

Anlage 1: Unabhängiger Evaluationsbericht

Evaluationsbericht

(gemäß Nr. 14.1 ANBest-IF)



| | |
|--|---|
| Konsortialführung: | Klinik für Anästhesiologie m. S. operative Intensivmedizin, Charité – Universitätsmedizin Berlin |
| Förderkennzeichen: | 01NVF16011 |
| Akronym: | ERIC |
| Projekttitel: | Enhanced Recovery after Intensive Care |
| ClinicalTrials.gov | NCT03671447 |
| Autoren – Konsortialpartner für die Evaluation: | <u>Evaluierendes Institut:</u> IBE – Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie der Ludwig-Maximilians-Universität München Lehrstuhl für Biometrie und Bioinformatik Marchioninstr. 15, 81377 München Prof. Dr. Ulrich Mansmann (Verantwortlicher Konsortialpartner) Dr. Christine Adrion, MPH <u>Verantwortlich für die gesundheitsökonomische Evaluation/ Akzeptanzerhebung:</u> Technische Universität Berlin Fachgebiet Management im Gesundheitswesen, Fakultät Wirtschaft und Management Straße des 17. Juni 135, 10623 Berlin Prof. Dr. Reinhard Busse (Leitung) Elke Berger, MPH Helene Eckhardt, MSc |
| Förderzeitraum: | 1. Mai 2017 – 31. Dezember 2020 |

Hinweise:

Aus Gründen der besseren Lesbarkeit wird auf die gleichzeitige Verwendung der Sprachformen männlich, weiblich und divers (m/w/d) verzichtet. Sämtliche Personenbezeichnungen gelten gleichermaßen für alle Geschlechter.

Inhaltsverzeichnis

| | |
|---|----|
| Tabellen..... | 6 |
| Abbildungen..... | 8 |
| Abkürzungen und Definitionen | 10 |
| 0 Einleitung & Protokoll-Information | 11 |
| 1 Ausgangslage und Ziele | 12 |
| 1.1 Ausgangslage..... | 12 |
| 1.2 Ziele, Fragestellungen und Hypothesen..... | 12 |
| 2 Darstellung der neuen Versorgungsform | 14 |
| 2.1 Kontrollphase: Regelversorgung mit Case-Care-Management | 15 |
| 2.2 Trainingsphase und Interventionsphase (neue Versorgungsform mit Case-Care-Management) | 16 |
| 3 Darstellung des Evaluationsdesigns | 19 |
| 3.1 Primäre Fragestellung und Arbeitshypothese | 19 |
| 3.2 Studiendesign und Setting..... | 19 |
| 3.3 Ein- und Ausschlusskriterien – Clusterebene und Patientenebene | 20 |
| 3.4 Randomisierung und Verblindung | 21 |
| 3.4.1 <i>Randomisierung</i> | 21 |
| 3.4.2 <i>Verblindungsgrad</i> | 22 |
| 3.5 Definierte Endpunkte (ICU-Phase und post-ICU Follow-up)..... | 23 |
| 3.5.1 <i>Primäre Outcomes</i> | 23 |
| 3.5.2 <i>Sekundäre Outcomes</i> | 25 |
| 3.6 Datenerhebung und Datenschutzaspekte: Fragebögen und Case Report Forms..... | 28 |
| 3.6.1 <i>E-Health-Plattform und REDCap-Forschungsdatenbank</i> | 28 |
| 3.6.2 <i>BARMER Wissenschafts-Data Warehouse (W-DWH) [TU Berlin]</i> | 29 |
| 3.6.3 <i>Hausarztfragebogen [TU Berlin]</i> | 29 |
| 3.7 Datenmanagement, Qualitätskontrolle und Datensicherheit..... | 30 |
| 3.7.1 <i>REDCap-basierte Datenbanken für ICU- sowie FU-Phase</i> | 30 |
| 3.7.2 <i>Datenzugriff für die Evaluation (Clinical Effectiveness)</i> | 31 |
| 3.8 Effektevaluation: Statistische Methoden | 32 |
| 3.8.1 <i>Initiale Fallzahlplanung</i> | 32 |
| 3.8.2 <i>Analyse-Strategien zur Clinical Effectiveness</i> | 33 |
| 3.8.2.1 <i>Konfirmatorische Analysen der primären Outcomes</i> | 34 |
| 3.8.2.2 <i>Explorative Analysen der sekundären Outcomes</i> | 36 |
| 3.9 Prozessevaluation | 37 |
| 3.10 Gesundheitsökonomische Evaluation – Methoden [TU Berlin]..... | 37 |
| 3.10.1 <i>Gesundheitsökonomisches Studiendesign und Perspektive</i> | 37 |
| 3.10.2 <i>Beschreibung von berücksichtigten Kosten und Effekten</i> | 37 |

| | | |
|----------|---|----|
| 3.10.2.1 | Kosten..... | 37 |
| 3.10.2.2 | Effekte..... | 41 |
| 3.10.3 | <i>Inkrementelle Kosteneffektivitätsrelation (IKER)</i> | 41 |
| 3.10.4 | <i>Diskontierung, Sensitivitätsanalysen und Umgang mit fehlenden Werten</i> | 41 |
| 3.10.5 | <i>Statistische Methoden (gesundheitsökonomische Evaluation)</i> | 42 |
| 3.10.5.1 | Kosten..... | 42 |
| 3.10.5.2 | Effekte..... | 42 |
| 3.11 | Akzeptanzerhebung – Methoden [TU Berlin] | 43 |
| 4 | Ergebnisse der Evaluation – ICU Phase | 44 |
| 4.1 | Teilnehmende Zentren und Patienten: grundlegende Informationen..... | 44 |
| 4.2 | Implementierung der Intervention: Geplantes vs. tatsächliches Datum für Treatment Switching 44 | |
| 4.3 | CONSORT-Flussdiagramm..... | 47 |
| 4.4 | Basisinformationen und Baseline Patienten-Charakteristika | 48 |
| 4.4.1 | <i>Basisinformationen, QI-Visitentage, und Versichertenstatus</i> | 48 |
| a) | Basisinformationen | 48 |
| b) | Versichertenstatus (BARMER vs. non-BARMER) | 49 |
| 4.4.2 | <i>Wechselsituationen</i> | 49 |
| a) | Wechsel des Clusters beim 2. ICU-Aufenthalt..... | 49 |
| b) | Crossover-Patienten vs. Wechsel-Patienten | 49 |
| 4.5 | Zentrums-spezifische Implementierung der Intervention und Rekrutierungsübersicht..... | 50 |
| 4.6 | Datenmanagement-Aspekte und Datenqualität | 52 |
| 4.6.1 | <i>Design-Variablen</i> | 52 |
| 4.6.2 | <i>Baseline Patienten-Charakteristika, weitere Variablen</i> | 52 |
| 4.6.3 | <i>Missing Baseline-Werte (wegen verspätetem Patienten-Einschluss)</i> | 52 |
| 4.7 | Deskriptive Analysen | 53 |
| 4.8 | Deskriptive Analysen des primären Efficacy-Endpunkts – pro QI..... | 55 |
| 4.8.1 | <i>QI 1 – Tägliche multiprofessionelle und interdisziplinäre klinische Visite mit Dokumentation von Tageszielen</i> | 55 |
| 4.8.2 | <i>QI 2 – Management von Sedierung, Analgesie & Delir</i> | 56 |
| 4.8.3 | <i>QI 3 – Patientenadaptierte Beatmung</i> | 57 |
| 4.8.4 | <i>QI 4 – Frühzeitige Entwöhnung von einer invasiven Beatmung (Weaning)</i> | 58 |
| 4.8.5 | <i>QI 6 – Maßnahmen zum Infektionsmanagement</i> | 59 |
| 4.8.6 | <i>QI 7 – Frühe enterale Ernährung</i> | 60 |
| 4.8.7 | <i>QI 8 – Dokumentation einer strukturierten Patienten- & Angehörigenkommunikation</i> | 61 |
| 4.8.8 | <i>QI 9 – Frühmobilisation</i> | 62 |
| 4.9 | Modellbasierte konfirmatorische Analysen für die einzelnen QIs..... | 63 |
| 4.9.1 | <i>QI 1 – Tägliche multiprofessionelle und interdisziplinäre klinische Visite mit Dokumentation von Tageszielen</i> | 66 |
| 4.9.2 | <i>QI 2 – Management von Sedierung, Analgesie & Delir</i> | 68 |

| | | |
|---------|---|-----|
| 4.9.3 | QI 3 – Patientenadaptierte Beatmung | 70 |
| 4.9.4 | QI 4 – Frühzeitige Entwöhnung von einer invasiven Beatmung (Weaning)..... | 71 |
| 4.9.5 | QI 6 – Maßnahmen zum Infektionsmanagement | 73 |
| 4.9.6 | QI 7 – Frühe enterale Ernährung..... | 74 |
| 4.9.7 | QI 8 – Dokumentation einer strukturierten Patienten- & Angehörigenkommunikation..... | 76 |
| 4.9.8 | QI 9 – Frühmobilisation | 77 |
| 4.10 | Zusammenfassende Bewertung der Ergebnisse zu den konfirmatorischen QI-Auswertungen | 79 |
| 5 | Ergebnisse der Evaluation – Post-ICU Follow-up Phase | 82 |
| 5.1 | Basisinformationen | 82 |
| 5.1.1 | Aspekte zur Datenvollständigkeit bzw. Fallzahlen im Follow-up..... | 82 |
| 5.1.2 | Nacherhebung | 82 |
| 5.1.3 | Datenmanagement-Aspekte: "Windowing" der FU-Visiten, Problem der zeitlichen Zuordnung | 82 |
| 5.2 | Analysen zu sekundären Endpunkten | 83 |
| 5.2.1 | Gesamt-Mortalität zu Monat 6 | 83 |
| 5.2.2 | Aktueller Gesundheitszustand anhand der Visuellen Analogskala (VAS) des EQ-5D-5L (FU 1 und FU 2)..... | 84 |
| 5.2.3 | Psychische Gesundheit – Angst und Depression: PHQ Fragebogen | 85 |
| 5.2.3.1 | Subdomäne Depression: Patient Health Questionnaire–2 (PHQ-2)..... | 85 |
| 5.2.3.2 | Subdomäne Angst: Generalized Anxiety Disorder Scale–2 (GAD-2)..... | 86 |
| 5.2.4 | Kognition – Gedächtnis, visuospatiale und visuostrukturelle Fähigkeiten | 86 |
| 5.2.4.1 | Mini-Cog..... | 86 |
| 5.2.5 | Körperliche Fähigkeiten..... | 86 |
| 5.2.5.1 | Subdomäne Mobilität: Timed-Up-and-Go (TUG)-Test | 86 |
| 5.2.5.2 | Subdomäne Muskelstärke: Hand-Grip-Strength (HGS)-Test | 87 |
| 6 | Ergebnisse der Gesundheitsökonomischen Evaluation [TU Berlin]..... | 88 |
| 6.1 | Gesundheitsökonomische Evaluation | 88 |
| 6.1.1 | Datenmanagement und Plausibilitätsanalysen | 88 |
| 6.1.2 | Deskriptive Analysen | 91 |
| 6.1.2.1 | Kosten..... | 91 |
| 6.1.2.2 | Effekte..... | 93 |
| 6.1.3 | Inferenzstatistische Analysen..... | 95 |
| 6.1.3.1 | Kosten..... | 95 |
| 6.1.3.2 | Effekte..... | 95 |
| 6.1.4 | Inkrementelle Kosteneffektivitätsrelation (IKER)..... | 95 |
| 7 | Ergebnisse der Akzeptanzerhebung [TU Berlin] | 96 |
| 8 | Schlussfolgerungen und Empfehlungen des Evaluators..... | 103 |
| 8.1 | Perspektive Wirksamkeitsevaluation | 103 |
| 8.2 | Perspektive Gesundheitsökonomische Evaluation [TU Berlin] | 105 |
| 8.3 | Perspektive der Akzeptanzerhebung [TU Berlin] | 106 |

| | | |
|--------|---|-----|
| 9 | Anhang | 107 |
| A.1 | REDCap Datenbanken: QI-Daten; FU-Daten..... | 108 |
| A.2 | SAPS II und SOFA Score..... | 108 |
| A.3 | PICS-bezogene sekundäre Outcomes: Auswertungsalgorithmen..... | 109 |
| A.3.1. | Subdomäne Mobilität: Timed-Up-and-Go (TUG)-Test | 109 |
| A.3.2. | Subdomäne Handkraft: Hand-Grip-Strength (HGS)-Test | 109 |
| A.4 | Ergänzende Tabellen & Abbildungen zur ICU-Phase | 110 |
| A.4.1. | Ergänzende Abbildung zur Rekrutierung | 110 |
| A.4.2. | QI-Auswertungen: Schätzer der logistischen Mixed Effects-Modelle (fixed effects und Varianzkomponenten) – Hauptmodell PM..... | 111 |
| A.4.3. | QI-Auswertungen: Schätzer der logistischen Mixed Effects-Modelle (fixed effects und Varianzkomponenten) – Modell SM.5 | 114 |
| A.4.4. | Ergänzende Abbildung & Tabellen zum Post-ICU Follow-up..... | 115 |
| A.5 | Technische Details – Effektevaluation | 119 |
| A.6 | Reporting Konventionen | 119 |
| B.1 | Anhang: Kostenkalkulation für die Interventionskosten [TU Berlin] | 120 |
| B.2 | Technische Details – gesundheitsökonomische Evaluation [TU Berlin] | 123 |
| B.3 | Inferenzstatistische Analysen und Sensitivitätsanalysen – gesundheitsökonomische Evaluation [TU Berlin]..... | 124 |
| | LITERATURVERZEICHNIS | 128 |

Tabellen

| | |
|---|-----|
| Tabelle 1. Offizielle Randomisierungsliste mit zugewiesener Sequenzgruppe je Cluster (zusammen mit Tabelle 0)..... | 22 |
| Tabelle 2. Konsensus-basiertes Set der 10 Qualitätsindikatoren der DIVI (3. Auflage 2017, siehe [5]), welches zur Definition der 8 binären primären Outcomes verwendet wurde..... | 24 |
| Tabelle 3. Sekundäre und explorative Wirksamkeits-Endpunkte..... | 26 |
| Tabelle 4. Gesundheitsökonomische Evaluation: Bewertung der berücksichtigten Kostenkategorien. | 39 |
| Tabelle 5. Cluster-spezifische Datumsangaben für 3-monatige Trainings-Periode (inkl. kurze Implementierungsphase) sowie Start der Intervention (Crossover-Datum für Treatment Switch)..... | 46 |
| Tabelle 6. Deskriptive Zusammenfassung für die Anzahl der QI-Tage pro Aufenthalt (zentrumsübergreifend, sowie stratifiziert nach Cluster)..... | 48 |
| Tabelle 7. "Table 1" – Baseline Charakteristika. Klinische und demographische Charakteristika der ITT-Population (1462 Patienten) zu Baseline oder am Tag der 1. QI Visite..... | 53 |
| Tabelle 8. Deskription zum Entlassungsstatus bzgl. <i>Index</i> ICU-Aufenthalt in der ITT-Population, stratifiziert nach Behandlungsbedingung (Kontrolle vs. Intervention)..... | 54 |
| Tabelle 9. Schätzer (fixed effects, mit 95% KIs) des primären Mixed Effects Modells (PM), für die einzelnen QIs. Ergebnisse auf logit-Skala..... | 64 |
| Tabelle 10. Schätzer (fixed effects, mit 95% KIs) des primären Mixed Effects Modells (PM), für die einzelnen QIs. Dargestellt sind die Odds Ratios (ORs) für die Adhärenz unter Interventions- vs. unter Kontrollbedingung (Routineversorgung)..... | 64 |
| Tabelle 11. QI 1 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen)... | 66 |
| Tabelle 12. QI 2 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen)... | 68 |
| Tabelle 13. QI 3 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen)... | 70 |
| Tabelle 14. QI 4 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen)... | 71 |
| Tabelle 15. QI 6 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen)... | 73 |
| Tabelle 16. QI 7 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen)... | 74 |
| Tabelle 17. QI 8 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen)... | 76 |
| Tabelle 18. QI 9 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen)... | 77 |
| Tabelle 19. Zeitliche Abweichung zwischen FU-Zeitpunkt und Kostenerhebung (in Tagen)..... | 88 |
| Tabelle 20. Beobachtete Kosten für den Indexaufenthalt (EUR) | 91 |
| Tabelle 21. Übersicht der beobachteten Kosten im HA-CRF 1 und HA-CRF 2 (EUR)..... | 92 |
| Tabelle 22. Gesamtkosten (EUR) | 93 |
| Tabelle 23. Deskriptive Beschreibung der QALYs für 6 Monate..... | 93 |
| Tabelle 24. Deskriptive Beschreibung des Endpunkts außerklinische Beatmung..... | 94 |
| Tabelle 25. Deskriptive Beschreibung des Endpunkts außerklinische Beatmung als beatmet entlassener Patienten..... | 94 |
| Tabelle 26. EQ VAS: Ergebnisse des linearen Mixed Effects Modells, adjustiert für Behandlung (Intervention vs. Kontrolle), kategorisierten SAPS II, Zeit seit Entlassung aus der Index-ICU. | 115 |
| Tabelle 27. PHQ-2 Score: Ergebnisse des linearen Mixed Effects Modells. | 116 |
| Tabelle 28. GAD-2 Score: Ergebnisse des linearen Mixed Effects Modells. | 116 |

| | |
|--|-----|
| Tabelle 29. Mini-Cog – binäres Assessment: Ergebnisse des binären Mixed Effects Modells. | 117 |
| Tabelle 30. TUG – binäres Assessment: Ergebnisse des binären Mixed Effects Modells. | 117 |
| Tabelle 31. HGS – binäres Assessment: Ergebnisse des binären Mixed Effects Modells. | 118 |
| Tabelle 32. Kosten: Ergebnisse der linearen gemischten Modelle mit logarithmierten Kosten. | 124 |
| Tabelle 33. Kosten: Ergebnisse der generalisierten linearen Modelle unter Annahme der Gammaverteilung als Sensitivitätsanalyse. | 124 |
| Tabelle 34. Kosten: Ergebnisse der generalisierten linearen Modelle unter Annahme der Weibullverteilung als Sensitivitätsanalyse. | 125 |
| Tabelle 35. Kosten des Indexaufenthalts: Ergebnisse der linearen gemischten Modelle mit logarithmierten Kosten als Sensitivitätsanalyse. | 125 |
| Tabelle 36. Kosten des Indexaufenthalts: Ergebnisse des generalisierten linearen Modells unter Annahme der Gammaverteilung als Sensitivitätsanalyse. | 126 |
| Tabelle 37. Kosten des Indexaufenthalts: Ergebnisse des generalisierten linearen Modells unter Annahme der Weibullverteilung als Sensitivitätsanalyse. | 126 |
| Tabelle 38. Qualitätsadjustierte Lebensjahre: Ergebnisse der linearen gemischten Modelle. | 127 |
| Tabelle 39. Außerklinische Beatmung: Ergebnisse der generalisierten gemischten Logit-Modelle. | 127 |

Abbildungen

| | |
|---|----|
| Abbildung 1. Schematischer Aufbau der neuen Versorgungsform ERIC (eigene Darstellung) während der ICU-Phase. Prozess des Telekonsils bzw. der Televisiten bis zur Ableitung der QI-bezogenen primären Outcomes für die Evaluation. | 14 |
| Abbildung 2. Überblick zu Ablauf und Datenerhebung drei und/oder sechs Monaten nach Entlassung aus der Intensivstation (modifizierte Darstellung anhand Spies <i>et al.</i> [7]). PHQ: Patient Health Questionnaire; EQ-5D-5L: EuroQol 5 Dimensions 5 Level; WHODAS: World Health Organization Disability Assessment Schedule; IES-R: Impact of Event Scale-Revised; mMRC: modified Medical Research Council Score. | 16 |
| Abbildung 3. Foto des im Rahmen der Interventionsphase eingesetzten Telemedizinergäts ("Visitenwagen") Typ LITE V3 (Hersteller: InTouch Health, jetzt Teladoc Health, Inc.); technische Spezifikationen siehe Herstellerbroschüre. Mikrophon: mono – direktional (hyper-cardioid) 10Hz – 40Hz, zwei 60W Mono-Lautsprecher; 15" LCD-Monitor, Schwenkbereich des Kopfteils: horizontal +/- 170°, vertikal +27°/-65° maximal; Videokamera; eigener Handhörer, Ablage, Griff, Erweiterungsfach, Netzstecker und -kabel, Radschloss und Korb. | 18 |
| Abbildung 4. Schematische Darstellung des Datenflusses zwischen Remote ICU (teilnehmende Krankenhäuser) und Tele-ICU der Charité aus informatorischer Perspektive. Anhand der dokumentierten patientenindividuellen QI- Rohdaten (festgelegte Key Performance Merkmale) erfolgte eine zentrale Bewertung und Ableitung der Adhärenz bzgl. der 8 QIs durch die Charité. | 28 |
| Abbildung 5. Relevante Kostenarten im Beobachtungszeitraum. | 37 |
| Abbildung 6. Schematische Illustration des im ERIC Projekt angewandten gestaffelten (sequentiellen) Rollout-Prozesses und zeitlichen Ablaufs (nach der Entscheidung, die Rekrutierung zu verlängern, inklusive der Verschiebung des dritten Crossover-Datums (ursprüngliches Datum dargestellt als vertikale gestrichelte Linie)). [In dieser Abbildung entsprechen die Nummern des Clusters den in diesem Bericht verwendeten Cluster-IDs. Die Höhe der Cluster entspricht der tatsächlichen ClustergroÙe]. | 45 |
| Abbildung 7. CONSORT Flow Chart. Der Fluss der Cluster und Patienten im Zeitverlauf der ERIC Stepped Wedge Cluster-randomisierten Studie. Nullpunkt für die Fallzahlen bzgl. ICU-Überlebende: Zeit seit Index-ICU-Entlassung. * Ein ICU-Todesfall war ein Wechselpatient, d.h. der Index-ICU-Aufenthalt erfolgte unter Kontroll-Bedingung, Patient beim 2. Aufenthalt unter Intervention verstorben. IQR: Interquartilsabstand. | 47 |
| Abbildung 8. Deskriptive Darstellung der individuellen Beobachtungsdauer aller 1462 Patienten (1553 ICU Aufenthalte) mit mindestens 1 QI Visite während ihres ICU Aufenthalts. Zuordnung der Patienten zum Cluster per Index-ICU-Aufenthalt. Verläufe in grün: QI-Beobachtung unter Kontroll-Bedingung; Verläufe in orange: QI-Beobachtung unter Interventions-Bedingung. Gelbe senkrechte Linien: Trainings-Periode. Schwarze gestrichelte senkrechte Linie: Start der Intervention (Telemedizin). S.Grp = Sequenzgruppe. | 50 |
| Abbildung 9. Heatmap für die Cluster-spezifische Rekrutierungsrate (Fall-Ebene) pro Kalendermonat, im Verlauf der 19-monatigen Rekrutierungsphase. Je dunkler die Farbe, desto mehr Patienten wurden in diesem Kalendermonat in die Studie eingeschlossen. Fehlende Angaben zur Anzahl der Patienten in einem bestimmten Monat bedeuten, dass in diesem Zeitraum keine Patienten rekrutiert wurden. | 51 |
| Abbildung 10. QI 1: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I). | 55 |
| Abbildung 11. QI 2: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I). | 56 |
| Abbildung 12. QI 3: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I). | 57 |

| | |
|---|-----|
| Abbildung 13. QI 4: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I)..... | 58 |
| Abbildung 14. QI 6: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I)..... | 59 |
| Abbildung 15. QI 7: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I)..... | 60 |
| Abbildung 16. QI 8: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I)..... | 61 |
| Abbildung 17. QI 9: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I)..... | 62 |
| Abbildung 18. Graphische Darstellung der Ergebnisse des primären Modells pro QI: Effekt der Intervention (zugrundeliegenden Annahme: "Step Change") und deren Wechselwirkung mit der Cluster-spezifischen Zeit auf die Adhärenz-Wahrscheinlichkeit im Zeitverlauf, d.h. vor vs. nach Start der Intervention..... | 81 |
| Abbildung 19. Kumulative Inzidenz für die interessierende Variable Behandlungsbedingung (Intervention vs. Kontrolle). Für diese graphische Darstellung wurde der mittlere Baseline SOFA- und mittlere SAPS II-Score zugrunde gelegt. Jede Kurve stellt die erwartete kumulative Inzidenz für die beiden (unbalancierten) Subpopulationen dar (marginaler Effekt der Behandlungsbedingung adjustiert für die beiden Scores für Erkrankungsschwere), geschätzt anhand des Cox-Frailty Modells. | 84 |
| Abbildung 20. EQ VAS für die beiden FU-Zeitpunkte (per Protokoll zu Monat 3 und Monat 6), N=828 Patienten. Die zeitlichen Abweichungen sind unter Kontroll-Bedingungen größer als unter Interventionsbedingungen. Definierter Nullpunkt: Index-ICU Entlassungsdatum. | 85 |
| Abbildung 21. Verteilung der Tage zwischen der Entlassung von der letzten Intensivbehandlung und 1. bzw. 2. Kostenerhebung (HA-CRF 1 und HA-CRF 2)..... | 90 |
| Abbildung 22. Nutzerfreundlichkeit der Televisitenroboter..... | 96 |
| Abbildung 23. Subjektive Nutzenbewertung der Televisitenroboter (1)..... | 97 |
| Abbildung 24. Subjektive Nutzenbewertung der Televisitenroboter (2)..... | 98 |
| Abbildung 25. Subjektive Nutzenbewertung der Televisitenroboter (3)..... | 98 |
| Abbildung 26. Subjektive Nutzenbewertung der Televisitenroboter (4)..... | 99 |
| Abbildung 27. Subjektive Nutzenbewertung der Televisitenroboter (5)..... | 99 |
| Abbildung 28. Ausmaß und Belastung infolge der Krisensituation (1)..... | 100 |
| Abbildung 29. Ausmaß und Belastung infolge der Krisensituation (2)..... | 100 |
| Abbildung 30. Einstellungen gegenüber Visitenrobotern. | 101 |
| Abbildung 31. Zufriedenheit mit den Televisitenrobotern (1). | 101 |
| Abbildung 32. Zufriedenheit mit den Televisitenrobotern (2). | 102 |
| Abbildung 33. Heatmap für die Cluster-spezifische Rekrutierung pro Kalendermonat (Fall-Ebene), stratifiziert nach den drei Sequenzgruppen (19 Monate Rekrutierungsphase). Je dunkler die Farbe, desto mehr Patienten wurden in diesem Kalendermonat in die Studie eingeschlossen. Fehlende Angaben zur Anzahl der Patienten in einem bestimmten Monat bedeuten, dass in diesem Zeitraum keine Patienten rekrutiert wurden..... | 110 |
| Abbildung 34. 6-Monats-Gesamtmortalität: Kaplan-Meier Kurven für die beiden Behandlungsbedingungen Telemedizin-basierte Intervention vs. Routineversorgung (Kontrolle). | 115 |

Abkürzungen und Definitionen

| | |
|------------|---|
| AHB | Anschlussheilbehandlung |
| (e)CRF | (electronic) Case report form |
| CI | confidence interval, Konfidenzintervall |
| CONSORT | Consolidated Standards of Reporting Trials |
| DIVI | Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin e. V. |
| DSGVO | Datenschutz-Grundverordnung der EU (EU-DSGVO) |
| EQ-5D-5L | EuroQoI 5 dimensions and 5 level Version |
| EQ VAS | EuroQoI Visual Analogue Scale |
| ERIC | Enhanced Recovery after Intensive Care |
| FAS | Full analysis set |
| FPF(L)V | First patient first (last) visit (erste (letzte) Visite des zuerst eingeschlossenen Patienten) |
| FU | Follow-up |
| G-BA | Gemeinsamer Bundesausschuss |
| HA-CRF | Hausarztfragebogen, Hausarzt-CRF |
| HGS Test | Hand Grip Strength Test (Test zur Messung der Handkraft) |
| HR | Hazard Ratio |
| ICH-GCP | International conference for harmonisation (of technical requirements for pharmaceuticals for human use) – good clinical practice guideline |
| ICU | Intensive care unit, Intensivstation |
| IES-R | Impact of Event Scale – revised |
| IKER | Inkrementelle Kosteneffektivitätsrelation |
| IQR | Interquartile Range, Interquartilsabstand |
| ITT | Intention to treat |
| KI | Konfidenzintervall |
| LPF(L)V | last patient first (last) visit |
| M(C)AR | Missing (Completely) at Random |
| mMRC | Modified British Medical Research Council (Dyspnea Scale) |
| OR | Odds Ratio |
| PatID | Patientenidentifikator |
| PHQ | Patient Health Questionnaire |
| PICS | Post intensive care syndrome |
| PP | per protocol |
| PTSD | Post-traumatic stress disorder, Posttraumatische Belastungsstörung |
| QI | Quality indicator |
| QALY | Quality-adjusted life year |
| QoL | Quality of life |
| REDCap | Research Electronic Data Capture |
| SAPS II | Simplified Acute Physiology Score II |
| SOFA Score | Sequential Organ Failure Assessment Score |
| SW-CRT | Stepped wedge cluster randomized trial |
| SWD | Stepped wedge design |
| Tele-ICU | telemedicine intensive care unit |
| trt | treatment (bedeutet bei ERIC: Interventionsbedingung) |
| TUG Test | Timed Up and Go Test |
| VAS | Visual analogue scale (des EQ Fragebogens) |
| W-DWH | Wissenschafts-Data Warehouse (der BARMER Versicherung) |
| WHODAS | WHO Disability Assessment Schedule |

0 Einleitung & Protokoll-Information

Das Protokoll der Interventionsstudie ERIC sowie der darin enthaltene statistische Analyseplan (SAP), Grundlage dieses Evaluationsberichts, wurden in folgendem wissenschaftlichen Artikel Open Access publiziert:

Protokoll-Veröffentlichung inkl. Statistischer Analyseplan und methodischen Online-Supplements

Adrion C, Weiss B, Paul N, Berger E, Busse R, Marschall U, Caumanns J, Rosseau S, Mansmann U, Spies C; *ERIC* study group.

Enhanced Recovery after Intensive Care (ERIC): study protocol for a German stepped wedge cluster randomised controlled trial to evaluate the effectiveness of a critical care telehealth program on process quality and functional outcomes.

BMJ Open 2020 Sep 25;10(9):e036096. doi: 10.1136/bmjopen-2019-036096.

Um selektives Reporting zu vermeiden orientiert sich dieses Dokument für den primären Endpunkt Abschnitt zur Wirksamkeitsevaluation an der CONSORT Reporting-Guideline sowie insbesondere an der entsprechenden Erweiterung für Stepped Wedge Cluster-randomisierte kontrollierte Studien (SW-CRT).[1, 2]

Protokollversionen

Amendment 04: Version 1.2, 14.12.2020. Dieses letzte Amendment beinhaltet die Melderegister-Abfrage und die Akzeptanzerhebung als Subprojekt.

Federführende Ethikkommission:

EK der Charité – Universitätsmedizin Berlin; Referenz-Nr. EA1/006/18

Offizieller Registereintrag

Die Studie wurde am 22 August 2018 prospektiv registriert unter ClinicalTrials.gov, Identifier NCT03671447, URL: <https://clinicaltrials.gov/ct2/show/NCT03671447>.

Studienwebseite: <https://www.eric-projekt.net/>

1 Ausgangslage und Ziele

1.1 Ausgangslage

Deutschland verfügt weltweit über eine der höchsten Dichten an Intensivbetten. Mehr als 2.2 Millionen Patienten werden jedes Jahr auf einer Intensivstation (Intensive Care Unit, ICU) behandelt, die Mehrzahl in kleinen und mittelgroßen Krankenhäusern. Dabei entfallen fast mehr als 10% der Gesamtausgaben des stationären Versorgungssektors auf Intensivfälle.

Umso wichtiger ist es, an dieser Stelle eine qualitativ hochwertige, an aktuellen Erkenntnissen ausgerichtete Therapie sicherzustellen. Die Intensivmedizin ist ein Hochrisikobereich, und neben dem Mortalitätsrisiko sind Langzeitfolgen ein großes, bislang wenig beachtetes Problem. Eine qualitativ hochwertige, an aktuellen wissenschaftlichen Erkenntnissen ausgerichtete Therapie ermöglicht es, auch schwere Erkrankungen zu überleben. Ein großer Teil dieser Überlebenden leidet jedoch auch Jahre nach Beendigung der Behandlung an funktionellen Langzeitbeeinträchtigungen. Dies betrifft Beeinträchtigungen der kognitiven Fähigkeiten und ein erhöhtes Risiko für die Entwicklung einer Demenz, Beeinträchtigungen der mentalen Gesundheit, insbesondere Depressionen, Angststörungen, Posttraumatische Belastungsstörungen (PTSD) und sowie Beeinträchtigungen der Mobilität mit verminderter Kraft und großem Hilfsbedarf bei der Erledigung alltäglicher Aufgaben. Diese Langzeitfolgen und Beeinträchtigungen werden als *Post Intensive Care Syndrome (PICS)* zusammengefasst und führen zu einer Verringerung der gesundheitsbezogenen Lebensqualität (HRQoL).[3, 4]

Bereits vor der COVID-19 Pandemie hat sich die Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin (DIVI) der evidenzbasierten Verbesserung der Qualität der intensivmedizinischen Therapie gewidmet und für Deutschland 10 *Qualitätsindikatoren (QIs)* der akuten, intensivmedizinischen Versorgung in einem standardisierten Verfahren anhand bestehender Leitlinien und der Studienlage definiert. Nach der Erstauflage 2010 sind diese in nunmehr dritter Auflage in der Version von 2017 vorhanden und passen sich zyklisch den neusten wissenschaftlichen Erkenntnissen an.[5] Ziel dieser QIs ist die Umsetzung einer leitliniengerechten Therapie am Patientenbett. Als Teil des von der DIVI initiierten, externen Peer-Reviews beinhalten die QIs Aspekte der Strukturqualität und Prozessqualität. Sie sind klar definiert und ihre Umsetzung messbar und quantifizierbar. Sie sind operationalisierbar, und für jeden QI wurden Benchmarks festgelegt. Obwohl gezeigt werden konnte, dass eine verbesserte Einhaltung (*Adhärenz*) der QIs Langzeitbeeinträchtigungen nach Intensivtherapie reduzieren kann, sind die QIs nur unzureichend verbreitet und implementiert. Telemedizin könnte ein möglicher Weg sein, die Verbreitung der QIs und Einhaltung dieser zu erhöhen. Letztlich kann die telemedizinische Mitbehandlung eine Umsetzung der veränderlichen Evidenzlage auf individualmedizinischer Ebene fördern und so die Versorgung von ICU-Patienten verbessern.

1.2 Ziele, Fragestellungen und Hypothesen

Die telemedizinische Mitversorgung intensivmedizinischer Patienten ist vor allem in den USA verbreitet. Bisherige Arbeiten konnten zeigen, dass sich durch Telemedizin die Mortalität während des Intensivstations- und Krankenhausaufenthalts reduzieren lässt sowie die Intensivstations- und Krankenhausaliedauer verkürzen lassen. Wie der Effekt genau vermittelt wird, ist weitgehend unklar. Es gibt aber Hinweise, dass Telemedizin das Potential besitzt, die Adhärenz zu spezifischen Empfehlungen zu erhöhen, und eine fallbasierte Allokation von medizinischen Ressourcen besser möglich ist.

Das primäre Ziel von *Enhanced Recovery after Intensive Care* (ERIC) war zu untersuchen, inwieweit sich durch eine telemedizinische Intervention in der ICU die auf täglicher Basis gemessene Adhärenz zu den einzelnen QIs der DIVI steigern lässt, und somit Telemedizin gezielt Prozessqualität steigern bzw. beeinflussen kann. Zur Überprüfung dieser Forschungshypothese wurde im Themenbereich "Neue Versorgungsformen" das Innovationsfond-Projekt ERIC als multizentrische, Cluster-randomisierte kontrollierte Studie im Stepped Wedge Design (SW-CRT) initiiert.

Als sekundäre Ziele untersuchte ERIC als intersektorales Projekt, inwieweit Überlebende einer Intensivtherapie mit einem Case-Care-Management nach ihrer stationären Behandlung in den ambulanten Sektor eingebunden werden können, mit Fokus auf PICS-Früherkennung und -Therapie. Ferner wurde untersucht, inwieweit Patienten, die in der neuen Versorgungsform behandelt werden, seltener von Langzeitbeeinträchtigungen betroffen waren im Vergleich zu unter Kontrollbedingungen behandelten Patienten.

Dazu zählt u.a. das Vorliegen einer außerklinischen Beatmung, die nicht nur die Lebensqualität der Betroffenen stark einschränkt, sondern auch mit hohen Versorgungskosten verbunden ist. So geht man insgesamt von Kosten in Höhe von zwei bis vier Milliarden Euro pro Jahr für die außerklinische Beatmung aus,[6, 7] wobei die Kosten pro außerklinisch beatmetem Patienten auf 15 000 zu 20 000 Euro pro Monat geschätzt werden.[8]

Zielgruppe der neuen Versorgungsform waren alle erwachsenen Patienten, die nach Kerndatensatz die Intensivkriterien erfüllen und auf einer Intensivstation in Deutschland behandelt werden.

ERIC wurde von folgenden fünf Konsortialpartnern durchgeführt:

1. Technische Universität Berlin (TU Berlin), Fachgebiet Management im Gesundheitswesen: gesundheitsökonomische Evaluation
2. Fraunhofer Institut für Offene Kommunikationssysteme (FOKUS): Etablierung der audiovisuellen Kommunikation der Tele-ICU, Entwicklung der E-Health-Plattform
3. Ernst von Bergmann Klinikum Bad Belzig gGmbH, Brandenburger Lungen- und Beatmungszentrum, Bad Belzig: Intersektorale, ambulante Versorgung mit Schwerpunkt auf langzeitbeatmete Patienten; Planung des intersektoralen Follow-ups (FU), Unterstützung des FUs bei Brandenburger Patienten
4. BARMER, Wuppertal: Als eine der größten gesetzlichen Krankenversicherungen (GKV) Deutschlands mit 8.4 Millionen Versicherten zuständig für die Ausarbeitung des Selektivvertrags, Unterstützung durch Daten des Wissenschafts-Data-Warehouse (W-DWH).
Entsprechend der ursprünglichen Planung war das ERIC-Projekt zunächst nur mit BARMER-versicherten Patienten gestartet.
5. Ludwig-Maximilians-Universität München (LMU), Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie (IBE): Konsortialpartner verantwortlich für das biometrische Konzept sowie die unabhängige Evaluation (Clinical Effectiveness).

Konsortialführer im Projekt und Studienleitung war die Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin der Charité – Universitätsmedizin Berlin.

2 Darstellung der neuen Versorgungsform

Der grundlegende Aufbau der neuen Versorgungsform ist in Abbildung 1 dargestellt und wird im Folgenden beschrieben.

Innersektorale Modifikation/ Intervention

Die primäre Versorgungsleistung und innersektorale Modifikation besteht in der Durchführung einer täglichen, Audio-Video-unterstützten QI-zentrierten telemedizinischen Visite (*QI Visite*), deren Dokumentation in einer E-Health-Plattform erfolgt. Partner der Televisite waren die teilnehmenden Intensivstationen des "ERIC-Verbundes" sowie das zentrale Tele-ICU Cockpit an der Charité.

Dabei waren die wesentlichen Eckpunkte der E-Health-Plattform:

1.) *Televisite:*

Ärzte sollten in die Lage versetzt werden, über Video standortunabhängig miteinander kommunizieren zu können. Durch die gleichzeitige Bereitstellung von Patientendaten sollen so Konferenzen, Konsile und Boards zwischen Krankenhäusern und dem zentralen intensivmedizinischen Telemedizin-Zentrum ("Tele-ICU Cockpit") möglich sein.

2.) *Bewertung der DIVI QIs für die akut-intensivmedizinische Versorgung:*

Die definierten QIs werden auf ärztlicher Seite unterstützt durch Geräte erhoben (z.B. Daten aus Beatmungsgerät) und sind Basis der durch das Telemedizin-Zentrum der Charité durchgeführten Bewertungen und Beratungen. Grundlage für das zentrale Rating sind Schlüsselperformance-Indikatoren, die in Kenntnis klinischer Situation etabliert wurden.

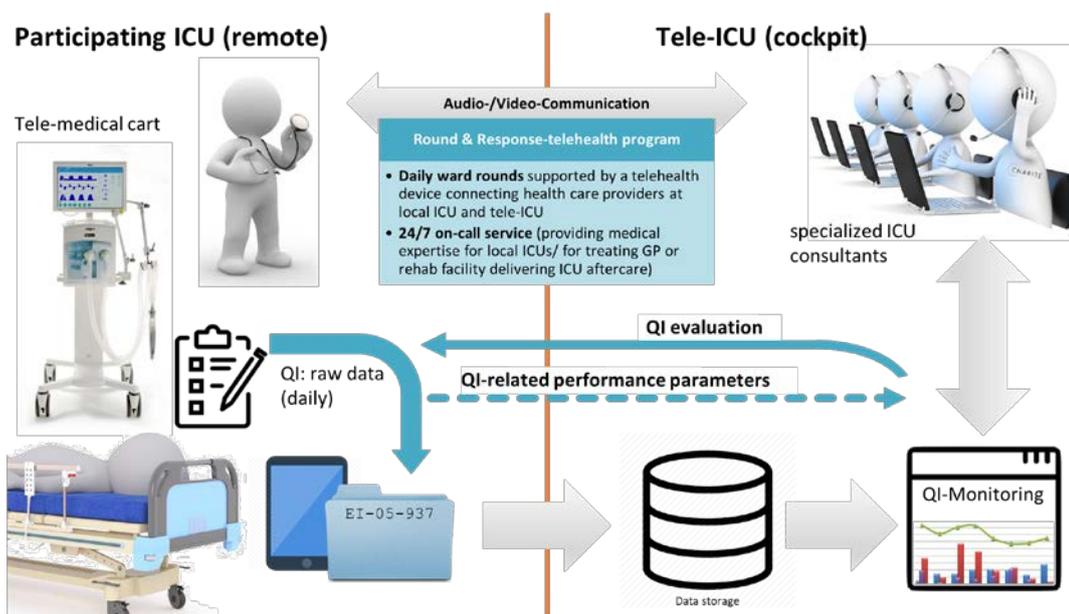


Abbildung 1. Schematischer Aufbau der neuen Versorgungsform ERIC (eigene Darstellung) während der ICU-Phase. Prozess des Telekonsils bzw. der Televisiten bis zur Ableitung der QI-bezogenen primären Outcomes für die Evaluation.

Die vier Säulen des ERIC-Konzepts

Als tragende Elemente des ERIC-Konzepts zur Weiterentwicklung der GKV-Regelversorgung gelten:

1. Audio-visuelle Televisite als "Round-and-Response"
2. QI-Zentrierung der patientenindividuellen täglichen Televisiten
3. Blended Learning-basierte QI-Schulungen
4. Strukturierte, intersektorale Nachsorge unter Einbeziehung der Hausärzte.

Für die intersektorale Versorgung eingeschlossener Patienten im Case-Care-Management während des Post-ICU Follow-up wurde mit dem Hausärzteverband Berlin und Brandenburg e.V. zusammengearbeitet.

Weitere Details zu den Abläufen der telemedizinischen QI-Visite als Kern der neuen Versorgungsform mit nachfolgendem Case-Care-Management sind im Ergebnisbericht dargestellt.

2.1 Kontrollphase: Regelversorgung mit Case-Care-Management

ICU-Phase:

In der Kontrollphase, in der alle teilnehmenden Cluster starteten, erhielten in die Studie eingeschlossene Patienten die intensivmedizinische Regelversorgung durch ihr behandelndes Krankenhaus. Zusätzlich wurde während der patientenindividuellen ICU-Phase täglich dokumentiert, ob die folgenden QIs der DIVI in Bezug auf die Behandlung eines einzelnen Studienpatienten erfüllt waren oder nicht:[5]

- QI 1: Tägliche multiprofessionelle und interdisziplinäre klinische Visite mit Dokumentation von Tageszielen
- QI 2: Management von Sedierung, Analgesie und Delir
- QI 3: Patientenadaptierte Beatmung
- QI 4: Frühzeitige Entwöhnung von einer invasiven Beatmung (Weaning)
- QI 6: Maßnahmen zum Infektionsmanagement
- QI 7: Frühe enterale Ernährung
- QI 8: Dokumentation einer strukturierten Patienten- und Angehörigenkommunikation
- QI 9: Frühmobilisation.

Die Adhärenz zu den einzelnen QIs wurde in einer elektronischen Plattform erfasst. Die Bewertung der QIs erfolgte anhand festgelegter Key Performance-Merkmale und durch geschultes Studienpersonal zentral ausschließlich an der Charité.

Die Erfüllung des DIVI QI 5 ("Überwachung der Maßnahmen zur Infektionsprävention") und QI 10 ("Leitung der Intensivstation") wird nicht täglich auf Patientenebene gemessen. Aus diesem Grund stellten diese keine primären Outcomes im Rahmen der Evaluation dar.

Post-ICU Follow-up (FU) Phase:

Drei und sechs Monate nach Entlassung von der Intensivstation erhielten Studienpatienten (sowohl Patienten der Kontroll-, als auch Patienten der Interventionsphase) in einem Case-Care-Management ein intersektorales Follow-up durch ihren behandelnden Hausarzt, in dem die Patienten auf mögliche funktionelle Langzeitbeeinträchtigungen untersucht wurden. Dieses Follow-up wurde im Rahmen des Projektes entwickelt und erstmals angewendet.[9]

Sofern das Follow-up Assessment nicht durch den eigenen Hausarzt durchgeführt werden konnte, wurden die Patienten durch Studienpersonal im Studienzentrum der Charité – Universitätsmedizin Berlin oder alternativ während eines Hausbesuchs untersucht.

Grundlage der Untersuchung im Post-ICU Follow-up war die Erfassung von funktionellen Langzeitbeeinträchtigungen und Symptombelastung in den folgenden PICS-bezogenen Domänen:

- mentale Gesundheit (Angststörungen und Depression),
- Kognition,
- Mobilität/physische Gesundheit (Muskelstärke),
- gesundheitsbezogene Lebensqualität.

Die für die Evaluation ausgewählten Messinstrumente (Patienten-Fragebögen; Tests) orientierten sich am sog. Core Outcome Measurement Set für PICS-Screening bei ICU-Überlebenden im ambulanten Setting,[10] welches im Rahmen des Projekts für dieses Patientenkollektiv entwickelt wurde (Details zum Konsensus-basierten Entwicklungsprozess siehe Spies *et al.* [9]). Die nachfolgende Abbildung gibt einen groben Überblick über den Ablauf und die Datenerhebung im Follow-up zum Zeitpunkt nach drei bzw. sechs Monaten.

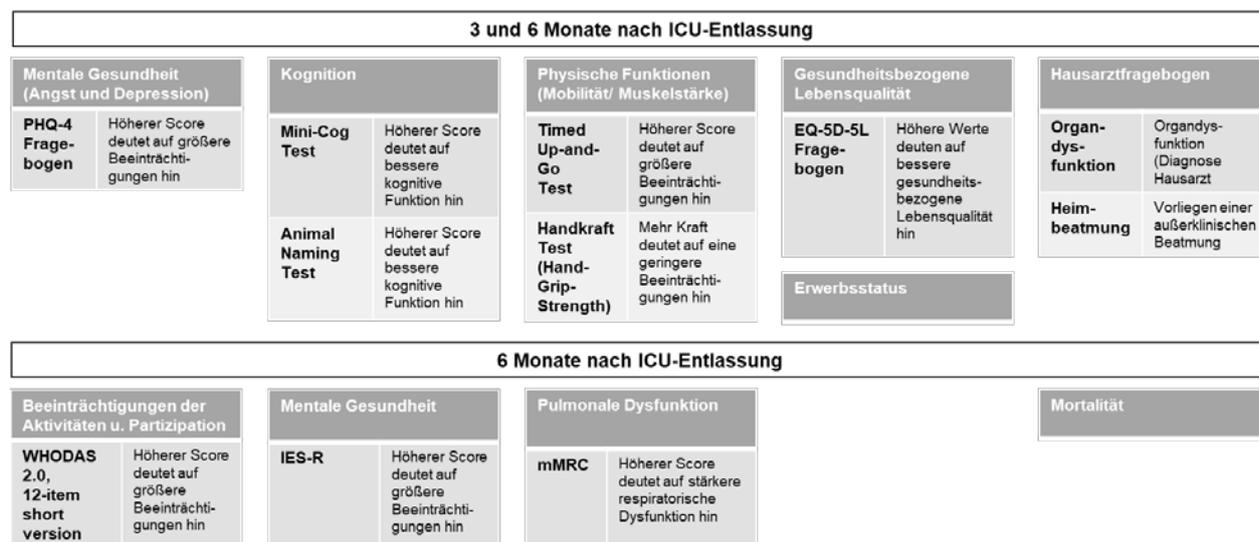


Abbildung 2. Überblick zu Ablauf und Datenerhebung drei und/oder sechs Monaten nach Entlassung aus der Intensivstation (modifizierte Darstellung anhand Spies *et al.* [7]). PHQ: Patient Health Questionnaire; EQ-5D-5L: EuroQol 5 Dimensions 5 Level; WHODAS: World Health Organization Disability Assessment Schedule; IES-R: Impact of Event Scale-Revised; mMRC: modified Medical Research Council Score.

2.2 Trainingsphase und Interventionsphase (neue Versorgungsform mit Case-Care-Management)

Jedes an ERIC teilnehmende Cluster durchlief vor dem Übergang von der Kontrollphase, in der im Rahmen der Regelversorgung behandelt wurde (im Folgenden bezeichnet als *Kontroll-Bedingung*), in die Interventionsphase ein ERIC-Trainingsprogramm zum Umgang mit dem telemedizinischen Visitenwagen und zu den QIs. Dieses Training begann drei Monate vor dem Wechsel in die Interventionsphase und basierte auf einem Blended-Learning-Konzept.

Es beinhaltete:

- Dezentrales Lernen in einem E-Learning-Kurs auf der ERIC-Lernplattform (<https://best-edx.charite.de>) mit E-Learning-Modulen für QI1 bis QI 9 (QI 10 nicht Bestandteil des Kurses). Jedes Modul ist multimedial aufgebaut und beinhaltet eingesprochene Inhalte, Videos und animierte Übungen. Am Ende jedes Moduls mussten zehn Multiple-Choice-Fragen beantwortet werden. Für jedes Modul wurde eine Lernzeit von 30 bis 45 Minuten veranschlagt.
- Strukturiertes, anwendungsbezogenes Präsenztraining durch praxiserfahrene Dozierende in Form von Simulationskursen, in denen fallbasiert verschiedene Simulationsszenarien geübt wurden. Es fanden vier Simulationsworkshops statt:
 1. Kommunikation und Team-Performance auf der Intensivstation,
 2. Management von Sedierung, Analgesie und Delir,
 3. Beatmung, frühzeitige Entwöhnung von der Beatmung und Frühmobilisation,
 4. Infektionsmanagement.
- Anwendungstraining während der Arbeit ("On-the-Job"). Die Multiplikatoren wurden bei der Umsetzung des Gelernten auf der eigenen Intensivstation durch Studienpersonal begleitet und unterstützt.

Multiplikatoren jedes Zentrums durchliefen das Trainingsprogramm, welche das erworbene Wissen auf ihrer Intensivstation an Kollegen weitergaben. Jede teilnehmende Intensivstation wurde zudem ermutigt, vor Beginn der Trainingsphase am Peer Review-Verfahren der DIVI (<https://www.divi.de/empfehlungen/qualitaetssicherung-intensivmedizin/peer-review>) teilzunehmen, basierend auf dem Plan-Do-Study-Act Konzept (PDSA).[11] Dieses Peer-Review ermöglicht einen interprofessionellen Austausch zwischen den Disziplinen einer Intensivstation und analysiert das Potential zur Verbesserung der Struktur-, Prozess- und Ergebnisindikatoren einer Intensivstation.[12-14]

Implementierung der experimentellen komplexen Intervention

Für die Durchführung der Televisiten setzte die Tele-ICU der Charité für die telemedizinische Versorgung externer Krankenhäuser Telepräsenz-Visitengeräte (Visitenwagen) vom Typ LITE V3 der Firma InTouch Health®/ Teladoc Health, Inc. ein (Abbildung 3). Nach Abschluss der Trainingsphase und unmittelbar vor Beginn der Intervention wurde jede teilnehmende ICU eines Clusters mit diesem telemedizinischen *Visitenwagen* ausgestattet.

Dieser Visitenwagen war zum einen manuell verschiebbar und konnte somit direkt zum Patientenbett gebracht werden, zum anderen war es ausgestattet mit einer hochauflösenden Kamera, höhenverstellbarem Bildschirm, Lautsprecher und Mikrofon. Diese Technologie ermöglichte eine direkte audio-visuelle Verbindung zu der am Campus Virchow-Klinikum der Charité – Universitätsmedizin Berlin etablierten Tele-ICU. Diese bestand aus zwei Arbeitsstationen mit jeweils einem Computer, mehreren Bildschirmen, Mikrofon, Kamera und Lautsprecher.

An jeder Arbeitsstation arbeitete täglich von 8 Uhr bis 17 Uhr, sieben Tage die Woche, ein erfahrener Facharzt für Anästhesiologie mit Zusatzbezeichnung Intensivmedizin, und (werktags) eine in der Intensivmedizin erfahrene Pflegefachkraft. Außerhalb dieser Zeiten stand ein ebenso qualifizierter Intensivmediziner in Rufbereitschaft zur Verfügung.

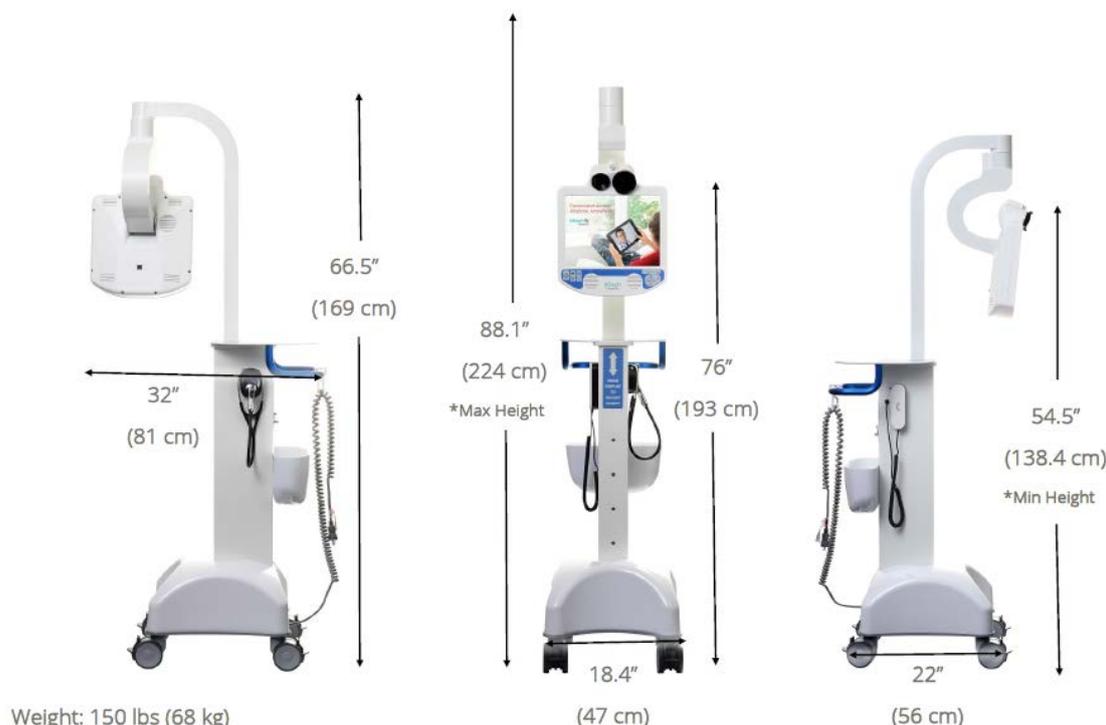


Abbildung 3. Foto des im Rahmen der Interventionsphase eingesetzten Telemedizingeräts ("Visitenwagen") Typ LITE V3 (Hersteller: InTouch Health, jetzt Teladoc Health, Inc.); technische Spezifikationen siehe Herstellerbroschüre¹. Mikrofon: mono – direktional (hyper-cardioid) 10Hz – 40Hz, zwei 60W Mono-Lautsprecher; 15" LCD-Monitor, Schwenkbereich des Kopfteils: horizontal +/- 170°, vertikal +27°/-65° maximal; Videokamera; eigener Handhörer, Ablage, Griff, Erweiterungsfach, Netzstecker und -kabel, Radschloss und Korb.

Ablauf der Intervention und Durchführung der Telemedizin-basierten QI-Visiten

Der Patienteneinschluss in die neue Versorgungsform erfolgte nach Aufnahme auf die ICU und Aufklärung des Intensivpatienten oder, falls dieser dazu nicht in der Lage war, des Angehörigen durch Ärzte der teilnehmenden Cluster. In die neue Versorgungsform eingeschlossene Patienten, die unter Interventionsbedingungen behandelt wurden, wurden täglich durch die Tele-ICU visitiert. Diese QI-Visiten erfolgten nach einem festgelegten Schema und bestanden grundsätzlich aus Telepräsenz und Dokumentation in der Plattform. Hierzu rief die Tele-ICU zu festgelegten Zeiten im teilnehmenden Zentrum an, um den Beginn der Televisite abzustimmen. Bei Zusage des teilnehmenden Zentrums wählte sich die Tele-ICU mittels einer sicheren, End-zu-End-verschlüsselten Verbindung auf das telemedizinische Präsenzgerät im teilnehmenden Zentrum ein. Direkt am Bett der Studienpatienten wurde mit dem Intensivmediziner, der Pflegefachkraft, dem ärztlichen Kollegen vor Ort sowie dem Patienten eine interdisziplinäre Visite durchgeführt (*QI-Visite*). Diese Visite fokussierte sich auf die Erfüllung der QIs der DIVI und bot eine darüberhinausgehende, medizinische Beratung.

Im Anschluss dokumentierten die Tele-ICU Fachärzte sowie ggf. die Ärzte der Partnerklinik die Inhalte der Televisite in der gemeinsamen Dokumentationsplattform. Es fand eine strenge Trennung der klinischen Routedokumentation und der Datennutzung zu wissenschaftlichen Forschungszwecken (Evaluation) statt. Die QI-Erfüllung (ja/nein) wurde für jeden Patienten täglich in der auf dem Gelände der Charité gehosteten Forschungsdatenplattform REDCap® (Research Electronic Data Capture) dokumentiert. Jede Visite dauerte ca. 20 Minuten, so dass maximal 24 in ERIC eingeschlossene Patienten pro Arbeitsstation täglich visitiert werden konnten.

¹ <https://cloud.kapostcontent.net/pub/8d517633-fd5f-48c8-b719-5bf632f0af3c/product-brochure-lite-reference-manual.pdf?kui=N3NYTZk7alpGrRkkGOroXA>, ab S.41ff [letzter Zugriff: 18.06.2021]

3 Darstellung des Evaluationsdesigns

3.1 Primäre Fragestellung und Arbeitshypothese

Die primäre Hypothese von ERIC gemäß Projektantrag lautete, dass die im Rahmen der neuen Versorgungsform erbrachten telemedizinisch unterstützten QI-Visiten, die täglich während des gesamten ICU-Aufenthalts eines Patienten stattfinden, zu einer anhand evidenzbasierter QIs gemessenen Verbesserung der Struktur-, Prozess- und Ergebnisqualität auf der ICU führen, im Vergleich zu der unter der Regelversorgung anhand der QIs gemessenen Qualität. Hierbei wird die Erfüllung der patientenindividuellen QIs auf täglicher Basis routinemäßig dokumentiert.[5]

3.2 Studiendesign und Setting

ERIC war eine nationale, multizentrische, pragmatische, Cluster-randomisierte kontrollierte Interventionsstudie (*quality improvement study*) im Open Cohort Stepped Wedge Design (SWD) und kontinuierlicher Rekrutierung (*continuous recruitment short exposure design*).[15, 16]

Das SWD kombiniert Elemente einer klassischen Cluster-randomisierten Studie (Randomisierung und Anwendung der Intervention auf Cluster- anstatt auf Patientenebene) sowie eines Vorher-Nachher-Vergleichs (jedes Cluster durchläuft Kontroll- und Interventionsphase in unidirektionaler Richtung).[17] Ein wesentlicher Vorteil des SWD stellt dessen höhere Akzeptanz und Praktikabilität bei den teilnehmenden Institutionen im Falle von experimentellen Interventionen mit sehr geringem Sicherheits-Risiko dar. So muss durch die zeitversetzte stufenweise Einführung in den Studienzentren die Intervention nur in einer begrenzten Anzahl an Einrichtungen gleichzeitig implementiert werden, was einen offensichtlichen Vorteil gegenüber einem Parallelgruppendesign darstellt.

Das Setting der Studie waren Intensivstationen von Krankenhäusern, die sich im Großraum Berlin und umliegenden Bundeslandes Brandenburg mit einer Bevölkerung von insgesamt etwa sechs Millionen Einwohnern befanden. In diesem Gebiet gibt es ca. 150 000 ICU-Aufnahmen pro Jahr (GBE-Bund Gesundheitsdaten [18]). Über den Studienzeitraum von 25 Monaten (first-patient-in bis last-patient-out), der ein sechsmonatiges Post-ICU Follow-up auf Patientenebene beinhaltete, wurden die teilnehmenden Krankenhauseinrichtungen zufällig auf drei vordefinierte Sequenzgruppen zugeteilt, um von der Kontroll- in die Interventionsbedingung zu drei unterschiedlichen Crossover Zeitpunkten zu wechseln. Nach dem Wechsel eines Clusters zur Interventionsbedingung blieb es für die restliche Dauer der Studie in der Intervention. Nach dem Abschluss der Rollout-Periode mit Übergang der der Sequenzgruppe 3 zugewiesenen Cluster zur Intervention waren alle Intensivstationen aller Cluster über eine 7-monatige Post-Rollout Periode vollständig exponiert, d.h. rekrutierte Patienten wurden unter Interventionsbedingung behandelt.

Aufgrund des Studiendesigns wurde angenommen, dass alle teilnehmenden Krankenhäuser gleichzeitig mit der Rekrutierung von Patienten starten und die Kontrollphase der Studie unmittelbar mit Studienstart für alle Cluster zeitgleich beginnt.

Die Studie wurde prospektiv bei ClinicalTrials.gov (ID: [NCT03671447](https://clinicaltrials.gov/ct2/show/study/NCT03671447)) registriert. Die Ethikkommission der Charité Universitätsmedizin, Berlin, als federführende Ethikkommission der ERIC Studie gab ihr positives Votum (Referenznummer: EA1/006/18) erstmalig am 26. Januar 2018. Die Ethikkommission der Medizinischen Hochschule Brandenburg (MHB) Theodor Fontane

schloss sich zu einem späteren Zeitpunkt an (Referenznummer: Z-01-20180828). Das zuletzt gültige Protokoll-Amendment, Version 1.2, ist datiert auf den 14.12.2020.

Die Finanzierung der Studie erfolgte durch den Innovationsfonds des Gemeinsamen Bundesausschusses (G-BA).

3.3 Ein- und Ausschlusskriterien – Clusterebene und Patientenebene

Cluster

Als geographische und organisatorisch und personell unabhängige Randomisierungseinheit bzw. Cluster wurden Krankenhauseinheiten ausgewählt, die folgende Kriterien erfüllen:

- Krankenhaus mit Intensivstationen für Erwachsene;
- Lage in der Metropolregion Berlin/Brandenburg;
- Unterschrift der entsprechenden Kooperationsverträge zur Teilnahme an der neuen Versorgungsform (u.a. Kooperationsvertrag mit der Charité - Universitätsmedizin Berlin);
- Akzeptanz der Cluster-Randomisierung und des damit vorgegebenen Zeitpunkts der Implementierung der Intervention auf institutioneller Ebene.

Patienten

Auf die Intensivstationen aufgenommene Patienten konnten in die Studie eingeschlossen werden, sofern diese die folgenden Einschlusskriterien erfüllten:

- Alter mindestens 18 Jahre;
- erwartete ICU-Aufenthaltsdauer von mindestens 24 Stunden;
- BARMER-Versicherung oder anderer gesetzlicher Krankenversicherungsstatus (ab 03/2019);
- schriftliche Einverständniserklärung durch den Patienten oder durch den gesetzlichen Vertreter.

Falls ein Patient zuvor bereits auf einer teilnehmenden Intensivstation aufgenommen worden war, war ein nachfolgender Studien-Einschluss möglich, mit QI-Dokumentation während des ICU-Aufenthalts unter der zum jeweiligen Zeitpunkt im Cluster vorherrschenden Behandlungsbedingung (Kontroll- bzw. Interventionsbedingung mit Tele-ICU). Aufgrund des Open Cohort Designs wurden in an ERIC teilnehmenden Institutionen kontinuierlich Patienten bei ICU-Aufnahme auf einen möglichen Studien-Einschluss gescreent. Die Prüfung der Einschlusskriterien fand durch Personal vor Ort statt, welches vor Rekrutierungsbeginn durch die Konsortialführung diesbezüglich fortgebildet wurde.

Abweichungen vom ursprünglich geplanten Vorgehen (Innovationsfond-Antrag): Änderungen der Einschlusskriterien

ERIC wurde als Projekt der besonderen Versorgung nach §140a, SGB V mit der BARMER Versicherung beantragt und umgesetzt. Allerdings war die Rekrutierung niedriger als erwartet und es ergab sich das Dilemma, dass Nicht-BARMER Patienten aus ethischen Gründen ebenso eingeschlossen werden mussten, aber nicht evaluiert werden konnten. Auch wenn die Rechtsform des Selektivvertrages grundsätzlich geeignet war, um Patienten der neuen Versorgungsform zuzuführen, und die Berechnungen zum Anteil von BARMER-Versicherten bereits im Antrag aufgeführt wurden, ergab sich in der Versorgungsform die Herausforderung, dass der Anteil im Rekrutierungszeitraum unerwartet niedrig war und viele Anfragen zu nicht-BARMER-Patienten an die Konsortialführung gestellt wurden.

Um das Projektziel und die geplante Patientenanzahl im beantragten Studienzeitraum zu erreichen, wurde im Verlauf des Projektes eine Erweiterung der Rechtsform beantragt und diesem Antrag stattgegeben, so dass gemäß §630a BGB ab Frühjahr 2019 alle GKV-versicherten Patienten in die Studie eingeschlossen und, in Abhängigkeit vom Crossover-Zeitpunkt des Clusters, gegebenenfalls in die neue Versorgungsform eingeschlossen werden konnten.

Die Rekrutierungsphase wurde als zweite Maßnahme zum Erreichen des Rekrutierungsziels auf 19 Monate verlängert, ohne Auswirkungen auf das Evaluationskonzept. Somit konnte die geplante Patienten-Fallzahl erreicht werden.

Es gab keine weiteren Abweichungen zum Protokoll-Paper und dem vom Projektträger genehmigten Evaluationskonzept (letzte Version 28.02.2020).

3.4 Randomisierung und Verblindung

3.4.1 Randomisierung

Die Randomisierung fand auf Vertragsebene, d.h. Klinikträgerebene statt, um einem Klinikträger den gleichen Projektplan für alle Häuser vorlegen zu können.

Die Randomisierung der insgesamt 12 Cluster, die zuvor im Rahmen des Projektantrags ihren Letter-of-Intent inklusive Einverständniserklärung abgegeben hatten, erfolgte wie bei einem SWD notwendig zu einem einzelnen Zeitpunkt vor Beginn der Patientenrekrutierung.

Für die Charité – Universitätsmedizin Berlin wurden aus organisatorischen Gründen zwei Cluster definiert (Campus Charité Mitte & Campus Virchow-Klinikum, sowie Charité Benjamin Franklin). Hiermit wurde dem personellen Austausch zwischen den Häusern eines einzelnen Clusters, d.h. dem sich zwischen einzelnen Standorten bewegenden und an ERIC teilnehmenden Ärzten und Pflegepersonal Rechnung getragen, um das Risiko von Kontamination zu minimieren.

Am 21.08.2018 wurde von einer unabhängigen Statistikerin des IBE der LMU, Kooperationspartner verantwortlich für die unabhängige Evaluation, die offizielle Randomisierungsliste generiert (Blockrandomisierung ohne Stratifizierung), welche die 12 Cluster den drei Sequenzgruppen zuwies und somit die *zeitliche Rangfolge* bei der Implementierung der telemedizinischen Intervention festlegte (Software: nQuery Advisor V7.0). Jede Randomisierungseinheit bzw. Cluster beinhaltete ein bis drei Intensivstationen (siehe nachfolgende Tabelle 1). Mit der per Randomisierung erfolgten Zuweisung der Häuser eines Clusters zu den 3 Sequenzgruppen wurde auch das Zeitintervall für die 3-monatige Cluster-spezifische *Trainingsperiode* auf der Haus-Ebene definiert, die 3 Monate vor dem geplanten Start der Intervention durchgeführt werden muss. Für die 3 definierten Sequenzgruppen wurden folgende Crossover-Zeitpunkte präspezifiziert:

Tabelle 0. Definierter Startzeitpunkt ("Crossover Datum") für den Rollout der Intervention, pro Sequenzgruppe.

| Allokation | Geplanter Start der Intervention |
|-----------------|----------------------------------|
| Sequenzgruppe 1 | 01.12.2018 |
| Sequenzgruppe 2 | 01.03.2019 |
| Sequenzgruppe 3 | 01.06.2019 |

Hierbei wurde angenommen, dass alle Cluster einer Randomisierungseinheit zu dem durch die Sequenzgruppen-Zugehörigkeit definierten Zeitpunkt gleichzeitig von Routineversorgung (Kontrolle) auf Tele-ICU-Bedingung (Intervention) umstellen können. Zu Abweichungen zwischen dem tatsächlichen Implementierungsstart und dem über die Randomisierung zugewiesenen Crossover-Zeitpunkt und der zugewiesenen Rangfolge der Implementierung finden sich Details im Ergebnissteil, Abschnitt 4.2.

3.4.2 Verblindungsgrad

Da die Prüfarzte vor Ort an der Durchführung und Umsetzung der telemedizinischen Intervention beteiligt waren, war es naturgemäß nicht möglich, das Krankenhaus- und Studienpersonal (sowohl ICU- als auch Follow-up Phase) und die für die Bewertung der patientenindividuellen, täglichen QIs zuständigen Rater zu verblinden. Der für die Evaluation zuständige Konsortialpartner (wissenschaftliche Mitarbeiter des IBE der LMU und der TU Berlin) hatte während der gesamten Studienlaufzeit keinen Zugriff auf die unverblindeten Datenbanken. Das Analysekonzept für die Evaluation basierte auf einem Testdatensatz und wurde a priori spezifiziert (vgl. veröffentlichtes Protokoll *BMJ Open* 2020;10:e036096).

Tabelle 1. Offizielle Randomisierungsliste mit zugewiesener Sequenzgruppe je Cluster (zusammen mit Tabelle 0).

| Cluster ID | ICU Größe (mittlere Bettenzahl) | Bundesland | Anzahl ICUs | Sequenzgruppe | Krankenhaus Bettenzahl |
|------------|------------------------------------|------------|-------------|---------------|------------------------|
| 1 | <15 | BB | 1 | 1 | <300 |
| 2 | >15 | BB | 1 | 2 | ≥700 |
| 3 | >15 | B | 3 | 3 | <300 |
| 4 | <15 | B | 1 | 3 | <300 |
| 5 | >15 | BB | 1 | 1 | ≤699 |
| 6 | <15 | BB | 1 | 2 | ≥700 |
| 7 | >15 | B | 2 | 1 | ≥700 |
| 8 | >15 | B | 1 | 2 | ≥700 |
| 9 | >15 | B | 1 | 3 | ≤699 |
| 10 | >15 | B | 1 | 1 | ≤699 |
| 11 | >15 | BB | 2 | 2 | ≥700 |
| 12 | <15 | B | 1 | 3 | ≤699 |

B = Berlin, BB = Brandenburg

3.5 Definierte Endpunkte (ICU-Phase und post-ICU Follow-up)

3.5.1 Primäre Outcomes

Um zu evaluieren, inwieweit die telemedizinische Intervention der Regelversorgung bei mindestens einem der 8 patienten-bezogenen QIs bzgl. dessen Adhärenz überlegen ist (Definition nach Kumpf *et al.*[5]), wurden 8 gleichberechtigte primäre binäre Wirksamkeitskriterien präspezifiziert, die jeweils aus mehreren QI-bezogenen Key Performance-Parametern abgeleitet werden (Tabelle 2).

Abweichungen vom ursprünglich geplanten Vorgehen (Innovationsfond-Antrag)

Zwei der QIs, QI 5 (Überwachung der Maßnahmen zur Infektionsprävention; Art des Indikators: Ergebnisqualität) sowie QI 10 (Leitung der Intensivstation; Art des Indikators: Strukturqualität) wurden in der Evaluation nicht berücksichtigt und auch nicht erfasst. Hierbei handelt es sich um Indikatoren, die nicht auf der Patientenebene im Rahmen der QI-Visiten auf täglicher Basis gemessen werden, sondern aufgrund ihres Strukturcharakters ICU-bezogen sind. Die Schätzung eines interpretierbaren Effekts der Intervention ist somit für QI 5 und QI 10 nicht möglich.

Jedes dieser 8 primären Zielkriterien kann dann wie folgt definiert werden:

- Adhärenz [erfüllt ja/nein] eines einzelnen krankenhausinternen QI, welcher auf Patientenebene täglich innerhalb eines 24-Stunden-Zeitfensters bewertet wird, beginnend mit dem Tag der ICU-Aufnahme (Tag nach ICU-Aufnahme) bis zur ICU-Entlassung (Tag vor ICU-Entlassung).

Ob ein einzelner QI für den Patienten i am Tag t erfüllt ist oder nicht, wurde anschließend von einem Facharzt mit Zusatzbezeichnung Intensivmedizin im Tele-ICU-Cockpit der Charité beurteilt anhand der zugrundeliegenden elektronischen QI-Dokumentation, d.h. der patientenspezifischen klinischen Parameter ("Rohdaten" bzw. Key-Performance Set, z.B. Beatmungseinstellungen, Medikamente, Antibiotika), die vom Arzt bzw. der Pflegefachkraft am Remote-Standort täglich erfasst wurden. Dieser zentrale Endpunkt-Bewertungsprozess (*endpoint adjudication*) zur Bestimmung des QI-Erfüllungsgrads erfolgte unabhängig davon, ob die jeweilige Intensivstation die Versorgung unter Kontroll- oder Interventionsbedingungen durchführte. Es wurden insgesamt sieben Rater (Ärzte/Oberärzte der Charité mit Zusatzbezeichnung Intensivmedizin) geschult, die die QI-Bewertung und somit die Ableitung des primären Outcomes durchführten.

Die QI-Bewertung erfolgte kontinuierlich fortlaufend. Die Entscheidung der QI-Erfüllung wurde anhand eines algorithmischen Bewertungskataloges der DIVI getroffen und wurde mit einer Entscheidungshilfe unterfüttert. Ebenso wurde der Visitenablauf formalisiert. Die acht untersuchten QIs sind operationalisierbar auf Patientenebene, auf täglicher Basis. Traf ein einzelner QI_k , $k=1, \dots, 8$, am Tag t nicht auf einen bestimmten Patienten i zu (z.B. Beatmung bei nicht-beatmeten Patienten), so wurde dieser immer als „erfüllt“ definiert, da von ihm kein möglicher "Schaden" (bei Nichterfüllung) ausging. Einzelne QIs bestehen aus mehreren Domänen bzw. Teilaspekten. Für die manuelle Ableitung des binären Outcomes (patientenindividuelle Adhärenz [ja/nein] eines einzelnen QIs an einem bestimmten Tag) gilt jeweils die UND-Verknüpfung, d.h. für eine QI-Erfüllung müssen alle zugrundeliegenden Bedingungen erfüllt sein.

Die Objektivität und Reliabilität wurden durch regelmäßige, i. d. R. wöchentliche, Schulungen und Checks gewährleistet. Die Bewertung der QI-Adhärenz wurde grundsätzlich nicht von den Primärbehandlern vorgenommen, sondern von einem unabhängigen Arzt.

Tabelle 2. Konsensus-basiertes Set der 10 Qualitätsindikatoren der DIVI (3. Auflage 2017, siehe [5]), welches zur Definition der 8 binären primären Outcomes verwendet wurde.

| Indikator Nr. | Qualitätsindikatoren Intensivpflege | Kriterien für die QI-Adhärenz |
|---------------|---|---|
| QI 1 | Tägliche multiprofessionelle und interdisziplinäre klinische Visite mit Dokumentation von Tageszielen | Tägliche ärztliche Visite mit einem multiprofessionellen Team und vorgegebenen Zielen |
| QI 2 | Management von Sedierung, Analgesie und Delir | Beurteilung 1) des Sedierungsgrads, 2) eines möglichen Delirs und 3) des Schmerzgrades mit geeigneten Scoring-Tools |
| QI 3 | Patientenadaptierte Beatmung | Bei mechanischer Beatmung Anwendung eines niedrigen Tidalvolumens, angemessener Ventilationsdruck |
| QI 4 | Frühzeitige Entwöhnung von einer invasiven Beatmung (Weaning) | Tägliche Bewertung des Weaning-Potenzials und standardisierte Spontanatmungsversuche |
| QI 5 * | <i>Überwachung der Maßnahmen zur Infektionsprävention</i> | — |
| QI 6 | Maßnahmen zum Infektionsmanagement | Frühzeitige, empirische Antiinfektiva-Therapie; frühzeitige mikrobiologische Testung; Vermeidung einer nicht indizierten Antiinfektiva-Therapie; therapeutisches Drug Monitoring |
| QI 7 | Frühe enterale Ernährung | Frühzeitige Ernährung mit patientenspezifischen Kalorienzielen; Anwendung von 50% des eingestellten Kalorienziels innerhalb der ersten 48 Stunden nach der Aufnahme auf die Intensivstation |
| QI 8 | Dokumentation einer strukturierten Patienten- und Angehörigenkommunikation | Dokumentierte Kommunikation mit Angehörigen oder dem Bevollmächtigten des Patienten; adäquater Inhalt einschließlich Berücksichtigung der persönlichen Präferenzen des Patienten |
| QI 9 | Frühmobilisation | Frühmobilisation innerhalb der ersten 72 Stunden nach Aufnahme auf die Intensivstation und dann tägliche Physiotherapie |
| QI 10 * | <i>Leitung der Intensivstation</i> | — |

* Die Qualitätsindikatoren QI 5 und QI 10 wurden nicht als primäre Outcomes definiert, da diese nicht auf Patientenebene im Rahmen der QI-Visiten auf täglicher Basis gemessen werden, und somit keine Schätzung eines interpretierbaren Effekts der Intervention möglich ist.

3.5.2 Sekundäre Outcomes

Gemäß Protokoll war geplant, sekundäre Outcomes der Post-ICU Follow-up Phase (*Sustainment Phase*) zu den Zeitpunkten 3 und 6 Monate nach ICU-Entlassung, d.h. nach Exposition des Patienten im Rahmen des Index-ICU-Aufenthalts (erster studienbezogener ICU-Aufenthalt), zu erfassen.

Sekundäre Endpunkte wurden präspezifiziert in Anlehnung an das Core Outcome Set Framework in Bezug auf die PICS-bezogenen Domänen mentale Gesundheit, Mobilität/physische Gesundheit, und Kognition, sowie die gesundheitsbezogene Lebensqualität und Mortalität.[4, 19]. Neben den Outcomes zu funktionellen Langzeitbeeinträchtigungen wurden für die gesundheitsökonomische Evaluation weitere sekundäre Outcomes definiert in Bezug auf die Inanspruchnahme der Gesundheitsversorgung und des sozioökonomischen Status.

Tabelle 3 benennt die in Abschnitt 2.1 genannten PICS-bezogenen Domänen und weitere sekundäre und explorative Outcomes, und beschreibt die zugehörigen Messinstrumente (inkl. Patienten-Fragebögen bzw. Tests; Hausarztfragebogen) zu deren Ableitung. Aufgelistet sind *Outcomes zur 'Clinical Effectiveness', gesundheitsökonomische Outcomes*, sowie auch die im Rahmen eines erweiterten Assessments eingesetzten Messinstrumente, die nicht Teil der Evaluation für diesen Bericht sind.

Für die beiden PICS-Domänen Kognition und körperliche Fähigkeiten wurden für die Analysen binäre sekundäre Outcome Measures ("auffällig" ja/nein) abgeleitet anhand der zugrundeliegenden rohen Punkt- bzw. Score-Werte der zugrundeliegenden Testinstrumente:

- Für die beiden Tests zur Kognition wurden etablierte Cut-offs zugrunde gelegt:
 - Mini-Cog Test "auffällig", falls sich in der Gesamtbewertung <3 Punkte ergeben (Range: 0 – 5 Punkte);
 - Animal Naming Test "auffällig", falls sich in der Gesamtbewertung <14 Punkte ergeben (Range: Anzahl der genannten Tiere ≥0).
- Für die resultierenden Punktwerte der beiden Tests zur Messung der körperlichen Funktionsfähigkeit wurden geschlechts- und/oder altersabhängige Cut-offs definiert entsprechend der Literatur.

Details zum Auswertungsalgorithmus zur Ableitung der dichotomisierten Outcomes für den Timed-Up-and-Go (TUG)-Test ("eingeschränkte Mobilität" ja/nein) sowie den Handkraft (HGS)-Test ("eingeschränkte Kraft" ja/nein) befinden sich im Anhang A.3 sowie in dem im Rahmen des ERIC Projekts entwickelten veröffentlichten Proposal zur Messung von PICS im ambulanten Setting (vgl. [9]).

Tabelle 3. Sekundäre und explorative Wirksamkeits-Endpunkte, die in einem Case-Care-Management im Rahmen von Follow-up Untersuchungen erhoben oder über andere Datenquellen erfasst wurden.[20]

| Domäne / Outcome | Zeitpunkt | Messinstrument |
|--|-----------------|---|
| Gesamtmortalität | bis zu 6 Monate | Anzahl der Todesfälle jeglicher Ursache, einschließlich der ICU-Mortalität (unter Verwendung von Krankenhaus- und der Follow-up Dokumentation zu Monat 3 und 6 (Hausärzte, Angehörige)). Zusätzlich erfolgte eine Abfrage der Sterbedaten in den örtlichen Melderegistern. |
| Psychische Gesundheit Subdomänen: Angst und Depression | Monat 3 und 6 | Die patientenberichtete Symptombelastung wurde mit dem papierbasierten Patient-Health-Questionnaire PHQ-4 (Screening-Tool für Angst und Depression) erfasst. Höhere Werte deuten auf eine höhere Belastung hin.[21] |
| Psychische Gesundheit Subdomäne: Posttraumatische Belastungsstörung (PTSD) † | Monat 6 | Die patientenberichtete Symptombelastung wurde mit dem papierbasierten Fragebogen Impact of Event Scale Revised (IES-R) erfasst. Höhere Werte deuten auf eine größere Belastung hin.[22] |
| Kognition Subdomänen: Gedächtnis, visuospatiale und visuokonstruktive Fähigkeiten | Monat 3 und 6 | Das funktionelle Outcome wurde mit dem Mini-Cog Test gemessen (2 Tests: Erinnerungs-Aufgabe mit 3 Wörtern; Aufgabe Zeichnen einer Uhr). Höhere Werte bedeuten eine bessere kognitive Funktion.[23-25] |
| Kognition Subdomäne: Sprachgewandtheit | Monat 3 und 6 | Das funktionelle Outcome wurde mit dem Animal-Naming Test gemessen. Höhere Werte bedeuten eine bessere kognitive Funktion.[26] |
| Körperliche Funktion Subdomäne: Mobilität | Monat 3 und 6 | Messung der Gehfähigkeit und des Sturzrisikos anhand des Timed Up & Go (TUG) Tests (Zeitdauer für Aufstehen und Gehen). Höhere Werte weisen auf eine größere Beeinträchtigung hin.[27] |
| Körperliche Funktion Subdomäne: Muskelkraft (Handkraft) | Monat 3 und 6 | Messung der Muskel- und Nervenfunktion des Patienten anhand des Hand-Grip-Strength (HGS) Tests mittels eines Hand-Dynamometers (maximale Wert der Handkraft [in kg] über 3 Versuche, für die dominante Hand).[28, 29] Höhere Werte weisen auf eine geringere Beeinträchtigung hin. |
| Gesundheitsbezogene Lebensqualität (HRQoL) * | Monat 3 und 6 | Die patientenberichtete HRQoL wurde anhand des EuroQol – 5 Dimensionen – 5 Level (EQ-5D-5L) Fragebogens erfasst: Deskriptives System (5 Items: Mobilität, Selbstversorgung, allgemeine Tätigkeiten, Schmerz/ körperliche Beschwerden, Angst/ Niedergeschlagenheit); Visuelle Analogskala (EQ VAS), eine Thermometer-ähnliche Bewertungsskala von 0 (schlechtester vorstellbarer Gesundheitszustand) bis 100 (bester vorstellbarer G.zustand). Höhere Werte (deskriptives System; VAS) deuten auf eine bessere HRQoL hin.[30, 31] |
| Organdysfunktion * | Monat 3 und 6 | Anzahl der Patienten mit Organfunktionsstörungen, Beurteilung durch Hausarzt oder Studienpersonal |
| Pulmonale Funktion und Symptome Subdomäne: Dyspnoe † | Monat 6 | Patientenberichtete Atemnot bei täglichen Aktivitäten, bewertet mit der Modified British Medical Research Council (mMRC) Dyspnea Scale [Range 0 – 4]. Höhere Werte weisen auf eine höhere Beeinträchtigung hin.[32, 33] |

| Domäne / Outcome | Zeitpunkt | Messinstrument |
|--|-------------------------------------|---|
| Ambulante Beatmung * | bis zu 6 Monate nach ICU-Entlassung | Gesamtdauer [in Tagen] der mechanischen Beatmung |
| Funktionsfähigkeit und Behinderung † | Monat 6 | Allgemeiner Disability-Score, gemessen mit dem WHO Disability Assessment Schedule (WHODAS 2.0) für Aktivitäts- und Teilhabeeinschränkung, 12-Item Kurzversion, Selbstausfüller-Fragebogen. Simple scoring erfolgt durch Aufsummierung der Werte pro Item. Höhere Score-Werte deuten auf eine höhere Beeinträchtigung hin.[34, 35] |
| Dauer des Aufenthalts auf der ICU (LOS) * | Monat 6 | Gesamtzahl der auf der ICU verbrachten Tage |
| Beschäftigungsstatus † | Monat 3 und 6 | Wiederaufnahme der Erwerbstätigkeit, oder Änderung des Beschäftigungsstatus (erfasst über Hausarztfragebogen) |

* relevant für Analysen zur gesundheitsökonomischen Evaluation.

† Erweitertes PICS-bezogenes Assessment. Diese Outcomes waren kein Bestandteil der Evaluation, da diese nicht im Projektantrag geplant waren. Diese Daten wurden für weitere Forschungsvorhaben erfasst.

Abweichungen vom ursprünglich geplanten Vorgehen (Innovationsfond-Antrag)

Die intersektoralen Qualitätsindikatoren zur Entlassung in eine ambulante Beatmung („QI-net“) und deren Auswertung waren abweichend zum Antrag nicht Teil der Evaluation. Grund dafür ist u.a., dass die ambulanten Daten der Krankenversicherung zu diesen Patienten noch nicht vollständig vorliegen, und eine Auswertung nur sinnvoll ist, wenn der präklinische Status der Patienten ausführlich bekannt ist. Die bislang verfügbaren Daten zur Langzeitbeatmung wurden in diesem Bericht unter der gesundheitsökonomischen Evaluation deskriptiv und inferenzstatistisch ausgewertet.

3.6 Datenerhebung und Datenschutzaspekte: Fragebögen und Case Report Forms

3.6.1 E-Health-Plattform und REDCap-Forschungsdatenbank

Die Datenerhebung und der Verarbeitungszweck der erhobenen Daten war zentrales Element von ERIC. Da die neuen Versorgungsformen sich der Rechtsgrundlage von klinischen Verfahren bedienen und gleichzeitig aber eine wissenschaftliche Evaluation gefordert war, werden einzelne Daten zu wissenschaftlichen Zwecken verarbeitet, aber auch zu klinischen Routinezwecken inkl. des gesetzlich geforderten Qualitätsmanagements. Grundsätzlich werden bei neuen Versorgungsformen klinische Routedaten von den zur Evaluation herangezogenen Daten unterschieden und unterschiedlich dokumentiert.

Diesem Umstand wurde nicht nur mit zwei Einwilligungserklärungen (Versorgungsform und wissenschaftlicher Zweck), sondern auch mit unterschiedlichen Dokumentationsweisen begegnet. Im Rahmen des Projektes wurde eine *E-Health-Plattform* ausgeschrieben und pilotiert, und eine *REDCap-Forschungsdatenbank* aufgesetzt und genutzt (siehe Abbildung 4).

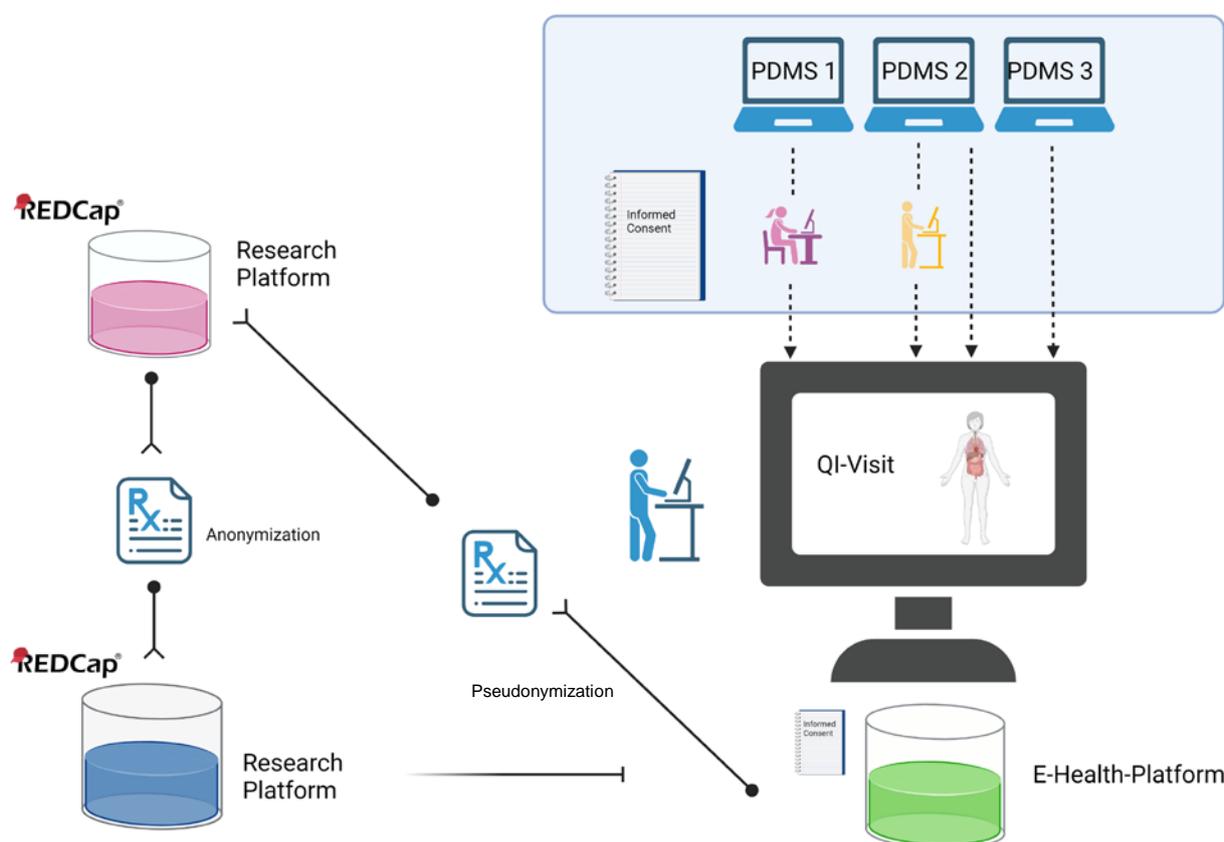


Abbildung 4. Schematische Darstellung des Datenflusses zwischen Remote ICU (teilnehmende Krankenhäuser) und Tele-ICU der Charité aus informatorischer Perspektive. Anhand der dokumentierten patientenindividuellen QI-Rohdaten (festgelegte Key Performance Merkmale) erfolgte eine zentrale Bewertung und Ableitung der Adhärenz bzgl. der 8 QIs durch die Charité.
PDMS: Patientendatenmanagementsystem.

Grundsätzlich können Daten aus der klinischen Routedokumentation nach Einwilligung des Patienten in nicht direkt personenbeziehbarer Form in die Forschungsdatenbank exportiert werden. Der Export kann dann hier mittels einer Patientenidentifikationsliste depseudonymisiert werden. In einem zweiten Schritt kann dieses Pseudonym entfernt und die demographische Daten so verändert werden (kategorisiert), dass eine anonyme Datenbank vorliegt, und diese damit auch

nicht mehr in den Geltungsbereich der Datenschutz-Grundverordnung (DSGVO) der EU fällt. Dieser Datenmanagement-Prozess wurde im Hybridprojekt ergänzt und durch alle beteiligten Konsortialpartner gestaltet und verantwortet. Offene Schnittstellen sind hier in jeder Datenbank gewährleistet.

3.6.2 BARMER Wissenschafts-Data Warehouse (W-DWH) [TU Berlin]

Die BARMER stellt pseudonymisierte Routinedaten zur Verfügung. Dies erfolgt durch einen Zugriff auf das Wissenschafts-Data-Warehouse (W-DWH) der BARMER. Hierbei handelt es sich um eine vollständig nach außen abgegrenzte technische Infrastruktur von virtuellen Rechnern, die einen Zugriff auf vollständig pseudonymisierte Routinedaten der BARMER erlauben. Die Pseudonymisierung umfasst sowohl die Versichertendaten als auch die der Leistungserbringer, so dass die höchstmögliche Sicherung persönlicher Daten gewährleistet ist.

Der Datenschutz wird durch ein Einzelplatzangebot, sowie ein spezifisches Rechtesystem vervollständigt.

Das W-DWH ermöglicht sektorenübergreifende Langzeitanalysen, da folgende Daten vorgehalten werden:

- Stammdaten
- Arzneimitteldaten
- Ambulante Arztdaten
- Krankenhausdaten (stationär und ambulant)
- Heil- und Hilfsmittel
- Pflege
- Arbeitsunfähigkeit / Krankengeld
- Zahndaten

Die Anmeldung von externen Nutzern außerhalb der BARMER-Netzwerke erfolgt nur über sichere Virtual Private Networks (virtuelles privates Kommunikationsnetz, VPN) und in Verbindung mit personifizierten VPN-Clients und Secure-ID-Token. Die Bewirtschaftung des W-DWH mit den aktuellen Rohdaten erfolgt regelmäßig im Jahresverlauf.

3.6.3 Hausarztfragebogen [TU Berlin]

Die Erhebung der Kosten seit Entlassung von der Intensivstation und des sekundären Endpunkts *Vorliegen außerklinischer Beatmung* erfolgte mithilfe eines papierbasierten Fragebogens (Case Report Forms (CRFs) bzw. HA-CRF 1 und HA-CRF 2), der zum FU 1 und 2 an die behandelnden Hausärzte versendet wurde. Waren Hausärzte nicht bekannt bzw. konnten diese nicht ausfindig gemacht werden, wurde der Fragebogen vom Studienpersonal während der FU-Assessments ausgefüllt.

Nach der genehmigten Erweiterung auf den Einschluss von Patienten über einen Behandlungsvertrag nach §630 BGB wurde der Fragebogen so angepasst und erweitert, dass auch für nicht bei der BARMER versicherte Patienten entsprechende Daten auf der Kostenseite erhoben werden können.

Abweichungen vom ursprünglich geplanten Vorgehen

Vereinzelte sollten Kostendaten auch von Versicherten der BARMER auf die gesamte Studienpopulation übertragen werden, z. B. Fahrtkosten (direkte nicht-medizinische Kosten). Da diese Daten jedoch zum Auswertungszeitpunkt noch nicht im BARMER W-DWH (siehe 3.6.2) vorgelegen haben, konnten diese nicht in die Analysen einbezogen werden. Da jedoch Transporte mit 2% generell nur einen geringen Teil der gesamten Gesundheitsausgaben in 2019 ausgemacht haben,[36] ist nicht davon auszugehen, dass die Ergebnisse durch deren fehlende Einbeziehung maßgeblich beeinflusst werden. Darüber hinaus sollen Krankenhausabrechnungsdaten herangezogen werden, um die Kosten des Indexaufenthalts zu ermitteln. Diese wurden jedoch nicht systematisch erfasst, sodass hier zur Bewertung der Kosten ein standardisierter Bewertungssatz herangezogen wurde.

Der Fragebogen (Anhang 11 des Ergebnisberichts) besteht aus vier Teilen. Im ersten Teil wurden Angaben zur Person abgefragt. Dazu gehören etwa Angaben zur Arbeitsfähigkeit, zum Aufenthaltsort, zum Pflegebedarf und dazu, ob eine Beatmungspflicht vorliegt. Der zweite Teil beinhaltet Fragen zu Rehabilitationsmaßnahmen, zu stationären Wiederaufnahmen, zur Unterbringung in stationären Pflegeeinrichtungen und zu in die Behandlung involvierten ärztlichen und nicht-ärztlichen Gesundheitsberufen. Der dritte Teil umfasst Angaben zur aktuellen Gesundheit und zu möglichen intensivstationären Folgen (PICS). In Teil vier wurden Dokumente aufgelistet, die gemeinsam mit dem ausgefüllten Hausarztfragebogen an das Studienzentrum geschickt werden sollten, darunter etwa Arztbriefe, Diagnoselisten sowie Medikamentenpläne. Darüber hinaus wurde erfasst, wer den Hausarztfragebogen ausgefüllt hat.

Lagen Fragebögen von verschiedenen ausfüllenden Personen vor, wurde der vom Studienpersonal ausgefüllte Fragebogen prioritär behandelt, gefolgt von Hausarzt, MFA, anderer und unbekannter ausfüllender Person. Hausarztfragebögen wurden nur von jenen Patienten berücksichtigt, die auch zum jeweiligen FU erschienen sind.

3.7 Datenmanagement, Qualitätskontrolle und Datensicherheit

3.7.1 REDCap-basierte Datenbanken für ICU- sowie FU-Phase

QI-Datenbank

Für die während eines ICU-Aufenthaltes (im Folgenden bezeichnet als *Fall*) auf täglicher Basis erfassten patientenindividuellen Studiendaten auf Fall-Ebene wurde eine QI-Datenbank mittels des eCRF-Systems REDCap® (Research Electronic Data Capture; Version 10.6.14, Vanderbilt University, Nashville, Tenn., USA) aufgesetzt, welches an der Charité am Standort der Konsortialführung gehostet wurde (<https://redcap.charite.de/eric/>). REDCap ist eine sichere, webbasierte eCRF-Plattform zur Verwaltung von Forschungsdaten und bietet 1) eine intuitive Schnittstelle für die validierte Dateneingabe, 2) Audit Trails zur Nachverfolgung von Datenmanipulations- und Datenexport-Aktivitäten, 3) automatisierte Exportverfahren für Daten-Downloads für etablierte Statistik-Programmpakete und 4) Verfahren für den Import von Daten aus externen Quellen.[16]

Um während des ERIC-Projektes eine hohe Datenqualität zu gewährleisten, erfolgte die Dateneingabe durch geschultes Personal und studentische Hilfskräfte an der Charité direkt in die

REDCap Forschungsplattform. Die QI-Daten betreffend erfolgten simple, automatisierte Plausibilitätschecks während der Dateneingabe in REDCap, sowie manuelle Checks vor Ort auf der ICU, um unplausible oder fehlende Daten bei ausgewählten Variablen rechtzeitig zu identifizieren. Weiterhin wurden vor Datenbankschluss die Bewertungen der 8 QIs weiteren umfassenden Qualitätsschecks unterzogen, indem automatisierte Reports erstellt wurden (fehlende/ unvollständige/ unplausible Eintragungen), und ca. 50% der QI-Visiten hinsichtlich der QI-Bewertung manuell überprüft wurden.

Verantwortlich für technische Implementierung der QI-Datenbank und das Datenmanagement der ICU-Phase war der Konsortialpartner FOKUS.

FU-Datenbank

Die post-ICU Follow-up Daten zu Monat 3 und 6 wurden zunächst über papierbasierte CRFs und Patienten-Fragebögen sowie einem Hausarzt-Fragebogen erfasst, bevor sie in einer separaten und von der QI-Datenbank technisch getrennt geführten REDCap FU-Datenbank digitalisiert wurden. Verantwortlich für Implementierung der FU-Datenbank, das Datenmanagement der FU-Phase und Zusammenführung mit den Daten der ICU-Phase war die Charité in Zusammenarbeit mit dem Konsortialpartner FOKUS sowie dem Institut für Medizinische Informatik der Charité.

Die ICU- und post-ICU-Mortalität wurden validiert, indem eine Melderegisterabfrage durchgeführt wurde, um insbesondere Post-ICU Todesfälle zu identifizieren.

Zwischen dem 04.12.2020 und dem 14.12.2020 (98% der Patienten) sowie einmalig am 03.02.2021 (2% der Patienten) fand eine Abfrage der Berliner und Brandenburger Melderegister, und in geringerer Zahl der Melderegister anderer Bundesländer statt. Hierbei konnte jeder ERIC-Patient identifiziert und, falls vorhanden, das Sterbedatum identifiziert, und somit der Survival-Status aller 1463 Patienten als verifiziert angenommen werden. Die Sterbedaten wurden zur Auswertung der sekundären Endpunkte in die FU-Datenbank übertragen. Bei Diskrepanzen zwischen der FU- und der QI-Datenbank, den Survival-Status oder das Sterbedatum von ICU-Patienten betreffend, wurden die anhand des Melderegisters überprüften Mortalitätsdaten der FU-Datenbank für Analysen verwendet.

Die finalisierten zur Datengenerierung genutzten CRFs und Patienten-Fragebögen sind ebenfalls Bestandteil des Ergebnisberichts (Anhang 11).

3.7.2 Datenzugriff für die Evaluation (Clinical Effectiveness)

Grundlage für die datenschutzrechtliche Bewertung des Projektes ist die während der Projektzeitlauf 2018 in Kraft getretene Datenschutz-Grundverordnung der EU. Alle IT-System-Installationen wurden im Einklang mit der DSGVO durchgeführt und vom zuständigen Datenschutzbeauftragten der Charité begleitet.

Die QI-Datenbank mit den für die Evaluation erforderlichen Daten wurde dem IBE als Datenbank-Klon ohne Personenbezug zur Verfügung gestellt (technische Umsetzung: FOKUS). Für die FU-Phase wurde dem IBE ein Datensatz mit ausgewählten Variablen per REDCap File Repository ohne Personenbezug zur Verfügung gestellt (siehe hierzu auch Anhang A.1).

3.8 Effektevaluation: Statistische Methoden

3.8.1 Initiale Fallzahlplanung

Aufgrund des Cluster-randomisierten Designs wurde eine eher pragmatische Fallzahlplanung auf Cluster-Basis (Klinikträger) durchgeführt, wobei gleiche Clustergrößen angenommen wurden, bei insgesamt begrenzter Anzahl an potenziellen Zentren und Ressourcen. Aufgrund der acht binären ko-primären Outcomes (d.h. keine hierarchische Ordnung der QIs) wurde eine Bonferroni-Korrektur für multiples Testen angewandt, bei einem globalen einseitigen Fehler 1. Art von $\alpha = 5\%$. Somit ergibt sich ein Signifikanzniveau von $\alpha/8 = 0.625\%$ für confirmatorisches Testen eines einzelnen QIs. Es wurde ein minimaler klinisch relevanter Unterschied von 10% in der QI-Adhärenz spezifiziert; diese zu entdeckende Differenz zwischen den beiden Behandlungsbedingungen orientierte sich an der geeigneter Literatur.[37, 38]

Ein 2-Stichproben χ^2 -Test mit einem einseitigen Signifikanzniveau von 0.625% hat eine Power von 82%, um einen Unterschied in der QI-Adhärenz in Gruppe 1 (Kontrollbedingung) von $\pi_1=60\%$ und einer Adhärenz in Gruppe 2 (Interventionsbedingung) von $\pi_2=70\%$ zu entdecken (Odds Ratio (OR) = 1.556), wenn die Fallzahl 530 Patienten pro Gruppe beträgt, bei Annahme unabhängiger Beobachtungen (Software: nQuery Advisor V7.0).

Um die Korrelation zwischen Patienten aus demselben Cluster zu berücksichtigen wurde ein Designeffekt (Varianzinflationsfaktor) von 1.35 geschätzt, zusammen mit einem Intracluster-Korrelationskoeffizienten von $ICC=0.117$ (abgeleitet aus unveröffentlichten Daten der ICUs der Charité zum QI 2), der die Korrelation zwischen Beobachtungen innerhalb desselben Clusters misst.[39, 40] Für diese CRT ergibt sich somit eine Gesamt-Fallzahl (Patientenebene) von 1431 GKV-versicherten Patienten.

Aufgrund praktischer Überlegungen zu Aufnahmezahlen und Kapazitäten der teilnehmenden Intensivstationen wurde folgendes Setting festgelegt: 12 Cluster (4 pro Sequenzgruppe) mit ca. 163 Fällen (d.h. ICU-Aufnahmen) pro Cluster und eine Gesamt-Fallzahl von 1956 Fällen (12x163). Hierbei wurden folgende weitere Annahmen getroffen:

a) ICU-Wiederaufnahmerate von 20%, b) 1 Cluster pro Sequenzgruppe als Aufschlag und Ausgleich für Dropout auf Cluster-Ebene oder für unzureichende Rekrutierung (Fallzahlplanung ergab 9 Cluster insgesamt, 3 pro Sequenzgruppe). Details hierzu sind im veröffentlichten Protokoll mit zugehörigem methodischem Supplement 4 genauer beschrieben.[20]

Bei dieser Fallzahlplanung wurden die Trainingsphase sowie variierende Cluster-Größen nicht berücksichtigt. Insgesamt gehen wir davon aus, dass die der Fallzahlplanung zugrundeliegenden Annahmen eher konservativ waren, um eine absolute Änderung der Adhärenz von 10% in mindestens einem der acht QIs zu entdecken.

3.8.2 Analyse-Strategien zur Clinical Effectiveness

Die primären Endpunkte werden nach dem Intention-to-treat-Prinzip (ITT) ausgewertet, und die Cluster werden als exponiert betrachtet (Behandlung unter Interventionsbedingung) ab dem tatsächlichen Datum des Treatment Switch (Crossover Datum). Anstatt einer Intention-to-treat Analyse im Sinne des über die Randomisierung definierten Zeitplans wurde für die Hauptanalyse eine sogenannte "as-implemented" Analysestrategie gewählt,[41] und Beobachtungen, die aufgrund der Verschiebung des dritten Steps um 3 Monate sowie in Cluster #6 (Sequenzgruppe 2) um 8 Monate entgegen der ursprünglichen Planung weiterhin unter der Kontroll-Bedingung erfasst wurden, unter der tatsächlichen Kontroll-Bedingung analysiert. Diese Vorgehensweise entspricht der Annahme einer durch die Randomisierung zugewiesenen Reihenfolge für den Rollout (anstatt eines über die Randomisierung zugewiesenen Crossover Datums, d.h. eines exakten Zeitpunkts für den Übergang von der Kontroll- in die Interventionsphase mit neuer Versorgungsform).[2]

Analysepopulation

Alle Patienten einer Randomisierungseinheit, die in der QI-Datenbank erfasst wurden und für die mindestens eine QI-Messung dokumentiert wurde, wurden in die Full Analysis Set (FAS) Population eingeschlossen.

Es war zu erwarten, dass die meisten Patienten entweder unter Kontroll- oder Interventionsbedingung während ihres *Index ICU-Aufenthalts* (definiert als erster studienbezogener ICU-Aufenthalt in einer der teilnehmenden Cluster) behandelt werden. Um Kontaminationen (innerhalb eines Clusters) zu vermeiden sollten "Crossover-Patienten" (Patienten, die beiden Bedingungen ausgesetzt waren, z.B. falls sie kurz vor dem zugewiesenen Crossover-Datum auf die ICU aufgenommen wurden) vermieden und gemäß Protokoll weiterhin unter der Kontrollbedingung behandelt werden. Wurde ein Patient zu einem späteren Zeitpunkt erneut auf die ICU aufgenommen – dokumentiert als neuer *Fall* – wurde er unter der zum jeweiligen Zeitpunkt am Cluster vorherrschenden Bedingung behandelt, die sich von der während des Index-ICU Aufenthalts unterscheiden konnte.

Abweichungen vom ursprünglich geplanten Vorgehen (Protokoll)

Patienten, die unter der Kontrollbedingung in die ICU aufgenommen wurden, später unter der Interventionsbedingung nach dem Crossover eines Clusters erneut aufgenommen wurden und somit beiden Bedingungen ausgesetzt waren, wurden von der Hauptanalyse nicht ausgeschlossen und waren Teil der FAS Population.

Die ursprüngliche Planung im Protokoll sah für diese Fälle einen Ausschluss aus der Hauptanalyse und weiterer explorativer Analysen der Follow-up Phase vor. Aufgrund der geringen Anzahl dieser "Crossover-Patienten" und "Wechselpatienten", der kurzen Exposition und des ITT Prinzips verblieben diese Patienten in der FAS Population.

Details zu den stattgefundenen Wechsel-Situationen befinden sich im Ergebnisteil, Abschnitt 4.2.

3.8.2.1 Konfirmatorische Analysen der primären Outcomes

Für den Wirksamkeitsvergleich bei den acht gleichberechtigten primären Outcomes wurden zweiseitige konfirmatorische Tests auf einem Bonferroni-adjustierten Signifikanzniveau von 0.625% angewandt.

Um mit der Struktur der Missings im Zeitverlauf umzugehen wurde ein Likelihood-basiertes logistisches Mixed-Effects Modell (GLMM) verwendet, für welches sich nicht nur unverzerrte Parameterschätzungen unter Missing at Random (MAR) ergeben, sondern welches auch relativ robuste Ergebnisse bei Verletzung der MAR Annahme liefert.[42-44].

Spezifikation des Hauptmodells und Sensitivitätsanalysen für einen einzelnen QI, und Betrachtung von Confounding des Interventionseffekts mit der Zeit

Als **primäres Modell** (Notation im weiteren Verlauf: PM) für alle 8 QIs wurde jeweils ein binäres Mixed Effects Modell mit logit-Link spezifiziert, aus dem die Schlussfolgerungen für die klinische Evaluation abgeleitet werden. Zusätzlich wurden fünf sekundäre Modelle als **Sensitivitätsanalysen** definiert (Notation: SM.1, SM.2, SM.3, SM.4, SM.5).

Das Hauptmodell sowie die Sensitivitätsanalysen wurden motiviert durch zwei zeitliche Aspekte (Confounding-Effekt der Zeit, Langzeittrends), die aufgrund des Stepped Wedge Designs und des sequentiellen Rollouts berücksichtigt werden müssen:[2, 45, 46]

- externe Zeit bzw. der *Effekt der Sequenzgruppe* (Referenz: mittlere Sequenzgruppe), definiert durch die bei der Randomisierung zugewiesene Reihenfolge des Treatment Switch (Implementierung der Intervention) und somit die Zeit zwischen den Clustern;
- interne Zeit (Zeit innerhalb der Cluster), definiert als die Cluster-spezifische *Expositions-Zeit*, d.h. die Zeit [in Monaten] seit (tatsächlichem) Start der Intervention; negativ vor Start der Intervention. Für diese stetige Zeitvariable wurde ein linearer Effekt in Bezug auf die Adhärenz-Wahrscheinlichkeit angenommen. Hiermit kann ein kumulativer Effekt der Intervention untersucht werden.

Übersicht QI-Auswertung: Spezifikation der binären Mixed Effects Modelle und Adjustierung für zeitliche Effekte

Primäres Modell

PM: $\sim \text{trt} * \text{t.0} + (1 | \text{cluster}) + (1 | \text{pid})$

Sensitivitätsanalysen

| | |
|-------------------------------|--|
| SM.1: Sekundäres Modell Nr. 1 | $\sim \text{trt} * \text{t.0} + (1 \text{pid})$ |
| SM.2: Sekundäres Modell Nr. 2 | $\sim \text{trt} * \text{seq.gr} + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.3: Sekundäres Modell Nr. 3 | $\sim \text{trt} * \text{seq.gr} + (1 \text{pid})$ |
| SM.4: Sekundäres Modell Nr. 4 | $\sim \text{trt} * (\text{t.0} + \text{seq.gr}) + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.5: Sekundäres Modell Nr. 5 | $\sim \text{trt} * \text{t.0} + (1 \text{cluster})$ |

trt: binäre Faktorvariable für Treatment, d.h. Behandlung unter Intervention (vs. Referenzkategorie: Kontrolle)
t.0: metrische Variable für die Cluster-spezifische Expositionszeit in Monaten (negativ vor Interventionsstart)
seq.grp: Faktorvariable für die 3 Sequenzgruppen (Referenzkategorie: mittlere Sequenzgruppe).

Es wurde nicht zusätzlich adjustiert für die *Kalenderzeit* (kontinuierliche Zeit seit Studienstart, z.B. studienunabhängige Zeittrends). Zeitliche Effekte fließen über die Sequenzgruppe in die Modelle S.2, S.3 und S.4 als diskreter Faktor mit ein, um den Effekt eines frühen Starts (Sequenzgruppe 1) oder späten Starts (Sequenzgruppe 3) im Vergleich zu einem mittleren Start zu untersuchen (Sequenzgruppe 3 startet zu einem durchschnittlichen Zeitpunkt bzgl. der Rekrutierungsphase).

Weiterhin wurden zwei Zufallseffekte (normalverteilte Random Intercepts) auf Cluster- sowie Patienten-Ebene betrachtet. Im Gegensatz zu SM.2 und SM.4 berücksichtigen die beiden Modelle SM.1 und SM.3 lediglich patientenspezifische Random Intercepts, d.h. keine Clusterbezogene Variabilität bzgl. Intercept (Baseline-Level). SM.5 betrachtet ausschließlich Cluster-spezifische Random Intercepts, d.h. keine Variabilität im Level auf Patienten-Ebene. Damit ist es möglich, mit der für einzelne QIs erwartbaren geringen Variabilität zwischen den Clustern umzugehen.

Spezifikation des primären Modells: Das primäre Modell betrachtete beide Zufallseffekte (d.h. sowohl Cluster- als auch patientenspezifische Intercepts), und als feste Effekte (fixed effects) einen internen Zeiteffekt (kontinuierliche Zeitvariable "time" (t_0), definiert als die Zeit seit Start der Intervention [in Monaten] auf Clusterebene, negativ vor Start der Intervention), den Interventionseffekt "trt" (Referenzkategorie: Kontrollbedingung), und dessen Wechselwirkung mit dem Zeiteffekt. Diese Zeit×Treatment Wechselwirkung beschreibt Lern- und Abschwächungseffekte oder kumulative Effekte nach der Implementierung der Intervention. Ein mögliches Szenario wäre ein starker Effekt der Intervention kurz nach dem Rollout, welcher sich im weiteren Verlauf abschwächt, je länger die Implementierung zurückliegt.

Wie zuvor beschrieben wurde konsequent eine *as-implemented* Analyse durchgeführt,[41] d.h. für die Expositionszeit wurde das tatsächliche Datum des Treatment Switch als Nullpunkt verwendet, nicht der bei der Randomisierung geplante Zeitpunkt. Somit wurden bei den "Late Adopter Sites", d.h. Cluster 6 (Sequenzgruppe 2; geplanter Start: 01.03.2019) sowie den drei Clustern von Sequenzgruppe 3 (geplanter Start: 01.06.2019), im Zeitraum des verzögerten Starts der Intervention (8 bzw. 3 Monate) QI-Bewertungen unter der tatsächlichen Kontroll-Bedingung ausgewertet. In einer *Intention-to-treat* Analyse im strengen Sinne würde die Analyse strikt den Vorgaben des Protokolls und den dort aufgeführten Zeitstrukturen (Datum für Treatment Switch) gemäß Randomisierungsplan folgen.

Die während der 3-monatigen Trainingsphase gemessenen QI-Bewertungen wurden als Beobachtungen unter Kontrollbedingungen analysiert, da ein vernachlässigbarer Effekt der Exposition zu erwarten war wegen Fehlen der telemedizinischen Komponente innerhalb dieses Zeitintervalls (Implementierung der Intervention erst am Ende der Trainingsphase abgeschlossen und anwendbar). Es wurde nicht zusätzlich adjustiert für Exposition nach Trainingsstart.

Für jeden QI wird das gleiche gemischte Modell "PM" mit zeitkontinuierlichem Ansatz gefittet. Als relative Effektmaße werden Odds Ratios (ORs) für die QI-Adhärenz unter Intervention vs. Kontrollbedingung mit zugehörigen 95% Konfidenzintervallen (KIs) berichtet, um Interventionseffekte zu quantifizieren. Es wurden keine multiplen Imputationsverfahren angewandt, stattdessen Missingness at Random (MAR) angenommen für die Fälle, bei denen der ICU-Aufenthalt nicht durch tägliche Messungen vom Tag der ICU-Aufnahme bis ICU-Entlassung abgedeckt war.[43]

Die statistischen Analysen wurden mit dem Softwarepaket R Version 4.0.4 unter RStudio Server Pro Version 1.3.1073-1 durchgeführt.

3.8.2.2 Explorative Analysen der sekundären Outcomes

Die folgenden Analysen werden aus verschiedenen Gründen *explorativ* genannt. Es wird nicht adjustiert für multiples Testen, alle aufgeführten p-Werte sind rein deskriptiv zu interpretieren. Erstens handelt es sich um sekundäre Endpunkte; zweitens findet zwischen ICU-Phase und Follow-up eine Selektion von Patienten statt (vgl. z.B. Colantuoni E *et al.* [47]). Nicht alle Patienten sind in der FU-Phase verfügbar, ein beträchtlicher Anteil ist verstorben, einige haben die Teilnahme am FU abgelehnt, manche konnten im vordefinierten Zeitintervall aus unterschiedlichen Gründen nicht gefunden, das FU-Assessment somit nicht per Protokoll durchgeführt werden. Es ist davon auszugehen, dass das reduzierte Patientenkollektiv im FU nicht das Patientenspektrum der ersten, für die Hauptfragestellung relevanten ICU-Studienphase repräsentiert.

Quantitative und binäre sekundäre Outcomes

Zur Untersuchung des Interventionseffekts in Bezug auf die im post-ICU Follow-up Zeitraum gemessenen sekundären Outcomes wurden explorative *Survivors* Analysen (im Sinne von Colantuoni E *et al.* [47]) durchgeführt. Als Analysepopulation (FAS) wurden nur die ICU-Überlebenden berücksichtigt. Weiterhin handelte es sich um *available-cases* Analysen, d.h. es wurden für diese rein explorativen Analysen alle vorhandenen Follow-up Assessments berücksichtigt, unter Annahme von Missing at Random (MAR).[43]

In Abhängigkeit vom Skalenniveau (metrisch oder binär) wurden sekundäre Outcomes mit (frequentistischen) linearen oder logistischen Mixed Effects Modellen analysiert, sofern diese zu beiden Zeitpunkten erhoben wurden. Ähnlich wie im Modell für die primären binären Outcomes der ICU-Phase wurden zwei Zufallseffekte (normalverteilte Random Intercepts) auf Cluster- sowie Patienten-Ebene betrachtet. Als fixed effects wurde die Intervention, die Zeit seit Index-ICU Entlassung (in Tagen) und die Wechselwirkung zwischen Zeit und Intervention spezifiziert. Die Adjustierung der gemessenen Symptombelastung für die seit ICU-Entlassung vergangene Zeit berücksichtigt die Tatsache, dass für einen relevanten Anteil der Patienten die FU-Assessments nicht im vorgesehenen Zeitfenster (1. FU geplant zu Monat 3, 2. FU zu Monat 6), sondern insbesondere für das 2. FU deutlich später als 6 Monate stattfanden.

Weiterhin wurde für die patientenspezifische Krankheitsschwere adjustiert, indem als fixed effect der zum Baseline-Zeitpunkt (erste vorhandene QI Visite) gemessene SAPS II und SOFA-Score mit aufgenommen wurde (siehe Anhang A.2).

Der anhand der Mixed Effects Modelle geschätzte, interessierende Interventionseffekt wurde für quantitatives Outcome als adjustierte Differenz der Mittelwerte zwischen beiden Behandlungsbedingungen (Intervention vs. Kontrolle), für binäres Outcome als adjustiertes OR berichtet.

Gesamt-Mortalität

Zur Analyse der Gesamtmortalität zu Monat 6 (definiert als Tag 180 nach Index ICU-Aufnahme, d.h. Studieneinschluss) wurde ein Cox-Modell mit Frailty-Term für die durch die einzelnen Cluster definierte hierarchische Struktur der Survival-Daten geschätzt.[48] Hierbei wurde adjustiert für den zum Baseline-Zeitpunkt gemessenen SAPS II sowie SOFA-Score (vgl. Anhang A.2), wobei als Baseline die erste QI-Visite als eine Form von Single Imputation (first observation carried backward) für die zum Zeitpunkt der ICU-Aufnahme (interpretierbar als "natürlicher" Baseline-Zeitpunkt) systematisch fehlende Messung verwendet wurde. Der interessierende Effektschätzer wurde als Hazard Ratio (HR) berichtet.

3.9 Prozessevaluation

Gemäß Projektantrag war die Prozessevaluation kein offizieller Bestandteil der Evaluation. Eine Prozessbeschreibung befindet sich im Ergebnisbericht.

3.10 Gesundheitsökonomische Evaluation – Methoden [TU Berlin]

3.10.1 Gesundheitsökonomisches Studiendesign und Perspektive

Die gesundheitsökonomische Evaluation erfolgte aus der Sozialversicherungsperspektive, wobei die Krankenversicherungsperspektive im Mittelpunkt stand. Hinsichtlich durchgeführter Rehabilitationsmaßnahmen wurde auch die Rentenversicherungsperspektive berücksichtigt und die Kosten gemeinsam mit denen aus der Krankenversicherungsperspektive dargestellt, da der Kostenträger der einzelnen Rehabilitationsmaßnahmen nicht bekannt war (relevant für die neurologische Frührehabilitation Phase C und die Anschlussheilbehandlung (AHB)). Die Perspektive der Pflegeversicherung wurde in die Krankenversicherungsperspektive integriert und ebenfalls nicht gesondert ausgewiesen.

Geplant waren Kosten-Wirksamkeits-Analysen mit den Effektparametern Mortalität und außerklinische Beatmung sowie eine Kosten-Nutzwert-Analyse mit dem Effektparameter der qualitätsadjustierten Lebensjahre (QALY).

3.10.2 Beschreibung von berücksichtigten Kosten und Effekten

3.10.2.1 Kosten

Bei der Kostenerhebung wurden direkte medizinische Kosten erfasst (siehe Abschnitt 3.6 (Datenerhebung) zu den direkten nicht-medizinischen Kosten). Dabei wurden zunächst aus der gewählten Perspektive relevante Kostenarten identifiziert (siehe Kostenarten in Abbildung 5).

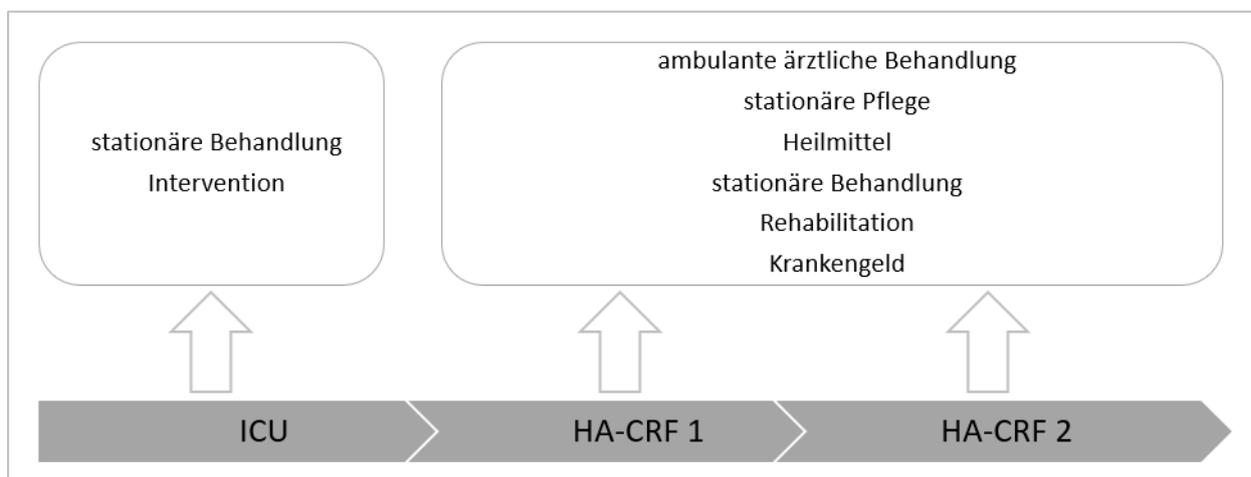


Abbildung 5. Relevante Kostenarten im Beobachtungszeitraum.

Abkürzungen: HA-CRF 1= Hausarztfragebogen 1; HA-CRF 2= Hausarztfragebogen 2; ICU=Intensivstation.

Im nächsten Schritt wurde die Einheit zur Messung des Ressourcenverbrauchs bestimmt (z. B. ICU-Aufenthalt in Tagen, siehe Tabelle 4) und der Ressourcenverbrauch quantifiziert. Daraufhin

wurde jeder Kostenart ein Preis zugeordnet und dieser Preis mit dem jeweils ermittelten Ressourcenverbrauch multipliziert. Zur Ermittlung der Gesamtkosten erfolgte dann eine Addition der für jede Kostenart ermittelten Kosten.

Abweichungen vom ursprünglich geplanten Vorgehen

Unter den direkten medizinischen Kosten war die Einbeziehung der Kostenart Hilfsmittel aufgrund des Antwortverhaltens im Hausarztfragebogen nicht möglich. Da auch der BARMER zum Auswertungszeitpunkt die entsprechenden Daten noch nicht für das Jahr 2020 vorlagen, konnte auch diese Datenquelle nicht herangezogen werden. Da Hilfsmittel jedoch mit 5% generell nur einen geringen Teil der gesamten Gesundheitsausgaben in 2019 ausgemacht haben,[49] ist nicht davon auszugehen, dass die Ergebnisse durch deren fehlende Einbeziehung maßgeblich beeinflusst werden. Die ambulanten Pflegekosten sowie Kosten für Behandlungspflege konnten aufgrund der Konzipierung des Hausarztfragebogens ebenfalls nicht in der Berechnung abgebildet werden, da der Zeitraum des Leistungsbezugs nicht bekannt war. Auch diesbezüglich ist nicht davon auszugehen, dass deren Fehlen einen maßgeblichen Einfluss auf die Ergebnisse hat, da in der gesundheitsökonomischen Evaluation die inkrementelle Betrachtung (Differenz der Kosten zw. Intervention und Kontrolle) von Interesse ist. Dahingegen konnten Transferleistungen (Krankengeld) berücksichtigt werden.

Tabelle 4. Gesundheitsökonomische Evaluation: Bewertung der berücksichtigten Kostenkategorien.

| Kostenart | Erhebungszeitpunkt | Einheit | Bewertungsgrundlage | Preis (Euro) |
|---------------------------------|--------------------|------------|---|--------------|
| Intervention ¹ | Indexaufenthalt | Patient | Kostenaufstellung | 651,79 |
| Intensivmedizinische Behandlung | Indexaufenthalt | Tage | Standardisierter Bewertungssatz ² | 1.441,03 |
| Rehabilitation | | | | |
| Phase B | HA-CRF 1, HA-CRF 2 | Tage | Standardisierter Bewertungssatz ² | 638,84 |
| Phase C | HA-CRF 1, HA-CRF 2 | Tage | Standardisierter Bewertungssatz ² | 131,26 |
| Anschlussheilbehandlung (AHB) | HA-CRF 1, HA-CRF 2 | Tage | Standardisierter Bewertungssatz ² | 131,26 |
| Unbekannte Rehabilitationsform | HA-CRF 1, HA-CRF 2 | Tage | Standardisierter Bewertungssatz ² | 131,26 |
| Stationäre Wiederaufnahmen | | | | |
| Normalstation | HA-CRF 1, HA-CRF 2 | Tage | Standardisierter Bewertungssatz ² | 638,84 |
| Intensivstation | HA-CRF 1, HA-CRF 2 | Tage | Standardisierter Bewertungssatz ² | 1.441,03 |
| Geriatric | HA-CRF 1, HA-CRF 2 | Tage | Standardisierter Bewertungssatz ² | 638,84 |
| Pflegekosten (stationär) | | | | |
| Pflegergrad 1 | HA-CRF 1, HA-CRF 2 | Tage | Sachleistungen nach Pfleegergrad ³ | n.a. |
| Pflegergrad 2 | HA-CRF 1, HA-CRF 2 | Tage | Sachleistungen nach Pfleegergrad ³ | 25,67 |
| Pflegergrad 3 | HA-CRF 1, HA-CRF 2 | Tage | Sachleistungen nach Pfleegergrad ³ | 42,07 |
| Pflegergrad 4 | HA-CRF 1, HA-CRF 2 | Tage | Sachleistungen nach Pfleegergrad ³ | 59,17 |
| Pflegergrad 5 | HA-CRF 1, HA-CRF 2 | Tage | Sachleistungen nach Pfleegergrad ³ | 66,83 |
| Pflegergrad unbekannt | HA-CRF 1, HA-CRF 2 | Tage | Sachleistungen nach Pfleegergrad ³ | 48,43 |
| Heilmittel | | | | |
| Physiotherapie | HA-CRF 1, HA-CRF 2 | Leistung | Heilmittelbericht ⁵ | 153,83 |
| Ergotherapie | HA-CRF 1, HA-CRF 2 | Leistung | Heilmittelbericht ⁵ | 349,82 |
| Logopädie | HA-CRF 1, HA-CRF 2 | Leistung | Heilmittelbericht ⁵ | 371,85 |
| Ambulante ärztliche Versorgung | | | | |
| Augenärzte | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 37,47 |
| Chirurgen | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 46,74 |
| fachärztliche Internisten | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 70,49 |
| Frauenärzte | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 32,46 |
| Hausärzte | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 21,61 |
| Hautärzte | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 20,35 |
| HNO-Ärzte | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 28,44 |
| Nervenärzte | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 48,17 |
| Orthopäden | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 27,38 |
| Psychotherapeuten | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 84,11 |
| Urologen | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 26,61 |
| Zahnärzte | HA-CRF 1, HA-CRF 2 | Arztbesuch | Standardisierter Bewertungssatz ² | 60,18 |
| Krankengeld | HA-CRF 1, HA-CRF 2 | Tage | Gräfe <i>et al.</i> 2019 [50] | 48,52 |

Anmerkungen: 1: Gilt nur für Patienten unter Intervention; 2: Standardisierte Bewertungssätze nach Bock *et al.* (2015) indexiert für das Jahr 2020; 3: 2020; 4: Durchschnittswert aus Sachleistungen für Pflegegrade 2 bis 5; 5: Waltersbacher (2020) indexiert für das Jahr 2020 [51]; Phase B und Phase C stehen jeweils für verschiedene Phasen der neurologischen Frührehabilitation. **Abkürzung:** HNO-Ärzte=Hals-Nasen-Ohren-Ärzte; n. a.= nicht zutreffend

Im Hinblick auf die Kalkulation der Interventionskosten wurde bei Kosten für die technische Ausstattung mit einer Lebensdauer von drei Jahren sowie eine Schulung für Pflege- und ärztliches Personal alle drei Jahre angenommen, die auf die jährlichen Kosten umgelegt worden sind.

Bei der Berechnung der Betriebskosten für das telemedizinische Zentrum wurde neben der Vergütung des telemedizinischen Personals auch ein Honorar für Hausärzte in Höhe von 40 Euro pro Patient für ein intersektorales Follow-Up berücksichtigt.

Die Vorgaben für intensivmedizinisch digital-gestützte (IDV) Zentren [36] besagen, dass pro angeschlossener Intensivstation eine tägliche Visite angeboten werden soll. Bei der vorliegenden Kostenkalkulation wurde jedoch davon ausgegangen, dass jeder angeschlossenen Intensivstation zwei Televisiten pro Tag angeboten werden. Bei einer Dauer von 25 Minuten pro Visite (inklusive Vor- und Nachbereitung) und einer im Median bei fünf Tagen liegenden Aufenthaltsdauer können so insgesamt 1460 Patienten im Jahr televisitiert werden.

Zu den Schulungskosten zählen Schulungen für das Personal im telemedizinischen Zentrum und den angebundenen Intensivstationen. Die Implementierungskosten beinhalten einmalige technische Anschaffungen, darunter die technische Ausstattung des telemedizinischen Zentrums sowie der Visitenroboter inklusive deren Inbetriebnahme und Installation von Software für die angeschlossenen Intensivstationen. Weitere Details zur Kostenkalkulation sind dem Anhang zu entnehmen.

Für den Indexaufenthalt² konnte nur die intensivmedizinische Behandlung mit Kosten bewertet werden, da zwar ein Aufnahme- aber kein Entlassungsdatum für den gesamten Krankenhausaufenthalt vorlag, und so Tage auf Normalstation nach dem ICU-Aufenthalt nicht berechnet werden konnten.

Wiederaufnahmen in Verbindung mit Mehrfacheinschlüssen wurden nicht berücksichtigt, da sich im exakten Test nach Fisher kein statistisch signifikanter Zusammenhang zwischen der Intervention und Anzahl der Wiederaufnahmen gezeigt hat ($p = 0,465$).

Im Hinblick auf die Ermittlung des Ressourcenverbrauchs wurde bei der ambulanten ärztlichen Versorgung je involviertem Fach- bzw. Hausarzt ein Patientenkontakt und bei den Heilmittelerbringern eine Leistung³ zugrunde gelegt. Für alle anderen Kostenkategorien wurde der tatsächliche Ressourcenverbrauch ermittelt.

Die Mehrheit der Kosten wurde mithilfe standardisierter Bewertungssätze nach Bock *et al.* (2015) kalkuliert.[52] Diese wurden auch von anderen Studien aus der Krankenversicherungsperspektive herangezogen.[50] Darüber hinaus wurden weitere Quellen zur Kostenbewertung herangezogen. Eine Übersicht der Kostenkategorien und ihrer Bewertung ist in Tabelle 4 dargestellt. Internationalen Standards [53] folgend wurden alle Bewertungssätze anhand der Inflationsrate für den Subindex Gesundheit [54] für das Indexjahr 2020 inflationsbereinigt. Hinsichtlich stationärer Wiederaufnahmen wurde neben dem Zeitraum auch abgefragt, ob eine Intensivbehandlung notwendig war. Wurde dies bejaht, erfolgte die Bewertung jeweils hälftig mit den Tagessätzen für die Intensivbehandlung (1.441,03 Euro) und für die Normalstation (683,84 Euro).

² Der Indexaufenthalt bezeichnet den ersten ICU-Aufenthalt eines Patienten.

³ Eine Leistung kann mehrere Behandlungen umfassen.

3.10.2.2 Effekte

Die interessierenden Effektparameter für die gesundheitsökonomische Evaluation sind die 6-Monats-Gesamtsterblichkeit, das Vorliegen einer außerklinischen Beatmung sowie die gesundheitsbezogene Lebensqualität und QALYs. Das Vorliegen einer außerklinischen Beatmung infolge eines intensivmedizinischen Aufenthalts stellt einen patientenrelevanten Endpunkt dar, der zudem mit hohen Versorgungskosten einhergeht.

Abweichungen vom ursprünglich geplanten Vorgehen

Die ursprüngliche Planung sah eine Auswertung der Dauer der außerklinischen Beatmung in Tagen vor. Aufgrund einer fehlenden Operationalisierbarkeit wurde dieser sekundäre Endpunkt stattdessen als binärer Endpunkt *Vorliegen einer außerklinischen Beatmung* ausgewertet.

Die qualitätsadjustierten Lebensjahre wiederum stellen ein Maß des Gesamtnutzens dar und werden deshalb ebenfalls im Rahmen einer Kosten-Nutzwert-Analyse betrachtet.[55] Die Valuation der mithilfe des EQ-5D-5L gemessenen gesundheitsbezogenen Lebensqualität erfolgt mit dem deutschen Valuationsset.[56] Verstorbenen Patienten wurde ein Nutzwert von 0 zugeordnet. Da eine Baseline-Erhebung bei der Studienpopulation nicht möglich war, wurde – dem Vorgehen anderer Studien folgend – den Patienten während ihres Aufenthalts auf der Intensivstation ein Nutzwert von 0 zugeordnet.[57, 58] Unter Verwendung dieser Baseline sowie der in FU 1 und FU 2 ermittelten Nutzwerte wurden die qualitätsadjustierten Lebensjahre mithilfe der AUC-Methodik ermittelt. Dabei wird eine lineare Veränderung der Nutzwerte zugrunde gelegt und der durchschnittliche Nutzwert über den Studienverlauf ermittelt.[51]

3.10.3 Inkrementelle Kosteneffektivitätsrelation (IKER)

Weiterhin war in allen Analysen, in denen entweder die Effekte oder die Kosten der Intervention signifikant höher oder geringer waren als die der Standardbehandlung, die Berechnung der inkrementellen Kosteneffektivitätsrelation (IKER) und deren zusammenfassende Darstellung in einer Kosten-Konsequenzen-Tabelle vorgesehen. Die IKER beschreibt das Verhältnis zwischen inkrementellen Kosten und inkrementellen Effekten und berechnet sich folgendermaßen:

$$IKER = \Delta \text{Kosten} / \Delta \text{Effekte}$$

3.10.4 Diskontierung, Sensitivitätsanalysen und Umgang mit fehlenden Werten

Da Intervention und Beobachtungszeitraum ein Jahr nicht überschreiten, wurde entsprechend der Empfehlungen des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) von einer Diskontierung abgesehen. Um die Robustheit der Ergebnisse zu prüfen, wurden Sensitivitätsanalysen durchgeführt.[55] Bei der Hauptanalyse handelt es sich um eine sogenannte *Complete Case Analysis*, in der nur unter der Annahme der MAR-Bedingung (missing at random) vollständig erhobene Datensätze einbezogen worden sind.[43]

Weiterhin wurden bei den Kostendaten als Sensitivitätsanalysen andere geeignete Modelle geschätzt. Es ist ebenfalls eine gesonderte Analyse nur der Kosten für den Indexaufenthalt (inklusive Interventionskosten bei Patienten unter Interventionsbedingung) erfolgt.

3.10.5 Statistische Methoden (gesundheitsökonomische Evaluation)

Zur Auswertung von Effekten und Kosten wurden deskriptive und inferenzstatistische Analysen vorgenommen. Dabei erfolgten die deskriptiven Analysen mit SPSS 27 und die inferenzstatistischen Auswertungen mit dem Statistikprogramm R (Version 4.0.5).

Für die inferenzstatistischen Analysen wurden verschiedene Regressionsmodelle angewendet, die in Abhängigkeit vom jeweiligen Modell mit den Befehlen `lmer`, `glm` und `survreg` mit dem Softwarepaket R geschätzt wurden. Bei den statistischen Analysen wurde ein Signifikanzniveau von 5% zugrunde gelegt. Technische Details sind dem Anhang B.2 zu entnehmen.

3.10.5.1 Kosten

Aufgrund der für Kostendaten (und ihre Residuen) klassischen schiefen Verteilung wurden die Kosten logarithmiert und mithilfe eines linearen gemischten Modells analysiert.[53] Um dem Cluster-randomisierten Studiendesign Rechnung zu tragen, wurde ein zufälliger Effekt für die Cluster im Modell berücksichtigt.

Zur Überprüfung der Robustheit der Ergebnisse wurden darüber hinaus generalisierte lineare Modelle unter Annahme der Gamma-Verteilung sowie generalisierte lineare Modelle unter Annahme der Weibull-Verteilung getestet, da diese neben der Logarithmierung zur Analyse von Kostendaten geeignet sind.[59]

3.10.5.2 Effekte

6-Monats-Gesamtsterblichkeit

Die verwendete Analyseverfahren ist unter Abschnitt 3.8.2.2 skizziert (Bestandteil der Effektevaluation).

Gesundheitsbezogene Lebensqualität (EQ VAS)

Die verwendete Analyseverfahren ist unter Abschnitt 3.8.2.2 skizziert (Bestandteil der Effektevaluation).

Qualitätsadjustierte Lebensjahre

Um dem Studiendesign Rechnung zu tragen, wurden zur Schätzung der qualitätsadjustierten Lebensjahre lineare gemischte Modelle herangezogen, mit denen sowohl Effekte auf Individual- als auch auf Clusterebene modelliert werden können. Dabei handelt es sich um Modelle, mit deren Hilfe sowohl fixe Effekte, die für alle Cluster gleich sind, als auch sogenannte zufällige Effekte, die für jede jedes Cluster einen spezifischen Effekt beschreiben, geschätzt werden. Darüber hinaus haben lineare gemischte Modelle den Vorteil, dass sie unbeobachtete Heterogenität auf Individual- und Clusterebene berücksichtigen können.[60] Bei diesem Endpunkt wurde das Studiencluster als zufälliger Effekt modelliert. In alle Modelle ist die Zeit seit der Entlassung als fester Effekt eingeflossen, um für Unterschiede im tatsächlichen Zeitpunkt des jeweiligen FU zu kontrollieren (siehe Abschnitt 5.1.3 der Effektevaluation).

Außerklinische Beatmung

Die außerklinische Beatmung als binärer Endpunkt wurde mithilfe eines generalisierten gemischten logistischen Modells analysiert. Hierbei wurde neben einem cluster- auch ein patientenspezifischer zufälliger Effekt modelliert. In dieses Modell sind alle Beobachtungen mit mindestens einem FU eingeflossen. Darüber hinaus ist die Zeit seit der Entlassung zur Kontrolle von zeitlichen Unterschieden des FU-Zeitpunkts in das Modell aufgenommen worden.

Die Modellauswahl und der Einbezug von Baseline-Charakteristika wie Alter, Geschlecht, SOFA-Score, SAPS II, Zeit zwischen HA-CRF und Entlassung von der Intensivstation und LOS wurde getestet und das finale Modell in einer Zusammenschau aus klinischer Sinnhaftigkeit und der Modellgüte gewählt. Die Beschreibung der finalen Modelle erfolgt bei den Ergebnissen.

3.11 Akzeptanzerhebung – Methoden [TU Berlin]

Es wurde eine anonyme Akzeptanzerhebung unter den teilnehmenden Zentren durchgeführt, bei der ermittelt werden sollte, wie die telemedizinische Intervention vom Personal vor Ort wahrgenommen wurde. Hierzu wurde durch das Robert-Koch-Institut (RKI) und die Charité – Universitätsmedizin Berlin ein an der internationalen Literatur orientierter Fragebogen im Rahmen des FUTURE-Projekts⁴ mit insgesamt 20 Variablen erstellt. Dieser ist dem Anhang des Ergebnisberichts zu entnehmen und fand nach leichter Modifikation auch für ERIC Anwendung, nachdem die Akzeptanzerhebung in das Amendment 4 des Ethikantrags inkludiert und am 14.12.2020 ein positives Ethikvotum erhalten worden ist.

Abweichung vom ursprünglich geplanten Vorgehen

Initial war eine Akzeptanzerhebung nicht vorgesehen. Die Akzeptanzerhebung erfolgte außerhalb der Projektlaufzeit und wurde nicht durch Projektmittel finanziert. Der Vollständigkeit halber wird diese aber in diesem Evaluationsbericht ebenfalls dargestellt.

Im Zusammenhang mit der Skalierung der Neuen Versorgungsform während der COVID-19-Pandemie hat das RKI im Rahmen des FUTURE-Projekts einen Fragebogen zur Akzeptanzerhebung entwickelt, der nach leichter Modifikation auch für ERIC genutzt werden konnte. Daher wurde die Gelegenheit genutzt, um auch im Rahmen von ERIC das Beratungsangebot mithilfe von Visitenrobotern zu evaluieren und zu optimieren.

Neben einem Block zu allgemeinen Fragen (Visiten-Teilnahme, Versorgungsstufe des Krankenhauses, Berufsgruppe) gab es fünf weitere Teilblöcke. Im ersten Teil wurde die Nutzerfreundlichkeit der Televisitenroboter evaluiert. Dies erfolgt in Anlehnung an die System Usability Scale (SUS) nach Brooke (1996) [61] und anhand von zehn Fragen, für deren Beantwortung eine 5-stufige Likert-Skala angelegt wurde. Die Spanne reichte dabei von „Stimmt gar nicht“ bis „Stimmt völlig“. Der Teil 2 enthielt Fragen zur subjektiven Nutzenbewertung. Im Teil 3 wurden die Teilnehmer*innen zum Ausmaß und zur Belastung der Krisensituation infolge der SARS-CoV-2-Pandemie befragt. Teil 4 enthielt Fragen zu den Einstellungen gegenüber den Visitenrobotern. Im Teil 5 wurde die Zufriedenheit mit der telemedizinischen Beratung mithilfe der Visitenroboter abgefragt. Im Teil 2 bis 5 wurde ebenfalls eine 5- bzw. 6-stufige Likert-Skala genutzt. Darüber hinaus konnten die Befragten im letzten Teil in Freitextform noch Optimierungsvorschläge geben.

Die Erhebung der Akzeptanz in den teilnehmenden Intensivstationen erfolgte zwischen dem 01.03.2021 und dem 20.05.2021 über einen online auszufüllenden Fragebogen, zu dem der Zugang durch die Konsortialführung an die Ansprechpartner*innen vor Ort weitergeleitet wurde.

⁴ FUTURE = Fachübergreifende, telemedizinische Unterstützung in Regionen mit erhöhtem Aufkommen kritisch kranker COVID-19 Patienten. Dabei handelt es sich um ein Folgeprojekt, das während der Pandemie gemeinsam mit dem RKI entwickelt und durchgeführt worden ist.

4 Ergebnisse der Evaluation – ICU Phase

4.1 Teilnehmende Zentren und Patienten: grundlegende Informationen

Die Randomisierung fand auf Krankenhausträgerebene statt. Alle 12 Cluster wurden randomisiert auf ein vordefiniertes Datum zur Implementierung der Telemedizin-basierten Intervention vor Ort, wobei jeweils 3 Monate vor Start der Intervention eine 3-monatige Trainings-Periode auf Cluster-Ebene initiiert wurde. Somit war der Start und das Ende der Trainings-Periode für ein einzelnes Cluster definiert über das zugewiesene Crossover-Datum, d.h. der Zeitpunkt des Wechsels eines Clusters von der Kontroll- zur Interventionsbedingung.

- Cluster #1 wurde zur Sequenzgruppe 1 randomisiert und initiiert. Der Cluster Dropout erfolgte nach Abschluss der Trainingsphase. Es wurden aus organisatorischen Gründen keinerlei Patienten in die Studie eingeschlossen.
- Cluster #12 wurde zur Sequenzgruppe 3 randomisiert, aber das Cluster hatte den Informed Consent (Clusterebene) noch vor Studienstart (FPFV) zurückgezogen („Cluster Withdrawal“).

Daher sind diese zwei Cluster nicht Teil des Full Analysis Set (FAS) nach dem ITT-Prinzip, definiert für die Wirksamkeitsevaluation.

First patient in, last patient out:

Die Datenerhebung mit Einschluss des ersten Patienten auf der ICU begann am 04.09.2018 in Cluster #8 (First patient first visit). Last Patient first visit (ICU Aufnahme) mit Ende des Rekrutierungszeitraums bzw. Post-Rollout Periode war am 31.03.2020 in Cluster #4 und #5 (Rekrutierungsende). Während der ICU-Phase war der Last patient last visit am 16.04.2020, welches gleichzeitig das späteste Datum der ICU-Entlassung unter den eingeschlossenen Studienpatienten darstellt.

In der Post-ICU Follow-up Phase war der Zeitpunkt der letzten FU-Erhebung am 17.11.2020 (Studienende). Der letzte Todesfall ereignete sich am 07.11.2020.

4.2 Implementierung der Intervention: Geplantes vs. tatsächliches Datum für Treatment Switching

Abbildung 6 zeigt den Randomisierungsplan, der die 3 Sequenzgruppen definiert, zusammen mit den *geplanten* 3 Crossover-Zeitpunkten (angedeutet als gestrichelte Linie für Sequenzgruppe 3 und Cluster #6 aus Sequenzgruppe 2) und den *tatsächlichen* Crossover-Zeitpunkten nach dem verzögerten Start der Intervention. Alle 3 Cluster (ID #3, #4, #9) aus Sequenzgruppe 3 können somit als *Late Adopter Sites* bezeichnet werden aufgrund des um 3 Monate verspäteten Crossover, welcher zwischen Konsortialpartnern und Konsortialführung zur Erhöhung der Fallzahl von Kontrollpatienten vereinbart wurde.

Auch bei Cluster #6 der Sequenzgruppe 2 handelt es sich um ein Late Adopter Zentrum, da der Start der Implementierung auf den 01.11.2019 verschoben wurde, also 14 Monate nach offiziellem Studienbeginn. Aufgrund der geringen Rekrutierung stimmten alle Konsortialpartner zu, die Dauer der Rekrutierung unter Kontrollbedingung um weitere 8 Monate gegenüber dem vorgeplanten Crossover-Datum zu verlängern. Insgesamt wurden in diesem Cluster 15 Pt. (11 unter der Kontroll-, 4 unter der Interventionsbedingung) rekrutiert, wobei der erste Cluster-spezifische ICU-Aufenthalt am 04.04.2019 stattfand.

Cluster #9 (n=96 Patienten insgesamt) aus Sequenzgruppe 3 begann erst am 19.05.2019, 8 Monate nach Studienstart und kurz vor dem dritten Crossover gemäß Randomisierungsplan mit der Rekrutierung unter Kontrollbedingung. Grund hierfür waren verzögerte Vertragsprüfungen und Freigaben auf Clusterebene.

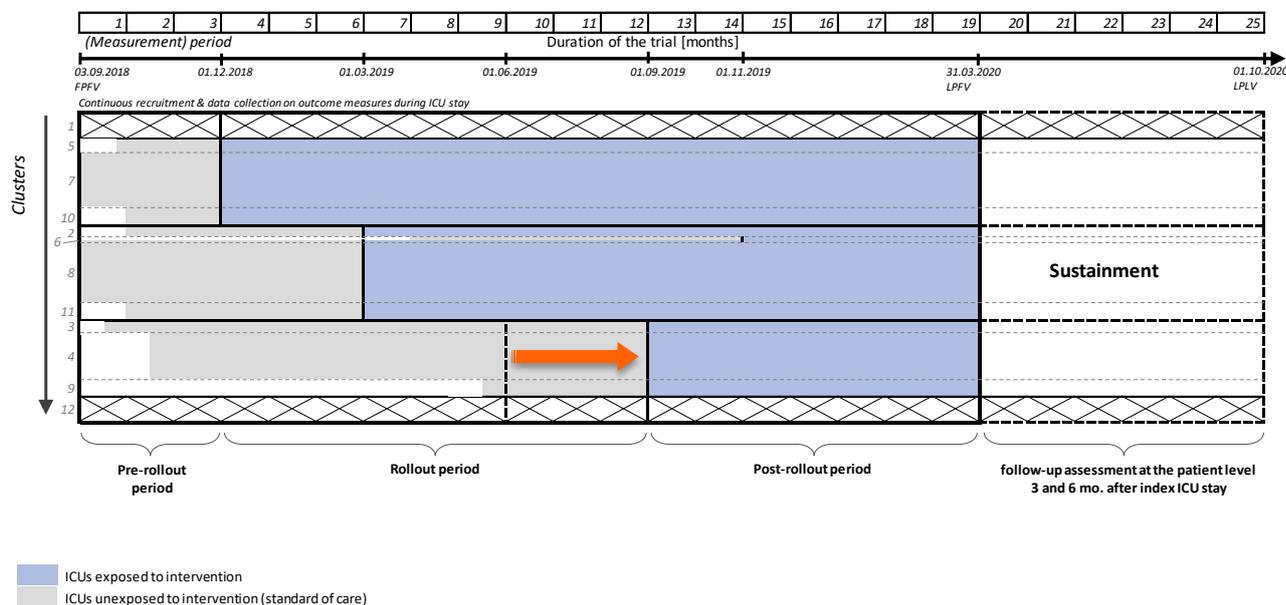


Abbildung 6. Schematische Illustration des im ERIC Projekt angewandten gestaffelten (sequentiellen) Rollout-Prozesses und zeitlichen Ablaufs (nach der Entscheidung, die Rekrutierung zu verlängern, inklusive der Verschiebung des dritten Crossover-Datums (ursprüngliches Datum dargestellt als vertikale gestrichelte Linie)). [In dieser Abbildung entsprechen die Nummern des Clusters den in diesem Bericht verwendeten Cluster-IDs. Die Höhe der Cluster entspricht der tatsächlichen Clustergröße].

Blaue Bereiche zeigen die Interventionsphase (Telemedizin), graue Bereiche die Kontrollphase, und der nicht schattierte Bereich zeigt die Post-ICU-Phase (dargestellt nach LPFV). In jeder Sequenzgruppe fand eine a priori geplante, 3-monatige Trainingsphase (Transitions-Phase) am Ende der Kontrollbedingung statt. Follow-up Assessments (sekundäre Outcomes) auf Patientenebene wurden 3 und 6 Monate nach dem ICU-Aufenthalt angestrebt. FPFV, first patient first visit; LPF(L)V, last patient first (last) visit.

Die nachfolgende Tabelle zeigt die studienrelevanten Zeitpunkte auf Clusterebene und gibt einen Überblick zur tatsächlichen Implementierung im Vergleich zu den durch die Randomisierung zugewiesenen Crossover-Zeitpunkten und der durch die Randomisierung definierten Rangfolge auf Clusterebene (siehe Abschnitt 3.4.1).

Tabelle 5. Cluster-spezifische Datumsangaben für 3-monatige Trainings-Periode (inkl. kurze Implementierungsphase) sowie Start der Intervention (Crossover-Datum für Treatment Switch).

| Cluster ID | Sequenz- gruppe | Start Training | Ende Training | Tatsächlicher Start der Intervention | Start der Intervention wie geplant? |
|------------|--------------------|----------------|---------------|--|---|
| 1 | 1 | 01.09.2018 | 30.11.2018 | -- | <i>Cluster Dropout</i> |
| 2 | 2 | 01.12.2018 | 28.02.2019 | 01.03.2019 | ja |
| 3 | 3 | 01.03.2019 | 31.05.2019 | 01.09.2019 | nein (verspätet) |
| 4 | 3 | 01.03.2019 | 31.05.2019 | 01.09.2019 | nein (verspätet) |
| 5 | 1 | 01.09.2018 | 30.11.2018 | 01.12.2018 | ja |
| 6 | 2 | 01.12.2018 | 28.02.2019 | 01.11.2019 | nein (verspätet*) |
| 7 | 1 | 01.09.2018 | 30.11.2018 | 01.12.2018 | ja |
| 8 | 2 | 01.12.2018 | 28.02.2019 | 01.03.2019 | ja |
| 9 | 3 | 01.03.2019 | 31.05.2019 | 01.09.2019 | nein (verspätet) |
| 10 | 1 | 01.09.2018 | 30.11.2018 | 01.12.2018 | ja |
| 11 | 2 | 01.12.2018 | 28.02.2019 | 01.03.2019 | ja |
| 12 | 3 | -- | -- | -- | <i>Rücknahme IC</i> |

Sequenzgruppe 1, 2, 3 (geplantes Datum des Interventionsstarts siehe Abschnitt 3.4.1).

Am Ende der Trainings-periode fand innerhalb weniger Tage (max. 1 Woche) die (technische) Implementierung der telemedizinischen Intervention in den ICUs eines Clusters statt.

Cluster #1: Cluster Dropout nach Trainingsphase, keine Rekrutierung, kein Start der Intervention;

Cluster #12: Rücknahme des Informed Consent vor Studienstart.

* Cluster #6 erheblich verspäteter Start der Intervention, und die Rekrutierung startete erst nach dem per Randomisierungsplan definierten eigentlichen Crossover-Datum 01.03.2019.

4.3 CONSORT-Flussdiagramm

13 Krankenhäuser wurden den 12 als Randomisierungseinheiten definierten Studien-Clustern zugeordnet. Die teilnehmenden Behandler schlossen im Studienverlauf insgesamt 1463 Patienten ein. Das CONSORT-Flussdiagramm für die ERIC Studie, angepasst an das Stepped Wedge Design, ist in Abbildung 7 dargestellt.

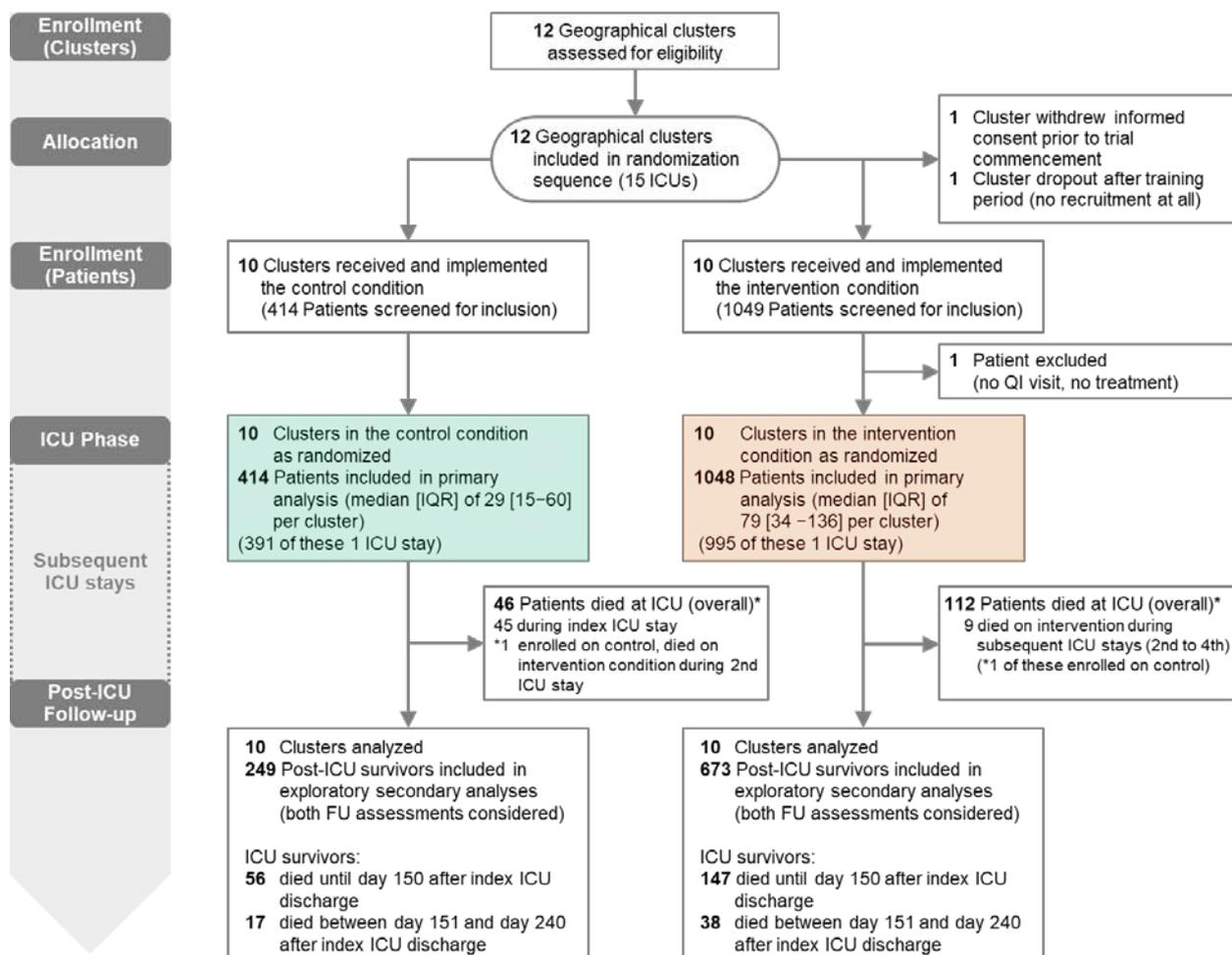


Abbildung 7. CONSORT Flow Chart. Der Fluss der Cluster und Patienten im Zeitverlauf der ERIC Stepped Wedge Cluster-randomisierten Studie. Nullpunkt für die Fallzahlen bzgl. ICU-Überlebende: Zeit seit Index-ICU-Entlassung. * Ein ICU-Todesfall war ein Wechselpatient, d.h. der Index-ICU-Aufenthalt erfolgte unter Kontroll-Bedingung, Patient beim 2. Aufenthalt unter Intervention verstorben. IQR: Interquartilsabstand.

Die in dieser Abbildung dargestellte Fallzahl für die ICU Phase beinhaltet auch ICU-Aufnahmen nach dem Index ICU-Aufenthalt (betrifft N=76 Patienten). Für diese Darstellung wurden für die Anzahl der ICU-Todesfälle auch Todesfälle auf der ICU nach dem Index ICU-Aufenthalt mit eingerechnet.

11 Patienten mit 2 Aufenthalten wurden unter Kontrolle eingeschlossen, beim 2. Aufenthalt unter Interventionsbedingungen behandelt ("Wechselpatienten") und werden in diesem Flow Chart unter der bei Studieneinschluss vorliegenden Behandlungsbedingung Kontrolle geführt.

4.4 Basisinformationen und Baseline Patienten-Charakteristika

4.4.1 Basisinformationen, QI-Visitentage, und Versichertenstatus

a) Basisinformationen

Zwischen FPFV und LPFV wurden 1463 Patienten (663 (45.3%) Frauen) in die Studie eingeschlossenen. Hierbei ergaben sich 1554 ICU-Aufenthalte, von denen 440 (28.3%) unter Kontrollbedingung, 1114 (71.7%) unter der Interventionsbedingung stattfanden.

Insgesamt wurden 14 783 QI-Visitentage für insgesamt **1462 Patienten** und **1553 ICU-Aufenthalte** mit mindestens 1 QI-Visite dokumentiert. Für einen einzigen Patienten (1 Aufenthalt, Cluster #4, Intervention) ergaben sich 0 QI-Tage, da unmittelbar am nächsten Tag die Entlassung auf Normalstation erfolgte.

Für 1386 (94.80%) der Patienten wurde ein einzelner ICU-Aufenthalt, für 76 Patienten wurden bis zu sechs Wiederaufnahmen auf die ICU im Rahmen der Studie dokumentiert (66 Patienten hatten im Zeitverlauf 2, 7 Patienten 3, 2 Patienten 4, und 1 Patient 6 ICU-Aufenthalte (in Cluster #8, jedes Mal unter der Kontrollbedingung)).

Tabelle 6. Deskriptive Zusammenfassung für die Anzahl der QI-Tage pro Aufenthalt (zentrumsübergreifend, sowie stratifiziert nach Cluster).

| | Patientenindividuelle Anzahl der QI-Tage | | | | | | |
|-------------------|--|---------|--------|-------|---------|--------|--------|
| | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | SD |
| Overall | 1.00 | 2.00 | 5.00 | 9.52 | 11.00 | 133.00 | 13.08 |
| Cluster 2 | 1.00 | 4.25 | 7.50 | 10.41 | 14.00 | 61.00 | 9.982 |
| Cluster 3 | 1.00 | 3.00 | 4.00 | 6.72 | 6.00 | 57.00 | 9.314 |
| Cluster 4 | 1.00 | 1.00 | 3.00 | 5.45 | 5.00 | 84.00 | 9.540 |
| Cluster 5 | 1.00 | 2.25 | 4.00 | 8.56 | 9.75 | 37.00 | 9.171 |
| Cluster 6 | 3.00 | 4.00 | 6.00 | 13.47 | 13.00 | 73.00 | 18.224 |
| Cluster 7 | 1.00 | 3.00 | 5.00 | 11.01 | 11.75 | 133.00 | 16.623 |
| Cluster 8 | 1.00 | 1.00 | 4.00 | 8.81 | 10.00 | 85.00 | 12.099 |
| Cluster 9 | 1.00 | 4.00 | 10.00 | 14.02 | 21.00 | 60.00 | 12.769 |
| Cluster 10 | 1.00 | 3.00 | 6.00 | 10.74 | 13.00 | 76.00 | 12.707 |
| Cluster 11 | 1.00 | 4.00 | 8.00 | 13.72 | 16.00 | 87.00 | 14.910 |

1st Qu.= unteres Quartil, 3rd Qu.= oberes Quartil, SD = Standardabweichung

Für die 1462 Patienten mit mindestens 1 QI Visite (1553 Aufenthalte) betrug die mediane Anzahl an QI-Tagen pro ICU-Aufenthalt 5 Tage (Interquartilsabstand: 2 bis 11 Tage). Wie Tabelle 6 zeigt, ist insgesamt eine relativ hohe Variabilität zwischen den Clustern zu beobachten. Die vergleichsweise hohe Streuung in Cluster #6 kann erklärt werden durch die geringe Fallzahl (15 Patienten). Cluster 4 weist eine relativ niedrige mediane Anzahl von 3 QI-Tagen (IQR: 1 – 5) pro Patient auf, insbesondere im Vergleich zu Cluster 9 mit einer medianen Anzahl von 10 QI-Tagen (4 – 21) pro Patient.

Die durchschnittliche Anzahl an ERIC-Patienten pro Tag in den ICUs der beteiligten Cluster zeigt ebenfalls eine hohe Variabilität, entsprechend der Größe des Clusters (z.B. Bettenzahl):

| Cluster 2 | Cluster 3 | Cluster 4 | Cluster 5 | Cluster 6 | Cluster 7 | Cluster 8 | Cluster 9 | Cluster 10 | Cluster 11 |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|------------|------------|
| 1.74 | 1.61 | 3.15 | 1.81 | 1.18 | 6.20 | 5.64 | 4.55 | 2.70 | 3.07 |

Dies kann auch in Relation gebracht werden zur verfügbaren Bettenzahl des jeweiligen Clusters (vgl. Tabelle 1). Genauere Details zum Cluster-spezifischen Rekrutierungsverhalten über die Zeit werden im Abschnitt 4.5 präsentiert.

b) Versichertenstatus (BARMER vs. non-BARMER)

Gemäß der ursprünglichen Planung sollten nur Patienten mit BARMER-Krankenversicherung in die Studie eingeschlossen werden. Die Ethikkommission verlangte jedoch von Anfang an, dass alle Patienten auf einer Station grundsätzlich Zugang zur Versorgung haben müssen, auch wenn diese nicht evaluiert werden können. Um Rekrutierungsschwierigkeiten zu begegnen und die geplante Patientenfallzahl erreichen zu können, wurde im Studienverlauf das entsprechende Einschlusskriterium geändert, so dass alle Patienten mit Gesetzlicher Krankenversicherung (GKV) rekrutiert werden konnten (Details siehe Ergebnis- und Schlussbericht). Der erste non-BARMER Patient wurde am 07.03.2019 eingeschlossen, kurz nach dem 2. Crossover-Zeitpunkt.

4.4.2 Wechselsituationen

a) Wechsel des Clusters beim 2. ICU-Aufenthalt

11 Patienten (mit jeweils 2 Aufenthalten) wechselten das Cluster; die übrigen Patienten mit >2 ICU-Aufenthalten wechselten das Cluster nicht. Weiterhin fand bei 4 dieser 11 Patienten, die das Cluster wechselten, zusätzlich ein Wechsel in der Behandlungsbedingung beim zweiten ICU-Aufenthalt statt.

b) Crossover-Patienten vs. Wechsel-Patienten

“Crossover-Patienten”

Als Crossover-Patient wurde ein Patient definiert, der unter Kontrollbedingung aufgenommen wurde (kurz vor Start der Intervention), aber nach Treatment-Switch auf Cluster-Ebene während desselben ICU-Aufenthalts auch unter Interventionsbedingung behandelt wurde.

Nur für Cluster 11 konnten durch die Konsortialführung bei einer ICU 3 Crossover-Patienten identifiziert werden, bei denen ein Wechsel der Behandlungsbedingung innerhalb ihres ICU-Aufenthalts stattfand. Diese Patienten wurden Ende Februar 2019 wenige Tage vor Start der Interventionsbedingung unter Kontrollbedingungen auf die ICU aufgenommen (Patient 1: 40 QI-Tage (5 Tage unter Intervention), Patient 2: 21 QI Tage (16 Tage unter Intervention, verstorben), Patient 3: 14 QI-Tage (4 Tage unter Intervention), verstorben). Diese Patienten mit nur 1 ICU-Aufenthalt im Beobachtungszeitraum wurden in der weiteren Analyse nicht ausgeschlossen und sind gemäß des Intention-to-treat Prinzips Teil des Full Analysis Sets (FAS).

“Wechsel-Patienten” (Interventionsbedingung bei Wiederaufnahme von Kontrollpatienten)

Bei 11 Patienten fand ein Wechsel der Behandlungsbedingung beim 2. ICU-Aufenthalt statt, bei 4 dieser Patienten inklusive eines Wechsels des Clusters (vgl. oben unter a)).

4.5 Zentrums-spezifische Implementierung der Intervention und Rekrutierungsübersicht

Die folgende Abbildung 8 zeigt 1462 Patienten (1553 ICU-Aufenthalte) mit mindestens einem QI-Visitentag auf Cluster-Ebene. Die QI-Verläufe sind farblich gekennzeichnet anhand der zugrundeliegenden Behandlungsbedingung (Kontrolle vs. Intervention). Folge-Aufenthalte eines Patienten sind über eine graue Linie verbunden mit dem Index-ICU-Aufenthalt.

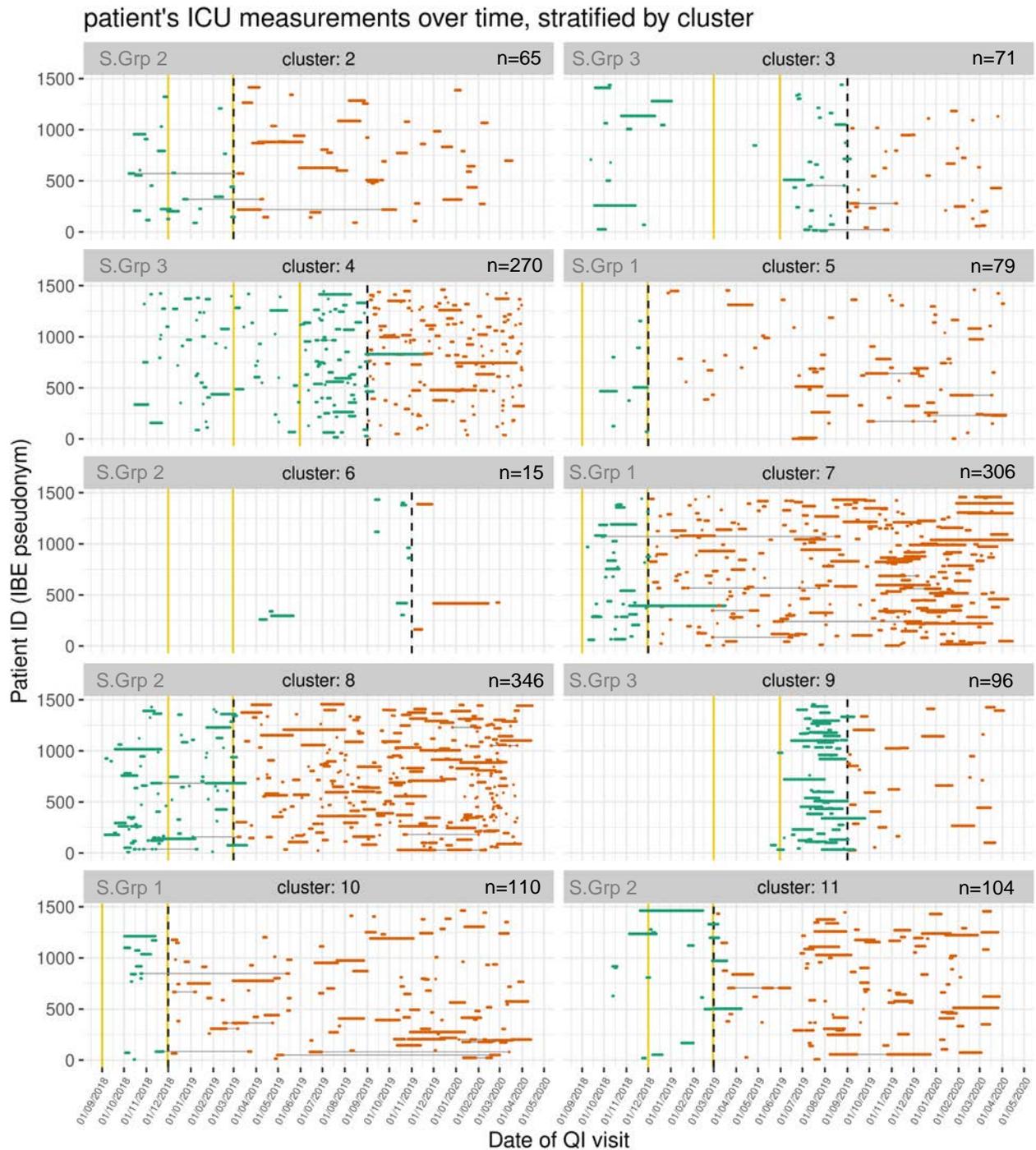


Abbildung 8. Deskriptive Darstellung der individuellen Beobachtungsdauer aller 1462 Patienten (1553 ICU Aufenthalte) mit mindestens 1 QI Visite während ihres ICU Aufenthalts. Zuordnung der Patienten zum Cluster per Index-ICU-Aufenthalt. Verläufe in grün: QI-Beobachtung unter Kontroll-Bedingung; Verläufe in orange: QI-Beobachtung unter Interventions-Bedingung. Gelbe senkrechte Linien: Trainings-Periode. Schwarze gestrichelte senkrechte Linie: Start der Intervention (Telemedizin). S.Grp = Sequenzgruppe.

Wechselpatienten innerhalb eines Clusters mit >1 Aufenthalt, die unter Kontrolle eingeschlossen und beim Folgeaufenthalt unter Intervention behandelt wurden, sind jeweils mit der tatsächlich durchgeführten Behandlungsbedingung farblich gekennzeichnet. Ein Patient des Clusters 4, Interventionsbedingung, wurde aus der Intention-to-treat (ITT) Population ausgeschlossen, da keine QI-Visite durchgeführt wurde (Patient wurde am Folgetag unmittelbar wieder auf die Normalstation verlegt).

Insgesamt ist eine starke Heterogenität in der Cluster-Größe erkennbar. Darüber hinaus kann sowohl eine über die Zeit variierende Rekrutierung zwischen den einzelnen Clustern, aber auch innerhalb eines einzelnen Clusters festgestellt werden (Abbildung 9, Anhang: Abbildung 33). Der deutlich verspätete Start der Rekrutierung in Cluster #6 und #9, oder die zeitweise aussetzende Rekrutierung in Cluster #3 (Sequenzgruppe 3) unter Kontrollbedingung wurden durch intermittierenden Personalmangel in den Remote-Sites, verspätete Vertragsunterzeichnung in Cluster #9, oder organisatorischen Schwierigkeiten bei der Identifikation bzw. Einschluss von Kontrollpatienten verursacht.

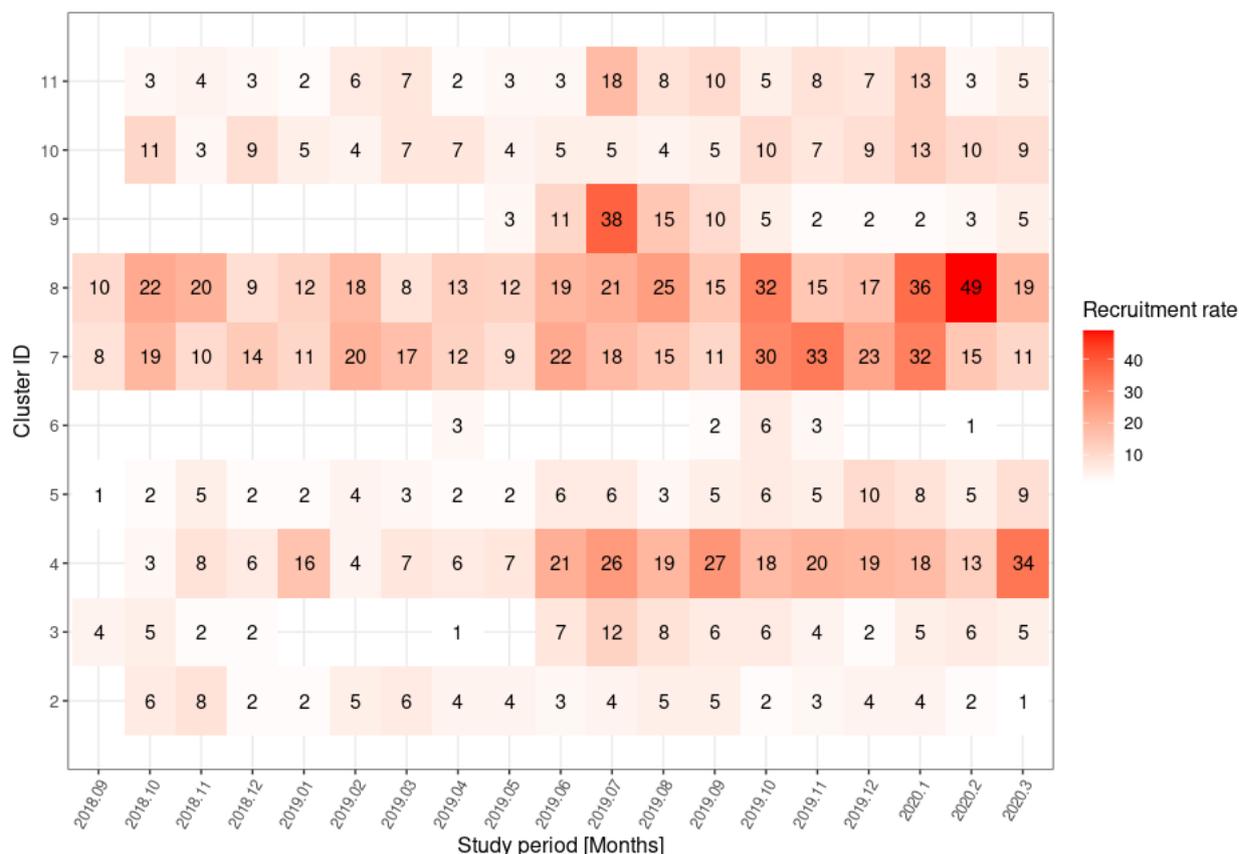


Abbildung 9. Heatmap für die Cluster-spezifische Rekrutierungsrate (Fall-Ebene) pro Kalendermonat, im Verlauf der 19-monatigen Rekrutierungsphase. Je dunkler die Farbe, desto mehr Patienten wurden in diesem Kalendermonat in die Studie eingeschlossen. Fehlende Angaben zur Anzahl der Patienten in einem bestimmten Monat bedeuten, dass in diesem Zeitraum keine Patienten rekrutiert wurden.

4.6 Datenmanagement-Aspekte und Datenqualität

4.6.1 Design-Variablen

In der offiziellen REDCap Datenbank waren die Designvariablen Cluster, Behandlungsbedingung (Kontrolle/ Intervention), Gesetzlicher Krankenversicherungsstatus des Patienten (BARMER vs. non-BARMER) über das bei Studieneinschluss vergebene Original-Fallpseudonym enthalten, welches behandlungsfallspezifisch ist, d.h. einen einzelnen ICU-Aufenthalt eines Patienten definierte (z.B. EC-08-001: ICU-Aufenthalt eines Patienten von Cluster #8, behandelt unter "c" = Kontrollbedingung, "E" für Ersatzkasse (BARMER)).

Als Patienten-Identifikator diente die eindeutige KV-Nr. (Krankenversicherungsnummer), die aufgrund der EU-DSGVO und der entsprechenden deutschen Datenschutzgesetze nur an der Charité über die klinische Routedokumentation zugänglich war. Daher wurde eine zweite REDCap-Datenbank von FOKUS abgeleitet, die ein neues Fallpseudonym und Patientenpseudonym (`patient_id_lmu`), aber nicht mehr das Original-Fallpseudonym enthielt. Bei diesem Prozess (Generierung der Datenbank für den Evaluator) wurden separate Variablen für Cluster, Behandlungsbedingung und Krankenversicherungs-Status auf Basis des ursprünglichen Fallpseudonyms generiert. Der Evaluator der LMU hatte aus Datenschutzgründen keinen Zugang zu personenbeziehbaren Daten.

4.6.2 Baseline Patienten-Charakteristika, weitere Variablen

Allgemeine Patientencharakteristika wurden am Tag der Krankenhausaufnahme bzw. der ICU-Aufnahme dokumentiert, letztere sollte idealerweise den Beginn der Dokumentation der QI-Visiten darstellen. Laut Protokoll sollte die erste QI-Visite am Tag der ICU-Aufnahme oder am nächsten Tag stattfinden, abhängig von der Tageszeit der Aufnahme auf die ICU. Baseline-Parameter wie z.B. SOFA Score, SAPS II wurden erst im Rahmen der QI-Visiten auf täglicher Basis dokumentiert.

4.6.3 Missing Baseline-Werte (wegen verspätetem Patienten-Einschluss)

Im Falle von Missings zum definierten Baseline-Zeitpunkt (ICU-Aufnahme) wurde der fehlende Wert ersetzt durch den zur ersten QI-Visite vorliegenden Messwert, d.h. es wurde *First Observation Carried Backward* (FOCB) als einfache Form eines Single Imputation Ansatzes angewandt (zugrundeliegende Annahme von Missing Completely at Random (MCAR)).

Für die Adjustierung verwendete, potentiell klinisch relevante Baseline-Variablen wie SAPS II oder SOFA wurden erst im Rahmen der QI-Visiten dokumentiert. Daher wurde die erste vorhandene Messung der ersten QI-Visite als Baseline-Messung definiert (FOCB) und für Regressionsanalysen verwendet.

4.7 Deskriptive Analysen

Die nachfolgende Tabelle beschreibt das eingeschlossene Patientenkollektiv, d.h. die Intention-to-treat (ITT) Population, anhand verschiedener Patientencharakteristika und klinischer Kenngrößen, die zum Baseline-Zeitpunkt erhoben wurden.

Tabelle 7. "Table 1" – Baseline Charakteristika. Klinische und demographische Charakteristika der ITT-Population (1462 Patienten) zu Baseline oder am Tag der 1. QI Visite.

Bei >1 Aufenthalt pro Patient wurde hier nur der Index-ICU Aufenthalt berücksichtigt.

| Variable | Control (n = 414) | Intervention (n = 1048) |
|---|----------------------|----------------------------|
| Demographics | | |
| Age [years], mean (SD) | 68.74 (14.41) | 67.25 (14.88) |
| Sex, n (%) | | |
| Female | 186 (44.9%) | 477 (45.5%) |
| Male | 228 (55.1%) | 571 (54.5%) |
| BARMER [†] SHI, n (%) | 299 (72.2%) | 452 (43.1%) |
| Admission status, n(%) | | |
| Medical | 208 (50.24) | 449 (42.84) |
| Emergency surgery | 105 (25.36) | 319 (30.44) |
| Elective surgery | 93 (22.46) | 259 (24.71) |
| Missing | 8 (1.93) | 21 (2.00) |
| Admission from, n (%) | | |
| Emergency medical services | 141 (32.0%) | 276 (24.8%) |
| Surgical | 167 (38.0%) | 499 (44.8%) |
| Normal ward | 87 (19.8%) | 173 (15.5%) |
| Other ICU | 20 (4.5%) | 86 (7.7%) |
| External | 25 (5.7%) | 79 (7.1%) |
| ICU primary admission diagnosis, n (%) | | |
| Respiratory | 49 (11.7%) | 102 (9.8%) |
| Sepsis/ Infection | 65 (15.5%) | 167 (16.0%) |
| Gastrointestinal | 47 (11.2%) | 104 (10.0%) |
| Cardiovascular | 101 (24.0%) | 281 (27.0%) |
| Trauma | 42 (10.0%) | 92 (8.8%) |
| Neurologic | 33 (7.9%) | 107 (10.3%) |
| Metabolic or endocrine | 20 (4.8%) | 36 (3.5%) |
| Oncologic | 55 (13.1%) | 140 (13.4%) |
| Other ‡ | 8 (1.9%) | 13 (1.2%) |
| SOFA score (at first QI visit) | | |
| mean (SD) | 4.32 (3.90) | 6.14 (3.88) |
| median (IQR) | 4.00 (1.00 – 7.00) | 6.00 (3.00 – 9.00) |
| BARMER SHI | | |
| mean (SD) | 4.20 (4.11) | 6.15 (3.71) |
| median (IQR) | 3.00 (0 – 7.00) | 6.00 (3.00 – 9.00) |

| Variable | Control (n = 414) | Intervention (n = 1048) |
|------------------------------------|-----------------------|----------------------------|
| <i>Other SHI</i> | | |
| mean (SD) | 4.64 (3.30) | 6.14 (4.01) |
| median (IQR) | 4.00 (2.00 – 7.00) | 6.00 (3.00 – 9.00) |
| SAPS II (at first QI visit) | | |
| mean (SD) | 30.58 (18.61) | 36.17 (18.90) |
| median (IQR) | 28.00 (16.00 – 42.25) | 35.00 (22.00 – 48.00) |
| Missing, n(%) | 2 (<0.01%) | 10 (0.01) |
| BARMER SHI | | |
| mean (SD) | 32.58 (18.88) | 36.88 (17.86) |
| median (IQR) | 30.00 (18.00 – 46.00) | 35.00 (23.00 – 48.00) |
| Missing, n(%) | 2 (<0.01%) | 4 (0.4%) |
| <i>Other SHI</i> | | |
| mean (SD) | 25.42 (16.91) | 35.62 (19.64) |
| median (IQR) | 22 (13.00 – 36.00) | 34.00 (20.00 – 48.75) |
| Missing, n(%) | 0 | 6 (0.6%) |

Bei 11 Patienten fand ein Wechsel der Behandlungsbedingung beim 2. ICU-Aufenthalt statt.

SHI: statutory health insurance (Gesetzliche Krankenversicherung, GKV).

† Die Studie wurde zunächst mit BARMER-versicherten Patienten gestartet. Der erste non-BARMER Patient wurde am 07.03.2019 eingeschlossen, d.h. kurz nach dem 2. Crossover-Zeitpunkt.

‡ Primärdiagnose zum Aufnahmezeitpunkt unter "Other" (sonstiges) z.B. Entfernung Plattenosteosynthese, Urolithiasis, medikamenteninduzierte Kieferosteonekrose, Pregabalin-Intoxikation, akutes Nierenversagen, katatone Schizophrenie, autoimmunhämolytische Anämie, Harnleiterstenose.

Tabelle 8. Deskription zum Entlassungsstatus bzgl. *Index* ICU-Aufenthalt in der ITT-Population, stratifiziert nach Behandlungsbedingung (Kontrolle vs. Intervention).

| Variable | Control (n = 414) | Intervention (n = 1048) | Total (n = 1462) |
|---|----------------------|----------------------------|---------------------|
| ICU discharge disposition, No. (%) | | | |
| Hospital ward | 295 (71.26) | 654 (62.40) | 949 (64.91) |
| Home | 4 (0.97) | 8 (0.76) | 12 (0.82) |
| Other ICU | 47 (11.35) | 212 (20.23) | 259 (17.72) |
| Rehabilitation facility | 19 (4.59) | 55 (5.25) | 74 (5.06) |
| Nursing facility | 2 (0.48) | 3 (0.29) | 5 (0.34) |
| Death at ICU | 45 (10.87) | 104 (9.92) | 149 (10.19) |
| Missing * | 2 (0.48) | 12 (1.15) | 14 (0.96) |

* Entlassungsstatus nicht dokumentiert.

ICU-Wiederaufnahmen sind in dieser Tabelle nicht berücksichtigt (d.h. pro Patient wurde jeweils nur der erste studienbezogene ICU-Aufenthalt verwendet).

4.8 Deskriptive Analysen des primären Efficacy-Endpunkts – pro QI

Für einen ersten Überblick hinsichtlich einer Änderung im Level der QI-Adhärenz ("Step Change") unter Kontrollbedingung vs. unter Intervention werden im Folgenden die Ergebnisse der explorativen Auswertung für die einzelnen 8 QIs separat dargestellt. Diese deskriptiven Zusammenhänge werden im Rahmen der modell-basierten konfirmatorischen Hauptanalyse im Detail untersucht (Abschnitt 4.9).

Wie im Methodenteil in Abschnitt 3.8.2.1 beschrieben, ist zu beachten, dass die 3-monatige Trainingsperiode für jedes Cluster analysiert wurde als Teil der Cluster-spezifischen Kontroll-Phase. (Durch die fehlende telemedizinische Unterstützung der QI-Visiten während der Trainingsphase konnte die Intervention nicht angewendet werden. Patienten gelten als behandelt unter Kontrollbedingung, wenngleich in diesem Intervall Personal für die nachfolgende Implementierung geschult wurde.)

4.8.1 QI 1 – Tägliche multiprofessionelle und interdisziplinäre klinische Visite mit Dokumentation von Tageszielen

Stratifiziert für die einzelnen Cluster ergab sich für die mittlere QI-Adhärenz, jeweils unter Kontroll- („res.C“) vs. unter Interventions-Bedingung („res.I“) folgendes Bild:

| ID | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| res.C | 0.947 | 0.689 | 0.990 | 0.836 | 0.967 | 0.623 | 0.701 | 0.990 | 0.992 | 0.911 |
| res.I | 0.988 | 0.964 | 0.989 | 0.772 | 0.973 | 0.935 | 0.979 | 0.942 | 0.983 | 0.983 |

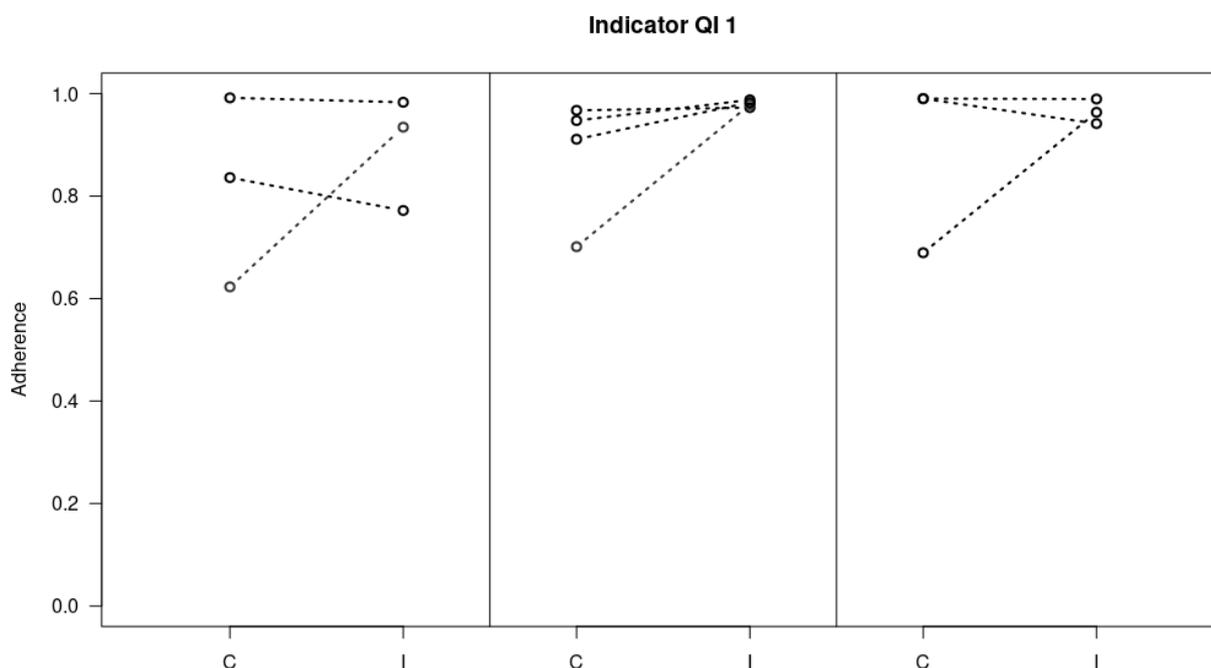


Abbildung 10. QI 1: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I).

4.8.2 QI 2 – Management von Sedierung, Analgesie & Delir

Stratifiziert für die einzelnen Cluster ergab sich für die mittlere QI-Adhärenz, jeweils unter Kontroll- („res.C“) vs. unter Interventions-Bedingung („res.I“) folgendes Bild:

| ID | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| res.C | 0.336 | 0.052 | 0.057 | 0.279 | 0.011 | 0.458 | 0.389 | 0.017 | 0.049 | 0.188 |
| res.I | 0.648 | 0.328 | 0.177 | 0.379 | 0.091 | 0.737 | 0.477 | 0.310 | 0.188 | 0.311 |

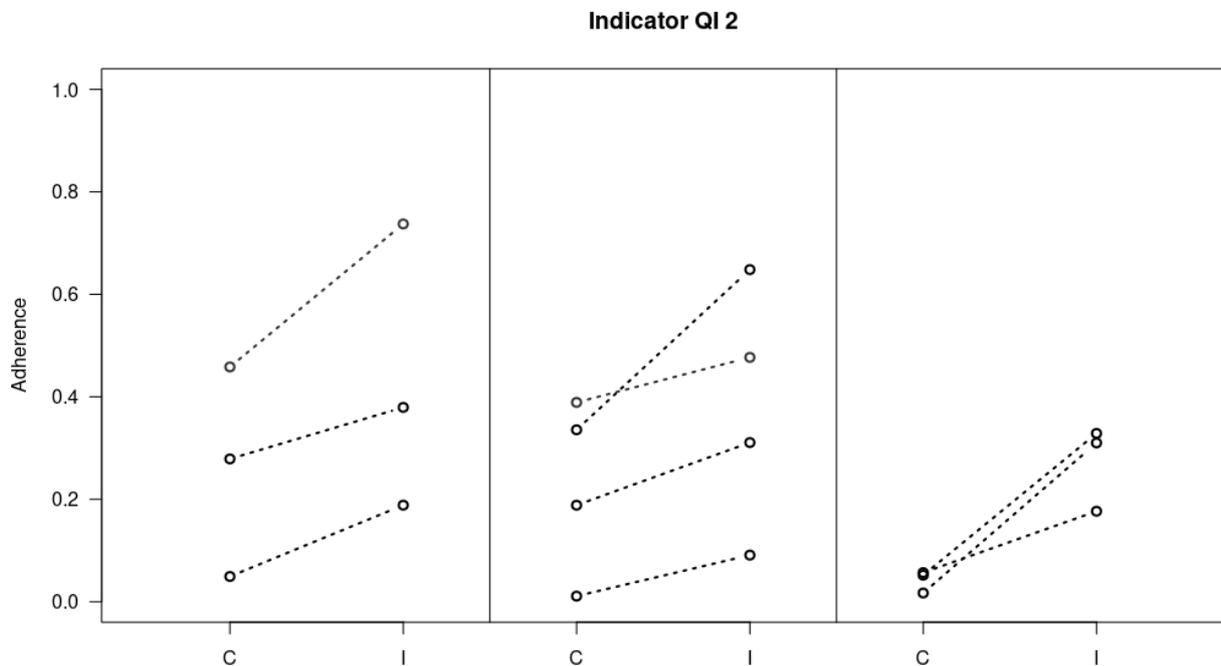


Abbildung 11. QI 2: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I).

Insgesamt ist für alle Cluster ein im Vergleich zu den anderen QIs relativ niedriges Level in Bezug auf die QI-Adhärenz unter Kontrolle festzustellen (globaler Mittelwert unter 10%), welches sich unter Intervention offenbar steigern ließ. Dieser rein deskriptive Zusammenhang wurde im Rahmen der modell-basierten konfirmatorischen Hauptanalyse quantifiziert und hinsichtlich Stabilität und konfundierender zeitlicher Effekte untersucht (Abschnitt 4.9).

4.8.3 QI 3 – Patientenadaptierte Beatmung

Stratifiziert für die einzelnen Cluster ergab sich für die mittlere QI-Adhärenz, jeweils unter Kontroll- („res.C“) vs. unter Interventions-Bedingung („res.I“) folgendes Bild:

| ID | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| res.C | 0.822 | 0.929 | 0.939 | 0.770 | 0.946 | 0.757 | 0.833 | 0.837 | 0.803 | 0.832 |
| res.I | 0.948 | 0.985 | 0.959 | 0.801 | 0.982 | 0.921 | 0.920 | 0.874 | 0.906 | 0.910 |

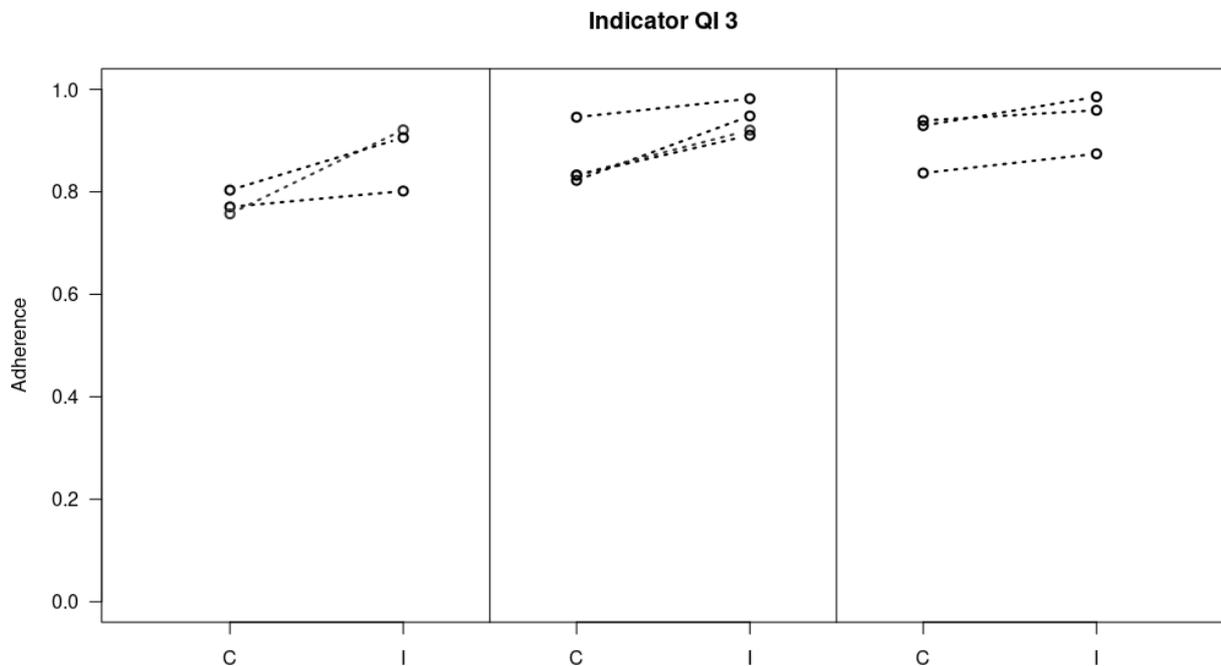


Abbildung 12. QI 3: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I).

4.8.4 QI 4 – Frühzeitige Entwöhnung von einer invasiven Beatmung (Weaning)

Stratifiziert für die einzelnen Cluster ergab sich für die mittlere QI-Adhärenz, jeweils unter Kontroll- („res.C“) vs. unter Interventions-Bedingung („res.I“) folgendes Bild:

| ID | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------|-------|-------|-------|-------|-------|-------|------|-------|-------|-------|
| res.C | 0.849 | 0.910 | 0.994 | 0.607 | 0.989 | 0.957 | 0.84 | 0.887 | 0.762 | 0.808 |
| res.I | 0.993 | 0.993 | 0.966 | 0.953 | 0.991 | 0.981 | 0.98 | 0.982 | 0.981 | 0.915 |

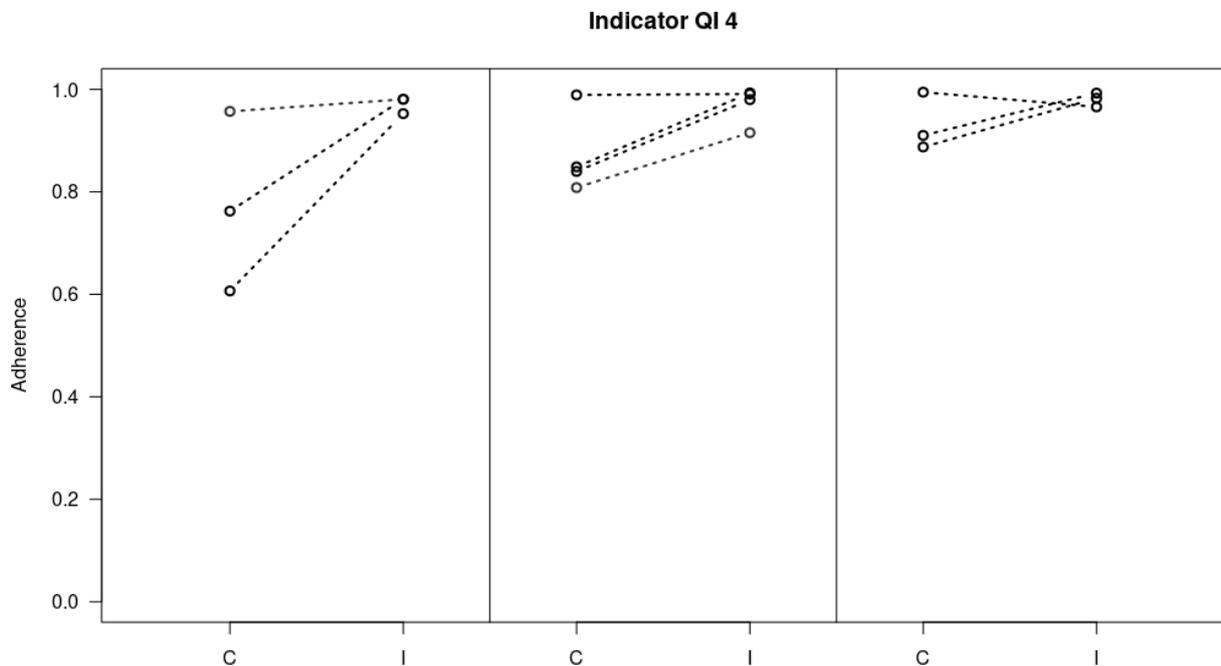


Abbildung 13. QI 4: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I).

4.8.5 QI 6 – Maßnahmen zum Infektionsmanagement

Stratifiziert für die einzelnen Cluster ergab sich für die mittlere QI-Adhärenz, jeweils unter Kontroll- („res.C“) vs. unter Interventions-Bedingung („res.I“) folgendes Bild:

| ID | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| res.C | 0.836 | 0.831 | 0.959 | 0.836 | 0.989 | 0.939 | 0.800 | 0.974 | 0.893 | 0.908 |
| res.I | 0.969 | 0.898 | 0.947 | 0.970 | 0.991 | 0.975 | 0.984 | 0.982 | 0.988 | 0.963 |

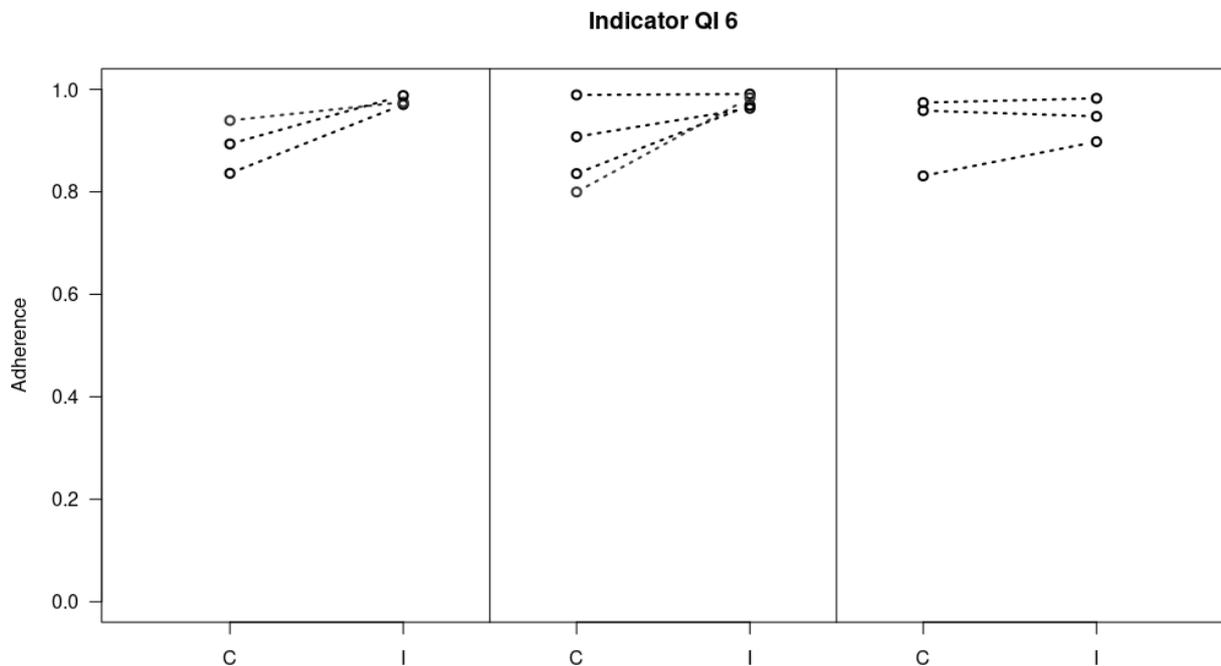


Abbildung 14. QI 6: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I).

4.8.6 QI 7 – Frühe enterale Ernährung

Stratifiziert für die einzelnen Cluster ergab sich für die mittlere QI-Adhärenz, jeweils unter Kontroll- („res.C“) vs. unter Interventions-Bedingung („res.I“) folgendes Bild:

| ID | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| res.C | 0.928 | 0.798 | 0.835 | 0.951 | 0.815 | 0.944 | 0.725 | 0.783 | 0.893 | 0.925 |
| res.I | 0.906 | 0.832 | 0.913 | 0.890 | 0.855 | 0.897 | 0.856 | 0.711 | 0.932 | 0.862 |

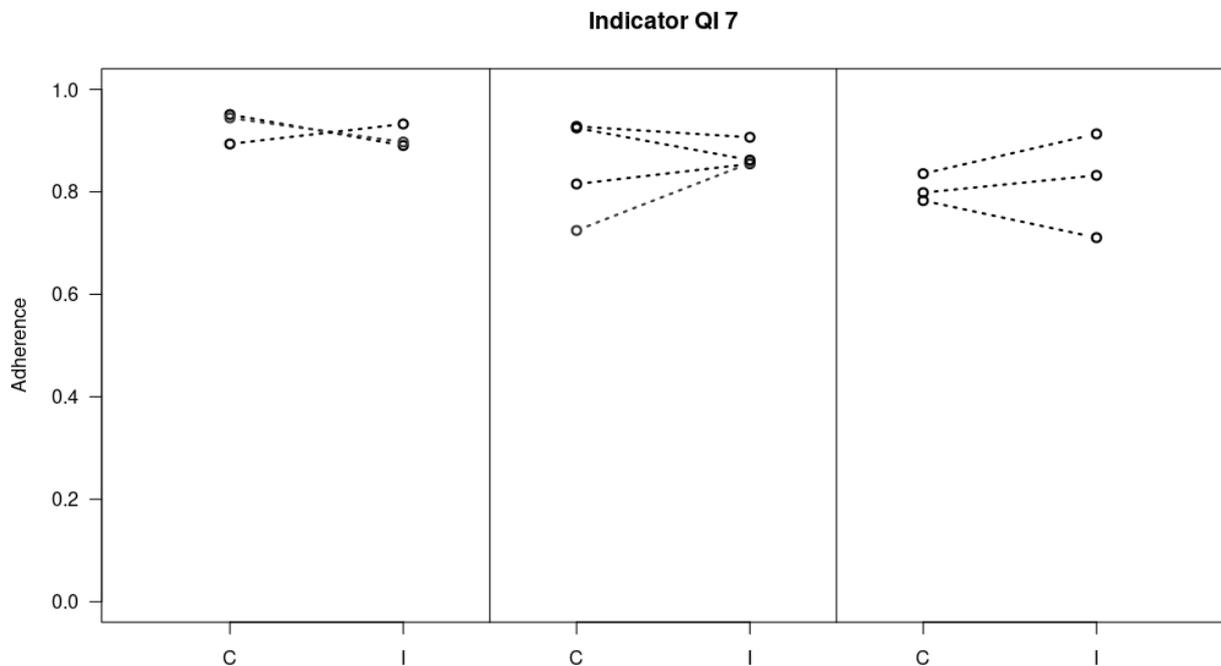


Abbildung 15. QI 7: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I).

4.8.7 QI 8 – Dokumentation einer strukturierten Patienten- & Angehörigenkommunikation

Stratifiziert für die einzelnen Cluster ergab sich für die mittlere QI-Adhärenz, jeweils unter Kontroll- („res.C“) vs. unter Interventions-Bedingung („res.I“) folgendes Bild:

| ID | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| res.C | 0.954 | 0.627 | 0.931 | 0.721 | 0.870 | 0.461 | 0.629 | 0.578 | 0.615 | 0.860 |
| res.I | 0.979 | 0.569 | 0.793 | 0.499 | 0.682 | 0.499 | 0.773 | 0.690 | 0.757 | 0.588 |

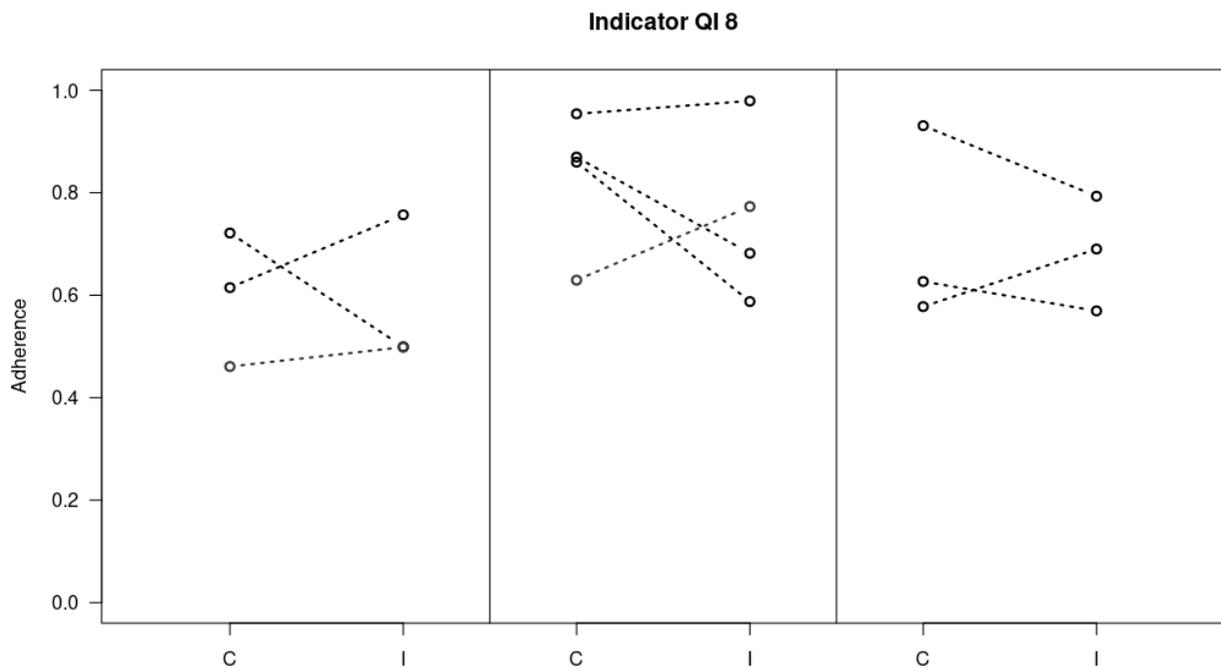


Abbildung 16. QI 8: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I).

4.8.8 QI 9 – Frühmobilisation

Stratifiziert für die einzelnen Cluster ergab sich für die mittlere QI-Adhärenz, jeweils unter Kontroll- („res.C“) vs. unter Interventions-Bedingung („res.I“) folgendes Bild:

| ID | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------|-------|-------|-------|-------|-----|-------|-------|-------|-------|-------|
| res.C | 0.987 | 0.858 | 0.964 | 0.984 | 1.0 | 0.810 | 0.684 | 0.626 | 0.910 | 0.932 |
| res.I | 0.995 | 0.891 | 0.929 | 0.930 | 1.0 | 0.752 | 0.760 | 0.933 | 0.934 | 0.917 |

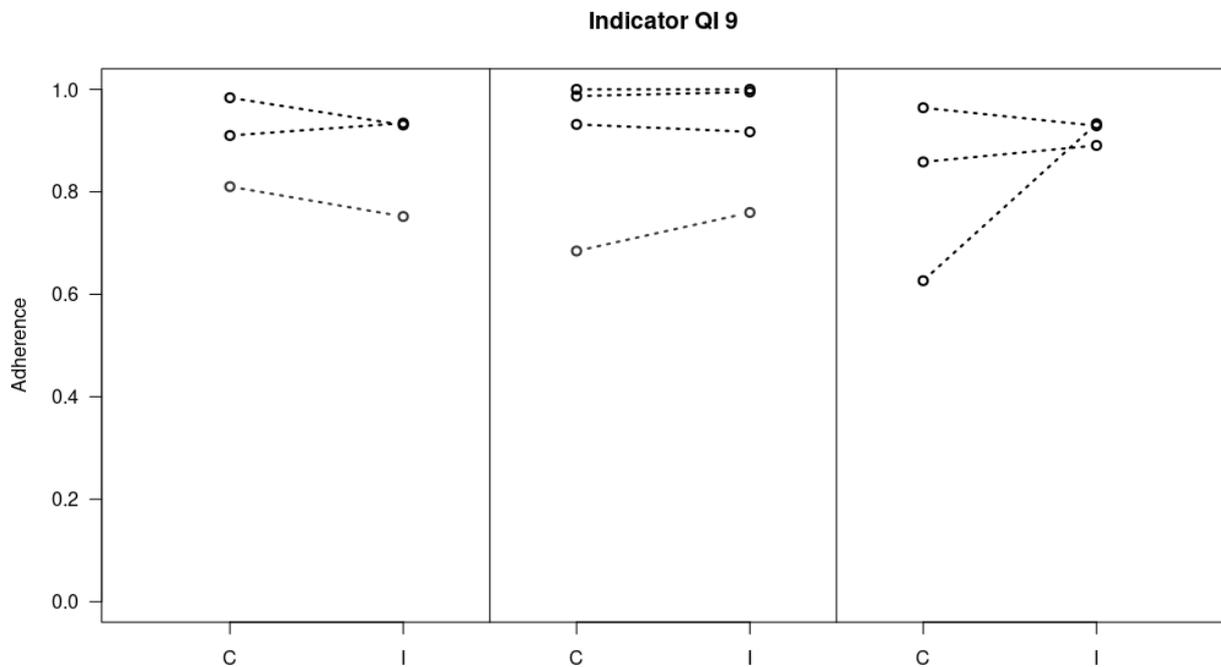


Abbildung 17. QI 9: Deskriptive Plots zur mittleren Adhärenz unter Kontroll- (C) und Interventions-Bedingung (I).

4.9 Modellbasierte confirmatorische Analysen für die einzelnen QIs

Den im Methodenteil Abschnitt 3.8.2.1 (Seite 34) im Detail beschriebenen Modellspezifikationen folgend werden in diesem Abschnitt die Ergebnisse hinsichtlich der Veränderung in der QI-Adhärenz durch die Interventions- im Vergleich zur Kontroll-Bedingung für die 8 untersuchten QIs separat dargestellt. Diese modellbasierten Analysen, welche auch (konfundierende) zeitliche Effekte berücksichtigen, erlauben die in der deskriptiven Darstellung postulierten Interventionseffekte zu quantifizieren und auf Robustheit zu untersuchen.

Die folgende Tabelle 9 zeigt die Schätzergebnisse (mit 95% KIs) des primären Modells PM für die Adhärenz-Wahrscheinlichkeit in Bezug auf einen einzelnen QI, jeweils auf der logit-Skala. Tabelle 10 zeigt als *relative* Effektschätzer die ORs als zentrales Hauptergebnis. Ein relativer Effekt in diesem Modell quantifiziert den Faktor, mit dem sich die Chance für QI-Adhärenz verändert; $OR > 1$ bedeutet einen positiven Effekt, $OR < 1$ bedeutet einen negativen Effekt. Details zu den Modellschätzungen der Hauptanalysen befinden sich in den im Anhang aufgelisteten Tabellen für die einzelnen QIs.

In den anschließenden Unterkapiteln zu den einzelnen QIs werden die Ergebnisse des Hauptmodells (PM) sowie der fünf alternativen Modelle (SM.1, SM.2, SM.3, SM.4, SM.5) im Rahmen von Sensitivitätsanalysen präsentiert.

Eine Zusammenfassung mit Interpretation und graphischer Darstellung der QI-Hauptergebnisse befinden sich im Abschnitt 4.10.

Tabelle 9. Schätzer (fixed effects, mit 95% KIs) des primären Mixed Effects Modells (PM), für die einzelnen QIs. Ergebnisse auf logit-Skala.

| Primäres Outcome | trt | | | time | | | trt x time | | |
|------------------|-------|--------|-------|--------|--------|--------|------------|--------|--------|
| | | 95% KI | | | 95% KI | | | 95% KI | |
| QI 1 | 0.474 | -0.044 | 0.992 | 0.332 | 0.241 | 0.423 | -0.234 | -0.333 | -0.136 |
| QI 2 | 1.673 | 1.350 | 1.996 | -0.009 | -0.075 | 0.057 | -0.018 | -0.087 | 0.051 |
| QI 3 | 0.810 | 0.359 | 1.261 | 0.049 | -0.037 | 0.136 | -0.065 | -0.156 | 0.025 |
| QI 4 | 2.203 | 1.338 | 3.068 | -0.183 | -0.327 | -0.040 | 0.090 | -0.068 | 0.248 |
| QI 6 | 1.481 | 0.702 | 2.260 | 0.116 | -0.009 | 0.242 | -0.154 | -0.296 | -0.012 |
| QI 7 | 0.457 | 0.152 | 0.762 | -0.005 | -0.062 | 0.052 | -0.021 | -0.083 | 0.041 |
| QI 8 | 1.915 | 1.532 | 2.299 | -0.204 | -0.276 | -0.133 | -0.072 | -0.148 | 0.004 |
| QI 9 | 1.151 | 0.878 | 1.424 | -0.182 | -0.240 | -0.124 | 0.101 | 0.039 | 0.162 |

Schätzer auf der logit-Skala.
Interventionseffekt (trt, 95% KI), Effekt der Cluster-spezifischen Expositionszeit in Monaten (time, 95% KI), Interaktionseffekt zwischen Expositionszeit und Intervention (trt x time, 95% KI).

Tabelle 10. Schätzer (fixed effects, mit 95% KIs) des primären Mixed Effects Modells (PM), für die einzelnen QIs.
Dargestellt sind die Odds Ratios (ORs) für die Adhärenz unter Interventions- vs. unter Kontrollbedingung (Routineversorgung).

| Primäres Outcome | trt | | | time | | | trt x time | | |
|------------------|-------|--------|--------|-------|--------|-------|------------|--------|-------|
| | | 95% KI | | | 95% KI | | | 95% KI | |
| QI 1 | 1.606 | 0.957 | 2.697 | 1.394 | 1.273 | 1.527 | 0.791 | 0.717 | 0.873 |
| QI 2 | 5.328 | 3.857 | 7.360 | 0.991 | 0.928 | 1.059 | 0.982 | 0.917 | 1.052 |
| QI 3 | 2.248 | 1.432 | 3.529 | 1.050 | 0.964 | 1.146 | 0.937 | 0.856 | 1.025 |
| QI 4 | 9.052 | 3.811 | 21.499 | 0.833 | 0.721 | 0.961 | 1.094 | 0.934 | 1.281 |
| QI 6 | 4.397 | 2.018 | 9.583 | 1.123 | 0.991 | 1.274 | 0.857 | 0.744 | 0.988 |
| QI 7 | 1.579 | 1.164 | 2.143 | 0.995 | 0.940 | 1.053 | 0.979 | 0.920 | 1.042 |
| QI 8 | 6.787 | 4.627 | 9.964 | 0.815 | 0.759 | 0.875 | 0.931 | 0.862 | 1.004 |
| QI 9 | 3.161 | 2.406 | 4.154 | 0.834 | 0.787 | 0.883 | 1.106 | 1.040 | 1.176 |

Interventionseffekt (trt, 95% KI), Effekt der Cluster-spezifischen Expositionszeit in Monaten (time, 95% KI), Interaktionseffekt zwischen Expositionszeit und Intervention (trt x time, 95% KI).

Bei den einzelnen QIs ergaben sich folgende (unadjustierte) p -Werte für den Interventionseffekt:

| <u>QI 1</u> | <u>QI 2</u> | <u>QI 3</u> | <u>QI 4</u> | <u>QI 6</u> | <u>QI 7</u> | <u>QI 8</u> | <u>QI 9</u> |
|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 0.073 | <0.000000001 | 0.0004347737 | 0.0000006003 | 0.0001955279 | 0.0033068394 | <0.000000001 | <0.000000001 |

Zur Korrektur des multiplen Testens wurde eine Bonferroni-Korrektur im Evaluationskonzept festgelegt, und es wurde keine hierarchische Ordnung für die einzelnen QIs zugrunde gelegt. Bei 8 simultanen Tests und einem globalen Signifikanzniveau von 0.05 gelten die Einzeltests als signifikant, falls $p < 0.05/8 = 0.00625$.

Gesamtresultat:

Damit zeigten sich für alle QIs bis auf den Indikator QI 1 signifikante Effekte der Intervention (Intervention "as implemented"), d.h. die Chance für Adhärenz bzgl. QI 2, QI 3, QI 4, QI 6, QI 7, QI 8, und QI 9 war unter Tele-ICU-Bedingungen höher im Vergleich zur Kontroll-Bedingung (Regelversorgung).

In den nachfolgenden Abschnitten werden die Analyseergebnisse des Hauptmodells (PM) für die einzelnen QIs im Detail und im Rahmen von Sensitivitätsanalysen (Notation: Modelle SM.1, SM.2, SM.3, SM.4, SM.5) hinsichtlich ihrer Robustheit diskutiert.

4.9.1 QI 1 – Tägliche multiprofessionelle und interdisziplinäre klinische Visite mit Dokumentation von Tageszielen

Tabelle 11. QI 1 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen).

| | Baseline log Odds 95% KI | log OR trt 95% KI | log OR time 95% KI | log OR trt x time 95% KI | log OR seq.grp 95% KI | log OR seq.grp x trt 95% KI |
|-------------|--------------------------------|-------------------------|--------------------------|--------------------------------|-----------------------------|-----------------------------------|
| PM | 3.743 2.765 4.721 | 0.474 -0.044 0.992 | 0.332 0.241 0.423 | -0.234 -0.333 -0.136 | - | - |
| SM.1 | 4.651 4.037 5.265 | 0.446 -0.200 1.091 | 0.311 0.199 0.423 | -0.322 -0.446 -0.198 | - | - |
| SM.2 | 2.862 2.192 3.532 | 1.959 1.583 2.335 | - | - | 1.139 0.297 1.98 | -0.232 -0.736 0.272 |
| SM.3 | 3.099 2.695 3.502 | 2.065 1.621 2.509 | - | - | 1.320 0.884 1.756 | -0.209 -0.757 0.339 |
| SM.4 | 3.717 2.962 4.472 | 0.467 -0.054 0.988 | 0.347 0.253 0.441 | -0.254 -0.358 -0.150 | 1.450 0.556 2.344 | -0.309 -0.833 0.215 |
| SM.5 | 2.457 1.575 3.339 | 0.921 0.675 1.167 | 0.158 0.120 0.196 | -0.071 -0.113 -0.029 | - | - |

Schätzer auf der logit-Skala. Interventionseffekt (trt, 95% KI), Effekt der Cluster-spezifischen Expositionszeit (time, 95% KI), Effekt der Sequenzgruppe (seq.grp, 95% KI), Interaktionseffekt zwischen Zeit und Intervention (trt x time, 95% KI) bzw. zwischen Sequenzgruppe und Intervention (seq.grp x time, 95% KI).

trt: binäre Faktorvariable für Treatment, d.h. Behandlung unter Intervention (Referenzkategorie: Kontrolle)

t.θ: metrische Variable für die Cluster-spezifische Expositionszeit in Monaten (negativ vor Beginn der Intervention)

seq.grp: Faktorvariable für Sequenzgruppe (Referenz: mittlere Sequenzgruppe).

PM: Primäres Modell ~ trt * t.θ + (1 | cluster) + (1 | pid)
 SM.1: Sekundäres Modell Nr. 1 ~ trt * t.θ + (1 | pid)
 SM.2: Sekundäres Modell Nr. 2 ~ trt * seq.gr + (1 | cluster) + (1 | pid)
 SM.3: Sekundäres Modell Nr. 3 ~ trt * seq.gr + (1 | pid)
 SM.4: Sekundäres Modell Nr. 4 ~ trt * (t.θ + seq.gr) + (1 | cluster) + (1 | pid)
 SM.5: Sekundäres Modell Nr. 5 ~ trt * t.θ + (1 | cluster)

Ergebnisse für QI 1

Im primären Modell zeigt sich kein signifikanter Interventionseffekt, OR = 1.61 (95% KI [0.957; 2.697]).

Es zeigt sich jedoch ein signifikanter positiver interner Zeittrend (OR für die Expositionszeit pro Monat 1.393, 95% CI [1.273; 1.527]), der sich nach Start der Intervention um den Faktor 0.791 abschwächt (95% KI [0.717; 0.873]).

Der Interventionseffekt und der interne Zeiteffekt (Expositions-Zeit) in SM.4 sind auch konsistent zu den Ergebnissen der beiden Modelle (PM, SM.1), wenn man den Effekt der Sequenzgruppe (externer Zeiteffekt), ebenfalls mitberücksichtigt.

In den Modellen SM.2 und SM.3, die allein den externen Zeiteffekt berücksichtigen, wird ein signifikanter Interventionseffekt gefunden.

Modell SM.2: OR = 7.092, 95% KI [4.870; 10.329]. Es zeigt sich ebenfalls ein positiver Sequenzeffekt (SM.2: OR = 3.124, 95% KI [1.346; 7.243]), aber keine Interaktion zwischen Intervention und Sequenzeffekt.

Auch Modell SM.5 liefert Hinweise auf einen signifikanten Interventionseffekt, OR = 2.512, 95% KI [1.964; 3.212], wenngleich auch deutlich niedriger als bei SM.2 und SM.3. Zusätzlich ergibt sich für SM.5 ein positiver interner Zeittrend, der sich nach Start der Intervention abschwächt um den Faktor 0.931, 95% KI [0.893; 0.971].

4.9.2 QI 2 – Management von Sedierung, Analgesie & Delir

Tabelle 12. QI 2 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen).

| | Baseline log Odds 95% KI | log OR trt 95% KI | log OR time 95% KI | log OR trt x time 95% KI | log OR seq.grp 95% KI | log OR seq.grp x trt 95% KI |
|-------------|--------------------------------|-------------------------|--------------------------|--------------------------------|-----------------------------|-----------------------------------|
| PM | -2.385 -3.280 -1.489 | 1.673 1.350 1.996 | -0.009 -0.075 0.057 | -0.018 -0.087 0.051 | - | - |
| SM.1 | -2.550 -2.881 -2.219 | 2.034 1.620 2.448 | 0.009 -0.071 0.089 | 0.051 -0.034 0.136 | - | - |
| SM.2 | -2.264 -3.024 -1.504 | 1.432 1.203 1.661 | - | - | -1.167 -2.137 -0.197 | 0.556 0.255 0.857 |
| SM.3 | -1.952 -2.201 -1.704 | 1.636 1.357 1.915 | - | - | -1.805 -2.116 -1.494 | 0.885 0.534 1.236 |
| SM.4 | -2.330 -3.113 -1.548 | 1.621 1.295 1.946 | -0.028 -0.096 0.040 | 0.010 -0.062 0.082 | -1.188 -2.169 -0.208 | 0.521 0.205 0.836 |
| SM.5 | -1.923 -2.603 -1.243 | 1.342 1.183 1.501 | -0.029 -0.066 0.008 | -0.002 -0.040 0.036 | - | - |

Schätzer auf der logit-Skala. Interventionseffekt (trt, 95% KI), Effekt der Cluster-spezifischen Expositionszeit (time, 95% KI), Effekt der Sequenzgruppe (seq.grp, 95% KI), Interaktionseffekt zwischen Zeit und Intervention (trt x time, 95% KI) bzw. zwischen Sequenzgruppe und Intervention (seq.grp x time, 95% KI).

PM: Primäres Modell ~ trt * t.0 + (1 | cluster) + (1 | pid)
 SM.1: Sekundäres Modell Nr. 1 ~ trt * t.0 + (1 | pid)
 SM.2: Sekundäres Modell Nr. 2 ~ trt * seq.gr + (1 | cluster) + (1 | pid)
 SM.3: Sekundäres Modell Nr. 3 ~ trt * seq.gr + (1 | pid)
 SM.4: Sekundäres Modell Nr. 4 ~ trt * (t.0 + seq.gr) + (1 | cluster) + (1 | pid)
 SM.5: Sekundäres Modell Nr. 5 ~ trt * t.0 + (1 | cluster)

Ergebnisse für QI 2

Für den QI 2 mit insgesamt niedrigem Level für alle Cluster unter Kontroll-Bedingungen zeigt sich über alle Modelle ein klarer signifikanter positiver Interventionseffekt. Für das primäre Modell ergibt sich ein OR = 5.328, 95% KI [3.857; 7.360], d.h. die Wahrscheinlichkeit für eine QI 2-Adhärenz

Akronym: ERIC
Förderkennzeichen: 01NVF16011

steigt unter Intervention um 432.8%. Es ist auch konsistent über die Modelle kein interner Zeiteffekt (Expositions-Zeit) und keine Interaktion zwischen Intervention und interner Zeit zu erkennen.

Es gibt einen externen Zeiteffekt, der mit fortschreitender Sequenz eine Reduktion der Erfolgsrate des QI 2 zeigt; für das Modell SM.2 ergibt sich ein OR = 0.311, 95% KI [0.118; 0.821], was eine Abschwächung in der Adhärenz um 68.87% in der Gruppe der zuletzt die Intervention implementierenden Cluster (Sequenzgruppe 3) im Vergleich zur mittleren Sequenzgruppe 2 bedeutet. Eine Interaktion zwischen Intervention und Sequenzgruppe ist ebenfalls zu erkennen (OR = 1.744, 95% KI [1.290; 2.356]).

4.9.3 QI 3 – Patientenadaptierte Beatmung

Tabelle 13. QI 3 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen).

| | Baseline log Odds 95% KI | log OR trt 95% KI | log OR time 95% KI | log OR trt x time 95% KI | log OR seq.grp 95% KI | log OR seq.grp x trt 95% KI |
|-------------|--------------------------------|-------------------------|--------------------------|--------------------------------|-----------------------------|-----------------------------------|
| PM | 2.993 2.517 3.469 | 0.810 0.359 1.261 | 0.049 -0.037 0.136 | -0.065 -0.156 0.025 | - | - |
| SM.1 | 3.140 2.772 3.507 | 0.849 0.386 1.311 | -0.008 -0.090 0.074 | -0.029 -0.118 0.060 | - | - |
| SM.2 | 2.794 2.794 2.795 | 1.089 1.089 1.090 | - | - | 0.891 0.890 0.891 | -0.508 -0.509 -0.508 |
| SM.3 | 2.875 2.564 3.186 | 0.911 0.583 1.239 | - | - | 0.811 0.470 1.151 | -0.486 -0.899 -0.072 |
| SM.4 | 2.918 2.550 3.286 | 0.986 0.517 1.456 | 0.022 -0.062 0.107 | -0.041 -0.134 0.052 | 0.821 0.477 1.164 | -0.551 -0.981 -0.121 |
| SM.5 | 1.869 1.483 2.254 | 1.195 0.979 1.411 | 0.026 -0.015 0.067 | -0.077 -0.120 -0.033 | - | - |

Schätzer auf der logit-Skala. Interventionseffekt (trt, 95% KI), Effekt der Cluster-spezifischen Expositionszeit (time, 95% KI), Effekt der Sequenzgruppe (seq.grp, 95% KI), Interaktionseffekt zwischen Zeit und Intervention (trt x time, 95% KI) bzw. zwischen Sequenzgruppe und Intervention (seq.grp x time, 95% KI).

PM: Primäres Modell ~ trt * t.0 + (1 | cluster) + (1 | pid)
 SM.1: Sekundäres Modell Nr. 1 ~ trt * t.0 + (1 | pid)
 SM.2: Sekundäres Modell Nr. 2 ~ trt * seq.gr + (1 | cluster) + (1 | pid)
 SM.3: Sekundäres Modell Nr. 3 ~ trt * seq.gr + (1 | pid)
 SM.4: Sekundäres Modell Nr. 4 ~ trt * (t.0 + seq.gr) + (1 | cluster) + (1 | pid)
 SM.5: Sekundäres Modell Nr. 5 ~ trt * t.0 + (1 | cluster)

Ergebnisse für QI 3

Für den QI 3 zeigt sich über alle Modelle ein klarer signifikanter positiver Interventionseffekt, OR = 2.248, 95% KI [1.432; 3.529].

Es ist auch konsistent über die Modelle kein interner Zeiteffekt (Expositions-Zeit) und keine Interaktion zwischen Intervention und interner Zeit zu erkennen (OR nicht signifikant von 1 verschieden).

Stattdessen gibt es ausreichend Evidenz für einen externen Zeiteffekt, der mit fortschreitender Sequenz eine Steigerung der Erfolgsrate des QI 3 zeigt (OR = 2.437, 95% KI [2.435; 2.438]). Eine reduzierende Interaktion zwischen Intervention und Sequenz ist ebenfalls zu erkennen, OR = 0.601, 95% KI [0.601; 0.602]).

4.9.4 QI 4 – Frühzeitige Entwöhnung von einer invasiven Beatmung (Weaning)

Tabelle 14. QI 4 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen).

| | Baseline log Odds 95% KI | log OR trt 95% KI | log OR time 95% KI | log OR trt × time 95% KI | log OR seq.grp 95% KI | log OR seq.grp × trt 95% KI |
|-------------|--------------------------------|-------------------------|--------------------------|--------------------------------|-----------------------------|-----------------------------------|
| PM | 5.140 4.362 5.919 | 2.203 1.338 3.068 | -0.183 -0.327 -0.040 | 0.090 -0.068 0.248 | - | - |
| SM.1 | 5.142 4.362 5.922 | 2.203 1.338 3.068 | -0.183 -0.327 -0.040 | 0.090 -0.068 0.249 | - | - |
| SM.2 | 5.151 4.399 5.904 | 1.315 0.751 1.878 | - | - | 0.992 0.413 1.572 | -0.796 -1.520 -0.073 |
| SM.3 | 5.064 4.331 5.797 | 1.346 0.792 1.900 | - | - | 0.984 0.415 1.553 | -0.833 -1.546 -0.121 |
| SM.4 | 4.757 4.015 5.499 | 2.438 1.569 3.308 | -0.155 -0.300 -0.010 | 0.051 -0.114 0.216 | 0.905 0.331 1.480 | -1.077 -1.843 -0.311 |
| SM.5 | 2.255 1.794 2.716 | 2.457 2.103 2.812 | 0.008 -0.037 0.053 | -0.109 -0.163 -0.056 | - | - |

Schätzer auf der logit-Skala. Interventionseffekt (trt, 95% KI), Effekt der Cluster-spezifischen Expositionszeit (time, 95% KI), Effekt der Sequenzgruppe (seq.grp, 95% KI), Interaktionseffekt zwischen Zeit und Intervention (trt × time, 95% KI) bzw. zwischen Sequenzgruppe und Intervention (seq.grp × time, 95% KI).

| | |
|-------------------------------|--|
| PM: Primäres Modell | $\sim \text{trt} * \text{t.0} + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.1: Sekundäres Modell Nr. 1 | $\sim \text{trt} * \text{t.0} + (1 \text{pid})$ |
| SM.2: Sekundäres Modell Nr. 2 | $\sim \text{trt} * \text{seq.gr} + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.3: Sekundäres Modell Nr. 3 | $\sim \text{trt} * \text{seq.gr} + (1 \text{pid})$ |
| SM.4: Sekundäres Modell Nr. 4 | $\sim \text{trt} * (\text{t.0} + \text{seq.gr}) + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.5: Sekundäres Modell Nr. 5 | $\sim \text{trt} * \text{t.0} + (1 \text{cluster})$ |

Ergebnisse für QI 4

Im primären Modell zeigt sich ein signifikanter Interventionseffekt $OR = 9.05$ (95% KI [3.811; 21.499]).

Es zeigt sich weiterhin ein signifikanter leicht negativer interner Zeittrend (OR pro Monat 0.832, 95% KI [0.721; 0.961]), der sich nach Interventionsstart nicht signifikant ändert.

Der geschätzte Interventionseffekt und der interne Zeiteffekt (Expositions-Zeit) in SM.4 sind konsistent zu den Ergebnissen der beiden Modelle (PM, SM.1), wenn man den externen Zeiteffekt (Effekt der Sequenzgruppe) ebenfalls mitberücksichtigt.

In SM.5, welches nur Cluster-spezifische Intercepts (keine Variabilität zwischen Patienten, d.h. ohne patientenspezifische Intercepts) berücksichtigt, ist der Interventionseffekt vergleichbar mit dem Schätzergebnis unter PM und SM.1, schwächt sich aber unter Intervention ab, je länger der Interventionsstart zurückliegt.

In den Modellen SM.2 und SM.3, die allein die Sequenzgruppe berücksichtigen, wird ein signifikanter Interventionseffekt gefunden, $OR = 3.725$ (95% KI [2.119; 6.540]) für SM.2. Es zeigt sich ebenfalls ein positiver Sequenzeffekt (SM.2: $OR = 2.696$, 95% KI [1.511; 4.816]), der in Modell SM.2 nach der Intervention um den Faktor 0.451 abgeschwächt wird ($OR = 0.451$, 95% KI [0.219; 0.930]).

4.9.5 QI 6 – Maßnahmen zum Infektionsmanagement

Tabelle 15. QI 6 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen).

| | Baseline log Odds 95% KI | log OR trt 95% KI | log OR time 95% KI | log OR trt × time 95% KI | log OR seq.grp 95% KI | log OR seq.grp × trt 95% KI |
|-------------|--------------------------------|-------------------------|--------------------------|--------------------------------|-----------------------------|-----------------------------------|
| PM | 5.108 4.304 5.912 | 1.481 0.702 2.260 | 0.116 -0.009 0.242 | -0.154 -0.296 -0.012 | - | - |
| SM.1 | 5.109 4.304 5.914 | 1.481 0.701 2.260 | 0.116 -0.010 0.242 | -0.154 -0.296 -0.012 | - | - |
| SM.2 | 4.428 3.713 5.143 | 1.621 1.090 2.152 | - | - | 0.632 0.099 1.164 | -0.959 -1.601 -0.317 |
| SM.3 | 4.440 3.716 5.164 | 1.665 1.131 2.198 | - | - | 0.630 0.096 1.165 | -1.014 -1.659 -0.368 |
| SM.4 | 4.551 3.851 5.251 | 2.000 1.206 2.795 | 0.155 0.034 0.277 | -0.259 -0.401 -0.116 | 0.746 0.232 1.259 | -1.481 -2.152 -0.809 |
| SM.5 | 2.547 2.149 2.944 | 1.679 1.346 2.012 | 0.058 0.021 0.095 | -0.102 -0.148 -0.055 | - | - |

Schätzer auf der logit-Skala. Interventionseffekt (trt, 95% KI), Effekt der Cluster-spezifischen Expositionszeit (time, 95% KI), Effekt der Sequenzgruppe (seq.grp, 95% KI), Interaktionseffekt zwischen Zeit und Intervention (trt × time, 95% KI) bzw. zwischen Sequenzgruppe und Intervention (seq.grp × time, 95% KI).

PM: Primäres Modell ~ trt * t.θ + (1 | cluster) + (1 | pid)
 SM.1: Sekundäres Modell Nr. 1 ~ trt * t.θ + (1 | pid)
 SM.2: Sekundäres Modell Nr. 2 ~ trt * seq.gr + (1 | cluster) + (1 | pid)
 SM.3: Sekundäres Modell Nr. 3 ~ trt * seq.gr + (1 | pid)
 SM.4: Sekundäres Modell Nr. 4 ~ trt * (t.θ + seq.gr) + (1 | cluster) + (1 | pid)
 SM.5: Sekundäres Modell Nr. 5 ~ trt * t.θ + (1 | cluster)

Ergebnisse für QI 6

Im primären Modell zeigt sich ein signifikanter Interventionseffekt, OR = 4.397 (95% KI [2.018; 9.583]). Dieser ist konsistent über die weiteren im Rahmen von Sensitivitätsanalysen betrachteten 4 Modelle.

Es zeigt sich kein klarer interner Zeittrend (in den Modellen PM, SM.1 und SM.4 nur marginal positiv). Dagegen wird der interne Zeittrend (Effekt der Exposition-Zeit) nach der Intervention abgeschwächt (SM.1: OR = 0.857, 95% KI [0.744; 0.988]). Dies erfolgt konsistent über die Modelle PM, SM.1 und SM.4.

Ein externer Zeittrend ist ebenfalls konsistent in den Modellen SM.2, SM.3, SM.4 und SM.5 zu erkennen (SM.2: OR = 1.881, 95% KI [1.104; 3.203]; SM.3: OR = 1.878, 95% KI [1.1008; 3.206]; SM.4: OR = 2.109 [1.261; 3.522]). Dieser Trend schwächt sich unter der Intervention ab, d.h. reduziert sich um 61.7% im Modell SM.2 (OR = 0.383 [0.202; 0.728]), um 63.7% im Modell SM.3 (OR = 0.363 [0.190; 0.692]), und um 77.3% in Modell SM.4 (OR = 0.227 [0.116; 0.445]).

4.9.6 QI 7 – Frühe enterale Ernährung

Tabelle 16. QI 7 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen).

| | Baseline log Odds 95% KI | log OR trt 95% KI | log OR time 95% KI | log OR trt x time 95% KI | log OR seq.grp 95% KI | log OR seq.grp x trt 95% KI |
|-------------|--------------------------------|-------------------------|--------------------------|--------------------------------|-----------------------------|-----------------------------------|
| PM | 2.124 1.820 2.427 | 0.457 0.152 0.762 | -0.005 -0.062 0.052 | -0.021 -0.083 0.041 | - | - |
| SM.1 | 2.002 1.767 2.237 | 0.537 0.216 0.858 | -0.018 -0.078 0.042 | 0.011 -0.054 0.076 | - | - |
| SM.2 | 2.207 1.979 2.436 | 0.179 -0.054 0.413 | - | - | -0.441 -0.711 -0.170 | 0.193 -0.104 0.490 |
| SM.3 | 2.203 1.987 2.419 | 0.189 -0.047 0.425 | - | - | -0.399 -0.645 -0.153 | 0.135 -0.154 0.424 |
| SM.4 | 2.156 1.895 2.418 | 0.475 0.130 0.819 | -0.023 -0.084 0.039 | -0.011 -0.077 0.055 | -0.445 -0.713 -0.178 | 0.074 -0.240 0.389 |
| SM.5 | 1.839 1.558 2.120 | 0.320 0.145 0.496 | 0.015 -0.016 0.047 | -0.039 -0.073 -0.005 | - | - |

Schätzer auf der logit-Skala. Interventionseffekt (trt, 95% KI), Effekt der Cluster-spezifischen Expositionszeit (time, 95% KI), Effekt der Sequenzgruppe (seq.grp, 95% KI), Interaktionseffekt zwischen Zeit und Intervention (trt x time, 95% KI) bzw. zwischen Sequenzgruppe und Intervention (seq.grp x time, 95% KI).

Akronym: ERIC
Förderkennzeichen: 01NVF16011

| | |
|-------------------------------|--|
| PM: Primäres Modell | $\sim \text{trt} * t.\theta + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.1: Sekundäres Modell Nr. 1 | $\sim \text{trt} * t.\theta + (1 \text{pid})$ |
| SM.2: Sekundäres Modell Nr. 2 | $\sim \text{trt} * \text{seq.gr} + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.3: Sekundäres Modell Nr. 3 | $\sim \text{trt} * \text{seq.gr} + (1 \text{pid})$ |
| SM.4: Sekundäres Modell Nr. 4 | $\sim \text{trt} * (t.\theta + \text{seq.gr}) + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.5: Sekundäres Modell Nr. 5 | $\sim \text{trt} * t.\theta + (1 \text{cluster})$ |

Ergebnisse für QI 7

Im primären Modell zeigt sich ein signifikanter Interventionseffekt $OR = 1.579$ (95% KI [1.164; 2.143]). Dieser ist konsistent über die weiteren in der Sensitivitätsanalyse betrachteten Modelle mit Expositions-Zeit (interner Zeit). In den Modellen SM.2 und SM.3 mit ausschließlich externem Zeiteffekt gibt es keine ausreichende Evidenz für einen Interventionseffekt (Effekt nicht mehr signifikant, d.h. $OR = 1.196$ (SM.2) sowie $OR = 1.208$ (SM.3) nicht mehr signifikant von 1 verschieden).

Es zeigen sich kein interner Zeittrend und keine Interaktion zwischen der internen Zeit und der Intervention.

Ein abnehmender externer Zeittrend ist ebenfalls konsistent in den Modellen SM.2, SM.3, SM.4 zu erkennen. Im Modell SM.2 ergibt sich für den Haupteffekt der Sequenzgruppe ein $OR = 0.643$ (95% KI [0.491; 0.844]), d.h. in der Gruppe 3 (später Treatment Switch) ist die Adhärenz im Vergleich zur mittlere Gruppe um 35.7% niedriger. Evidenz für eine Interaktion des externen Zeittrends mit der Intervention liegt nicht vor.

4.9.7 QI 8 – Dokumentation einer strukturierten Patienten- & Angehörigenkommunikation

Tabelle 17. QI 8 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen).

| | Baseline log Odds 95% KI | log OR trt 95% KI | log OR time 95% KI | log OR trt x time 95% KI | log OR seq 95% KI | log OR seq x trt 95% KI |
|-------------|--------------------------------|-------------------------|--------------------------|--------------------------------|-------------------------|-------------------------------|
| PM | 1.740 1.020 2.460 | 1.915 1.532 2.299 | -0.204 -0.276 -0.133 | -0.072 -0.148 0.004 | - | - |
| SM.1 | 1.520 1.213 1.827 | 2.449 2.017 2.881 | -0.266 -0.347 -0.185 | -0.040 -0.125 0.046 | - | - |
| SM.2 | 2.045 1.302 2.789 | -0.356 -0.637 -0.076 | - | - | 0.298 -0.651 1.247 | -0.051 -0.401 0.299 |
| SM.3 | 2.046 1.757 2.334 | -0.234 -0.548 0.080 | - | - | 0.278 -0.046 0.603 | 0.559 0.182 0.936 |
| SM.4 | 1.602 1.601 1.602 | 2.115 2.115 2.116 | -0.179 -0.180 -0.179 | -0.121 -0.122 -0.120 | 0.179 0.178 0.179 | -0.894 -0.894 -0.893 |
| SM.5 | 0.899 0.320 1.479 | 1.310 1.165 1.454 | -0.090 -0.118 -0.063 | -0.070 -0.100 -0.040 | - | - |

Schätzer auf der logit-Skala. Interventionseffekt (trt, 95% KI), Effekt der Cluster-spezifischen Expositionszeit (time, 95% KI), Effekt der Sequenzgruppe (seq.grp, 95% KI), Interaktionseffekt zwischen Zeit und Intervention (trt x time, 95% KI) bzw. zwischen Sequenzgruppe und Intervention (seq.grp x time, 95% KI).

PM: Primäres Modell ~ trt * t.θ + (1 | cluster) + (1 | pid)
 SM.1: Sekundäres Modell Nr. 1 ~ trt * t.θ + (1 | pid)
 SM.2: Sekundäres Modell Nr. 2 ~ trt * seq.gr + (1 | cluster) + (1 | pid)
 SM.3: Sekundäres Modell Nr. 3 ~ trt * seq.gr + (1 | pid)
 SM.4: Sekundäres Modell Nr. 4 ~ trt * (t.θ + seq.gr) + (1 | cluster) + (1 | pid)
 SM.5: Sekundäres Modell Nr. 5 ~ trt * t.θ + (1 | cluster)

Ergebnisse für QI 8

Im primären Modell zeigt sich ein signifikanter Interventionseffekt OR = 6.788 (95% KI [4.627; 9.964]). Dieser ist konsistent über die weiteren in der Sensitivitätsanalyse betrachteten Modelle mit interner Zeit (Expositions-Zeit). In den Modellen mit ausschließlich externem Zeiteffekt schlägt dieser Effekt um und zeigt eine prinzipielle Abnahme der Erfolgsrate unter Intervention an.

Es zeigen sich im Hauptmodell ein abnehmender interner Zeittrend (OR = 0.815, 95% KI [0.759; 0.875]), und nur im Modell SM.4 und SM.5 eine Interaktion zwischen der internen Zeit und der Intervention.

Ein externer Zeittrend ist ebenfalls konsistent in den Modellen SM.2, SM.3, SM.4 nicht zu erkennen. Evidenz für eine Interaktion des externen Zeittrends mit der Intervention zeigt sich nur in Modell SM.4.

4.9.8 QI 9 – Frühmobilisation

Tabelle 18. QI 9 – Schätzer der modellbasierten Analysen (Hauptanalyse und Sensitivitätsanalysen).

| | Baseline log Odds 95% KI | log OR trt 95% KI | log OR time 95% KI | log OR trt x time 95% KI | log OR seq.grp 95% KI | log OR seq.grp x trt 95% KI |
|-------------|--------------------------------|-------------------------|--------------------------|--------------------------------|-----------------------------|-----------------------------------|
| PM | 2.396 1.502 3.291 | 1.151 0.878 1.424 | -0.182 -0.240 -0.124 | 0.101 0.039 0.162 | - | - |
| SM.1 | 1.708 1.471 1.946 | 1.390 1.059 1.721 | -0.250 -0.322 -0.179 | 0.151 0.076 0.227 | - | - |
| SM.2 | 2.841 1.969 3.713 | 0.154 -0.047 0.355 | - | - | -0.427 -1.524 0.669 | 0.597 0.329 0.864 |
| SM.3 | 2.379 2.141 2.617 | 0.070 -0.185 0.326 | - | - | -0.230 -0.492 0.033 | 0.629 0.319 0.939 |
| SM.4 | 2.461 1.596 3.327 | 1.058 0.779 1.337 | -0.185 -0.242 -0.128 | 0.114 0.053 0.176 | -0.522 -1.602 0.559 | 0.414 0.139 0.690 |
| SM.5 | 2.202 1.286 3.118 | 1.008 0.834 1.182 | -0.118 -0.155 -0.082 | 0.052 0.013 0.091 | - | - |

Schätzer auf der logit-Skala. Interventionseffekt (trt, 95% KI), Effekt der Cluster-spezifischen Expositionszeit (time, 95% KI), Effekt der Sequenzgruppe (seq.grp, 95% KI), Interaktionseffekt zwischen Zeit und Intervention (trt x time, 95% KI) bzw. zwischen Sequenzgruppe und Intervention (seq.grp x time, 95% KI).

Akronym: ERIC
Förderkennzeichen: 01NVF16011

| | |
|-------------------------------|--|
| PM: Primäres Modell | $\sim \text{trt} * \text{t.0} + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.1: Sekundäres Modell Nr. 1 | $\sim \text{trt} * \text{t.0} + (1 \text{pid})$ |
| SM.2: Sekundäres Modell Nr. 2 | $\sim \text{trt} * \text{seq.gr} + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.3: Sekundäres Modell Nr. 3 | $\sim \text{trt} * \text{seq.gr} + (1 \text{pid})$ |
| SM.4: Sekundäres Modell Nr. 4 | $\sim \text{trt} * (\text{t.0} + \text{seq.gr}) + (1 \text{cluster}) + (1 \text{pid})$ |
| SM.5: Sekundäres Modell Nr. 5 | $\sim \text{trt} * \text{t.0} + (1 \text{cluster})$ |

Ergebnisse für QI 9

Im primären Modell zeigt sich ein signifikanter Interventionseffekt $OR = 3.161$ (95% KI [2.406; 4.154]). Dieser ist konsistent über die weiteren in der Sensitivitätsanalyse betrachteten Modelle SM.1, SM.4 und SM.5 mit interner Zeit (Expositions-Zeit). In den Modellen mit ausschließlich externem Zeiteffekt ist dieser Effekt nicht mehr vorhanden.

Es zeigen sich ein abnehmender interner Zeittrend (PM: $OR = 0.834$, 95% KI [0.787; 0.883]), und eine positive Interaktion zwischen der internen Zeit und der Intervention (PM: $OR = 1.106$, 95% KI [1.040; 1.176]).

Ein externer Zeittrend ist in den Modellen SM.2, SM.3 und SM.4 nicht zu erkennen. Evidenz für eine positive Interaktion des externen Zeittrends (Sequenzgruppe) mit der Intervention zeigt sich aber konsistent (SM.2: $OR = 1.817$, 95% KI [1.390; 2.372]).

4.10 Zusammenfassende Bewertung der Ergebnisse zu den konfirmatorischen QI-Auswertungen

Das Ergebnis der komplexen Modellierung zum QI-Erfüllungsgrad (Wahrscheinlichkeit für QI-Adhärenz) über die Zeit und nach Interventionsstart ist in nachfolgender Abbildung synoptisch dargestellt. Die Abbildung zeigt pro QI eine Zeitachse [in Monaten], auf der der Zeitpunkt 0 den Start der Intervention anzeigt. Zeit vor der Intervention ist negativ dargestellt, Zeit nach Start der Intervention hat positive Werte. Es werden die geschätzten Cluster-spezifischen Zeitverläufe zum entsprechenden QI-Erfüllungsgrad gezeigt. Der zum Zeitpunkt 0 zu beobachtende Sprung ("*level change*") zeigt pro Cluster den Interventionseffekt auf den Erfüllungsgrad des spezifischen QIs (siehe Parameterschätzer für Faktor *trt* in Tabelle 9 und Tabelle 10). Weiterhin zeigen sich um den Zeitpunkt 0 herum aufsteigende oder abfallende Kurven (siehe Tabelle 9 und Tabelle 10 die Kovariable *time*), die eine Verstärkung oder Abschwächung des Erfüllungsgrades im Laufe der Cluster-spezifischen Expositionszeit wiedergeben (siehe Parameterschätzer für *trt*time* in Tabelle 9 und Tabelle 10).

Abbildung 18 stellt die aus dem Hauptmodell (PM) resultierenden Schätzergebnisse aus Tabelle 9 (logit-Skala) und Tabelle 10 (Odds Ratio-Skala) graphisch dar, ergänzt um die Cluster-spezifischen Effekte. Die schwarze Kurve zeigt pro QI den (marginalen) globalen Effekt. Die 10 farbigen Kurven stellen die geschätzten Effekte auf Cluster-Level dar (d.h. fixed effects und Cluster-spezifische Intercepts).

In allen QIs zeigt sich eine starke Heterogenität zwischen den Clustern. Diese sind zurückzuführen auf die in den spezifischen QI Modellen geschätzten Zufallseffekte (Cluster-spezifische Intercepts). Anhand der Farben ist zu erkennen, dass nicht immer das gleiche Cluster die gleiche Position einnimmt, somit verschiedene Cluster in den einzelnen QIs verschiedenen Erfüllungsgrade aufweisen.

Alle acht QIs bis auf QI 1 zeigten signifikante Verbesserungen. Diese Interventionseffekte befinden sich in Tabelle 9 und Tabelle 10. Sie unterliegen auch zeitlichen Effekten, die sich einerseits als Verstärkung oder Abschwächung zeigen.

Die Ergebnisse zu **QI 1** zeigen, dass keine Evidenz für einen globalen Interventionseffekt vorliegt. Die Analyse zeigte jedoch unter beiden Behandlungsbedingungen eine signifikante Verbesserung im Erfüllungsrad über die Zeit, die sich dann aber nach Start der Intervention leicht abschwächt.

Für **QI 2** ergibt sich ein signifikanter Interventionseffekt hin zu einem besseren Erfüllungsgrad. Weiterhin liegt keine Evidenz für zeitliche Effekte beim Erfüllungsgrad vor, auch wenn die Grafik leichte zeitliche Tendenzen hinsichtlich abnehmender Erfüllungsgrade zeigt. Diese Abnahmen sind aber deutlich geringer als der Gewinn an Adhärenz durch den Interventionseffekt.

Analoges gilt für **QI 3**. Die Analyse zeigte einen signifikanten Interventionseffekt hin zu einem besseren Erfüllungsgrad. Es liegt jedoch keine Evidenz für zeitliche Effekte beim Erfüllungsgrad vor, auch wenn die Grafik leichte zeitliche Tendenzen hinsichtlich zunehmender Erfüllungsgrade vor der Intervention und abnehmender Erfüllungsgrade nach der Intervention zeigt. Wiederum sind diese Abnahmen deutlich geringer als der Gewinn an Erfüllungsgrad durch den Interventionseffekt.

Für **QI 4** gilt, dass alle Cluster einen sehr hohen Erfüllungsgrad bereits unter Kontrollbedingungen zeigten, und dennoch eine kleine, aber signifikante Verbesserung unter Intervention stattfand. Die Analyse zeigte aber auch einen über die ganze Studie andauernden signifikanten negativen

Zeittrend, sowohl unter Kontroll- als auch unter Behandlungsbedingung. Der signifikante positive Effekt der Sequenzgruppe (Zeitpunkt des Crossover) war geringer als der starke negative Effekt des Cluster-spezifischen Zeiteffekts ("time"). Da das Hauptmodell die ganze nicht erklärte Variabilität in die Zufallseffekte der Patienten integriert und keine Variabilität zwischen den Zentren geschätzt werden konnte, wurden für diese Abbildung die Ergebnisse aus dem Modell $\sim \text{trt} * \text{t} . \theta + (1 | \text{cluster})$ verwendet (Modell SM.5 der Sensitivitätsanalysen).

Für **QI 6** gilt ebenfalls, dass alle Cluster einen sehr hohen Erfüllungsgrad bereits unter Routineversorgung zeigten (Modell SM.5), und dennoch eine kleine aber signifikante Verbesserung unter Intervention stattfindet. Die Analyse zeigte eine vor der Intervention stattfindende Verbesserung der QI-Adhärenz. Danach erfolgt der Effekt der Intervention ("Level Change"). Nach Start der Intervention blieb die Adhärenz konstant. Der Trend hin zur Verbesserung wurde nach der Intervention durch einen Gegentrend kompensiert. Dies zeigt sich in der signifikanten Interaktion zwischen Zeit und Intervention ($\text{time} * \text{trt}$) und dem negativen Trend unter Intervention. Da das Hauptmodell die ganze nicht erklärte Variabilität in die Zufallseffekte der Patienten integriert und keine Variabilität zwischen den Zentren geschätzt wurde, wurden für diese Abbildung die Ergebnisse aus dem Modell $\sim \text{trt} * \text{t} . \theta + (1 | \text{cluster})$ verwendet (Modell SM.5 der Sensitivitätsanalysen).

Für **QI 7** ergab sich folgendes Bild: Die statistische Analyse zeigte einen signifikanten Interventionseffekt hin zu einem besseren Erfüllungsgrad. Es liegt jedoch keine Evidenz für zeitliche Effekte beim Erfüllungsgrad vor. Alle Zentren zeigten einen simultan hohen Erfüllungsgrad für QI 7.

Für **QI 8** zeigte sich ebenfalls ein signifikanter Interventionseffekt hin zu einem besseren Erfüllungsgrad. Es liegt auch Evidenz für einen negativen zeitlichen Effekt beim Erfüllungsgrad vor. Dieser wird nicht durch eine Zeittrend-Änderung ('Level Change') nach Start der Intervention kompensiert. Es erscheint auch naheliegend, dass der Interventionseffekt nach etwa 14 Monaten durch den konsistenten negativen Zeiteffekt aufgehoben ist.

Für **QI 9** zeigte sich ebenfalls ein signifikanter Interventionseffekt hin zu einem besseren Erfüllungsgrad. Es liegt auch Evidenz für einen negativen zeitlichen Effekt beim Erfüllungsgrad vor. Das Hauptmodell zeigte aber auch, dass dieser Trend nach Implementierung der Intervention zu einem gewissen Grad, aber nicht vollständig aufgehoben wird. Die Zeittrend-Änderung nach der Intervention wird nicht vollständig kompensiert. Es erscheint daher auch hier naheliegend, dass der Interventionseffekt nach etwa 14 Monaten durch den konsistenten negativen Zeiteffekt aufgehoben ist.

Zusammenfassend lässt sich folgendes feststellen:

1. Durch die Intervention finden bei allen QIs deutliche und signifikante Verbesserungen im Erfüllungsgrad statt. Ausgenommen hiervon ist QI 1 mit einem "marginalen" Ergebnis: OR=1.61, 95% KI [0.96; 2.70].
2. Die statistische Analyse legt nahe, dass bei einigen QIs (QI 2, QI3, QI 4, QI7, QI 8, QI 9) Abschwächungseffekte im Erfüllungsgrad über die Zeit erfolgen. Deshalb wäre zu prüfen, inwieweit sich jährlich wiederholende Schulungen zur QI Erfüllung eignen, den Interventionseffekt wieder voll herzustellen.
3. Die Bewertungen der auf täglicher Basis gemessenen patientenindividuellen QIs erfolgte durch zentrale Ratings an der Charité. Die statistische Analyse geht davon aus, dass diese mit hoher Reliabilität und Objektivität erfolgten. Diese für die Hauptfragestellung wichtigen

methodischen Aspekte sollten in einem zukünftigen Forschungsprojekt wissenschaftlich untersucht und quantifiziert werden, da sie auch bei der Bewertung der internen und externen Validität von Bedeutung sind.

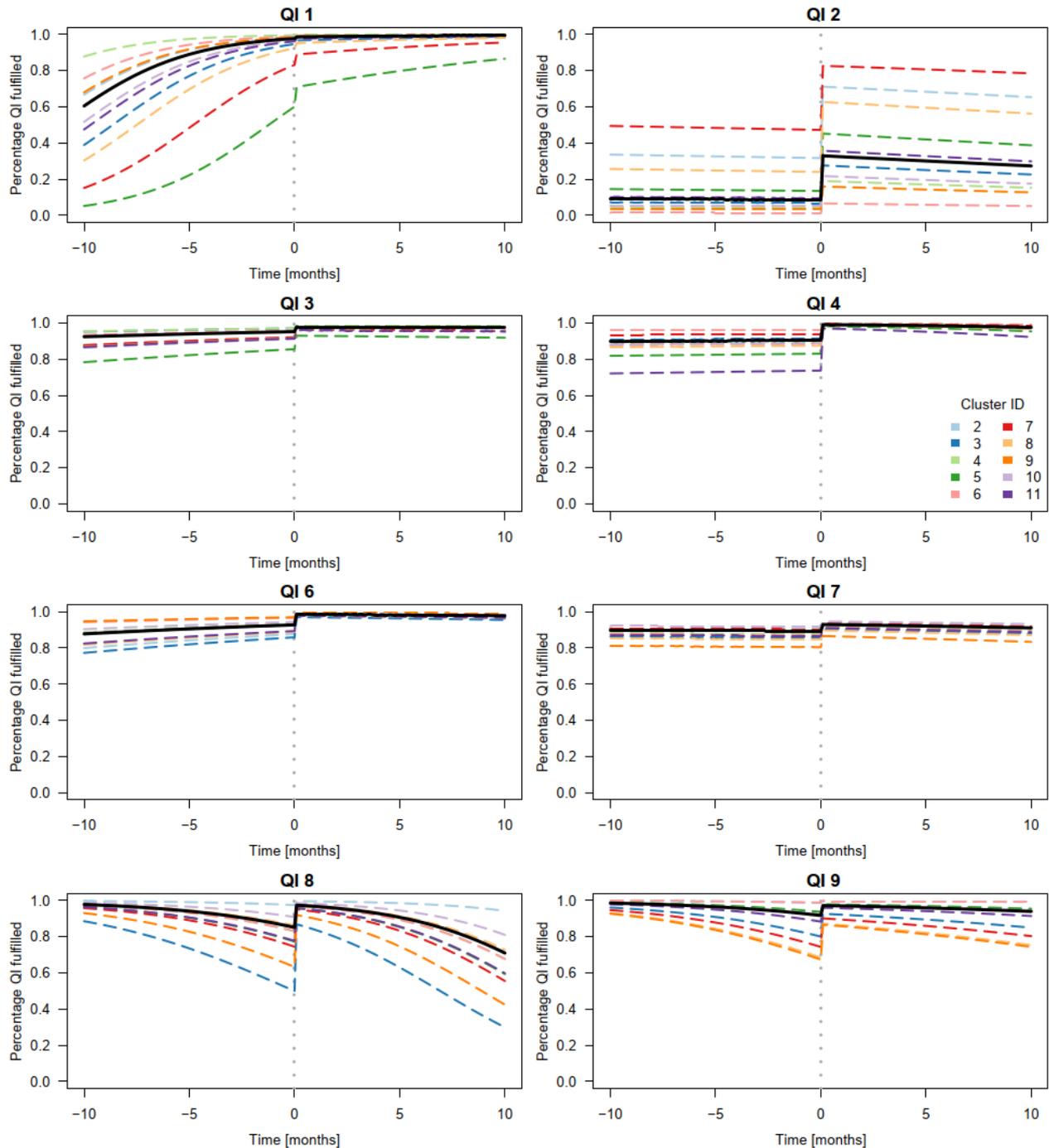


Abbildung 18. Graphische Darstellung der Ergebnisse des primären Modells pro QI: Effekt der Intervention (zugrundeliegende Annahme: "Step Change") und deren Wechselwirkung mit der Cluster-spezifischen Zeit auf die Adhärenz-Wahrscheinlichkeit im Zeitverlauf, d.h. vor vs. nach Start der Intervention. Schwarze Kurve: globaler (marginaler) Effekt (fixed effects ohne random effects) pro QI. Farbige gestrichelte Kurven: fixed effects und Cluster-spezifische Intercepts (patientenindividuelle Intercepts ignoriert). Deutlich werden das hohe Adhärenz-Level einzelner QIs bereits unter Kontroll-Bedingung, und die zeitlichen Effekte. QI 4 und QI 6: Parameterschätzer aus sekundärem Modell SM.5 (analog zu Hauptmodell, aber ohne Random Intercepts für Patient). x-Achse: Zeit in Monaten (30 Tage); y-Achse: Wahrscheinlichkeit für QI-Adhärenz.

5 Ergebnisse der Evaluation – Post-ICU Follow-up Phase

5.1 Basisinformationen

5.1.1 Aspekte zur Datenvollständigkeit bzw. Fallzahlen im Follow-up

Follow-up Daten wurden auf Fall-Ebene dokumentiert. Sofern Patienten >2 FU-Assessments erhielten, wurden für diese wenigen Patienten nur diese nach dem letzten Aufenthalt für die Analyse (pragmatisch) ausgewählt unter Berücksichtigung von im Protokoll geplanten zeitlichen Abläufen (insbesondere bzgl. des Zeitfensters der zwei FU-Visiten).

Als interessierende und im Protokoll präspezifizierte Zeitpunkte der sekundären und explorativen Endpunkte wurden Monat 3 und Monat 6 nach Index ICU-Entlassung definiert. Mit 149 Todesfällen während studienbezogener Index ICU-Aufenthalte würde sich für die FU-Phase eine maximale Analyse-Population von 1313 ergeben. Bei Ignorieren der Planmäßigkeit der FU-Visiten beinhaltete der FU-Datensatz Messungen von maximal 922 Überlebenden, die jeweils ein oder beide FU-Assessments (FU 1, FU 2) erhalten hatten, 249 Patienten behandelt unter Kontroll-Bedingung (Index ICU-Aufenthalt), 673 Patienten unter Intervention (Index ICU-Aufenthalt).

- FU 1: Insgesamt 814 Überlebende erhielten das erste FU-Assessment,
- FU 2: Insgesamt 786 Überlebende erhielten ein zweites FU-Assessment

Für die einzelnen Messinstrumente erhöhte sich ggfs. der Anteil an Missings.

5.1.2 Nacherhebung

Eine einzelne Nacherhebung fand statt am 19.09.2019 als Nachtermin zum FU 2 am 04.09.2019. In allen anderen Fällen fanden die Nacherhebungen statt zwischen 21.05. und 16.10.2020 (FU 1) bzw. 27.05. und 24.10.2020, aufgrund der durch die COVID-19 Pandemie-bedingten Einschränkungen.

5.1.3 Datenmanagement-Aspekte: "Windowing" der FU-Visiten, Problem der zeitlichen Zuordnung

Die Zeitdifferenz "Zeit seit Index ICU-Entlassung" versus Follow-up Visite FU 1 bzw. FU 2 kann wie folgt deskriptiv beschrieben werden:

| | Zeit in Tagen (seit Index ICU-Entlassung) | | | | | |
|-------------|---|---------|--------|-------|---------|-------|
| | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| FU 1 | 21.0 | 81.0 | 93.0 | 105.8 | 117.0 | 378.0 |
| FU 2 | 82.0 | 181.0 | 199.0 | 217.2 | 238.0 | 728.0 |

FU 1 und FU 2 fanden im Median nach 93 (geplant: 90) und 199 (geplant: 180) Tagen statt. Weiterhin zeigte sich, dass insbesondere das FU 2 bei etlichen Patienten zu spät stattfand. Bei 25 Patienten nach 1 Jahr, bei 101 Patienten nach 9 Monaten. 190 Patienten haben die Teilnahme an mindestens einem FU-Assessment verweigert, 59 unter Kontrolle, 131 unter Intervention.

Insgesamt zeigte sich für den tatsächlichen FU-Zeitpunkt, d.h. die Zeit seit Index ICU-Entlassung, eine hohe Variabilität, die im Rahmen der Post-ICU Diagnostik nicht ungewöhnlich ist, aber eine relevante Abweichung von den per Protokoll geplanten Zeitfenstern darstellt.

5.2 Analysen zu sekundären Endpunkten

Im Folgenden werden die Ergebnisse der rein explorativen Analysen zu den sekundären Wirksamkeitsendpunkten beschrieben. Alle aufgeführten p-Werte sind rein deskriptiv (siehe Methodenteil, Abschnitt 3.8.2.2). Eine kausale Interpretation ist nicht möglich.

Eine Dokumentation zur Symptombelastung zum Zeitpunkt der (Index-)Entlassung aus der ICU, oder zum Baseline-Level vor der Index-ICU Phase lag nicht vor.

5.2.1 Gesamt-Mortalität zu Monat 6

Über die 19-monatige ICU-Phase traten insgesamt 158 Todesfälle auf ICU auf, 46 unter Kontrollbedingung, 112 unter Interventionsbedingung (dies beinhaltet 9 Patienten, die bei nachfolgenden ICU-Aufenthalten verstorben waren, vgl. CONSORT Flow Chart).

Für das Post-ICU Follow-up Assessment verbliebe somit eine maximale Analyse-Population von 1304 ICU-Überlebenden (Kontrolle: 368, Intervention: 936 Patienten), d.h. 89.19% der 1462 in der ICU-Phase eingeschlossenen Studienpatienten (natürlicher Nullpunkt), ohne Berücksichtigung verschiedener Dropout-Szenarien in den beiden interessierenden Follow-up Intervallen (z.B. Ablehnung der FU-Untersuchungen), Tod außerhalb der ICU und Loss-to Follow-up.

6-Monats Gesamtmortalität (6 month all-cause mortality)

Bis zur Überprüfung des Überlebensstatus anhand des Melderegistereintrags aller 1432 ERIC-Patienten 12/2020 wurden insgesamt 490 Todesfälle in der Datenbank dokumentiert (Gesamtmortalität, inklusive ICU-Mortalität). Jeder Todesfall nach Tag 180 (bezogen auf die Index-ICU Aufnahme) wurde somit rechts-zensiert (Patient überlebt bis zu Tag 180).

Bis Tag 180 post-Enrolment wurden insgesamt **385 Todesfälle** beobachtet, 107/414 (25.85%) Patienten eingeschlossen unter Kontrolle, 278/1048 (26.53%) Patienten unter Intervention.

In einer einfachen Kaplan-Meier-Schätzung fand sich keine Evidenz für einen Unterschied in der Überlebenswahrscheinlichkeit zwischen den beiden Behandlungsbedingungen, das HR für Intervention- vs. Kontroll-Bedingung betrug 1.033 (95% KI 0.827 bis 1.292; $p=0.773$). Als geschätzte Überlebenswahrscheinlichkeit ergab sich 0.742 (0.701 bis 0.785) unter Kontrolle vs. 0.735 (0.709 bis 0.762) unter Intervention nach 180 Tagen (siehe Anhang, Abbildung 34).

Adjustierte Survival-Kurven für das Cox Proportional Hazards-Modell

Um für einen potentiellen Einfluss der zur Baseline gemessenen Erkrankungsschwere zu adjustieren wurde ein Cox Proportional Hazards-Modell gefittet, adjustiert für SAPS II und SOFA Score (jeweils gemessen zur ersten QI-Visite), mit Frailty-Term (Random Intercept) für Cluster. Es konnte kein Unterschied in der 180-Tage Gesamtmortalität zwischen den beiden Behandlungsbedingungen festgestellt werden, d.h. es gibt keinen signifikanten Effekt der Interventionsbedingung bzgl. des 180-Tage Gesamtüberlebens, HR = 0.847 (0.668 bis 1.073; $p=0.170$). Hierbei war die Anpassungsgüte eines Cox-Modells adjustiert für SAPS II und SOFA höher als bei einem konkurrierenden Modell nur mit Adjustierung für SAPS II ($p=0.001$, X^2 Test). Abbildung

19 zeigt die kumulative Inzidenz unter beiden Behandlungsbedingungen (Routineversorgung vs. Intervention) bei mittlerem Baseline SOFA-Score und SAPS II.

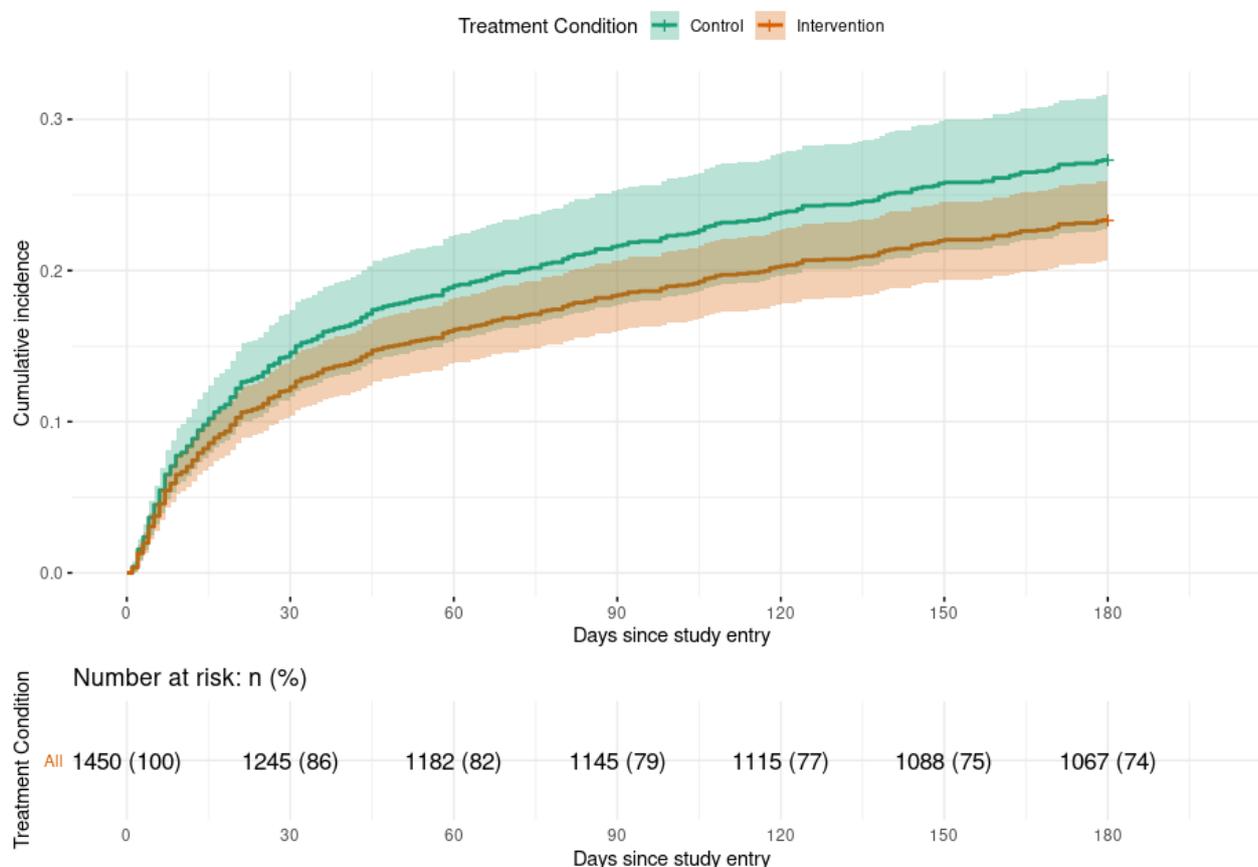


Abbildung 19. Kumulative Inzidenz für die interessierende Variable Behandlungsbedingung (Intervention vs. Kontrolle). Für diese graphische Darstellung wurde der mittlere Baseline SOFA- und mittlere SAPS II-Score zugrunde gelegt. Jede Kurve stellt die erwartete kumulative Inzidenz für die beiden (unbalancierten) Subpopulationen dar (marginaler Effekt der Behandlungsbedingung adjustiert für die beiden Scores für Erkrankungsschwere), geschätzt anhand des Cox-Fraily Modells.

5.2.2 Aktueller Gesundheitszustand anhand der Visuellen Analogskala (VAS) des EQ-5D-5L (FU 1 und FU 2)

Abbildung 20 zeigt einen Verlaufplot für die Messungen unter Kontrolle vs. Intervention. Zur Quantifizierung des Interventionseffekts wurde ein Linear Mixed Effects Modell geschätzt (Random Intercept für Patient sowie Cluster). Wie im Methodenteil spezifiziert, wurde hierbei die volle Information berücksichtigt, d.h. 808 Patienten trugen mit einem oder zwei Assessments bei. Als zeitliche Komponente wurde die exakte FU-Zeit (in Tagen), gerechnet ab erster ICU-Entlassung (Index ICU-Entlassung) definiert. Es wurde adjustiert für die seit dem Ende des Index ICU-Aufenthalts vergangene Zeit sowie für den anhand der Quartile kategorisierten Baseline SAPS II (21-33, 34-48, >48 vs. ≤20 Punkte).

Der globale mittlere EQ VAS beträgt 61.05 Punkte (55.66 bis 66.44). Es gab keine Evidenz für einen Interventionseffekt. Die "Lebensqualität heute" (gemessen anhand der VAS) bei Überlebenden behandelt unter Kontrolle unterscheidet sich nicht von denen unter Interventionsbedingungen, mittlere Differenz = -0.22 Punkte (-6.36 bis 5.92; p=0.943). Die Zeit seit Entlassung nach Index-ICU Aufenthalt spielt keine Rolle, d.h. der VAS bleibt weitestgehend konstant über die

betrachtete patientenspezifische FU-Periode. Für Überlebende mit einem SAPS II >48 Punkte verschlechterte sich der selbstberichtete Gesundheitszustand signifikant um 6.55 Punkte (-11.18 bis -1.91; $p=0.006$) im Vergleich zu geringen Werten ≤ 20 . Details zum Modellfit zeigt Tabelle 26 im Anhang.

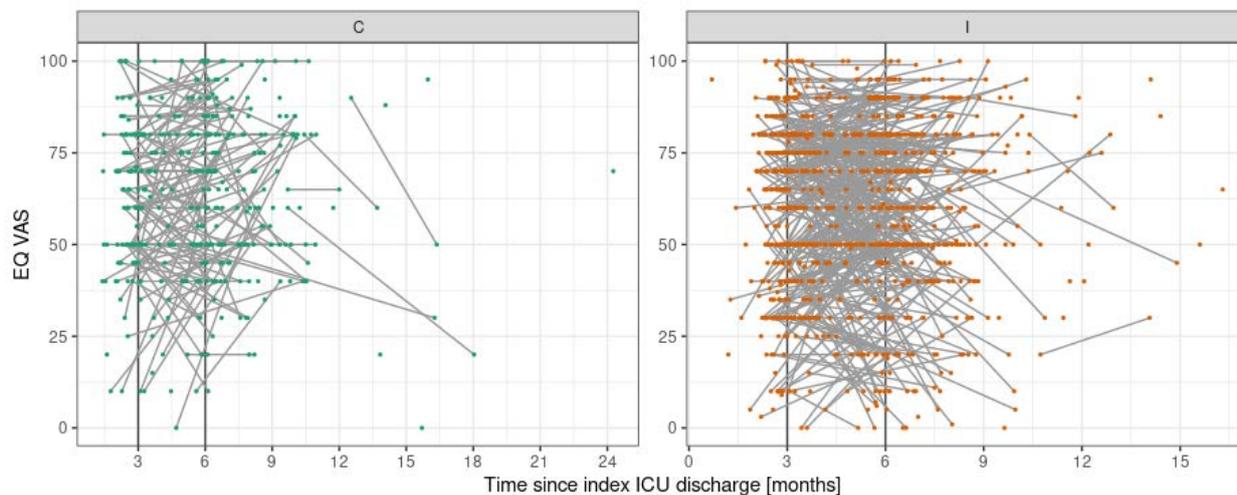


Abbildung 20. EQ VAS für die beiden FU-Zeitpunkte (per Protokoll zu Monat 3 und Monat 6), $N=828$ Patienten. Die zeitlichen Abweichungen sind unter Kontroll-Bedingungen größer als unter Interventionsbedingungen. Definierter Nullpunkt: Index-ICU Entlassungsdatum.

5.2.3 Psychische Gesundheit – Angst und Depression: PHQ Fragebogen

5.2.3.1 Subdomäne Depression: Patient Health Questionnaire–2 (PHQ-2)

Der PHQ-2 (als Teil des umfassenderen Depressionsmoduls des Gesundheitsbogens für Patienten (PHQ-D) ist ein kurzes Screening-Instrument zur Diagnostik von Depressivität, wobei nur die zwei Hauptkriterien der Major Depression nach DSM-IV verwendet werden. Erfragt wird die Häufigkeit der subjektiven Beeinträchtigung durch Interessens- und Freudeverlust (Frage 1) sowie durch Niedergeschlagenheit, Schwermut oder Hoffnungslosigkeit (Frage 2) in den letzten 2 Wochen. Aus der Summation beider Items wurde ein Skalenwert mit Wertebereich von 0–6 Punkten errechnet. Höhere Punktwerte entsprechen einer höheren Symptombelastung. Ein Wert von >2 Punkten (Cut-off Wert) kann als Indikator für pathologische depressive Ausprägungen gewertet werden.

Für 873 Patienten lagen zu FU 1 und/oder FU 2 Messungen anhand des PHQ-2 Fragebogens zu Depression vor.

Es gab keine Evidenz für einen Interventionseffekt. Die Symptombelastung zu Depression, quantifiziert durch den PHQ-2 Summenscore, unterschied sich bei Überlebenden behandelt unter Kontrolle nicht von denen behandelt unter Intervention. Die mittlere Differenz betrug 0.14 Punkte (-0.27 bis 0.55; $p=0.489$). Adjustiert für die Zeit seit Index-ICU-Entlassung (in Tagen) sowie Baseline SAPS II (unter Annahme eines linearen Zusammenhangs) betrug der globale mittlere Summenscore 1.511 (1.14 bis 1.88) Punkte, blieb weitestgehend konstant über die betrachtete patientenspezifische FU-Periode, sowohl für Patienten unter Kontrolle wie unter Intervention. Die Zeit seit Entlassung nach Index-ICU Aufenthalt spielt keine Rolle.

5.2.3.2 Subdomäne Angst: Generalized Anxiety Disorder Scale-2 (GAD-2)

Aus der Summation beider Items zur Domäne Angst wurde ein Skalenwert mit Range 0–6 Punkten errechnet. Höhere Punktwerte entsprechen höheren Belastungen. Ein Wert von >2 Punkten (Cut-off Wert) war als relevanter Indikator für pathologische Angstaussprägungen zu werten.

Für 869 Patienten lagen zu FU 1 und/oder FU 2 Messungen anhand des GAD-2 Fragebogens zur Angststörungen vor.

Es gab keine Evidenz für einen Interventionseffekt. Die Angst-Symptombelastung, quantifiziert durch den GAD-2 Summenscore, unterschied sich für Überlebenden behandelt unter Kontrolle nicht von denen behandelt unter Interventionsbedingungen, die mittlere Differenz beträgt 0.25 Punkte (-0.13 bis 0.63; $p=0.203$). Adjustiert für die Zeit seit Index ICU-Entlassung (in Tagen) betrug der globale mittlere Summenscore 1.47 (1.15 bis 1.79) Punkte, blieb weitestgehend konstant über die betrachtete patientenspezifische FU-Periode, sowohl für Patienten unter Kontrolle wie unter Intervention. Die Zeit seit Entlassung nach Index-ICU Aufenthalt spielt hierbei keine Rolle.

Es konnte keine Assoziation zwischen dem Baseline SAPS II und der Angstbelastung im FU festgestellt werden. Weder im Modell mit linearem Effekt noch in einem alternativen Modell mit kategorisiertem Effekt konnte ein Zusammenhang gefunden werden.

5.2.4 Kognition – Gedächtnis, visuospatiale und visuostrukturelle Fähigkeiten

5.2.4.1 Mini-Cog

Als sekundäres Outcome wurde die dichotomisierte Bewertung "auffällig (<3 Punkte)" vs. "nicht auffällig" verwendet. Missings in der Bewertung, kodiert mit "unbekannt", wurden im Sinne eines *worst score*-Ansatzes auf "auffällig" gesetzt (19 Patienten), d.h. es wurde angenommen, dass für diese Patienten das schlechteste Outcome vorliegt (bei binärem Outcome somit "auffällig"). Hiermit konnten Assessments von 812 Patienten zu FU 1 und/oder FU 2 berücksichtigt werden, um ein logistisches gemischtes Modell mit patientenspezifischen Intercepts zu schätzen, welches adjustiert für die seit Index-ICU Entlassung vergangene Zeit in Monaten.

Es existiert keine ausreichende Evidenz für einen positiven Effekt der Intervention, d.h. das Risiko für ein auffälliges Testergebnis ist nach ICU-Aufenthalt unter Interventions- nicht niedriger als unter der Kontroll-Bedingung, OR = 2.02 (0.73 bis 5.61; $p=0.175$). Es gibt keine Evidenz für einen zeitlichen Effekt, OR = 0.99 (0.88 bis 1.13; $p=0.934$), der ggfs. nach Intervention das Risiko für kognitive Einschränkungen reduziert.

5.2.5 Körperliche Fähigkeiten

5.2.5.1 Subdomäne Mobilität: Timed-Up-and-Go (TUG)-Test

Der Timed-Up-and-Go Test (engl.: Zeitdauer für Aufstehen und Gehen) ist ein einfacher Mobilitätstest zur Beurteilung der Beweglichkeit, Gelenkfunktion bzw. des Körpergleichgewichts und des daraus resultierenden Sturzrisikos insbesondere bei älteren Menschen. Das Ergebnis des TUG-Tests wurde altersstandardisiert dichotomisiert, die körperlichen Fähigkeiten des Patienten klassifiziert als "eingeschränkte Mobilität" vs. "keine eingeschränkte Mobilität" (siehe Anhang).

Missings in der Bewertung, kodiert mit "unbekannt", wurden im Sinne eines *worst score*-Ansatzes auf "auffällig" gesetzt, und daher für 4 Patienten Mobilitätseinschränkungen für diese explorative Analyse zugrunde gelegt.

Von 647 Patienten (bei 1304 ICU-Überlebenden) konnten Assessments zu FU 1 und/oder FU 2 berücksichtigt werden, um ein logistisches gemischtes Modell (mit patientenspezifischen Intercepts) zu schätzen, welches adjustiert für die seit Index-ICU Entlassung vergangene Zeit in Monaten.

Es existiert keine Evidenz für einen positiven Effekt der Intervention, d.h. das Risiko für eingeschränkte Mobilität (TUG-Test "auffällig") ist nach ICU-Aufenthalt unter Interventions- nicht niedriger als unter der Kontroll-Bedingung, OR = 2.15 (0.17 bis 27.26; p=0.554). Es gibt keine Evidenz für einen zeitlichen Effekt, OR = 1.08 (0.80 bis 1.46; p=0.622), der ggfs. nach Intervention das Risiko für kognitive Einschränkungen reduziert.

Eine der Limitationen dieses Messinstruments ist, dass der TUG-Test grundsätzlich selbständiges Aufstehen voraussetzt. Dies bedeutet einen Floor-Effekt, der Fortschritte beim Gehen, solange das Aufstehen nicht selbständig möglich ist, nicht darstellbar macht. Weiterhin ist davon auszugehen, dass die im FU untersuchte Population nicht die in der ICU-Studienphase untersuchte Population repräsentiert. Eine Aussage zu kausalen Zusammenhängen anhand dieser explorativen Analyse ist nicht zulässig.

5.2.5.2 Subdomäne Muskelstärke: Hand-Grip-Strength (HGS)-Test

Das Ergebnis des HGS-Tests wurde altersstandardisiert und geschlechtsspezifisch dichotomisiert, die Muskelstärke des Patienten klassifiziert als "eingeschränkte Handkraft" (ja vs. nein). Es wurden keine Untersuchungen dokumentiert, bei denen die Bewertung mit "unbekannt" dokumentiert wurde.

Von 758 Patienten (bei 1304 ICU-Überlebenden) konnten Assessments zu FU 1 und/oder FU 2 berücksichtigt werden, um ein logistisches gemischtes Modell (mit patientenspezifischen Intercepts) zu schätzen, welches adjustiert für die seit Index-ICU Entlassung vergangene Zeit in Monaten.

Es existiert keine Evidenz für einen positiven Effekt der Intervention, d.h. das Risiko für eingeschränkte Muskelstärke (HGS Test "auffällig") ist nach ICU-Aufenthalt unter Interventions- nicht niedriger als unter der Kontroll-Bedingung, OR = 1.35 (0.10 bis 18.75; p=0.825). Es gibt keine Evidenz für einen zeitlichen Effekt, OR = 0.94 (0.66 bis 1.32; p=0.711), der ggfs. nach Intervention das Risiko für Einschränkungen in der Muskelstärke reduziert.

6 Ergebnisse der Gesundheitsökonomischen Evaluation [TU Berlin]

6.1 Gesundheitsökonomische Evaluation

6.1.1 Datenmanagement und Plausibilitätsanalysen

In Vorbereitung auf die Analysen wurden Plausibilitätsanalysen durchgeführt. Wurden dabei Überschneidungen zwischen Angaben aus dem Hausarztfragebogen 1 und 2 identifiziert, so wurden doppelte Angaben nur einfach und für den HA-CRF 2 einbezogen. Es gab jedoch auch sich überschneidende Angaben, die nicht deckungsgleich waren (z. B. hinsichtlich des Zeitraums einer Rehabilitationsmaßnahme). In diesen Fällen wurde jeweils der längere Zeitraum gewählt. Bei Patienten, die sich bereits vor der intensivmedizinischen Behandlung in einer Einrichtung der stationären Pflege befanden, wurden nur die Tage seit der Entlassung von der Intensivstation, nicht jedoch die Zeiten davor berücksichtigt. Gleiches gilt für den Bezug von Krankengeld. Eine Anschlussheilbehandlung dauert in der Regel 21 Tage. Wurden mehr als doppelt so viele Tage angegeben, stellte das einen unplausiblen Wert dar, der auf 42 Tage gesetzt wurde.

Bei Beobachtungen, wo zwar Angaben im Hausarztfragebogen, jedoch kein Datum vorhanden war, wurde das Datum der Nachvalidierung genutzt. Das betraf bei Beobachtungen mit mind. einem FU insgesamt 65 Patienten im HA-CRF 1 und keine im HA-CRF 2.

Da die Kostenerhebung mithilfe der HA-CRFs nicht zeitgleich mit der Erfassung des EQ-5D-5L erfolgt ist, wurden in Vorbereitung auf die Kosten-Nutzwert-Analyse die Abweichungen zwischen den Erhebungszeitpunkten untersucht. Dabei zeigte sich für das FU 1, dass die Kostendaten mithilfe des Hausarztfragebogen bei 165 Patienten vor der Erfassung des EQ-5D-5L und wiederum bei 61 Patienten danach erfolgt ist. Im FU 2 lag die Kostenerhebung bei 22 Patienten vor der Erfassung des EQ-5D-5L und bei 135 Patienten danach. Da es sich im Durchschnitt mit 12,9 bzw. 14,4 Tagen um eine moderate Abweichung handelt und der Median jeweils bei null Tagen liegt (siehe Tabelle 19), wird dies für die vorgesehene Kosten-Nutzwert-Analyse als unproblematisch erachtet.

Abweichung vom ursprünglich geplanten Vorgehen

Aufgrund der beschriebenen zeitlichen Differenz zwischen den FU und den HA-CRF wurde von einer Extrapolation auf 12 Monate abgesehen, da dies in zusätzliche Unsicherheiten in die ohnehin schon mit Unsicherheiten behafteten Modellierungen resultieren würde.

Tabelle 19. Zeitliche Abweichung zwischen FU-Zeitpunkt und Kostenerhebung (in Tagen)

| | FU 1 | FU 2 |
|---------------------------------|-------|-------|
| Mittelwert | 12,90 | 14,38 |
| Median | 0,00 | 0,00 |
| SD | 38,55 | 50,46 |
| max. Anzahl Tage vor FU | 242 | 267 |
| max. Anzahl Tage nach FU | 226 | 215 |

Darüber hinaus weicht das Datum der Kostenerhebung mithilfe der HA-CRF 1 und HA-CRF 2 von den geplanten 90 bzw. 180 Tagen nach der letzten Entlassung von der Intensivstation ab.

Die Histogramme (Abbildung 21) illustrieren die Tage, die zwischen der Entlassung des letzten Aufenthalts auf der Intensivstation und der ersten bzw. zweiten Kostenerhebung gelegen haben. Für die erste Kostenerhebung zeigt sich, dass die Kostenerhebung im Mittel 116,5 Tage nach der Entlassung erfolgt ist; im Median lag dieser Wert bei 100. Abweichungen von den vorgesehenen 90 Tagen sind angesichts des Studienkollektivs als moderat zu bezeichnen. Die zweite Kostenerhebung mithilfe des Hausarztfragebogens wurde im Mittel nach 226,5 Tagen durchgeführt, wobei der Median bei 209 Tagen lag. Auch wenn die Abweichung für den zweiten Erhebungszeitpunkt größer ist als für den ersten, so wird auch hier von einer moderaten und daher unproblematischen Abweichung von den vorgesehenen 180 Tagen ausgegangen.

Nichtsdestotrotz können die beschriebenen zeitlichen Abweichungen beim FU bzw. bei der Kostenerhebung zu Verzerrungen geführt haben.

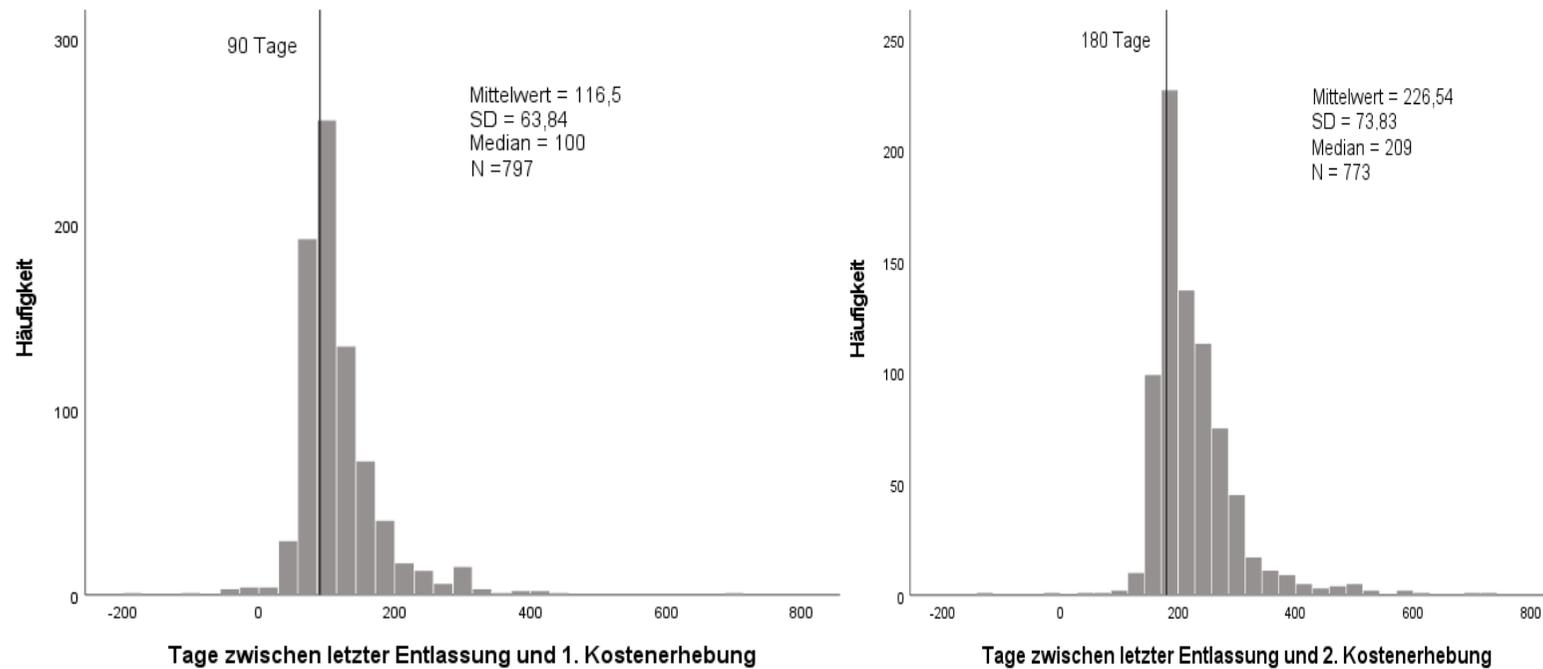


Abbildung 21. Verteilung der Tage zwischen der Entlassung von der letzten Intensivbehandlung und 1. bzw. 2. Kostenerhebung (HA-CRF 1 und HA-CRF 2)

6.1.2 Deskriptive Analysen

6.1.2.1 Kosten

Die beobachteten Kosten für den Indexaufenthalt (inklusive Interventionskosten bei Patienten unter Intervention) sind in Tabelle 20 dargestellt. Darin einbezogen wurden Patienten, bei denen mindestens ein FU vorgelegen hat. Dabei zeigen sich mit 14.231,93 Euro für die Intervention höhere Kosten als für die Kontrolle, bei der der Indexaufenthalt mit durchschnittlich 12.319,63 Euro bewertet wurde.

Tabelle 20. Beobachtete Kosten für den Indexaufenthalt (EUR)

| Intervention | | | Kontrolle | | |
|--------------|-----------|-----------|-----------|-----------|-----------|
| n | MW | SD | n | MW | SD |
| 644 | 14.231,93 | 18.040,51 | 245 | 12.319,63 | 17.474,35 |

Anmerkungen: Für Patienten mit mind. einem FU.

Die **Interventionskosten** in Höhe 651,79 Euro je Patient beinhalten die Betriebs-, Schulungs- und Implementierungskosten.

Die durchschnittlichen Kosten für den Nachbeobachtungszeitraum (d. h. nach der Entlassung entstandene Kosten) sind für Intervention und Kontrolle in Tabelle 21 dargestellt. Dabei ergeben sich sowohl für den Zeitraum zwischen Entlassung von der Intensivstation und der ersten Kostenerhebung (HA-CRF 1), als auch für den Zeitraum zw. erster und zweiter Kostenerhebung (zwischen HA-CRF 1 und HA-CRF 2) jeweils geringere Kosten in der Intervention. In der ersten Kostenerhebung betrifft das alle Kostenarten und auch in der zweiten liegt lediglich die durchschnittliche Höhe des Krankengelds in der Intervention über der in der Kontrolle. In der ersten Kostenerhebung zeigt sich, insbesondere durch die stationären Wiederaufnahmen bedingt, ein höherer Kostenunterschied zwischen Intervention und Kontrolle als in der zweiten Kostenerhebung. Limitierend sind hier die unter 6.1.1. beschriebenen zeitlichen Abweichungen bei der Kostenerhebung zu nennen, die zu Verzerrungen geführt haben können.

Tabelle 21. Übersicht der beobachteten Kosten im HA-CRF 1 und HA-CRF 2 (EUR)

| Kostenkategorie | HA-CRF 1 | | | | | | HA-CRF 2 | | | | | |
|---------------------------------|----------------------|-----------------|-----------|-------------------|------------------|-----------|----------------------|-----------------|-----------|-------------------|-----------------|-----------|
| | Intervention (n=568) | | | Kontrolle (n=212) | | | Intervention (n=527) | | | Kontrolle (n=207) | | |
| | n | MW | SD | n | MW | SD | n | MW | SD | n | MW | SD |
| Rehabilitation | 171 | 3.062,88 | 8.652,50 | 56 | 2.288,51 | 6.703,24 | 107 | 1.860,63 | 6.621,83 | 40 | 1.924,50 | 5.706,41 |
| Stationäre Wiederaufnahmen | 210 | 5.153,00 | 11.605,69 | 75 | 8.161,13 | 23.450,46 | 153 | 5.780,44 | 22.753,10 | 65 | 6.429,28 | 29.202,45 |
| Pflegekosten (stationär) | 20 | 146,70 | 1.040,07 | 8 | 181,37 | 1.800,70 | 26 | 277,38 | 1.606,82 | 10 | 300,88 | 1.618,84 |
| Heilmittel | 174 | 113,79 | 223,40 | 64 | 117,48 | 232,01 | 200 | 115,23 | 207,52 | 74 | 123,83 | 233,30 |
| Ambulante ärztliche Versorgung | 422 | 45,90 | 40,25 | 170 | 54,47 | 49,06 | 402 | 62,76 | 52,33 | 159 | 63,94 | 51,10 |
| Krankengeld | 29 | 316,23 | 1.933,56 | 12 | 433,12 | 2.185,13 | 50 | 1.228,90 | 4.491,26 | 19 | 1.041,58 | 3.997,61 |
| Durchschnitt pro Patient | | 8.838,51 | | | 11.236,07 | | | 9.325,35 | | | 9.884,00 | |

Anmerkungen: Für eine bessere Vergleichbarkeit beziehen sich die Mittelwerte jeweils auf alle Patienten mit entsprechender Kostenerhebung; das n bezieht sich auf Patienten mit entsprechenden Kostenangaben.

Die Gesamtkosten für Intervention und Kontrolle innerhalb des Beobachtungszeitraums ergeben sich aus der Addition der Kosten des Indexaufenthalts sowie der ersten und zweiten Kostenerhebung (siehe Tabelle 22).

Tabelle 22. Gesamtkosten (EUR)

| Zeitraum | Intervention | Kontrolle |
|------------------------------|---------------------|------------------|
| Indexaufenthalt ¹ | 14.231,93 | 12.319,63 |
| Kostenerhebung 1 | 8.838,51 | 11.236,07 |
| Kostenerhebung 2 | 9.325,35 | 9.884,00 |
| Gesamt | 32.395,79 | 33.439,70 |

Anmerkungen: 1: Inklusive Interventionskosten in der Interventionsgruppe.

Demnach ergeben sich im gesamten Beobachtungszeitraum Gesamtkosten in Höhe von 32.395,79 bzw. 33.439,70 Euro für Patienten unter Interventions- bzw. Kontrollbedingungen.

6.1.2.2 Effekte

Qualitätsadjustierte Lebensjahre

Die deskriptive Beschreibung der qualitätsadjustierten Lebensjahre (6 Monate) ist in Tabelle 23 dargestellt. Bezogen auf ein halbes Jahr spiegelt der Wert 0,5 die volle Gesundheit innerhalb des gesamten Zeitraums wieder. In der Kontrolle zeigen sich mit 0,174 durchschnittlich mehr qualitätsadjustierte Lebensjahre als mit 0,155 in der Intervention. Limitierend sind hier die zeitlichen Abweichungen bei der Erhebung der den QALYs zu Grunde liegenden Nutzwengewichte mithilfe des EQ-5D-VL anzumerken, die zu Verzerrungen geführt haben können.

Tabelle 23. Deskriptive Beschreibung der QALYs für 6 Monate

| Intervention (n=666) | | Kontrolle (n=263) | |
|-----------------------------|-------|--------------------------|-------|
| MW | SD | MW | SD |
| 0,155 | 0,150 | 0,174 | 0,156 |

Außerklinische Beatmung

Das Vorliegen einer außerklinischen Beatmung ist in Tabelle 24 dargestellt. Hier zeigt sich in beiden FU ein leicht höherer Anteil an außerklinisch beatmeten Patienten in der Intervention gegenüber der Kontrolle. Betrachtet man im FU 1 die Gruppe jener Patienten, die nicht beatmet sind, so ist dieser mit 88,4 % in der Intervention deutlich höher als mit 76,4 % in der Kontrolle. Im FU 2 wiederum ist der Anteil in der Intervention etwas höher als in der Kontrolle (87,9 % vs. 89,6 %). Bemerkenswert ist der hohe Anteil an Patienten mit unbekanntem Beatmungsstatus bei der Kontrolle im FU 1. Dieser lag mit 19,3 % deutlich über dem in der Intervention (6,1 %). Im FU 2 ist der Anteil vergleichbar. Bleiben die Patienten mit unbekanntem Beatmungsstatus unberücksichtigt, gleichen sich die Werte beim FU 1 an (siehe Tabelle 24b).

Tabelle 24. Deskriptive Beschreibung des Endpunkts außerklinische Beatmung

| | FU 1 | | | | FU 2 | | | |
|-----------|----------------------|------|-------------------|------|----------------------|------|-------------------|------|
| | Intervention (n=553) | | Kontrolle (n=212) | | Intervention (n=527) | | Kontrolle (n=207) | |
| | n | % | n | % | n | % | n | % |
| ja | 30 | 5,5 | 9 | 4,2 | 41 | 7,8 | 13 | 6,3 |
| nein | 489 | 88,4 | 162 | 76,4 | 463 | 87,9 | 186 | 89,6 |
| unbekannt | 34 | 6,1 | 41 | 19,3 | 23 | 4,4 | 8 | 3,9 |

Anmerkungen: Fälle mit vorhandenem CRF, aber fehlenden Angaben wurden als unbekannt gewertet.

Tabelle 24b. Deskriptive Beschreibung des Endpunkts außerklinische Beatmung für Fälle mit bekanntem Beatmungsstatus

| | FU 1 | | | | FU 2 | | | |
|------|----------------------|------|-------------------|------|----------------------|------|-------------------|------|
| | Intervention (n=519) | | Kontrolle (n=171) | | Intervention (n=504) | | Kontrolle (n=199) | |
| | n | % | n | % | n | % | n | % |
| ja | 30 | 5,8 | 9 | 5,3 | 41 | 8,1 | 13 | 6,5 |
| nein | 489 | 94,2 | 162 | 94,7 | 463 | 90,9 | 186 | 93,5 |

Auch wenn es sich dabei nur um wenige Fälle handelt, soll an der Stelle noch der Beatmungsstatus jener Patienten betrachtet werden, die als beatmet von der Intensivstation entlassen worden sind (siehe Tabelle 25).

Von den 888 ICU-Überlebenden sind insgesamt 120 Patienten mit dem Status „beatmet“ entlassen worden. In der Kontrolle waren es 20 und in der Intervention 100 Patienten. Davon sind in der Kontrolle 20 % und in der Intervention 25 % vor dem FU 1 (90 Tage nach der Entlassung des letzten Aufenthalts) verstorben. Von den Überlebenden sind in der Kontrolle 18,8 % und in der Intervention 14,7 % beatmet. Mit 49,3 % sind unter Interventionsbedingungen deutlich mehr Patienten erfolgreich von der Beatmung entwöhnt als in der Kontrolle, wo dieser Anteil 37,5 % beträgt. Betrachtet man das FU 2, wird der Unterschied noch größer. Hier sind von den bis zum FU 2 überlebenden Patienten 49,3 % in der Intervention von der Beatmung entwöhnt, wohingegen der Anteil in der Kontrolle auf 20 % gesunken ist. Bemerkenswert ist auch hier der Anteil an Patienten mit unbekanntem Beatmungsstatus, der in der Kontrolle jeweils deutlich höher ist als in der Intervention.

Tabelle 25. Deskriptive Beschreibung des Endpunkts außerklinische Beatmung als beatmet entlassener Patienten.

| | FU 1 | | | | FU 2 | | | |
|------------------------|----------------------|-------------|------------------|-------------|---------------------|-------------|------------------|-------------|
| | Intervention (n=100) | | Kontrolle (n=20) | | Intervention (n=75) | | Kontrolle (n=16) | |
| | n | % | n | % | n | % | n | % |
| verstorben | 25 | 25,0 | 4 | 20,0 | 4 | 5,3 | 1 | 6,3 |
| beatmet ¹ | 11 | 14,7 | 3 | 18,8 | 14 | 19,7 | 3 | 20,0 |
| entwöhnt ¹ | 37 | 49,3 | 6 | 37,5 | 35 | 49,3 | 3 | 20,0 |
| unbekannt ¹ | 27 | 36,0 | 7 | 43,8 | 22 | 31,0 | 8 | 53,3 |

Anmerkungen: 1: Die Prozentangaben beziehen sich auf die Überlebenden.

6.1.3 Inferenzstatistische Analysen

6.1.3.1 Kosten

Bei der Analyse der Gesamtkosten zeigte das lineare gemischte Modell der logarithmierten Kosten die beste Modellgüte im Vergleich zu alternativen getesteten Modellen. Dabei wurden 727 Beobachtungen eingeschlossen und im Hauptmodell (M1; siehe Anhang, Tabelle 32) für die Intervention, das Alter, das Geschlecht (Referenz: männlich), den SAPS II und die Zeit seit der letzten Entlassung von der Intensivstation adjustiert. Das Cluster wurde als zufälliger Effekt modelliert. Das Modell zeigt keinen Effekt der Intervention (95% KI: -0,25; 0,92; $p=0,263$) im Hinblick auf die Kosten. Dahingegen zeigt sich ein signifikanter Einfluss des SAPS II. Die Ergebnisse bestätigen sich in einem alternativen Modell (M2; siehe Anhang Tabelle 32) und in den Sensitivitätsanalysen unter Anwendung anderer für Kostendaten geeigneter Regressionsverfahren (siehe Anhang, Tabellen 33 bis 34).[59] Neben den Gesamtkosten wurden die Kosten des Indexaufenthalts inklusive der Interventionskosten gesondert analysiert. Auch hier zeigte sich in keinem der Modelle ein statistisch signifikanter Interventionseffekt (siehe Anhang, Tabellen 35 bis 37).

6.1.3.2 Effekte

Qualitätsadjustierte Lebensjahre

Die qualitätsadjustierten Lebensjahre wurden mithilfe eines linearen gemischten Modells analysiert. Dabei wurden 897 Beobachtungen eingeschlossen. Im Hauptmodell (M1; siehe Anhang, Tabelle 38) wurde für die Intervention, das Alter, das Geschlecht (Referenz: männlich) und den SAPS II adjustiert. Das Cluster wurde als zufälliger Effekt berücksichtigt. Der globale Mittelwert beträgt 0,44. Das Modell zeigt keinen Effekt der Intervention (95% KI: -0,03; 0,01; $p=0,239$) im Hinblick auf die qualitätsadjustierten Lebensjahre. Dahingegen zeigt sich ein signifikanter Einfluss des SAPS II, des Alters, des Geschlechts und der Dauer des Indexaufenthalts. Die Ergebnisse bestätigen sich in einem alternativen Modell (M2) (siehe Tabelle 38 im Anhang).

Außerklinische Beatmung

Das Vorliegen einer außerklinischen Beatmung als binärer Endpunkt wurde mithilfe eines generalisierten gemischten Logit-Modells analysiert. Dabei wurde neben dem Cluster auch ein patientenspezifischer zufälliger Effekt innerhalb des Clusters modelliert. In dieses Modell sind insgesamt Angaben von 817 Patienten mit mindestens einem FU eingeflossen. Darüber hinaus ist die Zeit zwischen der Erhebung des Hausarztfragebogens und der letzten Entlassung zur Kontrolle von zeitlichen Unterschieden des Erhebungszeitpunkts in das Modell aufgenommen worden. Dabei zeigt sich im Hauptmodell (M1; siehe Anhang, Tabelle 39) mit einem Odds Ratio von 0,01 eine gegenüber Kontrollbedingungen deutlich verringerte Beatmungswahrscheinlichkeit in der Intervention, die jedoch statistisch nicht signifikant ist (95% KI: 0,00; 3,26; $p=0,118$). Auch die Zeit seit der Entlassung sowie der SAPS II, das Alter und das Geschlecht zeigen keinen signifikanten Einfluss auf die Beatmungswahrscheinlichkeit. Das bestätigt sich auch in einem alternativen Modell (M2) (siehe Anhang, Tabelle 39).

6.1.4 Inkrementelle Kosteneffektivitätsrelation (IKER)

Von einer Berechnung der IKER wurde aufgrund des nicht nachweisbaren (nicht signifikanten) Unterschieds der Kosten und Effekte zwischen Kontrolle und Intervention abgesehen.

7 Ergebnisse der Akzeptanzerhebung [TU Berlin]

Insgesamt haben 19 Teilnehmer an der Befragung teilgenommen, die alle angegeben haben, zum ärztlichen Personal zu gehören. 10 Befragte gaben an, in einem Krankenhaus der Zentralversorgung (>500 Betten) zu arbeiten, wohingegen fünf in einem Krankenhaus der Regelversorgung (300 bis 500 Betten) und die verbleibenden vier in einem Krankenhaus der Grundversorgung (<300 Betten) tätig sind.

Teil 1: Nutzerfreundlichkeit der Televisitenroboter

Die Mehrheit der Befragten gab an, dass Sie den Televisitenroboter gern häufig nutzen würden. Nur wenige empfanden den Roboter als unnötig komplex oder umständlich zu nutzen. Der Aussage, dass der Televisitenroboter einfach zu nutzen sei, stimmten 80% Befragten voll und die verbleibenden 20% ebenfalls zu. Eine Person, die bei der Benutzung anleitet, wird von fast allen Befragten als nicht notwendig angesehen. Etwa die Hälfte der Befragten stimmte der Aussage zu, dass die verschiedenen Funktionen in dem Televisitenroboter gut integriert waren. Insgesamt fühlten sich etwa zwei Drittel der Befragten sehr sicher im Umgang mit dem Roboter.

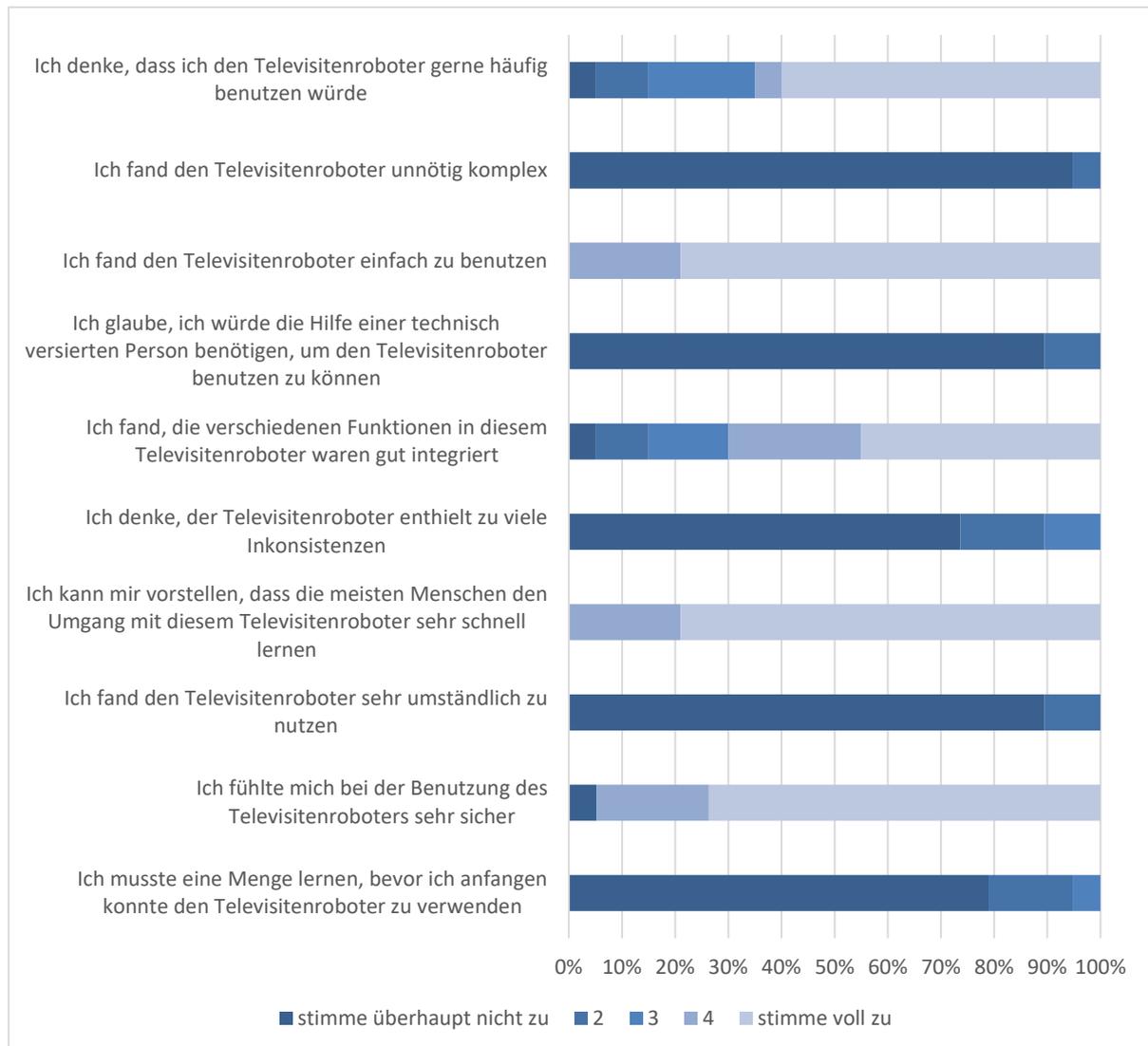


Abbildung 22. Nutzerfreundlichkeit der Televisitenroboter.

Teil 2: Subjektive Nutzenbewertung

Der Großteil der Befragten (90%) gab an, dass die telemedizinische Beratung sie in schwierigen Entscheidungen unterstützt hat. Mehr als drei Viertel hatten in der telemedizinischen Beratung ausreichend Zeit für die eigenen Anliegen und konnten bei Bedarf auf die telemedizinische Beratung zurückgreifen. Eine Verringerung der Arbeitsbelastung resultierte nur bei wenigen Teilnehmern aus der Nutzung der telemedizinischen Beratung. Für eine evidenzbasierte Entscheidungsfindung sowie bei schwierigen Entscheidungen empfanden jedoch zwei Drittel die telemedizinische Beratung als hilfreich. Fast 80% der teilnehmenden Ärzte gaben an, dass die telemedizinische Beratung das Potential habe, auch in anderen Krankenhäusern evidenzbasierte Behandlungen positiv beeinflussen zu können.

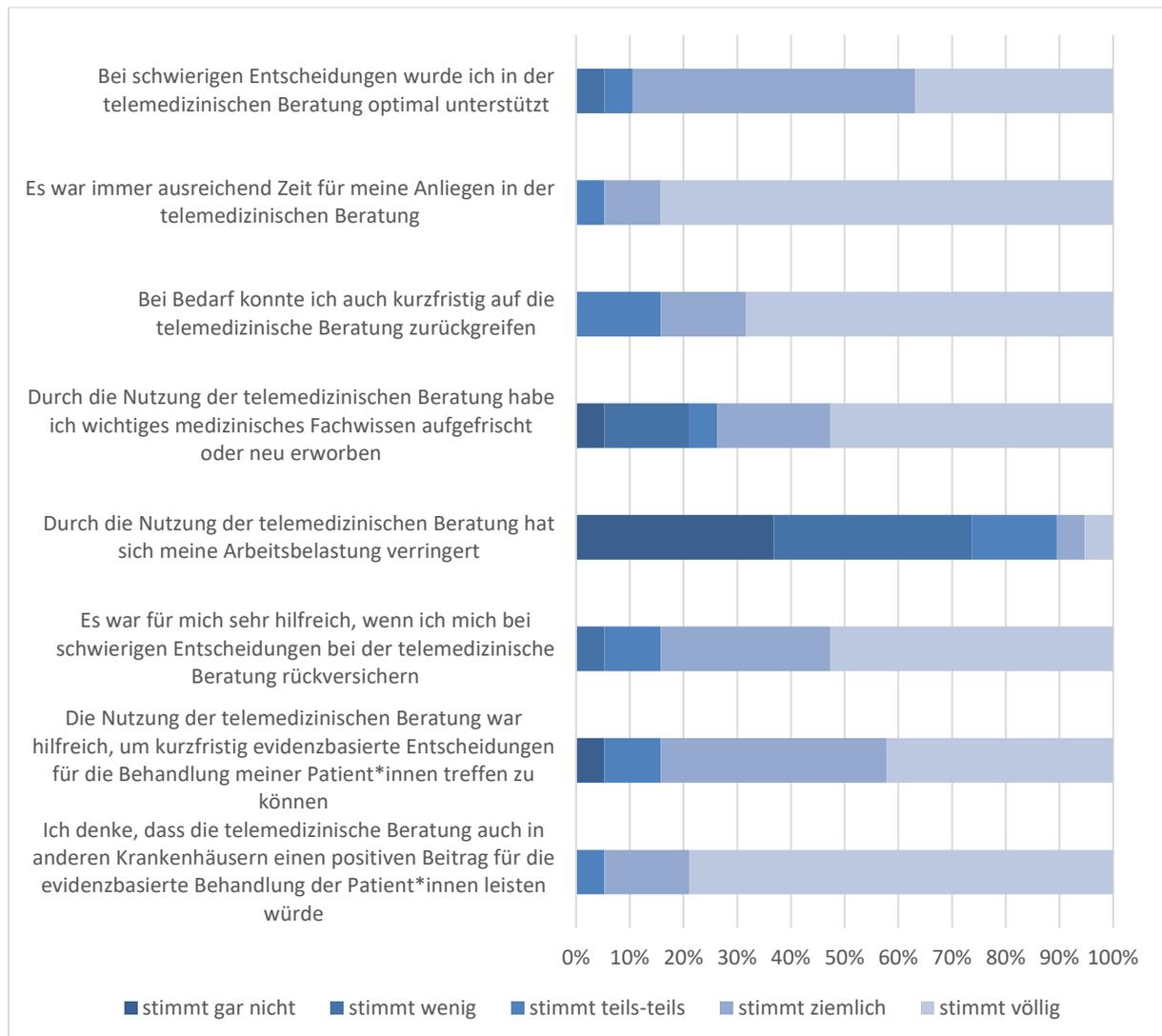


Abbildung 23. Subjektive Nutzenbewertung der Televisitenroboter (1).

Mehr als die Hälfte der Befragten gab an, dass eventuelle Unsicherheiten oder Behandlungsfehler immer angesprochen werden konnten. Alle Befragten empfanden den Telemediziner stets als freundlich und konstruktiv und fast alle teilnehmenden Ärzte gaben an, dass die Telemediziner auf einen respektvollen Umgang mit den Patienten geachtet haben. Dass die Patienten die Televisiten sehr gut angenommen haben, berichtete über ein Drittel der Befragten. 80% der Befragten gaben an, die gemeinsam ausgearbeiteten Behandlungspläne umzusetzen, wohingegen 20% diesem Punkt nur teils-teils zugestimmt haben.

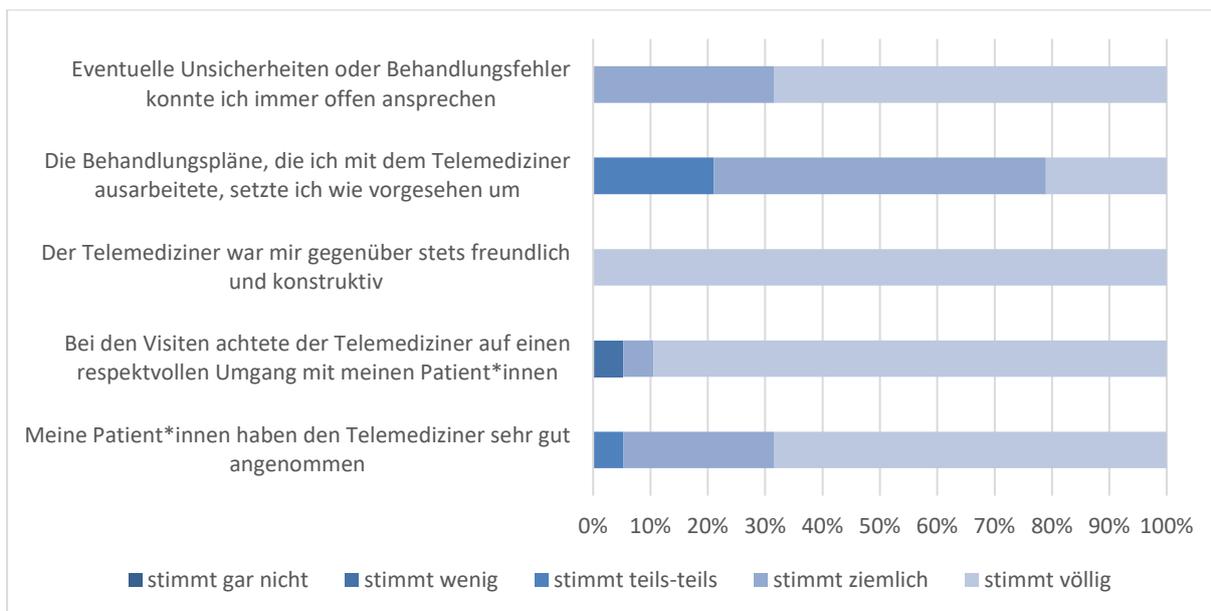


Abbildung 24. Subjektive Nutzenbewertung der Televisitenroboter (2).

Etwa die Hälfte der Befragten gab an, dass Vitaldaten und Testergebnisse mit Hilfe der Kamera problemlos übermittelt werden konnten. 80% der Befragten stimmten völlig oder ziemlich zu, dass die Bildqualität gut und fehlerfrei war. Ein Drittel der befragten Ärzte empfand auch die Tonqualität als gut und fehlerfrei und nur für etwa ein Viertel der Befragten traf dies nur teilweise zu. Die Verständigung mit dem Telemediziner funktionierte bei fast allen Befragten technisch stets reibungslos.

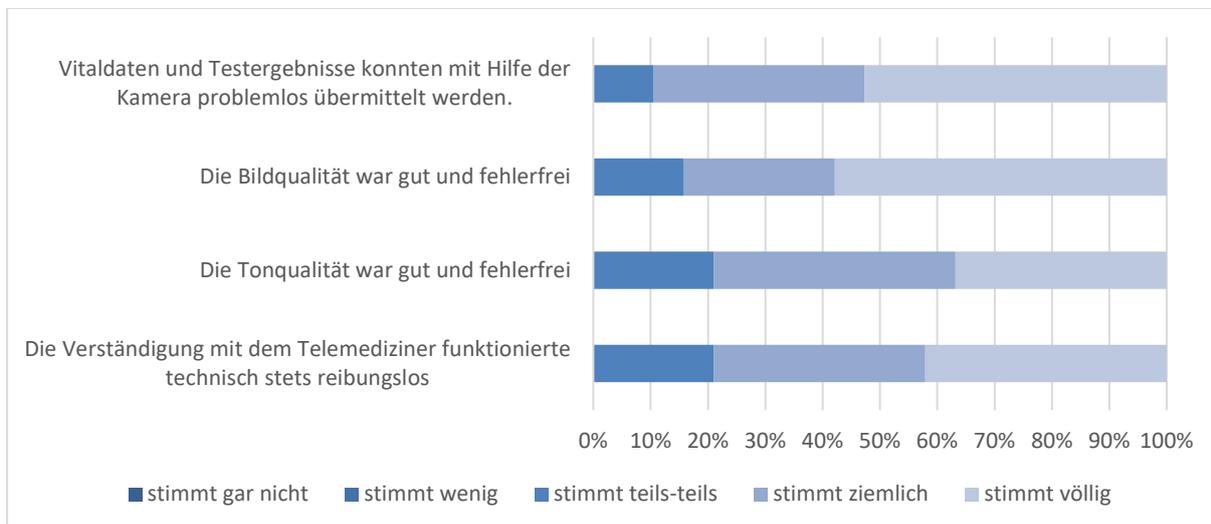


Abbildung 25. Subjektive Nutzenbewertung der Televisitenroboter (3).

Bei hohem Patientenaufkommen war die Nutzung der telemedizinischen Beratung für 80% der Befragten eher wenig oder nur teilweise hilfreich. Für etwa die Hälfte war es bei der Behandlung der Infektionskrankheit COVID-19 hilfreich, auf die Einschätzung eines darin erfahrenen Telemediziners zurückgreifen zu können. Die Mehrheit (80%) stimmte wenig oder nur teilweise zu, dass bei hohem Patientenaufkommen keine Zeit für die Durchführung einer telemedizinischen Beratung blieb.

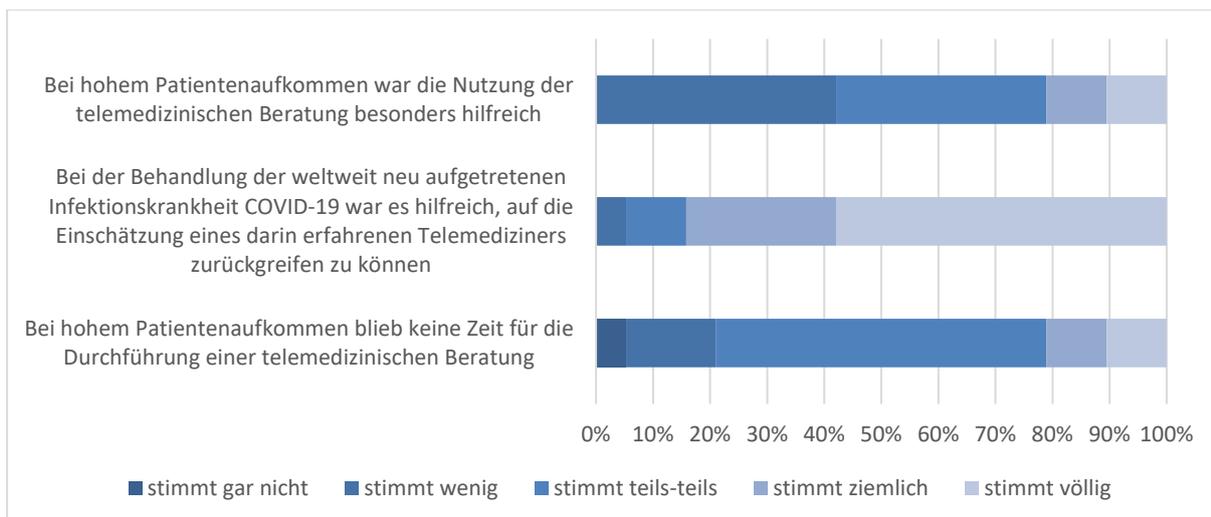


Abbildung 26. Subjektive Nutzenbewertung der Televisitenroboter (4).

Etwa 80% der Befragten gaben an, dass die telemedizinische Beratung bei der Behandlung ihrer Patienten geholfen hat. Den übrigen Befragten hatte sie etwas geholfen.

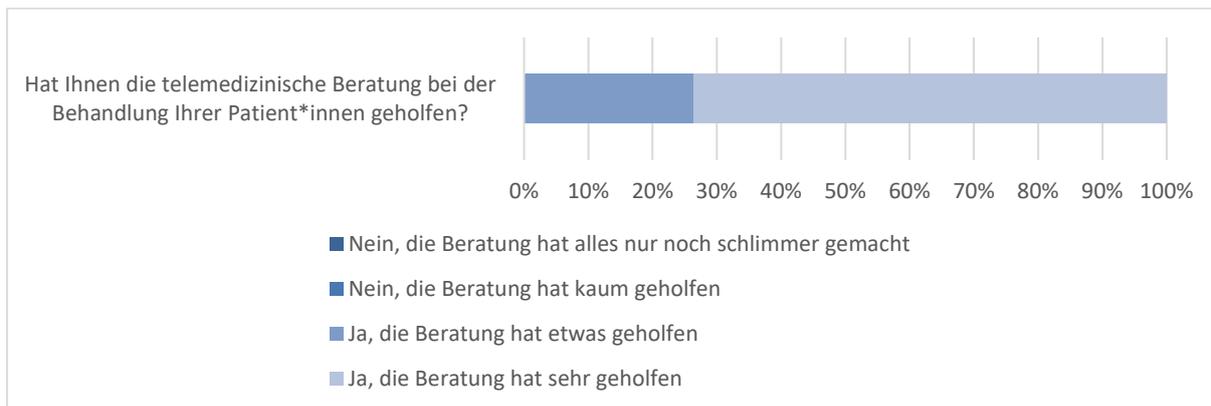


Abbildung 27. Subjektive Nutzenbewertung der Televisitenroboter (5).

Teil 3: Ausmaß und Belastung infolge der Krisensituation

Während der Teilnahme am ERIC-Projekt war von etwas mehr als der Hälfte der Befragten die Station infolge der SARS-CoV-2-Pandemie überdurchschnittlich stark ausgelastet. Fast die Hälfte der Befragten stimmte der Aussage, dass ihre Station infolge der SARS-CoV-2-Pandemie so stark ausgelastet war, dass zumindest zeitweise eine Krisensituation zu bewältigen war, gar nicht oder wenig zu. Wiederum etwas mehr als die Hälfte stimmten ziemlich oder völlig zu, dass auf der eigenen Station die Situation während der Teilnahme am ERIC-Projekt (zumindest zeitweise) infolge eines sehr hohen COVID-19-Patientenaufkommens außergewöhnlich angespannt war.

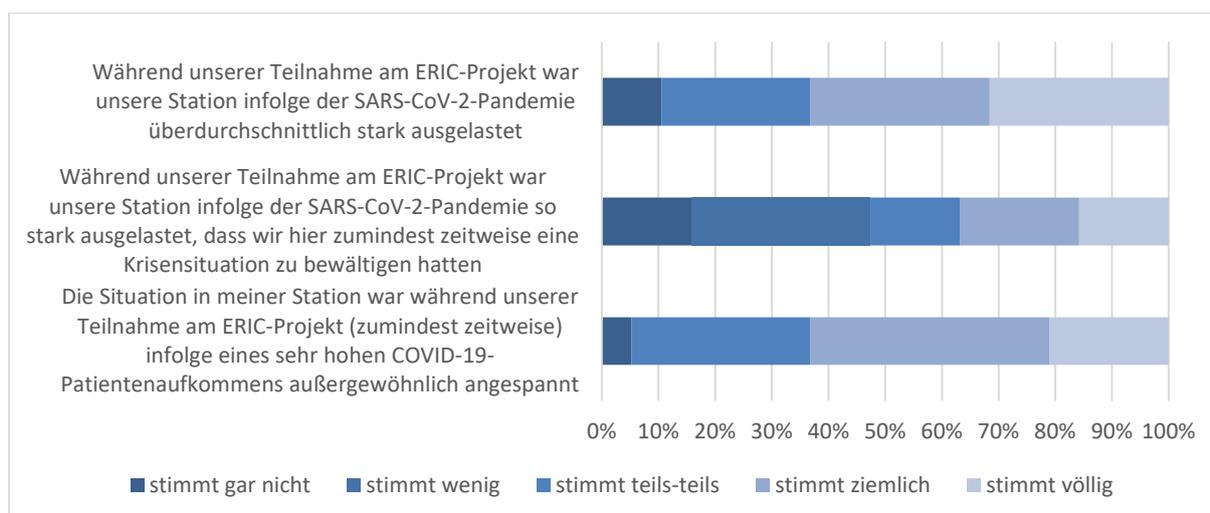


Abbildung 28. Ausmaß und Belastung infolge der Krisensituation (1).

Für etwa die Hälfte der Befragten war Arbeitsbelastung infolge eines sehr hohen COVID-19-Patientenaufkommens während der Teilnahme am ERIC-Projekt außergewöhnlich hoch.

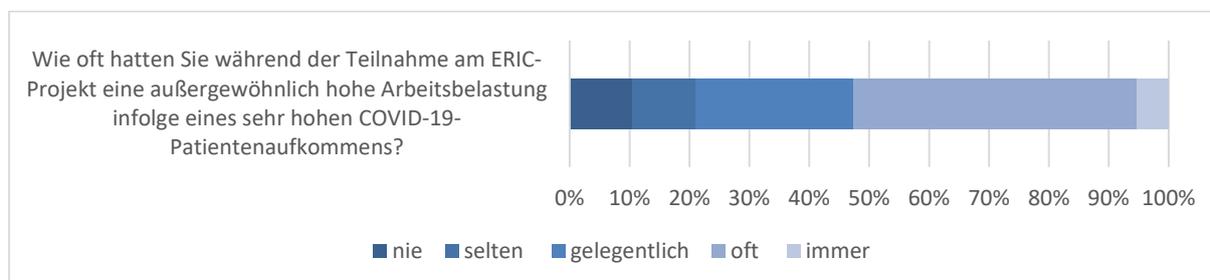


Abbildung 29. Ausmaß und Belastung infolge der Krisensituation (2).

Teil 4: Einstellungen gegenüber Visitenrobotern

Fast 50% der Befragten hatten anfänglich keine Bedenken, eine telemedizinische Beratung anhand von Visitenrobotern durchzuführen. Vielmehr stand der Großteil der telemedizinischen Beratung anhand von Visitenrobotern grundsätzlich positiv gegenüber. Über die Hälfte der Befragten stimmten der Aussage, dass sie im Verlauf der Nutzung der telemedizinischen Beratung anhand von Visitenroboter zunehmend skeptisch geworden sind, gar nicht zu. Fast 90% lehnen die Beratung anhand von Visitenrobotern nicht grundsätzlich ab und haben keine ethischen Bedenken bei deren Nutzung. Über 60% der teilnehmenden Ärzte haben keine datenschutzrechtlichen Bedenken bei der Nutzung von Visitenrobotern.

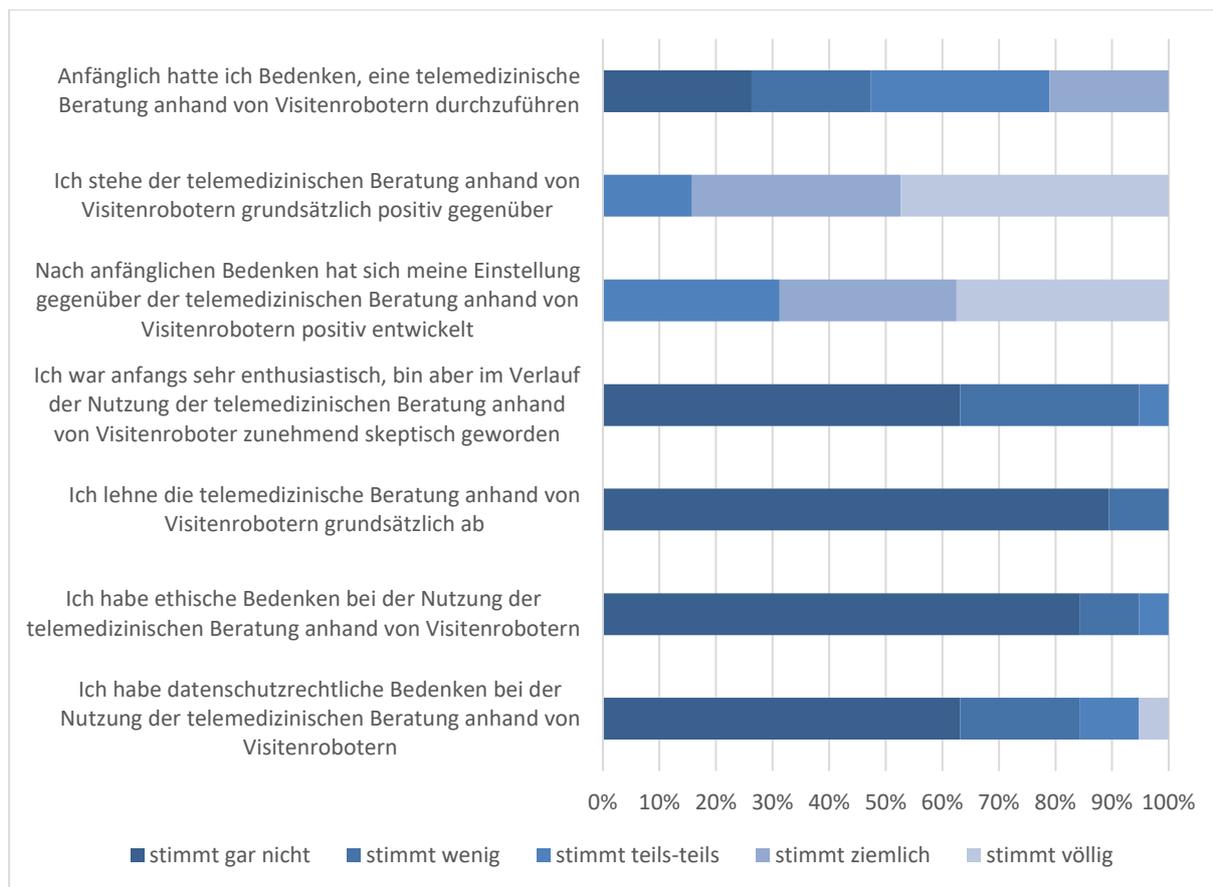


Abbildung 30. Einstellungen gegenüber Visitenrobotern.

Teil 5: Zufriedenheit und Optimierungsbedarf

Zufriedenheit:

Etwa zwei Drittel sind ziemlich oder sehr zufrieden mit der telemedizinischen Beratung anhand von Visitenrobotern.

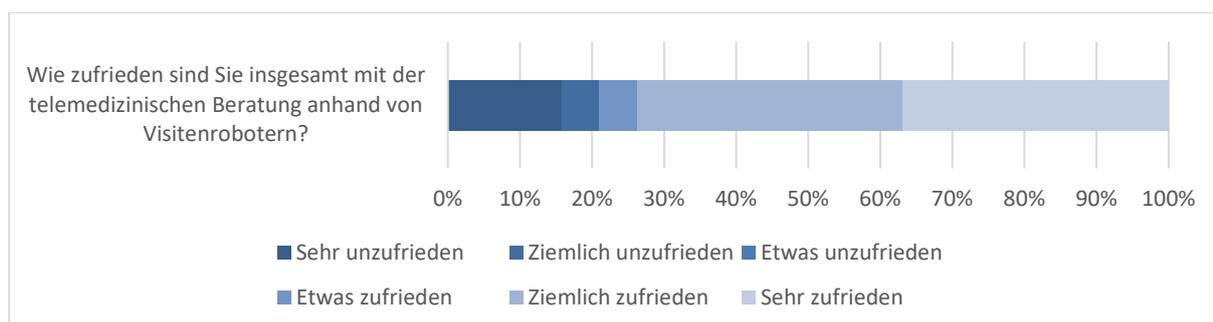


Abbildung 31. Zufriedenheit mit den Televisitenrobotern (1).

Etwas mehr als die Hälfte der Befragten gab an, dass die telemedizinische Beratung anhand von Visitenrobotern gut funktioniert und nicht verändert werden muss. Der Rest der Befragten gibt an, dass die telemedizinische Beratung anhand von Visitenrobotern im Großen und Ganzen gut funktioniert, aber in einigen Punkten verändert werden muss.

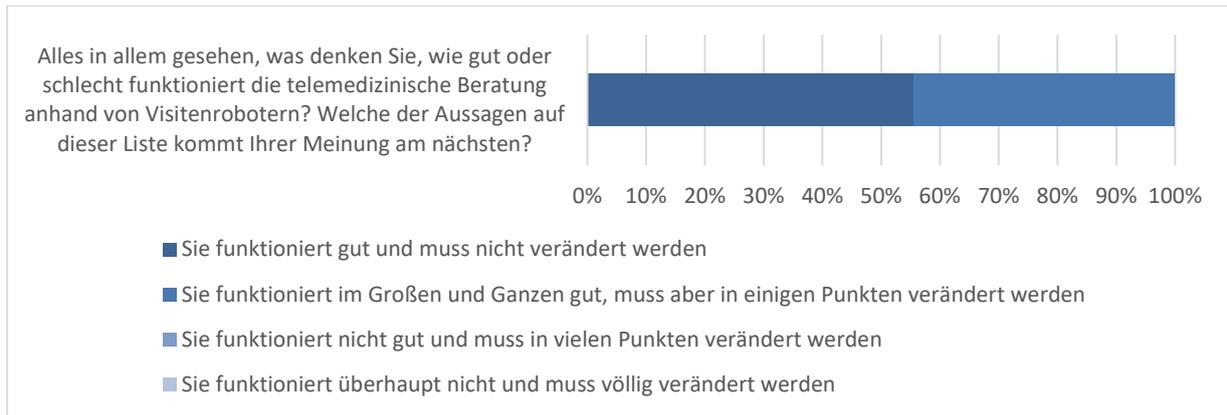


Abbildung 32. Zufriedenheit mit den Televisitenrobotern (2).

Viele der Befragten haben angegeben, keine Verbesserungsvorschläge zu haben. Die gemachten Optimierungsvorschläge betreffen oftmals die Koordinierung der Visitermine, wobei sowohl der Wunsch nach mehr Flexibilität als auch nach festerer Planung geäußert wurde. Einige wenige Verbesserungsvorschläge betrafen technische Eigenschaften des Televisitenroboters, darunter etwa der Wunsch nach direkter Übertragung von Röntgen-/CT-Bildern oder Laborergebnissen. Darüber hinaus wurde vorgeschlagen, das Konzept auf IMC-Stationen auszuweiten, den Visitenroboter für strukturierte Übergaben zwischen den Schichten oder zur Weiterbildung zu nutzen oder auch direkt mit dem PDMS sowie für Konsultationen mit anderen Disziplinen wie der Mikrobiologie zu nutzen.

8 Schlussfolgerungen und Empfehlungen des Evaluators

In der Gesamtschau der zuvor dargestellten Ergebnisse und Limitationen lässt sich die Evaluation des Innovationsfondsprojektes ERIC aus unterschiedlicher Perspektive zusammenfassen.

8.1 Perspektive Wirksamkeitsevaluation

1. *Design.* Das im Antrag gewählte Stepped Wedge Cluster-randomisierte kontrollierte Studiendesign konnte mit Abweichungen und zeitlichen Verschiebungen inklusive Verlängerung der Rekrutierungsperiode von 12 auf 19 Monate umgesetzt werden (vgl. Abbildung 6 [schematische Illustration des sequentiellen Rollout-Prozesses und zeitlicher Verlauf der Implementierung der Intervention] sowie Abbildung 8 [Deskriptive Darstellung der patientenindividuellen QI-Intervalle (ICU-Aufenthalte) über die gesamte Studiendauer]). Die initial berechnete Patienten-Fallzahl konnte nach Öffnung für alle GKV-Versicherten mit den randomisierten Clustern trotz heterogener Rekrutierungsrate (zwischen den einzelnen Clustern sowie über die Zeit) erreicht werden.
2. *Patientenpopulation.* Es ergab sich ein Ungleichgewicht zwischen der unter Kontrollbedingung (Regelversorgung) und der unter Interventionsbedingung (Tele-ICU) versorgten Anzahl an Patienten (414 (28.3%) vs. 1048 (71.7%)). Unterschiede im zur ersten QI-Visite gemessenen SOFA-Score sowie SAPS II geben Hinweise auf eine höhere Krankheitsschwere bei unter Intervention eingeschlossenen Patienten, was eine erhöhte Mortalität und Rate an funktionellen Langzeitbeeinträchtigungen im Vergleich zur Gruppe der unter Routinebedingungen behandelten Patienten erwarten lässt.
3. *Implementierung der Intervention.* Nach einer Transition-Periode (3-monatige Trainingsperiode inkl. Implementierungsphase von wenigen Tagen) konnte die Intervention mit Einsatz des Televisiten-Wagens in 10 Clustern begonnen werden. Abbildung 8 zeigt für jedes Cluster die tatsächliche Abfolge der Phasen.
4. *QI Outcomes.* Während es in der internationalen Literatur keinen Konsens darüber gibt, wie Prozessqualität in der Intensivmedizin zu messen ist, wurden von der DIVI für Deutschland evidenzbasierte multiprofessionelle QIs für die akut-intensivmedizinische Behandlung etabliert. Als primäres Outcome Measure wurde die tägliche patientenzentrierte Erfolgsrate eines einzelnen QIs der DIVI (Version 3, Auflage 2017, siehe [5]) präspezifiziert. Die Beurteilung der QI-Adhärenz erfolgte durch ein unabhängiges Expertengremium (7 Ärzte/Oberärzte der Charité mit Zusatzbezeichnung Intensivmedizin als unabhängige Rater) zentral an der Charité. Der Bewertungsprozess dieser evidenzbasierten und multiprofessionellen QIs für die akut-intensivmedizinische Behandlung kann als objektiv und reliabel betrachtet werden.
5. *Wirksamkeit.* In 7 der 8 patientenzentrierten QIs zeigte die telemedizinische Intervention in den konfirmatorischen Analysen einen statistisch signifikanten, deutlich positiven Effekt auf den QI-Erfüllungsgrad, der bei QI 2 und QI 8 als klinisch relevant beurteilt werden kann. Die anderen QIs zeigten trotz vorliegendem hohen Erfüllungsgrad unter Kontrolle auch signifikante Verbesserungen unter Intervention (QI 3, QI 4, QI 6, QI 7, QI 9). Dieser Interventionseffekt erwies sich in Sensitivitätsanalysen als robust.
6. *Heterogenität der Zentren bei der QI-Adhärenz:* Es zeigten sich unterschiedliche Levels in den Erfüllungsgraden zwischen den Zentren. Diese Heterogenität im Qualitätsniveau blieb auch nach Verbesserung der QI-Adhärenz durch die Intervention erhalten.

7. *Zeitliche Trends.* Gleichzeitig zeigten sich temporale Trends sowohl in der Kontrollphase als auch nach Exposition, die nahelegen, dass nach 12 Monaten, z.B. durch Weggang des geschulten ärztlichen und Pflegepersonals oder anderer Mechanismen auf institutioneller Ebene, erneute (E-learning-basierte) QI-Schulungen angebracht erscheinen (vgl. Abbildung 18 zu den QI-Hauptergebnissen).
8. *Generalisierbarkeit der Effekte auf QIs:* Die ERIC-Studie betrachtet verschiedene Cluster: Zentren der Sekundär- und Tertiärversorgung. Die teilnehmenden Zentren liegen sowohl in einem ländlichen Raum wie auch in einer Metropolregion. Über das gesamte Spektrum von Clustern zeigte sich der Effekt der Telemedizin-basierten Intervention auf die QIs. Dies plausibilisiert die Verallgemeinerbarkeit des Interventionseffektes auf weitere Bereiche der Versorgung. Hinzu kommt, dass trotz Abweichungen vom ursprünglich geplanten, ausgeglichenen (balancierten) Design robuste Effekte beobachtet werden konnten.
9. *Mortalität.* Es gab keine Evidenz dafür, dass sich die 6-Monats-Gesamtmortalität (ICU sowie 6-monatiges post-ICU Follow-up) in der Gesamtpopulation zwischen Regelversorgung und Tele-ICU Bedingung unterscheidet (Abbildung 19). Sie lag, adjustiert für Baseline SAPS II und SOFA-Score als Surrogat für die Krankheitsschwere zu Beginn des Index ICU-Aufenthalts, unter beiden Behandlungsbedingungen bei ca. 25% zu Monat 6.
10. *Post-ICU Follow-up.* In der Gruppe der ICU-Überlebenden und im Rahmen der intersektoralen FU-Assessments untersuchten Patienten konnte in explorativen Analysen keine Evidenz für Unterschiede in PICS-bezogenen sekundären Outcomes zwischen ICU-Überlebenden nach Behandlung unter Kontroll- vs. Interventionsbedingung gezeigt werden. Dies kann als Indiz dafür gewertet werden, dass die in ERIC implementierte Telemedizin-basierte Intervention neben einer Verbesserung der ICU-Behandlungsqualität auch einen günstigen Effekt auf funktionelle Beeinträchtigungen und Mortalität über die Entlassung aus der Intensivstation hinaus hat.

Zusammenfassend stellt der klinische Evaluator die Wirksamkeit der telemedizinischen Intervention hinsichtlich der QI-Adhärenz auf Patientenebene fest. Dies gilt für 8 der DIVI-Qualitätsindikatoren, QI 2 bis QI 9 (Management von Sedierung, Analgesie und Delir; Patientenadaptierte Beatmung; Weaning; Maßnahmen zum Infektionsmanagement; Frühe enterale Ernährung; Dokumentation einer strukturierten Patienten- und Angehörigenkommunikation; Frühmobilisation). Dieses Ergebnis hat ein hohes Potential, auch auf Institutionen außerhalb des Studienkollektivs verallgemeinert zu werden.

Diskussion möglicher Limitationen:

Es ist mit einem Selektionseffekt auf Clusterebene zu rechnen. Die teilnehmenden Zentren der Region Berlin und Brandenburg waren in vielen QI-Domänen bereits auf einem hohen, evidenzbasierten Niveau. Hauptsächlich nahmen ICUs teil, die schon engagiert in der Umsetzung von evidenzbasierten Prinzipien im Rahmen der akut-intensivmedizinischen Behandlung waren. Cluster starteten mit hoher QI-Adhärenz bereits vor Übergang in die neue Versorgungsform, was Effekte auf sekundäre Outcomes abmildert.

Darüber hinaus existieren mögliche Selektionseffekt auf Patientenebene. Bereits zu Anfang der Televisite gab es den ethischen Druck, auch andere, besonders kritisch kranke Patienten telemedizinisch vorzustellen, welche besonders von der telemedizinischen Intervention profitieren könnten. In der Evaluation wurde dies erst ab dem Zeitpunkt der Öffnung des Einschlusses für alle GKV Patienten (§630 BGB) ermöglicht. Hierdurch wurden im Vergleich zur Kontrollphase

Patienten mit höherem Krankheitsschweregrad vorgestellt, worauf ein höherer SAPS II und SOFA-Score, gemessen zur ersten QI-Visite, hindeutet.

8.2 Perspektive Gesundheitsökonomische Evaluation [TU Berlin]

Im Rahmen der gesundheitsökonomischen Evaluation wurden auf der Effektseite die sekundären Endpunkte 6-Monats-Gesamsterblichkeit, die gesundheitsbezogene Lebensqualität bzw. qualitätsadjustierte Lebensjahre und das Vorliegen einer außerklinischen Beatmung untersucht. Auf der Kostenseite wurden direkte medizinische Kosten über den gesamten Beobachtungszeitraum vom ersten Aufenthalt auf der Intensivstation bis zur zweiten Nacherhebung nach sechs Monaten aus der Sozialversicherungsperspektive betrachtet.

Im Hinblick auf die Kosten zeigten sich deskriptiv einige Unterschiede zwischen Kontroll- und Interventionsgruppe. So entstanden für den Aufenthalt auf der Intensivstation im Mittel geringere Kosten bei Patienten, die in die Regelversorgung (d.h. Kontrollbedingung) eingeschlossen worden sind (12.319,63), als bei denen in der Intervention (14.231,93 Euro). Dabei sind die Interventionskosten bereits berücksichtigt. Im Hinblick auf die Folgekosten nach der Entlassung zeigen sich Unterschiede, wobei Patienten in der Kontrollgruppe bis zum ersten Erhebungszeitpunkt deutlich mehr Kosten verursacht haben als Patienten in der Interventionsgruppe. Die größten Kostenunterschiede waren dabei im Bereich der stationären Wiederaufnahmen zu verzeichnen. Andererseits entstanden bei Patienten in der Interventionsgruppe höhere Kosten im Bereich von Rehabilitationsmaßnahmen. Zwischen dem ersten und zweiten Follow-up-Erhebungszeitpunkt nähern sich die Kosten in Kontrolle und Intervention einander an, wobei Patienten in der Interventionsgruppe weiterhin weniger Kosten verursachen als jene, die unter Kontrollbedingungen eingeschlossen worden sind. Die Unterschiede lassen sich auch hier zum größten Teil durch stationäre Wiederaufnahmen erklären.

Die inferenzstatistischen Analysen hingegen zeigten, dass der Unterschied zwischen den beiden Patientengruppen statistisch nicht signifikant war. Im Hinblick auf die Kosten lassen die Ergebnisse darauf schließen, dass die Intervention und damit eine höhere Adhärenz zu fast allen untersuchten Qualitätsindikatoren nicht mit Mehrkosten einhergeht. Dabei können die Interventionskosten in Höhe von 651,79 Euro pro Patient für die gesetzliche Krankenversicherung reduziert werden, wenn die Implementierungskosten im Rahmen des Krankenhauszukunftsfonds finanziert würden (siehe Ergebnisbericht).

In der Gesamtschau der deskriptiven Analysen zeigten sich auf der Effektseite kaum Unterschiede im Hinblick auf die Sterblichkeit, wobei die Patienten in der Kontrolle eine höhere Anzahl qualitätsadjustierter Lebensjahre im Beobachtungszeitraum aufwiesen als Patienten in der Intervention. Für die außerklinische Beatmung zeigten sich leichte Unterschiede zum zweiten Follow-up-Erhebungszeitpunkt, wobei eine Subgruppenanalyse von beatmet entlassenen Patienten ergab, dass in der Intervention ein mehr als doppelt so großer Anteil der Patienten erfolgreich von der Beatmung entwöhnt werden konnten. Die inferenzstatistischen Analysen zeigten auch hier keine signifikanten Unterschiede.

Möglicherweise hätte bei vergleichbarer Krankheitsschwere und ohne die in 8.1 beschriebenen Selektionseffekte ein deutlicherer Einfluss der Intervention auf die im Rahmen der gesundheitsökonomischen Evaluation betrachteten sekundären Endpunkte gezeigt werden können.

Darüber hinaus ist eine direkte Übersetzung von verbesserter Struktur- und Prozessqualität (hier: die QI) zu verbesserter Ergebnisqualität in der Praxis häufig schwierig – das betrifft vor allen Dingen klinische Endpunkte.[62] Es ist durchaus möglich, dass sich bei anderen als den gewählten Effektparametern (qualitätsadjustierte Lebensjahre, 6-Monats-Mortalität und das Vorliegen einer außerklinischen Beatmung) ein Unterschied zwischen den Interventionsarmen gezeigt hätte. So weist etwa die Untersuchung von Mahdavi *et al.* (2018) darauf hin, dass sich Struktur- und Prozessindikatoren eher auf der Ebene von Patientenerfahrungen auswirken.[62] Weiterhin sind die teilnehmenden Intensivstationen zumeist auf einem hohen Qualitätsniveau in die neue Versorgungsform eingetreten, was einen Effekt der Intervention auf die hier untersuchten patientenrelevanten Endpunkte vermutlich auch erschwert hat.

Bei einer Übertragung in die Regelversorgung wären jedoch auch Intensivstationen mit initial weniger hoher QI-Adhärenz inkludiert, wo ein größerer inkrementeller Effekt erwartet werden kann. Dies wäre auch und insbesondere im Hinblick auf die außerklinische Beatmung relevant – sowohl für die Betroffenen selbst und ihre Lebensqualität als auch für die gesetzliche Krankenversicherung, die mit jedem von der außerklinischen Beatmung entwöhnten Patienten monatliche Ausgaben bis zu schätzungsweise 20.000 Euro vermeiden könnte.

Vor diesem Hintergrund und der Tatsache, dass die Intervention keine Mehrkosten bei verbesserter Prozessqualität mit sich bringt, erfolgt aus gesundheitsökonomischer Perspektive eine Empfehlung für die Übertragung der Intervention in die Regelversorgung. Diese wird durch die Angaben in der Akzeptanzerhebung etwa in Bezug auf eine Förderung evidenzbasierter Entscheidungsfindung nur noch unterstrichen.

8.3 Perspektive der Akzeptanzerhebung [TU Berlin]

Insgesamt weist die durchgeführte Erhebung zur Akzeptanz von Televisiten mit Visitenrobotern im Rahmen des Projekts ERIC darauf hin, dass eine recht hohe Nutzerfreundlichkeit sowie Zufriedenheit bei den Ärzten auf den beteiligten Intensivstationen vorlagen. Besonders bemerkenswert ist, dass ein Großteil der Befragten die telemedizinische Beratung als hilfreich im Hinblick auf evidenzbasierte Entscheidungsfindung empfunden haben und dass über ein Drittel davon ausgeht, dass die telemedizinische Beratung das Potential hat, auch in anderen Krankenhäusern zu einer evidenzbasierten Behandlung beitragen zu können. Wenngleich sich auch infolge der Arbeitsbelastung durch die SARS-CoV-2-Pandemie nur wenige an den Televisiten beteiligte Ärzte an der Umfrage beteiligt haben, so stützt deren subjektive Einschätzung doch die Ergebnisse der Evaluation der primären Endpunkte, die für viele QIs eine Steigerung der Adhärenz zeigen konnte. Dies und die Vorschläge für eine Ausweitung des Konzepts auf andere Stationen bzw. Bereiche sprechen dafür, dass es sich bei ERIC um eine erfolgreiche Intervention handelt.

9 Anhang

- A.1 REDCap Datenbanken: QI-Daten; FU-Daten
- A.2 SAPS II und SOFA Score
- A.3 PICS-bezogene sekundäre Outcomes: Auswertungsalgorithmen
- A.4 Technische Details – Effektevaluation
- A.5 Reporting Konventionen
- B.1 Technische Details – gesundheitsökonomische Evaluation [TU Berlin]

A.1 REDCap Datenbanken: QI-Daten; FU-Daten

Es wurde eine geklonte Version der offiziellen, in REDCap implementierten ICU Datenbank für die Evaluation abgeleitet, mit den genannten Modifikationen zur Pseudonymisierung. Diese Datenbank wurde zum Export für die Analysen in REDCap Version 10.0.26 (<https://redcap.charite.de/eric/>) zur Verfügung gestellt. Hierfür wurde von FOKUS ein neues Patienten-Pseudonym `patient_id_1mu` und ein neues Fall-Pseudonym generiert, bestehend aus 32 Nummern und Buchstaben. Der remote Zugriff durch das IBE erfolgte via des VPN-Clients OpenVPN GUI v11.15.0.0 (<https://openvpn.net/>). Zudem wurde für den Evaluator ein Ausschnitt der FU-Datenbank für ausgewählte sekundäre Endpunkte und weitere analyserelevante Variablen über das REDCap File Repository als csv-Datei in nicht-personenbeziehbarer Form zur Verfügung gestellt.

A.2 SAPS II und SOFA Score

Der SAPS II (Simplified Acute Physiology Score), mit Minimum=0 und Maximum=163, misst die Erkrankungsschwere eines Patienten und dient zur Prognose der Krankenhaus-Mortalität.[63]

Der SOFA (Sequential Organ Failure Assessment) Score, mit Minimum=0 und Maximum=24 (Tod), dient zur Beurteilung des Grades der Organdysfunktion und zur Prädiktion der ICU Mortalität, basierend auf Laborergebnissen und klinischen Daten (6 täglich erhobene Parameter (ZNS, respiratorisches System, Blutgerinnung, Leber, Niere und kardiovaskuläres System). Die Letalitätsrate steigt mit der Anzahl dysfunktioneller Organsysteme.[64, 65]

Für das gesamte ITT-Patientenkollektiv wurden zur ersten QI-Visite folgende Score-Werte beobachtet (deskriptive Zusammenfassung):

SOFA:

| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|-------|---------|--------|-------|---------|--------|
| 0.000 | 3.000 | 5.000 | 5.627 | 8.000 | 20.000 |

SAPS II:

| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | Missings |
|------|---------|--------|-------|---------|--------|----------|
| 0.00 | 20.00 | 33.00 | 34.58 | 47.00 | 107.00 | 12 |

A.3 PICS-bezogene sekundäre Outcomes: Auswertungsalgorithmen

Im Folgenden werden für ausgewählte Messinstrumente ergänzende Informationen zur Operationalisierung bzw. Ableitung der sekundären Outcomes für die explorativen Analysen zur Post-ICU Follow-up Phase dargestellt.

A.3.1. Subdomäne Mobilität: Timed-Up-and-Go (TUG)-Test

Der Patient sitzt auf einem Stuhl mit Armlehne und wird aufgefordert aufzustehen, drei Meter zu gehen, umzukehren und sich wieder zu setzen. Hilfsmittel wie Gehhilfen sind dabei erlaubt, Hilfe von anderen Personen nicht. Währenddessen wird die Zeit (in Sekunden) gemessen.

binäre Bewertung des TUG-Testergebnisses

Die Auswertung des Performance-Tests (eingeschränkt ja/nein) erfolgte in Abhängigkeit der benötigten Zeit, bei Verwendung von altersspezifischen Cut-offs.

| Alter | | | Einschränkung |
|--------------|--------------|--------------|--------------------------------|
| ≤60 | 61 - 80 | ≥81 | |
| ≤10 Sekunden | ≤15 Sekunden | ≤20 Sekunden | Keine eingeschränkte Mobilität |
| >10 Sekunden | >15 Sekunden | >20 Sekunden | Eingeschränkte Mobilität |

Cut-offs basierend auf Untersuchungen zu altersstratifizierten Normwerten, siehe z.B. Kenny RA *et al. J Am Geriatr Soc* 2013; 61 Suppl 2:S279-90. doi: 10.1111/jgs.12195; Ibrahim A *et al. J Multidiscip Healthc* 2017;10:409-416. doi:10.2147/JMDH.S142520; or Kear BM *et al.* [66])

A.3.2. Subdomäne Handkraft: Hand-Grip-Strength (HGS)-Test

Für das Gesamt-Assessment wurde jeweils der beste von 3 Versuchen gewertet. Das binäre sekundäre Outcome (eingeschränkt ja/nein) wurde anhand alters- und geschlechtsspezifischer Cut-offs abgeleitet.

binäre Bewertung des HGS-Testergebnisses

| Geschlecht | Alter | | | Einschränkung |
|------------|---------|---------|---------|----------------------------|
| | ≤60 | 61 – 80 | ≥81 | |
| Männer | ≥25 kg | ≥23 kg | ≥19 kg | Keine eingeschränkte Kraft |
| | < 25 kg | < 23 kg | < 19 kg | Eingeschränkte Kraft |
| Frauen | ≥15 kg | ≥14 kg | ≥13 kg | Keine eingeschränkte Kraft |
| | < 15 kg | < 14 kg | < 13 kg | Eingeschränkte Kraft |

Cut-offs basierend auf Untersuchungen zu altersstratifizierten Normwerten (Kenny RA *et al. J Am Geriatr Soc* 2013; 61 Suppl 2:S279-90. doi: 10.1111/jgs.12195.)

A.4 Ergänzende Tabellen & Abbildungen zur ICU-Phase

A.4.1. Ergänzende Abbildung zur Rekrutierung

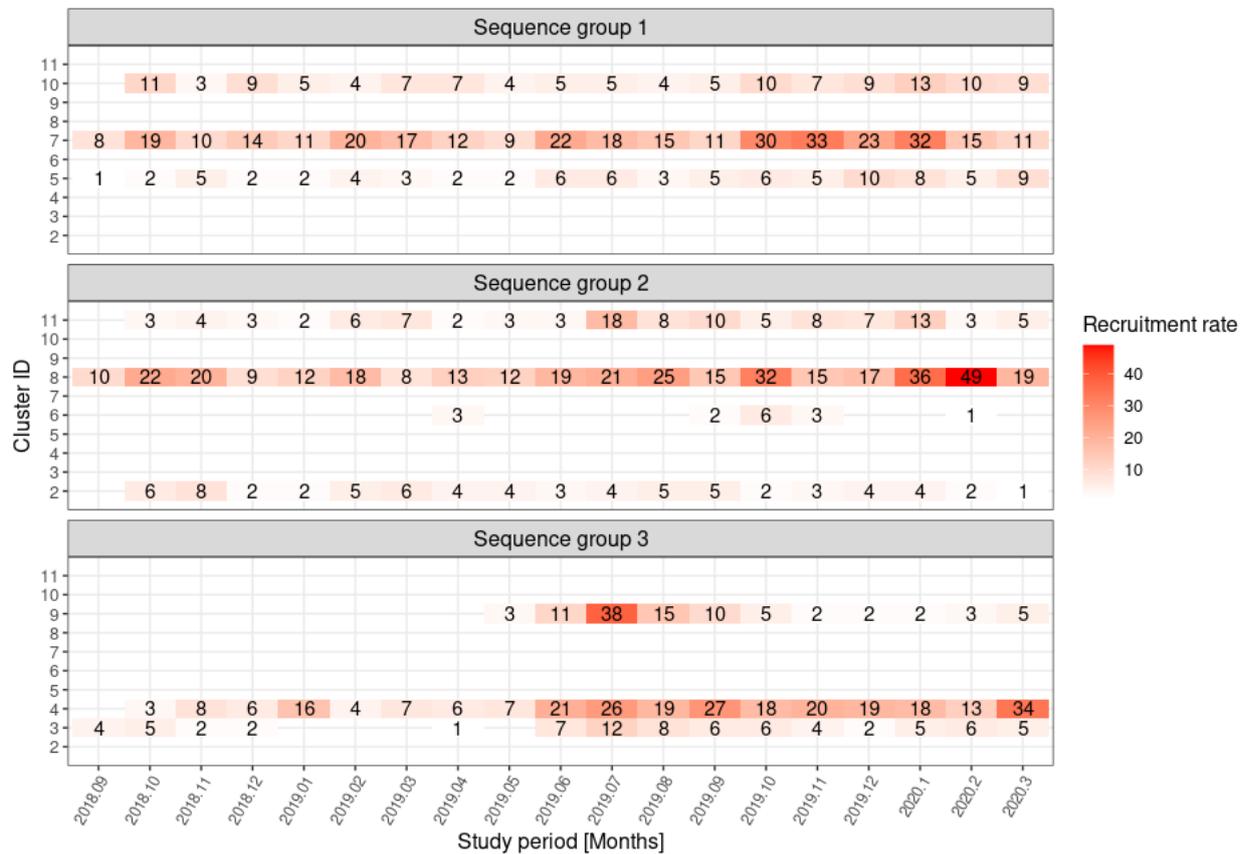


Abbildung 33. Heatmap für die Cluster-spezifische Rekrutierung pro Kalendermonat (Fall-Ebene), stratifiziert nach den drei Sequenzgruppen (19 Monate Rekrutierungsphase). Je dunkler die Farbe, desto mehr Patienten wurden in diesem Kalendermonat in die Studie eingeschlossen. Fehlende Angaben zur Anzahl der Patienten in einem bestimmten Monat bedeuten, dass in diesem Zeitraum keine Patienten rekrutiert wurden.

A.4.2. QI-Auswertungen: Schätzer der logistischen Mixed Effects-Modelle (fixed effects und Varianzkomponenten) – Hauptmodell PM

Die Definitionen der Variablen-Labels für die nachfolgenden Tabellen sind näher beschrieben in Abschnitt 3.8.2.1 im Methodenteil.

| Parameter | QI 1 erfüllt | | |
|--|---------------|----------------|------------------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 42.23 | 15.87 – 112.32 | <0.001 |
| trt (<i>reference: control</i>) | 1.61 | 0.96 – 2.70 | 0.073 |
| time | 1.39 | 1.27 – 1.53 | <0.001 |
| trt * time | 0.79 | 0.72 – 0.87 | <0.001 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T00_pid (<i>patient-specific intercepts</i>) | 2.62 | | |
| T00_cluster (<i>cluster-specific intercepts</i>) | 2.02 | | |
| ICC | 0.59 | | |
| Marginal R ² / Conditional R ² | 0.128 / 0.638 | | |

trt: Interventionseffekt; time: Effekt der Cluster-spezifischen Expositionszeit; trt x time: Interaktionseffekt zwischen Zeit und Intervention; σ^2 : residual variance (For logistic regression the residual variance is fixed to $\pi^2/3$); CI: Confidence interval; ICC: intra-class correlation coefficient; R² Werte sind marginale und konditionale R² Statistiken, basierend auf Nakagawa *et al.*[67]

| Parameter | QI 2 erfüllt | | |
|--|---------------|-------------|------------------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 0.09 | 0.04 – 0.23 | <0.001 |
| trt (<i>reference: control</i>) | 5.33 | 3.86 – 7.36 | <0.001 |
| time | 0.99 | 0.93 – 1.06 | 0.793 |
| trt * time | 0.98 | 0.92 – 1.05 | 0.607 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T00_pid (<i>patient-specific intercepts</i>) | 1.48 | | |
| T00_cluster (<i>cluster-specific intercepts</i>) | 1.89 | | |
| ICC | 0.51 | | |
| Marginal R ² / Conditional R ² | 0.057 / 0.534 | | |

trt: Interventionseffekt; time: Effekt der Cluster-spezifischen Expositionszeit; trt x time: Interaktionseffekt zwischen Zeit und Intervention; σ^2 : residual variance (bei logistischer Regression ist die residual variance definiert als $\pi^2/3$); CI: Confidence interval; ICC: intra-class correlation coefficient; R² Werte sind marginale und konditionale R² Statistiken, basierend auf Nakagawa *et al.*[67]

| Parameter | QI 3 erfüllt | | |
|--|---------------|---------------|------------------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 19.94 | 12.39 – 32.11 | <0.001 |
| trt (reference: control) | 2.25 | 1.43 – 3.53 | <0.001 |
| time | 1.05 | 0.96 – 1.15 | 0.264 |
| trt * time | 0.94 | 0.86 – 1.03 | 0.158 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T00 _{pid} (patient-specific intercepts) | 2.32 | | |
| T00 _{cluster} (cluster-specific intercepts) | 0.26 | | |
| ICC | 0.44 | | |
| Marginal R ² / Conditional R ² | 0.021 / 0.451 | | |

trt: Interventionseffekt; time: Effekt der Cluster-spezifischen Expositionszeit; trt x time: Interaktionseffekt zwischen Zeit und Intervention; σ^2 : residual variance (bei logistischer Regression ist die residual variance definiert als $\pi^2/3$); CI: Confidence interval; ICC: intra-class correlation coefficient; R² Werte sind marginale und konditionale R² Statistiken, basierend auf Nakagawa *et al.*[67]

| Parameter | QI 4 erfüllt | | |
|--|--------------|----------------|------------------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 170.79 | 78.40 – 372.07 | <0.001 |
| trt (reference: control) | 9.05 | 3.81 – 21.49 | <0.001 |
| time | 0.83 | 0.72 – 0.96 | 0.012 |
| trt * time | 1.09 | 0.93 – 1.28 | 0.264 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T00 _{pid} (patient-specific intercepts) | 10.10 | | |
| T00 _{cluster} (cluster-specific intercepts) | 0.00 | | |
| Marginal R ² / Conditional R ² | 0.107 / NA | | |

trt: Interventionseffekt; time: Effekt der Cluster-spezifischen Expositionszeit; trt x time: Interaktionseffekt zwischen Zeit und Intervention; σ^2 : residual variance (bei logistischer Regression ist die residual variance definiert als $\pi^2/3$); CI: Confidence interval; ICC: intra-class correlation coefficient; R² Werte sind marginale und konditionale R² Statistiken, basierend auf Nakagawa *et al.*[67]

| Parameter | QI 6 erfüllt | | |
|--|--------------|----------------|------------------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 165.33 | 73.99 – 369.45 | <0.001 |
| trt (reference: control) | 4.40 | 2.02 – 9.58 | <0.001 |
| time | 1.12 | 0.99 – 1.27 | 0.070 |
| trt * time | 0.86 | 0.74 – 0.99 | 0.033 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T00 _{pid} (patient-specific intercepts) | 10.87 | | |
| T00 _{cluster} (cluster-specific intercepts) | 0.00 | | |
| Marginal R ² / Conditional R ² | 0.118 / NA | | |

trt: Interventionseffekt; time: Effekt der Cluster-spezifischen Expositionszeit; trt x time: Interaktionseffekt zwischen Zeit und Intervention; σ^2 : residual variance (bei logistischer Regression ist die residual variance definiert als $\pi^2/3$); CI: Confidence interval; ICC: intra-class correlation coefficient; R² Werte sind marginale und konditionale R² Statistiken, basierend auf Nakagawa *et al.*[67]

| Parameter | QI 7 erfüllt | | |
|--|---------------|--------------|--------------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 8.36 | 6.17 – 11.33 | <0.001 |
| trt (<i>reference: control</i>) | 1.58 | 1.16 – 2.14 | 0.003 |
| time | 1.00 | 0.94 – 1.05 | 0.863 |
| trt * time | 0.98 | 0.92 – 1.04 | 0.512 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T ₀₀ pid (<i>patient-specific intercepts</i>) | 1.13 | | |
| T ₀₀ cluster (<i>cluster-specific intercepts</i>) | 0.10 | | |
| ICC | 0.27 | | |
| Marginal R ² / Conditional R ² | 0.004 / 0.275 | | |

trt: Interventionseffekt; time: Effekt der Cluster-spezifischen Expositionszeit; trt x time: Interaktionseffekt zwischen Zeit und Intervention; σ^2 : residual variance (bei logistischer Regression ist die residual variance definiert als $\pi^2/3$); CI: Confidence interval; ICC: intra-class correlation coefficient; R² Werte sind marginale und konditionale R² Statistiken, basierend auf Nakagawa *et al.*[67]

| Parameter | QI 8 erfüllt | | |
|--|---------------|--------------|------------------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 5.70 | 2.77 – 11.71 | <0.001 |
| trt (<i>reference: control</i>) | 6.79 | 4.63 – 9.96 | <0.001 |
| time | 0.82 | 0.76 – 0.88 | <0.001 |
| trt * time | 0.93 | 0.86 – 1.00 | 0.065 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T ₀₀ pid (<i>patient-specific intercepts</i>) | 2.53 | | |
| T ₀₀ cluster (<i>cluster-specific intercepts</i>) | 1.13 | | |
| ICC | 0.53 | | |
| Marginal R ² / Conditional R ² | 0.159 / 0.602 | | |

trt: Interventionseffekt; time: Effekt der Cluster-spezifischen Expositionszeit; trt x time: Interaktionseffekt zwischen Zeit und Intervention; σ^2 : residual variance (bei logistischer Regression ist die residual variance definiert als $\pi^2/3$); CI: Confidence interval; ICC: intra-class correlation coefficient; R² Werte sind marginale und konditionale R² Statistiken, basierend auf Nakagawa *et al.*[67]

| Parameter | QI 9 erfüllt | | |
|--|---------------|--------------|------------------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 10.98 | 4.49 – 26.87 | <0.001 |
| trt (<i>reference: control</i>) | 3.16 | 2.41 – 4.15 | <0.001 |
| time | 0.83 | 0.79 – 0.88 | <0.001 |
| trt * time | 1.11 | 1.04 – 1.18 | 0.001 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T ₀₀ pid (<i>patient-specific intercepts</i>) | 0.55 | | |
| T ₀₀ cluster (<i>cluster-specific intercepts</i>) | 1.89 | | |
| ICC | 0.43 | | |
| Marginal R ² / Conditional R ² | 0.028 / 0.442 | | |

trt: Interventionseffekt; time: Effekt der Cluster-spezifischen Expositionszeit; trt x time: Interaktionseffekt zwischen Zeit und Intervention; σ^2 : residual variance (bei logistischer Regression ist die residual variance definiert als $\pi^2/3$); CI: Confidence interval; ICC: intra-class correlation coefficient; R² Werte sind marginale und konditionale R² Statistiken, basierend auf Nakagawa *et al.*[67]

A.4.3. QI-Auswertungen: Schätzer der logistischen Mixed Effects-Modelle (fixed effects und Varianzkomponenten) – Modell SM.5

Für den Qualitätsindikator **QI 4** und **QI 6** konnten im Hauptmodell (Random Intercept für Cluster sowie Patient) die Varianzkomponente $T_{00 \text{ cluster}}$ für die Cluster-spezifischen Intercepts nicht geschätzt werden (fehlende Variabilität auf Cluster-Ebene). Daher sind in den folgenden zwei Tabellen die Parameterschätzer des Modells SM.5 der Sensitivitätsanalysen dargestellt, welche für die zentrale Abbildung 18 für QI 4 und QI 6 verwendet wurden.

| Parameter | QI 4 erfüllt | | |
|--|---------------|--------------|------------------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 9.53 | 6.01 – 15.12 | <0.001 |
| trt (reference: control) | 11.67 | 8.19 – 16.64 | <0.001 |
| time | 1.01 | 0.96 – 1.05 | 0.725 |
| trt * time | 0.90 | 0.85 – 0.95 | <0.001 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| $T_{00 \text{ cluster}}$ | 0.47 | | |
| ICC | 0.12 | | |
| Marginal R ² / Conditional R ² | 0.149 / 0.254 | | |

trt: Interventionseffekt; time: Effekt der Cluster-spezifischen Expositionszeit; trt x time: Interaktionseffekt zwischen Zeit und Intervention; σ^2 : residual variance (bei logistischer Regression ist die residual variance definiert als $\pi^2/3$); CI: Confidence interval; ICC: intra-class correlation coefficient; R² Werte sind marginale und konditionale R² Statistiken, basierend auf Nakagawa *et al.*[67]

| Parameter | QI 6 erfüllt | | |
|--|---------------|--------------|------------------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 12.77 | 8.58 – 19.00 | <0.001 |
| trt (reference: control) | 5.36 | 3.84 – 7.48 | <0.001 |
| time | 1.06 | 1.02 – 1.10 | 0.002 |
| trt * time | 0.90 | 0.86 – 0.95 | <0.001 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| $T_{00 \text{ cluster}}$ | 0.33 | | |
| ICC | 0.09 | | |
| Marginal R ² / Conditional R ² | 0.107 / 0.188 | | |

trt: Interventionseffekt; time: Effekt der Cluster-spezifischen Expositionszeit; trt x time: Interaktionseffekt zwischen Zeit und Intervention; σ^2 : residual variance (bei logistischer Regression ist die residual variance definiert als $\pi^2/3$); CI: Confidence interval; ICC: intra-class correlation coefficient; R² Werte sind marginale und konditionale R² Statistiken, basierend auf Nakagawa *et al.*[67]

A.4.4. Ergänzende Abbildung & Tabellen zum Post-ICU Follow-up

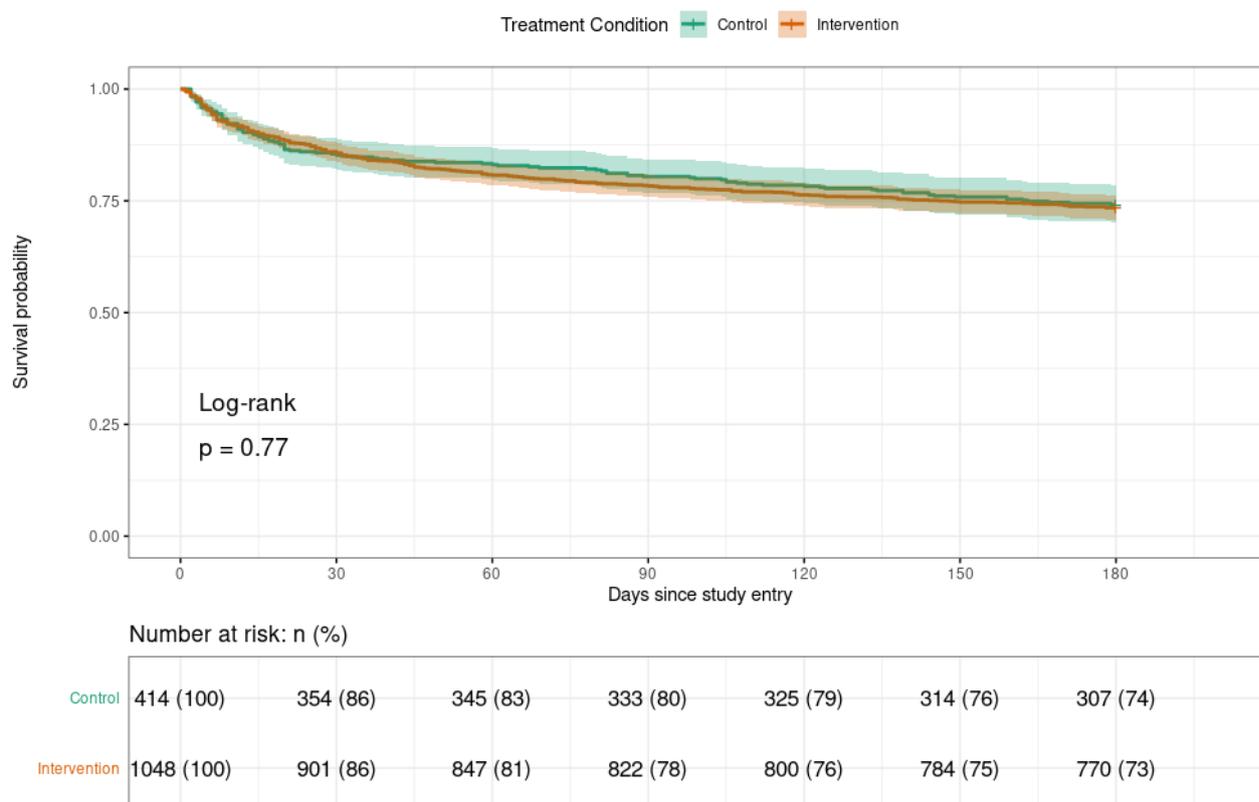


Abbildung 34. 6-Monats-Gesamtmortalität: Kaplan-Meier Kurven für die beiden Behandlungsbedingungen Telemedizin-basierte Intervention vs. Routineversorgung (Kontrolle).

Tabelle 26. EQ VAS: Ergebnisse des linearen Mixed Effects Modells, adjustiert für Behandlung (Intervention vs. Kontrolle), kategorisierten SAPS II, Zeit seit Entlassung aus der Index-ICU.

| Parameter | EQ VAS | | |
|--|---------------|----------------|--------|
| | Estimates | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 61.05 | 55.66 – 66.44 | <0.001 |
| trt | -0.22 | -6.36 – 5.92 | 0.943 |
| time | 0.02 | -0.01 – 0.04 | 0.211 |
| SAPS II.cat 21-33 | -0.29 | -4.11 – 3.53 | 0.882 |
| SAPS II.cat 34-48 | -3.26 | -7.36 – 0.84 | 0.119 |
| SAPS II.cat [>48] | -6.55 | -11.18 – -1.91 | 0.006 |
| trt * time | -0.01 | -0.04 – 0.02 | 0.465 |
| Random Effects | | | |
| σ^2 | 282.76 | | |
| T00 _{patid} (patient-specific intercepts) | 268.31 | | |
| ICC | 0.49 | | |
| N _{patid} | 808 | | |
| Observations | 1272 | | |
| Marginal R ² / Conditional R ² | 0.014 / 0.494 | | |

trt: binäre Faktorvariable für Treatment (Referenzkategorie: Kontrolle), time: Zeit seit Index-Entlassung (in Tagen), trt x time: Interaktionseffekt zwischen Zeit und Intervention, SAPS II.cat: kategorisierter SAPS II (Referenzkategorie ≤ 20 Punkte), σ^2 : residual variance, CI: Confidence interval, ICC: intra-class correlation coefficient.

Psychische Gesundheit – Subdomäne Depression

Tabelle 27. PHQ-2 Score: Ergebnisse des linearen Mixed Effects Modells.

| Parameter | PHQ-2 Score | | |
|--|---------------|--------------|--------|
| | Estimates | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 1.53 | 1.14 – 1.91 | <0.001 |
| trt (<i>reference: control</i>) | 0.14 | -0.27 – 0.55 | 0.504 |
| time | -0.00 | -0.00 – 0.00 | 0.831 |
| Baseline SAPS II | 0.01 | -0.00 – 0.01 | 0.080 |
| trt * time | 0.00 | -0.00 – 0.00 | 0.917 |
| Random Effects | | | |
| σ^2 | 1.49 | | |
| T00 patid (<i>patient-specific intercepts</i>) | 1.28 | | |
| T00 cluster (<i>cluster-specific intercepts</i>) | 0.01 | | |
| ICC | 0.46 | | |
| N _{cluster} | 10 | | |
| N _{patid} | 863 | | |
| Observations | 1460 | | |
| Marginal R ² / Conditional R ² | 0.006 / 0.468 | | |

trt: binäre Faktorvariable für Treatment (Referenzkategorie: Kontrolle), time: Zeit seit Index ICU-Entlassung;
 σ^2 : residual variance, CI: Confidence interval, ICC: intra-class correlation coefficient.

Psychische Gesundheit – Subdomäne Angst

Tabelle 28. GAD-2 Score: Ergebnisse des linearen Mixed Effects Modells.

| Parameter | GAD-2 Score | | |
|--|---------------|--------------|--------|
| | Estimates | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 1.47 | 1.15 – 1.79 | <0.001 |
| trt (<i>reference: control</i>) | 0.25 | -0.13 – 0.63 | 0.203 |
| time [days] | -0.00 | -0.00 – 0.00 | 0.432 |
| trt * time | -0.00 | -0.00 – 0.00 | 0.427 |
| Random Effects | | | |
| σ^2 | 1.26 | | |
| T00 patid (<i>patient-specific intercepts</i>) | 1.32 | | |
| ICC | 0.51 | | |
| N _{cluster} | 10 | | |
| N _{patid} | 869 | | |
| Observations | 1476 | | |
| Marginal R ² / Conditional R ² | 0.004 / 0.513 | | |

trt: binäre Faktorvariable für Treatment (Referenzkategorie: Kontrolle), time: Zeit seit Index ICU-Entlassung;
 σ^2 : residual variance, CI: Confidence interval, ICC: intra-class correlation coefficient.

Domäne Kognition: Mini-Cog

Tabelle 29. Mini-Cog – binäres Assessment: Ergebnisse des binären Mixed Effects Modells.

| Parameter | Mini-Cog Gesamtbewertung | | |
|--|--------------------------|-------------|-------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 0.23 | 0.09 – 0.58 | 0.002 |
| trt (<i>reference: control</i>) | 2.02 | 0.73 – 5.61 | 0.175 |
| time [months] | 0.99 | 0.88 – 1.13 | 0.934 |
| trt * time | 0.85 | 0.72 – 0.99 | 0.042 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T ₀₀ patid (<i>patient-specific intercepts</i>) | 5.61 | | |
| ICC | 0.63 | | |
| N _{patid} | 812 | | |
| Observations | 1345 | | |
| Marginal R ² / Conditional R ² | 0.013 / 0.635 | | |

trt: binäre Faktorvariable für Treatment (Referenzkategorie: Kontrolle), time: Zeit seit Index ICU-Entlassung;
 σ^2 : residual variance, CI: Confidence interval, ICC: intra-class correlation coefficient.

Subdomäne Mobilität: Timed-Up-and-Go (TUG)-Test

Tabelle 30. TUG – binäres Assessment: Ergebnisse des binären Mixed Effects Modells.

| Parameter | TUG Gesamtbewertung | | |
|--|---------------------|--------------|--------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 0.00 | 0.00 – 0.00 | <0.001 |
| trt (<i>reference: control</i>) | 2.15 | 0.17 – 27.26 | 0.554 |
| time [months] | 1.08 | 0.80 – 1.46 | 0.622 |
| trt * time | 0.94 | 0.66 – 1.35 | 0.756 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T ₀₀ patid (<i>patient-specific intercepts</i>) | 348.24 | | |
| ICC | 0.99 | | |
| N _{patid} | 647 | | |
| Observations | 1008 | | |
| Marginal R ² / Conditional R ² | 0.000 / 0.991 | | |

trt: binäre Faktorvariable für Treatment (Referenzkategorie: Kontrolle), time: Zeit seit Index ICU-Entlassung;
 σ^2 : residual variance, CI: Confidence interval, ICC: intra-class correlation coefficient.

Subdomäne Muskelstärke: Hand-Grip-Strength (HGS)-Test

Cell Contents

| | |
|-----------------|---|
| | N |
| N / Row Total | |
| N / Col Total | |
| N / Table Total | |

| Handkraft-Test – Bewertung | | | |
|----------------------------|----------|--------|-----------|
| Eingeschränkte Kraft | | | |
| Behandlung | 0 (nein) | 1 (ja) | Row Total |
| Kontrolle | 212 | 56 | 268 |
| | 0.791 | 0.209 | 0.233 |
| | 0.244 | 0.199 | |
| | 0.184 | 0.049 | |
| Intervention | 658 | 225 | 883 |
| | 0.745 | 0.255 | 0.767 |
| | 0.756 | 0.801 | |
| | 0.572 | 0.195 | |
| Column Total | 870 | 281 | 1151 |
| | 0.756 | 0.244 | |

Tabelle 31. HGS – binäres Assessment: Ergebnisse des binären Mixed Effects Modells.

| Parameter | HGS Gesamtbewertung | | |
|---|---------------------|--------------|--------|
| | Odds Ratio | 95% CI | p |
| Fixed Effects | | | |
| (Intercept) | 0.00 | 0.00 – 0.00 | <0.001 |
| trt (<i>reference: control</i>) | 1.35 | 0.10 – 18.75 | 0.825 |
| time [months] | 0.94 | 0.66 – 1.32 | 0.711 |
| trt * time | 0.99 | 0.67 – 1.46 | 0.959 |
| Random Effects | | | |
| σ^2 | 3.29 | | |
| T_{00} patid (<i>patient-specific intercepts</i>) | 228.23 | | |
| ICC | 0.99 | | |
| N_{patid} | 758 | | |
| Observations | 1151 | | |
| Marginal R^2 / Conditional R^2 | 0.000 / 0.986 | | |

trt: binäre Faktorvariable für Treatment (Referenzkategorie: Kontrolle), time: Zeit seit Index ICU-Entlassung;
 σ^2 : residual variance, CI: Confidence interval, ICC: intra-class correlation coefficient.

A.5 Technische Details – Effektevaluation

Im Folgenden werden einige wichtige methodische und softwaretechnische Referenzen aufgeführt, die insbesondere für die primäre Wirksamkeitsanalyse relevant sind:

Die statistischen Analysen wurden mit der Open-Source Softwarepaket R (<http://www.r-project.org>) durchgeführt:[68]

R version 4.0.4 (2021-02-15, Platform: x86_64-pc-linux-gnu (64-bit), unter:
Debian GNU/Linux 11 (bullseye)), unter RStudio Server Pro (Version 1.3.1073-1),

sowie zusätzlich auf einer lokalen Instanz:

R version 4.0.5 (2021-03-31), Platform: x86_64-w64-mingw32/x64 (64-bit), unter:
Windows 10 x64 (build 19043).



Session Info (relevanter Ausschnitt):

Matrix products: default

locale:

```
[1] LC_COLLATE=German_Germany.1252 LC_CTYPE=German_Germany.1252
LC_MONETARY=German_Germany.1252
[4] LC_NUMERIC=C LC_TIME=German_Germany.1252
```

attached base packages:

```
[1] stats graphics grDevices utils datasets methods base
```

other attached packages:

```
[1] lme4_1.1-25 Matrix_1.2-18 ggeffects_0.16.0 plyr_1.8.6 Hmisc_4.4-1
ggplot2_3.3.2
[7] Formula_1.2-4 survival_3.2-7 lattice_0.20-41
```

Wichtige verwendete R-packages für die Modell-basierten Analysen zu den QIs:

- Bates D, Maechler M, Bolker B and Walker S (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-25.
URL <http://CRAN.R-project.org/package=lme4>.
- Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
doi:10.18637/jss.v067.i01.

A.6 Reporting Konventionen

Das globale Signifikanzniveau wurde auf 0.05 festgelegt. Alle verwendeten Tests waren zweiseitig. p -Werte ≥ 0.001 werden mit 3 Dezimalstellen angegeben; p -Werte kleiner als 0.001 werden angegeben als " <0.001 ".

Der Mittelwert, die Standardabweichung (SD) und alle anderen statistischen Kenngrößen, die keine Quantile sind, werden mit einer Dezimalstelle mehr als die Originaldaten angegeben. Quantile oder Minimum und Maximum verwenden die gleiche Anzahl von Dezimalstellen wie die Originaldaten. Geschätzte Parameter mit anderer Skala im Vergleich zu den Rohdaten (z.B. Regressionskoeffizienten) werden mit drei Dezimalstellen berichtet.

B.1 Anhang: Kostenkalkulation für die Interventionskosten [TU Berlin]

PERSONAL HUB pro Workstation

| Gruppierung/Stufe | Beschreibung | Einsatz | FTE | Kosten pro Monat (EUR) |
|-------------------------------|--|--------------------------|------|------------------------|
| Ä3/2 | Qualifizierter Arzt, Oberarzt 2. Stufe | Visiten in 10 ICUs Mo-Fr | 1,30 | 13.502,50 |
| Ä3/2 | Qualifizierter Arzt, Oberarzt 2. Stufe | Visiten in 10 ICUs Sa-So | 0,52 | 5.401,00 |
| RD | Abdeckung 24/7 durch Rufdienst mit entsprechender Ä3/2 Qualifikation | RD Abdeckung 10 ICUs | NN | 6.462,08 |
| E9/4 | Spezialpflegekraft für pflegerische Fragestellungen, z.B. Bauchlagerung etc. | Teilnahme an Visiten | 0,50 | 2.390,79 |
| E 13/4 | Techniker zum Support der Hardware vor Ort | Technischer Support | 0,33 | 2.203,95 |
| Gesamtkosten pro Monat | | | | 29.960,32 |
| Gesamtkosten pro Jahr | | | | 359.523,81 |

| Parameter | |
|--------------|------------------------------|
| Gehaltsstufe | Haushaltsbrutto Jahresgehalt |
| Ä3/2 | 124.638 € |
| E9/4 | 57.379 € |
| E13/4 | 81.377 |

| ** Rufdienstherleitung | Wert |
|--|-------------|
| Rufdienst je Tag (h) | 16 |
| Tage im Jahr (d) | 365 |
| Vergütung je Rufbereitschaftsstunde des Stundentgeltes (%) | 12,50% |
| Aktivzeit je Tag (h) | 1 |
| Vergütete Stunden je Tag insgesamt (h) | 3 |
| Stunden pro Jahr (h) | 1095 |
| Nettojahresarbeitszeit in Studen (h) | 1760 |
| Stundenentgelt (h) | 70,8172912 |
| Gesamtkosten pro Jahr (h) | 77544,93387 |

Akronym: ERIC
 Förderkennzeichen: 01NVF16011

Technik Ausstattung pro Workstation mit 10 ICUs gemäß Angebot LITE vom 09. März 2020

| Beschreibung | Einheiten/Frequenz | Preis (EUR) | Spalte1 | Spalte2 |
|---|--------------------|-------------|---------|---------|
| TELEMEDICINE CART | | | | |
| Lite Telemedicine Cart | 10/einmalig | 140.168,40 | | |
| Installation | 10/einmalig | 12.859,50 | | |
| Shipping & Handling | 10/einmalig | 15.000,00 | | |
| Support & Maintenance | 10/jährlich | 17.360,30 | | |
| AUDIO VIDEO | | | | |
| Installation | 10/einmalig | 3.674,20 | | |
| License | 10/jährlich | 6.123,60 | | |
| Support & Maintenance | 10/jährlich | 32.450,00 | | |
| QI SOFTWARE PLATFORM | | | | |
| Installation | 10/einmalig | 3.674,20 | | |
| License | 10/jährlich | 59.521,10 | | |
| Support & Maintenance | 10/jährlich | 22.044,90 | | |
| Einmalige Kosten gesamt (unter Annahme von dreijähriger Nutzungsdauer) | | 175.376,30 | | |
| Jährliche Kosten gesamt | | 137.499,90 | | |
| Gesamtkosten pro Jahr | | 195.900,21 | | |

Technik Ausstattung pro Workstation im Hub

| Beschreibung | Einheiten/Frequenz | Preis (EUR) |
|--|--------------------|-------------|
| PC-Technik, Telefontechnik, Sichtschutz, Kameratechnik zentral, inkl. Installation | 1/jährlich | 30.000,00 |
| Gesamtkosten pro Jahr | | 30.000,00 |

Akronym: ERIC
 Förderkennzeichen: 01NVF16011

Schulungskosten Hub

| Beschreibung | Einheiten/Frequenz | Preis (EUR) |
|------------------------------|--------------------|-----------------|
| Arzt | 10/jährlich | 500,00 |
| Pflegekraft | 10/jährlich | 150,00 |
| Gesamtkosten pro Jahr | | 6.500,00 |

Schulungskosten Hub

| Beschreibung | Einheiten/Frequenz | Preis (EUR) |
|-------------------------------------|---------------------|------------------|
| Arzt | 150/alle drei Jahre | 500,00 |
| Pflegekraft | 150/alle drei Jahre | 150,00 |
| Gesamtkosten alle drei Jahre | | 97.500,00 |
| Gesamtkosten pro Jahr | | 32.467,50 |

Honorar Hausarzt

| Beschreibung | Einheiten/Frequenz | Preis Patient (EUR) |
|------------------------------|---------------------|---------------------|
| Arzt Tele-ICU | je Patient/jährlich | 40,00 |
| Gesamtkosten pro Jahr | | 58.400,00 |

Gesamtkosten pro Jahr

| Kostenart | Gesamtkosten (EUR) |
|-------------------------------------|--------------------|
| Personal | 359.523,81 |
| Technik | 225.900,21 |
| Technik zzgl. USt. | 268.821,25 |
| Honorar Hausärzte | 58.400,00 |
| Schulungen | 38.967,50 |
| Gesamtkosten pro Jahr | 951.612,77 |
| resultierende Kosten pro Patient*in | 651,79 |
| resultierende Kosten pro Televisite | 130,36 |

| Annahmen | |
|---|-----------|
| Anzahl angebundener remote ICUs | 10 |
| Anzahl Televisiten pro Tag | 20 |
| Dauer der einzelnen Televisite (min) | 25 |
| Anzahl Televisiten pro Jahr/Hub | 7300 |
| Anzahl Televisiten pro Patient | 5 |
| Anzahl Patienten pro Jahr/Hub | 1460 |
| Anzahl Patienten pro Jahr/Remote ICU | 146 |

B.2 Technische Details – gesundheitsökonomische Evaluation [TU Berlin]

Im Folgenden ist der für die gesundheitsökonomische Evaluation relevante Ausschnitt der Session Infos im Statistikprogramm R dargestellt.

R version 4.0.2 (2020-06-22)

Platform: x86_64-w64-mingw32/x64 (64-bit)

Running under: Windows 10 x64 (build 19042)

Matrix products: default

locale:

[1] LC_COLLATE=German_Germany.1252 LC_CTYPE=German_Germany.1252

[3] LC_MONETARY=German_Germany.1252 LC_NUMERIC=C

[5] LC_TIME=German_Germany.1252

attached base packages:

[1] stats graphics grDevices utils datasets methods base

other attached packages:

[1] broom.mixed_0.2.6 sjlabelled_1.1.8 sjmisc_2.8.7 sjPlot_2.8.8

[5] flextable_0.6.6 mfx_1.2-2 betareg_3.1-4 MASS_7.3-51.6

[9] lmtest_0.9-38 zoo_1.8-9 sandwich_3.0-1 survival_3.2-11

[13] lmerTest_3.1-3 lme4_1.1-27 Matrix_1.2-18

B.3 Inferenzstatistische Analysen und Sensitivitätsanalysen – gesundheitsökonomische Evaluation [TU Berlin]

Tabelle 32. Kosten: Ergebnisse der linearen gemischten Modelle mit logarithmierten Kosten.

| Parameter | M1 | | | M2 | | |
|----------------------------|---------|-------------|--------|---------|-------------|--------|
| | β | KI | p | β | KI | p |
| (Intercept) | 9,74 | 9,13; 10,35 | <0,001 | 9,47 | 8,99; 9,95 | <0,001 |
| Intervention (Referenz: K) | 0,31 | -0,27; 0,89 | 0,299 | 0,33 | -0,25; 0,92 | 0,263 |
| Zeit | -0,00 | -0,00; 0,00 | 0,327 | -0,00 | -0,00; 0,00 | 0,445 |
| SAPS II | 0,02 | 0,01; 0,02 | <0,001 | 0,01 | 0,01; 0,02 | <0,001 |
| Intervention * Zeit | -0,00 | -0,00; 0,00 | 0,579 | -0,00 | -0,00; 0,00 | 0,568 |
| Alter | -0,00 | -0,01; 0,00 | 0,078 | | | |
| Geschlecht (Referenz: m) | 0,11 | -0,04; 0,26 | 0,163 | | | |

Random Effects

| | | | |
|---------------------------------------|---------------|-------------------|---------------|
| σ^2 | 1,07 | | 1,07 |
| τ_{00} | 0,07 | fall_studycluster | 0,08 |
| ICC | 0,06 | | 0,07 |
| N | 10 | fall_studycluster | 10 |
| Beobachtungen | 727 | | 727 |
| Marginal R ² / Conditional | 0,076 / 0,129 | | 0,066 / 0,128 |

Abkürzungen: ICC=Intraklassenkorrelation; K=Kontrolle; KI=95 %-Konfidenzintervall; m=männlich; N=Anzahl; p = p-Wert; SAPS 2= Simplified Acute Physiology Score; β =Regressionskoeffizient

Tabelle 33. Kosten: Ergebnisse der generalisierten linearen Modelle unter Annahme der Gammaverteilung als Sensitivitätsanalyse.

| Parameter | M1 | | | M2 | | |
|----------------------------|---------|--------------|--------------|---------|--------------|--------------|
| | β | KI | p | β | KI | p |
| (Intercept) | 4,59 | 3,08; 6,17 | <0,001 | 5,26 | 4,21; 6,36 | <0,001 |
| Intervention (Referenz: K) | 1,85 | -4,25; 7,51 | 0,536 | 1,78 | -4,61; 7,69 | 0,571 |
| Cluster | -1,25 | -2,45; -7,46 | 0,042 | -1,32 | -2,58; -9,26 | 0,038 |
| SAPS II | -3,59 | -4,86; -2,17 | <0,001 | -3,55 | -4,84; -2,09 | <0,001 |
| Alter | 1,60 | -1,67; 3,21 | 0,065 | | | |
| Geschlecht (Referenz: m) | -6,69 | -1,26; -1,05 | 0,023 | | | |
| Beobachtungen | 727 | | | 727 | | |
| R ² Nagelkerke | 0,121 | | | 0,095 | | |

Anmerkung: Aufgrund des nicht signifikanten Interventionseffekts wird auf die Darstellung der marginalen Effekte verzichtet.

Abkürzungen: K=Kontrolle; KI=95 %-Konfidenzintervall; m= männlich; p=p-Wert; β = Regressionskoeffizient

Tabelle 34. Kosten: Ergebnisse der generalisierten linearen Modelle unter Annahme der Weibullverteilung als Sensitivitätsanalyse.

| Parameter | M1 | | | M2 | | |
|-------------------------------|----------|--------------------|------------------|----------|--------------------|------------------|
| | β | KI | p | β | KI | p |
| (Intercept) | 27452,01 | 15004,47; 50225,88 | <0,001 | 20805,44 | 12783,54; 33861,22 | <0,001 |
| Intervention (Referenz: K) | 1,13 | 0,65; 1,99 | 0,659 | 1,17 | 0,67; 2,06 | 0,585 |
| Zeit | 1,00 | 1,00; 1,00 | 0,125 | 1,00 | 1,00; 1,00 | 0,183 |
| Cluster | 1,03 | 1,00; 1,06 | 0,045 | 1,04 | 1,01; 1,07 | 0,016 |
| SAPS II | 1,01 | 1,01; 1,02 | <0,001 | 1,01 | 1,01; 1,02 | <0,001 |
| Intervention * Zeit | 1,00 | 1,00; 1,00 | 0,378 | 1,00 | 1,00; 1,00 | 0,336 |
| Log(scale) | 0,98 | 0,93; 1,04 | 0,526 | 0,99 | 0,94; 1,05 | 0,810 |
| Alter | 1,00 | 0,99; 1,00 | 0,064 | | | |
| Geschlecht (Referenz: m) | 1,22 | 1,05; 1,41 | 0,008 | | | |
| Beobachtungen | 727 | | | 727 | | |
| R ² Nagelkerke | 0,098 | | | 0,094 | | |

Abkürzungen: K=Kontrolle; KI=95 %-Konfidenzintervall; m= männlich; p=p-Wert; β = Regressionskoeffizient

Tabelle 35. Kosten des Indexaufenthalts: Ergebnisse der linearen gemischten Modelle mit logarithmierten Kosten als Sensitivitätsanalyse.

| Parameter | M1 | | | M2 | | |
|--|---------------|-------------------|------------------|---------------|-------------------|------------------|
| | β | KI | p | β | KI | p |
| (Intercept) | 8,88 | 8,48; 9,27 | <0,001 | 8,64 | 8,37; 8,92 | <0,001 |
| Intervention (Referenz: K) | 0,13 | -0,09; 0,34 | 0,245 | 0,15 | -0,06; 0,37 | 0,169 |
| Zeit | -0,00 | -0,00; -0,00 | 0,009 | -0,00 | -0,00; -0,00 | 0,029 |
| SAPS II | 0,02 | 0,01; 0,02 | <0,001 | 0,02 | 0,01; 0,02 | <0,001 |
| Intervention * Zeit | 0,00 | -0,00; 0,00 | 0,165 | 0,00 | -0,00; 0,00 | 0,176 |
| Alter | -0,00 | -0,01; -0,00 | 0,024 | | | |
| Geschlecht (Referenz: m) | 0,19 | 0,08; 0,30 | 0,001 | | | |
| Random Effects | | | | | | |
| σ^2 | 0,82 | | | 0,84 | | |
| T ₀₀ | 0,07 | fall_studycluster | | 0,07 | fall_studycluster | |
| ICC | 0,07 | | | 0,08 | | |
| N | 10 | fall_studycluster | | 10 | fall_studycluster | |
| Beobachtungen | 1120 | | | 1120 | | |
| Marginal R ² / Conditional R ² | 0,134 / 0,198 | | | 0,117 / 0,188 | | |

Abkürzungen: K=Kontrolle; KI=95 %-Konfidenzintervall; m= männlich; p=p-Wert; β = Regressionskoeffizient

Tabelle 36. Kosten des Indexaufenthalts: Ergebnisse des generalisierten linearen Modells unter Annahme der Gammaverteilung als Sensitivitätsanalyse.

| Parameter | M1 | | | M2 | | |
|-------------------------------|---------|--------------|--------|---------|--------------|--------|
| | β | KI | p | β | KI | p |
| (Intercept) | 1,21 | 9,51; 1,47 | <0,001 | 1,33 | 1,16; 1,51 | <0,001 |
| Intervention (Referenz: K) | -3,40 | -1,36; 6,02 | 0,498 | -3,46 | -1,41; 6,48 | 0,510 |
| Cluster | -3,46 | -5,40; -1,53 | 0,001 | -3,62 | -5,63; -1,63 | <0,001 |
| SAPS II | -8,75 | -1,04; -6,88 | <0,001 | -8,79 | -1,06; -6,82 | <0,001 |
| Alter | 3,26 | 3,65; -8,60 | 0,023 | | | |
| Geschlecht (Referenz: m) | -1,76 | -2,69; -8,59 | <0,001 | | | |
| Beobachtungen | 1422 | | | 1422 | | |
| R ² Nagelkerke | 0,189 | | | 0,157 | | |

Abkürzungen: K=Kontrolle; KI=95 %-Konfidenzintervall; m= männlich; p=p-Wert; β = Regressionskoeffizient

Tabelle 37. Kosten des Indexaufenthalts: Ergebnisse des generalisierten linearen Modells unter Annahme der Weibullverteilung als Sensitivitätsanalyse.

| Parameter | M1 | | | M2 | | |
|-------------------------------|---------|------------------|--------|---------|-------------------|--------|
| | β | KI | p | β | KI | p |
| (Intercept) | 6461,05 | 4958,28; 8419,28 | <0,001 | 8508,96 | 5673,43; 12761,67 | <0,001 |
| Intervention (Referenz: K) | 1,00 | 0,80; 1,25 | 0,997 | 1,02 | 0,81; 1,28 | 0,860 |
| Zeit | 1,00 | 1,00; 1,00 | 0,010 | 1,00 | 1,00; 1,00 | 0,010 |
| Cluster | 1,06 | 1,03; 1,08 | <0,001 | 1,05 | 1,03; 1,08 | <0,001 |
| SAPS II | 1,02 | 1,01; 1,02 | <0,001 | 1,02 | 1,01; 1,02 | <0,001 |
| Intervention * Zeit | 1,00 | 1,00; 1,00 | 0,416 | 1,00 | 1,00; 1,00 | 0,590 |
| Log(scale) | 0,98 | 0,94; 1,02 | 0,276 | 0,96 | 0,92; 1,00 | 0,069 |
| Alter | | | | 0,99 | 0,99; 1,00 | 0,005 |
| Geschlecht (Referenz: m) | | | | 1,29 | 1,15; 1,45 | <0,001 |
| Beobachtungen | 1120 | | | 1120 | | |
| R ² Nagelkerke | 0,152 | | | 0,128 | | |

Abkürzungen: K=Kontrolle; KI=95 %-Konfidenzintervall; m= männlich; p=p-Wert; β = Regressionskoeffizient

Tabelle 38. Qualitätsadjustierte Lebensjahre: Ergebnisse der linearen gemischten Modelle.

| Parameter | M1 | | | M2 | | |
|--|---------------|-------------------|--------|---------------|-------------------|--------|
| | β | KI | p | β | KI | p |
| (Intercept) | 0,44 | 0,39; 0,49 | <0,001 | 0,24 | 0,21; 0,27 | <0,001 |
| Intervention (Referenz: K) | -0,01 | -0,03; 0,01 | 0,239 | -0,00 | -0,03; 0,02 | 0,653 |
| Cluster | 0,00 | -0,00; 0,01 | 0,424 | 0,00 | -0,00; 0,01 | 0,424 |
| SAPS II | -0,00 | -0,00; -0,00 | <0,001 | -0,00 | -0,00; -0,00 | <0,001 |
| Alter | -0,00 | -0,00; -0,00 | <0,001 | | | |
| Geschlecht (Referenz: m) | 0,02 | 0,00; 0,04 | 0,026 | | | |
| LOS Indexaufenthalt | -0,00 | -0,00; -0,00 | <0,001 | | | |
| Random Effects | | | | | | |
| σ^2 | 0,02 | | | 0,02 | | |
| T ₀₀ | 0,00 | fall_studycluster | | 0,00 | fall_studycluster | |
| ICC | 0,02 | | | 0,01 | | |
| N | 10 | fall_studycluster | | 10 | fall_studycluster | |
| Beobachtungen | 897 | | | 897 | | |
| Marginal R ² / Conditional R ² | 0,191 / 0,205 | | | 0,107 / 0,118 | | |
| Abkürzungen: ICC=Intraklassenkorrelation; K=Kontrollgruppe; KI=95 %-Konfidenzintervall; LOS=Aufenthaltsdauer; m=männlich; N=Anzahl; p =p-Wert; SAPS II= Simplified Acute Physiology Score; β=Regressionskoeffizient | | | | | | |

Tabelle 39. Außerklinische Beatmung: Ergebnisse der generalisierten gemischten Logit-Modelle.

| Parameter | M1 | | | M2 | | |
|--|---------------|----------------------|--------|---------------|----------------------|--------|
| | OR | KI | p | OR | KI | p |
| (Intercept) | 0,00 | 0,00; 0,00 | <0,001 | 0,00 | 0,00; 0,00 | <0,001 |
| Intervention (Referenz: K) | 0,01 | 0,00; 3,26 | 0,118 | 0,01 | 0,00; 3,23 | 0,116 |
| Zeit | 3,49 | 0,35; 34,72 | 0,286 | 3,50 | 0,35; 34,81 | 0,286 |
| SAPS II | 1,02 | 0,97; 1,07 | 0,545 | 1,02 | 0,97; 1,07 | 0,523 |
| Intervention * Zeit | 12,38 | 0,64; 239,60 | 0,096 | 12,40 | 0,64; 240,96 | 0,096 |
| Alter | 1,01 | 0,94; 1,07 | 0,856 | | | |
| Geschlecht (Referenz: m) | 1,09 | 0,18; 6,71 | 0,925 | | | |
| Random Effects | | | | | | |
| σ^2 | 3,29 | | | 3,29 | | |
| T ₀₀ | 254,20 | fall_studycluster:ID | | 254,68 | fall_studycluster:ID | |
| ICC | 0,99 | | | 0,99 | | |
| N | 10 | fall_studycluster | | 10 | fall_studycluster | |
| N | 817 | ID | | 817 | ID | |
| Beobachtungen | 1438 | | | 1438 | | |
| Marginal R ² / Conditional R ² | 0,011 / 0,987 | | | 0,011 / 0,987 | | |
| Abkürzungen: ICC=Intraklassenkorrelation; K=Kontrolle; KI=95 %-Konfidenzintervall; m=männlich; N=Anzahl; OR= Odds Ratio; p =p-Wert; SAPS II= Simplified Acute Physiology Score; β=Regressionskoeffizient | | | | | | |

LITERATURVERZEICHNIS

Im Folgenden werden einige wichtige methodische und softwaretechnische Referenzen aufgeführt, die insbesondere für die primäre Wirksamkeitsanalyse relevant sind.

1. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG: **CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials**. *BMJ* 2010, **340**:c869.
2. Hemming K, Taljaard M, McKenzie JE, Hooper R, Copas A, Thompson JA, Dixon-Woods M, Aldcroft A, Doussau A, Grayling M *et al*: **Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration**. *BMJ* 2018, **363**:k1614.
3. Needham DM, Davidson J, Cohen H, Hopkins RO, Weinert C, Wunsch H, Zawistowski C, Bemis-Dougherty A, Berney SC, Bienvenu OJ *et al*: **Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference**. *Crit Care Med* 2012, **40**(2):502-509.
4. Elliott D, Davidson JE, Harvey MA, Bemis-Dougherty A, Hopkins RO, Iwashyna TJ, Wagner J, Weinert C, Wunsch H, Bienvenu OJ *et al*: **Exploring the scope of post-intensive care syndrome therapy and care: engagement of non-critical care providers and survivors in a second stakeholders meeting**. *Crit Care Med* 2014, **42**(12):2518-2526.
5. Kumpf O, Braun JP, Brinkmann A, Bause H, Bellgardt M, Bloos F, Dubb R, Greim C, Kaltwasser A, Marx G *et al*: **Quality indicators in intensive care medicine for Germany - third edition 2017**. *Ger Med Sci* 2017, **15**:Doc10.
6. Grohe C SB: **Außerklinische Beatmung - eine wachsende Herausforderung**. *Dtsch Med Wochenschr* 2017, **142**:1181.
7. Karagiannidis C SS, Callegari J ., et al.: **Epidemiologische Entwicklung der außerklinischen Beatmung: Eine rasant zunehmende Herausforderung für die ambulante und stationäre Patientenversorgung**. *Pneumologie* 2019, **73**:670-667.
8. BKK Dachverband und SBK: **Erster Qualitätsvertrag zur Beatmungsentwöhnung (Weaning)**; 2019.
9. Spies CD, Krampe H, Paul N, Denke C, Kiselev J, Piper SK, Kruppa J, Grunow JJ, Steinecke K, Gülmez T *et al*: **Instruments to measure outcomes of post-intensive care syndrome in outpatient care settings – Results of an expert consensus and feasibility field test**. *Journal of the Intensive Care Society* 2021, **22**(2):159-174.
10. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, Williamson PR: **Core Outcome Set-STANDards for Development: The COS-STAD recommendations**. *PLoS Med* 2017, **14**(11):e1002447.
11. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE: **Systematic review of the application of the plan–do–study–act method to improve quality in healthcare**. *BMJ Quality & Safety* 2014, **23**(4):290-298.
12. Braun JP, Bause H: **[Peer review in ICU]**. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 2012, **106**(8):566-570.
13. McNicholas C, Lennox L, Woodcock T, Bell D, Reed JE: **Evolving quality improvement support strategies to improve Plan–Do–Study–Act cycle fidelity: a retrospective mixed-methods study**. *BMJ Quality & Safety* 2019, **28**(5):356-365.
14. Brenner T, Bingold T, Braun J, Bause H, Dubb R, Henninger A, Kaltwasser A, Kumpf O, Krueger W, Marx10 G: **Quality assurance concepts in intensive care medicine–**

- Medical peer review, forms of certification & benchmarking tools.**
ANASTHESIOLOGIE & INTENSIVMEDIZIN 2019, **60**:209-222.
15. Hooper R, Copas A: **Stepped wedge trials with continuous recruitment require new ways of thinking.** *J Clin Epidemiol* 2019.
 16. Hooper R: **Key concepts in clinical epidemiology: Stepped wedge trials.** *J Clin Epidemiol* 2021, **137**:159-162.
 17. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ: **The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting.** *BMJ : British Medical Journal* 2015, **350**:h391.
 18. **Online-Datenbank der Gesundheitsberichterstattung (GBE) des Bundes - The Information System of the Federal Health Monitoring** [www.gbe-bund.de]
 19. Needham DM, Sepulveda KA, Dinglas VD, Chessare CM, Aronson Friedman L, Bingham Iii CO, Turnbull AE: **Core Outcome Measures for Clinical Research in Acute Respiratory Failure Survivors: An International Modified Delphi Consensus Study.** *Am J Respir Crit Care Med* 2017.
 20. Adrion C, Weiss B, Paul N, Berger E, Busse R, Marschall U, Caumanns J, Rosseau S, Mansmann U, Spies C: **Enhanced Recovery after Intensive Care (ERIC): study protocol for a German stepped wedge cluster randomised controlled trial to evaluate the effectiveness of a critical care telehealth program on process quality and functional outcomes.** *BMJ open* 2020, **10**(9):e036096.
 21. Lowe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, Schneider A, Brahler E: **A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population.** *J Affect Disord* 2010, **122**(1-2):86-95.
 22. Weiss D: **The Impact of Event Scale: Revised.** In: *Cross-cultural assessment of psychological trauma and PTSD.* edn. Edited by Wilson JP, Tang CCS-K: Springer Science & Business Media; 2007: 219-238.
 23. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A: **The Mini-Cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly.** *Int J Geriatr Psychiatry* 2000, **15**(11):1021-1027.
 24. Borson S, Scanlan JM, Chen P, Ganguli M: **The Mini-Cog as a screen for dementia: validation in a population-based sample.** *J Am Geriatr Soc* 2003, **51**(10):1451-1454.
 25. Borson S, Scanlan JM, Watanabe J, Tu SP, Lessig M: **Improving identification of cognitive impairment in primary care.** *Int J Geriatr Psychiatry* 2006, **21**(4):349-355.
 26. Sager MA, Hermann BP, La Rue A, Woodard JL: **Screening for dementia in community-based memory clinics.** *WMJ* 2006, **105**(7):25-29.
 27. Podsiadlo D, Richardson S: **The timed "Up & Go": a test of basic functional mobility for frail elderly persons.** *J Am Geriatr Soc* 1991, **39**(2):142-148.
 28. Bohannon RW: **Hand-grip dynamometry predicts future outcomes in aging adults.** *J Geriatr Phys Ther* 2008, **31**(1):3-10.
 29. Roberts HC, Denison HJ, Martin HJ, Patel HP, Syddall H, Cooper C, Sayer AA: **A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach.** *Age Ageing* 2011, **40**(4):423-429.
 30. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X: **Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L).** *Qual Life Res* 2011, **20**(10):1727-1736.
 31. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, Swinburn P, Busschbach J: **Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L**

- across eight patient groups: a multi-country study.** *Qual Life Res* 2013, **22(7)**:1717-1727.
32. Mahler DA, Wells CK: **Evaluation of clinical methods for rating dyspnea.** *Chest* 1988, **93(3)**:580-586.
33. Celli BR, Cote CG, Marin JM, Casanova C, Montes de Oca M, Mendez RA, Pinto Plata V, Cabral HJ: **The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease.** *N Engl J Med* 2004, **350(10)**:1005-1012.
34. Andrews G, Kemp A, Sunderland M, Von Korff M, Ustun TB: **Normative data for the 12 item WHO Disability Assessment Schedule 2.0.** *PLoS One* 2009, **4(12)**:e8343.
35. Üstün TB, Kostanjsek N, Chatterji S, Rehm J: **Measuring health and disability: Manual for WHO disability assessment schedule WHODAS 2.0:** World Health Organization; 2010.
36. **Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Zentrums-Regelungen: Zentren in einem intensivmedizinischen digital-gestützten Versorgungsnetzwerk (IDV-Zentren) – Anhang zu den Anlagen 5 und 7**
[\[https://www.g-ba.de/downloads/39-261-4720/2021-02-18_Zentrums-Regelungen_IDV-Zentren_BAnz.pdf\]](https://www.g-ba.de/downloads/39-261-4720/2021-02-18_Zentrums-Regelungen_IDV-Zentren_BAnz.pdf)
37. Luetz A, Balzer F, Radtke FM, Jones C, Citerio G, Walder B, Weiss B, Wernecke KD, Spies C: **Delirium, sedation and analgesia in the intensive care unit: a multinational, two-part survey among intensivists.** *PLoS One* 2014, **9(11)**:e110935.
38. Nachtigall I, Tamarkin A, Tafelski S, Deja M, Halle E, Gastmeier P, Wernecke KD, Bauer T, Kastrup M, Spies C: **Impact of adherence to standard operating procedures for pneumonia on outcome of intensive care unit patients.** *Crit Care Med* 2009, **37(1)**:159-166.
39. Hussey MA, Hughes JP: **Design and analysis of stepped wedge cluster randomized trials.** *Contemp Clin Trials* 2007, **28(2)**:182-191.
40. Rutterford C, Copas A, Eldridge S: **Methods for sample size determination in cluster randomized trials.** *Int J Epidemiol* 2015, **44(3)**:1051-1067.
41. Lenguerrand E, Winter C, Siassakos D, MacLennan G, Innes K, Lynch P, Cameron A, Crofts J, McDonald A, McCormack K *et al*: **Effect of hands-on interprofessional simulation training for local emergencies in Scotland: the THISTLE stepped-wedge design randomised controlled trial.** *BMJ Quality & Safety* 2020, **29(2)**:122-134.
42. Molenberghs G, Thijs H, Jansen I, Beunckens C, Kenward MG, Mallinckrodt C, Carroll RJ: **Analyzing incomplete longitudinal clinical trial data.** *Biostatistics* 2004, **5(3)**:445-464.
43. O'Kelly M, Ratitch B: **Clinical Trials with Missing Data: A Guide for Practitioners:** John Wiley & Sons; 2014.
44. White IR, Carpenter J, Horton NJ: **Including all individuals is not enough: lessons for intention-to-treat analysis.** *Clin Trials* 2012, **9(4)**:396-407.
45. Nickless A, Voysey M, Geddes J, Yu L-M, Fanshawe TR: **Mixed effects approach to the analysis of the stepped wedge cluster randomised trial—Investigating the confounding effect of time through simulation.** *PLoS One* 2018, **13(12)**:e0208876.
46. Barker D, McElduff P, D'Este C, Campbell MJ: **Stepped wedge cluster randomised trials: a review of the statistical methodology used and available.** *BMC Med Res Methodol* 2016, **16(1)**:69.

47. Colantuoni E, Scharfstein DO, Wang C, Hashem MD, Leroux A, Needham DM, Girard TD: **Statistical methods to compare functional outcomes in randomized controlled trials with high mortality.** *BMJ* 2018, **360**:j5748.
48. Balan TA, Putter H: **A tutorial on frailty models.** *Stat Methods Med Res* 2020, **29**(11):3424-3454.
49. Bundesamt S: **Gesundheitsausgaben in Deutschland in Mio. € Gliederungsmerkmale: Jahre, Art der Einrichtung, Art der Leistung, Ausgabenträger.** In.; 2021.
50. Gräfe V, Berger T, Hautzinger M, Hohagen F, Lutz W, Meyer B, Moritz S, Rose M, Schröder J, Späth C *et al*: **Health economic evaluation of a web-based intervention for depression: the EVIDENT-trial, a randomized controlled study.** *Health economics review* 2019, **9**(1):16.
51. Manca A, Hawkins N, Sculpher MJ: **Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility.** *Health Econ* 2005, **14**(5):487–496.
52. Bock J-O, Brettschneider C, Seidl H, Bowles D, Holle R, Greiner W, König HH: **Ermittlung standardisierter Bewertungssätze aus gesellschaftlicher Perspektive für die gesundheitsökonomische Evaluation.** *Das Gesundheitswesen* 2015, **77**(1):53–61.
53. Drummond MF, Sculpher M, Claxton K, Stoddart GL, Torrance GW: **Methods for the economic evaluation of health care programmes**, 4th edition edn. Oxford: Oxford University Press; 2015.
54. Bundesamt S: **Verbraucherpreisindex: Deutschland, Jahre, Klassifikation der Verwendungszwecke des Individualkonsums (COICOP 2-/3-/4-/5-/10 Steller/Sonderpositionen).** In.; 2021.
55. Gesundheitswesen IfQuWi: **Allgemeine Methoden - Version 6.0.** In.; 2020: 291.
56. Ludwig K, Graf von der Schulenburg J-M, Greiner W: **German Value Set for the EQ-5D-5L.** *Pharmacoeconomics* 2018, **36**(6):663–674.
57. Thalanany MM, Mugford M, Hibbert C, Cooper NJ, Truesdale A, Robinson S, Tiruvoipati R, Elbourne DR, Peek GJ, Clemens F *et al*: **Methods of data collection and analysis for the economic evaluation alongside a national, multi-centre trial in the UK: conventional ventilation or ECMO for Severe Adult Respiratory Failure (CESAR).** *BMC Health Serv Res* 2008, **8**:94.
58. Gates S, Lall R, Quinn T, Deakin CD, Cooke MW, Horton J, Lamb SE, Slowther A-M, Woollard M, Carson A *et al*: **Prehospital randomised assessment of a mechanical compression device in out-of-hospital cardiac arrest (PARAMEDIC): a pragmatic, cluster randomised trial and economic evaluation.** *Health technology assessment (Winchester, England)* 2017, **21**(11):1–176.
59. Malehi AS, Pourmotahari F, Angali KA: **Statistical models for the analysis of skewed healthcare cost data: a simulation study.** *Health economics review* 2015, **5**:11.
60. Fahrmeir L, Kneib T, Lang S: **Regression: Modelle, Methoden und Anwendungen.** Berlin: Springer; 2007.
61. Brooke J: **SUS - A quick and dirty usability scale.** In: *Usability Evaluation in Industry. Volume 189*, edn. London: Taylor & Francis; 1996.
62. Mahdavi M, Vissers J, Elkhuizen S, van Dijk M, Vanhala A, Karampli E, Faubel R, Forte P, Coroian E, van de Klundert J: **The relationship between context, structure, and processes with outcomes of 6 regional diabetes networks in Europe.** *PLoS One* 2018, **13**(2):e0192599.

63. Le Gall JR, Lemeshow S, Saulnier F: **A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study.** *JAMA* 1993, **270**(24):2957-2963.
64. Vincent JL, de Mendonça A, Cantraine F, Moreno R, Takala J, Suter PM, Sprung CL, Colardyn F, Blecher S: **Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine.** *Crit Care Med* 1998, **26**(11):1793-1800.
65. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL: **Serial evaluation of the SOFA score to predict outcome in critically ill patients.** *JAMA* 2001, **286**(14):1754-1758.
66. Kear BM, Guck TP, McGaha AL: **Timed Up and Go (TUG) Test: Normative Reference Values for Ages 20 to 59 Years and Relationships With Physical and Mental Health Risk Factors.** *J Prim Care Community Health* 2017, **8**(1):9-13.
67. Nakagawa S, Johnson PCD, Schielzeth H: **The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded.** *Journal of The Royal Society Interface* 2017, **14**(134):20170213.
68. R Development Core Team: **R: A language and environment for statistical computing.** In. Vienna, Austria: R Foundation for Statistical Computing, <http://www.R-project.org/>; 2021.