten somit 11.233,13EUR in Rechnung gestellt werden.

Für die Abrechnung ist es unerheblich, ob der Patient 7 oder 42 Tage im Krankenhaus lag. Sollten entsprechend alle Patient*innen eines hypothetischen Krankenhauses eine Liegedauer von 7 Tagen haben, könnte dieses Krankenhaus nach 42 Tagen ca. 67 TEUR in Rechnung stellen. Ein Krankenhaus mit einer Verweildauer in Höhe von 42 Tagen entsprechend nur ca. 11 TEUR. Lediglich unterhalb von 7 und überhalb von 43 Tagen wäre die Abrechnung mit Ab- oder Zuschlägen verbunden.

Die Höhe der jeweiligen Bewertungsrelation einer DRG ergibt sich auf Basis der kalkulierten Gesamtkosten je Fall in ca. 250 Krankenhäusern in Deutschland, die diese Kalkulationen für das InEK durchführen. Die kalkulierte Fallmenge liegt bei ca. 4 Millionen Fällen.

Neben den DRGs erhalten Krankenhäuser bei besonderen Krankheitsbildern (bspw. für Bluter) weiterhin Zusatzentgelte. Daneben sieht das deutsche Vergütungssystem eine Vielzahl von Zu- und Abschlägen vor. Hierzu gehört beispielsweise ein DRG-Systemzuschlag, der der Pflege und Weiterentwicklung des DRG-Systems dient. Sowohl Zusatzentgelte als auch Zu- und Abschläge können jedoch aufgrund des Umfangs nicht weiter ausgeführt werden.

Weiterhin ist zu berücksichtigen, dass ab dem 01. Januar 2020 durch Krankenhäuser ein individuelles Pflegebudget nach dem Selbstkostendeckungsprinzip vereinbart wird. Diese sind damit der oben genannten leistungsorientierten Vergütung entzogen worden.

1.3.4 Besonderheiten der Intensivmedizin

Da Intensivpatient*innen durch das DRG-System schwer abbildbar waren, wurde dieses beispielsweise in Australien (dem Ursprungsland der deutschen DRGs) nicht für die Finanzierung der Intensivmedizin verwendet.[23] Hintergrund für die schwere Abbildung ist,

das intensivmedizinische Leistungen häufig nicht an spezielle Hauptdiagnosen gekoppelt sind. Dennoch hat man sich jedoch für die Abbildung der Intensivmedizin innerhalb des deutschen DRG-Systems entschieden.

Anfangs erfolgte die spezielle Vergütung dieser Patient*innen zunächst auf Grundlage der Beatmungsdauer. Hierfür wurde an den Anfang des in Kapitel 1.3.1 dargestellten Algorithmus die Beatmungsdauer eingesetzt, sodass Fälle mit einer Beatmungsdauer von mehr als 95 Stunden grundsätzlich in die sogenannten A-DRGs (beispielsweise A04A: „Beatmung > 1799 Stunden mit intensivmedizinischer Komplexbehandlung > 2940 / 5520 / 5520 Aufwandspunkte oder mit hochkomplexem Eingriff“) fallen.

Die Beatmungsdauer wurde 2005 ergänzt durch einen Operationen- und Prozedurenschlüssel (OPS 8-980 Komplexbehandlung Intensivmedizin (Basisprozedur)). [23] Die Kodierung dieses OPS ist abhängig von verschiedenen Strukturkriterien, bspw. einer 24-Stunden-Überwachung sowie der Behandlungsleitung durch Fachärzt*innen mit Zusatzweiterbildung „Intensivmedizin“.

Im Jahr 2013 beantragte das DIVI den OPS-Code „OPS 8-98f“ für die aufwendige intensivmedizinische Komplexbehandlung, der weitergehende Mindestanforderungen an die abrechnenden Intensivstationen im Vergleich zum OPS 8-890 stellt.

Hintergrund der Beantragung der genannten OPS-Kodierungen war jeweils die Schaffung eines eigenen Vergütungsbestandteils für Krankenhäuser mit sehr hohen Vorhaltekosten und sehr schweren Fällen, sog. Maximalversorger. In beiden Fällen ist dieses Ziel nicht erreicht worden, da diese Kodierungen mittlerweile auch von vielen kleineren Krankenhäusern genutzt werden.[23]

1.4 Zielstellung der Arbeit

Wie in Kapitel 1.3 dargestellt, gibt es keine finanziellen Anreize für die Anwendung der Qualitätsindikatoren in der Intensivmedizin. Durch die Vergütung langer Beatmungszeiten besteht ein (selbstverständlich ausschließlich) finanzieller Anreiz eher in Richtung von schlechter Qualität.

Die Anpassung von Anreizen ist ein relevantes Thema im deutschen Gesundheitswesen: Beispielsweise mit der Einführung des DRG-Systems sollten auch Verweildauern verkürzt werden. Die Krankenhäuser wurden nicht mehr anhand belegter Betten mit Tagespauschalen, sondern mit einer Behandlungspauschale für die jeweilige Diagnose bzw. Krankheit finanziert. Mit dieser Umstellung war jedoch der (negative) Anreiz von zu kurzen Verweildauern, sogenannter ‚blutige Entlassungen‘ verbunden. Entsprechend stand der Gesetzgeber vor der Herausforderung einen Ausgleich herzustellen. Die Forderung nach der Einbeziehung der Qualität erbrachter Leistungen in die Vergütung medizinischer Leistungen wird seit Langem gefordert. Erst 2016 wurden jedoch durch den deutschen Gesetzgeber erste Grundlagen für eine qualitätsabhängige Vergütung mit Zu- und Abschlägen gelegt.

Daher wird zusammenfassend die These aufgestellt, dass Qualitätsindikatoren zwar zu besserem klinischen jedoch zu schlechterem ökonomischen Outcome in einem sehr stark DRG-orientierten Gesundheitssystem führen und damit Fehlanreize im aktuellen System gesetzt werden. Zur Messung des ökonomischen Outcomes wurden dabei sowohl der Case Mix pro Tag als auch die Differenz von DRG-Erträgen und Aufwendungen (Gewinn des Krankenhauses pro Tag) nach den standardisierten Berechnungen des Instituts für das Entgeltsystem im Krankenhaus (InEK) herangezogen.

2 Material und Methoden

2.1 Studiendesign

Wir haben an der Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin mit Standorten am Charité Campus Mitte sowie dem Campus Virchow Klinikum eine retrospektive Kohortenstudie durchgeführt. Hierfür wurden alle mit einer DRG abgerechneten Patient*innen mit Aufnahme und Entlassung im Jahr 2016 einbezogen, die nicht bereits Teil einer Studie waren. Hintergrund hierfür war, dass Patient*innen mit Studienteilnahme die Überwachung von Delir-, Schmerz und Sedierung im Rahmen der Studie erhalten und daher kein Vergleich von zwei Gruppen bzw. metrische Zusammenhangsanalysen hätte durchgeführt werden können.

Es wurden Patient*innen unter 18 Jahren sowie im Krankenhaus verstorbene Patient*Innen, Patient*innen mit keinem dokumentierten Tag zwischen Aufnahme und Entlassung sowie Patient*innen mit internen Intensivstations-Verlegungen ausgeschlossen. Gründe für die Ausschlüsse waren die Besonderheiten der Kinder- und Jugendmedizin, mögliche Verzerrungen von zeitlichen Variablen durch Berücksichtigung verstorbener Patient*innen sowie die Auswertungsverzerrungen durch Verlegungen. Weiterhin konnten für die Delir-Auswertung einige Patient*innen nicht einbezogen werden, da die Messung aufgrund physiologischer Einschränkungen (aufgrund einer zu starken Sedierung) nicht möglich war.

Vor der Datenauswertung wurde die Studie bei der Ethikkommission der Charité eingereicht und dort freigegeben (EA 2/092/14). Weiterhin wurde die Studie bei clinicaltrials.gov unter der Nummer NCT 02265263 registriert.

Die verwendeten Daten sind Routinedaten der klinischen Versorgung, die den Krankenhaus-Informationssystemen COPRA, Berlin und SAP, Walldorf entnommen wurden.

2.2 Variablen

2.2.1 Unabhängige Variablen

Die wesentlichen unabhängigen Variablen waren die Einhaltungquoten der jeweiligen Überwachung gemäß dem Qualitätsindikator „Monitoring von Schmerz, Sedierung und Delir“ der DIVI (siehe Kapitel 1.2.2).

Der Qualitätsindikator wurde für die Auswertung von Schmerz, Sedierung und Delir separiert erhoben, sodass für alle Patient*innen jeweils ein individueller Prozentwert für die Schmerz-, Sedierungs- und Delir-Überwachung vorliegt.

Patient*innen erreichten einen Wert von 100%, wenn diese während ihres gesamten Aufenthalts in jeder Schicht (alle 8 Stunden) mindestens einmal überwacht wurden. Jede Schicht ohne durchgeführte Überwachung reduzierte den Prozentwert entsprechend anteilig. Der Tag der Aufnahme und Entlassung wurde jeweils nicht mit einbezogen.

Wenn ein Fall somit zusätzlich zu Aufnahme- und Entlasstag zwei Tage intensivmedizinisch behandelt wurde und an diesen zwei Tagen in fünf von sechs Schichten überwacht wurde, hat dieser einen Compliance-Wert von 83,30% (siehe Abbildung 3).

Aufnahmetag			Aufnahmetag + 1			Aufnahmetag + 2			Entlasstag (Aufnahmetag + 3)		
Schicht 1	Schicht 2	Schicht 3	Schicht 1	Schicht 2	Schicht 3	Schicht 1	Schicht 2	Schicht 3	Schicht 1	Schicht 2	Schicht 3
			ja	ja	nein	ja	ja	ja			

▶ 83,30%

Abbildung 3: Darstellung der Berechnung (Quelle: Eigene Darstellung)

In der klinischen Praxis werden die drei Werte Schmerz, Sedierung und Delir nach einem Algorithmus erhoben: Im ersten Schritt wird der Sedierungs-Zustand der Patient*innen anhand eines Score-Systems geprüft. Als Score wird die Richmond Agitation Sedation Scale (RASS) verwendet. Diese reicht von +4 (aggressiv) über 0 (wach und ruhig) hin zu -5 (nicht weckbar). Bei Patient*innen mit einem RASS von -3 oder höher wird das Vorhandensein eines Delirs geprüft. Hierfür wird die Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) verwendet. Im letzten Schritt erfolgt die Evaluierung des Schmerzes. Hierfür wird den

Patient*innen eine Schmerz-Skala gezeigt, auf der diese die Intensität ihrer Schmerzen einschätzen müssen. Als Skala hierfür dient die Visual Agitation Scale (VAS).

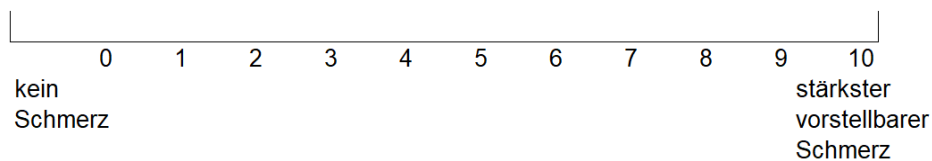


Abbildung 4: Beispiel für eine Visual Agitation Scale (Quelle: Eigene Darstellung)

Wenn Patient*innen hierzu aufgrund eines Delirs oder zu starker Sedierung nicht in der Lage waren, wurde die Schmerzintensität mit der Behavioural Pain Scale (BPS) erhoben. Hierbei handelt es sich um eine Methode der Fremdeinschätzung des Schmerzzustandes anhand von Gesichtsausdruck (entspannt bis Grimassieren), Bewegung der oberen Extremität (keine Bewegung bis ständiges Anziehen) und Anpassung an das Beatmungsgerät (Toleranz bis kontrollierte Beatmung nicht möglich).

Neben Schmerz-, Sedierungs- und Delir-Überwachung, wurden für die multiple lineare Analyse sowie den Machine-Learning-Algorithmus Boruta weitere Variablen ausgewertet. Diese waren das Alter und der Krankheits-Schweregrad der Patient*innen. Zur Bewertung des Schweregrades wurde der Acute Physiology and Chronic Health Evaluation II (folgend APACHE II oder APACHE) genutzt. Beide Variablen wurden aufgrund der Relevanz für die abhängigen Variablen Schmerz-, Sedierungs- und Delir-Überwachung auf Grundlage von klinischer Erfahrung ausgewählt.

2.2.2 Abhängige Variablen

Als abhängige Variablen wurden für das klinische Outcome die Krankenhaus-Verweildauer, ITS-Verweildauer sowie die (maschinelle) Beatmungsdauer herangezogen.

Für das ökonomische Outcome wurden das DRG-Relativgewicht pro Krankenhaus-Verweildauertag (Case Mix pro Tag) sowie der Gewinn pro Krankenhaus-Verweildauertag (Gewinn pro Tag) genutzt. Die Berechnung einer relativen Kennzahl für die ökonomischen

Werte wurde gewählt, da davon ausgegangen wird, dass jedes Bett auf der Intensivstation durchgängig belegt werden kann.

Das DRG-Relativgewicht pro Krankenhaus-Verweildauertag ist eine Umsatzgröße, die Relevanz für den Marktanteil eines Krankenhauses hat. Gewinn pro Krankenhaus-Verweildauertag ist eine weitere Kennzahl mit wesentlicher Anreizfunktion für das Management von Krankenhäusern. Der Gewinn pro Tag wurde durch die Subtraktion der Fallerlöse und Aufwendungen ermittelt. Die Aufwendungen bzw. Kosten pro Fall wurden vom Controlling der Charité ermittelt, da die Charité im Jahr 2016 eins von bundesweit 254 sogenannten Kalkulationskrankenhäusern des InEK war. Die Kalkulationskrankenhäuser unterstützen das InEK bei der Pflege und Weiterentwicklung des deutschen DRG-Systems durch die Erhebung von Aufwendungen bzw. Kosten nach Vorgaben des InEK.

2.3 Analyse

Es wurden sowohl klassische statistische Ansätze (univariate und multiple lineare Analyse) als auch ein Machine-Learning-Algorithmus (Boruta) sowie ein moderner Ansatz für die Zusammenhangsanalyse (MIC) verwendet.

Während im Rahmen der univariaten Analyse mithilfe des Mann-Whitney-U-Test ausschließlich getestet wird, ob die zentralen Tendenzen zweier unabhängiger Stichproben verschieden sind, beschreibt die multiple lineare Analyse lineare Zusammenhänge und Korrelationen. Die Nutzung eines nicht-linearen Machine-Learning-Algorithmus beschreibt hingegen auch andersartige Zusammenhänge.

Weiterhin kann der Boruta-Algorithmus auch Prädiktor für weitere (bisher nicht erhobene) Daten sein. Dies begründet sich darin, dass der Boruta-Algorithmus eine Kreuzvalidierung durchführt. Im Rahmen dieser Kreuzvalidierung wird im ersten Schritt ein Modell auf Basis

einer Teilmenge der Daten erstellt und folgend die Restmenge beurteilt. So wird die Wahrscheinlichkeit der Erfassung von Fehlkorrelationen verringert.

2.3.1 Klassische statistische Ansätze

Für die klassischen statistischen Analysen wurde SPSS Version 24.0.0.0 (IBM SPSS Statistics) verwendet. Im Rahmen der univariaten Analyse haben wir für die Durchführung des Mann-Whitney-U-Testes die Patient*innen in zwei Gruppen aufgeteilt: Zum einen die Gruppe mit einer durchgängigen Überwachung in allen Schichten während der gesamten IST-Verweildauer (100%) und eine weitere Gruppe mit allen Werten unterhalb von 100%. Um zwischen unterschiedlichen Schweregraden zu differenzieren und hieraus resultierende Verzerrungen zu vermeiden, sind die Gruppen weiter unterteilt worden in eine leicht-, mittelschwer- und schwerkranke Gruppe (APACHE ≤ 10 , APACHE 11-20 sowie APACHE > 20). Weiterhin wurde der Zusammenhang zwischen der Überwachungsintensität bzw. Einhaltung des Qualitätsindikators sowie der klinischen und ökonomischen Werte anhand eines multiplen linearen Regressionsmodells überprüft.

2.3.2 Machine-Learning-Algorithmus und Zusammenhangsanalyse

Für den Machine-Learning-Algorithmus sowie die moderne Zusammenhangsanalyse wurden RStudio Version 1.1.419 (R-Foundation for Statistical Computing) und das „Boruta“ Paket verwendet.

Boruta ist eine random-forest-basierte Klassifikationsmethode zur Identifikation relevanter Variablen. Ein Random Forest[24] ist ein Modell, das zahlreiche (häufig tausende) Entscheidungsbäume auf Basis der eingepflegten Daten erstellt, um Vorhersagen treffen zu können. Diese Modelle können nicht-lineare und nicht-monotone Beziehungen zwischen den unabhängigen und abhängigen Variablen erfassen, die mit linearen Modellen nicht erfasst werden könnten.

Der Boruta-Algorithmus baut die Random Forests aus dem Datensatz auf. Dann mischt er die Werte jeder Variablen nach dem Zufallsprinzip, eine nach der anderen. Damit prüft Boruta, ob die Klassifizierungsleistung des Random Forest abnimmt, wenn der potenzielle statistische Zusammenhang einer unabhängigen Variablen mit dem Kriterium durch diesen Prozess eliminiert wird. Diese Methode ist auch als permutationsbasierte Merkmalsrelevanz bekannt.

Der Boruta-Algorithmus sagt ausschließlich etwas über die relative Stärke der Relevanz einzelner unabhängiger Variablen für die jeweilige abhängige Variable aus. Daher kann die Stärke der Relevanz zwischen verschiedenen abhängigen Variablen nicht direkt verglichen werden. Um die Stärke der Relevanz der Variablen zu untersuchen, wurde zusätzlich der Maximal Information Coefficient (MIC) [25] genutzt. Wie ein Random Forest ist MIC nicht limitiert auf spezielle Funktionstypen (linear, non-linear oder non-monoton). MIC Werte liegen zwischen 0 und 1 und können identisch zu R-Quadrat interpretiert werden.

3 Ergebnisse

3.1 Patient*innen-Charakteristik

Es wurden 1.323 Patient*innen für Schmerz- und Sedierungs-Monitoring sowie 1.266 Patient*innen für Delir-Monitoring einbezogen (siehe Abbildung 5). Die Gruppengröße divergiert, da für die Differenz keine Testung auf ein Delir aufgrund einer zu tiefen Sedierung stattfinden konnte.

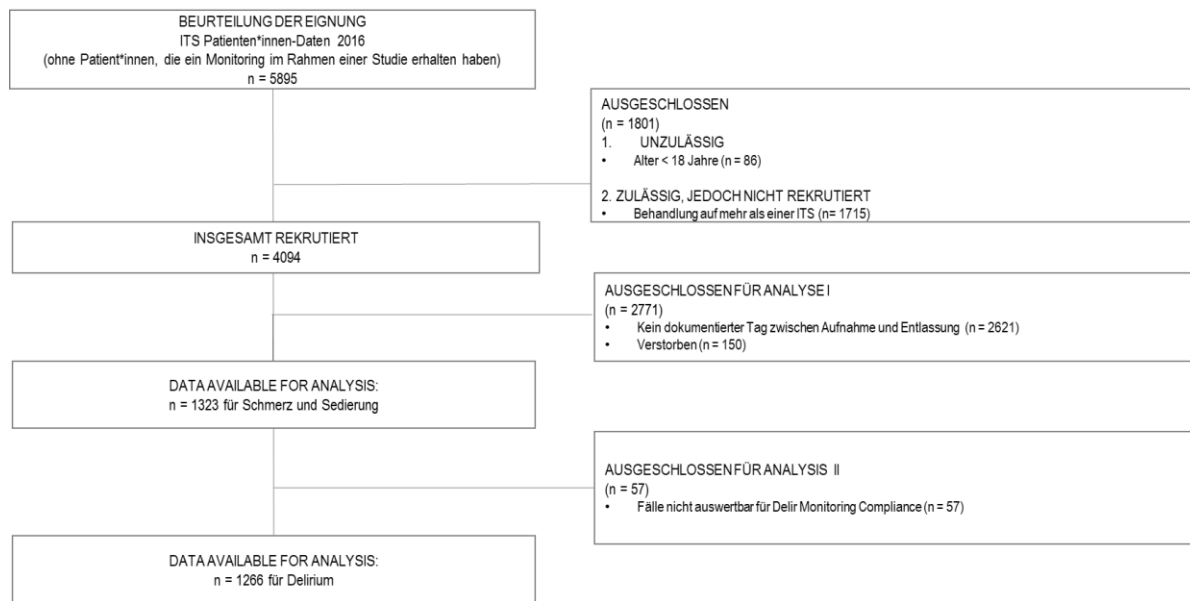


Abbildung 5: Consort-Flussdiagramm (Quelle: Eigene Darstellung)

Das Medianalter der Patient*innen betrug 68 Jahre, 84,4% wurden im Rahmen des stationären Aufenthaltes operiert und 60% waren männlich. Die Intensivstations(ITS)-Verweildauer betrug im Median 4 Tage, die Krankenhaus-Verweildauer hingegen 12 Tage und die (maschinelle) Beatmungsdauer betrug 32 Stunden. Das Schmerz-Monitoring lag im Median bei 91,7%, das Sedierungs-Monitoring bei 90,4% und das Delir-Monitoring bei 100% (Tabelle 1). Die ökonomischen Größen lagen bei 0,44 Case Mix pro Tag und 11,05 EUR Gewinn pro Tag im Median. Der Schweregrad (APACHE II-Score) bei Aufnahme betrug im Durchschnitt (Median) 14, was einer wahrscheinlichen Überlebensrate von 85% entspricht.

Tabelle 1: Patient*innen-Charakteristik

Charakteristik	(n = 1323)
Operation (%)	84,4
Männlich (%)	60,0
Alter, Jahre [Median(IQR)]	68(56 - 76)
Krankenhaus Verweildauer, Tage [Median(IQR)]	12(7 -19)
Verweildauer Intensivstation, Tage [Median(IQR)]	4(3 - 8)
Beatmungsdauer, Stunden [Median(IQR)]	32(0.0 -109)
Schmerz Monitoring Überwachung, % [Median(IQR)]	91,7(80,9 - 100)
% [100% erfüllt]	60,7
Sedierungs Monitoring Überwachung, % [Median(IQR)]	90,4(76,2 to 100)
% [100% erfüllt]	67,0
Delir Monitoring Überwachung, % [Median(IQR)]	100(92,3 - 100)
% [100% erfüllt]	64,6
Gewinn pro Tag, EUR [Median(IQR)]	11,05(-240,6 - 445,3)
DRG-Relativgewicht pro Tag, Punkte [Median(IQR)]	0,44(0,26 - 0,71)
APACHE II bei Aufnahme, Punkte [Median(IQR)]	14(9 - 22)

APACHE II = Schweregrad der Erkrankung, IQR = Interquartilsabstand (25. und 75. Perzentile),
DRG = Diagnosis-Related-Group

3.2 Univariate und multiple lineare Analyse

3.2.1 Univariate Analyse

Für Schmerz-Monitoring (Tabellen 2 und 3) sind, in den meisten Fällen, statistisch signifikante Verbesserungen der klinischen Werte und schlechtere ökonomische Werte über alle Krankheits-Schweregrade festgestellt worden. Ausnahmen waren der Case Mix pro Tag in der APACHE>20-Gruppe, Gewinn pro Tag in der APACHE<=10-Gruppe und der APACHE11-20-Gruppe sowie Gewinn pro Tag in der APACHE>20-Gruppe. Alle Ausnahmen waren statistisch nicht signifikant.

Für Sedierungs-Monitoring (Tabellen 2 und 3) sind in den meisten Fällen statistisch signifikante Verbesserungen der klinischen Werte und schlechtere ökonomische Werte über

alle Krankheits-Schweregrade festgestellt worden. Ausnahmen waren die Krankenhaus- sowie ITS-Verweildauer in der APACHE \leq 10-Gruppe, Gewinn pro Tag in der APACHE11-20-Gruppe sowie Gewinn pro Tag in der APACHE $>$ 20-Gruppe. Alle Ausnahmen waren statistisch nicht signifikant.

Für Delir-Monitoring (Tabellen 2 und 3) sind in den meisten Fällen statistisch signifikante Verbesserungen der klinischen Werte und schlechtere ökonomische Werte über alle Krankheits-Schweregrade festgestellt worden. Ausnahmen waren die Krankenhaus-Verweildauer in der APACHE \leq 10-Gruppe, Case Mix pro Tag in der APACHE $>$ 20-Gruppe, Gewinn pro Tag in der APACHE \leq 10-Gruppe, der APACHE 11-20-Gruppe und der APACHE $>$ 20-Gruppe. Alle Ausnahmen waren statistisch nicht signifikant.

Tabelle 2: Überwachung und klinische Effekte

			Fälle	Krankenhausverweildauer		Verweildauer Intensivstation		Beatmungsdauer	
			Nr.	Median (IQR)	P	Median (IQR)	P	Median (IQR)	P
Einhaltung Überwachung Schmerz	Alle	nicht erreicht	803	13*(8 - 22)	<0,001	6*(4 - 11)	<0,001	60*(0 - 167)	<0,001
		erreicht	520	10*(6 - 15)	<0,001	3*(2 bis 4)	<0,001	0*(0 - 55,75)	<0,001
	APACHE <=10	nicht erreicht	230	11*(19 - 7)	0,003	5*(8 - 3)	<0,001	29*(0 - 107,75)	<0,001
		erreicht	189	9*(6 - 14)	0,003	2*(2 - 4)	<0,001	0*(0 - 23,50)	<0,001
	APACHE 11-20	nicht erreicht	319	13*(8 - 22)	<0,001	6*(3 - 10)	<0,001	48*(0 - 134)	<0,001
		erreicht	212	10*(7 - 15)	<0,001	3*(2 - 4)	<0,001	0*(0 - 55)	<0,001
APACHE >20	nicht erreicht	254	16*(9 - 25)	<0,001	7*(5 - 17)	<0,001	122,5*(34,75 - 316)	<0,001	
	erreicht	119	10*(7 - 15)	<0,001	4*(3 - 5)	<0,001	45*(0 - 86)	<0,001	
Einhaltung Überwachung Sedierung	Alle	nicht erreicht	886	13*(8 - 21)	<0,001	5*(3 - 10)	<0,001	51*(0 - 149)	<0,001
		erreicht	437	10*(6 - 15,5)	<0,001	3*(2 - 4)	<0,001	0*(0 - 59)	<0,001
	APACHE <=10	nicht erreicht	270	10(6 - 18)	0,161	4(2 - 8)	<0,001	3*(0 - 92,75)	<0,001
		erreicht	149	10(6 - 14)	0,161	3(2 - 4)	<0,001	<0,001*(0 - 29)	<0,001
	APACHE 11-20	nicht erreicht	349	13*(8 - 21)	<0,001	5*(3 - 9)	<0,001	39*(0 - 123)	<0,001
		erreicht	182	10*(7 - 17,25)	<0,001	3*(2 - 4)	<0,001	<0,001*(0 - 60,25)	<0,001
APACHE >20	nicht erreicht	267	16*(10 - 25)	<0,001	7*(5 - 16)	<0,001	110*(38 - 302)	<0,001	
	erreicht	106	9,5*(6,75 - 14,25)	<0,001	4*(2,75 - 6)	<0,001	34*(0 - 86)	<0,001	
Einhaltung Überwachung Delir	Alle	nicht erreicht	448	15*(8 - 24)	<0,001	7*(4 - 13,75)	<0,001	85*(8 - 222)	<0,001
		erreicht	818	11*(7 - 17)	<0,001	4*(2 - 6)	<0,001	16*(0 - 76)	<0,001
	APACHE <=10	nicht erreicht	116	11(6 - 21)	0,050	6*(3 - 11)	<0,001	54,5*(0 - 146,75)	<0,001
		erreicht	278	10(6 - 15)	0,050	3*(2 - 5)	<0,001	0*(0 - 49,25)	<0,001
	APACHE 11-20	nicht erreicht	183	14*(8 - 23)	0,001	6*(4 - 12)	<0,001	69*(0 - 176)	<0,001
		erreicht	326	11*(8 - 17)	0,001	4*(2 - 6)	<0,001	11*(0 - 69,5)	<0,001
APACHE >20	nicht erreicht	149	18*(10 - 27,5)	<0,001	9*(6 - 20)	<0,001	149*(59 - 444,5)	<0,001	
	erreicht	214	12*(7 - 17,25)	<0,001	5*(3-7)	<0,001	51*(1 - 112,75)	<0,001	

*= P<0.05 (Mann-Whitney-U-Test), APACHE II = Schweregrad der Erkrankung, IQR = Interquartilsabstand (25. und 75. Perzentile)

Tabelle 3: Überwachung und ökonomische Effekte

			Fälle	Case Mix pro Tag		Gewinn pro Tag	
			Nr.	Median (IQR)	P	Median (IQR)	P
Einhaltung Überwachung Schmerz	Alle	nicht erreicht	803	0,48*(0,29 - 0,73)	<0,001	27,40*(-223,61 - 496,30)	0,015
		erreicht	520	0,39*(0,23 - 0,64)	<0,001	-33,08*(-270,44 - 370,62)	0,015
	APACHE <=10	nicht erreicht	230	0,43*(0,25 - 0,70)	0,005	-21,12(-286,75 - 511,47)	0,055
		erreicht	189	0,32*(0,21 - 0,60)	0,005	-113,22(-280,76 - 205,22)	0,055
	APACHE 11-20	nicht erreicht	319	0,45*(0,265 - 0,70)	0,038	-9,75(-269,19 - 445,29)	0,469
		erreicht	212	0,36*(0,23 - 0,64)	0,038	-28,31(-274,76 - 419,81)	0,469
APACHE >20	nicht erreicht	254	0,55(0,36 - 0,80)	0,102	124,70(-138,42 - 615,73)	0,679	
	erreicht	119	0,49(0,31 - 0,71)	0,102	137,73(-207,29 - 532,74)	0,679	
Einhaltung Überwachung Sedierung	Alle	nicht erreicht	886	0,49*(0,29 - 0,77)	<0,001	50,33*(-217,26 - 534,16)	<0,001
		erreicht	437	0,34*(0,22 - 0,61)	<0,001	-71,39*(-278,35 - 302,10)	<0,001
	APACHE <=10	nicht erreicht	270	0,44*(0,26 - 0,73)	<0,001	26,11*(-252,61 - 586,55)	<0,001
		erreicht	149	0,30*(0,19 - 0,51)	<0,001	-144,31*(-344,96 - 74,24)	<0,001
	APACHE 11-20	nicht erreicht	349	0,47*(0,26 - 0,73)	<0,001	8,72(-243,62 bis 459,53)	0,134
		erreicht	182	0,34*(0,21 - 0,60)	<0,001	-77,43(-286,23 - 327,93)	0,134
APACHE >20	nicht erreicht	267	0,57*(0,38 - 0,79)	0,019	130,49(-141,11 - 643,58)	0,526	
	erreicht	106	0,46*(0,30 - 0,71)	0,019	139,49(-180,65 - 503,26)	0,526	
Einhaltung Überwachung Delir	Alle	nicht erreicht	448	0,52*(0,32 - 0,78)	<0,001	45,02(-230,23 - 481,08)	0,13
		erreicht	818	0,39*(0,24 - 0,66)	<0,001	-9,23(-252,77 - 387)	0,13
	APACHE <=10	nicht erreicht	116	0,49*(0,28 - 0,81)	<0,001	-45,40(-299,38 - 501,20)	0,325
		erreicht	278	0,33*(0,21 - 0,60)	<0,001	-70,95(-303,41 - 280,49)	0,325
	APACHE 11-20	nicht erreicht	183	0,47*(0,28 - 0,73)	0,004	5,24(-251,34 - 477,18)	0,172
		erreicht	326	0,37*(0,22 - 0,62)	0,004	-42,63(-277,46 - 350,86)	0,172
APACHE >20	nicht erreicht	149	0,55(0,38 - 0,79)	0,196	116,41(-156,88 - 518,23)	0,321	
	erreicht	214	0,51(0,32 - 0,76)	0,196	155,17(-132,74 - 644,72)	0,321	

*= P<0.05 (Mann-Whitney-U-Test), APACHE II = Schweregrad der Erkrankung, IQR = Interquartilsabstand (25. und 75. Perzentile)

3.2.2 Multiple lineare Analyse

Die multiple lineare Analyse hat verschiedene Zusammenhänge aufgezeigt (Tabelle 4). Ein höherer Schweregrad (APACHE II) war mit einer höheren Krankenhaus-Verweildauer verbunden. Eine höhere Sedierungs-Überwachung und Schweregrad (APACHEII) war mit einer höheren ITS-Verweildauer verbunden. Eine höhere Schmerz-Überwachung war mit einer niedrigeren ITS-Verweildauer verbunden. Eine höhere Sedierungs-Überwachung und höherer Schweregrad (APACHEII) war mit einer höheren Beatmungsdauer verbunden. Ein höheres Alter war mit einer niedrigeren Beatmungsdauer verbunden. Eine höhere Schmerz-Überwachung, Alter und Schweregrad (APACHEII) waren mit einem höheren Case Mix pro Tag verbunden. Eine höhere Delir-Überwachung war mit einem niedrigeren Case Mix pro Tag verbunden. Eine höhere Schmerz-Überwachung, Alter und Schweregrad (APACHEII) war mit einem höheren Gewinn pro Tag verbunden. Das angepasste R-Quadrat war für alle Regressionen sehr gering.

Tabelle 4: Multiple lineare Regression

	Klinische Effekte									Ökonomische Effekte					
	Krankenhaus VWD (angepasstes R-Quadrat = .021; F = 6.409; p < .001)			ITS VWD (angepasstes R-Quadrat = .047; F = 13.413; p < .001)			Beatmungsdauer (angepasstes R-Quadrat = .059; F = 16.907; p < .001)			Case Mix pro Tag (angepasstes R-Quadrat = .042; F = 12.037; p < .001)			Gewinn pro Tag (angepasstes R-Quadrat = .020; F = 6.062; p < .001)		
	B	95% CI	p	B	95% CI	p	B	95% CI	p	B	95% CI	p	B	95% CI	p
Überwachung Schmerz	-0,018	-0,041 bis 0,005	0,430	-0,045	-0,063 bis -0,027	0,013	-0,693	-1,11 bis -0,275	0,097	0,170	0,109 bis 0,231	0,005	2,894	1,720 bis 4,068	0,014
Überwachung Sedierung	0,022	0,002 bis 0,042	0,263	0,049	0,033 bis 0,065	0,002	1,011	0,642 bis 1,38	0,006	-0,103	-0,157 bis -0,05	0,055	-0,576	-1,614 bis 0,461	0,578
Überwachung Delir	-0,010	-0,032 bis 0,013	0,660	-0,006	-0,023 bis 0,012	0,751	-0,598	-1,011 bis -0,185	0,148	-0,194	-0,254 bis -0,133	0,001	-0,363	-1,525 bis 0,799	0,755
Schweregrad (APACHEII)	0,220	0,18 bis 0,261	0,000	0,228	0,196 bis 0,26	0,000	6,313	5,563 bis 7,062	0,000	0,505	0,396 bis 0,614	0,000	8,413	6,305 bis 10,521	0,000
Alter	-0,028	-0,051 bis -0,005	0,230	-0,014	-0,032 bis 0,004	0,443	-0,955	-1,381 bis -0,53	0,025	0,176	0,114 bis 0,238	0,005	2,431	1,233 bis 3,629	0,043

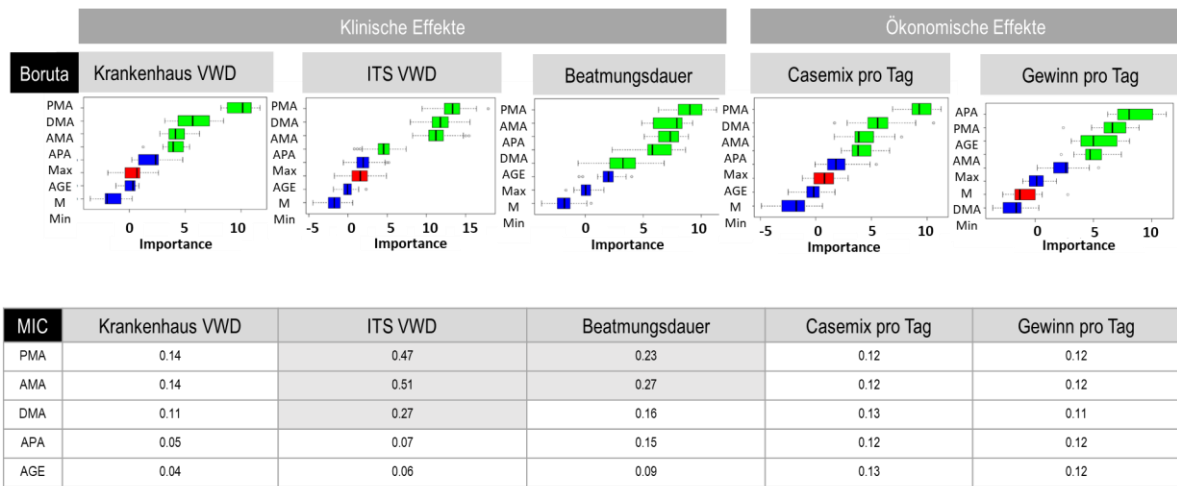
3.3 Boruta und MIC

3.2.1 Boruta

Die Machine-Learning-Analyse zeigte, dass Schmerz-Monitoring die wichtigste Prädiktor-Variable für klinische Werte (Krankenhaus-Verweildauer, ITS-Verweildauer sowie Beatmungsdauer) sowie den Case Mix pro Tag sind (Boruta in Abbildung 2). Weiterhin ist die Sedierungs-Überwachung relevanter für klinische Werte und den Case Mix pro Tag als der Schweregrad (APACHE). APACHE war relevanter für den Gewinn pro Tag als andere Variablen. Delir-Überwachung war relevant für klinische und ökonomische Werte und relevanter als der APACHE für die Krankenhaus-Verweildauer, ITS-Verweildauer sowie den Case Mix pro Tag, jedoch nicht für die Beatmungsdauer. Delir-Überwachung kann keine Vorhersage für den Gewinn pro Tag liefern. Das Alter ist relevant für die Beatmungsdauer, aber weniger relevant als Schmerz-, Sedierungs- und Delir-Überwachung sowie den Schweregrad. Das Alter war auch relevant für den Gewinn pro Tag, jedoch weniger relevant als der Schweregrad sowie Schmerz- und Sedierungs-Überwachung.

3.2.2 MIC

Der stärkste Zusammenhang wurde zwischen der Überwachung von Schmerz, Sedierung sowie Delir und der ITS Verweildauer mit den MIC-Werten 0,47, 0,51 und 0,27 identifiziert (MIC in Abbildung 2). Ein stärkerer Zusammenhang als 0,2 wurde zwischen Schmerz- und Sedierungs-Überwachung mit der Beatmungsdauer festgestellt. Die Stärke des Zusammenhangs zwischen Schmerz-, Sedierungs- sowie Delir-Überwachung mit Krankenhausverweildauer war kleiner als 0,15. Die Stärke des Zusammenhangs zwischen Schmerz-, Sedierungs- und Delir-Überwachung mit Case Mix pro Tag war kleiner als 0,14.



PMA = Schmerz Überwachung, DMA = Delir Überwachung, AMA = Sedierung Überwachung, APA = APACHE bei Aufnahme, AGE = Alter bei Aufnahme, Max = shadowMax, M = shadowMean, Min = shadowMin, MIC = maximal information coefficient, Background with shadow = MIC > 0,2

Abbildung 6: Ergebnisse Boruta and MIC (Quelle: Boruta und eigene Darstellung)

4 Diskussion

4.1 Interpretation der Ergebnisse

Die Studie hat gezeigt, dass die Einhaltung von Leitlinien im Hinblick auf Schmerz-, Sedierungs- und Delir-Überwachung mit besseren klinischen Werten (Krankenhaus- und ITS-Verweildauer, Beatmungsdauer) und schlechterem ökonomischen Outcome (Case Mix pro Tag und Gewinn pro Tag) verbunden ist.

Dieses Ergebnis konnte jedoch nicht in der multiplen linearen Regression bestätigt werden. Die multiple lineare Regression hat jedoch aufgrund eines sehr geringen R-Quadrats keine starke Aussagekraft. Ein möglicher Grund für den wenig ausgeprägten Effekt in der Studie kann die insgesamt hohe Leitlinien-Einhaltung in der analysierten Klinik sein. Desto stärker die Überwachung von Schmerz, Sedierung und Delir bereits in klinischen Prozessen integriert ist, umso geringer könnte der damit verbundene Effekt ausfallen.

Zusätzlich zur klassischen statistischen Analyse, wurde der Machine-Learning-Algorithmus Boruta sowie ein moderner Zusammenhangskoeffizient (MIC) verwendet. Der Boruta-Algorithmus hat identifiziert, dass die Überwachung von Schmerz, Sedierung und Delir sowohl für klinische als auch für ökonomische Werte relevant ist. Schmerz-Überwachung war der wichtigste Prädiktor für klinische Werte und den Case Mix pro Tag. Die Analyse hat gezeigt, dass Sedierungs- und Delir-Überwachung in einigen Fällen weniger relevant als Schmerz-Überwachung ist, jedoch meistens relevanter als der Schweregrad (APACHEII) und das Alter der Patient*innen. Während das Alter insgesamt irrelevant war, zeigte der Boruta-Algorithmus die Relevanz des Schweregrads als Prädiktor für klinische und ökonomische Werte. Die klinische Perspektive des APACHE deckt sich mit Studien, wonach der Schweregrad im Zusammenhang mit Krankenhaus- und ITS-Verweildauern sowie Sterberaten steht.[26]

Während Boruta ausschließlich die relative Relevanz der verschiedenen unabhängigen Variablen für die jeweilige abhängige Variable beschreibt, analysierte MIC die Stärke des Zusammenhangs. Der stärkste Zusammenhang wurde zwischen Schmerz-, Sedierungs- und Delir-Überwachung mit der ITS-Verweildauer festgestellt.

Die Ergebnisse stehen im Einklang zu den meisten vorangegangenen Studien, die eine positive Verknüpfung von Leitlinien-Einhaltung und verbesserten klinischen Werten festgestellt haben. Beispielsweise haben Luetz und Kolleg*innen eine Verbindung zwischen Delirium-Monitoring und Krankenhaussterblichkeit für beatmete Patient*innen festgestellt, Mansouri et al. fanden eine substantielle Reduzierung der Beatmungsdauer, ITS-Verweildauer und Sterblichkeit durch protokollbasiertes Schmerz-, Sedierungs- sowie Delir-Management und Dale et al. fanden eine Reduzierung des Auftretens eines Deliriums, der Beatmungsdauer sowie von Krankenhaus- und ITS-Verweildauer.[27–29]

Neben den positiven klinischen Effekten zeigten sich anhand des deutschen DRG-Systems in der univariaten Analyse negative ökonomischen Effekte anhand des Case Mix pro Tag sowie des Gewinns pro Tag. Anhand der multiplen linearen Analyse konnten diese Ergebnisse jedoch nicht bestätigt werden. Dabei muss jedoch berücksichtigt werden, dass das angepasste R-Quadrat der multiplen Analysen in allen Fällen sehr gering war.

Die Ergebnisse geben damit Anlass, sich mit der Vergütung von Krankenhausleistungen auseinander zu setzen. Eine aktuelle Studie zeigt auf, dass Art und Höhe der Vergütung einen starken Einfluss auf die ausgewählte Behandlungsstrategie haben [30] und unterstreicht damit die Notwendigkeit weitergehender Überlegungen und Studien.

Die Relevanz des DRG-Systems ist in Deutschland höher als in anderen Gesundheitssystemen: Während in Deutschland ca. 80% der Krankenhausleistungen anhand von DRG finanziert werden, ist der Anteil mit 40 bis 50% in anderen Ländern deutlich

geringer.[31] Damit unterscheidet sich das deutsche System deutlich von anderen Gesundheitssystemen, welche zusätzlich Sicherstellungszuschläge, Zuschläge für Notversorgung und besondere Regelungen für Maximalversorger enthalten.

In Deutschland entstehen die Relativgewichte durch die Kalkulation von Behandlungsfällen durch sogenannte Kalkulationskrankenhäuser ohne Berücksichtigung der jeweils dort umgesetzten Qualitätsstandards. Dies führt zu dem Risiko eines „Race-to-the-bottom“, hin zum günstigsten (und nicht qualitativ besten) Krankenhaus.[1]

Bis 2020 gab es keine Zuschläge für Fixkosten und Qualitätsmaßnahmen bzw. Abschläge für nicht umgesetzte Qualitätsmaßnahmen. Die Diskussion um Fehlanreize hat jedoch in den vergangenen Jahren so stark zugenommen, dass der deutsche Gesetzgeber bereits mit ersten Maßnahmen reagiert hat. Beispielweise wurden die Kosten für Pflegekräfte aus dem DRG-System herausgelöst, separiert budgetiert und werden aktuell auf Basis des Selbstkostendeckungsprinzips finanziert.

Weiterhin hat der deutsche Gesetzgeber mit dem Krankenhausstrukturgesetz (KHSG) im Jahr 2016 beschlossen, dass der Gemeinsame Bundesausschuss (G-BA) einen Katalog von Leistungen beschließen soll, die sich für eine qualitätsabhängige Vergütung mit Zu- und Abschlägen eignen.

Zusätzlich werden seit Juli des Jahres 2019 Qualitätsverträge getestet, die mögliche Verbesserungen durch die Vereinbarung von höherwertigen Qualitätsanforderungen und Anreizen erproben. Hierin sind bereits Besonderheiten der Intensivmedizin als Leistungsbereiche (Prävention des postoperativen Delirs bei der Versorgung von älteren Patient*innen und Respiratorentwöhnung von langzeitbeatmeten Patient*innen) enthalten.

Mit den Reformen der letzten Jahre hat der deutsche Gesetzgeber damit begonnen sich von einem reinen DRG-System zu trennen und auch andere Vergütungsbestandteile stärker in das

bestehende Modell zu integrieren. Dabei ist zu beachten, dass jedes Vergütungssystem verschiedene Ziele verfolgt und Anreize setzt. Es gibt kein System, das keine Fehlanreize schafft. Herausforderung für den Gesetzgeber und die beteiligten Stakeholder wird es sein, Ziele und Anreize fortlaufend angemessen anzupassen. Hierfür wird eine begleitende Forschung mit medizinischem und ökonomischem Fokus notwendig sein, die viele emotionale und interessengesteuerte Argumentationen falsifizieren, relativieren oder verifizieren kann.

4.2 Stärken und Limitationen der Studie

Diese Arbeit ist die erste Studie, die einen Machine-Learning-Algorithmus zur Bewertung der Relevanz von Schmerz-, Sedierungs- und Delir-Überwachung auf klinische und ökonomische Effekte genutzt und damit eine neue Perspektive in die Diskussion gebracht hat. Weiterhin ist es die erste Studie, die auch ökonomische Effekte anhand von Case Mix pro Tag und Gewinn pro Tag aus der Perspektive des Krankenhauses untersucht.

Diese Studie hat verschiedene Limitationen. Die erste Limitation besteht in der Nutzung von Routinedaten aus klinischen EDV-Systemen. Aufgrund von fortlaufenden Validierungsprüfungen, kann die Datenqualität zwar insgesamt als hoch, unkorrekte oder unvollständige Einzelwerte jedoch nicht ausgeschlossen werden.

Eine weitere Limitation kann in der Fokussierung von ökonomischen Effekten auf Case Mix und Gewinn pro Tag gesehen werden. Die Gesamtkosten wurden nicht separiert analysiert. Bei der identifizierten geringeren ITS- und Krankenhaus-Verweildauer sinken wahrscheinlich auch die Gesamtkosten. Die ökonomische Sichtweise basiert jedoch auf einer optimalen Ressourcennutzung des Krankenhauses sowie deren ökonomisches Anreizsystem, welches stärker durch die analysierten Variablen beeinflusst wird.

Weiterhin ist der Case Mix ein sehr kontextabhängiges System. Die Auswertung könnte in einem anderen Jahr aufgrund der permanenten Systemänderungen (insbesondere durch Änderungen der Bewertung und des Kodier-Algorithmuses) anders ausfallen.

Die Analyse der ökonomischen Effekte gilt ausschließlich für Deutschland. Eine Aussage über andere Gesundheitssysteme kann nicht getroffen werden. Zwar haben viele Länder auch ein auf DRG basiertes System, nutzen jedoch eigene Datenbanken und unterschiedliche Zusatzvergütungen oder Abschläge.

Die genannte Studie zum Zusammenhang von APACHE und ITS-Verweildauer verwendete den APACHE IV, während das Studienkrankenhaus APACHE II nutzt.[26] Daneben gibt es weitere Faktoren, neben Schmerz-, Sedierungs- und Delir-Überwachung, welche mit klinischen oder ökonomischen Effekten verbunden sind. Diese Faktoren konnten jedoch nicht quantifiziert werden.

Eine weitere Limitation ist die Gruppenaufteilung der univariaten Analyse. Alle Patient*innen mit einer geringeren Überwachungsquote als 100% wurden als "Qualitätsindikator nicht erfüllt" betrachtet. Der Grund hierfür lag darin, dass das Studienkrankenhaus über eine hohe Implementierungsquote verfügt. Die nationale Leitlinie der DIVI führt aus, dass das Monitoring in mindestens 70% aller Schichten stattfinden muss.[32] Die 70%-Grenze ist aus anderen Kontexten als effektive Größe für das Erreichen eines signifikanten Effektes im Rahmen der Einführung von Qualitätsindikatoren bekannt.[33] Unser Ergebnis zeigt die Möglichkeit weiterer Studien im Hinblick auf die Notwendigkeit höherer Zielvorgaben.

Zielvorgaben sind aktuell weiterhin auch abhängig von den jeweils genutzten Leitlinien. Beispielsweise verlangt die spanische Leitlinie alle sechs Stunden ein Monitoring, jedoch ausschließlich für maschinell beatmete Patient*innen, und gibt einen Zielwert von 95% vor.[34] Unabhängig von der genauen Ausgestaltung des Managements von Schmerz, Sedierung und

Delir in den einzelnen Leitlinien, verbessern diese grundsätzlich die Sensibilität der Beschäftigten für dieses Thema und unterstützen damit die Verbesserung der Qualität in intensivmedizinischen Einrichtungen.

Weiterhin ist anzumerken, dass Schmerz-, Sedierungs- und Delir-Management separiert ausgewertet wurden. Eingangs wurde dargestellt, dass es sich dabei um ein Gesamtkonzept handelt und eine vollständig isolierte Betrachtung nicht sinnvoll ist. Aus Gründen der Transparenz ist jedoch die separierte Auswertung bevorzugt worden.

Es besteht Bedarf weitere prospektive Studien sowohl zur Schmerz-, Sedierungs- und Delir-Überwachung als auch zu den Effekten anderer Qualitätsindikatoren durchzuführen, um die Ergebnisse dieser Studie zu validieren sowie einen stärkeren Fokus auf das ökonomische Outcome zur Verbesserung der Anreize im DRG-System zu entwickeln. Zukünftige Studien sollten weiterhin Kohorten mit stärkeren Differenzen in der Überwachungsintensität zur Bestätigung dieser Studie berücksichtigen.

5. Schlussfolgerung

Die Einhaltung der Leitlinien zur Schmerz-, Sedierungs- und Delir-Überwachungen auf Intensivstationen sind relevant für das klinische und ökonomische Outcome. In der univariaten Analyse ist die durchgängige Überwachung mit besserem klinischem und schlechterem ökonomischem Outcome verbunden. Dies konnte jedoch mit der multiplen linearen Analyse nicht bestätigt werden.

LITERATURVERZEICHNIS

1. Riessen R, Hermes C, Bodmann KF, Janssens U, Markewitz A. [Reimbursement of intensive care services in the German DRG system: Current problems and possible solutions]. *Medizinische Klinik, Intensivmedizin und Notfallmedizin*. 2018;113:13–23. doi:10.1007/s00063-017-0390-x.
2. Halpern NA, Pastores SM, Greenstein RJ. Critical care medicine in the United States 1985-2000: an analysis of bed numbers, use, and costs. *Critical care medicine*. 2004;32:1254–9.
3. Kumpf O. Qualitätsindikatoren in der Intensivmedizin : Hintergrund und praktischer Nutzen. *Medizinische Klinik, Intensivmedizin und Notfallmedizin* 2019. doi:10.1007/s00063-019-00630-w.
4. Novaes MA, Aronovich A, Ferraz MB, Knobel E. Stressors in ICU: patients' evaluation. *Intensive care medicine*. 1997;23:1282–5.
5. Mehta S, Spies C, Shehabi Y. Ten tips for ICU sedation. *Intensive care medicine*. 2018;44:1141–3. doi:10.1007/s00134-017-4992-9.
6. Vincent J-L, Shehabi Y, Walsh TS, Pandharipande PP, Ball JA, Spronk P, Longrois D, Strøm T, Conti G, Funk G-C, Badenes R, Mantz J, Spies C, Takala J. Comfort and patient-centred care without excessive sedation: the eCASH concept. *Intensive care medicine*. 2016;42:962–71. doi:10.1007/s00134-016-4297-4.
7. Balzer F, Weiss B, Kumpf O, Treskatsch S, Spies C, Wernecke KD, Krannich A, Kastrup M. Early deep sedation is associated with decreased in-hospital and two-year follow-up survival. *Critical care (London, England)*. 2015;19:197. doi:10.1186/s13054-015-0929-2.
8. Tanaka LMS, Azevedo LCP, Park M, Schettino G, Nassar AP, Réa-Neto A, Tannous L, Souza-Dantas VC de, Torelly A, Lisboa T, Piras C, Carvalho FB, Maia MdO, Giannini FP, Machado FR, Dal-Pizzol F, Carvalho AGR de, dos Santos RB, Tierno PFGMM, Soares M, Salluh JIF. Early sedation and clinical outcomes of mechanically ventilated patients: a prospective multicenter cohort study. *Critical care (London, England)*. 2014;18:R156. doi:10.1186/cc13995.

9. Ely EW, Shintani A, Truman B, Speroff T, Gordon SM, Harrell FE, Inouye SK, Bernard GR, Dittus RS. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. *JAMA*. 2004;291:1753–62. doi:10.1001/jama.291.14.1753.
10. Girard TD, Jackson JC, Pandharipande PP, Pun BT, Thompson JL, Shintani AK, Gordon SM, Canonico AE, Dittus RS, Bernard GR, Ely EW. Delirium as a predictor of long-term cognitive impairment in survivors of critical illness. *Critical care medicine*. 2010;38:1513–20. doi:10.1097/CCM.0b013e3181e47be1.
11. Slooter AJC, van de Leur RR, Zaal IJ. Delirium in critically ill patients. *Handb Clin Neurol*. 2017;141:449–66. doi:10.1016/B978-0-444-63599-0.00025-9.
12. Svenningsen H, Egerod I, Videbech P, Christensen D, Frydenberg M, Tønnesen EK. Fluctuations in sedation levels may contribute to delirium in ICU patients. *Acta Anaesthesiol Scand*. 2013;57:288–93. doi:10.1111/aas.12048.
13. Zoremba N, Coburn M, Schälte G. Delir beim Intensivpatienten : Eine multiprofessionelle Herausforderung. *Anaesthesist* 2018. doi:10.1007/s00101-018-0497-3.
14. Spies Cea, editor. *SOPs in Intensivmedizin und Notfallmedizin*: Thieme; 2013.
15. Luetz A, Balzer F, Radtke FM, Jones C, Citerio G, Walder B, Weiss B, Wernecke KD, Spies C. Delirium, sedation and analgesia in the intensive care unit: a multinational, two-part survey among intensivists. *PloS one*. 2014;9:e110935. doi:10.1371/journal.pone.0110935.
16. Saller T, V VD, Hofmann-Kiefer K. [Knowledge and implementation of the S3 guideline on delirium management in Germany]. *Der Anaesthesist*. 2016;65:755–62. doi:10.1007/s00101-016-0218-8.
17. Trupkovic T, Kinn M, Kleinschmidt S. Analgesia and sedation in the intensive care of burn patients: results of a European survey. *Journal of intensive care medicine*. 2011;26:397–407. doi:10.1177/0885066610393442.
18. Wang J, Peng ZY, Zhou WH, Hu B, Rao X, Li JG. A National Multicenter Survey on Management of Pain, Agitation, and Delirium in Intensive Care Units in China. *Chinese medical journal*. 2017;130:1182–8. doi:10.4103/0366-6999.205852.

19. Rundshagen I. [Pain, agitation and delirium. Amended 2013 guidelines of the American College of Critical Care Medicine]. *Der Anaesthesist*. 2013;62:914–8. doi:10.1007/s00101-013-2232-4.
20. Balas MC, Burke WJ, Gannon D, Cohen MZ, Colburn L, Bevil C, Franz D, Olsen KM, Ely EW, Vasilevskis EE. Implementing the awakening and breathing coordination, delirium monitoring/management, and early exercise/mobility bundle into everyday care: opportunities, challenges, and lessons learned for implementing the ICU Pain, Agitation, and Delirium Guidelines. *Critical care medicine*. 2013;41:S116-27. doi:10.1097/CCM.0b013e3182a17064.
21. Carrothers KM, Barr J, Spurlock B, Ridgely MS, Damberg CL, Ely EW. Contextual issues influencing implementation and outcomes associated with an integrated approach to managing pain, agitation, and delirium in adult ICUs. *Critical care medicine*. 2013;41:S128-35. doi:10.1097/CCM.0b013e3182a2c2b1.
22. Radtke FM, Heymann A, Franck M, Maechler F, Drews T, Luetz A, Nachtigall I, Wernecke KD, Spies CD. How to implement monitoring tools for sedation, pain and delirium in the intensive care unit: an experimental cohort study. *Intensive care medicine*. 2012;38:1974–81. doi:10.1007/s00134-012-2658-1.
23. Burchardi H. Einführung des Fallpauschalensystems in Deutschland : Ein geschichtlicher Rückblick. *Medizinische Klinik, Intensivmedizin und Notfallmedizin*. 2018;113:5–12. doi:10.1007/s00063-017-0365-y.
24. Wright MN, Ziegler A. ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *J. Stat. Soft.* 2017. doi:10.18637/jss.v077.i01.
25. Reshef DN, Reshef YA, Finucane HK, Grossman SR, McVean G, Turnbaugh PJ, Lander ES, Mitzenmacher M, Sabeti PC. Detecting novel associations in large data sets. *Science*. 2011;334:1518–24. doi:10.1126/science.1205438.
26. Vasilevskis EE, Kuzniewicz MW, Cason BA, Lane RK, Dean ML, Clay T, Rennie DJ, Vittinghoff E, Dudley RA. Mortality probability model III and simplified acute physiology

- score II: assessing their value in predicting length of stay and comparison to APACHE IV. *Chest*. 2009;136:89–101. doi:10.1378/chest.08-2591.
27. Mansouri P, Javadpour S, Zand F, Ghodsbin F, Sabetian G, Masjedi M, Tabatabaee HR. Implementation of a protocol for integrated management of pain, agitation, and delirium can improve clinical outcomes in the intensive care unit: a randomized clinical trial. *Journal of critical care*. 2013;28:918–22. doi:10.1016/j.jcrc.2013.06.019.
28. Dale CR, Kannas DA, Fan VS, Daniel SL, Deem S, Yanez, N. D., 3rd, Hough CL, Dellit TH, Treggiari MM. Improved analgesia, sedation, and delirium protocol associated with decreased duration of delirium and mechanical ventilation. *Annals of the American Thoracic Society*. 2014;11:367–74. doi:10.1513/AnnalsATS.201306-210OC.
29. Luetz A, Weiss B, Boettcher S, Burmeister J, Wernecke KD, Spies C. Routine delirium monitoring is independently associated with a reduction of hospital mortality in critically ill surgical patients: A prospective, observational cohort study. *Journal of critical care*. 2016;35:168–73. doi:10.1016/j.jcrc.2016.05.028.
30. Xi X, Wang E, Lu Q, Chen P, Wo T, Tang K. Does an economic incentive affect provider behavior?: Evidence from a field experiment on different payment mechanisms. *Journal of medical economics*. 2018:1–25. doi:10.1080/13696998.2018.1539399.
31. Klauber M, Geraedts M, Friedrich J, Wasem J, editors. *Krankenhaus-Report 2017: Schwerpunkt: Zukunft gestalten*. Stuttgart: Schattauer Verlag; 2017.
32. Kumpf O, Braun JP, Brinkmann A, Bause H, Bellgardt M, Bloos F, Dubb R, Greim C, Kaltwasser A, Marx G, Riessen R, Spies C, Weimann J, Wobker G, Muhl E, Waydhas C. Quality indicators in intensive care medicine for Germany - third edition 2017. *German medical science : GMS e-journal*. 2017;15:Doc10. doi:10.3205/000251.
33. Nachtigall I, Tamarkin A, Tafelski S, Deja M, Halle E, Gastmeier P, Wernecke KD, Bauer T, Kastrup M, Spies C. Impact of adherence to standard operating procedures for pneumonia on outcome of intensive care unit patients. *Critical care medicine*. 2009;37:159–66. doi:10.1097/CCM.0b013e3181934f1b.

34. Martín MC, Cabré L, Ruiz J, Blanch L, Blanco J, Castillo F, Galdós P, Roca J, Saura RM.
Indicadores de calidad en el enfermo crítico. *Med Intensiva*. 2008;32:23–32.

EIDESSTATTLICHE VERSICHERUNG UND ANTEILSERKLÄRUNG

Eidesstattliche Versicherung

„Ich, Marc Deffland, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: Klinische und ökonomische Effekte aus der Anwendung von Qualitätsindikatoren in der Intensivmedizin am Beispiel der Überwachung von Schmerz, Sedierung und Delir selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

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

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

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

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

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

01.10.2020

Marc Deffland

Anteilerklärung an den erfolgten Publikationen

Marc Deffland hatte folgenden Anteil an der folgenden Publikation: “Marc Deffland (MD), Claudia Spies (CS), Björn Weiss (BW), Niklas Keller (NK), Mirjam Jenny (MJ), Jochen Kruppa (JK), Felix Balzer (FB), Effects of pain, sedation and delirium monitoring on clinical and economic outcome: a retrospective study, PloSOne, 2020”

Bei der von mir in Erstautorenschaft erstellten Publikation handelt es sich um eine Arbeit aus der Schnittstelle zwischen Intensivmedizin und Ökonomie, welche medizinische als auch ökonomische Fachkenntnisse erforderte. Im Rahmen der veröffentlichten Studie wurde untersucht, ob Routine-Monitoring das klinische und ökonomische Outcome verbessert. Die Studie wurde von Frau Professorin Spies konzipiert.

Im ersten Schritt habe ich mit der Forschungslage zu Qualitätsindikatoren und Leitlinien im Kontext von Routine-Monitoring von Delir, Schmerz und Sedierung auf Intensivstationen und dessen klinischen und ökonomischen Effekten beschäftigt. Im Anschluss wurde die Arbeits-Hypothese und die auszuwertenden Variablen mit Prof. Dr. Balzer und Prof. Dr. Spies abgestimmt. Folgend wurde gemeinsam mit Prof. Dr. Felix Balzer geprüft, aus welchen Systemen die jeweiligen Daten in welcher Form exportiert werden können.

Im nächsten Schritt habe ich die Rohdaten aus unterschiedlichen Systemen für Auswertungen vorbereitet und in einer Datenbank zusammengefasst. Hierbei habe ich auch das Consort Diagramm erstellt, welches als Abbildung 1 in der Publikation veröffentlicht wurde.

Diese Datenbank habe ich für eigenständige statistische Auswertungen mit SPSS genutzt. Die

Ergebnisse der statistischen Auswertungen wurden mit Prof. Dr. Balzer und Prof. Dr. Spies besprochen. Abschließend habe ich meine statistischen Ergebnisse Dr. Jochen Kruppa vorgestellt und mit ihm diskutiert. Aus meinen statistischen Auswertungen sind die Tabellen 1 bis 4 der Publikation entstanden.

Neben diesen klassischen Auswertungen wurden auch ein Machine-Learning-Ansatz sowie ein moderner Index für die Zusammenhangsanalyse genutzt. Diese Auswertungen wurden nach einer Erläuterung des Forschungsprojektes und der bisherigen Ergebnisse von Dr. Mirjam Jenny und Dr. Niklas Keller vorgenommen. Diese nicht von mir erstellten statistischen Auswertungen sind in Abbildung 2 der Publikation abgebildet.

Nach den statistischen Auswertungen habe ich den ersten vollständigen Entwurf des Papers (inkl. Der Erstellung der Grafiken und Tabellen) geschrieben. Diese wurde durch Frau Prof. Dr. Spies, Dr. Björn Weiß, Dr. Niklas Keller, Dr. Jochen Kruppa und Prof. Dr. Balzer Review gelesen und anschließend von mir überarbeitet und eingereicht.

Unterschrift, Datum und Stempel des erstbetreuenden Hochschullehrers

Unterschrift des Doktoranden

AUSZUG JOURNAL SUMMARY LIST

Hinweis: Hinweis: PLoS One liegt unter den besten 35%. Ein begründeter Antrag (Stellungnahme des Betreuers) an den Vorsitzenden der Promotionskommission liegt den eingereichten Unterlagen bei.

Journal Data Filtered By: **Selected JCR Year: 2018** Selected Editions: SCIE,SSCI
 Selected Categories: **"MULTIDISCIPLINARY SCIENCES"** Selected Category
 Scheme: WoS

Gesamtanzahl: 69 Journale

Rank	Full Journal Title	Total Cites	Journal Impact Factor	Eigenfactor Score
1	NATURE	745,692	43.070	1.285010
2	SCIENCE	680,994	41.037	1.070190
3	National Science Review	1,842	13.222	0.006500
4	Science Advances	21,901	12.804	0.110010
5	Nature Communications	243,793	11.878	1.103290
6	Nature Human Behaviour	1,230	10.575	0.006550
7	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA	661,118	9.580	1.022190
8	Science Bulletin	3,569	6.277	0.009840
9	Scientific Data	3,240	5.929	0.015610
10	Frontiers in Bioengineering and Biotechnology	1,994	5.122	0.006540
11	Journal of Advanced Research	2,691	5.045	0.004780
12	Research Synthesis Methods	1,932	5.043	0.005420
13	GigaScience	2,674	4.688	0.012510
14	Annals of the New York Academy of Sciences	46,385	4.295	0.025840
15	Scientific Reports	302,086	4.011	1.061540
16	Journal of the Royal Society Interface	12,933	3.224	0.029190
17	NPJ Microgravity	203	3.111	0.000670
18	PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A-MATHEMATICAL PHYSICAL AND ENGINEERING SCIENCES	19,227	3.093	0.028200

Rank	Full Journal Title	Total Cites	Journal Impact Factor	Eigenfactor Score
19	FRACTALS-COMPLEX GEOMETRY PATTERNS AND SCALING IN NATURE AND SOCIETY	1,429	2.971	0.001120
20	Journal of Radiation Research and Applied Sciences	860	2.963	0.001860
21	MIT Technology Review	929	2.893	0.001910
22	JOURNAL OF KING SAUD UNIVERSITY SCIENCE	1,120	2.835	0.001670
23	PROCEEDINGS OF THE ROYAL SOCIETY A-MATHEMATICAL PHYSICAL AND ENGINEERING SCIENCES	18,683	2.818	0.018940
24	PLoS One	650,727	2.776	1.706770
25	COMPLEXITY	2,753	2.591	0.003890
26	Royal Society Open Science	4,118	2.515	0.017150
27	PeerJ	11,911	2.353	0.045900
28	SCIENCE AND ENGINEERING ETHICS	1,719	2.275	0.003450
29	INTERNATIONAL JOURNAL OF BIFURCATION AND CHAOS	7,008	2.145	0.007390
30	Symmetry-Basel	2,097	2.143	0.002590
31	SCIENTIFIC AMERICAN	6,609	1.946	0.003540
32	Science of Nature	508	1.839	0.002000
33	PROCEEDINGS OF THE JAPAN ACADEMY SERIES B-PHYSICAL AND BIOLOGICAL SCIENCES	1,532	1.833	0.001960
34	Journal of Taibah University for Science	779	1.640	0.001240
35	Frontiers in Life Science	241	1.622	0.000500
36	ARABIAN JOURNAL FOR SCIENCE AND ENGINEERING	3,838	1.518	0.005840
37	SCIENCE PROGRESS	521	1.500	0.000400

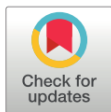
RESEARCH ARTICLE

Effects of pain, sedation and delirium monitoring on clinical and economic outcome: A retrospective study

Marc Deffland¹, Claudia Spies^{1*}, Bjoern Weiss¹, Niklas Keller¹, Mirjam Jenny^{2,3,4}, Jochen Kruppa^{5,6}, Felix Balzer¹

1 Department of Anaesthesiology and Intensive Care Medicine (CCM, CVK), Charité–Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt Universität zu Berlin and Berlin Institute of Health, Berlin, Germany, **2** Science Communication Unit, Robert Koch Institute, Berlin, Germany, **3** Harding Center for Risk Literacy, Faculty of Health Sciences Brandenburg, University of Potsdam, Potsdam, Germany, **4** Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany, **5** Institute of Biometry and Clinical Epidemiology, Charité–Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany, **6** Berlin Institute of Health (BIH), Berlin, Germany

* claudia.spies@charite.de

**OPEN ACCESS**

Citation: Deffland M, Spies C, Weiss B, Keller N, Jenny M, Kruppa J, et al. (2020) Effects of pain, sedation and delirium monitoring on clinical and economic outcome: A retrospective study. PLoS ONE 15(9): e0234801. <https://doi.org/10.1371/journal.pone.0234801>

Editor: Tobias Eckle, University of Colorado Denver, UNITED STATES

Received: January 10, 2020

Accepted: June 2, 2020

Published: September 2, 2020

Copyright: © 2020 Deffland et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data cannot be shared publicly because they contain potentially identifying or sensitive patient information, and thus legal restrictions have been imposed by the Ethics Committee and the data protection commissioner of the Charité – Universitätsmedizin Berlin. Requests may be sent to dairresearchdata@charite.de.

Funding: The research leading to these results has received funding from the European Union Seventh Framework Program [FP7/2007-2013] under grant

Abstract

Background

Significant improvements in clinical outcome can be achieved by implementing effective strategies to optimise pain management, reduce sedative exposure, and prevent and treat delirium in ICU patients. One important strategy is the monitoring of pain, agitation and delirium (PAD bundle). We hypothesised that there is no sufficient financial benefit to implement a monitoring strategy in a Diagnosis Related Group (DRG)-based reimbursement system, therefore we expected better clinical and decreased economic outcome for monitored patients.

Methods

This is a retrospective observational study using routinely collected data. We used univariate and multiple linear analysis, machine-learning analysis and a novel correlation statistic (maximal information coefficient) to explore the association between monitoring adherence and resulting clinical and economic outcome. For univariate analysis we split patients in an adherence achieved and an adherence non-achieved group.

Results

In total 1,323 adult patients from two campuses of a German tertiary medical centre, who spent at least one day in the ICU between admission and discharge between 1. January 2016 and 31. December 2016. Adherence to PAD monitoring was associated with shorter hospital LoS (e.g. pain monitoring 13 vs. 10 days; $p < 0.001$), ICU LoS, duration of mechanical ventilation shown by univariate analysis. Despite the improved clinical outcome, adherence to PAD elements was associated with a decreased case mix per day and profit per day

agreement n° 602461. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. There was no additional external funding received for this study.

Competing interests: MD has nothing to declare. CS reports grants from Aridis Pharmaceutical Inc., grants from B. Braun Melsungen AG, grants from Drägerwerk AG & Co. KGaA, grants from Grünenthal GmbH, grants from Infectopharm GmbH, grants from Sedana Medical Ltd., grants from Deutsche Forschungsgemeinschaft / German Research Society, grants from Deutsches Zentrum für Luft- und Raumfahrt e. V. (DLR) / German Aerospace Center, grants from Einstein Stiftung Berlin/ Einstein Foundation Berlin, grants from European Society of Anaesthesiology, grants from Gemeinsamer Bundesausschuss / Federal Joint Committee (G-BA), grants from Inneruniversitäre Forschungsförderung / Inner University Grants, grants from Projektträger im DLR / Project Management Agency, grants from Stifterverband/ Non-Profit Society Promoting Science and Education, grants from WHOCC, grants from Baxter Deutschland GmbH, grants from Biotest AG, grants from Cytosorbents Europe GmbH, grants from Edwards Lifesciences Germany GmbH, grants from Fresenius Medical Care, grants from Grünenthal GmbH, grants from Masimo Europe Ltd., grants from Medtronic GmbH, grants from Pfizer Pharma PFE GmbH, personal fees from Georg Thieme Verlag, grants from Dr. F. Köhler Chemie GmbH, grants from Sintetica GmbH, grants from European Commission, grants from Stifterverband für die deutsche Wissenschaft e.V. / Philips, grants from Stiftung Charité, grants from AGUETTANT Deutschland GmbH, grants from AbbVie Deutschland GmbH & Co. KG, grants from Amomed Pharma GmbH, grants from InTouch Health, grants from Copra System GmbH, grants from Correvio GmbH, grants from Max-Planck-Gesellschaft zur Förderung der Wissenschaften e. V., outside the current work; In addition, Prof. Spies has a patent 10 2014 215 211.9 licensed, a patent 10 2014 215 212.9 licensed, a patent 10 2018 114 364.8 licensed, and a patent 10 2018 110 275.5 pending. Dr. Bjoern Weiss reports personal fees from Orion Pharma Ltd. Outside the current work; in addition Dr. Weiss reports grants from the European Society of Intensive Care Medicine and the Robert-Koch-Institute outside the current work. NK has nothing to declare. MJ reports talks or teaching for which she received modest honoraria by the following organizations: German National Association of Statutory Health Insurance Physicians, Dr. Risch Laboratories (Buchs, Switzerland), Health Ministry Baden-

shown by univariate analysis. Multiple linear analysis did not confirm these results. PAD monitoring is important for clinical as well as economic outcome and predicted case mix better than severity of illness shown by machine learning analysis.

Conclusion

Adherence to PAD bundles is also important for clinical as well as economic outcome. It is associated with improved clinical and worse economic outcome in comparison to non-adherence in univariate analysis but not confirmed by multiple linear analysis.

Trial registration

clinicaltrials.gov [NCT02265263](https://doi.org/10.1186/1745-6215-263), Registered 15 October 2014.

Introduction

Since the late 1990s, there has been increasing awareness of the impact of clinical management of pain, sedation and delirium in intensive care units (ICUs) [1]. This type of comprehensive management is often referred to as PAD management (pain, agitation/sedation and delirium) and has become an integral part of routine intensive care. There are several reasons why this management is paramount for patients:

Critically ill adults frequently experience pain, which is known to be a major stressor in the ICU and a driver of distress and agitation [2, 3]. Therefore, analgesia is the first priority to minimize sedation. This achieves broad consensus, for example, in the eCASH statement [4].

Deep sedation of ICU patients is associated with adverse outcomes, including longer durations of mechanical ventilation (MV) and increased risk of mortality [5, 6]. Therefore, an approach that minimises sedation is recommended.

Delirium is a common brain dysfunction in critically ill patients and is associated with cognitive dysfunction [7, 8]. Delirium is linked to longer hospital and ICU Length of Stay (LoS) and long-term cognitive impairment [9]. The development of delirium in ICU patients may be contribute by fluctuations in sedation levels and a non-adequate pain management [10, 11].

Significant improvements in patient short and long term outcome can be achieved by implementing effective strategies to optimise pain management, reduce sedative exposure, and prevent and treat delirium in ICU patients [12, 13]. The most effective measures for implementation are bundles and clinical concepts, such as the ABCDEF approach (Assess, prevent, and manage pain; Both spontaneous awakening and breathing trials; Choice of Analgesia and Sedation; Delirium assess, prevent, and manage; Early Mobility and Exercise; Family engagement/empowerment), which aims to improve the management of pain, sedation, and delirium [14].

The implementation of these bundles and especially a diagnostic screening for PAD using validated instruments, which is referred to as PAD monitoring, is recommended in national and international guidelines [15–17].

Although routine monitoring should be part of daily clinical practice, studies have shown low implementation rates [18–21] and demonstrate that implementation of guidelines in clinical practice is a considerable challenge in intensive care [22]. Studies on implementation strategies and barriers currently focus on training, behaviour and organisational structures [23–25]. An additional reason for low implementation rates could be a lack of financial

Württemberg (Germany), mediX AG(Zurich, Switzerland), VDE e.V. (Germany), Geburtshilfliche Gesellschaft zu Hamburg (Obstetrics Association Hamburg), Landesgesundheitsprojekte e.V. (Germany), Gesundheitsforen Leipzig GmbH (Germany), Deutsche Menopause Gesellschaft e.V. (German Menopause Association), Deutscher Berufsverband für Pflegeberufe (DBfK) Nordwest e. V. (German Nursing Association), Swiss Rheumatism Symposium, Sana Herzchirurgie Stuttgart GmbH (Germany). The Harding Center (MJ's former employer) has received funding from: AOK Insurance (Germany), Helsana Insurance (Switzerland), BKK Viactiv Insurance (Switzerland), German Football Association, Microsoft Research New Work, David Harding (Winton Philanthropies), German Federal Institute for Risk Assessment, ERGO Insurances (Germany), Goethe University of Frankfurt (Germany), German Ministry for Transport and digital Infrastructure, German Ministry for Justice and Consumer Protection, German Aerospace Center. JK has nothing to declare. FB reports grants from Berlin Institute of Health, grants from Einstein Foundation, personal fees from Axon Publishing, grants from Vifor Pharma, grants and personal fees from Clearflow, personal fees and non-financial support from Medtronic outside the current work. I hereby confirm that all patents and personal fees that have been reported by the authors are entirely outside the current work. There are no patents, products in development or marketed products associated with this research to declare. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

incentivisation of these measures by the reimbursement system. We therefore hypothesise that although adherence to the PAD bundle is linked to improved clinical outcome, there is no sufficient financial benefit to implementing these methods in a Diagnosis Related Group (DRG)-based reimbursement system. This absence of a financial benefit is in turn associated with a worse economic outcome for the hospital, which means lower daily revenues (case mix) and profits.

Material and methods

The institutional review board (“Ethics Committee of Charité–Universitätsmedizin Berlin”) approved the analysis and waived informed consent (EA 2/092/14). We accessed the data between June 2017 and August 2018. We allocated an alias to data immediately after export. The treatment range was between January 2016 and December 2016. All data come from our hospital.

We set up a retrospective cohort study analysing both clinical and economic outcome. In addition to classical statistical approaches, we used a machine-learning algorithm (Boruta) and the maximal information coefficient (MIC) to analyse the importance (Boruta) and strength (MIC) of effects of monitoring adherence on clinical and economic outcome.

We used routinely collected data and included DRG-invoiced ICU patients who were admitted to and discharged from one of the centre's ICUs in 2016 and who did not receive PAD monitoring as part of a clinical trial. We excluded non adults, re-admissions to more than one ICU ward, patients without a documented day between admission and discharge, deceased patients and cases with no possible CAM-ICU monitoring (see explanation on the main predictor variables).

The study has been registered at ClinicalTrials.gov, number NCT 02265263. The local ethics committee approved the analysis and waived informed consent (EA 2/092/14).

Data sources

Routine clinical data were acquired from the two electronic patient data management systems used at the hospital (COPRA, Berlin, Germany and SAP, Walldorf, Germany).

Measures

Main predictor variables. The main predictor variables were adherence to pain, sedation and delirium monitoring. Pain, sedation and delirium monitoring were aligned in an algorithm: Starting with sedation monitoring, all patients with a Richmond Agitation Sedation Scale (RASS) of -3 or greater were monitored for delirium with the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) [26]. Patients with negative CAM-ICU results were assessed for pain using the Visual Agitation Scale (VAS) [27]. In case of a positive CAM-ICU, or if screening for sedation revealed a RASS of -4 or less, patients were screened for pain using the Behavioural Pain Scale (BPS) [28].

The adherence to pain, sedation and delirium monitoring for a patient was calculated as follows: the number of “adhered shifts” was divided by the total number of ICU shifts in which the patient was treated. Shifts on the day of admission and the day of discharge were not considered. Accordingly, the adherence could assume a value between 0 and 100%. A patient's ICU shift was rated as “adhered” if the patient was monitored at least once per shift. Patients with an Adherence of 100% were classified as “achieved” for univariate analysis. If the patient received no delirium monitoring, the shift was only rated as adhered when there was no RASS assessment or CAM-ICU was not possible. Patients for whom CAM-ICU monitoring was not

possible were not considered in the analysis of CAM-ICU monitoring adherence. We evaluated pain, sedation and delirium separately, not as bundle.

Outcome variables. The clinical outcome variables were hospital LoS, ICU LoS and duration of mechanical ventilation. The economic variables included case mix per day and profit per day because of their incentive effect for hospitals. Profit and turnover or case mix influence the management of a hospital: Profit per day addresses two points, a day view and the difference between turnover and costs. Turnover is a variable concerning the market share of a hospital. Therefore, we use profit per day and not total costs or costs per day.

The case mix (measured in case mix points) was derived from Diagnosis Related Groups (DRG). These points multiplied with a base rate (measured in EUR) are the substantial part of hospital revenue for a hospital case. Profit per day was calculated by case mix multiplied with base rate plus other receipts minus the case costs documented for the German nationwide institute of hospital revenue and costs calculations (InEK: Institut für das Entgeltsystem im Krankenhaus).

Covariates. The control variables for the multiple linear analysis and machine-learning algorithm were determined a priori based on available literature and clinical experience. They included age and the Acute Physiology and Chronic Health Evaluation II (APACHE II) scale score.

Analysis

We used both a classical statistical approach and a machine-learning analysis approach (the Boruta algorithm with MIC). While the classical approach uses all available data to explain or best describe particular linear associations and correlations, the use of cross-validation via a non-linear machine learning method allowed us to identify robust predictor variables for our outcome measures that are potentially non-linearly related to the latter. Cross-validation is the process of training the model on a subset of the data and then allowing it to assess the remainder of the cases in the dataset. This reduces the chance of model overfitting, e.g., capturing spurious correlations. In short, we supplemented the classical statistical analyses by alternative approaches for the identification of variables that are non-linearly related to our outcome and that are likely to also be predictive in new patient populations, i.e., that are likely to generalise.

For all classical statistical analyses, we used SPSS Version 24.0.0.0 (IBM SPSS Statistics). For univariate analysis, we split patients into groups. To differentiate between the monitored (monitoring adherence achieved) and not-monitored (monitoring adherence not achieved) patients, we set a 100% adherence quote. To differentiate between disease severity we used APACHE < = 10-group, APACHE 11-20-group and APACHE >20-group. For testing the association between monitoring adherence and outcome, we also used linear regression models.

For the machine-learning analysis, we used RStudio version 1.1.419 (R Foundation for Statistical Computing), and the “Boruta” package (for details, see [29]). Boruta is a random-forest-based method of feature selection. A random forest [30] is an ensemble model that constructs a multitude (often thousands) of decision trees based on the data and then makes a committee prediction (e.g., the algorithm predicts whatever the majority of the individual decision trees predict). These models can capture non-linear and non-monotonic relationships between the input variables and outcome criteria that linear models would not be able to capture. The Boruta algorithm builds these random forests from the dataset. It then randomly shuffles the values of each variable one by one to test whether the forest’s classification performance declines when a variable’s potential statistical relationship with the criterion is eliminated by this process of value randomisation. This method is also known as permutation-based variable importance.

The Boruta algorithm only provides the relative strength of association of the different input variables; therefore, the association strength cannot be easily compared across the different outcome criteria. To identify the strength of the associations, we used the maximal information coefficient (MIC) [31], a measure of information entropy that, like random forests, is not limited to specific types of functions (linear, non-linear, or non-monotone). MIC values can range from 0 to 1 and tend to be similar to R^2 in size and interpretation.

Results

We included 1323 patients for pain and agitation/sedation monitoring and 1266 patients for delirium monitoring (Fig 1). The groups differed because some patients had insufficient RASS data: For example, a patient with a RASS of -3 or less during their hospital stay could not be monitored for delirium (see the explanation of the main predictor variables).

The median age of the admitted patients was 68 years, 84.4% received surgery, and 60% were male. The median ICU LoS was 4 days, while the median hospital LoS was 12 days, and the median duration of mechanical ventilation was 32 hours. The median monitoring adherence was 91.7% for pain, 90.4% for agitation/sedation and 100% for delirium (Table 1). The median economic outcomes were 0.44 case mix per day and 11.05 EUR profit per day. The median APACHE II score at admission was 14.

Relating monitoring and outcome

a) Two-group-analyses. For pain monitoring adherence (Tables 2 and 3), in most cases, we observed statistically significant improvement of clinical outcome and worse economic outcome for all patients and for each disease severity patient group. Exceptions were case mix per day for the APACHE >20 group, profit per day for the APACHE \leq 10 group and the APACHE 11–20 group and profit per day for the APACHE >20 group. All exceptions were not significant.

For sedation monitoring adherence (Tables 2 and 3), in most cases, we observed statistically significant improvement of clinical outcome and worse economic outcome for all patients and for each disease severity patient group. Exceptions were hospital LoS for the APACHE \leq 10-group, the ICU LoS for the APACHE \leq 10 group, profit per day for the APACHE 11–20 groups and profit per day for the APACHE >20 group. All exceptions were not significant.

For delirium monitoring adherence (Tables 2 and 3), in most cases, we observed significantly better clinical outcome and worse economic outcome for all patients and for each

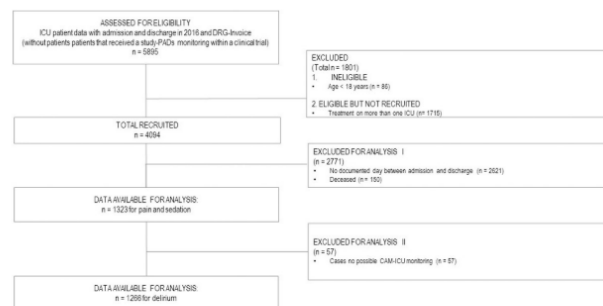


Fig 1. Consort diagram.

<https://doi.org/10.1371/journal.pone.0234801.g001>

Table 1. Patient characteristics.

Characteristics	(n = 1323)
Surgery (%)	84.4
Male (%)	60.0
Age, y [median (Q1 to Q3)]	68(56 to 76)
Hospital LoS, d [median (Q1 to Q3)]	12(7 to 19)
ICU LoS, d [median (Q1 to Q3)]	4(3 to 8)
Duration of MV, h [median (Q1 to Q3)]	32(0.0 to 109)
Pain monitoring adherence,	
% [median (Q1 to Q3)]	91.7(80.9 to 100)
% [adherence achieved]	60.7
Agitation/sedation monitoring adherence,	
% [median (Q1 to Q3)]	90.4(76.2 to 100)
% [adherence achieved]	67.0
Delirium monitoring adherence,	
% [median (Q1 to Q3)]	100(92.3 to 100)
% [adherence achieved]	64.6
Profit per day, EUR [median (Q1 to Q3)]	11.05(-240.6 to 445.3)
Case mix per day, points [median (Q1 to Q3)]	0.44(0.26 to 0.71)
APACHE II score on admission, points [median (Q1 to Q3)]	14(9 to 22)

LoS = length of stay, APACHE II = acute physiology and chronic health evaluation, MV = mechanical ventilation, y = years, h = hours

<https://doi.org/10.1371/journal.pone.0234801.t001>

disease severity patient group. Exceptions were hospital LoS for the APACHE ≤ 10 group, case mix per day for the APACHE >20 group, profit per day for the APACHE ≤ 10 and APACHE 11–20 group and profit per day for the APACHE >20 group. All exceptions were not significant.

b) Multiple linear regression. Multiple linear regression found different associations (Table 4). An increase of APACHE II score was associated with an increased hospital LoS. An increase of sedation monitoring adherence and APACHE II score was associated with an increase in days of ICU LoS. An increase of pain monitoring adherence was associated with a decreased ICU LoS. An increase of sedation monitoring adherence and APACHE score was associated with an increase in duration of MV. An increased age was associated with a decreased duration of MV. An increase of pain monitoring adherence, age and APACHE score was associated with an increased case mix per day. An increase in delirium monitoring adherence was associated with a decreased case mix per day. An increase of pain monitoring adherence, age and APACHE score was associated with an increased profit per day. Adjusted R-square was very low for all multiple linear regression.

Predicting outcome based in monitoring adherence

a) Relative predictability. The machine-learning analysis showed that pain monitoring adherence was the most important predictor of clinical outcome (hospital LoS, ICU LoS, duration of MV) and case mix per day (see Boruta in Fig 2). Furthermore, sedation monitoring was more important to clinical outcome and case mix per day than the APACHE II score. The APACHE II was more important to profit per day than the other variables. However, delirium monitoring was important to clinical and economic outcome and was more important than the APACHE II score for hospital LoS, ICU LoS and case mix per day, but not for duration of

Table 2. Relating monitoring and clinical outcome.

			Cases			Hospital LoS		ICU LoS		Duration of MV	
			No.	Med. (Q1 to Q3)	p	Med. (Q1 to Q3)	p	Med. (Q1 to Q3)	p		
Pain monitoring adherence	All patients	Not achieved	803	13* (8 to 22)	<0.001	6* (4 to 11)	<0.001	60* (<0.001 to 167)	<0.001		
		Achieved	520	10* (6 to 15)	<0.001	3* (2 to 4)	<0.001	0* (<0.001 to 55.75)	<0.001		
	APACHE <= 10	Not achieved	230	11* (19 to 7)	0.003	5* (8 to 3)	<0.001	29* (0.001 to 107.75)	<0.001		
		Achieved	189	9* (6 to 14)	0.003	2* (2 to 4)	<0.001	0* (<0.001 to 23.50)	<0.001		
	APACHE 11–20	Not achieved	319	13* (8 to 22)	<0.001	6* (3 to 10)	<0.001	48* (<0.001 to 134)	<0.001		
		Achieved	212	10* (7 to 15)	<0.001	3* (2 to 4)	<0.001	0* (<0.001 to 55)	<0.001		
	APACHE >20	Not achieved	254	16* (9 to 25)	<0.001	7* (5 to 17)	<0.001	122,5* (34.75 to 316)	<0.001		
		Achieved	119	10* (7 to 15)	<0.001	4* (3 to 5)	<0.001	45* (<0.001 to 86)	<0.001		
Sedation/agitation monitoring adherence"	All patients	Not achieved	886	13* (8 to 21)	<0.001	5* (3 to 10)	<0.001	51* (<0.001 to 149)	<0.001		
		Achieved	437	10* (6 to 15.5)	<0.001	3* (2 to 4)	<0.001	0* (0 to 59)	<0.001		
	APACHE <= 10	Not achieved	270	10(6 to 18)	0.161	4(2 to 8)	<0.001	3* (<0.001 to 92.75)	<0.001		
		Achieved	149	10(6 to 14)	0.161	3(2 to 4)	<0.001	<0.001* (<0.001 to 29)	<0.001		
	APACHE 11–20	Not achieved	349	13* (8 to 21)	<0.001	5* (3 to 9)	<0.001	39* (<0.001 to 123)	<0.001		
		Achieved	182	10* (7 to 17.25)	<0.001	3* (2 to 4)	<0.001	<0.001* (0.001 to 60.25)	<0.001		
	APACHE >20	Not achieved	267	16* (10 to 25)	<0.001	7* (5 to 16)	<0.001	110* (38 to 302)	<0.001		
		Achieved	106	9.5* (6.75 to 14.25)	<0.001	4* (2.75 to 6)	<0.001	34* (<0.001 to 86)	<0.001		
Delirium monitoring adherence	All patients	Not achieved	448	15* (8 to 24)	<0.001	7* (4 to 13.75)	<0.001	85* (8 to 222)	<0.001		
		Achieved	818	11* (7 to 17)	<0.001	4* (2 to 6)	<0.001	16* (<0.001 to 76)	<0.001		
	APACHE <= 10	Not achieved	116	11(6 to 21)	0.050	6* (3 to 11)	<0.001	54,5* (<0.001 to 146.75)	<0.001		
		Achieved	278	10(6 to 15)	0.050	3* (2 to 5)	<0.001	<0.001* (<0.001 to 49.25)	<0.001		
	APACHE 11–20	Not achieved	183	14* (8 to 23)	0.001	6* (4 to 12)	<0.001	69* (<0.001 to 176)	<0.001		
		Achieved	326	11* (8 to 17)	0.001	4* (2 to 6)	<0.001	11* (<0.001 to 69.5)	<0.001		
	APACHE >20	Not achieved	149	18* (10 to 27.5)	<0.001	9* (6 to 20)	<0.001	149* (59 to 444.5)	<0.001		
		Achieved	214	12* (7–17.25)	<0.001	5* (3–7)	<0.001	51* (1–112.75)	<0.001		

* = p<0.05 (Mann-Whitney-U-Test), LoS = length of stay, APACHE II = acute physiology and chronic health evaluation, MV = mechanical ventilation, Med. = median / (quartile 1 to quartile 3)

<https://doi.org/10.1371/journal.pone.0234801.t002>

MV. Delirium was further not predictive of profit per day. Age was not important for hospital LoS, ICU LoS or case mix per day. Age was important for the duration of MV but was less important than PAD monitoring and APACHE II score. Age was also important for profit per day but less important than the APACHE II and pain monitoring and agitation monitoring.

b) Associations between monitoring adherence and outcome. The strongest associations were shown between PAD monitoring and ICU LoS, with MICs of 0.47, 0.51 and 0.27,

Table 3. Relating monitoring and economic outcome.

			Cases			Case mix per day		Profit per day	
			No.	Med. (Q1 to Q3)	p	Med. (Q1 to Q3)	p		
Pain monitoring adherence	All patients	Not achieved	803	0.48* (0.29 to 0.73)	<0.001	27.40* (-223.61 to 496.30)	0.015		
		Achieved	520	0.39* (0.23 to 0.64)	<0.001	-33.08* (-270.44 to 370.62)	0.015		
	APACHE < = 10	Not achieved	230	0.43* (0.25 to 0.70)	0.005	-21.12(-286.75 to 511.47)	0.055		
		Achieved	189	0.32* (0.21 to 0.60)	0.005	-113.22(-280.76 to 205.22)	0.055		
	APACHE 11–20	Not achieved	319	0.45* (0.265 to 0.70)	0.038	-9.75(-269.19 to 445.29)	0.469		
		Achieved	212	0.36* (0.23 to 0.64)	0.038	-28.31(-274.76 to 419.81)	0.469		
	APACHE >20	Not achieved	254	0.55(0.36 to 0.80)	0.102	124.70(-138.42 to 615.73)	0.679		
		Achieved	119	0.49(0.31 to 0.71)	0.102	137.73(-207.29 to 532.74)	0.679		
Sedation/agitation monitoring adherence"	All patients	Not achieved	886	0.49* (0.29 to 0.77)	<0.001	50.33* (-217.26 to 534.16)	<0.001		
		Achieved	437	0.34* (0.22 to 0.61)	<0.001	-71.39* (-278.35 to 302.10)	<0.001		
	APACHE < = 10	Not achieved	270	0.44* (0.26 to 0.73)	<0.001	26.11* (-252.61 to 586.55)	<0.001		
		Achieved	149	0.30* (0.19 to 0.51)	<0.001	-144.31* (-344.96 to 74.24)	<0.001		
	APACHE 11–20	Not achieved	349	0.47* (0.26 to 0.73)	<0.001	8.72(-243.62 to 459.53)	0.134		
		Achieved	182	0.34* (0.21 to 0.60)	<0.001	-77.43(-286.23 to 327.93)	0.134		
	APACHE >20	Not achieved	267	0.57* (0.38 to 0.79)	0.019	130.49(-141.11 to 643.58)	0.526		
		Achieved	106	0.46* (0.30 to 0.71)	0.019	139.49(-180.65 to 503.26)	0.526		
Delirium monitoring adherence	All patients	Not achieved	448	0.52* (0.32 to 0.78)	<0.001	45.02(-230.23 to 481.08)	0.13		
		Achieved	818	0.39* (0.24 to 0.66)	<0.001	-9.23(-252.77 to 387)	0.13		
	APACHE < = 10	Not achieved	116	0.49* (0.28 to 0.81)	<0.001	-45.40(-299.38 to 501.20)	0.325		
		Achieved	278	0.33* (0.21 to 0.60)	<0.001	-70.95(-303.41 to 280.49)	0.325		
	APACHE 11–20	Not achieved	183	0.47* (0.28 to 0.73)	0.004	5.24(-251.34 to 477.18)	0.172		
		Achieved	326	0.37* (0.22 to 0.62)	0.004	-42.63(-277.46 to 350.86)	0.172		
	APACHE >20	Not achieved	149	0.55(0.38 to 0.79)	0.196	116.41(-156.88 to 518.23)	0.321		
		Achieved	214	0.51(0.32–0.76)	0.196	155.17(-132.74–644.72)	0.321		

* = p<0.05 (Mann-Whitney-U-Test), LoS = length of stay, APACHE II = acute physiology and chronic health evaluation, MV = mechanical ventilation, Med. = median / (quartile 1 to quartile 3)

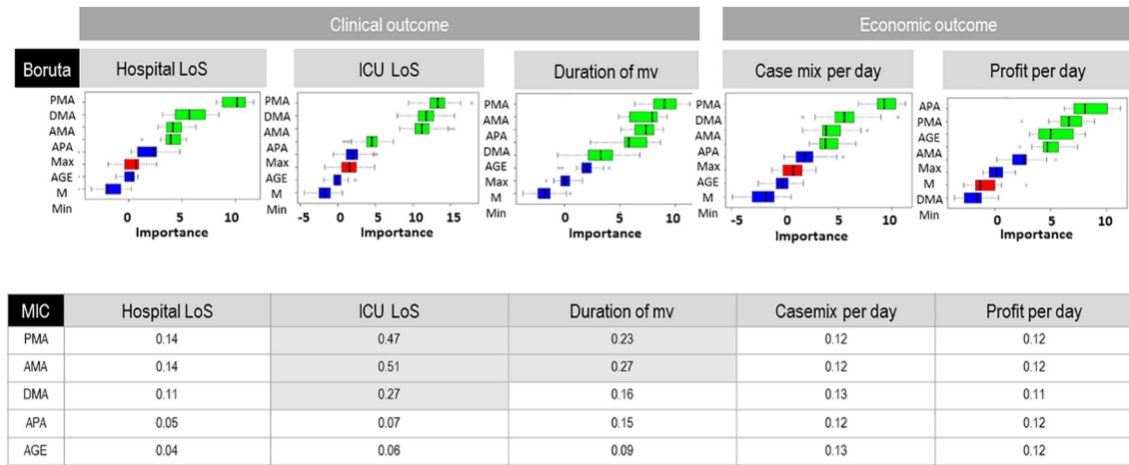
<https://doi.org/10.1371/journal.pone.0234801.t003>

Table 4. Multiple linear regression.

	clinical outcome									economic outcome					
	Hospital LoS (adjusted R-square = .021; F = 6.409; p < .001)			ICU LoS (adjusted R-square = .047; F = 13.413; p < .001)			Duration of MV (adjusted R-square = .059; F = 16.907; p < .001)			Case mix per day (adjusted R-square = .042; F = 12.037; p < .001)			Profit per day (adjusted R-square = .020; F = 6.062; p < .001)		
	B	95% CI	p	B	95% CI	p	B	95% CI	p	B	95% CI	p	B	95% CI	p
pain monitoring adherence	-0,018	-0.041 to 0.005	0,430	-0,045	-0.063 to -0.027	0,013	-0,693	-1.11 to -0.275	0,097	0,170	0.109 to 0.231	0,005	2,894	1.720 to 4.068	0,014
sedation monitoring adherence	0,022	0.002 to 0.042	0,263	0,049	0.033 to 0.065	0,002	1,011	0.642 to 1.38	0,006	-0,103	-0.157 to -0.05	0,055	-0,576	-1.614 to -0.461	0,578
delir monitoring adherence	-0,010	-0.032 to 0.013	0,660	-0,006	-0.023 to 0.012	0,751	-0,598	-1.011 to -0.185	0,148	-0,194	-0.254 to -0.133	0,001	-0,363	-1.525 to 0.799	0,755
APACHE	0,220	0.18 to 0.261	0,000	0,228	0.196 to 0.26	0,000	6,313	5.563 to 7.062	0,000	0,505	0.396 to 0.614	0,000	8,413	6.305 to 10.521	0,000
age	-0,028	-0.051 to -0.005	0,230	-0,014	-0.032 to 0.004	0,443	-0,955	-1.381 to -0.53	0,025	0,176	0.114 to 0.238	0,005	2,431	1.233 to 3.629	0,043

* = p<0.05 (XXX), LoS = length of stay, APACHE II = acute physiology and chronic health evaluation, MV = mechanical ventilation

<https://doi.org/10.1371/journal.pone.0234801.t004>



PMA = pain monitoring adherence, DMA = delirium monitoring adherence, AMA = sedation monitoring adherence, APA = APACHE at admission, AGE = Age at admission, Max = shadowMax, M = shadowMean, Min = shadowMin, MIC = maximal information coefficient, Background with shadow = MIC > 0,2

Fig 2. Results of Boruta and MIC for effects of monitoring adherence to clinical and economic outcome. The green predictor variables are deemed predictive by the Boruta algorithm while the red are not.

<https://doi.org/10.1371/journal.pone.0234801.g002>

respectively (see MIC in Fig 2). A strength of association higher than 0.2 was found between pain and sedation management and the duration of MV (see MIC in Fig 2). The strength of the association between pain, sedation and delirium monitoring and hospital LoS were all lower than 0.15. The strength of the association between pain, agitation/sedation and delirium monitoring and case mix per day were all lower than 0.14.

Discussion

We revealed that adherence to PAD bundles was associated with improved clinical outcome (hospital LoS, ICU LoS, duration of MV) but worse economic outcome (case mix per day, profit per day). Unfortunately, the result could not be confirmed in a multiple linear regression. A cause for the less pronounced effect in our cohort could be the overall high adherence and, thus, a smaller effect size.

In addition to the classical statistical evaluation, we used the machine-learning algorithm Boruta and a novel statistic (MIC). The Boruta algorithm revealed that monitoring adherence was important for clinical and economic outcome. Pain monitoring was the most important predictor of clinical outcome and case mix per day. The analysis showed that sedation and delirium monitoring are less important than pain monitoring but in most cases are more important than APACHE II score and age. While age was irrelevant, Boruta showed that the APACHE II score is an important predictor of clinical and economic outcome. This matches studies showing that the APACHE II score can predict hospital and ICU LoS as well as mortality rates [32].

While the Boruta algorithm only describes the relative strength of the associations of different input variables, MIC analysis identifies the strength of the individual associations. The strongest associations were shown between pain, agitation/sedation and delirium monitoring and ICU LoS.

Our clinical data are in line with most previous studies showing that adherence to the PAD bundle is independently associated with improved clinical outcome. For example, Luetz and colleagues found an independent association between delirium monitoring and in-hospital mortality for ventilated patients, Mansouri et al. found a substantial reduction in the duration of MV, ICU LoS, and mortality through protocol-directed PAD management and Dale et al. found decreases in delirium, duration of MV as well as ICU and hospital LoS [33–35].

Despite the favourable clinical outcome, the economic outcome within the German DRG system was associated with a decrease in both case mix per day and profit per day shown by univariate analysis but not conformed by multiple linear analysis. But it has to be considered, adjusted R-square was very low for all multiple linear regression. The focus of implementation of a DRG system was primarily keeping the quality on the existing level and reining back the costs. The focus on better quality and a change of incentive structure is an actual trend (see below).

The relevance of DRG for Germany is more important than in other health systems: While in other countries, the rate is much lower (40%-50%), in Germany, 80% of hospital finances are covered by DRG [36]. There were no surcharges for fixed costs and quality until 2020, due to a DRG system. Although there are no studies showing false incentives by DRG system, there is a trend concerning the implementation of surcharges for quality and financing of some fixed costs by surcharges. For example, Germany implemented a surcharge for nursing expenses since 2020. A recent study showed that the type and amount of reimbursement has a strong influence on the chosen treatment strategy [37]. This could be a way to optimise incentives in a DRG dominated system.

Strengths and limitations of this study

Our study has several strengths. This is the first study to use machine-learning analyses to examine the importance of monitoring in terms of both clinical and economic outcome. Furthermore, this study is the first to evaluate the economic effects of all PAD bundle monitoring values in a DRG system by case mix and profit per day.

Our study has several limitations. The first limitation of our results is that we used routine data from our clinical systems. Because of permanent validation procedures, the quality of the data can be assumed to be high, but incorrect entries in individual datasets cannot be excluded.

A further limitation is seen by measuring economic outcome only with turnover (casemix) and profit per day. We did not separately analyse the total costs. With shorter ICU and hospital LoS the total costs are probably lower than with regular LoS. Our economic view based just on the figures with an (economic) incentive for hospital (see explanation on outcome variables).

Additionally, case mix is a highly context-sensitive system. Hence, the results might not be applicable to different international health care systems. Additionally, the German national system is constantly changing, and the results for years other than 2016 could be different.

The economic analysis is only for Germany. A predication for other health systems is not possible because of the different health reimbursement systems. Many countries have a DRG system but using own databases and cost accounting guidelines.

The above cited study regarding APACHE and ICU LoS used the APACHE IV, while our hospital uses the APACHE II [32]. Furthermore, we assume in our economic analysis that reductions in LoS could allow additional patients to be treated. That is the reason for using case mix and profit per day and not per case. Factors other than the PAD bundle certainly also influenced clinical and economic outcome, but these factors could not be quantified.

A further limitation is that we treated anything less than full adherence as non-adherence in univariate analysis. The reason for this choice was the high implementation rates at our hospital. The national guideline recommends that monitoring should occur during a minimum of 70% of all shifts [17]. The 70% threshold is known from other contexts (antibiotics stewardship) to be an effective margin for reaching a significant effect by implementing standard operating procedures [38]. However, greater implementation still has effects, as our clinical data show. Using a lower percentage threshold would decrease the effect size and is therefore the more conservative approach. Our results might be a reason to conduct further studies concerning the need for higher target values.

Adherence to monitoring and analysis of effects depends on the guideline used by the institution because there are often differences in the details of different guidelines; e.g., the Spanish guideline requires sedation monitoring every six hours only for mechanically ventilated patients and gives a target value of 95% [39]. The Germany national guideline requires sedation monitoring every eight hours for every patient and gives a standard of 70% [17]. However, independent of the individual recommendations of a guideline, all advice concerning PAD management increases the sensitivity of employees in intensive care and helps to increase quality of care.

There is a need to conduct prospective studies on the topic of PAD monitoring adherence to validate our results and focus more closely on economic outcome to improve incentives for quality in a DRG-based system. Further studies should also aim at cohorts with larger differences regarding PADs adherence for confirmatory analysis.

Conclusion

Adherence to PAD bundle is important for clinical as well as economic outcome. It is associated with improved clinical and worse economic outcome in comparison to non-adherence in univariate analysis but not confirmed by multiple linear analysis.

Supporting information

S1 Fig. Boxplots for pain monitoring adherence.

(PPTX)

S2 Fig. Boxplots for sedation monitoring adherence.

(PPTX)

S3 Fig. Boxplots for delirium monitoring adherence.

(PPTX)

Author Contributions

Conceptualization: Claudia Spies.

Formal analysis: Marc Deffland, Niklas Keller, Mirjam Jenny.

Methodology: Marc Deffland, Claudia Spies, Bjoern Weiss, Niklas Keller, Mirjam Jenny, Jochen Kruppa, Felix Balzer.

Supervision: Claudia Spies, Felix Balzer.

Validation: Mirjam Jenny, Jochen Kruppa.

Writing – original draft: Marc Deffland.

Writing – review & editing: Claudia Spies, Bjoern Weiss, Niklas Keller, Jochen Kruppa, Felix Balzer.

References

1. Crippen DW. The role of sedation in the ICU patient with pain and agitation. *Crit Care Clin*. 1990; 6:369–92. PMID: [1971521](https://pubmed.ncbi.nlm.nih.gov/1971521/)
2. Novaes MA, Aronovich A, Ferraz MB, Knobel E. Stressors in ICU: patients' evaluation. *Intensive care medicine*. 1997; 23:1282–5. <https://doi.org/10.1007/s001340050500> PMID: [9470087](https://pubmed.ncbi.nlm.nih.gov/9470087/)
3. Mehta S, Spies C, Shehabi Y. Ten tips for ICU sedation. *Intensive care medicine*. 2018; 44:1141–3. <https://doi.org/10.1007/s00134-017-4992-9> PMID: [29151124](https://pubmed.ncbi.nlm.nih.gov/29151124/)
4. Vincent J-L, Shehabi Y, Walsh TS, Pandharipande PP, Ball JA, Sprong P, et al. Comfort and patient-centred care without excessive sedation: the eCASH concept. *Intensive care medicine*. 2016; 42:962–71. <https://doi.org/10.1007/s00134-016-4297-4> PMID: [27075762](https://pubmed.ncbi.nlm.nih.gov/27075762/)
5. Balzer F, Weiss B, Kumpf O, Treskatsch S, Spies C, Wernecke KD, et al. Early deep sedation is associated with decreased in-hospital and two-year follow-up survival. *Critical care (London, England)*. 2015; 19:197. <https://doi.org/10.1186/s13054-015-0929-2> PMID: [25928417](https://pubmed.ncbi.nlm.nih.gov/25928417/)
6. Tanaka LMS, Azevedo LCP, Park M, Schettino G, Nassar AP, Réa-Neto A, et al. Early sedation and clinical outcomes of mechanically ventilated patients: a prospective multicenter cohort study. *Critical care (London, England)*. 2014; 18:R156. <https://doi.org/10.1186/cc13995> PMID: [25047960](https://pubmed.ncbi.nlm.nih.gov/25047960/)
7. Ely EW, Shintani A, Truman B, Speroff T, Gordon SM, Harrell FE, et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. *JAMA*. 2004; 291:1753–62. <https://doi.org/10.1001/jama.291.14.1753> PMID: [15082703](https://pubmed.ncbi.nlm.nih.gov/15082703/)
8. Girard TD, Jackson JC, Pandharipande PP, Pun BT, Thompson JL, Shintani AK, et al. Delirium as a predictor of long-term cognitive impairment in survivors of critical illness. *Critical care medicine*. 2010; 38:1513–20. <https://doi.org/10.1097/CCM.0b013e3181e47be1> PMID: [20473145](https://pubmed.ncbi.nlm.nih.gov/20473145/)
9. Slooter AJC, van de Leur RR, Zaal JJ. Delirium in critically ill patients. *Handb Clin Neurol*. 2017; 141:449–66. <https://doi.org/10.1016/B978-0-444-63599-0.00025-9> PMID: [28190430](https://pubmed.ncbi.nlm.nih.gov/28190430/)
10. Svenningsen H, Egerod I, Videbech P, Christensen D, Frydenberg M, Tønnesen EK. Fluctuations in sedation levels may contribute to delirium in ICU patients. *Acta Anaesthesiol Scand*. 2013; 57:288–93. <https://doi.org/10.1111/aas.12048> PMID: [23294103](https://pubmed.ncbi.nlm.nih.gov/23294103/)
11. Zoremba N, Coburn M, Schälte G. Delir beim Intensivpatienten: Eine multiprofessionelle Herausforderung. *Anaesthesist* 2018. <https://doi.org/10.1007/s00101-018-0497-3> PMID: [30298270](https://pubmed.ncbi.nlm.nih.gov/30298270/)
12. Chanques G, Jaber S, Barbotte E, Violet S, Sebbane M, Perrigault PF, et al. Impact of systematic evaluation of pain and agitation in an intensive care unit. *Critical care medicine*. 2006; 34:1691–9. <https://doi.org/10.1097/01.CCM.0000218416.62457.56> PMID: [16625136](https://pubmed.ncbi.nlm.nih.gov/16625136/)
13. Kastrup M, Dossow V von, Seeling M, Ahlborn R, Tamarkin A, Conroy P, et al. Key performance indicators in intensive care medicine. A retrospective matched cohort study. *The Journal of international medical research*. 2009; 37:1267–84. <https://doi.org/10.1177/147323000903700502> PMID: [19930832](https://pubmed.ncbi.nlm.nih.gov/19930832/)
14. Pun BT, Balas MC, Barnes-Daly MA, Thompson JL, Aldrich JM, Barr J, et al. Caring for Critically Ill Patients with the ABCDEF Bundle: Results of the ICU Liberation Collaborative in Over 15,000 Adults. *Critical care medicine* 2018. <https://doi.org/10.1097/CCM.0000000000003482> PMID: [30339549](https://pubmed.ncbi.nlm.nih.gov/30339549/)
15. Tsuruta R, Fujita M. Comparison of clinical practice guidelines for the management of pain, agitation, and delirium in critically ill adult patients. *Acute medicine & surgery*. 2018; 5:207–12. <https://doi.org/10.1002/ams2.337> PMID: [29988658](https://pubmed.ncbi.nlm.nih.gov/29988658/)
16. Devlin JW, Skrobik Y, Gélinas C, Needham DM, Slooter AJC, Pandharipande PP, et al. Clinical Practice Guidelines for the Prevention and Management of Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU. *Critical care medicine*. 2018; 46:e825–e873. <https://doi.org/10.1097/CCM.0000000000003299> PMID: [30113379](https://pubmed.ncbi.nlm.nih.gov/30113379/)
17. Kumpf O, Braun JP, Brinkmann A, Bause H, Bellgardt M, Bloos F, et al. Quality indicators in intensive care medicine for Germany—third edition 2017. *German medical science: GMS e-journal*. 2017; 15: Doc10. <https://doi.org/10.3205/000251> PMID: [28794694](https://pubmed.ncbi.nlm.nih.gov/28794694/)
18. Luetz A, Balzer F, Radtke FM, Jones C, Citerio G, Walder B, et al. Delirium, sedation and analgesia in the intensive care unit: a multinational, two-part survey among intensivists. *PLoS one*. 2014; 9:e110935. <https://doi.org/10.1371/journal.pone.0110935> PMID: [25398099](https://pubmed.ncbi.nlm.nih.gov/25398099/)
19. Saller T, V VD, Hofmann-Kiefer K. [Knowledge and implementation of the S3 guideline on delirium management in Germany]. *Der Anaesthesist*. 2016; 65:755–62. <https://doi.org/10.1007/s00101-016-0218-8>

20. Trupkovic T, Kinn M, Kleinschmidt S. Analgesia and sedation in the intensive care of burn patients: results of a European survey. *Journal of intensive care medicine*. 2011; 26:397–407. <https://doi.org/10.1177/0885066610393442> PMID: 21257635
21. Wang J, Peng ZY, Zhou WH, Hu B, Rao X, Li JG. A National Multicenter Survey on Management of Pain, Agitation, and Delirium in Intensive Care Units in China. *Chinese medical journal*. 2017; 130:1182–8. <https://doi.org/10.4103/0366-6999.205852> PMID: 28485318
22. Rundshagen I. [Pain, agitation and delirium. Amended 2013 guidelines of the American College of Critical Care Medicine]. *Der Anaesthesist*. 2013;62:914–8. <https://doi.org/10.1007/s00101-013-2232-4>
23. Balas MC, Burke WJ, Gannon D, Cohen MZ, Colburn L, Bevil C, et al. Implementing the awakening and breathing coordination, delirium monitoring/management, and early exercise/mobility bundle into everyday care: opportunities, challenges, and lessons learned for implementing the ICU Pain, Agitation, and Delirium Guidelines. *Critical care medicine*. 2013; 41:S116–27. <https://doi.org/10.1097/CCM.0b013e3182a17064> PMID: 23989089
24. Carrothers KM, Barr J, Spurlock B, Ridgely MS, Damberg CL, Ely EW. Contextual issues influencing implementation and outcomes associated with an integrated approach to managing pain, agitation, and delirium in adult ICUs. *Critical care medicine*. 2013; 41:S128–35. <https://doi.org/10.1097/CCM.0b013e3182a2c2b1> PMID: 23989090
25. Radtke FM, Heymann A, Franck M, Maechler F, Drews T, Luetz A, et al. How to implement monitoring tools for sedation, pain and delirium in the intensive care unit: an experimental cohort study. *Intensive care medicine*. 2012; 38:1974–81. <https://doi.org/10.1007/s00134-012-2658-1> PMID: 22945432
26. Sessler CN, Gosnell MS, Grap MJ, Brophy GM, O'Neal PV, Keane KA, et al. The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care unit patients. *Am J Respir Crit Care Med*. 2002; 166:1338–44. <https://doi.org/10.1164/rccm.2107138> PMID: 12421743
27. Ahlers SJGM, van Gulik L, van der Veen AM, van Dongen HPA, Bruins P, Belitser SV, et al. Comparison of different pain scoring systems in critically ill patients in a general ICU. *Critical care (London, England)*. 2008; 12:R15. <https://doi.org/10.1186/cc6789> PMID: 18279522
28. Payen JF, Bru O, Bosson JL, Lagrasta A, Novel E, Deschaux I, et al. Assessing pain in critically ill sedated patients by using a behavioral pain scale. *Critical care medicine*. 2001; 29:2258–63. <https://doi.org/10.1097/00003246-200112000-00004> PMID: 11801819
29. Kursa MB, Rudnicki WR. Feature Selection with the Boruta Package. *J. Stat. Soft.* 2010. <https://doi.org/10.18637/jss.v036.i11>
30. Wright MN, Ziegler A. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *J. Stat. Soft.* 2017. <https://doi.org/10.18637/jss.v077.i01>
31. Reshef DN, Reshef YA, Finucane HK, Grossman SR, McVean G, Turnbaugh PJ, et al. Detecting novel associations in large data sets. *Science*. 2011; 334:1518–24. <https://doi.org/10.1126/science.1205438> PMID: 22174245
32. Vasilevskis EE, Kuzniewicz MW, Cason BA, Lane RK, Dean ML, Clay T, et al. Mortality probability model III and simplified acute physiology score II: assessing their value in predicting length of stay and comparison to APACHE IV. *Chest*. 2009; 136:89–101. <https://doi.org/10.1378/chest.08-2591> PMID: 19363210
33. Mansouri P, Javadpour S, Zand F, Ghodsbini F, Sabetian G, Masjedi M, et al. Implementation of a protocol for integrated management of pain, agitation, and delirium can improve clinical outcomes in the intensive care unit: a randomized clinical trial. *Journal of critical care*. 2013; 28:918–22. <https://doi.org/10.1016/j.jcrc.2013.06.019> PMID: 24011845
34. Dale CR, Kannas DA, Fan VS, Daniel SL, Deem S, Yanez N. D. 3rd, et al. Improved analgesia, sedation, and delirium protocol associated with decreased duration of delirium and mechanical ventilation. *Annals of the American Thoracic Society*. 2014; 11:367–74. <https://doi.org/10.1513/AnnalsATS.201306-210OC> PMID: 24597599
35. Luetz A, Weiss B, Boettcher S, Burmeister J, Wernecke KD, Spies C. Routine delirium monitoring is independently associated with a reduction of hospital mortality in critically ill surgical patients: A prospective, observational cohort study. *Journal of critical care*. 2016; 35:168–73. <https://doi.org/10.1016/j.jcrc.2016.05.028> PMID: 27481754
36. Klauber M, Geraedts M, Friedrich J, Wasem J, editors. Krankenhaus-Report 2017: Schwerpunkt: Zukunft gestalten. Stuttgart: Schattauer Verlag; 2017.
37. Xi X, Wang E, Lu Q, Chen P, Wo T, Tang K. Does an economic incentive affect provider behavior?: Evidence from a field experiment on different payment mechanisms. *Journal of medical economics*. 2018; 1–25. <https://doi.org/10.1080/13696998.2018.1539399> PMID: 30351230
38. Nachtigall I, Tamarkin A, Tafelski S, Deja M, Halle E, Gastmeier P, et al. Impact of adherence to standard operating procedures for pneumonia on outcome of intensive care unit patients. *Critical care medicine*. 2009; 37:159–66. <https://doi.org/10.1097/CCM.0b013e3181934f1b> PMID: 19050600

39. Martín MC, Cabré L, Ruiz J, Blanch L, Blanco J, Castillo F, et al. Indicadores de calidad en el enfermo crítico. *Med Intensiva*. 2008; 32:23–32. [https://doi.org/10.1016/s0210-5691\(08\)70899-4](https://doi.org/10.1016/s0210-5691(08)70899-4) PMID: 18221710

LEBENS LAUF

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

PUBLIKATIONSLISTE

a) Mit Impact Factor

- Effects of pain, sedation and delirium monitoring on clinical and economic outcome: a retrospective study. Marc Deffland, Claudia Spies, Björn Weiss, Niklas Keller, Mirjam Jenny, Jochen Kruppa, Felix Balzer. In: PloSOne (<https://doi.org/10.1371/journal.pone.0234801>)
- Critical care guidelines on pain, agitation and delirium management: Which one to use? A systematic literature search and quality appraisal with AGREE II. Max Rosenthal, Julius J. Grunow, Prof. Dr. Claudia D. Spies, Rudolf Mörgeli, Nicolas Paul, Marc Deffland, Prof. Dr. Alawi Luetz, Dr. Anika Mueller, Dr. Sophie K. Piper, PD Dr. Bruno Neuner, Dr. Monika Nothacker, and Dr. Björn Weiss. In: Journal of Critical Care (<https://www.sciencedirect.com/science/article/pii/S0883944120305670>)

b) Ohne Impact Factor

- "Alea iacta est. Das neue COSO-Modell. Down- oder Upgrade für Compliance-Systeme?"; In: "Corporate Compliance Zeitschrift des Beck Verlages (CCZ)", 5/2020
- "Wenn das Finanzamt zweimal klingelt"; Ko-Autoren: Babisch, Mandy, Müller, Thomas; In: "KU Gesundheitsmanagement", 06/2020
- "Update zum Unternehmensstrafrecht"; Ko-Autoren: Gehring, Christoph Leo; In: "Der Krankenhaus-Justitiar", 02/2020
- "Das Unternehmensstrafrecht kommt. Auswirkungen für die Interne Revision?!"; Ko-Autoren: Deckena, Joscha, Schönheit, Rebecca und Hochmuth, Tim; In: Zeitschrift Interne Revision (ZIR), 04/2019
- „Ein Krankenhaus hinter "Schwedischen Gardinen"? - Potentielle Änderungen der strafrechtlichen Haftung von Krankenhäusern"; Ko-Autoren: Deckena, Joscha und Schönheit, Rebecca; In: "das Krankenhaus", 12/2018
- „Zwei Kerzen auf der Geburtstagstorte - Ein Grund zum Feiern?"; Ko-Autor: Gehring, Christoph Leo; In: comply, 03/2018
- „Die Überwachung der Internen Revision: Ein Quick-Check-Modell für Aufsichtsräte"; Ko-Autor: Hauptmann, Gitta; In: Zeitschrift für Corporate Governance (ZCG), 06/17
- „Das Berliner Compliance-Modell und Orientierungshilfen der Healthcare-Compliance"; Ko-Autor: Hochmuth, Tim; In: Fachzeitschrift Risk, Fraud & Compliance (ZRFC), 04/2017
- „Unterstützung für Schäubles neue Spezialeinheit"; Ko-Autor: Hochmuth, Tim; In: Fachzeitschrift Risk, Fraud & Compliance (ZRFC), 02/2017
- "Wer, wenn nicht wir? Compliance in einer Anstalt öffentlichen Rechts"; Ko-Autor: Bethke, Kai; In: comply, 04/2016
- "Methoden zur Risikobewertung. Wer sucht, der findet - und sollte auch bewerten!"; Ko-Autor: Hochmuth, Tim; In: Der Krankenhausjustitiar, 04/2016
- "Methoden zur Risikoidentifikation. Wer sucht, der findet - und sollte auch bewerten!"; Ko-Autor: Hochmuth; In: Der Krankenhausjustitiar, 03/2016
- "Continuous Auditing – In zwei Stunden zur ersten Prüfung!"; Ko-Autor: Hochmuth; In: Zeitschrift Interne Revision (ZIR), 03/2016
- "Ein Compliance-Manager in der Notaufnahme"; In: comply, 03/2015
- "Compliance in der Medizin"; In: Euteneier (Hrsg.): Handbuch Klinisches Risikomanagement; Springer, Berlin, 2015
- "Compliance-Programme"; In: Euteneier (Hrsg.): Handbuch Klinisches Risikomanagement; Springer, Berlin, 2015
- "Whistleblower-Systeme"; In: Euteneier (Hrsg.): Handbuch Klinisches Risikomanagement; Springer, Berlin, 2015
- "Organisation eines ganzheitlichen Risikomanagements"; Ko-Autor: Löber, Nils; In: KU Gesundheitsmanagement, 03/2015
- "Rechnungslegung von KMU nach Handels- und Steuerrecht"; Ko-Autor: von Sicherer; Studienbrief des Hochschulverbundes Distance Learning, Brandenburg, 2013
- "Hohe Fehlerrisiken mit Ergebnisrelevant. Berechnung von Erlösausgleichen unter Berücksichtigung der Veränderung des Ausgleichsmechanismus durch die Konvergenzphase"; In: KU Gesundheitsmanagement, 04/2012
- "Materielles Steuerrecht - Ertragssteuern"; Ko-Autor: Zeis, Jürgen; Studienbrief des Hochschulverbundes Distance Learning, Brandenburg, 2012
- "Abschlussprüfung"; Ko-Autor: Penter, Volker; Arnold, Christoph; In Penter, Volker/Siefert, Bernd (Hrsg.) "Kompendium Krankenhaus-Rechnungswesen", Baumann Fachverlage, Kulmbach 2010
- "Investitionspauschalen - Bemessung an technischer Ausstattung und Leistungsfähigkeit"; In KPMG Gesundheitsbarometer, Heft 04/2010
- "KPMG-Umfrage zu Qualitätsmaßstäben im Krankenhaus"; In KPMG Gesundheitsbarometer, Heft 08/2009
- "An den Nachbarn orientieren"; Ko-Autoren: Penter, Volker; Buchheim, Regine; In: WohlfahrtIntern, Heft 04/2008
- "Transparenz der Gemeinnützigkeit im europäischen Vergleich"; Ko-Autor: Penter, Volker; In Neue Caritas, Heft 11/2008
- "Transparenzstandards im Dritten Sektor: Bestandsaufnahme der Rechnungslegungspflichten gemeinnütziger Vereine in Europa", GRIN Verlag, München, 2008

DANKSAGUNG

Ein erster und besonderer Dank geht an dieser Stelle an meinen Doktorvater, **Herr Professor Dr. Felix Balzer**, Leiter des Lehrstuhls „E-Health and Shared Decision Allocation“ am Einstein Center Digital Future und der Charité – Universitätsmedizin Berlin, für die optimistische, konsequente, kritische und stets motivierende Unterstützung dieser Arbeit.

Für die Ermutigung zur Durchführung dieser Arbeit und kontinuierlichen sehr wertvollen Unterstützung danke ich **Frau Professor Dr. Claudia Spies**, Klinikdirektorin der Klinik für Anästhesiologie und operative Intensivmedizin Campus Charité Mitte und Campus Virchow-Klinikum, Charité - Universitätsmedizin Berlin.

Weiterhin möchte ich mich bei den Co-Autoren unserer Publikation, **Dr. Björn Weiß**, **Dr. Niklas Keller**, **Dr. Miriam Jenny** sowie **Dr. Jochen Kruppa** aus der Charité – Universitätsmedizin Berlin, dem Berlin Institute of Health sowie dem Harding-Zentrum für Risikokompetenz des Themas für die gute Zusammenarbeit bei der bedanken. Gleiches gilt meinem ersten Zweitbetreuer **Dr. Matthes Seeling**, der jedoch aufgrund beruflicher Veränderungen nicht bis zum Abschluss dabei bleiben konnte.

Ein besonderer persönlicher Dank geht an meine Ehefrau **Rebecca Schönheit**, die nicht müde wurde, mich immer wieder zu ermutigen und mir damit viel Kraft für diese Arbeit gegeben hat.

Zuallerletzt danke ich meinen **Eltern**, ohne deren Unterstützung weder die Erstellung dieser Arbeit noch das Studium der Betriebswirtschaftslehre sowie das Studium der Wirtschaftspädagogik für mich möglich gewesen wäre.