

Anhang 2

zum gemeinsamen Bericht von GKV-Spitzenverband, PKV und DKG
zur Einführung eines pauschalierenden Entgeltsystems für psychiatrische und psychosomatische
Einrichtungen gemäß § 17d Absatz 4 Satz 8 KHG
vom 28.06.2019

- **Gesetzliche Evaluation von Modellvorhaben nach § 64b SGB V – EVA 64 (Stand: April 2019)**

Zusammenfassung des aktuellen Standes der Evaluation von Modellvorhaben nach §64b SGB V (EVA64)

Zur Vorlage an das Bundesministerium für Gesundheit
im Rahmen des §17d KHG (4)

Auf Basis der Ersten und Zweiten Zwischenberichte von:
DIAKO Nordfriesland gGmbH,
Johanniter-Krankenhaus Geesthacht GmbH,
imland Klinik Rendsburg GmbH,
Westküstenkliniken Brunsbüttel und Heide gGmbH,

und der Ersten Zwischenberichte von:
LWL-Universitätsklinikum Bochum,
Krankenhaus und Poliklinik Rüdersdorf GmbH,
Klinikum Hanau GmbH,
Klinikum Itzehoe – Zentrum für Psychosoziale Medizin,
St. Marien-Hospital Hamm gGmbH,
Psychiatrische Klinik Lüneburg gemeinnützige GmbH,
Heinrich-Braun-Klinikum Zwickau gGmbH,
Rudolf Virchow Klinikum Glauchau gGmbH,
Südharz Klinikum Nordhausen gGmbH

Bundesweit einheitliche Wissenschaftliche Evaluation von Modellvorhaben nach
§64b SGB V (EVA64)

Datum: 24.04.2019

Autoren:

Dr. Anne Neumann, M.P.H.¹, Roman Kliemt, M.A.², Dipl.-Psych. Fabian Baum¹, Dr. rer. nat. Olaf Schoffer¹, Dr. rer. medic. Stefanie March, M.A.³, Dipl.-Phys. Martin Seifert¹, PD Dr. rer. biol. hum. Enno Swart³, Dr. Dennis Häckl², Prof. Dr. med. Andrea Pfennig, M.Sc.⁴, Prof. Dr. med. Jochen Schmitt, M.P.H.¹

¹ Zentrum für Evidenzbasierte Gesundheitsversorgung, Universitätsklinikum Carl Gustav Carus an der Technischen Universität Dresden

² WIG2 Wissenschaftliches Institut für Gesundheitsökonomie und Gesundheitssystemforschung Leipzig

³ Institut für Sozialmedizin und Gesundheitsökonomie, Medizinische Fakultät, Otto-von-Guericke-Universität Magdeburg

⁴ Klinik und Poliklinik für Psychiatrie und Psychotherapie, Universitätsklinikum Carl Gustav Carus an der Technischen Universität Dresden

Diese Zusammenfassung ist keine vertraglich vereinbarte Leistung gemäß der Rahmenvereinbarung zur bundesweiten und einheitlichen Evaluation von Modellvorhaben nach §64b SGB V gegenüber dem Krankenkassenverbund (Auftraggeber). Die Zusammenfassung dient zur Vorbereitung des Berichts des GKV-SV an den Bundestag im Rahmen von §17d KHG (4).

1. Hintergrund und Methoden

Ein wissenschaftliches Konsortium bestehend aus der Medizinischen Fakultät Carl Gustav Carus der TU Dresden (Zentrum für Evidenzbasierte Gesundheitsversorgung sowie Klinik und Poliklinik für Psychiatrie und Psychotherapie), dem Wissenschaftlichen Institut für Gesundheitsökonomie und Gesundheitssystemforschung (WIG² GmbH) in Leipzig sowie dem Institut für Sozialmedizin und Gesundheitsökonomie (ISMG) der Otto-von-Guericke-Universität Magdeburg wurde mit der bundesweit einheitlichen Evaluation von insgesamt 18 Modellvorhaben nach §64b SGB V beauftragt. Ziele der Modellverträge nach §64b sind die Implementierung und Weiterentwicklung einer verbesserten Patientenversorgung durch eine sektorenübergreifende Behandlungsplanung und –durchführung. Konkret sollen Alternativen zur vollstationären Behandlung - mit verbesserter Anpassung der Dauer und Intensität der Behandlung an den konkreten Behandlungsbedarf betroffener Versicherter - gestärkt werden, eine kontinuierliche Behandlung und Stabilisierung der Patienten unter Berücksichtigung ihres sozialen und beruflichen Umfelds erfolgen, die Akzeptanz der patientenorientierten psychiatrischen, psychotherapeutischen und soziotherapeutischen Interventionen erhöht werden sowie Leistungstransparenz und eine kosten-effektivere Verwendung der vorhandenen Mittel zur Versorgung der Versicherten erreicht werden.

In der vorliegenden Zusammenfassung wird der aktuelle Stand der Evaluation basierend auf Auswertungen der bislang vorliegenden und abgenommenen Ersten Zwischenberichte von 13 Modellkliniken sowie der bislang vorliegenden und abgenommenen Zweiten Zwischenberichte von vier Modellkliniken in kurzer Form dargelegt. Die Auswertungen der Ersten Zwischenberichte fokussieren auf einen Vergleich des 1. Modelljahres mit dem Jahr vor Einschluss in die Evaluation. Die Zweiten Zwischenberichte erlauben eine Untersuchung der patientenindividuell ersten drei Jahre ab Einschluss in die Evaluation für Patienten, die im 1. Modelljahr behandelt wurden. Es wurden Regressionsmodelle für die Zielparameter (siehe unten) berechnet und daraus Behandlungseffekte zwischen Patienten der Modellkliniken (IG) und der Kontrollkliniken (KG) im Zeitverlauf als Difference-in-Difference (DiD) abgeleitet. Die DiD-Methode erlaubt einen Vergleich der Unterschiede zwischen IG und KG im Zeitverlauf vom Zeitraum vor dem Modellstart (Präzeitraum) zum 1. Jahr nach Modellstart unter Berücksichtigung der Ausgangspositionen von IG und KG.

Basis für die vorliegende Zusammenfassung sind Routinedaten der Gesetzlichen Krankenversicherung von insgesamt 26.398 erwachsene Patienten aus 13 Modellkliniken sowie strukturähnlichen Kontrollkliniken (Erste Zwischenberichte, mit 1-Jahres-Follow-up) bzw. 9.954 erwachsene Patienten aus vier Modellkliniken sowie strukturähnlichen Kontrollkliniken (Zweite Zwischenberichte, mit 3-Jahres-Follow-up). Aufgrund der vergleichsweise geringeren Fallzahl für die Kinder- und Jugendpsychiatrie (KJP) (zwei der 13 bereits untersuchten Modellkliniken, nur Erste Zwischenberichte; n=1.664) und der damit verbundenen geringeren Aussagekraft werden in der vorliegenden Zusammenfassung nur Ergebnisse für erwachsene Patienten dargestellt. Die bislang vorläufig evaluierten Modellkliniken verteilen sich über die Bundesländer Brandenburg, Hessen, Schleswig-Holstein, Nordrhein-Westfalen, Niedersachsen, Sachsen und Thüringen. Der Beobachtungszeitraum umfasst jeweils ein Jahr ab Modellstart für die Ersten Zwischenberichte und bis zu drei Jahre ab Modellstart für die Zweiten Zwischenberichte, wobei der Modellstart für die bisher vorläufig evaluierten Modelle zwischen Januar 2013 und Dezember 2014 lag.

Vier Modellvorhaben begannen aus vorheriger Regelversorgung. Drei Modelle bauten auf einem Vertrag für Integrierte Versorgung nach § 140a SGB V und sechs auf einem bereits vorbestehenden Modell zur Entwicklung pauschalierter Vergütungen nach §24 Bundespflegesatzverordnung auf. Es ist daher nicht unwahrscheinlich, dass in Kliniken mit solchen Vorläuferverträgen, in denen teils bereits eine §64b-artige Struktur gebahnt wurde, im betrachteten Präzeitraum bereits ein interventionsähnlicher Effekt wirkte. Im Hinblick auf die Interpretation der Ergebnisse ist daher zu berücksichtigen, dass in diesen Kliniken Effekte vorweggenommen werden könnten, die für die Einführung des Modellvorhabens angenommen werden. Diese Problematik ist besonders für Patienten relevant, welche bereits vor dem Modellstart in der betreffenden Einrichtung/Modellklinik in psychiatrischer Behandlung waren („klinikbekannte Patienten“, siehe unten). Weiterhin ist zu beachten, dass voraussichtlich Modellvorhaben, die aus der Regelversorgung herausgehen, eine gewisse Zeit benötigen, um die Modellstruktur zu etablieren und Effekte für Patienten des ersten Modelljahres möglicherweise geringer ausfallen. Alle Modelle mit einem Zweiten Zwischenbericht wiesen Vorläuferverträge auf.

Um potenziell bereits existierende Effekte durch Vorläuferverträge bestmöglich von Modelleffekten zu trennen, wurde in der Evaluation daher eine Unterteilung des Patientenkollektivs in klinikneue und klinikbekannte Patienten vorgenommen: Als klinikneu wurden solche Patienten definiert, welche mit dem Einschluss in die Evaluationsstudie zum ersten Mal in der psychiatrischen Fachabteilung bzw. den jeweils angeschlossenen Bereichen (z.B. Psychiatrische Institutsambulanz) der Referenzklinik behandelt wurden (bezogen auf zwei Jahre vor Einschluss). Analog dazu wurden als klinikbekannt die Patienten identifiziert, welche bereits vor Einschluss in die Evaluationsstudie (bezogen auf zwei Jahre vor Einschluss) in diesen Bereichen der Referenzklinik behandelt worden waren.

Für klinikneue Patienten ist davon auszugehen, dass im Präzeitraum auch in Modellkliniken mit Vorläufervertrag noch kein interventionsähnlicher Effekt wirkte, womit anhand dieser Patienten trotz der unterschiedlichen Ausgangsbedingungen der Modelle eine gemeinsame Evaluation aller Modelle ermöglicht wird. Für die klinikneuen Patienten lassen sich damit in Bezug auf die Effektivität, die psychiatrischen Versorgungskosten sowie die Effizienz der Modellversorgung gezieltere Aussagen im Hinblick auf die abgeleiteten Hypothesen treffen als für die klinikbekannten Patienten. Ergebnisse zu den klinikbekannten Patienten sind mit großer Vorsicht zu interpretieren. Neben den für einige Modelle möglicherweise vorgezogenen interventionsähnlichen Effekten kann bei klinikbekannten Patienten nicht ausgeschlossen werden, dass eine vor Modellstart in derselben Klinik begonnene Behandlung nach Modellstart unverändert gemäß Regelversorgung fortgesetzt wurde.

In der Beschreibung der Patienten mit Einschluss im ersten Modelljahr und Beobachtung über drei Jahre (Zweite Zwischenberichte) werden zusätzlich die klinikbekannten Patienten dargestellt, um u.a. Effekte durch Vorläuferverträge abschätzen zu können. Die Evaluation berücksichtigte Patienten, die im Beobachtungszeitraum aufgrund einer von 16 vordefinierten psychischen Erkrankungen¹ in einer der Modell- oder Kontrollkliniken behandelt wurden.

¹ F00: Demenz; F01: Vaskuläre Demenz; F02: Demenz bei anderenorts klassifizierten Krankheiten; F03: Nicht näher bezeichnete Demenz; F07: Persönlichkeits- und Verhaltensstörung aufgrund einer Krankheit, Schädigung oder Funktionsstörung des Gehirns; F10: Psychische und Verhaltensstörungen durch Alkohol; F20-F29: Schizophrenie, schizotype und wahnhaftige Störungen; F30-F39: Affektive Störungen; F43: Reaktionen auf schwere Belastungen und Anpassungsstörungen; F45: Somatoforme Störungen; F40-F48: Neurotische, Belastungs- und somatoforme Störungen; F50: Essstörungen; F60.31: Spezifische Persönlichkeitsstörungen vom Typ Borderline; F70-F79: Intelligenzstörung; F84:

Hauptzielparameter der Evaluation sind die kumulierte Dauer stationärer Aufenthalte und - für die Teilgruppe der erwerbstätigen Patienten - die kumulierte Dauer der Arbeitsunfähigkeit von Versicherten. Diese sind gut mit Routinedaten der Gesetzlichen Krankenversicherungen abzubilden und erlauben Aussagen hinsichtlich vorgesehener Ziele der Modellvorhaben. Die Dauer stationärer Aufenthalte steht als Parameter für das Teilziel einer verbesserten sektorenübergreifenden Patientenversorgung und einer Stärkung stationsäquivalenter Behandlungsmöglichkeiten. Die Dauer der Arbeitsunfähigkeit steht als Parameter für das Teilziel des Erhalts der sozialen Teilhabe. Es wurde die Hypothese aufgestellt, dass die Modellintervention zu einer Verringerung der kumulierten Dauer stationärer Aufenthalte sowie zu einer Verringerung der kumulierten Dauer von Arbeitsunfähigkeitsepisoden im Vergleich zum Vorzeitraum führen sollte. Weitere (sekundäre) Zielparameter sind Häufigkeit des Einsatzes ambulanter Leistungen, Kontinuität der Behandlung nach Entlassung aus der Klinik, Abbrüche des Kontakts schwer psychisch Kranker zum psychiatrischen Versorgungssystem, Arzt- und/oder Krankenhaus-Hopping, Wiederaufnahmeraten, Mortalität, Progression psychischer Erkrankungen, Leitlinienadhärenz, Patientenmix und Leistungserbringungen sowie Kosten.

Über die Einzelergebnisse der 13 Ersten Zwischenberichte wurde zusätzlich eine Meta-Analyse gerechnet, um statistisch zusammengefasste oder „gepoolte“ (modellübergreifende) Effekte für diese 13 Modelle ableiten zu können. Dies war bei den Zweiten Zwischenberichten aufgrund der bisher nur vier bereits ausgewerteten Modelle nicht sinnvoll. Weitere Ausführungen zum Studiendesign und den Methoden sind im Methodenband der Berichte zu finden.

2. Ergebnisse für das erste Modelljahr über 13 Modelle

Weil davon auszugehen ist, dass die untersuchten klinikneuen Patienten im Präzeitraum nur in geringem Maße Leistungen durch andere Kliniken in Anspruch genommen haben, ist für die IG und die KG von einem Anstieg der Dauer stationärer Aufenthalte pro Person auszugehen. Tatsächlich nahm diese über die Zeit für alle Modelle in der IG (prä: 1,0 bis 3,8 Tage; Mittelwert: 2,0 Tage | 1. Jahr: 6,4 bis 30,0 Tage; Mittelwert: 15,6 Tage) und der KG (prä: 0,4 bis 3,6 Tage; Mittelwert: 2,2 Tage | 1. Jahr: 13,9 bis 25,9 Tage; Mittelwert: 21,3 Tage) zu. Dabei war die Dauer vollstationärer Aufenthalte im ersten Jahr nach Modellstart für elf Modelle in der IG kleiner als in der KG und für zwei Modelle in der IG größer als in der KG. Auch bei Betrachtung der DiD-Schätzung war der Anstieg der Dauer vollstationärer Aufenthalte hypothesenkonform in der IG über die Zeit signifikant geringer als in der KG. Besonders stark war der Effekt einer verringerten Aufenthaltsdauer in den Kliniken mit modellähnlichem Vorläufervertrag ausgeprägt, dagegen weniger bei solchen ohne Vorvertrag. Der gepoolte Schätzer über alle 13 Modellvorhaben betrug -5,43 vollstationäre Tage (95%-Konfidenzintervall: - 7,41; - 3,44 vollstationäre Tage). Dies bedeutet, dass über alle 13 untersuchten Modellvorhaben hinweg bei den klinikneuen Patienten der IG der Anstieg der vollstationären Behandlungstage je Patient im Schnitt ca. 5 Tage geringer ausfiel als bei der KG. Die Ergebnisse liefern hier also erste Hinweise für eine Bestätigung der Hypothese und damit die Erreichung eines zentralen Modellziels.

Für klinikneue erwerbstätige erwachsene Patienten (n=6.863) nahm die kumulierte Dauer der Arbeitsunfähigkeit (AU) pro Person über die Zeit für alle Modelle in der IG (prä: 15,3 bis 33,8 Tage; Mittelwert: 20,8 Tage | 1. Jahr: 57,8 bis 83,6 Tage; Mittelwert: 67,6 Tage) und in der KG (prä: 10,3 bis 29,5 Tage; Mittelwert: 18,7 Tage | 1. Jahr: 57,4 bis 82,6 Tage; Mittelwert: 65,5 Tage) zu. Dabei war die kumulierte Dauer der AU im ersten Jahr nach Modellstart für sechs Modelle in der IG kleiner als in der KG und für sieben Modelle in der IG größer als in der KG. Entsprechend gab es für die Gruppe der klinikneuen erwachsenen Patienten über die Gesamtheit der in die Analyse eingegangenen Modelle hinsichtlich der Dauer von AU-Episoden keinen statistisch signifikanten Effekt. Auf die AU-Dauer hatte zwar fast jedes einzelne Modellvorhaben in der DiD-Schätzung einen signifikanten Effekt, die Effekte wirkten jedoch in unterschiedliche Richtungen, sodass der gepoolte Schätzer nicht signifikant wurde. Dabei fanden sich hypothesenkonforme Effekte einer verringerten AU-Dauer eher in den Kliniken mit Vorläufervertrag. Der gepoolte Schätzer über alle 13 Modellvorhaben betrug -0,02 AU-Tage (95% Konfidenzintervall: - 5,51; 5,47 AU-Tage).

Für die sekundären Zielparameter Patientenanteil mit stationärem Aufenthalt, Häufigkeit des Einsatzes ambulanter Leistungen, Kontinuität der Behandlung nach Entlassung aus der Klinik, Anzahl an AU-Episoden, Komorbidität, Mortalität sowie Leitlinienadhärenz ließen sich über die kurze Nachbeobachtungszeit noch keine Aussagen zur Modellwirksamkeit ableiten, da hier überwiegend keine signifikanten Effekte vorlagen, die Effekte teils stark streuten bzw. die festgestellten statistisch signifikanten Effekte uneinheitlich waren. Gerade die Ergebnisse zu den sekundären Zielparametern sind als nicht abschließend zu bewerten, weil den bisherigen Analysen teilweise geringe Fallzahlen und damit eine geringe statistische Aussagekraft („Power“) zugrunde lagen. Für weitere Ausführungen zu den sekundären Zielparametern siehe unten („3. Ergebnisse für die ersten drei Modelljahre über vier Modelle“).

Für die Kosten kann anhand der Ersten Zwischenberichte keine abschließende Aussage getroffen werden. Für klinikneue Patienten zeigte sich über den Beobachtungszeitraum bezüglich der kumulierten psychiatrischen Versorgungskosten über alle Modelle hinweg je Patient ein um 449,02 € geringerer, jedoch statistisch nicht signifikanter Kosteneffekt der IG gegenüber der KG im Zeitverlauf. Signifikante Effekte zeigten sich in der Gruppe der klinikneuen Patienten für die fünf Modelle, wobei drei Modelle geringere Kostenanstiege in der IG im Vergleich zur KG und in zwei Modellen höhere Kostenanstiege ergaben. Für die übrigen Modelle fanden sich keine signifikanten Kosteneffekte.

Ursächlich für die drei geringeren Kostenanstiege waren nach derzeitigem Stand eine geringere Zahl und Dauer vollstationärer und somit kostenintensiver Behandlungen in der IG gegenüber der KG. Maßgeblich verantwortlich für die zwei höheren Kosteneffekte waren durch Ausgleich der Vorzeiträume verursachte, deutlich höhere Pflegesätze in der IG, wodurch sich bei gleicher bzw. sogar wesentlich geringerer Anzahl vollstationärer Tage dennoch höhere Kosten ergaben.

Dennoch lässt sich trotz der teilweise sehr geringen Gruppengrößen in Verbindung mit der durch einzelne Hochkostenfälle verursachten großen Streuung, der teilweise vorhandenen Verzerrung durch Budgetausgleiche sowie der beträchtlichen Heterogenität der Effekte und ihrer Richtungen ein Trend zu einer kostengünstigeren, zumindest aber gleich teuren Versorgung erkennen.

Gleiches gilt für die als Verhältnis von eingesparten vollstationären Behandlungstagen bzw. eingesparten AU-Tagen zu den Kosten von IG und KG ermittelten Effizienz. Anhand der durchgeführten

Analyse lässt sich bezogen auf das Vermeiden vollstationärer Behandlungstage für elf Modelle eine effektivere und davon für sieben Modelle zudem eine kostengünstigere Versorgung ablesen. Ein positives Outcome im Sinne vermiedener AU-Tage ergab sich für neun Modelle, wobei in sieben Modellen außerdem eine günstigere Versorgung stattfand. Mit um 82,69 € je vermiedenem vollstationären Behandlungstag eingesparten Kosten lässt anhand der vorläufigen Evaluationsergebnisse einen Trend zu einer kosteneffektiven Mittelverwendung erkennen.

3. Ergebnisse für die ersten drei Modelljahre über vier Modelle

Bei dem primären Zielkriterium kumulative stationäre Behandlungsdauer und Anzahl stationärer Aufenthalte aufgrund von Indexdiagnosen von linikneuen Patienten mit Referenzfall im ersten Modelljahr zeigte sich für alle vier Modelle ein vergleichbares Muster im Verlauf der durchschnittlichen vollstationären Behandlungsdauer zwischen der IG und der KG mit vergleichbar niedrigem Wert im Jahr vor Modelleinschluss (Prä-Wert), Peak im ersten patientenindividuellen Jahr und Stabilisierung auf einem niedrigen Niveau ab dem zweiten Jahr. Die Höhe des Peaks unterschied sich dabei zwischen IG und KG deutlich im ersten Jahr. Der Peak war in der IG (7,6 bis 19,1 vollstationäre Tage, Mittelwert 14,3 vollstationäre Tage) deutlich geringer als in der KG (21,1 bis 24,4 vollstationäre Tage, Mittelwert 22,7 vollstationäre Tage). Bei allen vier Modellen war das Muster der teilstationären Behandlungsdauer ähnlich dem der vollstationären Behandlungsdauer mit vergleichbar niedrigem Ausgangsniveau, Peak im ersten Modelljahr und Stabilisierung auf niedrigem Niveau ab dem zweiten Modelljahr. Jedoch war hier der Peak im ersten Jahr in der IG (8,3 bis 17,6 teilstationäre Tage, Mittelwert 12,5 teilstationäre Tage) deutlich höher als in der KG (4,6 bis 6,8 teilstationäre Tage, Mittelwert 6,0 teilstationäre Tage). Die Ergebnisse unterstreichen hier die Tendenzen für eine Bestätigung der Hypothese analog der ersten Berichte um eine Verringerung der vollstationären Behandlungstage. Weiterhin ist in den vier hier untersuchten Zweiten Zwischenberichten eine Verlagerung von vollstationären zu teilstationären Behandlungstagen ersichtlich.

Auch bei den linikbekannten Patienten zeigten sich in allen vier Modellen zwischen beiden Gruppen vergleichbare Muster mit einer Abnahme der durchschnittlichen vollstationären Behandlungsdauer über die Zeit. Dabei war die vollstationäre Behandlungsdauer über den Beobachtungszeitraum in der IG bei allen vier Modellen um durchschnittlich fünf Tage geringer als in der KG (2,9 bis 9,9 vollstationäre Tage). Das Muster der Dauer teilstationärer Aufenthalte war zwischen IG und KG bei allen Modellen vergleichbar. Jedoch wies ein Modell in der IG stets höhere Werte für teilstationäre Tage (durchschnittlich 6,5 teilstationäre Tage) im Vergleich zur KG auf mit einem Peak im ersten patientenindividuellen Jahr (9,6 mehr teilstationäre Tage in der IG im Vergleich zur KG).

Beim zweiten primären Zielkriterium Arbeitsunfähigkeit zeigten sich bei allen vier Modellen zwischen der IG und der KG ähnliche Muster. Das Muster der durchschnittlichen Anzahl an AU-Tagen aufgrund Indexdiagnosen pro linikneuem Patient war bei drei Modellen in beiden Gruppen vergleichbar mit niedrigem Ausgangsniveau im patientenindividuellen Prä-Wert, Peak im patientenindividuellen ersten Jahr und Stabilisierung auf einem Wert im Bereich des Ausgangsniveaus ab dem zweiten Jahr. Der Peak im ersten patientenindividuellen Jahr war in der IG bei zwei Modellen um ca. 13 bzw. 14 AU-Tage geringer im Vergleich zur KG und bei einem Modell in der IG um ca. 10 AU-Tage größer im Vergleich zur KG. Bei einem Modell war der Verlauf der AU-Tage über die Zeit fast identisch.

Bei den klimbekannt Patienten verliefen die Werte für AU-Tage zwischen IG und KG parallel, wobei bei zwei Modellen die Verläufe nahezu identisch waren, bei einem Modell die IG im Vergleich zur KG ca. 5 AU-Tage weniger im ersten patientenindividuellen Jahr aufwies und bei einem Modell die IG durchschnittlich 12 AU-Tage weniger als die KG hatte.

Bei den sekundären Zielkriterien zeigten sich für die Inanspruchnahme ambulanter Leistungen klimbekannt Patienten über die Beobachtungszeit in beiden Gruppen vergleichbare bzw. fast identische Muster einer leicht rückläufigen ambulanten Kontaktrate in der Psychiatrischen Institutsambulanz (PIA) aufgrund Indexdiagnose. Zu beachten gilt jedoch, dass die Intensität der PIA-Kontakte mit den vorliegenden Routinedaten nur bedingt einzuschätzen ist.

Auch das Muster einer über die Beobachtungszeit weitgehend unveränderten ambulanten Inanspruchnahme von Vertragsärzten für Psychiatrie bzw. Kinder- und Jugendpsychiatrie oder einem Psychotherapeuten war zwischen IG und KG vergleichbar. Dabei lag die IG hier im Durchschnitt aller Patienten in zwei Modellen mit ca. 1,5 bis 2 Kontakten unter dem Niveau der KG und ein Modell mit ca. 1,5 Kontakten über dem Niveau der KG. Bei einem Modell entwickelten sich die Verläufe der Kontakte bei einem anderen Vertragsarzt mit Indexdiagnose parallel zwischen IG und KG, jedoch hatte die IG in jedem untersuchten Jahr jeweils ca. 3 Kontakte pro Patient mehr im Vergleich zur KG.

Auch bei den klimbekannt Patienten zeigten sich in drei Modellen keine relevanten Unterschiede in der durchschnittlichen Anzahl der PIA-Kontakte mit nahezu gleichen durchschnittlichen PIA-Kontakten (zwei Modelle) über die Zeit bzw. leichtem Rückgang der Anzahl der PIA-Kontakte über die Zeit (ein Modell). Bei der ambulanten Inanspruchnahme von Vertragsärzten für Psychiatrie bzw. Kinder- und Jugendpsychiatrie oder einem Psychotherapeuten zeigten sich auch vergleichbare Muster über die Zeit zwischen IG und KG in allen vier Modellen. Dabei lag bei einem Modell die IG hier im Durchschnitt aller Patienten ca. 2 Kontakte pro Patient über dem Niveau der KG.

Bei dem Zielkriterium Behandlungskontinuität zeigten sich sowohl für klimbekannt als auch für klimbekannt Patienten bei allen Modellen keine deutlichen Veränderungen im Zeitverlauf und zwischen den Gruppen, abgesehen von Schwankungen innerhalb des Streubereichs. Bei einem Modell stieg die Kontinuität eines ambulanten Kontakts innerhalb von 30 und 90 Tagen nach Krankenhausaufenthalt bei klimbekannt Patienten. Dieser Anstieg verlief in der IG und der KG analog.

Vom Arzt- und Krankenhausshopping sind in allen vier untersuchten Modellen nur wenige betroffen. Die Muster verliefen zwischen IG und KG identisch mit wenig Veränderung über die Zeit. Auch das Zielkriterium Komorbidität war zwischen IG und KG bei allen vier Modellen vergleichbar bzw. nahezu identisch.

Bei dem Zielkriterium Progressionsrate zeigten sich bei allen untersuchten vier Modellen vergleichbare Muster. Dabei war bei einem Modell ein deutlicher Niveauunterschied mit geringerer Progression zu schweren bzw. rezidivierenden depressiven Störungen in der IG gegenüber der KG für klimbekannt Patienten festzustellen.

Bei den weiteren sekundären Parametern Wiederaufnahme (Zeitreihenbetrachtung nicht zweckmäßig), Mortalität (geringe absolute Werte) und Leitlinienadhärenz (oder leitlinien-nahe

Behandlung) (geringe Fallzahlen bzw. mit Routinedaten nur bedingt abbildbar) zeigten sich keine systematischen Unterschiede zwischen IG und KG.

Die Kosten der psychiatrischen Versorgung linikneuer Patienten zeichneten im Längsschnitt über vier Jahre (ein Jahr vor und drei Jahre nach Interventionsbeginn) folgendes Bild: sowohl in der IG als auch in der KG lagen die Kosten im Präzeitraum auf einem geringen Niveau, stiegen im ersten Jahr deutlich an und fielen anschließend auf ein auf ca. das Doppelte der Kosten des Präzeitraumes liegendes Niveau ab. Im gepoolten Mittel ergaben sich je Patient im Präzeitraum in der IG ($\bar{x}_{\text{gepoolt}} = 985,83 \text{ €}$; $\bar{x}_{\text{min}} = 644,22 \text{ €}$; $\bar{x}_{\text{max}} = 1400,58 \text{ €}$) um 221,73 € geringere psychiatrische Behandlungskosten gegenüber der KG ($\bar{x}_{\text{gepoolt}} = 1207,56 \text{ €}$; $\bar{x}_{\text{min}} = 767,39 \text{ €}$; $\bar{x}_{\text{max}} = 1456,93 \text{ €}$). Für das erste Jahr nach Interventionsbeginn stiegen die Kosten in der IG im Mittel auf 7.106,97 € ($\bar{x}_{\text{min}} = 5.652,95 \text{ €}$; $\bar{x}_{\text{max}} = 8.682,95 \text{ €}$) und in der KG auf 7.048,04 € ($\bar{x}_{\text{min}} = 6.111,56 \text{ €}$; $\bar{x}_{\text{max}} = 7.904,29 \text{ €}$).

Dem entgegengesetzt zeigte sich bei den linikbekannten Patienten ein paralleler Kostenverlauf mit einem bereits im Präzeitraum in beiden Gruppen hohen Ausgangsniveau, das im ersten Jahr nach Interventionsbeginn minimal anstieg und anschließend auf ca. 70% des Präniveau fiel. Dabei waren in der IG im gepoolten Mittel sowohl im Präzeitraum ($\bar{x}_{\text{IG}} = 5.508,80 \text{ €}$; $\bar{x}_{\text{KG}} = 5.973,52 \text{ €}$) als auch im ersten Jahr nach Interventionsbeginn ($\bar{x}_{\text{IG}} = 5.873,08 \text{ €}$; $\bar{x}_{\text{KG}} = 6.159,13 \text{ €}$) sowie in den zwei darauffolgenden Nachbeobachtungsjahren geringere psychiatrische Versorgungskosten zu verzeichnen.

Die Kosteneffektivität wurde für die vier Modelle der Zweiten Zwischenberichte für die Patienten, deren Interventionsbeginn im dritten Jahr nach Modellbeginn lag, ausgewiesen. Bezogen auf den Parameter der vollstationären Behandlungstage zeigte sich ein um 6,5 Tage geringerer Anstieg vollstationärer Tage in der IG gegenüber dem der KG. Zudem stiegen die Kosten der IG um 262,65 € weniger als die der KG. Somit ergab sich ein inkrementelles Kosteneffektivitätsverhältnis von -40,67 €, d. h. es wurden je zusätzlich vermiedenem vollstationären Behandlungstag 40,67 € eingespart.

4. Fazit

Diese Zusammenfassung gibt einen Überblick über die bisherigen Ergebnisse der Modelle im Rahmen der Evaluation EVA64 und liefert eine gesamtheitliche Bewertung hinsichtlich der erwarteten Effekte. Dabei ist zu beachten, dass hier zum einen auf Ergebnisse der ersten 13 von insgesamt 18 Ersten Zwischenberichte mit Fokus auf klinikneue Patienten zurückgegriffen wird, welche das erste Jahr nach Modellstart betrachten. Zum anderen werden vier von insgesamt 18 Zweiten Zwischenberichten zusammenfassend dargestellt. Daher können auf Basis der bisher vorliegenden Daten nur vorläufige Aussagen zu den Hypothesen getroffen werden. Für eine abschließende Bewertung ist die Betrachtung aller Modelle mit mehrjährigen Nachbeobachtungszeiträumen notwendig.

Bezüglich der Dauer stationärer Aufenthalte ist abzuleiten, dass es Hinweise zur Bestätigung der aufgestellten Hypothese einer verringerten Aufenthaltsdauer in Modellkliniken gegenüber der Kontrollgruppe gibt. Der gepoolte Schätzer bei der Betrachtung des jeweils ersten Jahres bei den 13 Ersten Zwischenberichten gibt an, dass bei den klinikneuen Patienten der Modellkliniken der Anstieg der vollstationären Behandlungstage im Schnitt ca. 5 Tage geringer ausfiel im Vergleich zu den jeweiligen Kontrollgruppen. Das beschriebene zeitliche Muster bei der Untersuchung der ersten drei Jahre nach Einschluss in die Evaluation weist darauf hin, dass die Einführung der Modellprojekte bei

den vier Modellvorhaben der Zweiten Zwischenberichte besonders im patientenindividuellen ersten Jahr nach Modellstart verkürzend auf die vollstationäre Behandlungsdauer gewirkt hat und stattdessen zu längeren Behandlungszeiten im teilstationären Bereich geführt haben. Unterschiede in der Inanspruchnahme ambulanter Kontakte zwischen beiden Gruppen können die gefundenen Effekte in der vollstationären Behandlungsdauer nicht bzw. nicht vollständig erklären. Weiterhin kann die Inanspruchnahme in der PIA mit Routinedaten nur bedingt abgebildet werden, da die Inhalte und Intensität der PIA-Kontakte mit den vorliegenden Daten nicht vollständig abbildbar sind. Es scheint also in den vier bisher untersuchten Modellen mit Zweiten Zwischenberichten in der Modellklinik vornehmlich eine Umsteuerung vom vollstationären auf den teilstationären Bereich zu geben. Zu beachten ist jedoch, dass die vier hier untersuchten Modellkliniken ausschließlich aus Schleswig-Holstein kommen und einen Vorläufervertrag aufweisen. Die Muster in den anderen Modellkliniken können abweichen und müssen noch untersucht werden.

Bezüglich der klinikbekannten Patienten, weist das bereits im Präzeitraum ersichtliche unterschiedliche Niveau der vollstationären und teilstationären Behandlungstage darauf hin, dass die Vorverträge in den vier hier beschriebenen Modellkliniken der Zweiten Zwischenberichte schon vor der Einführung des Modellprojektes nach §64b SGB V auf eine geringere vollstationäre und eine erhöhte teilstationäre Behandlungsdauer wirkten.

Beim Zielparame-ter Arbeitsunfähigkeit von erwerbstätigen Versicherten wurden für die einzelnen Modelle der 13 Ersten Zwischenberichte zumeist signifikante Effekte gefunden. Die Richtung der Effekte war dabei jedoch sehr unterschiedlich. Dabei wurde besonders bei Modellen mit Vorläufervertrag ein hypothesenkonformer Effekt im Sinne einer reduzierten Arbeitsunfähigkeitsdauer festgestellt. Dies kann möglicherweise daran liegen, dass Vorläufermodelle die Umstellung der Behandlung auf Modellbedingungen erleichtern, sodass Modelleffekte bei Kliniken mit Vorläufervertrag schneller zum Tragen kommen als bei Kliniken ohne Vorvertrag. Bei den vier Zweiten Zwischenberichten ließen sich auch Tendenzen in unterschiedliche Richtungen erkennen, so dass hier noch keine abschließende Beurteilung möglich ist. Für diesen Parameter muss einschränkend erwähnt werden, dass die Arbeitsfähigkeit und ggf. die Wiedereingliederung in den Arbeitsmarkt von psychisch Erkrankten nur bis zu einem gewissen Grad beeinflusst werden kann. Der Arbeitsmarkt für diese Patienten unterliegt vielen verschiedenen Faktoren, die durch eine veränderte Behandlung nur teilweise geleitet werden können.

Die weitgehende Vergleichbarkeit der Muster der weiteren Zielparame-ter zeigt, dass die Qualität der Behandlung zwischen IG und KG vergleichbar ist und somit trotz der Einsparung von vollstationären Behandlungstagen keine Verringerung der Qualität ersichtlich ist, soweit das mit Routinedaten zu beurteilen ist. Weiterhin ist jedoch anzumerken, dass die Abbildung einiger sekundärer Zielkriterien (z.B. Leitlinienadhärenz oder Progression der psychischen Erkrankung) mit Routinedaten nur eingeschränkt aussagekräftig sind. Studien mit direkter Datenerhebung am Patienten, wie bspw. „PsychCare“, werden bezüglich der Versorgungsqualität zusätzliche wichtige Erkenntnisse liefern.

Eine abschließende übergreifende Bewertung bezüglich der Kosten ist derzeit noch nicht möglich. So lassen sich neben der insgesamt hohen Heterogenität der Effekte auf Modellebene in einigen Modellvorhaben Preiseffekte erkennen, bei denen Ausgleichszahlungen aus dem Zeitraum vor Modellbeginn die Kosten des ersten Jahres beeinflussen. Das Herausrechnen der Ausgleichszahlungen ist jedoch unter der momentan verwendeten Methodik nicht möglich. Für die Abschlussberichte ist

eine Anpassung der Methodik geplant, um diese Verzerrungen adäquat zu berücksichtigen. Eine Aussage zur Kosteneffektivität ist für das erste Jahr nach Modellbeginn aus genannten Gründen ebenfalls nur sehr eingeschränkt möglich.

Bei den vier Schleswig-Holsteiner Modellen, für die auch die Kohorte der Patienten, deren Interventionsbeginn im dritten Modelljahr lag, berücksichtigt werden konnten, lassen sich allerdings etwas sicherere Aussagen treffen, da hier die ausgleichsinduzierten Preiseffekte nicht mehr in so großem Maß wirken bzw. über die Entgeltsystematik herausgerechnet werden können. So zeigten sich im gepoolten Mittel der vier Modelle eine kostengünstigere Versorgung und eine geringere Anzahl vollstationärer Behandlungstage und somit eine kosteneffektive Verwendung der Mittel. Es muss jedoch einschränkend erwähnt werden, dass zwar in allen vier Modellen weniger vollstationäre Behandlungstage anfielen, die Kosten jedoch heterogener verteilt waren: so ließen sich in zwei Modellen geringere und in zwei Modellen höhere Versorgungskosten verzeichnen. Da sich aber lediglich für die Modelle mit geringeren Kosten statistische Signifikanz ergab, kann jedoch insgesamt von einem Trend zur kostengünstigeren und kosteneffektiveren Versorgung ausgegangen werden.