

THEORETISCHE UND PRAKTI- SCHE ÜBERLEGUNGEN FÜR DIE EVALUATION KOMPLEXER INTER- VENTIONEN ZUR VERBESSERUNG DER BEHANDLUNGSQUALITÄT

Manuela Pretto

Komplexe Interventionen wie z.B. klinische Behandlungspfade leisten einen Beitrag zu verbesserten Patientenergebnissen – falls sie denn gut implementiert sind. Sie stellen hohe Anforderungen an die Evaluation: Im Vorfeld müssen Überlegungen zu Kontext und zu angenommenen Ursache-Wirkungs-Zusammenhängen stattfinden und wichtige Stakeholder sollten dabei einbezogen sein. Bevor nach der Wirkung einer Maßnahme gefragt werden kann, muss zuerst deren Implementierung überprüft werden. Andernfalls können Ergebnisse der Wirkungsanalyse nur teilweise mit der Intervention in Verbindung gebracht, und Evaluationsergebnisse können nur schwer für die Verbesserung der Maßnahme genutzt werden.

Theoretical and practical considerations for the evaluation of complex interventions to improve quality of care

Complex interventions such as clinical treatment pathways make a contribution to improved patient outcomes - if they are as well implemented. They make high demands

on the evaluation: At first, considerations must be connected with the context and with cause-and-effect correlations, and key stakeholders should be included here. At first the implementation must be proved before thinking about the effect of a intervention. Otherwise, results of the impact analysis can only partially associated with the intervention, and evaluation results can be used hardly for the improvement of the measure..

Korrespondenzadresse:

Manuela Pretto

Universitätsspital Basel
Abteilung Praxisentwicklung, Ressort
Pflege/MTT
Spitalstrasse 21/
Petersgraben 4
CH-4031 Basel
Manuela.Pretto@usb.ch

Irena Anna Frei
IrenaAnna.Frei@usb.ch

Verena Friedrich

Katharina Fierz

Norbert Suhm

Seite 354-362
Eingereicht am: 14.06.2013
Akzeptiert am: 14.06.2013
DOI: 10.3936/1258

Einleitung Problemstellung und Fragestellung

Viele medizinische Maßnahmen sind komplexe Interventionen. Sie bestehen aus mehreren Einzelkomponenten, die sich wechselseitig bedingen und von deren Zusammenspiel der Erfolg einer solchen Intervention abhängt (Mühlhauser, Lenz & Meyer, 2012). Beispiele dafür sind Behandlungspfade zur Verbesserung der Behandlungskontinuität oder Krankheitsmanagementprogramme. Zur Evaluation von solch komplexen Interventionen mit ihren vielschichtigen Auswirkungen sind simple Vorher-Nachher-Untersuchungen in der Regel ungenügend und ein ganzheitlicher methodischer Ansatz ist erforderlich. Evaluation im Sinne einer wissenschaftlichen Dienstleistung, welche systematisch und mit wissenschaftlichen Methoden einen Gegenstand überprüft und bewertet, und deren Ergebnisse einen Nutzen bringen sollen, stellt einen solchen Ansatz dar (Beywl, Kehr, Mäder & Niestroj, 2008; Rossi, Lipsey & Freeman, 2004; Westermann, 2002; Widmer, 2006).

Die Kontinuität in der Behandlung von Patienten ist für die Behandlungsqualität entscheidend. Da ein einzelnes interdisziplinäres Behandlungsteam nicht rund um die Uhr verfügbar sein kann, sind integrierte Systemansätze wie klinische Behandlungspfade oder Krankheitsmanagementprogramme gefordert, um die Kontinuität und Qualität der Behandlung sicherzustellen (Biem, Hadjistavropoulos, Morgan, Biem, & Pong, 2003). Klinische Behandlungspfade für spezifische Patientengruppen können durch Standardisierung von Abläufen, der Koordination von Tätigkeiten unterschiedlicher Berufsangehöriger (wer macht was?), durch praxisrelevante Richtlinien, oder durch Unterstützung von Patienten in Bezug auf deren Selbstma-

nagement einen wichtigen Beitrag zu Behandlungskontinuität und damit verbunden zu verbesserten Behandlungsergebnissen leisten (Barbieri et al., 2009; Biem et al., 2003). Ein Behandlungspfad mit definierten Abläufen und Richtlinien allein ist allerdings nicht ausreichend, um das Verhalten von Patienten und Leistungserbringern in der Praxis zu verändern: Behandlungspfade müssen (aus unterschiedlichen Perspektiven) systematisch evaluiert werden, um sicherzustellen bzw. zu überprüfen, ob sie die verfolgten Ziele und Wirkungen auch tatsächlich erreichen. Ein **Problem** besteht darin, dass Studien zur Evaluation von Behandlungspfaden häufig nur deren Auswirkungen im Sinne der Wirksamkeit (efficacy), also verbesserte Resultate bei Patienten zu vergleichbaren oder tieferen Kosten betrachten. Andere wichtige Dimensionen der Evaluation wie z.B. die Implementierung, die Kontinuität von Therapien, oder die Informationsübermittlung zwischen Leistungserbringern werden meist ausgeblendet, wären jedoch relevant, wenn es um das Verständnis von Ursache-Wirkungs-Zusammenhängen und um die Verbesserung solch komplexer Interventionen geht (Barbieri et al., 2009; Glasgow, Vogt & Boles, 1999).

Das **Ziel** dieses Artikels ist es, Herausforderungen in der Evaluation von komplexen Interventionen wie z.B. einem Behandlungspfad aus Sicht der Evaluationsforschung zu betrachten und daraus Hilfestellungen für die praktische (Evaluations-)Arbeit abzuleiten. Die Berücksichtigung von Kontextfaktoren, der Einbezug von zentralen Stakeholdern, der Stand der Implementierung und (implizit) angenommene Wirkungsketten werden im Zusammenhang mit der Formulierung von Evaluations-Zwecken und -fragenstellungen thematisiert. Ein zugrundeliegendes Wirkungsmodell, mögliche Evaluationsfragenstellungen sowie Chancen und Gefahren

unterschiedlicher Evaluationsansätze werden diskutiert. (Die Begriffe „Behandlungspfad“, „Programm“ und „komplexe Intervention“ werden in diesem Artikel aus dem Evaluationsblickwinkel heraus synonym verwendet).

Die zentralen Fragestellungen, die im vorliegenden Artikel beantwortet und mit Beispielen illustriert werden, sind:

- Welche Überlegungen zu Kontext und Wirkungsmodellen sollten im Vorfeld einer Behandlungspfad-Evaluation gemacht werden, damit die Evaluation und deren Ergebnisse nutzbringend sind?
- Was sind mögliche Betrachtungsebenen und Fragestellungen an die Evaluation und wie stehen diese miteinander in Verbindung? (Ein praktischer Ansatz)
- Welche Chancen und Risiken bestehen bezüglich verschiedener Evaluationsansätze und was ist in der Beurteilung von Resultaten zu beachten?

Kontext und Wirkungsmodelle

Bei der Evaluation von komplexen Interventionen wie einem Behandlungspfad gilt es vorab, sich über den Kontext und das entsprechende Wirkungsmodell Gedanken zu machen (vgl. Frage eins).

Unter dem **Kontext** wird die Kombination aller Rahmenbedingungen verstanden, welche die Interventionen umgeben, z.B. die institutionelle Einbindung, das soziale und politische Klima, die personellen Charakteristika der Beteiligten und Betroffenen (Stakeholder) oder wirtschaftliche

Rahmenbedingungen. Zum Kontext gehört auch die (Evaluations- und Lern-) Kultur in einem Betrieb/einer Klinik und wie systematisch auf einen Sachverhalt hingeschaut und dieser überprüft wird. Diese Faktoren haben einen maßgeblichen Einfluss auf die Einführung einer komplexen Intervention und deren Auswirkungen. Kontextwissen ist daher notwendig, um die Evaluation realistisch in Bezug auf bestehende Möglichkeiten (z.B. vorhandene Ressourcen) und abgestimmt auf den Kontext ausgestalten zu können. Kontextfaktoren sind zu beachten, um die Evaluation in Relation zu den genannten Punkten planen, durchführen und deren Resultate vermitteln zu können (Widmer, Landert & Bachmann, 2000).

Ein Beispiel: Kontextfaktoren in Bezug auf einen Behandlungspfad für spezifische chirurgische Eingriffe können sein: Die Art und Positionierung des Akutspitals resp. der nachfolgenden Rehabilitationsklinik, die Zusammenarbeit mit Hausärzten, Spitex, Reha-Kliniken und weiteren Partnern, die interne Führungs- und Lernkultur, Organisationsformen (z.B. Organigramm, Operationsplanung, Position der Pflege), die Art und Ausprägung der interdisziplinären Zusammenarbeit resp. die Zusammensetzung von therapeutischen Teams (Medizin, Pflege, Therapien), finanzielle Ressourcen, etc. Diese und weitere Kontextfaktoren beeinflussen nicht nur die Behandlungsrichtlinie und die Art ihrer Implementierung, sondern auch, wie diese wirksam werden kann.

Ein **Wirkungsmodell** bildet das logische Ineinandergreifen von Zielsetzungen, Handlungen, erbrachten Leistungen, Reaktionen der Zielgruppen und Wirkungen bei den Betroffenen ab und verdeutlicht damit implizit angenommene Wirkungsketten, die hinter den Interventionen stehen. Je nach Autor wird anstelle

von Wirkungsmodell z.B. auch von „Logischem Modell“ oder von „Programmbaum“ gesprochen (Beyl et al., 2008; Hense, 2010; Kellogg Foundation, 2004; Rossi et al., 2004). Ein Wirkungsmodell fragt danach, was über welchen Weg mit welchen Mitteln/Produkten bei wem erreicht werden soll: Von welchen Annahmen, Bedingungen und welcher Ausgangslage wird ausgegangen? Dazu gehören nebst Kontextfaktoren auch Know-how und Ressourcen, die in ein Programm fließen (Input). Weiter fragt das Wirkungsmodell danach, welche Probleme auf welche Art angegangen werden sollen, wer Beteiligte und Betroffene sind und welche Leistungen oder Produkte erbracht werden, resp. aus dem Programm hervorgehen (Aktivitäten). Schließlich definiert das Wirkungsmodell auch Resultate und beabsichtigte Wirkungen, die bei Zielgruppen, also z.B. Patienten, erwartet werden (outcomes) und welche übergeordneten gesellschaftlichen Auswirkungen angestrebt werden (impact). Zur Planung der Evaluation ist ein Wirkungsmodell eine unabdingbare Hilfe: Es dient der Klärung von Kontext, Evaluationsgegenstand, Zweck und Fragestellungen und sollte im Vorfeld der Evaluation in Absprache mit Auftraggebern und zentralen Stakeholdern erstellt werden.

Ein Beispiel: Kontinuität in der medizinisch-pflegerischen Behandlung ist vor allem bei Menschen mit chronischen Krankheiten, in komplexen Patientensituationen und bei limitierten Selbstpflegefähigkeiten wichtig (Biem et al., 2003; van Servellen, Fongwa & Mockus D'Errico, 2006). Betagte Menschen nach einem Sturz mit Frakturfolge gehören aufgrund ihres gebrechlichen Gesundheitszustandes zu dieser Gruppe von pflegebedürftigen, verletzlichen und komplexen Patient/-innen: In der Literatur werden Zusammenhänge zwischen vorbestehendem Gesundheitszu-

stand und (Langzeit-) Resultaten aufgezeigt, und es wird deutlich, dass Begleiterkrankungen und perioperative Komplikationen die Komplexität und das Risiko der Behandlung erhöhen und deren Resultate verschlechtern (Abrahamsen, van Staa, Ariely, Olson, & Cooper, 2009; French, Bass, Bradham, Campbell & Rubenstein, 2008; Roche, Wenn, Sahota & Moran, 2005). Ein Behandlungspfad für solche Patienten muss demnach standardisierte Vorgehen sowohl zur Diagnostik/ Behandlung der Fraktur als auch zur Erkennung potentieller Komplikationen und Präventionsmaßnahmen für Folgefrakturen beinhalten. Spezifische geriatrische Assessments sollen standardisiert durchgeführt, im interdisziplinären Team besprochen und auf den Patienten angepasst werden. (Pretto et al., 2009; Suhm et al., 2009). Damit soll erreicht werden, dass postoperative Komplikationen vermieden resp. rasch erkannt und behandelt werden. Die festgesetzte Behandlung wird über die Institutionsgrenzen hinweg fortgeführt, der Behandlungspfad soll potentielle Informationsbrüche zwischen verschiedenen Leistungserbringern und über Institutionsgrenzen hinweg überbrücken und damit die Behandlungsqualität und Patientensicherheit erhöhen: Der Patient erlebt Kontinuität in Form einer kohärenten, zusammenhängenden Betreuung. Die unterschiedlichen Leistungserbringer erleben Kontinuität auf zwei Arten: Einerseits in Form von genügend Wissen und Informationen über ihre Patienten, um ihre Fachkompetenz bestmöglich einzubringen, andererseits als Vertrauen darauf, dass ihre Behandlungsbeiträge aufgenommen und fortgesetzt werden (Biem et al., 2003; Haggerty et al., 2003; Kripalani et al., 2007). Nebst der Verkürzung der Spital-Aufenthaltsdauer werden durch einen solchen Behandlungspfad bessere Resultate bei Patient/innen erwartet (z.B. weniger Folgefrakturen, re-

duzierte Re-Hospitalisationsraten, tiefere Mortalität und ein größerer Anteil an Patient/-innen, welche ein Jahr nach der Fraktur immer noch im eigenen Heim leben (siehe Abbildung 1, „Beispiel für ein Wirkungsmodell“). Das Wirkungsmodell zeigt die angenommenen Ursache-Wirkungsketten auf und macht deutlich, dass zur Erreichung von Zielen und gewünschten Outcomes bei Patienten alle Fachpersonen, welche in die Behandlung involviert sind, den Behandlungspfad (inkl. Standards) kennen und umsetzen müssen. Die so dargestellte komplexe Intervention „Behandlungspfad“ ist hilfreich, wenn man sich nun der Frage nach Betrachtungsebenen und Fragestellungen der Evaluation zuwendet.

Mögliche Betrachtungsebenen und Fragestellungen an die Evaluation

Bei der Festlegung von Betrachtungsebenen und Fragestellungen an die Evaluation gilt es, verschiedene Elemente zu beachten, wobei das oben dargestellte Wirkungsmodell eine gute Grundlage bietet. Wir müssen uns die Frage stellen, wozu die Evaluation dienen soll (Zweck) und über welche Aspekte des Behandlungspfades (Strukturen, Prozesse, Outcomes) die Evaluation zweckdienliche Informationen liefern kann, wobei der Entwicklungsstand eines Programms eine entscheidende Rolle bei der Fest-

legung von Evaluationszweck und Fragestellungen an die Evaluation spielt (vgl. Frage zwei). Das heißt, je nach Stand des Projektes machen unterschiedliche Fragen an die Evaluation Sinn, und nicht jede Frage kann und soll (z.B. zu Beginn, während der Einführung, oder nach Implementierung eines Programms) gestellt werden. Der Evaluationszweck richtet die Evaluation aus und gibt eine Orientierung, wozu die Evaluationsergebnisse genutzt werden (Beywl et al., 2008). Eine funktionale Unterscheidung von Evaluationszwecken steht in Abhängigkeit der Programmreife und unterscheidet Phasen vor, während und nach der Einführung eines Programms (Fierz, 2010; Rossi et al., 2004). Diese sind nachfolgend dargestellt und mit Beispielen illustriert.

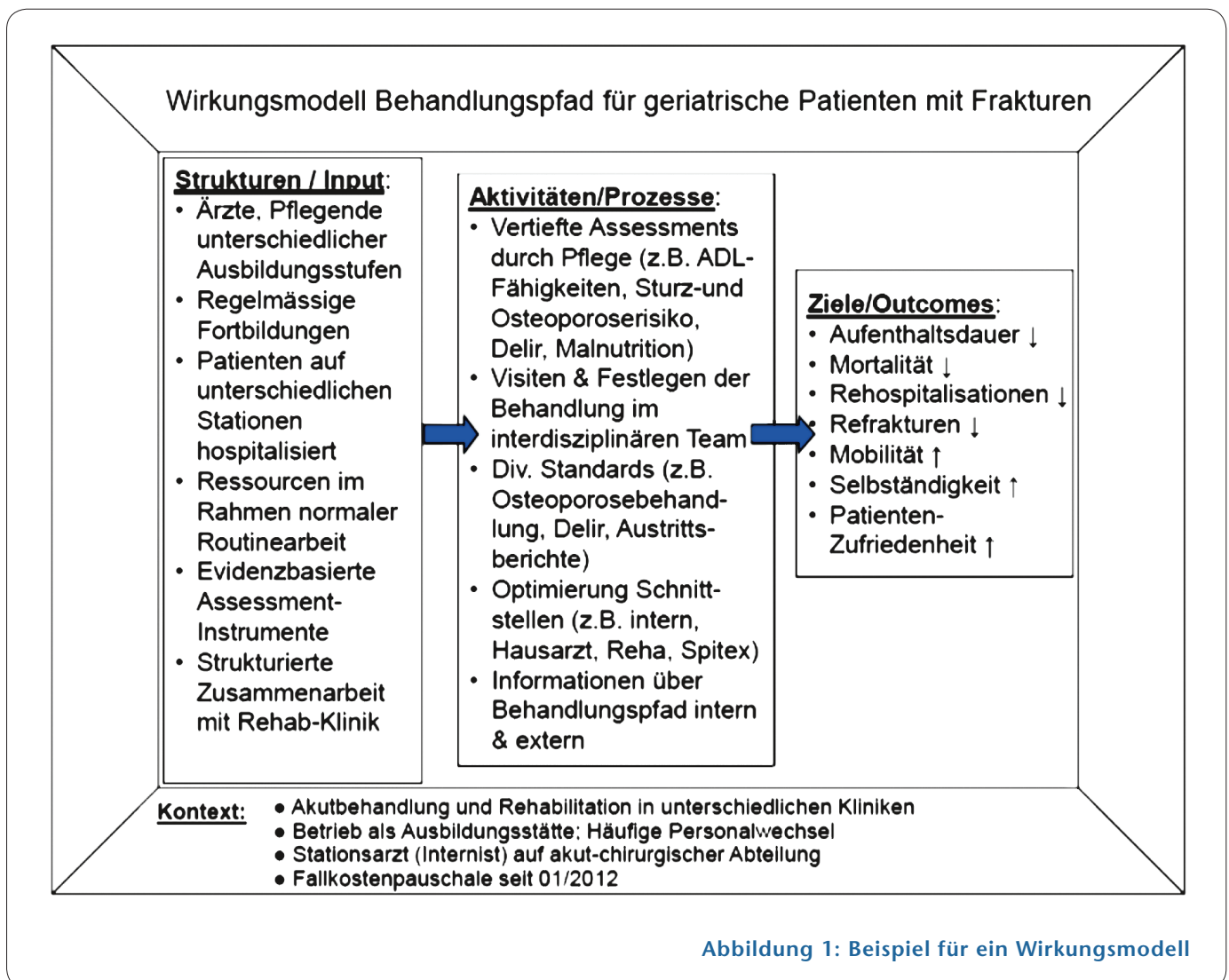


Abbildung 1: Beispiel für ein Wirkungsmodell

Evaluation in der Planungsphase, vor Einführung der Intervention: Hier wird von **proaktiver und klärender Evaluation** gesprochen. Vor Einführung eines neuen Programms muss bspw. ermittelt werden, ob ein Bedarf danach besteht. Es wird gefragt, wie groß das Problem ist, welche Personen davon betroffen sind, welche deren Bedürfnisse sind oder welche Angebote benötigt werden. In einem zweiten Schritt muss die Detailkonzeption evaluiert und Fragen nach dem Programmdesign gestellt werden. **Mögliche Fragen** sind z.B.: „Für wen soll die Intervention sein?“, „Sind die Ziele klar?“, „Wie soll die Intervention im Detail aussehen?“ , „Sind Maßnahmen wie geplant umsetzbar?“, „Wie werden Zielgruppen erreicht?“, „Wie soll das Programm implementiert werden?“ (Beywl et al., 2008; Rossi et al., 2004).

Beispiele: Vor der Einführung eines Behandlungspfades könnte man sich z.B. fragen: Wie viele entsprechende Patienten werden bei uns betreut und wer kann vom Behandlungspfad profitieren? Welche Abläufe und Standards sind heute bereits vorhanden (Ist-Analyse)? Welche Interventionen werden benötigt (evidenzbasiertes Wissen)? Welche Fachpersonen im Betrieb können/sollen zur Ausarbeitung des Behandlungspfades einbezogen werden? Wie werden der Behandlungspfad und darin enthaltene Standards bei beteiligten Fachpersonen, die damit arbeiten sollen (Stakeholder), in- und außerhalb des Spitals bekannt gemacht? Wie werden die Maßnahmen konkret umgesetzt? Wer trägt welche Verantwortung für die Behandlungskontinuität und Umsetzung der vorgesehenen Maßnahmen? Welche (personellen und finanziellen) Ressourcen sind vorhanden, um den Behandlungspfad zu implementieren und zu überprüfen? Welches (Vor-) Wissen haben Pflegendende und Ärzte in Bezug auf spezifische Assessments und Interventionen

im Rahmen des Behandlungspfades resp. wo sind Fortbildungen notwendig? Welche Strukturen und Abläufe sind nötig, um relevante Informationen (von Ärzten, Pflegenden und TherapeutenInnen) im Rahmen des Behandlungspfades an die nächste Institution weiterzuleiten? Wo könnten in der Implementierung des Behandlungspfades Probleme auftreten und wie kann diesen vorgebeugt werden?

Evaluation parallel zur Implementierung eines Programms: Hier wird von „Interaktiver Evaluation“, von „Process or Implementation Evaluation“ oder von Umsetzevaluierung gesprochen. Diese Art von Evaluation erfolgt parallel zur Umsetzung eines Programms und gibt Hinweise, welche Elemente beibehalten, welche optimiert oder ausgetauscht werden sollten, und in welchem Ausmaß Zielpersonen in der geplanten Weise mitarbeiten (Implementationskontrolle, compliance). Es wird deshalb auch von „formativer Evaluation“ oder von „dynamischer Evaluation“ gesprochen. In gewissen Fachgebieten wird hier auch die „Aktionsforschung“ genannt (McCormack & Manley, 2009). **Mögliche Fragen**, welche in dieser Phase gestellt werden können, sind z.B.: „Stimmt der tatsächliche Ablauf des Programms mit dem geplanten überein?“, „Ist das Programm gut organisiert und arbeiten die Berufsangehörigen wie vorgesehen zusammen?“, „Werden Interventionen wie geplant umgesetzt und falls nicht, aus welchen Gründen?“ (Implementierung), „Gibt es Betroffene, die die Intervention benötigen, aber nicht erhalten – falls ja, warum nicht?“ (Zielgruppenerreichung). In dieser Phase gehört (spätestens) auch die Überprüfung der Ressourcen, z.B. personell, administrativ, finanziell oder Know-how, welche die Implementierung des Programms sicherstellen sollen, und die Frage nach adäquaten Ressourcen zur Unterstüt-

zung wichtiger Programmfunktionen (Beywl et al., 2008; Fierz, 2010; Rossi et al., 2004).

Einen hilfreichen Rahmen für einen umfassenden Evaluationsansatz bietet das „RE-AIM-Modell“, welches zur Evaluation von Gesundheits- und Präventionsprogrammen eingesetzt wird und welches verschiedene Dimensionen – nebst der Wirksamkeit v.a. die Umsetzung – fokussiert (Glasgow, Klesges, Dziewaltowski, Estabrooks & Vogt, 2006; Glasgow et al., 1999). „RE-AIM“ steht für R = reach (erreicht die Intervention die Zielgruppe?), E = efficacy (ist die Intervention wirksam?), A = adoption (wird die Intervention von relevanten Organisationen/Stakeholdern übernommen?), I = implementation (wird die Intervention wie intendiert umgesetzt?) und M = maintenance (wird die Intervention über einen längeren Zeitraum aufrechterhalten?).

Beispiele: Während der Umsetzung eines Behandlungspfades könnte bspw. gefragt werden: Wie viele Patienten (relativ zur Gesamtzielgruppe) wurden gemäß den Vorgaben des Behandlungspfades behandelt? In welchem Ausmaß finden die im Behandlungspfad vorgegebenen Assessments und (Präventions-) Maßnahmen statt? In welchem Ausmaß sind diese Interventionen gemäß Behandlungspfad Teil der Routine-Behandlung geworden? Kennen alle Mitarbeitenden den Behandlungspfad und werden die einzelnen Behandlungspfad-Anteile von allen Mitarbeitenden (auch von neuen) umgesetzt? Wie (vollständig und verständlich) werden Informationen zur Behandlung des Patienten an nachfolgende Institutionen und Leistungserbringer weitergeleitet und wie werden Behandlungen weitergeführt?

Die **wirkungsfeststellende Evaluation**, auch **summative Evaluation** oder Wirkungsevaluation genannt,

ist meistens von zentralem Interesse für außenstehende Fachpersonen, für Geldgeber, für Führungsverantwortliche. Diese Art der Überprüfung fragt nach der Wirksamkeit eines Programms bei der Zielgruppe. Die wichtigste Voraussetzung für eine wirkungsfeststellende Evaluation ist jedoch ein reifes Programm mit bewährten Konzepten (hier im Beispiel ein Behandlungspfad inkl. Standards), und es empfiehlt sich, mit dieser Art von Evaluation zu warten, bis Programm/Interventionen etabliert sind. Erst wenn ein Programm steht, wie gewünscht läuft und sich eine Routine etabliert hat, ist es sinnvoll, dessen Wirkungen zu evaluieren. **Mögliche Fragen**, welche hier gestellt werden können, sind z.B.: „Werden durch das Programm bei der Zielgruppe (z.B. Patienten) beabsichtigte Wirkungen (Ziele des Programms) erreicht?“, „Inwieweit sind Veränderungen in der Zielgruppe auf das Programm zurückzuführen?“

Wird die Wirkungsevaluation zu früh durchgeführt, können Störfaktoren (z.B. ungenügende Programm-Implementierung, fehlende Routine, fehlende Systematik) die Resultate verwässern oder Wirkungen verschleiern, sodass z.B. Unterschiede zwischen Interventions- und Vergleichsgruppe nicht nachgewiesen werden können. Umgekehrt gilt es ebenso zu vermeiden, dass festgestellte „Wirkungen“ voreilig einem Programm zugeschrieben werden, obwohl sie möglicherweise mit – ungenügend beachteten – Kontextfaktoren zusammenhängen: Evaluationen verwenden mehrheitlich quasi-experimentelle Forschungsdesigns, z.B. Vorher-Nachher-Designs, welche Schlüsse zur Kausalität nur begrenzt zulassen. So könnte z.B. eine verkürzte Aufenthaltsdauer vielmehr mit der Einführung des DRG-Finanzierungssystems als mit einem neuen Behandlungspfad zusammen-

hängen. Anders als bei experimentellen Designs/randomisierten Studien, bei welchen möglichst viele Störfaktoren ausgeschaltet werden, kann bei den nicht-experimentellen Designs ein Gruppen-Unterschied nicht zweifelsfrei auf ein (komplexes und multifaktorielles) Programm oder eine einzelne Intervention zurückgeführt werden (Fierz, 2010; Rossi et al., 2004; Westermann, 2002).

Beispiele: Nach der Implementierung eines Behandlungspfades für geriatrische Patienten mit Frakturen könnte eine Wirkungsevaluation nach folgenden **Outcome- und Impact-Parametern** fragen:

Wie und in welchem Ausmaß haben sich Patientenresultate durch die Umsetzung des Behandlungspfades verändert (Vorher-Nachher-Design)? Konkret: Wie verändert sich die Anzahl der Patienten mit postoperativen Komplikationen (z.B. Delir, Malnutrition)? Wie verändert sich die durchschnittliche Hospitalisationsdauer in der Akut- und Rehabilitationsklinik? Wie verändern sich Mortalitätsraten (im Spital, nach 30 Tagen, nach 1 Jahr)? Wie verändert sich der Anteil der Patienten, welche ein Jahr nach der Fraktur neu im Heim resp. noch immer zuhause leben? Wie verändert sich die Anzahl der Frakturen aufgrund eines erneuten Sturzes resp. die Anzahl notwendiger Re-Hospitalisationen aufgrund der initialen Fraktur? Wie verändern sich Mobilität, funktionale und kognitive Fähigkeiten von Patienten ein Jahr nach der Fraktur im Vergleich zu vor der Fraktur? Wie gut oder schlecht sind diese Resultate verglichen mit Erkenntnissen aus der Literatur?

Im Rahmen der Impact-Evaluation könnte interessieren, welche Veränderungen sich durch den Behandlungspfad auf übergeordneter Ebene zeigen: Wie stellt sich ein Kosten/Nutzen-Vergleich von heute resp.

früher behandelten Hüftfraktur-Patienten dar? Wie hat sich die Praxis der Zuweisungen, die Zusammenarbeit zwischen Stakeholdern in und außerhalb des Spitals verändert? Inwiefern wird die Klinik als Altersfraktur-Kompetenzzentrum anerkannt?

Diskussion: Chancen und Risiken der Evaluation

Eine sorgfältige Vorbereitung der Evaluation unter Einbezug der relevanten Stakeholder und des Kontextes ist die Voraussetzung, um eine qualitativ gute Evaluation mit zuverlässigen und aussagekräftigen Resultaten zu erhalten (Widmer et al., 2000). Werden die vorgängigen Überlegungen auf die Evaluation eines Behandlungspfades bezogen, zeigen sich verschiedene Chancen und Risiken, welche nachfolgend beleuchtet werden (vgl. Frage drei).

In der skizzierten Vorgehensweise – Umsetzungsevaluation im ersten Schritt und nachfolgende Wirkungsevaluation als zweiter Schritt – besteht die **Chance**, dass gezielte Verbesserungen in der Programm-Implementierung vorgenommen werden können. Schwächen in der Programmumsetzung oder der Programmkonzeption werden erkannt und Verbesserungsmaßnahmen können einfließen bevor eine Wirkungsevaluation stattfindet. Die Darstellung des Programms in einem Wirkungsmodell ist zudem hilfreich, um implizit angenommene Wirkungsketten transparent zu machen, um Umsetzungs- und Wirkungsevaluation zu planen und um die richtigen Fragen zum richtigen Zeitpunkt zu stellen (vgl. Beispiele von Fragestellungen oben). Die Evaluation wird so Teil des Qualitätsmanagementprozesses und stellt Informationen darüber zur Verfügung, aus wessen Sicht, für wen, unter welchen Bedingungen und zu welchem Preis «etwas funktioniert». Eine Evaluation

ist von begrenztem Nutzen, wenn sie ausschließlich im Nachhinein Rückmeldungen liefert (Bundesamt für Gesundheit, 2005). Ein Programm oder Behandlungspfad wie im oben aufgeführten Beispiel würde durch Befragung von Stakeholdern im Rahmen der Evaluation zudem einen impliziten Evaluationsnutzen bringen, indem der Behandlungspfad bekannter gemacht wird. Der wichtigste Vorteil dieser Herangehensweise, wo die Wirkungsevaluation erst nach der überprüft-guten Implementierung von Interventionen erfolgt, ist die größere Wahrscheinlichkeit, dass Ergebnisse zuverlässig sind, Veränderungen durch Intervention festgestellt werden können und diese Veränderungen auch tatsächlich auf die Interventionen/den Behandlungspfad zurückgeführt werden können, resp. nicht andere Ursachen haben (z.B. ungenügende Beachtung von Kontextfaktoren).

Ein großes Risiko – und ein in Evaluationen oft begangener Fehler – ist die verfrühte Konzentration auf Outcomes oder Impact eines Programms (Fierz, 2010). Auf der einen Seite besteht bei einer zu früh angesetzten Outcome-Evaluation das Risiko darin, dass die Intervention selbst, deren Maßnahmen und in ihr wirksame Mechanismen im Dunkeln bleiben. Damit wird die Chance vertan, potentielle Verbesserungen zur Behandlungspfad-Implementierung zu erkennen und gezielt umzusetzen. Die beschriebenen Vorher-Nachher-Designs von Evaluationen liefern – auch wenn sie mit großer methodischer Sorgfalt durchgeführt werden – meistens nur Informationen zur Wirksamkeit von Maßnahmen im Sinne der Zielerreichung, während die Prozesse ignoriert werden, welche zu diesen Effekten geführt haben. Da aber gerade diese Prozesse von erheblicher praktischer Relevanz sind, wenn es darum geht, Maßnahmen zu verbessern und in der Brei-

te zu implementieren, sind diese als „Black-Box-Evaluationen“ bezeichneten Analysen für die Praxis oft nur von geringem Wert (Hense, 2010).

Ein weiteres Risiko der vorzeitigen Wirkungsevaluation könnte bei einem Behandlungspfad auch darin bestehen, dass dessen Wirkungen bei suboptimaler Implementierung nicht sichtbar gemacht werden können. Ohne Evaluation der Programmumsetzung müsste ev. festgestellt werden, dass die beabsichtigte Wirkung nicht erreicht wurde, oder sogar, dass das Gegenteil der erhofften Wirkung eingetreten ist, ohne zu wissen, ob dies mit dem Behandlungspfad selbst im Zusammenhang steht oder nicht.

Kontextfaktoren, die nicht in die Evaluation einbezogen werden, können ebenfalls ein Risiko darstellen, da sie sich auf Ergebnisse auswirken können ohne dass dies bedacht wird. Die Einführung des DRG-Finanzierungssystems führt z.B. beim vorliegenden Beispiel dazu, dass sowohl in der Akut- als auch in der Rehabilitationsklinik zunehmend Druck entsteht, die Aufenthaltsdauer zu verkürzen. Im klinischen Alltag wird auf eine schnelle Verlegung des Patienten in die Rehabilitationsklinik resp. zurück ins Pflegeheim gedrängt. Die akut-Hospitalisationsdauer sinkt also möglicherweise aufgrund dieses Drucks und nicht als Zeichen besserer Resultate durch den Behandlungspfad. Es wäre sogar vorstellbar, dass eine kürzere Akut-Aufenthaltsdauer nicht im beabsichtigten Gewinn für Patienten im Sinne eines besseren Gesundheitszustandes resultiert, sondern sich lediglich Probleme in nachfolgende Institutionen verlagern, und dass innerhalb des Behandlungspfad ein eigentlicher Zielkonflikt besteht (Ziel „kürzere Aufenthaltsdauer in Akutklinik“ und gleichzeitig „weniger Komplikationen“): Bei hoher Komorbidität sind bei betagten Hüftfraktur-Patienten Delirien, Infektionen und andere

Probleme eher Regel denn Ausnahme, und erschweren die stationäre Behandlung (Specht-Leible et al., 2003). Die damit verbundene Belastung für betreuende Fachkräfte und Anforderungen steigen also auch in den Rehabilitationskliniken, wo Personal weniger für Akutsituationen ausgebildet ist. Dadurch könnten vermehrt Probleme entstehen, welche mittelfristig Outcomes negativ beeinflussen (z.B. anhaltendes Delir, erneute Stürze, Re-Hospitalisationen) und damit verbunden könnten sich – als Konsequenz der verkürzten Hospitalisationsdauer – auch langfristige Outcomes wie z.B. Mobilität verschlechtern.

Schlussfolgerungen und Relevanz für die Pflegeforschung

Bei der Evaluation von Behandlungspfaden und komplexen Interventionsprogrammen gilt es, wesentliche Aspekte der Evaluation bereits in der Vorbereitung und Planung zu berücksichtigen. Nur so können Evaluationsergebnisse zuverlässig interpretiert, ein Nutzen erreicht und damit den geforderten Qualitätskriterien für eine gute Evaluation entsprochen werden (Bundesamt für Gesundheit, 2005; Widmer et al., 2000):

Ein Wirkungsmodell ist im Vorfeld einer Evaluation immer und unter Einbezug der zentralen Stakeholder zu erarbeiten. Dabei sind die Kontextfaktoren eines Programms resp. Behandlungspfad zu berücksichtigen und Überlegungen zu Ursache-Wirkungs-Beziehungen sowie zur Implementierung vorzunehmen. Auf diese Weise wird sichergestellt, dass für die Evaluation zur richtigen Zeit die richtigen Fragen gestellt werden (Fierz, 2010).

Die Evaluation der Programmimplementierung ist zentral und hat pa-

rallel zur Programmeinführung zu erfolgen, um dieses fortlaufend zu optimieren. Die optimale Umsetzung des Programms sollte sichergestellt sein, bevor eine Wirkungsevaluation vorgenommen wird. Das RE-AIM-Modell bietet einen hilfreichen Rahmen für eine umfassende, strukturierte Behandlungspfad-Evaluation (Glasgow et al., 1999).

Bei der Wirkungsevaluation und insbesondere bei der Betrachtung von Patienten-Outcomes in Vorher-Nachher-Designs ist Vorsicht geboten, wenn Resultate in Verbindung mit dem Programm gebracht werden: Beeinflussende äußere Faktoren und der Stand der Programmimplementierung müssen bei der Interpretation von Ergebnissen (Veränderungen und Nicht-Veränderungen) beachtet werden. Resultate sollten als Gesamtbild, also primäre und sekundäre Outcomes (z.B. Mortalität, Mobilität, Komplikationsrate, Lebenssituation nach einem Jahr) betrachtet werden; das Herausgreifen einzelner weniger Outcomes führt zu einem verfälschten Bild.

Der vorliegende Artikel macht deutlich, wie eine gute Evaluation im Sinne der nationalen Qualitätskriterien des Bundesamtes für Gesundheit (SEVAL-Standards) einen Nutzen für die beteiligten Stakeholder bringen kann (Widmer et al., 2000). Die Evaluation von Programmen/Behandlungspfaden muss dazu jedoch über eine ausschließliche Wirkungsevaluation hinausgehen, indem sie Stakeholder, Kontextfaktoren und Implementierungsstand mit einbezieht.

Literatur

Abrahamsen, B., van Staa, T., Ariely, R., Olson, M., & Cooper, C. (2009). Excess mortality following hip fracture: a systematic epidemiological review. *Osteoporosis Int.*

Barbieri, A., Vanhaecht, K., Van Herck, P., Sermeus, W., Faggiano, F., Marchisio, S., & Panella, M. (2009). Effects of clinical pathways in the joint replacement: a meta-analysis. [Comparative Study Meta-Analysis]. *BMC Med*, 7, 32.

Beywl, W., Kehr, J., Mäder, S., & Nestroj, M. (2008). Evaluation Schritt für

Schritt: Planung von Evaluationen. Münster: Heidelberger Institut Beruf und Arbeit hiba GmbH, Münster.

Biem, H. J., Hadjistavropoulos, H., Morgan, D., Biem, H. B., & Pong, R. W. (2003). Breaks in continuity of care and the rural senior transferred for medical care under regionalisation. *Int J Integr Care*, 3, e03.

Bundesamt für Gesundheit. (2005). Wirksamkeitsüberprüfung von Massnahmen in der öffentlichen Gesundheitspolitik - Strategischer Rahmen, Jan 2011, from <http://www.bag.admin.ch/evaluation/index.html?lang=de>

Fierz, K. (2010). „Evaluiert doch einfach einmal“ ... die Kunst, eine gute (Evaluations-)Frage zu stellen. *Pflege*, 6(2010) - Editorial

French, D. D., Bass, E., Bradham, D. D., Campbell, R. R., & Rubenstein, L. Z. (2008). Rehospitalization after hip fracture: predictors and prognosis from a national veterans study. *J Am Geriatr Soc*, 56(4), 705-710.

Glasgow, R. E., Klesges, L. M., Dzewaltowski, D. A., Estabrooks, P. A., & Vogt, T. M. (2006). Evaluating the impact of health promotion programs: using the RE-AIM framework to form summary measures for decision making involving complex issues. [Research Support, Non-U.S. Gov't]. *Health Educ Res*, 21(5), 688-694.

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: the RE-AIM framework. [Research Support, U.S. Gov't, P.H.S. Review]. *Am J Public Health*, 89(9), 1322-1327.

Haggerty, J. L., Reid, R. J., Freeman, G. K., Starfield, B. H., Adair, C. E., & McKendry, R. (2003). Continuity of care: a multidisciplinary review. *BMJ*, 327(7425), 1219-1221.

Hense, J. (2010). Programmtheorien und logische Modelle in der Evaluation. Lehrtext im Rahmen Studiengang DAS in Evaluation, Universität Bern

Kellogg Foundation. (2004). Logic Model Development Guide, from <http://www.wkkf.org/knowledge-center/resources/2006/02/WK-Kellogg-Foundation-Logic-Model-Development-Guide.aspx> Zugriff am 2.10.2012

Kripalani, S., LeFevre, F., Phillips, C. O., Williams, M. V., Basaviah, P., & Baker, D. W. (2007). Deficits in communication and information transfer between hos-

pital-based and primary care physicians: implications for patient safety and continuity of care. *JAMA*, 297(8), 831-841.

McCormack, B., & Manley, K. (2009). Die Evaluation von Praxisentwicklung. In I. A. Frei & R. Spirig (Eds.), *In: Praxisentwicklung in der Pflege* (pp. 81-103). Bern: Verlag Hans Huber.

Mühlhauser, I., Lenz, M., & Meyer, G. (2012). Bewertung von komplexen Interventionen: Eine methodische Herausforderung. *Dtsch Arztebl* 109(1-2).

Pretto, M., Muri-John, V., Suhm, N., Frank, M., Kressig, R. W., & Spirig, R. (2009). Frakturen bei Betagten: Multidisziplinäre Behandlung und Betreuung. [*Clinical Updates*]. *Krankenpflege*(9), 32-33.

Roche, J. J., Wenn, R. T., Sahota, O., & Moran, C. G. (2005). Effect of comorbidities and postoperative complications on mortality after hip fracture in elderly people: prospective observational cohort study. *Bmj*, 331(7529), 1374.

Rossi, P., Lipsey, M., & Freeman, H. (2004). *Evaluation - A Systematic Approach* (7th Edition ed.). California: SAGE Publications, Thousand Oaks.

Specht-Leible, N., Schultz, U., Kraus, B., Meeder, P. J., Quentmeier, A., Ewerbeck, V., . . . Oster, P. (2003). [Case management and functional outcome in persons aged 65 years and over with hip fracture]. *Unfallchirurg*, 106(3), 207-214.

Suhm, N., Pretto, M., Frank, M., Conzelmann, M., Tyndall, A., Vogt, T., . . . Spirig, R. (2009). Moderne Alterstraumatologie: Interdisziplinäre & Interprofessionelle Behandlung im „Kompetenznetzwerk Altersfrakturen“. *The Medical Journal*, 3:16-20.

Van Servellen, G., Fongwa, M., & Moccus D'Errico, E. (2006). Continuity of care and quality care outcomes for people experiencing chronic conditions: A literature review. *Nurs Health Sci*, 8(3), 185-195.

Westermann, R. (2002). Merkmale und Varianten von Evaluationen: Überblick und Klassifikation. *Zeitschrift für Psychologie, Sonderdruck*, 210(1), 4-26.

Widmer, T. (2006). Qualität der Evaluation - Wenn Wissenschaft zur praktischen Kunst wird. In Stockmann & R. (Hrsg.) (Eds.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (3. Auflage ed., Vol. Band 1). Münster: Waxmann Verlag, GmbH, Münster.

Widmer, T., Landert, C., & Bachmann, N. (2000). *Evaluations-Standards der*

Schweizerischen Evaluationsgesellschaft (SEVAL-Standards) <http://www.seval.ch/de/standards/index.cfm> Zugriff am 07.12.2011