

„Verbindlichkeit ist die Vo setzung für effektive Steu

Im Interview: Prof. Dr. Jonas Schreyögg, Gesundheitsökonom an der Universität Hamburg, wissenschaftlicher Direktor des für die Studie verantwortlichen Hamburg Center for Health Economics (HCHE) und stellvertretender Vorsitzender des Sachverständigenrates Gesundheit

Die Techniker Krankenkasse hat ihre Evaluation der TK HzV-Bund-Verträge in 13 Bundesländern Deutschlands vorgestellt und prompt entbrannte ein Deutungskampf. Krankenkassen und ärztliche Verbände lesen das Ergebnis als Beleg, dass Hausarztprogramme den Versicherten keinen Vorteil bringen. Der Mann, der die Studie verantwortet, widerspricht. Prof. Dr. Jonas Schreyögg, Direktor des für die Studie verantwortlichen Hamburg Center for Health Economics (HCHE), erklärt im Gespräch mit „Monitor Versorgungsforschung", warum die HzV im untersuchten Modell nicht nicht, sondern „nur bedingt" steuert, weshalb er sie in der derzeitigen Form für ein Übergangsphänomen hält und wie ein verbindliches Primärversorgungsmodell unser Gesundheitssystem umbauen könnte.

>> Herr Professor Schreyögg, der Bundesvorsitzende des Virchow-Bundes, Dr. Dirk Heinrich – dessen Verband seit 17 Jahren Kooperationspartner des baden-württembergischen HzV-Modells ist –, wirft Ihnen vor, Ihre Studie (1) bilde „nur die halbe Wahrheit" ab, weil sie die Versorgungsrealität in Baden-Württemberg bewusst ausblende. Was entgegnen Sie ihm?

„Die eigentliche Krux ist die Konstruktion der HzV, so wie sie angelegt ist."

Diesen Vorwurf weise ich zurück. Dass Baden-Württemberg nicht enthalten ist, war keine wissenschaftliche oder gar politische Entscheidung, sondern bedingt sich alleine durch die Aufgabenstellung, auf die wir uns als Evaluatoren beworben und den Zuschlag bekommen haben. Die TK hat uns beauftragt, ihren TK-HzV-Bundvertrag, den sie in 13 Bundesländern separat abgeschlossen hat, zu evaluieren. Wichtig ist dabei zu wissen, dass es sich beim TK-HzV-Bundvertrag in den 13 Bundesländern um eine einheitliche und deshalb vergleichbare Intervention handelt, die Gegenstand der von uns gewonnenen Evaluationsausschreibung war.

Und Baden-Württemberg gehört eben nicht dazu.

Nur ganz bedingt. Der Grund ist, dass es neben den TK-HzV-Bundverträgen Sonderkonstruktionen gibt – etwa einen Add-on-Vertrag in Mecklenburg-Vorpommern, eigene Verträge in Sachsen-Anhalt und eben auch in Baden-Württemberg. Der baden-württembergische Vertrag ist jedoch anders aufgebaut als der Bundvertrag und – das möchte ich hinzufügen – zugleich anders als der HzV-Vertrag der AOK Baden-Württemberg, auf den Dr. Heinrich rekurriert. Will heißen, aus methodischen Gründen hätte man die KV-Regionen Baden-Württemberg, Mecklenburg-Vorpommern und Sachsen-Anhalt auch gar nicht in einen Pool mit den Bund-Verträgen geben können, da die Verträge von der bundesweiten Struktur der TK-HzV-Bundverträge abweichen und somit eine andere Intervention darstellen. Wer das trotzdem zusammenwerfen würde, vergleicht Unvergleichbares. Einfach ausgedrückt: Damit werden schlichtweg Äpfel mit Birnen verglichen. Was im Übrigen beiden Modellen – weder der TK noch der AOK Baden-Württemberg – gerecht werden würde. Ich persönlich finde übrigens das HzV-Modell der AOK Baden-Württemberg sehr spannend – es ist quasi ein Innovationslabor, in dem man immer wieder neue Sachen ausprobiert. Gerade deshalb ist es nicht mit dem in der Versorgungspraxis anders aufgebauten TK-Bundvertrag vergleichbar.

raus-
erung"

Die TK verkündet als „klares Ergebnis", Teilnehmende würden „nicht besser durch Versorgung gelotst", hätten nicht weniger Facharztkontakte und Klinikaufenthalte und verursachten keine geringeren Ausgaben. Auf welche Ihrer Befunde stützen sich diese Aussagen – und tragen Ihre Daten diese?

Die HzV reduziert – wenn wir auf die untersuchte Gesamtpopulation schauen – im ambulanten Bereich die Hausarztkontakte ganz leicht. Zugleich sehen wir aber mehr spezialfachärztliche Kontakte; und im stationären Bereich keine signifikante Steuerungswirkung. Wir finden auch bei Arzneimitteln, Arbeitsunfähigkeit und Krankengeld keinen wesentlichen positiven Einfluss. Unter dem Strich und auf die untersuchte Gesamtpopulation bezogen führt die Teilnahme in sämtlichen Leistungsbereichen sogar zu gesteigerten Ausgaben. Der erwartete Mechanismus – weniger Hospitalisierungen, weniger ungesteuerte Spezialistenbesuche – greift in dieser Form nicht. Anders ist das aber bei Subpopulationen, in denen wir teils einen Steuerungseffekt sehen, vor allem in der Gruppe derer, die sich an den Überweisungsweg über die Hausärzte halten. Der Grund für die mangelnde Steuerungswirkung liegt mit hoher Wahrscheinlichkeit jedoch nicht bei den Hausärzten selbst, sondern in der Konstruktion des Modells, weshalb die HzV eben nur einen bedingten Steuerungseffekt erzeugt.

Genau hier setzt ein konzeptioneller Einwand an: Ihre Kontrollgruppe ist die Regelversorgung – aber die ist kein Nullzustand, denn die Grundsätze hausärztlicher Tätigkeit sind nach Paragraf 73 SGB V längst normiert. Ist damit das Ausbleiben einer Differenz womöglich kein Beleg für Wirkungslosigkeit, sondern ein Deckeneffekt in einem schon gut funktionierenden System?

Das ist eine berechtigte und wichtige Frage – sie führt zur konzeptionellen Schlüsselstelle der Evaluation. Unser Vergleichsmaßstab ist bewusst die Regelversorgung ohne HzV-Einschreibung, denn wir wollen ja den zusätzlichen, kausalen Effekt der Einschreibung isolieren. Wenn sich die Regelversorgung über die Jahre selbst verbessert – etwa weil sich Steuerung auch dort einspielt –, dann fangen wir das ab, weil wir die Unterschiede im Zeittrend zwischen beiden Gruppen betrachten. Ein reiner Deckeneffekt würde aber bedeuten, dass es nichts mehr zu steuern gibt. Das sehen wir nicht: Wir beobachten in der HzV en gros eben keinen Rückgang der Spezialistenkontakte ohne Überweisung und keinen Rückgang der stationären Aufnahmen. Das eigentliche Problem ist, dass die HzV den entscheidenden Hebel, die verbindliche Steuerung, gar nicht in der Hand hat.

Warum nicht?

Weil die Versicherten nicht verpflichtet sind, sich an den vorgesehenen Weg zu halten. Die Hausärzt:innen können nicht steuern, wenn die Patient:innen direkt zum Spezialisten gehen dürfen. Wir haben dazu eigens eine Grafik in den Bericht aufgenommen, die zeigt, wie hoch der Anteil derer ist, die sich tatsächlich an den Pfad halten – und der ist sehr begrenzt. Die eigentliche Krux ist also die Konstruktion der HzV, so wie sie angelegt ist. Das ist übrigens kein TK-Spezifikum, so ist die HzV in ganz Deutschland gebaut: Sie geht davon aus, dass sich die Versicherten freiwillig steuern lassen.

Sie werten aber auch mehr Spezialistenkontakte als ausgebliebene Steuerung. Wäre in einem gut funktionierenden Primärarztsystem

nicht gerade die geplante, koordinierte Überweisung ein Qualitätsmerkmal – und der Anstieg damit ein Erfolg statt ein Defizit?

Im Prinzip ja. Gesteuerte Überweisungen sind erwünscht, ungesteuerte Direktzugänge sind das Problem. Genau deshalb schauen wir nicht nur auf die Zahl der Kontakte, sondern auch darauf, ob sie über eine Überweisung zustande kommen. Und da sehen wir einen geringen Rückgang der ungesteuerten Spezialistenkontakte, aber insgesamt eine Zunahme der Spezialistenkontakte. Der internationale Vergleich ist hier aufschlussreich: In Frankreich, wo ein Spezialistenkontakt ohne Überweisung mit einem deutlich höheren Eigenanteil verbunden ist, sinkt die Zahl der Facharztkontakte. In den Niederlanden ist eine fachärztliche Konsultation ohne hausärztliche Überweisung praktisch nicht erstattungsfähig. Doch gibt es in der deutschen HzV weder eine diesbezügliche finanzielle Incentivierung noch einen Ausschluss der fachärztlichen Versorgung ohne Überweisung.

Sie betonen, als Wissenschaftler erhöben Sie keinen Anspruch, Ihre Erkenntnisse der TK-HzV-Evaluierung auf andere Verträge zu verallgemeinern. Was gilt?

Beides lässt sich auflösen. Schaut man die TK-Bundverträge im Detail an, sind deren Elemente relativ ähnlich zu sehr vielen anderen HzV-Verträgen, schlichtweg weil sie sich über die Jahre stark angeglichen haben. Das von uns untersuchte TK-HzV-Modell ist demnach kein singulärer Vertrag, sondern einer, wie ihn viele andere Kassen mit gleichen oder ähnlichen Leistungspartnern auch abgeschlossen haben, und insofern gut vergleichbar. Verallgemeinerbar ist deshalb der strukturelle Befund: Ohne Verbindlichkeit bleibt die Steuerung schwach.

„Man kann Verbindlichkeit auch als den wirklichen Game-Changer bezeichnen."

Bleiben wir bei der möglichen Übertragbarkeit der Erkenntnisse. Die TK-Klientel ist im Schnitt jünger und gesünder als die der GKV insgesamt. Misst Ihre Studie damit nicht ausgerechnet dort, wo koordinierte Hausarztversorgung am wenigsten zusätzlichen Nutzen stiften kann?

Das Durchschnittsalter der HzV-Teilnehmenden liegt bei knapp 50 Jahren. Ja, die TK hat im engeren Sinne keine repräsentative Bevölkerungsabdeckung – das wissen wir aus vielen Innovationsfondsprojekten. Doch ist die TK-Population durch das starke Wachstum der Kasse näher an einen repräsentativen Aufbau herangerückt. Die AOK bildet quasi das andere Ende des Spektrums; idealerweise müsste man also beide Populationen zusammenführen. Das wäre das Optimum, doch entwertet das unsere Evaluation nicht, weil wir nicht bei der Gesamtschau stehen geblieben sind. Wir gehen in die Tiefe – in einzelne Altersgruppen und in die chronisch Erkrankten – und sehen dort, dass die Steuerung beispielsweise bei Älteren eher greift als bei Jüngeren. Genau deshalb spreche ich von einem bedingten Steuerungseffekt. Eine Verbesserung der Versorgung bei zugleich geringeren Ausgaben zeigt sich übrigens klar erst bei den über 80-Jährigen.

Auch hier wird es sich um einen grundlegenden Modellfehler handeln: Die HzV war ursprünglich doch nicht als reines Gatekeeping-System gedacht, sondern als strukturiertes Chronic-Care-Programm. Das von Ihnen eingesetzte Evaluationsdesign unterstellt, dass Teilnehmende und Nicht-Teilnehmende nach dem Balancing vergleichbar sind. Sie räumen jedoch auch ein, dass Einstellung und Motivation nicht erfasst sind. Gesundheitsbewusstere Menschen schreiben sich häufiger ein – ein klassischer Selektionseffekt. Wie valide ist Ihr Difference-in-Differences-Ansatz dann überhaupt?

Genau diesen Selektionseffekt zu neutralisieren, ist eines der wichtigsten Ziele der Studie. Wir kombinieren von daher ein Entropy Balancing mit einer Difference-in-Differences-Analyse, konkret dem Stacked DiD-Ansatz. Das Entscheidende ist der Vorher-Nachher-Vergleich gegen die Kontrollgruppe: Wenn jemand sich einschreibt, weil er oder sie motivierter ist, dann war das auch schon vor Einschreibung so. Dann hat er oder sie auch schon vor der HzV besser mit seinem Arzt kooperiert. Genau das wird in dem Modell sichtbar, da wir die Differenz der Unterschiede zwischen beiden Gruppen über die Zeit miteinander vergleichen – eben die Differenz (der beiden Gruppen) in der Differenz (nachher zu vorher). War die Motivation vorher schon da, ist sie über den Vortrend erkennbar; entsteht sie erst durch die HzV, ist sie Teil der Intervention. Beides berücksichtigen wir. Bedingung sind parallele Trends vor der Intervention; das ist die Hauptannahme für das DiD-Modell. Wo die Trends nicht von selbst parallel laufen, erzwingen wir das über das Balancing – und wir prüfen zudem die Robustheit mit dem Test von Rambachan und Roth.

Sie haben damit faktisch eine Art „RCT light" gebaut. Wie groß ist die Datenbasis, und wo lag die methodische Hauptschwierigkeit?

Die TK ist mit über zwölf Millionen Versicherten die größte deutsche Krankenkasse; über eine Million der TK-Versicherten sind in die HzV-Bundverträge eingeschrieben, an denen 10.000 Ärztinnen und Ärzte beteiligt sind. Für die Evaluation konnten wir über eine halbe Million Neueinschreibungen aus den Jahren 2015 bis 2022 heranziehen. Und in der Kontrollgruppe stehen uns je nach Jahr mindestens acht Millionen Beobachtungen aus der Regelversorgung zur Verfügung. Die Hauptschwierigkeit ist die sequentielle Einschreibung: Versicherte treten Jahr für Jahr bei, nicht zu einem Stichtag. Für jeden Jahrgang braucht man passende Kontrollpersonen, und man darf dieselben nicht immer wieder verwenden, sonst entsteht ein Bias. Genau das verhindert der gestaffelte Schätzer. Dafür braucht man einen riesigen Datensatz – den haben wir. Auch haben wir sehr transparent gemacht, wie wir selektiert haben.

Wie aufwendig war die Evaluation?

Der Aufwand war erheblich: Ab der Datenbereitstellung und -auswertung haben wir mit mehreren Personen rund ein Jahr gearbeitet. Das ist keine Studie, die man in wenigen Wochen herunterrechnet – schon weil die Methodik mit sequentieller Einschreibung und parallelen Trends sehr datenhungrig und rechenintensiv ist.

Im deutschen Forschungsdatenzentrum, dem FDZ, lägen noch

mehr Daten. Hätte das Ihre Studie aussagekräftiger gemacht?

Theoretisch ja, praktisch sind wir davon weit entfernt. So wie das FDZ derzeit konstruiert ist, gilt nur eine 60-Tage-Nutzung – in dieser Taktung ließe sich eine solche Auswertung nicht stemmen. Hinzu kommt ein inhaltliches Problem: Um über das FDZ sauber zu evaluieren, müsste man alle Verträge aller Kassen anschauen und prüfen, welche wirklich identisch sind. Viele dieser Verträge sind öffentlich, aber eine Übersicht über alle HzV-Verträge in Deutschland kenne ich zumindest nicht. Wir haben deshalb im Anhang der Studie zumindest den TK-Bundvertrag mit dem Vertrag der AOK Baden-Württemberg verglichen. Oberflächlich wirken viele Elemente ähnlich; im Detail zeigen sich entscheidende Unterschiede, etwa bei der Vergütung und beim Facharztprogramm. Eine vollständige Vertragsübersicht aller HzV-Verträge wäre überfällig – sie würde die Transparenz deutlich erhöhen, denn rund um diese Verträge wird viel spekuliert. Für das FDZ bräuchte man aber zusätzlich die Zustimmung jeder Kasse, denn man muss ja die Kassen identifizieren, um den jeweiligen HzV-Vertrag zuzuordnen. Ergo ist eine Analyse über das FDZ nicht realistisch.

Es gibt 13 TK-Bundverträge in 14 KV-Bezirken, es gibt aber 17 KVen. Wären die drei Regionen ohne Bundvertrag nicht die perfekte Kontrollgruppe?

Es wäre verlockend, aber es funktioniert nicht. Denn auch diese Regionen haben eine Intervention und taugen damit nicht als unbeeinflusste Kontrolle. Unbeeinflusst ist allein die Regelversorgung – und die ist in den vier Regionen nichts anderes als in den Bundvertrags-Ländern. Sie sprechen damit aber auch einen der wichtigsten Punkte an: die Wahl und Definition der Kontrollgruppe. Es gibt hier drei Strategien: nur Nicht-HzV-Versicherte in Nicht-HzV-Praxen, Nicht-HzV-Versicherte in HzV-Praxen, oder eine Mischung aus beiden. Wir haben alle drei gerechnet. Die Unterschiede sind allerdings so marginal, dass man sie kaum sieht; bei manchen Indikatoren stimmt das Ergebnis sogar in der zweiten Nachkommastelle noch überein. Alle drei sind übrigens in der Studie bzw. im Studienanhang publiziert.

Dr. Albrecht Kloepfer vom iX-Media-Institut hat in einem Editorial die aktuelle Debatte umgedeutet: Die eigentliche Frage laute nicht „HzV ja oder nein“, sondern „Unter welchen Vertragsbedingungen entfaltet ein 73b-Modell Steuerungswirkung?“. Baden-Württemberg funktioniere nur, weil der verbindliche 73b-Hausarztvertrag durch freiwillige Facharztverträge nach 73c und 140a flankiert werde.

Ich teile zumindest die Stoßrichtung. Genau deshalb sage ich, beide Evaluationen sind komplementär. Ein wichtiger Unterschied zwischen dem TK-Bundvertrag und dem AOK-Vertrag in Baden-Württemberg liegt nicht im Hausarztvertragsmodell selbst, sondern im Facharztprogramm und der damit stärkeren Einbindung der spezialfachärztlichen Versorgung. Wir führen das in der Diskussion sogar konkret aus: Dass die baden-württembergische Evaluation eine Reduktion der Polymedikation findet und wir nicht, könnte gerade an dieser engeren Abstimmung zwischen Haus- und Facharztebene liegen. Unsere Studie ist kein Verdikt über die Idee der hausärztlichen Koordination, sondern ein Befund über eine konkrete Vertragsimplementierung ohne Verbindlichkeit. Genau diese strukturelle Erkenntnis – dass es auf die Verbindlichkeit und auf die intrasektorale Verzahnung ankommt – ist sicherlich verallgemeinerbar.



Prof. Dr. Jonas Schreyögg

ist seit 2010 Inhaber des Lehrstuhls für Management im Gesundheitswesen und seit 2011 Wissenschaftlicher Direktor des Hamburg Center for Health Economics (HCHE) der Universität Hamburg. Nach seinem Studium der Betriebs- und Volkswirtschaftslehre an der Technischen Universität Berlin hat er hier auch seine Promotion am Fachgebiet für Finanzwissenschaft und Gesundheitsökonomie (Prof. Dr. Klaus-Dirk Henke) zum Dr. rer. oec. abgeschlossen. Nach seiner Promotion war er als Postdoc in Stanford.

Ihre Studie wird politisch in beide Richtungen gezogen. Stört Sie das?

So ist das Geschäft, das muss man nüchtern sehen. Ich bemühe mich, die wissenschaftliche Sicht differenziert darzustellen, bevor das Ganze ins Postfaktische abgleitet. Kassen und Leistungserbringerverbände haben legitime Interessen und Agenden. Doch es ist die Aufgabe der Wissenschaft, das politisch Gewollte und das empirisch Belegte sauber zu trennen.

Wenn Sie die wichtigsten Erkenntnisse auf den Punkt bringen: Was folgt daraus für ein künftiges Primärversorgungssystem?

Die wichtigste Erkenntnis ist tatsächlich die schon im aktuellen Koalitionsvertrag erwähnte Verbindlichkeit. Durch sie steigt die Wahrscheinlichkeit eines Steuerungseffekts deutlich – bei der Versorgungsqualität wie bei der Wirtschaftlichkeit. Man kann Verbindlichkeit auch als den wirklichen Game-Changer bezeichnen. International ist das eindeutig: Ich kenne kein System der Welt, in dem gesagt wird: Ihr macht jetzt Patientensteuerung, aber für die Versicherten ist das freiwillig. Konkret heißt das für ein verpflichtendes Primärversorgungsmodell: Man muss die Verbindlichkeit herstellen, entweder über die verbindliche Regelung oder über finanzielle Anreize beim ungesteuerten Facharztbesuch, damit der Zugang zur spezialfachärztlichen Versorgung über den Hausarzt läuft. Der zweite Punkt ist, die chronisch Erkrankten und Älteren explizit in den Blick zu nehmen, etwa indem man Disease-Management-Programme enger anbindet und bei The-

men wie der PRISCUS-Liste ansetzt. Drittens braucht es eine stärkere Vernetzung im ambulanten Sektor – auch mit standardisiertem digitalem Befundaustausch und digitalen Überweisungen, die der Hausarzt nach Dringlichkeit priorisiert.

Sie verweisen in Ihrer Studie oft ins Ausland. Welche konkreten Bausteine welcher Länder würden Sie übernehmen?

Die von den Niederlanden, England, Dänemark oder Frankreich. Gemeinsam ist bei diesen funktionierenden Systemen die Verbindlichkeit: In den Niederlanden senkt die Registrierung beim Hausarzt die Zuzahlungen und erleichtert die Terminvergabe, während der Facharzt nur mit Überweisung zugänglich ist; in Dänemark und Norwegen ist der Zugang ähnlich an die Überweisung gekoppelt. Daraus folgt für mich eine Registrierungspflicht – die im SGB V ohnehin angelegt ist, aber nicht praktiziert wird. Sie verstärkt die Steuerung, verbessert die Arzt-Patienten-Kommunikation und reduziert Mehrfachinanspruchnahmen. Üblich ist in solchen Systemen außerdem, bestimmte Fachgruppen wie Gynäkologie, Pädiatrie oder Augenheilkunde vom Überweisungsgebot auszunehmen. Was die HzV selbst angeht, bin ich mir sicher: Die derzeitige HzV an sich ist ein Übergangsphänomen hin zu einem umfassenden Primärversorgungssystem.

Sie sagen bewusst „Primärversorgung", der aktuelle Koalitionsvertrag spricht von „Primärarztmodellen". Semantik oder ein inhaltlicher Dissens?

Das ist ein ganz wichtiger inhaltlicher Punkt, den man – denke ich – im Bundesgesundheitsministerium ähnlich sieht. Man nennt das Vorhaben dort „Primärversorgung", weil auch andere Berufsgruppen beteiligt sein werden. Progressive Gemeinschaftspraxen machen das ohnehin – es ist in Deutschland nur sehr heterogen verteilt. Ein modernes Primärversorgungsmodell denkt das Team mit, nicht nur den Arzt.

Die TK fordert eine standardisierte digitale Ersteinschätzung als Gatekeeping-Instrument. Kritiker:innen warnen, solche Triage-Systeme würden vor allem von Älteren, bildungsferneren und chronisch Kranken seltener genutzt – also gerade von den Bedürftigsten – und verschärften die soziale Ungleichheit. Wo beginnt für Sie die sinnvolle Telematik und wo die Gefahr?

Eine digitale Ersteinschätzung ist sinnvoll, aber wir dürfen sie nicht zu losgelöst von den Primärversorgern denken – genau das adressiert auch die Gerechtigkeitsfrage. Wir wissen: 90 Prozent der Menschen in Deutschland nutzen mindestens einen Kinder- und Jugend- bzw. Hausarzt; zählt man die mit, die Gynäkolog:innen in Hausarztfunktion nutzen, sind es rund 95 Prozent. Im Grunde hat damit faktisch fast jeder in Deutschland einen Primärversorger, an den er sich wenden kann. Würde man stattdessen ein anonymes System aufsetzen, sehe ich – Auslandsevaluationen zeigen das – die Gefahr einer zusätzlichen Inanspruchnahme; zudem bleiben gerade vulnerable Gruppen außen vor. Die 116117 hat für Notfälle ihre Berechtigung, ebenso die 112, aber für Nicht-Notfälle würde ich die Ersteinschätzung an den Primärversorger koppeln. In vielen Ländern, in denen ich bereits gelebt habe – etwa in den USA, England oder Singapur – erreicht der Erstanruf zunächst eine Community Nurse bei einem Primärversorger, die strukturiert abfragt und einschätzt, ob und mit welchem Level eine Inanspruchnahme nötig ist. Das lässt sich auch digital gestalten, aber bitte eingebettet in die bestehende Versorgungsbeziehung.

Ein Primärversorgungssystem braucht eine Steuerungsinstanz, die haus- und fachärztliche Versorgung verbindlich integriert. In Baden-Württemberg leistet das eine Kasse mit hohem Marktanteil. Wer soll das bundesweit tun – die KBV, der G-BA? Und was wird aus dem Kontrahierungszwang, der die HzV in strukturschwachen Regionen heute an Hausärzt:innen bindet?

Ich würde das nicht zentralistisch in einer Behörde verorten. Mein Vorschlag ist ein verbindliches Basis-Primärversorgungsmodell für alle – für alle Kassen, alle Versicherten und alle Hausärztinnen und Hausärzte –, das aber darauf aufbauend wettbewerblich und vertraglich ausgestaltet wird, so wie wir es heute in der HzV sehen. Vertragspartner können, wie bisher in der HzV, sowohl für das Basismodell als auch für die HzV je nach Region der Hausärzteverband oder die KV sein. Es hat sich bewährt, dass man regional differenzieren kann, weil Regionen nun einmal sehr unterschiedlich aufgestellt sind. Wichtig ist mir: Man sollte die etablierten Strukturen nicht einreißen, sondern weiterführen. Ein abrupter Wegfall von Bindungsinstrumenten in unterversorgten Gebieten wäre kontraproduktiv – die Reihenfolge

Wie misst man Wirkung ohne randomisiertes Experiment? DiD mit Entropy Balancing

Eine Differenz-von-Differenzen-Analyse (DiD) vergleicht nicht einfach Eingeschriebene mit Nicht-Eingeschriebenen, sondern deren Veränderung über die Zeit: Wie entwickelt sich die Interventionsgruppe vor und nach dem Beitritt – relativ zur zeitlichen Entwicklung der Kontrollgruppe? Zeitkonstante Unterschiede lassen sich so herausrechnen. Das Problem: In die HzV schreiben sich Versicherte Jahr für Jahr ein (gestaffelte Einführung). Klassische Modelle verzerren die Schätzung dann, weil sie bereits Behandelte fälschlich als Kontrollen heranziehen. Der Stacked DiD-Ansatz vermeidet das: Er ermittelt für jede Einschreibekohorte und jedes Jahr einen eigenen, sauberen Effekt und nutzt nur Personen, die nicht der Intervention ausgesetzt sind, für die Kontrollgruppe. Vorgeschaltet ist ein Entropy Balancing – ein Gewichtungsverfahren, das die Kontrollgruppe so austariert, dass ihre messbaren Merkmale (Alter, Vorerkrankungen regionale Merkmale) exakt denen der Eingeschriebenen entsprechen – man nennt das auch synthetische Kontrollgruppe. Erst dadurch werden bei bestimmten Indikatoren die Verläufe vor dem Beitritt vergleichbar.

Kann man der Annahme trauen? Der Rambachan-&-Roth-Test

Jede DiD-Analyse steht und fällt mit der Annahme „paralleler Trends": Ohne die Intervention hätten sich beide Gruppen gleich weiterentwickelt. Beweisen lässt sich das nie, denn die hypothetische Entwicklung der Eingeschriebenen ohne HzV ist nicht beobachtbar. Prüfbar ist allein die Zeit davor – der Vortrend. Hier setzt das Verfahren von Rambachan und Roth (2023) an. Statt zu unterstellen, die Trends seien perfekt parallel, fragt es: Wie stark dürfte diese Annahme verletzt sein, bevor das Ergebnis kippt? Aus den beobachteten Abweichungen vor dem Beitritt wird ein Spielraum möglicher Verzerrungen danach abgeleitet. Bleibt der Effekt auch unter plausiblen Verletzungen bestehen, gilt er als robust. Die Kennzahl M^- benennt, bis zu welchem Ausmaß einer Trendverletzung der Effekt noch standhält; fällt sie zu niedrig aus, ist er nicht belastbar. In der TK-HzV-Studie bestehen diese Prüfung die allermeisten Indikatoren. Wenn die Gruppen aber kleiner werden, z. B. bei den Spezialistenkontakten der adhärenten Gruppe ($M^- = 0{,}79$), kann auch ein Indikator diesen Test nicht bestehen, dann entsteht kein robuster Effekt und man kann keine Aussage ableiten.

Inhaltlicher Hinweis zur Einordnung: Das M^- (Mbar) ist in der Methode die „Breakdown"-Schwelle – das Ausmaß einer Trendverletzung relativ zur größten beobachteten Vortrend-Abweichung, bis zu dem der Effekt signifikant bleibt. Die Studie wertet Werte unterhalb ihres geforderten Niveaus als „kein robuster Effekt".

Zitationshinweis

Schreyögg, J., Stegmaier, P.: „Verbindlichkeit ist die Voraussetzung für effektive Steuerung“, in „Monitor Versorgungsforschung“ (04/26), S. 12–17. http://doi.org/10.24945/MVF.04.26.1866-0533.2841

muss stimmen: erst die verbindliche Basis bauen, dann Bewährtes überführen.

Damit sind wir beim Übergangsrisiko: Strukturen werden in Deutschland leichter abgebaut als aufgebaut. Wie verhindern Sie ein Steuerungsvakuum, wenn die HzV verschwindet, bevor das Nachfolgesystem steht – ausgerechnet zulasten der Älteren und chronisch Kranken, für die Sie positive Effekte messen?

Indem man eben kein Parallelsystem aufbaut und keine Lücke zulässt. Wir können den Versicherten nicht sagen: Alle, die nicht in der HzV sind, sind in einer verbindlichen Primärversorgung und müssen den Versorgungsweg einhalten – und in der HzV kann alles so weitergehen wie bisher. So ist es im Koalitionsvertrag auch nicht angedacht. Es kann nur so sein, dass man ein Basis-Primärversorgungsmodell hat, das von Anfang an für alle gilt. Darauf ließe sich dann eine freiwillige HzV oben aufsetzen, etwa wenn eine Kasse ein Facharztprogramm selektivvertraglich oder Elemente für bestimmte Patientengruppen ergänzen will, weil es den Effekt verstärkt und sich so einen Vorteil im Wettbewerb verschaffen kann. Der Übergang ist dann kein Sprung ins Leere, sondern eine Überführung – die bestehenden Verträge und Beziehungen bleiben das Gerüst.

Wo fühlen Sie sich derzeit am gründlichsten missverstanden?

Am stärksten dort, wo behauptet wird, unsere Studie zeige, die HzV sei überflüssig. Das Gegenteil ist der Fall. Ich sehe sie als Stärkung eines künftigen Primärversorgungsmodells, so wie es im Koalitionsvertrag steht. Der zweite Punkt ist eine verbreitete Angst der Hausärzt:innen und ihrer Verbände, dass ein Primärversorgungsmodell alles in den Kollektivvertrag überführen würde. Ein Primärversorgungsmodell als Basis ist viel mehr als eine HzV – es ist verbindlich und hat viel mehr Elemente. Und Vertragspartner können aus meiner Sicht beide sein: die KVen oder der Hausärzteverband.

Wenn Sie das eben Gesagte auf einen Satz verdichten müssten: Was ist die Bedingung, ohne die kein Primärversorgungssystem in Deutschland gelingen wird?

Sinnvoll wird es sein, dass man ein Basis-Primärversorgungsmodell für alle einführt. Wenn ich es auf ein Wort zuspitzen soll: Dieses eine Wort – „verpflichtend“ – ist sehr ausschlaggebend. Erst die Verbindlichkeit macht aus einem gut gemeinten Programm ein System, das die Versorgung verbessert und gleichzeitig die Leistungsausgaben senken kann. So wie wir es heute betreiben, ist die Steuerung de facto in das Belieben der Versicherten gestellt – das ist ökonomisch ineffizient und für die Hausärztinnen und Hausärzte unbefriedigend, weil sie einen Steuerungsauftrag haben, den sie gar nicht durchsetzen können. Mein Petitum: Bauen wir eine verbindliche Basis und setzen darauf wettbewerbliche Elemente, bewahren wir die Erfahrungen aus der HzV, auch aus Baden-Württemberg, ohne ihre Konstruktionsfehler fortzuschreiben. Das wäre für mich der eigentliche Lernertrag unserer Evaluation.

Herr Professor Schreyögg, danke für das Gespräch. <<

Das Interview führte MVF-Chefredakteur Peter Stegmaier.

1: https://t1p.de/m0a9b

Lungenkrebs in zertifizierten Zentren behandeln

>> Wer Lungenkrebs überlebt, hängt nicht nur von der Diagnose ab – sondern auch davon, wo man behandelt wird. Zu diesem Schluss kommt eine aktuelle Studie einer Autorengruppe um PD Dr. Thomas Elter (Uniklinik Köln). Die jüngst im Elsevier-Journal „Lung Cancer“ publizierte Studie (1) basiert auf Abrechnungsdaten der BARMER und analysiert den Versorgungsalltag von mehr als 8.000 Patient:innen mit nicht-kleinzelligem Lungenkarzinom. Die Ergebnisse sind ernüchternd: Leitliniengerechte Behandlung ist keine Selbstverständlichkeit. Und wie auch schon in der WiZen-Studie von Prof. Dr. med. Jochen Schmitt, MPH (Universität Dresden), gezeigt: Der Unterschied zwischen spezialisierten Zentren und nicht-zertifizierten Kliniken kann buchstäblich lebensentscheidend sein.

Die aktuelle Studie untersuchte zwei Patientengruppen getrennt voneinander. Die erste umfasste ca. 900 Patient:innen in frühen Krankheitsstadien, die chirurgisch behandelt wurden. Als Qualitätsmerkmal galt hier, ob eine angemessene bildgebende Diagnostik – konkret ein FDG-PET/CT oder eine Kombination aus Szintigrafie und CT bzw. MRT – innerhalb von 42 Tagen vor dem Eingriff stattgefunden hatte. Das Ergebnis: Nur gut 61% der Patient:innen erhielten diese zeitgerechte Bildgebung. Jene, bei denen sie durchgeführt wurde, lebten statistisch deutlich länger als jene ohne adäquate voroperative Diagnostik. Besonders ausgeprägt war der Überlebensvorteil bei Patient:innen, die spezifisch ein FDG-PET/CT erhalten hatten, gegenüber denen, bei denen lediglich Szintigrafie mit CT oder MRT zum Einsatz kam.

Zertifizierte Lungenkrebszentren schnitten in einem wichtigen Punkt besser ab: Ihre Patient:innen bekamen häufiger das als Goldstandard geltende FDG-PET/CT als Patient:innen in nicht-zertifizierten Einrichtungen. Auf das Gesamtüberleben hatte die Zertifizierung des Behandlungsorts jedoch keinen direkt nachweisbaren eigenständigen Effekt – der Überlebensvorteil war vielmehr an die Qualität der Diagnostik selbst geknüpft, nicht allein an den Ort ihrer Durchführung.

Noch deutlichere Unterschiede traten in der zweiten Patientengruppe zutage: bei rund 1.600 Patient:innen mit fortgeschrittenem Lungenkarzinom, die ausschließlich systemisch behandelt wurden. Für sie ist die molekulare Biomarkertestung entscheidend, denn nur wer die passenden Genmutationen aufweist, kommt für zielgerichtete Therapien infrage – Behandlungen, die das Überleben im Vergleich zu nicht-zielgerichteten Therapien nach den Studiendaten nahezu halbieren können. Das Problem: Die Testrate lag in nicht-spezialisierten Versorgungsbereichen dramatisch unter der der Teilnehmenden am Nationalen Netzwerk Genomische Medizin Lungenkarzinom, kurz nNGM. Während im nNGM-Programm fast alle Patient:innen getestet wurden, waren es im übrigen Versorgungsbereich höchstens rund drei Viertel.

Die Autor:innen betonen, dass Abrechnungsdaten zwar keine vollständige klinische Information liefern und die beobachteten Zusammenhänge nicht als Kausalbeweise gelten können. Dennoch bieten sie einen unverstellten Blick auf die Versorgungsrealität jenseits der großen Studienzentren. Und dieser Blick zeigt: Während die Qualität der Lungenkrebsversorgung in Deutschland insgesamt gestiegen ist, bleibt die Lücke zwischen Spitzenzentren und der Regelversorgung besorgniserregend groß. Die Schlussfolgerung der Studienautor:innen fällt entsprechend klar aus: Lungenkrebspatient:innen sollten ausschließlich in zertifizierten Zentren behandelt werden. <<

1: https://t1p.de/hm55e