

## Evaluationsbericht (gemäß Nr. 14.1 ANBest-IF)

<b>Konsortialführung:</b>	Universitätsklinikum Essen
<b>Förderkennzeichen:</b>	01NVF20013
<b>Akronym:</b>	DigiSep
<b>Projekttitle:</b>	Optimierung der Sepsis-Therapie auf Basis einer patientenindividuellen digitalen Präzisionsdiagnostik
<b>Autorinnen und Autoren:</b>	Juliana Schmidt, Universität Bielefeld Maren Steinmann, Universität Bielefeld Prof. Dr. Wolfgang Greiner, Universität Bielefeld Prof. Dr. Meinhard Kieser, Institut für Medizinische Biometrie Heidelberg Dr. Manuel Feiñt, Institut für Medizinische Biometrie Heidelberg Paul Thalmann, Institut für Medizinische Biometrie Heidelberg
<b>Förderzeitraum:</b>	01.09.2021 - 31.08.2024
<b>Ansprechperson:</b>	Prof. Dr. Meinhard Kieser, Institut für Medizinische Biometrie Heidelberg, [T] +49 62 21 56 43 71, [M] Meinhard.Kieser@imbi.uni-heidelberg.de Prof. Dr. Wolfgang Greiner, Universität Bielefeld, [T] +49 521 106 6989, [M] wolfgang.greiner@uni-bielefeld.de

Das dieser Veröffentlichung zugrundliegende Projekt DigiSep wurde mit Mitteln des Innovationsausschusses beim Gemeinsamen Bundesausschuss unter dem Förderkennzeichen 01NVF20013 gefördert.

## Zusammenfassung

**Hintergrund:** Die hier vorliegende, randomisiert-kontrollierte, multizentrische Interventionsstudie mit dem Titel „Optimization of sepsis therapy based on patient-specific digital precision diagnostics using next generation sequencing“ (DigiSep-Trial) hatte das Ziel, die Sepsis-Behandlung patientenzentriert und digital gestützt durch zeitnahe und verlässliche Identifikation des Erregers mittels Next Generation Sequencing (NGS)-basierter Präzisionsdiagnostik (in Ergänzung zur parallel stattfindenden mikrobiologischen Routinediagnostik) und einer darauf aufbauenden antiinfektiven Therapiesteuerung zu optimieren.

**Methodik:** Das Institut für Medizinische Biometrie am Universitätsklinikum Heidelberg führte die Evaluation der klinischen Daten und Outcomes durch. Details der Analyse wurden im einem statistischen Analyseplan präspezifiziert. Die Universität Bielefeld führte die gesundheitsökonomische Evaluation (Kosten nach Leistungsbereichen, Gesamtkosten, Inanspruchnahme, Kosten-Effektivitäts-Analyse), formative Evaluation (Akzeptanz) sowie die Evaluation der gesundheitsbezogenen Lebensqualität (EQ-5D-5L, VR-36) durch.

**Ergebnisse:** Die primäre Analyse der DigiSep-Studie lieferte kein statistisch signifikantes Ergebnis womit die Studie hinsichtlich der klinischen Wirksamkeit als nicht erfolgreich angesehen werden muss. Es zeigte sich jedoch über alle Patientensubgruppen ein gleichmäßiger positiver Effekt der Intervention. In den sekundären Endpunkten konnte über alle Endpunkte hinweg ein positiver Effekt der Intervention erreicht werden, der sogar teilweise statistische Signifikanz erzielte.

In der gesundheitsökonomischen Evaluation wurden keine statistisch signifikanten Unterschiede in den Gesamtkosten oder der Inanspruchnahme ermittelt. Hingegen lieferten sowohl die deterministische als auch probabilistische Kosten-Effektivitäts-Analyse Hinweise auf eine dominante Intervention (kostengünstiger als auch wirksamer in Bezug auf die Senkung der Mortalität nach 28 Tagen). Hinsichtlich der gesundheitsbezogenen Lebensqualität gab es kurzfristig Verbesserungen in der Interventionsgruppe, jedoch längerfristig keine signifikanten Unterschiede im Vergleich zur Kontrollgruppe. Die formative Evaluation hob hervor, dass die Präzisionsdiagnostik von den Teilnehmenden als wertvolle Ergänzung angesehen wird, jedoch organisatorische und personelle Herausforderungen bei der Implementierung bestehen. Diese könnten im Regelbetrieb möglicherweise reduziert werden, bedürfen jedoch weiterer Optimierungen.

**Diskussion:** Die Fallzahlkalkulation wurde ausschließlich auf Basis des primären Endpunkts durchgeführt, was die Aussagekraft und Interpretierbarkeit der Analysen der sekundären Endpunkte und insbesondere der gesundheitsökonomischen Analysen limitiert. Zusammenfassend liefert die Studie Hinweise auf eine potenziell wirksame Intervention mit Kosteneinsparungen, jedoch ist die Evidenz aufgrund der aufgeführten Punkte und Limitationen begrenzt. Eine breite Implementierung der neuen Versorgungsform sollte daher erst nach weiteren Untersuchungen und Optimierungen erfolgen.

**Schlagworte:** Intensivmedizin, Sepsis, Septischer Schock, Next Generation Sequencing

## **Inhaltsverzeichnis**

I	Abkürzungsverzeichnis .....	4
II	Abbildungsverzeichnis.....	5
III	Tabellenverzeichnis .....	6
1	Ziele der Evaluation .....	7
2	Darstellung des Evaluationsdesigns .....	8
3	Ergebnisse der Evaluation .....	25
4	Diskussion der Projektergebnisse .....	48
5	Schlussfolgerungen und Empfehlungen des Evaluators .....	52
IV	Literaturverzeichnis.....	53
V	Anlagen.....	58

## I Abkürzungsverzeichnis

AIC	Akaike-Informationskriterium
ANOVA	Varianzanalyse
ANV	Akutes Nierenversagen
CEA	Kosten-Effektivitäts-Analyse
CEAC	Kosten-Effektivitäts-Akzeptanzkurve
DigiSep-Trial	Optimization of sepsis therapy based on patient-specific digital precision diagnostics using next generation sequencing
DOOR	Desirability of Outcome Ranking
FAS	Full Analysis Set
GAMLSS	Verallgemeinerte additive Modelle für Lage-, Skalen- und Formparameter
G-BA	Gemeinsamer Bundesauschuss
GKV	Gesetzliche Krankenversicherung
HKP	Häusliche Krankenpflege
IA	Innovationsausschuss
ICER	Inkrementelles Kosten-Effektivitäts-Verhältnis
ICU	Intensive Care Unit
IG	Interventionsgruppe
IMC	Intermediate Care
IQR	Interquartilsabstand
IQWIG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KG	Kontrollgruppe
KI	Konfidenzintervall
LLM	Longitudinales lineares Mixed-Modell
LOS	Length of Stay
MCS	Psychische Summenskala
MW	Mittelwert

NGS	Next Generation Sequencing
OCO	Overall Clinical Outcome
PCS	Physische Summenskala
PSA	Probabilistische Sensitivitätsanalyse
PZN	Pharmazentralnummer
QALY	Quality-Adjusted-Life Years
RADAR	Response Adjusted for Duration of Antibiotic Risk
RCT	Klinisch prospektive, multizentrische, randomisierte, kontrollierte Interventionsstudie
SAP	Statistischer Analyseplan
SD	Standardabweichung
SF-36	36-Item Short Form Surveys
UTAUT	Unified Theory of Acceptance and Use of Technology
VAS	Visuelle Analogskala
VR-36	Veterans Rand-36 Item Health Survey

## II Abbildungsverzeichnis

Abbildung 1: GANTT-Chart DigiSep.....	12
Abbildung 2: Patienten-Flow-Chart mit Ausschluss (Mehrfachgrund möglich) .....	25
Abbildung 3: Überlebenskurven bis 180 Tage, group=Gruppe, Survival probability=Überlebenswahrscheinlichkeit, Time(days) = Zeit in Tagen, Number at risk = Anzahl unter Risiko .....	30
Abbildung 4: Gesundheitsbezogene Lebensqualität im Vergleich zum Vorjahr - VR-36 (Gesamtkohorte).....	42
Abbildung 5: Gesundheitsbezogene Lebensqualität im Vergleich zum Vorjahr - VR-36 (GKV- Kohorte) .....	42
Abbildung 6: Nutzung der Präzisionsdiagnostik in der Zukunft (y-Achse) in Abhängigkeit der gegenwärtigen Nutzung (x-Achse) .....	44
Abbildung 7: Sozialer Einfluss .....	45
Abbildung 8: Leistungserwartung an die Präzisionsdiagnostik .....	46
Abbildung 9: Aufwandserwartungen an die Präzisionsdiagnostik .....	47
Abbildung 10: Unterstützende Rahmenbedingungen .....	48

### III Tabellenverzeichnis

Tabelle 1: Erhebung klinischer Daten an den jeweiligen Studienzentren .....	10
Tabelle 2: Übereinstimmungsbewertung Experten .....	16
Tabelle 3: Outcomes und Erhebungszeitpunkte* .....	17
Tabelle 4: Patientencharakteristika Studienkohorte .....	26
Tabelle 5: Durchgeführte Therapieänderungen in der Interventionsgruppe.....	28
Tabelle 6: Deskriptive Beschreibung des primären Endpunkts.....	29
Tabelle 7: Abgeleitete Empfehlungen der Experten .....	32
Tabelle 8: Baseline Charakteristika der GKV-Kohorte.....	33
Tabelle 9: Primäre und sekundäre Endpunkte (GKV-Kohorte).....	33
Tabelle 10: Ergebnisse des GAMLSS für die Gesamtkosten .....	36
Tabelle 11: Gesundheitsbezogene Lebensqualität - EQ-5D-5L VAS .....	39
Tabelle 12: Gesundheitsbezogene Lebensqualität - VR-36.....	40

## 1 Ziele der Evaluation

### Ausgangslage:

Die Sepsis stellt eine lebensbedrohliche Erkrankung dar, bei der die Reaktion des Körpers auf eine Infektion zu einer Schädigung der körpereigenen Organe führt. Ausgelöst wird die Sepsis hauptsächlich von Bakterien; aber auch Pilze, Viren oder Parasiten können ein septisches Krankheitsbild hervorrufen. Trotz weitgreifender Fortschritte in der Sepsisforschung sind neue therapeutische Ansätze rar und die Mortalität der Patienten bleibt weiterhin inakzeptabel hoch. Zudem ist der Keimnachweis nach wie vor nicht Gegenstand der zuletzt in 2016 aktualisierten Sepsis-3 Definition [1]. Neben einer frühestmöglichen Fokussanierung (Beseitigung des Infektionsherds) stellt die rein kalkulierte (sog. empirische) Therapie mit Breitspektrumantiinfektiva in der Frühphase der Erkrankung einen Eckpfeiler des gegenwärtigen Behandlungskonzeptes dar [2]. Eine zuverlässige Erregeridentifizierung und patientenindividuelle Therapie hat aber entscheidenden Einfluss auf den Heilungsverlauf sowie das Überleben von Patienten mit Sepsis/septischem Schock [3]. Kulturbasierte Methoden stellen den Goldstandard im Rahmen der Infektionsdiagnostik dar, wenngleich diese mit relevanten Nachteilen assoziiert sind [4]. Eine patientenindividuelle, sog. Gezielte Behandlung ist daher oft nicht gewährleistet; die daraus resultierende antiinfektive Übertherapie erhöht das Risiko für Antiinfektiva-assozierte Medikamentennebenwirkungen und begünstigt die Selektion multiresistenter Keime. Ebenso ist das Auftreten von Langzeitschäden (wie z.B. die Persistenz der Nierenersatztherapiepflichtigkeit) unmittelbar vom initialen Management (einschließlich der Adäquanz der antiinfektiven Therapie) abhängig, so dass eine gezielte Diagnostik unmittelbaren und sektorenübergreifenden Einfluss auf die Nachsorge des vormals septischen Patienten hat [5].

Sepsis und septischer Schock gehören zu den kostenintensivsten Krankheitsbildern in der intensivmedizinischen Versorgung. Patienten, die an einer Sepsis/septischen Schock erkranken, verursachen im Vergleich zu anderen intensivpflichtigen Patienten (Intensive Care Unit (ICU)- / Intermediate Care (IMC)-Patienten) deutlich erhöhte direkte Tageskosten (durchschnittlich 1.090 € vs. 745 €) [6]. Diese Kosten steigen mit zunehmendem Schweregrad der Sepsis und gehen mit einer deutlich verlängerten Verweildauer in einem Überwachungsbereich (ICU/IMC-Length of Stay (LOS)) einher [7]. Die Kosten, die der Versorgung von Patienten mit Sepsis zugeschrieben werden können, belaufen sich pro Fall auf ca. 27.500 € an direkten bzw. 60.000 € indirekten Kosten [7, 8]. Allein in Deutschland führen Sepsisbehandlungen jährlich zu Ausgaben in Höhe von ca. 1,7 Mrd. €, was rund einem Drittel des intensivmedizinischen Gesamtbudgets entspricht. Dabei entfallen etwa 70 % der Gesamtbelastung auf die vorgenannten indirekten Kosten, wie Produktivitätsausfälle durch Arbeitsunfähigkeit, Frühverrentung oder Todesfälle. Schätzungen zufolge betragen die sepsis-assozierten Produktivitätsverluste in Deutschland ca. 5-7,7 Mrd. € pro Jahr [7, 8]. Ein entscheidender Kostentreiber der intensivstationären Versorgung sind Fixkosten, die eng mit der stationären Verweildauer korrelieren. Eine verlängerte Verweildauer ist häufig Folge einer inadäquaten antiinfektiven Therapie [9]. Eine schnellere und zugleich patientenindividuellere antiinfektive Therapie besäße somit das Potential, nicht nur die Heilungsaussichten zu verbessern, sondern zugleich auch zu Einsparungen zu führen. Dazu zählen reduzierte

Verweildauern, geringere Medikamentenkosten, kürzere Beatmungsdauern (direkte Kosten) sowie eine geringere Notwendigkeit für Nierenersatztherapie (indirekte Folgekosten) [10, 11].

#### **Zielsetzung:**

Durch patientenindividuelle Erregerinformationen kann eine gezielte und optimierte Therapiesteuerung von Patienten mit Sepsis/septischem Schock erfolgen. Die hier vorliegende, randomisiert-kontrollierte, multizentrische Interventionsstudie mit dem Titel „Optimization of sepsis therapy based on patient-specific digital precision diagnostics using next generation sequencing“ (DigiSep-Trial) hatte das Ziel, die Sepsis-Behandlung patientenzentriert und digital gestützt durch zeitnahe und verlässliche Identifikation des Erregers mittels Next Generation Sequencing (NGS)-basierter Präzisionsdiagnostik (in Ergänzung zur parallel stattfindenden mikrobiologischen Routinediagnostik) und einer darauf aufbauenden antiinfektiven Therapiesteuerung zu optimieren.

Die daraus resultierende, schnellere Diagnosestellung und optimierte Therapieentscheidung können des Weiteren auch aus gesundheitsökonomischer Perspektive die Versorgung weiter verbessern und die Behandlung effizienter sowie kostensparender gestalten. Es wird angenommen, dass etwa durch eine verringerte stationäre Verweildauer, insbesondere der intensivpflichtigen stationären Verweildauer, verringter Morbidität, geringerer Häufigkeit eines akuten Nierenversagens, reduziertem Antibiotikaverbrauch sowie verringriger Krankheitskosten durch Nachsorge- und Rehabilitationsbehandlungen insgesamt geringere Kosten aus Perspektive der Gesetzlichen Krankenversicherung (GKV-Perspektive) entstehen. Zudem wird ebenfalls postuliert, dass die neue Versorgungsform langfristig positive Auswirkungen auf die gesundheitsbezogene Lebensqualität der Patienten hat.

## **2 Darstellung des Evaluationsdesigns**

### **Studiendesign**

Die vorgestellte Untersuchung stellt eine klinisch prospektive, multizentrische, randomisierte, kontrollierte Interventionsstudie (RCT) dar.

### **Ein- und Ausschlusskriterien**

In DigiSep wurden Patienten auf ICU & IMC-Stationen von teilnehmenden klinischen Studienzentren eingeschlossen, die an einer Sepsis/septischem Schock (Sepsis-3) erkrankt sind. Unter überwachungspflichtigen Stationen wurden in dem hier dargestellten Projekt bewusst sowohl ICU-, als auch IMC-Stationen zusammengefasst, um eine Vergleichbarkeit der erhobenen Outcomedaten unabhängig von der zugrundeliegenden Versorgungsstufe erzielen zu können.

Die entsprechenden Ein- & Ausschlusskriterien für die Studienteilnahme sind wie folgt definiert:

Einschlusskriterien:

- Sepsis, Beginn <24h
  - Lebensbedrohliche Organdysfunktion infolge einer fehlregulierten Wirtsantwort auf dem Boden einer vermuteten oder nachgewiesenen Infektion
  - Detektion der Organdysfunktion anhand einer Änderung des SOFA-Scores  $\geq 2$  Punkte-
  - Alternative: Änderung des quick (q) SOFA-Scores  $\geq 2$  Punkte als Hinweis für eine Sepsis
- oder septischer Schock, Beginn <24h
  - Persistierende Hypotonie trotz adäquater Volumensubstitution, welche den Einsatz von Vasopressoren notwendig macht, um einen arteriellen Mitteldruck von  $> 65$  mmHg aufrechtzuerhalten
  - Serumlaktat  $> 2$  mmol/l (18 mg/dl)
- schriftliches Einverständnis durch den Studienteilnehmer oder dessen gesetzlichen Vertreter
- Alter  $\geq 18$  Jahre
- Gesetzliche Krankenversicherung, da Förderung durch Innovationsfonds des Gemeinsamen Bundesausschusses (G-BA)

Ausschlusskriterien:

- Alter <18 Jahre
- Verweigerung der Studienteilnahme
- Verlegung von der ICU/IMC <72h nach Studieneinschluss wahrscheinlich
- palliatives Therapieziel
- bestehende Therapieziellimitierung
- Tod des Patienten ist bereits bei Studieneinschluss absehbar
- Private Krankenversicherung, da Förderung durch Innovationsfonds des G-BA
- Bereits eingeschlossene Patienten konnten kein weiteres Mal eingeschlossen werden.

Die Hauptanalysepopulation, im folgenden Full Analysis Set (FAS) genannt, umfasste alle Patienten und die Auswertung erfolgte nach dem Intention-to-treat-Prinzip, d.h. Patienten wurden in der Gruppe (Interventionsgruppe (IG), Kontrollgruppe (KG)) ausgewertet in die sie randomisiert wurden, unabhängig davon welche Behandlung sie erhielten. Patienten, welche Einschlusskriterien nicht erfüllten oder ein Ausschlusskriterium erfüllten, aber dennoch in der Studie eingeschlossen wurden, wurden nicht in die Analyse einbezogen.

### Studienablauf:

Im Rahmen der Studie erfolgte sowohl für die IG, als auch für die KG eine einmalige Erfassung der Basisdaten zum Zeitpunkt des Sepsisbeginns (= Onset). Die Sepsis-Diagnoseerstellung erfolgte innerhalb von 24 Stunden gemäß der Sepsis-3-Difinition (siehe Einschlusskriterien).

Die klinische Datenerhebung erfolgte analog zur vorgenannten Probengewinnung ebenfalls zum Zeitpunkt des Sepsisbeginns (= Onset) sowie 3, 7 und 14 Tage später. Die abschließende klinische Outcome-Ergebnung für den primären Endpunkt erfolgte zum Zeitpunkt 28 Tage (= 28 d) nach Sepsisbeginn. Zudem wurden einige klinische Outcomes als sekundäre Endpunkte nach 90 sowie 180 Tagen erhoben. Darüber hinaus erfolgte bei Studieneinschluss, nach 90 Tagen sowie 180 Tagen eine Erhebung der gesundheitsbezogenen Lebensqualität (mittels EQ-5D-5L & VR-36). Die gesundheitsökonomische Evaluation der Kosten sowie der Evaluation der Leistungsinanspruchnahmen erfolgten für den patientenindividuellen Beobachtungszeitraum.

Die Zeitpunkte für die Datenerhebungen wurden nach folgenden (auf Erfahrungen beruhenden) Kriterien ausgewählt: 28 Tage: Patient ist wahrscheinlich nicht mehr auf der ICU/ICM; 90 und 180 Tage: Patient ist wahrscheinlich aus der ReHa entlassen. Die Analyse der sekundären, klinischen Endpunkte ist in zwei Auswertungszeitpunkte unterteilt: Die erste, frühere Auswertung geht mit der Auswertung des primären Endpunkts 28 Tage nach Studieneinschluss einher. Die zweite, spätere Auswertung der sekundären Endpunkte erfolgte 180 Tage nach Studieneinschluss und somit nach dem Studienende jedes Patienten. Für jeden sekundären Endpunkt wurde bestimmt, ob dieser zum ersten und/oder zweiten Zeitpunkt ausgewertet werden sollte.

In der Kontrollgruppe erfolgte die antiinfektive Therapiesteuerung (Eskalation, Deeskalation, Absetzen) ausschließlich auf Basis der Kultur-basierten mikrobiologischen Standarddiagnostik des jeweils teilnehmenden Studienzentrums. Im Gegensatz dazu wurden bei Patient:innen der Interventionsgruppe zusätzlich zu den vorgenannten Ergebnissen der mikrobiologischen Standarddiagnostik auch noch die Ergebnisse der NGS-Diagnostik zurückgemeldet, so dass diese auch noch in die Therapiesteuerung mit einbezogen werden konnten.

Die geplanten, klinischen Studienzeitpunkte sind in nachfolgender Tabelle 1 zusammengefasst:

*Tabelle 1: Erhebung klinischer Daten an den jeweiligen Studienzentren*

Erhebungszeitpunkte	Check	Visite 1	Visite 2	Visite 3	Visite 4	Visite 5	Visite 6	Visite 7
Zeitfenster		Onset	3d	7d	14d	28d	90d	180d
Kontrolle der Ein- & Ausschlusskriterien	•							
Einverständniserklärung	•							
Anamnesedaten		•						
Klinische Datenerhebung		•	•	•	•			
Next Generation Sequencing (NGS)		•	•					
Blutkulturen		•	•					

PCT-Messung		(•)*	(•)*					
<b>Klinische</b>					•	•	•	
<b>Outcomeerhebung</b>								
<b>Lebensqualitätserhebung</b>		•				•	•	

\*wenn möglich

Die Projektdurchführung konnte planmäßig in 36 Monaten erfolgen. Mit entsprechender Laufzeitverschiebung ergab sich dementsprechend ein Förderzeitraum vom 01.09.2021 bis zum 31.08.2024 (siehe Abb. 1). Für die Studienphase wurden alle Meilensteine zum Großteil planmäßig bzw. mit minimalen Verzögerungen erreicht. So konnte der erste Patient am 16.03.2022 am Universitätsklinikum Essen in die Studie eingeschlossen werden. Die Fallzahlplanung (Ziel Q3/2022: 108 Patienten) wurde aufgrund des verzögerten Beginns an den verschiedenen Studienzentren verfehlt und es wurde eine angepasste Fahlzahlplanung dem Projekträger vorgelegt. Das Rekrutierungsziel von 150 Patienten konnte u.a. Mitte Februar 2023 erreicht werden (Plan 11/2022).

Die Gesamtdauer der Rekrutierungsphase betrug planmäßig 18 Monate, um die Fallzahl von 410 Studienteilnehmern gemäß Auswertungskonzept zu erreichen. Um das planmäßige Rekrutierungsziel erreichen zu können, wurden im Projektverlauf weitere Studienzentren als Kooperationspartner in die Studie mit aufgenommen. Ziel dieser Maßnahme war es, zeitnah eine noch größere Gesamtzahl an Studienzentren zu initiieren, um die Rekrutierungsquote deutlich zu erhöhen und das bestehende Rekrutierungsdefizit auszugleichen. Mit der Gewinnung von insgesamt 6 weiteren Studienzentren konnte zum einen das Rekrutierungsziel planmäßig erreicht, und zum anderen Ausfälle von Kooperationspartnern kompensiert werden. Die Aufnahme weiterer Kooperationspartner wurde seitens des Projekträgers genehmigt. Das Rekrutierungsziel konnte am 04.08.2023 erfolgreich erreicht werden. Damit konnte die Studie knapp 3 Wochen vor dem eigentlich geplanten Rekrutierungsende vorzeitig beendet werden. Basierend auf dem Rekrutierungsende am 08.08.2023 konnte im November 2023 die 90d-Follow-up-Phase abgeschlossen werden, während die 180d-Follow-up-Phase in Q1/2024 planmäßig zu Ende gebracht werden konnte. Der Datenbankschluss für die vollständige Erfassung der Outcome-Daten zu den Zeitpunkten Visite (V) 6 / Tag 90 sowie V7 / Tag 180 erfolgte am 23.04.2024.

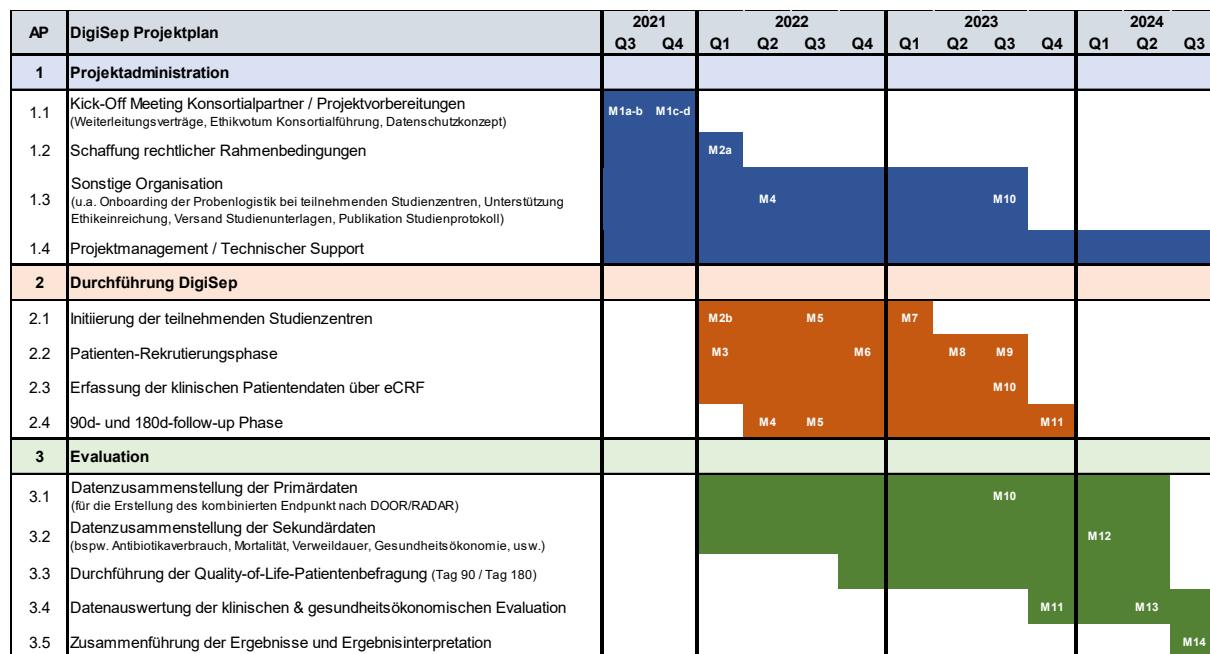


Abbildung 1: GANTT-Chart DigiSep

### Randomisierung:

Eine Randomisierung erfolgte auf Patientenebene. Die Patienten wurden 1:1 in eine der beiden Studienarme (Interventionsgruppe/Kontrollgruppe) randomisiert. Es erfolgte eine Blockrandomisierung, stratifiziert nach Zentrum. Die Randomisierung erfolgte mit Hilfe eines Internet-basierten Tools (Randomizer.at). Die Bedienung des Randomisierungstools erfolgte durch das eingewiesene Studienpersonal (Studienarzt und/oder Studienassistenten) der jeweiligen teilnehmenden Studienzentren.

### Fragestellung:

*Klinische Evaluation (Institut für Medizinische Biometrie Heidelberg)*

Die Effektevaluation untersuchte die Fragestellung, ob bei Patienten mit vermuteter bzw. nachgewiesener Sepsis/septischem Schock (Kriterien gemäß Sepsis-3) die Kombination aus digitaler Präzisionsdiagnostik, Expertenaustausch und Kultur-basierter Standarddiagnostik der rein Kultur-basierten Standarddiagnostik überlegen ist. Konkret wurde untersucht ob die DigiSep-Intervention als neue Versorgungsform (Optimierung und Qualitätssicherung in der Sepsis-Therapie und Nachsorge) bei den in die IG randomisierten Patienten zu signifikanten Verbesserungen bezüglich der definierten Endpunkte im Vergleich zur KG führt.

Als primärer Endpunkt wurde der DOOR/RADAR Score, basierend auf dem Overall Clinical Outcome (OCO) gemäß des DOOR (Desirability of Outcome Ranking)-Prinzips mit Adjustierung gemäß RADAR (Response Adjusted for Duration of Antibiotic Risk) gewählt [12]. Das innovative DOOR/RADAR-Konzept wurde zur Anwendung gebracht, um gleichzeitig mehrere Outputs gewichtet in die Evaluation des primären Endpunkts mit einfließen lassen zu können

(Reduktion der Verweildauer auf der Intensivstation, Tage unter antiinfektiver Therapie, Mortalität, Auftreten eines Nierenersatztherapie-pflichtigen akuten Nierenversagens (ANV)).

Die primäre Hypothese der Studie ist, dass es einen Unterschied zwischen der IG und der KG, jeweils des Full Analysis Set (FAS), bezüglich des DOOR/RADAR-Scores gibt.

Die Nullhypothese  $H_0$ :

$$P(DOOR/RADAR_{IG} > DOOR/RADAR_{KG}) = P(DOOR/RADAR_{IG} < DOOR/RADAR_{KG})$$

wurde gegen die Alternativhypothese  $H_1$ :

$$P(DOOR/RADAR_{IG} > DOOR/RADAR_{KG}) \neq P(DOOR/RADAR_{IG} < DOOR/RADAR_{KG})$$

mit einem zweiseitigen Signifikanzniveau von  $\alpha = 0.05$  unter Verwendung eines Mann-Whitney-U-Tests getestet.

#### *Gesundheitsökonomische Evaluation (Universität Bielefeld)*

Die gesundheitsökonomische Analyse basierte auf der Annahme, dass die zusätzliche Nutzung der Präzisionsdiagnostik (IG) im Vergleich zur ausschließlichen Nutzung der Standarddiagnostik (KG) geringere Kosten verursacht.

Das zentrale Ziel der gesundheitsökonomischen Evaluation bestand darin, die Inanspruchnahme von Leistungen der GKV sowie die daraus resultierenden Kosten zwischen den beiden Studiengruppen (IG und KG) darzustellen und zu vergleichen. Die Analyse wurde entsprechend den Empfehlungen des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) aus der GKV-Perspektive durchgeführt [13]. Neben einer Betrachtung der Gesamtkosten wurden spezifische Leistungsbereiche separat analysiert, darunter: ambulante und stationäre Leistungen, Arzneimittel, Heilmittel, Hilfsmittel, Rehabilitation, häusliche Krankenpflege (HKP), Rettungseinsätze, Arbeitsunfähigkeit, Krankentransporte sowie Pflegeleistungen.

#### *Gesundheitsbezogene Lebensqualität (Universität Bielefeld)*

Darüber hinaus basierte die Evaluation der gesundheitsbezogenen Lebensqualität auf der Annahme, dass durch die zusätzliche Nutzung der Präzisionsdiagnostik im Vergleich zur ausschließlichen Nutzung der Standarddiagnostik in der KG eine höhere gesundheitsbezogene Lebensqualität erzielt werden kann.

#### *Formative Evaluation (Universität Bielefeld)*

Eine formative Prozessevaluation untersuchte zudem Parameter, die Aufschluss über die Akzeptanz sowie fördernde und hemmende Faktoren für die Implementierung der Präzisionsdiagnostik aus Sicht der behandelnden Ärzte geben.

#### **Darstellung der primären und sekundären Endpunkte inkl. der Erhebungs-/Nachbeobachtungszeitpunkte und Erhebungsinstrumente:**

#### *Klinische Evaluation (Institut für Medizinische Biometrie Heidelberg)*

Folgende Einteilung für den DOOR/RADAR Score wurde zugrunde gelegt:

1. Gruppe: Überleben\* + ICU/IMC LOS <12 Tage ohne ANV (akutes Nierenversagen)\*\*
2. Gruppe: Überleben\* + ICU/IMC LOS < 12 Tage mit ANV\*\*
3. Gruppe: Überleben\* + ICU/IMC LOS  $\geq$  12 Tage ohne ANV\*\*
4. Gruppe: Überleben\* + ICU/IMC LOS  $\geq$  12 Tage mit ANV\*\*
5. Gruppe: Tod des Patienten\*

\* erfasst nach 28 Tagen bzw. bis zur Verlegung auf Normalstation.

\*\* mit der Notwendigkeit für die Durchführung einer Nierenersatztherapie.

Danach wird die Dauer der antiinfektiven Therapie einbezogen, und es ergibt sich der DOOR/RADAR-Score wie folgt:

DOOR/RADAR – Score:

$$\left\{ \begin{array}{ll} = 0 & \text{falls } \text{OCO} \leq 4 \\ = 5 & \text{falls } \text{OCO} = 5 \end{array} \right. + \left[ \frac{\text{Anzahl Tage mit Antiinfektiva}}{\text{maximale Anzahl Tage mit Antiinfektiva} + 1 \text{ Tag}} \right]$$

Um der Verabreichung verschiedener Antiinfektiva Rechnung zu tragen, wurde ein modifizierter DOOR/RADAR-Score für die ersten 7 Tage berechnet. Dieser bezieht sich auf die Antiinfektivatage. Ein Antiinfektivatag ist definiert als ein Tag mit einem verabreichten Antiinfektivum, d.h. dass bspw. ein Patient, welcher Antiinfektiva A über 4 Tage und Antiinfektiva B über 3 Tage verabreicht bekommen hat, mit insg. 7 Antiinfektivatagen gewertet wird.

Weiter wurden folgende sekundäre, klinische Endpunkte erhoben (siehe auch Anlage 3 DigiSep - Statistischer Analyseplan) :

Langzeitmortalität (Visite 5/6/7):

Der Endpunkt „Langzeitmortalität“ wurde definiert als die Zeit zwischen Randomisierung und Versterben bzw. Studienende.

Beatmungsdauer (Visite 5/6/7):

Der Endpunkt „Beatmungsdauer“ wurde definiert als die kumulative Beatmungsdauer über den gesamten Studienverlauf, relativ zur Überlebensdauer.

Zeitdauer bis zur Schockresolution (Visite 5/6/7):

Der Endpunkt „Zeitdauer bis zur Schockresolution“ wurde definiert als die Zeit zwischen Beginn des septischen Schocks und Ende des septischen Schocks, erfasst über die kumulative Dauer der Katecholaminpflichtigkeit, relativ zur Überlebenszeit.

#### Andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung (Visite 5/6/7):

Der Endpunkt „Andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung“ wurde definiert als die kumulative Zeit der Nierenersatztherapiepflichtigkeit über den gesamten Studienverlauf, relativ zur Überlebenszeit.

#### Kumulativer Antiinfektivaverbrauch (Visite 5/6/7):

Der Endpunkt „Kumulativer Antiinfektivaverbrauch“ wurde definiert als die Anzahl der Tage, in welchen der Patient ein Antiinfektivum verabreicht bekommen hat. Hierbei wird die Anzahl an Tagen für jedes Antiinfektivum separat erfasst und aufsummiert. Diese Größe wird dann relativ zur Überlebenszeit bewertet.

#### Gesamtverweildauer im Krankenhaus (Visite 5/6/7):

Der Endpunkt „Gesamtverweildauer im Krankenhaus“ wurde definiert als die Zeit zwischen Randomisierung und Entlassung aus dem Krankenhaus. Die Gesamtverweildauer im Krankenhaus wird zum einen relativ erfasst, als Anteil der Tage seit Randomisierung, die der Patient im Krankenhaus verweilt hat.

Zusätzlich wurde auch die Gesamtverweildauer auf ICU analog zur Krankenhausverweildauer untersucht.

#### *Gesundheitsbezogene Lebensqualität (Universität Bielefeld) - (Visite 1/6/7):*

Zur Erfassung der gesundheitsbezogenen Lebensqualität wurde sowohl der EQ-5D-5L als auch der Veterans Rand-36 Item Health Survey (VR-36) eingesetzt.

##### EQ-5D-5L

Der EQ-5D-5L ist ein generisches Instrument zur Messung der gesundheitsbezogenen Lebensqualität [14]. Der Fragebogen wurde sowohl für den Selbstbericht als auch eine Fremdeinschätzung, sofern ein Selbstbericht nicht möglich ist, entwickelt und validiert [15–17]. Das deskriptive System des Fragebogens besteht aus fünf Dimensionen (Beweglichkeit/ Mobilität, Alltägliche Tätigkeiten, Für sich selbst sorgen, Schmerzen/ Körperliche Beschwerden, Angst/ Niedergeschlagenheit) mit je fünf Antwortstufen (keine, leichte, mäßige, große, extreme Probleme oder nicht in der Lage) sowie einer visuellen Analogskala (VAS). Das deskriptive System wird neben den Gesundheitsprofilen zur Berechnung eines präferenzbasierten Indexwertes verwendet. Für die vorliegende Analyse wurden die Präferenzwerte der deutschen Allgemeinbevölkerung von Ludwig et al. (2018) verwendet [18]. Der Indexwert hat damit eine Range von -0,661 bis 1, wobei negative Indexwerte als ein Zustand schlechter als der Tod zu bewerten sind. Sofern ein Patient zum entsprechenden Erhebungszeitpunkt bereits verstorben ist, wird diesem ein Indexwert von 0 zugewiesen [14]. Die VAS umfasst eine Einschätzung des momentanen Gesundheitszustands anhand einer Skala von 0 (schlechtester vorstellbarer Gesundheitszustand) bis 100 (bester vorstellbarer Gesundheitszustand) [14].

##### VR-36

Bei dem VR-36 handelt es sich um ein generisches Indexinstrument zur Erfassung der gesundheitsbezogenen Lebensqualität. Der VR-36 wurde auf Grundlage des 36-Item Short Form Surveys (SF-36) entwickelt. Der VR-36 wurde nicht nur in den USA, sondern auch für den

Einsatz in Deutschland validiert, wodurch seine Anwendbarkeit in unterschiedlichen kulturellen Kontexten gewährleistet ist [19–22]. Der VR-36 umfasst insgesamt 37 Items. 35 Items lassen sich acht Dimension (Körperliche Funktionsfähigkeit, körperliche Rollenfunktion, emotionale Rollenfunktion, körperliche Schmerzen, allgemeine Gesundheitswahrnehmung, soziale Funktionsfähigkeit, Vitalität und psychisches Wohlbefinden) zuordnen, welche sich gemäß des Scoring Manuals zu einzelnen Dimensionsscores transformieren und auf Basis der US-Allgemeinbevölkerung standardisieren lassen. Darüber hinaus lassen sich aus einzelnen Items sowohl eine physische (PCS) als auch eine psychische (MCS) Summenskala bilden. Die Scores für die MCS und PCS ergeben sich als gewichtete Summe der acht standardisierten Dimensionen und stellen jeweils einen möglichen Wertebereich von 0 bis 100 dar, wobei ein Wert von 100 der besten Ausprägung des Gesundheitszustands entspricht. Die Items 36 und 37 im VR-36 beurteilen die körperliche Gesundheit und emotionalen Probleme (z.B. Angstgefühle, Niedergeschlagenheit, Reizbarkeit) zum jeweiligen Befragungszeitpunkt im Vergleich zum Vorjahr. Die Patienten können ihre körperliche Gesundheit und emotionalen Probleme im Vergleich zum Vorjahr auf einer fünfstufigen Skala von „viel besser“ bis „viel schlechter“ bewerten [22].

*Formative Evaluation (Universität Bielefeld):*

Die formative Evaluation wurde als Online-Befragung mittels Unipark durchgeführt und über die beteiligten Studienzentren sowie projektbeteiligte Ärzt:innen verteilt. Das Sampling erfolgte gezielt unter den im Projekt eingebundenen Fachkräften, die per E-Mail zur Teilnahme eingeladen und gebeten wurden, die Befragung auch an weitere projektbeteiligte Ärzt:innen weiterzuleiten. Neben der Studienteilnahme wurden keine weiteren Ein- und Ausschlusskriterien definiert. Die Teilnahme war freiwillig und die Datenerhebung erfolgte pseudonymisiert.

*Retrospektive Evaluation des Expertenpanels (Institut für Medizinische Biometrie Heidelberg)*

Darüber hinaus fand eine retrospektive Evaluation der NGS-Probenergebnisse durch ein unabhängiges Expertenpanel statt. Die Evaluation des 3-köpfigen Expertenpanels fand mittels eines Fragebogens bestehend aus 6 Items statt.

Pro Item wurden die Raten der vollkommenen Übereinstimmung der 3 Bewerter, Mehrheitsübereinstimmung der 3 Bewerter und die Rate der Nichtübereinstimmung berichtet. Die Bewertung ist wie folgt:

*Tabelle 2: Übereinstimmungsbewertung Experten*

+++;---	Vollkommene Übereinstimmung
++; +-; ++fehlend; --fehlend; fehlend, fehlend; +fehlend, fehlend	Mehrheitsübereinstimmung
+fehlend; fehlend, fehlend, fehlend	Keine Übereinstimmung

Zusammenfassend wurde dann die Änderung der Therapieempfehlung und die Art der Therapieempfehlung über eine Übereinstimmung abgeleitet.

Tabelle 3: Outcomes und Erhebungszeitpunkte\*

Outcome	Datenquelle/ Datenerhebung	Zeitpunkt/-raum
Primäres Outcome		
DOOR/RADAR-Score	eCRF	28 Tage
Sekundäre Outcomes		
Langzeitmortalität	GKV-Routinedaten oder eCRF	90 und 180 Tage
Ausmaß der Organdysfunktion/-versagen:	eCRF	28, 90 und 180 Tage
Beatmungszeit		
Zeitdauer bis zur Schockresolution		
Andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung		
Gesamtverweildauer im Krankenhaus	GKV-Routinedaten und eCRF	28 Tage, 90 und 180 Tage
Kumulativer Antibiotikaverbrauch	eCRF	28, 90 und 180 Tage
Zeitpunkt des Beginns der zielgerichteten Antibiotikatherapie	eCRF	28, 90 und 180 Tage
Leistungsinanspruchnahme (ambulant und stationär)	GKV-Routinedaten	innerhalb von 180 Tagen
Arbeitsausfall	GKV-Routinedaten	innerhalb von 180 Tagen
Versichertenkosten	GKV-Routinedaten	innerhalb von 180 Tagen
Gesundheitsbezogene Befragung der Lebensqualität	Erhebung mittels EQ-5D-5L und VR-36	0, 90 und 180 Tage 90 und 180 Tage
Akzeptanzevaluation der Anwender	Fragebogen	Einmalig im Projektverlauf
Retrospektive Analyse durch unabhängiges Expertenpanel	Fragebogen	Nach Studienablauf

\* Auch wenn die Erhebung zu Tag 28,90 und 180 erfolgten, so wurden die Analysen für die Zeitintervalle bis Tag 28 sowie bis Tag 180 durchgeführt.

## Fallzahlberechnung

Für die Planung von DigiSep wurde eine differenzierte Outcome-Analyse der Studienpatienten von 4 großen deutschen Sepsis-Studien (VISEP, MAXSEP, HYPRESS, SISPCT) zur Verfügung gestellt, die sich folgendermaßen auf die u.g. OCO-Graduierungen von DigiSep verteilt hätten: OCO 1: 33%, OCO 2: 3 %, OCO 3: 22 %, OCO 4: 14 %, OCO 5: 28 %. Es wurde davon ausgegangen, dass im Mittel eine antiinfektive Therapiedauer von 7 Tagen in Stufe 1 und 2 des OCO und von 15 Tagen in Stufe 3 und 4 mit einer Standardabweichung von 10 Tagen notwendig ist [4, 10]. Nimmt man nun an, dass sich bei 42% der Patienten in der DigiSep-IG das OCO um 2 Stufen verbessert und im Mittel die Antiinfektivagabe um 2 Tage verkürzt wird. Somit ergab sich bei 180 Patienten pro Studienarm eine Power von 90%, unter der Annahme, dass mit einem relativen Effekt  $p$  von 0.6 ein Patient aus der IG einen niedrigeren DOOR/RADAR-Score als ein Patient aus der KG aufweist. Unter der Annahme einer 12.5%-igen Drop Out-Rate wurde eine finale Fallzahl von insg. 410 Patienten (n=205 pro Arm) errechnet.

## Datenerhebung:

### *Primärdaten/ Klinische Evaluation (Institut für Medizinische Biometrie Heidelberg)*

Für die Erhebung der Primärdaten wurde ein elektronischer Case Report Form (eCRF) genutzt, der in der webbasierten Software „REDCap“ ([www.project-redcap.org](http://www.project-redcap.org)) implementiert wurde. Die Erhebung der Daten erfolgte über das Internet mittels einer verschlüsselten Verbindung (HTTPS-Protokoll). Nur berechtigte Personen hatten Zugang zu den Daten. Um die Datenqualität zu gewährleisten, wurden die Daten auf Vollständigkeit und Plausibilität überprüft. Dazu wurden vorab die Validierungsregeln in einem Datenvalidierungsplan festgelegt. Entstehende Nachfragen zu den Daten wurden an die Studienzentren zurückgemeldet und müssen von diesen beantwortet werden. Änderungen an den Daten wurden in einem Audit Trail protokolliert. Alle im eCRF erhobenen Daten wurden zur Auswertung in das statistische Analysesystem SAS importiert und zur Analyse aufbereitet.

Bei DigiSep handelte es sich um eine nicht-verblindete Studie, so dass das datenerhebende Klinikpersonal, das datenverarbeitendes Auswertungspersonal sowie die beteiligten Patient:innen über die Gruppenzugehörigkeit Kenntnis hatten.

Ebenfalls lag bei dem optionalen infektiologischen Expertenkonsil keine Verblindung vor, da für die ausführliche Beratung die vorhandenen NGS-Befunde der Studienpatienten vorlagen mussten. Lediglich die Experten für das retrospektive Expertenpanel hatten keine Kenntnis davon, ob die Patienten der IG oder der KG zugehörig waren.

Die Adhärenz der durchzuführenden NGS-Testung in den teilnehmenden Zentren wurde durch eine verpflichtende Abfrage im eCRF sichergestellt und geprüft. Ein Bestandteil der Initiierung der teilnehmenden Studienzentren war u.a. die Einweisung in die Studiendurchführung inkl. NGS-Diagnostik. Zudem wurden vorab 1-2 NGS-Probetestungen inkl. Probenlogistik durchgeführt.

### *Gesundheitsökonomische Evaluation (Universität Bielefeld)*

Die Grundlage der gesundheitsökonomischen Evaluation bildeten GKV-Routinedaten, die von den drei Konsortial-Krankenkassen (BARMER, TK, AOK Rheinland-Hamburg) bereitgestellt wurden. Nach der Standardisierung der Datenformate erfolgte eine Prüfung der Datensätze

auf Vollständigkeit und Plausibilität. Dabei wurden potenziell fehlende oder doppelte Datensätze sowie implausible Werte identifiziert und, falls erforderlich, korrigiert. Anschließend wurden die für die Evaluation relevanten Variablen aufbereitet. Hierbei wurde ein einheitlicher individueller Beobachtungszeitraum definiert, der bei allen Patienten den Zeitpunkt des initialen Studieneinschlusses sowie einen Nachbeobachtungszeitraum von 180 Tagen umfasste.

Die identifizierenden Stammdaten der Versicherten wurden ausschließlich zur Zuordnung der GKV-Routinedaten zur Studien-ID verwendet und von der Vertrauensstelle verwaltet. An die Universität Bielefeld wurden schließlich pseudonymisierte Daten übermittelt. Die Übermittlung dieser pseudonymisierten Daten erfolgte mit Zustimmung der Patienten. Die übermittelten Stammdaten (Alter, Geschlecht, Wohnregion) wurden ausschließlich für statistische Analysen verwendet. Rückschlüsse auf die Identität der Versicherten oder deren Zugehörigkeit zu einer bestimmten Krankenkasse waren ausgeschlossen.

Die bereitgestellten Daten umfassten die folgenden Kategorien:

- Stammdaten: Geburtsjahr, Geschlecht, Wohnregion
- Verlauf des Versichertenstatus: Eintritt, Austritt, Todesfall (sofern zutreffend), Personengruppe, Versichertenstatus
- Diagnose- und Leistungsdaten: ambulante ärztliche Versorgung, Arzneimittelversorgung, stationäre Krankenhausversorgung, Hilfsmittelversorgung, Heilmittelversorgung, Arbeitsunfähigkeit und Krankengeld, Verlauf von Pflegegrad und Pflege, HKP, Krankentransporte.

Die Kosten der Intervention wurden anhand des im Rahmen der DigiSep Studie vereinbarten Selektivvertrags und der angefallenen Studienkosten bestimmt. Im Rahmen der Studie wurde für die KG ebenfalls eine NGS-Diagnostik durchgeführt, wenngleich diese nicht in Therapieentscheidungen berücksichtigt wurde, sodass im Rahmen des Projekts 723,4 € pro KG-Patient berechnet wurden. Für die IG wurden 1.499,21 € pro Probe festgelegt und zwei Proben pro Patient entnommen (Szenario I: Projektkosten). Eine weitere Analyse beruhte auf der Annahme, dass im Regeleinsatz eine Probe pro Patient notwendig sein wird, deren Kosten sich auf 1.499,21 € pro Patient belaufen (Szenario II: Regelversorgung).

#### *Gesundheitsbezogene Lebensqualität (Universität Bielefeld)*

Die Befragungsinstrumente wurden zu zwei (VR-36) bzw. drei (EQ-5D-5L) verschiedenen Zeitpunkten über den Projektverlauf hinweg an die eingeschriebenen Sepsispatienten übergeben bzw. versandt. Der EQ-5D-5L wurde aufgrund seiner kompakten und generischen Eigenschaften bereits bei Aufnahme in das Projekt (T0) an die Patienten oder einem möglichen Proxy (Ehepartner, Kind, Elternteil, sonstige Verwandtschaft oder Dritte: Pflegepersonal/Betreuungspersonal) übergeben und nach dem Ausfüllen wieder eingesammelt. Waren entsprechende Personen nicht persönlich anzutreffen, wurden diese telefonisch kontaktiert. Die Fragebögen zu allen weiteren Erhebungszeitpunkten (T1 nach 90 Tagen, T2 nach 180 Tagen) wurden den Patienten postalisch und datenschutzkonform über die Studienzentren zugestellt. Der Rückversand des Fragebogens erfolgte durch die Patienten selbst. Die Digitalisierung wurde an der Universität Bielefeld durch eine räumlich und organisatorisch von den Evaluatoren abgegrenzte Person durchgeführt. Hierdurch wurde

sichergestellt, dass die Evaluatoren zu keinem Zeitpunkt Informationen über Klarnamen oder Adressen der Versicherten verfügten.

#### *Formative Evaluation (Universität Bielefeld)*

Die Universität Bielefeld führte eine formative Begleitevaluation durch, um die Komponenten der Studie hinsichtlich ihrer Anwendbarkeit und Akzeptanz zu untersuchen. Dafür wurde die Nutzungsintention und das Nutzungsverhalten des teilnehmenden ärztlichen Personals anhand von vier Determinanten bestimmt: Leistungserwartung (Performance Expectancy), Aufwandserwartung (Effort Expectancy), Sozialer Einfluss (Social Influence) und unterstützende Faktoren (Facilitating Conditions). Die Datenerhebung erfolgte mithilfe eines eigens für das Projekt entwickelten Fragebogens, der sich an einem etablierten und validierten Instrument zur Messung der Akzeptanz und Nutzung neuer Technologien (UTAUT - Unified Theory of Acceptance and Use of Technology) orientierte [23]. Der Fragebogen kann Anlage 1 entnommen werden.

#### **Auswertungsstrategie:**

##### *Klinische Outcomes (Institut für Medizinische Biometrie Heidelberg)*

###### *Deskription*

Stetige Variablen und Scores werden mit Mittelwert, Standardabweichung, Median, Q1, Q3, Minimum und Maximum, kategoriale Variablen werden mit absoluten und relativen Häufigkeiten pro Behandlungsgruppe und insgesamt beschrieben. Die Anzahl fehlender Werte wird jeweils mit angegeben; relative Häufigkeiten beziehen sich auf die Anzahl der Patienten ohne fehlende Werte. Stetige Variablen werden zwischen den beiden Behandlungsgruppen per t-Tests verglichen, ordinale und kategoriale Variablen per Chi-Quadrat-Test.

###### *Primärer Endpunkt*

Der Endpunkt ist ein künstliches Konstrukt, dass somit keine Interpretation als intervallskalierte Variable zulässt und so auch empfohlen wird [24]. Somit musste die Analyse per nichtparametrischer Methode, dem Mann-Whitney-U Test, durchgeführt werden. Die primäre Analyse basierte auf allen randomisierten Patienten (ITT-Population). Dabei wurde die Nullhypothese

$$H_0: P(\text{DOOR/RADAR}_{IG} > \text{DOOR/RADAR}_{KG}) = P(\text{DOOR/RADAR}_{IG} < \text{DOOR/RADAR}_{KG})$$

wurde gegen die Alternativhypothese

$$H_1: P(\text{DOOR/RADAR}_{IG} > \text{DOOR/RADAR}_{KG}) \neq P(\text{DOOR/RADAR}_{IG} < \text{DOOR/RADAR}_{KG})$$

mit einem zweiseitigen Signifikanzniveau von  $\alpha = 0.05$  getestet.

Fehlende Werte wurden mit Methoden der multiplen Imputation ersetzt. Die Imputation basierte auf den Variablen Alter, Sequential Organ Failure Assessment Score (SOFA)-Score bei Sepsisbeginn und Überlebensstatus. Es wurden Sensitivitätsanalysen zur Überprüfung der Annahmen des Imputationsmodells, sowie ein „best-case“ und ein „worst case“-Szenario

durchgeführt. Außerdem erfolgte eine Analyse der Fälle ohne fehlende Werte (complete case Analyse). Der primäre Endpunkt wurde zudem dichotomisiert (gutes/schlechtes Outcome, 1 versus 2-5) mittels logistischer Regression ausgewertet. Die Gruppe und zusätzliche Confounder (Alter, Geschlecht, SOFA Score, Komorbiditäten Charlson-Komorbiditäts-Index) wurden in das logistische Regressionsmodell als Kovariaten aufgenommen. Weiter wurde der modifizierte DOOR/RADAR-Score analog zur primären Analyse ausgewertet und die primäre Analyse in verschiedenen präspezifizierten Subgruppen (siehe Anlage 3 SAP-Analyseplan, Seite 19/26 im Kapitel 7.3 ) durchgeführt.

#### *Sekundäre Endpunkte*

Die sekundären Endpunkte wurden mittels deskriptiver statistischer Methoden beschrieben und es wurden sowohl 95% Konfidenzintervalle (KI), als auch deskriptive p-Werte für die jeweiligen Effektschätzer berechnet. Die 90- und 180-Tage-Mortalität wurde zusätzlich mittels Methoden zur Überlebenszeitanalyse (Kaplan-Meier, Cox-proportional-hazards-Regression) evaluiert.

Einzelheiten der Analyse der klinischen Endpunkte wurden in einem statistischen Analyseplan (SAP) präzisiert, der vor Datenbankschluss finalisiert wurde. Dieser ist als Anhang 1 zum Evaluationsbericht beigefügt.

Alle Analysen werden einer validierten R-Umgebung durchgeführt, basierend auf R-Version  $\geq 4.2.0$ . Die folgenden externen Packages werden für den angegebenen Zweck verwendet:

- Für Imputation: MICE [25]
- Für allgemeine Plots: ggplot2 [26]
- Für Survival Plots: survminer [27]
- Für Survival Analysis: survival [28]
- Für Datenaufbereitung: Tidyverse [29].

#### *Gesundheitsökonomische Outcomes (Universität Bielefeld)*

Die Kosten- und Leistungsanalyse wurde nach Studienarm (IG und KG) und einzelnen Leistungsbereichen stratifiziert durchgeführt. Die Kostenanalyse wurde für alle Leistungsbereiche pro Patient durchgeführt. Dabei wurden Nullkosten mitberücksichtigt und als solche in die Evaluation integriert. Die Leistungsinanspruchnahme der ambulanten Versorgung konzentrierte sich auf die Anzahl der Arztkontakte. Für die stationäre Versorgung sowie Rehabilitation wurde die Anzahl der Fälle sowie die Verweildauer herangezogen. Die Auswertung der Arzneimittel, der Heil- und Hilfsmittel, der HKP, der Pflege sowie der Krankentransporte umfasste die Anzahl der Verordnungen (Pharmazentralnummer (PZN) bzw. Positionsnummer). Für die Arbeitsunfähigkeit wurden die Fälle herangezogen. Die Gesamtkosten des jeweiligen Beobachtungszeitraums ergaben sich aus den Kosten der einzelnen Leistungsbereiche. Ausreißer in den Kosten der Leistungsbereiche sowie in den Gesamtkosten wurden auf Basis des 99 %-Perzentils entfernt.

Zunächst wurde eine deskriptive Analyse der Verteilungsmaße (Mittelwert, Median, Standardabweichung, Interquartilsabstand) durchgeführt und visualisiert. Da die Normalverteilungsannahme verletzt wurde, wurde für die Gruppenvergleiche wiederum der

zweiseitige Wilcoxon-Rangsummentest (Mann-Whitney-U) verwendet [30, 31]. Bei der IG und KG wird von zwei unverbundenen Stichproben ausgegangen.

Für weiterführende Datenanalysen sind darüber hinaus lineare Modelle essenziell [32]. Die lineare Regression versucht, die Beziehung zwischen zwei oder mehr Variablen mit einer linearen Funktion zu beschreiben. Dadurch wird ermöglicht, die Beziehungen zwischen abhängigen und unabhängigen Variablen mathematisch und visuell darzustellen [33]. Verallgemeinerte additive Modelle für Lage-, Skalen- und Formparameter (GAMLSS) ermöglichen Anpassungen von Regressionsmodellen, bei denen die Verteilung der abhängigen Variable nicht zur Exponentialfamilie gehören muss und schiefe sowie kurtotische kontinuierliche und diskrete Verteilungen einschließt. GAMLSS ermöglicht es, alle Parameter der Verteilung der abhängigen Variable als lineare/nichtlineare Funktionen der unabhängigen Variablen zu modellieren [34]. Im Auswahlprozess der Verteilungsannahme sowie der unabhängigen Variablen, die im Modell berücksichtigt werden sollten, wurde ein systematischer Ansatz verfolgt. Zunächst wurde für den Endpunkt Gesamtkosten (inklusive Interventionskosten) ein einfaches Basismodell berechnet, das nur aus einem Intercept bestand. Das Basismodell, welches keine unabhängigen Variablen enthielt, diente als Referenzpunkt und Grundlage für die Auswahl der am besten geeigneten Verteilung der abhängigen Variable Gesamtkosten. Auf dieser Grundlage wurde eine systematische Evaluierung verschiedener Verteilungsannahmen durchgeführt. Abhängig von den Verteilungseigenschaften des Endpunkts wurden verschiedene Verteilungen geprüft und die Modelle anhand des Akaike-Informationskriteriums (AIC) [35] sowie durch visuelle Vergleiche beurteilt. Dabei wurde die Lognormalverteilung als die geeignete Verteilung identifiziert. Im nächsten Schritt wurde ein Modell erstellt, welches die Abhängigkeit der Gesamtkosten von einer Vielzahl unabhängiger Variablen modellierte, die sich bestmöglich am primären Outcome orientierten. Die Wahl der unabhängigen Variablen erfolgte aufgrund theoretischer Überlegungen sowie der Verfügbarkeit vollständiger Daten. Auf diese Weise ergab sich nachfolgendes Grundmodell:

$$Y_i = \beta_0 + \beta_1 * \text{studygroup}_i + \beta_2 * \text{age}_i + \beta_3 * \text{sex}_i + \beta_4 * \text{cci}_i + \beta_5 * \text{ventilation}_i \\ + \beta_6 * \text{dialysis}_i + \beta_7 * \text{observation duration}_i + \beta_8 * \text{n rehospitalisation}_i \\ + \beta_9 * \text{icu duration}_i + \beta_{10} * \text{inpatient duration}_i + \beta_{11} * \text{preyear costs}_i$$

Für das Outcome Gesamtkosten stellt  $Y_i$  die logarithmierten Gesamtkosten aus den genannten Leistungsbereichen des Nachbeobachtungszeitraums je Patient  $i$  dar (vgl. Anlage 4, Abbildung 1 Modellgüte). Durch die Nutzung des Logarithmus wurde sichergestellt, dass keine negativen Werte geschätzt werden (welches für Kostendaten implausibel wäre) [36]. Die geschätzten Koeffizienten  $\beta$  mussten zur Interpretation exponenziert werden und lassen sich dann als prozentuale Änderungen in den Gesamtkosten interpretieren, wenn sich die zugehörige unabhängige Variable um eine Einheit ändert. Das optimale Modell wurde anhand einer Backward Selection identifiziert, um Multikollinearität sowie Overfitting zu reduzieren [35]. Die Güte des Modells wurde mittels Residuenanalysen überprüft. Dafür kamen typische Residuenplots zum Einsatz. Außerdem wurde die Verteilung der Residuen mittels Histogramme und Q-Q Plots dargestellt (vgl. Anlage 4, Abbildung 2 Modellgüte). Darüber hinaus wurde das R<sup>2</sup> bestimmt, um die Erklärkraft des Modells einordnen zu können [31]. Auf

dezidierte Regressionsanalyse der einzelnen Leistungsbereiche wurde, aufgrund nicht signifikanter Ergebnisse, verzichtet.

Eine deterministische Kosten-Effektivitäts-Analyse (CEA) sowie eine probabilistische Sensitivitätsanalyse (PSA) wurde für die GKV-Population durchgeführt. In der deterministischen Analyse wurden die Gesamtkosten sowie Interventionskosten berücksichtigt, während als Effekt die Anzahl der durchschnittlich vermiedenen Todesfälle ab Tag 28 diente. Das inkrementelle Kosten-Effektivitäts-Verhältnis (ICER) setzte die inkrementellen Kosten in Relation zu den durchschnittlich vermiedenen Todesfällen ab Tag 28 (Kosten pro vermiedenem Todesfall):

$$ICER = \frac{Cost_{Intervention} - Cost_{Control}}{Effect_{Intervention} - Effect_{Control}}$$

Um die Stichprobenunsicherheit und die Unsicherheit der Kostenschätzung und des Effekts zu berücksichtigen, kam zusätzlich eine PSA mit dem Bootstrapping-Verfahren zum Einsatz [37], bei der 1.000 neue Stichproben generiert wurden, um die Variation der Schätzungen zu erfassen. Jedes der 5.000 Datensets lieferte mittlere Kosten und Todesfälle in der IG und KG, aus denen inkrementelle Kosten und Effekte berechnet wurden. Diese Ergebnisse wurden in einem Kosten-Effektivitäts-Diagramm visualisiert. Basierend auf den 1.000 Bootstrapping-Ergebnissen wurde eine Kosten-Effektivitäts-Akzeptanzkurve (CEAC) geschätzt, die zeigt, wie wahrscheinlich es ist, dass die Intervention bei verschiedenen angenommenen Zahlungsbereitschaften kosteneffektiv ist. Die Zahlungsbereitschaft definiert den maximalen Betrag, den eine Gemeinschaft bereit ist, für eine Wirksamkeitssteigerung oder eine zusätzliche Nutzeneinheit (hier: ein vermiedener Todesfall) zu zahlen [38, 39]. Die Wahrscheinlichkeit, dass die Intervention als kosteneffektiv angesehen wird, hängt von der angenommenen Zahlungsbereitschaft ab. Die CEAC diente zur Darstellung dieser Wahrscheinlichkeiten.

**Abweichung vom Evaluationskonzept:**

Die gesellschaftliche Analyseperspektive stellt den umfassendsten Ansatz dar, indem sie alle gesundheitsbezogenen Kosten berücksichtigt, unabhängig davon, wer diese trägt. Hierzu zählen neben den direkten medizinischen Kosten auch indirekte (Arbeitsunfähigkeit, Erwerbsunfähigkeit und Wegfall von Restlebensarbeitszeit bei vorzeitigem Tod), intersektorale Kosten (monetäre Effekte auf Bereiche wie Bildung, soziale Sicherheit, Haushalt und strafrechtliche Belange) oder Selbstzahlerleistungen der Patienten. In der vorliegenden Analyse wurde jedoch auf die gesellschaftliche Perspektive verzichtet, da keine Daten zu Kosten anderer Sozialversicherungsträger, anderer Bereiche oder Selbstzahlerleistungen vorlagen. Eine isolierte Betrachtung von Produktivitätsverlusten durch Arbeitsunfähigkeit wäre zudem unzureichend, da sie nur einen Teilaspekt der gesellschaftlichen Perspektive darstellt.

**Abweichung vom Evaluationskonzept:**

In der vorliegenden gesundheitsökonomischen Evaluation wurde die CEA nicht auf Basis des primären Endpunkts (DOOR/RADAR) (Tabelle 8) oder einer Analyse der gesundheitsbezogenen Lebensqualität (Anlage 2, Tabelle 11) durchgeführt, da die beobachteten Effekte innerhalb der GKV-Kohorte als zu gering in Relation zu den Kosten erachtet wurden, um sinnvolle Zusatzinformationen für Entscheidungsträger zu liefern.

Um die gesundheitsbezogene Lebensqualität im Zeitverlauf und im Vergleich zwischen den beiden Gruppen darzustellen, wurden der EQ-5D-5L Index, die EQ-5D-5L VAS sowie die VR-36 PCS und MCS Scores deskriptiv ausgewertet. Im Weiteren wurden auch die einzelnen Dimensionen des EQ-5D-5L deskriptiv analysiert. Vor der Durchführung weiterführender Tests auf Index- bzw. Scoreebene wurde für diesen Endpunkt mittels Shapiro-Wilk-Test und Histogrammen die Normalverteilungsannahme getestet. Da für beide Instrumente keine Normalverteilung vorlag, wurde der Wilcoxon-Rangsummentest (Mann-Whitney-U) gewählt, um Unterschiede zwischen der IG und KG zu identifizieren [30]. Bei der IG und KG wird von zwei unverbundenen Stichproben ausgegangen. Der zeitliche Verlauf wurde anhand der Veränderung zu den entsprechenden Erhebungszeitpunkten (T1 und T2) analysiert und anhand relativer Mittelwerte und Mediane dargestellt sowie eine mehrfaktorielle Varianzanalyse (ANOVA) mit Messwiederholungen durchgeführt [36].

Auf Grundlage der EQ-5D-5L Indexwerte wurden ebenfalls Quality-Adjusted-Life Years (QALY) berechnet [40]. QALYs werden berechnet, indem die Dauer, die eine Person in einem bestimmten Gesundheitszustand verbringt, mit einem gewichteten Indexwert für die Lebensqualität in diesem Gesundheitszustand multipliziert wird. Auf diese Weise werden die beiden zentralen Elemente – gesundheitsbezogene Lebensqualität und Überleben – berücksichtigt. Diese Methodik ermöglicht eine vergleichende Bewertung von Gesundheitsinterventionen hinsichtlich ihrer Auswirkungen auf gesundheitsbezogene Lebensqualität und Lebensdauer [41]. In der inferenzstatistischen Analyse der QALY-Unterschiede (IG und KG) wurde ein longitudinales lineares Mixed-Modell (LLM) verwendet, welches die gleichen unabhängigen Variablen einbezog wie die Regressionsanalyse der Gesamtkosten [42, 43].

#### *Formative Evaluation (Universität Bielefeld)*

Bei der Auswertung der Fragebögen wurde ein rein deskriptiver Ansatz (relative Häufigkeiten) gewählt, um die Daten übersichtlich und intuitiv zugänglich zu machen. Wo möglich, wurden Chi-Quadrat Tests auf (Un-)Abhängigkeit der kategorialen Variablen durchgeführt. Auf weitere inferenzstatistische Tests wurde aufgrund der kleinen Stichprobengröße nicht zurückgegriffen. Bei der visuellen Aufbereitung lag der Fokus darauf, potenzielle Zusammenhänge zwischen demographischen/berufsbezogenen Merkmalen (bspw. Alter, Geschlecht, Berufserfahrung) und den einzelnen Fragebogenitems zu untersuchen. Anstelle von Kreuztabellen wurde die Funktion prop.table in R verwendet, um die proportionalen Verteilungen der Antworten innerhalb der jeweiligen Kategorien zu berechnen. Diese Verteilungen wurden grafisch aufbereitet, um die Ergebnisse anschaulich darzustellen und einen klaren Überblick über die Verteilungsmuster zu ermöglichen.

Die Datenaufbereitung und Analysen erfolgten mithilfe der Statistiksoftware R (Version 4.4.0) unter Nutzung der folgenden Pakete: janitor, dplyr, ggplot2, eq5d, boot, nlme, utils, readxl, lubridate, gamm4, ggpointdensity, rio, tidyverse [26, 29, 44–55].

### 3 Ergebnisse der Evaluation

#### Klinische Evaluation/ Primärdaten (Institut für Medizinische Biometrie Heidelberg)

##### Fallzahlerreichung

Insgesamt wurden 409 Patienten in die Studie eingeschlossen und randomisiert.

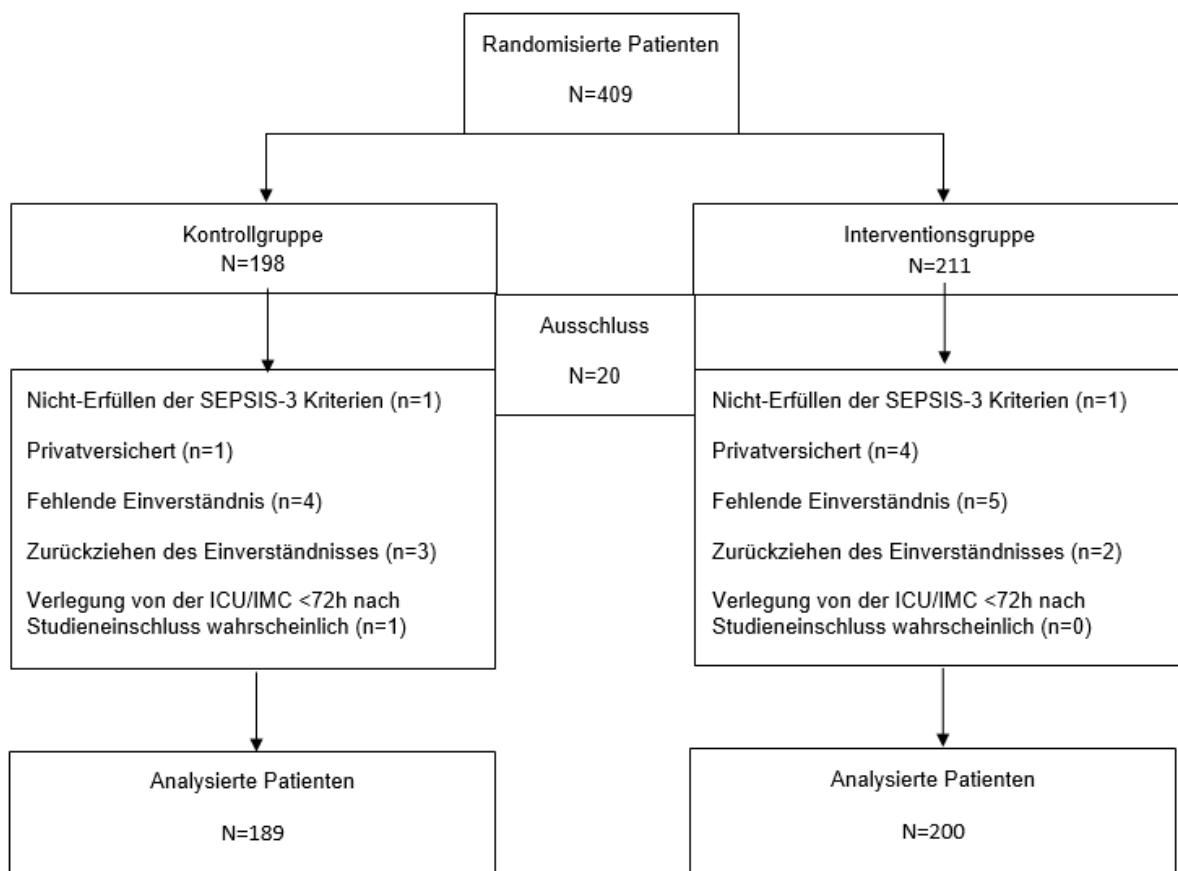


Abbildung 2: Patienten-Flow-Chart mit Ausschluss (Mehrfachgrund möglich)

Abbildung 2 bildet die Rekrutierung der Patienten ab. Insgesamt wurden 20 Patienten (entspricht 4.9%) aus der FAS ausgeschlossen (11 in der IG und 9 in der KG). Somit betrug die Studienpopulation insgesamt 389 Patienten, wovon 200 in der IG und 189 in der KG waren. Zusammen mit 6 Patienten, bei denen der primäre Endpunkt nicht gemessen werden konnte, ergibt sich somit eine Drop-Out Rate von 6,4% die deutlich unter der angenommenen Rate von 12,5% lag.

##### Beschreibung der Gesamtkohorte

In Tabelle 4 sind die wichtigsten Charakteristika der Gesamtkohorte aufgeteilt nach Studiengruppe dargestellt. Durch die Randomisierung wurde erfolgreich eine strukturelle Homogenisierung der Studiengruppen erzielt.

Tabelle 4: Patientencharakteristika Studienkohorte

Variable	Interventionsgruppe (N = 200)	Kontrollgruppe (N = 189)
Alter (Jahren)	64,1±14,3	64,3±15,6
Männliches Geschlecht - Anzahl (%)	120 (60,0%)	118 (62,4%)
Body Mass Index (kg/m2)	27,9±6,9	28,8±7,4
Charlson Comorbidity Index (Punkte)	2,5±2,2	2,3±2,1
Immunsuppression - Anzahl (%)		
Neutropenie	4 (2,0%)	2 (1,1%)
Stammzelltransplantation	1 (0,5%)	1 (0,5%)
Einnahme von Kortikosteroiden	5 (2,5%)	9 (4,8%)
T-Zellen-Immunsuppression	10 (5,0%)	8 (4,2%)
Immundefizienz	1 (0,5%)	2 (1,1%)
Septischer Schock - Anzahl (%)	151 (75,5%)	149 (78,8%)
Sequential Organ Failure Assessment Score (SOFA) (Punkte)	8,5 ±2,7	8,8±2,6
Labordaten		
Laktat (mg/dl)	26,6±29,3	26,7±26,8
C-reactive protein (mg/dl)	19,8±24,6	21,9±26,8
Procalcitonin (ng/ml)	36,8 ± 77,9	29,5 ±54,2
Leukozytenzahl (103/µl)	16,0 ± 9,7	16,7 ± 11,6
Quelle der Infektion - Anzahl (%)		
Ambulant	119 (59,5%)	110 (58,2%)
Nosokomial	81 (40,5%)	79 (41,8%)
Ort der Infektion - Anzahl (%)		
Pulmonal / obere oder untere Atemwege / thorakal	78 (45,6%)	79 (50,6%)
Intraabdominaler / biliärer / gastrointestinaler Trakt	55 (32,2%)	46 (29,5%)
Primäre Bakteriämie / Katheterstelle	11 (6,4%)	7 (4,5%)
Knochen-, Gelenk- und Weichteilinfektionen / Infektionen der Operationsstelle	28 (16,4%)	18 (11,5%)
Zentrales Nervensystem	3 (1,8%)	1 (0,6%)

Kardiovaskulär	5 (2,9%)	1 (0,6%)
Urogenitaltrakt	34 (19,9%)	22 (14,1%)
Andere	4 (2,3%)	4 (2,6%)
Art der Fokussanierung - Anzahl (%)		
Total	102 (51,0%)	88 (46,6%)
Chirurgisch	72 (36,0%)	66 (34,9%)
Interventionell	25 (12,5%)	18 (9,5%)
Entfernen intravasaler Katheter	14 (7,0%)	12 (6,3%)

Das mittlere Alter der eingeschlossenen Patienten betrug 64 Jahre (IG:  $64,1 \pm 14,3$  Jahre vs. KG:  $64,3 \pm 15,6$  Jahre) und diese waren in knapp 2/3 der Fälle männlichen Geschlechts (IG: 120/200 (60,0%) vs. KG: 118/189 (62,4%)). Das Ausmaß der bestehenden Vorerkrankungen war ebenso in beiden Studiengruppen vergleichbar, was anhand des Charlson-Komorbiditäts-Index belegt werden konnte (IG:  $2,5 \pm 2,2$  vs. KG:  $2,3 \pm 2,1$ ). Hinsichtlich der Schwere des septischen Erkrankungsbildes war in beiden Gruppen der septische Schock dominierend (IG: 151/200 (75,5%) vs. KG: 149/189 (78,8%)). Ebenso wies der Sequential Organ Failure Assessment (SOFA)-Score bei Studieneinschluss auf eine hohe Erkrankungsschwere in beiden Studiengruppen hin (IG:  $8,5 \pm 2,7$  vs. KG:  $8,8 \pm 2,6$ ).

Analog zu den allgemeinen Patientencharakteristika wiesen auch die infektiologisch-mikrobiologischen Patientencharakteristika keine signifikanten Unterschiede zwischen den beiden Studiengruppen auf. Die dem Studieneinschluss zugrundeliegende Sepsis war in der Mehrzahl der Fälle ambulant erworben (IG: 119/200 (59,5%) vs. KG: 110/189 (58,2%)) und hatte einen pulmonalen Fokus als Ursache (IG: 78/200 (45,6%) vs. KG: 79/189 (50,6%)). Zudem stellten Erkrankungen im Bereich des Abdomens (IG: 55/200 (32,2%) vs. KG: 46/189 (29,5%)) des Urogenitaltraktes (IG: 34/200 (19,9%) vs. 22/189 (14,1%)) sowie der Knochen- & Weichteile (IG: 28/200 (16,4%) vs. KG: 18/189 (11,5%)) weitere relevante Sepsis-Foki dar. Eine Fokussanierung konnte in knapp der Hälfte der Fälle erfolgen (IG: 102/200 (51,0%) vs. KG: 88/189 (46,6%)), wobei der chirurgischen Fokussanierung die größte Bedeutung zukam (IG: 72/200 (36,0%) vs. KG: 66/189 (46,6%)). Zudem wurden interventionelle Maßnahmen (IG: 25/200 (12,5%) vs. KG: 18/189 (9,5%)) sowie Katheterentfernungen (IG: 14/200 (7,0%) vs. KG: 12/189 (6,3%)) im Sinne einer Fokussanierung durchgeführt.

### Durchgeführte Therapieänderung

Diese Information wurde erst ab Januar 2023 erhoben, als die Rekrutierung schon im Gange war. Somit konnten hier nur Daten von 139 der 200 Interventionspatienten analysiert werden. Die Notwendigkeit einer Erfassung der Therapieadhärenz an die übermittelten NGS-Befunde wurde bei der initialen Studienplanung nicht bzw. nur unzureichend gewürdigt, weswegen diese Erfassung erst zeitverzögert nach Beginn der Rekrutierung in die eCRF-basierte Datenerhebung mit einbezogen wurde. Im Nachhinein hat sich leider genau dieser Aspekt als absolut spielentscheidend herauskristallisiert, da die NGS-Befundadhärenz unzureichend war

und eine NGS-basierte Therapieanpassung nur in etwas mehr als 10% vorgenommen wurde, obwohl diese bei mehr als 30% indiziert gewesen wäre.

*Tabelle 5: Durchgeführte Therapieänderungen in der Interventionsgruppe*

Variable	Total (N=139)
<b>Durchgeführte Therapieänderung Tag 0</b>	
Ja	8 (5,76%)
Nein	131 (94,24%)
<b>Durchgeführte Therapieänderung Tag 3</b>	
Ja	12 (10,08%)
Nein	107 (89,92%)
(Missing)	20
<b>Durchgeführte Therapieänderung insgesamt</b>	
Ja	17 (12,23%)
Nein	122 (87,77%)

Tabelle 5 zeigt, dass bei insgesamt 17 Patienten (12,23%) der IG eine Therapieänderung dokumentiert wurde. Diese Anpassungen stellten in 29,41% der Fälle eine Eskalation, in 52,94 % der Fälle eine Deeskalation und in 17,65% der Fälle ein komplettes Absetzen dar. Die therapeutischen Konsequenzen der rein Kultur-basierten mikrobiologischen Standarddiagnostik wurden nicht systematisch erfasst bzw. ausgewertet. Vielmehr ging es um den Mehrwert der NGS-Diagnostik, wenn diese zusätzlich zur Standraddiagnostik zur Anwendung gebracht wird.

### Evaluation des primären Endpunktes

Für den primären Endpunkt wurde ein Mann-Whitney-U Test mit multipler Imputation durchgeführt. Bei 6 Patienten war der primäre Endpunkt fehlend. Tabelle 6 zeigt die deskriptive Beschreibung des primären Endpunkts.

Tabelle 6: Deskriptive Beschreibung des primären Endpunkts

Variable	IG	KG	Total
	(N=200)	(N=189)	(N=389)
<b>DOOR/RADAR</b>			
N	197	186	383
Missing	3 (1,5%)	3 (1,6%)	6 (1,5%)
Mittelwert	3,21	3,49	3,35
SD	1,54	1,51	1,53
Median	3,62	3,93	3,66
Q1 - Q3	1,41 -- 4,97	1,62 -- 5,00	1,48 -- 5,00
min - max	1,03 -- 5,00	1,00 -- 5,00	1,00 -- 5,00

Der p-Wert der gepoolten Modelle über alle Imputationen hinweg erzielt einen Wert von p=0,149. Somit wurde die statistische Signifikanz des primären Endpunktes verfehlt. Die Wahrscheinlichkeit, dass ein Patient in der IG einen niedrigeren Wert als in der KG aufweist, wurde mit p=0,543 [0,486-0,601] geschätzt (ursprüngliche Fallzahlplanungsannahme p=0,6). In Sensitivitätsanalysen der multiplen Imputation (Variation der Anzahl der Imputation und Donoren) konnte mit einer Range von P=0,135-0,153 die Stabilität des p-Wertes bestätigt werden.

Im Best Case Szenario (bzgl. der fehlenden Werte) wurde ein p-Wert von p=0,058 erreicht, im Worst Case ein p-Wert von p=0,330. Somit hätte auch im Best Case keine statistische Signifikanz erreicht werden können.

Bezüglich der logistischen Regression mit binärem Endpunkt DOOR/RADAR  $\geq/ < 2$  betrug der p-Wert des Gruppenunterschieds p=0,276 mit einem Oddsratio von O=1,32 [0,80-2,18]. Weitere Parameter des Modells können ebenso der Anlage 2, Tabelle 1 entnommen werden.

Im modifizierten DOOR/RADAR Score ergab sich mittels multipler Imputation ein p-Wert von p=0,106 bei einem Mittelwert von  $3,28 \pm 1,43$  in der IG vs  $3,54 \pm 1,38$  in der KG.

Somit können über alle Sensitivitätsanalysen zur primären Analyse konsistente Ergebnisse beobachtet werden: der Effekt weist eine tendenziell positive Wirkung der Intervention auf, erreicht aber nicht statistische Signifikanz.

Nur in der Subgruppe der Patienten mit Fokussierung konnte eine Überlegenheit der Intervention auch mit statistischer Signifikanz nachgewiesen werden (p=0,030). Über die meisten Subgruppen (siehe Subgruppenanalyse in Anlage 2, Tabelle 2) hinweg konnte aber ein homogener Treatment-Effekt erreicht werden. Somit zeigte sich der Therapieeffekt konsistent über fast alle Subgruppen hinweg, jedoch ohne statistische Signifikanz zu erreichen.

### Evaluation der sekundären Endpunkte

Langzeitmortalität:

Die Langzeitmortalität ist in Abbildung 3 über den Kaplan-Meier Schätzer dargestellt, mit abgeleiteten Schätzern zu Tag 28, 90 und 180 in Anlage 2, Tabelle 3. Zum Zeitpunkt 28 Tage waren 99 Patienten verstorben (50 IG vs. 49 KG), zum Zeitpunkt 90 Tage waren es 124 (59 IG vs. 65 KG) und zum Zeitpunkt 180 Tage waren insgesamt 149 Patienten verstorben (64 IG vs. 75 KG).

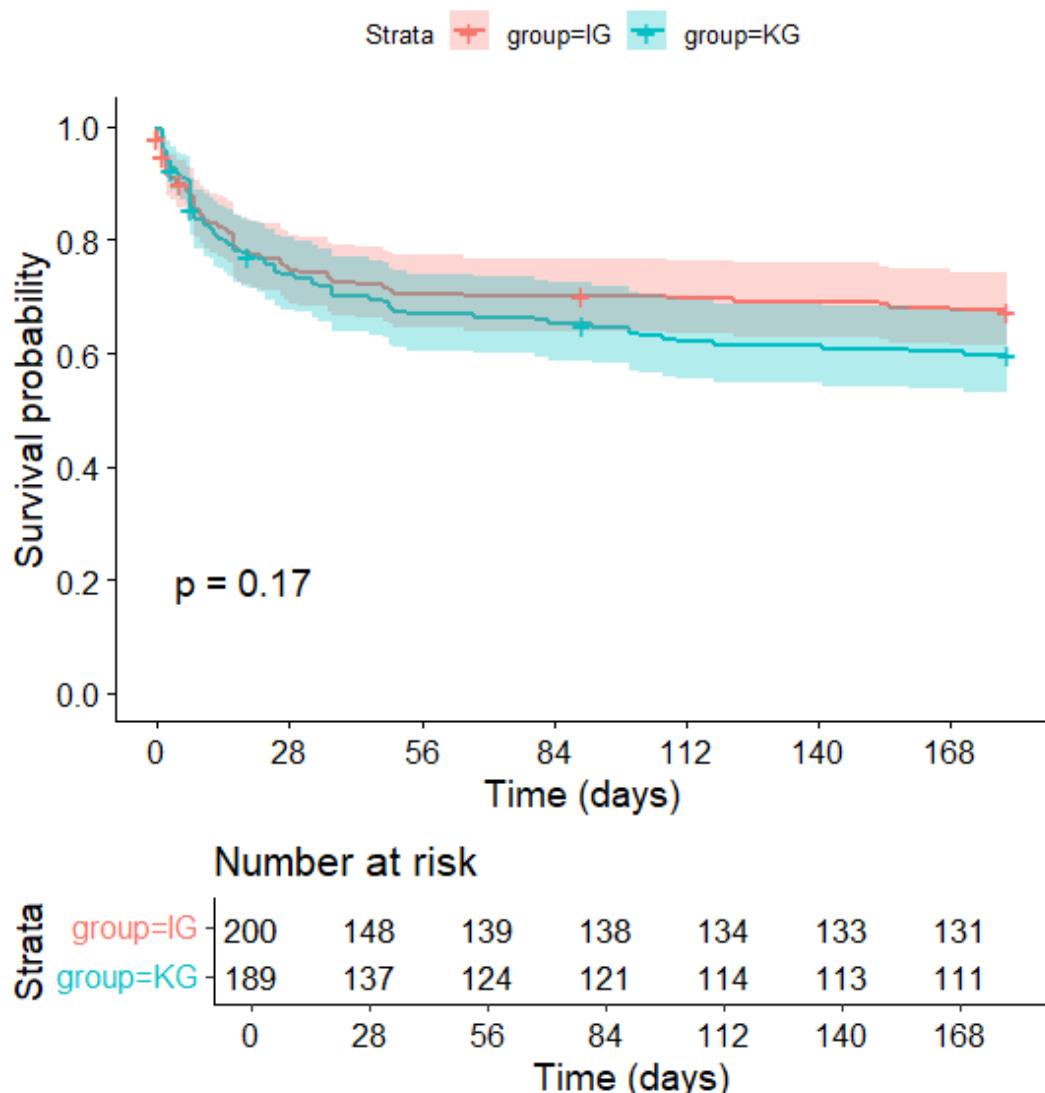


Abbildung 3: Überlebenskurven bis 180 Tage, group=Gruppe, Survival probability=Überlebenswahrscheinlichkeit, Time(days)=Zeit in Tagen, Number at risk = Anzahl unter Risiko

Somit konnte zu Tag 180 eine um 7,7% höhere Überlebenswahrscheinlichkeit in der IG festgestellt werden. Der Unterschied im Gesamtüberleben über den kompletten Studienverlauf konnte mit  $p=0,17$  (bzgl. eines Log-Rank Tests) aber keine statistische Signifikanz erreichen.

## Weitere sekundäre Endpunkte

Die weiteren sekundären Endpunkte wurden zu zwei Zeitpunkten ausgewertet: einmal am Tag 28 und erneut am Studienende (Tag 180). Allerdings könnten die Ergebnisse an Tag 180 verstärkt durch Ausreißer verzerrt sein, weshalb der Auswertung am Tag 28 eine höhere Bedeutung bzgl. der klinischen Evaluation beigemessen werden sollte.

In Anlage 2, Tabelle 4 und 5 sind die Ergebnisse der sekundären Endpunkte zu Tag 28 bzw. 180 Tagen dargestellt. In sämtlichen sekundären Endpunkten zeigte die IG eine Überlegenheit gegenüber der KG, teils sogar mit statistischer Signifikanz. Die Analyse der sekundären Endpunkte machte deutlich, dass die NGS-basierte Diagnostik mit einer signifikanten Verkürzung der absoluten (IG:  $6,6 \pm 9,4$  Tage vs. KG:  $9,3 \pm 10,6$  Tage /  $p=0,025^*$ ) sowie relativen Beatmungsdauer (IG:  $0,35 \pm 0,41$  vs. KG:  $0,44 \pm 0,42$  /  $p=0,034^*$ ) und einer signifikant reduzierten Zeitdauer bis zur Schockresolution (IG:  $6,9 \pm 7,4$  vs. KG:  $8,8 \pm 8,5$  /  $p=0,046^*$ ) assoziiert ist. Zudem zeigte sich in der Interventionsgruppe eine (nicht-signifikante) Verkürzung der Krankenhaus- sowie der Intensivstationsverweildauer um jeweils knapp 2 Tage.

Insbesondere in der Analyse in der Subgruppe der überlebenden Patienten traten deutliche Vorteile der IG hervor. Bis zum Tag 28 nach Sepsis-Onset zeigte sich kein Überlebensunterschied zwischen den beiden Studiengruppen (IG: 75,5% (48 Events) vs. KG: 74,1% (49 Events) /  $p=0.770$ ). Eine Ausnahme bildete der Dauer der antiinfektiven Therapie, in der kein relevanter Unterschied festgestellt werden konnte.

In der 180-Tages-Auswertung (Anlage 2, Tabelle 5) wurden die Effekte, die am Tag 28 beobachtet wurden, erwartungsgemäß durch Ausreißer verzerrt, was zu einer höheren Varianz und einer Abschwächung der Effekte führte. Dennoch bleibt die Überlegenheit der Intervention über alle sekundären Endpunkte hinweg konsistent: In jedem Endpunkt erzielte die IG deskriptiv bessere Werte als die KG. Auch an Tag 180 zeigte sich hier ein Überlebensvorteil in Höhe von 8%, der durchaus als klinisch relevant eingestuft werden muss.

Auch wenn die Rate der tatsächlich durchgeführten Therapieänderungen auf Basis der NGS-Diagnostik in der Interventionsgruppe (IG) erst ab Januar 2023 erfasst wurde, so dürfte die in dieser Subgruppe erhobene Therapieänderungsrate in Höhe von 12,23% durchaus repräsentativ für die gesamte Interventionsgruppe (IG) sein. Die Häufigkeit der durchgeführten Änderungen in der Interventionsgruppe (IG) bleibt damit deutlich hinter den aus vorhergehenden Studien resultierenden Erwartungen zurück, auf Basis derer man eine Therapieänderungsrate in Höhe von mindestens 30% erwartet hätte. Auch wenn sich die Gründe für diese Diskrepanz im Nachhinein nicht durch eine spezifische Auswertung klären lassen, so dürfte diese unzureichende Würdigung aus der mangelnden Erfahrung der teilnehmenden Kliniker mit der neuen Technologie resultieren. Weiterführende Forschungsprojekte, die sich mit der klinischen Anwendung der NGS-Diagnostik beschäftigen, müssen daher einen starken Fokus auf NGS-Schulungsmaßnahmen sowie verpflichtende Expertenkonsile legen, im Rahmen derer dann jeder einzelne NGS-Befund tagesaktuell beraten und in eine adäquate therapeutische Konsequenz umgesetzt wird.

NGS-Proben konnten binnen der ersten 14 Tage des intensivmedizinischen Zeitraums abgenommen werden. Die NGS-basierte Therapieänderungen sind während dieses

Behandlungszeitraums bei den Studienteilnehmern der Interventionsgruppe (IG) vorgenommen worden.

### **Retrospektive Evaluation des Expertenpanels**

Die Ergebnisse der retrospektiven Auswertung des Expertenpanels sind in Tabelle 7 und Anlage 2, Tabelle 6 dargestellt.

*Tabelle 7: Abgeleitete Empfehlungen der Experten*

Variablen	Total (N=389)
<b>Therapieänderung empfohlen</b>	
Ja	125 (32,38%)
Nein	261 (67,62%)
<b>Art der Therapieänderung</b>	
Absetzen_Verkürzen	2 (0,52%)
Eingrenzen	70 (18,13%)
Erweitern	26 (6,74%)
Erweitern_und_Eingrenzen	1 (0,26%)
Keine_Mehrheit	26 (6,74%)
Maintain	261 (67,62%)

In 125 der 389 (32,38%) Fälle sprachen sich die Experten mehrheitlich für eine Therapieänderung aus, was deutlich über dem Wert der tatsächlich dokumentierten Therapieänderungen liegt (12,23%) liegt. Auch wenn die Rate der tatsächlich durchgeföhrten Therapieänderungen auf Basis der NGS-Diagnostik in der Interventionsgruppe (IG) erst ab Januar 2023 erfasst wurde, so dürfte die in dieser Subgruppe erhobene Therapieänderungsrate in Höhe von 12,23% durchaus repräsentativ für die gesamte Interventionsgruppe (IG) sein. Die Häufigkeit der durchgeföhrten Änderungen in der Interventionsgruppe (IG) bleibt damit deutlich hinter den aus vorhergehenden Studien resultierenden Erwartungen zurück, auf Basis derer man eine Therapieänderungsrate in Höhe von mindestens 30% erwartet hätte. Auch wenn sich die Gründe für diese Diskrepanz im Nachhinein nicht durch eine spezifische Auswertung klären lassen, so dürfte diese unzureichende Würdigung aus der mangelnden Erfahrung der teilnehmenden Kliniker mit der neuen Technologie resultieren. Weiterführende Forschungsprojekte, die sich mit der klinischen Anwendung der NGS-Diagnostik beschäftigen, müssen daher einen starken Fokus auf NGS-Schulungsmaßnahmen sowie verpflichtende Expertenkonsile legen, im Rahmen derer dann jeder einzelne NGS-Befund tagesaktuell beraten und in eine adäquate therapeutische Konsequenz umgesetzt wird.

### Gesundheitsökonomische Evaluation (Universität Bielefeld)

#### Beschreibung der GKV-Kohorte

Die gesundheitsökonomische Evaluation der Studie beruht auf den GKV-Routinedaten und wurde aus der GKV-Perspektive durchgeführt (vgl. Kapitel 2).

Von der Gesamtkohorte konnten für 33,41 % Patienten ebenfalls GKV-Routinedaten übermittelt werden. In der GKV-Kohorte konnten somit von 130 übermittelten Fällen 129 in die weiteren Analysen aufgenommen werden, mit einer Verteilung von 75 in der IG und 54 in der KG. Die Überprüfung der Baseline-Charakteristika dieser spezifischen GKV-Kohorte ergibt keine signifikanten Unterschiede in den wesentlichen Merkmalen zwischen den Patienten der IG und KG (Tabelle 8).

Tabelle 8: Baseline Charakteristika der GKV-Kohorte

Characteristika	Interventionsgruppe (N = 75)	Kontrollgruppe (N = 54)	P Wert
Alter (Jahren)	62,27±14,5	63,81±15,6	0,573*
Männliches Geschlecht - Anzahl (%)	40 (53,0%)	33 (61,0%)	0,379**
Body Mass Index (kg/m2)	27,22±6,37	29,18±7,37	0,124*
Charlson Comorbidity Index (Punkte)	244±2,41	2,17±2,04	0,488*
Immunsuppression - Anzahl (%)			
Neutropenie	0 (0,0%)	1 (2,0 %)	0,484**
Stammzelltransplantation	0 (0,0%)	0 (0,0 %)	-
Einnahme von Kortikosteroiden	3 (4,0 %)	3 (6,0 %)	0,615**
T-Zellen-Immunsuppression	4 (5,0%)	2 (4,0 %)	0,457**
Immundefizienz	0 (0,0 %)	0 (0,0 %)	0,394**

\*t-test \*\*chi-quadrat-Test

Im primären Endpunkt (DOOR/RADAR Score) sowie weiteren sekundären Endpunkten nach 28 Tagen, konnten keine signifikanten Effekte in der GKV-Kohorte identifiziert werden (Tabelle 9).

Tabelle 9: Primäre und sekundäre Endpunkte (GKV-Kohorte)

	Interventions- gruppe (N=75)	Kontroll- gruppe (N=54)	Effekt (95% KI)	P Wert
Primärer Endpunkt				
DOOR-RADAR-Score	3,40±1,42	3,38±1,45	0,02 [-0,50; 0,53]	0,829*

Sekundäre Endpunkte zu Tag 28				
Beatmungsdauer				
Relativ	0,37±0,42	0,49±0,44	-0,12 [-0,27; -0,04]	0,132*
Absolut bzgl. Überlebender	7,68±10,26	9,68±11,06	-2,01 [-6,37; 2,37]	0,366*
Zeit bis zur Schockresolution				
Relativ	0,37±0,38	0,46±0,39	-0,09 [-0,23; 0,05]	0,225*
Absolut bzgl. Überlebender	6,60±7,64	8,49±8,38	-1,89 [-5,18; 1,49]	0,256*
Dauer der Nierenersatztherapie				
Relativ	0,17±0,33	0,10±0,28	0,07 [-0,05; 0,17]	0,265*
Absolut bzgl. Überlebender	3,25±7,94	3,1±7,79	0,15 [-3,05; 3,34]	0,927*
Krankenhausverweildauer				
Relativ	0,84±0,26	0,84±0,27	0,0 [-0,09; 0,1]	0,901*
Absolut bzgl. Überlebender	22,5±7,63	22,1±8,0	0,4 [-2,80; 3,61]	0,804*
Intensivverweildauer				
Relativ	0,67±0,36	0,65±0,37	0,02 [-0,11; 0,15]	0,757*
Absolut bzgl. Überlebender	16,54±10,25	15,29±9,95	1,25 [-2,85; 5,35]	0,546*
Dauer der antiinfektiven Therapie				
Relativ	0,64±0,31	0,62±0,34	0,02 [-0,10; 0,14]	0,748*
Absolut bzgl. Überlebender	15,74±8,07	14,39±8,94	1,35 [-2,15; 4,84]	0,446*

\*t-test

## Gesundheitsökonomische Evaluation

### Krankheitskosten

Anlage 2, Tabelle 7 zeigt die durchschnittlichen sowie medianen Krankheitskosten (in Euro) in der IG und KG innerhalb des patientenindividuellen Beobachtungszeitraum. Da die Daten nicht normalverteilt sind, wurde auf den Wilcoxon-Rangsummentest zurückgegriffen. Insgesamt zeigt sich, dass in der IG die mittleren Kosten pro Patient (96.719,22 € (Standardabweichung (SD): 104.559,40)) niedriger sind als die der KG (99.196,95 € (SD: 111.391,80)), während sich der Median gegenläufig verhält, sodass die medianen Kosten der IG (59.503,06 € (IQR: 95.109,13)) höher liegen als die der KG (57.865,41 € (IQR: 109.062,2)). Die Unterschiede sind jedoch in Relation zu der Gesamthöhe der Kosten marginal und der Gruppenunterschied statistisch nicht signifikant ( $p = 0,883$ ). Die gegenläufigen Tendenzen weisen auf weitere Ausreißer in der KG hin, wenngleich im Rahmen der deskriptiven Analyse der Datensatz bereits um Ausreißer bereinigt wurde.

Mit Blick auf die einzelnen Kostenbereiche wird ebenfalls ersichtlich, dass die teilnehmenden Patienten der IG und KG ähnliche Kosten aufweisen. In keinem der Leistungsbereiche wurden signifikante Gruppenunterschiede identifiziert. Die Standardabweichung der Mittelwerte sowie Interquartilsabstände der Mediane weisen auf eine hohe Streuung der Kosten pro Patient hin.

In Anlage, 2 Tabelle 8 werden die Gesamtkosten der IG und KG über den individuellen Beobachtungszeitraum unter Berücksichtigung der Interventionskosten betrachtet. Im Rahmen der Studie sind in der IG 1.499,21 € je Probe angefallen, was bei zweifacher Probenentnahme 2.998,42 € pro Patient entspricht, während in der KG 723,40 € je Patient für die Probenverarbeitung kalkuliert wurden (Szenario I, Projektkosten). Die Annahme, dass lediglich für Interventionspatienten zusätzliche Kosten für die Erbringung der Intervention entstehen (1.499,21 € je Patient), wurden ebenfalls als Analyseszenario (Szenario II, mögliche Regelversorgung) inkludiert. Bei den Szenarien unterscheiden sich die mittleren Kosten pro Patient marginal in Relation zu den berechneten Gesamtkosten. Während in Szenario I Interventionspatienten mittlere Gesamtkosten von 98.717,64 € (SD: 104.559,45) verursachen, liegen die mittleren Gesamtkosten von Kontrollpatienten bei 99.920,35 € (SD: 111.391,81). In Szenario II liegt die Differenz der mittleren Gesamtkosten der Gruppen bei 978,52 € zugunsten der IG. Auch bei der Betrachtung der Gesamtkosten unter Inklusion verschiedener Szenarien der Interventionskosten zeigt sich, dass keine signifikanten Gruppenunterschiede bestanden. Die Annahmen aus Szenario II (Interventionskosten IG: 1.499,21 €, Interventionskosten KG: 0 €) bildeten die Grundlage für alle weiterführenden Kostenanalysen.

#### *Inanspruchnahme*

Anlage 2, Tabelle 9 stellt die Leistungsinanspruchnahme nach Bereich dar. Es lässt sich weitgehend feststellen, dass die Versicherten in der IG und KG eine ähnliche Inanspruchnahme der verschiedenen Leistungsbereiche aufwiesen und kein signifikanter Unterschied zwischen beiden Gruppen bestand ( $p > 0,05$ ). Lediglich im Bereich der stationären Versorgung wurden Gruppenunterschiede zwischen der IG (0,51 (SD: 0,94)) und der KG (0,20 (SD: 0,56)) in der mittleren Anzahl der Rehospitalisierungen identifiziert ( $p = 0,021$ ).

#### *Regressionsmodell*

In Tabelle 10 sind die Studienergebnisse der Regressionsanalyse hinsichtlich der Gesamtkosten dargestellt. Der Intercept beträgt 9.328,11 (95 %-KI: 8.212,27–15.411,07) und stellt den geschätzten Baseline-Wert der abhängigen Variablen (Gesamtkosten) dar. Die Zugehörigkeit zur Gruppe (IG und KG) hatte keinen signifikanten Einfluss auf die abhängige Variable (1,08; 95 %-KI: 0,84–1,39;  $p > 0,05$ ). Der Schätzer zeigt jedoch deskriptiv, dass die KG 8 % höhere Gesamtkosten verzeichnete. Demgegenüber führte die Nierenersatzpflicht zu einer signifikanten Erhöhung der Gesamtkosten um 52 % (1,52; 95 %-KI: 1,21–2,33;  $p = 0,01$ ). Die Dauer der Intervention zeigte einen signifikanten positiven Effekt: Jeder zusätzliche Tag, den Patienten noch nicht verstorben waren, erhöhte die abhängige Variable, d.h. die Gesamtkosten, um 0,4 % (1,004; 95 %-KI: 1,002–1,006;  $p < 0,001$ ). Ähnlich wurde mit jedem zusätzlichem Tag auf der Intensivstation (gemessen als Verweildauer auf der Intensivstation) ein Anstieg von 1,3 % an Kosten assoziiert (1,013; 95 %-KI: 1,007–1,020;  $p < 0,001$ ). Auch die

gesamte stationäre Verweildauer zeigte einen signifikanten Effekt: Mit jedem zusätzlichen Tag stieg die Zielgröße um 0,9 % (1,009; 95 %-KI: 1,005–1,014;  $p < 0,001$ ). Vorjahreskosten beeinflussten die abhängige Variable minimal, jedoch signifikant (1,000003; 95 %-KI: 1,000–1,000;  $p = 0,01$ ). Das Modell erklärte 65,1 % der Gesamtvarianz ( $R^2 = 0,651$ ), was auf eine gute Modellgüte hinweist (vgl. Anlage 4, Abbildung 2).

*Tabelle 10: Ergebnisse des GAMLSS für die Gesamtkosten*

	Effektschätzer (% Veränderung)	95 %-Konfidenzintervall	p-Wert
Intercept	9.328,11 (-)	8.212,27 – 15.411,07	<b>0,001***</b>
Gruppe (Kontrolle)	1,08 (+8%)	0,84 – 1,39	> 0,05
Nierenersatzpflicht	1,52 (+52%)	1,21 – 2,33	<b>0,01**</b>
Interventionsdauer (Vorzeitige Mortalität)	1,004 (+0,4%)	1,002 – 1,006	<b>0,001***</b>
Verweildauer Intensivstation	1,013 (+1,3%)	1,007 – 1,020	<b>0,001***</b>
Stationäre Verweildauer	1,009 (+0,9%)	1,005 – 1,014	<b>0,001***</b>
Vorjahreskosten	1,000003 (+0,0003%)	1,000 – 1,000	<b>0,01**</b>
Signifikanz: *** $p \leq 0,001$ , ** $p \leq 0,01$ , * $p \leq 0,05$			

## Kosten-Effektivitäts-Analyse

### Deterministische Analyse

Die deterministische CEA basierte auf dem Inkrement der Kosten und der Wirksamkeit zwischen IG und KG. Für die Gesamtkosten (inklusive die Interventionskosten) ergab sich ein Kostenvorteil von 978,52 € zugunsten der IG (vgl. Anlage 2, Tabelle 8). Gleichzeitig wurde eine Reduktion der mittleren Todesfälle um 0,1659 zugunsten der IG ermittelt (Berechnung Effekt s. Anlage 4, Tabelle 1), basierend auf dem Outcome der Mortalität nach 28 Tagen. Das bedeutet, dass die IG sowohl kostengünstiger als auch wirksamer in Bezug auf die Senkung der Mortalität nach 28 Tagen war (sog. dominante Intervention).

Aufgrund der Kosteneinsparungen bei gleichzeitig höherer Wirksamkeit in Bezug auf die Mortalität nach 28 Tagen konnte an dieser Stelle auf die deterministische Berechnung, beispielsweise die Erstellung eines ICER, verzichtet werden, da die Interpretation der Kosteneffektivität auf Basis des Effektschätzers sowie den eingesparten Kosten der IG die Dominanz der Intervention gegenüber der KG nahelegte. Ein ICER wird üblicherweise herangezogen, um das Verhältnis zwischen zusätzlichen Kosten und zusätzlicher Wirksamkeit zu bewerten, was hier obsolet war, da keine zusätzlichen Kosten anfallen. Stattdessen reichen

die Interpretation der Wirksamkeit und der eingesparten Kosten aus, um die Überlegenheit der Intervention gegenüber der KG darzustellen.

#### *Probabilistische Sensitivitätsanalyse*

Anlage 2, Abbildung 1 zeigt die Ergebnisse für den Interventionseffekt und die Gesamtkosten pro Patient auf Grundlage der Daten aus den 1.000 Bootstrap-Replikationen, hier dargestellt als Kosten-Effektivitäts-Diagramm. Auf der x-Achse sind die mittleren vermiedenen Todesfälle aufgetragen. Auf der y-Achse ist die Veränderung der Gesamtkosten pro Patient abzulesen. Der rote Punkt in der Grafik stellt den deterministischen Schätzer des ICER dar. Die blauen Punkte stellen die Ergebnisse der Bootstrapping-Analyse dar, die die Unsicherheit in der Schätzung des ICER visualisieren.

In dem süd-östlichen Quadranten des Kosten-Effektivitäts-Diagrammes sind die Ergebnisse zu sehen, in denen die Intervention sowohl wirksam im Sinne einer Reduktion der Mortalität nach 28 Tagen als auch kostensenkend war. Da sich der rote Punkt im süd-östlichen Quadranten befindet, deutet dies darauf hin, dass die Intervention geringere Kosten verursachte und gleichzeitig eine höhere Wirksamkeit in Bezug auf die Mortalität nach 28 Tagen erzielte. In diesem Fall spricht man von einer dominanten Intervention. In insgesamt 51,1 % der 1.000 Bootstrap-Replikationen war die Intervention dominant, d. h. überlegen gegenüber der KG. Im nord-westlichen Quadranten sind die Beobachtungen zu sehen, in denen die Intervention sowohl schlechter in der Wirksamkeit, als auch teurer war. In diesem Quadranten liegen 166 Beobachtungen, somit war die Intervention in 1 % der Fälle dominant. Der Anteil der Beobachtungen aus den 1.000 Replikationen, in denen die Intervention eine Verschlechterung der Wirksamkeit zu geringeren Kosten aufzeigt, beträgt 0,7 % (süd-westlicher Quadrant) und der Anteil der Beobachtungen, in denen die Schätzungen eine Wirksamkeit der Intervention zu höheren Kosten zeigten, beläuft sich auf 47,2 % (nord-östlicher Quadrant). Insgesamt unterstreicht das Kosten-Effektivitäts-Diagramm die Überlegenheit der Intervention in der Wirksamkeit, jedoch ebenfalls die geringen Unterschiede in den Kosten. Die große Streuung der Bootstrap-Replikationen weist auf eine hohe Unsicherheit der Ergebnisse hin.

Die CEAC veranschaulicht die Wahrscheinlichkeit, dass eine Intervention kosteneffektiv ist, in Abhängigkeit von der Zahlungsbereitschaft pro vermiedenem Todesfall (Anlage 2, Abbildung 2). Auf der x-Achse wird die maximale Zahlungsbereitschaft in Euro dargestellt. Die y-Achse zeigt die entsprechende Wahrscheinlichkeit in Prozent, mit der die Intervention unter dieser Zahlungsbereitschaft als kosteneffektiv eingestuft wird. Die CEAC steigt mit zunehmender Zahlungsbereitschaft. Bei einer niedrigeren Zahlungsbereitschaft liegt die Wahrscheinlichkeit der Kosteneffektivität bei etwa 50–60 %, was bedeutet, dass in diesem Bereich die Entscheidung für oder gegen die Intervention unsicher ist. Mit steigender Zahlungsbereitschaft nimmt diese Wahrscheinlichkeit zu und erreicht ab einer Zahlungsbereitschaft von 200.000 € pro vermiedenem Todesfall eine Wahrscheinlichkeit der Kosteneffektivität von > 90 %.

## **Gesundheitsbezogene Lebensqualität**

### **EQ-5D-5L**

Für den EQ-5D-5L wurden Rücklaufquoten von 99,48 % (T0), 55,02 % (T1, 90 Tage) und 47,39 % (T2, 180 Tage) erzielt. Zu T0 wurde der EQ-5D-5L von 34,1 % der Patienten selbst ausgefüllt. 9,2 % der Fragebögen wurden durch Ehepartner (eigenständig oder telefonisch), 7,2 % wurden durch die Kinder (eigenständig oder telefonisch), 1,7 % wurden durch Elternteile (eigenständig oder telefonisch), 3,4 % durch sonstige Verwandtschaft und 44,4 % wurden durch Dritte (Pflegepersonal/Betreuungspersonal, eigenständig oder telefonisch) stellvertretend für den Patienten ausgefüllt (Anlage 4, Tabelle 2).

Die Ergebnisse der einzelnen Dimensionen des EQ-5D-5L zeigen, dass die schwersten Einschränkungen zu T0 in den Dimensionen Beweglichkeit/Mobilität und Alltägliche Tätigkeiten auftraten, wobei über 50 % der Teilnehmenden die höchsten Problemausprägungen (Level 4 und 5) berichteten. Im Zeitverlauf (T1, T2) verschob sich die Verteilung insbesondere in den Dimensionen Schmerzen/körperliche Beschwerden sowie Angst/Niedergeschlagenheit hin zu niedrigeren Kategorien (Level 1 und 2), was auf eine Verbesserung der Gesundheitszustände hindeutet. Trotz dieser positiven Entwicklungen verbleiben Beweglichkeit/Mobilität und Für sich selbst sorgen als besonders belastete Dimensionen, auch wenn die Häufigkeit schwerster Einschränkungen (Level 5) in diesen Bereichen ebenfalls abnahm. Obwohl leichte Unterschiede in den prozentualen Häufigkeiten zwischen den Gruppen sichtbar sind, zeigen diese insgesamt ein ähnliches Muster in der Verteilung der Problemausprägungen über die Zeit (Anlage 4, Tabelle 3).

Anlage 2, Tabelle 10 zeigt die Ergebnisse der gesundheitsbezogenen Lebensqualität der Patienten mittels EQ-5D-5L Index zwischen IG und KG der Gesamtkohorte sowie der GKV-Kohorte. Anlage 4, Abbildungen 3 und 4 können ergänzende grafischen Darstellungen entnommen werden. Für die gesundheitsbezogene Lebensqualität, basierend auf dem EQ-5D-5L Index, wurden zwischen der IG und KG vergleichbare Ergebnisse festgestellt; es wurden lediglich geringe Unterschiede ermittelt. Die deskriptiven Ergebnisse deuten darauf hin, dass die Intervention in der Gesamtkohorte kurzfristig (T1) eine Verbesserung der gesundheitsbezogenen Lebensqualität erzielen könnte, jedoch keine langfristigen Effekte (T2) zeigte. Während in der GKV-Kohorte zu keinem der Erhebungszeitpunkte signifikante Unterschiede der Indexwerte zwischen IG und KG zu identifizieren waren, erwiesen sich die Unterschiede zwischen der IG (Mittelwert (MW): 0,312 (SD: 0,386)) und KG (MW: 0,208 (SD: 0,373)) zu T1 (90 Tage) als statistisch signifikant ( $p = 0,047$ ). Sowohl in der Gesamtkohorte als auch in der GKV-Kohorte lässt sich deskriptiv erkennen, dass die mittlere gesundheitsbezogene Lebensqualität im Zeitverlauf geringfügig abnahm. Auffällig ist dabei eine deutliche Verschlechterung des Medians, der ab T1 einen Wert von 0 erreichte. Die IQR weisen insgesamt eine hohe Streuung auf, was auf eine heterogene Verteilung des EQ-5D-5L Index hindeutet. Eine zusätzliche Analyse der verbundenen Stichprobe ist Anlage 4, Tabelle 4 zu entnehmen. Die Ergebnisse der mehrfaktoriellen ANOVA der Gesamt- und GKV-Kohorte bestätigten die oben beschriebenen Ergebnisse und zeigen, dass der EQ-5D-5L Index im Zeitverlauf signifikante Veränderungen aufwies (Gesamt:  $p = 0,045$ ; GKV:  $p = 0,008$ ), während zwischen den Gruppen (IG bzw. KG) keine signifikanten Unterschiede festgestellt werden konnten (Gesamt: 0,704; GKV:  $p = 0,907$ ). Ebenso ergab die Interaktion zwischen Gruppe (IG

bzw. KG) und Zeit keinen signifikanten Effekt (Gesamt:  $p = 0,422$ ; GKV:  $p = 0,142$ ), was auf ähnliche zeitliche Veränderungen in beiden Gruppen (IG bzw. KG) hindeutet (Anlage 4, Tabelle 5).

Anlage 2, Tabelle 11 zeigt die Ergebnisse des LLM für QALY-Unterschiede zwischen der IG und KG. Insgesamt sind die Gruppenunterschiede (IG und KG) für beide Kohorten (Gesamtkohorte und GKV-Kohorte) als marginal bis sehr gering einzuschätzen. Die LLM-Analyse ergab für die Gesamtkohorte eine QALY-Differenz von 0,023 (95%-KI: -0,081 bis 0,127) und für die GKV-Kohorte eine Differenz von 0,056 (95%-KI: -0,200 bis 0,223). Anhand des 95%-KI waren in beiden Analysen keine statistisch signifikanten Unterschiede zwischen der IG und KG festzustellen.

Die Studienergebnisse der EQ-5D-5L VAS für die IG und KG über die drei Messzeitpunkte hinweg sind in Tabelle 11 dargestellt. Anlage 4, Abbildungen 5 und 6 können ergänzende grafischen Darstellungen entnommen werden. Zu allen Messzeitpunkten zeigten sich keine statistisch signifikanten Unterschiede der VAS-Werte zwischen der IG und KG, weder in der Gesamtkohorte noch in der GKV-Kohorte. Die Mittelwerte und Mediane stiegen im Verlauf (T0 bis T2) in beiden Gruppen (IG und KG) leicht an, jedoch blieben die Unterschiede zu allen Zeitpunkten nicht signifikant. Eine zusätzliche Analyse der verbundenen Stichprobe ist Anlage 4, Tabelle 6 zu entnehmen. Über die Messzeitpunkte hinweg wurden in beiden Gruppen (IG und KG) allgemeine Verbesserungen festgestellt. Die mehrfaktorielle ANOVA mit Messwiederholungen zeigte in der Gesamtkohorte einen signifikanten Effekt der Zeit ( $p < 0,001$ ), während weder der Gruppeneffekt ( $p = 0,113$ ) noch die Interaktion zwischen Gruppe und Zeit ( $p = 0,369$ ) signifikant waren. In der GKV-Kohorte ergeben sich keine signifikanten Ergebnisse, sodass es keinen Hinweis auf einen systematischen Unterschied der Gruppe (IG bzw. KG) über die Zeit hinweg gibt (Anlage 4, Tabelle 7).

Tabelle 11: Gesundheitsbezogene Lebensqualität - EQ-5D-5L VAS

	Gesamtkohorte		GKV-Kohorte	
	Interventionsgruppe	Kontrollgruppe	Interventionsgruppe	Kontrollgruppe
EQ-5D-5L VAS				
T0 - n	177	170	67	50
MW (SD)	31,49 (23,35)	28,91 (22,17)	32,79 (24,15)	26,44 (16,89)
Median (IQR)	30 (35)	25 (29)	30 (37,5)	30 (20)
p-Wert	0,319		0,230	
T1 - n	68	59	27	12
MW (SD)	53,78 (22,74)	47,24 (24,18)	52,00 (21,93)	55,17 (24,84)
Median (IQR)	50 (32,75)	45 (37,5)	50 (28,5)	59,5 (37,5)
p-Wert	0,112		0,647	
T2 - n	52	49	20	12

	Gesamtkohorte		GKV-Kohorte					
	Interventionsgruppe	Kontrollgruppe	Interventionsgruppe	Kontrollgruppe				
MW (SD)	57,35 (25,29)	55,98 (25,06)	47,60 (22,15)	56,83 (28,11)				
Median (IQR)	62,5 (40)	55 (45)	48,5 (32,5)	60 (35)				
p-Wert	0,822		0,182					
IQR: Interquartilsabstand, MW: Mittelwert, n: Stichprobengröße, SD: Standardabweichung								
Test auf Unterschiede aufgrund fehlender Normalverteilung: Wilcoxon-Rangsummentest für unverbundene Stichproben								
Signifikanz: *** $p \leq 0,001$ , ** $p \leq 0,01$ , * $p \leq 0,05$								

## VR-36

Für den VR-36 wurden Rücklaufquoten von 55,42 % (T1, 90 Tage) und 47,39 % (T2, 180 Tage) erzielt. Tabelle 12 zeigt die Ergebnisse der gesundheitsbezogenen Lebensqualität der Patienten mittels VR-36 zwischen IG und KG der Gesamtkohorte sowie der GKV-Kohorte. Anlage 4, Abbildungen 7 und 8 können ergänzende grafischen Darstellungen entnommen werden.

In der GKV-Kohorte ist auffällig, dass in der IG PCS und MCS schlechter als in der KG bewertet wurden. Signifikante Unterschiede zwischen IG und KG ergaben sich hierbei bei der Bewertung des PCS zum zweiten Befragungszeitpunkt ( $p \leq 0,01$ ). Weder im PCS ( $p = 0,57$ ) noch im MCS ( $p = 0,32$ ) zeigten sich signifikante Unterschiede zwischen den beiden Messzeitpunkten bei denjenigen Patienten in der GKV-Kohorte die zu beiden Messzeitpunkten den VR-36 vollständig ausgefüllt haben (verbundene Stichprobe). Allerdings muss insbesondere in der Subgruppe der GKV-Kohorte die geringe Anzahl der Patienten mit gültigen Antworten berücksichtigt werden. Dies verringert die statistische Power. Detaillierte Darstellungen der Patienten, die zu beiden Messzeitpunkten den VR-36 vollständig ausgefüllt haben (verbundene Stichprobe), sind Anlage 4, Abbildungen 9 und 10 zu entnehmen.

Tabelle 12: Gesundheitsbezogene Lebensqualität - VR-36

	Gesamtkohorte		GKV-Kohorte	
	Interventionsgruppe	Kontrollgruppe	Interventionsgruppe	Kontrollgruppe
VR-36 MCS				
T1 - n	59	46	23	7
MW (SD)	40,40 (12,24)	38,15 (14,21)	38,14 (12,88)	45,79 (13,81)
Median (IQR)	38,39 (19,35)	36,34 (20,16)	35,05 (14,12)	49,76 (13,72)

	Gesamtkohorte		GKV-Kohorte					
	Interventionsgruppe	Kontrollgruppe	Interventionsgruppe	Kontrollgruppe				
p-Wert	0,33		0,13					
T2 - n	47	41	18	11				
MW (SD)	42,22 (12,84)	40,09 (14,64)	38,42 (14,21)	45,52 (13,32)				
Median (IQR)	43,67 (22,24)	39,95 (21,05)	33,53 (18,20)	48,26 (14,76)				
p-Wert	0,52		0,15					
VR-36 PCS								
T1 - n	59	46	23	7				
MW (SD)	32,44 (10,28)	31,40 (9,25)	28,76 (11,39)	37,09 (8,20)				
Median (IQR)	33,05 (17,35)	29,62 (14,21)	28,02 (17,19)	37,86 (7,10)				
p-Wert	0,45		0,09					
T2 - n	47	41	18	11				
MW (SD)	35,48 (12,19)	37,14 (9,57)	29,29 (8,59)	42,15 (10,64)				
Median (IQR)	35,05 (16,24)	36,45 (15,51)	28,27 (11,94)	45,90 (15,09)				
p-Wert	0,44		≤ 0,01**					
GKV: Gesetzliche Krankenversicherung, IQR: Interquartilsabstand, MCS: Psychische Summenskala, MW: Mittelwert, n: Stichprobengröße, PCS: Physische Summenskala SD: Standardabweichung, VR-36: Veterans RAND 36-Item Health Survey								
Test auf Unterschiede aufgrund fehlender Normalverteilung: Wilcoxon-Rangsummentest für unverbundene Stichproben								
Signifikanz: *** p ≤ 0,001, ** p ≤ 0,01, * p ≤ 0,05								

Die Items 36 und 37 im VR-36 beurteilen die körperliche Gesundheit und emotionalen Probleme (z.B. Angstgefühle, Niedergeschlagenheit, Reizbarkeit) der Patienten zum jeweiligen Befragungszeitpunkt im Vergleich zum Vorjahr. Die meisten Patienten beurteilten ihre emotionalen Probleme und körperliche Gesundheit zu beiden Befragungszeitpunkten „ungefähr gleich“, „etwas schlechter“ oder „viel schlechter“ verglichen zum Vorjahr. Es gab in der Gesamtkohorte keine auffälligen Unterschiede zwischen IG und KG (Abbildung 4). In der GKV-Kohorte (Abbildung 5) fiel auf, dass mehr Patienten der IG ihre körperliche Gesundheit und emotionalen Probleme „viel schlechter“, im Vergleich zum Vorjahr, als in der KG

bewerteten. Dennoch muss wieder die geringe Stichprobe in der Subgruppe berücksichtigt werden.

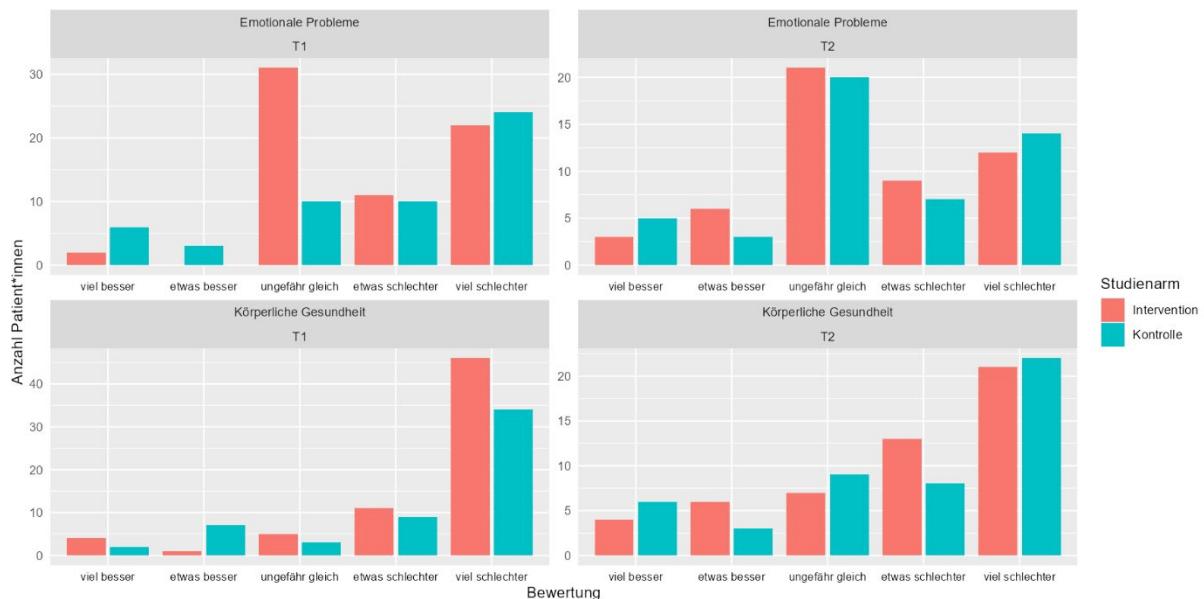


Abbildung 4: Gesundheitsbezogene Lebensqualität im Vergleich zum Vorjahr - VR-36 (Gesamtkohorte)

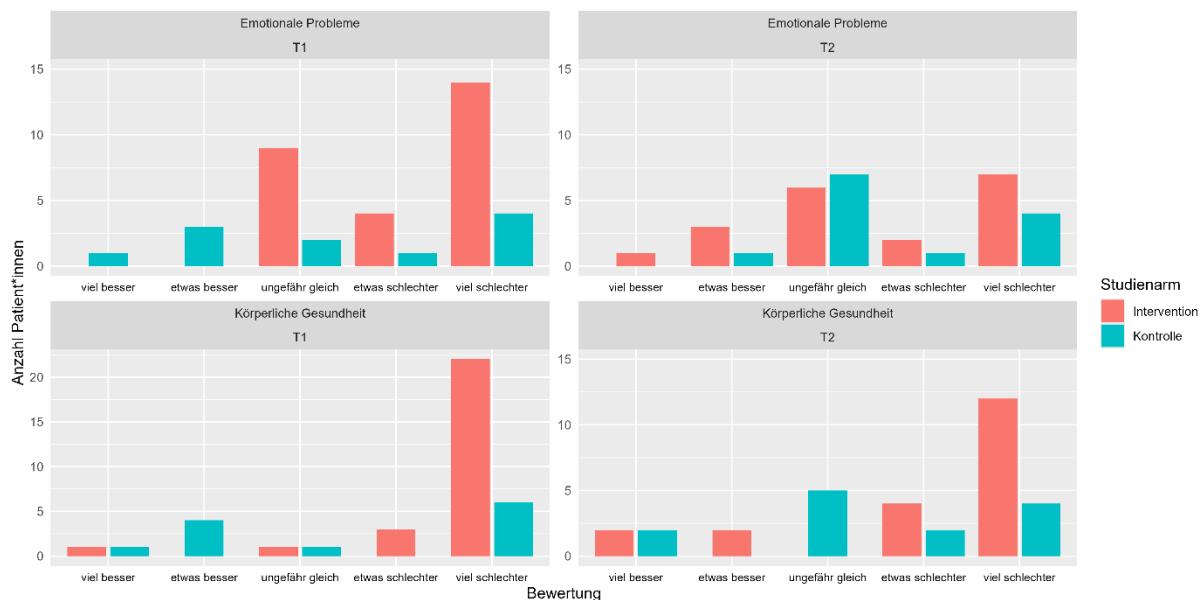


Abbildung 5: Gesundheitsbezogene Lebensqualität im Vergleich zum Vorjahr - VR-36 (GKV-Kohorte)

## Formative Evaluation

Die Befragung richtete sich an medizinisches Fachpersonal mit verschiedenen beruflichen Hintergründen, insbesondere aus dem intensivmedizinischen Bereich. Von den Teilnehmenden ( $n = 44$ ) war die Mehrheit männlich (61,36 %). Das mittlere Alter betrug 43,7 Jahre, mit einer Altersspanne von 28 bis 67 Jahren. Das mittlere Alter bei Frauen betrug 41,5 Jahre (Spanne: 29-58 Jahre), während das mittlere Alter bei Männern 45,6 Jahre (Spanne: 28-67 Jahre) betrug. Ein Großteil der Befragten (75 %) konnte der Berufsgruppe der Oberärzte

bzw. Oberärztinnen zugeordnet werden. Der Rest verteilte sich auf die Berufsgruppen der Fachärzte (11,36 %), Assistenzärzte (6,82 %), Chefärzte (4,55 %) sowie sonstige (2,27 %).

Ein Großteil der Befragten gab an, die Ergebnisse der Präzisionsdiagnostik während der Studie für initiale Therapieentscheidungen genutzt zu haben. Dabei setzten 25 % der Teilnehmenden die Diagnostik konsequent bei allen Patienten der IG ein. Andere wiederum berichteten, sie nur bei einem Teil der Patienten genutzt zu haben. Im Mittel lag der Anteil der Patienten, für welche die Präzisionsdiagnostik zur initialen Therapieentscheidung genutzt wurde, bei ca. 28 %, variierte jedoch stark nach Angabe der Befragten (10-50 %). Einige Teilnehmende (39 %) verzichteten gänzlich darauf, die Diagnostik für initiale Entscheidungen heranzuziehen. Die Nutzung der Präzisionsdiagnostik zur initialen Therapieentscheidung variierte dabei deutlich zwischen den verschiedenen Berufsgruppen. Oberärzte nutzten die Diagnostik am häufigsten, wobei 27 % angaben, diese konsequent bei allen Patienten der IG eingesetzt zu haben, während weitere 39 % sie nur bei einem Teil der Patienten verwendeten. Assistenzärzte und Chefärzte hingegen griffen kaum auf die Diagnostik zurück. Fachärzte und Teilnehmende aus der Kategorie "Sonstige" zeigten ebenfalls eine vergleichsweise geringe Nutzung. Der Chi-Quadrat-Test bestätigte einen signifikanten Zusammenhang zwischen der Berufsgruppe und der Nutzung der Diagnostik.

Die Bereitschaft zur Nutzung der Präzisionsdiagnostik für Patienten mit Sepsis oder septischem Schock hängt stark von spezifischen Bedingungen ab. Viele Befragte würden die Diagnostik einsetzen, wenn eine klinische Verschlechterung oder ein septischer Schock trotz Therapie unklar bleibt. Weitere Faktoren umfassen unklare Erreger- und Infektionsnachweise, fehlende Keimnachweise bei fulminantem Schock, oder langwierige Therapieversuche ohne Erfolg. Gleichzeitig spielen organisatorische Rahmenbedingungen eine wichtige Rolle: Eine vollständige Kostenübernahme durch Krankenkassen, klare Logistik und eine Reduktion der Kosten wurden als entscheidend genannt. Zusätzlich wurde die zeitliche Verfügbarkeit der Diagnostik betont, etwa Ergebnisse innerhalb von Stunden oder einer maximalen Bearbeitungszeit von 24 Stunden direkt am Standort.

Viele der Befragten (43,2 %) sprachen sich dafür aus, die Diagnostik in Zukunft bei allen Patienten mit Sepsis oder septischem Schock einzusetzen. Die Analyse zeigt eine klare Tendenz, dass die Nutzung der Präzisionsdiagnostik während der Studie die Bereitschaft zur zukünftigen Anwendung beeinflusst (siehe Abbildung 6). Besonders Teilnehmende, die die Diagnostik konsequent bei allen Patienten einsetzen, befürworten die zukünftige Nutzung uneingeschränkt (63,6 %). Währenddessen zeigen Personen, die die Diagnostik während der Studie nur selektiv oder gar nicht genutzt haben, eine geringere Bereitschaft zur uneingeschränkten zukünftigen Nutzung. Der Chi-Quadrat-Test deutet auf einen schwachen Zusammenhang zwischen diesen Variablen hin ( $p = 0,066$ ), jedoch ohne statistische Signifikanz.

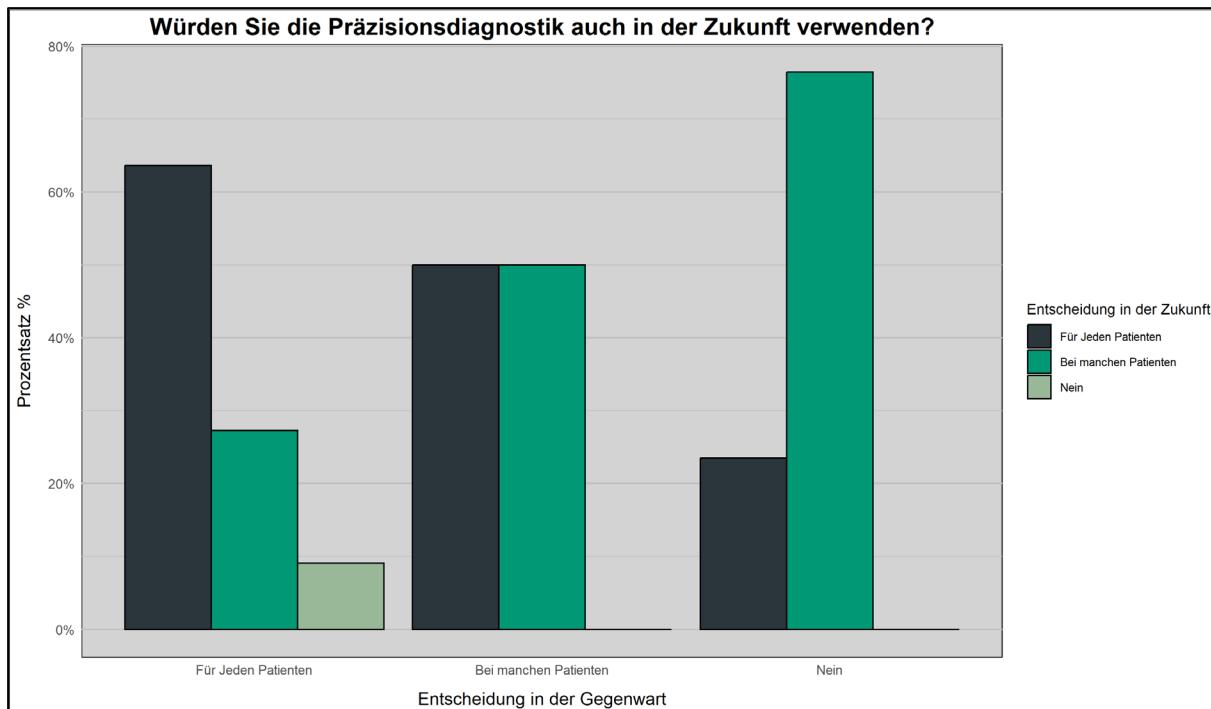


Abbildung 6: Nutzung der Präzisionsdiagnostik in der Zukunft (y-Achse) in Abhängigkeit der gegenwärtigen Nutzung (x-Achse)

54,5 % der Befragten zeigten eine differenzierte Haltung und sahen die Anwendung nur unter bestimmten Bedingungen vor, etwa bei Verfügbarkeit zusätzlicher Ressourcen oder einer spezifischen klinischen Indikation. Ein kleiner Teil der Befragten lehnte eine zukünftige Nutzung außerhalb des Studiensestings ab (17,7%). Die Anwendung der Präzisionsdiagnostik außerhalb der Intensivstation wird vor allem dann als sinnvoll angesehen, wenn diagnostische Lücken bestehen. Dies betrifft Fälle, in denen die konventionelle Diagnostik versagt, persistierende Infektionen ohne positive Blutkulturen auftreten oder seltene Erreger beteiligt sind. Weitere spezifische Anwendungsfälle umfassen beispielsweise die Liquordiagnostik, chirurgische Foki mit unklarer Anamnese, und schwerwiegende Infektionen in ausgewählten Patientengruppen. Zudem betonten die Befragten die Bedeutung einer effizienten Logistik und Kostenreduktion, insbesondere in der ambulanten Medizin. Hinsichtlich der klinischen Relevanz wurde die Anwendung durch erfahrene Intensivmediziner sowie die Ergänzung der Diagnostik zu bestehenden Verfahren hervorgehoben. Zusammengefasst zeigt sich, dass die Präzisionsdiagnostik potenziell breite Einsatzmöglichkeiten bietet, diese jedoch stark von spezifischen klinischen, organisatorischen und logistischen Rahmenbedingungen abhängen.

Die Einschätzung des sozialen Umfelds zeigt eine überwiegend positive Haltung gegenüber der weiteren Nutzung der Präzisionsdiagnostik (siehe Abbildung 7). In Bezug auf den Einfluss von Menschen, die das Verhalten der Befragten prägen, stimmten die Mehrheit der Teilnehmenden zu, dass diese Personen der Meinung seien, die Diagnostik weiterhin zu nutzen. Etwa 75 % gaben an, dass diese Aussage entweder voll zutreffend oder eher zutreffend ist. Ein kleinerer Anteil, weniger als ein Viertel, war hingegen der Meinung, dass dies nicht zutrifft oder eher nicht zutrifft. Ähnlich positiv fiel die Einschätzung der Kollegen

aus. Auch hier stimmte eine deutliche Mehrheit von etwa 75 % der Befragten zu, dass ihre Kollegen denken, sie sollten die Präzisionsdiagnostik weiterhin nutzen („trifft voll zu“ oder „trifft eher zu“). Allerdings gab es ebenfalls einen kleinen Anteil von etwa 25 %, der diese Einschätzung nicht teilt oder nur teilweise teilt.

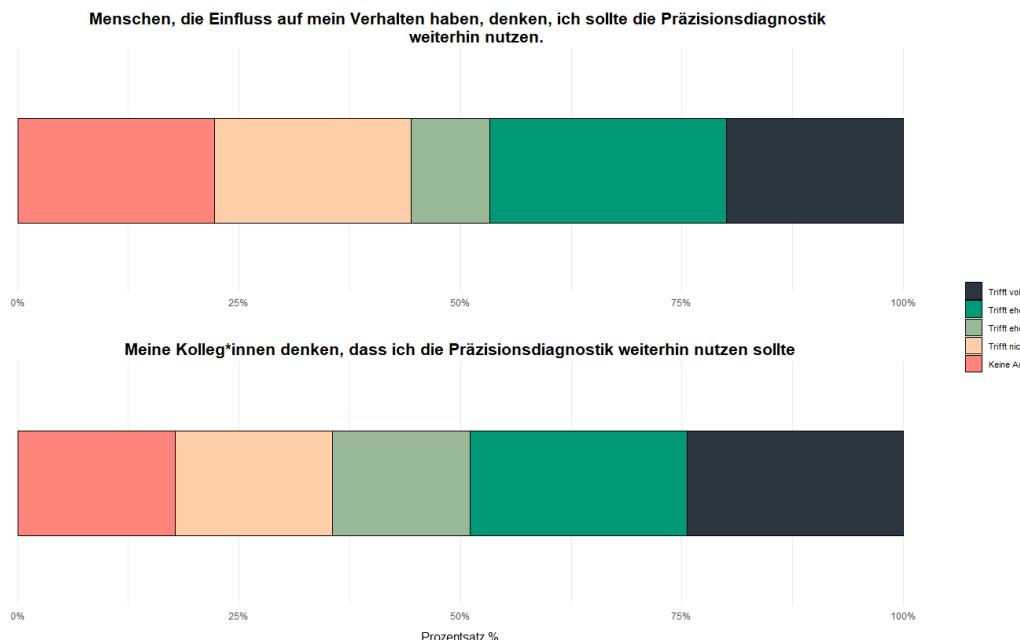


Abbildung 7: Sozialer Einfluss

Die Befragten bewerteten die Leistungserwartungen an die Präzisionsdiagnostik insgesamt positiv. Eine klare Mehrheit stimmte zu, dass die Diagnostik vergleichbare Ergebnisse wie die Standarddiagnostik erzielt (siehe Abbildung 8). Rund 75 % gaben an, dass diese Aussage entweder voll zutreffend oder eher zutreffend ist. Besonders positiv wurde der Beitrag der Präzisionsdiagnostik zur langfristigen Sicherung von Therapieerfolgen gesehen, dem über 80 % der Befragten zustimmten. Darüber hinaus betonten viele Teilnehmende den Mehrwert der Präzisionsdiagnostik für die Verbesserung der medizinischen Versorgung: Über drei Viertel der Befragten bewerteten diese Aussage mit „trifft voll zu“ oder „trifft eher zu“. Auch die Integration in den klinischen Alltag wurde von einer Mehrheit als gut umsetzbar wahrgenommen. Die Unterstützung bei der Therapiesteuerung wurde besonders hervorgehoben, wobei fast alle Befragten dieser Aussage zustimmten.

Allerdings zeigte sich bei der Zeitersparnis ein differenzierteres Bild. Zwar sah ein relevanter Anteil der Befragten die Präzisionsdiagnostik als hilfreich an, jedoch stimmte hier nur eine kleinere Gruppe vollständig zu. Dies deutet darauf hin, dass die Diagnostik zwar als leistungsfähig wahrgenommen wird, der zeitliche Nutzen jedoch nicht von allen als vorrangig erlebt wird.

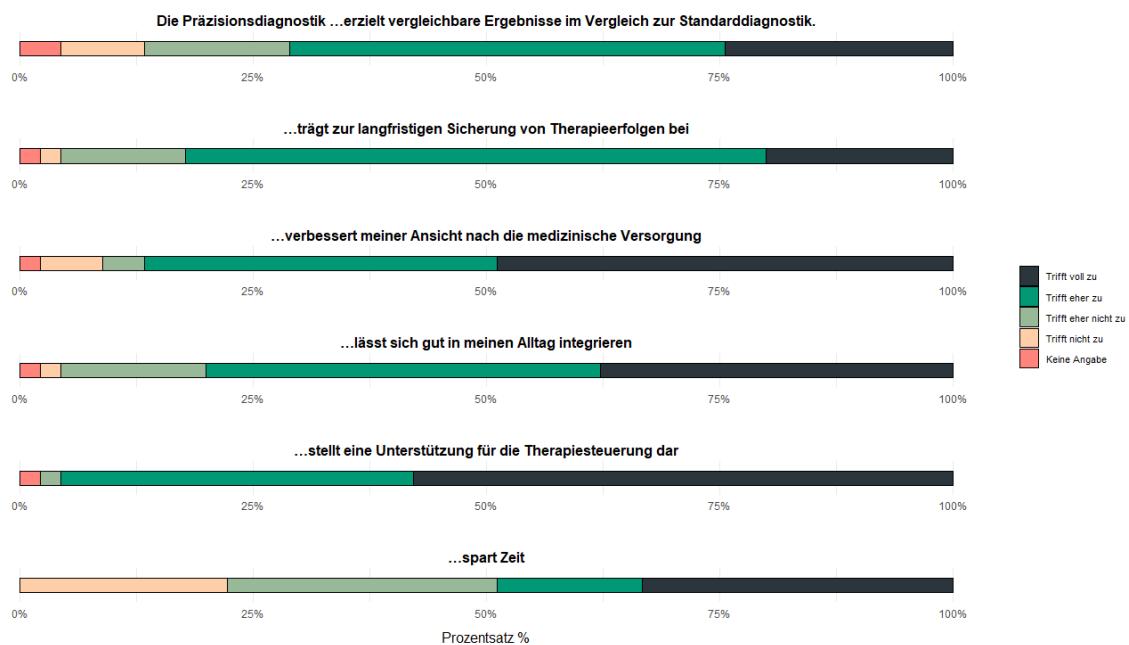


Abbildung 8: Leistungserwartung an die Präzisionsdiagnostik

Die Befragten zeigten insgesamt eine positive Wahrnehmung der Präzisionsdiagnostik, insbesondere im Hinblick auf die Erfüllung der Aufwandserwartungen (siehe Abbildung 9). Über drei Viertel der Teilnehmenden gaben an, dass die Technologie ihre Erwartungen entweder „voll“ oder „eher“ erfüllt hat. Gleichzeitig wurde jedoch deutlich, dass die Nutzung der Präzisionsdiagnostik mit erheblichen organisatorischen und ressourcenbezogenen Herausforderungen verbunden ist. Fast die Hälfte der Befragten empfand einen erhöhten Verwaltungs- und Organisationsaufwand, und ähnlich viele berichteten, dass die Integration der Technologie in den klinischen Alltag schwierig sei.

Zusätzlich wurde ein hoher Schulungsaufwand als wesentliche Herausforderung wahrgenommen, dem über 60 % der Befragten zustimmten. Darüber hinaus stimmte die Mehrheit der Teilnehmenden zu, dass die Nutzung der Diagnostik mit einem hohen allgemeinen Ressourcenaufwand verbunden ist, was sowohl personelle als auch finanzielle Aspekte umfassen könnte.

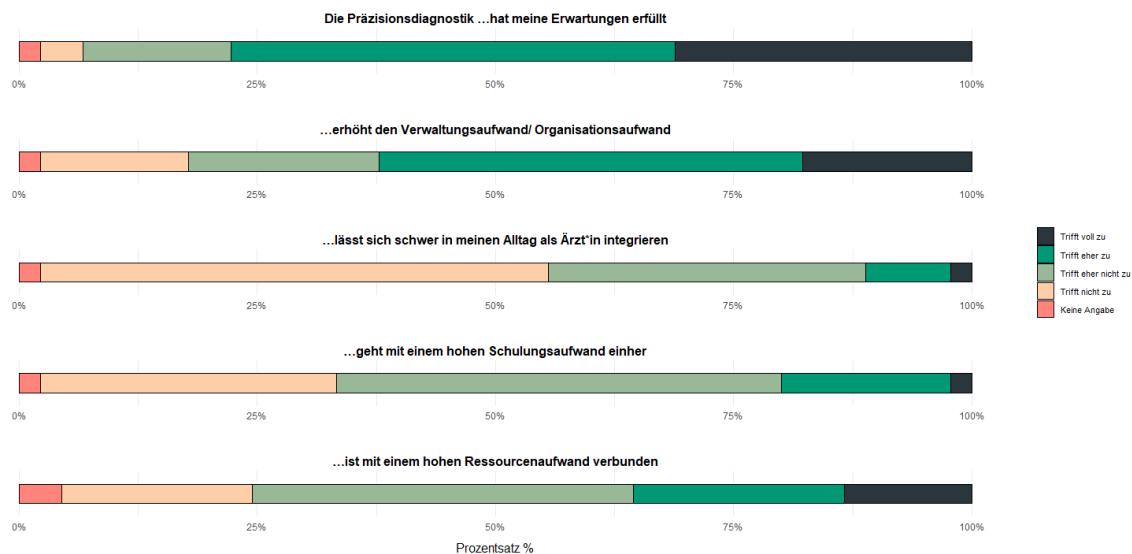


Abbildung 9: Aufwandserwartungen an die Präzisionsdiagnostik

Die Befragten bewerteten die organisatorischen und technischen Rahmenbedingungen der Präzisionsdiagnostik differenziert (siehe Abbildung 10). Eine überwiegende Mehrheit gab an, alle notwendigen Informationen erhalten zu haben, um die Diagnostik anzuwenden, was auf eine umfassende Unterstützung in der Einführung hinweist. Gleichzeitig wurde der Zugriff auf die Ergebnisse von der Mehrheit als einfach bewertet, was auf funktionierende technische Strukturen hinweist.

Jedoch berichteten viele Befragte auch von Herausforderungen: Rund 50 % stimmten zu, dass es häufig zu Problemen beim organisatorischen Ablauf gekommen ist. Zusätzlich gaben etwa 30 % der Teilnehmenden an, dass es häufig zu technischen Fehlern beim Abruf der Ergebnisse kam. Trotz dieser Schwierigkeiten bewertete eine Mehrheit die Verfügbarkeit schneller Hilfe im Falle von Problemen positiv, was zeigt, dass unterstützende Strukturen vorhanden sind, um Anwender in schwierigen Situationen zu unterstützen.

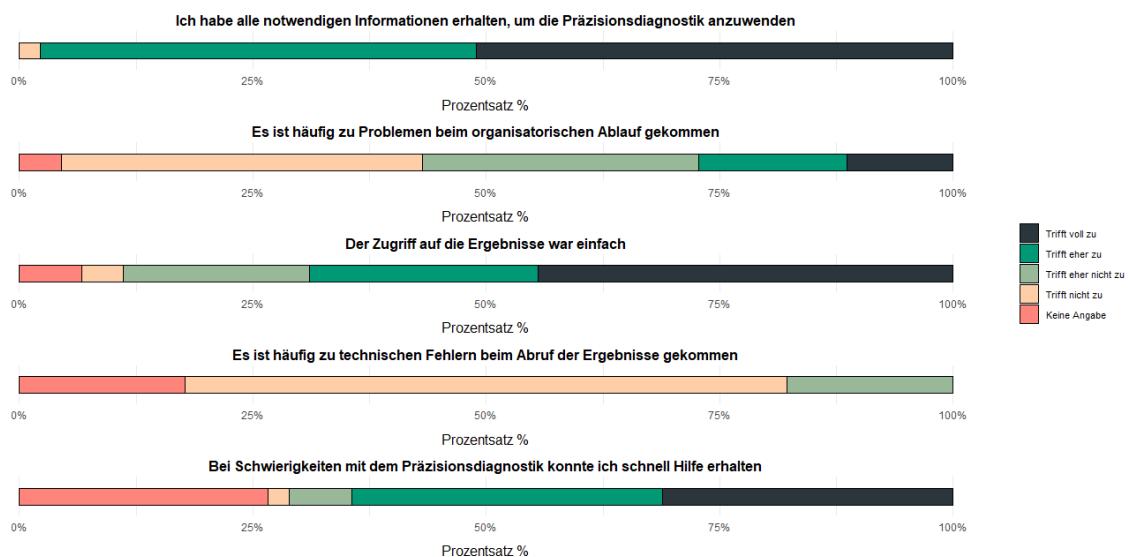


Abbildung 10: Unterstützende Rahmenbedingungen

Ein weiteres Element der Studie war das fakultative Expertenkonsil, dessen Nutzung ebenfalls im Rahmen der formativen Evaluation untersucht wurde. Da nur eine Person das Expertenkonsil während der Studie in Anspruch nahm, ist die Aussagekraft dieses Teils der formativen Evaluation eingeschränkt und alle weiterführenden Fragen wurden von den Teilnehmenden nicht beantwortet.

## 4 Diskussion der Projektergebnisse

### Diskussion der Studienergebnisse:

#### *Klinische Evaluation (Institut für Medizinische Biometrie Heidelberg)*

Die DigiSep Studie stellt eine klinisch prospektive, multizentrische, randomisierte, kontrollierte Interventionsstudie (RCT) dar und entspricht somit einem sehr hohen Evidenzlevel. Die Strukturgleichheit der beiden Behandlungsgruppen konnte durch die Randomisierung erreicht werden. Somit ist die Voraussetzung gegeben, dass in den Endpunkten erkenntliche Unterschiede auf die Intervention zurückzuführen sind.

Im primären Endpunkt, dem DOOR/RADAR-Score, konnten in der IG durchschnittliche bessere Werte als in der KG erreicht werden, jedoch reichte die beobachtete Differenz nicht aus, um statistische Signifikanz zu erreichen: Der gepoolte p-Wert nach multiplen Imputation lag bei  $p=0,149$  und der Schätzer für die Wahrscheinlichkeit eines niedrigeren DOOR/RADAR-Scores in der IG lag mit  $p=0,543$  [95%-CI: 0,486-0,601] unter dem, in der Fallzahlplanung als klinisch relevant angenommenen,  $p=0,6$ . Somit wurde das primäre Studienziel der DigiSep-Studie verfehlt.

Das Verfehlen der statistischen Signifikanz im primären Endpunkt lässt sich höchstwahrscheinlich auf zwei Hauptgründe zurückführen.

- Mortalität bis Tag 28: Die Analyse des Überlebens bis Tag 28 zeigt nur einen geringen Unterschied zwischen der IG und der KG (75,5% vs. 74,1%). Da der DOOR/RADAR-Score

die Mortalität beinhaltet und generell eine relativ hohe Mortalität in DigiSep vorliegt (ca. 25% bis zu Tag 28) wird hierdurch der Effekt verschleiert, da dadurch potentielle Unterschiede, die im Krankheitsverlauf der Patienten zu sehen waren, überlagert werden (da alle Verstorbenen den Wert 5 erhalten). Die Maßnahmen der DigiSep-Intervention waren nicht in der Lage, die Kurzzeitsterblichkeit zu beeinflussen. Wenn man sich den Effekt in der Subgruppe der überlebenden Patienten ansieht, ist dieser einerseits größer als im Gesamtkollektiv (0,33 vs. 0,28) und mit einem p-Wert von p=0,059 sehr nahe an der statistischen Signifikanz, trotz geringerer Power aufgrund einer geringeren Fallzahl.

- Geringe dokumentierte Therapieanpassungen: Eine Therapieänderung wurde lediglich bei 12,14 % der Patienten in der IG dokumentiert. Dies steht in erheblichem Widerspruch zur von Experten vorgeschlagenen Therapieänderungsrate von 32,38 % für das DigiSep-Patientenkollektiv (32,49 % für die Interventionsgruppe). Daraus lässt sich schließen, dass in zahlreichen Fällen eine empfohlene Therapieanpassung nicht umgesetzt wurde, obwohl Experten ein Potenzial zur Verbesserung der Patientenversorgung durch eine Modifikation der Therapie identifiziert hatten. Dies könnte dazu geführt haben, dass der potenziell erzielbare Behandlungseffekt nicht vollständig realisiert wurde.

Subgruppenanalysen zeigten erfreulicherweise, dass der Interventionseffekt relativ gleichmäßig in allen Subgruppen auftritt und somit nahezu alle Patienten von der Intervention profitierten.

Bei den sekundären Endpunkten, die teilweise Aspekte des DOOR/RADAR-Scores abdecken, konnte über alle Endpunkte hinweg ein positiver Effekt der Intervention erreicht werden. Dieser erzielte teilweise sogar statistische Signifikanz, obwohl die Studie für diese Endpunkte nicht gepowert war und statistisch signifikante Effekte eher in einem kombinierten Endpunkt (wie dem primären Endpunkt) zu erwarten waren. Hervorzuheben ist auch, dass trotz eines kaum vorhandenen Effekts zu Tag 28, in der Langzeitmortalität bis Tag 180 eine Verbesserung von 7,7% erzielt werden konnte (67,5% vs. 59,8%), was nochmals das Potential der Intervention unterstreicht, auch die Mortalität positiv zu beeinflussen, wenn auch nur im Langzeitverlauf.

Zusammenfassend ist die DigiSep Studie aufgrund des Nichteinreichens der statistischen Signifikanz im primären Endpunkt zwar als nicht erfolgreich zu bewerten, dennoch zeigt eine detaillierte Analyse, dass ein relevanter Interventionseffekt vorhanden zu sein scheint, der konsistent über alle Patienten und Endpunkte hinweg wirkt. Die Ergebnisse deuten darauf hin, dass der primäre Endpunkt, der DOOR/RADAR-Score, in seiner derzeitigen Definition für die vorliegende Intervention Optimierungsbedarf aufweist. Zudem legt die geringe Rate dokumentierter Therapieänderungen nahe, dass das volle Potenzial der Intervention noch nicht ausgeschöpft wurde. Eine gezielte Schulung des medizinischen Personals könnte möglicherweise dazu beitragen, die Wirksamkeit der Intervention zu steigern.

#### *Gesundheitsökonomische Evaluation (Universität Bielefeld)*

Neben den beschriebenen klinischen Wirksamkeitsanalysen der neuen Versorgungsform wurde eine gesundheitsökonomische Evaluation durchgeführt. Die Studienergebnisse der

gesundheitsökonomischen Evaluation zeigten insgesamt keine signifikanten Unterschiede der Gesamtkosten zwischen der IG und KG, auch nicht unter Berücksichtigung der Interventionskosten. Die weiterführende Regressionsanalyse bestätigte ebenfalls, dass die deskriptiven Unterschiede der Gesamtkosten nicht auf tatsächliche Gruppenunterschiede zurückzuführen waren. Statistisch signifikante Unterschiede ließen sich zudem nicht für die einzelnen Leistungsbereiche identifizieren, wenngleich deskriptive Analysen auf mittlere Kosteneinsparungen in der IG hinwiesen. Die Evaluation zeigte eine vergleichbare Leistungsinanspruchnahme der beiden Studiengruppen, was die mangelnde statistische Signifikanz der Kostenunterschiede zum Teil erklären konnte. Deskriptiv waren die Gruppenunterschiede in den Leistungsbereichen sowie bei den Gesamtkosten nur gering in Relation zum insgesamten Kostenniveau. So lagen beispielsweise die mittleren stationären Kosten pro Patient in der IG bei 74.995,07 € (SD: 74.565,08 €) und in der KG bei 74.011,30 € (SD: 77.049,74 €). Eine weitere Studie aus dem deutschen Versorgungskontext stellte mittlere stationäre Kosten von 60.455 € für den initialen Aufenthalt fest [56]. Wenngleich die hier vorliegenden stationären Kosten höher lagen, ist an dieser Stelle zu beachten, dass die entsprechende Studie lediglich Patienten mit katheter-assoziierten Blutstrominfektionen inkludierte und unter dem Kostenpunkt der stationären Kosten nicht nur der initiale Krankenhausaufenthalt, sondern ebenfalls Rehospitalisierungen berücksichtigt wurden. Dies war in zuvor benannter Studie nicht der Fall. Die mittleren Gesamtkosten pro Patient (exkl. Interventionskosten) in der IG lagen bei 96.719,22 € beziehungsweise bei 99.196,95 € in der KG. Im Vergleich dazu wurden die Gesamtkosten von Sepsispatienten im ersten Jahr nach Sepsis in der Studie von Fleischmann-Struzek et al. (2021) mit 18.051 € (95%-KI: 28.090 €) beziffert [57]. Wenngleich für diese Studie ebenfalls GKV-Routinedaten herangezogen wurden (AOK) und somit die Vergleichbarkeit der Datengrundlagen gegeben war, bezogen sich die Kosten lediglich auf Sepsisüberlebende. In der vorliegenden Digi-Sep Studie wurden dagegen die Kosten der verstorbenen Patienten inkludiert. Allgemeine Erkenntnisse zu hohen Kosten zum Lebensende, nach welchen ein erheblicher Anteil der gesamten Gesundheitsausgaben einer Person in den letzten Lebensmonaten anfällt, die in den vergangenen Jahrzehnten mehrfach festgestellt werden konnten [58, 59], könnten somit ein möglicher Erklärungsansatz für die hier nachgewiesenen höheren Kosten sein. Die genannten Studien zeigen übereinstimmend, dass ein großer Teil der Krankenhauskosten in den letzten sechs Monaten vor dem Tod anfällt, vor allem durch ICU-Aufenthalte, invasive Maßnahmen und lebensverlängernde Behandlungen.

Die deterministische CEA zeigte, dass die Intervention sowohl kostensenkend (978,52 € geringere Kosten) als auch wirksamer (0,166 weniger Todesfälle nach 28 Tagen) war. Allerdings wies die PSA auf eine hohe Unsicherheit hin. Die Streuung der Bootstrap-Replikationen war erheblich, und die Intervention war nur in 51,1 % der Replikationen sowohl kostensenkend als auch wirksam. Die Intervention schien in der Langzeitmortalität wirksam zu sein und könnte gleichzeitig zu Kosteneinsparungen führen. Die fehlenden signifikanten Unterschiede in den Kosten sowie die hohe Variabilität der Ergebnisse legen jedoch nahe, dass die Unterschiede der Gesamtkosten sowohl von der Höhe der Interventionskosten als auch der Anzahl der durchgeführten Interventionen abhängen. Angesichts dieser Unsicherheiten konnte auf Basis der vorliegenden Daten kein eindeutiger finanzieller Vorteil der Intervention auf die Gesamtkosten festgestellt werden. Dennoch legten die Ergebnisse zur klinischen Wirksamkeit, insbesondere hinsichtlich der Reduktion der Langzeitmortalität nach 180 Tagen,

nahe, dass die Intervention im klinischen Kontext vorteilhaft sein könnte. Um eine belastbare gesundheitsökonomische Bewertung vorzunehmen, sind jedoch weitere Studien erforderlich, die primär auf CEA ausgelegt sind und größere Stichproben umfassen.

#### *Evaluation der gesundheitsbezogenen Lebensqualität (Universität Bielefeld)*

In der IG zeigten sich in der Gesamtkohorte kurzfristige Verbesserungen der gesundheitsbezogenen Lebensqualität auf Grundlage des EQ-5D-5L Index, jedoch konnten keine langfristigen Effekte nachgewiesen werden. Die deskriptiven und inferenzstatistischen Analysen deuteten auf eine insgesamt ähnliche zeitliche Entwicklung sowohl in der IG als auch in der KG hin. Im Vergleich zur Allgemeinbevölkerung in Deutschland (Index: 0.88 (SD: 0.18); VAS: 71.59 (SD: 21.36)) [60] sowie im Vergleich allgemeinen älteren Population ab 65 Jahren (Index: 0.84 (SD: 0.22); VAS: 73.2 (SD: 18.5)) [61], lagen die hier identifizierten Indexwerte der Sepsispopulation deutlich darunter. In Anbetracht der schwerwiegenden kurz- und langfristigen gesundheitlichen Einschränkungen sowie der Mortalitätsrate war dieser Unterschied kaum verwunderlich. Darüber hinaus zeigten sich ebenfalls keine statistisch signifikanten Unterschiede der QALYs.

#### *Formative Evaluation (Universität Bielefeld)*

Die Ergebnisse der formativen Evaluation zeigten, dass die Präzisionsdiagnostik von vielen Teilnehmenden als sinnvolle und wertvolle Ergänzung zur Standarddiagnostik angesehen wird. Wenngleich der primären Endpunkt in den vorliegenden Analysen verfehlt wurde, schätzten die teilnehmenden Ärzt:innen die Intervention als geeignet, die Qualität der Therapie zu verbessern und langfristige Erfolge zu sichern. Dennoch blieb der hohe organisatorische und personelle Aufwand eine Herausforderung, die bei einer breiteren Implementierung berücksichtigt werden muss. Jedoch war nicht auszuschließen, dass die beschriebenen Herausforderungen sowie der organisatorische Mehraufwand dem Projektcharakter (zusätzlicher administrativer Aufwand für die Erhebung und Dokumentation primärer und sekundärer Endpunkte) zugeschrieben werden konnte, welcher in der Regelversorgung potenziell geringer ausfallen würde. Insgesamt verdeutlichten die formativen Studienergebnisse das Potenzial der Präzisionsdiagnostik, gleichzeitig aber auch die Notwendigkeit, ihre Anwendung durch optimierte Abläufe und Schulungen zu erleichtern.

#### *Limitationen (Institut für Medizinische Biometrie Heidelberg und Universität Bielefeld):*

Die Aussagekraft der Evaluation ist durch methodische Limitationen eingeschränkt. Die Fallzahlkalkulation wurde ausschließlich auf Basis des primären Endpunkts durchgeführt, was die Aussagekraft und Interpretierbarkeit der Analysen der sekundären Endpunkte und insbesondere der gesundheitsökonomischen Analysen limitiert.

Darüber hinaus standen Sozialversicherungsdaten nur für einen Teil der Gesamtkohorte zur Verfügung, wodurch die Stichprobengröße für die gesundheitsökonomischen Analysen weiter eingeschränkt war. Dies erhöhte die Anfälligkeit der Ergebnisse gegenüber Ausreißern erschwert robuste Schlussfolgerungen. Zudem sind Kostendaten per se mit einer hohen Variabilität behaftet, da sie stark von individuellen Faktoren wie Komorbiditäten, Behandlungsstrategien und regionalen Versorgungsstrukturen abhängen, die aufgrund ihrer mangelnden Verfügbarkeit nicht alle in die Regressionsanalyse mit aufgenommen werden konnten. Die hohe Variabilität der Kosten könnte darauf zurückzuführen sein, dass die

untersuchte Patientenpopulation per se eine erhebliche Heterogenität aufweist. Ebenso könnten unterschiedliche Behandlungsstrategien zwischen den beteiligten Zentren zu divergierenden Kostenstrukturen führen, da Variationen in der Anwendung der Intervention sowie in den begleitenden therapeutischen Maßnahmen nicht vollständig ausgeschlossen werden können.

Wenngleich für die Analyse der gesundheitsbezogenen Lebensqualität valide Instrumente zum Einsatz kamen und der Rücklauf insbesondere zu T0 als hoch eingeschätzt werden kann, kam es innerhalb der Fragebögen zu unvollständigen Angaben, was die entsprechende Berechnung der Scores oder von QALYs in diesen Fällen unmöglich machte. Diese Herausforderungen bestehen allgemein bei patientenberichteten Fragebögen und in anderen Studien [42, 62–64]. Eine methodisch verlässliche Alternative, um Unterschiede in QALYs zu untersuchen stellen LLM dar, welche alle verfügbaren Informationen nutzen, sofern Baseline-Werte vollständig sind und somit der Stichprobenreduktion entgegenkommen [42, 43]. Darüber hinaus wurde der EQ-5D-5L zu T0 in einigen Fällen durch Proxyberichte erhoben, bei denen Dritte (sog. Proxys wie Angehörige oder medizinisches Personal) anstelle der Patienten selbst Angaben gemacht haben, was die Vergleichbarkeit und Genauigkeit der Baseline-Daten einschränken könnte. Validierungsstudien haben jedoch gezeigt, dass Erhebungen der gesundheitsbezogenen Lebensqualität durch Proxys eine valide Alternative zum Selbstbericht sind und eine gute Annäherung der gesundheitsbezogenen Lebensqualität ermöglichen [65–68]. Schließlich war die Beobachtungszeit der Studie relativ kurz, wodurch langfristige Auswirkungen auf die gesundheitsbezogene Lebensqualität nicht erfasst werden konnten, insbesondere in Bezug auf mögliche späte Folgen oder nachhaltige Verbesserungen. Diese Limitationen sollten bei der Interpretation der Ergebnisse und der Planung zukünftiger Studien berücksichtigt werden.

## 5 Schlussfolgerungen und Empfehlungen des Evaluators

Die primäre Analyse der DigiSep-Studie lieferte kein statistisch signifikantes Ergebnis womit die Studie hinsichtlich der klinischen Wirksamkeit als nicht erfolgreich angesehen werden muss. Es zeigte sich jedoch über alle Patienten subgruppen ein gleichmäßiger positiver Effekt der Intervention. In den sekundären Endpunkten konnte über alle Endpunkte hinweg ein positiver Effekt der Intervention erreicht werden, der sogar teilweise statistische Signifikanz erzielte. Die Ergebnisse der Studie zeigen, dass die neue Versorgungsform sowohl in ihrer Wirksamkeit als auch in ihren gesundheitsökonomischen Auswirkungen Potenzial aufweist, jedoch keine eindeutigen und statistisch signifikanten Vorteile gegenüber der Standardversorgung nachgewiesen werden konnten.

Die deterministische CEA deutete auf mögliche Kosteneinsparungen hin, jedoch wiesen die Ergebnisse aufgrund hoher Unsicherheiten in der PSA keine belastbare Aussagekraft auf.

Hinsichtlich der gesundheitsbezogenen Lebensqualität gab es kurzfristig Verbesserungen in der IG, jedoch längerfristig keine signifikanten Unterschiede im Vergleich zur KG.

Die formative Evaluation hob hervor, dass die Präzisionsdiagnostik von den Teilnehmenden als wertvolle Ergänzung angesehen wird, jedoch organisatorische und personelle

Herausforderungen bei der Implementierung bestehen. Diese könnten im Regelbetrieb möglicherweise reduziert werden, bedürfen aber weiterer Optimierungen.

Die Aussagekraft der gesundheitsökonomischen Evaluation war durch methodische Limitationen wie eine geringe Stichprobengröße, heterogene Kostenverteilungen und unvollständige Angaben in den Lebensqualitätsfragebögen eingeschränkt. Diese Limitationen verdeutlichen den Bedarf an methodischen Anpassungen in zukünftigen Studien mit größeren Kohorten und längeren Beobachtungszeiträumen, um belastbarere Ergebnisse zu erzielen.

Zusammenfassend liefert die Studie Hinweise auf eine potenziell wirksame Intervention mit Kosteneinsparungen, jedoch ist die Evidenz aufgrund der aufgeführten Punkte und Limitationen begrenzt. Eine breite Implementierung der neuen Versorgungsform sollte daher erst nach weiteren Untersuchungen und Optimierungen erfolgen.

#### IV Literaturverzeichnis

1. Singer M, Deutschman CS, Seymour CW, et al (2016) The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 315:801. <https://doi.org/10.1001/jama.2016.0287>
2. Rhodes A, Evans LE, Alhazzani W, et al (2017) Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. *Intensive Care Med* 43:304–377. <https://doi.org/10.1007/s00134-017-4683-6>
3. Brunkhorst FM, Weigand MA, Pletz M, et al (2020) S3-Leitlinie Sepsis – Prävention, Diagnose, Therapie und Nachsorge: Zusammenfassung starker Empfehlungen. *Med Klin Intensivmed Notfmed* 115:178–188. <https://doi.org/10.1007/s00063-020-00671-6>
4. SepNet Critical Care Trials Group (2016) Incidence of severe sepsis and septic shock in German intensive care units: the prospective, multicentre INSEP study. *Intensive Care Med* 42:1980–1989. <https://doi.org/10.1007/s00134-016-4504-3>
5. Tiru B, DiNino EK, Orenstein A, et al (2015) The Economic and Humanistic Burden of Severe Sepsis. *PharmacoEconomics* 33:925–937. <https://doi.org/10.1007/s40273-015-0282-y>
6. Moerer O, Plock E, Mgbor U, et al (2007) A German national prevalence study on the cost of intensive care: an evaluation from 51 intensive care units. *Crit Care* 11:R69. <https://doi.org/10.1186/cc5952>
7. Fleischmann C, Thomas-Rueddel DO, Hartmann M, et al (2016) Hospital Incidence and Mortality Rates of Sepsis: An Analysis of Hospital Episode (DRG) Statistics in Germany From 2007 to 2013. *Deutsches Ärzteblatt international*. <https://doi.org/10.3238/ärztebl.2016.0159>
8. Moerer O, Burchardi H (2006) Kosten der Sepsis. *Anaesthesist* 55:36–42. <https://doi.org/10.1007/s00101-006-1039-y>

9. Shorr AF, Micek ST, Welch EC, et al (2011) Inappropriate antibiotic therapy in Gram-negative sepsis increases hospital length of stay\*: Critical Care Medicine 39:46–51. <https://doi.org/10.1097/CCM.0b013e3181fa41a7>
10. Richter DC, Frey O, Röhr A, et al (2019) Therapeutic drug monitoring-guided continuous infusion of piperacillin/tazobactam significantly improves pharmacokinetic target attainment in critically ill patients: a retrospective analysis of four years of clinical experience. Infection 47:1001–1011. <https://doi.org/10.1007/s15010-019-01352-z>
11. Grumaz S, Grumaz C, Vainshtein Y, et al (2019) Enhanced Performance of Next-Generation Sequencing Diagnostics Compared With Standard of Care Microbiological Diagnostics in Patients Suffering From Septic Shock. Critical Care Medicine 47:e394–e402. <https://doi.org/10.1097/CCM.0000000000003658>
12. Evans SR, Rubin D, Follmann D, et al (2015) Desirability of Outcome Ranking (DOOR) and Response Adjusted for Duration of Antibiotic Risk (RADAR). Clin Infect Dis 61:800–806. <https://doi.org/10.1093/cid/civ495>
13. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) Allgemeine Methoden - Version 7.0
14. Devlin N, Parkin D, Janssen B (2020) Methods for Analysing and Reporting EQ-5D Data. Springer International Publishing, Cham
15. Michalowsky B, Hoffmann W, Mohr W, et al (2022) Comparing the psychometric properties of EQ-5D-3L and EQ-5D-5L proxy ratings by informal caregivers and a health professional for people with dementia. Health Qual Life Outcomes 20:140. <https://doi.org/10.1186/s12955-022-02049-y>
16. Toh HJ, Yap P, Wee SL, et al (2021) Feasibility and validity of EQ-5D-5L proxy by nurses in measuring health-related quality of life of nursing home residents. Qual Life Res 30:713–720. <https://doi.org/10.1007/s11136-020-02673-5>
17. Feng Y-S, Kohlmann T, Janssen MF, Buchholz I (2021) Psychometric properties of the EQ-5D-5L: a systematic review of the literature. Qual Life Res 30:647–673. <https://doi.org/10.1007/s11136-020-02688-y>
18. Ludwig K, Graf Von Der Schulenburg J-M, Greiner W (2018) German Value Set for the EQ-5D-5L. PharmacoEconomics 36:663–674. <https://doi.org/10.1007/s40273-018-0615-8>
19. Kazis LE, Miller DR, Skinner KM, et al (2006) Applications of Methodologies of the Veterans Health Study in the VA Healthcare System: Conclusions and Summary. Journal of Ambulatory Care Management 29:182–188. <https://doi.org/10.1097/00004479-200604000-00011>
20. Kazis LE, Miller DR, Clark JA, et al (2004) Improving the Response Choices on the Veterans SF-36 Health Survey Role Functioning Scales: Results From the Veterans

- Health Study. *Journal of Ambulatory Care Management* 27:263–280. <https://doi.org/10.1097/00004479-200407000-00010>
21. Buchholz I, Feng Y-S, Buchholz M, et al (2021) Translation and adaptation of the German version of the Veterans Rand—36/12 Item Health Survey. *Health Qual Life Outcomes* 19:137. <https://doi.org/10.1186/s12955-021-01722-y>
  22. Kazis LE (2000) The Veterans SF-36 health status questionnaire: Development and application in the Veterans health administration. *Medical Outcomes Trust Monitor* 5:1–2
  23. Venkatesh V, Morris MG, Davis GB, Davis FD (2003) Unified Theory of Acceptance and Use of Technology
  24. Ong SWX, Petersiel N, Loewenthal MR, et al (2023) Unlocking the DOOR—how to design, apply, analyse, and interpret desirability of outcome ranking endpoints in infectious diseases clinical trials. *Clinical Microbiology and Infection* 29:1024–1030. <https://doi.org/10.1016/j.cmi.2023.05.003>
  25. Van Buuren S, Groothuis-Oudshoorn K (2006) mice: Multivariate Imputation by Chained Equations. 3.17.0
  26. Wickham H (2016) *ggplot2: elegant graphics for data analysis*, Second edition. Springer international publishing, Cham
  27. Kassambara A, Kosinski M, Biecek P (2016) survminer: Drawing Survival Curves using “ggplot2.” 0.5.0
  28. Therneau TM (2001) *survival: Survival Analysis*. 3.8-3
  29. Wickham H, Averick M, Bryan J, et al (2019) Welcome to the Tidyverse. *JOSS* 4:1686. <https://doi.org/10.21105/joss.01686>
  30. Van Den Brink WP, Van Den Brink SGJ (1989) A comparison of the power of the t test, Wilcoxon’s test, and the approximate permutation test for the two-sample location problem. *Brit J Math & Statis* 42:183–189. <https://doi.org/10.1111/j.2044-8317.1989.tb00907.x>
  31. Hedderich J, Sachs L (2020) *Angewandte Statistik: Methodensammlung mit R*, 17., überarbeitete und ergänzte Auflage. Springer Spektrum, Berlin
  32. Moran JL, Solomon PJ, Peisach AR, Martin J (2007) New models for old questions: generalized linear models for cost prediction. *Evaluation Clinical Practice* 13:381–389. <https://doi.org/10.1111/j.1365-2753.2006.00711.x>
  33. Marill KA (2004) Advanced Statistics: Linear Regression, Part II: Multiple Linear Regression. *Academic Emergency Medicine* 11:94–102. <https://doi.org/10.1197/j.aem.2003.09.006>

34. Stasinopoulos DM, Rigby RA (2007) Generalized Additive Models for Location Scale and Shape (GAMLSS) in R. *J Stat Soft* 23:.. <https://doi.org/10.18637/jss.v023.i07>
35. Bozdogan H (1987) Model Selection and Akaike's Information Criterion (AIC): The General Theory and its Analytical Extensions. *Psychometrika* 52:345–370. <https://doi.org/10.1007/BF02294361>
36. Fahrmeir L, Heumann C, Künstler R, et al (2023) Statistik: Der Weg zur Datenanalyse. Springer Berlin Heidelberg, Berlin, Heidelberg
37. Glick HA, Doshi JA, Sonnad SS, Polsky D (2014) Economic Evaluation in Clinical Trials. Oxford University Press
38. Schöffski O, Schumann A, Kuhlmann A, Schwarzbach C (2012) 6 Das Schwellenwertkonzept: Theorie sowie Umsetzung beim IQWiG und anderen Institutionen. In: Schöffski O, Graf Von Der Schulenburg J-M (eds) Gesundheitsökonomische Evaluationen. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 111–153
39. Greiner W, Damm O (2012) 3 Die Berechnung von Kosten und Nutzen. In: Schöffski O, Graf Von Der Schulenburg J-M (eds) Gesundheitsökonomische Evaluationen. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 23–42
40. Sassi F (2006) Calculating QALYs, comparing QALY and DALY calculations. *Health Policy and Planning* 21:402–408. <https://doi.org/10.1093/heapol/czl018>
41. Whitehead SJ, Shehzad A (2010) Health outcomes in economic evaluation: the QALY and utilities. *British Medical Bulletin* 96:5–21
42. Rösel I, Serna-Higuita LM, Al Sayah F, et al (2022) What difference does multiple imputation make in longitudinal modeling of EQ-5D-5L data? Empirical analyses of simulated and observed missing data patterns. *Qual Life Res* 31:1521–1532. <https://doi.org/10.1007/s11136-021-03037-3>
43. Ben AJ, Van Dongen JM, Alili ME, et al (2023) The handling of missing data in trial-based economic evaluations: should data be multiply imputed prior to longitudinal linear mixed-model analyses? *Eur J Health Econ* 24:951–965. <https://doi.org/10.1007/s10198-022-01525-y>
44. Firke S (2016) janitor: Simple Tools for Examining and Cleaning Dirty Data. 2.2.1
45. Wickham H, François R, Henry L, et al (2014) dplyr: A Grammar of Data Manipulation. 1.1.4
46. Morton F, Nijjar JS (2019) eq5d: Methods for Analysing “EQ-5D” Data and Calculating “EQ-5D” Index Scores. 0.15.6
47. Davison AC, Hinkley DV (1997) Bootstrap Methods and their Application, 1st ed. Cambridge University Press

48. Pinheiro J, Bates D, DebRoy S, et al nlme: Linear and Nonlinear Mixed Effects Models
49. Bengtsson H (2005) R.utils: Various Programming Utilities. 2.12.3
50. Wickham H, Bryan J (2015) readxl: Read Excel Files. 1.4.3
51. Graelmund G, Wickham H (2011) Dates and Times Made Easy with lubridate. *Journal of Statistical Software* 40:1–25
52. Rigby RA, Stasinopoulos DM (2005) Generalized Additive Models for Location, Scale and Shape. *Journal of the Royal Statistical Society Series C: Applied Statistics* 54:507–554. <https://doi.org/10.1111/j.1467-9876.2005.00510.x>
53. Otto J, Kahle D (2022) ggdensity: Interpretable Bivariate Density Visualization with “ggplot2.” 1.0.0
54. Becker J, Chan C, Schoch D, Leeper TJ (2013) rio: A Swiss-Army Knife for Data I/O. 1.2.3
55. Wickham H, Vaughan D, Girlich M (2014) tidyverse: Tidy Messy Data. 1.3.1
56. Leistner R, Hirsemann E, Bloch A, et al (2014) Costs and prolonged length of stay of central venous catheter-associated bloodstream infections (CVC BSI): a matched prospective cohort study. *Infection* 42:31–36. <https://doi.org/10.1007/s15010-013-0494-z>
57. Fleischmann-Struzek C, Rose N, Freytag A, et al (2021) Epidemiology and Costs of Postsepsis Morbidity, Nursing Care Dependency, and Mortality in Germany, 2013 to 2017. *JAMA Netw Open* 4:e2134290. <https://doi.org/10.1001/jamanetworkopen.2021.34290>
58. Riley GF, Lubitz JD (2010) Long-Term Trends in Medicare Payments in the Last Year of Life. *Health Services Research* 45:565–576. <https://doi.org/10.1111/j.1475-6773.2010.01082.x>
59. Lubitz JD, Riley GF (1993) Trends in Medicare Payments in the Last Year of Life. *N Engl J Med* 328:1092–1096
60. Grochtdreis T, Dams J, König H-H, Konnopka A (2019) Health-related quality of life measured with the EQ-5D-5L: estimation of normative index values based on a representative German population sample and value set. *Eur J Health Econ* 20:933–944. <https://doi.org/10.1007/s10198-019-01054-1>
61. Marten O, Greiner W (2021) EQ-5D-5L reference values for the German general elderly population. *Health Qual Life Outcomes* 19:76. <https://doi.org/10.1186/s12955-021-01719-7>
62. Rombach I, Gray AM, Jenkinson C, et al (2018) Multiple imputation for patient reported outcome measures in randomised controlled trials: advantages and disadvantages of

imputing at the item, subscale or composite score level. *BMC Med Res Methodol* 18:87. <https://doi.org/10.1186/s12874-018-0542-6>

63. Grady KL, Jones PG, Cristian-Andrei A, et al (2017) Causes and Consequences of Missing Health-Related Quality of Life Assessments in Patients Who Undergo Mechanical Circulatory Support Implantation: Insights From INTERMACS (Interagency Registry for Mechanically Assisted Circulatory Support). *Circ: Cardiovascular Quality and Outcomes* 10:. <https://doi.org/10.1161/CIRCOUTCOMES.116.003268>
64. Faria R, Gomes M, Epstein D, White IR (2014) A Guide to Handling Missing Data in Cost-Effectiveness Analysis Conducted Within Randomised Controlled Trials. *PharmacoEconomics* 32:1157–1170. <https://doi.org/10.1007/s40273-014-0193-3>
65. Hofhuis J, Hautvast JLA, Schrijvers AJP, Bakker J (2003) Quality of life on admission to the intensive care: can we query the relatives? *Intensive Care Med* 29:974–979. <https://doi.org/10.1007/s00134-003-1763-6>
66. Gifford JM, Husain N, Dinglas VD, et al (2010) Baseline quality of life before intensive care: A comparison of patient versus proxy responses: *Critical Care Medicine* 38:855–860. <https://doi.org/10.1097/CCM.0b013e3181cd10c7>
67. Golicki D, Niewada M, Buczak J, et al (2015) Validity of EQ-5D-5L in stroke. *Qual Life Res* 24:845–850. <https://doi.org/10.1007/s11136-014-0834-1>
68. Toh HJ, Yap P, Wee SL, et al (2021) Feasibility and validity of EQ-5D-5L proxy by nurses in measuring health-related quality of life of nursing home residents. *Qual Life Res* 30:713–720. <https://doi.org/10.1007/s11136-020-02673-5>

## V Anlagen

- Anlage 1: Standardisierter Fragebogen der formativen Evaluation
- Anlage 2: Zum Evaluationsbericht Detaillierte Projektergebnisse – Ergänzungen zu Kapitel 3
- Anlage 3: DigiSep - Statistischer Analyseplan
- Anlage 4: Weiterführende Projektergebnisse – Gesundheitsökonomische Evaluation und gesundheitsbezogene Lebensqualität

## Fragebogen

### 1 Demografische Angaben

Bitte beantworten Sie zunächst einige Angaben zu Ihrer Person.

Bitte geben Sie Ihr Alter in Jahren an:

Bitte geben Sie Ihr Geschlecht an:

- weiblich
- männlich
- divers
- keine Angabe

Ich bin...

- PJ-Student\*in
- Assistenzärzt\*in
- Fachärzt\*in
- Oberärzt\*in
- Chefärzt\*in
- Sonstiges

Wie viele Jahre Berufserfahrung haben Sie in der intensivmedizinischen Versorgung?

- 0 - 5 Jahre
- 6 - 10 Jahre
- 11 - 15 Jahre
- 16 - 20 Jahre
- > 21 Jahre

### 2 Erfahrungen

Haben Sie während der Studie die Ergebnisse der Präzisionsdiagnostik für initiale Therapieentscheidungen verwendet?

- Ja, für jede\*n Patient\*in der Interventionsgruppe
- Ja, aber nicht bei allen Patient\*innen
- Nein

### 3.1 Standardseite

Bitte schätzen Sie für wie viel Prozent der Patient\*innen Sie die Präzisionsdiagnostik für initiale Therapieentscheidungen verwenden.

### 4 Standardseite

Würden Sie die Präzisionsdiagnostik auch in Zukunft (d.h. außerhalb des Studiensesttings) verwenden?

- Ja, für jede\*n Patient\*in mit Sepsis/septischem Schock
- Ja, aber nur unter bestimmten Bedingungen bzw. bei bestimmten Rahmenbedingungen
- Nein

## 5.1 Filter

**Unter welchen Bedingungen bzw. Rahmenbedingungen würden Sie die Präzisionsdiagnostik für jede\*n Patient\*in mit Sepsis/septischem Schock verwenden?**

## 6 Standardseite

**Sehen Sie einen Nutzen der Präzisionsdiagnostik in anderen Settings (z.B. andere Stationen, ambulanter Bereich)?**

- Ja
- Ja, aber nur unter bestimmten Bedingungen bzw. bei bestimmten Rahmenbedingungen
- Nein

## 7.1 Filter

**Unter welchen Bedingungen bzw. Rahmenbedingungen sehen Sie einen Nutzen der Präzisionsdiagnostik in anderen Settings?**

## 8 Sozialer Einfluss

**Im Nachfolgenden soll es um Ihre Einschätzung der Perspektive Ihres sozialen Umfelds gehen.**

**Inwiefern stimmen Sie den folgenden Aussagen zu?**

	Trifft voll zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu	Keine Angabe
Menschen, die Einfluss auf mein Verhalten haben, denken, ich sollte die Präzisionsdiagnostik weiterhin nutzen.	<input type="radio"/>				
Meine Kolleg*innen denken, dass ich die Präzisionsdiagnostik weiterhin nutzen sollte.	<input type="radio"/>				
	<input type="radio"/>				

## 9 Leistungserwartung

**An dieser Stelle sind wir an Informationen zu Ihren Erfahrungen mit der Präzisionsdiagnostik interessiert.**

**Inwiefern stimmen Sie den folgenden Aussagen zu?**

Die Präzisionsdiagnostik...

	Trifft voll zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu	Keine Angabe
...erzielt vergleichbare Ergebnisse im Vergleich zur Standarddiagnostik.	<input type="radio"/>				
...trägt zur langfristigen Sicherung von Therapieerfolgen bei.	<input type="radio"/>				
...verbessert meiner Ansicht nach die medizinische Versorgung.	<input type="radio"/>				
...lässt sich gut in meinen Alltag integrieren.	<input type="radio"/>				
...stellt eine Unterstützung für die Therapiesteuerung dar.	<input type="radio"/>				

...spart Zeit.  
Die Präzisionsdiagnostik...

	Trifft voll zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu	Keine Angabe
...hat meine Erwartungen erfüllt.	<input type="radio"/>				
...erhöht den Verwaltungsaufwand/ Organisationsaufwand.	<input type="radio"/>				
...lässt sich schwer in meinen Alltag als Ärzt*in integrieren.	<input type="radio"/>				
...geht mit einem hohen Schulungsaufwand einher.	<input type="radio"/>				
...ist mit einem hohen Ressourcenaufwand verbunden.	<input type="radio"/>				

#### Inwiefern stimmen Sie den Aussagen hinsichtlich der Rahmenbedingungen der Präzisionsdiagnostik zu?

	Trifft voll zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu	Keine Angabe
Ich habe alle notwendigen Informationen erhalten, um die Präzisionsdiagnostik anzuwenden.	<input type="radio"/>				
Es ist häufig zu Problemen beim organisatorischen Ablauf gekommen.	<input type="radio"/>				
Der Zugriff auf die Ergebnisse war einfach.	<input type="radio"/>				
Es ist häufig zu technischen Fehlern beim Abruf der Ergebnisse gekommen.	<input type="radio"/>				
Bei Schwierigkeiten mit dem Präzisionsdiagnostik konnte ich schnell Hilfe erhalten.	<input type="radio"/>				

## 10 Expertenkonsil

Zuletzt möchten wir durch die nachfolgenden Fragen mehr über Ihre Erfahrung(en) mit dem fakultativen Expert\*innenkonsil erfahren.

**Haben Sie das Angebot des fakultativen Expert\*innenkonsils während der Studie in Anspruch genommen?**

- Ja  
 Nein

### 11.1 Filter

**Für wie viele Patient\*innen haben Sie im Rahmen der Studie das Expert\*innenkonsil in Anspruch genommen?**

Angabe:

Nr. Studienpatient\*innen mit Expertenkonsil

Nr. alle Studienpatient\*innen im Studienzentrum

#### Das Expert\*innenkonsil...

	Trifft voll zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu	Keine Angabe
(28) ...verbessert meiner Ansicht nach die medizinische Versorgung.	<input type="radio"/>				
(29) ...lässt sich gut in meinen Alltag integrieren.	<input type="radio"/>				
(30) ...stellt eine Unterstützung für die Therapiesteuerung dar.	<input type="radio"/>				
(31) ...hat meine Erwartungen erfüllt.	<input type="radio"/>				
(32) ... ist mit einem hohen Ressourcenaufwand verbunden.	<input type="radio"/>				

## 12.1 Filter

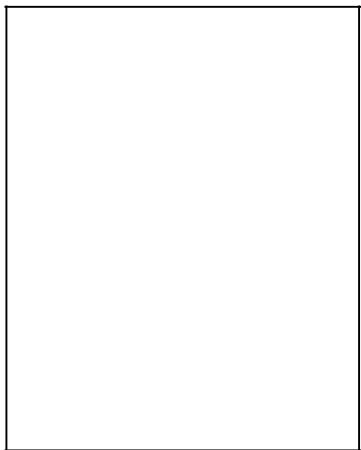
**Wieso haben Sie das Expertenkonsil nicht in Anspruch genommen?**

- Dafür gab es keinen Bedarf.

- Das Angebot war mir nicht bekannt.
- Dafür hatte ich keine Zeit.
- Sonstiges

## 13 Freitextfeld

**Ihr Feedback, Ihre Kommentare oder sonstige Ergänzungen im Umgang mit der Präzisionsdiagnostik sind an dieser Stelle herzlich willkommen:**

A large, empty rectangular box with a thin black border, intended for the respondent to type their feedback or comments.

## 14 Endseite

**Herzlichen Dank für Ihre Teilnahme!**

Verantwortlich für diese Umfrage:

Juliana Schmidt, M. Sc.

Tel.: +49 521 / 106 4350

E-Mail: [juliana.schmidt@uni-bielefeld.de](mailto:juliana.schmidt@uni-bielefeld.de)

Maren Steinmann, M. Sc.

Tel.: +49 521/1064679

E-Mail: [maren.steinmann@uni-bielefeld.de](mailto:maren.steinmann@uni-bielefeld.de)

Prof. Dr. Wolfgang Greiner

Tel.: +49 521 / 106 6989

E-Mail: [wolfgang.greiner@uni-bielefeld.de](mailto:wolfgang.greiner@uni-bielefeld.de)



**Anlage 2 zum Evaluationsbericht Detaillierte Projektergebnisse – Ergänzungen zu Kapitel 3**

**Inhaltsverzeichnis**

Vertiefende Ergebnisse zu Kapitel 3 Evaluation des primären Endpunktes .....	2
Weiterführende Ergebnisse zu Kapitel 3 Evaluation der sekundären Endpunkte .....	3
Weiterführende Ergebnisse zu Kapitel 3 Retrospektive Evaluation des Expertenpanels .....	5
Weiterführende Ergebnisse zu Kapitel 3 Gesundheitsökonomische Evaluation .....	8
Ergänzungen zu Kapitel 3 Gesundheitsbezogene Lebensqualität.....	15

## Vertiefende Ergebnisse zu Kapitel 3 Evaluation des primären Endpunktes

Tabelle 1: Binäres logistisches Regressionsmodell Primärer Endpunkt dichotom

Variable	Odds-Ratio (KI)	p-Wert
Gruppe (Ref: IG)	1,321 [0,80-2,18]	0,276
Alter	0,993 [0,98-1,01]	0,422
Geschlecht (Ref: m)	1,593 [0,94-2,71]	0,085
SOFA-Score	1,270 [1,14-1,42]	<0,001
Charlson Comorbidity Index	1,128 [0,99-1,29]	0,071

Tabelle 2: Subgruppenanalyse primärer Endpunkt

Subgruppe	Mittelwert DOOR/RADAR Interventionsgruppe	Mittelwert DOOR/RADAR Kontrollgruppe	P Wert (MWU, ggf. imputiert)
Patienten ohne zeitverzögerte Abnahme	3,24±1,53	3,51±1,51	0,146
Patienten mit zeitverzögerter Abnahme	2,73±1,69	3,07±1,43	0,753
Septischer Schock: V1 - Nein und V2 - Nein	2,83±1,50	3,15±1,64	0,564
Septischer Schock: V1 - Ja und V2 - Nein	2,47±1,35	2,86±1,49	0,158
Septischer Schock: V1 - Nein und V2 - Ja	3,43±1,83	3,05±1,68	0,448
Septischer Schock: V1 - Ja und V2 - Ja	3,70±1,36	3,89±1,28	0,415
Fokus			
Pneumonie, sonstige obere oder untere Atemwege, Thorakal	3,61±1,46	3,67±1,45	0,859
Gastrointestinal, Gallenwege, Intraabdominell / Peritonitis	3,28±1,55	3,48±1,49	0,777
Bakterämie, Katheterinfektion	3,79±1,35	2,82±1,69	0,213
Knochen / Weichteile, Chirurg. Wundinfektion	3,51±1,35	3,72±1,46	0,438
Urogenital	2,5±1,53	2,96±1,67	0,273
Infektion			
Grad der Sicherung: Klinisch vermutet	3,12±1,56	3,48±1,52	0,098
Grad der Sicherung: Mikrobiologisch nachgewiesen	3,83±1,23	3,53±1,47	0,701
Ursprung der Infektion: Ambulant	3,22±1,59	3,49±1,52	0,343

Ursprung der Infektion: Nosokomial (>48h nach KH-Aufnahme)	3,21±1,48	3,49±1,49	0,328
Immunsuppression liegt vor	3,61±1,54	4,25±1,09	0,380
Immunsuppression liegt nicht vor	3,18±1,54	3,47±1,52	0,151
Charlson Comorbidity Index > 4	3,46±1,56	3,97±1,29	0,277
Charlson Comorbidity Index =< 4	3,17±1,53	3,40±1,53	0,260
Modifizierter SOFA kleiner/gleich als Median des SOFA	2,82±1,43	3,16±1,54	0,095
Modifizierter SOFA größer als Median des SOFA	3,79±1,48	3,98±1,38	0,878
Blutkultur positiv	3,32±1,49	3,51±1,51	0,269
Blutkultur negativ	2,65±1,68	3,35±1,51	0,207
Fokussanierung: Ja	2,93±1,49	3,43±1,53	0,030
Fokussanierung: Nein	3,51±1,54	3,54±1,49	0,812
Fokussanierung: Ja - chirurgische Fokussanierung	3,03±1,47	3,37±1,54	0,176
Fokussanierung: Ja - Entfernen intravasaler Katheter	2,73±1,62	4,19±1,10	0,028
Fokussanierung: Ja - interventionelle Fokussanierung	2,66±1,46	3,35±1,59	0,240
Patienten mit Therapieänderung in IG	3,74±1,26	3,49±1,51	0,741
Überlebende	2,62±1,32	2,95±1,40	0,059

### Weiterführende Ergebnisse zu Kapitel 3 Evaluation der sekundären Endpunkte

*Tabelle 3: Geschätzte Überlebensraten (mit KI) zu verschiedenen Zeitpunkten*

Überleben zu	Interventionsgruppe	Kontrollgruppe
Tag 28	0,747 [0,689;0,810]	0,739 [0,678;0,804]
Tag 90	0,701 [0,640;0,768]	0,652 [0,588;0,724]
Tag 180	0,675 [0,613;0,744]	0,598 [0,532;0,673]

Tabelle 4: Sekundäre Endpunkte zu Tag 28. \*t-Test

	Interventionsgruppe (N=200)	Kontroll- gruppe (N=189)	Effekt (95% KI)	P Wert
<b>Sekundäre Endpunkte zu Tag 28</b>				
Beatmungsdauer				
Relativ	0,35±0,41	0,44±0,42	-0,09 [-0,17; -0,01]	0,034*
Absolut bzgl. Überlebender	6,6±9,4	9,3±10,6	-2,7 [-5,03;-0,34]	0,025*
Zeit bis zur Schockresolution				
Relativ	0,40±0,38	0,45±0,38	-0,05 [-0,13; 0,03]	0,208*
Absolut bzgl. Überlebender	6,9±7,4	8,8±8,5	-1,9 [-3,75; -0,04]	0,046*
Dauer der Nierenersatztherapie				
Relativ	0,15±0,32	0,18±0,34	-0,03 [-0,10; 0,03]	0,299*
Absolut bzgl. Überlebender	2,5±7,0	4,2±8,4	-1,7 [-3,52; 0,11]	0,061*
Krankenhausverweildauer				
Relativ	0,79±0,29	0,84±0,26	-0,05 [-0,11; 0,00]	0,064*
Absolut bzgl. Überlebender	20,3±8,6	22,1±8,0	-1,8[-3,69; 0,19]	0,077*
Intensivverweildauer				
Relativ	0,60±0,37	0,66±0,36	-0,06 [-0,13; 0,02]	0,130*
Absolut bzgl. Überlebender	13,8±9,8	15,9±10,0	-2,1 [-4,42; 0,10]	0,072*
Dauer der antiinfektiven Therapie				
Relativ	0,62±0,31	0,64±0,32	-0,02[-0,08; 0,05]	0,582*
Absolut bzgl. Überlebender	14,8±8,1	15,5±8,2	-0,7 [-2,60; 1,20]	0,467*

Tabelle 5: Sekundäre Endpunkte zu Tag 180. \*t-Test

	Interventionsgruppe (N=200)	Kontroll- gruppe (N=189)	Effekt (95% KI)	P Wert
<b>Sekundäre Endpunkte zu Tag 180</b>				
Beatmungsdauer				
Relativ	0,23±0,37	0,30±0,40	-0,07 [-0,15; -0,00]	0,053*
Absolut bzgl. Überlebender	9,2±17,1	14,1±27,2	-4,9 [-10,8; 0,93]	0,098*
Zeit bis zur Schockresolution				
Relativ	0,26±0,39	0,31±0,39	-0,05 [-0,13; 0,03]	0,221*
Absolut bzgl. Überlebender	7,9±11,4	10,7±16,6	-2,8 [-6,4; 0,95]	0,145*
Dauer der Nierenersatztherapie				
Relativ	0,11±0,27	0,11±0,27	0,00 [-0,06; 0,05]	0,838*
Absolut bzgl. Überlebender	3,6±13,3	3,9±11,9	-0,3 [-3,5; 2,8]	0,831*
Krankenhausverweildauer				
Relativ	0,45±0,26	0,50±0,40	-0,05 [-0,13; 0,03]	0,198*
Absolut bzgl. Überlebender	42,6±45,4	44,1±41,4	-1,5 [-12,5; 9,5]	0,785*
Intensivverweildauer				
Relativ	0,33±0,39	0,40±0,41	-0,07 [-0,15; 0,02]	0,123*
Absolut bzgl. Überlebender	19,4±23,5	23,0±26,7	-3,6 [-10,1; 2,8]	0,264*
Dauer der antiinfektiven Therapie				
Relativ	0,32±0,37	0,38±0,39	-0,06 [-0,14; 0,02]	0,129*
Absolut bzgl. Überlebender	19,2±21,2	22,8±25,5	-3,6 [-9,6; 2,4]	0,236*

### Weiterführende Ergebnisse zu Kapitel 3 Retrospektive Evaluation des Expertenpanels

Tabelle 6: Auflistung der Art der Mehrheitsfindung der Experten

Variable	IG (N=197)	KG (N=189)	Total (N=386)
<b>V1_Blutkultur_therapierelevant</b>			
Vollkommen	157 (79,70%)	163 (86,24%)	320 (82,90%)
Mehrheit	35 (17,77%)	20 (10,58%)	55 (14,25%)

Variable	IG (N=197)	KG (N=189)	Total (N=386)
Keine_Mehrheit	5 (2,54%)	6 (3,17%)	11 (2,85%)
<b>V1_NGS_plausibel</b>			
Mehrheit	84 (42,64%)	83 (43,92%)	167 (43,26%)
Vollkommen	113 (57,36%)	104 (55,03%)	217 (56,22%)
Keine_Mehrheit	0 (0,00%)	2 (1,06%)	2 (0,52%)
<b>V1_NGS_therapierelevant</b>			
Mehrheit	37 (18,78%)	24 (12,70%)	61 (15,80%)
Vollkommen	93 (47,21%)	92 (48,68%)	185 (47,93%)
Keine_Mehrheit	67 (34,01%)	73 (38,62%)	140 (36,27%)
<b>V2_Blutkultur_therapierelevant</b>			
Vollkommen	98 (49,75%)	103 (54,50%)	201 (52,07%)
Mehrheit	57 (28,93%)	52 (27,51%)	109 (28,24%)
Keine_Mehrheit	42 (21,32%)	34 (17,99%)	76 (19,69%)
<b>V2_NGS_plausibel</b>			
Vollkommen	116 (58,88%)	110 (58,20%)	226 (58,55%)
Mehrheit	48 (24,37%)	52 (27,51%)	100 (25,91%)
Keine_Mehrheit	33 (16,75%)	27 (14,29%)	60 (15,54%)
<b>V2_NGS_therapierelevant</b>			
Mehrheit	63 (31,98%)	64 (33,86%)	127 (32,90%)
Keine_Mehrheit	66 (33,50%)	66 (34,92%)	132 (34,20%)
Vollkommen	68 (34,52%)	59 (31,22%)	127 (32,90%)
<b>VbeliebigerZeitpunkt_Blutkultur_therapierelevant</b>			
Keine_Mehrheit	192 (97,46%)	188 (99,47%)	380 (98,45%)
Mehrheit	2 (1,02%)	0 (0,00%)	2 (0,52%)
Vollkommen	3 (1,52%)	1 (0,53%)	4 (1,04%)
<b>VbeliebigerZeitpunkt_NGS_plausibel</b>			
Keine_Mehrheit	186 (94,42%)	185 (97,88%)	371 (96,11%)
Mehrheit	6 (3,05%)	2 (1,06%)	8 (2,07%)
Vollkommen	5 (2,54%)	2 (1,06%)	7 (1,81%)

Variable	IG (N=197)	KG (N=189)	Total (N=386)
<b>VbeliebigerZeitpunkt _NGS_therapierelevant</b>			
Keine_Mehrheit			
Keine_Mehrheit	187 (94,92%)	186 (98,41%)	373 (96,63%)
Mehrheit	5 (2,54%)	1 (0,53%)	6 (1,55%)
Vollkommen	5 (2,54%)	2 (1,06%)	7 (1,81%)
<b>Therapieänderung empfohlen</b>			
Mehrheit			
Mehrheit	102 (51,78%)	87 (46,03%)	189 (48,96%)
Vollkommen	95 (48,22%)	102 (53,97%)	197 (51,04%)

## Weiterführende Ergebnisse zu Kapitel 3 Gesundheitsökonomische Evaluation

Tabelle 7: Krankheitskosten in € (exkl. Interventionskosten)

	Interventionsgruppe	Kontrollgruppe	p-Wert
<b>Ambulante Versorgung</b>			
n	74	54	
MW (SD)	395,65 (641,79)	336,57 (452,52)	
MD (IQR)	162,76 (459,26)	129,48 (503,15)	0,759
<b>Stationäre Versorgung</b>			
n	75	53	
MW (SD)	74.995,07 (74.565,08)	74.011,30 (77.049,74)	
MD (IQR)	47.116,45 (80.333,74)	39.509,87 (81.550,22)	0,786
<b>Arzneimittel</b>			
n	74	54	
MW (SD)	1.850,17 (5.900,88)	1.995,00 (5.160,45)	
MD (IQR)	314,16 (1.347,64)	453,005 (1.965,53)	0,741
<b>Heilmittel</b>			
n	75	53	
MW (SD)	131,29 (338,78)	138,21 (370,42)	
MD (IQR)	0 (0)	0 (0)	0,905
<b>Hilfsmittel</b>			
n	74	54	
MW (SD)	633,92 (1.436,23)	488,63 (976,53)	
MD (IQR)	0 (571,52)	0 (489,12)	0,709
<b>Rehabilitation</b>			
n	74	54	
MW (SD)	2.069,26 (7.244,98)	124,74 (839,05)	
MD (IQR)	0 (0)	0 (0)	0,053
<b>Pflege</b>			
n	75	53	
MW (SD)	2.808,44 (4.278,18)	1.412,58 (2.660,78)	
MD (IQR)	106,8 (4.309,60)	0 (1.943,83)	0,089
<b>Häusliche Krankenpflege</b>			

n	75	53	
MW (SD)	175,65 (681,68)	447,80 (2.024,50)	
MD (IQR)	0 (0)	0 (0)	0,219
<b>Krankengeld</b>			
n	75	53	
MW (SD)	1.560,38 (3.922,92)	1.218,35 (3.838,59)	
MD (IQR)	0 (0)	0 (0)	0,483
<b>Krankentransporte</b>			
n	74	54	
MW (SD)	3.930,26 (18.083,60)	10.419,00 (60.111,05)	
MD (IQR)	450,32 (1.399,84)	14,72 (834,00)	0,058
<b>Gesamt</b>			
n	74	54	
MW (SD)	96.719,22 (104.559,40)	99.196,95 (111.391,80)	
MD (IQR)	59.503,06 (95.109,13)	57.865,41 (109.062,2)	0,883

IQR: Interquartilsabstand, MD: Median, MW: Mittelwert, n: Stichprobengröße, SD: Standardabweichung

Test auf Unterschiede aufgrund fehlender Normalverteilung: Wilcoxon-Rangsummentest für unverbundene Stichproben

Signifikanz: \*\*\* p ≤ 0,001, \*\* p ≤ 0,01, \* p ≤ 0,05

Tabelle 8: Krankheitskosten in € (inkl. Interventionskosten)

	Interventionsgruppe (n = 74)	Kontrollgruppe (n = 54)	p-Wert
Kosten exkl. Interventionskosten			
MW (SD)	96.719,22 (104.559,40)	99.196,95 (111.391,80)	
MD (IQR)	59.503,06 (95.109,13)	57.865,41 (109.062,2)	0,883
Szenario I: Kosten inkl. Interventionskosten (Projektkosten)			
IG: 1.499,21 € p.P. vs. KG: 723,4 € p.P.			
MW (SD)	98.717,64 (104.559,40)	99.920,35 (111.391,81)	
MD (IQR)	62.502,27 (95.109,13)	58.588,81 (109.062,21)	0,830
Szenario II: Kosten inkl. Interventionskosten (Mögliche Regelversorgung)			
IG: 1.499,21 € p.P. vs. KG: 0 € p.P.			
MW (SD)	98.218,43 (104.559,45)	99.196,95 (111.391,81)	
MD (IQR)	61.003,06 (95.109,13)	57.865,41 (109.062,21)	0,778

IQR: Interquartilsabstand, MD: Median, MW: Mittelwert, n: Stichprobengröße, p.P.: pro Probe, SD: Standardabweichung

Test auf Unterschiede aufgrund fehlender Normalverteilung: Wilcoxon-Rangsummentest für unverbundene Stichproben

Signifikanz: \*\*\* p ≤ 0,001, \*\* p ≤ 0,01, \* p ≤ 0,05

Tabelle 9: Leistungsinanspruchnahme über den patientenindividuellen Beobachtungszeitraum – 180 Tage

	Interventionsgruppe	Kontrollgruppe	p-Wert
<b>Ambulante Versorgung - Arztkontakte</b>			
n	74	54	
MW (SD)	10,03 (11,35)	9,07 (10,39)	
MD (IQR)	7 (13)	5 (16)	0,682
<b>Stationäre Versorgung - Rehospitalisierung</b>			
n	74	54	
MW (SD)	0,51 (0,94)	0,20 (0,56)	
MD (IQR)	0 (1)	0 (0)	0,021*
<b>Stationäre Versorgung – Verweildauer initialer Aufenthalt (Tage)</b>			
n	74	54	
MW (SD)	30,95 (26,31)	32,22 (31,39)	
MD (IQR)	26,5 (36,25)	25,5 (34,75)	0,792
<b>Stationäre Versorgung – Verweildauer (Tage)</b>			
n	74	54	
MW (SD)	35,28 (28,96)	35,09 (34,21)	
MD (IQR)	27,5 (38,65)	26 (40,25)	0,562
<b>Arzneimittel – Verordnungen (PZN)</b>			
n	74	54	
MW (SD)	17,554 (21,774)	16,352 (20,753)	
MD (IQR)	12 (24,75)	9,5 (24,25)	0,813
<b>Heilmittel – Verordnungen (Positionsnummer)</b>			
n	75	53	
MW (SD)	11,04 (29,05)	0 (2)	
MD (IQR)	6,32 (15,11)	0 (0)	0,515
<b>Hilfsmittel – Verordnungen (Positionsnummer)</b>			
n	74	54	
MW (SD)	28,24 (81,46)	33,41 (107,81)	
MD (IQR)	1,5 (9)	3,5 (10)	0,584
<b>Rehabilitation</b>			
n	74	54	

MW (SD)	0,14 (0,34)	0,04 (0,19)	
MD (IQR)	0 (0)	0 (0)	0,062
<b>Pflege</b>			
n	74	54	
MW (SD)	7,05 (9,33)	6,83 (17,96)	
MD (IQR)	2 (12,75)	0 (6)	0,232
<b>Häusliche Krankenpflege</b>			
n	74	54	
MW (SD)	10,18 (48,03)	9,65 (49,17)	
MD (IQR)	0 (0)	0 (0)	0,070
<b>Arbeitsunfähigkeit</b>			
n	75	54	
MW (SD)	0,34 (0,60)	0,31 (0,61)	
MD (IQR)	0 (1)	0 (0,75)	0,682
<b>Arbeitsunfähigkeit mit Krankengeld</b>			
n	75	54	
MW (SD)	0,20 (0,40)	0,17 (0,38)	
MD (IQR)	0 (0)	0 (0)	0,635
<b>Krankentransporte</b>			
n	74	54	
MW (SD)	38,84 (88,08)	34,78 (107,58)	
MD (IQR)	2,5 (20,5)	0,5 (12)	0,067

IQR: Interquartilsabstand, MD: Median, MW: Mittelwert, n: Stichprobengröße, SD: Standardabweichung, PZN: Pharmazentralnummer

Test auf Unterschiede aufgrund fehlender Normalverteilung: Wilcoxon-Rangsummentest für unverbundene Stichproben

Signifikanz: \*\*\* p ≤ 0,001, \*\* p ≤ 0,01, \* p ≤ 0,05

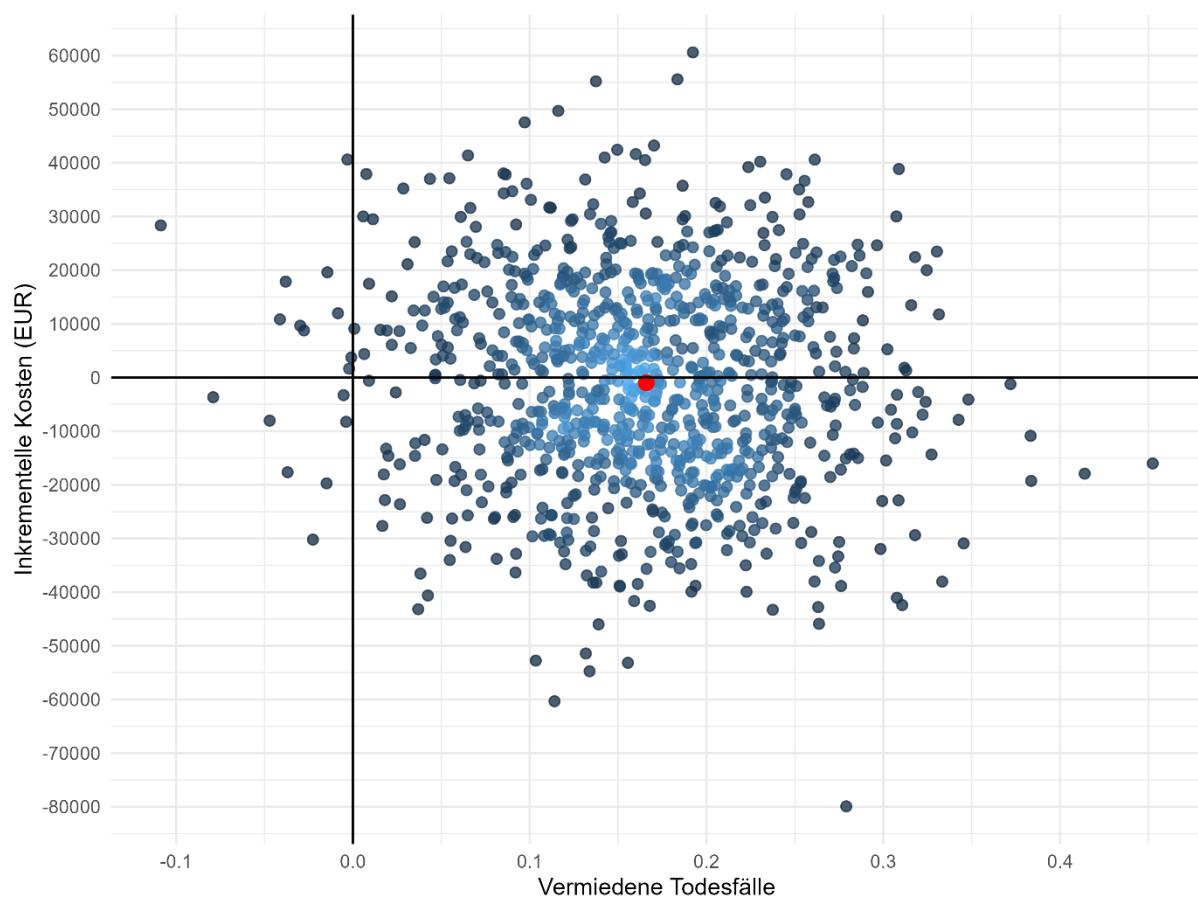


Abbildung 1: Kosten-Effektivitäts-Diagramm

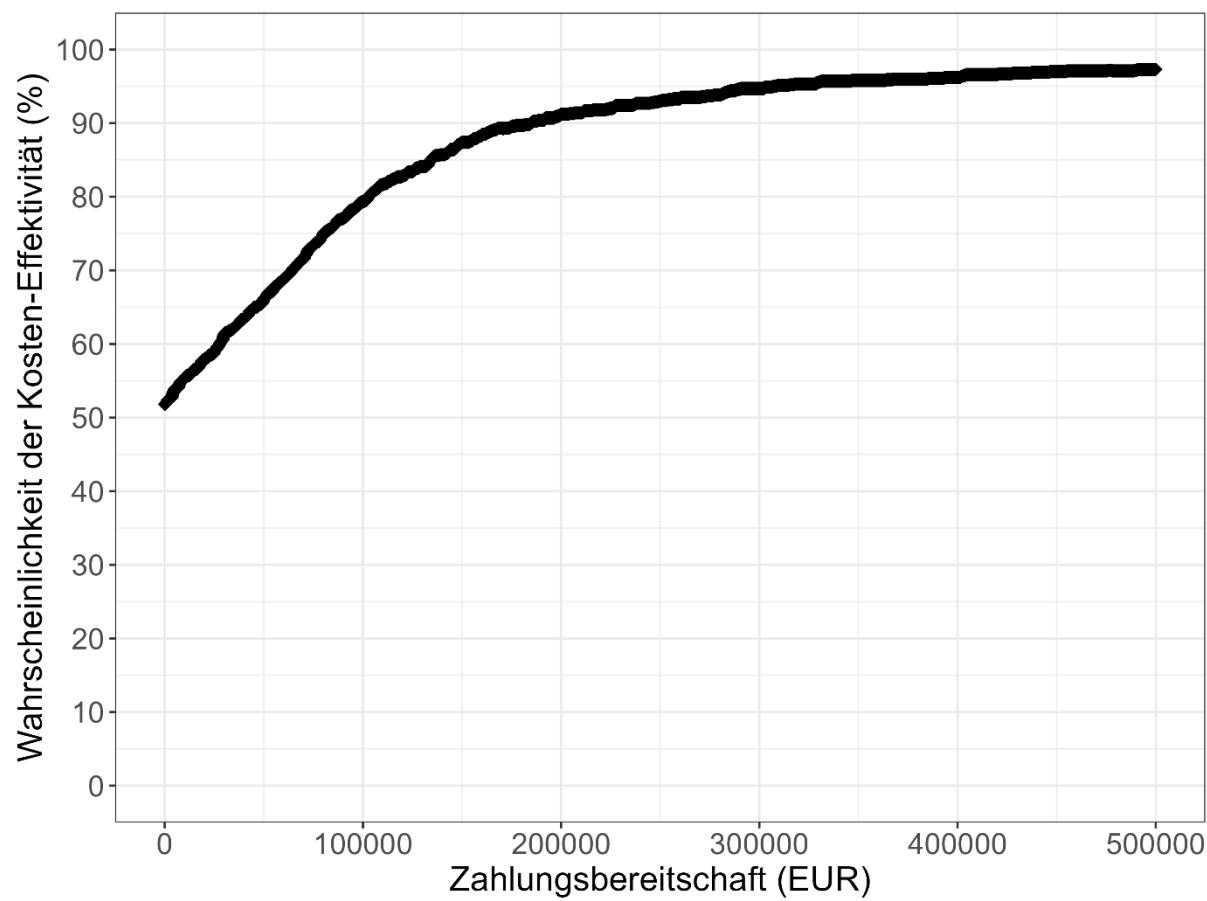


Abbildung 2: Kosten-Effektivitäts-Akzeptanzkurve

## Ergänzungen zu Kapitel 3 Gesundheitsbezogene Lebensqualität

Tabelle 10: Gesundheitsbezogene Lebensqualität - EQ-5D-5L Index

	Gesamtkohorte		GKV-Kohorte					
	Interventionsgruppe	Kontrollgruppe	Interventionsgruppe	Kontrollgruppe				
EQ-5D-5L Index								
T0 - n	180	170	68	51				
MW (SD)	0,294 (0,374)	0,247 (0,394)	0,326 (0,324)	0,210 (0,402)				
Median (IQR)	0,307 (0,374)	0,307 (0,377)	0,307 (0,266)	0,307 (0,275)				
p-Wert	0,191		0,169					
T1 - n	128	127	48	32				
MW (SD)	0,312 (0,386)	0,208 (0,373)	0,266 (0,362)	0,255 (0,401)				
Median (IQR)	0 (0,752)	0 (0,473)	0,058 (0,544)	0 (0,611)				
p-Wert	0,047*		0,725					
T2 - n	145	139	59	37				
MW (SD)	0,220 (0,370)	0,236 (0,386)	0,166 (0,322)	0,247 (0,418)				
Median (IQR)	0 (0,475)	0 (0,603)	0 (0,213)	0 (0,770)				
p-Wert	0,926		0,509					
IQR: Interquartilsabstand, MW: Mittelwert, n: Stichprobengröße, SD: Standardabweichung								
Test auf Unterschiede aufgrund fehlender Normalverteilung: Wilcoxon-Rangsummentest für unverbundene Stichproben								
Signifikanz: *** p ≤ 0,001, ** p ≤ 0,01, * p ≤ 0,05								

Tabelle 11: Quality-Adjusted Life Years

	Gesamtkohorte		GKV-Kohorte					
	Interventionsgruppe	Kontrollgruppe	Interventionsgruppe	Kontrollgruppe				
EQ-5D-5L QALY (LLM)								
n	180	170	68	51				
QALY-Differenz (95 %-KI)	0,023 (-0,081; 0,127)		0,056 (-0,200; 0,223)					
KI: Konfidenzintervall, LLM: Longitudinal Linear Mixed Model, n: Stichprobengröße, QALY: Quality-Adjusted Life Years								
Statistische Signifikanz anhand des 95%-Konfidenzintervalls festgestellt								

## **Statistischer Analyse Plan für abschließende Auswertung**

**DigiSep**

**Optimierung der Sepsis-Therapie auf Basis einer patientenindividuellen digitalen  
Präzisionsdiagnostik**

**Optimization of sepsis therapy based on patient-specific digital precision diagnostics  
using next generation sequencing**

**(DigiSep-Trial)**

**Trial registration no.:**

**NCT04571801**

### Unterschriften

*Stellvertretender Biometriker*

*Autor des statistischen Analyseplans*

Paul Thalmann  
Institut für Medizinische Biometrie  
Im Neuenheimer Feld 130.3  
69120 Heidelberg  
Telefon: +49-(0)6221 5634053  
E-Mail: thalmann@imbi.uni-heidelberg.de

11.12.2023

[REDACTED]  
Datum / Unterschrift

*Verantwortlicher Biometriker*

Dr. sc. hum. Manuel Feißt  
Institut für Medizinische Biometrie  
Im Neuenheimer Feld 130.3  
69120 Heidelberg  
Telefon: +49-(0)6221 5632633  
E-Mail: feisst@imbi.uni-heidelberg.de

5.12.23

[REDACTED]  
Datum / Unterschrift

*Konsortialführer, Studienleitung*

Univ.-Prof. Dr. med. Thorsten Brenner, MHBA,  
Klinik für Anästhesiologie und Intensivmedizin  
Universitätsklinikum Essen (UKE)  
Hufelandstraße 55  
45147 Essen  
Telefon: +49 201 723 1401  
E-Mail: Thorsten.Brenner@uk-essen.de

Klinik für Anästhesiologie und Intensivmedizin

Univ.-Prof. Dr.  
Univers

Hufeland  
Fon: +  
Fax: +

04.12.2023

[REDACTED]  
Datum / Unterschrift

## Inhaltsverzeichnis

Unterschriften .....	2
1 Ziele der Studie .....	5
2 Studiendesign.....	5
3 Analysepopulationen .....	5
4 Definition der zu analysierenden Endpunkte .....	6
4.1 Primärer Endpunkt.....	6
4.2 Sekundäre Endpunkte .....	7
4.2.1 Langzeitmortalität.....	7
4.2.2 Beatmungsdauer.....	7
4.2.3 Zeitdauer bis zur Schockresolution .....	7
4.2.4 Andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung.	8
4.2.5 Gesamtverweildauer im Krankenhaus .....	8
4.2.6 Kumulativer Antiinfektivaverbrauch.....	8
4.2.7 Zeitpunkt des Beginns zielgerichtete Antiinfektivatherapie .....	8
4.2.8 Retrospektive Evaluation des Expertenpanels.....	9
4.2.9 Lebensqualität.....	9
5 Datenbehandlung .....	9
6 Statistische Methoden .....	9
6.1 Deskriptive Methoden.....	9
6.2 Allgemeine Anamnese und Basisdaten .....	10
6.3 Sepsiskriterien.....	10
6.4 Neuroscores .....	10
6.5 Studiencheck.....	11
6.6 Weitere Blutkulturen .....	12

6.7	Physiologie, Katecholamine, Bilanzierung, Routinelabor.....	12
6.8	Fokussanierung, Antiinfektiva, Mikrobiologie, Virologie .....	15
7	Analyse des primären Endpunkts .....	17
7.1	Primäre Analyse .....	17
7.2	Sensitivitätsanalysen .....	18
7.3	Subgruppenanalysen:.....	19
8	Analyse der sekundären Endpunkte.....	20
8.1	Langzeitmortalität: .....	20
8.2	Beatmungsdauer: .....	21
8.3	Zeitdauer bis zur Schockresolution:.....	21
8.4	Andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung: ....	21
8.5	Gesamtverweildauer im Krankenhaus:.....	21
8.6	Kumulativer Antiinfektivaverbrauch: .....	21
8.7	Retrospektive Evaluation des Expertenpanels .....	22
9	Analyse der Sicherheitsdaten.....	22
10	Software.....	22
11	Appendix.....	23
11.1	Visitenmatrix.....	23
11.2	Modifizierter SOFA-Score .....	24
11.3	Fragebogen Expertenrunde .....	25
12	Literatur.....	26

## 1 Ziele der Studie

Die hier vorliegende randomisiert-kontrollierte, multizentrische Interventionsstudie mit dem Titel „*Optimization of sepsis therapy based on patient-specific digital precision diagnostics using next generation sequencing*“ (DigiSep-Trial) soll die Sepsis-Behandlung patientenzentriert und digital gestützt durch zeitnahe und verlässliche Identifikation des Erregers mittels Next Generation Sequencing (NGS)-basierter Präzisionsdiagnostik (in Ergänzung zur parallel stattfindenden mikrobiologischen Routinediagnostik) und einer darauf aufbauenden antiinfektiven Therapiesteuerung optimieren. Zusätzlich besteht für die behandelnden Ärzte die Möglichkeit, bei Bedarf ein webbasiertes infektiologisches (fachübergreifendes) Expertenkonsil in Anspruch zu nehmen, im Rahmen dessen NGS-Befunde und daraus resultierende Therapieentscheidungen gemeinsam besprochen werden können. Hierdurch können komplexe Befunde mit unabhängigen Experten diskutiert und Therapieentscheidungen ggf. weiter optimiert werden. Im Optimalfall werden dadurch Patienten schneller, sicherer, effizienter und kostensparender behandelt.

## 2 Studiendesign

Die vorgestellte Untersuchung stellt eine klinisch prospektive, multizentrische, randomisierte, kontrollierte Interventionsstudie (RCT) dar.

## 3 Analysepopulationen

In DigiSep werden Patienten auf Intensive Care Unit (ICU) & Intermediate Care (IMC)-Stationen von Maximalversorgern (MVs) eingeschlossen, die an einer Sepsis/septischem Schock (Sepsis-3) erkrankt sind. Unter überwachungspflichtigen Stationen werden in dem hier dargestellten Projekt bewusst sowohl ICU-, als auch IMC-Stationen zusammengefasst, um eine Vergleichbarkeit der erhobenen Outcomedaten unabhängig von der zugrundeliegenden Versorgungsstufe erzielen zu können.

Die entsprechenden Ein- & Ausschlusskriterien für die Studienteilnahme sind wie folgt definiert:

Einschlusskriterien:

- Sepsis/septischer Schock (gemäß Sepsis-3), Beginn <24h
- schriftliches Einverständnis durch den Studienteilnehmer oder dessen gesetzlichen Vertreter
- Alter ≥18 Jahre
- Gesetzliche Krankenversicherung, da Förderung durch Innovationsfonds des G-BA

Ausschlusskriterien:

- Alter <18 Jahre
- Verweigerung der Studienteilnahme
- Verlegung von der ICU/IMC <72h nach Studieneinschluss wahrscheinlich
- palliatives Therapieziel
- bestehende Therapieziellimitierung
- Tod des Patienten ist bereits bei Studieneinschluss absehbar
- Private Krankenversicherung, da Förderung durch Innovationsfonds des G-BA

Bereits eingeschlossene Patienten können kein weiteres Mal eingeschlossen werden.

Die Hauptanalysepopulation, im folgenden Full Analysis Set (FAS) genannt, umfasst alle Patienten und die Auswertung erfolgt nach dem Intention-to-treat-Prinzip, d.h. Patienten werden in der Gruppe (Interventionsgruppe (IG), Kontrollgruppe (KG)) ausgewertet in die sie randomisiert wurden, unabhängig davon welche Behandlung sie erhielten. Patienten, welche Einschlusskriterien nicht erfüllten oder ein Ausschlusskriterium erfüllten, aber dennoch in der Studie eingeschlossen wurden, werden nicht in die Analyse einbezogen.

## 4 Definition der zu analysierenden Endpunkte

### 4.1 Primärer Endpunkt

DOOR/RADAR Score, basierend auf Overall Clinical Outcome (OCO) gemäß des DOOR (Desirability of Outcome Ranking)-Prinzips mit Adjustierung gemäß RADAR (Response Adjusted for Duration of Antibiotic Risk). Folgende Einteilung wird zugrunde gelegt:

1. Gruppe: Überleben\* + ICU/IMC LOS (length of stay) <12 Tage ohne ANV (akutes Nierenversagen)\*\*
2. Gruppe: Überleben\* + ICU/IMC LOS < 12 Tage mit ANV\*\*
3. Gruppe: Überleben\* + ICU/IMC LOS  $\geq$  12 Tage ohne ANV\*\*
4. Gruppe: Überleben\* + ICU/IMC LOS  $\geq$  12 Tage mit ANV\*\*
5. Gruppe: Tod des Patienten\*

\* erfasst nach 28 Tagen bzw. bis zur Verlegung auf Normalstation. Eine Verlegung auf Normalstation wird als Überleben gewertet. Falls keine vollständigen Daten bezüglich des Überlebens oder Verlegung auf Normalstation vorliegen (Zensierung) wird mit diesen fehlenden Werten gemäß Kapitel 7.1 umgegangen.

\*\* mit der Notwendigkeit für die Durchführung einer Nierenersatztherapie.

Danach wird die Dauer der antiinfektiven Therapie einbezogen, und es ergibt sich der DOOR/RADAR-Score wie folgt:

DOOR/RADAR – Score:

$$\begin{cases} = \text{OCO} + \left[ \frac{\text{Anzahl Tage mit Antiinfektiva}}{\text{maximale Anzahl Tage mit Antiinfektiva} + 1 \text{ Tag}} \right], & \text{falls OCO} \leq 4 \\ = 5 & \text{falls OCO} = 5 \end{cases}$$

## 4.2 Sekundäre Endpunkte

Die Analyse der sekundären Endpunkte ist in zwei Auswertungszeitpunkte unterteilt: Die erste, frühere Auswertung geht mit der Auswertung des primären Endpunkts 28 Tage nach Studieneinschluss einher. Die zweite, spätere Auswertung der weiteren sekundären Endpunkte erfolgt 180 Tage nach Studieneinschluss und somit nach dem Studienende jedes Patienten. Für jeden sekundären Endpunkt wird bestimmt, ob dieser zum ersten und/oder zweiten Zeitpunkt ausgewertet wird.

### 4.2.1 Langzeitmortalität

Der Endpunkt „Langzeitmortalität“ ist definiert als die Zeit zwischen Randomisierung und Versterben bzw. Studienende. Patienten die „lost-to-follow-up“ sind, werden zensiert. Die Langzeitmortalität wird zum ersten und zweiten Zeitpunkt ausgewertet.

### 4.2.2 Beatmungsdauer

Der Endpunkt „Beatmungsdauer“ ist definiert als die kumulative Beatmungsdauer über den gesamten Studienverlauf, relativ zur Überlebensdauer. Patienten ohne Beatmungspflichtigkeit werden als 0 bewertet. Falls die Überlebensdauer zensiert ist, ist die Beatmungsdauer als fehlender Wert zu werten. Die Subgruppe der überlebenden Patienten wird absolut in Tagen erfasst. Die Beatmungsdauer wird zum ersten und zweiten Zeitpunkt ausgewertet.

### 4.2.3 Zeitdauer bis zur Schockresolution

Der Endpunkt „Zeitdauer bis zur Schockresolution“ ist definiert als die Zeit zwischen Beginn des septischen Schocks und Ende des septischen Schocks, erfasst über die kumulative Dauer der Katecholaminpflichtigkeit, relativ zur Überlebenszeit. Patienten ohne Katecholaminpflichtigkeit werden als 0 bewertet. Falls die Überlebensdauer zensiert ist, ist die „Zeitdauer bis zur Schockresolution“ als fehlender Wert zu werten. Die Subgruppe der

überlebenden Patienten wird absolut in Tagen erfasst. Die Zeitdauer bis zur Schockresolution wird zum ersten und zweiten Zeitpunkt ausgewertet.

#### **4.2.4 Andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung**

Der Endpunkt „Andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung“ ist definiert als die kumulative Zeit der Nierenersatztherapiepflichtigkeit über den gesamten Studienverlauf, relativ zur Überlebenszeit. Patienten ohne Nierenersatztherapiepflichtigkeit werden als 0 bewertet. Falls die Überlebenszeit censiert ist, ist die „Andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung“ als fehlender Wert zu werten. Die Subgruppe der überlebenden Patienten wird absolut in Tagen erfasst. Die andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung wird zum ersten und zweiten Zeitpunkt ausgewertet.

#### **4.2.5 Gesamtverweildauer im Krankenhaus**

Der Endpunkt „Gesamtverweildauer im Krankenhaus“ ist definiert als die Zeit zwischen Randomisierung und Entlassung aus dem Krankenhaus. Die Gesamtverweildauer im Krankenhaus wird zum einen relativ erfasst, als Anteil der Tage seit Randomisierung, die der Patient im Krankenhaus verweilt hat. Verstirbt der Patient im Krankenhaus, wird der Anteil als 100% gewertet. Zum anderen wird für die Subgruppe der überlebenden Patienten die Gesamtverweildauer im Krankenhaus absolut in Tagen erfasst. Die Gesamtverweildauer im Krankenhaus wird zum ersten und zweiten Zeitpunkt ausgewertet.

#### **4.2.6 Kumulativer Antiinfektivaverbrauch**

Der Endpunkt „Kumulativer Antiinfektivaverbrauch“ ist definiert als die Anzahl der Tage, in welchen der Patient ein Antiinfektivum verabreicht bekommen hat. Hierbei wird die Anzahl an Tagen für jedes Antiinfektivum separat erfasst und aufsummiert. Diese Größe wird dann relativ zur Überlebenszeit bewertet. Zusätzlich wird für die Subgruppe der überlebenden Patienten der kumulative Antiinfektivaverbrauch absolut in Tagen erfasst. Der kumulative Antiinfektivaverbrauch wird nur zum ersten Zeitpunkt ausgewertet.

#### **4.2.7 Zeitpunkt des Beginns zielgerichtete Antiinfektivatherapie**

Der genaue Zeitpunkt des Beginns einer zielgerichteten Antiinfektivatherapie wurde nicht erfasst. Dadurch kann dieser Endpunkt nicht ausgewertet werden.

#### **4.2.8 Retrospektive Evaluation des Expertenpanels**

Die Evaluation des 3-köpfigen Expertenpanels findet über einen Fragebogen bestehend aus 6 Items statt.

Pro Item werden auch die Raten der vollkommenen Übereinstimmung der 3 Bewerter, Mehrheitsübereinstimmung der 3 Bewerter und die Rate der Nichtübereinstimmung berichtet. Die Bewertung ist wie folgt:

+++;---	Vollkommende Übereinstimmung
++-;      +-;      ++fehlend;      --fehlend; -fehlend, fehlend; +fehlend, fehlend	Mehrheitsübereinstimmung
+fehlend; fehlend, fehlend, fehlend	Keine Übereinstimmung

#### **4.2.9 Lebensqualität**

Teil der gesundheitsökonomischen Auswertung, durchgeführt vom Lehrstuhl für Gesundheitsökonomie & Gesundheitsmanagement der Universität Bielefeld

### **5 Datenbehandlung**

Patienten, die die schriftliche Einwilligung vor Datenbankschluss zurückgezogen und nicht der weiteren Verwendung der Daten zugestimmt haben, werden aus der Studie ausgeschlossen.

Die Glasgow Coma Scale (GCS) kann bei Intensivpatienten häufig nicht einheitlich erhoben werden (aktueller GCS zum Zeitpunkt der Datenerhebung mit ggf. analgosedierender Medikation vs. GCS zum letztmöglichen Zeitpunkt ohne analgosedierende Medikation). Daraus folgt, dass statt dem SOFA-Score ein modifizierter SOFA-Score berechnet wird (siehe 11.2), welcher die GCS exkludiert.

### **6 Statistische Methoden**

#### **6.1 Deskriptive Methoden**

Stetige Variablen und Scores werden mit Mittelwert, Standardabweichung, Median, Q1, Q3, Minimum und Maximum, kategoriale Variablen werden mit absoluten und relativen Häufigkeiten pro Behandlungsgruppe und insgesamt beschrieben. Die Anzahl fehlender Werte wird jeweils mit angegeben; relative Häufigkeiten beziehen sich auf die Anzahl der Patienten ohne fehlende Werte. Stetige Variablen werden zwischen den beiden Behandlungsgruppen per t-Tests verglichen, ordinale und kategoriale Variablen per Chi-Quadrat-Test.

## 6.2 Allgemeine Anamnese und Basisdaten

Folgende Daten werden zu V1 erhoben.

### Demographische Daten:

- Alter
- Geschlecht (männlich/weiblich)
- BMI (wird aus Größe und Gewicht berechnet)
- Art der Zuweisung (operativer Notfall, nicht-operativer Notfall, elektiv)
- Aufenthaltsort vor ITS-Aufnahme (Notaufnahme, Schockraum, Normalstation (intern), andere ITS/IMC (intern), anderes KH, anderer Ort)

### Komorbidität:

- Charlson-Komorbiditätsindex (CCI)

### Immunstatus:

- Anamnese für Neutropenie (ja, nein, unbekannt)
- Z. n. Allogener Stammzelltransplantation (ja, nein, unbekannt)
- Längerdauernde Kortikosteroideinnahme (ja, nein, unbekannt)
- Behandlung mit anderen T-Zell-Immunsuppressiva (ja, nein, unbekannt)
- Angeborene schwere Immundefizienzsyndrome (ja, nein, unbekannt)

Die Variablen werden mit deskriptiven Methoden berichtet.

## 6.3 Sepsiskriterien

Die folgenden Daten werden zu V1 und V2 erhoben:

### Sepsiskriterien

- Vorliegen einer Infektion (ja, nein)
- Grad der Sicherung (klinisch vermutet, mikrobiologisch nachgewiesen)
- Ursprung der Infektion (ambulant, nosokomial)
- Lokalisation bekannt (ja (inkl. spezifischer Angabe), nein)
- Organdysfunktion (ja, nein)
- Laktat (mmol/l, mg/dl (inkl. Maximalwert binnen 24 Stunden))
- Notwendigkeit, den Kreislauf durch Vasopressoren soweit zu stabilisieren, dass ein mittlerer arterieller Druck von 65 mmHg erreicht wird, und ein Anstieg der Laktatkonzentration auf über 2 mmol/l (18 mg/dl) trotz ausreichender Flüssigkeitszufuhr (ja, nein)
- Zeitspanne mutmaßlicher Sepsisbeginn und Randomisierung (in Stunden)

Die Variablen werden zu beiden Zeitpunkten mit deskriptiven Methoden berichtet.

## 6.4 Neuroscores

Die folgenden Daten werden zu V1 bis V5 erhoben (falls vorzeitiger Studienabbruch: bis V4):

- RASS (ordinal: -5, -4, ..., 4)
- CAM-ICU
  - CAM-ICU erhebbar (ja, nein)
  - Merkmal 1: akuter Beginn oder schwankender Verlauf (positiv, negativ)
  - Merkmal 2: Aufmerksamkeitsstörung (positiv, negativ)
  - Merkmal 3: Unorganisiertes Denken (positiv, negativ)
  - Merkmal 4: Bewusstseinsstörung (positiv, negativ)

- Gesamt CAM-ICU (positiv, negativ)
- GCS
  - GCS erhebbar (ja, nein)
  - Zeitpunkt GCS (aktueller Wert, vor Sedierungsbeginn)
  - Augen öffnen (spontan, nach Aufforderung, auf Schmerzreiz, nicht)
  - Verbale Kommunikation ([orientiert, klar], verwirrt, einzelne Wörter, einzelne Laute, keine)
  - Motorische Reaktion (nach Aufforderung, gezielte Abwehrreaktion, ungezielte Abwehrbewegung, Beugebewegung, Streckbewegung, keine)
  - GCS-Score (ordinal: 3,4, ..., 15)

Die Variablen werden zu jedem Zeitpunkt mit deskriptiven Methoden berichtet.

## 6.5 Studiencheck

Die folgenden Daten werden zu V1 in der IG erhoben, falls „Abnahme der Blutproben für die NGS erfolgt“ mit „Ja“ beantwortet. Absolute und relative Häufigkeiten der Variable „Abnahme Blutproben für die NGS erfolgt“ werden berichtet.

- Änderung des antiinfektiven Therapieregimes aufgrund der NGS-Befunde von V1/Tag 0 (ja, nein)
  - Durchgeführte Änderungen (Erweiterung des antiinfektiven Keimspektrums, Verschmälerung des antiinfektiven Keimspektrums, Antiinfektiva komplett absetzen, Sonstiges (Freitext))

Die folgenden Daten werden zu V2 in der IG erhoben falls „Abnahme der Blutproben für die NGS erfolgt“ mit „Ja“ beantwortet:

- Änderung des antiinfektiven Therapieregimes aufgrund der NGS-Befunde von V2/Tag 3 (ja, nein)
  - Durchgeführte Änderungen (Erweiterung des antiinfektiven Keimspektrums, Verschmälerung des antiinfektiven Keimspektrums, Antiinfektiva komplett absetzen, Sonstiges (Freitext))

Die folgenden Daten werden zu V1 bis V4 erhoben:

- Abnahme der Blutproben für erweitertes Immunmonitoring und Transkriptomanalysen erfolgt? (ja, nein)

Die Variablen werden zu allen Zeitpunkten mit deskriptiven Methoden berichtet.

## 6.6 Weitere Blutkulturen

Die folgenden Daten werden zu V2 bis V4 erhoben:

- Wurden weitere Blutkulturen (jeweils 2 Sets: 2x aerob & 2x anaerob) seit der letzten planmäßigen Visite abgenommen? (ja (Datum), nein)
- Erfolgte die Abnahme außerplanmäßiger Blutproben für das erweiterte Immunmonitoring und Transkriptomanalysen seit der letzten regulären Probenabnahme? (ja, nein)

Die folgenden Daten werden zu V2 bis V4 nur in der IG erhoben:

- Wurden weitere NGS-Proben seit der letzten planmäßigen Visite abgenommen? (ja, nein)
  - Änderung des antiinfektiven Therapieregimes aufgrund der NGS-Befunde? (ja, nein)
    - Durchgeführte Änderungen (Erweiterung des antiinfektiven Keimspektrums, Verschmälerung des antiinfektiven Keimspektrums, Antiinfektiva komplett absetzen, Sonstiges (Freitext))

Die Variablen werden zu allen Zeitpunkten mit deskriptiven Methoden berichtet.

## 6.7 Physiologie, Katecholamine, Bilanzierung, Routinelabor

Die folgenden Daten werden zu V2 bis V5 erhoben. Falls unterschiedliche Einheiten vorliegen, werden diese zusammengeführt.

### Physiologie

- Herzfrequenz (ja, nein),
- Herzfrequenz (Minimum und Maximum in 1/min),
- Systolischer Blutdruck (ja, nein)
- Systolischer Blutdruck (Minimum und Maximum in mmHg),
- Mittlerer arterieller Blutdruck (ja, nein)
- Mittlerer arterieller Blutdruck (Minimum und Maximum in mmHg),
- Körpertemperatur (ja, nein)
- Körpertemperatur (Minimum und Maximum in °C)
- Atemfrequenz (ja, nein)
- Atemfrequenz (Minimum und Maximum in 1/min)
- Bicarbonat (ja, nein)
- Bicarbonat (Minimum und Maximum in mmol/l),
- pH (ja, nein)
- pH (Minimum und Maximum in pH-Wert)
- Horowitz-Quotient (ja, nein);
- Horowitz-Quotient (Messwert in (kPa, mmHg)),
- PaO<sub>2</sub> (ja, nein)
- PaO<sub>2</sub> (Messwert in (kPa, mmHg))
- FiO<sub>2</sub> (ja, nein)
- FiO<sub>2</sub> (Messwert in %)
- PaCO<sub>2</sub> (ja, nein)
- PaCO<sub>2</sub> (Messwert in (kPa, mmHg)),
- Art der Beatmung (keine, nicht-invasiv, invasiv)
- Urinausscheidung (ja, nein)

- Menge der Urinausscheidung (ml)
- Dauer der Urinausscheidung (h)
- Nierenersatztherapie (ja, nein)
- Nierenersatztherapie: Dialyseverfahren (CVVH (kontinuierliche venovenöse Hämofiltration), CVVHD (kontinuierliche venovenöse Hämodialyse), CVVHDF (kontinuierliche venovenöse Hämodiafiltration), IHD (intermittierende Hämodialyse), SLED (sustained low efficiency dialysis), Sonstiges (nähere Beschreibung (Freitext))
- Art der Antikoagulation (keine, unfraktioniertes Heparin, Citrat-Antikoagulation, andere (Freitext))
- ECMO (ja, nein)
- ECMO (Art: venoarteriell, venovenös)

### Katecholamine

- Einsatz von Vasopressoren und Inotropika (ja, nein)
- Einsatz von Adrenalin (ja, nein)
- Einsatz von Adrenalin (Maximale Dosierung in (µg/kg/min, µg/min, mg/h))
- Einsatz von Noradrenalin (ja, nein)
- Einsatz von Noradrenalin (Maximale Dosierung in (µg/kg/min, µg/min, mg/h))
- Einsatz von Dopamin (ja, nein)
- Einsatz von Dopamin (Maximale Dosierung in (µg/kg/min, µg/min, mg/h))
- Einsatz von Dobutamin (ja, nein)
- Einsatz von Dobutamin (Maximale Dosierung in (µg/kg/min, µg/min, mg/h))
- Einsatz von Vasopressin (ja, nein)
- Einsatz von Vasopressin (Maximale Dosierung in (µg/kg/min, µg/min, mg/h, IE/h))
- Einsatz von Hydrocortison (ja, nein)
- Einsatz von Hydrocortison (Maximale Dosierung (mg/Tag))
- Einsatz von Dexamethason (ja, nein)
- Einsatz von Dexamethason (Maximale Dosierung (mg/Tag))

### Bilanzierung

#### Volumentherapie

- Wurde eine Volumentherapie durchgeführt? (ja, nein)
- Kristalloide Lösungen (ja, nein)
- Kristalloide Lösungen (Gesamtmenge (ml/24h)),
- Humanalbumin 20% (ja, nein)
- Humanalbumin 20% (Gesamtmenge (ml/24h))
- Humanalbumin 4-5% (ja, nein)
- Humanalbumin 4-5% (Gesamtmenge (ml/24h))
- Erythrozytenkonzentrate (ja, nein)
- Erythrozytenkonzentrate (Gesamtmenge (ml/24h))
- Frischplasmen (ja, nein)
- Frischplasmen (Gesamtmenge (ml/24h)))
- Thrombozytenkonzentrate (ja, nein)
- Thrombozytenkonzentrate (Gesamtmenge (ml/24h))
- Gelatinepräparate (ja (Gesamtmenge (ml/24h)), nein)
- Gelatinepräparate (ja (Gesamtmenge (ml/24h)), nein)
- HES 6% [200/0,5] (ja, nein)
- HES 6% [200/0,5] (Gesamtmenge (ml/24h))
- HES 10% [200/0,5] (ja, nein)

- HES 10% [200/0,5] (Gesamtmenge (ml/24h))
- HES 6% [130/0,4] (ja, nein)
- HES 6% [130/0,4] (Gesamtmenge (ml/24h))
- HES [70/0,5] (ja, nein)
- HES [70/0,5] (Gesamtmenge (ml/24h))
- Dextrane (ja, nein)
- Dextrane (Gesamtmenge (ml/24h))
- andere kolloidale Lösungen (ja, nein)
- andere kolloidale Lösungen (Gesamtmenge (ml/24h))

#### Ausfuhr

- Urinausfuhr (ja, nein)
- Urinausfuhr (Gesamtmenge (ml/24h))
- Dialyse (ja, nein)
- Dialyse (Gesamtmenge Ultrafiltration (ml/24h))
- Drainagen (ja, nein)
- Drainagen (Gesamtmenge (ml/24h))
- Blutverlust (ja, nein)
- Blutverlust (Gesamtmenge (ml/24h))
- Abführen/Stoma (ja, nein)
- Abführen/Stoma (Gesamtmenge (ml/24h))
- Reflux/Erbrechen (ja, nein)
- Reflux/Erbrechen (Gesamtmenge (ml/24h))

#### Routinelabor

- Natrium (mehrmalig, einmalig, keine)
- Natrium (Minimum & Maximum bei mehrmalig, Wert bei einmalig in (mmol/l, mg/dl))
- Kalium (mehrmalig, einmalig, keine)
- Kalium (Minimum & Maximum bei mehrmalig, Wert bei einmalig in (mmol/l, mg/dl))
- Harnstoff (mehrmalig, einmalig, keine)
- Harnstoff (Minimum & Maximum bei mehrmalig, Wert bei einmalig in (mmol/l,  $\mu$ mol/l, mg/dl))
- Kreatinin (mehrmalig, einmalig, keine)
- Kreatinin (Minimum & Maximum bei mehrmalig, Wert bei einmalig in (mg/dl,  $\mu$ mol/l))
- GFR ([Ja, nach Cockcroft-Gault], [Ja, nach MDRD], [nicht erhoben]),
- GFR (mehrmalig, einmalig, keine)
- GFR (Minimum & Maximum bei mehrmalig, Wert bei einmalig in ml/min)
- Gesamt-Bilirubin (mehrmalig, einmalig, keine)
- Gesamt-Bilirubin (Minimum & Maximum bei mehrmalig, Wert bei einmalig in (mg/dl,  $\mu$ mol/l))
- Laktat (mehrmalig, einmalig, keine)
- Laktat (Minimum & Maximum bei mehrmalig, Wert bei einmalig in (mg/dl,  $\mu$ mol/l))
- Albumin (mehrmalig, einmalig, keine)
- Albumin (Minimum & Maximum bei mehrmalig, Wert bei einmalig in (g/l, g/dl))
- CRP (mehrmalig, einmalig, keine)
- CRP (Minimum & Maximum bei mehrmalig, Wert bei einmalig in (mg/l, mg/dl,  $\mu$ mol/l))
- PCT (mehrmalig, einmalig, keine)
- PCT (Minimum & Maximum bei mehrmalig, Wert bei einmalig in ( $\mu$ g/l, ng/ml))
- Gesamteiweiß (mehrmalig, einmalig, keine)
- Gesamteiweiß (Minimum & Maximum bei mehrmalig, Wert bei einmalig) in (g/l, g/dl)

- ALAT (GPT) (mehrmalig, einmalig, keine)
- ALAT (GPT) (Minimum & Maximum bei mehrmalig, Wert bei einmalig) in (µmol/l, U/l)
- ASAT (GOT) (mehrmalig, einmalig, keine)
- ASAT (GOT) (Minimum & Maximum bei mehrmalig, Wert bei einmalig) in (µmol/l, U/l)
- Alkalische Phosphatase (AP) (mehrmalig, einmalig, keine)
- Alkalische Phosphatase (AP) (Minimum & Maximum bei mehrmalig, Wert bei einmalig) in (µmol/l, U/l)
- Hämatokrit (mehrmalig, einmalig, keine)
- Hämatokrit (Minimum & Maximum bei mehrmalig, Wert bei einmalig) in %
- Leukozyten (mehrmalig, einmalig, keine)
- Leukozyten (Minimum & Maximum bei mehrmalig, Wert bei einmalig) in (gpt/l, x 103/µl, /nl)
- Thrombozyten (mehrmalig, einmalig, keine)
- Thrombozyten (Minimum & Maximum bei mehrmalig, Wert bei einmalig) in (gpt/l, x 103/µl, /nl)
- Quick (mehrmalig, einmalig, keine)
- Quick (Minimum & Maximum bei mehrmalig, Wert bei einmalig) in %
- aPTT (mehrmalig, einmalig, keine)
- aPTT (Minimum & Maximum bei mehrmalig, Wert bei einmalig) in sec
- INR (mehrmalig, einmalig, keine)
- INR (Minimum & Maximum bei mehrmalig, Wert bei einmalig)

Die Variablen werden zu allen Zeitpunkten mit deskriptiven Methoden berichtet. Vergleiche für kontinuierliche Variablen über alle Zeitpunkte werden anhand einer repeated measures ANOVA durchgeführt und grafisch über Boxplot-Verläufe dargestellt.

## 6.8 Fokussanierung, Antiinfektiva, Mikrobiologie, Virologie

Die folgenden Daten sind „repeatable instruments“

### Fokussanierung

- Durchgeführt? (ja, nein)
- Zeitpunkt (Tag seit Studieneinschluss)
- Art der Fokussanierung (chirurg. Fokussanierung, Entfernen intravasaler Katheter, interventionelle Fokussanierung)
- Nähere Beschreibung (Freitext)

### Antiinfektiva

- Substanzwahl (Antibiotika, Virostatika, Antimykotika)
- Art & nähere Beschreibung (Auswahl und ggf. Freitext)
- Beginn der Gabe (-3d, -2d, -1d, 0d (= Sepsis-Onset), +1d, +2d, +3d, +4d, +5d, +6d, +7d)
- Antiinfektiva-Gabe andauernd? (ja, nein)
- Ende der Gabe (-3d, -2d, -1d, 0d (= Sepsis-Onset), +1d, +2d, +3d, +4d, +5d, +6d, +7d, +>7d)
- Strategie der Antiinfektiva-Gabe (durchweg empirisch, durchweg gezielt, zunächst empirisch begonnen, dann aber gezielt weitergeführt)

Mikrobiologie:

- Probenmaterial (Blutkultur, respiratorisches Sekret, frisches Drainagematerial, altes Drainagematerial, OP-Abstrich, Wundabstrich, Punktat, Katheterspitze, frischer Urin, alter Urin, Liquor, Gewebe, Hygieneabstrich, Stuhl)
- Studienbedingte Abnahme von Blutkulturen (2x aerob & 2x anaerob) zum Zeitpunkt Sepsis-Onset (ja, nein)
- Studienbedingte Abnahme von Blutkulturen (2x aerob & 2x anaerob) zum Zeitpunkt 72h nach Sepsis-Onset (ja, nein)
- Probentyp (aerob, anaerob)
- Ort der Abnahme (arteriell, venös)
- Zugang (peripher, zentral)
- Ort der Abnahme (chirurgische Drainage, interventionelle Drainage)
- Art (Easyflow, Robinson, Saug-Spülendrainage, T-Drainage, sonstiges)
- Art (Abszess- oder Verhaltdrainage, PTCD, sonstiges)
- Art des Materials (ZVK, Arterie, EVD, PTCD, sonstiges)
- Ort der Abnahme (rektal, nasal)
- Ist der Zeitpunkt der Probenentnahme gegeben (ja, nein)
- Tag der Abnahme (-3d, -2d, -1d, 0d (= Sepsis-Onset), +1d, +2d, +3d, +4d, +5d, +6d)
- Befund positiv (ja, nein)
- Einteilung (VRE, 3-MRGN, 4-MRGN)
- Einteilung (MRSA)
- Art des Erregers (Enterococcus faecium, Enterococcus faecalis)
- Art des Erregers (Escherichia spp., Pseudomonas spp., Klebsiella spp., Citrobacter spp., Enterobacter spp., Serratia spp., Proteus spp., Acinetobacter spp., sonstige
- Kokken (ja, nein):  
Wenn ja, dann gebe Art des Erregers, Nachweis und Therapie an (vgl. CRF)
- Enterobactericeae (ja, nein):  
Wenn ja, dann gebe Art des Erregers, Nachweis, Therapie, ESBL-Phänotyp, Ceftriaxon- und Cefotaxim-Resistenz und MHK mit Wert an (vgl. CRF)
- Non-Fermenter (ja, nein):  
Wenn ja, dann gebe Art des Erregers, Nachweis und Therapie an (vgl. CRF)
- Pilze (ja, nein):  
Wenn ja, dann gebe Art des Erregers, Nachweis und Therapie an (vgl. CRF)
- Relevante Erreger (ja, nein) (kommt aus Stuhlprobe):  
Wenn ja, dann gebe Art des Erregers, Nachweis und Therapie an (vgl. CRF)

Virologie:

- Probenmaterial (Blut, respiratorisches Sekret, Abstrich, Liquor, Stuhl)
- Art der Diagnostik (CMV-PCR, VZV-PCR, HSV1-PCR, pp65-Ag, Asp.-Ag, SARS-CoV2-PCR, SARS-CoV2-AG, RSV-PCR, Para-Influenza-PCR, Influenza-PCR, sonstiges)
- Ist der Zeitpunkt der Probenentnahme gegeben (ja, nein)
- Tag der Abnahme (-3d, -2d, -1d, 0d (= Sepsis-Onset), +1d, +2d, +3d, +4d, +5d, +6d)

- Befund positiv (ja, nein)
- Norovirus (ja, nein)
- Rotavirus (ja, nein)
- Astrovirus (ja, nein)
- Sonstiges (ja, nein)
- Angabe je nach Art der Diagnostik in [Anzahl Kopien, keine Einheit oder 1/500000 Zellen]

Die Variablen werden mit deskriptiven Methoden berichtet. Weiter wird für die Mikrobiologie und Virologie die Gesamtanzahl der Proben pro Tag und die Anzahl der Proben pro Patient mit deskriptiven Methoden berichtet.

## 7 Analyse des primären Endpunkts

### 7.1 Primäre Analyse

Die primäre Hypothese der Studie ist, dass es einen Unterschied zwischen der IG und der KG, jeweils des FAS, bezüglich des DOOR/RADAR-Scores gibt (Definition siehe 4.1).

Die Nullhypothese

$$H_0: P(DOOR/RADAR_{IG} > DOOR/RADAR_{KG}) = P(DOOR/RADAR_{IG} < DOOR/RADAR_{KG})$$

wird gegen die Alternativhypothese

$$H_1: P(DOOR/RADAR_{IG} > DOOR/RADAR_{KG}) \neq P(DOOR/RADAR_{IG} < DOOR/RADAR_{KG})$$

mit einem zweiseitigen Signifikanzniveau von  $\alpha = 0.05$  unter Verwendung eines Mann-Whitney-U-Tests getestet. Zusätzlich zu dem (asymptotischen) p-Wert werden normalapproximative 95% Konfidenzintervalle angegeben.

#### *Umgang mit fehlenden Werten*

Für fehlende Werte in den Variablen, welche die Komponenten des DOOR/RADAR-Scores bilden, wird eine multiple Imputation für die jeweiligen Einzelvariablen durchgeführt. Die verwendete Methode ist das „predictive mean matching“, welche die Variablen Alter, SOFA-Score bei Sepsisbeginn und Überlebensstatus einbezieht. Die Anzahl der Spender für das predictive mean matching ist  $d = 5$ .

Bei der multiplen Imputation werden  $m = 10$  imputierte Datensätze generiert. Da eine angemessene Anzahl von Burn-in-Iterationen ohne Kenntnis des Missing-Patterns nicht sicher angegeben werden kann, werden 50 Burn-in-Iterationen als Ausgangspunkt verwendet und gegebenenfalls angepasst. Der für die multiple Imputation verwendete Seed ist 230842.

Die Analysen bezüglich des primären Endpunkts werden separat für jeden erstellten Datensatz durchgeführt und gemäß Rubin's Rule (Rubin 1987) kombiniert.

## 7.2 Sensitivitätsanalysen

Die multiple Imputation wird durch verschiedene Sensitivitätsanalysen unterstützt:

- Zum einen werden „best case“ und „worst case“ Analysen durchgeführt, um der Beurteilung der Imputation einen Rahmen zu geben. Zusätzlich wird eine “complete case” Analyse durchgeführt, bei der alle Patienten mit fehlenden Werten für den DOOR/RADAR-Score ausgeschlossen werden.
- Zudem wird die Anzahl Spender für das predictive mean matching variiert (zu  $d = 3$  und  $d = 10$ ), um die Robustheit der Ergebnisse zu gewährleisten. Eine zusätzliche Variation der Anzahl an imputierten Datensätzen  $m$  (zu  $m = 5$  und  $m = 20$ ) unterstützt diese ebenfalls.

Die Primäranalyse wird durch eine logistische Regression ergänzt. Hierbei wird der ursprünglich ordinale DOOR/RADAR-Score in eine binäre Variable transformiert, indem zwischen einem DOOR/RADAR-Score von  $<2$  und einem DOOR/RADAR-Score von  $\geq 2$  unterschieden wird. Als exogene Variablen werden neben der Gruppenzugehörigkeit (IG, KG) auch das Alter, Geschlecht, SOFA-Score und der Charlson-Komorbiditätsindex in das Modell aufgenommen. Für alle Kovariaten werden Odds-Ratios inklusive 95%-Konfidenzintervalle und entsprechende p-Werte angegeben.

Um der Verabreichung verschiedener Antiinfektiva Rechnung zu tragen, wird ein modifizierter DOOR/RADAR-Score für die ersten 7 Tage berechnet. Dieser bezieht sich auf die Antiinfektivatage. Ein Antiinfektivatag ist definiert als ein Tag mit einem verabreichten Antiinfektiva, d.h. dass bspw. ein Patient, welcher Antiinfektiva A über 4 Tage und Antiinfektiva B über 3 Tage verabreicht bekommen hat, mit 7 Antiinfektivatagen gewertet wird.

modifizierter DOOR/RADAR – Score:

$$\left\{ \begin{array}{l} = \text{OCO} + \left[ \frac{\text{Antiinfektivatage}}{\text{Anzahl verabreichter Antiinfektiva} * \text{maximale Anzahl Tage mit Antiinfektiva} + 1 \text{ Tag}} \right], \\ \text{falls OCO} \leq 4 \\ = 5, \text{falls OCO} = 5 \end{array} \right.$$

Anschließend wird eine Analyse analog zur primären Analyse (siehe 7.1) durchgeführt.

### 7.3 Subgruppenanalysen:

Die Primäranalyse des DOOR/RADAR-Scores wird ebenfalls für verschiedene Subgruppen durchgeführt. Kategorien können kombiniert werden, wenn die Subgruppen zu klein für valide Analysen sind.

Für fehlende Werte werden die imputierten Datensätze aus der multiplen Imputation der primären Analyse verwendet und gemäß Rubin's Rule (Rubin 1987) kombiniert.

Die Subgruppen sind:

- Zeitpunkt Abnahme NGS:
  - Patienten ohne zeitverzögerte Abnahme
  - Patienten mit zeitverzögerter Abnahme
- Erkrankungsschwere im Verlauf:
  - Septischer Schock: V1- Nein und V2 - Nein
  - Septischer Schock: V1- Ja und V2 - Nein
  - Septischer Schock: V1- Nein und V2 - Ja
  - Septischer Schock: V1- Ja und V2 - Ja
- Fokus (Ja oder Vermutet):
  - Pneumonie, sonstige obere oder untere Atemwege, Thorakal
  - Gastrointestinal, Gallenwege, Intraabdominell / Peritonitis
  - Bakterämie, Katheterinfektion
  - Knochen / Weichteile, Chirurg. Wundinfektion
  - ZNS
  - Kardiovaskulär
  - Urogenital
  - Sonstiges
- Vorliegen Infektion
  - Grad der Sicherung
    - Klinisch vermutet
    - Mikrobiologisch nachgewiesen
  - Ursprung der Infektion
    - Ambulant
    - Nosokomial
- Immunsuppression
  - Immunsuppression liegt vor
  - Immunsuppression liegt nicht vor
- Charlson Comorbidity Index zu Baseline
  - Charlson Comorbidity Index > 4
  - Charlson Comorbidity Index ≤ 4

- Charlson Comorbidity Index kleiner/gleich als Median
- Charlson Comorbidity Index größer als Median
- Charlson Comorbidity Index größer als 3.Quartil
- Modifizierter SOFA-Score zu Baseline:
  - Kleiner/gleich als Median
  - Größer als Median
  - Größer als 3.Quartil
- Blutkultur
  - Positiv (Patienten mit mindestens einem positiven Befund in der Mikrobiologie)
  - Negativ
- Fokussanierung
  - Ja
  - Nein
  - Art der Fokussanierung
    - chirurg. Fokussanierung
    - Entfernen intravasaler Katheter
    - Interventionelle Fokussanierung
- Therapieänderung: Wurde für einen Patienten eine Therapieänderung aufgrund NGS-Diagnostik durchgeführt oder vorgeschlagen, ungeachtet des Zeitpunkts, ist dieser der Subgruppe „Veränderung“ zuzuordnen
  - Durchgeführte Veränderung in IG vs. Vorgeschlagene Änderung in KG (nachträglich durch Expertenpanel)
  - Vorgeschlagene Veränderung in IG (nachträglich durch Expertenpanel) vs. Vorgeschlagene Änderung in KG (nachträglich durch Expertenpanel)
  - Keine durchgeführte Veränderung in IG vs. Keine vorgeschlagene Änderung in KG
  - Keine vorgeschlagene Veränderung in IG (nachträglich durch Expertenpanel) vs. Keine vorgeschlagene Änderung in KG

## 8 Analyse der sekundären Endpunkte

Die Analyse aller sekundären Endpunkte erfolgt auf Basis des FAS.

### 8.1 Langzeitmortalität:

Die Langzeitmortalität wird 28 und 180 Tage nach Studieneinschluss analysiert.

Deskriptiv werden für IG und KG sowie das gesamte Patientenkollektiv die Anzahl an Events bis zum jeweiligen Zeitpunkt angegeben. Anschließend werden Kaplan-Meier Plots erstellt, welche zwischen der IG und KG unterscheiden. Zusätzlich wird ein deskriptiver Log-Rank Test berechnet, dessen p-Wert ebenfalls angegeben wird. Das 28-Tage, 90-Tage und 180-

Tagesüberleben werden über den Kaplan-Meier-Schätzer mit zugehörigem 95%-Konfidenzintervall pro Gruppe und insgesamt beschrieben.

### **8.2 Beatmungsdauer:**

Die Beatmungsdauer wird 28 und 180 Tage nach Studieneinschluss analysiert.

Die abgeleiteten Variablen (siehe Kapitel 4.2) werden mit deskriptiven Methoden ausgewertet. Die Differenz zwischen den beiden Gruppen wird über den Mittelwert mit zugehörigem 95% Konfidenzintervall berichtet.

### **8.3 Zeitdauer bis zur Schockresolution:**

Die Zeitdauer bis zur Schockresolution wird 28 und 180 Tage nach Studieneinschluss analysiert.

Die abgeleiteten Variablen (siehe Kapitel 4.2) werden mit deskriptiven Methoden ausgewertet. Die Differenz zwischen den beiden Gruppen wird über den Mittelwert mit zugehörigem 95% Konfidenzintervall berichtet.

### **8.4 Andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung:**

Die andauernde Nierenersatztherapiepflichtigkeit in der Nachbeobachtung wird 28 und 180 Tage nach Studieneinschluss analysiert.

Die abgeleiteten Variablen (siehe Kapitel 4.2) werden mit deskriptiven Methoden ausgewertet. Die Differenz zwischen den beiden Gruppen wird über den Mittelwert mit zugehörigem 95% Konfidenzintervall berichtet.

### **8.5 Gesamtverweildauer im Krankenhaus:**

Die Gesamtverweildauer im Krankenhaus wird 28 Tage und 180 Tage nach Studieneinschluss analysiert.

Die abgeleiteten Variablen (siehe Kapitel 4.2) werden mit deskriptiven Methoden ausgewertet. Die Differenz zwischen den beiden Gruppen wird über den Mittelwert mit zugehörigem 95% Konfidenzintervall berichtet.

### **8.6 Kumulativer Antiinfektivaverbrauch:**

Der kumulative Antiinfektivaverbrauch wird 28 Tage nach Studieneinschluss analysiert.

Die abgeleitete Variable (siehe Kapitel 4.2) wird mit deskriptiven Methoden ausgewertet. Die Differenz zwischen den beiden Gruppen wird über den Mittelwert mit zugehörigem 95% Konfidenzintervall berichtet.

## 8.7 Retrospektive Evaluation des Expertenpanels

Pro Item wird die Rate an vollkommenen Übereinstimmungen, Mehreinheitsübereinstimmung und Nichtübereinstimmung mit deskriptiven Methoden berichtet. Alle Prozentsätze werden insgesamt und auch getrennt nach Behandlungsgruppe mit exakten Konfidenzintervallen (Pearson-Clopper) berichtet.

Danach wird bzgl. der Fälle mit einer Übereinstimmung (vollkommen oder Mehrheit) jedes Item des Fragebogens mit deskriptiven Methoden analysiert.

Zusätzlich wird in der IG der Vergleich zwischen empfohlener und tatsächlicher Therapieänderung über eine Vierfeldertafel dargestellt und mit einem Chi-Quadrat-Test evaluiert.

## 9 Analyse der Sicherheitsdaten

Folgende Sicherheitsdaten werden erhoben:

- Ist ein UE aufgetreten (ja, nein)
- Ereignis (Text)
- Beginn des Ereignisses (Datum)
- Intensität (leicht, mittel, schwer)
- SUE (ja, nein)
- Ausgang (Wiederhergestellt, Besserung, nicht wiederhergestellt, wiederhergestellt mit Folgen, Tod, unbekannt)
  - Wenn „Ausgang = Wiederhergestellt“ oder „Ausgang = Wiederhergestellt mit Folgen“: Beendigung des Ereignisses (Datum)

Aus dem erhobenen Beginn des Ereignisses und dem Ende wird die Dauer des Ereignisses berechnet und berichtet. Die Variablen werden mit deskriptiven Methoden berichtet. Zudem wird auch die Anzahl an Patienten mit UE und SUE sowie die Anzahl der UE und SUE pro Patient jeweils mit deskriptiven Methoden untersucht. Zusätzlich wird ein Listing aller aufgetretenen UEs erstellt.

## 10 Software

Alle Analysen werden einer validierten R-Umgebung durchgeführt, basierend auf R-Version  $\geq 4.2.0$ . Die folgenden externen Packages werden für den angegebenen Zweck verwendet:

- Für Imputation: MICE
- Für allgemeine Plots: ggplot2
- Für Survival Plots: survminer
- Für Survival Analysis: survival
- Für Datenaufbereitung: Tidyverse

## 11 Appendix

### 11.1 Visitenmatrix

Data Collection Instrument	Screening	Expertenkonsil Tag 0 - 14	V1   Tag 0	V2   Tag 3	V3   Tag 7	V4   Tag 14	V5   Tag 28	V6   Tag 90	V7   Tag 180	Log Pages
Einwilligung	●									
Ein-/ Ausschlusskriterien	●									
Expertenkonsil		●								
Randomisierung			●							
Demographie			●							
Komorbiditaet			●							
Immunstatus			●							
Sepsiskriterien			●	●						
Outcome							●	●	●	
Neuroscores			●	●	●	●	●	●*		
Studiencheck			●	●	●	●				
Weitere Blutkulturen			●	●	●	●				
Physiologie			●	●	●	●	●	●*		
Katecholamine			●	●	●	●	●	●*		
Bilanzierung			●	●	●	●	●	●*		
Routinelabor			●	●	●	●	●	●*		
Lebensqualität			●					●*	●*	
Fokussierung					●*					
Antiinfektiva					●*					
Mikrobiologie					●*					
Virologie					●*					
Signatur								●		
Unerwünschte Ereignisse									●*	
Protokollverletzung									●*	

\* Repeatable Instruments

\* nur ausfüllbar wenn Studie nicht vorzeitig abgebrochen

## 11.2 Modifizierter SOFA-Score

modifizierter SOFA-Score		0	1	2	3	4
<b>Herz/ Kreislauf</b>	<b>MAP [mmHg] / Katecholamine [µg/kg/min]*</b>	$\geq 70$	$< 70$	Dopamin $< 5^*$ oder Dobutamin	Dopamin $5-15^*$ oder Adrenalin $\leq 0,1^*$ oder Noradrenalin $\leq 0,1^*$	Dopamin $>15^*$ oder Adrenalin $>0,1^*$ oder Noradrenalin $>0,1^*$
<b>Lunge</b>	<b>paO<sub>2</sub>/FiO<sub>2</sub></b>	$\geq 400$	$< 400$	$< 300$	$< 200$	$< 100$
<b>Leber</b>	<b>Bilirubin [mg/dl]</b>	$< 1,2$	$1,2 - 1,9$	$2,0 - 5,9$	$6,0 - 11,9$	$> 12$
<b>Gerinnung</b>	<b>Thrombozyten <math>\times 10^3/\mu\text{l}</math></b>	$\geq 150$	$< 150$	$< 100$	$< 50$	$< 20$
<b>Niere</b>	<b>Kreatinin [mg/dl]</b> <b>Urin [ml/d]</b>	$< 1,2$	$1,2 - 1,9$	$2,0 - 3,4$	$3,5 - 4,9$	$> 5$
					$< 500$	$< 200$

Für jedes Organsystem werden 0-4 Punkte vergeben, die Spannweite des modifizierten SOFA-Scores ist damit 0-20. Liegen für einen Zeitraum mehrere Messungen vor, so wird das jeweilige Maximum zur Berechnung des modifizierten SOFA-Scores verwendet.

### 11.3 Fragebogen Expertenrunde

Sind die NGS-Befunde plausibel?	Tag 0 (Sepsis Onset)	Ja	Nein
Sind die NGS-Befunde plausibel?	Tag 3	Ja	Nein
Würden Sie aufgrund der NGS-Befunde erwägen, die empirische antiinfektive Therapie unverändert fortzusetzen?	Würden Sie aufgrund der NGS-Befunde erwägen, die empirische antiinfektive Therapie unverändert fortzusetzen?	Ja	Nein
In welcher Art und Weise würden Sie eine Änderung der empirischen antiinfektiven Therapiestrategie in Betracht ziehen?	a. Würden Sie auf Basis der NGS-Befunde eine Einengung des antiinfektiven Wirkspektrums vornehmen?	Ja	Nein
In welcher Art und Weise würden Sie eine Änderung der empirischen antiinfektiven Therapiestrategie in Betracht ziehen?	b. Würden Sie auf Basis der NGS-Befunde eine Erweiterung des antiinfektiven Wirkspektrums vornehmen?	Ja	Nein
In welcher Art und Weise würden Sie eine Änderung der empirischen antiinfektiven Therapiestrategie in Betracht ziehen?	c. Würden Sie negative NGS-Befunde dazu veranlassen, in dem hier vorliegenden Fall die gesamte antiinfektive Therapie zu verkürzen oder gar abzusetzen?	Ja	Nein

## 12 Literatur

Rubin, D.B., 1987. Multiple Imputation for Nonresponse in Surveys, 1st ed, Wiley Series in Probability and Statistics. Wiley. <https://doi.org/10.1002/9780470316696>

## **Anlage 4 - Evaluationsbericht DigiSep (01N VF20013)**

### **Tabellenverzeichnis**

Tabelle 1: Berechnung inkrementelle Kosten und inkrementeller Effekt.....	3
Tabelle 2: Beantwortungsmodus EQ-5D-5L (T0) .....	3
Tabelle 3: Ergebnisse der EQ-5D-5L Dimensionen.....	4
Tabelle 4: Gesundheitsbezogene Lebensqualität - EQ-5D-5L Index (Verbundene Stichprobe)7	7
Tabelle 5: Mehrfaktorielle ANOVA mit Messwiederholungen - EQ-5D-5L Index .....	7
Tabelle 6: Gesundheitsbezogene Lebensqualität - EQ-5D-5L VAS (Verbundene Stichprobe) . 9	9
Tabelle 7: Mehrfaktorielle ANOVA mit Messwiederholungen - EQ-5D-5L VAS.....	9

### **Abbildungsverzeichnis**

Abbildung 1: Modelldiagnostik des GAMLSS für Gesamtkosten (Grundmodell) .....	2
Abbildung 2: Modelldiagnostik des GAMLSS für Gesamtkosten (Finales Modell) .....	2
Abbildung 3: Gesundheitsbezogene Lebensqualität – EQ-5D-5L Index (Gesamtkohorte) .....	6
Abbildung 4: Gesundheitsbezogene Lebensqualität – EQ-5D-5L Index (GKV-Kohorte) .....	6
Abbildung 5: Gesundheitsbezogene Lebensqualität – EQ-5D-5L VAS (Gesamtkohorte).....	8
Abbildung 6: Gesundheitsbezogene Lebensqualität – EQ-5D-5L VAS (GKV-Kohorte) .....	8
Abbildung 7: Gesundheitsbezogene Lebensqualität – VR-36 (Gesamtkohorte).....	10
Abbildung 8: Gesundheitsbezogene Lebensqualität – VR-36 (GKV- Kohorte).....	10
Abbildung 9: Gesundheitsbezogene Lebensqualität – VR-36 (Gesamtkohorte, verbundene Stichprobe) .....	11
Abbildung 10: Gesundheitsbezogene Lebensqualität – VR-36 (GKV-Kohorte, verbundene Stichprobe).....	11

Abbildung 1: Modelldiagnostik des GAMLSS für Gesamtkosten (Grundmodell)

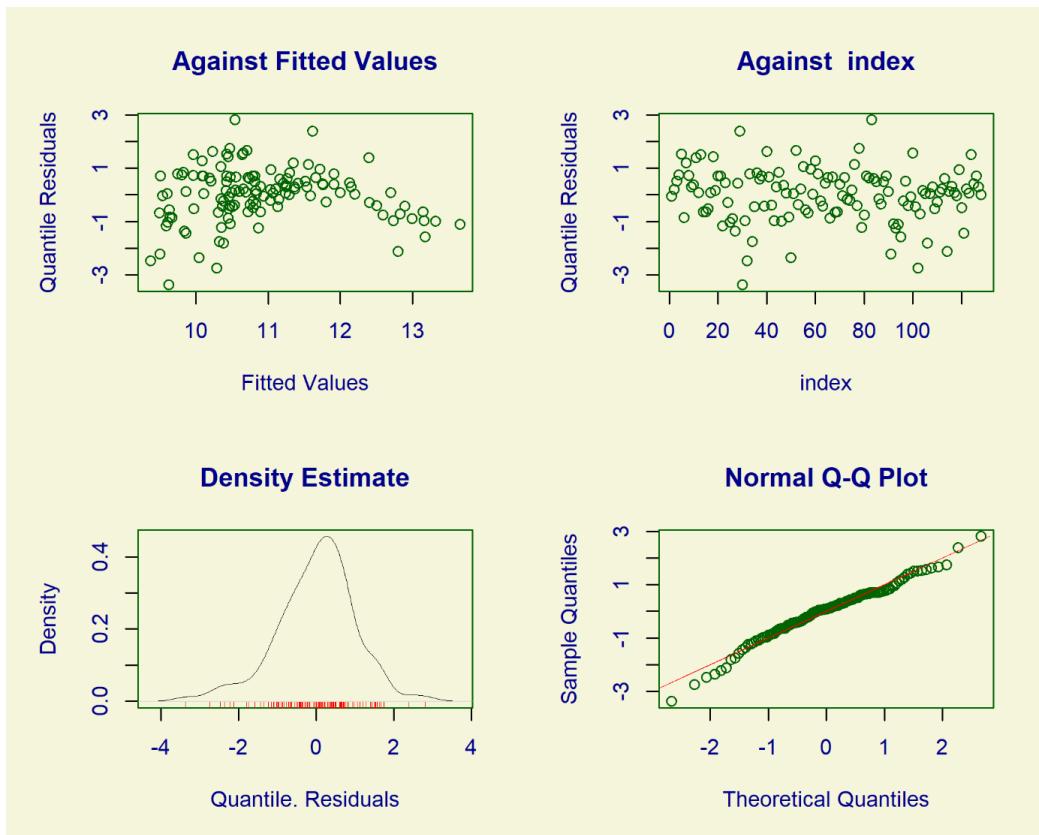


Abbildung 2: Modelldiagnostik des GAMLSS für Gesamtkosten (Finales Modell)

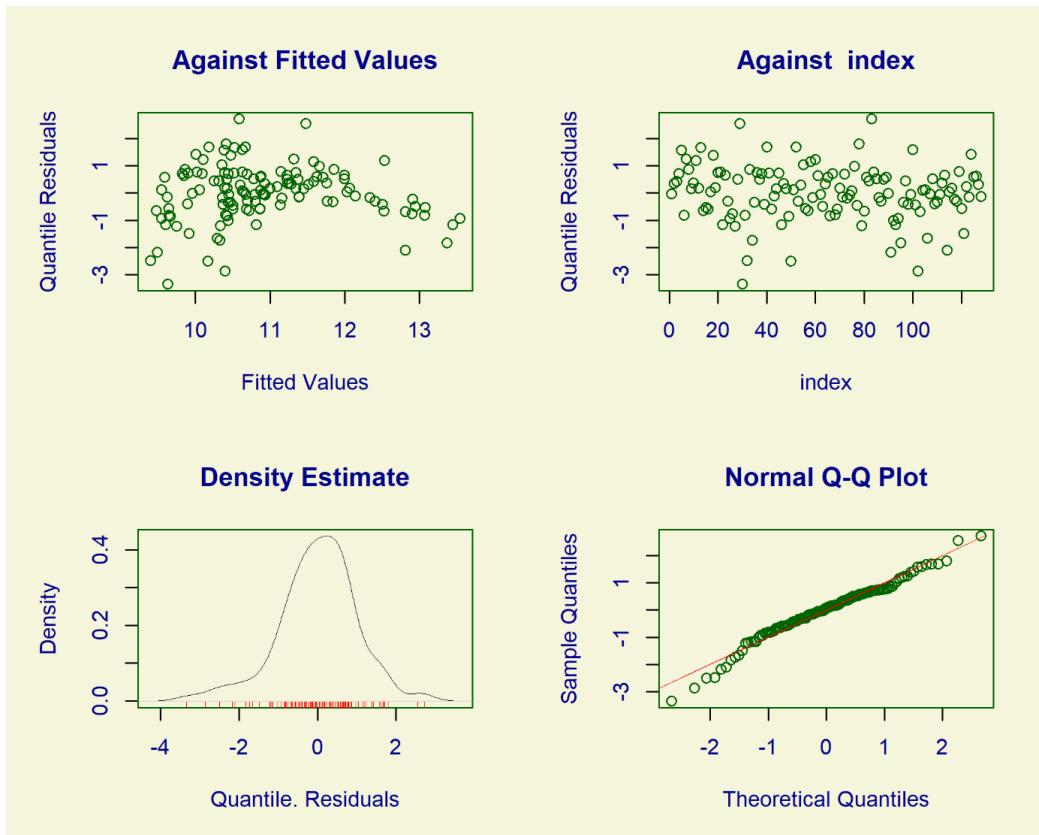


Tabelle 1: Berechnung inkrementelle Kosten und inkrementeller Effekt

	n	Total		Inkrementell	
		MW Todesfälle	MW Kosten	Todesfälle	Kosten
Interventionsgruppe	56	0,0536	98.218,43 €	0,1659	- 978,53 €
Kontrollgruppe	41	0,2195	99.196,95 €		

MW: Mittelwert, n: Stichprobengröße

Erläuterung verhinderte Todesfälle

Interventionsgruppe: 3 Todesfälle / 56 Patienten „at risk“ zu Tag 28 = 0,0535

Kontrollgruppe: 9 Todesfälle / 41 Patienten „at risk“ zu Tag 28 = 0,2195

Tabelle 2: Beantwortungsmodus EQ-5D-5L (T0)

Beantwortet durch... (n = 349)	n (%)
Selbstbeurteilung	119 (34,10)
Ehepartner	23 (6,59)
Ehepartner (telefonisch)	9 (2,58)
Kind	16 (4,58)
Kind (telefonisch)	9 (2,58)
Elternteil	5 (1,43)
Elternteil (telefonisch)	1 (0,29)
Sonstige Verwandschaft	8 (2,29)
Sonstige Verwandschaft (telefonisch)	4 (1,15)
Dritte: Pflegepersonal/Betreuungspersonal	148 (42,41)
Dritte: Pflegepersonal/Betreuungspersonal (telefonisch)	7 (2,01)

Tabelle 3: Ergebnisse der EQ-5D-5L Dimensionen

	Interventionsgruppe					Kontrollgruppe				
	1 – keine Probleme n (%)	2 – leichte Probleme n (%)	3 – mäßige Probleme n (%)	4 – große Probleme n (%)	5 – extreme Probleme n (%)	1 – keine Probleme n (%)	2 – leichte Probleme n (%)	3 – mäßige Probleme n (%)	4 – große Probleme n (%)	5 – extreme Probleme n (%)
<b>Beweglichkeit/ Mobilität</b>										
T0	14 (7,73)	14 (7,73)	17 (9,39)	32 (17,68)	104 (57,46)	18 (10,34)	9 (5,17)	8 (4,60)	22 (12,64)	117 (67,24)
T1	13 (19,40)	16 (23,88)	14 (20,90)	11 (16,42)	13 (19,40)	9 (15,00)	12 (20,00)	14 (23,33)	13 (21,67)	12 (20,00)
T2	17 (32,69)	14 (26,92)	4 (7,69)	10 (19,23)	7 (13,46)	14 (27,45)	13 (25,49)	12 (23,53)	9 (17,65)	3 (5,88)
<b>Für sich selbst sorgen</b>										
T0	25 (13,89)	12 (6,67)	14 (7,78)	32 (17,78)	97 (53,89)	22 (12,64)	4 (2,30)	17 (9,77)	25 (14,37)	106 (60,92)
T1	31 (45,59)	7 (10,29)	14 (20,59)	6 (8,82)	10 (14,71)	20 (33,33)	10 (16,67)	5 (8,33)	14 (23,33)	11 (18,33)
T2	25 (47,17)	11 (20,75)	7 (13,21)	4 (7,55)	6 (11,32)	26 (50,98)	9 (17,65)	7 (13,73)	6 (11,76)	3 (5,88)
<b>Alltägliche Tätigkeiten</b>										
T0	21 (11,60)	11 (6,08)	12 (6,63)	19 (10,50)	118 (65,19)	19 (10,98)	6 (3,47)	7 (4,05)	15 (8,67)	126 (72,83)
T1	7 (10,29)	14 (20,59)	19 (27,94)	14 (20,59)	14 (20,59)	7 (11,86)	7 (11,86)	15 (25,42)	11 (18,64)	19 (32,20)
T2	12 (22,64)	12 (22,64)	14 (26,41)	5 (9,43)	10 (18,87)	9 (17,65)	17 (33,33)	9 (17,65)	9 (17,65)	7 (13,73)
<b>Schmerz/ Körperliche Beschwerden</b>										
T0	56 (31,11)	32 (17,78)	51 (28,33)	31 (17,22)	10 (5,56)	59 (34,30)	30 (17,44)	40 (23,26)	28 (16,28)	15 (8,72)
T1	15 (22,39)	23 (34,33)	22 (32,84)	6 (8,96)	1 (1,49)	8 (13,56)	16 (27,12)	22 (37,29)	11 (18,64)	2 (3,39)
T2	10 (18,87)	19 (35,85)	13 (24,53)	7 (13,21)	4 (7,55)	10 (19,61)	19 (37,25)	15 (29,41)	7 (13,73)	0 (0,00)
<b>Angst/ Niedergeschlagenheit</b>										
T0	81 (45,25)	31 (17,32)	37 (20,67)	21 (11,73)	9 (5,03)	70 (40,70)	36 (20,93)	38 (22,09)	19 (11,05)	9 (5,23)
T1	34 (50,75)	12 (17,91)	15 (22,39)	6 (8,96)	0 (0,00)	24 (40,00)	12 (20,00)	9 (15,00)	12 (20,00)	3 (5,00)



Abbildung 3: Gesundheitsbezogene Lebensqualität – EQ-5D-5L Index (Gesamtkohorte)

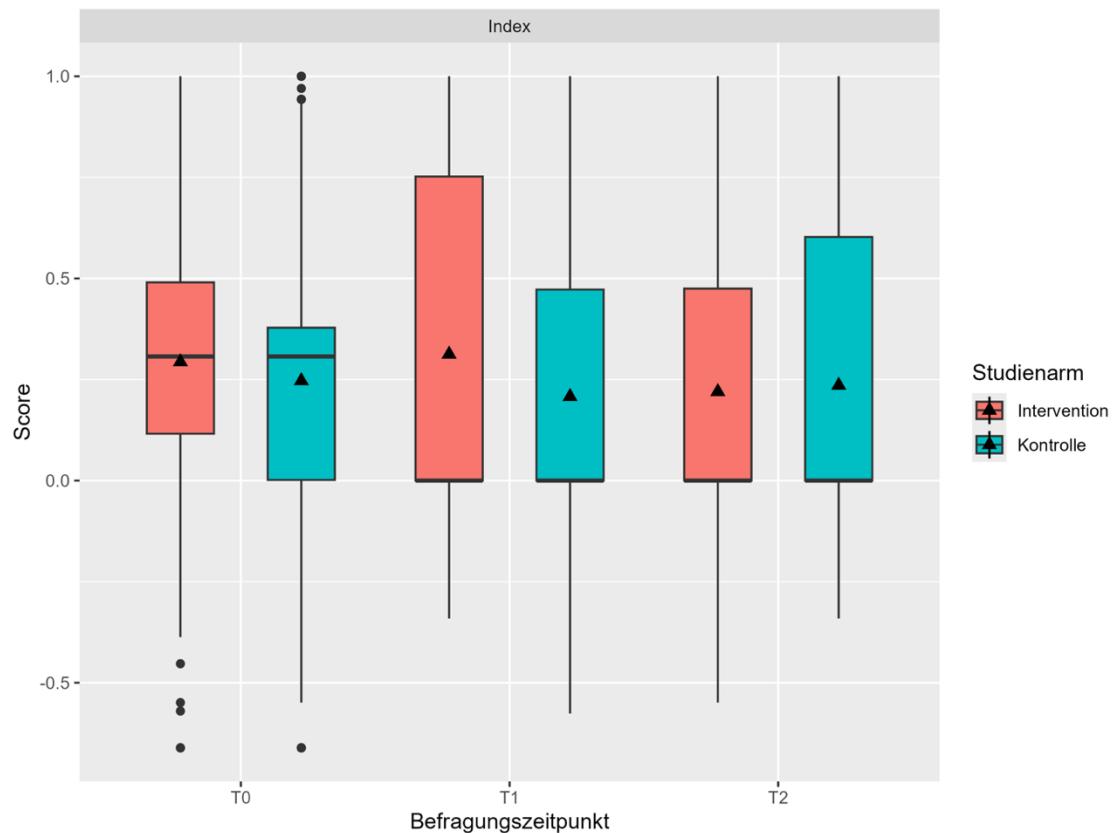


Abbildung 4: Gesundheitsbezogene Lebensqualität – EQ-5D-5L Index (GKV-Kohorte)

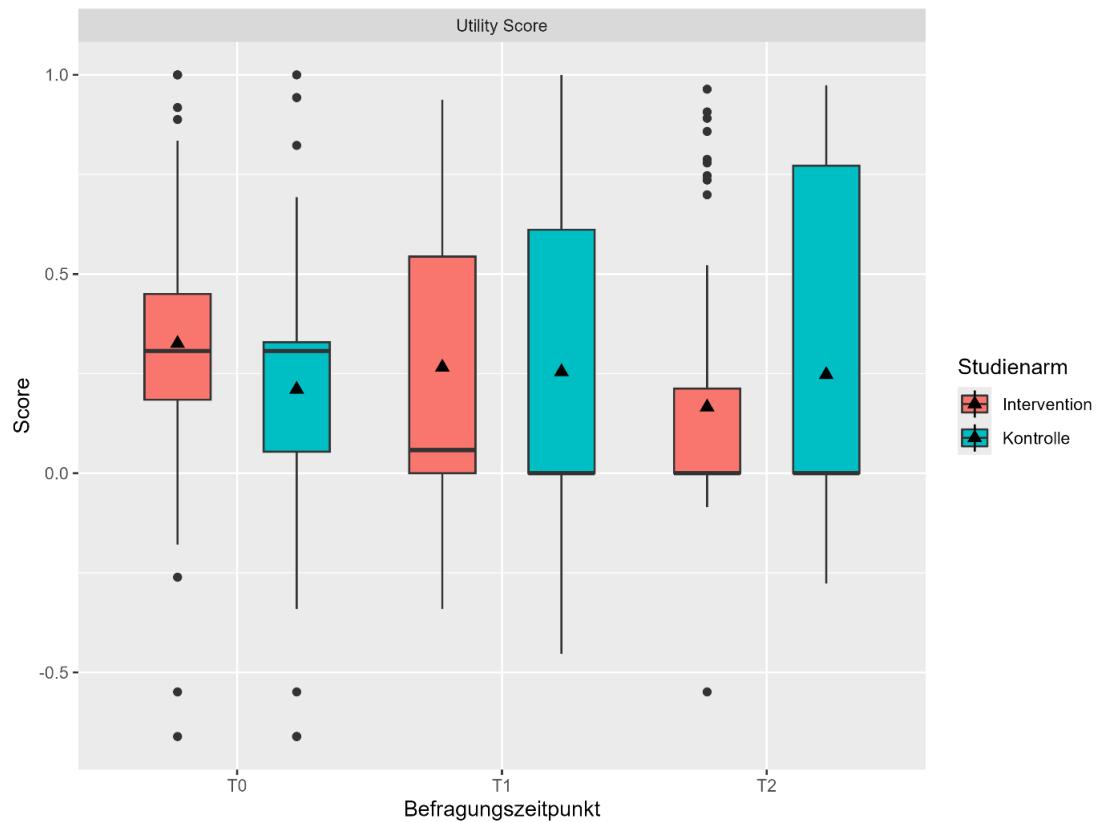


Tabelle 4: Gesundheitsbezogene Lebensqualität - EQ-5D-5L Index (Verbundene Stichprobe)

	Gesamtkohorte (n = 209)		GKV-Kohorte (n = 71)	
	Intervention (n = 102)	Kontrolle (n = 107)	Intervention (n = 42)	Kontrolle (n = 29)
<b>EQ-5D-5L Index</b>				
T0				
MW (SD)	0,315 (0,384)	0,287 (0,381)	0,390 (0,285)	0,279 (0,403)
Median (IQR)	0,307 (0,401)	0,307 (0,323)	0,307 (0,376)	0,307 (0,256)
T1				
MW (SD)	0,280 (0,387)	0,231 (0,382)	0,229 (0,349)	0,266 (0,415)
Median (IQR)	0 (0,667)	0 (0,595)	0 (0,516)	0 (0,636)
T2				
MW (SD)	0,211 (0,370)	0,242 (0,394)	0,181 (0,320)	0,230 (0,419)
Median (IQR)	0 (0,369)	0 (0,674)	0 (0,314)	0 (0,77)
IQR: Interquartilsabstand, MW: Mittelwert, n: Stichprobengröße, SD: Standardabweichung				

Tabelle 5: Mehrfaktorielle ANOVA mit Messwiederholungen - EQ-5D-5L Index

Effekt	Df	Summe der Quadrate	Mittlere Quadrate	F-Wert	p-Wert
<b>Gesamtkohorte</b>					
Gruppe	1	0,04	0,04	0,145	0,704
Residuen	207	50,52	0,24		
Zeit	1	0,56	0,56	4,085	0,045
Gruppe x Zeit	1	0,09	0,09	0,648	0,421
Residuen	207	28,58	0,14		
Residuen (within)	209	12,12	0,06		
<b>GKV-Kohorte</b>					
Gruppe	1	0,003	0,003	0,014	0,907
Residuen	69	16,52	0,24		
Zeit	1	0,74	0,74	7,434	0,008
Gruppe x Zeit	1	0,22	0,22	2,203	0,142
Residuen	69	6,86	0,10		
Residuen (within)	71	3,55	0,05		
<i>Df: Freiheitsgrade</i>					
<i>Signifikanz: *** p ≤ 0,001, ** p ≤ 0,01, * p ≤ 0,05</i>					

Abbildung 5: Gesundheitsbezogene Lebensqualität – EQ-5D-5L VAS (Gesamtkohorte)

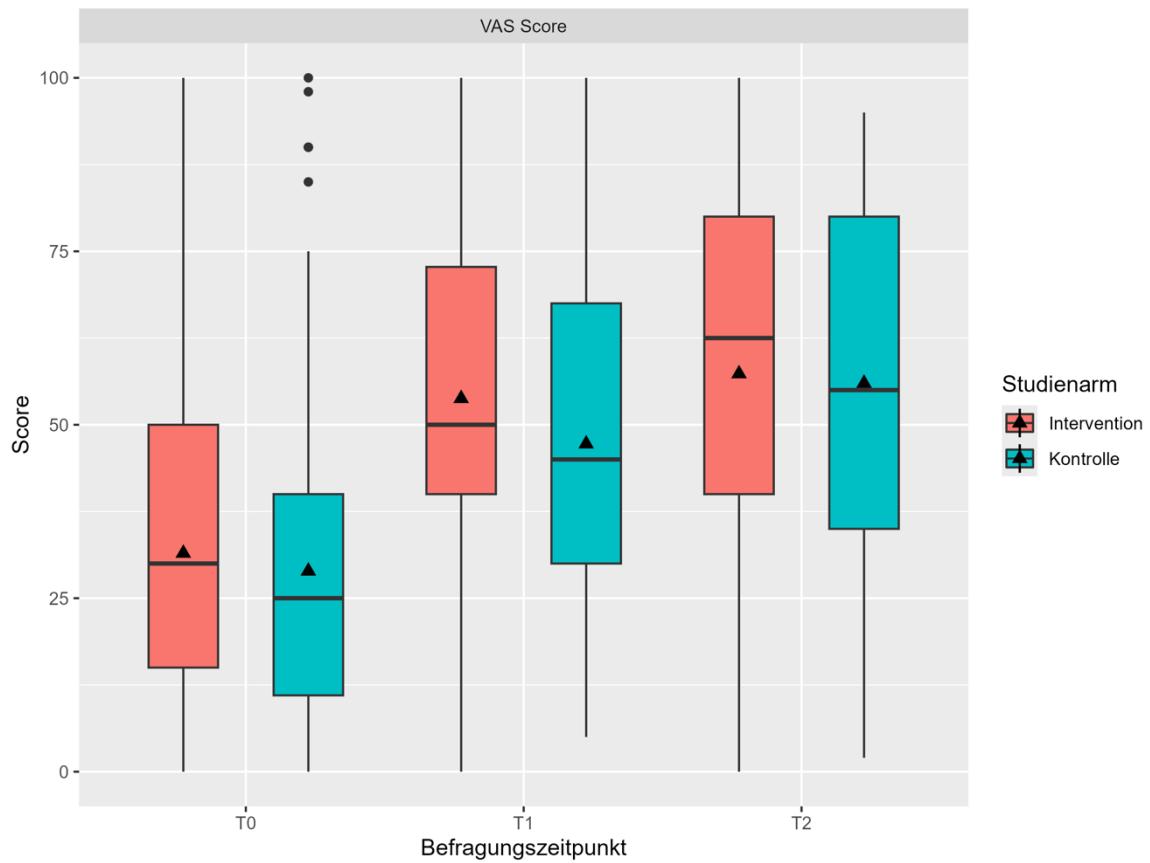


Abbildung 6: Gesundheitsbezogene Lebensqualität – EQ-5D-5L VAS (GKV-Kohorte)

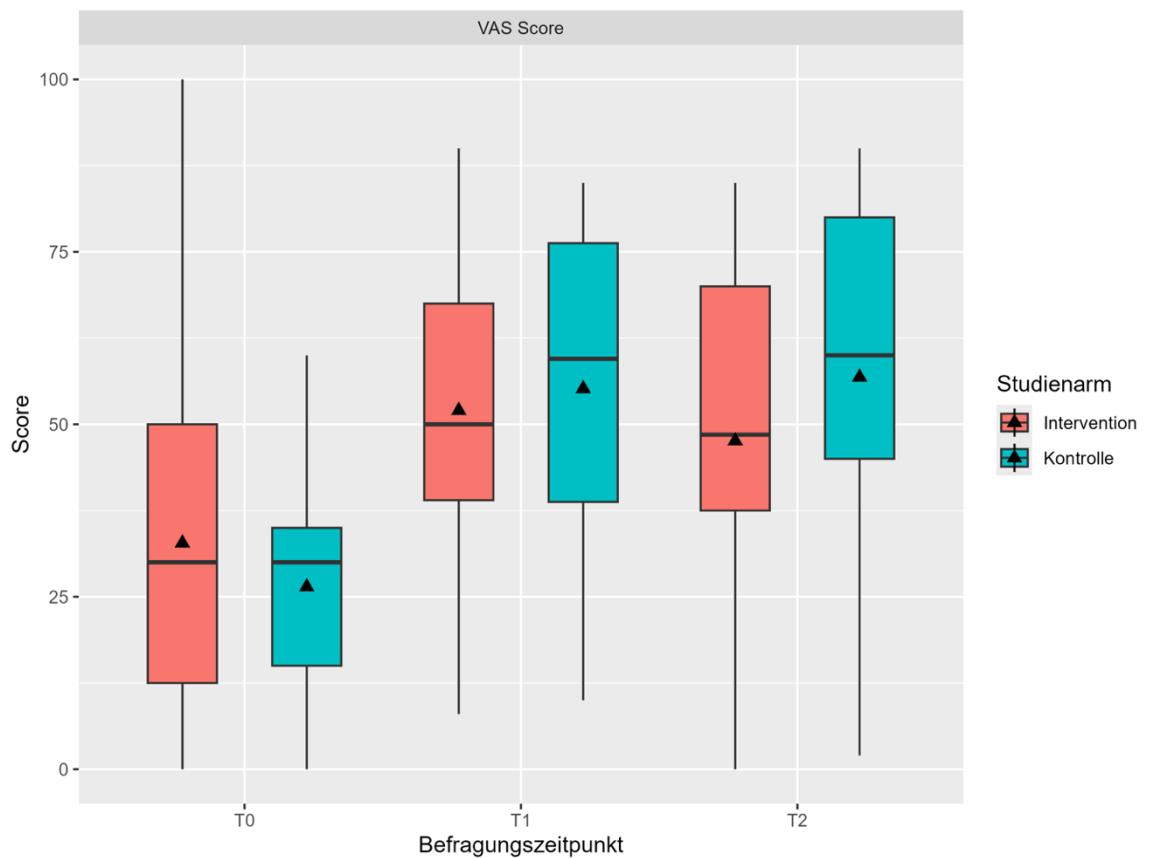


Tabelle 6: Gesundheitsbezogene Lebensqualität - EQ-5D-5L VAS (Verbundene Stichprobe)

	Gesamtkohorte		GKV-Kohorte	
	Intervention (n = 33)	Kontrolle (n = 38)	Intervention (n = 13)	Kontrolle (n = 8)
<b>EQ-5D-5L VAS</b>				
T0				
MW (SD)	43,73 (27,94)	32,05 (26,04)	49,85 (30,21)	32,38 (15,97)
Median (IQR)	35,00 (40,00)	27,50 (32,75)	50,00 (57,00)	32,5 (22,5)
T1				
MW (SD)	54,94 (21,74)	50,66 (22,71)	47,15 (18,31)	50,00 (30,00)
Median (IQR)	50,00 (30,00)	47,50 (40,25)	48,13 (27,12)	40,00 (40,00)
T2				
MW (SD)	60,36 (24,36)	56,52 (25,18)	49,38 (22,78)	49,00 (30,11)
Median (IQR)	65,00 (30,00)	60,00 (45,00)	50,00 (30,00)	50 (48,75)
IQR: Interquartilsabstand, MW: Mittelwert, n: Stichprobengröße, SD: Standardabweichung				

Tabelle 7: Mehrfaktorielle ANOVA mit Messwiederholungen - EQ-5D-5L VAS

Effekt	Df	Summe der Quadrate	Mittlere Quadrate	F-Wert	p-Wert
<b>Gesamtkohorte</b>					
Gruppe	1	2.307	2.306,50	2,578	0,113
Residuen	69	61.730	894,60		
Zeit	1	15.405	15.405	23,218	<b>&lt; 0,001***</b>
Gruppe x Zeit	1	542	542	0,818	0,369
Residuen	69	45.780	663		
Residuen (within)	71	20.514	288,9		
<b>GKV-Kohorte</b>					
Gruppe	1	471	470,6	0,464	0,504
Residuen	19	19.251	1.013,2		
Zeit	1	384	384,0	0,675	0,422
Gruppe x Zeit	1	723	722,9	1,271	0,274
Residuen	19	10.810	568,9		
Residuen (within)	21	4.772	227,2		
<i>Df: Freiheitsgrade</i>					
<i>Signifikanz: *** p ≤ 0,001, ** p ≤ 0,01, * p ≤ 0,05</i>					

Abbildung 7: Gesundheitsbezogene Lebensqualität – VR-36 (Gesamtkohorte)

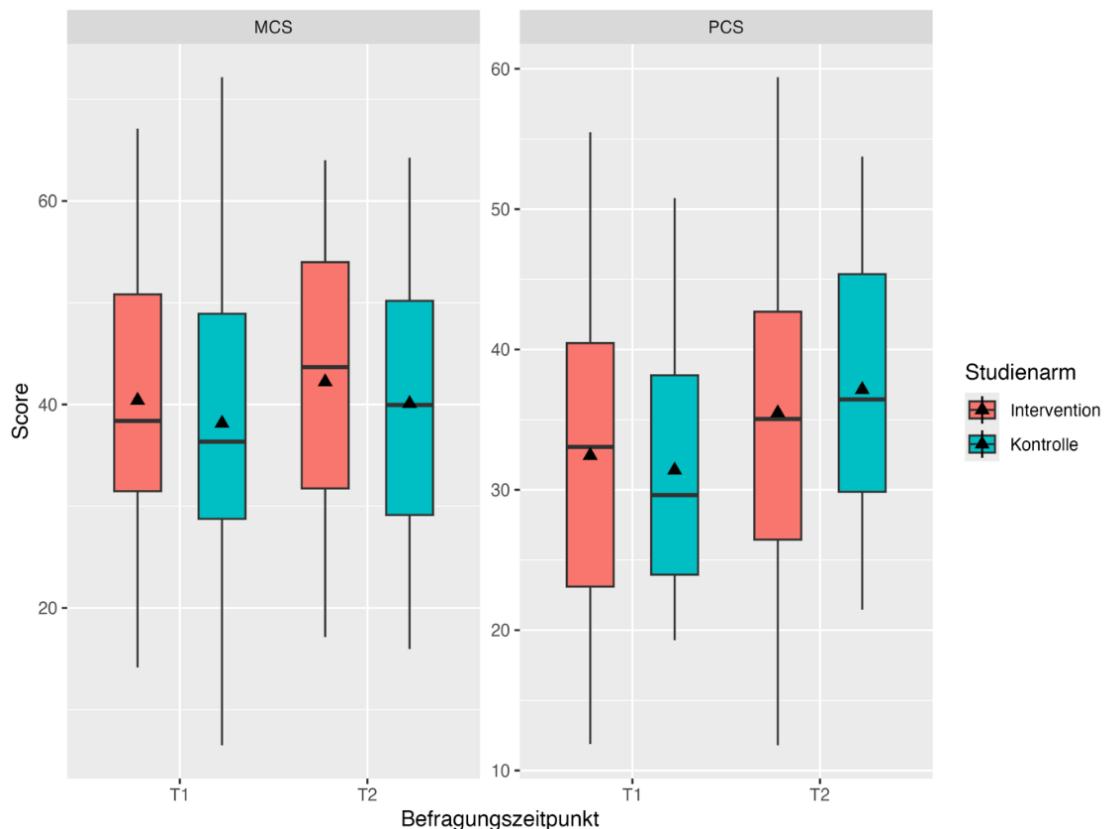


Abbildung 8: Gesundheitsbezogene Lebensqualität – VR-36 (GKV- Kohorte)

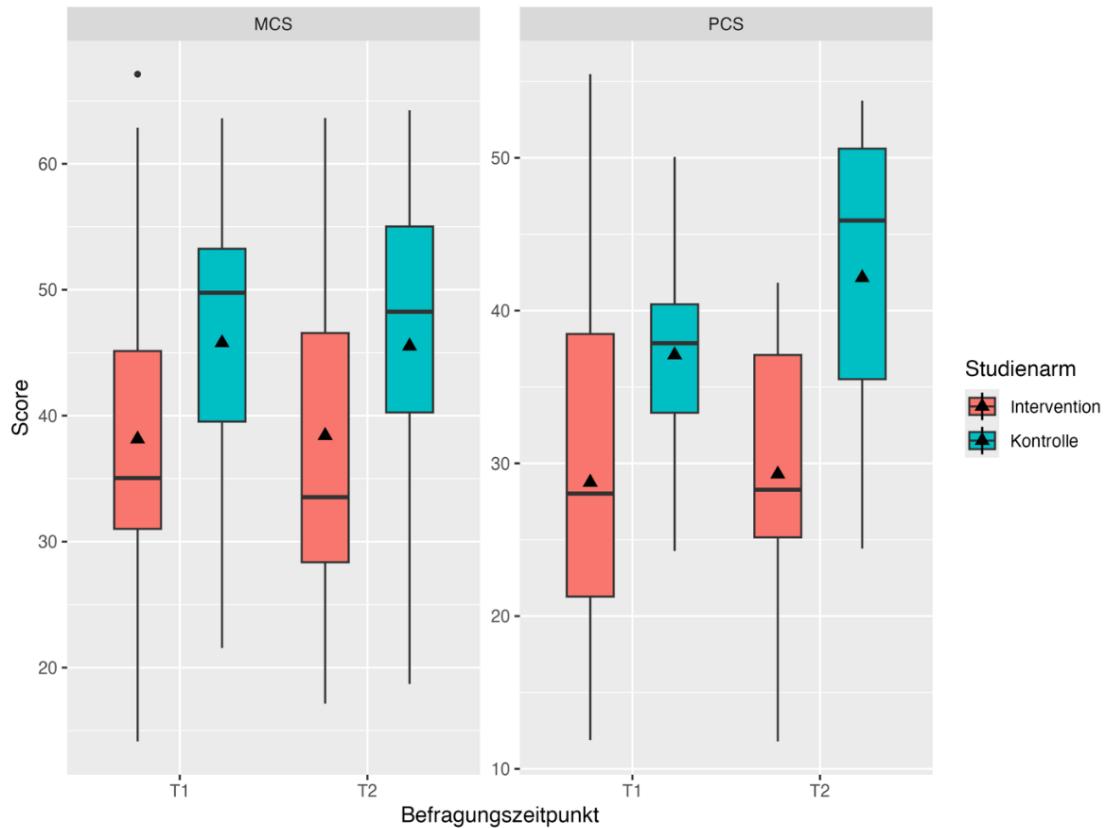


Abbildung 9: Gesundheitsbezogene Lebensqualität – VR-36 (Gesamtkohorte, verbundene Stichprobe)

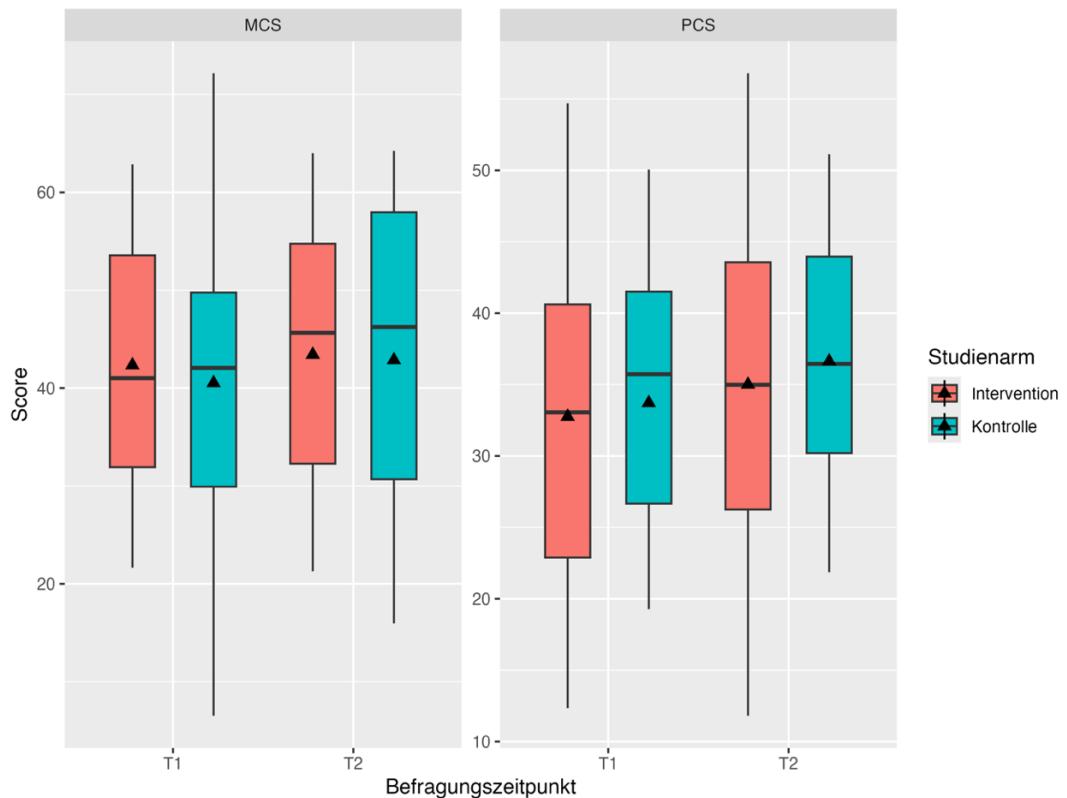


Abbildung 10: Gesundheitsbezogene Lebensqualität – VR-36 (GKV-Kohorte, verbundene Stichprobe)

