

Entwicklung und Evaluation komplexer Interventionen weiterdenken

NICK BERTRAM,
SEBASTIAN BINDER,
LINDA KERKEMEYER,
VOLKER E. AMELUNG

Dr. med.dent. Nick Bertram, MPH ist Projektmanager am inav – privates Institut für angewandte Versorgungsforschung GmbH, Berlin

Sebastian Binder ist Projektmanager Versorgungsforschung am inav – privates Institut für angewandte Versorgungsforschung GmbH, Berlin

Linda Kerkemeyer ist Senior Managerin Gesundheitsökonomie am inav – privates Institut für angewandte Versorgungsforschung GmbH, Berlin

Prof. Dr. Volker Eric Amelung ist Inhaber der Schwerpunktprofessur für Internationale Gesundheitssystemforschung an der Medizinischen Hochschule Hannover sowie Geschäftsführer und Leiter des inav – privates Institut für angewandte Versorgungsforschung GmbH, Berlin. Prof. Amelung ist Mitherausgeber der Zeitschrift *Gesundheits- und Sozialpolitik*

Eine der Voraussetzungen für die Projektförderung im Rahmen des Innovationsfonds ist ein tragfähiges Evaluationskonzept. Dieses ist wesentlich für die Beurteilung, ob eine neue Versorgungsform in die Regelversorgung übernommen werden kann. Für die Wissenschaft birgt die Evaluation dieser Projekte eine enorme Herausforderung. Möglichkeiten, dieser mit adäquaten Methoden zu begegnen, werden im Folgenden erörtert.

Der Innovationsfonds als Motor komplexer Interventionen

Mit der Verabschiedung des GKV-Versorgungsstärkungsgesetzes im Jahr 2015 wurde der Innovationsfonds mit einer Laufzeit von – zunächst – 2016 bis 2019 als gesundheitspolitisches Instrument zur Förderung der Integrierten Versorgung und Versorgungsforschung in Deutschland ins Leben gerufen. Im besonderen Fokus stehen neue Versorgungsformen, die die sektorenübergreifende Versorgung verbessern sollen und ein hinreichendes Potential haben, dauerhaft in die Regelversorgung aufgenommen zu werden. Die Trägerschaft des Innovationsfonds hat der Gemeinsame Bundesausschuss (G-BA) inne, die Administration obliegt dem eigens dafür eingerichteten Innovationsausschuss beim G-BA und die Projektträgerschaft übernimmt das Deutsche Zentrum für Luft- und Raumfahrt (DLR) (1).

In den ersten Förderwellen waren die Themenschwerpunkte des Innovationsfonds u.a. Versorgungsmodelle in strukturschwachen oder ländlichen Gebieten, Modellprojekte zur Arzneimitteltherapie und Arzneimitteltherapiesicherheit, Versorgungsmodelle unter Nutzung von

Telemedizin, Telematik und E-Health, Modelle mit Delegation und Substitution von Leistungen oder sozialleistungsträgerübergreifende Versorgungsmodelle (2).

Eine der Förderungsvoraussetzungen im Rahmen des Innovationsfonds stellt ein tragfähiges Evaluationskonzept dar, denn die Evaluationsergebnisse sollen nach Abschluss der Projektlaufzeit dem G-BA Erkenntnisse darüber liefern, ob die neue Versorgungsform in die Regelversorgung übernommen werden sollte bzw. als Grundlage für strukturelle Veränderungen des gesetzlichen Rahmens dienen könnte (2). Für die Wissenschaft jedoch stellt die Evaluation dieser Projekte eine enorme Herausforderung dar, denn die Beurteilung derartiger komplexer Interventionen befindet sich vorerst noch in einem Anfangsstadium.

Was sind komplexe Interventionen?

Komplexe Interventionen (*complex interventions*) sind insbesondere im Gesundheits- und Sozialwesen und im Zusammenhang mit Public Health-Maßnahmen weit verbreitet. Allen ist gemeinsam, dass sie Einfluss, Auswirkungen und Folgen auf die Gesundheit haben können (3).

Komplexen Interventionen wird in der Literatur meist der Terminus der einfachen bzw. singulären Interventionen (*simple interventions*) gegenübergestellt. Jedoch sollte festgehalten werden, dass eine scharfe Trennung beider Termini äußerst schwierig ist. Einfache Interventionen beziehen sich zumeist auf neue Arzneimittel, Diagnoseverfahren, Medizinprodukte oder chirurgische Eingriffe, da diese in der Regel von nur einem Leistungserbringer erbracht werden und einer genau definierten Population mit entsprechend zugrundeliegender Indikation zugutekommen. Diese Interventionen werden als einfache lineare Wege betrachtet. Intervention und Ergebnis sind unmittelbar – und im besten Falle – kausal miteinander in Verbindung zu bringen. Sie betrachten archetypisch lediglich eine Erkrankung und entsprechend auch nur ein Outcome (4). In der Literatur wird als Paradebeispiel einfacher Interventionen immer wieder die Erprobung neuer Arzneimittel gegenüber dem jeweiligen Goldstandard in der Arzneimitteltherapie in einer randomisierten kontrollierten Studie (RCT; *randomized controlled trial*) angeführt. Jedoch sollte bereits an dieser Stelle festgehalten werden, dass bei derartiger Betrachtung und Definition komplexe Sachverhalte – wie bspw. das Einnahmeverhalten oder potentielle Me-

dikamentenwechselwirkungen – zumeist ausgeblendet werden. Finden diese jedoch Beachtung, so erscheinen selbst einfache Interventionen nicht mehr ganz so einfach.

Bei komplexen Interventionen ist das Verhältnis zwischen Intervention und Ergebnis nicht derart „einfach“, da u.a. mehrere Erkrankungen und/oder mehrere Outcomes in die Betrachtungen einbezogen werden.

Erste Versuche, komplexe Interventionen zu definieren sowie eine Anleitung zur Entwicklung und Bewertung solcher Interventionen zu formulieren, gab es im Jahr 2000 mit der Publikation des *Framework for Design and Evaluation of Complex Interventions to Improve Health* vom *Medical Research Council (MRC)* (5). Dieses erfuhr im Jahr 2008 eine Aktualisierung und Erweiterung (3). Das MRC definiert eine Intervention als komplex, wenn sie eine oder mehrere der folgenden Merkmale erfüllt:

- Vielzahl interagierender Komponenten,
- Hohe Anzahl (Mischung) involvierter (Patienten-)Gruppen, Leistungserbringer, Organisationen oder weitere Akteure unterschiedlicher sozialer, wirtschaftlicher, kultureller, ökologischer oder politischer Kontexte und Settings,

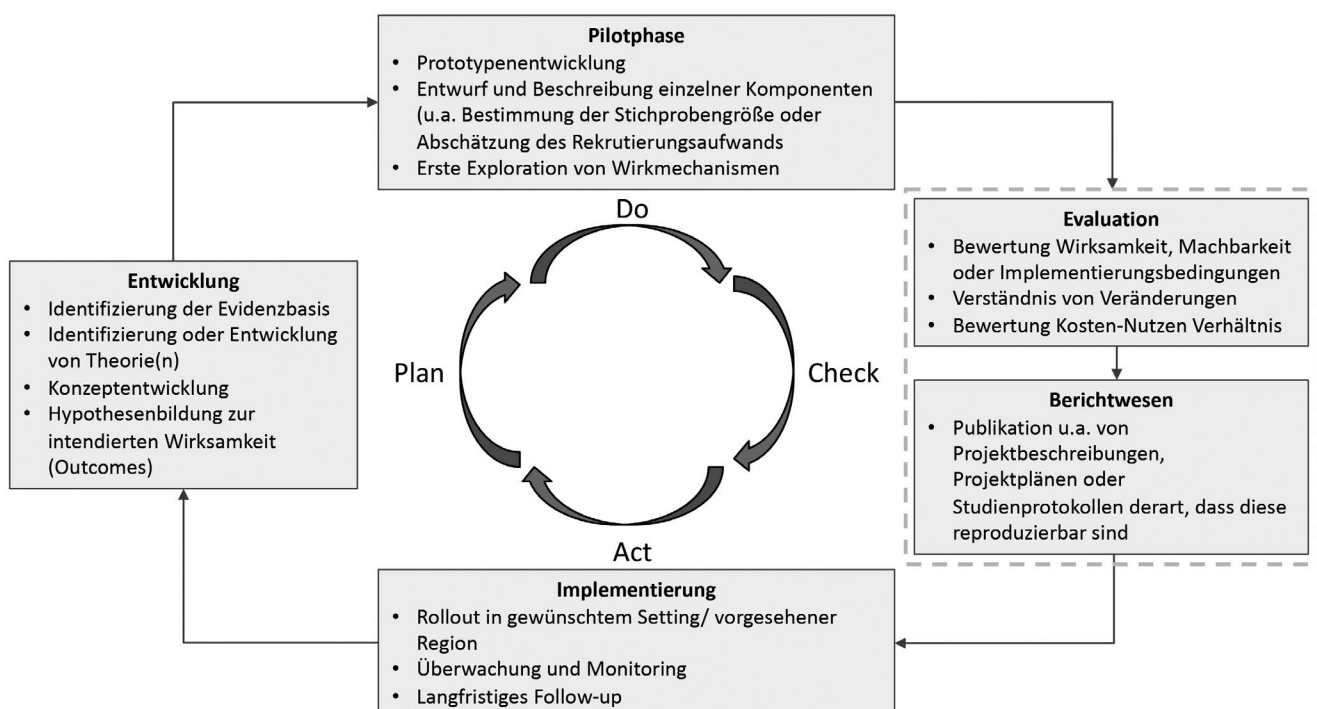
- Interdisziplinarität und Multisektoralität,
- Mehrere Indikationen mit entsprechend intendierten Outcomes,
- Lernfähigkeit und Anpassungsvermögen der Intervention durch Rückkopplungs-/Feedbackschleifen sowie
- Beeinträchtigung der Wirksamkeit der Intervention durch das Verhalten derjenigen, die die Intervention erbringen bzw. erhalten.

Dabei sollte festgehalten werden, dass Komplexität nicht unbedingt ein Charakteristikum der Intervention selbst darstellen muss. Denn *komplex* kann bspw. auch das Setting sein, in dem die Intervention implementiert wird (6).

Einflüsse auf die Ergebnisse komplexer Interventionen haben neben der Intervention selbst auch weitere Kontextfaktoren wie Hierarchien oder Berufsgrenzen im jeweiligen Setting, soziale, geografische und umweltbedingte Barrieren sowie die Auswirkungen anderer, gleichzeitig stattfindender organisatorischer Veränderungen (7).

Viele Maßnahmen der Gesundheitsversorgung stellen komplexe Interventionen dar. Sie bestehen aus mehreren Einzelkomponenten, die sich wechselseitig bedingen. Als Beispiele hierfür können

Abbildung 1: Phasenmodell zur Entwicklung und Evaluation komplexer Interventionen



Quelle: Eigene Darstellung in Anlehnung an Craig et al. und Mühlhauser et al. (3,8)

u.a. Stroke Units, Disease-Management-Programme oder Projekte zur Verbesserung der Krankenhaushygiene angeführt werden (8).

Der epidemiologische Übergang als Folge und Ursache des demografischen Wandels führt zu einer Veränderung in der Bevölkerungsstruktur und –zusammensetzung mit entsprechender Auswirkung auf die Häufigkeit von Krankheiten und Todesursachen. Diese Herausforderungen erfordern zunehmend komplexe Interventionen. Je größer die Schwierigkeit ist, präzise zu definieren, welche Bestandteile einer Intervention tatsächlich wirksam sind und wie sie sich aufeinander beziehen, desto größer ist die Wahrscheinlichkeit, dass es sich um eine komplexe Intervention handelt (9).

Entwicklung einer komplexen Intervention

Das MRC hat im Jahr 2000 – und in überarbeiteter und erweiterter Form im Jahr 2008 – analog zur Entwicklung von Medikamenten ein sog. Phasenmodell zur Entwicklung und Evaluation von komplexen Interventionen vorgeschlagen. Es umfasst die folgenden Phasen (3,5):

- Entwicklung (*development*),
- Pilotphase (*testing/piloting*),
- Evaluation (*evaluation*),
- Berichtswesen (*reporting*) und
- Implementierung (*implementation*).

Dieses Modell lässt sich auch gut mit dem bekannten Shewhart-Zyklus (*Plan-Do-Check-Act-Cycle*) (10) vergleichen und beschreiben (Abbildung 1).

Bei der Entwicklung, Pilotierung und Evaluation einer komplexen Intervention ist die Recherche oder – sofern kein geeigneter oder aktueller Review vorhanden ist – die eigenständige Durchführung eines systematischen Reviews zur Generierung evidenzbasierter Theorie zur Thematik notwendig (11). Darunter ist nicht nur der traditionelle systematische Review zu verstehen, wie er maßgeblich seitens Cochrane Collaboration gefordert wird. So sollten im Hinblick auf komplexe Interventionen gängige Reviewmethoden erweitert und Limitierungen auf bestimmte Studiendesigns und Forschungsmethoden aufgehoben werden. Neben klassischen RCTs sollten demnach alternative Studiendesigns wie kontrollierte clusterrandomisierte

Interventionsstudien oder – wenn eine Randomisierung aus gewichtigen ethischen oder logistischen Gründen im Rahmen der Forschungsfrage(n) nicht möglich ist – quasi-experimentelle oder Beobachtungsstudien als potenziell gleichberechtigt herangezogen werden. Essentiell dabei ist, dass bei begründetem Verzicht auf eine Randomisierung fundierte methodische Maßnahmen zur Adjustierung von effektverzerrenden Elementen in Studienplanung und -analyse mitgedacht werden (12). Darüber hinaus können neben rein quantitativen Analysemethoden auch genuin qualitative oder mixed methods Ansätze Anwendung finden, um umfassendere Erkenntnisgewinne über komplexe Interventionen zu erzielen. In Vorbereitung der geplanten Evaluation sollte der Wissenstand zur Thematik systematisch – beispielsweise durch Methoden wie das sog. *Snowballing* der Literatur relevanter Studien oder die Kontaktaufnahme zu betreffenden Autoren und Experten – aufbereitet werden, um bestehende Wissenslücken effektiv schließen oder eine vermeintliche Wiederholung methodischer Schwächen verhindern zu können (8,13,14).

Alle Phasen sollten transparent sein und insbesondere so publiziert werden, dass diese als zusammenhängender, vollständiger Prozess oder einzelne Mechanismen der einzelnen Phasen für andere Akteure reproduzierbar sind und bei Entscheidungsprozessen unterschiedlicher Stakeholder unterstützend wirken können (3).

Probleme bei der Evaluation komplexer Interventionen

Komplexe Interventionen erfordern die Analyse verschiedener Bestandteile (9):

- Synergien zwischen interagierenden Komponenten,
- Rückkopplungs-/Feedbackschleifen und deren Auswirkungen,
- Interaktionen zwischen unterschiedlichen gesundheitlichen und nicht-gesundheitlichen Outcomes,
- Zugrundeliegende Prozesse und deren Umstellungen.

Was zunächst einfach und logisch erscheint, ist jedoch mit einigen nicht unerheblichen Schwierigkeiten verbunden. Dies lässt sich auch darauf zurückführen, dass oftmals ein rein medizinisches For-

schungsverständnis zugrunde gelegt wird, wie es aber – aufgrund der Heterogenität der Interventionsansätze – bei komplexen Interventionen nicht der Fall sein kann.

Unter praktischen und methodischen Gesichtspunkten stellt sich zunächst die Frage nach dem *richtigen* Studiendesign. So gelten bei *einfachen* Interventionen RCTs als Goldstandard bei der Evaluation, weil diese – allseits akzeptiert – als beste Methode für die Untersuchung der Effektivität von Interventionen gelten sowie eine verzerrte Effektschätzungen und damit (Selektions-)Bias vermeiden helfen (vgl. u.a. 15–17). Hingegen sind *komplexe* Interventionen häufig Interventionen, die keinen Kontrollgruppenansatz und damit per se auch keine Randomisierung im eigentlichen Sinne zulassen. Dies trifft bspw. zu, wenn Maßnahmen zeitgleich auf die gesamte Bevölkerung ausgerollt werden sollen. Des Weiteren sprechen nachfolgende Gründe gegen die Durchführung eines RCT:

- Ethik (z.B. der Kontrollgruppe eine positive Intervention vorenthalten)
- Praktikabilität (z.B. können komplexe Interventionen selten en bloc implementiert werden, sondern müssen vielmehr sequentiell ausgerollt werden) und Machbarkeit (z.B. Fokussierung auf einzelne Individuen und nicht ein ganzes System)
- Wirtschaftlichkeit (z.B. begrenzte finanzielle Ressourcen bei Forschungsvorhaben)
- Methodik (z.B. Bildung von Kontrollgruppen bei Settingansätzen sehr schwierig, aufwändig und oft ohne direkt resultierenden Nutzen)

Dies führt gleichzeitig auch dazu, dass experimentelle Ansätze nur schwer oder gar nicht möglich sind. Hier können dann lediglich gute, nicht-experimentelle Ansätze zum Einsatz kommen (12,18).

Die organisatorische und logistische Komplexität derartiger Interventionen führt darüber hinaus zu enormen Schwierigkeiten bei der Standardisierung von Erhebungsinstrumenten und Daten und der damit einhergehenden Abbildung des Einflusses des Settings. Zudem ist eine ad hoc Bewertung von betrachteten Outcomes, wie im Kontext von *einfachen* Interventionen, äußerst selten möglich, da komplexe Interventionen häufig über einen langen Zeitraum durchgeführt werden bzw. es oft lange

dauert, bis diese ihre *volle Wirkung* entfalten (3).

Erfahrungen aus der Literatur zeigen, wie schwierig es ist, bspw. die Wirksamkeit oder den Nutzen (Schaden) komplexer Interventionen zu beurteilen. So bleiben der Beitrag einer jeden Einzelkomponente von komplexen Interventionen sowie der Einfluss ihrer Interaktionen mit dem Setting, den agierenden Personen und Organisationen auf das Gesamtergebnis meist unklar und können nicht adäquat abgebildet werden (8). So kann jeder Erfolgswachweis einer Intervention dadurch verkompliziert werden, dass neben der eigentlichen (komplexen) Intervention weitere Faktoren auf diese Einfluss nehmen können, die derart nicht betrachtet oder angedacht waren (18). Darüber hinaus sind Effekte wahrscheinlich kleiner und variabler, sobald Interventionen bevölkerungs- und nicht mehr nur personen- bzw. indikationsbezogen umgesetzt werden (3).

In der Evaluation komplexer Interventionen bestehen oft Herausforderungen bei der Festlegung von *harten* Outcomes, die alle Effekte erfassen. Die traditionelle Fokussierung auf einen primären Endpunkt und weitere sekundäre Outcomes erscheint in diesem Kontext auch nicht mehr zeitgemäß (7). Auch ist der *Untersuchungsgegenstand* nicht zwangsläufig über die Dauer der Intervention konstant und kann entsprechend nur begrenzt standardisiert werden (19).

Evaluation komplexer Interventionen

Dienten Evaluationen in der Vergangenheit primär dem Nachweis von Wirksamkeit von Interventionen (Ist eine signifikante (positive) Veränderung als Folge der Intervention eingetreten?), so rückt mittlerweile immer stärker in den Fokus, dass Evaluationen ein bedeutend größeres Spektrum aufweisen als reine Wirksamkeitsbeurteilungen. So tritt vermehrt anstelle des *Beweisens* der ursprüngliche Evaluationszweck des *Lernens* in den Vordergrund des Evaluationsinteresses, da das Verstehen von Veränderungsprozessen mitunter einen viel größeren Benefit produzieren kann als reine Wirksamkeitsbeweise mit bisweilen nur geringer Effektstärke (19,20).

Bevor die Evaluation von komplexen Interventionen erfolgen kann, gilt es u.a. folgende Punkte zu beachten (21):

- (Möglichst) genaue Benennung der Ziele der Evaluation,
- Beschreibung, was bereits über die Intervention bekannt ist,
- Benennung der Zielgruppe der Evaluation (wichtigste Stakeholder oder Forscher),
- Darstellung von Studiendesign und Forschungsmethoden sowie
- Beschreibung der Parameter, die erhoben werden sollen, sowie der Daten, die für die Auswertung verwendet bzw. benötigt werden.

Darüber hinaus sollten auch stets folgende Fragen im Auge behalten werden (22):

- Wie wirkt sich die Evaluation der Intervention auf politische Entscheidungen aus?
- Verschaffen die Ergebnisse der Evaluation gegenüber den vorhandenen wissenschaftlichen Erkenntnissen einen Mehrwert?
- Ist es praktikabel und überhaupt möglich, die Intervention in der zur Verfügung stehenden Zeit oder mit den zur Verfügung stehenden finanziellen Mitteln zu bewerten?

Nicht jede komplexe Intervention erfordert eine komplexe Analyse, es sei denn, die Forschungsfrage verlangt danach (9).

Eine zentrale Frage bei der Bewertung einer komplexen Intervention ist die praktische Wirksamkeit, d.h. ob die Intervention in der täglichen Routine/im Alltag funktionieren kann. In diesem Fall ist es wichtig, die gesamte Bandbreite der Effekte zu verstehen, wie sie sich zwischen allen von der Intervention Betroffenen und zwischen den Settings, im Laufe der Zeit und bezüglich der Ursachen unterscheiden (3).

Um diese Fragen adäquat zu beantworten, gehören alle *aktiven* Komponenten und patientenrelevanten Endpunkte der komplexen Intervention beschrieben und begutachtet sowie gegenseitige Abhängigkeiten zwischen Ergebnisparametern, Prozessen und Zielen analysiert (8).

Bei der Durchführung von Evaluationen komplexer Interventionen sind stets Zusatzbedarfe einzukalkulieren, die sich u.a. durch die Wahl *ungewöhnlicher* Studiendesigns zur Abbildung von Prozessen, Sicherung der Validität durch interne/externe Vergleichsgruppen, Auswertung von Clustern, Durchführung von Matching-Verfahren, Moderatoren-

und Mediatorenkontrolle und der daraus resultierenden Notwendigkeit der Einplanung ausreichender (personeller und monetärer) Ressourcen ergeben (3,23).

Die Evaluation komplexer Interventionen kann eine Kombination aus Dachevaluation und komponentenspezifischen Einzelevaluationen sein (24).

Ergebnis-, Prozess- und gesundheitsökonomische Evaluation – oder mindestens einzelne Teile davon – sollten in jedem Fall in der Evaluation komplexer Interventionen berücksichtigt werden.

1) Ergebnisevaluation

Die Ergebnis- bzw. summative Evaluation dient dem Vergleich zwischen postuliertem Ziel einer (komplexen) Intervention und dem letztlich erreichten Zielzustand und gibt entsprechend Auskunft bezüglich vorab formulierter Outcomes (25). Hierfür bietet sich die Festlegung auf Indikatoren an, die bspw.

- Vorher-Nachher-Vergleiche,
- Soll-Ist-Vergleiche,
- Vergleiche von Entwicklungen über die Zeit,
- Gruppenvergleiche,
- Vergleiche mit hypothetischen Situationen oder
- Vergleiche mit einer theoretischen Referenz

zulassen (25). Bei komplexen Interventionen im Gesundheitswesen sind bspw. Mortalitätsraten, Überlebenszeiten (Kaplan-Meier-Methode) oder (modifizierte) Years of (potential) life lost – Y(P)LL – gängige Kennzahlen in der Ergebnisevaluation (26).

2) Prozessevaluation

Die Prozessevaluation ist bei komplexen Interventionen, insbesondere bei denjenigen, die den Umbau etablierter Strukturen mit entsprechend einhergehender Veränderung von Prozessen zum Ziel haben, unabdingbar. Sie dient der Klärung der Frage, ob die durchgeführte Intervention wie geplant erbracht wird und hilft – so dies negiert werden kann – bspw. Implementierungsfehler oder Schwierigkeiten in der Umsetzung im Versorgungsalltag zu suchen und aufzudecken.

Vor jeder Prozessevaluation ist eine genaue Beschreibung, aus welchen Komponenten die komplexe Intervention besteht

und welche Veränderungen aufgrund von Rückkopplungs-/Feedbackschleifen im Laufe der Zeit neu hinzutreten, zwingend erforderlich. Nur dann kann nach Ende der Intervention abschließend geklärt werden, ob die Veränderungen die gewünschten Outcomes produziert haben oder aber Schwächen im Design des Versorgungskonzeptes oder eine unvollständige Umsetzung zur Nichterreichung bestimmter Ziele geführt haben.

Auch soll die Prozessevaluation darüber Auskunft geben, *wie* und *warum* die Umsetzung erreicht wurde, also bspw. durch spezielle Trainings und besonderen Support von beteiligten Leistungserbringern oder die Schaffung spezieller Management- oder Kommunikationsstrukturen (15). Hier sollte ein entsprechender *Vorher-Nachher*-Vergleich oder ein Vergleich mit der *Standardbehandlung* (Regelversorgung) erfolgen (3).

In Abhängigkeit der Intervention kann dies entweder durch die Überwachung einzelner Implementierungsschritte erfolgen oder auch durch die Messung von potentiellen Mediatoreffekten. Neben unterschiedlichen quantitativen Methoden gehören qualitative (Einzel-)Interviews aber auch Fokusgruppen mit Patienten/Betroffenen, beteiligten Akteuren/Stakeholdern der Intervention und Experten zu den gängigen Methoden (13).

Das Verständnis einer komplexen Intervention ist als einzigartig im Interventionskontext anzusehen. Dies bedeutet, dass eine Evaluation der Wirkmechanismen im Sinne gängiger Kausalitätstheorien schwierig ist und die Prozessevaluation somit entsprechend häufig nur als Fallstudie – sprich Beschreibung der neuen und veränderten Prozesse – umgesetzt werden kann (18).

3) Gesundheitsökonomische Evaluation

Gesundheitsökonomische Evaluationen dienen der Identifizierung und Quantifizierung von Kosten und Nutzen der komplexen Intervention bspw. im Vergleich zur Regelversorgung und verwenden dafür u.a. Daten zu Ressourcenverbrauch, Kosten und Outcomes. Bei Gesundheitsökonomischen Evaluationen lassen sich nicht-vergleichende und vergleichende Studientypen unterscheiden. Zu den

nicht-vergleichenden Studienformen gehören die Kosten-Analyse (*cost analysis*) und die Krankheitskosten-Analyse (*cost-of-illness analysis*). Zu den vergleichenden Studienformen zählen die Kosten-Nutzen-Analyse (*cost-benefit analysis*), die Kosten-Wirksamkeits-Analyse (*cost-effectiveness analysis*) und die Kosten-Nutzwert-Analyse (*cost-utility analysis*) (27,28). In der Kosten-Nutzen-Analyse

Bei komplexen Interventionen können RCTs aus unterschiedlichsten Gründen zumeist nicht unter Realbedingungen durchgeführt werden.

werden alle Kosten- und Nutzenkomponenten einer Intervention in monetären Einheiten abgebildet. Bei der Kosten-Wirksamkeits-Analyse werden die Effekte in natürlichen Maßeinheiten (bspw. Verlängerung des Lebens in gewonnenen Lebensjahren) gemessen. In der Kosten-Nutzwert-Analyse wird der Nutzen einer Intervention aus Sicht des Patienten bewertet, wobei unterschiedliche mehrdimensionale Nutzengrößen wie bspw. Lebensqualität und Lebensdauer zu einem Index zusammengeführt werden (29). Der Effekt einer Intervention lässt sich dann u.a. durch QALYs (*quality adjusted life year*; qualitätskorrigierte Lebensjahre) darstellen (30).

Das Ziel von gesundheitsökonomischen Evaluationen ist es, knappe Ressourcen – sei es nun monetärer, personeller oder zeitlicher Art oder an Ausstattung und Wissen – sinnvoll einzusetzen. Sie erlauben es, potenzielle Auswirkungen von Interventionen zu antizipieren und bestenfalls sogar die tatsächlichen Effekte von (Ver-)Änderungen mit entsprechenden Auswirkungen auf das Gesundheitssystem zu messen. So können gesundheitsökonomische Evaluationen wie die Kosten-Nutzen-Analyse, die Kosten-Wirksamkeits-Analyse und die Kosten-Nutzwert-Analyse ein Instrument für Entscheidungsprozesse im Gesundheitswesen darstellen: Bietet die Intervention einen Mehrwert in Bezug auf die Kosten und den Nutzen verglichen mit dem etablierten Standard (*best practice*) (31,32)?

Mögliche Studiendesigns bei der Evaluation komplexer Interventionen

Bei der Evaluation komplexer Interventionen stehen dem Wissenschaftler prinzipiell viele verschiedene (experimentelle und nicht-experimentelle) Studiendesigns zur Verfügung. Die Auswahl des Studiendesigns hängt letztlich von der jeweiligen Art der komplexen Intervention, den unterschiedlichen zugrundeliegenden Fragestellungen sowie den verschiedenen Umständen, Kontexten und Settings, in denen die Intervention implementiert werden soll, ab (3).

Obwohl sich das MRC dafür ausgesprochen hat, wann immer möglich RCTs – im besten Fall mit Randomisierung auf individueller Ebene (*individually randomised trials*) – durchzuführen, um die interne Validität einer Intervention sicherzustellen, kann bei komplexen Interventionen dieser Goldstandard aus unterschiedlichsten Gründen zumeist nicht unter Realbedingungen durchgeführt werden (3).

Folgende Designs sind der Literatur als Alternativen zum *klassischen* RCT für die Evaluation komplexer Interventionen zu entnehmen:

1) Clusterrandomisierte Studie (*cluster randomised trial*)

Bei einer clusterrandomisierten Studie werden nicht einzelne Personen zufällig auf Interventions- und Kontrollgruppe aufgeteilt, sondern es werden Einheiten (*cluster*) randomisiert (33). Dies kann bspw. auf Ebene von Arztpraxen oder Krankenhäusern geschehen oder aber auch auf höheren Ebenen wie Regionen oder Landkreisen. Die Bildung dieser Cluster, die sich durch bestimmte, gleichartige oder ähnliche Merkmale auszeichnen, findet *vor* der Randomisierung statt. Die Intervention findet schließlich auf Clusterebene statt.

2) Kohortenstudie (*cohort study*)

Die Kohortenstudie als beobachtendes Studiendesign (*observational study*) möchte per definitionem einen Zusammenhang zwischen einer oder mehreren Expositionen und dem Auftreten von Krankheiten untersuchen, wobei exponierte und nichtexponierter Personen über einen bestimmten Zeitraum prospektiv beobachtet werden (34).

Bei komplexen Interventionen kann dieses Studiendesign durchaus sinnvoll eingesetzt werden, wenn eine Randomisierung nicht möglich ist, weil bspw. das Rollout einer Intervention in einer ganzen Region stattfinden soll. In diesem Fall würde die Population dieser Region die Interventionsgruppe und eine Population aus einer anderen, strukturähnlichen Region die Kontrollgruppe darstellen.

Um Interventions- und Kontrollgruppe möglichst gut vergleichbar zu machen, bieten sich unterschiedliche Matching-Verfahren an. Durch die Bildung bzw. den Vergleich von statistischen Zwillingen (*matched pairs*) lässt sich ein möglicher Bias reduzieren und die Effekte der Intervention können besser geschätzt werden (35). Ideal wäre ein **Exact Matching**, welches sich aufgrund einer begrenzten Stichprobe und damit einhergehend begrenzten Daten oftmals nicht realisieren lässt. Bei einem Exact Matching Ansatz werden Personen ausgeschlossen, denen kein Partner

Anstelle von Beweisen muss wieder der ursprüngliche Zweck von Evaluationen, das Lernen, treten.

zugeordnet werden kann, was wiederum zu enormen Störgrößen (*confoundern*) führt (36,37). Um möglichst viele Personen in ein Matching einbeziehen zu können, kann bspw. das sog. **Propensity Score Matching** angewendet werden, das gleichzeitig etwaige Konfundierungseffekte weitestgehend eliminieren hilft, weil nach bekannten Confoundern, z.B. Alter oder Geschlecht, die als Hauptinflussfaktoren auf die Morbidität und dementsprechend die Inanspruchnahme von gesundheitlichen Leistungen gelten, gematcht wird. Die Prämisse des Exact Matchings wird dabei durch die Schaffung einer Kontrollgruppe mit *bestmöglicher Strukturgleichheit* ersetzt (38).

3) Natürliches Experiment (*natural experiment*)

Natürliche Experimente sind Quasi-Experimente und bieten sich gut für die nicht-randomisierte Auswertung komplexer Interventionen an. Sie sind dadurch

gekennzeichnet, dass Personen aufgrund von natürlichen und nicht kontrollierbaren Ereignissen nach dem Zufallsprinzip entweder der Interventions-/Experiment- oder Kontrollgruppe zugeordnet und ohne weiteres Eingreifen seitens der Forschung beobachtet werden. Im Idealfall ähnelt die natürliche Zuteilung der Personen einer Randomisierung. Die Durchführung natürlicher Experimente ist insbesondere dann sinnvoll, wenn eine klar umschriebene Exposition/Intervention für eine bestimmte Subpopulation vorliegt bzw. vorgesehen ist, der eine andere Subpopulation nicht ausgesetzt ist (3,39).

4) Mixed methods Design

Das mixed methods Design ist ein relativ neuer Ansatz bei der Evaluation komplexer Interventionen. Es kombiniert quantitative und qualitative Forschungsmethoden miteinander. Dabei wird davon ausgegangen, dass die Kombination beider Ansätze im jeweiligen Kontext bessere Ergebnisse liefert als nur eine dieser Methoden allein und dass dadurch die entsprechende(n) Forschungsfrage(n) besser beantwortet werden kann. Somit können die Stärken der jeweiligen Methode genutzt und betont sowie etwaige Schwächen eliminiert werden. So können qualitative Methoden

wie bspw. Expertengespräche mit Forschern und Praktikern, Fokusgruppen oder Interviews mit der Zielpopulation quantitative Forschungsergebnisse sinnvoll ergänzen (23,40). Neben dieser Komplementarität zählen auch Praktikabilität, Stärkung von Glaub- und Vertrauenswürdigkeit sowie schrittweise Erhöhung des Wissens – bedingt durch Rückkopplungs-/Feedbackschleifen – zu den Vorteilen von mixed methods Ansätzen (41).

Im Vergleich zu anderen Studiendesigns ermöglicht es der mixed methods Ansatz die Forschungsfrage(n) zu modifizieren oder zu ändern. Darüber hinaus können im Laufe der Zeit auch neue Fragen – bspw. aufgrund von Erkenntnissen der Rückkopplungs-/Feedbackschleifen – hinzutreten (23). In diesem Zusammenhang erfahren auch sog. **zusammengesetzte Endpunkte** (*composite endpoints*) eine zunehmende Verwendung. Diese werden nach einem zuvor genau definierten Algorithmus aus mehreren Endpunkten (z.B.

Todesfälle, Herzinfarkte oder Schlaganfälle) zusammengesetzt und in toto sowie individuell in die abschließende Auswertung einbezogen. Durch zusammengesetzte Endpunkte wird die Ereignisrate erhöht und eine Reduktion von Beobachtungszeit und Fallzahl ermöglicht (42,43).

Die Zusammenführung von Daten aus quantitativer und qualitativer Erhebungen kann in jeder Phase der Evaluation erfolgen, z.B. während der Datenerhebung, in der Datenanalyse oder auch erst in der Interpretationsphase (23).

Da der mixed methods Ansatz – dessen Durchführung ein multiprofessionelles Forschungsteam voraussetzt – eine noch recht junge und anspruchsvolle Disziplin ist, sind erst wenige Veröffentlichungen zur Erarbeitung und Entwicklung von geeigneten Forschungsfragen oder (guten) Praxisbeispielen vorhanden. Dies bedeutet auch, dass sich Forschungsparadigmen noch in der Entwicklung befinden (23).

Weitere Studiendesigns, die u.a. vom MRC (3) für die Evaluation komplexer Interventionen für geeignet gehalten werden, sind

- Wartegruppensdesign (*wait list control group design*) bzw. **stepped wedge Design**,
- **Comprehensive cohort Design**,
- **Interrupted time series Design**,
- **Multiple Baseline Design** oder
- **N-of-1 Design**.

Fazit

Allein in den ersten drei Förderwellen der neuen Versorgungsformen des Innovationsfonds wurde 81 Projekten ein Förderzuschlag erteilt (44), von denen viele die Definition einer komplexen Intervention gemäß MRC erfüllen dürften (u.a. mehrere, sich wechselseitig bedingende Einzelkomponenten oder Etablierung in komplexen Kontexten und Settings).

Die Innovationsfondsprojekte der ersten Förderwellen befinden sich bereits in der Umsetzungsphase, wobei diese durch eine intensive Kommunikation mit dem Projektträger begleitet werden. Im Zentrum der Dialoge stehen neben administrativen vor allem auch Fragen zu Evidenz oder dem richtigen Evaluationsdesign. Dies unterstreicht, wie wichtig ein breitgefächertes Diskurs über komplexe Interventionen in der deutschen Forschungsgemeinschaft ist.

So herrscht vielfach immer noch die Meinung vor, RCTs – in der evidenzbasierten Medizin als Goldstandard beim Nachweis von Wirksamkeit angesehen – seien auch bei komplexen Interventionen unverzichtbar. Das RCTs bei der Evaluation solcher komplexen Interventionen oft, sei es aus ethischen, wirtschaftlichen oder praktikablen Gründen, nicht angewendet werden können und darüber hinaus der Komplexität dieser Interventionen nicht gerecht werden, wird dabei häufig übersehen. Ebenso wird oftmals außer Acht gelassen, dass vor Beginn einer komplexen Intervention mitunter noch gar nicht alle Komponenten feststehen oder sich Bestandteile wegen Rückkopplungs-/Feedbackschleifen ändern können. Eine ad hoc Festlegung auf vorab fest definierte Zielkriterien führt fast zwangsläufig dazu, dass wichtige Aspekte übersehen und vielleicht irrelevante Parameter starr gemessen und gesammelt werden. Die wesentlichen Aspekte jedoch, die die komplexe Intervention auszeichnen, können somit

übersehen und gar nicht betrachtet werden. Die Tatsache, dass sich Effekte von einzelnen Komponenten solcher komplexen Interventionen aufgrund enger Verknüpfungen oft nicht ausdifferenzieren lassen, zeigt, wie schwierig sich die Evaluation komplexer Interventionen gestalten kann.

Gerade im Zusammenhang mit der Evaluation der komplexen Innovationsfondsprojekte bietet sich in Deutschland aus wissenschaftlicher Perspektive die Möglichkeit, sich vom Mantra bestimmter vermeintlicher Standards zu lösen und nach neuen, pragmatischeren Wegen und Methoden zu suchen, um der Evaluation komplexer Interventionen gerecht zu werden. Denn weder national noch international gibt es gegenwärtig einen Konsens über *best practice*, weshalb die Formulierung einer *one-size-fits-all*-Lösung nicht stattfinden kann.

Wie das MRC richtig formuliert hat, können die besten verfügbaren Methoden, auch wenn sie theoretisch nicht optimal sind, nützliche Ergebnisse liefern.

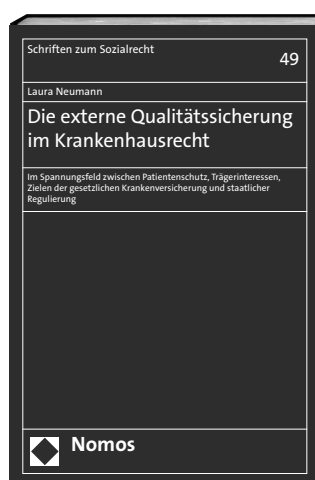
Eine Evaluation muss gut durchdacht durchgeführt werden, damit keine vielversprechenden Interventionen als ineffizient abgelehnt werden. Insbesondere aufgrund des knappen Zeitrahmens, in denen die Innovationsfondsprojekte durchgeführt und evaluiert werden (drei bis vier Jahre Projektlaufzeit), ist es schwierig vorherzusagen, ob tatsächlich vielen der Projekte das Potential beschiedenen werden kann, in die Regelversorgung überführt zu werden, da komplexe Interventionen eigentlich ein langfristiges Follow-up benötigen, welches Rückschlüsse darüber erlaubt, ob kurzfristige Veränderungen auch längerfristig bestehen bleiben.

Anstelle von *Beweisen* muss wieder der ursprüngliche Zweck von Evaluationen, das *Lernen*, treten, damit aus komplexen Interventionen der größtmögliche Nutzen für die Neubewertung der Regelversorgung und für künftige Vorhaben gezogen werden kann. ■

Literatur

- Innovationsausschuss beim Gemeinsamen Bundesausschuss; Der Innovationsfonds und der Innovationsausschuss beim Gemeinsamen Bundesausschuss [Internet]. 2017 [cited 2017 Mar 6]. Available from: <https://innovationsfonds.g-ba.de>
- Innovationsausschuss beim Gemeinsamen Bundesausschuss; Förderung neuer Versorgungsformen [Internet]. 2018 [cited 2018 Oct 18]. Available from: <https://innovationsfonds.g-ba.de/versorgungsformen>
- Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. 2008 Sep;337:a1655.
- Petticrew M. When are complex interventions “complex”? When are simple interventions “simple”? *Eur J Public Health*. 2011 Aug;21(4):397–8.
- Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ*. 2000 Sep;321(7262):694–6.
- Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. *BMJ*. 2008 Jun;336(7656):1281–3.
- Datta J, Petticrew M. Challenges to evaluating complex interventions: a content analysis of published papers. *BMC Public Health*. 2013;13(1):568.
- Mühlhauser I, Lenz M, Meyer G. Entwicklung, Bewertung und Synthese von komplexen Interventionen – eine methodische Herausforderung. In: Robert Koch-Institut, Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit, editors. *Evaluation komplexer Interventionsprogramme in der Prävention: Lernende Systeme, lehrreiche Systeme?* Beiträge zur Gesundheitsberichterstattung des Bundes. Berlin: RKI; 2012. p. 43–55.
- Tsiachristas A, Rutten-van Molken MP. Evaluating Complex Interventions. In: Amelung VE, Stein V, Goodwin N, Balicer R, Nolte E, Suter E, editors. *Handbook Integrated Care*. Heidelberg: Springer; 2017. p. 297–313.
- Best M, Neuhauser D, Walter A, Shewhart, 1924, and the Hawthorne factory. *Qual Saf Health Care*. 2006;15(2):142–3.
- Blettner M, Dierks M-L, Donner-Banzhoff N, Hertrampf K, Klusen N, Köpke S, et al. Überlegungen des Expertenbeirats zu Anträgen im Rahmen des Innovationsfonds. *Z Evid Fortbild Qual Gesundhwes*. 2018;130:42–8.
- Hernán MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, et al. Observational Studies Analyzed Like Randomized Experiments. *Epidemiology*. 2008 Nov;19(6):766–79.
- Richards DA, Hallberg I. The complex interventions framework. In: Richards DA, Rahm Hallberg I, editors. *Complex interventions in health: an overview of research methods*. London, New York: Routledge; 2015. p. 1–15.
- Walach H. Zirkulär statt hierarchisch – methodische Überlegungen zur Evaluation komplexer therapeutischer Maßnahmen. In: Bödeker W, Kreis J, editors. *Evidenzbasierung in Gesundheitsförderung und Prävention*. Bremerhaven: Wirtschaftsverlag NW, Verlag für neue Wissenschaft; 2006. p. 33–56.
- Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. 2015 Mar;350:h1258.
- Weßling H. *Theorie der klinischen Evidenz – Versuch einer Kritik der evidenzbasierten Medizin*. 7th ed. Berlin, Münster, Wien, Zürich, London: LIT Verlag; 2011.
- Rossi PH, Lipsey MW, Freeman HE. *Evaluation: A Systematic Approach*. Thousand Oaks: Sage Publications; 2003.
- Bödeker W. Wirkungen und Wirkungsnachweis bei komplexen Interventionen. In: Robert-Koch-Institut, Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit, editors. *Evaluation komplexer Interventionsprogramme in der Prävention: Lernende Systeme, lehrreiche Systeme?* Beiträge zur Gesundheitsberichterstattung des Bundes. Berlin: RKI; 2012. p. 33–41.
- Elkeles T. Evaluation von Gesundheitsförderung und die Forderung nach Evidenzbasierung: fünf Thesen zur Anwendbarkeit auf Gesundheit. *Zeitschrift für Eval*. 2006;(1):39–70.
- Stevens W. Economic evaluation of health promotion interventions. In: Thorogood M, Coombes Y, editors. *Evaluating health promotion: practice and methods*. 2nd ed. Oxford: Oxford Medical Publications; 2004. p. 41–55.

- 21.** Lamont T, Barber N, de Pury J, Fulop N, Garfield-Birkbeck S, Lilford R, et al. New approaches to evaluating complex health and care systems. *BMJ*. 2016;352:i154.
- 22.** Ogilvie D, Cummins S, Petticrew M, White M, Jones A, Wheeler K. Assessing the evaluability of complex public health interventions: Five questions for researchers, funders, and policymakers. *Milbank Q*. 2011;89(2):206–25.
- 23.** Borglin G. The value of mixed methods for researching complex interventions. In: *Complex Interventions in Health: An Overview of Research Methods*. London, New York: Routledge; 2015. p. 29–45.
- 24.** Kuhn J, Lampert T, Ziese T. Einführung ins Thema: Komplexe Interventionen – komplexe Evaluationen? In: *Evaluation komplexer Interventionsprogramme in der Prävention: Lernende Systeme, Lehrreiche Systeme? Beiträge zur Gesundheitsberichterstattung des Bundes*. Berlin: RKI; 2012. p. 9–14.
- 25.** Scriven M. Die Methodologie der Evaluation. In: Wulf C, editor. *Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen*. München: R. Piper & Co. Verlag; 1972. p. 60–91.
- 26.** Schulte T, Pimperl A, Fischer A, Dittmann B, Wendel P, Hildebrandt H. Ergebnisqualität Gesundes Kinzigtal – quantifiziert durch Mortalitätskennzahlen – Eine quasi-experimentelle Kohortenstudie: Propensity Score-Matching von Eingeschriebenen vs. Nicht-Eingeschriebenen der Integrierten Versorgung Gesundes Kinzigtal auf Basis [Internet]. 2014 [cited 2017 Mar 7]. Available from: <https://optimedis.de/files/Publikationen/Studien-und-Berichte/2014/Mortalitaetsstudie-2014/Mortalitaetsstudie-2014.pdf>
- 27.** Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the economic evaluation of health care programme*. 3rd ed. Oxford: Oxford University Press; 2005.
- 28.** Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—Explanation and Elaboration: A Report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value Heal*. 2013;16(2):231–50.
- 29.** Schöffski O. Grundformen gesundheitsökonomischer Evaluation. In: *Gesundheitsökonomische Evaluationen*. 4th ed. Berlin, Heidelberg: Springer; 2012. p. 43–70.
- 30.** Payne K, Thompson A. Economic evaluations of complex interventions. In: Richards DA, Rahm Hallberg I, editors. *Complex interventions in health: an overview of research methods*. London, New York: Routledge; 2015. p. 248–57.
- 31.** Tsiachristas A, Rutten-van Mölken MP. Economic Evaluation of Integrated Care. In: Amelung VE, Stein V, Goodwin N, Balicer R, Nolte E, Suter E, editors. *Handbook Integrated Care*. Heidelberg: Springer; 2017. p. 315–32.
- 32.** Schöffski O. Einführung. In: *Gesundheitsökonomische Evaluationen*. 4th ed. Berlin, Heidelberg: Springer; 2012. p. 3–11.
- 33.** Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012;345(sep04 1):e5661–e5661.
- 34.** Gordis L. Using Epidemiology to Identify the Cause of Disease. In: *Epidemiology*. 5th ed. Philadelphia PA: Saunders Elsevier; 2014. p. 177–305.
- 35.** Gensler S, Skiera B, Böhm M. Einsatzmöglichkeiten der Matching Methode zur Berücksichtigung von Selbstselektion. *J für Betriebswirtschaft*. 2005;55(1):37–62.
- 36.** Stuart EA. Matching Methods for Causal Inference: A Review and a Look Forward. *Stat Sci*. 2010;25(1):1–21.
- 37.** Rosenbaum PR, Rubin DB. Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *Am Stat*. 1985;39(1):33.
- 38.** Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*. 1983 Apr;70(1):41.
- 39.** DiNardo J. Natural experiments and quasi-natural experiments. In: Durlauf SN, Blume LE, editors. *The New Palgrave Dictionary of Economics*. 2nd ed. Basingstoke: Palgrave Macmillan; 2008. p. 856–9.
- 40.** Hardeman W, Sutton S, Griffin S, Johnston M, White A, Wareham NJ, et al. A causal modelling approach to the development of theory-based behaviour change programmes for trial evaluation. *Health Educ Res*. 2005;20(6):676–87.
- 41.** Creswell JW. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 2nd ed. Thousand Oaks: Palgrave Macmillan; 2003.
- 42.** Hofman CS, Makai P, Boter H, Buurman BM, de Craen AJM, Olde Rikkert MGM, et al. Establishing a composite endpoint for measuring the effectiveness of geriatric interventions based on older persons' and informal caregivers' preference weights: a vignette study. *BMC Geriatr*. 2014;14:51.
- 43.** Kleist P. Composite Endpoints for Clinical Trials. *Int J Pharm Med*. 2007;21(3):187–98.
- 44.** Innovationsausschuss beim Gemeinsamen Bundesausschuss. Geförderte Projekte Neue Versorgungsformen [Internet]. 2018 [cited 2018 Oct 10]. Available from: <https://innovationsfonds.g-ba.de/projekte/neue-versorgungsformen/>



Die externe Qualitätssicherung im Krankenhausrecht

Im Spannungsfeld zwischen Patientenschutz, Trägerinteressen, Zielen der gesetzlichen Krankenversicherung und staatlicher Regulierung

Von RAin Dr. Laura Neumann

2019, ca. 330 S., brosch., ca. 86,– €

ISBN 978-3-8487-5327-7

eISBN 978-3-8452-9462-9

(Schriften zum Sozialrecht, Bd. 49)

Erscheint ca. Januar 2019

nomos-shop.de/39889

Die Qualitätssicherung steht im Vordergrund der aktuellen Gesundheitsgesetzgebung. Die Autorin stellt Ziele und Instrumente der Qualitätssicherung im Krankenhaussektor dar und untersucht deren Umsetzung durch das Krankenhaus- und Sozialrecht einschließlich verfassungsrechtlicher Bezüge.



Unser Wissenschaftsprogramm ist auch online verfügbar unter: www.nomos-elibrary.de

Bestellen Sie jetzt telefonisch unter (+49)7221/2104-37.

Portofreie Buch-Bestellungen unter www.nomos-shop.de

Alle Preise inkl. Mehrwertsteuer



Nomos