

Im Blickpunkt

# Entwicklung, Bewertung und Synthese von komplexen Interventionen – eine methodische Herausforderung

Ingrid Mühlhauser<sup>1,\*</sup>, Matthias Lenz<sup>1</sup>, Gabriele Meyer<sup>2</sup>

<sup>1</sup>Universität Hamburg, MIN-Fakultät, Gesundheitswissenschaften, Hamburg

<sup>2</sup>Universität Witten/Herdecke, Fakultät für Gesundheit, Department für Pflegewissenschaft, Witten

## Zusammenfassung

**Hintergrund:** Viele medizinische Maßnahmen sind komplexe Interventionen. Sie bestehen aus Einzelkomponenten, die sich wechselseitig bedingen und unterschiedlich mit Implementierungskontexten interagieren, z.B. *Stroke Units*. Im Vergleich zu Einzelmaßnahmen, wie der Behandlung mit einem Medikament, ist die Beurteilung von Wirksamkeit, Nutzen und Schaden komplexer Interventionen schwerer zu erschließen. Differenzierte methodische Verfahren zur Entwicklung, Bewertung und Synthese komplexer Interventionen werden international diskutiert.

**Methoden:** Systematische Bestandsaufnahme methodischer Leitfäden zur Entwicklung und Evaluation komplexer Interventionen mittels Recherchen in wissenschaftlichen Literaturdatenbanken und Webressourcen internationaler Institute. Aufbereitung des Diskussionsstandes und der Vorschläge zu Entwicklung, Bewertung und Synthese von komplexen Interventionen.

**Ergebnisse:** Fünf methodische Leitfäden wurden identifiziert. Eine strukturierte Anleitung zur Entwicklung komplexer Interventionen im

Gesundheits- und Medizinbereich bietet vor allem die Leitlinie des *UK Medical Research Councils*. Die Synthese komplexer Interventionen in systematischen Übersichtsarbeiten und ihren Meta-Analysen gemäß üblicher Methoden liefert unvollständige Informationen. Andere Ansätze der Synthese wie Realist Evaluation oder statistische Verfahren zur Bestimmung der aktiven Komponenten komplexer Interventionen werden diskutiert. Allerdings können diese post hoc-analytischen Verfahren kein adäquater Ersatz für eine sorgfältige prospektive Entwicklung und Exploration der komplexen Intervention und der Interdependenzen mit kontextuellen Faktoren sein.

**Schlussfolgerung:** Komplexe Interventionen machen eine mehrstufige Entwicklung, Benutzung verschiedener Methoden, Berichterstattung über alle Entwicklungsphasen und neue Formen der Synthese notwendig. Die Darstellung der Gesamtevidenz für einzelne komplexe Interventionen kann relevanter und nachvollziehbarer sein, als die Synthese unterschiedlicher komplexer Interventionen mit den üblicherweise angewendeten (meta-analytischen) Verfahren systematische Übersichtsarbeiten.

**Schlüsselwörter:** Komplexe Interventionen, Evaluationsstudien, systematische Reviews, Metaanalysen

\*Korrespondenzadresse: Univ.-Prof. Dr. med. Ingrid Mühlhauser, Universität Hamburg, Fakultät für Mathematik, Informatik, Naturwissenschaften (MIN), Institut für Pharmazie, Gesundheitswissenschaften, Martin-Luther-King Platz 6, D-20146 Hamburg. Tel.: +49 40/42838-3988; Fax: +49 40/42838-3732. E-Mail: [Ingrid.Muehlhauser@uni-hamburg.de](mailto:Ingrid.Muehlhauser@uni-hamburg.de) (I. Mühlhauser).

# Development, appraisal and synthesis of complex interventions – a methodological challenge

## Summary

**Background:** Many medical interventions are of a complex nature. They comprise interdependent components differently interacting within various complex settings, e.g., stroke units. Appraising the efficacy, benefit and harm of complex interventions is far more difficult than appraising single interventions like specific drug treatments. Detailed methodological procedures for the development, appraisal and synthesis of complex interventions are increasingly discussed internationally.

**Methods:** Systematic inventory of methodological guidance for the development and evaluation of complex interventions through searching scientific literature databases and web resources of international institutes. Review of the current state of discussion and suggestions for the development, appraisal and synthesis of complex interventions.

**Results:** Five methodological guidance papers have been identified. In particular, the UK Medical Research Council's guidance offers a structured in-

roduction to the development and evaluation of complex interventions in health care and medicine. Synthesis of complex interventions using customary methods of systematic reviews is not satisfying. Other approaches to synthesis like realist evaluation as well as statistical procedures exploring the active components of complex interventions have been discussed. However, post hoc analytical procedures could never adequately replace careful prospective development and exploration of complex interventions and interdependencies with contextual factors.

**Conclusion:** Complex interventions require multi-stage development, use of different methods, reporting on all developing phases and new approaches for synthesis. Presentation of the complete evidence on a specific complex intervention might be more useful than synthesis of a variety of different complex interventions by customarily applied methods of (meta-analytical) systematic review.

**Key words:** complex interventions, evaluation studies, systematic reviews, meta-analyses

## Einleitung

Viele medizinische Maßnahmen sind komplexe Interventionen. Sie bestehen aus mehreren Einzelkomponenten, die sich wechselseitig bedingen und ihrerseits in komplexe Kontexte implementiert werden. Beispiele dafür sind *Stroke Units*, *Disease Management* Programme oder Projekte zur Verbesserung der Krankenhaushygiene. Ähnliche komplexe Interventionen gibt es in assoziierten Berufs- und Handlungsfeldern. Zum Beispiel Prävention von Sturz und Dekubitus in der Pflege, Ernährungs- und Sportprogramme in Schulen, Prävention posttraumatischer Störungen, Früherkennung von Kindesmisshandlung und -verwahrlosung, Verringerung von Jugendkriminalität, Prävention von Unfällen im Straßenverkehr oder Web-basierte Lernprogramme.

Die Evaluation von Einzelmaßnahmen, wie die Behandlung mit einem Medikament, in randomisierten kontrollierten Studien (RCT) und deren systematische Übersichtsarbeiten ist vergleichsweise einfach. Die Wirksamkeit und der Nutzen und Schaden von komplexen Interventionen hingegen ist sehr viel schwerer zu erschließen. Der Beitrag der Einzelkomponenten zum Gesamtergebnis und die Interaktionen mit dem Setting bleiben häufig unklar. Seit einigen Jahren werden darum differenzierte methodische Verfahren zur

Entwicklung, Bewertung und Synthese von komplexen Interventionen diskutiert.

Ziel dieses Artikels ist es, den Unterschied zwischen Einzelinterventionen und komplexen Interventionen herauszuarbeiten sowie methodische Ansprüche an die Entwicklung, Bewertung und Synthese von komplexen Interventionen zur Diskussion zu stellen.

## Methodische Leitfäden zur Entwicklung und Evaluation komplexer Interventionen

Im Juni 2011 wurde systematisch nach Methodenpapieren zur Entwicklung und Evaluation komplexer Interventionen recherchiert. Die Datenbanken PubMed, Embase und PsycINFO wurden unter Verwendung von Suchbegriffen wie „complex intervention\*“, „multifaceted intervention\*“ in Kombination mit „methods“ durchsucht. Die vollständige Recherchestrategie ist auf Anfrage bei den Autoren erhältlich. Die Internetseiten der *Campbell Collaboration*, der *Cochrane Collaboration*, des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), des Deutschen Instituts für Medizinische Dokumentation und Information (DIMDI), des britischen *Social Care Institute for Excellence* (SCIE), des

britischen *Centre for Reviews and Dissemination* (CRD), des britischen *Medical Research Council* (UKMRC), des britischen *National Institute for Health and Clinical Excellence* (NICE), des *International Network of Agencies for Health Technology Assessment* (INAHTA), des US-amerikanischen *Institute of Medicine* (IOM) und der *Agency for Healthcare Research and Quality* (AHRQ) wurden gesichtet. Eingeschlossen wurden Methodenpapiere bzw. methodische Leitfäden zur Entwicklung und Evaluation komplexer Interventionen, die die methodischen Herausforderungen deren Komplexität übergreifend berücksichtigen. Ausgeschlossen wurden Publikationen, die sich speziell mit Teilmethodiken (z.B. Methodik kontrollierter Studien oder Fokusgruppen) beschäftigen, sowie Übersichtsarbeiten, die sich allgemein mit Forschungsmethoden auseinandersetzen.

Die systematische Datenbankrecherche erzielte 1261 Treffer. Aus eigenen Publikationsbeständen und nach Screening der Internetseiten kamen sieben Publikationen hinzu. Nach Screening von Titeln und Abstracts wurden 1225 Treffer ausgeschlossen und 43 Publikationen im Volltext begutachtet. Das Flussdiagramm nach PRISMA und die Ergebnistabellen der Recherche sind auf Anfrage bei den Autoren erhältlich. Davon entsprachen 38 Arbeiten nicht den Einschlusskriterien. Hierzu zählten neben

forschungsmethodischen Übersichtsartikeln [1–10] Publikationen, die sich mit spezifischen methodischen Aspekten beschäftigen: 1) quantitative Methoden in Übersichtsarbeiten [11,12], 2) qualitative neben quantitativen Methoden bzw. „Mixed Methods“ [13–25], 3) „Realist Review“ [26,27] und 4) Evaluation komplexer Interventionen im Rahmen von RCTs und Prozessevaluation [28–30].

Fünf relevante methodische Leitfäden wurden identifiziert: 1) Cochrane „Handbook for Systematic Reviews of Interventions“ [31], 2) Leitlinien „Developing and evaluating complex interventions“ des UKMRC [32], 3) „Guidance for undertaking reviews in health care“ des CRD [33], 4) „Systematic research reviews“ des SCIE [34] und 5) „Systematic Reviews in Social Sciences“ von Petticrew & Roberts [35].

Die fünf Leitfäden überschneiden sich inhaltlich. Eine strukturierte Anleitung zur Entwicklung komplexer

Interventionen im Gesundheits- und Medizinbereich bietet vor allem die Leitlinie des UKMRC [32]. Die Methodenpapiere des CRD [33] und des SCIE [34] verweisen auf diese Leitlinie. Das Cochrane Handbook [31] richtet sich vor allem an Review-Autoren. Petticrew & Roberts diskutieren in ihrem als Buch erschienenen Leitfaden „Systematic Reviews in Social Sciences“ [35] wesentliche Aspekte aus den anderen Methodenpapieren aus sozialwissenschaftlicher Perspektive. Mark Petticrew ist zudem Ko-Autor der Leitfäden des UKMRC, CRD und SCIE. Alle fünf Leitfäden sehen die Berücksichtigung qualitativer und quantitativer Methoden für die Entwicklung und Evaluation komplexer Interventionen in sehr ähnlicher Weise vor. Die zentralen methodischen Aspekte haben wir aus den fünf Leitfäden extrahiert und in einen konzeptionellen Rahmen integriert (Abb. 1). Der konzeptionelle Rahmen soll die Zusammenhänge zwischen den wesentlichen methodischen Aspekten

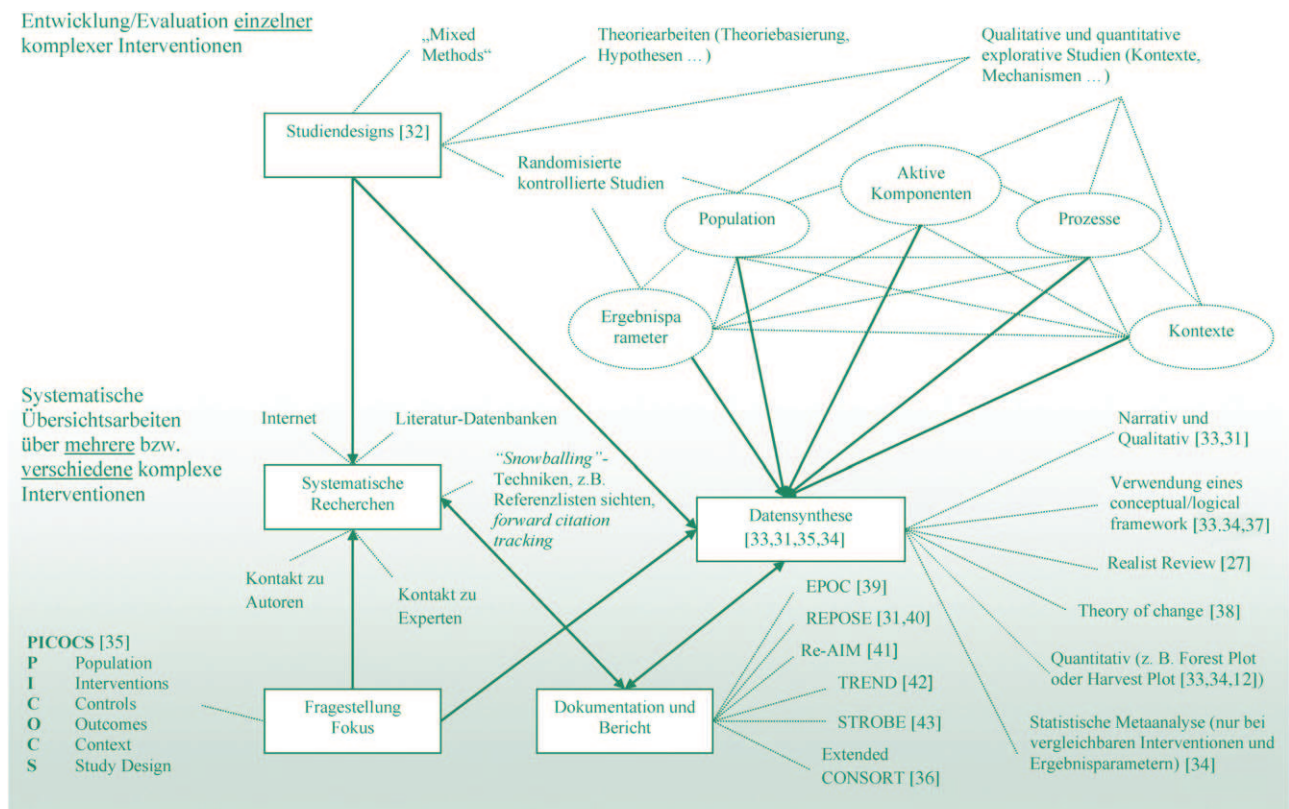
visualisieren. Er beinhaltet zusätzlich Referenzen zu methodischen Teilverfahren, z.B. CONSORT Statement [36].

## Sind nicht alle Interventionen komplex?

Ein Arzneimittel lässt sich als singuläre Intervention definieren. Mit der Zulassung durch die Arzneimittelbehörde sind die Entwicklungs- und Prüfphasen I bis III abgeschlossen, das Arzneimittel liegt somit in standardisierter Form vor. Die Wirksamkeit wird in RCTs durch verblindeten Vergleich mit einem Placebo-Präparat oder Standardmedikament nachgewiesen.

Es folgen kontrollierte Phase IV Studien zur Implementierung der neuen medizinischen Behandlung einschließlich Dokumentation der Sicherheit und Nutzen-Kosten Analysen.

Genau genommen besteht auch eine Arzneimittelbehandlung aus mehreren Komponenten. Die klinischen Effekte



\* CONSORT=Consolidated Standards of Reporting Trials; EPOC=Effective Practice and Organisation of Care; RE-AIM=Reach, Efficacy, Adoption, Implementation, and Maintenance; REPOSE=Reporting of primary empirical research Studies in Education; STROBE=Strengthening the Reporting of Observational Studies in Epidemiology; TREND=Transparent Reporting of Evaluations with Nonrandomized Designs

Abb. 1 Eigener Vorschlag zu einem konzeptionellen Rahmen der Entwicklung und Evaluation komplexer Interventionen (vgl. [37–43]).

sind von verschiedenen Variablen abhängig, wie z.B. pharmazeutische Zubereitung, Applikationsform, Dosierung und Dauer der Behandlung, Patientenpopulation, spezifischem Krankheitsbild und Begleittherapie. Hinzu kommen Interaktionen mit anderen Arzneimitteln oder pathophysiologische Veränderungen wie eine Niereninsuffizienz, wie sie bei fast allen Medikamentenbehandlungen zu berücksichtigen sind.

Zudem kann auch in qualitativ hochwertigen RCTs die klinische Wirksamkeit eines Arzneimittels durch weitere Begleitfaktoren beeinflusst werden. So war z.B. in der Women's Health Initiative die postmenopausale Hormonbehandlung mit niedrigeren Blutzucker- und Cholesterinwerten assoziiert. Frauen in der Placebogruppe hatten folglich häufiger Statine erhalten, was zu einer Interaktion auf den primären kardiovaskulären Endpunkt führte [44]. Medikamentenspezifische Beschwerden oder Komplikationen, wie das gehäufte Auftreten von gynäkologischen Blutungen unter Hormonbehandlung, können zu einer Intensivierung von Diagnostik und Versorgung führen. Auch die Effektstärke des Placebos bzw. der Standardtherapie beeinflusst das Ergebnis. Die in einem bestimmten RCT klinisch nachweisbare Wirksamkeit eines spezifischen Arzneimittels hängt somit vom Design der Studie, der methodischen Qualität der Studie, den Eigenschaften der Patientenpopulation und kontextuellen Faktoren ab. Die Bedingungen und Voraussetzungen bei der Evaluation von Medikamenten können jedoch weitestgehend standardisiert werden.

## Wann sind komplexe Interventionen komplex?

Komplexe Interventionen bestehen aus mehreren interdependenten Komponenten. Ein Beispiel sind strukturierte Behandlungs- und Schulungsprogramme zur Insulintherapie von Patienten mit Diabetes Typ 1 [45]. Auch hier gibt es Arzneimittel-spezifische Faktoren. So hängt die Blutzuckerwirksamkeit von der Pharmakokinetik des Insulinpräparats ab, ob die Insuline variabel

zu mischen sind und wann sie wie oft, mit welchem Spritz-Ess-Abstand, in welcher Dosierung verabreicht werden. Andere Einflussfaktoren sind jedoch von vergleichsweise größerer Bedeutung. Da die Insuline von den Patienten selbst appliziert und dosiert werden, wird die Wirksamkeit und Sicherheit der Insulinbehandlung entscheidend von den Möglichkeiten und Fähigkeiten des Patienten bestimmt, diese erfolgreich durchzuführen.

Zur Komponente Arzneimittel kommen also Komponenten wie Therapie-schemata, Patientenschulung, Qualifikation und Motivation des Schulungs- und Behandlungsteams und Gesundheitssystembedingungen. Interdependenzen ergeben sich, da die Interventionskomponenten voneinander abhängen. Krankheits- und behandlungsbezogenes Wissen ist unverzichtbar, es reicht jedoch nicht aus. Die vermittelten Inhalte müssen Evidenz-basiert und handlungsrelevant sein. Blutzucker-/Stoffwechselfbstkontrollen sind unverzichtbar, jedoch kein Selbstzweck.

Nur wenn die Patienten die Messungen korrekt durchführen, die Werte interpretieren können und durch adäquate Therapieanpassung reagieren, sind sie sinnvoll und nützlich. Eine Flexibilisierung des Tagesablaufs einschließlich variabler Nahrungszufuhr mit dem Ziel einer guten Blutzuckereinstellung ist nur durch eine angemessene, durch Patienten selbst gesteuerte Insulindosierung möglich. Ob die Therapie tatsächlich erfolgreich ist, hängt davon ab, ob und wie Stoffwechselfbstkontrollen durchgeführt werden, ob die Patienten passende (wirksame und sichere) Dosierungsregeln erhalten haben bzw. sich selbst erschließen, ob die Materialien ausreichend verfügbar sind und ob die Patienten die Motivation, die Kompetenzen und das Selbstvertrauen besitzen, die Therapien korrekt durchzuführen. Die Einstellungen und Haltungen des Schulungspersonals bzw. Behandlungsteams und die Voraussetzungen des jeweiligen Gesundheitssystems sind grundlegende Determinanten. Die Tabelle 1 illustriert die

**Tabelle 1.** Komponenten, die den Erfolg eines Diabetes-Schulungsprogramms zur Verbesserung der Blutzuckereinstellung bestimmen können – ein Beispiel.

Komponenten	Inhaltliche Determinanten
<b>Definition Patientenschulung</b> Zugrundeliegende Theorie/n?	<ul style="list-style-type: none"> <li>• Was verstehen die Autoren darunter?</li> <li>• Was verstehen die Studienutzer darunter?</li> </ul>
<b>Ziele der Patientenschulung</b> Therapieziele?	<ul style="list-style-type: none"> <li>• Compliance vs. Selbstmanagement?</li> <li>• Arzt- vs. Patienten-definiert?</li> </ul>
<b>Inhalte Patientenschulung</b>	<ul style="list-style-type: none"> <li>• Evidenz-basiert?</li> <li>• Relevant?</li> <li>• Vollständig?</li> </ul>
<b>Komponenten Patientenschulung</b>	<ul style="list-style-type: none"> <li>• Stoffwechselfbstkontrolle – welche, wie?</li> <li>• Strikte Diät vs. freie Diät?</li> <li>• Selbstanpassung der Therapie durch Patienten?</li> <li>• Therapieverfahren?</li> </ul>
<b>Qualifikation Patientenschulung</b>	<ul style="list-style-type: none"> <li>• Wer schult?</li> <li>• Ausbildung der Schulenden?</li> <li>• Motivation, Einstellung und Überzeugungen des Behandlungsteams?</li> </ul>
<b>Organisation Patientenschulung</b>	<ul style="list-style-type: none"> <li>• Einzelschulung vs. Gruppenschulung</li> <li>• Ambulant – stationär?</li> <li>• Wie viele Unterrichtseinheiten?</li> </ul>
<b>Bedingungen Patientenschulung</b>	<ul style="list-style-type: none"> <li>• Gesundheitssystem?</li> <li>• Was ist verfügbar?</li> <li>• Was wird bezahlt?</li> </ul>
<b>Definition „Verbesserung der Blutzuckereinstellung“</b>	<ul style="list-style-type: none"> <li>• Individualisierte Therapieziele?</li> <li>• Intensität der medikamentösen Behandlung?</li> <li>• Liberalisierung von Diät und Lebensstil?</li> </ul>
<b>Heterogenität der Patientengruppen</b>	<ul style="list-style-type: none"> <li>• Ersts Schulung nach Diagnose</li> <li>• Wiederholungsschulungen</li> <li>• Bei Spätschäden</li> </ul>

Komponenten, die den Erfolg eines Diabetes-Schulungsprogramms zur Verbesserung der Blutzuckereinstellung bestimmen können (Tab. 1).

Einerseits kann bereits eine einzelne Komponente einer solchen multimodalen Intervention zum Misserfolg eines Programms führen. Andererseits sind es typischerweise mehrere unverzichtbare Komponenten, deren Zusammenwirken für den Erfolg einer komplexen Intervention verantwortlich sind.

Im Unterschied zur Entwicklung und Evaluation eines definierten Arzneimittels, macht die Multidimensionalität der komplexen Intervention eine sehr viel umfassendere Beschreibung und Begründung der übrigen Komponenten erforderlich, die die Gesamtheit der Intervention ausmachen. Während die Daten zur Zulassung eines Arzneimittels systematisch erhoben und dokumentiert werden müssen, fehlen bei komplexen Interventionen zumeist ausreichend detaillierte und strukturierte Informationen zu ihrer theoretischen Fundierung und Entwicklung [6,9,46].

## Komplexe Endpunkte

Ergebnisvariablen können eher singuläre, kombinierte oder auch komplexe Endpunkte sein.

Ein Beispiel für einen singulären Endpunkt ist die Gesamtmortalität. Ein häufig benutzter kombinierter Endpunkt ist „nicht-tödlicher Herzinfarkt, Schlaganfall oder Tod durch koronare Herzkrankheit“ [47].

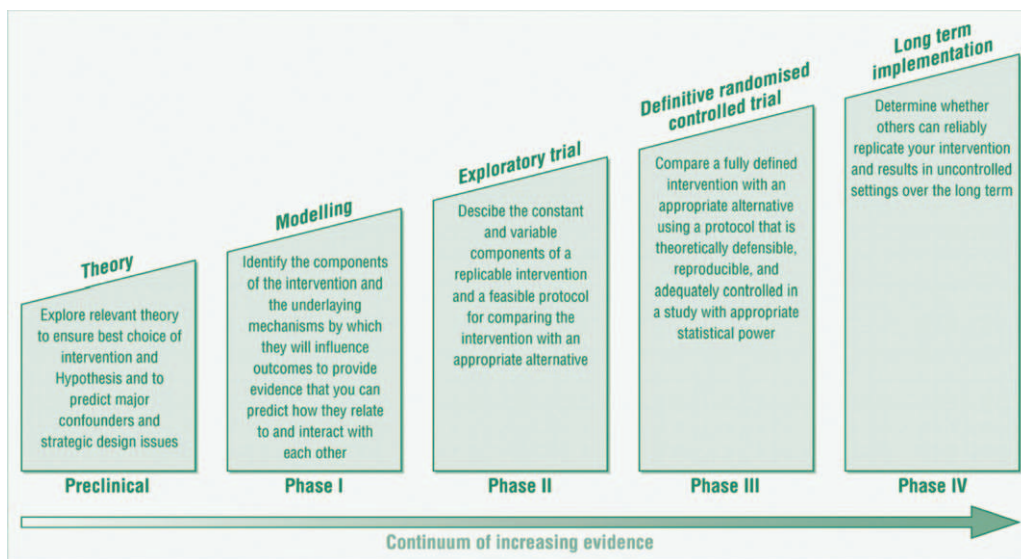
Im Gegensatz zum kombinierten Endpunkt ist ein komplexer Endpunkt direkt von anderen Erfolgsparametern abhängig. Ein Beispiel ist der HbA1c-Wert, wenn er zur Wirksamkeitsbewertung von strukturierten Diabetes-Behandlungs- und Schulungsprogrammen benutzt wird. Der HbA1c-Wert macht nur in der Zusammenschau mit anderen Erfolgsparametern Sinn. So kann bei Patienten mit Diabetes Typ 2 ein bestimmter HbA1c-Zielwert unterschiedlich erreicht werden, z.B. mit einer Intensivierung [48] oder auch Reduzierung der medikamentösen Behandlung [49], je nachdem ob die komplexe Intervention eine gleichzeitige Gewichtsreduzierung vorsieht oder nicht. Bei Patienten mit Diabetes Typ 1 kann eine Verbesserung der HbA1c-Werte mit einer Zunahme oder Abnahme an schweren Unterzuckerungen einhergehen [50]. Auch kann die Intervention eine Reduzierung der Arztkontakte aber auch Intensivierung der Betreuung vorsehen [50]. Der HbA1c-Wert kann sowohl ein zwischen Patient und Arzt individuell definiertes Therapieziel implizieren oder

aber ein für alle Patienten pauschal festgelegtes Therapieziel. Ein individuell festgelegter HbA1c-Zielwert kann z.B. bei fortgeschrittenem Alter oder Begleiterkrankungen durchaus höher liegen als der in Leitlinien als optimal definierte Grenzwert. Der komplexe Endpunkt HbA1c-Wert kann für ein Individuum also nicht interpretiert werden, ohne diese Begleitaspekte zu berücksichtigen.

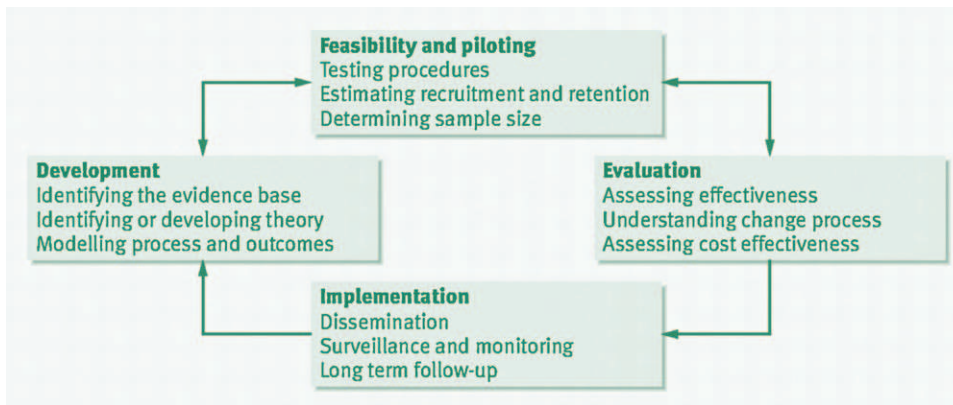
Auf der Methodenebene systematischer Übersichtsarbeiten dürften die Einzelkomponenten von komplexen Endpunkten nicht – wie sonst üblich – separiert voneinander betrachtet werden. Auch das GRADE System [51], das sich inzwischen nicht nur für die Leitlinienerstellung, sondern auch für die Anfertigung von systematischen Reviews und Health Technology Assessments etabliert hat, bietet bislang keine methodische Lösung wie mit komplexen Endpunkten zu verfahren ist.

## Entwicklung und Evaluation von komplexen Interventionen

Für die Entwicklung und Evaluation von komplexen Interventionen liegen seit 1999 Vorschläge vom britischen Medical Research Council [32,52] vor.



**Abb. 2** Sequenzielle Phasen der Entwicklung und Evaluation komplexer Interventionen (extrahiert aus [52], reproduziert mit Genehmigung der BMJ Publishing Group).



**Abb. 3** Kernelemente der Entwicklung und Evaluation komplexer Interventionen (extrahiert aus [32], reproduziert mit Genehmigung der BMJ Publishing Group).

Analog zur Entwicklung eines Arzneimittels wurden klinische Phasen I bis IV definiert (Abb. 2). Die präklinische oder auch theoretische Phase dient zunächst der Identifizierung der bereits vorliegenden Evidenz, der Entwicklung erster Konzepte und der Bildung von Hypothesen über die intendierte Wirksamkeit der komplexen Intervention. Die Phase I beinhaltet den Entwurf und die Beschreibung einzelner Komponenten der komplexen Intervention, der Prototypenentwicklung und ersten Exploration von Wirkmechanismen und Erfolgsparametern. Die Phase II besteht in der Evaluation der Machbarkeit. Ermittelt werden beispielsweise unterschiedliche Implementierungsbedingungen, die Stabilität der komplexen Intervention in unterschiedlichen Kontexten und Effekte durch Modifikationen einzelner Komponenten in variierenden Kontexten. Ziel ist es auch, die Bedingungen für ein RCT zu explorieren (z.B. Wirkhypothesen, Schätzungen der Stichprobe, Kontrollintervention). In der Phase III wird die komplexe Intervention als Ganzes mit einer angemessenen Alternative im RCT verglichen. Die Phase IV besteht schließlich in der Evaluation der Langzeitimplementierung. Untersucht werden die Übertragbarkeit, Reproduzierbarkeit und Langzeitwirksamkeit unter realen Bedingungen. Die Phasen beinhalten unterschiedliche qualitative und quantitative Forschungsmethoden, bauen aufeinander auf und bilden so ein Kontinuum ansteigender Evidenz (*continuum of increasing evidence* im Folgenden *increasing evidence*) [52].

Vor der Revision des Konzeptes im Jahr 2008 hat die MRC-Arbeitsgruppe die Kritik zum Modell und die aufgezeigten Limitierungen analysiert [32]. So impliziert das Konstrukt der *increasing evidence* eine der Praxis wenig entsprechende Linearität im Entwicklungs- und Evaluationsprozesses. Auch wurde die Variabilität von Versorgungskontexten nicht ausreichend berücksichtigt. Der revidierte Ansatz [32] beinhaltet vier Entwicklungs- und Evaluationsphasen (Abb. 3), die präklinische Phase und die Phase I wurden zur ersten Phase zusammengefasst, mögliche Evaluationsmethoden der einzelnen Phasen wurden spezifiziert, die Zirkularität der Prozesse und Kontextfaktoren sind berücksichtigt. Der Kasten 1 illustriert ein eigenes Beispiel der Entwicklung und Evaluation einer komplexen Intervention aus der Altenpflege (Kasten 1).

### Limitierungen der üblichen Verfahren der Synthese von Evidenz aus komplexen Interventionen

Mit der Zulassung eines Arzneimittels ist dieses in seiner Zusammensetzung standardisiert und lässt sich als definierte singuläre Intervention in klinischen Studien einsetzen. Die klinischen Studien unterscheiden sich dann im Wesentlichen durch Studien-, Patienten- und Kontextfaktoren. Die Intervention selbst ist jedoch weitgehend identisch. Das ursprüngliche primäre Ziel systematischer Übersichtsarbeiten und

deren Meta-Analysen ist es, die Evidenz aus den verfügbaren RCTs zu einer Medikamentenbehandlung bei einer bestimmten Indikation quantitativ zusammenzufassen. Damit sollten die Erkenntnisse aus den oft viel zu kleinen klinischen Studien genutzt werden und Ressourcen verantwortungsvoll eingesetzt werden. Wirksame Behandlungen sollten frühzeitiger erkannt werden, der zu erwartende Effekt abgeschätzt werden, unnötige weitere Studien vermieden werden. Patienten sollten somit möglichst rasch wirksame und sichere Therapien erhalten und andererseits nicht nochmals unterlegenen Therapien (einschließlich Placebo) in experimentellen Studien ausgesetzt werden. Systematische Übersichtsarbeiten sind in den letzten Jahrzehnten zunehmend an die Stelle traditioneller, selektiver Reviews (*literature reviews, narrative reviews*) getreten. Sie haben den Umgang mit klinischen Studien erheblich beeinflusst und zur Ausdifferenzierung empirisch geprüfter methodischer Standards klinischer Studien geführt. Der direkte Einfluss von systematischen Übersichtsarbeiten auf die Verbesserung der Patientenergebnisse ist jedoch schwer abzuschätzen. Aus kritischen Analysen zum Verzicht auf systematische Übersichtsarbeiten lässt sich jedoch indirekt auf ihren Nutzen schließen [53,54]. Nicht nur für die praktische Versorgung sind systematische Übersichtsarbeiten unentbehrlich geworden. Sie sollten den Ausgangspunkt einer jeden klinischen Studie darstellen. Handlungsleitend ist die Frage, ob es

### **Kasten 1: Entwicklung komplexer Interventionen gemäß Leitlinie des UKMRC – ein Beispiel.**

In einem sechsjährigen, vom BMBF geförderten Vorhaben im Rahmen des Pflegeforschungsverbundes Nord, haben wir versucht, eine nachhaltig wirksame komplexe Intervention zur Reduktion von freiheitsentziehenden Maßnahmen (FEM) in Pflegeheimen gemäß UKMRC-Zyklus zu entwickeln. Wir haben zunächst eine epidemiologische Studie [55] durchgeführt zur Bestimmung der Häufigkeit der Anwendung von FEM und der FEM-Arten (Gurte, Tische am geriatrischen Stuhl, beidseitige Bettgitter etc.) sowie Surveys zur Exploration von Haltungen, Einstellungen und Meinungen Pflegenden und Angehöriger von Pflegeheimbewohnern gegenüber FEM [56,57]. Es folgte als notwendige Voraussetzung zur Modellierung der Intervention die Anfertigung eines systematischen Übersichtsartikels [58] zur Identifikation wirksamer Interventionskomponenten. In einem nächsten Schritt wurde eine evidenzbasierte Leitlinie samt Implementierungshilfen entwickelt [59]. In der folgenden explorativen Phase wurden die Schulungs- und Implementierungsmaterialien und -programme in Fokusgruppen mit Pflegenden und Bewohnervertretern auf ihre Verständlichkeit und Machbarkeit untersucht. Zur Prüfung der Machbarkeit und Akzeptanz der Interventionsstudie und der Erhebungsinstrumente erfolgte eine Pilotphase mit vier randomisiert zugewiesenen Einrichtungen [60]. Es schloss sich ein Cluster-RCT an mit 36 Alten- und Pflegeheimen und ca. 4000 Bewohnern. Die Interventionsgruppe erhielt die komplexe Intervention bestehend aus (1) Leitlinien-gestütztem Schulungsprogramm (1,5 h Schulung für alle Pflegenden in der Einrichtung, sechs Stunden Seminar für je zwei FEM-Beauftragte pro Einrichtung), (2) strukturierter Begleitung der FEM-Beauftragten über drei Monate, (3) Bereitstellung von Leitlinie und Informationsmaterial, 4) Reminder wie Poster, Becher oder Stifte mit einem projekteigenen Emblem. Die Kontrollgruppe erhielt eine Standardinformation [60]. Auf der Prozessevaluationssebene wurden a) die Schulungen in den Einrichtungen der Interventionsgruppe (n=58) mit Pflegenden (n=569) anhand einer Erhebung zu Wissen und Selbstwirksamkeit evaluiert. Die FEM-Beauftragten wurden am Ende der Studie nach sechs Monaten interviewt, um die Barrieren und begünstigende Faktoren zu explorieren sowie das Ausmaß der Implementierung und Nachhaltigkeit. Da die Intervention klinisch wirksam und sicher zu einer Reduktion von FEM geführt hat [61], wurden zum Zwecke der Dissemination die Einrichtungen der Kontrollgruppe im Anschluss an die Studie geschult und angeleitet, sowie die Leitlinie samt aller Implementierungsmaterialien kostenlos im Internet zur Verfügung gestellt (<http://www.leitlinie-fem.de>). Eine Implementierungsstudie befindet sich in Vorbereitung. Die Prozessevaluation liefert wiederum wertvolle Information zur Weiterentwicklung der Intervention zur Überwindung von hinderlichen Barrieren der Umsetzung.

sinnvoll und geboten ist, eine weitere Studie durchzuführen. Renommierte medizinische Journale fordern inzwischen den Nachweis einer systematischen Übersichtsarbeit [62]. Methodiker rufen nachdrücklich dazu auf, nicht nur zur Bedarfsbestimmung einer klinischen Studie, sondern auch danach zur Einbettung der eigenen Ergebnisse eine systematische Übersichtsarbeit anzufertigen [63].

Systematische Übersichtsarbeiten und deren Meta-Analysen haben ihre Limitierungen, im Wesentlichen durch die Qualität und Vollständigkeit der Berichterstattung der verfügbaren Daten. Die methodischen Probleme werden zunehmend erkannt und verbesserte Verfahren diskutiert, zum Beispiel zum Umgang mit nicht publizierten Daten oder mit der Heterogenität der Studien [64]. Für Meta-Analysen von Medikamentenstudien wurde kürzlich

gefordert, die Gesamtheit der Evidenz zu betrachten, zum Beispiel die Ergebnisse aus unterschiedlichen Indikationsgebieten [65]. Der Schwerpunkt des methodologischen Diskurses liegt aktuell auf der Verbesserung statistischer Verfahren.

Kaum im Diskurs beachtet ist hingegen, ob die etablierten Verfahren komplexen Interventionen gerecht werden, d.h. die ursprünglichen, oben dargelegten Ziele systematischer Übersichtsarbeit erfüllt werden können.

In einer methodischen Übersichtsarbeit haben wir zu illustrieren versucht, dass systematische Übersichtsarbeiten derzeit nicht geeignet sind, komplexe Interventionen kritisch zu bewerten [6]. Über systematische Recherchen (PubMed, Cumulative Index to Nursing and Allied Health (CINAHL), Cochrane Library und HTA-Datenbanken) wurden zunächst systematische Reviews über

Schulungs- und Selbstmanagementprogramme im Bereich Diabetes und Hypertonie identifiziert. Die der Ko-Autorin IM bekannte öffentlich verfügbare kumulative Evidenz von drei Schulungs- und Selbstmanagementprogrammen wurde als Referenzstandard benutzt, um zu untersuchen, ob die Programme und die dazugehörigen verfügbaren Publikationen angemessen berücksichtigt und bewertet wurden.

Vierzehn systematische Reviews wurden analysiert, darunter drei Cochrane Reviews. Zwölf der 14 Reviews bezeichneten die Programme korrekterweise als „multidimensional“ oder „komplex“. Die Ergebnisse unserer Analysen zeigten dennoch erhebliche Defizite. In Reviews mit vergleichbaren Fragestellungen wurden verschiedene Publikationen identischer Referenzprogramme analysiert. Identische Programme wurden zwischen verschiedenen aber auch

innerhalb derselben Reviews unterschiedlich klassifiziert. In sechs Reviews wurden mittels meta-analytischer Verfahren Einflüsse von Einzelkomponenten (z.B. Dauer der Schulung) über unterschiedliche Schulungsprogramme hinweg analysiert. Deren klinische Heterogenität (z.B. unterschiedliche zugrunde liegende pädagogische Ansätze und Studiensettings) wurde nicht berücksichtigt. Interdependenzen zwischen Komponenten und Kontexten wurden nicht evaluiert. Auch die komplexen Interdependenzen zwischen komplexen Endpunkten (z.B. HbA1c und Hypoglykämien) und Kontexten und Zielen der Schulungsprogramme wurden nicht berücksichtigt.

Auf Basis unserer Ergebnisse haben wir schließlich sechs Kernpunkte zu einer angemessenen Evaluation komplexer Interventionen formuliert: 1) Alle Publikationen zur Entwicklung, Evaluation und Implementierung einer komplexen Intervention sollten berücksichtigt werden; 2) Die Literaturrecherchen sollten entsprechend angepasst werden (keine Limitierungen auf Studiendesign, spezifische Zielgruppen und Publikationszeiträume; „Snowballing“-Techniken wie Referenzlisten sichten, *forward citation tracking*, systematischer Autorenkontakt); 3) Die aktiven Komponenten sollten beschrieben und begutachtet werden; 4) Komplexe Interventionen sollten nicht Kategorien zugeordnet werden, die sich auf abhängige Komponenten beziehen; 5) Alle patientenrelevanten Endpunkte sollten angemessen berücksichtigt werden; 6) Das Poolen von Endpunkten über verschiedene komplexe Interventionen ist wegen inhärenter klinischer Heterogenität meist inadäquat. Stattdessen sollten gegenseitige Abhängigkeiten zwischen Ergebnisparametern, Prozessen und Zielen analysiert und berichtet werden.

## „Further research is needed“

Häufig lautet die Schlussfolgerung einer systematischen Übersichtsarbeit „*further research is needed*“. Dieser Appell erscheint unlogisch, wenn mehrere Dutzend Studien bereits vorliegen, die

sich durch klinische und statistische Heterogenität auszeichnen. Am Beispiel Sturzprävention sei dies illustriert. Sturzprävention für Senioren besteht oft aus Risikoeinschätzung und Angebot Risiko-adaptierter Interventionen. Die häufig als multifaktorielle Interventionen bezeichneten Maßnahmenpakete können neben Schulung und Information, Wohnraumanpassung und körperlichem Training, Medikamentenreview, Anpassung der Sehhilfen oder der Schuhe, Angebot von Hilfsmitteln oder Hüftprotektoren beinhalten. Ein Cochrane Review zum Thema [66,67] wurde kürzlich Setting-spezifisch in Sturzprävention im stationären Umfeld, d.h. Alten- und Pflegeheim und Krankenhaus [66], und Sturzprävention in der Häuslichkeit [67] unterteilt. Die Anzahl der eingeschlossenen RCTs ist beeindruckend: n = 41 Studien mit 25422 Teilnehmern sowie n = 111 Studien mit 55303 Teilnehmern. Die Schlussfolgerungen sind hingegen eher vage. Gillespie et al. [67] resümieren aufgrund der klinischen und statistischen Heterogenität der eingeschlossenen Studien, dass die Kontexte in denen die vermutlich wirksamen Interventionen ihren klinischen Nutzen entfalten können, zu untersuchen bleiben. Auch Cameron et al. [66] konstatieren ausgesprochene Inkonsistenz der Studienergebnisse und weiteren Forschungsbedarf. Für die Bestimmung der aktiven Komponenten innerhalb komplexer Interventionen werden statistische Verfahren, z.B. Metaregressionsanalysen [68] diskutiert. Diese post hoc-analytischen Verfahren können zweifelsohne kein Ersatz sein für eine sorgfältige prospektive Entwicklung und Exploration der komplexen Intervention und der Interdependenzen mit kontextuellen Faktoren. Sie haben Observationscharakter und generieren maximal Hypothesen.

## Sind andere Formen der Evidenz-Synthese besser geeignet für komplexe Interventionen?

Naheliegender ist es, sich anstatt der (ausschließlichen) Meta-Analyse explanatorischen Formen der Synthese

zuzuwenden. Ein Modell zur rigorosen und strukturierten narrativen Synthese von Daten aus RCTs für die Fälle systematischer Reviews, in denen eine Meta-Analyse nicht möglich oder empfehlenswert ist, wurde kürzlich vorgelegt [69].

In der Pflegewissenschaft und Public Health erfreut sich die Realist Evaluation bzw. der Realist Review seit einigen Jahren großer Beliebtheit [70–72]. Daher gilt es, diese Form der Evidenzsynthese hier eingehender zu diskutieren. Vorgesprochen werden auch die Kombination von Realist Review und systematischem Review nach Cochrane oder Campbell Methoden [73].

Beim Realist Review stehen nicht die Fragen „*Does it work*“ oder „*What works*“ im Mittelpunkt, sondern die Frage „*What works for whom in what circumstances*“. Der Realist Review [27] möchte diese komplexe Frage beantworten und unterscheidet sich nicht nur in seinem Fokus, sondern auch in seinen Methoden von einem konventionellen systematischen Review. Realist Reviews verlaufen weniger linear und sequentiell als konventionelle Reviews. Anfänglich definierte Fragen führen zu mehr Fragen und daher auch zur Auswertung diverser Informationsquellen mit verschiedenen Studiendesigns in dem Versuch zu erklären, warum, wann und wie die Intervention funktioniert. Ein Realist Review beginnt mit der einschlägigen Lektüre zur Identifikation der Faktoren mit dem größten vermeintlichen erklärendem Wert. Ein provisorisches explanatorisches Modell inkludiert die Faktoren und Behauptungen und ist handlungsleitend für den folgenden Review [74]. Der Realist Review geht davon aus, dass der Wirkungsweise einer komplexen Intervention ein Modell zugrundeliegt, das erklärt, wie ein messbarer Effekt verursacht wird. Der entscheidende Aspekt ist der Kontext, in dem das Programm bzw. die komplexe Intervention stattfindet. **C**(ontext) - **M**(echanism) - **O**(utcome) Konfigurationen werden untersucht. Die identifizierte Literatur wird auf das provisorische explanatorische Modell angewendet und verfeinert dieses iterativ oder modifiziert es. Am Ende des Prozesses steht ein revidiertes explanatorisches Modell [74].

## **Kasten 2: 13. Jahrestagung des DNEbM zum Thema „Komplexe Interventionen – Entwicklung durch Austausch“, 15. bis 17.03.2012 in Hamburg.**

Die methodischen Probleme medizinischer komplexer Interventionen stellen sich in ähnlicher Weise für Interventionen aus dem Pflege-, Ernährungs-, Erziehungs- oder Politikbereich, um einige zu nennen. Die 13. Jahrestagung des Deutschen Netzwerks für Evidenzbasierte Medizin (<http://www.ebm-kongress.de/>) soll die Möglichkeit eröffnen, sich über Wissenschaftsgrenzen hinweg über die aktuellen methodischen Verfahren zur Bewertung von Nutzen und Schaden von komplexen Interventionen zum Erhalt und der Verbesserung von Gesundheit und Lebensbedingungen auszutauschen. Die wertvollen Erkenntnisse aus anderen Wissenschaftsbereichen sollen zur kritischen Reflektion anregen, zu neuen Ideen inspirieren und gemeinsame Projekte anbahnen. Es sind sowohl Wissenschaftler als auch Nutzer wissenschaftlicher Ergebnisse geladen, ihre wissenschaftlichen Ergebnisse zu präsentieren, Projekte zur Diskussion zu stellen oder ganz einfach aus den Erfahrungen anderer zu lernen.

Selbstredend ist ein Realist Review eine post hoc Synthese von Daten aus unterschiedlichen Quellen, auch inkonsistenter interner Validität. Ein demgemäß abgeleitetes Erklärungsmodell zur Wirksamkeit kann unserer Meinung nach nicht bedeuten, dass die Implementierung der in Frage stehenden komplexen Intervention nicht unter kontrollierten Bedingungen erfolgen soll.

Die Ergebnisse eines Realist Reviews können bei hohem Arbeitsaufwand erstaunlich marginal ausfallen. Dies sei an einem Beispiel illustriert: Wong et al. [75] intendieren, Theorie-basierte Kriterien für die Entwicklung und Evaluation Web-basierter medizinischer Ausbildungskurse zu generieren. Nach umfangreichen Literaturrecherchen werden aus 12586 Treffern schließlich 249 Studien jedwedem Designs eingeschlossen und anhand des aufwändigen Realist Evaluation Verfahrens analysiert. Zwei Theorien, *Davis's Technology Acceptance Model* and *Laurillard's model of interactive dialogue*, werden identifiziert, die die Variation in den Ergebnissen und in der Zufriedenheit mit dem Angebot erklären. Das Ergebnis des Realist Reviews ist wenig spektakulär. Erstens sollen Web-basierte Kurse ihre Zielgruppe aktiv mit einbinden. Demnach steigt die Akzeptanz Web-basierter medizinischer Ausbildungskurse, wenn die Lernenden einen Vorteil empfinden gegenüber Nicht-Internet-basierten Angeboten, die technische Benutzung einfach und das Angebot kompatibel mit den eigenen Werten und Normen ist. Zweitens wird Interaktion hoch geschätzt.

Lernende möchten mit einem Tutor oder anderen Lernenden in einen Dialog treten und ein formatives Feedback erhalten. Wong et al. [75] reklamieren in der Diskussion ihres Artikels die limitierte Güte der ausgewerteten Studien. Wie jede andere Sekundäranalyse, kann die Güte der Synthese nur so gut sein wie die ihr zugrunde liegenden Daten. Nach Lektüre des Realist Reviews drängt sich dem Leser die Frage auf, ob nicht ein geringerer Aufwand zu vergleichbaren Ergebnissen geführt hätte. Beispielsweise durch strukturierte Befragung von Experten, die Web-basierte Kurse entwickelt haben.

### **Schlussfolgerung**

Die prospektive Entwicklung und Evaluation komplexer Interventionen gemäß der UKMRC Leitlinie sowie die Bereitstellung aller Ergebnisse aus Pilotphasen und Begleitevaluationen kann nicht durch post hoc Ansätze ersetzt werden. Die Zusammenfassung der Forschung zu einer sorgfältig und langwierig entwickelten und erfolgreich evaluierten komplexen Intervention mit detaillierter Beschreibung der theoretischen Grundlagen und Bedingungskomponenten, der Erfolgs-befördernden und hinderlichen Faktoren stellt ebenso eine Form der Evidenz-Synthese dar. Sie ist für den Nutzer und Entscheidungsfinder im Gesundheitswesen relevanter und nachvollziehbarer als die grobe und teilweise irreführende Schlussfolgerung aus einer gemäß den üblichen Methoden

angefertigten, nicht das gesamte Kontinuum der Evidenz sichtenden systematischen Übersichtsarbeit oder aus einer post hoc Synthese von Evidenz aus diversen Quellen unterschiedlicher Qualität wie im Realist Review.

Die Gesamtheit der Evidenz ist unverzichtbar, wenn darüber entschieden werden soll, ob eine Intervention in ein anderes Gesundheitssystem übertragbar ist. Herkömmliche systematische Übersichtsarbeiten im Interventionsbereich fokussieren auf RCTs. Deren Stellenwert soll hier keinesfalls in Abrede gestellt werden. Auf dem Kontinuum der Evaluation komplexer Interventionen erfolgt ein RCT jedoch erst nach einschlägigen Vorarbeiten in einem weit gediehenen Stadium der Interventionsentwicklung und -evaluation (Abb. 2 und 3). Die Gesamtheit der Evidenz für eine komplexe Intervention ist in systematischen Übersichtsarbeiten, die den standardisierten Methoden folgen, nicht zu beurteilen, da nicht danach recherchiert wird und die einzelnen Publikationen oft auch mit den üblichen Suchstrategien in den aktuellen Datenbanken nicht identifizierbar sind [6]. Insofern liefern systematische Übersichtsarbeiten zu komplexen Interventionen unvollständige Informationen. RCTs zu komplexen Interventionen, die sorgfältig und langwierig vorbereitet sind, werden mit unzureichend entwickelten komplexen Interventionen kombiniert. Studien, die die Wirksamkeit eines Programms erstmals experimentell evaluieren, werden mit Studien, die das Programm in einen anderen Kontext übertragen, synthetisiert [6].

Es drängt sich die Frage auf, ob systematische Übersichtsarbeiten ihrer ursprünglichen Intention bei komplexen Interventionen noch gerecht werden können und tatsächlich zu einem Mehrwert an Wissen führen, oder aber eher zum Selbstzweck werden.

Diese und ähnliche Fragen sind bisher, gemessen an ihrer Bedeutung, nicht hinlänglich international kritisch gewürdigt. Die kommende Jahrestagung des Deutschen Netzwerks für evidenzbasierte Medizin in Hamburg hat daher als Kongress-Hauptthema „Komplexe Interventionen – Entwicklung durch Austausch“ bestimmt (Kasten 2).

## Literaturverzeichnis

- [1] Blackwood B. Methodological issues in evaluating complex healthcare interventions. *J Adv Nurs* 2006;54:612–22.
- [2] Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, et al. Designing and evaluating complex interventions to improve health care. *BMJ* 2007;334:455–9.
- [3] Egan M, Bamba C, Petticrew M, Whitehead M. Reviewing evidence on complex social interventions: appraising implementation in systematic reviews of the health effects of organisational-level workplace interventions. *J Epidemiol Community Health* 2009;63:4–11.
- [4] Glasziou P, Chalmers I, Altman DG, Bastian H, Boutron I, Brice A, et al. Taking healthcare interventions from trial to practice. *BMJ* 2010;341:c3852.
- [5] Herbert RD, Bo K. Analysis of quality of interventions in systematic reviews. *BMJ* 2005;331:507–9.
- [6] Lenz M, Steckelberg A, Richter B, Muhlhauser I. Meta-analysis does not allow appraisal of complex interventions in diabetes and hypertension self-management: a methodological review. *Diabetologia* 2007;50:1375–83.
- [7] Pfaff H, Albert US, Bornemann R, et al. *Methoden für die organisationsbezogene Versorgungsforschung*. Gesundheitswesen 2009;71:777–90.
- [8] Raspe H. Unterschiede in der Bewertung von medikamentösen und nichtmedikamentösen Maßnahmen? *Z Evid Fortbild Qual Gesundhwes* 2009;103:253–60.
- [9] Shepperd S, Lewin S, Straus S, Clarke M, Eccles MP, Fitzpatrick R, et al. Can we systematically review studies that evaluate complex interventions? *PLoS Med* 2009;6:e1000086.
- [10] Walach H, Falkenberg T, Fonnebo V, Lewith G, Jonas WB. Circular instead of hierarchical: methodological principles for the evaluation of complex interventions. *BMC Med Res Methodol* 2006;6:29.
- [11] Lancaster GA, Campbell MJ, Eldridge S, et al. Trials in primary care: statistical issues in the design, conduct and evaluation of complex interventions. *Stat Methods Med Res* 2010;19:349–77.
- [12] Ogilvie D, Fayer D, Petticrew M, et al. The harvest plot: a method for synthesising evidence about the differential effects of interventions. *BMC Med Res Methodol* 2008;8:8.
- [13] Andrew S, Halcomb EJ. Mixed methods research is an effective method of enquiry for community health research. *Contemp Nurse* 2006;23:145–53.
- [14] Curry LA, Nembhard IM, Bradley EH. Qualitative and mixed methods provide unique contributions to outcomes research. *Circulation* 2009;119:1442–52.
- [15] Kelle U, Krones T. “Evidence based Medicine” und “Mixed Methods” - wie methodologische Diskussionen in der Medizin und den Sozialwissenschaften voneinander profitieren könnten. *Z Evid Fortbild Qual Gesundhwes* 2010;104:630–5.
- [16] Lewin S, Glenton C, Oxman AD. Use of qualitative methods alongside randomised controlled trials of complex healthcare interventions: methodological study. *BMJ* 2009;339:b3496.
- [17] Miller CL, Druss BG, Rohrbaugh RM. Using qualitative methods to distill the active ingredients of a multifaceted intervention. *Psychiatric Services* 2003;54:568–71.
- [18] Miller SI, Fredericks M. Mixed-methods and evaluation research: trends and issues. *Qual Health Res* 2006;16:567–79.
- [19] Nastasi BK, Hitchcock J. Challenges of evaluating multilevel interventions. *Am J Community Psychol* 2009;43:360–76.
- [20] Palinkas LA, Aarons GA, Horwitz S, Chamberlain P, Hurlburt M, Landsverk J. Mixed method designs in implementation research. *Adm Policy Ment Health* 2011;38:44–53.
- [21] Paterson C, Baarts C, Launso L, Verhoef MJ. Evaluating complex health interventions: a critical analysis of the ‘outcomes’ concept. *BMC Complement Altern Med* 2009;9:18.
- [22] Pluye P, Gagnon MP, Griffiths F, Johnson-Lafleur J. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *Int J Nurs Stud* 2009;46:529–46.
- [23] Schifferdecker KE, Reed VA. Using mixed methods research in medical education: basic guidelines for researchers. *Med Educ* 2009;43:637–44.
- [24] Whitley R. Mixed-methods studies. *J Ment Health* 2007;16:697–701.
- [25] Yoshikawa H, Weisner TS, Kalil A, Way N. Mixing qualitative and quantitative research in developmental science: Uses and Methodological Choices. *Dev Psychol* 2008;44:344–54.
- [26] Hunt KS, Sridharan S. A realist evaluation approach to unpacking the impacts of the sentencing guidelines. *AJE* 2010;31:463–85.
- [27] Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review - a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy* 2005;10(Suppl 1):21–34.
- [28] Emsley R, Dunn G, White IR. Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. *Stat Methods Med Res* 2010;19:237–70.
- [29] Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Process evaluation in randomised controlled trials of complex interventions. *BMJ* 2006;332:413–6.
- [30] Perera R, Heneghan C, Yudkin P. Graphical method for depicting randomised trials of complex interventions. *BMJ* 2007;334:127–9.
- [31] Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. [www.cochrane-handbook.org](http://www.cochrane-handbook.org), accessed 15.6.2011.
- [32] Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: new guidance. UK Medical Research Council, 2008. [www.mrc.ac.uk/complexinterventionsguidance](http://www.mrc.ac.uk/complexinterventionsguidance), accessed 16.6.2011.
- [33] Centre for Reviews and Dissemination. *Systematic Reviews: CRD’s guidance for undertaking reviews in health care*. 2009. [www.york.ac.uk/inst/crd/pdf/Systematic\\_Reviews.pdf](http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf), accessed 16.6.2011.
- [34] Social Care Institute for Excellence (SCIE). *SCIE systematic research reviews: guidelines (2nd edition, 2010)*. [www.scie.org.uk/publications/researchresources/rr01.pdf](http://www.scie.org.uk/publications/researchresources/rr01.pdf), accessed 16.6.2011.
- [35] Petticrew M, Roberts H. *Systematic reviews in the social sciences: a practical guide*. Malden: Blackwell; 2006.
- [36] Boutron I, Moher D, Altman DG, Schulz KF, Ravaut P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008;148:295–309.
- [37] Briss PA, Zaza S, Pappaioanou M, et al. Developing an evidence-based guide to community preventive services - methods. The Task Force on Community Preventive Services. *Am J Prev Med* 2000;18(1 Suppl):35–43.
- [38] Coote A, Allen J, Woodhead D. *Finding out what works: understanding complex, community-based initiatives*. London: King’s Fund; 2004.
- [39] Cochrane Effective Practice and Organisation of Care (EPOC) Review Group. *Data collection checklist*. Ottawa: 2000.
- [40] Newman M, Elbourne D. Improving the usability of educational research: Guidelines for the REPORTing of primary empirical research

- Studies in Education (The REPOSE Guidelines). *Evaluation and Research in Education* 2005;18:201–12.
- [41] Glasgow RE, Klesges LM, Dziewaltowski DA, Estabrooks PA, Vogt TM. Evaluating the impact of health promotion programs: using the RE-AIM framework to form summary measures for decision making involving complex issues. *Health Educ Res* 2006;21:688–94.
- [42] Des Jarlais DC, Lyles C, Crepez N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004;94:361–6.
- [43] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening of reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61:344–9.
- [44] Writing Group for the Women's Health Initiative Investigators. Risks and Benefits of Estrogen plus progesterin in healthy postmenopausal women. *JAMA* 2002;288:321–33.
- [45] Mühlhauser I, Berger M. Patient education - evaluation of a complex intervention. *Diabetologia* 2002;45:1723–33.
- [46] Warsi A, Wang PS, LaValley MP, Avorn J, Solomon DH. Self-management education programs in chronic disease: a systematic review and methodological critique of the literature. *Arch Intern Med* 2004;164:1641–9.
- [47] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden. Entwurf für Version 4.0 vom 09.03.2011. [www.iqwig.de/download/IQWiG\\_Entwurf\\_Methoden\\_Version\\_4-0.pdf](http://www.iqwig.de/download/IQWiG_Entwurf_Methoden_Version_4-0.pdf), accessed 1.8.2011.
- [48] The Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med* 2008;358:2545–59.
- [49] Kronsbein P, Jörgens V, Mühlhauser I, Scholz V, Venhaus A, Berger M. Evaluation of a structured treatment and teaching programme on non-insulin-dependent diabetes. *Lancet* 1988;2(8625):1407–11.
- [50] Bott S, Bott U, Berger M, Mühlhauser I. Intensified insulin therapy and the risk of severe hypoglycaemia. *Diabetologia* 1997;40:926–32.
- [51] Atkins D, Best D, Briss P, Guyatt G, Henry D, Hill S, et al., GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490.
- [52] Campbell M, Fitzpatrick R, Haines A, Kinmonth A, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;321:694–6.
- [53] Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction *JAMA* 1992;268:240–8.
- [54] Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC, et al. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992;327:248–54.
- [55] Meyer G, Köpke S, Haastert B, Mühlhauser I. Restraint use among nursing home residents: cross-sectional study and prospective cohort study. *J Clin Nurs* 2009;18:981–90.
- [56] Hamers JPH, Meyer G, Köpke S, Lindemann R, Groven R, Huizing AR. Attitudes of Dutch. German and Swiss nursing staff towards physical restraint use in nursing home residents, a cross-sectional study *Int J Nurs Stud* 2009;46:248–55.
- [57] Haut A, Kolbe N, Strupeit S, Mayer H, Meyer G. Attitudes of relatives of nursing home residents towards physical restraints. *J Nurs Scholarsh* 2010;42:448–56.
- [58] Möhler R, Richter T, Köpke S, Meyer G. Interventions for preventing and reducing the use of physical restraints in long-term geriatric care. *Cochrane Database Syst Rev* 2011;2:CD007546.
- [59] Köpke S, Meyer G, Haut A, Gerlach A. Methodenpapier zur Entwicklung einer Praxisleitlinie zur Vermeidung von freiheits-einschränkenden Maßnahmen in der beruflichen Altenpflege. *Z Evid Fortbild Qual Gesundh wesen* 2008;102:45–53.
- [60] Haut A, Köpke S, Gerlach A, Mühlhauser I, Haastert B, Meyer G. Evaluation of an evidence-based guidance on the reduction of physical restraints in nursing homes: a cluster-randomised controlled trial [ISRCTN34974819]. *BMC Geriatrics* 2009;9:42.
- [61] Meyer G, Haut A, Gerlach A, Köpke S. Evaluation of a guidance on the reduction of physical restraints in nursing homes: a cluster-RCT. *Abstract J Clin Nurs* 2010;19(Suppl. 1):58.
- [62] Clark S, Horton R. Putting research into context - revisited. *Lancet* 2010;376:10–1.
- [63] Clarke M, Hopewell S, Chalmers I. Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting. *Lancet* 2010;376:20–1.
- [64] Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;342:d549.
- [65] Ioannidis JP, Karassa FB. The need to consider the wider agenda in systematic reviews and meta-analyses: breadth, timing, and depth of the evidence. *BMJ* 2010;341:c4875.
- [66] Cameron ID, Murray GR, Gillespie LD, Robertson MC, Hill KD, Cumming RG, et al. Interventions for preventing falls in older people in nursing care facilities and hospitals. *Cochrane Database Syst Rev* 2010;(1):CD005465.
- [67] Gillespie LD, Robertson MC, Gillespie WJ, et al. Interventions for preventing falls in older people living in the community. *Cochrane Database Syst Rev* 2009;(2):CD007146.
- [68] Bower P, Gilbody S, Richards D, Fletcher J, Sutton A. Collaborative care for depression in primary care. Making sense of a complex intervention: systematic review and meta-regression *Br J Psychiatry* 2006;89:484–93.
- [69] Rodgers M, Sowden A, Petticrew M, et al. Testing methodological guidance on the conduct of narrative synthesis in systematic reviews. *Evaluation* 2009;15: 047–71.
- [70] Kane SS, Gerretsen B, Scherpbier R, Dal Poz M, Dieleman M. A realist synthesis of randomised control trials involving use of community health workers for delivering child health interventions in low and middle income countries. *BMC Health Serv Res* 2010;10:286.
- [71] Rycroft-Malone J, Fontenla M, Bick D, Seers K. A realistic evaluation: the case of protocol-based care. *Implement Sci* 2010;5:38.
- [72] Walshe C, Luker KA. 'District nurses' role in palliative care provision: a realist review. *Int J Nurs Stud* 2010;47:1167–83.
- [73] Van der Knaap LM, Leeuw FL, Bogaerts S, Nijsen LTJ. Combining Campbell Standards and the realist evaluation approach: The best of two worlds? *AJE* 2008;29: 48–57.
- [74] Leeman J, Chang YK, Lee EJ, Voils CI, Crandell J, Sandelowski M. Implementation of antiretroviral therapy adherence interventions: a realist synthesis of evidence. *J Adv Nurs* 2010;66:1915–30.
- [75] Wong G, Greenhalgh T, Pawson R. Internet-based medical education: a realist review of what works, for whom and in what circumstances. *BMC Med Educ* 2010;10:12.