

Probleme der Evaluation einer regionalen integrierten Vollversorgung am Beispiel *Gesundes Kinzigtal*

Achim Siegel (1), Ulrich Stößel (1), Ingrid Schubert (2), Antje Erler (3)

(1) Evaluations-Koordinierungsstelle Integrierte Versorgung, Abt. f. Medizinische Soziologie, Albert-Ludwigs-Universität Freiburg, Freiburg im Breisgau, Deutschland

(2) PMV forschungsgruppe, Universität zu Köln, Köln, Deutschland

(3) Institut für Allgemeinmedizin, Johann-Wolfgang-Goethe-Universität Frankfurt, Frankfurt am Main, Deutschland

Korrespondierender Autor: Dr. Achim Siegel, M.P.H., Evaluations-Koordinierungsstelle Integrierte Versorgung, Abt. f. Medizinische Soziologie, Universität Freiburg, Hebelstr. 29, 79104 Freiburg. Tel. +49761203-5518 oder -5530, Email: achim.siegel@medsoz.uni-freiburg.de

Danksagung: Wir danken der Gesundes Kinzigtal GmbH, der AOK Baden-Württemberg und der LKK Baden-Württemberg für gute Kooperation und finanzielle Unterstützung der Evaluations-Koordinierungsstelle Integrierte Versorgung (EKIV).

Keywords:

Deutsch: Evaluation, Evaluationsmethodik, Integrierte Versorgung, Managed Care

Englisch: evaluation, evaluation methodology, integrated health care systems, managed care

Zusammenfassung

Die Integrierte Versorgung Gesundes Kinzigtal (IVGK) ist eines der wenigen integrierten Vollversorgungssysteme in Deutschland. Die externe Evaluation ist modular aufgebaut und wird von der Abteilung für Medizinische Soziologie an der Universität Freiburg koordiniert. Im folgenden Beitrag wird ein Überblick über die laufenden Evaluationsprojekte gegeben, und es werden exemplarisch fünf methodisch relevante Probleme erörtert, die bei der Evaluation regionaler Versorgungssysteme typischerweise auftreten. Diese bestehen in einer anfangs geringen Größe der Interventionspopulation, die eine Analyse auf Basis epidemiologischer Maßzahlen (wie z.B. Inzidenzen) erschwert. Zudem können wiederholte Befragungen zu einer „Überforschung“ der Interventionspopulation und damit zu abnehmender Beteiligung führen. Dem Problem der geringen Größe der Interventionspopulation kann oft mit einer Vollerhebung begegnet werden; um das Problem der Überforschung zu vermeiden, ist der Einsatz möglichst kurzer Erhebungsinstrumente, die zeitversetzte Koordination verschiedener Primärdatenerhebungen und die Nutzung von Sekundärdaten (z.B. GKV-Routinedaten) zu empfehlen. Ein weiteres typisches Problem bei der Evaluation kleinräumiger Versorgungssysteme besteht in der Zuordnung und Messung der Einzeleffekte von Gesundheits- bzw. Krankheitsmanagementprogrammen: Da ein relativ großer Teil der Versicherten an mehr als nur einem Programm teilnimmt, ist die Interventionsbedingung bei der Evaluation eines einzelnen Programms durch multiple Programmeffekte kontaminiert, was bei der Auswertung und ggf. der Interpretation der Ergebnisse berücksichtigt werden muss. Schließlich lassen sich gerade bei der Evaluation regionaler Versorgungssysteme nicht immer randomisierte Studiendesigns umsetzen oder durchsetzen. Dies führt dazu, dass die Evaluationsergebnisse eine (im Vergleich zu Ergebnissen aus randomisierten Studien) geringere Beweiskraft haben.

Abstract (engl.)

Gesundes Kinzigtal Integrated Care (GKIC) is one of the few population-based integrated care systems in Germany. Its external evaluation is coordinated by the

Department of Medical Sociology at Freiburg University and has a modular design. In this article we outline the different evaluation modules and address typical problems which arise in the evaluation of population-based integrated health care systems. One of these problems consists in the small size of the intervention population in a recently established pilot care system which makes the use of epidemiologic measures (such as e.g. incidence figures) difficult. Total population surveys could alleviate this problem, but when the same intervention population is questioned repeatedly for different evaluation modules, the danger of 'over-surveying' arises. This may lead to a decrease in participation in further studies. These problems may be managed by using *short* survey instruments, by conducting surveys successively (not simultaneously) using time shifts, and by analysing claims data from health insurers. A further problem in the evaluation of comparably small systems of care is how to attribute outcomes (effects) to certain health promotion or disease management programmes: As participants usually take part in more than one programme, the intervention effect is contaminated by multiple programme effects. These will have to be adjusted for in the final evaluation of the data. Finally, as a randomised controlled study (RCT) design frequently cannot be used in the evaluation of a region-wide pilot intervention, the evidence generated by such a (non-randomised) study is weaker than the evidence generated by an RCT.

1. Grundzüge des integrierten Vollversorgungssystems Gesundes Kinzigtal (IVGK) und der Stellenwert der Evaluation

Integrierte Versorgungssysteme verfolgen den Anspruch, die Schnittstellenprobleme der in Sektoren gegliederten Normalversorgung zu überwinden. Dabei wird vor allem versucht, die Kooperation der verschiedenen Leistungserbringer zu verbessern – sowohl innerhalb einer Disziplin und eines Sektors als auch disziplin- und sektorübergreifend. Davon erwartet man sich eine höhere Versorgungsqualität. Ein Hauptziel integrierter Versorgungssysteme ist jedoch auch, ökonomischer zu wirtschaften als die Normalversorgung, d.h. die Versorgungseffizienz zu erhöhen.

Integrierte Versorgungsverträge wurden in Deutschland von 2004 bis 2008 durch eine Anschubfinanzierung aus Mitteln der Regelversorgung gesetzlich gefördert. In diesem Zeitraum entstanden insgesamt 6407 Integrationsverträge [1]. Hierbei überwogen deutlich *indikationsbezogene* Verträge, die sich nur auf eine oder einige wenige Indikationen konzentrierten, wie z.B. Hüft- oder Kniegelenk-*endoprothetik* oder kardiologische Eingriffe mit anschließenden Rehabilitationsmaßnahmen und ggf. ambulanter Nachsorge. Nur wenige der neu entstandenen Integrationsverträge hatten den *Aufbau integrierter Vollversorgungssysteme* (synonym: *populationsbezogene integrierte Versorgungssysteme*) zum Ziel, die (fast) alle Indikationen einer definierten Wohnbevölkerung einbeziehen [2].

Die Integrierte Versorgung *Gesundes Kinzigtal* (IVGK) in Südbaden ist eines der wenigen integrierten Vollversorgungssysteme in Deutschland (vgl. den Beitrag von Hildebrandt et al. in diesem Heft). In klassischer Weise verfolgt die IVGK ein doppeltes Ziel: Zum einen soll die Versorgungsqualität verbessert, zum anderen ökonomischer gewirtschaftet werden als in der Normalversorgung. Die Managementgesellschaft *Gesundes Kinzigtal GmbH* hat dabei mit den Vertragspartnern AOK und LKK Baden-Württemberg ein komparatives Einspar-Contracting vereinbart. Das bedeutet: Die Managementgesellschaft finanziert sich und ihre Aktivitäten nach Ende der Anschubfinanzierung aus den Einsparungen, die im Kinzigtal im Vergleich zur Normalversorgung erzielt werden [3]. Diese sollen prinzipiell aus drei Quellen entstehen:

1. einer besseren Schnittstellenorganisation zwischen verschiedenen Leistungserbringern und Sektoren (vgl. z. B. [4]),
2. einer Verringerung der Morbidität durch gezielte Prävention und aus
3. einem günstigeren Einkauf externer Produkte (z.B. Medikamente).

Die Vertragspartner haben verschiedene Vorkehrungen dafür getroffen, dass die Versorgungsqualität *nicht* zugunsten wirtschaftlicher Interessen vernachlässigt wird:

Zum ersten wurde das in der ambulanten ärztlichen Versorgung geltende Vergütungssystem (Einzelleistungsvergütung mit Budgetobergrenzen) bewusst belassen, d.h. für die einzelnen Leistungserbringer wurde *keine* Capitation-Vergütung etabliert [4,5].

Zum zweiten wurde der Versorgungsvertrag so gestaltet, dass weder für die Managementgesellschaft noch für die Leistungserbringer Anreize entstehen, die eine Risikoselektion oder eine Unterversorgung mit medizinisch-therapeutisch notwendigen Leistungen zur Folge haben könnten [5,6].

Zum dritten soll eine modular aufgebaute Evaluation, die von der Abteilung für Medizinische Soziologie an der Universität Freiburg inhaltlich vorstrukturiert und koordiniert wird, verschiedene Aspekte des Versorgungssystems begleitend untersuchen. Eine Übersicht über die derzeit laufenden Studien zeigt Tabelle 1; für weiterführende Informationen siehe <http://www.ekiv.org/de/index.php>.

[hier Tabelle 1 einfügen]

2. Methodisch relevante Probleme der Evaluation kleinräumiger Versorgungssysteme am Beispiel der IVGK

Die Evaluation von Versorgungsmodellen wie der IVGK wirft eine Reihe von methodisch relevanten Problemen auf, von denen die wichtigsten im Folgenden dargestellt werden sollen. In Abschnitt 2.1 wird ein Problem erläutert, das für die Evaluation kleinräumiger Versorgungssysteme von zentraler Bedeutung ist, nämlich die zumindest in der Aufbauphase eines Systems geringe Interventionspopulation. Aus diesem Problem ergeben sich drei weitere Probleme, die in den Abschnitten 2.2 bis 2.4 skizziert werden. Abschließend (Abschnitt 2.5) wird das methodologisch wohl

schwerwiegendste Problem beschrieben, nämlich das Problem der Nicht-Umsetzbarkeit bzw. Nicht-Durchsetzbarkeit eines randomisierten Studiendesigns.

2.1. Die kritische Größe der Interventionspopulation

Betrachtet man das Einzugsgebiet der IVGK genauer, dann wird auch gleich das erste größere Problem klar, das sich bei der Evaluation stellt: die kritische Größe der Interventionspopulation. Im Einzugsbereich der IVGK leben etwa 60.000 Menschen; etwa die Hälfte davon sind bei der AOK Baden-Württemberg oder der LKK Baden-Württemberg versichert, nämlich rund 31.000 Personen. Diese Größe muss als untere Grenze betrachtet werden, ab der eine Evaluation der Versorgungsqualität überhaupt sinnvoll ist, sofern man dabei auch die Wirksamkeit von Präventionsprogrammen auf Basis epidemiologischer Maßzahlen evaluieren möchte: Folgt man gängigen Inzidenzschätzungen [5,7], so sind im Kinzigtal bei ca. 31.000 Versicherten pro Jahr u. a. folgende Inzidenzen zu erwarten:

- 24 tödlich verlaufende Herzinfarkte¹,
- 55 ischämische Schlaganfälle²,
- 19 Lungenkrebsfälle,
- 45 Schenkelhalsfrakturen aufgrund einer Osteoporose-Erkrankung,
- 4 Diabetesfälle mit Dialysepflicht,
- 12 Amputationen wegen eines diabetischen Fußsyndroms,
- 4 Erblindungen aufgrund einer Diabetes-Erkrankung.

Bei einem Versorgungssystem, das der Verringerung der Morbidität durch Prävention einen so großen Stellenwert einräumt wie die IVGK, sind diese Zahlen als untere Grenze für eine epidemiologisch sinnvolle Evaluation anzusehen, damit eventuelle Verringerungen der Inzidenzen – die ggf. einen Erfolg des Versorgungskonzepts anzeigen – als statistisch zuverlässig beurteilt werden können. Untersuchte man hier nur Stichproben von 3.000, 5.000 oder 10.000 Versicherten, so bekäme man zu geringe Häufigkeiten, um stabile Schätzungen vorzunehmen. Sinnvollerweise sind hier also von vornherein *Vollerhebungen* anzustreben.

¹ Diese Schätzung basiert auf Angaben des Statistischen Bundesamts zur Mortalität nach Herzinfarkt aus dem Jahr 2001 (65.228 Verstorbene nach Herzinfarkt in Deutschland). Da die Herzinfarktinzidenz und -mortalität beträchtlich nach Regionen variiert und in Baden-Württemberg unterdurchschnittlich ist, liegt eine realistische Schätzung für das Kinzigtal unterhalb von 24 Fällen pro Jahr.

² Laut Kolominisky-Rabas und Heuschmann [8] kann die Inzidenz des ischämischen Schlaganfalls in Deutschland auf 182/100.000 Einwohner geschätzt werden.

2.2. Subpopulationen innerhalb der Interventionspopulation, die jeweils einer unterschiedlichen Interventionsintensität unterliegen

Das Problem der kritischen Größe der Interventionspopulation wird noch komplizierter, wenn man sich fragt: Wer bildet im Kinzigtal die Interventionspopulation? In die IVGK eingeschrieben und damit aktive Teilnehmer integrierter Versorgungs- und Präventionsprogramme waren bis zum 31.12.2007 etwas mehr als 2.000 Personen. Gegenwärtig (26. September 2010) sind es 7.059 Personen. Zwar sind unter den IV-Eingeschriebenen kranke oder krankheitsgefährdete Personen deutlich überproportional vertreten [6]; dennoch ist der aus praktischen Gründen relativ langsam voranschreitende Einschreibungsprozess ein limitierender Faktor für epidemiologisch orientierte Evaluationsstudien, *sofern man sich lediglich auf die IV-eingeschriebenen Versicherten beschränkt.*

Das Problem wird allerdings durch die folgenden beiden Aspekte relativiert: Zum einen sind etwas mehr als 60% der niedergelassenen Ärzte im Einzugsbereich der IVGK Vertragspartner. Wissenszugewinne dieser Netzärzte und die Vorteile, die sich aufgrund der intensiveren Kooperation mit anderen Netzärzten und Institutionen ergeben, werden auch denjenigen Patienten zugute kommen, die sich nicht in die IVGK eingeschrieben haben. Die IVGK hat damit auch einen mittelbaren Einfluss auf die Versorgung vieler *nicht* eingeschriebener Versicherter (wobei bestimmte Präventionsangebote den eingeschriebenen Versicherten vorbehalten bleiben). So profitieren von insgesamt 31.000 im Kinzigtal wohnenden AOK- und LKK-Versicherten schätzungsweise ca. 19.000 Versicherte mindestens indirekt von der IVGK. Da rund 60% der niedergelassenen Ärzte im Kinzigtal IVGK-Partner sind, kann in einer ersten Schätzung davon ausgegangen werden, dass ein ebenso hoher Anteil aller AOK- und LKK-Versicherten in der Region – d.h. rund 19.000 Versicherte – von diesen Ärzten betreut werden und somit einen mindestens indirekten Nutzen von der IVGK haben. (Einen indirekten Nutzen haben im Übrigen auch Versicherte anderer Kassen, die von IVGK-Partnern betreut werden, denn es muss davon ausgegangen werden, dass die IVGK-Partnerärzte ihre neuen, in Zusammenarbeit mit der IVGK erworbenen Fach- und Koordinationskompetenzen grundsätzlich allen ihren Patienten zugute kommen lassen.)

Zum anderen übernimmt die Gesundes Kinzigtal GmbH eine Budgetmitverantwortung für *alle* im Kinzigtal wohnenden AOK- und LKK-Versicherten, unabhängig davon, ob diese in die IVGK eingeschrieben sind und unabhängig davon, ob sie von Netzärzten behandelt werden oder nicht. Im Hinblick auf die Frage der

Budgetverantwortung sind deshalb *alle* 31.000 im Kinzigtal wohnenden AOK- und LKK-Versicherten als Interventionspopulation anzusehen, von der jedoch nur eine Teilmenge von den Behandlungsprinzipien und Programmen der IVGK erreicht wird. Die Gesamtzahl der im Kinzigtal wohnenden AOK- und LKK-Versicherten bildet also gleichsam eine Art „intention-to-treat“-Population für die IVGK, die allerdings nie zur Gänze erreicht werden wird.

Hinsichtlich der Evaluation eines solchen Systems sollte daher folgende Konsequenz gezogen werden: Im Idealfall sollten in jeder Evaluationsstudie die Daten *aller* Subpopulationen der Interventionspopulation erhoben und je nach Interventionsintensität abgestuft analysiert werden. Im Fall der IVGK-Evaluation erhält man im Idealfall also differenzierte Ergebnisse für folgende drei Subpopulationen:

- (a) die in die IVGK eingeschriebenen Versicherten,
- (b) die *nicht* in die IVGK eingeschriebenen, aber von IVGK-Netzärzten betreuten Versicherten
- (c) die *nicht* in die IVGK eingeschriebenen und *nicht* von IVGK-Netzärzten betreuten, aber im Kinzigtal wohnenden Versicherten.

An dieser Stelle kann nicht näher darauf eingegangen werden, welche Konsequenzen sich aus dieser Differenzierung der Interventionspopulation(en) für die Bestimmung einer angemessenen Kontrollgruppe ergeben. Hier muss die Feststellung genügen, dass in den drei genannten IVGK-Evaluationsstudien mit kontrolliertem Design (vgl. Tab. 1) *unterschiedliche* Kontrollgruppen herangezogen werden, die sich aber stets aus AOK- bzw. LKK-Versicherten zusammensetzen.

2.3. Die Gefahr der „Überforschung“ einer relativ kleinen Interventionspopulation

Wegen der zumindest in der Anfangsphase geringen Größe der Interventionspopulation im engeren Sinne (d.h. der eingeschriebenen Versicherten) müssen aus statistischen Gründen häufig Vollerhebungen vorgenommen werden. So wurde in der SDM-Studie, die als kontrollierte Kohortenstudie konzipiert war, die Gruppe der Eingeschriebenen in Form einer Vollerhebung rekrutiert, denn unmittelbar vor Beginn der ersten Befragungswelle (31.03.2007) waren nur ca. 1200 Personen in die IVGK eingeschrieben. Eine Befragung (per Fragebogen) aller 1.200 IV-Versicherten war in der Erstbefragung unbedingt anzustreben, da ein kleiner Interventionseffekt erwartet

wurde und für die dritte Befragungswelle im Jahr 2009 nur noch mit einer Response-Rate von ca. 25% (der zu Anfang Befragten) gerechnet werden konnte.

Hieraus resultiert jedoch ein weiteres Problem: Starten mehrere Befragungsstudien in etwa demselben Zeitraum, so häuft sich für die betreffende (Sub-)Population zwangsläufig die Zahl der Befragungen. Damit droht die Gefahr der „Überforschung“ der Interventionspopulation [9] – eventuell mit der Folge, dass viele Befragte sich nicht mehr beteiligen und im Extremfall die Evaluation ihre Akzeptanz verliert. Das Problem wird im Fall *Gesundes Kinzigtal* durch die soziodemographischen Besonderheiten der eingeschriebenen AOK- und LKK-Versicherten noch verschärft: Das eher niedrige formale Bildungsniveau und das höhere Lebensalter von durchschnittlich 58 Jahren der bis Frühjahr 2008 eingeschriebenen Versicherten [6] bedingen, dass viele Befragte einen 10-seitigen Fragebogen als Überforderung wahrnehmen und evtl. nicht an der Befragung teilnehmen.

Lösungen bzw. Teillösungen bestehen in der strikten Bemühung um Kürze bei Interviews und Fragebögen, in der zeitversetzten Koordinierung verschiedener Primärdatenerhebungen und im Ausweichen auf Sekundärdaten-Analysen (z. B. von GKV-Routinedaten – vgl. [10,11]) oder Behandlungsdaten der Netzärzte. Diese können dann im Idealfall, sofern für die Fragestellung notwendig, durch wenige, möglichst im Behandlungskontext zu erhebende Primärdaten ergänzt werden. Es hat sich gezeigt, dass es in jedem Fall hilfreich ist, vor umfangreicheren Primärdatenerhebungen (z. B. per Fragebogen) alle am Versorgungssystem beteiligten Stellen und Vertragspartner – Ärzte und Versorgungseinrichtungen, Geschäftsstelle der Kinzigtal GmbH, Krankenkassen, wenn möglich auch die Versicherten selbst (im Fall der IVGK in Form von „Mitgliederversammlungen“) – über die bevorstehende Aktion zu informieren, damit die Beteiligten bei den Versicherten um Verständnis werben oder zumindest Rückfragen beantworten können.

2.4. Multiple Programmeffekte bei IV-Versicherten, die an mehreren Präventions- und/oder Krankheitsmanagementprogrammen teilnehmen

Ein weiteres Problem der Evaluation besteht in der Zuordnung der Einzeleffekte von Gesundheits- oder Krankheitsmanagementprogrammen: Ein relevanter Teil der IV-Eingeschriebenen nimmt an mehr als nur einem strukturierten Behandlungs- oder Präventionsprogramm teil. So überschneiden sich z.B. die Teilnehmerkreise der AGil-Intervention (vgl. Abschnitt 1) zum Teil mit den Teilnehmerkreisen der

Programme zur Osteoporose-Prävention und zum metabolischen Syndrom. Zudem werden viele Personen, die an diesen Gesundheitsprogrammen teilnehmen, zusätzlich noch im Rahmen strukturierter Behandlungsprogramme (DMP) für Diabetes mellitus Typ II oder KHK behandelt (zu den Teilnehmerzahlen vgl. den Beitrag von Hildebrandt et al. in diesem Heft). Technisch gesprochen ist hier also die Interventionsbedingung (Teilnahme an einem bestimmten Programm) kontaminiert durch mögliche Effekte, die aus der gleichzeitigen Teilnahme an anderen Programmen resultieren. Berücksichtigte man die Möglichkeit derartiger multipler Programmeffekte *nicht*, würde die Effektivität des zu evaluierenden einzelnen Programms vermutlich überschätzt.

Multiple Programmeffekte können zwar prinzipiell bei der Auswertung einer Studie kontrolliert und die Ergebnisse entsprechend adjustiert werden; Aussicht auf eine befriedigende Kontrolle multipler Programmeffekte besteht jedoch nur dann, wenn relativ viele Versicherte an dem zu evaluierenden Programm teilnehmen und zumindest ein Großteil der Programmteilnehmer nicht noch weitere outcome-relevante Programme durchläuft. Daher muss sorgfältig überlegt werden, ob (und ggf. welche) einzelnen Programme im Rahmen eines integrierten Versorgungssystems überhaupt hinsichtlich ihrer spezifischen Effektivität evaluiert werden sollen.

In relativ kleinen integrierten Versorgungssystemen mit einer Vielzahl an Programmen bietet es sich an, den Fokus der Evaluation zusätzlich auf eher „summarische“ Outcome-Analysen zu legen und das Ensemble der einzelnen Präventions- und Behandlungsprogramme als eine einzige zusammengesetzte, komplexe Interventionsvariable zu konzipieren. (Zur Diskussion über spezifische Evaluationsprobleme bei „komplexen Interventionen“ vgl. [12–19].) Man sollte sich in diesem Fall jedoch der Problematik bewusst sein, dass dann nur noch die *kombinierte* Wirkung sehr vieler einzelner Wirkfaktoren beschrieben werden kann und ihr individueller Beitrag nicht zufriedenstellend zu klären ist.

2.5. Interne Validität der Evaluationsstudien

Ein Problem, das grundsätzlich die Validität von Evaluationsstudien und die Aussagekraft ihrer Ergebnisse berührt, ist die Frage der Vergleichsstichprobe, welche zur Kontrolle des Interventionseffekts herangezogen wird. Die größtmögliche Beweiskraft (engl.: „evidence“) wird bekanntlich durch Studiendesigns erzielt, in deren Kontext die Probanden randomisiert auf Interventions- und Kontrollgruppe verteilt werden [20, 21]. Bei einer Vielzahl von Studien im Kontext des Versorgungsalltags erweist sich

die Variante einer cluster-randomisierten Zuteilung der Probanden auf die verschiedenen Behandlungsbedingungen als sehr praktikabel [22–24]; Randomisierungseinheit ist hierbei nicht der einzelne Patient bzw. Proband, sondern die übergeordnete Versorgungsinstanz (z.B. Arztpraxis, Klinikabteilung) mit den jeweils dazu gehörenden Patienten.

Zur Untersuchung bestimmter Evaluationsfragen hätte auch im Kinzigtal im Prinzip ein randomisiertes bzw. cluster-randomisiertes Studiendesign angewandt werden können: Hätte man z.B. den Effekt der AGil-Intervention unter Bedingungen eines ländlich-kleinstädtisch geprägten integrierten Versorgungssystems untersuchen wollen, so hätten die potentiellen Teilnehmer der AGil-Intervention (IV-eingeschriebene Kinzigtal-Versicherte ab 60 Jahren) randomisiert auf Interventionsbedingung (AGil-Programmteilnahme) bzw. Kontrollbedingung (keine AGil-Programmteilnahme bzw. Wartekontrollgruppe) aufgeteilt werden können. Bei der Umsetzung hätte es verschiedene praktische Probleme gegeben – z.B. bzgl. der Rekrutierung einer hinreichend großen Stichprobe innerhalb der gegebenen Projektlaufzeit; diese hätten sich jedoch evtl. pragmatisch lösen lassen. Der Hauptgrund, keine randomisierte Studie durchzuführen, war jedoch folgender: Die Gesundes Kinzigtal GmbH sah sich in der Pflicht, *allen* IV-Versicherten im Kinzigtal *so früh wie möglich* Gesundheitsprogramme anzubieten, die eine hinreichende Aussicht auf einen Präventionserfolg bzw. eine Verbesserung ihrer Gesundheit versprachen. Zudem war das AGil-Programm zuvor (in einem großstädtischen Kontext) erfolgreich durchgeführt worden und hatte u. a. 2005 den Deutschen Präventionspreis erhalten. Ein Studiendesign, das für einen Teil der IV-Versicherten (hier: die Probanden der Kontrollgruppe bzw. Wartekontrollgruppe) einen *Aufschub* der Versorgung mit einem aussichtsreichen Präventionsprogramm bedeutet hätte, war für das IVGK-Management nicht akzeptabel. Die Preisgabe eines randomisierten Designs bedeutet jedoch ein Opfer, das methodologisch sehr schwer wiegt – denn man verzichtet damit von vornherein darauf, einer Studie die größtmögliche Aussagekraft zu sichern.

Für andere Evaluationsfragen der IVGK-Evaluation kam unter den gegebenen historischen und organisatorischen Rahmenbedingungen von vornherein kein randomisiertes oder cluster-randomisiertes Design in Frage. Dies betrifft z. B. die Frage, inwieweit ein regionales IV-System (wie im Kinzigtal) zum Abbau von Phänomenen der Über-, Unter- und Fehlversorgung in einer ländlich-kleinstädtisch geprägten Versorgungsregion beizutragen vermag. Eine solche Frage wäre nur dann mithilfe

eines randomisierten Studiendesigns beantwortbar, wenn mehrere gleichartige Regionen – die gleichermaßen für den Aufbau eines regionalen IV-Systems in Frage kämen – randomisiert der Interventionsbedingung (Aufbau eines regionalen IV-Systems) oder der Kontrollbedingung (Fortsetzung von „usual care“ bzw. Wartekontrollgruppe) zugewiesen würden. Hierfür fehlten im Kontext der Kinzigtal-Evaluation die Voraussetzungen. Die Frage, inwieweit ein IV-System des Typs *Gesundes Kinzigtal* zum Abbau von Über-, Unter- und Fehlversorgung beiträgt, wird im Rahmen der Kinzigtal-Evaluation (in Form der ÜUF-Studie – vgl. Abschnitt 1) daher auf der Grundlage einer bevölkerungsbezogenen, quasi-experimentellen kontrollierten Studie zu beantworten versucht. Darin werden die Daten aller AOK- und LKK-Versicherten im Kinzigtal (Interventionsgruppe) mit den Daten einer für das übrige Baden-Württemberg repräsentativen Zufallsstichprobe AOK- und LKK-Versicherter (Vergleichsgruppe) verglichen. Dabei werden die Werte der Zielvariablen in der Vergleichsgruppe stets hinsichtlich prognostisch wichtiger Variablen (Alter, Geschlecht, ggf. weitere Merkmale) standardisiert bzw. adjustiert, und zwar gemäß einer Verteilung, wie sie in der Interventionsgruppe vorgefunden wurde.

Werden keine randomisierten Studiendesigns angewandt, müssen Abstriche an der internen Validität, d.h. an der Aussagekraft der Studienergebnisse in Kauf genommen werden. Denn selbst wenn eine befriedigende Adjustierung der (nicht-randomisiert bestimmten) Kontrollgruppe in Bezug auf alle *bekannt*en Kovariablen und Confounder gelänge, ist bei einem nicht-randomisierten Design immer noch damit zu rechnen, dass die ungleiche Verteilung *unbekannt*er Confounder auf Interventions- und Kontrollgruppe zu systematischen Fehlschlüssen führen kann [21]. Dies ist bei der Interpretation der Ergebnisse solcher Studien zu berücksichtigen; daher sollten etwaige „positive Ergebnisse“ zurückhaltend bewertet werden.³

3. Fazit

Anhand des Beispiels der Evaluation der Integrierten Versorgung *Gesundes Kinzigtal* (IVGK) wurden fünf grundsätzliche Probleme benannt, die sich bei der Evaluation regionaler Versorgungssysteme ergeben. Es wurde deutlich, dass manche Lösungen neue Probleme erzeugen. So resultiert z. B. aus der Strategie der Vollerhebung in

³ Im Rahmen seines Sondergutachtens 2009 erstellte der Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen auch einen systematischen Review zur Frage, ob Managed-Care-Systeme geeignet sind, die Versorgungsqualität erhöhen. In methodologischer Hinsicht kam der Rat interessanterweise zu dem Schluss, „dass Studien mit einem geringerwertigen Studiendesign den positiven Effekt von Managed-Care-Systemen nicht überschätzen, es also keinen Bias zuungunsten strengerer Studiendesigns gibt.“ [25: p. 448].

einer bestimmten Subgruppe der Interventionspopulation die Gefahr ihrer „Überforschung“. Dieses Problem kann aber wiederum durch bestimmte Maßnahmen und Untersuchungsansätze zumindest teilweise entschärft werden. Weitere typische Probleme, die die Aussagekraft entsprechender Evaluationsstudien einschränken können, resultieren aus der Wahrscheinlichkeit multipler Programmeffekte in der Interventionspopulation sowie im – zuweilen notwendigen – Rückgriff auf nicht-randomisierte Studiendesigns, deren Beweiskraft prinzipiell geringer ist als diejenige randomisierter oder cluster-randomisierter Studien.

4. Literatur

[1] Gemeinsame Registrierungsstelle zur Unterstützung der Umsetzung des § 140d SGB V. Entwicklung der Integrierten Versorgung in der Bundesrepublik Deutschland 2004-2008. Bericht gemäß § 140d SGB V auf der Grundlage der Meldungen von Verträgen zur integrierten Versorgung, Datenbank-Stand: 19.10.2009, Autor: Franz-Josef Grothaus.

[2] Weatherly JN, Seiler R, Meyer-Lutterloh K, Schmid E, Läger R, Amelung VE. Leuchtturmprojekte Integrierter Versorgung und Medizinischer Versorgungszentren. Innovative Modelle in der Praxis. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2007.

[3] Hildebrandt H, Hermann C, Knittel R, Richter-Reichhelm M, Siegel A, Witzenrath W. Gesundes Kinzigtal Integrated Care: Improving Population Health by a Shared Health Gain Approach and a Shared Savings Contract. International Journal of Integrated Care 2010; 10 (April-June): 1-15. Download via: <http://www.ijic.org/index.php/ijic/article/view/539/1051>.

[4] Hildebrandt H, Richter-Reichhelm M, Trojan A, Glaeske G, Hesselmann H. Die Hohe Kunst der Anreize: Neue Vergütungsstrukturen im deutschen Gesundheitswesen und der Bedarf an Systemlösungen. Sozialer Fortschritt 2009; 58 (7): 154-260.

[5] Hermann C, Hildebrandt H, Richter-Reichhelm M, Schwartz FW, Witzenrath W. Das Modell „Gesundes Kinzigtal“. Managementgesellschaft organisiert Integrierte Versorgung einer definierten Population auf Basis eines Einsparcontractings. Gesundheits- und Sozialpolitik 2006; 5-6: 11-29.

[6] Siegel A, Stößel U, Geßner D, Knittel R, Beckebans F, Hildebrandt H, Hermann C, Trojan A, Pfaff H (2008): Kooperation und Wettbewerb im integrierten Versorgungssystem „Gesundes Kinzigtal“. In: Amelung VE, Sydow J, Windeler A, editors. Vernetzung im Gesundheitswesen. Wettbewerb und Kooperation. Stuttgart: Kohlhammer; 2008. p. 223-235.

- [7] Schauder P, Berthold H, Eckel H, Ollenschläger G, editors. Zukunft sichern: Senkung der Zahl chronisch Kranker. Verwirklichung einer realistischen Utopie. Köln: Deutscher Ärzte-Verlag; 2006.
- [8] Kolominsky-Rabas P, Heuschmann P. Inzidenz, Ätiologie und Langzeitprognose des Schlaganfalls. Fortschritte der Neurologie und Psychiatrie 2002; 70: 657-662.
- [9] Baumberger J. Die drei Gefahren für Managed-Care-Systeme. In: Preuß KJ, Rübiger J & Sommer J, editors. Managed Care. Evaluation und Performance-Measurement integrierter Versorgungsmodelle. Stand der Entwicklung in der EU, der Schweiz und den USA. Stuttgart/New York: Schattauer; 2002. p. 264-266.
- [10] Ihle P, Köster I, Herholz H, Rambow-Bertram P, Schardt T, Schubert I. Versichertenstichprobe AOK Hessen/KV Hessen – Konzeption und Umsetzung einer personenbezogenen Datenbasis aus der Gesetzlichen Krankenversicherung. Gesundheitswesen 2005; 67: 638-645.
- [11] Schubert I, Köster I, Küpper-Nybelen J, Ihle P. Versorgungsforschung mit GKV-Routinedaten. Nutzungsmöglichkeiten versichertenbezogener Krankenkassendaten für Fragestellungen der Versorgungsforschung. Bundesgesundheitsbl – Gesundheitsforsch – Gesundheitsschutz 2008; 51 (10): 1095-1105.
- [12] Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P. Framework for design and evaluation of complex interventions to improve health. BMJ 2000; 321: 694-696.
- [13] Slade M & Priebe S. Are randomised controlled trials the only gold that glitters? British Journal of Psychiatry 2001; 179: 286-287.
- [14] Hawe P, Shiell A, Riley T. Complex interventions: how “out of control” can a randomised controlled trial be? BMJ 2004; 328: 1561-1563.
- [15] Blackwood B. Methodological issues in evaluating complex healthcare interventions. Journal of Advanced Nursing 2006; 54 (5): 612-622.
- [16] May C. A rational model for assessing and evaluating complex interventions in health care. BMC Health Services Research 2006; 6: 86.
- [17] Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. In: BMC Medical Research Methodology 2006; 6: 54.
- [18] Morfeld M & Wirtz M. Methodische Ansätze in der Versorgungsforschung. Das Beispiel Evaluation der Disease-Management-Programme. Bundesgesundheitsbl – Gesundheitsforsch – Gesundheitsschutz 2006; 49: 120-129.
- [19] Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. BMJ 2008; 336: 1281-1283.
- [20] Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. Controlled Clinical Trials 1995; 16: 62-73.

[21] Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche 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.

[22] Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004; 328: 702-708.

[23] Chenot F. Cluster-randomisierte Studien: eine wichtige Methode in der allgemeinmedizinischen Forschung. *ZEFQ* 2009; 103: 475-480.

[24] Schrappe M. Methoden der Qualitäts- und Patientensicherheitsforschung. In: Pfaff H, Neugebauer EAM, Glaeske G, Schrappe M, editors. *Lehrbuch Versorgungsforschung*. Stuttgart: Schattauer; 2011; p. 323-328.

[25] Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen. Koordination und Integration. *Gesundheitsversorgung in einer Gesellschaft des längeren Lebens*. Bundestagsdrucksache 16/13770. Berlin 2009; Download via: <http://dip21.bundestag.de/dip21/btd/16/137/1613770.pdf>.

Tabellen

Tabelle 1: Externe Evaluationsstudien zur Integrierten Versorgung Gesundes Kinzigtal (IVGK) im Zeitraum 2007-2011

Modul	Studienleitung	Forschungsfragen	Studiendesign
SDM-Studie	Prof. Dr. Dr. Martin Härter, Universitätsklinikum Hamburg-Eppendorf	Verändert sich durch eine SDM-bezogene Fortbildung der Netzärzte und durch andere Merkmale der IVGK die Patientenbeteiligung an Therapieentscheidungen, die Zufriedenheit mit der ärztlichen Versorgung und die gesundheitsbezogene Lebensqualität?	Dreiarmlige kontrollierte Kohortenstudie (Primärdaten)
ÜUF-Studie	Dr. Ingrid Schubert, PMV forschungsgruppe, Universität zu Köln	Wie verändert sich im Zuge der IVGK-Intervention im Zeitraum 2004 (Basisjahr) bis 2011 die administrative Prävalenz bzw. Inzidenz ausgewählter Erkrankungen im Kinzigtal im Vergleich zum übrigen Baden-Württemberg? Können Phänomene der Über-, Unter- und Fehlversorgung stärker reduziert werden?	Quasi-experimentelle kontrollierte Studie mit GKV-Routinedaten
PEGL-Studie	Dr. Matthias Nübling, Gesellschaft für empirische Beratung, Denzlingen	Wie verändern sich im Verlauf der IVGK-Etablierung die Einstellungen der Leistungspartner zu (bzw. ihre Zufriedenheit mit) verschiedenen prozessualen und strukturellen Aspekten?	Kohortenstudie/ Trendstudie (Primärdaten)
AGil-Studie	Prof. Dr. Olaf von dem Knesebeck, Universitätsklinikum Hamburg-Eppendorf	Wie verändern sich gesundheitsbezogene Lebensqualität, Inanspruchnahme von Versorgungsleistungen und Versorgungskosten von Teilnehmern des AGil-Programms?	Kohortenstudie (Primärdaten) und retrospektive kontrollierte Kohortenstudie (Sekundärdaten)

Erläuterungen zu Tabelle 1: SDM – Shared-Decision-Making; ÜUF – Über-, Unter- und Fehlversorgung; PeGL – Prozessevaluation aus Sicht der GK-Leistungspartner; AGil – Aktive Gesundheitsförderung bei alten Menschen im Kinzigtal.