

G. Klever-Deichert, A. Gerber, MA Schröer, E. Plamper

## **International erfolgreiche Interventionen der Prävention und Gesundheitsförderung und ihre Übertragbarkeit auf Deutschland**

Entwicklung und exemplarische Anwendung eines Bewertungsinstruments

Abschlussbericht

---

### **Auf einen Blick...**

- Für die Risikofaktoren ergaben sich zwischen rund 12.500 und 5.000 Treffer. Nach Berücksichtigung von Ein- und Ausschlusskriterien verblieben rund 10 % der Studien. Studien zu homogenen Handlungsfeldern wurden vertiefend analysiert.
- Die Analyse ergab, dass die Studien- und Interventionsqualität nicht mit den Effekten der Studien korrelieren.
- Eine Auswahl einzelner Studien wurde dargestellt, die hohe Werte der Studien- und der Interventionsqualität erzielten. Sofern möglich, wurden daraus Empfehlungen zur Übertragung auf Deutschland gemacht.
- Aus der Analyse anhand des Kriterienkatalogs ergibt sich kein Automatismus zur Empfehlung von Interventionen. Der Katalog wird vielmehr als Screeninginstrument zu Qualitätsstandards und Wirksamkeit verstanden, der eine entsprechende Auswahl von Interventionen liefert, die einer ergänzenden Bewertung durch Expertengremien bedürfen.
- Der Kriterienkatalog ist potenziell entwicklungsfähig. Kriterien, die derzeit nur als beschreibende Indikatoren erhoben werden, können zu qualitätsbewertenden Indikatoren umgewandelt werden. Außerdem kann der Kriterienkatalog um eine Kosten-Nutzen-Evaluation von Interventionen erweitert werden.



FORSCHUNGSBERICHTE DES  
INSTITUTS FÜR GESUNDHEITSOÖKONOMIE UND KLINISCHE  
EPIDEMIOLOGIE DER UNIVERSITÄT ZU KÖLN



## **STUDIEN ZU GESUNDHEIT, MEDIZIN UND GESELLSCHAFT**

NR. 09/ 2007 VOM 12.11.2007 \*\*\* ISSN 1862-7412 \*\*\* WWW.IGKE.DE/SGMG

---

# **International erfolgreiche Interventionen der Prävention und Gesundheitsförderung und ihre Übertragbarkeit auf Deutschland**

## **Entwicklung und exemplarische Anwendung eines Bewertungsinstruments**

---

Ein Projekt des Institut für Gesundheitsökonomie und Klinische Epidemiologie (IGKE)  
der Universität zu Köln, gefördert von der Hans Böckler Stiftung (Projekt-Nr. S-2006-832-4).  
Mai 2006- Juli 2007

Klever-Deichert G  
Gerber A  
Schröer MA  
Plamper E

Korrespondierende Autorin: milly.schroeer@uk-koeln.de



Die Reihe „Studien zu Gesundheit, Medizin und Gesellschaft“ umfasst Arbeits- und Forschungsberichte des Instituts für Gesundheitsökonomie und Klinische Epidemiologie der Universität zu Köln sowie kooperierender Wissenschaftler und Institutionen.

Die Berichte und weitere Informationen zu den Forschungsberichten können im Volltext abgerufen werden unter

<http://www.igke.de/SGMG>

Bitte zitieren Sie vorliegenden Bericht als

*Klever-Deichert G, Gerber A, Schröer MA, Plamper E. International erfolgreiche Interventionen der Prävention und Gesundheitsförderung und ihre Übertragbarkeit auf Deutschland. Studien zu Gesundheit, Medizin und Gesellschaft 2006; Köln: Ausgabe 09/2007 vom 12.11.2007.*

# Inhaltsverzeichnis

<b>1</b>	<b>Zusammenfassung</b> .....	<b>7</b>
1.1	Hintergrund und Fragestellungen .....	9
1.2	Können Empfehlungen für Präventionsprogramme ausgesprochen werden? .....	10
1.3	Empfehlungen.....	13
1.4	Methode des Projekts.....	14
<b>2</b>	<b>Einleitung: Ziel und Gegenstand des vorliegenden Forschungsprojekt</b> .....	<b>17</b>
<b>3</b>	<b>Evaluation von Interventionen in Prävention und Gesundheitsförderung</b> .....	<b>21</b>
3.1	Prävention und Gesundheitsförderung.....	21
3.2	Evidenzbasierte Medizin .....	23
3.3	Evidence-based Health Promotion als eigenständige Methode .....	27
3.4	Evaluation und Indikatoren der Qualität: Nachvollziehbarkeit und Konsens.....	34
3.5	Von der Bewertung zur Empfehlung .....	37
<b>4</b>	<b>Der IGKE-Kriterienkatalog zur Evaluation von Interventionen</b> .....	<b>39</b>
4.1	Rationale des IGKE-Kriterienkatalogs .....	39
4.2	Der IGKE- Kriterienkatalog .....	41
4.3	Manual zum IGKE-Kriterienkatalog .....	47
4.4	Auswertung der qualitätswirksamen Fragen des IGKE-Kriterienkatalogs.....	59
4.5	Die Umsetzung in Empfehlungen .....	61
4.6	Weiterentwicklungen des IGKE-Kriterienkatalogs.....	62
<b>5</b>	<b>Methodisches Vorgehen</b> .....	<b>65</b>
5.1	Systematische Literatur- und Datenbankrecherche.....	65
5.2	Der IGKE-Kriterienkatalog .....	70
<b>6</b>	<b>Bewegung</b> .....	<b>71</b>
6.1	Stand der Forschung.....	71

6.2	Ergebnisse.....	79
6.3	Diskussion .....	99
<b>7</b>	<b>Depression.....</b>	<b>102</b>
7.1	Stand der Forschung.....	102
7.2	Ergebnisse.....	106
7.3	Diskussion .....	119
<b>8</b>	<b>Ernährung.....</b>	<b>122</b>
8.1	Stand der Forschung.....	122
8.2	Ergebnisse.....	126
8.3	Diskussion .....	137
<b>9</b>	<b>Rauchen.....</b>	<b>140</b>
9.1	Stand der Forschung.....	140
9.2	Ergebnisse.....	143
9.3	Diskussion .....	159
<b>10</b>	<b>Gesamtdiskussion und Fazit.....</b>	<b>161</b>
10.1	Zur Situation der Prävention .....	161
10.2	Katalog zur Priorisierung von Präventionsvorhaben.....	161
10.3	Zur Übertragbarkeit und Wiederholbarkeit von Studienergebnissen.....	162
10.4	Verhaltensprävention versus Verhältnisprävention.....	165
10.5	Empfehlung auf Basis der Evaluation.....	166
10.6	Übertragung der Ergebnisse auf Deutschland .....	168
10.7	Ertrag für die Priorisierung von Interventionen.....	170
<b>11</b>	<b>Literatur .....</b>	<b>172</b>
11.1	Literatur Evaluation .....	172
11.2	Literatur Bewegung.....	177
11.3	Literatur Depression .....	188
11.4	Literatur Ernährung .....	195
11.5	Literatur Rauchen.....	201



# 1 Zusammenfassung

In Deutschland konzentriert sich die Diskussion in der Prävention und Gesundheitsförderung derzeit wesentlich auf die Mittelherkunft, also darauf, wer die Angebote finanziert. Kaum eine Diskussion gibt es um die entscheidendere Frage, für welche Projekte die Mittel ausgegeben werden sollten. Diese Frage hat zwei Facetten:

1. Welche Krankheitsbilder oder Zustände sollen verhindert oder in ihrem Verlauf abgeschwächt oder verzögert werden?
2. Welche Programme sind überhaupt wirksam?

Die erste Frage muss politisch beantwortet werden, indem nationale Präventionszielen auf der Grundlage von betroffenen Gruppen, Häufigkeiten und auch erwartbaren Folgekosten einer Erkrankung gesetzt werden.

Die zweite Frage soll durch das vorliegende Gutachten beantwortet werden. Grundlegender Gedanke war, dass die Erfahrungen des Auslands systematisch analysiert werden, um wirksame Programme übernehmen zu können und unwirksame nicht erneut aufzulegen. Diese Idee entstammt der Evidenz-basierten Medizin. Auch dort werden klinische Studien systematisch gesichtet und aufbereitet, um schließlich Vorschläge (Leitlinien) an den Arzt geben zu können, welche Therapie Wirksamkeit verspricht

Die Übertragung der Methode der Evidenz-basierten Medizin auf die Prävention ist nicht unumstritten. So ist es schwierig, in einem sehr heterogenen Umfeld Programme zu übertragen, etwa von einem Stadtteil auf einen anderen. Ebenso decken Präventionsprogramme eine Vielzahl von Typen ab, vom Aufhängen von Plakaten bis zum Coaching einzelner Versicherter durch einen Arzt. Die Vergleichbarkeit beziehungsweise Identifizierung der für den Erfolg oder Misserfolg ausschlaggebenden Kernpunkte ist dann schwierig. Drittens schließlich ist der Zusammenhang zwischen Programm und späterem Auftreten einer Erkrankung beziehungsweise eines Zustands generell schwierig, da die betrachteten Zeiträume sehr lang und die Fallzahl (beispielsweise Neuerkrankungen) sehr klein sind.

Ein weiteres Problem in der Bewertung von Präventionsprogrammen besteht auch darin, dass man auf das Einhalten einer so genannten guten Strukturqualität vertraut, und damit eine gewünschte Wirkung des Programms garantiert sieht. Die Effekte selbst werden dann oftmals gar nicht mehr gemessen beziehungsweise geprüft. Solche Strukturmerkmale sind beispielsweise die Beteiligung der Betroffenen an der Entwicklung des Programmes, die Ausbildung von Multiplikatoren und Kriterien der Nachhaltigkeit. Sofern diese Qualitätskriterien gesichert sind, sollte auch das Programm durchgeführt werden. Unser Ansatz hingegen berücksichtigt gleichermaßen die Effektivität, das heißt: Programme, die lediglich gute Strukturqualität bieten, ohne die Effektivität nachweisen zu können, sind letztendlich nicht zwangsläufig empfehlenswert.

Das Gutachten untersuchte systematisch Studien in den Bereichen Bewegung, Ernährung, Rauchen und Depression. Es konnten keine generellen Aussagen zu Merkmalen getroffen werden, welche die Wahrscheinlichkeit erhöhen würden, dass ein Programm auch wirkungsvoll ist. Wesentlich dafür waren mehrere Gründe:

- Die gefundenen Programme waren sehr heterogen, so dass abschließend nicht festgestellt werden konnte, welche Faktoren der Studien Wirksamkeit (oder Nicht-Wirksamkeit) verursachten.
- Viele Programme wurden nicht oder nicht ausreichend evaluiert. Das wissenschaftliche Niveau der Evaluation war oftmals so schwach, dass die Ergebnisse des Programms nicht zweifelsfrei auf das Programm zurück geführt werden konnten.

Einige Erkenntnisse zeigen aber zumindest einen Trend. So sind zur Förderung der **Bewegung** Plakate mit der Aufforderung zum Gebrauch der Treppe statt des Aufzugs in Betrieben nicht effektiv. Bei Jugendlichen kann eine Zunahme der Bewegung nur mit kombinierten Angeboten, welche Schule und/oder Eltern einbeziehen, erreicht werden, nicht mit Programmen, die rein edukativ ansetzen. Die Änderung des **Ernährungsverhaltens** konnte von isolierten Programmen kaum nachgewiesen werden, lediglich die Kombination mit verstärkter Bewegung unter Einbeziehung der Familien/ Eltern zeigt Effektivität. Die Änderung des **Rauchverhaltens** bei Schwangeren wird am ehesten dann erreicht, wenn als Autoritäts- oder Fachpersonen wahrgenommene Gruppen im Programm beteiligt werden, etwa Ärzte oder Hebammen. Auch im Bereich

Depression scheint sich zu zeigen, dass rein edukative Ansätze kaum Wirkung zeigen, jedoch Programme mit verhaltenstherapeutischen Ansätzen eher effektiv sind.

Als Schlussfolgerung aus diesen Erkenntnissen kann gezogen werden:

- Generell muss zwingend eine Evaluation von Programmen durchgeführt werden. Ansonsten können effektive Programme und nicht-effektive Programme nicht unterschieden werden.
- Die Methoden zur Evaluation von Präventionsprogrammen müssen weiter entwickelt werden. Dazu gehören Definitionen von Zielgrößen, Randomisierung, Kontrollgruppen, Nachbeobachtungszeiträume und die Benennung von Störgrößen.
- Die überprüfbaren Ziele der Programme müssen im Vorfeld festgelegt werden.
- Da der Erfolg der Programme derzeit unsicher ist, sollten alle Programme zeitlich befristet werden.
- Da die erfolgversprechenden Strukturen von Programmen derzeit nicht bekannt sind, sollte möglichst ein Wettbewerb um beste Effektivität und Kosten-Effektivität entstehen. Dies bedeutet, dass eine Begrenzung auf Programme, die lediglich Strukturqualität aufweisen, nicht angebracht ist.

## **1.1 Hintergrund und Fragestellungen**

Das Krankheitsgeschehen in industrialisierten Ländern ist durch eine Zunahme chronischer Krankheiten geprägt, die zum großen Teil lebensstilbedingt sind. Über die Notwendigkeit einer Stärkung der Prävention besteht in Deutschland vor diesem Hintergrund Einigkeit. Obwohl die Knappheit der verfügbaren Ressourcen eine Priorisierung präventiver Maßnahmen erfordert, bleibt das Wie der Priorisierung gegenwärtig unbeantwortet.

Deutschland wurde durch die mit dem Beitragsentlastungsgesetz vom 01.11.1996 erfolgte zwischenzeitliche Herausnahme primärpräventiver Leistungen aus dem SGB V beim Aufbau präventiver Strukturen zurückgeworfen. Erst mit der erneuten Aufnahme von Satzungsleistungen zur Primärprävention in § 20 SGB V durch das Gesundheitsreformgesetz 2000 und der Ankündigung eines Präventionsgesetzes wurde der Prävention wieder vermehrte Aufmerksamkeit gewidmet. Es ist zu erwarten, dass auch ein neuer Gesetzesentwurf keinen abschließenden Katalog von

Präventionsmaßnahmen vorsehen wird, sondern vielmehr Anforderungen festlegt, welche die von den Präventionsträgern zu erbringenden Präventionsleistungen zu erfüllen haben. Vor diesem Hintergrund hatte unser Projekt folgende Fragestellungen:

- Können im Ausland durchgeführte und als effektiv bewertete Projekte der Prävention und Gesundheitsförderung auf Deutschland übertragen werden?
- Für welche Projekte der Prävention und der Gesundheitsförderung konnte in den ausgewählten Ländern ein Effekt nachgewiesen werden?
- Welche Erfahrungen liegen im Ausland mit der Messung der Effektivität und Qualität von Präventionsmaßnahmen vor?
- Lassen sich diese Präventionsmaßnahmen nach Kriterien untersuchen und kategorisieren, um Rückschlüsse auf mögliche Gründe für den Erfolg spezifischer Maßnahmen zu formulieren?

## **1.2 Können Empfehlungen für Präventionsprogramme ausgesprochen werden?**

Generelle Erfolgsfaktoren für Präventionsprojekte können aus der Literatur nicht abgeleitet werden. Jedoch können einzelne, herausragende Ergebnisse ausgewählter Studien in den vier Handlungsfeldern

- Bewegung
- Depression
- Ernährung
- Rauchen

dargestellt werden:

### **1.2.1 Bewegung**

Im Handlungsfeld Frauen und Mädchen und Bewegung existieren zwei effektive bzw. teilweise effektive Studien, die in den Bereich stark empfehlenswert gruppiert wurden. Sie waren multistraigisch ausgerichtet, schlossen eine vorherige Befragung der Zielgruppe ein und integrierten mög-

lichst alle Beteiligten wie Schule, Gemeinde etc. (Wen et al. 2002, Pate et al. 2005). Unter den Studien im Bereich eingeschränkter Empfehlung zeigte eine Studie mit einem sehr engen verhaltenspräventiven Ansatz keinen Effekt (Robbins et al. 2006). Zwei weitere Studien mit einem gemischten nachweisbaren Effekt haben ebenfalls einen sehr engen Ansatz und sind von der Anlage her eher dem Feld efficacy (Schofield et al. 2005) bzw. einem Pilotcharakter (Williams et al. 2005) zuzuordnen.

Für den Bereich Betrieb und Bewegung zeigt sich, dass alle Aktionen, die mit Plakaten etc. das Treppensteigen isoliert steigern wollen, nicht mehr durchgeführt werden sollten, da sie scheitern (Adams and White 2002, Marshall et al. 2002). Nur zwei der acht Studien zum Treppensteigen im betrieblichen Setting erreichten in unserer Evaluation das nötige Qualitätsniveau. Die einzige Studie, die nach Qualität in den Bereich empfehlenswert kam, und einen positiven Effekt hat, entspricht zwar eher einem von oben verordneten Programm (Harrell et al. 1996). Sie weist damit aber in der Tendenz auf, dass Bewegung im Betrieb, hier zwar das spezielle Klientel von Polizei-angestellten, „einfach“ durchgeführt werden muss, so dass man im Betrieb alle Mitarbeiter erreicht. Weiterhin zeigten die weiteren gescheiterten Studien Faktoren, die man als Misserfolg-faktoren beurteilen könnte: a) Arbeitsbefreiung wird nicht gewährt (Gerdle et al. 1992). b) Die Intervention ist isoliert auf Beratung oder Fitnessstestung als Anreiz ausgelegt (Aittasalo et al. 2004).

## **1.2.2 Depression**

Die zur Prävention der Depression analysierten Studien konzentrieren sich auf wenige Programme. Am häufigsten wurden die Programme Leichtigkeit im sozialen Alltag (LISA), das Programm FRIENDS sowie Resourceful Adolescent Programm (RAP) evaluiert, die auch alle unter den Studien mit überdurchschnittlichen Qualitätsscores vertreten sind. Für alle drei Programme ist mindestens eine der Studien mit überdurchschnittlichem Qualitätsscore teilweise effektiv. Besonders hervorzuheben sind die beiden mit überdurchschnittlichen Qualitätsscores durchgeführten Studien zu dem Programm RAP, die beide als teilweise effektiv bewertet worden sind. Weitere Studien für dieses Programm mit unterdurchschnittlicher Studienqualität liegen nicht vor (Merry et al. 2004; Shochet et al. 2001). Die Studie von Merry et al. 2004 zu dem Programm RAP erreicht als einzige Studie den Qualitätslevel stark empfehlenswert.

Für das Programm LISA (Pössel et al. 2004; Groen et al. 2003; Pössel et al. 2006a) und für das Programm FRIENDS (Lowry-Webster et al. 2003) liegen auch Studien mit unterdurchschnittlicher Studienqualität vor. Die Studien für das Programm FRIENDS sind aber alle zumindest teilweise effektiv und weisen damit eine einheitliche Tendenz auf. Für das Programm LISA liegen dagegen sowohl ineffektive als auch teilweise effektive Studien vor.

Es gibt Anhaltspunkte, dass die Ausgestaltung der Interventionen über die in den Qualitätsscores erfassten Aspekte hinaus großen Einfluss auf die Ergebnisse haben (Pössel et al. 2006 b). Nach unserer Auffassung ist im Bereich der Prävention von Depressionen und Angststörungen noch weiterer Forschungsbedarf gegeben. Trotz der positiven Tendenzen der kognitiv-verhaltenstherapeutischen Ansätze sind gegenwärtig keine klaren Empfehlungen möglich. Sie scheinen jedoch dem rein edukativen Ansatz überlegen, für den die zwei analysierten Studien – weder die mit unterdurchschnittlichen noch die mit überdurchschnittlichen Qualitätsscores – keine Wirksamkeit zeigen konnten. (Clark et al. 1993) Eine flächendeckende Umsetzung kognitiv verhaltenstherapeutischer Interventionen zur universalen Prävention der Depression erscheint gegenwärtig aufgrund der offenen Fragen zur optimalen Ausgestaltung verfrüht.

### **1.2.3 Ernährung**

Für die 24 evaluierten Studien über alle Qualitätsstufen hinweg ergab sich, dass bei den acht Interventionen, die kombiniert Ernährung und Bewegung in den Blick nahmen, nur der Ernährungsbereich einen positiven Effekt hatte, der Bewegungsbereich hingegen einen negativen. Alle drei Interventionen, die in den Bereich empfehlenswert fielen, hatten entweder einen Ansatz, der sowohl Ernährung und Bewegung einbezog (Gortmaker et al. 1999a) oder einen Ansatz, in dem die Eltern eng eingebunden waren (Reynolds et al. 2000, Townsend et al. 2006).

### **1.2.4 Rauchen**

Die internationale Studienlage zur Prävention des Rauchens bei Schwangeren erlaubt keine eindeutigen Empfehlungen. In der Tendenz sind effektive Interventionen mit hoher Qualität in der internationalen Studienlage multimodal angelegt, jedoch sind auch einige multimodale Programme gescheitert. Erfolgreiche Programme wurden eher durch Hebammen, Ärzte oder Pflegepersonal während der routinemäßigen Schwangerschaftsuntersuchungen durchgeführt.

### 1.3 Empfehlungen

Unsere Studie hat gezeigt, dass die Situation zum Nachweis von Effektivität in Prävention und Gesundheitsförderung auch im internationalen Umfeld in etwa den Stand der kurativen Medizin von vor 25 Jahren aufweist. Belastbare und ausreichend erprobte Kataloge zur Bewertung und anschließenden Priorisierung von Präventionsangeboten fehlen weitgehend. Vorhandene Kataloge haben eine spezifische Zielrichtung (etwa Beratung der Projektdurchführenden).

In Deutschland wurden in den letzten Jahren Gesundheitsziele erarbeitet.<sup>1</sup> (GVG 2003) Aufgrund des erwartbaren und notwendigen Ressourcenvolumens für Prävention und Gesundheitsförderung ist nun zu entscheiden, welchen Beitrag Prävention und Gesundheitsförderung dazu leisten können. Nach unserer Auffassung müssen die Erfahrungen der Evidenz-basierten Medizin aufgegriffen werden, um rationale Entscheidungen zu treffen.

Um die Rationalität aufbauen zu können muss der Gesetzgeber eine verbindliche Evaluation der geförderten Interventionen, gemessen an den Zielen, vorsehen. Da die derzeit nicht ausreichende empirische Studienlage dazu führt, dass absehbar auch nicht wirksame Interventionen durchgeführt werden, sollten darüber hinaus alle Interventionen zunächst zeitlich befristet werden.

Auf Grund der Heterogenität der Erhebungsmethoden und der Effektparameter sollte der Gesetzgeber dafür Sorge tragen, dass bundesweit einheitliche Evaluationsstandards erarbeitet werden. Nach dem Ablauf einer genügend großen Zahl von Interventionen und nachfolgender Evaluation können dann Empfehlungen ausgearbeitet werden, welche Interventionstypen Erfolg versprechend sind. Der Gesetzgeber sollte daher auch vorsehen, dass in einem angemessenen Zeitraum eine umfassende Zusammenführung der durchgeführten Interventionen beziehungsweise ihrer Evaluationen erfolgt.

Die bisherigen Entwürfe zum Präventionsgesetz schrieben für Maßnahmen der Verhaltensprävention Effektivitätsnachweise, für Maßnahmen der Verhältnisprävention ein angemessenes Qualitätsmanagement vor. Es hat sich gezeigt, dass international zwischen Qualität des Studiendesigns, der Intervention und dem Effekt keine Korrelation besteht. Dies ist wissenschaftstheoretisch zunächst nicht erstaunlich. Es führt aber in der Konsequenz dazu, dass eine Beschränkung auf eine reine Einhaltung vorgegebener Qualitätskriterien, wie es der Gesetzgeber im

---

<sup>1</sup> Die sechs Gesundheitsziele auf Bundesebene sind: „Diabetes mellitus Typ2: Erkrankungsrisiko senken, Erkrankte früh erkennen und behandeln“, „Brustkrebs: Mortalität vermindern, Lebensqualität erhöhen“, „Tabakkonsum reduzieren“, „Gesund aufwachsen“, „Patient(inn)ensouveränität stärken“, „Depressive Erkrankungen: verhindern, früh erkennen, nachhaltig behandeln“.

Präventionsgesetz für Interventionen der Verhältnisprävention vorgesehen hatte, nicht ausreichend ist. Mit anderen Worten reicht es beispielsweise nicht aus, eine Partizipation der Betroffenen vorzusehen oder eine Schulung von Multiplikatoren. Trotz der Einhaltung dieser Qualitätskriterien kann eine Intervention gleichermaßen am Ende einen Effekt aufweisen oder wirkungslos bleiben.

## 1.4 Methode des Projekts

Der empirische Teil des Projektes gliedert sich in die Teile systematische Literaturrecherche sowie Bewertung der gefundenen Studien und Entwicklung eines Kriterienkatalogs zur Evaluation von Interventionen.

Für die Bewertung der Studien bzw. Interventionen wurde ein Kriterienkatalog entwickelt, der auf einer kritischen Durchsicht und Zusammenschau der derzeit international gebräuchlichen Qualitätsinstrumente beruht. Dazu wurden unter anderen die Instrumente PREFFI (PREvention EFFect-management-Instrument), Closing the Gap, IDM (Interactive Domain Model), QSP (Qualitätssicherung Prävention), Quintessenz etc. herangezogen. Die Erstellung des Kriterienkatalogs erfolgte in einem begründeten Verfahren nach dokumentierten Kriterien. Anschließend wurden Korridore für Empfehlungsstärken festgelegt, wie sie ähnlich in der Evidenz-basierten Medizin und auch der Public-Health-Forschung in Konsensusprozessen entwickelt werden.

Der Kriterienkatalog bewertet Programme hinsichtlich ihrer Qualität und Effektivität. Er besteht insgesamt aus 68 Fragen. Die im systematischen Suchverfahren gefundenen Studien wurden im Gutachten hinsichtlich Qualität des Studiendesigns und Qualität der Intervention evaluiert. Unter **Qualität des Studiendesigns** werden Gütekriterien aus der evidenzbasierten Medizin herangezogen wie statistische Aussagekraft, Vergleichbarkeit von Interventions- und Kontrollgruppe etc. Diese Gütekriterien würden für eine Implementierung wegfallen, wenn für eine Intervention nach einer Studienphase die Wirksamkeit belegt worden ist. Unter **Qualität der Intervention** werden Gütekriterien der Best practice sowie des Qualitätsmanagements zusammen geführt, die die inhaltliche Gestaltung der Intervention bzw. eines Programms bewerten können. Für jedes Qualitätskriterium wurden nach einem festgelegten Verfahren zwischen null und einem Punkt vergeben. Die maximale Punktzahl im Score war abhängig vom Studiendesign. Die unterschied-

lich hohen erreichbaren Punktzahlen ergeben sich, da einige Kriterien nicht anwendbar waren, etwa die Güte der Randomisierung.

Gegenstand des Projektes waren Interventionen der Prävention und Gesundheitsförderung für die Risikofaktoren Bewegung, Depression, Ernährung und Rauchen. In die Untersuchung wurden präventive und gesundheitsfördernde Interventionen aus den USA, Kanada, Australien, Neuseeland, Großbritannien, Frankreich, Niederlanden, den skandinavischen Ländern, Österreich und der Schweiz einbezogen. Die Recherche wurde auf Publikationen der Jahre 1990 bis 2006 beschränkt. Recherchiert wurde in den Literaturdatenbanken Medline, Amed, Biosis, Embase, Samed, Cinahl, Goyal Health und Psychinfo. Eingeschlossen wurden Publikationen, die nicht nur eine Intervention, sondern auch deren Evaluation enthielten. Ausgeschlossen wurden Studien, welche ausschließlich auf die Multiplikatorebene, also z. B. Ärzte/innen ausgerichtet waren. D.h., die Maßnahme sollte in ihrer Wirkung auf die Betroffenen evaluiert sein, deren Gesundheit durch die präventiven Maßnahmen beeinflusst werden sollte. Ein weiteres Ausschlusskriterium war, dass die Evaluation nicht nur Veränderungen im Wissen oder in den Einstellungen gemessen hatte, sondern auch Veränderungen im Verhalten oder klinischen beziehungsweise physischen Parametern.

Die Zahl der in der Literaturrecherche erzielten Treffer zu den verschiedenen Handlungsfeldern lag zwischen ca. 12.400 (Rauchen) und ca. 5.000 Studien (Depressionen). Nach Berücksichtigung der Ein- und Ausschlusskriterien verblieben immer noch rund 1.500 bzw. 500 Studien für die genannten Handlungsfelder.

Auf Grund dieser großen Menge an Studien zu Interventionen innerhalb der Felder Bewegung, Depression, Ernährung und Rauchen mussten Fokusbereiche ausgewählt werden. Es wurden daher im Bereich Bewegung das Setting Betrieb ausgewählt (26 Interventionen) sowie Intervention bei Mädchen und Frauen (14 Interventionen). Im Handlungsfeld Depression wurden 18 Studien der universellen Prävention der Depression in Schulen bewertet. Desweiteren wurden 24 Studien zur Ernährung in der Zielgruppe Kinder und Jugendliche im Setting Schule sowie 34 Interventionen zur Raucherentwöhnung bei Schwangeren analysiert.

Die Verbindung der Ergebnisse der Qualitätsevaluation mit den Ergebnissen der erzielten Effekte der Intervention machte es möglich, Empfehlungen auszusprechen. Dabei wurden Interventionen, die mehr als 80% der Punkte in den beiden Dimensionen Qualität des Studiendesigns und

Qualität der Intervention erzielen, als „in hohem Maße empfehlenswert“ eingestuft. Das heißt wenn die Intervention einen Effekt hatte, liegt eine hohe Wahrscheinlichkeit vor, dass dieser so wieder eintritt und umgekehrt, wenn diese Intervention keinen Effekt hat, liegt eine hohe Wahrscheinlichkeit vor, dass er bei erneuter Durchführung ausbleibt. Interventionen deren Qualität einen Wert von 60% erreichte, wurden als „stark empfehlenswert“ bezeichnet. Interventionen die noch über dem Durchschnittswert von der Qualität des Studiendesigns und der Qualität der Intervention lagen, konnten nur „eingeschränkt empfohlen“ werden. Alle darunterliegenden Interventionen wurden „nicht weiter in die Empfehlung“ aufgenommen. Bei dieser Einstufung ist zu berücksichtigen, dass für eine Empfehlung auch ein positiver Effekt nachgewiesen sein muss. Bei einer nicht-effektiven Intervention, die mit hohen Qualitätsscores durchgeführt wurde, ist die Wahrscheinlichkeit hoch, dass sie in eine Negativliste aufgenommen werden muss. Nicht-effektive Studien mit geringeren Qualitätsscores können dagegen bei einer Wiederholung der Studie mit besserer Qualität gegebenenfalls zu anderen (besseren) Ergebnissen führen. Daher bedeutet eine fehlende Empfehlung nicht, dass die Intervention generell zu verwerfen ist. Es bedeutet lediglich, dass die Qualität der Studiendurchführung keine Verlässlichkeit für eine Empfehlung bietet.

## 2 Einleitung: Ziel und Gegenstand des vorliegenden Forschungsprojekt

Im Jahr 2005 verabschiedete der Bundestag ein „Präventionsgesetz“, BT-Drucksache 15/4833. Nach dem Scheitern des Beschlusses im Bundesrat wurde, nicht zuletzt aufgrund der zwischenzeitlich beschlossenen Neuwahlen für September 2005, eine weitere Initiative für ein Präventionsgesetz zunächst verschoben. Die nach den Wahlen entstandene Koalition aus CDU/CSU und SPD verständigte sich im Koalitionsvertrag auf eine Neufassung blieb diese bisher (August 2007) allerdings schuldig. Dennoch erscheint es weiterhin dringlich, Kriterien für eine Priorisierung zukünftiger Interventionen in Prävention und Gesundheitsförderung zu erarbeiten.

Vor dem Hintergrund der Diskussion um das Präventionsgesetz hat das Institut für Gesundheitsökonomie und Klinische Epidemiologie (IGKE) der Universität zu Köln das Forschungsprojekt „Internationaler Erfolge in Prävention und Gesundheitsförderung und ihre Übertragbarkeit auf Deutschland“ im Auftrag der Hans Böckler Stiftung durchgeführt. Im Entwurf des Gesetzes – wenn auch zum damaligen Zeitpunkt gerade an den parlamentarischen Hürden gescheitert - war keine Entscheidung über das Wo und Wie von Prävention vorgegeben. Vielmehr beinhaltete der Entwurf ein Drei-Ebenen-Modell, in dem die konkreten Ansatzpunkte und Maßnahmen der Prävention erst konkretisiert werden sollten: Auf Bundesebene sollten Präventionsziele vereinbart werden, auf Landesebene sollten Länder und Sozialversicherungsträger in gemeinsamen Entscheidungsgremien Bedarf und Organisation von Präventionsmaßnahmen konkretisieren. Auf der Ebene der Sozialversicherungsträger sollten insbesondere Maßnahmen der Verhaltensprävention und der Betrieblichen Gesundheitsförderung umgesetzt werden.

Für Maßnahmen der Verhaltensprävention schrieb der Gesetzentwurf Effektivitätsnachweise vor, für Maßnahmen der Verhältnisprävention war ein angemessenes Qualitätsmanagement vorgesehen. In Stellungnahmen zum Gesetzentwurf wurde angemahnt, zur Priorisierung von Interventionen Erfahrungen und Erkenntnisse des Auslandes heranzuziehen, um dann bei Start des Präventionsgesetzes über eine entsprechende Informationsbasis zu verfügen. Diese Anregung haben wir aufgegriffen.

Daraus leitet sich die Fragestellung ab: „Gibt es in anderen Ländern Wirksamkeitsnachweise zu präventiven bzw. gesundheitsfördernden Interventionen, die auf Deutschland übertragbar sind?“ Ziel war dabei nicht herauszufinden, ob eine generelle Wirksamkeit vorliegt, wie etwa der Zusammenhang, dass regelmäßige Bewegung das Herzinfarktrisiko mindern kann, sondern die Frage, mit welchen Interventionen man eine solche regelmäßig durchzuführende Bewegung bei welchen Zielgruppen wie fördern kann.

In der Literatur zu Prävention und Gesundheitsförderung wird teilweise betont, dass die evidenzbasierte Medizin allein nur eingeschränkt Anwendung finden kann, um die Empfehlung für oder gegen Interventionen aussprechen zu können.

Es wird ferner vertreten, dass die Qualität von Präventionsprojekten die Ergebnisse maßgeblich beeinflusst. Der vorliegenden Analyse liegt die Auffassung zugrunde, dass sowohl die Wirksamkeit als auch die Qualität berücksichtigt werden müssen, um eine Empfehlung für oder gegen eine Intervention abgegeben zu können. Dabei ist innerhalb des Kriteriums zu trennen zwischen der Qualität der Studiendurchführung („formale Validität“) und der Qualität der Intervention („inhaltliche Validität“).

In einer systematischen Literaturrecherche wurden Studien anhand bestimmter Ein- und Ausschlusskriterien identifiziert. Da zu der Bewertung der Interventionen und der darauf aufbauenden Evaluationsstudien ein Bewertungsmaßstab erforderlich war, wurde unter Verwendung bereits bestehender Qualitätsinstrumente ein Kriterienkatalog entwickelt, mit dem die gefundenen Interventionen und Studien bewertet wurden. Der Kriterienkatalog basiert auf den Prinzipien der evidenzbasierten Medizin (EbM), berücksichtigt aber die Besonderheiten der Prävention und Gesundheitsförderung. Der Kriterienkatalog sollte geeignet sein, effektive Präventionsprogramme zu identifizieren und somit eine Rangfolge für vorhandene oder geplante Interventionen in Deutschland aufzustellen.

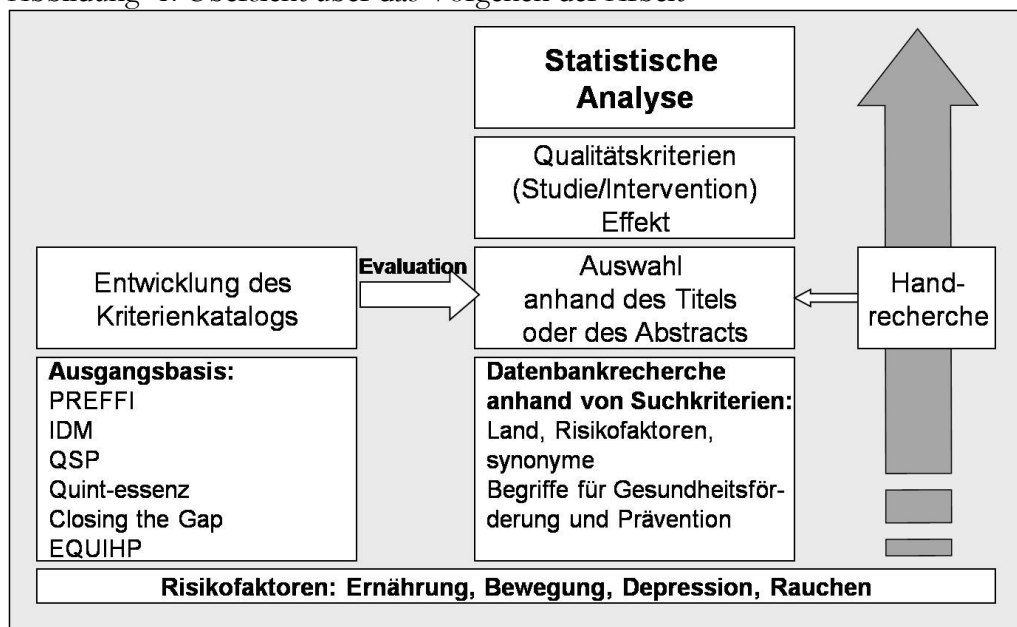
Bezüglich der berücksichtigten Qualitätsstandards orientiert sich der Kriterienkatalog an den Qualitätskriterien, die in der einschlägigen Literatur für erfolgreiche Interventionen anerkannt sind. In Hinblick auf die Studienqualität orientiert sich das Analyseinstrument an den Kriterien der EbM bzw. der Biostatistik. Anders als im klassischen Verständnis der EbM sehen wir das gewählte Studiendesign nicht als Ausschlusskriterium. D.h., auch Interventionen, welche z.B. nicht als randomisierte kontrollierte Studien (RCT) evaluiert wurden, können zu einer positiven Empfehlung der evaluierten Intervention führen. Denn auch andere als randomisierte kontrollier-

te prospektiv geplante Studiendesigns können sich je nach Charakter der Intervention, z. B. eine landesweite Kampagne, als maximal erreichbarer Studienstandard aus der Fragestellung ergeben. Je nach Studiendesign ist jedoch eine unterschiedliche Zahl an Qualitätskriterien zu berücksichtigen. So ergibt es sich, dass eine nachvollziehbare Randomisierung als Qualitätskriterium nur auf die sog. randomisiert kontrollierten Studien anwendbar ist.

Zusätzlich zu den beiden Dimensionen Qualität der Studiendurchführung und Qualität der Intervention ist der Effekt zu beachten. Eine uneingeschränkt empfehlenswerte Intervention wäre somit gekennzeichnet, durch eine hohe Qualität der Intervention (beispielsweise wurde eine Partizipation vorgesehen), eine hohe Qualität einer evaluierenden Studie (beispielsweise wurde randomisiert eine Kontrollgruppe vorgesehen, welche die Intervention nicht erlebte) und schließlich der Nachweis eines Effektes relativ zur Kontrollgruppe.

Da bekannt ist, dass Rauchen, Fehlernährung und Bewegungsmangel die Entstehung von Krankheiten begünstigen und Depression zugleich Krankheit und auch Risikofaktor für weitere Krankheiten sind, wurden diese Bereiche ausgewählt, um empfehlenswerte Interventionen der Prävention und Gesundheitsförderung zu benennen.

Abbildung 1: Übersicht über das Vorgehen der Arbeit



Ein Ziel des vorliegenden Projektes war es, Interventionen in Gesundheitsförderung und Prävention aus ausgewählten Ländern zu sichten und zu bewerten im Hinblick auf die Frage, ob diese Interventionen in Deutschland übernommen werden sollten. Gegenstand war damit in einem ersten Schritt die Entwicklung eines Bewertungsinstruments, das Interventionen unterschiedlicher Qualität voneinander trennen sollte.

## 3 Evaluation von Interventionen in Prävention und Gesundheitsförderung

In diesem Kapitel wird ein Aufriss der Forschungslage zur Evaluation von Interventionen in Prävention und Gesundheitsförderung gegeben. Dabei dienen als Leitfaden folgende zwei Fragen, die Gegenstand wissenschaftlicher Diskussionen sind:

- Wie sind Prävention und Gesundheitsförderung einander zuzuordnen?
- Können Interventionen der Prävention und der Gesundheitsförderung nach Kriterien der Evidenzbasierten Medizin ausgewertet werden?

Da beide Problemstellungen eng miteinander zusammen hängen, wird zunächst geklärt, wie wir Prävention und Gesundheitsförderung einander zuordnen. Eine Intervention in den Bereichen Prävention und Gesundheitsförderung definieren wir zunächst wie folgt: „any policy of public service practice that may affect other people’s lives“ (Oliver et al. 2005). Diese allgemeine Umschreibung lässt es zu, dass ein weites Spektrum von Eingriffen in das Verhalten und/ oder die Verhältnisse auf individueller, lokaler, regionaler und nationaler Ebene als Intervention aufgefasst und in der vorliegenden Untersuchung nebeneinander gestellt werden kann.

### 3.1 Prävention und Gesundheitsförderung

Es können vier Modelle unterschieden werden, nach denen das Verhältnis von Prävention und Gesundheitsförderung zueinander bestimmt werden kann:

#### 3.1.1 Modell der Substitution

Klassisch wird Gesundheitsförderung (Health Promotion) definiert als das Zusammenführen von der „Stärkung [...] persönlicher und sozialer Gesundheitskompetenz“ (Empowerment, Schutzfaktoren) „verbunden mit einer systematischen Politik, die auf [...] den Abbau von gesundheitlicher Ungleichheit abzielt“ (Intersektoralität) (Kickbusch 2003, S. 182). Sie wird damit der Prävention gegenübergestellt, die an den Kategorien Risiko(faktoren), Krankheit und Individuum anknüpft (s. auch Hurrelmann 2004 et al., S. 14). In Anlehnung an Syme (1986) und Antonovsky (1987)

und die Ottawa-Charta (1986) (Prozesshaftigkeit) wird ein „Defizitmodell“ der Prävention durch ein „Unterstützungsmodell“ der Gesundheitsförderung abgelöst. Auf der Basis der dem jeweiligen Ansatz inhärenten „Logik“ ergeben sich grundsätzlich unterschiedliche Interventionsformen mit verschiedenartigen Wirkungsprinzipien (Hurrelmann 2000).

### **3.1.2 Modell der Integration**

Ein anderes Modell der Zuordnung von Gesundheitsförderung und Prävention wird folgendermaßen beschrieben: “Health promotion comprises efforts to enhance positive health and prevent ill-health among individuals and the communities, through education, prevention and protection measures.” (zitiert bei Wong 2002 in Anlehnung an Tannahill). Prävention wird als integrativer Bestandteil eines übergreifenden Konzepts von Gesundheitsförderung verstanden.

### **3.1.3 Modell der Komplementarität**

Im Modell der Komplementarität werden Gesundheitsförderung und Prävention als sich ergänzend angesehen. „Sowohl der krankheitsorientierte Ansatz der Prävention als auch der ressourcenorientierte Ansatz der Gesundheitsförderung zielen – [...] – auf die verbesserte Gesundheit des einzelnen sowie der Bevölkerung und sollten als einander ergänzend betrachtet werden.“ (Walter und Schwartz 2003; auch Hurrelmann et al. 2004). Maßnahmen können demnach Elemente aus beiden Ansätzen umfassen.

### **3.1.4 Modell der Utopie**

Gesundheitsförderung kann in ihrer Reingestalt wie die Definition der Gesundheit der WHO als „Utopie“ verstanden werden. Damit soll nicht gesagt sein, dass der Begriff auf Grund dessen, dass er unmöglich umsetzbar ist, belang- oder bedeutungslos wird. Der Begriff wird im Blochschen Sinne als „konkrete Utopie“ verstanden, also nicht als alltagssprachliche Abwertung im Sinne des Unrealisierbaren, sondern als tastender und experimentierender „Prozess der Verwirklichung“ (Bloch 1985). Gesundheitsförderung ist somit als nicht-erreichbares Ziel (ou-topos) zu denken, dient aber zugleich als Messlatte für konkrete Interventionen.

Wir schließen uns einer Auffassung von Prävention und Gesundheitsförderung an, nach der die beiden einander komplementär zu sehen sind und dabei auch eine Schnittmenge zulassen. Prä-

vention und Gesundheitsförderung sind damit nicht zwei Vorgehensweisen, die sich gegeneinander ausschließen und eine per se unterschiedliche Form der Evaluation verlangen, sondern sie können mit denselben Instrumenten untersucht werden. Der Übergang von einer klar abgrenzbaren Präventionsleistung wie einer Impfung oder einem Nikotinpflaster zur Raucherentwöhnung zu einer komplexen gesundheitsfördernden Intervention wie einer Veränderung der Strukturen in einem Stadtteil ist als Kontinuum zu sehen. Eine eindeutige Grenze ist nicht zu ziehen, allenfalls können – wie durch die Beispiele belegt – Pole eines kontinuierlichen Übergangs definiert werden. Des Weiteren haben Interventionen aus Prävention und Gesundheitsförderung sowohl Charakteristika und Elemente in ihrer Durchführung, die sie mit anderen Interventionen aus Therapie gemein haben, z. B. Partizipation der Zielgruppe, als auch Charakteristika und Elemente, die sie von Interventionen aus der Therapie trennen. Unabhängig davon, welchem Modell der Verhältnisbestimmung von Gesundheitsförderung<sup>2</sup> und Prävention man den Vorzug erteilt, bleibt das Problem bestehen, dass zwischen Interventionen priorisiert werden muss. Die Notwendigkeit, Effekte zu messen, stellt sich in doppelter Weise: 1. Um zu einer Auswahl aus einem Bündel von Interventionen zu kommen, muss man ihre (potenzielle) Wirksamkeit berücksichtigen. 2. Führt man dann eine Intervention durch, muss wiederum deren Erfolg durch Effekte in der konkreten Anwendung belegt werden. Dabei darf die Gegenüberstellung von Prävention, die am Individuum ansetzend als Risikoreduktion oder vermiedene Krankheitslast messbar ist, und Gesundheitsförderung, deren Erfolg mangels Operationalisierbarkeit (welche Indikatoren und wie sind diese zu erheben?) schwerer zu erheben ist, nicht dazu führen, dass eine Hierarchisierung von Interventionen generell abgelehnt wird. Denn die Priorisierung ist für die Allokation verfügbarer Ressourcen unabdingbar.

## 3.2 Evidenzbasierte Medizin

Im Folgenden wird Evidenzbasierte Medizin näher erläutert. Insbesondere die Frage, ob Kriterien der sogenannten EbM in die Evaluation von Interventionen aus Prävention und Gesundheitsförderung aufgenommen werden können, wird kontrovers diskutiert (z. B. Elkeles 2006 versus Bödeker 2006). Wir meinen, dass sich ein Großteil der Differenzen auf ein oft vertretenes

---

<sup>2</sup> Selbst wenn man die Position von Gadamer teilt, dass man Gesundheit nicht messen kann, da sie „ein Zustand der inneren Angemessenheit und der Übereinstimmung mit sich selbst ist, die man nicht durch eine andere Kontrolle überbieten kann“ (1993, 138f), so gibt es doch sogenannte Statthalter, Gesundheitsindikatoren, die messbar sind (vgl. z. B. Beck 2003).

falsches Verständnis der EbM sowohl auf Seiten der Gegner als auch der Befürworter von EbM zurückführen lässt. Wir entwerfen ein Bild der EbM, das EbM nicht einfach mit der randomisierten kontrollierten Studie (= RCT, randomized controlled trial) gleich setzt. Damit halten wir uns auch an die Definition der Begründer der EbM selbst (s. dazu auch die ausführliche Darstellung in Kreis (2006), S. 160ff).<sup>3</sup>

### **3.2.1 Was ist Evidenzbasierte Medizin?**

Zunächst ist die Evidenzbasierte Medizin (EbM) eine Art und Weise, Medizin zu betreiben, die drei Bereiche medizinischen Handelns, nämlich (1) Wissensbestand außerhalb des einzelnen Anwenders, (2) Erfahrungswissen des einzelnen Anwenders und (3) Vorstellungen und Wünsche der Patienten gleichwertig nebeneinander stellt.

Im engeren Sinne ist EbM definiert als die Anwendung des extern besten verfügbaren Wissens von einem Anwender auf Basis dessen Erfahrungswissen und in Absprache mit der Person, an der dieses Wissen zu Anwendung gebracht werden soll. Dabei werden (wissenschaftlich gewonnene) Ergebnisse in Hinsicht auf ihre Anwendbarkeit an einzelnen Patienten oder Kollektiven bewertet. Es handelt sich um einen prozesshaften Vorgang, „systematically finding, appraising and using contemporaneous research findings ...“ (Rosenberg und Donald 1995), der zunächst vornehmlich auf klinische Fragestellungen angewandt wurde. Allerdings wurde eines der ersten Klassifikationsschemata dezidiert für den Bereich Prävention entwickelt (Canadian Task Force on the Periodic Health Examination 1979).

Dass die Überlegungen, wie Forschung bewertet und hierarchisiert werden kann, bisher kein endgültig und allseits akzeptiertes System hervorgebracht haben, ist auch Vertretern der EbM durchaus bewusst. Das belegen Publikationen wie die von Glasziou et al. (2004) sowie die kontinuierliche Überarbeitung entsprechender Leitlinien (SIGN etc.).

### **3.2.2 Wie kann Wissen zusammen geführt werden? Das systematische Review und neue Entwicklungen**

In einem mittlerweile als klassisches Verfahren der Evidenzbasierten Medizin (EbM) etablierten systematischen Review werden die efficacy (generelle Wirksamkeit) bzw. effectiveness (kontext-

---

<sup>3</sup> Auch Bödeker (2006), S. 5: Schon Sackett und Rosenberg heben hervor, dass nicht nur RCTs in systematische Reviews eingehen müssen. S. 6: Auch die Cochrane Collaboration erlaubt ihren Fachgruppen die Integration anderer Studiendesigns.

vermittelte bzw. alltägliche Wirksamkeit) verschiedener Studien als ein Ergebnis zusammen geführt. Ein systematisches Review ist eine systematische, qualitätsbewertete zusammenfassende Darstellung von Studienergebnissen (s. Timmer und Antes (2006), S. 17; Guidelines z. B. bei Jackson et al. 2005), die narrativ ist bzw. qualitativ, wenn sie nicht zu einer Metaanalyse zusammengefasst werden können bzw. werden sollen. Ein systematisches Review versucht vom Ansatz her, eine „Quersumme“ über mehrere Studien zu bilden. Das Ergebnis besteht darin, dass alle Studien zu einer Aussage zusammen gezogen werden:

Die entsprechende(n) Intervention(en) zeigen insgesamt k/einen Effekt. Auch wenn die Studienqualität (z. B. Egger et al. 2003) herangezogen wird, bleibt die Qualität der Intervention zumeist unberücksichtigt. Fragen wie z.B. warum Studien (Interventionen) einen positiven oder negativen Effekt haben, bleiben damit unbeantwortet (siehe Abbildung 2).

Abbildung 2: Einteilung von Studien in einem systematischen Review

		EFFEKT	
		positiv	negativ
Q U A L I T Ä T	positiv		
	negativ	Kommt nicht In das Review	Kommt nicht In das Review

Das Systematische Review ist für die vorliegende Untersuchung nicht ausreichend, da die Fragestellung lautet: Wenn es Studien guter Qualität gibt, die in effektive und nicht-effektive getrennt werden können, sollte man dann nicht untersuchen, ob und worin sich diese unterscheiden?

Ein weiteres Problem ist, dass die systematischen Reviews im Bereich Prävention und Gesundheitsförderung häufig sehr global fragen. Ein Beispiel ist die Frage: Helfen umweltverändernde Interventionen am Arbeitsplatz?

Somit verfolgt das vorliegende Forschungsvorhaben kein klassisches systematisches Review vorzulegen. Es versucht auf Basis einer systematischen Suche und Bewertung von Interventionen einer Positivliste von „hinlänglich wirksamen Interventionen“ (Trojan 2006a) wie auch einer entsprechenden Negativliste von Interventionen näher zu kommen, für die man mit hoher Gewissheit das Ausbleiben einer Wirkung annehmen kann.

Wir entwickeln damit eine Synthese aus dem Vorgehen der Best-Practice-Forschung (Gute/ Beste Praxis) und der EbM. Wie oben dargestellt, berücksichtigt das systematische Review fast nie die Qualität der Intervention.

Diese aber ist Gegenstand der sog. Best-Practice-Evaluationen. Deren Philosophie ist es, Interventionen mit positivem Effekt auf Kriterien wie Partizipation der Zielgruppe, Prozessdokumentation u. a. zu „screenen“ und dann Interventionen als „Leuchttürme“ vorzustellen, die diese Kriterien (zum Teil) erfüllen (Abbildung 3).

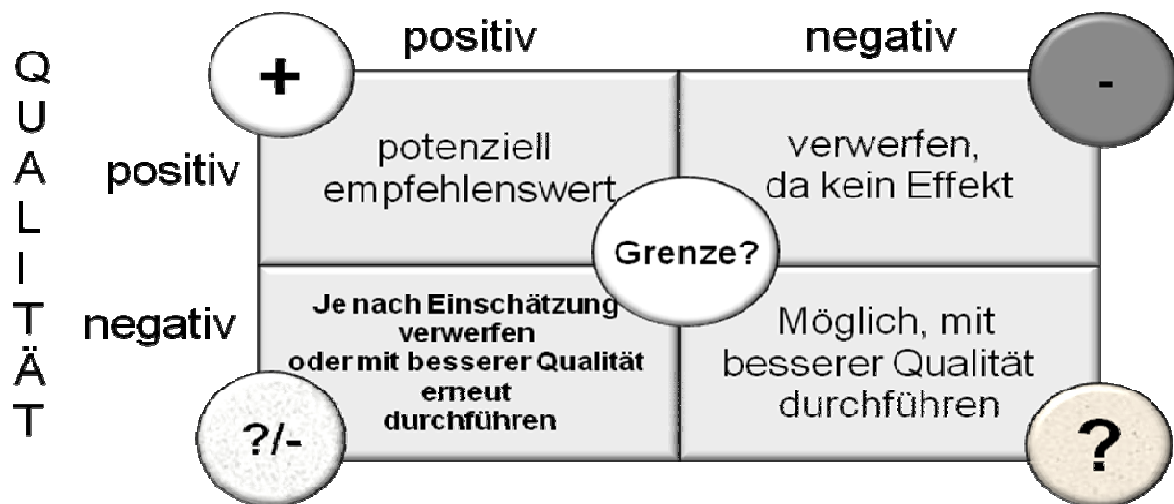
Abbildung 3: Einteilung von Studien nach dem Vorgehen von Best Practice

		EFFEKT	
		positiv	negativ
Q U A L I T Ä T	positiv		Kommt nicht in die Liste
	negativ	Kommt nicht in die Liste	Kommt nicht in die Liste

Das Problem ist dabei allerdings, dass damit keine systematische und umfassende Suche durchgeführt wird und des Weiteren die Kriterien des Studiendesigns keine Berücksichtigung finden. „Filtering out studies for exclusion without examining them in detail may deprive both reviewers and users of important evidence and insights.“ (Ogilvie et al. 2005, S. 890).

Um somit die Nachteile der beiden Vorgehensweise möglichst zu vermeiden, aber auch die Vorteile voll zu integrieren, haben wir diese beiden Zugänge zusammengeführt. Dies nennen wir „ordinale Metaanalyse“, da Effekt und die beiden Qualitätsdimensionen mathematisch einbezogen werden, aber kein Gesamteffekt über Interventionen hinweg errechnet wird. Qualität der Studiendurchführung wie der Intervention werden gleichermaßen berücksichtigt (s. Abbildung 4). Es werden die Qualität der Studie und die Qualität der Intervention als zwei separate Dimensionen bewertet.

Abbildung 4: Einteilung von Interventionen in der „ordinalen Metaanalyse“



### 3.3 Evidence-based Health Promotion als eigenständige Methode

In dem Bewusstsein, dass „Public Health“ ein corpus mixtum aus Wissenschaft, Praxis und intuitiver Anwendung ist und damit ein weites Spektrum von Zugängen eröffnet, kann evidenzbasierte Evaluation nicht einfach die Anwendung der bisherigen Verfahren sein, sondern findet in dem Wissen um kulturelle Besonderheit und soziale Differenzierung statt (vgl. McQueen 2002). Zu effektiver Praxis gehört ebenso wie die Expertise des Ausführenden und die Bedürfnisse der Patienten (Sackett et al 1996; so auch bei Tang et al 2003) der Konsensusprozess mit Beteiligten aus der Bevölkerung, der Gesundheitspolitik und der Wissenschaft.

„Few topics in the field of health promotion have engendered as much heated debate as that of evidence. The spectrum of opinion is broad and diverse” (McQueen 2001). Das Spektrum, ob eine Evaluation nach Kriterien der EbM auf Interventionen der Prävention oder der Gesundheitsförderung Anwendung finden soll, reicht von totaler Ablehnung bis zu Verteidigung in positivistischer Manier. Anhänger einer vermittelnden Position versuchen, „the wider role of evidence in terms of knowledge development in the field“ zu entwickeln (McQueen 2001). So findet sich auch Evidence-based Health Promotion (EbHP) als eigenständige Methode, Interventionen präventiver wie gesundheitsfördernder Natur zu bewerten mit dem Ziel, begrenzte Mittel auf möglichst effektive und ggf. effiziente Interventionen zu allozieren (s. auch IUHPE 1999).

### **3.3.1 Therapie versus Prävention und Gesundheitsförderung:**

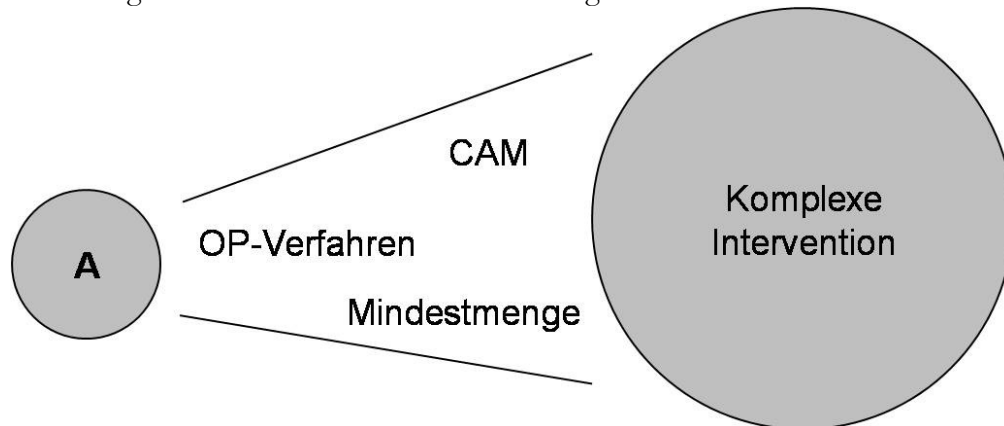
#### **Diastase oder Kontinuum**

Es stellt sich die Frage, ob Interventionen der Prävention bzw. der Gesundheitsförderung sich so von anderen Interventionen unterscheiden, dass sie grundsätzlich einer Bewertung auf Basis eines an der EbM orientierten Verfahrens nicht zugänglich wären. Dazu sei auf die folgende Fragestellung verwiesen: Es kann von großer Bedeutung sein zu wissen, ob bei einer Erkrankung an einem Tumor z. B. die Operation oder das Zuwarten bessere Ergebnisse bringt. Hier ergeben sich in der Untersuchung viele der Probleme, die von manchen als Ausschlusskriterien einer Auswertung nach EbM angeführt werden. Z. B. kann es sein, dass ein bestimmter Typ von Menschen es vorzieht zu warten, ein anderer eher gescreent zu werden, ein dritter eher operiert zu werden. Dennoch bleibt die Ausgangsfrage. Weiterhin wird unter dem Vorbehalt der Nichtvergleichbarkeit, der Singularität eines Kontextes behauptet, dies lasse die Bildung von Kontrolldesigns nicht zu. Nach Auffassung der Autoren gilt dies auch für jede Medikamentenstudie, denn es muss entschieden werden, ob in Studien berücksichtigte Altersgruppen, Ethnie oder weitere Indikationen bei der Frage der Übertragbarkeit vernachlässigt werden dürfen.

Auch die von Trojan (2006a, 2006b) aufgeworfenen Fragen, z.B. welche Themen erforscht werden sollen bzw. wie pharmakologische gegen nichtpharmakologische Interventionen getestet werden, stellen sich ebenso für die „klassische“ EbM. Wer entscheidet, wenn Testpersonen generell lieber ein Medikament wollen? Es stellt sich dann wie bei vielen Interventionen der Prävention und Gesundheitsförderung die Frage der Freiwilligkeit der Teilnehmenden auch. Darüber hinaus sind die Fragen, wer bezahlt oder wer entscheidet, was gute (Studien)Qualität ist (s. Ja-

dadscore gegen viele andere Scores) sowohl für eng medizinische als auch für weit gesundheitsfördernde Interventionen von Belang.

Abbildung 5: Das Kontinuum verschiedenartiger Interventionen



Legende: In der Theoriebildung gibt es generell zwei Vorstellungen, wie man damit umgeht, dass es unterschiedlich komplexe Interventionen gibt. Dies kann dichotom aufgefasst werden, aber auch als Kontinuum, so dass die Frage, ob eine Randomisierung, Verblindung der Auswerter etc. angemessen sind, nicht durch eine Dichotomisierung der Gegenstände per se ausgeschlossen bzw. gefordert werden kann. CAM = Complementary and Alternative Medicine.

Festzuhalten ist, dass sich der Untersuchungsgegenstand pharmakologischer Studien klarer abgrenzen lässt. Dennoch gilt: Alle Probleme, die es in der Gesundheitsförderung angeblich unmöglich machen, dass so etwas wie evidenzbasierte Wissensbildung zur Anwendung kommt, gelten nach Auffassung der Autoren auch im medizinischen Alltag. Daher gilt, dass eine klare Dichotomie von EbM vs. Gesundheitsförderung/ Prävention heuristisch hilfreich sein mag, aber Medizin und Gesundheitsförderung/ Prävention letztlich als zwei Pole eines Kontinuums anzusehen sind.<sup>4</sup>

Die Ablehnung einer Hierarchisierung nach evidenzbasierten Kriterien mit dem Argument, randomisiert kontrollierte Studien (RCTs) als Goldstandard seien in der Prävention nicht anwendbar (IoM 2001, S. 279f), beruht zum einen, auf einem positivistisch ausgelegten Zerrbild von EbM. Zum anderen vernachlässigt es, dass EbM im Bereich Evidence-based Patient Choice (EbPC,

<sup>4</sup> Wissenschaftstheoretisch kann man dies an der Abgrenzung der Konzepte von Gesundheit und Krankheit in normativ und naturalistisch verdeutlichen. Die Neufassung dieser Dichotomie im Sinne einer dialektischen Auflösung ihrer selbst kann bei Khushf (2007) nachverfolgt werden.

Edwards und Elwyn 2001) wie auch im Bereich der Bewertung von Informationsmaterial für Patienten (Moyer und Elliott 2004) methodisch differenzierter an Fragestellungen herangeht.

### 3.3.2 Die Wahl der Kriterien und Zielparameter

Dass es bei der Übertragung und Anwendung von Methoden der EbM aus dem Sektor der Beurteilung therapeutischer und diagnostischer Maßnahmen in den Bereich präventiver und gesundheitsfördernder Interventionen bzw. Programme Probleme gibt, ist hinlänglich bekannt (u. a. Rychetnik und Frommer ohne Jahr). Nichtsdestotrotz bleibt die Frage, wie Interventionen aber auch ganze Programme bewertet werden können. Wie ausführlich dargestellt, sind sich Vertreter der EbM bewusst, dass RCTs als „Goldstandard“ in vielen Situationen, und zu diesen gehören neben Prävention und Gesundheitsförderung auch Bereiche wie Screening, Chirurgie, Pädiatrie, Psychiatrie, Alternative Medizin, weder angemessen (appropriate) noch durchführbar (feasible) sind (Wong 2002). Umgekehrt darf aber auch nicht vorschnell aus Kostengründen, Bequemlichkeit oder dem Interesse rasche Ergebnisse vorzulegen, übersehen werden, dass in vielen Kontexten ein RCT planbar und durchführbar ist. Wenn man hier die Vergleichbarkeit z. B. verschiedener Schulen, an denen eine Intervention durchgeführt wird, so weit in Frage stellt, dass sie keinen gemeinsamen RCT zulassen, dann bleibt jede Intervention wissenschaftstheoretisch ein singulärer nicht übertragbarer Akt und muss sich jeweils isoliert als effektiv bzw. nicht-effektiv erweisen.

Abbildung 6: Mittel-Ziel-Hierarchie als Ergebnismodell zur Darstellung von Gesundheitsförderung (in Anlehnung an Nutbeam 1998)

<b>Ergebnisse</b>	<b>Lebensqualität, Unabhängigkeit, Gleichheit, Mortalität, Morbidität, Behinderung</b>
<b>Instrumentalziele</b>	<b>Lebensstil, Gesundheitssystem, Umgebung</b>
<b>Instrumente</b>	<b>Gesundheitskompetenz (-verstehen), soziale Aktivierung, politische Maßnahmen</b>
<b>Aktion</b>	<b>Erziehung, Ermöglichung, Verfechtung</b>

Evidence-based Health Promotion soll hinsichtlich der Messung der Zielerreichung von gesundheitsfördernden Projekten Indikatoren bereit stellen. „Oberstes Ziel in Gesundheitsförderung und Prävention ist immer die Verbesserung der Gesundheit der Zielpopulationen. Die letztlich angestrebten objektiven gesundheitlichen Erträge können in den meist kurzzeitig angelegten Beobachtungsperioden von zwei bis maximal fünf Jahren häufig allerdings nicht oder nicht hinreichend erhoben werden. Es müssen deshalb „intermediäre“ Outcome-Parameter identifiziert und eingesetzt werden, die eine Abschätzung der wahrscheinlichen gesundheitlichen Entwicklung der Teilnehmer erlauben.“ (Schwartz 2000).

Dazu werden Kriterien entwickelt, wobei sich schon bei der Hierarchie der Kriterien zeigt, dass sich Interdependenzen erkennen lassen und somit keine aufsteigende Reihenfolge strictu sensu vorliegt. In der von Nutbeam entwickelten Mittel-Ziel-Hierarchie sind hinsichtlich der Ergebnisse (Basis- oder Fundamentalziele) auch Instrumentalziele zu erkennen, die der Messbarkeit zugänglich sind. Zum einen sind konkrete Indikatoren wie Mortalität und Morbidität auch als Ergebnisse (Outcomes) der Gesundheitsförderung in die Mittel-Ziel-Hierarchie eingeordnet, zum anderen sind sie zugleich messbare Instrumentalziele. Ebenso scheinen die Instrumente zugleich auch Zielvorgaben zu sein. In Bezug auf Prävention und Gesundheitsförderung sind manche Parameter zugleich Instrumente, Instrumentalziele und Basisziele.

Die Anforderung, ein auf kontextuelle Umsetzung hin ausgerichtetes Evaluationsschema zu entwickeln, bleibt also bestehen. Diese Diskussion (s. Rychetnik und Frommer 2000) wird durchaus auch im Bereich derjenigen, die EbM fördern, kontrovers gesehen. Von den mehrfach zitierten Rychetnik et al. wird betont, dass soziale und kulturelle Fragen („contextual social and political factors“) immer in die Evaluation einbezogen werden müssen.

Zur Evaluation gehören deshalb neben der Darstellung genauer Angaben zur Publikation, der Intervention, des Bewertungskontextes (-horizontes) und der Methode zur Bewertung der Intervention vor allem auch die Aufarbeitung des „context (or setting)“ der Intervention, insbesondere auch mit der Fragestellung, ob die „contextual factors are critical or integral to the intervention“. (Rychetnik und Frommer 2000). Es muss auch gewährleistet sein, dass der Kontext so detailliert dargestellt wird, dass die Übertragbarkeit der Intervention auf andere Kontexte abzuschätzen ist. Übertragbarkeit (transferability) kann immer nur vor dem Hintergrund von Gege-

benheiten sowohl kultureller als auch politischer Natur sowie des Gesundheitssystems bewertet werden.

In Australien wurde vom National Public Health Partnership (NPHP) in einem ersten Schritt ein Instrumentarium zur evidenzgesteuerten Bewertung von präventiven und gesundheitsfördernden Maßnahmen entwickelt. Darauf wurde in einem zweiten Schritt dieses Schema revidiert an Hand der Applikation auf Settings, "evidence-agenda map". Ziel des Schemas war es u. a. (Rychetnik und Wise 2004), "implement interventions that are proven to be effective and safe; efficient, feasible ...; interventions to address social, economic and environmental determinants of health." „Different groups in different contexts will identify different health promotion agendas.“ (Rychetnik und Wise 2004, S. 254). Dabei werden drei Ziele im Verhältnis zu fünf Dimensionen von Prävention und Gesundheitsförderung angesetzt (s. Abbildung 7).

Abbildung 7: Evidenz-Agenda-Matrix (in Anlehnung an Rychetnik und Wise 2004)

"Evidenz"	Stufen der Prävention bzw. Gesundheitsförderung		
	Verhaltens- und Risikomodifikation	Veränderung der sozialen, ökonomischen und ökologischen Bedingungen von Gesundheit	Reduktion von Ungleichheit im Bereich Gesundheit
Ausmaß und Ätiologie	<b>Bewertung der relevanten verfügbaren Evidenz</b>		
Effektivität lokaler Programme			
Effektivität übergeordneter Initiativen, bzw. Umsetzung auf größere Bevölkerungsteile			
Nachweis für weitere Veränderung außerhalb des "Gesundheitssektors"			
Kosteneffektivität der Maßnahmen			

### 3.3.3 Schaden und Nichtwirksamkeit in Prävention und Gesundheitsförderung

Es gibt eine stillschweigende Übereinkunft, dass Interventionen der Prävention und der Gesundheitsförderung, die nicht helfen, wenigstens nicht schaden können. Wir müssen nicht darauf verweisen, dass Impfungen – wenn auch in viel geringerem Maße als angenommen – und Screeninguntersuchungen auch schädliche Folgen für die Zielgruppen haben können. Auch für Interventionen in den Bereichen Bewegung und Ernährung kann es zu negativen Folgen für unterschiedliche Subgruppen kommen. Darauf verweisen auch Harden et al. (2004) zu derselben Frage vor dem Hintergrund der Reviews des EPPI zu psychischer Gesundheit, Ernährung und Bewegung und unterscheidet „effective, ineffective, and harmful interventions“. Ein Beispiel sind z. B. Jugendzentren, die im guten Glauben eingerichtet wurden, den Jugendlichen Freiräume zu bieten. Viele dieser Zentren sind „nur“ Aufenthaltsort für Jugendliche aus den sog. sozialen Randgruppen. Zumindest vertane Zeit beim Aufbau der Programme und bei der Teilnahme können als Kosten angesehen werden. Diese Zeit (und natürlich auch anderweitige Ressourcen) hätte in anderen Verwendungen einen positiven Nutzen stiften können.

Auch aus diesem Grund ist eine systematische Zusammenstellung und eine Bewertung von Interventionen nötig, um unerwünschte Effekte aufzudecken, z. B. dass bei einer Intervention mit Bewegung über ganze Jahrgänge ein Effekt feststellbar ist, aber bei genauerer Analyse nur für die sowieso schon Schlanken erkennbar ist. Dies konnte auch die so niederländische Intervention Food Steps<sup>5</sup> belegen, bei der sich Effekte durch das Treppensteigen auf die zuvor schon Schlanken ausgewirkt haben. Damit kann diese Intervention sogar als schädlich bewertet werden, da sie den Unterschied zwischen den Normalgewichtigen und den Übergewichtigen noch erhöht hat.

---

<sup>5</sup> Vorstellung durch W van Mechelen anlässlich der Europäischen Konferenz „Gesundheitliche Prävention“ vom 25. bis zum 27. Februar 2007 in Badenweiler. Poppel M van, Engbers W, Mechelen W van. FoodSteps. Reshaping an office environment. Does it make sense? S. dazu Engbers et al. 2007b. Darin wurde durch Zeichen etc. auf die Benutzung der Treppen aufmerksam gemacht. Ein signifikanter Effekt hat sich nur für die Männer ergeben. Der Body-Mass-Index hat sich signifikant nur bei denjenigen mit einem BMI unter 25kg/qm verändert.

### **3.4 Evaluation und Indikatoren der Qualität:**

#### **Nachvollziehbarkeit und Konsens**

Grundlage der EbM muss immer das systematische und explizite Vorgehen sein, das (möglichst) den gesamten Wissensbestand einbeziehen sollte. Dieser muss dann nach Kriterien, die auf der Basis eines Konsenses gewonnen werden, eingeteilt und bewertet werden. Die Einteilung kann sich nach formalen Kriterien der Studiendurchführung richten. Zusätzlich kann auch eine Bewertung an Hand von Qualitätskriterien berücksichtigt werden. Diese können entweder aus der systematischen Suche und Bewertung des Wissensbestands gewonnen werden oder auf Basis von Wertvorstellungen. So wurde das Instrument zur Priorisierung von Interventionen in Finnland ganz bewusst auf der Grundlage von demokratischen Prinzipien im menschlichen Miteinander entwickelt. Dies bedeutet, dass Gesundheitsförderung als moralische Aufgabe gerade den Respekt vor und damit die aktive Beteiligung einer Zielgruppe an einer Intervention fordert.

In der wissenschaftlichen Diskussion wird Qualität im Bereich Prävention und Gesundheitsförderung unterschiedlich abgegrenzt. Eine Abgrenzung unterscheidet Prozess-, Struktur- und Ergebnisqualität. Eine andere Abgrenzung bezieht die Konzeptqualität in die Qualitätssicherung ein. Qualität einer Studiendurchführung dient in der Evidenzbasierten Medizin (EbM) als Hilfe, sich Gewissheit darüber zu schaffen, ob die Ergebnisse einer Studie verlässlich sind. In diesem Sinne sind die Indikatoren der Qualitätssicherung in die Evaluation von publizierten Interventionen einzubeziehen.

Es stellen sich die Fragen, wie aus diesen unterschiedlichen Definitionen von Qualität ein Indikator für eine Intervention/Studie gewonnen werden kann bzw. wie Studienqualität und Interventionsqualität zueinander in Beziehung gesetzt werden können. Weiterhin muss gefragt werden, welche Kriterien wurden normativ gewonnen – was nicht verwerflich ist, denn Qualität beinhaltet immer eine Norm, da eine Zielerfüllung definiert wird und dies dann an Indikatoren festgemacht wird. Allerdings ist es auch möglich, diese Indikatoren empirisch im Hinblick auf ihre Sensitivität, Qualität bzw. Qualitätsdefizite zu prüfen. Dies kann wiederum nach Methoden und Kriterien der EbM erfolgen. Ein Beispiel ist die, wenn auch nicht eindeutig als systematisches Review erkennbare, Darstellung von Wallerstein (2006) zur Evidenz von Empowerment. Sie kann an wenigen Beispielen nachweisen, dass Interventionen, die zusätzlich Strategien zu Empowerment integrie-

ren, bei Projekten der Entwicklungsarbeit zu einer umfassenden Veränderung von Strukturen in den Zieldörfern führen.

Allerdings ist der Streit Werte/Tatsachen (fact/values) der EbM nicht fremd, da es auch für die Bewertung von RCTs eine Vielzahl von Scores gibt, die unterschiedliche Kriterien für Qualität heranziehen und diese dann unterschiedlich gewichten. Dies reicht von dem einfachsten Jadadscore mit nur fünf Items bis zu hochgradig komplexen Scores mit einer Vielzahl von Fragen und z. T. sogar Gewichtungen der Antworten. Wie Timmer und Antes (2006, S.21) bestätigen, ist der Streit, ob ein Gesamtscore für die Güte gebildet werden soll oder die Güte der Einzelkriterien ausschlaggebend ist, nicht beigelegt. Und er wird auch mit Methoden der EbM nicht beigelegt sein.

Weiterhin muss die Frage der Willkürlichkeit von Bewertungsinstrumenten geklärt werden. Proper et al. (2002) nennen ihr "rating system" freiweg "arbitrary", aber zugleich "considered a suitable method for distinguishing objectively between different study designs ... and to assess the methodological quality of studies ... in a reproducible manner." (S. 82). Und weiter "...cut-off points used for categorizing a study as a high-quality study and for concluding consistent results were chosen arbitrarily." Ihre Sensitivitätsanalyse mit einer 10% Veränderung des „cut-off points“, Studien ein- bzw. auszuschließen, hat aber keine Veränderung ergeben. Damit zeigt sich erneut, dass EbM auf Nachvollziehbarkeit und Konsens angelegt ist, aber die Wahl und die Wertigkeit von Kriterien damit immer Gegenstand eines Kommunikations- wie Entscheidungsprozesses bleibt.

Wichtig ist jedenfalls zu bedenken, dass Qualität nicht direkt gemessen werden kann, sondern immer nur über Indikatoren, die Hinweise auf gute bzw. schlechte Qualität geben. Für diese Indikatoren gelten bestimmte Anforderungen, die in den Akronymen SMART bzw. RUMBA-Regel konkretisiert werden.

#### **RUMBA=**

**R** relevant, d. h. wichtig für einen Problembereich,

**U** nderstandable, d. h. der Indikator sollte sowohl von den Leistungserbringern als auch von den Patienten verstanden werden,

**M** easurable, d. h. der Indikator sollte mit hoher Reliabilität und Validität messbar sein,

**B** ehaviourable, d. h. der Indikator sollte sich auf Bereiche beziehen, die durch eine Verhaltensänderung verbessert werden können,

**A** chievable, d. h. das Einhalten des Indikators sollte möglich sein.

Qualitätskriterien bleiben damit aber immer „somewhat arbitrary“ (s. Harden et al. 2004, S. 795f) sowie unscharf. Dies bedeutet aber im Umkehrschluss, dass sie dennoch möglichst „clear, sufficient, explicit“ ausgewählt und formuliert werden sollten, damit sie dem Kriterium der Nachvollziehbarkeit Genüge tun. Es darf nicht mehr einfach eine Studie mit dem Argument schlechter Qualität ausgeschlossen werden, sondern dies muss an zwar u.U. strittigen, aber nachprüfbaren Kriterien festgemacht werden.

Folgende Frage zur Interventionsqualität wird im Rahmen einer Synthese eines EbM-Ansatzes kritisch diskutiert: Sollen Empowerment und Partizipation als Elemente der Interventionsqualität herangezogen werden, auch wenn ihre Überlegenheit in Hinblick auf die erreichte Wirksamkeit gegenüber Studien ohne die beiden Elemente noch nicht belegt ist? Das Vorgehen im vorliegenden Forschungsprojekt basiert auf folgende Überlegungen:

Da sich insbesondere die Gesundheitsförderung auf dem Boden der Ottawa-Charta verortet, gehören Empowerment und Partizipation zum Selbstverständnis dieses Denkens. Ebenso versteht sich z. B. die finnische Instrument zur Priorisierung: Im Rahmen eines demokratisch verhafteten und die Demokratie weiter stärkenden Verständnisses von Interventionen müssen diese inhärent auf Empowerment und Partizipation ausgehen.

In diesem Sinne könnte gleichermaßen gefragt werden, ob die Demokratie bestimmte andere Ziele besser erreicht als die Aristokratie. Genau dies würde aber gar nicht zum Gegenstand von Studien gemacht werden, denn a priori ist hier schon eine Entscheidung auf Basis von Werten gefallen.

Weiterhin basiert auch die EbM immer auf Wertentscheidungen. Wann wird ein Ergebnis einer Untersuchung als signifikant angesehen. Dies kann einen p-Wert von 0,05, 0,01 oder sogar kleiner als 0,01 bedeuten. Damit sind hier immer konsensuell-willkürliche Entscheidungen auch in der EbM, die auf einer Einschätzung basieren, ab wann man ein Ergebnis als gewiss bzw. sicher auffassen darf.

Da wir aber dennoch den Einwand ernst nehmen, werden wir im Verlauf der Arbeit eine Sensitivitätsanalyse der bewerteten Studien vornehmen. D.h. neben der Bewertung mit dem kompletten

Kriterienkatalog wird zusätzlich eine Bewertung unter Abzug der Kriterien Empowerment und Partizipation erfolgen. Dies entspricht z. B. einer Sensitivitätsanalyse nach Proper et al. (2003).

### 3.5 Von der Bewertung zur Empfehlung

Selbst wenn ein systematisches Review erstellt worden ist, bedeutet das nicht, dass damit schon unmittelbar umzusetzende Handlungsoptionen erkennbar wären. An die Bewertung und Hierarchisierung von Studien schließen sich weitere Bewertungsverfahren an, aus denen sich dann Empfehlungen gewinnen lassen.

Das folgende Ratingsystem wurde im American Journal of Health Promotion zur Gesamtbewertung der Evidenz für die Wirksamkeit von Interventionen der Prävention (dort S. 41) vorgelegt:

- **Conclusive:** Die Datenlage ist schlüssig, da sie auf Basis einer umfangreichen Zahl von Studien mit gutem Untersuchungsdesign und randomisierter Kontrollgruppe gewonnen wurde. Die Experten stimmen fast universell überein in ihrer Bewertung der Studienlage.
- **Acceptable:** Die Datenlage wird als akzeptabel bezeichnet und entspricht in Umfang und Art der vorhergehenden. Allerdings stimmt nur eine Mehrheit der Experten in der Bewertung der Datenlage überein.
- **Indicative:** Die Datenlage wird als hinweisend auf ein Vorgehen angesehen. Es liegt ein gutes Studiendesign vor, aber keine oder wenige mit Kontrollen. Eine Mehrheit der Experten sieht einen kausalen Zusammenhang, bewertet diesen aber als vorläufig, da keine RCTs vorliegen und daher alternative Erklärungen für den Zusammenhang möglich sind.
- **Suggestive:** Die Kategorie suggestiv basiert auf konsistenten Ergebnissen. Diese kommen aber nicht aus Studien mit gutem Design. Die Mehrheit der Experten sieht die Befundlage als plausibel, aber begrenzt an.
- **Weak:** Die schwache Datenlage ergibt sich auf Basis nicht experimenteller und/ oder ungenügend operationalisierter Forschungsbefunde. Experten sprechen der Wirksamkeit eine gewisse Plausibilität zu.

Eine andere Umsetzung in vier Stufen findet sich bei Richardson (2005), die in Tabelle 1 dargestellt wird.

Tabelle 1: Darstellung der vier Grade von Evidenz, die einen Umsetzungsanspruch einer Intervention begründen (Four grades of evidence in support of a health claim)

Health claim	Grade of evidence	Qualifying language
Yes	Convincing	“Experts agree that scientific evidence supports ...” Modal verb “will”
Yes	Probable	„... Although there is scientific evidence supporting the claim, the evidence is not conclusive Modal verb: “can”
Yes	Possible	“Some scientific evidence suggests ... However, the evidence is limited and not conclusive Modal verb “may”
No	Insufficient	“There is little scientific evidence supporting this claim.”

Fazit: “The evidence never speaks for itself, but is always open to interpretation.“, was selbst für den Streit um den Effekt unterschiedlicher Statine und das Brustkrebscreening gilt. Damit folgt aber nicht, das Ziel einer systematischen und umfassenden Bewertung aufzugeben, sondern: „None the less, we suggest that it is preferable to reach conclusions, however tentative, that are based on the best available evidence rather than simply stating that no evidence is available.“ Im Anschluss muss dann ein Expertengremium die aufbereitete “Evidenz” in Empfehlungen gießen.

## 4 Der IGKE-Kriterienkatalog zur Evaluation von Interventionen

In diesem Kapitel werden grundlegende inhaltliche Überlegungen ausgeführt, die zur Entstehung des Kriterienkatalogs beigetragen haben. Weiterhin werden der Katalog selbst und das dazu gehörige Manual vorgestellt. Abschließend folgen die Anweisungen zur Auswertung sowie Hinweise zu einer Weiterentwicklung des Kriterienkatalogs, die sich aus der Evaluation von Studien ergeben haben.

Das Manual ist aus der Erarbeitung eines Kriterienkatalogs unter Berücksichtigung der international gebräuchlichen Qualitätskriterien-Instrumente wie PREFFI, IDM etc. und interner Diskussionen und Erfahrungen mit dem Kriterienkatalog entstanden. Es dient der Erklärung der im Fragenkatalog gegebenen Antwortkategorien.

### 4.1 Rationale des IGKE-Kriterienkatalogs

Der IGKE-Kriterienkatalog zur Evaluation von Interventionen soll es Versorgungsträgern und Geldgebern ermöglichen, erfolgversprechende von nicht erfolgversprechender Intervention in Prävention und Gesundheitsförderung zu trennen. Er hat die Funktion eines Screeninginstruments, das primär zur Bewertung schon durchgeführter Interventionen anwendbar ist, um aus diesen Kandidaten für eine Übertragbarkeit auf Deutschland auszuwählen. Das Instrument ist auch auf Interventionen im Antragsstadium anwendbar, wenn Geldgeber zwischen verschiedenen Interventionen, die sich um eine Förderung bewerben, entscheiden müssen. Es ist aber gleichfalls denkbar und zulässig, dass es auf durchzuführende Intervention (im Projektstadium) angewandt wird. Auf Basis der drei Dimensionen Effekt, Qualität der Studiendurchführung und Qualität der Intervention soll hinlänglich Gewissheit darüber erzielt werden, ob eine Intervention bei erneuter Durchführung einen reproduzierbaren Effekt erwarten lassen kann.

#### 4.1.1 Dimension Effekt

Studien können positive oder negative Effekte haben. Allerdings gibt es auch viele Beispiele von Studien, die entweder für einzelne untersuchte Subgruppen oder hinsichtlich einzelner Parameter

Effekte nachweisen können, die für andere Subgruppen oder Parameter ausbleiben. Diese Studien werden von uns in die Kategorie teils-teils eingestuft und erfordern eine weitere Prüfung, wenn sie auf Basis der beiden Qualitätsdimensionen in die Felder potenziell empfehlenswerter Interventionen fallen. Damit wird vor allem dem Umstand Rechnung getragen, dass im Bereich Prävention und Gesundheitsförderung oftmals Stellvertreterparameter zur Bewertung des Outcomes herangezogen werden, die unterschiedlich bewertet werden können in Bezug auf das, was Versorgungsträger sich z. B. von einer Intervention erwarten.

### **4.1.2 Dimension Studienqualität**

Zur Studienqualität bzw. Qualität der Studiendurchführung gehören unserer Ansicht nach Kriterien, die wie in der Evidenzbasierten Medizin je nach Einhaltung Grade der Sicherheit des Studienergebnisses zulassen. Es fließen also Kriterien wie die Darstellung der Randomisierung und die Vergleichbarkeit von Interventions- und Kontrollgruppen, die Berechnung von Power etc. Diese Kriterien sind bei einer Umsetzung einer Intervention nach dem Stadium der Überprüfung durch die Studienform nicht mehr relevant.

### **4.1.3 Dimension Interventionsqualität**

Zur Interventionsqualität hingegen gehören alle Kriterien, die sich auf die Idee des Programms beziehen. Sie bleiben auch nach erfolgreichem Durchlaufen eines Studienstadiums relevante Kriterien. Dazu gehören beispielsweise die Schulung von Multiplikatoren, die Berücksichtigung des Kontexts, die Kompetenzzuweisung. Dabei werden Kriterien zu Prävention und Gesundheitsförderung aus den genannten Instrumenten zusammen geführt. Hierbei spielt im Rahmen unserer Anwendung vor allem auch eine Rolle, ob sich ein entsprechender Qualitätsindikator operationalisierbar abfragen lässt.

Der IGKE-Kriterienkatalog (s. nächsten Abschnitt 4.2) wurde im webbasierten Statistikprogramm Grafstat2 zur Bearbeitung der Studien eingetragen. Die qualitätswirksamen Fragen werden mit dem Kürzel QS (Qualität der Studiendurchführung) und QI (Qualität der Intervention) gekennzeichnet.

Nachfolgend wird die inhaltlich und wörtlich identische Version als Word-Dokument integriert.

## 4.2 Der IGKE- Kriterienkatalog



UNIKLINIK  
KÖLN

### IGKE- Kriterienkatalog

1. Kennung _____ 0-7000
2. Titel der Publikation _____
3. Autor der Publikation _____
4. Publikationsjahr _____ 1990-2006
5. Bewertung der Publikation am _____
6. Quelle der Publikation <input type="checkbox"/> Literaturrecherche <input type="checkbox"/> handverlesen
7. Titel weiterer Publikationen _____
8. Autoren weiterer Publikationen _____
9. Publikationsjahr(e) weiterer Publikationen _____
10. Quelle der weiteren Publikationen <input type="checkbox"/> Literaturrecherche <input type="checkbox"/> handverlesen
11. In welchem Land wurde die Intervention durchgeführt? (DI) <input type="checkbox"/> Australien <input type="checkbox"/> Dänemark <input type="checkbox"/> Frankreich <input type="checkbox"/> Neuseeland <input type="checkbox"/> Schweiz <input type="checkbox"/> Österreich <input type="checkbox"/> Deutschland <input type="checkbox"/> Großbritannien <input type="checkbox"/> Norwegen <input type="checkbox"/> USA <input type="checkbox"/> Kanada <input type="checkbox"/> Finnland <input type="checkbox"/> Niederlande <input type="checkbox"/> Schweden
12. In welchem Bereich bzw. Bereichen werden Interventionen durchgeführt? (DI) <input type="checkbox"/> Rauchen <input type="checkbox"/> Ernährung <input type="checkbox"/> Bewegung <input type="checkbox"/> Stress <input type="checkbox"/> Depression
13. Wurden die Fragestellungen und/oder Hypothesen und/oder Ziele formuliert? (QS) <input type="checkbox"/> ja <input type="checkbox"/> nein <input type="checkbox"/> keine Angaben
14. Benennung der Fragestellung oder der Hypothesen oder der Ziele. (QS) _____
15. Welches Studiendesign wurde gewählt? (DS) <input type="checkbox"/> RCT <input type="checkbox"/> CT <input type="checkbox"/> Intervention ohne Kontrolle
16. Wurde die Art der Randomisierung beschrieben? (QS) <input type="checkbox"/> ja <input type="checkbox"/> nein <input type="checkbox"/> an anderer Stelle

17. Wenn ja, fallen Besonderheiten auf, die einen Verdacht des Einflusses auf den Effekt nicht ausschließen? (QS)
_____
18. Wurde eine qualitative Begleitforschung zum Verlauf und/oder zur Wirkung durchgeführt (multimethodische Triangulation)? (DS)
<input type="checkbox"/> ja <input type="checkbox"/> nein
19. Sind die Kontrollgruppen und Interventionsgruppen vergleichbar? (QS)
<input type="checkbox"/> ja <input type="checkbox"/> nein <input type="checkbox"/> nicht zutreffend <input type="checkbox"/> nicht erkennbar
20. Häufigkeit bzw. Intensität der Interventionen? (DI)
<input type="checkbox"/> einmal <input type="checkbox"/> multipel/kontinuierlich
21. Wie viele Monate wurden die multiplen/kontinuierlichen Interventionen durchgeführt? (DI)
_____ Monate
22. Wie lange wurde die Nachbeobachtung durchgeführt? (DS)
_____ Monate
23. Welche Effektmaße wurden verwendet? (ggf. begrenzen) (DS)
_____
24. Art der Datenerhebung? (DS)
<input type="checkbox"/> Effektmaße per objektiver Messung <input type="checkbox"/> Effektmaße per Selbsteinschätzung
<input type="checkbox"/> keine Angaben
25. Wurden die Messinstrumente validiert (Pretest)? (QS)
<input type="checkbox"/> ja <input type="checkbox"/> nein <input type="checkbox"/> zum Teil <input type="checkbox"/> keine Angaben
26. Wie stark und signifikant (p-Wert) waren die gemessenen Effekte? (DS)
_____
27. Welche Art von Interventionen wurde durchgeführt? (DI)
_____
28. Art der Interventionsdurchführung? (DI)
<input type="checkbox"/> Information <input type="checkbox"/> Aktion passiv <input type="checkbox"/> freie Aktivität (ohne Vorgaben)
<input type="checkbox"/> Beratung <input type="checkbox"/> vorgegebene Aktivität

<p>29. Welche Art der Information wurde gewählt? (DI)</p> <p><input type="checkbox"/> Massenmedien      <input type="checkbox"/> mündlich      <input type="checkbox"/> Internet</p> <p><input type="checkbox"/> schriftliches Material      <input type="checkbox"/> Ausstellung</p>
<p>30. Welches Setting der Beratung wurde gewählt? (DI)</p> <p><input type="checkbox"/> Einzelberatung      <input type="checkbox"/> Gruppenberatung</p>
<p>31. Welcher Ansatz dominiert? (DI)</p> <p><input type="checkbox"/> Individuum      <input type="checkbox"/> Setting      <input type="checkbox"/> Bevölkerung</p>
<p>32. Welche Art des Ansatzes wurde gewählt? (DI)</p> <p><input type="checkbox"/> Verhalten      <input type="checkbox"/> Umwelt</p>
<p>33. Wo wurden die Interventionen durchgeführt? (DI)</p> <p><input type="checkbox"/> Arztpraxis    <input type="checkbox"/> Kindergarten    <input type="checkbox"/> Pflegeheim      <input type="checkbox"/> Öffentlicher Raum</p> <p><input type="checkbox"/> Betrieb      <input type="checkbox"/> Gemeinde      <input type="checkbox"/> Altenheim      <input type="checkbox"/> sonstiges _____</p> <p><input type="checkbox"/> Schule      <input type="checkbox"/> Verein      <input type="checkbox"/> Krankenhaus</p>
<p>34. Wurden Instrumente zur Qualitätssicherung verwendet? (QI)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein      <input type="checkbox"/> keine Angaben</p>
<p>35. Welche Qualitätsinstrumente wurden verwendet? (QI)</p> <p><input type="checkbox"/> EQUIP      <input type="checkbox"/> IDM      <input type="checkbox"/> BZgA (Closing the Gap)</p> <p><input type="checkbox"/> PREFFI      <input type="checkbox"/> OSP      <input type="checkbox"/> Quint-Essenz</p> <p><input type="checkbox"/> keine Angaben      <input type="checkbox"/> sonstige _____</p>
<p>36. Wie alt war die Zielgruppe überwiegend? (DI)</p> <p><input type="checkbox"/> 0-6 Jahre    <input type="checkbox"/> 7-16 Jahre    <input type="checkbox"/> 17-25 Jahre</p> <p><input type="checkbox"/> 26-65 Jahre    <input type="checkbox"/> &gt; 80 Jahre</p>
<p>37. Welches Geschlecht hatte die Zielgruppe? (DI)</p> <p><input type="checkbox"/> weiblich    <input type="checkbox"/> männlich    <input type="checkbox"/> keine Angaben</p>
<p>38. Welche ethnische Abstammung hatte die Zielgruppe? (DI)</p> <p><input type="checkbox"/> keine Angaben    <input type="checkbox"/> mehr als eine Ethnie    <input type="checkbox"/> eine Ethnie _____</p>
<p>39. Wurde die Zielgruppe an Hand von Vorerkrankungen ausgewählt? (DI)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein</p>

40. Welche Vorerkrankungen? (DI)

\_\_\_\_\_

41. Welchen Ansatz hatte die Intervention? (DI)

- schichtspezifisch     altersspezifisch     religionsspezifisch  
 genderspezifisch     kulturspezifisch     unspezifisch     sonstiges \_\_\_\_\_

42. Auf welcher räumlichen Ebene waren die Interventionen angesiedelt? (DI)

- lokal     regional     national

43. Ist die Zielgruppe repräsentativ in bezug auf das Land in dem die Intervention durchgeführt wird? (DI)

- ja     nein     offen

44. Welche wissenschaftliche Evidenz wurde im Vorfeld der Interventionen berücksichtigt? (QI)

- Literatur     Pretest     anderes Projekt     unklar \_\_\_\_\_

45. Welche Angaben wurden zur Evidenz der berücksichtigten Literatur gemacht? (QI)

- Metaanalyse     Beobachtung     keine Evidenz  
 Expertenmeinung     RCT/CT     Systematisches Review  
 Review     unklar

46. Wurde eine Powerberechnung zur Größe der Stichprobe durchgeführt? (QS)

- ja     nein

47. Wie viele Probanden der Interventionsgruppe(n) wurden zum letzten Messzeitpunkt ausgewertet? (DS)

\_\_\_\_\_ absolute Zahl

48. Wie viele Probanden der Kontrollgruppe wurden ausgewertet? (DS)

\_\_\_\_\_ absolute Zahl

49. Welches Auswertungsverfahren wurde verwendet?

- ITT     per Protocol     keine Angaben

50. Wurden die statistischen Methoden zur Analyse des Zielkriteriums transparent beschrieben? (QS)

- ja     nein

<p>51. Ist die Auswertung in der Studienplanung als Subgruppenanalyse ausgelegt? (DS)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein</p>
<p>52. Wurden die zugrundeliegenden Theorien und Konzepte der Intervention explizit benannt? (DI)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein</p>
<p>53. Wurden die projektrelevanten Voraussetzungen der Zielgruppe analysiert? (DI)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein      <input type="checkbox"/> keine Angaben      <input type="checkbox"/> nicht zutreffend</p>
<p>54. Wurde die primäre Zielgruppe bei der Konzeptentwicklung beteiligt? (QI)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein      <input type="checkbox"/> keine Angaben      <input type="checkbox"/> nicht anwendbar</p>
<p>55. Wurde eine Strategie zur Stärkungen der Selbstwirksamkeit (Empowerment) in die Intervention integriert? (QI)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein      <input type="checkbox"/> keine Angaben</p>
<p>56. Wurden Multiplikatoren an der Konzeptentwicklung beteiligt? (QI)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein      <input type="checkbox"/> nicht zutreffend</p>
<p>57. Wurden Multiplikatoren bei der Durchführung der aktuellen Intervention systematisch beteiligt? (QI)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein      <input type="checkbox"/> nicht zutreffend</p>
<p>58. Wurden die Multiplikatoren qualifiziert? (QI)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein      <input type="checkbox"/> nicht zutreffend</p>
<p>59. Wie wurde der Bedarf für Interventionen analysiert?</p> <p><input type="checkbox"/> nach Studienlage    <input type="checkbox"/> eigene Berechnung    <input type="checkbox"/> keine Angabe</p>
<p>60. Woran wurde der Bedarf gemessen? (DI)</p> <p><input type="checkbox"/> Kosten      <input type="checkbox"/> Prävalenz/Inzidenz    <input type="checkbox"/> nach Krankheitsschwere</p>
<p>61. Wurde der Kontext berücksichtigt? (QI)</p> <p><input type="checkbox"/> ja      <input type="checkbox"/> nein      <input type="checkbox"/> keine Angaben</p>

62. Gab es klare Kompetenzzuweisungen während der Durchführung der Maßnahmen?

ja       nein       keine Angaben

63. Wurde der Prozess dokumentiert? (QI)

ja       nein       keine Angaben

64. Wurde der Prozess evaluiert? (QI)

ja       nein       keine Angaben

65. Wurden die Interventionen als effektiv bewertet? (DI)

ja       nein       teils teils       mögliche sonstige Einflüsse

66. Wurde die Organisation durch die Intervention verändert? (DI)

ja       nein       keine Angaben       nicht relevant

67. Wurde die Intervention in der Organisation dauerhaft implementiert? (DI)

ja       nein       keine Angaben       nicht relevant

68. Wurde die Intervention über die Organisation hinaus dauerhaft implementiert? (DI)

ja       nein       keine Angaben       nicht relevant

### 4.3 Manual zum IGKE-Kriterienkatalog<sup>6</sup>

In dem Manual werden die Fragen des im Kapitel 5.2 abgedruckten IGKE-Kriterienkatalogs nach laufender Nummer erläutert. Dazu wird das abgefragte Kriterium zunächst in Kurzform angeführt. Sinn und Ziel des Kriteriums werden erklärt und Hilfestellungen zur Beantwortung der jeweiligen Frage gegeben.

Die Fragen 1 bis 10 sind zur Identifizierung von Studie und Auswerter/in vorangestellt. Die weiteren 58 Fragen sind den vier Kategorien Deskription der Studie (DS), Deskription der Intervention (DI), Qualität der Studiendurchführung (QS) und Qualität der Intervention (QI) zugeordnet. Auf die Darstellung entfallen insgesamt 38 Fragen, zur Qualität werden insgesamt 20 Fragen gestellt. Diese teilen sich dann jeweils auf die Kategorie Studie und Intervention auf (s. Tabelle 3).

Tabelle 3: Aufteilung der Fragen des Kriterienkatalogs auf die Kategorien Darstellung und Qualität sowie Studien(durchführung) und Intervention

	Studie	Intervention
Darstellung	DS: 10	DI: 28
Qualität	QS: 8	QI: 12

Da die Reihenfolge der Fragen sich aus Gründen der Benutzerfreundlichkeit möglichst am üblichen Aufbau der Publikationen orientiert, werden die den genannten vier Kategorien zugeordneten Fragen nicht gebündelt abgehandelt. Die Kürzel im IGKE-Kriterienkatalog verweisen auf die jeweilige Zuteilung.

Die Fragen werden bis auf die Fragen Nr. 2 bzw. 7 (Titel der Publikationen) auf Deutsch beantwortet, also übersetzt. Sollte es aus irgendeinem Grund wichtig sein, auch auf das Original zu verweisen, dann müsste in Klammern eine Angabe in der Sprache des Texts gemacht werden, z. B. bei dem Effektmaß: Länge der Laufstrecke (walking distance).

Das Abstract sollte nicht in die Auswertung einbezogen werden. Das Abstract dient nur als „Appetizer“ bzw. Orientierung für die Leser/innen. Wenn also eine Frage im Abstract explizit be-

---

<sup>6</sup> Stand: 22. Januar 2007 auf Basis der Veränderungen des Fragebogens vom 18. Januar 2007 und den Ergänzungen zur Auswertung der Qualitätsfragen vom 13. März 2007. Die Reihenfolge der Fragen orientiert sich am Fragebogen im Format Grafstat 2 in der Fassung vom 18. Januar 2007. Das Manual kann fortlaufend aus der Erfahrung heraus verändert bzw. ergänzt werden. Dies sollte jedoch immer in Rücksprache mit (möglichst) allen mit dem Kriterienkatalog Arbeitenden erfolgen.

antwortet wird, aber im Text der Publikation nicht ausgeführt wird, ist dies ein Nachteil in der Qualität der Publikation.

**Frage 1 Kennung 1000-4999<sup>7</sup>**

Tabelle 4: Übersicht über die von uns vergebenen Kennnummern für die vier Handlungsfelder

Bereich	Kennung
Bewegung	1000-1999
Depression	2000-2999
Ernährung	3000-3999
Rauchen	4000-4999

**Frage 2 Titel der Publikation**

Es müssen maximal die ersten zehn Wörter eingegeben werden, es kann aber auch der gesamte Titel eingegeben werden. Der Titel des Originals (keine deutsche Übersetzung) muss eingegeben werden.

Es ist sinnvoll, bei mehreren Publikationen zu einer Intervention bzw. einer Studie, die jüngste Publikation an dieser Stelle einzustellen, damit dann die älteren Publikationen als Antwort zu den Fragen 7 bis 10 eingetragen werden. Ggf. muss eine Umstellung erfolgen, wenn erst im Verlauf eine noch jüngere Publikation zu einer Intervention gefunden wird. Es wird zusätzlich eine Exceltabelle erstellt, in der neben Kennung (1), Titel der Publikation (2), Autoren der Publikation (3) und Publikationsjahr (4) auch die Quelle der Publikation (6) aufgenommen wird. Außerdem werden ggf. weitere Publikationen (7-10) mit in diese Tabelle aufgenommen, wenn die Bewertung einer Intervention nur unter Berücksichtigung mehrerer Publikationen umfassend möglich ist.

**Frage 3 Autor(en) der Publikation**

Format: Meyer A, Müller B, Schmitt C

Sollten es mehr als drei Autoren sein, dann kann nach den zweiten Autoren et al. angefügt werden. Es darf hier wie in Frage 8 auch eine Study Group genannt werden, da über diese Autorenschaft auch der Originalbeitrag wieder aufgefunden werden kann.

**Frage 4 Publikationsjahr**

**Frage 5 Datum der Auswertung**

Format: Jahr/Monat/Tag

---

<sup>7</sup> Diese Verteilung bezieht sich auf die von uns gewählten Kennnummern

**Frage 6      Quelle der Publikation**

Es wird unterschieden, ob die Publikation direkt Ergebnis der systematischen Literatursuche war oder ob die Publikation über Referenzen, Handsuche etc. gefunden wurde.

**Frage 7      Titel weiterer Publikationen**

Da manche Interventionen nicht ausgewertet werden können, wenn nur eine Publikation dazu ausgewertet wird, gilt es hier, die weiteren Publikationen zu Methodenbeschreibung, Subgruppenanalysen, qualitativer Forschung etc. einzutragen. Ggf. ist mit der Auswertung einer Intervention zu warten, bis alle dazu gehörigen Beiträge vorliegen.

Können ggf. ältere Studien, die für die Auswertung wichtig sind, nicht voll eingesehen werden, dann müsste nach Eingabe des Titels in Frage 7 die Bemerkung erfolgen: *nicht eingesehen*. Dies ist ein auch von anderen Autoren eingesetztes Verfahren (vgl. z. B. Proper et al. 2003), wenn Texte auch durch intensives Nachforschen innerhalb einer gewissen Zeit nicht verfügbar gemacht werden können.

Wenn der Sonderfall vorliegt, dass eine Studie identisch mehrfach publiziert ist ohne weitere Subgruppenanalysen, dann gilt, dass nur die jüngste Publikation in die Frage 2 eingetragen werden muss und die älteren gar nicht genannt werden.

**Frage 8      Autoren weiterer Publikationen**

**Frage 9      Publikationsjahre der weiteren Publikationen**

**Frage 10      Quelle der weiteren Publikationen**

**Frage 11      In welchem Land wurde die Intervention durchgeführt?**

Mehrfachnennungen sind möglich.

**Frage 12      Bereich der Intervention**

Mehrfachnennungen sind möglich.

Es reicht nicht aus, dass Ergebnisse aus verschiedenen Bereichen genannt werden, wenn z. B. eine Intervention zum Rauchen gemacht wird, aber auch Auswertungen zu einem veränderten Bewegungsverhalten in der Studie präsentiert werden. D. h. es müssen explizit in der Beschreibung der Intervention einer oder mehrere der anzukreuzenden Bereiche genannt werden.

**Frage 13      Wurden die Fragestellungen und/oder Hypothesen und/oder Ziele formuliert?**

Hier wird in ein formales Kriterium und ein inhaltliches Kriterium differenziert, die beide erfüllt sein müssen: Formales Kriterium: Es müssen explizit eine Frage (auch in indirekter Rede), eine

Hypothese oder ein Ziel formuliert worden sein. Es reicht nicht, wenn der Eindruck entsteht, dass eine Frage oder Hypothese angedacht worden ist, die dann indirekt aus dem Text erschlossen wird. Allerdings wird den Autoren die Möglichkeit gelassen, die Form zu wählen, also Frage, Hypothese oder Ziele.

Inhaltliches Kriterium: Frage, Hypothese oder Ziel dürfen nicht global formuliert sein, sondern es muss der Bezug zu den Endpunkten formuliert werden, d. h. z. B.: Kann die Beratung von Rauchenden in Arztpraxen die Rauchstopprate erhöhen?

**Frage 14 Benennung von Fragestellungen, Hypothesen, Zielen.**

Fragestellungen oder Hypothesen müssen hier benannt werden, wie sie von den Autoren der Studien ausgeführt wurden. Diese sollen übersetzt, nicht in Originalsprache eingetragen werden. Zudem muss explizit ein Ziel (objective/ goal/ aim of the study/intervention) genannt worden sein.

**Frage 15 Studiendesign**

**Frage 16 Art der Randomisierung beschrieben?**

Es muss dezidiert genannt werden, ob z. B. eine Cluster- oder eine Blockrandomisierung erfolgte. Diese Frage gilt in der Auswertung nur für RCTs.

**Frage 17 Besonderheiten, die bei der Randomisierung auffallen und einen Verdacht auf einen Einfluss auf den Effekt nicht ausschließen lassen**

Diese Frage schließt sich an die Frage 16 an. Hier könnte ggf. angeführt werden, dass die Randomisierung wichtige Einflussfaktoren nicht berücksichtigt hat wie z. B. soziale Faktoren: Es wurden in einer Studie beispielweise nur Altenheime in reichen Stadtteilen in die Intervention und nur Altenheime in armen Stadtteilen in die Kontrolle randomisiert, ohne vorher eine Stratifizierung vorzunehmen. Des Weiteren kann angesprochen werden, wenn die von den Autoren genannte Art der Randomisierung nicht der Art der Randomisierung entspricht, die sie dann tatsächlich angewendet haben. Andere Besonderheiten wären z. B., dass die Randomisierung nachträglich verändert wurde. Eine umfassende Liste kann hier nicht gegeben werden, im Zweifelsfall sollten Einflüsse genannt werden, die dann nach der Evaluation ggf. nochmals herangezogen werden sollten, wenn die Studien zu den besonders guten oder besonders schlechten Studien gezählt werden. Dies führt zum Abzug für RCTs. Auch dies ist ein in der Literatur bekanntes Verfahren, s. u. a. Proper et al. 2003.

**Frage 18      Qualitative Begleitforschung zu Verlauf und/oder Wirkung (multimethodische Triangulation)**

Es reicht nicht aus, dass Studienteilnehmer irgendwie (!) zu ihrem Wohlbefinden befragt worden sind, sondern es muss ein qualitatives Design eingehalten werden. Insbesondere soll hier geklärt werden, ob Studiendesigns, die nicht auf dem Level eines RCT durchgeführt worden sind, durch entsprechende qualitative Begleitforschung ihre Validität bzw. Glaubhaftigkeit erhöhen und damit die Sicherheit des Ergebnisses stärken. Die Begleitforschung kann ggf. auch in anderen Publikationen gemacht worden sein; der Bezug darauf muss aber erkennbar sein.

**Frage 19      Vergleich Kontroll- und Interventionsgruppen**

An Hand der Baseline-Analyse soll die Vergleichbarkeit geprüft werden. Wenn wichtige Angaben fehlen, muss die Kategorie „nicht erkennbar“ angekreuzt werden und ggf. mit einem kurzen Kommentar auf das Fehlen entscheidender Größen hingewiesen werden. Insbesondere sollte eine tabellarische Übersicht aus den Studien zur Kontrolle der Vergleichbarkeit herangezogen werden. Die Kategorie nicht zutreffend ist anzukreuzen, wenn keine kontrollierte Studie vorliegt.

**Frage 20      Häufigkeit bzw. Intensität der Intervention**

Hier wird unterschieden, ob es sich um ein einziges gegenüber mehreren Einzelereignissen (z. B. jährlicher Tag des Rauchens oder jährlich einmaliges Treffen mit Arzt) oder einer kontinuierlich durchgeführten Interventionen (z. B. tägliches Sportprogramm in der Schule) handelt. Da eine exakte Grenze zwischen einer Kategorie kontinuierlich versus multiple nicht festzulegen ist, werden diese beiden zusammen gefasst. Ggf. kann hier auch beides angekreuzt werden, wenn innerhalb einer kontinuierlichen Intervention z. B. ein Tag mit einer besonderen Veranstaltung hervorsteht.

**Frage 21      Zeitraum der kontinuierlichen bzw. multiplen Intervention**

Dieser Zeitraum bezieht sich auf die Intervention. Bei einmaligen auf einen Zeitpunkt bezogenen Interventionen wäre per definitionem 0 anzugeben. Die Angaben erfolgen in Monaten. Teile von Monaten sind auf Dezimalstellen umzurechnen, z. B. 10,5 Monate, wenn es 10½ Monate oder zehn Monate und 15 Tage sind.

**Frage 22      Dauer der Nachbeobachtung**

Dieser Zeitraum bezieht sich nur auf den Zeitraum nach Abschluss der Interventionen (unabhängig davon, wie lange die Intervention selbst gedauert hat). Es liegt also eine Beobachtung nach Ablauf eines weiteren Zeitraums vor. Auch bei einmaligen zeitpunktbezogenen Interventio-

nen beginnt die Nachbeobachtung erst in dem Moment, wo die erste Evaluation abgeschlossen ist. So ist ein erster Vergleich von Kontrolle und Intervention z. B. sechs Monate nach einer Fortbildung die Auswertung der Wirkung. Dann kann die Phase der Nachbeobachtung beginnen, wenn z. B. ein weiteres Jahr später beobachtet wird, ob sich der Effekt noch nachweisen lässt. Eine Nachbeobachtung soll erheben, ob der Effekt einer Intervention über den Zeitpunkt der (ersten) Evaluation hinaus anhält. Cave Sprache: Follow-up ist nicht identisch mit Nachbeobachtung, sondern kann im Englischen auch die Phase während einer dauerhaften Intervention nach der ersten Aufteilung in Interventions- und Kontrollgruppe meinen.

### **Frage 23      Effektmaße**

Es sollten nicht alle Effektmaße genannt werden, ggf. auf die fünf wichtigsten beschränken und ggf. die Zahl angeben, wie viele weitere Effektmaße erhoben wurden. Nach wie vor steht im Vordergrund, dass es Effektmaße sind, die nicht (nur) die Veränderung von Einstellungen etc. betreffen, sondern Intermediärparameter umfassen, die in enger Korrelation zu einem Ziel Krankheit/ Gesundheit stehen.

### **Frage 24      Art der Datenerhebung**

Hier soll insbesondere differenziert werden, ob es sich um sog. „objektive“ oder sog. „subjektive“ Daten handelt, mit denen der Effekt beschrieben wurde. Wenn z. B. die Veränderung des Bewegungsverhaltens mittels eines Pedometers erhoben wird, ist dies eine objektive Methode. Ein Fragebogen ist keine objektive Methode, wenn die Personen dabei gefragt werden, um wie viel sie ihre tägliche Wegstrecke gesteigert haben, da die Veränderung der Wegstrecke nicht objektiv gemessen worden ist.

### **Frage 25      Validierung der Messinstrumente**

In Abgrenzung der Vorarbeiten zur Intervention (s. Frage 44 und 45) geht es hier um die Frage, ob eine Validierung der Messinstrumente vorliegt, z. B. ein Pretest mit dem eingesetzten Fragebogen durchgeführt wurde. Die Validierung kann im Rahmen der vorliegenden Studie oder einer Vorstudie durch die Studiendurchführenden selbst erfolgen. Ebenso gilt es als validiertes Verfahren, wenn die Studiendurchführenden auf ein anderswo validiertes Instrument zurückgreifen. Dies muss explizit benannt werden. Grenzfälle wie das Problem, dass ein Fragebogen ggf. nur in einem anderen Land vorliegt, müssen je nach Einschätzung mit zum Teil oder mit nein beantwortet werden.

**Frage 26      Signifikanzniveau der Effektmaße**

Hier sollten die zahlenmäßigen Veränderungen eingetragen werden sowie deren p-Werte, z. B. nach folgendem Format: Reduktion der Anzahl der Raucher in der Interventionsgruppe von 22% auf 18%, in der Kontroll-gruppe von 22% auf 21% ( $p = 0,01$ ).

**Frage 27      Art der Interventionen**

Hier sollte die Intervention genau beschrieben werden, z. B. in folgendem Format: Veränderung des Schulessens zu höherem Anteil von Obst und Gemüse in Verbindung mit Aufklärung von Schülern zu Ernährung und schulweitem Bewegungsprogramm (15 Min. Hüpfen jeden Morgen).

**Frage 28      Art der Interventionsdurchführung**

Hier gilt es insbesondere neben den Interventionen „Information“ und „Beratung“ die drei Stufen der Aktivität zu differenzieren: „Aktion passiv“ beinhaltet z. B. eine Maßnahme für Demente, die z. B. mehr Gemüse erhalten, ohne dass die Dementen daran aktiv mitwirken. Eine „vorgegebene Aktivität“ meint, dass z. B. eine Schulklasse drei Mal in der Woche Sport machen muss, die Kinder also selbst aktiv sind, aber nicht darüber bestimmen, wie ihr Programm aussieht, sondern ein vorgegebenes Programm abarbeiten. Die Kategorie „freie Aktivität“ grenzt dazu die Möglichkeit ab, dass eine Intervention darin besteht, dass die Kinder dazu angeregt werden, sich mit ihrer Ernährung auseinanderzusetzen. Dabei gestalten die Kinder dann völlig frei, ob sie eher Informationen wollen oder selbst etwas kochen oder ggf. die Verbindung zur Bewegung ziehen, die ihnen auch fehlt. Information und Beratung sind nie scharf zu trennen, als Hilfe können aber die folgenden Fragen dienen: In der Beratung müssen die Fragen: Wo stehst du? Wo willst du hin? als Leitschnur dienen. Für eine Information braucht es dieses direkte Eingehen auf die Teilnehmenden nicht.

**Frage 29      Art der Information**

Es geht in dieser Frage darum, wie die Zielgruppe der Intervention mit Information erreicht worden ist.

**Frage 30      Setting der Beratung**

Auch hier ist eine Doppelnennung möglich, es sollte aber, wenn eine Form dominiert, auch diese angekreuzt werden, z. B. wenn mehrheitlich eine Gruppenberatung für Rauchende stattgefunden hat, aber sehr wenig Einzelgespräche, kann das als Gruppenberatung verzeichnet werden.

**Frage 31      Ansatz personale Ebene**

Diese Kategorien gelten nicht als disjunkt, es sollte aber versucht werden, einen dominierenden Ansatz zu benennen. So ist z. B. eine Studie, die über die Kirchengemeinde läuft, als Settingansatz zu begreifen, auch wenn Elemente von individueller Beratung eingeschlossen sind. Ein individuumsbezogener Ansatz dagegen würde nicht von einem Setting ausgehen.

### Frage 32 Durchführungsebene

Die Zuordnung erfolgt an Hand des Schemas von Rosenbrock (2004). Es gibt durchaus Interventionen die sowohl der Kategorie Verhalten als auch der Kategorie Verhältnis zugleich zugeordnet werden können. Auch hier gilt, bei einer Intervention, bei der deutlich entweder das Verhalten oder die Verhältnisse im Vordergrund stehen, sollte eine Antwort ausgewählt werden.

Anbei eine tabellarische Übersicht mit Interventionen, die zu den Fragen 31 und 32 als Hilfe dienen kann.

Tabelle 5: Beispiele für die Einteilung von Interventionen nach Ebene Individuum, Setting, Bevölkerung sowie nach Art des Ansatzes Verhalten und Verhältnis (in Anlehnung an Rosenbrock 2004)

Interventionsebene	Verhalten	Verhältnis
	Fokus: Information, Motivation, Aufklärung, Beratung ohne verhältnispräventive Elemente	Kontextbezug bzw. Weiterentwicklung von Rahmenbedingungen des Verhaltens
Individuum	Ärztliche Gesundheitsberatung für einen Patienten in der Praxis	Wiederkehrender „präventiver Hausbesuch“ mit Bezug/ Beratung zu Lebensbedingungen und Lebensstil, ein einmaliger Hausbesuch wäre wie die ärztliche Gesundheitsberatung als Verhaltensprävention einzustufen
Setting	Edukative Angebote in Schulen, Informationsangebote in Betrieben	Gesundheitsförderung in Schulen etc. durch Organisationsentwicklung, Schulhofgestaltung u.v.a. mehr
Bevölkerung	Motivationskampagnen ohne Kontextbezug (Esst mehr Obst etc.)	Anti-Tabak-Kampagne unter Einschluss verhältnispräventiver Elemente (Abbau von Automaten, Erhöhung Tabaksteuer etc.)

**Frage 33 Ort der Durchführung der Intervention**

Mehrfachnennungen sind möglich. Es soll der Ort genannt werden, an dem die Zielgruppe „erwischt“ wird. Mit Gemeinde ist die politische Einheit Kommune gemeint. Kirchengemeinde wäre der Kategorie Sonstiges zuzuordnen.

**Frage 34 Qualitätssicherung**

Wurden überhaupt Instrumente zur Qualitätssicherung eingesetzt.

**Frage 35 Instrumente der Qualitätssicherung**

Hier sollten ggf. weitere Instrumente genannt werden, so hat z. B. das finnische Zentrum für Gesundheitsförderung ein eigenes Instrument entwickelt.

**Frage 36 Alter der Zielgruppe überwiegend.**

Mehrfachantworten möglich. Auch hier ist keine scharfe Grenze zu ziehen, wenn die Teilnehmenden aus zwei der von uns gewählten Alterskategorien kommen. Eine ungefähre Grenze ist 10%, d.h. wenn z. B. ca. 8% auch aus einer älteren oder jüngeren Altersgruppe sind, kann das vernachlässigt werden, bei 15% nicht.

**Frage 37 Geschlecht der Zielgruppe**

**Frage 38 Ethnische Abstammung**

**Frage 39 Vorerkrankungen der Zielgruppe**

Diese Frage zielt auf die sekundäre Prävention ab, sofern sie nicht durch medikamentöse oder andere dem medizinischen Bereich zuzurechnende Verfahren aus unserer Untersuchung ausgenommen worden ist.

**Frage 40 Welche Vorerkrankungen**

**Frage 41 Spezifität des Ansatzes**

Genderspezifisch beispielsweise bedeutet nicht, dass entweder nur Frauen oder nur Männer in ein Projekt aufgenommen worden sind oder verblieben sind. Spezifische Belange von Frauen oder Männern werden berücksichtigt. Als Beispiel wäre eine Intervention zu nennen, die sich auf Basis der Lebenssituation 20- bis 40jähriger Frauen auch den Fragen widmet, wie diese Frauen ihre Zeit so einteilen können, dass sie sich neben Kindern, Haushalt und Beruf auch noch Zeit für Bewegung nehmen. Es geht weiterhin darum zu erkennen, ob ein Ansatz als kulturspezifisch dann ggf. nicht übertragbar ist, da er sich so stark an einer Kultur ausrichtet. Beispiel: Singen chinesischer Lieder als Präventionsmaßnahme. Zentral wäre dann nicht das Singen selbst, sondern die chinesischen Lieder. Religionsspezifisch z. B. heißt nicht, dass eine Intervention im Setting

Kirchengemeinde stattfindet: Die Kirchengemeinde ist zwar Zugangsweg, aber eine Ernährungsintervention könnte genau so gut den Zugangsweg Verein benutzen. Die zu stellene Kontrollfrage lautet: Kann die Intervention auch anderswo mit einer anderen Zielgruppe durchgeführt werden?

**Frage 42 Räumliche Ebene**

Mehrfachantworten möglich

**Frage 43 Repräsentativität der Zielgruppe**

Dies muss aus der Publikation hervorgehen. Eine eigene Recherche unsererseits kann nicht durchgeführt werden. Die Zielgruppe sollte in den USA z. B. sowohl weiße als auch schwarze und sog. hispanische Bevölkerungsgruppen umfassen.

**Frage 44 Wissenschaftliche Evidenz im Vorfeld**

Insbesondere der Pretest verweist hier auf einen Pretest zur Intervention, nicht zur Messmethode (s. oben Frage 25).

**Frage 45 Evidenz der berücksichtigten Literatur**

Frage 45 ist eine Unterfrage zur Antwortkategorie A (Literatur) der Frage 44. Literatur ist hier weit gefasst. Eine engere Fassung ist möglich (s. dazu mehr in Kapitel 7.5).

**Frage 46 Powerberechnung**

Die folgenden zwei Fragen beziehen sich auf kontrollierte bzw. nicht-kontrollierte Studien. Im zweiten Fall läuft die Frage 48 einfach leer.

Der Frage, dass Personen auf dem Weg von der Auswahl bis zur Auswertung verloren gegangen sind, wird durch Frage 49 Auswertungsverfahren Rechnung getragen.

**Frage 47 Probandenzahl Interventionsgruppe ausgewertet zum letzten Meßzeitpunkt**

Um einen Größenvergleich von Studien anzustellen wird die Probandenzahl mehrerer Interventionsarme zusammengefasst. Dies bezieht sich nicht auf die Zusammenfassung der Ergebnisse.

**Frage 48 Probandenzahl Kontrollgruppe ausgewertet**

**Frage 49 Auswertungsverfahren**

Im Fall, dass sowohl ITT als auch per protocol durchgeführt wurden, wird nur ITT angekreuzt.

**Frage 50 Statistische Methoden transparent beschrieben.**

Hier geht es nicht um die Bewertung der statistischen Methoden; dies kann ggf. nachgeholt werden, wenn die Studien in die Einzelbetrachtung der Positiv- bzw. Negativliste fallen.

**Frage 51 Subgruppenanalyse in Studienplanung angelegt**

Es reicht nicht, dass Subgruppenanalysen durchgeführt wurden, nur um im Nachhinein statistisch signifikante Effekte zu finden, sondern dies muss in der Studienplanung schon genannt worden sein.

**Frage 52 Zugrundeliegende Theorien und Konzepte der Intervention explizit benannt**

**Frage 53 Projektrelevante Voraussetzungen der Zielgruppe**

Gemeint sind Fähigkeiten, die zur Teilnahme an der Intervention unabdingbare Voraussetzung sind. Dies kann z.B. die Fähigkeit des Lesens sein, wenn die Information in der Verbreitung von Printmedien liegt. Ausdrücklich nicht gefragt ist an dieser Stelle nach der Selbstwirksamkeit, deren Veränderung ggf. mit einer Intervention erst erzielt werden soll (Siehe dazu Frage 55).

**Frage 54 Beteiligung der primären Zielgruppe bei der Projektentwicklung**

Unter primärer Zielgruppe verstehen wir die Teilnehmenden im Gegensatz zu den Multiplikatoren, die auch an der Intervention teilnehmen, aber nicht als Zielpopulation ausgewählt wurden, bei der etwas verändert werden soll. Die Kategorie nicht relevant ist anzukreuzen, wenn eine Intervention z. B. für Demenzerkrankte geplant wird, die nicht an der Konzeptentwicklung beteiligt sein können. Allerdings gilt es zu überlegen, in wie fern Angehörige dieser Demenzerkrankten partizipieren sollten. Dies wird in Zukunft als weitere Kategorie einzurichten sein (Beteiligung von sog. Proxies).

Die Zielgruppe muss auch die Möglichkeit haben/ gehabt haben, verändernd auf das Design der Intervention einzuwirken.

**Frage 55 Selbstwirksamkeit**

Der deutsche Begriff kann sowohl für den übergeordneten Begriff des Empowerment als auch für den engeren Begriff von self-efficacy stehen. Im weiten Sinne geht es hier nur darum, ob die Veränderung von Selbstwirksamkeit nicht nur bedacht, sondern auch in die Auswertung aufgenommen worden ist. Da es ein Qualitätskriterium ist, soll es nicht nur um den Input gehen (Strategie oder Elemente einer Strategie zur Stärkung der Selbstwirksamkeit), sondern eine Veränderung der Selbstwirksamkeit muss auch erhoben worden sein (Outcome). Über das Instrument der

Erhebung (Fragebogen, Beobachtung) ist damit noch nichts gesagt, auch wenn es zur Messung der self-efficacy im engeren Sinn ein festes Methodeninstrumentarium gibt.

**Frage 56      Beteiligung von Multiplikatoren an der Konzeptentwicklung**

**Frage 57      Beteiligung von Multiplikatoren an der aktuellen Durchführung**

**Frage 58      Qualifizierung der Multiplikatoren**

Es gibt Studientypen wie z. B. Massenkampagnen, bei denen die Beteiligung von Multiplikatoren nicht zutrifft. Dies wird über eine weitere Kategorie nicht zutreffend berücksichtigt. Damit erreichen diese Studien bei der Qualität Intervention eine geringere Maximalpunktzahl, sind somit nicht in der Bewertung der Qualität der Intervention benachteiligt.

**Frage 59      Analyse des Bedarfs**

**Frage 60      Messung des Bedarfs**

**Frage 61      Berücksichtigung des Kontexts**

Es sollten strukturelle und kulturelle Rahmenbedingungen (die Lebenswelt) berücksichtigt sein. Dies ist nicht identisch mit der Frage des Ansatzes. So kann eine Studie im Kontext der schwarzen Bevölkerung der USA angesiedelt sein und diese in ihrem Kirchengemeindekontext erreichen, aber der Ansatz kann dennoch auf andere Settings übertragbar sein, wäre also nicht kultur- bzw. religionsspezifisch.

**Frage 62      Verantwortungszuweisung während der Durchführung**

**Frage 63      Dokumentation des Prozesses**

**Frage 64      Evaluation des Prozesses**

Diese Frage geht über die Evaluation der Ergebnisse hinaus. Gefragt wird, ob auch das Prozessgeschehen einer Bewertung unterzogen wird und ob es ggf. zu Rückkopplungsprozessen im Laufe der Interventionsphase kommt.

**Frage 65      Bewertung der Intervention (durch Autoren der Studie) als effektiv**

Diese Frage ist losgelöst von der Frage 14 zu beantworten. Auch wenn weder eine klare Fragestellung noch ein klares Ziel benannt wurden, kann eine Intervention effektiv sein.

Allgemeines zu 66-68: In den folgenden Fragen ist von Organisationen die Rede. Damit sind abgegrenzte Settings gemeint, in denen interveniert wird. Eine Krankenkasse, die dauerhaft Schulungen implementiert, ist damit nicht gemeint, sondern die Veränderung muss z. B. am Ort des Geschehens zu einer Veränderung der Struktur (Gemeinde, Schule) führen.

**Frage 66      Veränderung der Organisation durch Intervention während der Intervention**

Dies ist zeitlich begrenzt auf die Dauer der Intervention. Ausschließlich nicht damit gemeint ist, wenn die Multiplikatoren beteiligt werden oder wenn Verantwortung für Teile des Projekts übernommen werden, es sei denn, sie werden dafür vom Fachunterricht freigestellt. Beide Punkte werden nicht als Veränderung der Organisationsstruktur angesehen.

**Frage 67      Dauerhafte Implementierung der Intervention in der Organisation**

Dies bezieht sich zeitlich auf den Zeitraum nach Beendigung der Intervention. Wenn es sich um ein einmaliges Geschehen handelt, wäre auch die Wiederholung eines Aktionstages in jedem kommenden Jahr eine dauerhafte Implementierung.

**Frage 68      Dauerhafte Implementierung der Intervention über die Organisation hinaus**

Die Intervention wird auch in anderen Schulen eines Landes z. B. umgesetzt.

## **4.4    Auswertung der qualitätswirksamen Fragen des IGKE-Kriterienkatalogs**

Alle in der gegenwärtigen Fassung des IGKE-Kriterienkatalogs qualitätswirksamen Fragen sind erkennbar an Hand der Kürzel QS sowie QI. Es werden jedoch nicht alle mit QI gekennzeichneten Fragen mit Qualitätspunkten bewertet. So ist Frage 35 als Erläuterung der Frage 34 zu verstehen, wird aber nicht bepunktet.

### **4.4.1 Die qualitätswirksamen Fragen und ihre Bepunktung**

Frage 13: ja = 1, nein etc. = 0.

Frage 16 (betrifft nur RCTs): ja und an anderer Stelle = 1, nein = 0.

Frage 17: Führt bei RCTs bei Frage 16 zum Punktabzug.

Frage 19: ja = 1, nein und nicht erkennbar = 0, nicht zutreffend bedenken, da nur für CTs (einschließlich RCTs) anwendbar. Frage 25: ja = 1, zum Teil = 0,5, nein = 0.

Frage 34: ja = 1, nein = 0.

Frage 44: Literatur, Pretest, anderes Projekt = 1, unklar = 0.

Maximal kann 1 Punkt erreicht werden, selbst wenn neben Literatur auch z. B. ein Pretest vorliegt.

Frage 46: ja = 1, nein = 0.

Frage 49: ITT = 1, Per Protocol = 0,5, keine Angaben = 0.

Maximal kann 1 Punkt erreicht werden, wenn ITT und per Protokoll analysiert wurde.

Frage 50: ja = 1, nein = 0.

Frage 54: ja = 1, nein und keine Angaben = 0, nicht anwendbar in den sehr seltenen Fällen als Reduktion des Maximalscores. Dies wurde bisher in keiner der von uns ausgewerteten Studien (mehr als 120) angewendet.

Frage 55: ja = 1, ein und keine Angaben = 0.

Frage 56: ja = 0,5, nein und keine Angaben = 0.

Frage 57: Ja = 0,5, nein und keine Angaben = 0.

Frage 58: Ja = 0,5, nein und keine Angaben = 0.

Wenn eine Intervention keine Multiplikatoren benötigt, dann trifft die Kategorie nicht zutreffend zu. Der Maximalscore wird um insgesamt 1,5 Punkte reduziert.

Frage 59: Nach Studienlage und eigene Berechnungen = 1, kein Angaben = 0.

Maximal kann 1 Punkt erreicht werden, selbst wenn beide Möglichkeiten vorliegen.

Frage 61: ja = 1, nein = 0.

Frage 62: ja = 1, nein und keine Angaben = 0.

Frage 63: ja = 1, nein und keine Angaben = 0.

Frage 64: ja = 1, nein = 0.

## 4.4.2 Die Scorebildung

### Score Qualität Studien(design)

Zur Ermittlung des Qualitätsscores der Studiendesigns werden die Fragen 13, 16/17, 19, 25, 46, 49 und 50 zusammen gefasst.

Es können für RCTs maximal 7/7, für CTs maximal 6/6 und für Nicht-CTs maximal 5/5 Punkte erreicht werden. Die tatsächlich erreichten Punkte werden dann als Anteile von 100% berechnet.

Da die Fragen 14 als inhaltliche Bestimmung von 13 zu gelten hat, wird dafür kein Punkt vergeben.

### **Score Qualität Intervention**

Zur Ermittlung des Qualitätsscores der Interventionsqualität werden die Fragen 33, 44, 54-59 und 61-64 zusammen gefasst.

Es können maximal 10,5/ 10,5 erreicht werden. Dies gilt für nahezu alle Studien, da nur in seltenen Fällen die Frage 54 mit Bezug auf die primäre Zielgruppe nicht anwendbar ist.

Da die Frage 45 als inhaltliche Bestimmung von Frage 44 zu gelten hat, wird dafür kein Punkt vergeben.

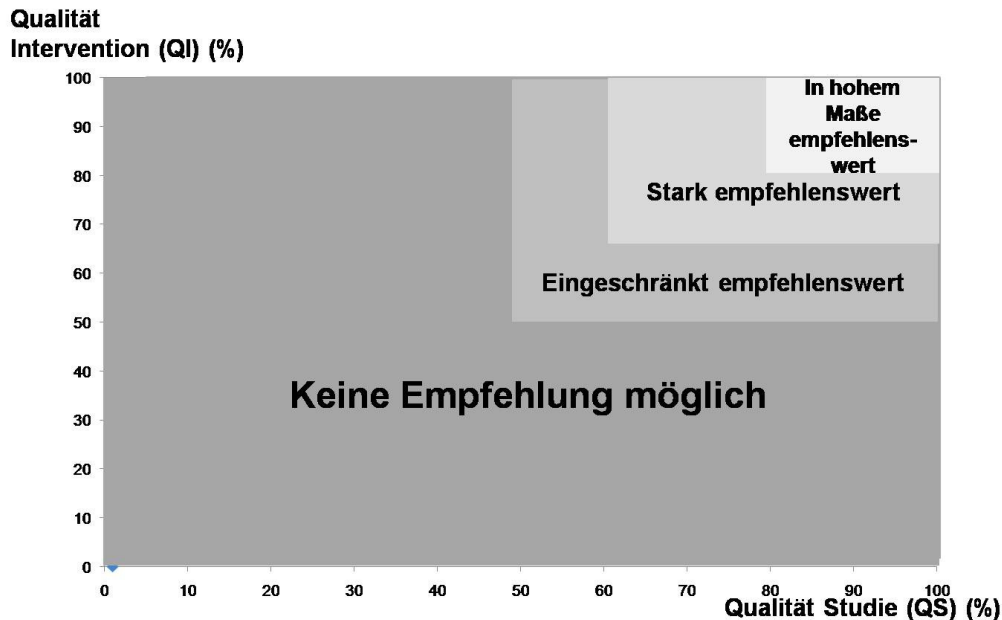
Für alle Studien, bei denen Multiplikatoren nicht eingesetzt werden müssen, ist der Maximalscore 9/9.

## **4.5 Die Umsetzung in Empfehlungen**

Nachdem alle Studien zu einem Handlungsfeld nach Qualität der Studiendurchführung und Qualität der Intervention bewertet wurden und daraus Scores errechnet wurden, mussten sich aus diesen Werten Empfehlungen ableiten lassen. Wie zuvor gezeigt wurde, ist dies ein übliches Verfahren der evidenzbasierten Medizin, im Expertenkonsens auf Grundlage der Bewertung von Studien Empfehlungen zur Umsetzung auszusprechen.

Wir entschieden uns für das folgende Vorgehen: Studien mit sehr hoher Qualität sollten eine Art absolute Empfehlung – nicht zu verwechseln mit einem Automatismus – erhalten gegenüber den Studien, die weniger Qualitätspunkte erreichten. Da ein Schnitt gefunden werden musste, wurden vier Empfehlungsgrade differenziert. Dabei wurden zumindest teilweise effektive Interventionen, die mehr als 80% der Punkte in den beiden Dimensionen Qualität der Studiendurchführung und Qualität der Intervention erzielten, als „in hohem Maße empfehlenswert“ eingestuft. Interventionen deren Qualität einen Wert von 60% erreichte, wurden als „stark empfehlenswert“ bezeichnet. Interventionen die noch über dem Durchschnittswert von der Qualität der Studiendurchführung und der Qualität der Intervention lagen, konnten nur „eingeschränkt empfohlen“ werden. Alle darunterliegenden Interventionen wurden „nicht weiter in die Empfehlung“ aufgenommen.

Abbildung 8: Empfehlungsgrade nach der Bewertung der Studien



Legende: Aufgetragen wird die Qualität der Studiendurchführung gegen die Qualität der Intervention. Damit lässt sich erkennen, in welches Feld die Studien fallen. Die Grenzen der beiden höchsten Empfehlungsgrade (rechts oben) wurden auf 80% bzw. 60% in beiden Dimensionen konsentiert. Die dritte Grenze ist variabel und ergibt sich aus dem Mittelwert beider Qualitätsdimensionen. Daher erklärt sich die große L-förmige Fläche des Bereichs nicht empfehlenswerter Studien.

## 4.6 Weiterentwicklungen des IGKE-Kriterienkatalogs

Aus der Anwendung des IGKE-Kriterienkataloges zur Evaluation von Interventionen heraus haben sich Aspekte ergeben, wie der Katalog weiter entwickelt und an verschiedene Studien- wie Interventionsdesigns adaptiert werden kann.

In Zukunft muss neben der qualitativen Begleitforschung (Frage 18) auch die Berücksichtigung quantitativer wie gleichermaßen qualitativer Befragungen von Zielgruppen oder Stellvertretern der Zielgruppe zu Hemmnissen oder fördernden Faktoren im Rahmen einer Intervention in die Evaluation aufgenommen werden. Dazu sollten zwei Fragen gewählt werden, da unsere Frage 18 zum Studiendesign gehört, die von uns vorgeschlagene Frage aber eher zu den Inhalten der Intervention. Diese Fragestellung sollte auch getrennt von der Frage nach der Analyse projektrelevanter Voraussetzungen wie der Beteiligung der Zielgruppe an der Entwicklung des Konzepts (im Sinne einer Partizipation) (Fragen 53 und 54) berücksichtigt werden.

Derzeit wurde die Frage 44 in Verbindung mit der Frage 45 sehr weit gefasst, um nicht per se innovationsfeindlich zu sein. Es besteht die Möglichkeit, insbesondere die Antwortkategorien A-C der Frage 45 höher zu bewerten als die weiteren Antwortkategorien, wenn es Ziel ist, hier die Hierarchisierung der EbM auch für die schon einer zu bewertenden Intervention vorausgehenden Datenlage einzusetzen. Dies kann sinnvoll sein, um die beiden Klippen der Innovationsfeindlichkeit wie der Schwäche vorausgehender Evidenz gleichzeitig zu umschiffen. Das heißt, streng genommen dürften Pretests oder auch Interventionen, die sich also solche zu erkennen geben, nicht berücksichtigt werden, sondern müssten einer von diesem Evaluationsverfahren gesonderten Bewertung unterzogen werden. Somit wären sie per se nicht benachteiligt, wenn sie innovativ und damit ohne „Evidenz“ in Literatur, Pretest oder anderem Projekt sind. Umgekehrt darf es aber auch nicht geschehen, dass unter dem Deckmantel der Innovation versäumt wird, einen Pretest zu machen.

Da es in einigen Studien mehrarmige Interventionen gab, sollten die Fragen 47 und 48 erweitert werden. Weil es Studien gibt, wie z. B. eine Studie, die zu zwei Zeitpunkten an unterschiedlichen Gruppen gleichen Alters den Effekt einer Intervention messen will, wäre hier eine Auswertung nach intention to treat oder per protocol nicht möglich. Um diese Studien nicht in der Auswertung zu benachteiligen, wird eine Kategorie nicht zutreffend zu ergänzen sein, so dass diese Studien einen um diesen Punkt niedrigeren Maximalscore im Bereich Qualität Intervention erreichen können.

In Zukunft sollte differenziert werden, ob genuin<sup>2</sup> im Setting verortete Multiplikatoren, z. B. Lehrer in der Schule, oder andere Multiplikatoren wie z. B. speziell ausgebildete Schwestern oder Psychologiestudierende bei manchen Interventionen im Bereich Depression eingesetzt wurden. Die Debatte kann hier inhaltlich pro und contra Lehrer geführt werden. Wichtig ist nur, dass die Multiplikatoren nach ihrem Verhaftetsein im Setting unterschieden werden.

Die Kategorie teils teils in der Frage 65 gereicht einer Studie zunächst nicht zum Schaden, sondern führt dazu, dass diese Studie von einem Expertengremium auf die Wichtigkeit der verschiedenen Effektmaße oder der Widersprüche zwischen verschiedenen Gruppen hin beurteilt werden muss.

Das Screeninginstrument kann mit seiner Analyse immer nur zu einer Vorauswahl an Interventionen führen. Diese Vorauswahl muss dann unserer Ansicht nach durch ein Expertengremium bewertet werden, wenn sich zwischen Studien oder innerhalb von Studien (s. vorherigen Ab-

schnitt) Inkonsistenzen ergeben. Denn auch wenn einzelne Studien den Bereich der Empfehlbarkeit erreichen können, müssen die Widersprüche zwischen Studienergebnissen ggf. auch durch eine klassische Metaanalyse beigelegt werden, wenn dies möglich ist, oder durch eine genauere Analyse der Intervention ausgeräumt werden. Eine genauere Betrachtung verschiedener Intervention kann möglicherweise zu einer Erklärung abweichender Ergebnisse beitragen.

## **5 Methodisches Vorgehen**

Der empirische Teil der vorliegenden Arbeit besteht aus zwei Teilen: Einer systematischen Literaturrecherche und der Entwicklung eines Kriterienkatalogs zur Evaluation von Studien. Kapitel 3.1 erläutert das Vorgehen der systematischen Datenbankrecherche. Kapitel 3.2 beschreibt die Entstehung des IGKE-Kriterienkatalogs.

### **5.1 Systematische Literatur- und Datenbankrecherche**

#### **5.1.1 Auswahl der Länder**

Für die systematische Literaturrecherche wurden dreizehn Länder ausgewählt, die einen mit Deutschland vergleichbaren sozioökonomischen Status aufweisen. Interventionen in Prävention und Gesundheitsförderung der folgenden Länder wurden gesichtet und evaluiert: USA, Kanada, Australien, Neuseeland, Großbritannien, Frankreich, Niederlande, Dänemark, Finnland, Schweden, Norwegen, Österreich und Schweiz. Gegenstand des Projektes waren Interventionen der Prävention und Gesundheitsförderung für die Risikofaktoren Ernährung, Bewegung, Rauchen und Depression. Dabei wird Prävention umfassend verstanden, d.h. die Untersuchung konzentrierte sich nicht explizit auf Primär-, Sekundär oder Tertiärprävention.

#### **5.1.2 Recherchevorgehen**

Ziel der systematischen Literatur- bzw. Datenbankrecherche war es, alle relevanten internationalen Studien zu sichten und nachfolgend zu evaluieren. Es wurde in acht Datenbanken recherchiert. Innerhalb der Datenbanken bestehen Unterschiede in der Art der Verschlagwortung bzw. der Suchmöglichkeiten. Dieses machte eine separate Recherche je Datenbank erforderlich, um auf die Besonderheiten eingehen zu können.

## **Auswahl der Suchmaschinen für die Datenbankrecherche**

Prinzipiell sind Datenbankrecherchen über verschiedene Suchmaschinen wie über das Deutsche Institut für Medizinische Dokumentation und Information (DIMDI) oder über die Suchmaschine „Medpilot“ der Deutschen Zentralbibliothek für Medizin (ZBmed) möglich. Es wurde entschieden, für sieben Datenbanken (Medline 90, Amed, Biosis, Embase 90, Global Health, Somed und Psychinfo) über das DIMDI zu recherchieren. Lediglich für die Datenbank CINAHL, die über das DIMDI nicht abzurufen ist, erfolgte der Zugang über das Datenbank-Informationssystem. Darüber hinaus wurden ergänzend Recherchen im Internet, z.B. Google, Google Scholar, etc. und per Hand durchgeführt.

## **Suchstrategie für die Datenbankrecherche**

Nicht alle Datenbanken sind nach dem gleichen Thesaurus verschlagwortet. Auch sind nicht alle Begriffe als Schlagwort (Controlled Term = CT) gelistet, so dass alternativ für diese Begriffe Volltext-Recherchen durchgeführt werden mussten. Da aber gerade die großen Datenbanken bei der Volltext-Recherche unverhältnismäßig hohe Trefferzahlen lieferten, wurde entschieden, die Recherche in diesen Datenbanken weitgehend mit Controlled Terms durchzuführen, zumal die Suche mit Controlled Terms sich durch eine höhere Präzision und einen höheren Recall der Ergebnisse auszeichnet. Unter Präzision wird der Anteil der gefundenen relevanten Dokumente aus allen gefundenen Dokumenten verstanden, sprich wie treffgenau ist die Suche; unter Recall ist der Anteil der gefundenen relevanten Dokumente aus allen relevanten Dokumenten gemeint. So wurde z.B. für Medline 90 nur der Ländername als FT recherchiert, da nicht davon auszugehen ist, dass regelmäßig das Land, in dem eine Intervention durchgeführt wurde, als Controlled Term genannt wird. Die Synonyme für Prävention bzw. die Risikofaktoren wurden dagegen als Controlled Term recherchiert.

Mit dem Ziel, eine möglichst umfassende Recherche durchzuführen, die aber gleichzeitig noch handhabbar sein muss, wurde die Suchstrategie für die einzelnen Datenbanken konkretisiert. Die Suchstrategie wurde anhand der MeSH-Schlagwortbäume komplettiert. Sowohl für die jeweiligen Risikofaktoren (smoking, exercise etc.) als auch für jedes Land und für jeden Risikofaktor vorgesehenen Suchbegriffe „prevention“ and „health promotion“ wurde nach (weiteren) Synonymen

gesucht. Die Begriffe „health education“ and „community health services“, die beide im MeSH-Datenbaum als Schlagwort erfasst sind, wurden als Suchbegriffe aufgenommen.

Andere Thesauri, z.B. für SOMED (lögd), und Embase (M-Tree) konnten nicht identifiziert werden. Die o.g. Suchbegriffe wurden daher auch auf Datenbanken mit anderen Thesauri übertragen. Um die Untersuchung handhabbar zu halten, konzentrierte sich die Analyse auf Studienergebnisse seit 1990, zumal sich gesellschaftliche und politische Strukturen gegenüber Zeiträumen davor stark verändert haben.

Konkret wurde folgende Suchstrategie für die verschiedenen über das DIMDI zu recherchierenden Datenbanken festgelegt (Tabelle 2):

Tabelle 2: Vorgehensweise der systematischen Suchstrategie über das „Deutsches Institut für Medizinische Dokumentation und Information“ (DIMDI)

Arbeitsschritte		Operationalisierung
# 1	Auswahl der Datenbank	<u>Datenbanken</u> : Amed, Medline 90, Medline Alert, SOMED, Global Health, Biosis, Embase, Embase Alert
# 2	Filtereingabe: Zeitraum 1990 - 2006	
# 3, verknüpft mit Boolesche Operator <b>OR</b>	Ländernamen bzw. Ländernamen	<b>14 ausgewählte Länder**</b>
# 4, verknüpft mit Boolesche Operator <b>OR</b>	Schlüsselwörter „Prävention und Gesundheitsförderung“***	Prevention?, Health Promotion, Community Health Services, Health
# 5, verknüpft mit Boolesche Operator <b>OR</b>	Schlüsselwörter „Risikofaktor“***	<u>Bewegung</u> : exercice, physical education, sports <u>Depression</u> : depression, depressive disorder, bipolar disorder <u>Ernährung</u> : diet

		<u>Rauchen</u> : smoking, smoking cessation
#6, # 3, # 4, # 5 verknüpft mit Boolesche Operator <b>AND</b>	Trefferzahl	
# 7	Selektion nach Titel, Abstract/Zusammenfassung	

Legende: \* mehrere Länder werden mit zwei oder drei Ländernamen gesucht; je nach Datenbank werden Ländernamen als Textfeld oder Schlagwort recherchiert.

\*\*\* United States/USA, Canada, Australia, New Zealand, Great Britain/United Kingdom, France, Netherlands/Niederland, Finland/Suomi, Sweden/Sverige, Norge/Norway, Danmark/Denmark, Austria/Oesterreich, Switzerland/Suisse/Schweiz, Deutschland/Germany.

\*\* je nach Datenbank werden Schlüsselwörter als Textfeld oder Schlagwort recherchiert.

Boolesche Operatoren: **OR**, findet alle Dokumente, die entweder das eine oder das andere Wort beinhalten; **AND** findet alle Dokumente, die beide Wörter beinhalten.

### **Ergänzende Literatur-Recherche**

Neben dieser systematischen Datenbankrecherche wurden mit Prävention befasste Institutionen in den verschiedenen Ländern angeschrieben und um Hinweise auf erfolgreiche Projekte und Maßnahmen der Prävention gebeten. Als Feedback wurden zwar Hinweise auf nationale Präventionsstrukturen und Gesetzgebung im Bereich der Prävention gegeben, konkrete Hinweise auf einzelne Präventionsprojekte blieben überwiegend aus.

Darüber hinaus wurde über das Internet (Google bzw. Scholar google) im Schneeballsystem anhand der Literaturangaben der identifizierten Studien, aber auch durch das Durchforsten von Publikationslisten nationaler und internationaler Institutionen nach weiteren Studien gesucht. Die Suche im Schneeballsystem erfolgte nur begleitend zur Analyse der bereits identifizierten Projekte.

### **Ein-und Ausschlusskriterien**

Über die Literaturrecherche hinaus wurde ein Katalog von Ein- und Ausschlusskriterien entwickelt, welcher als weiterer „Filter“ herangezogen wurde, um aus den in der systematischen Datenbankrecherche angezeigten „Treffern“ die Projekte zu identifizieren, die in die weitere Analyse einbezogen werden konnten. Die Entscheidung für die Ein- und Ausschlusskriterien erfolgte dabei u.a. vor dem Ziel, mit dem Projekt eine Datenbasis für potenzielle Veränderungen im Ge-

sundheitszustand der deutschen Bevölkerung durch Prävention und Gesundheitsförderung zu erhalten.

Die Suchkriterien (Ein- und Ausschlusskriterien, s.u.) waren möglichst weit gefasst. Berücksichtigt wurden alle Maßnahmen mit Interventionen, ganz gleich auf welcher Ebene (Individuum, Setting, Bevölkerung) diese durchgeführt wurden. Gesucht wurde nur nach Interventionen, die effectiveness untersuchten (vs. efficacy). D.h. nicht, dass Projekte die als RCT durchgeführt wurden, automatisch ausgenommen wurden. Denn je nach Ausgestaltung wird auch in RCTs die effectiveness gemessen. Werden unter anderem Vereine, Schulen oder Kindergärten mit dem gleichen Interventionsansatz als Settings nach vorangehender Randomisierung Interventions- und Kontrollgruppen zugewiesen, so handelt es sich um ein RCT, der effectiveness untersucht. Es wurden Interventionen, die sowohl effektiv als auch nicht effektiv waren, eingeschlossen.

Verworfen wurden dagegen unter anderen folgende Interventionen:

- alle präventive pharmakotherapeutische Interventionen sowie diagnostische Maßnahmen (Screening),
- Interventionen, die vornehmlich anderen Risikofaktoren dienen (z. B. Bewegung zur Prävention von Osteoporose),
- Studien, die nur die efficacy, aber nicht die effectiveness untersuchen,
- Studien zur Prävention des Plötzlicher Kindstodes (SIDS),
- Studien zur HIV/ AIDS-Prävention,
- Nicht evaluierte Studien,
- Projekte die auf Veränderung allein bei den Stakeholdern/ Multiplikatoren zielen.

Aus allen gesammelten „Treffern“ wurde anhand des Titels zunächst eine Vorauswahl vorgenommen. Für ausgewählte Titel wurde das Abstract bzw. direkt der Volltext bestellt. Angesichts der hohen Trefferzahl wurden nicht alle Titel von zwei Personen gesichtet. Um aber in Zweifelsfällen eine Zweitmeinung zu berücksichtigen, wurden in diesen Fällen die Abstracts von einer zweiten Person gesichtet, um dann gemeinsam über die Aufnahme in die weitere Analyse zu entscheiden. Spätestens durch die Lektüre der Artikel wurde dann ersichtlich, ob die aufgestellten Ein- bzw. Ausschlusskriterien für den Einbezug in die Analyse der Übertragbarkeit erfüllt sind. D.h., dass Studien, bei denen z.B. keine Intervention erfolgte oder keine Evaluation durchgeführt wurde, nicht in die weitere Analyse einbezogen wurden.

## 5.2 Der IGKE-Kriterienkatalog

Der IGKE Kriterienkatalog wurde im Zeitraum von Juli 2006 bis Januar 2007 von den Autoren entwickelt. Er basiert auf sechs Qualitätsinstrumenten, die in diesem Zeitraum verfügbar waren: PREFFI, IDM, QSP, Quint-essenz, Closing the Gap, EQUIHO.

Damit berücksichtigt der Kriterienkatalog die gegenwärtig national und international anerkannten Qualitätskriterien für eine erfolgreiche Prävention und Gesundheitsförderung. Ziel bei der Konstruktion des Kriterienkataloges war es, neben der Qualität auch die Wirksamkeit zu berücksichtigen.

## 6 Bewegung

In diesem Kapitel werden zunächst Daten zur Prävalenz des Bewegungsmangels sowie der daraus resultierenden Folgen, also Erkrankungen und deren Kosten, dargestellt. Dann folgt eine kurze Übersicht über den Forschungsstand zu Interventionen im Bereich Bewegung an Hand von systematischen Reviews. Im Weiteren werden Interventionen zu Bewegung, die ausschließlich Frauen und Mädchen einschließen, sowie Interventionen auf der Ebene von Betrieben nach der Qualitätsbewertung des Kriterienkatalogs vorgestellt und auf ihre Übertragbarkeit hin bewertet.

### 6.1 Stand der Forschung

#### Epidemiologische Daten zur Bewegung in Europa und in Deutschland

62,6% der Bevölkerung erfüllen im gewichteten Durchschnitt in der Europäischen Union die Bewegungsempfehlungen, d. i. 14 MET-Stunden<sup>8</sup> oder mehr in den letzten sieben Tagen vor ihrer Befragung. Andere Untersuchungen gehen eher davon aus, dass 2/3 der Bevölkerung in der Europäischen Union die den Empfehlungen entsprechende Zeit in Bewegung/ Woche nicht erreicht.

Nach Daten des Robert Koch-Instituts (vgl. auch Mensink 2003 zu Verteilung nach Altersgruppen) ergibt sich die folgende Verteilung der Bewegung in der deutschen Bevölkerung (Tabelle 6).

Tabelle 6: Sport und Bewegung in der deutschen Bevölkerung

Sport/ Bewegung	Männer	Frauen
Gar nicht	37,3%	38,4%
< 2h/ Woche	20,9%	28,4%
> 2h/ Woche	41,7%	33,2%

In der Altersgruppe von 18 bis 29jährigen geben 75% an, sich an sportlichen Aktivitäten zu beteiligen. In höheren Altersgruppen nimmt dieser Prozentsatz ab auf ca. 55 bis 65% bei den 30 bis 69-Jährigen und schließlich auf 45% der über 70jährigen.

---

<sup>8</sup> Mit der metabolischen Einheit MET (=metabolic equivalent) lässt sich der Energieverbrauch verschiedener Tätigkeiten und Bewegungsformen einfach ermitteln. Es gilt die Formel: 1 MET in kcal/ min. = 1 kcal Energieverbrauch pro kg Körpergewicht und Stunde an. Bezugspunkt mit einem Energieverbrauch von 1 MET ist das Sitzen.

Neueste Daten für die Lage der Kinder und Jugendlichen wurden im September 2006 auf Basis der KiGGS-Daten (= Studie zur Gesundheit von Kinder und Jugendlichen in Deutschland, <http://www.kiggs.de/>) vorgestellt (Lampert et al. 2006). 75% der Kinder zwischen drei und zehn Jahren spielen demnach täglich im Freien. Mehr als 50% dieser Altersgruppe sind in einem Sportverein aktiv. Es zeigt sich aber, dass Kinder aus Familien mit sozioökonomisch niedrigem Status oder Migrationshintergrund zwei bis drei Mal seltener Sport ausüben. Für die Jugendlichen ergibt sich, dass 54% mindestens drei Mal in der Woche sportlich aktiv sind, 23% täglich. Jungen verbringen 8, Mädchen 4,5 Stunden pro Woche bei sportlicher Aktivität. Die Unterschiede nach Sozialstatus und Migrationshintergrund bestehen in dieser Altersgruppe nur für Mädchen (Lampert et al. 2006).

### **Die Folgen mangelnder Bewegung**

Als Folgen mangelnder Bewegung sind Adipositas, Diabetes mellitus Typ 2, koronare Herzkrankheit und Osteoporose belegt. Beobachtungsstudien haben schon in den 90er Jahren den Zusammenhang eines bewegungsarmen Lebensstils mit der Genese eines Typ 2 Diabetes (Johnson and Ballin 1996) erhärtet.

Seit 1984 hat in den USA die Zahl der Todesfälle durch kardiovaskuläre Erkrankungen bei Frauen die bei Männern übertroffen (American Heart Association 2002, 2003), insbesondere auch bei sozioökonomisch schwachen und ethnischen Minderheiten. Für Großbritannien werden 37% der Krankheitslast durch koronare Herzerkrankung auf mangelnde Bewegung zurück geführt (British Heart Foundation 2003).

Auch wenn es kaum Längsschnittstudien gibt, so ist belegt, dass wenig oder nicht aktive Kinder und Jugendliche dieses Verhalten als Erwachsene tendenziell beibehalten (Malina 1996).

Aktuell kursieren unterschiedliche Zahlen zum Anteil der Übergewichtigen und Adipösen in Deutschland. Daraus ergeben sich auch unterschiedliche Rangpositionen im Vergleich zu den anderen europäischen Ländern. Aktuell geht man etwa davon aus, dass unter den Erwachsenen in Deutschland 55% der Frauen und 65% der Männer übergewichtig sind, was im europäischen Vergleich nach Angaben der WHO etwa ein Platz im Mittelfeld bedeutet (Anteil der Adipösen in Europa nach Ländern: Frauen 32 bis 79% , Männer 28 bis 78%).

In Deutschland hat die Adipositas auch bei Kindern zugenommen. Wie die Erhebung des Kinder- und Jugendgesundheits surveys des Robert Koch-Institutes (RKI) aus den Jahren 2004 und

2005 zeigt, sind 15% der Kinder zwischen drei und 17 Jahren übergewichtig. 6,3% sind sogar adipös (Kurth und Schaffrath Rosario 2007). Das macht insgesamt 1,9 Mio. Kinder und Jugendliche, die übergewichtig sind, 800.000, die adipös sind. Auch hier zeigt sich das Muster wie in anderen Ländern, dass Kinder aus sozial benachteiligten Familien oder aus Migrantenfamilien stärker betroffen sind.

### **Abschätzung von Kosten durch mangelnde Bewegung**

Eine exakte Abschätzung ist schwierig. Es kann aber davon ausgegangen werden, dass ein Teil der jährlichen Kosten für die Behandlung von Herz-Kreislauf-Erkrankungen von 35 Mrd. € sowie für Erkrankungen des Muskel- und Skelettsystems von 25 Mrd. € durch Bewegungsmangel zustande kommt (vgl. Bundesregierung 2007). Eine neue Studie aus Großbritannien geht davon aus, dass im Jahr 2002 3% der sog. DALYs (disability adjusted life years) auf das Konto mangelnder Bewegung gehen, was 1,06 Mio. £ an direkten Kosten im National Health System verursacht hat (Allender et al. 2007). Anderson et al. (2005) haben abgeschätzt, dass die Risikofaktoren Bewegungsmangel, Ernährung und Übergewicht zu 27% der nationalen Gesundheitsausgaben in den USA betragen.

### **Wissenschaftliche Datenlage zum Einfluss von Bewegung auf die Krankheitslast**

Aus dem dargelegten epidemiologischen Zusammenhang von mangelnder Bewegung und Genese ausgewählter Krankheiten ergibt sich, dass ein Mehr an Bewegung die Entstehung von Zuständen wie Adipositas und Krankheiten wie Diabetes mellitus Typ 2 oder koronare Herzkrankheit verhindern kann. Dies wird auch in internationalen Studien zur „efficacy“, also der generellen Wirksamkeit unabhängig von Alltagsbedingungen, belegt: Für den Diabetes mellitus Typ 2 kann eine Reduktion um bis zu 58% des Auftretens erreicht werden (s. Tuomilehto et al. 2001, Knowler et al. 2002). Auch Hu et al. (1999) bestätigen, dass forsches Gehen für 30 Min. pro Tag eine Reduktion um 30-40% bewirkt, an einem Diabetes mellitus Typ 2 zu erkranken. Kohl et al. (1992) haben schon vor fast 20 Jahren postuliert, dass schon „some level of regular exercise“ zur Vermeidung von koronarer Herzkrankheit und Schlaganfall förderlich sind. Auch für Kinder ist belegt, dass der allgemeine Gesundheitszustand bei Bewegung besser ist (Riddoch 1998). Je nach Altersgruppe sind außerdem positive Effekte auf das psychische Wohlbefinden, den Selbstwert, Depression und Angst belegt. Auch kann Bewegung zu besserem Sozial- sowie Lernverhalten

und Konzentrationsfähigkeit beitragen (Calfas und Taylor 1994, Gruber 1986, Mutrie und Parfitt 1998).

Insbesondere Manson et al. (1996) haben mit der Nurses' Health Study mit einer Gesamtzahl von 72.488 eingeschlossenen Frauen nachgewiesen, wie durch Bewegung das Risiko für eine koronare Herzkrankheit gesenkt werden kann. Dies kann auch adjustiert für Alter, Rauchen, Body-Mass-Index und weitere Kovariablen statistisch belegt werden. Konkret konnte das Quintil mit dem höchsten Energieverbrauch durch Bewegung gegenüber dem mit dem geringsten Energieverbrauch das Risiko auf etwa 2/3 reduzieren bei einer wöchentlichen Gesamtzeit von drei oder mehr Stunden forschen Gehens. Diese Reduktion ist gleichwertig der durch regelmäßige intensive sportliche Bewegung erreichbaren Reduktion von 30 bis 40% des Risikos gegenüber sich nicht bewegenden Frauen gleicher Altersstufen. Ähnliche Ergebnisse liegen auch bei Andersen et al. (2000) aus Dänemark vor. Es wurden über insgesamt 14,5 Jahren in einem prospektiven Studiendesign 30.600 Personen im Alter von 20 bis zu 93 Jahren in Kopenhagen eingeschlossen. Es konnte eine um 40% niedrigere Gesamtmortalität unter denen beobachtet werden, die mit dem Fahrrad zur Arbeit fahren. Dies konnte bei Männern und Frauen gleichermaßen über alle untersuchten Altersgruppen erhoben werden.

### **Wie viel Bewegung ist nötig?**

Zunächst war lange gefordert worden, sich „vigorous“ zu bewegen. Mitte der 90er Jahre hat sich dann ein Wechsel abgezeichnet, dass auch Bewegung im Alltagsablauf angerechnet wurde. Diese Entwicklung geht zurück auf einen Beitrag von Pate et al. (1995), in dem es heißt: „moderate „accumulated“ physical activity ist zu berücksichtigen. Es wird empfohlen, dass regelmäßige moderate Bewegung „preventive and health-enhancing“ sei (US Department of Health and Human Services 1996). Zu den anzurechnenden Alltagsaktivitäten gehören Treppensteigen, Fahrradfahren, Rasenmähen, Fensterputzen oder Schneeschippen (vgl. z. B. Martin et al. 2001).

Weiterhin muss man sich nicht mehr am Stück bewegen, sondern es dürfen Bewegungsepisoden von 10 Minuten addiert werden. Darüber hinaus wird gefordert, dass man drei Mal pro Woche 20 bis 60 Minuten leichtes Schwitzen und beschleunigtes Atmen erreichen soll, dabei aber noch sprechen können sollte (vgl. u. a. Martin et al. 2001).

Derzeit kann als state of the art folgende zusammenfassende Empfehlung auf Basis der „efficacy“ nach Rutten (2006) gelten: “The concept of health enhancing physical activity (HEPA) ...

expanded the context of processes to be considered. In particular, the contexts of work and transportation have been added to the traditional context of leisure time (sport) activities. Accordingly, public health recommendations shifted towards an accumulated 30 minutes of almost all kind of moderate physical activity ....” Dies entspricht auch den Aussagen der WHO (2002) im European Health Report. Oder kurz gefasst: “Regular walking provides significant health benefits” (Oguma und Shinoda-Tagawa 2004).

### **Welche Evidenz gibt es für konkrete Interventionen international und national**

Für die Zeit vor Mitte der 90er Jahre verweisen wir auf den schon oben zitierten Review des US Department of Health and Human Services (1996).

Eden et al. (2002) haben Effekte im Rahmen der Beratungen in Hausarztpraxen (primary care) in ihrem systematischen Review untersucht. Eingeschlossen wurden acht Studien. Ergebnisse hinsichtlich „efficacy“ sind widersprüchlich, da – wie die Autoren feststellen – „most studies had at least one methodologic limitation, it was difficult to rigorously assess the efficacy“ (Eden et al. 2002).

Kahn et al. (2002) wiederum haben 96 von 253 gefundenen Studien in ihre weitere Untersuchung eingeschlossen und drei Bereichen zugeordnet: 1) Reine Informationsangebote, 2) Verhalten und soziale Unterstützung, 3) Umwelt und politische Maßnahmen. Für die drei Bereiche werden dann folgende Effekte angegeben:

Ad 1) Informationsangebote werden noch weiter unterteilt in: a) Poster, die bei der Wahl Treppe vs. Rolltreppe/ Aufzug auf die Wahl der Treppe hinweisen. Der Effekt beläuft sich auf einen Anstieg der Treppenbenutzung zwischen 5,5 und 128,6% (sieben Studien). Dagegen ist kritisch anzuführen aus Blamey und Mutrey (2004): „There is limited knowledge to date about the ideal type and size of sign and the ideal message.“, so dass die Frage, wie man Menschen zur Treppenbenutzung motiviert, noch völlig ungeklärt ist. b) Gemeindeweite, gut sichtbare Kampagnen haben einen Effekt zwischen – 2,9 und 9,4% auf den Anteil der sich bewegenden Personen (zehn Studien). c) Die massenmediale Verbreitung allein hat laut Kahn et al. bis zum Jahr 2002 keine ausreichenden Nachweise erbracht, ob und in welcher Höhe sie wirksam ist.

Ad 2) a) Interventionen zur familienbasierten Unterstützung sind bisher ohne ausreichende Nachweise geblieben. b) Gemeindenahe soziale Unterstützung (wie Walkinggruppen, Buddysysteme) haben einen Effekt auf Zeit, die in Aktivität verbracht wird, zwischen 19,4 und 45,6%, auf

Häufigkeit der Aktivität zwischen 14,6 und 57,6% und auf die Ausdauer der Teilnehmenden von 4,7 bis 6,1% (neun Studien). c) Individuell zugeschnittene Gesundheitsprogramme haben einen Effekt auf Zeit, die in Aktivität verbracht wird, zwischen 16,7 und 83,3% (18 Studien).

Ad 3) Interventionen, die sich aus einer Mischung aus der Förderung/ Verbesserung des Zugangs zu Fahrradwegen, aus betrieblicher Förderung, aus Unterrichtseinheiten u.a. (11 Studien) zusammensetzen haben einen Effekt auf die Ausdauer zwischen 2,8 und 9,6%, auf den Energieverbrauch -2,5 und 24%, und auf die Zeit, die mit Freizeitsport verbracht wird, zwischen -6 und 8,5%.

Foster und Hillsdon (2004) haben in ihr systematisches Review über Veränderung der Umwelt und Effekte auf Bewegung aus über 14.000 zunächst in der Suche gefundenen Titeln 17 Studien ausgewählt, die den Einschlusskriterien entsprachen (dort 757). Auch hier sind die Effekte selbst zwischen den Settings - widersprüchlich und nicht übertragbar (Treppen in Betrieben vs. in U-Bahn-Stationen). Außerdem sind nach unserer Meinung viele der Interventionen doch Verhaltensmaßnahmen, wie z. B. ein Poster, das zum Treppensteigen vs. Rolltreppe (s. o.) animiert, denn damit werden die Verhältnisse nicht verändert. Eine Veränderung der Umwelt bestünde aus einem Stilllegen der Rolltreppe.

Der Stand der Forschung, der durch diese systematischen Reviews dokumentiert wurde, kann folgendermaßen kritisch bewertet werden: „Although this review (gemeint ist Hillsdon et al. 2002) is a very useful addition to existing evidence, like the review by Kahn et al., it presents high-level findings that illustrate the many gaps in our knowledge of what is effective and how such interventions can be generalized to wider populations.“ (Blamey und Mutrey 2004) Und weiter heißt es ebenda: „Buchner and Miles (2002) suggested that systematic reviews of physical activity have to date not uncovered a few specific interventions that, if prioritized and implemented, would dramatically influence population levels of physical activity.“ Systematische Reviews geben damit aber keine Auskunft darüber, was warum wo besser bzw. schlechter funktioniert. Dies ist aber die Fragestellung unserer Untersuchung.

Vielfach werden als Ergebnisparameter nach Interventionen Einstellungen oder Wissen gemessen. Dass dies aber oft gar nicht zu einer Verhaltensänderung führt, ist bekannt. Auch wenn eine einzelne Studie nicht ausreicht zu belegen, wie gerade im Bereich Bewegung Einstellungen und Wissen am Ende einer Intervention nicht mit dem Effekt auf das Bewegungsverhalten korrelieren, wird dennoch exemplarisch eine aktuell publizierte Studie herangezogen, die den „garstigen“

Graben zwischen dem Wissen und den Einstellungen auf der einen Seite und dem Verhalten auf der anderen Seite eindrucksvoll belegt. So haben Zizzi et al. (2006) 165 Schüler/innen aus mehreren High Schools des US-Bundesstaates West Virginia auf freiwilliger Basis in die Studie eingeschlossen. Sie erhielten entweder einen Schrittzähler für drei Wochen (a) oder sollten zusätzlich zu dem Schrittzähler individuelle Ziele aufstellen, wie viel Schritte durchschnittlich pro Tag erreicht werden sollten (b). Insgesamt konnten die beiden Hypothesen, dass die Schrittzahl mit der Zeit ansteigen würde und dass die Gruppe, die sich Ziel setzte, einen höheren Anstieg verzeichnen würde, nicht bestätigt werden. Auch der Anstieg gegenüber Baseline war nicht signifikant. Dennoch haben 70% der Teilnehmenden in qualitativen Interviews ausgesagt, dass ihr Bewusstsein und ihre Motivation zu Gesundheit und Bewegung durch die Schrittzähler positiv beeinflusst worden seien. Allerdings haben nur 25% der Schüler überhaupt die empfohlene altersbezogene Schrittzahl von 11 bis 13.000 pro Tag erreicht.

### 6.1.1 Bewegung im Betrieb

Es wurden 1994 im NHIS (National Health Interview Survey 1994, zitiert nach Grosch et al. 1998) 5219 Personen befragt, die repräsentativ für etwa 53 Mio. Arbeitnehmer in den USA stehen.

Tabelle 7: Zugang und Teilnahme zu betrieblichen Interventionen in den USA (Grosch et al. 1998)

Einrichtung	Zugang	Teilnahme
Schließfächer für Kleidung	21,3%	42,2%
Trainingsraum	19,7%	32,4%
Duschen	19,6%	31,0%
Geräte für Zirkeltraining	17,4%	28,3%
Ausrüstung für Krafttraining	16,6%	27,5%
Geh- oder Joggingpfad	14,0%	36,1%
Fahrradständer	7,4%	9,8%
Schwimmbecken	6,7%	24,4%
Trimpfad	5,2%	12,2%
Fahrradweg	4,7%	16,7%

Tabelle 8: Zugang und Teilnahme an konkreten innerbetrieblichen Bewegungsprogrammen

Bewegungsprogramm	Zugang	Teilnahme
Aerobics	12,2%	14,4%
Gehltreff	9,7%	28,1%
Zuzahlung zu Mitgliedsbeitrag in Studio	6,5%	18,7%
Sport- oder Übung als Wettbewerb	6,3%	26,7%
Jogging/ Laufgruppe	5,7%	16,6%
Krafttraining	5,0%	20,8%
Turnübungen, nicht Aerobics	4,7%	17,6%
Fahrradgruppe	3,2%	14,6%
Schwimmen	3,2%	10,2%
Volle Bezahlung des Mitgliedsbeitrags in Studio	2,2%	22,3%

Die Tabelle 7 stellt für zehn ausgewählte innerbetriebliche Einrichtungen dar, für wie viele der 53 Mio. Arbeitnehmer diese verfügbar sind und wie viele derjenigen, die somit diese nutzen können, diese auch real nutzen.

Allein die Gegebenheiten bzw. das Setzen von Anreizen (insbesondere finanziell) scheinen die Mehrheit der Mitarbeiter/innen noch nicht zu einem entsprechenden Bewegungsverhalten zu motivieren.

Im Hinblick auf die Ergebnisparameter wie Arbeitszufriedenheit oder zu Fehlzeiten kommt das systematische Review von Proper et al. (2002) zu einem ernüchternden Ergebnis: „In conclusion, there is limited evidence for the effectiveness of physical activity programs at worksites with respect to absenteeism from work and inconclusive evidence for effectiveness regarding job satisfaction, job stress, and employee turnover, and there is no evidence for effectiveness in respect of productivity. These conclusions are probably due to a lack of randomized controlled trials of high methodological quality.“ Allerdings lässt sich dieses Ergebnis des systematischen Reviews auch dadurch erklären, dass die Interventionen möglicherweise im Hinblick auf die von den Autoren nicht untersuchte Interventionsqualität Mängel haben.

Vor diesem Hintergrund wurden die beiden Themenbereiche Bewegung bei Frauen und Mädchen und im Betrieb ausgewählt. Mädchen bewegen sich signifikant weniger, was z. B. insbesondere in den USA nachgewiesen werden konnte (z. B. Grunbaum et al. 2001, Kimm et al. 2002).

Weiterhin sind die Effekte betrieblicher Interventionen zu Bewegung nicht auf die von uns ausgewählten Zielparameter wie tatsächliche Zeit in Bewegung, Veränderung von Gewicht oder klinischen Parametern etc. untersucht worden, so dass dieser zweite Fokus in der vorgestellten Analyse gewählt wurde.

## 6.2 Ergebnisse

Für die Risikofaktoren Bewegung wurden die Handlungsfelder Bewegungsförderung für Frauen und Mädchen sowie Interventionen zu Bewegung im Betrieb analysiert. Ausgewählte Studien zu diesen Themenfeldern werden weiter unten dargestellt.

Tabelle 9: Übersicht über die eingeschlossenen Studien im Handlungsfeld Bewegung

Kenn-Nummer	Autor (Jahr)	Projekt- oder Interventionsname (ggf.)	Land	Design
1002	Schofield et al. (2005)		Australien	RCT
1006	Wen et al. (2002)		Australien	non-CT
1008	Robbins et al. (2006)	Girls on the Move	USA	RCT
1009	Williams et al. (2005)		USA	RCT
1010	Felton et al. (2005)		USA	RCT
1018	Aldana et al. (2005)		USA	non-CT
1023	Carels et al. (2004)		USA	RCT
1026	Ransdell et al. (2004)	Get Fit	USA	RCT
1032	Pohjonen et al. (2001)		Finnland	CT
1033	Emmons et al. (1999)	The Working Healthy Project	USA	RCT
1034	Kerr et al. (2004)		USA	non-CT
1035	Ewart et al. (1998)		USA	RCT
1036	Gamble et al. (1993)		Großbritannien	CT
1037	Mutrie et al. (2002)		Großbritannien	RCT
1038	Vuori et al. (1994)		Finnland	RCT
1039	Fahrenwald et al. (2004)	Moms on the Move	USA	RCT

Kenn-Nummer	Autor (Jahr)	Projekt- oder Interventionsname (ggf.)	Land	Design
1040	Aittasalo et al. (2004)		Finnland	RCT
1041	Pegus et al. (2002)	Heart at Work	USA	CT
1042	Glasgow et al. (1997)	Take Heart II	USA	CT
1043	Gerdle et al. (1995)		Schweden	RCT
1044	Marshall et al. (2002)		Australien	non-CT
1045	Kerr et al. (2001)		Großbritannien	non-CT
1046	Boutelle et al. (2001)		USA	non-CT
1047	Adams et al. (2002)		Großbritannien	non-CT
1048	Russell et al. (1999)		USA	non-CT
1049	Titze et al. (2001)		Schweiz	non-CT
1050	Coleman et al. (2001)		USA	CT
1051	Sorensen et al. (2005)	Healthy directions – small business	USA	RCT
1052	Badland et al. (2005)		Australien	CT
1053	Atlantis et al. (2006)		Australien	RCT
1054	Lee et al. (1997)		Australien	RCT
1055	Pate et al. (2005)		USA	RCT
1056	Harrell et al. (1996)		USA	RCT
1057	Fisher et al. (1995)		USA	CT
1058	Grønningsæter et al. (1992)		Norwegen	RCT
1059	Kuller et al. (2001)	Women's Health Lifestyle Project	USA	RCT
1060	Pritchard et al. (1997)		Australien	RCT

Die in den Studien enthaltenen Effektmaße, werden im Folgenden in einer Übersicht angegeben. Auch innerhalb der Effektmaße besteht deutliche Heterogenität der Parameter. So werden die Effektmaße entweder mit unterschiedlichen Endpunkten (körperliche Fitness z. B.) oder mit verschiedenen Methoden/ Instrumenten erhobenen Gewicht (z. B. Gewicht) erhoben.

Tabelle 10: Ausgewählte Effektmaße im Fokus Bewegung im Betrieb

Ken- nung	Aerobe Fitness	Anteil Personen in Bewegung	Anteil Zeit in Bewe- gung	Body Mass Index/ Gewicht	Körperli- che Fitness	Treppen- steigen	Weitere
1018	X			X			
1032					X		Anteil Körper- fett
1033		X					
1034						X	
1036	(x)			X			
1037			X				
1038	X	(x)					
1040		X					Schritt- zahl
1041				(x)			
1042			(x)				
1043	(x)			X	(x)		
1044						X	
1045						X	
1046						X	
1047						X	
1048						X	
1049						X	
1050						X	
1051		(x)					
1052							Schritt- zahl
1053	X			X			X
1054							
1056	X				X		X
1057				X			X
1058	X						X
1060				X			X

Tabelle 11: Ausgewählte Effektmaße im Fokus Bewegung Frauen und Mädchen

Kennung	Aerobe Fitness	Anteil Personen in Bewegung	Anteil Zeit in Bewe- gung	Body Mass In- dex/ Gewicht	Körperli- che Fitness	Schrittzahl	Weitere
1002			X			x	
1006		X		x			

1008			X				
1009		X				x	
1010			X				
1023				x	x		
1026	(x)		X				
1032					(x)		x
1035	(x)			x			
1039			X			x	
1043							
1054							
1055							
1059							

## 6.2.1 Bewegung Frauen und Mädchen

Insgesamt konnten vierzehn Studien in die Auswertung aufgenommen werden. Deren Evaluation wird an Hand von je vier Kriterien der Studien- wie der Interventionsqualität tabellarisch dargestellt.

Tabelle 12: Qualität der Studiendurchführung im Fokus Bewegung Frauen und Mädchen<sup>9</sup>

Kennung	Studien- design	Rando- misation (16)	Ver- gleich- barkeit (19)	Messin- strumen- te validiert (25)	Power (46)	Sum-me	Erzielba- rer Maximal- Score	Pro- zentua- ler Score
1002	RCT	1	0	1	0	3,5	7	50
1006	Non-CT	X	X	1	1	3	5	60
1008	RCT	1	0	1	0	4	7	57
1009	RCT	1	0	1	0	3,5	7	50
1010	RCT	1	0	1	0	3	7	43
1023	RCT	1	1	0	0	3	7	43
1026	RCT	0	1	0,5	0	2,5	7	36
1032	CT	X	1	0,5	0	4	6	67
1035	RCT	0	0	0	0	1,5	7	21
1039	RCT	1	1	1	1	5,5	7	79
1043	RCT	0	1	0	0	3	7	43
1054	RCT	1	0	0	0	0	7	0
1055	RCT	1	1	1	0	5,5	7	79
1059	RCT	1	1	1	0	5	7	71

<sup>9</sup> Aus Gründen des Platzes werden für alle Studien nur jeweils vier Kriterien der Qualität des Studiendesigns sowie der Qualität der Intervention dargestellt.

Legende: Sollte eine Frage nach Manual nicht zutreffen, so wird ein X gesetzt. Ansonsten wird nach Manual zugeteilt in 1, 0,5 oder 0.

Tabelle 13: Qualität der Intervention im Fokus Bewegung Frauen und Mädchen

Kennung	Studien- design	Partizi- pation (54)	Epower- ment (55)	Prozess doku- men- tiert (63)	Prozess evaluiert (64)	Summe	Erzielba- rer Maximal- Score	Pro- zentu- aler Score
1002	RCT	0	0	1	0	4	10,5	38
1006	non-CT	1	0	1	0	7,5	10,5	71
1008	RCT	0	0	1	0	6	10,5	57
1009	RCT	0	0	1	0	4	10,5	38
1010	RCT	0	1	1	1	8,5	10,5	81
1023	RCT	0	1	0	0	2	10,5	19
1026	RCT	0	0	0	0	1	9	11
1032	CT	0	0	0	0	1,5	10,5	14
1035	RCT	0	0	0	0	0	9	0
1039	RCT	0	1	1	0	5	10,5	48
1043	RCT	0	0	1	0	2,5	9,5	26
1054	RCT	0	0	1	0	4,5	10	45
1055	RCT	0	1	1	0	7	10,5	67
1059	RCT	0	0	0	0	2,5	9,5	59

Folgende Korrelationen zwischen Qualität und Effekt konnten erhoben werden:

Tabelle 14: Korrelation von Effekt, Qualität der Studiendurchführung und Qualität der Intervention im Fokus Bewegung bei Frauen und Mädchen

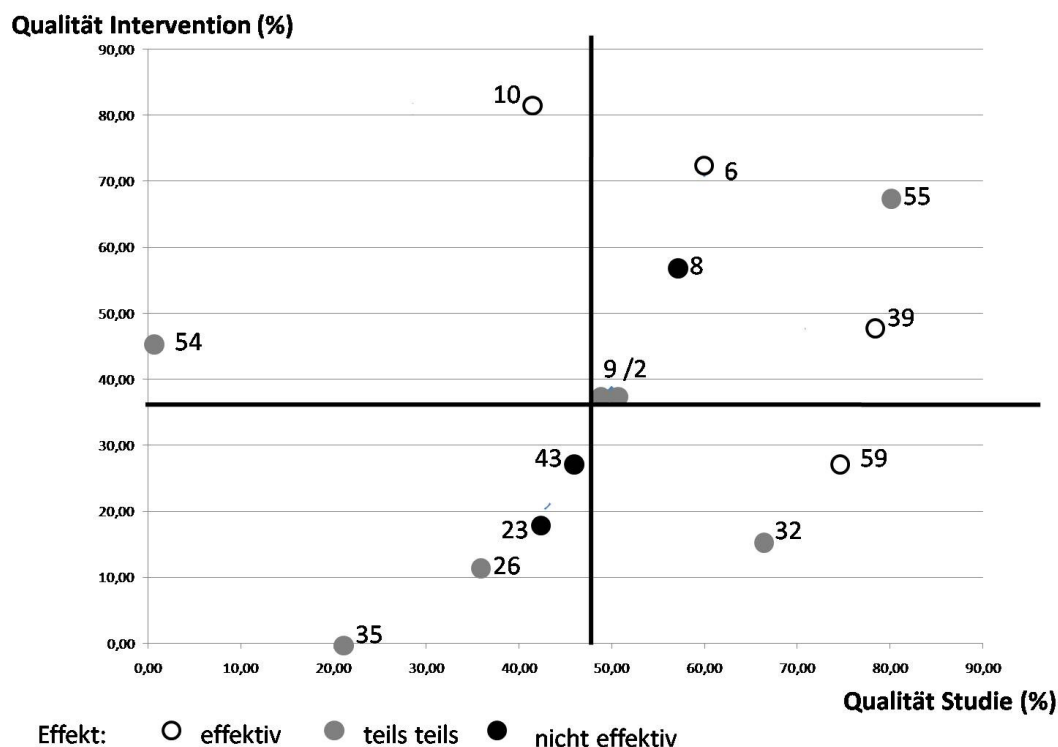
Zu korrelierende Größen	Spearmankorrelation
Qualität gesamt und Effekt	0,47
Qualität der Studiendurchführung und Qualität der Intervention	0,29
Qualität der Studiendurchführung und Effekt	0,36
Qualität der Intervention und Effekt	0,33

Es zeigt sich, dass die Qualität der Intervention sowie die Qualität der Studiendurchführung nur mit einer Korrelation nach Spearman von 0,33 bzw. 0,36 mit dem Effekt zusammen hängen. D.

h. eine nach unserem Kriterienkatalog hohe Qualität in einem der beiden Gebiete ist keine Gewähr, dass sich ein positiver Effekt ergibt.

Der Durchschnittswerte für die Qualität der Studiendurchführung liegt bei 49,8%, für die Qualität der Intervention bei 38,7%. Damit werden die Grenzen festgelegt zwischen der Kategorie keine Aussage über Empfehlungen zu machen und den drei Kategorien abgestufter Empfehlungsgrade (vgl. Abbildung 8).

Abbildung 9: Bewegung Frauen und Mädchen



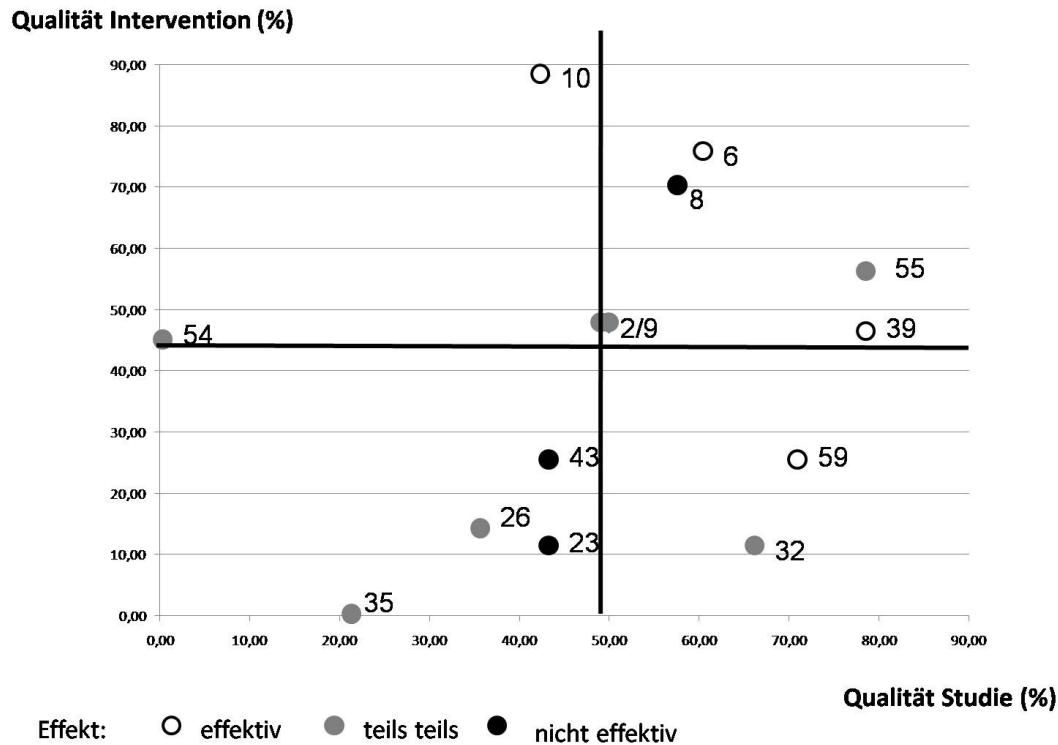
Legende: Aufgetragen wurden die Scores für Qualität des Studiendurchführung (x-Achse) und der Qualität der Intervention (y-Achse) nach Bewertung.

### Darstellung der Sensitivitätsanalyse

Da insbesondere die Kriterien Partizipation sowie Empowerment werteträgend sind (vgl. Kapitel 3.4), legen wir die Abbildung 10 auf Basis einer Berechnung von Studienqualität und Interventionsqualität ohne diese beiden Kriterien vor. Für die Sensitivitätsanalyse hat sich nachvollziehbar nur der Durchschnittswert für die Qualität der Intervention verändert; er ist angestiegen auf 43,4%.

Es zeigt sich keine Abweichung von der o.g. Darstellung hinsichtlich der Studien, über die überhaupt Empfehlungen ausgesprochen werden können. Dies ist dem geschuldet, dass überhaupt nur wenige der betrachteten Studien Partizipation (eine Studie) bzw. Empowerment (vier Studien) in die Intervention integriert haben.

Abbildung 10: Bewegung Frauen und Mädchen: Auswertung ohne Empowerment und Partizipation (Sensitivitätsanalyse)



## 6.2.2 Darstellung einzelner Studien (Bewegung Mädchen und Frauen)

Im Folgenden werden die Studien, die in die drei Bereiche fallen, über die Empfehlungen (s. Abbildung 9) ausgesprochen werden können, näher betrachtet. Insgesamt fallen sechs Studien in den Bereich empfehlenswerter Interventionen.

Keine Studie fällt in den Bereich „in hohem Maße empfehlenswert“. Eine Studie ist „stark empfehlenswert“, weist also zu Studien und Interventionsqualität Scores über 60% auf. Die anderen fünf Studien liegen über dem Durchschnitt und fallen in die Kategorie „eingeschränkt empfehlenswert“.

### *Schofield et al 2005 (1002)*

Land: Australien.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: Intervention 1 mit 23, Intervention 2 mit 21 und Kontrolle mit 24 16jährigen Mädchen.

Intervention: Es wurde ein über sechs Wochen dauerndes Unterrichtsprogramm zur Steigerung von Bewegung mit Selbstkontrolle der Mädchen über Pedometer bzw. Minutenzählung entwickelt. Zielgruppe waren sich wenig bewegende Mädchen der Sekundarstufe II.

Studiendauer: 1,5 Monate bzw. 3 Monate.

Nachbeobachtung: 1,5 Monate.

Im eigentlichen Sinne liegt keine Nachbeobachtungszeit vor. Zwar dauerte die Intervention nur sechs Wochen, dennoch wurden in den weiteren sechs Wochen alle zwei Wochen Erinnerungsschreiben an die teilnehmenden Mädchen der Interventionsgruppen versandt.

Effektmaße: Ausgewählt wurden 1) Vier-Tages-Schrittzahl, 2) Gesamtsumme moderat bis intensiver Bewegung und 3) die Gesamtsumme intensiver Bewegung.

Effekte und p-Werte: 1) Zunahme der Vier-Tages-Schrittzählung in der Endauswertung mit kleinem Effekt von 0,13 ( $p=0,03$  nur zwischen Pedometer- und Kontrollgruppe, nicht im Vergleich von Pedometer- und Minutenzählgruppe sowie von Minutenzähl- und Kontrollgruppe). 2) In den Gruppen konnte eine signifikante Veränderung bei Pedometer- und Minutenzählgruppe zwischen 0 und zwölf Wochen (also erst nach Ende der eigentlichen Intervention) sowie zwischen sechs und zwölf Wochen ( $p=0,00$  bis  $0,01$ ) festgestellt werden, nicht in der Kontrollgruppe. 3) Der Anteil von moderat bis intensiver und intensiver Bewegung hat sich nicht signifikant verändert. 4) Ebenso ergab sich keine signifikante Veränderung bei BMI und aerober Kapazität.

Besondere Stärken und Schwächen: Kleine Studie, eher Vergleich der efficacy (generellen Wirksamkeit) zweier Interventionen. Die Erwartungen in Relation zur Dauer der Studie und der Nachbeobachtungsphase scheinen unrealistisch, denn das Programm ist kürzer als 2 Monate. Umgekehrt ist die Nachbeobachtung trotz Kürze der Studie und damit Darstellung eines verzögerten Effekts eine Stärke der Studie.

Empfehlung und Übertragung: Pedometer und Minutenzählen können ein Anreiz für eine selektive Intervention eine Gruppe von sich wenig bewegenden Mädchen der Sekundarstufe II sein, ihr Bewegungsverhalten zu ändern. Allerdings bedarf es einer längeren Beobachtung.

### ***Wen et al. 2002 (1006)***

Land: Australien

Studiendesign: non-CT.

Größe der Interventions- und Kontrollgruppe, bzw. der Gruppen im Vorher-Nachher-Design: 1762 Frauen wurden vor der Intervention, 1801 nach der Intervention angerufen.

Intervention: Es wurde eine mit der politischen Gemeinde zusammen geplante multistrategische Intervention für Frauen von 20 bis 50 zur Anregung von Bewegung durchgeführt. Darin enthalten waren Information über Presse/Medien, die bessere Ausschilderung von Wegen, Walkinggruppen sowie die Kinderbetreuung während der Zeit der Walkinggruppen. Weitere Sportarten wurden angeboten. Dazu gab es einzelne Veranstaltungen wie die Vorstellung der Initiative anlässlich eines lokalen Jahrmarkts. Eine soziale Marketingkampagne über die lokalen Printmedien getragen.

Studiendauer: zwei Jahre.

Nachbeobachtung: keine.

Als Effektmaße wurden zum einen der Prozentsatz an Frauen, die einen sitzenden Lebensstil haben bzw. die sich leicht, moderat und intensiv bewegen, und zum anderen die Anzahl der Stunden, die in den letzten 14 Tagen mit Gehen verbracht worden sind, gewählt.

Effekte und p-Werte: 1) Es konnte eine signifikante Zunahme der Frauen, die sich leicht bewegen, von 41,6 auf 47,9% (Konfidenzintervalle im Prä- Post- Vergleich überdecken sich nicht.) bei gleichzeitiger Abnahme der sich nicht bewegenden Frauen von 21,6 auf 15,2% (Konfidenzintervalle im Prä- Post-Vergleich überdecken sich nicht) erhoben werden. 2) Der Anteil der Frauen, die in den letzten 14 Tagen keine Stunde mit Gehen verbracht haben, hat sich von 30,8 auf 20,0% ( $p < 0,001$ ) reduziert.

Besondere Stärken oder Schwächen: Initial wurden zwölf Fokusgruppen mit 75 beteiligten Frauen zu fördernden und hemmenden Faktoren ins Leben gerufen. Des Weiteren wurden Interviews mit für das Projekt relevanten Gemeindeangestellten geführt und qualitativ ausgewertet. Abschließend wurden erneut Interviews mit 75 Frauen geführt und qualitativ ausgewertet. Universelle Intervention.

Empfehlung und Übertragung: Die Studie baut auf die Erfolgsfaktoren Partizipation, multistrategisches Vorgehen, Integration der Entscheidungsträger, Ausrichtung an Bedürfnissen und kann damit auch einen Effekt erzielen. Diese Studie kann zur Umsetzung empfohlen werden.

***Robbins et al. 2006 (1008)***

Land: USA.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: Insgesamt 77 Mädchen der Klassen 6 bis 8, die sich wenig bewegen, 45 Mädchen vs. 32 in Interventions- bzw. Kontrollgruppe.

Intervention: „Girls on the Move“ als individualisiertes Computerprogramm für Mädchen in Kombination mit Beratung durch eine pädiatrisch geschulte Kraft (nurse practitioner= Pflegekraft, darf aber auf Basis von Erfahrung und Prüfungen ärztliche Tätigkeiten ausüben).

Studiendauer: Drei Monate.

Nachbeobachtung: keine.

Effektmaße: 1) Bewegung (Zeit), 2) Selbstwirksamkeit, 3) Freude bei der Bewegung.

Effekte und p-Werte: keine signifikante Veränderung der uns interessierenden Effektparameter: Zeit in Bewegung.

Besondere Stärken und Schwächen: Zeit in Bewegung nur durch subjektive Angaben erhoben. Obwohl zuvor adolescenten Mädchen zu Barrieren wie fördernden Faktoren im Bezug auf Bewegung befragt wurden, ergab sich kein Effekt. Ebenso wurden als Ergebnisparameter die Selbstwirksamkeit, die Freude bei Bewegung und die soziale Unterstützung untersucht, Parameter, die ansonsten oft vernachlässigt werden. Selektive Intervention.

Empfehlung und Übertragung: Bei guter Studien- und Interventionsqualität kein Effekt, somit eher unwahrscheinlich, dass ein individualisiertes Programm ohne weitere Änderung der Verhältnisse Effekte zeigen kann.

***Williams et al. 2005 (1009)***

Land: USA.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 16 und 19 Frauen im Alter.

Intervention: Frauen zwischen 50 und 68 Jahren wurden beraten. Mit ihnen wurden individuelle Zielsetzungen vereinbart. Dann wurde die Einhaltung der Zielsetzung als vertragliche Übereinkunft formuliert.

Studiendauer: 1,75 Monate.

Nachbeobachtung: keine.

Effektmaße: 1) Erreichen einer halben Stunde forsches Gehen pro Tag, 2) Erreichen von durchschnittlich 10000 Schritten pro Tag.

Effekte und p-Wert: 1) Steigerung der Schrittzahl in beiden Gruppen signifikant. 2) Die Schrittzahl von 10000 pro Tag wird nicht erreicht. 3) Interventionsgruppe erreicht Ziel halbe Stunde forsches Gehen pro Tag (81 vs. 31%,  $p = 0,006$ )

Besondere Stärken und Schwächen: Das Programm ist kürzer als 2 Monate.

Empfehlung und Übertragung: Das Ziel einer durchschnittlichen minimalen Schrittzahl von 10000 pro Tag konnte nicht erreicht werden. Die Studiendauer ist sehr kurz, um darauf aufbauend Empfehlungen zur Umsetzung auszusprechen.

### ***Fahrenwald et al. 2004 (1039)***

Land: USA.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 22 zu 22.

Intervention: Eine interaktive Broschüre mit Pros und Cons speziell für Mütter aus sozioökonomisch schwachen Verhältnissen wurde ausgegeben. Zusätzlich wurden vier Telefonanrufe im Rhythmus von je zwei Wochen mit Gespräch über wichtigste Pro und Cons aus Sicht der Mütter geführt.

Studiendauer: 2 Monate.

Nachbeobachtung: 0,5 Monate.

Effektmaße: 1) Stadien der Veränderung, 2) Bewegung und Bewegungsverhalten, 3) Konstrukte wie Selbstwirksamkeit

Effekt und p-Wert: 1) In der Interventionsgruppe konnte eine signifikant höhere Veränderung aus Stadium der contemplation und preparation ( $p < 0,001$ ) verzeichnet werden. 2) Die täglich und wöchentlich durch Bewegung verbrauchte Energie nimmt zu in der Interventionsgruppe ( $p = 0,001$ ), ebenso die Zeit, die in moderater Bewegung ( $p = 0,001$ ) verbracht wird, sowie die Zahl der Schritte pro Tag von 5.825 auf 9.180 ( $p < 0,001$ ), 3) Selbstwirksamkeit etc. sowie auch Neueinschätzung der Umgebung haben sich signifikant verändert.

Besondere Stärken und Schwächen: Es handelt sich eher um eine Pilotstudie. Die Durchführung erfolgt im Rahmen des WIC (Women, Infants, and Children) Programms, das für Frauen mit niedrigem Einkommen aufgelegt wurde. Selektive Intervention.

Empfehlung und Übertragung: Zur Zeit noch im Pilot- bzw. Preteststadium, daher vergleichbare Wirkung plausibel, aber noch nicht hinlänglich erwiesen. Effekt kann auch auf gute Vernetzung mit dem übergeordneten Programm zurück gehen.

### ***Pate et al. 2005 (1055)***

Land: USA.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 863 zu 741 Mädchen der achten und neunten Klasse.

Intervention: LEAP= lifestyle education for activity program. Insbesondere wurde der Sportunterricht völlig verändert, dabei Mädchen und Jungen getrennt unterrichtet. Die Intervention umfasste daneben weitere fünf Komponenten: Gesundheitserziehung, Veränderung der Schulumgebung, Einschluss der Schulgesundheitsshelfer, Promotion durch Lehrer, Einschluss von Familien und Gemeinde.

Studiendauer: 10 Monate.

Nachbeobachtung: keine.

Effektmaße: Anteil von Mädchen in intensiver Bewegung mehr als 30 min. x Tag in retrospektiver subjektiver Erhebung (Fragebogen zu den letzten drei Tagen), BMI.

Effekte und p-Werte: BMI ohne Veränderung. Anteil der Mädchen, die sich 30 min. pro Tag intensiv bewegen, unterscheidet sich signifikant zwischen Interventions- und Kontrollgruppe nach Intervention: 45% vs. 36% ( $p= 0,05$ ).

Besondere Stärken und Schwächen: große Studie. Das Programm ist mädchenspezifisch. Eine lokale Adaptation eines Programms wird gewünscht. Es bestehen auch Elemente zur Stärkung der Selbstwirksamkeit (z. B. Dishman et al. (2004)). Es gibt weitere nach Interventionsqualität sehr hochstehende Einzelstudien, die an das Programm angelagert wurden (Studie 1010). Es wurden Lehrer, Familien und Gemeinden bei Ausrichtung auf das Setting Schule integriert.

Empfehlung und Übertragung: Erfolgreicher settingbezogener Ansatz. Das widersprüchliche Ergebnis der Effektparameter BMI und Bewegung kann unterschiedlich erklärt werden: a) Fett

wurde in Muskel umgewandelt, der schwerer ist, oder b) Es gab keine Gewichtsabnahme. Hier muss entschieden werden, welcher Parameter zielführend sein soll, bzw. ob es weitere quantitative bzw. qualitative Analysen und Auswertungen gibt, die die Gründe darlegen, z. B. weiterhin wenig Bewegung unter den Adipösen, so dass kein Effekt auf BMI oder Messung der entsprechenden weiteren Parameter (Trizepsfalten etc.), um Veränderung der Körperzusammensetzung zu messen.

### 6.2.3 Bewegung Betrieb

Tabelle 15: Qualität der Studiendurchführung im Fokus Bewegung im Betrieb

Kennnummer	Studien-Design	Randominisation (16)	Vergleichbarkeit (19)	Messinstrumente validiert (25)	Power (46)	Summe	Erzielbarer Maximalscore	Prozentualer Score
1018	non-CT	X	X	0	0	1	5	20
1032	CT	X	1	0,5	0	4	6	67
1033	RCT	1	1	0	0	2,5	7	36
1034	non-CT	X	X	0	0	1	5	20
1036	CT	X	0	0	0	1	6	17
1037	RCT	1	0	0,5	1	3	7	43
1038	RCT	0	0	0	0	1	7	14
1040	RCT	0	1	1	0	3	7	43
1041	CT	X	0	0	0	1,5	6	25
1042	CT	X	1	0,5	1	3,5	6	58
1043	RCT	0	1	0,5	0	3	7	43
1044	non-CT	X	X	0	1	2	5	40
1045	Non-CT	X	X	0	0	0	5	0
1046	Non-CT	X	X	1	0	2	5	40
1047	Non-CT	X	X	1	1	4	5	80
1048	Non-CT	X	X	0,5	0	1,5	5	30
1049	Non-CT	X	X	0,5	0	1,5	5	30
1050	CT	X	0	0	0	0	6	0
1051	RCT	1	1	0	0	5	7	71
1052	CT	X	0	0	0	1	6	17
1053	RCT	1	0	0,5	1	4,5	7	64
1054	RCT	1	0	0	0	0 <sup>10</sup>	7	0

<sup>10</sup> Das Ergebnis 0 ist plausibel, da der Punkt für die Beschreibung der Randomisation wieder abgezogen wurde, da die Randomisation nicht so durchgeführt worden war, dass alle Einflussvariablen auszuschließen waren. Beispiel: Bei betrieblichen oder schulischen Interventionen wurde innerhalb eines Betriebs, einer Schule randomisiert, so dass nicht auszuschließen ist, dass Effekte von der Interventionsgruppe in die Kontrollgruppe überschwapen.

1056	RCT	0	1	1	0	3	7	43
1057	CT	X	0	0	0	0	6	0
1058	RCT	0	0	0,5	0	2	7	29
1060	RCT	1	1	1	0	3	7	43

Legende: Sollte eine Frage nach Manual nicht zutreffen, so wird ein x gesetzt. Ansonsten wird nach Manual zugeteilt in 1, 0,5 oder 0.

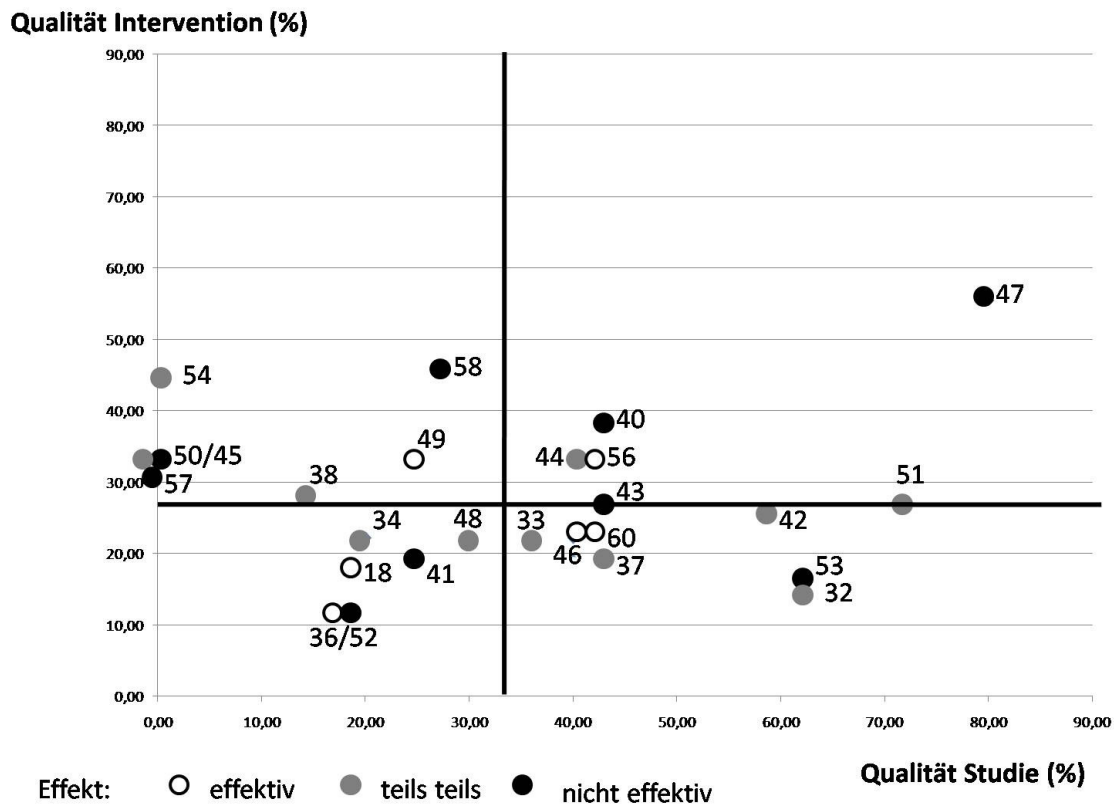
Tabelle 16: Qualität der Intervention im Fokus Bewegung im Betrieb

Kennnummer	Studien-Design	Partizipation (54)	Empowerment (55)	Prozess dokumentiert (63)	Prozess evaluiert (64)	Summe	Erzielbarer Maximalscore	Prozentualer Score
1018	Non-CT	0	0	0	0	2,5	10,5	19
1032	CT	0	0	0	0	1,5	10,5	14
1033	RCT	1	0	0	0	2	9	22
1034	Non-CT	0	0	1	0	2	9	22
1036	CT	0	0	0	0	1	9	11
1037	RCT	0	0	1	0	2	10,5	19
1038	RCT	0	0	1	0	3	10,5	29
1040	RCT	0	0	1	1	4	10,5	38
1041	CT	0	0	0	0	2	10,5	19
1042	CT	0	0	1	0	2,5	10	25
1043	RCT	0	0	1	0	2,5	9,5	26
1044	Non-CT	0	0	1	1	3	9	33
1045	Non-CT	0	0	1	1	3	9	33
1046	Non-CT	0	0	1	0	2	9	22
1047	Non-CT	1	0	1	1	5	9	56
1048	Non-CT	0	0	1	0	2	9	22
1049	Non-CT	0	0	1	0	3,5	10,5	33
1050	CT	0	0	1	0	3	9	33
1051	RCT	1	0	1	0	2,5	10	25
1052	CT	0	0	0	0	1	9	11
1053	RCT	0	0	1	0	2	10,5	19
1054	RCT	0	0	1	0	4,5	10	45
1056	RCT	0	0	0	0	3,5	10,5	33
1057	CT	1	0	0	0	3	10	30
1058	RCT	0	1	1	0	5	10,5	48
1060	RCT	0	0	0	0	2	9	22

Legende: Sollte eine Frage nach Manual nicht zutreffen, so wird ein x gesetzt. Ansonsten wird nach Manual zugeteilt in 1, 0,5 oder 0.

Diese Bewertung wurde wiederum graphisch umgesetzt mit den Dimensionen Qualität der Studie, Qualität der Intervention wie Effekt (s. Abbildung 11). Als Durchschnittswerte wurden 33,5% für die Qualität des Studiendesigns und 27,4% für die Qualität der Intervention errechnet.

Abbildung 11: Bewegung im Betrieb



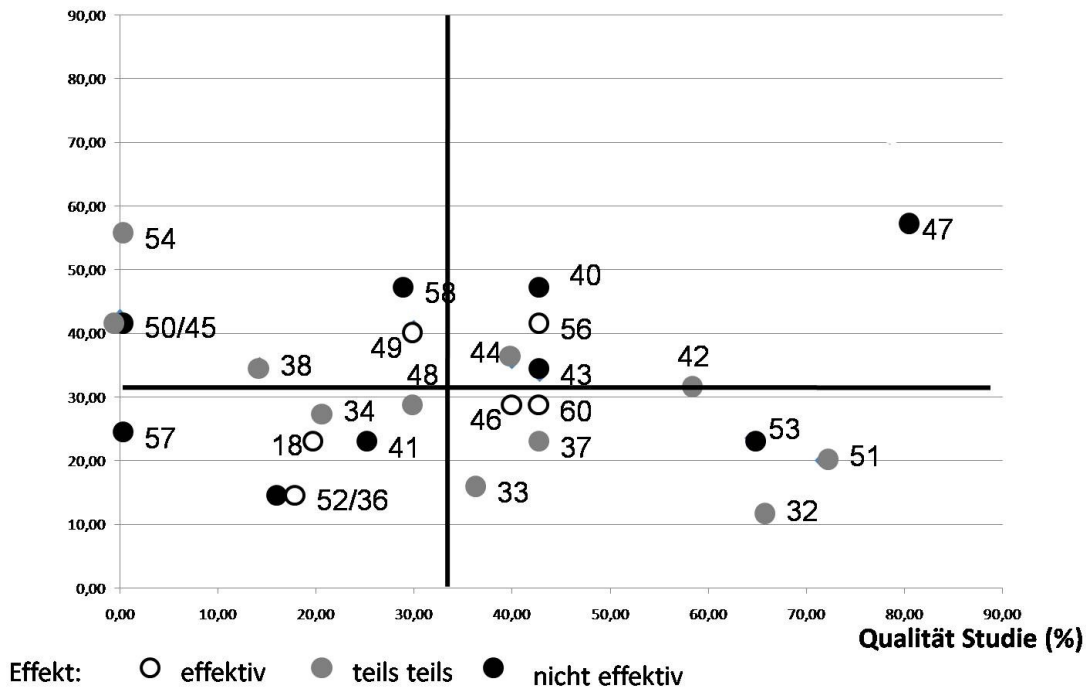
Legende: Aufgetragen wurden die Scores für Qualität der Studiendurchführung (x-Achse) und der Qualität der Intervention (y-Achse) nach Bewertung.

### Darstellung der Sensitivitätsanalyse

Auch für den Bereich Bewegung im Betrieb wurde die Qualität der Intervention auch ohne Empowerment und Partizipation berechnet. Der neue Durchschnitt ergab 31,5% für die Qualität der Intervention. Im Bereich Betrieb kommt eine weitere Studie, nämlich 1043, auf Grund der Sensitivitätsanalyse in den Bereich von eingeschränkt zu empfehlenden Studien; die Studie 1042 liegt mit 31% knapp unter dem Durchschnitt für die Interventionsqualität.

Abbildung 12: Sensitivitätsanalyse Bewegung Betrieb

**Qualität Intervention (%)**



## 6.2.4 Darstellung einzelner Studien (Bewegung Betrieb)

### *Aittasalo et al. 2004 (1040)*

Land: Finnland.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: Es wurden die Interventionen Beratung und Beratung plus Fitnesstest mit jeweils 52 Teilnehmenden einer Kontrollgruppe mit ebenfalls 52 Personen gegenübergestellt.

Intervention: In sieben sog. occupational health centres mit insgesamt acht Betriebskrankenschwestern in neun Betrieben wurde Beratung durchgeführt: Interventionsgruppe 1 erhielt nur Beratung, Interventionsgruppe 2 bekam zusätzlich einen initialen Fitnesstest. Studiendauer: 12 Monate.

Nachbeobachtung: keine.

Effektmaße: Anzahl und Zeit der Bewegungseinheiten, Schrittzahl pro Tag, Zeit im Sitzen

Effekte und p-Werte: alle drei Gruppen haben Steigerung in allen Effektmaßen, so dass bis auf Zeit im Sitzen zwischen Interventionsgruppe 1 (nur Beratung) und Interventionsgruppe 2 (Beratung und Fitnessstest) keine signifikanten Unterschiede. Die Zeit im Sitzen ist in der Interventionsgruppe 1 signifikant höher mit 7,5 bis 43,9% bei  $p=0,006$ , allerdings werden keine absoluten Angaben gemacht.

Besondere Stärken oder Schwächen: Die Randomisierung hat innerhalb der neun teilnehmenden Firmen stattgefunden, so dass Einflüsse der Interventionsgruppen auf die Kontrollgruppe möglich sind, wenn allein schon über die Testung oder die Beratung gesprochen wird.

Empfehlung und Übertragung: Weder isoliertes Testen der Fitness noch Beratung und Testen der Fitness zeigen einen Effekt gegenüber dem Nichtstun. Auch wenn durch die Art der Randomisierung ein Einfluss auf die Kontrollgruppe nicht ausgeschlossen werden kann, ist die Qualität der Durchführung ein Hinweis darauf, dass die Interventionen hinlänglich als nicht effektiv erwiesen worden sind.

### ***Gerdle et al. 1995 (1043)***

Land: Schweden.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 32 versus 45.

Intervention: 2 x Woche eine Stunde Fitness mit versch. Elementen wie Kraft, Beweglichkeit, Ausdauer, Fitness. Für die Teilnahme wird keine Arbeitsbefreiung gewährt.

Studiendauer: 12 Monate.

Nachbeobachtung: keine.

Effektmaße: verschiedene Indizes zu muskuloskeletalen Beschwerden etc., Arbeitszufriedenheit, Gewicht, Bewegungsumfang, Ausdauerleistung anhand des Sauerstoffverbrauchs.

Effekte und p-Werte: Keine statistisch signifikanten Veränderungen, auf Grund des Programms signifikante Verschlechterung in der Interventionsgruppe bei Indizes zu Einschätzung der Führung, des eigenen Einflusses sowie der Ergonomie.

Stärken und Schwächen: Alle Teilnehmerinnen kamen aus einem Betrieb, damit wurde die Randomisation konterkariert durch möglichen Austausch der Teilnehmenden. Der Stress nahm zu, da die Teilnehmenden nicht von der Arbeit befreit wurden.

Empfehlung und Übertragung: Auf jeden Fall zeigt die Studie, dass Bewegung im Betrieb durch entsprechende Arbeitsbefreiung unterstützt werden muss. Auch andere Interventionsstudien zeigen zum Teil zunächst negativen Effekt auf Einschätzung der eigenen Situation, so dass dieser Effekt über längere Zeit untersucht werden muss.

***Marshall et al. 2002 (1044)***

Land: Australien.

Studiendesign: non-CT.

Größe der Interventionsgruppe: keine Angabe zur Anzahl der Mitarbeiter, aber 158350 Infrarot-kamerakontakte beim Treppenaufstieg.

Studiendauer: 1 Monat (in 2 Abschnitten zu je 0,5 Monaten).

Nachbeobachtung: 0,75 Monate.

Effektmaße: Treppenaufstieg in % gegen Baseline durch objektive Messung und subjektive Befragung (ex post).

Effekte und p-Werte: Zunahme des Treppensteigens um 1% in einer ersten Interventionsphase, dann stärkere Abnahme und kein Wiederanstieg in zweiter Interventionsphase ( $p = 0,02$ ), im Trend Anstieg der Treppenbenutzung unter den Befragten ( $n = 53$  vorher, 40 nachher), aber in der letzten Kontrollphase objektiv gemessen geringeres odds ratio, die Treppe zu benutzen.

Besondere Stärken und Schwächen: Kürze der Intervention. Positiv ist hervorzuheben, dass die eigene Intervention nochmals getestet wurde und ein wenn auch sehr kurzer Nachbeobachtungszeitraum einbezogen wurde. Der erste positive Effekt ist minimal, Im Ansatz gut ist weiterhin, dass nur der Aufstieg, nicht der Abstieg als Effektmaß aufgenommen wurde, da der Abstieg keinen nachgewiesenen bzw. gegenüber dem Aufsteigen vernachlässigenswerten Trainingseffekt hat.

Empfehlung und Übertragung: Steigerung der Treppenbenutzung mittels Zeichen/ Poster scheint als Intervention isoliert wenig ratsam, auch wenn subjektiv eine Steigerung gemessen werden konnte.

***Adams et al. 2002 (1047)***

Land: Großbritannien.

Studiendesign: non-CT.

Größe der Interventionsgruppe: Baseline und Woche eins und Woche vier mit insgesamt 1750, 1770 und 1773 Beobachtungen zu Treppensteigen vs. Aufzugbenutzung.

Intervention: Nach der Befragung der Zielgruppe und daraufhin konzipierten Postern wurde diese in einem fünfstöckigen Gebäude einer med. Fakultät aufgehängt. Die Benutzung der Treppen sollte damit gesteigert werden.

Dauer der Intervention: Ein Monat.

Nachbeobachtung: keine.

Effektmaße: Treppennutzung.

Effekte und p-Werte: Keine signifikante Veränderung, aber Abbruch, da 29 der 39 Poster innerhalb des ersten Monats verschwunden, auch wenn sie z. T. als S/W ersetzt wurden.

Besondere Stärken und Schwächen: In die Planung wurden Personen aus der Zielgruppe einbezogen. Eine qualitative Befragung zu den Postern wurde durchgeführt, nachdem diese mutwillig entfernt worden waren. Dabei ergab sich u. a., dass die Poster eher ein schlechtes Gewissen machen als dass sie motivieren würden.

Empfehlung und Übertragung: Studie kombiniert Intervention mit Prozessevaluation und –dokumentation und gibt somit eine qualitativ hochwertige Datenbasis ab, auf Grund derer die Intervention Treppenbenutzung in Betrieben kaum ein effektives Mittel ist, Aktivität der Mitarbeiter bzw. Studierenden zu steigern.

### ***Harrell et al. 1996 (1056)***

Land: USA.

Studiendesign: RCT.

Interventions- und Kontrollgruppe: 752 zu 752 Personen.

Intervention: Wellness- und Fitnessprogramm mit vier Stunden Vortrag zu Gesundheit, Ernährung und Fitness, zwölf Stunden Test und 27 Stunden überwachtem Training in 25 Polizeistationen in North Carolina.

Studiendauer: 2,25 Monate.

Nachbeobachtungszeit: keine.

Effektmaße Anteil Körperfett, Sauerstoffkapazität, Stärke allgemein, Anzahl der Situps, Beweglichkeit allgemein.

Effekte und p-Werte: Kontrollgruppe versus Interventionsgruppe -1,2 vs. -5,6 Körperfettanteil ( $p= 0,001$ ), Sauerstoffkapazität +13,4 vs. +21,5 ( $p= 0,001$ ), Kraft kein Effekt, Beweglichkeit +7,4 vs. + 8,3 ( $p= 0,014$ ), Situps 16,4 vs. 26 ( $p= 0,001$ )

Stärken und Schwächen: Im Projekt geht es kaum um Motivation, sondern eher um Effekt des Programms. Programm wurde in den einzelnen Polizeistationen von Peers durchgeführt, die dazu während einer Weiterbildung ausgebildet wurden. Diese Multiplikatoren wurden kontrolliert während einer der Trainingseinheiten an ihren Polizeistationen (Prozessevaluation).

Empfehlung und Übertragung: Möglicherweise muss Bewegung einfach während der Arbeitszeit „gemacht“ werden.

## **6.3 Diskussion**

### **6.3.1 Frauen und Mädchen und Bewegung**

Von den sechs ausgewählten Studien hatten zwei einen positiven Effekt, drei je nach Effektmaß ein unterschiedliches Ergebnis und eine einen negativen, sprich keinen Effekt.

Zwei der sechs ausgewählten Studien zeigen ein multistrategisches Vorgehen mit Einschluss vieler Institutionen (1006, 1055), einem settingorientierten Ansatz und Veränderungen der Umwelt (Gemeinde bzw. Schule). Die Studie 1006 ist erfolgreich, die Studie 1055 in Teilen, da für das Effektmaß Gewicht keine Veränderung zu beobachten war (s. oben dazu). Auch die Studie von Fahrenwald (1039), wenn auch nur mit kleiner Zahl von Personen durchgeführt, kann wahrscheinlich den Erfolg darauf zurückführen, dass die Intervention innerhalb eines bestehenden umfassenden Programms für allein stehende Frauen mit Kindern aus sozial schwachen Verhältnissen durchgeführt wurde.

Für die Übertragung eignet sich insbesondere die Studie 1006. Diese Studie umfasst Verhaltens- und Verhältnisveränderungen und integriert die politischen und administrativen Einrichtungen einer Gemeinde. Diese sind in vergleichbarer Form auch in Deutschland gegeben. Die Studie 1008 dagegen zeigt, dass verhaltensverändernde Prävention ohne entsprechende Veränderung der Verhältnisse und Inklusion der verschiedenen Entscheidungsträger wenig Effekte zeigt. Dem gegenüber belegt die Studie 1055 in Verbindung mit der Teilstudie 1010, dass dies wichtige Voraussetzungen zum Gelingen einer schulbasierten Intervention für Mädchen ist. Bei allen anderen

Studien müsste vor der Übertragung geklärt werden, in wie fern die sich widersprechenden Ergebnisse einer Übertragung widersprechen.

Bei aller Vorsicht in der Beurteilung auf Grundlage unserer schmalen Datenbasis zeichnet sich wohl der Trend ab, dass man für den Bereich Bewegung von Frauen und Mädchen sagen kann, dass die erfolgreichen Studien bzw. Interventionen, die zugleich die Qualitätsstandards erfüllen, durch ein multistrategisches Vorgehen, die Beteiligung vieler Akteure und möglichst die Partizipation der Zielgruppe durch vorherige Befragungen als Erfolgsfaktoren gekennzeichnet sind.

### **6.3.2 Betrieb und Bewegung**

Drei der ausgewählten Studien mit entsprechender Qualität in der Studiendurchführung und der Intervention haben ein negatives Ergebnis, eine ein teils-teils Ergebnis mit negativem Effekt und eine Studie hat einen positiven Interventionseffekt.

Insbesondere die Studien zu Treppensteigen im Betrieb bestätigen die Probleme, die in einem Review von Eves und Webb (2006) angesprochen wurden. Die in den Bereich Empfehlung fallenden beiden Interventionen haben ein negatives bzw. gemischtes Ergebnis. Insbesondere die Studie von Marshall et al. (2002) zeigt erneut, wie sich objektive und subjektive Ergebnisse widersprechen können. Im Grunde müsste die Studie daher sogar als gescheitert angesehen werden. Zwar liegen weitere Studien zum Treppensteigen vor, u. a. auch Titze et al. (2001), aber alle diese Studien haben das Auf- und Absteigen nicht getrennt, so dass schon auf Ebene der Effektparameter die Frage ist, soll einfach nur mehr Bewegung, wenn Personen z. B. nur mehr treppab gehen, aber nicht treppauf, oder soll auch ein entsprechend höherer Energieverbrauch erreicht werden.

### **6.3.3 Diskussion**

In Zukunft sollten sich Planer von Studien bzw. Interventionen besser auf wenige Effektmaße, Endpunkte, Methoden und Instrumente einigen, um die Vergleichbarkeit sowie die Möglichkeit der übergreifenden Bewertung von Effekten zu erleichtern bzw. überhaupt erst zu gestatten.

Empowerment und Partizipation ändern an der Bewertung der betrachteten Studien wenig, ebenso auch die Berücksichtigung von Qualitätsinstrumenten, da sie nur sehr selten eingesetzt worden sind. Diese Qualitätskriterien laufen derzeit faktisch leer, können aber in Zukunft zur Diskrimi-

nierung von Studien beitragen, für die eine höhere Gewissheit angenommen werden kann, dass ihr Ergebnis verlässlich wiederholbar ist.

### **6.3.4 Fazit für beide Handlungsfelder**

Im Handlungsfeld Bewegung im Betrieb hat sich eher eine Negativliste herauskristallisiert. Im Handlungsfeld Frauen und Mädchen gibt es zwei gute Vorbilder für in Deutschland mögliche Interventionen aus Australien (Frauen zwischen 20 und 50) und den USA (Mädchen in der Sekundarstufe I). Beide Interventionen arbeiten multistrategisch im Setting und beziehen verschiedene Akteure (Gemeinde, Familien etc.) mit ein.

# 7 Depression

## 7.1 Stand der Forschung

Standen bei älteren Krankheitsbegriffen somatische Erkrankungen im Vordergrund, sind im Spektrum der dominierenden Krankheiten neuerer Krankheitsbegriffe Depressionen zunehmend in den Vordergrund gerückt. Die Global Burden of Disease Study der WHO kam für 2000 zu dem Ergebnis, dass Depressionen mit einem Anteil von 4,5% der disability adjusted life years (DALYs) zu den Krankheiten mit der größten Krankheitslast zählen. 12,1% der in Krankheit verbrachten Lebensjahre (YLD: Years lived with disabilities) sind gemäß dieser Studie den Depressionen zuzurechnen. 1990 hatten die entsprechenden Werte noch 3,7% (DALYs) bzw. 10,7% (YLD) betragen. Üstun et al. (2004) Die WHO geht von einem weiteren Anstieg der Häufigkeit depressiver Erkrankungen aus, sodass Depressionen im Jahr 2020 weltweit die zweithäufigste Volkskrankheiten nach den ischämischen Herzerkrankungen, in den sogenannten entwickelten Staaten sogar die häufigste Krankheit sein könnten (Murray und Lopez 1997).

Depressionen zählen nicht nur zu den Erkrankungen mit der größten Krankheitslast in der Bevölkerung, sondern sind auch mit einer hohen somatischen Gesamtmorbidität assoziiert (Kapfhammer 2007). Die Wechselwirkung von Depressionen und körperlichen Erkrankungen wird daher zunehmend diskutiert (Hinz und Schwarz 2001). In diesem Sinne kommt der Depression im Rahmen der vorliegenden Untersuchung eine doppelte Bedeutung zu. Auf der einen Seite handelt es sich bei Depressionen um ein Krankheitsbild, welches in verschiedenen Ausprägungen auftritt. Auf der anderen Seite stellen Depressionen einen Risikofaktor für die physische Gesundheit dar. Mit der Prävention der Depression wird damit gleichzeitig auch den somatischen Krankheitsfolgen der Depression entgegengewirkt.

### **Prävalenz**

Epidemiologische Studien belegen, dass depressive Störungen in der Allgemeinbevölkerung häufig vorkommen. Nach einer WHO-Schätzung liegt die altersstandardisierte Prävalenz weltweit für Männer bei 1.607 Fällen pro 100.000 Personen, für Frauen bei 2.552 pro 100.000. Die Schätzungen zeigen große regionale Unterschiede mit höheren Prävalenzen in den entwickelten Ländern (Murray und Lopez 1997).

Im National Comorbidity Survey wird für die US-amerikanische Bevölkerung eine Lebenszeitprävalenz von rund 17% ermittelt (Kessler et al. 1994). Für die Länder der EU liegt die 12-Monats-Prävalenz der Depression bei 5–8%. Aufgrund des episodenhaften Auftretens der Krankheit liegt die entsprechende Lebenszeitprävalenz mit 12–16% deutlich höher (Andlin-Sobocki und Wittchen 2005.) Die höhere Prävalenz für die USA wird auch durch die Global Burden of Disease Study gestützt. Anders als für andere psychischen Störungen, sind in den letzten 10 Jahren in Europa steigende Raten der Depression (und Suchterkrankungen) festzustellen (Wittchen und Jacobi 2005).

In der Gesamtpopulation einer für Deutschland repräsentativen Studie wurden für 15,8% der Teilnehmer in einem Screeningverfahren erhöhte Depressionswerte, bei starker Alters- und Geschlechtsabhängigkeit festgestellt. Da es zur Diagnosestellung dieser Screeningergebnisse zunächst einer diagnostischen Abklärung bedarf, können diese Werte jedoch nicht als Prävalenz interpretiert werden (Hinz und Schwarz 2001). Die Prävalenz innerhalb eines Jahres liegt in Deutschland für die Bevölkerung im Alter von 16–65 Jahren bei rund 11%. Davon leiden 4 % bereits zum wiederholten Male an depressiven Episoden (Bramsfeld und Schwartz 2007; Wittchen und Jacobi 2006) Der überwiegende Teil der psychischen Störungen manifestiert sich in der Kindheit und Adoleszenz, wird jedoch selten früh erkannt und adäquat behandelt. (Wittchen und Jacobi 2005) Gemäß epidemiologischen Studien leiden schon 15–20% der unter 18jährigen mindestens an einer depressiven Episode, wobei die Prävalenz ab dem 13. Lebensjahr deutlich ansteigt (Birmaher et al. 1996; Wade et al. 2002).

## **Folgen**

Das Risiko, nach einem ersten Auftreten einer Depression erneut depressive Episoden zu erleiden ist hoch. Innerhalb von 5 Jahren erleiden rund 75% der Betroffenen eine zweite Erkrankungsepisode (Bramsfeld und Schwartz 2007). In den seltensten Fällen treten Depressionen dabei allein auf. Vielmehr gehen sie mit anderen psychischen Erkrankungen einher (Wittchen und Jacobi 2005. Bei verschiedenen Volkskrankheiten werden durch Depressionen zudem Chronifizierungsprozesse verstärkt und Morbidität sowie Mortalität erhöht (GVG 2006). So ist z.B. die höhere Mortalität depressiver Menschen an kardiovaskulären Erkrankungen von Bedeutung. Ebenso wurde ein höheres Risiko Depressiver, eine KHK zu entwickeln, festgestellt (Wulsin et al. 1999, Ferketich et al. 2000; Barth et al. 2004).

Aber nicht nur die höhere Mortalität bei chronischen Krankheiten führt zu einer insgesamt höheren Sterblichkeit depressiver Menschen. Zum einen haben depressive Menschen ein erhöhtes Risiko, an Unfällen zu versterben (Joukaama et al. 2001). Zum anderen besteht zwischen Depression und Suizidalität ein enger Zusammenhang. 30% der depressiv Erkrankten haben Suizidversuche begangen. Der Anteil der Depressiven die in akuten Phasen der Erkrankung unter Suizidgedanken leiden liegt mit 60–70% noch deutlich höher. (Möller 2003) Schätzungen zufolge begehen 3-4% der depressiv erkrankten Menschen im Laufe ihres Lebens einen Suizid. Bei Personen, die aufgrund schwerer Depressionen stationär behandelt wurden, liegt die Wahrscheinlichkeit sogar bei 15% (Statistisches Bundesamt 1998).

Aufgrund ihrer psychischen und sozialen Belastung für Betroffene und ihre Angehörigen zählen Depressionen zu den Erkrankungen mit erheblichem Einfluss auf Gesellschaft und Gesundheitswesen.

### **Kosten**

Depressionen sind mit hohen Kosten verbunden, die sowohl durch direkte Behandlungskosten als auch durch indirekte Kosten verursacht werden. Untersuchungen für die USA kommen zu dem Ergebnis, dass der Anteil der indirekten Kosten an den Gesamtkosten ein Vielfaches der direkten Kosten ausmacht. Einer Langzeitstudie zufolge betragen die reinen Behandlungskosten nur 8,1% der Gesamtkosten der Depression (Booth et al. 1997; Greenberg et al. 2003). Zu dieser Relation trägt der hohe Anteil der Produktivitätsausfälle aufgrund frühzeitiger Berentung aber auch der hohe Anteil der durch Depressionen verursachten Arbeitsunfähigkeitstage (AU-Tage) bei. In Deutschland haben z.B. zwischen 1997 und 2004 bei insgesamt rückläufigem Krankenstand die durch psychische Erkrankungen bedingten AU-Tage um 70% zugenommen (Weber et al. 2006).

Für die USA liegen schon seit längerem Kostenanalysen der Depression vor. In den verschiedenen Untersuchungen unterscheiden sich die Kosten nicht zuletzt aufgrund unterschiedlicher Studienansätze deutlich. Auch angesichts der verschiedenen Systeme lassen sich US-Daten nicht auf Europa und Deutschland übertragen. Vor diesem Hintergrund sind in jüngerer Zeit auch Kosten-Studien für Europa und Deutschland durchgeführt worden. Danach belaufen sich die in Kaufkraftparitäten gemessenen Kosten der Depression je Fall in 2004 für Europa auf 3.826 €. Der entsprechende Wert für Deutschland beträgt gemäß der Studie 7.102 € (Andlin-Sobocki et al.

2005). Von den für Europa insgesamt ermittelten Kosten für affektive Störungen entfallen rund 27% auf die direkten Kosten. Zu den Kosten der Depression liegen verschiedene Untersuchungen für Deutschland vor. Das Kompetenznetz Depressionen geht davon aus, dass Depression in Deutschland Gesamtkosten von rund 17 Mrd. € verursachen, so dass bei 4 Mio. Behandlungsbedürftigen in allen Altersgruppen im Jahr durchschnittliche Kosten von 4.250 € anfallen (Pössel, Schneider et al. 2006). Das Statistische Bundesamt hat für das Jahr 1994 direkte Kosten der Depression von 1,3 Mrd. Euro ermittelt. Dies entsprach 0,7% der gesamten Gesundheitsausgaben (Statistisches Bundesamt 1998). Eine neuere Studie kommt für 2002 auf 1,6 Mrd. Euro direkter Behandlungskosten (Friemel et al. 2005). Die durchschnittlichen direkten Kosten pro Patient werden in der Literatur zwischen 686 € und 2.541 € beziffert (Friemel et al. 2005; Salize et al. 2004).

### **Internationale und nationale Evidenz**

In den letzten zwei Dekaden hat das Interesse an Interventionen zur Prävention depressiver Erkrankungen zugenommen (Harrington und Clark 1998; Sheffield et al. 2006). Jané-Llopis et al. sehen in ihrem Review Konzepte der Prävention von Depressionen als einen Baustein in einer umfassenden Entwicklung verschiedener Ansätzen zur Förderung der psychischen Gesundheit. Die Mehrzahl der Präventionsprogramme zur Depression knüpft an den persönlichen Fähigkeiten der Individuen an und zielt durch relativ kurze Interventionen auf die Stärkung persönlicher Fähigkeiten und der Resilienz. In einer Metaanalyse von Durlak und Wells zu Präventionsprogrammen bei Kindern und Jugendlichen zeigen verhaltenstherapeutische und kognitiv-verhaltenstherapeutische Programme gegenüber anderen Präventionsprogrammen doppelt so hohe Effektstärken (Durlak und Wells 1997).

Gegenwärtig wurde der Einfluss solcher Programme aber erst in relativ wenigen Studien untersucht. Dabei konzentrierten sich zunächst viele Studien auf die selektive und indizierte Prävention, bei der Personen mit hohem Risiko bzw. Symptomträger an Präventionsprogrammen teilnahmen (Jané-Llopis et al. 2005; Clark et al. 1995). Noch geringer war die Zahl der Studien, in der die Frage untersucht wurde, ob Präventionsprogramme geeignet sind, die Inzidenz neuer Fälle psychischer Erkrankungen zu reduzieren. In eine entsprechende Metaanalyse wurden ausschließlich selektive bzw. indizierte Präventionsprogramme einbezogen (Cuijpers et al. 2005).

In einer Metaanalyse zu universalen und gezielten (d.h. selektiven und indizierten) Präventionsprogrammen betonen Merry et al. (2003) die unzureichende methodischen Ansprüche der insgesamt durchgeführten Studien. Genannt werden z.B. die geringe Studiengröße und die zu kurzen Nachbeobachtungen. Für die universale Prävention sehen sie auf der Basis von vier einbezogenen Studien keine Effektivität für die Prävention der Depression. Die Schlussfolgerungen für die fünf berücksichtigten Studien der selektiven bzw. indizierten Prävention fallen in dem Cochrane-Review mit einem kleinen, aber signifikanten Effekt unmittelbar nach der Intervention etwas günstiger aus. Aber auch dieser Effekt war nach einem zwölfmonatigen Follow up nicht länger signifikant. Nach Erscheinen der Metaanalyse durch Merry et al. in 2003 sind weitere universalpräventive Studien mit geringerem, aber signifikantem Effekt publiziert worden (Pössel et al. 2004; Merry et al. 2004). Eine weitere Metaanalyse kommt zu dem Ergebnis, dass universale, selektive und indizierte Präventionsprogramme vergleichbare Effekte aufweisen (Jane-Llopis 2002 zitiert nach Pössel 2006a). Eine andere Metaanalyse betont die Überlegenheit selektiver und indizierter Prävention. Dies wird damit begründet, dass aufgrund des geringeren Risikos in Studien der universellen Prävention Teilnehmer der Kontrollgruppe weniger Symptome entwickeln (Horowitz und Garber 2006).

## **7.2 Ergebnisse**

### **7.2.1 Allgemeine Darstellung der ausgewählten Studien**

Die der nachfolgenden Analyse zugrundeliegende Literaturrecherche erbrachte über alle Datenbanken und alle Länder insgesamt 5.005 Treffer. Nach Berücksichtigung der Ein- und Ausschlusskriterien blieben rund 10% der Publikationen übrig. Dies waren 457 Publikationen. Die Vielzahl dieser Treffer mache eine Einschränkung der zu evaluierenden Studien erforderlich. Die hohe Trefferzahl war aber nur ein Grund für die Auswahl. Analysen, welche Projekte über alle möglichen Interventionen zusammen auswerten, verdecken letztlich mehr als sie offenlegen.

Innerhalb des Handlungsfeldes Depression konnten verschiedene Schwerpunkte identifiziert werden, die eine gewisse Homogenität in dem prinzipiell weitreichenden Feld erkennen ließen. Themenfelder im Handlungsfeld Depression waren Depressionen bei Kindern und Jugendlichen, die Wahrnehmung der Depression in der Öffentlichkeit (Abbau von Stigmata und Vorurteilen),

Depressionen im Alter, Prävention der Selbsttötung, Prävention der postpartalen Depression, die Qualität der Primärversorgung Depressiver sowie die Prävention von Depressionen über das Internet.

Wurde ursprünglich angenommen, dass Depressionen nur bei Erwachsenen auftreten, ist inzwischen erkannt, dass ein beachtlicher Teil der Depressionen schon im Kindes- und Jugendalter auftritt (Costello et al. 1996). Insgesamt sind Depressionen stark alters- und geschlechtsabhängig (Hinz und Schwartz 2001). Angesichts der zunehmenden Prävalenz ab dem 13. Lebensjahr und der auch bei Jugendlichen zunehmenden Inanspruchnahme medizinischer Leistungen aufgrund depressiver Störungen wird der Prävention im Kindes- und Jugendalter zunehmend Aufmerksamkeit geschenkt (GVG 2006; Grobe und Dörning 2003). Andrews und Kollegen folgern angesichts der Häufigkeit der ersten Episode der Depression vor dem 21. Lebensjahr, dass erfolgreiche Prävention bei jungen Menschen einsetzen muss (Andrews et al. 2002). Auch in der vorliegenden Analyse wird daher auf Präventionsstudien bei Kindern und Jugendlichen fokussiert. Dazu wurden die Studien universeller Prävention bei Kindern und Jugendlichen anhand des Kriterienkataloges evaluiert. D.h., es geht in diesen Projekten um die Prävention der Depression bei solchen Jugendlichen, die kein erhöhtes Risiko aufweisen und auch nicht zu den Symptomträgern zählen. Da Angstsymptome ein zentraler Risikofaktor für Depressionen sind, werden auch Studien zur Prävention der Angstsymptomatik in die Analyse einbezogen. Zahlreiche Studien fokussieren sowohl auf die Prävention der Angst als auch der Depression, so dass die Berücksichtigung beider Studienziele nur konsequent erscheint.

In der Literatur werden die Vorteile und vergleichsweise geringen Nachteile der universalen Prävention bei Kindern und Jugendlichen gegenüber der selektiven und indizierten Prävention betont (Pössel, Schneider et al. 2006b; Shochet et al. 2001). Die universalen Präventionsprogramme beziehen ganze Bevölkerungsgruppen unabhängig von ihrem individuellen Risiko ein, wodurch z.B. die Erfordernis eines Screenings entfällt. Zudem wird eine mögliche Stigmatisierung der Teilnehmer vermieden.

Von den in der Literaturrecherche ermittelten Treffern entfielen 18 Studien auf Programme der universalen Prävention bei Kindern und Jugendlichen. Davon waren rd. 78% randomisierte kontrollierte Studien (RCTs). Der Rest entfiel auf kontrollierte Studien (CTs). Alle Programme wurden im Setting Schule durchgeführt.

Die folgende Tabelle 17 gibt einen Überblick über die in die Analyse einbezogenen Studien.

Tabelle 17: Übersicht über die eingeschlossenen Studien im Handlungsfeld Depression

Kennnummer	Autor (Jahr)	Projektname bzw. Intervention	Land	Design
2001	Merry et al. (2004)	RAP-KiWi	Neuseeland	RCT
2002	Pössel et al. (2004)	LISA-T	Deutschland	RCT
2003	Groen et al. (2003)	LARS & LISA	Deutschland	CT
2004	Pössel et al. (2003)	LISA-T	Deutschland	CT
2005	Pössel, Horn et al. (2006)	LISA-T & LISA-S	Deutschland	CT
2006	Lowry-Webster et al. (2003)	FRIENDS	Australien	RCT
2007	Shochet et al. (2001)	RAP-A & RAP-F	Australien	CT
2008	Barrett et al. (2006)	FRIENDS	Australien	RCT
2009	Barrett und Turner (2001)	FRIENDS for Children	Australien	RCT
2010	Sheffield et al. (2006)	---	Australien	RCT
2012	Pössel et al. (2005)	---	Deutschland	RCT
2013	Clarke et al. (1993)	---	USA	RCT
2014	Clarke et al. (1993)	---	USA	RCT
2015	Kenardy et al. (2003)	---	Australien	RCT
2016	Spence et al. (2003)	Problem Solving for Life Program	Australien	RCT
2017	Oria et al. (2001)	Youth Leadership class	USA	CT
2018	Dubow et al. (1993)	Problem Solving Skill Training	USA	RCT
2019	Manz et al. (2001)	GO Gesundheit und Optimismus	Deutschland	RCT

Auch wenn es sich bei diesen Studien um Studien mit meist kognitiv-verhaltenstherapeutischen Interventionen handelt, unterscheiden sich die in den Studien herangezogenen Effektmaße deutlich. Dies gilt sowohl für die Art als auch für die Menge der berücksichtigten Maße. So kann die Depressions- bzw. Angstsymptomatik über verschiedene Skalen gemessen werden. Einige Studien verwenden mehrere dieser Skalen. Darüber hinaus werden Effektmaße z.B. für die verfügbare soziale Unterstützung (Social support) oder auch für automatisch negative Gedanken (ATQ) einbezogen. Bei Studien mit einer großen Zahl an Effektmaßen werden in die Analyse nur aus-

gewählte Maße einbezogen. Als effektiv werden dann auch nur solche Studien bewertet, bei denen im Hinblick auf alle *ausgewählten* Parameter eine signifikante Verbesserung durch die Intervention belegt werden konnte. Sind nur signifikante Verbesserungen einzelner Parameter belegt, werden die Studien als teils-teils effektiv bewertet. Tabelle 18 gibt einen Überblick über ausgewählte Effektparameter.

Tabelle 18: Ausgewählte Effektmaße im Handlungsfeld Depression

Kenn-Nummer	CES-D	ADS	BDI	BAI	BHS	CDI	ADIS-C	DSM-IV	ASI	ATQ:	SCAS	CDR-R	RCMAS	RADS
2001			x											x
2002	x									X				
2003		X												
2004		X								X				
2005		X												
2006						x					x		x	
2007					x	x								x
2008						x					x		x	
2009						x					x		x	
2010	x				x	x	x				x			
2012	x									X				
2013	x													
2014	x													
2015	x								x					
2016			x				x	x						
2017						x								
2018														
2019			x	x										

Effektmaße: CES-D: Center for Epidemiological Studies Depression Scale; ADS: Allgemeine Depressions Scala; BDI: Beck Depression Scale; BAI: Beck Anxiety Index; BHS : Beck Hopelessness Scale; CDI: Childhood Depression Inventory; ADIS-C: Scala für Angststörungen; DSM-IV: Dysthymia; ASI: Anxiety Sensitivity Index; ATQ: Automatic Thought Questionnaire; SCAS: Spence Children Anxiety Scale; CDR: Clinical Symptoms of Depression; RCMAS: Childrens Manifest Anxiety Scale; RADS: Reynolds Adolescent Depression Scale.

Neben der Effektivität wird in der Analyse die Qualität der Studien evaluiert, wobei Interventions- und Studienqualität unterschieden werden. Die nachfolgenden Tabellen zeigen den Gesamtscore der in den Studien erreichten Interventions- (Tabelle 19) bzw. Studienqualität (Tabelle 20) und geben die Qualitätspunkte für ausgewählte Parameter der Intervention wieder.

Tabelle 19: Qualität der Studiendurchführung im Handlungsfeld Depression

Kenn-Nummer	Studien-design	Rando-mi-sation (16)	Ver-gleich-barkeit (19)	Me-ßinstru-mente validiert (25)	Power (46)	Summe	Erzielbarer Maximal-Score	Prozen-tualer Score
2001	RCT	1	1	1	1	7	7	100
2002	RCT	1	1	0	0	4,5	7	64
2003	CT	X	0	0,5	0	3	6	50
2004	CT	X	0	0	0	2	6	33
2005	CT	X	0	0	0	2,5	6	42
2006	RCT	1	0	0	0	3	7	43
2007	CT	X	1	0,5	0	4	6	67
2008	RCT	1	0	0,5	0	4	7	57
2009	RCT	1	1	0	0	4,5	7	64
2010	RCT	1	1	0	1	5,5	7	79
2012	RCT	1	0	0	1	4,5	7	64
2013	RCT	0	1	1	0	4	7	57
2014	RCT	0	1	1	0	4	7	57
2015	RCT	1	0	0,5	0	4	7	57
2016	RCT	1	1	0,5	0	3	7	43
2017	CT	X	0	0	0	2	6	33
2018	RCT	1	0	0	0	3,5	7	50
2019	RCT	1	0	0	0	2	7	29

Tabelle 20: Qualität der Intervention im Handlungsfeld Depression

Kennnummer	Studien-design	Parti-zipation (54)	Em-powerment (55)	Prozess-doku-mentiert (63)	Prozess-evaluiert (64)	Summe	Erziel-barerMa-ximal-score	Pro-zentu-aler Score
2001	RCT	0	1	1	1	7,5	10,5	71
2002	RCT	0	1	1	1	7	10,5	67
2003	CT	0	1	0	1	5	10,5	48
2004	CT	0	1	0	0	4,5	10	45
2005	CT	0	1	0	0	4,5	10	45
2006	RCT	0	1	1	1	6	10,5	57
2007	CT	0	1	1	1	6	10,5	57
2008	RCT	0	1	1	1	7	10,5	67
2009	RCT	0	1	1	1	6	10,5	57
2010	RCT	0	1	0	0	4	10,5	38
2012	RCT	0	1	1	1	6,5	10	65
2013	RCT	0	0	0	0	3	10,5	29
2014	RCT	0	1	1	1	5,5	9,5	58
2015	RCT	0	1	1	1	5	10,5	48
2016	RCT	0	1	1	1	6	10,5	57
2017	CT	0	1	0	0	4	10,5	38
2018	RCT	0	1	1	1	5	10	50
2019	RCT	0	1	0	0	3,5	9,5	37

Die nachfolgende Abbildung 13 veranschaulicht die in den einzelnen Studien erreichte Studien- bzw. Interventionsqualität in Relation zum Effekt.<sup>11</sup> Von den insgesamt 18 Studien zur universalen Prävention der Depression bei Kindern und Jugendlichen, die anhand des Kriterienkataloges evaluiert wurden, sind neun nicht effektiv, acht erreichen nur eine teilweise Effektivität und eine Studie ist effektiv. Von allen Studien liegen elf in dem aufgrund des unterdurchschnittlichen Qualitätsscores als nicht empfehlenswert definierten Bereich. Dazu zählt auch die einzige in Hinblick auf alle ausgewählten Effektmaße effektiv bewertete Studie (Lowry-Webster et al. 2001). Von den sieben Studien mit überdurchschnittlichen Qualitätsscores Studien fällt eine in den potenziell eingeschränkt empfehlenwerten Bereich und sechs erreichen mit Scores der Interventions- und Studienqualität zwischen 60 und 80% den stark empfehlenswerten Bereich. Von diesen Studien sind fünf teilweise effektiv. Zwei erreichen bei überdurchschnittlicher Qualität keine Effektivität.

<sup>11</sup> In der Abbildung sind im Interesse einer übersichtlichen Darstellung nur die letzte bzw. die letzten beiden Ziffern der Kennnummer genannt.

Die Korrelation zwischen der Gesamtqualität und der Effektivität ist mit einem Spearman-Korrelationskoeffizient von 0,38 eher schwach ausgeprägt. Die Korrelation zwischen Studienqualität und dem Effekt fällt mit einem Koeffizienten von 0,25 noch geringer aus. Dagegen liegt die Korrelation von Interventionsqualität und Effekt mit 0,39 deutlich höher.

Tabelle 21: Korrelation von Effekt, Qualität der Studiendurchführung und Qualität der Intervention im Handlungsfeld Depression

Zu korrelierende Größen	Spearmankorrelation
Qualität gesamt und Effekt	0,38
Qualität der Studie (QS) und Qualität der Intervention (QI)	0,52
Qualität der Studie (QS) und Effekt	0,24
Qualität der Intervention (QI) und Effekt	0,39

In der vorliegenden Analyse führen hohe Qualitätswerte bei (teilweise) effektiven Studien nicht automatisch zu einer Empfehlung. Vielmehr sind die Ergebnisse aus verschiedenen homogenen Studien in einer Gesamtschau zu berücksichtigen. Die Ergebnisse dienen als Informationsgrundlage, um in Expertengruppen qualitätsgestützte Entscheidungen treffen zu können.

Abbildung 13: Interventions- und Studienqualität

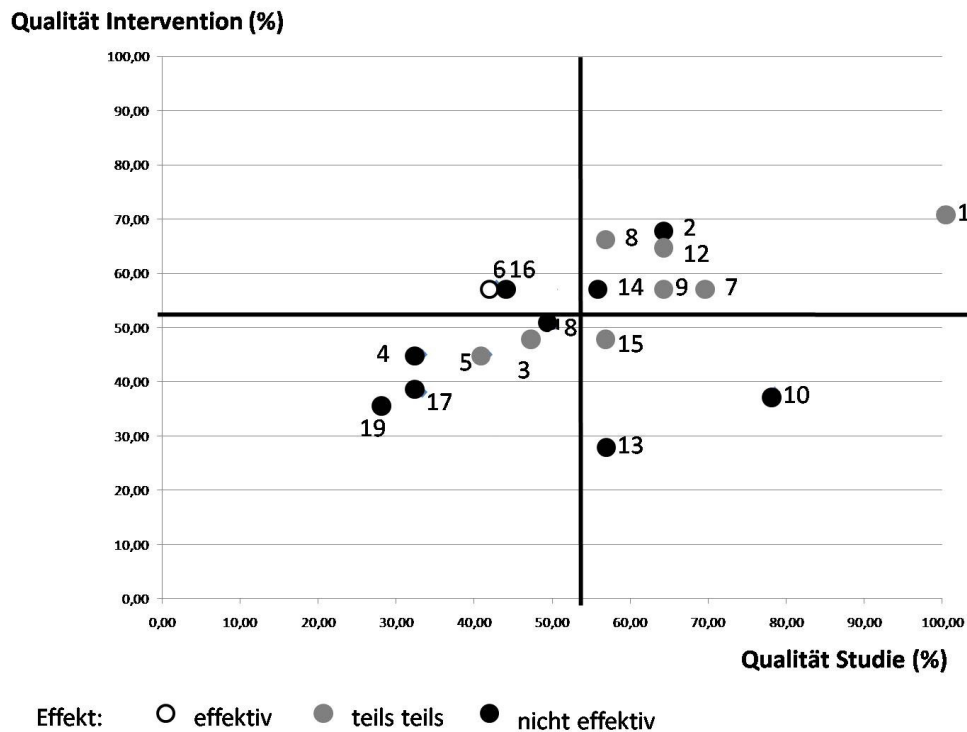
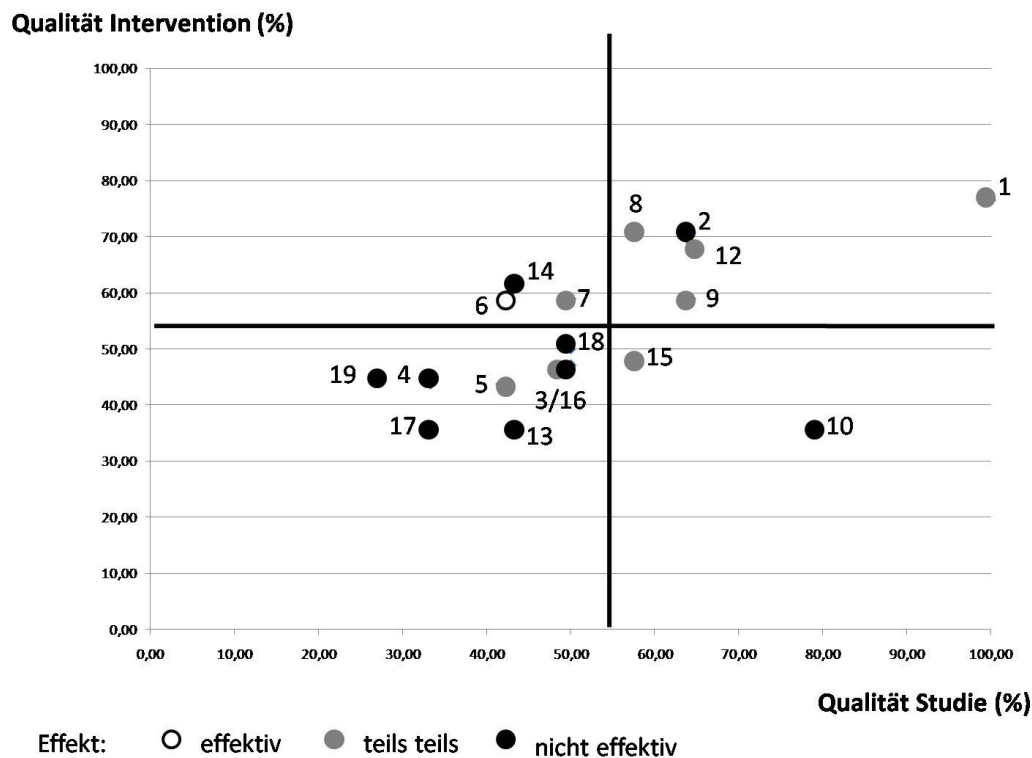


Abbildung 14 zeigt die Ergebnisse einer Sensitivitätsanalyse, in der zwei Kriterien der Interventionsqualität, nämlich *Partizipation* und *Empowerment* nicht berücksichtigt wurden (zur Begründung für diese Sensitivitätsanalyse s. Kapitel 3.4). Durch die Änderung verschiebt sich das relative Gewicht der Qualitätskriterien mit der Folge, dass die insgesamt erreichten Scores zweier Studien nicht mehr ausreichen, um zu den potenziell empfehlenswerten Studien zu zählen. Nur noch fünf der ursprünglich sieben Studien fallen damit in diesen Bereich.

Abbildung 14: Sensitivitätsanalyse der Qualität



## 7.2.2 Darstellung einzelner Studien

Im Folgenden werden die sieben gemäß Gesamtanalyse potenziell empfehlenswerten Studien vorgestellt.

### *Merry et al. 2004 (2001)*

Merry und Kollegen haben eine randomisierte kontrollierte Studie zur Prävention von Depressionen bei Jugendlichen durchgeführt, in der sie das von Shochet und Kollegen (Shochet 2001) entwickelte Resourceful Adolescent Program (RAP) für Neuseeland adaptieren (RAP-KiWi). In der Studie wurde für 192 Schüler der Klassen 9 und 10 eine kognitiv-verhaltenstherapeutische Intervention von Lehrern durchgeführt. Für die 189 Schüler der Placebogruppe wurde ein Placeboprogramm entwickelt, in dem jedoch keine kognitiv-verhaltenstherapeutischen Elemente enthalten waren. Die Lehrer wurden im Vorfeld geschult.

Die Interventions- bzw. Placeboprogramme umfassten jeweils elf Sitzungen, die ein- bzw. zwei Mal wöchentlich durchgeführt wurden. Die letzte Nachbeobachtung erfolgte nach 18 Monaten. Als Effektmaße wurden mit dem *Beck Depression Inventory II* (BDI) und dem *Reynolds Adolescent Depression Scale* (RADS) zwei Instrumente zur Selbstbewertung der depressiven Symptomatik eingesetzt.

Im Vergleich zur Baseline wurde für die BDI-II unmittelbar nach der Intervention eine absolute Risikoreduktion von 3% festgestellt. Die Number Needed to Treat lag bei 33. Auch im 18 Monats-Follow-up blieben die Depressionswerte in beiden Gruppen unter den Ausgangswerten. Für die BDI gab es jedoch keine signifikanten Differenzen, während die Veränderungen der RADS für die Interventionsgruppe signifikant größer ausfiel ( $p=0,03$ ). Damit gilt die Studie als teilweise effektiv.

Eine Besonderheit der Studie liegt darin, dass die Intervention im Vergleich zu einer *aktiven* Placebogruppe untersucht wird. Merry et al. kritisieren, dass trotz nachgewiesenem Placeboeffekt in vielen Studien auf eine aktive Kontrollgruppe verzichtet wird.

### ***Pössel et al. 2004 (2002)***

In einem in Deutschland durchgeführten RCT untersuchen Pössel und Kollegen die Wirkung des Präventionsprogramms *Leichtigkeit im Sozialen Alltag (LISA)* im Vergleich zu einer Kontrollgruppe ohne Intervention. Insgesamt wurden 324 Jugendliche Achtklässler einbezogen, von denen 177 der Interventionsgruppe und 147 der Kontrollgruppe angehörten. Die Schüler der Kontrollgruppe erhielten normalen Unterricht. Gegenstand des Programms LISA ist die soziale Informationsverarbeitung durch kognitive Verhaltenstherapie. LISA umfasst zehn Sitzungen zu je 90 Minuten die einmal wöchentlich durchgeführt werden, so dass die Intervention ca. 2,5 Monate dauerte. Die letzte Nachbeobachtung erfolgte nach sechs Monaten.

Als Effektmaße für die kurzfristigen Ziele wurden der *Questionnaire of social support (FESU)* und der *Automatic Thought Questionnaire (ATQ)* eingesetzt. Für die langfristigen Ziele wurde die *Center for Epidemiological Studies Depression Scale (CES-D)* genutzt. Zwar war bei Niedrigdepressiven in der Kontrollgruppe ein signifikanter Anstieg der selbstberichteten depressiven Symptomatik zu beobachten, der in der Interventionsgruppe nicht eintrat. In der Trainingsgruppe gab es zudem eine klare Veränderung von den höheren Risikogruppen mit klinisch relevanten Symptomen zu der Gruppe mit minimalen Symptomen. Zu keinem Messzeitpunkt bestanden aber zwischen der Tai-

nings- und der Kontrollgruppe für Teilnehmer mit gleichem Risikostatus signifikante Unterschiede in der selbstberichteten Depressionssymptomatik. Für die dysfunktional automatischen Gedanken und die Effektmaße der sozialen Unterstützung war kein Interventionseffekt nachweisbar.

### ***Shochet et al. 2001 (2007)***

Shochet und Kollegen haben in einer kontrollierten Studie die präventive Wirkung des Resourceful Adolescent Program (RAP) an 260 Schülern der 9. Klasse untersucht. Die Schüler wurden neben dem reinen Programm für Jugendliche (RAP-A) einer Programmversion für Familien (RAP-F) bzw. einer Kontrollgruppe ohne Intervention zugewiesen. Auf RAP-A entfielen 53, auf RAP-F 51 und auf die Kontrollgruppe 90 Teilnehmer. Das Programm RAP-A umfasst elf einmal wöchentlich stattfindende Sitzungen. Das Programm RAP-F umfasst zusätzlich drei Gruppensitzungen mit den Eltern der Schüler. Die letzte Nachbeobachtung erfolgte nach zehn Monaten.

Als Effektmaße wurden in der Studien die Reynolds Adolescent Depression Scale (RADS), der Children Depression Inventory (CDI) und die Beck Hopelessness Scale (BHS) herangezogen. Während beide Interventionsgruppen gegenüber der Kontrollgruppe zum letzten Messzeitpunkt signifikante Verbesserungen der CDI und BHS aufwiesen, zeigten sich für die RADS keine Interventionseffekte. Die Intervention inklusive der Elternsitzungen brachte keinen zusätzlichen positiven Effekt.

### ***Barrett et al. 2006 (2008)***

Bei dieser neueren, in Australien durchgeführten randomisiert kontrollierten Studie von Barrett und Kollegen steht insbesondere die Frage des optimalen Alters der Prävention von Angst- und Depressionen im Vordergrund. Dazu wurden sowohl in der Interventions- als auch in der Kontrollgruppe Sechst- und Neuntklässler einbezogen. Von den bei Studienbeginn 737 Schülern wurden nach zwölf Monaten noch 669 berücksichtigt, von denen 379 der Interventionsgruppe und 290 der Kontrollgruppe angehörten. Die Zahl der Probanden zum hier vorgestellten 36 Monats Follow-up wurde nicht explizit ausgewiesen.

Die Intervention besteht in dem von Lehrern unter der Assistenz von Psychologiestudenten durchgeführten FRIENDS-Programm. FRIENDS steht dabei als Akronym für: F - feeling worried?, R - relax and feel good, I - inner thoughts, E – explore plans, N nice work – self reward, D

don't forget to practice, S - stay calm. Das Programm ist als kognitiv- verhaltenstherapeutische Intervention konzipiert und wird zur Prävention von Angstsymptomen eingesetzt. Über zehn Wochen findet dazu jeweils eine Sitzung wöchentlich statt. Bereits 2003 wurden die Ergebnisse nach zwölf Monaten publiziert. (Lock und Barrett 2003) Neben der Spence Child Anxiety Scale (SCAS) und der Revised Children Manifested Anxiety Scale (RCMAS) als Effektmaße für Angst- bzw. Angstsymptome wurden mit dem Childhood Depression Inventory (CDI) auch die Auswirkung auf die Depressionssymptomatik analysiert. Für die Childhood Depression Inventory gab es keine signifikanten Ergebnisse. Die Auswertung nach 36 Monaten ergab nur für die Sechsklässler der Interventionsgruppe signifikante Verbesserungen der Angstsymptome und der Angstwerte gegenüber der Kontrollgruppe ( $p < 0,05$ ). Die Autoren leiten aus den Analysen ab, dass der optimale Zeitpunkt für die Prävention von Angstwerten bei ungefähr neun bis zehn Jahren liegt. Subgruppenanalysen nach Geschlecht zeigten allerdings, dass signifikant bessere Angstwerte der Interventionsgruppe für Mädchen nur im Follow up nach zwölf bzw. 24 Monaten bestanden, aber nicht über 36 Monate andauerten. Da für Mädchen ein größeres Angstisiko beobachtet wurde, empfehlen die Autoren trotz der sehr positiven kurzfristigen Ergebnisse auch für Mädchen, die Intervention im Abstand von zwei bis drei Jahren zu wiederholen. Bezüglich der Depressionswerte ergaben sich keine signifikanten Unterschiede zwischen den Gruppen.

### ***Barrett, Turner 2001 (2009)***

Barrett und Kollegen haben das von der Arbeitsgruppe um Barrett entwickelte australische FRIENDS-Programm schon in einem früheren RCT evaluiert. In der australischen Studie wurden 489 Kinder im Alter von zehn bis zwölf Jahren einbezogen, die einem von drei Studienarmen zugeordnet wurden. Einer von Psychologen geleitete Interventionsgruppe gehörten 152, einer von Lehrern geleiteten Interventionsgruppe 253 und einer Kontrollgruppe 84 Schüler an. Psychologen und Lehrer wurden als Trainer im Vorfeld für die Intervention qualifiziert.

Die Interventionsphase dauerte zehn Wochen, in denen jeweils eine 75minütige Sitzung stattfand. Zusätzlich fanden in dieser Zeit vier Abendsitzungen für die Eltern der teilnehmenden Schüler statt. Eine Nachbeobachtung über die Post-Auswertung erfolgt nicht.

Als Effektmaße wurde das Spence Children Anxiety Scale (SCAS), das Revised Children's Manifest Anxiety Scale (RCMAS) und das Children's Depression Inventory (CDI) herangezogen. Die statistische Analyse zeigte, dass die Interventionsgruppen eine deutlich stärkere Reduktion selbst-

berichteter Angstsymptome aufwiesen als die Kontrollgruppe ( $p < 0,05$ ). Zwischen den beiden Interventionsgruppen bestand dabei kein Unterschied. Bezüglich der selbstberichteten Depressionswerte gab es keine Überlegenheit der Interventionsgruppen. Die von Lehrern geleitete Interventionsgruppe wies sogar gegenüber den beiden anderen Gruppen einen signifikant größeren Anstieg der selbstberichteten Depressionswerte auf.

Eine Besonderheit dieser Studie liegt darin, dass durch das dreiarmlige Studiendesign die Relevanz der Qualifikation der Gruppenleiter untersucht wurde.

### ***Pössel et al. 2005 (2012)***

In der Studie werden die Ergebnisse einer in Deutschland durchgeführten randomisiert kontrollierten Studie des Programms *Leichtigkeit im Sozialen Alltag* (LISA) vorgestellt. Von ursprünglich 347 randomisierten Jugendlichen wurden Daten von 279 Jugendlichen im Alter von 14 Jahren ausgewertet. 163 gehörten der Interventionsgruppe, 116 der Kontrollgruppe an. Das Programm umfasst zehn Sitzungen, die einmal pro Woche stattfanden. Nach einem Follow-up von drei Monaten wurde der Effekt der Intervention evaluiert. Dabei wurde die Wirkung des Programms auf die automatischen negativen Gedanken (ATQ), die soziale Unterstützung (FESU-N), die Selbstwirksamkeit und die Depressionswerte (CES-D) gemessen. Für ATQ wurden keine signifikanten Differenzen im Gruppenvergleich festgestellt. Die Größe des Netzwerkes der Jugendlichen erwies sich im drei Monats Follow-up in der Trainingsgruppe signifikant größer als in der Kontrollgruppe ( $p < 0,05$ ). Darüber hinaus waren im Follow-up in der Trainingsgruppe die Werte der Depression für Jugendliche mit geringer Selbstwirksamkeit signifikant geringer als in der Kontrollgruppe ( $p < 0,01$ ).

### ***Clark et al. 1993 (2014)***

Die Studiengruppe um Clark hat in den USA ein schulbasiertes kognitives Gruppenprogramm für Jugendliche der Klassen neun und zehn durchgeführt. Es handelt sich bei der Studie um ein RCT. Von den insgesamt 300 in der Auswertung berücksichtigten Schülern gehörten 151 zur Interventions- und 149 zur Kontrollgruppe. Neben der reinen Wissensvermittlung umfasste die Intervention ein verhaltenstherapeutisches Skill-Training, welches in fünf Sitzungen durchgeführt wurde.

Die Studiendauer wurde nicht explizit erwähnt. Es erfolgte eine Nachbeobachtung nach zwölf Wochen, in der die Programmeffekte anhand der Center for Epidemiological Study Depression scale (CES-D) gemessen wurden. Es wurde kein Effekt auf die Depressivität der Jugendlichen festgestellt.

Als problematisch kann die kurze Dauer der Intervention und der kurze Follow-up-Zeitraum betrachtet werden.

### **7.3 Diskussion**

Die Programme zur universellen Prävention der Depression basieren bei allen Unterschieden im Detail auf kognitiven verhaltenstherapeutischen Interventionen. Insgesamt weisen diese Studien einen hohen Qualitätsstandard auf. Die durchschnittlichen Qualitätsscores liegen über 50 %. Studien, die in anderen Handlungsfeldern noch als potenziell empfehlenswert gelten, fallen damit bei der Depression schon in den als nicht empfehlwerten Bereich.

Die in der vorliegenden Analyse einbezogenen Studien konzentrieren sich auf wenige Programme. Am häufigsten wurden die Programme Leichtigkeit im sozialen Alltag (LISA; vier Mal), FRIENDS (drei Mal), Resourceful Adolescent Programm (RAP; zwei Mal) evaluiert. Unter den sieben Studien mit überdurchschnittlichen Qualitätsscores kommen diese Programme jeweils zwei Mal vor. Darüber hinaus ist nur die edukative Studie von Clark im Bereich der aufgrund ihrer Qualitätsscores potenziell empfehlwerten Studien. Die Interventionen der Studien sind überwiegend „teilweise effektiv“ (fünf von sieben). Die teils-teils-Bewertung ergibt sich aufgrund der Vielzahl der einbezogenen Effektmaße, von denen nicht alle signifikante Ergebnisse erreichen. Daraus kann nicht abgeleitet werden, dass die Wirksamkeit generell schlechter ist als in effektiv bewerteten Studien, da hier zum Teil mehr Effektmaße einbezogen wurden. Um diese Verzerrung zu verhindern, sollte in Zukunft die Vielzahl und Art der Effektmaße vereinheitlicht werden. Nur so kann die Bedeutung der Intervention i.e.S. für die Ergebnisse von dem Einfluss verschiedener Effektmaße isoliert werden.

Beide Studien mit dem Programm RAP sind als teilweise effektiv bewertet und es liegen keine weiteren Studien für dieses Programm mit unterdurchschnittlicher Studienqualität vor. Die Studie von Merry et al. 2004 zu dem Programm RAP erreicht als einzige Studie mit Qualitätswerten zwischen 60 und 80% den Qualitätslevel (potenziell) stark empfehlenswert.

Für das Programm LISA liegen drei Studien (Pössel et al. 2004; Groen et al. 2003; Pössel et al. 2006a) und für das Programm FRIENDS eine weitere Studie mit unterdurchschnittlicher Studienqualität vor (Lowry-Webster et al. 2003). Die Studien für das Programm RAP und FRIENDS weisen einheitliche Tendenzen auf. Für das Programm LISA liegen dagegen sowohl ineffektive als auch teilweise effektive Studien vor.

Insgesamt ist festzuhalten, dass die Forschung zur universalen Prävention der Angst und Depression relativ jung ist. Bezüglich der Ausgestaltung der Intervention in den verschiedenen Studien bestehen über die in den Qualitätsscores erfassten Aspekte hinaus Unterschiede in als relevant erachteten Aspekten der Programme. Es gibt Anhaltspunkte, dass die Ausgestaltung der Interventionen großen Einfluss auf die Ergebnisse haben, auch wenn die hierzu vorliegenden Hinweise zum gegenwärtigen Zeitpunkt noch nicht als empirisch gesichert gelten können (Pössel et al. 2006b). So sind z.B. die Frage der geeigneten Trainer (Lehrer oder Psychologen), die optimale Gruppenstärke sowie die Dauer der Intervention bzw. des Follow-up in Ihrer Bedeutung für die Ergebnisse noch nicht abschließend erforscht. So wird in der Literatur z.B. ein verzögerter Präventionseffekt diskutiert, da sich die in den Programmen vermittelten Kompetenzen im Alltag erst festigen müssen (Barrett et al. 2006). In der vorliegenden Analyse ist zudem der Interventionsqualität z.B. bei der Frage der Trainer schon dann Genüge getan, wenn Lehrer *oder* Psychologen qualifiziert und in die Programmentwicklung einbezogen wurden. Ob die Entscheidung für Lehrer oder Psychologen als Trainer einen Einfluss auf die Ergebnisse haben, ist bisher aber erst in einer Studie analysiert worden, in der zu geringe statistische Power abschließende Aussagen zur Überlegenheit einer Gruppe nicht zulassen (Barrett und Turner 2001; Pössel 2006b). Nicht zuletzt ist zu bedenken, dass nur die Studie von Merry et al. 2004 eine aktive Kontrollgruppe einbezieht. Die Ergebnisse der anderen Studien vernachlässigen damit einen möglichen Placeboeffekt.

Insgesamt ist im Bereich der Prävention von Depressionen und Angststörungen noch weiterer Forschungsbedarf gegeben. Dazu sind idealerweise verschiedene Ausgestaltungen der Programme im direkten Vergleich zu untersuchen (Pössel 2006b; Barrett et al. 2006). Angesichts der bisher erzielten Erfolge bei Maßnahmen der universalen Prävention von Angststörungen und Depressionen ist diese Forschung zu forcieren. Da alle betrachteten Programme stark auf das Verhalten der Probanden zielen und keine kontextuellen Faktoren einbeziehen, ist davon auszugehen, dass Studienergebnisse verschiedener Länder übertragbar sind. Eine flächendeckende Umsetzung

erscheint gegenwärtig aufgrund der offenen Fragen zur optimalen Ausgestaltung verfrüht. Werden Programme der universellen Prävention mit der für die offenen Fragen nötigen statistischen Power vereinzelt durchgeführt und umfassend evaluiert, können weitere wichtige Erkenntnisse gewonnen werden. Selbst wenn Interventionen der universellen Prävention eine geringere Effektstärke als selektive und indizierte Präventionsprogramme haben sollten, können sie bei niedrigen Präventionskosten kosteneffektiv sein (Horowitz und Garber 2006).

## 8 Ernährung

Im folgenden Kapitel wird zunächst eine allgemeine Übersicht über das Themenfeld Ernährung gegeben. Es wird auf Prävalenz und Folgen von Fehlernährung, die entstehenden Kosten sowie auf die internationale wissenschaftliche Datenlage zu Prävention und Gesundheitsförderung im Bereich Ernährung eingegangen. Anschließend werden die Ergebnisse im Handlungsfeld Ernährung dargestellt. Zunächst wird ein Überblick über alle evaluierten Studien bezüglich Autor, Studiendesign und Land, in dem die Studie durchgeführt wurde, gegeben. Im Anschluss werden die Qualität der Studien und die Qualität der Interventionen exemplarisch an verschiedenen Kriterien dargestellt. Es folgt eine kritische Diskussion der Ergebnisse.

### 8.1 Stand der Forschung

#### Prävalenz

Der Risikofaktor Ernährung gewinnt in Wissenschaft und Gesellschaft zunehmend an Bedeutung (Kroke und Boeing 2003). Obwohl bereits weltweit evidenzbasierte Ernährungsempfehlungen vorliegen, gibt es in der Gesellschaft Probleme im Essverhalten (Bönnhoff 2005).

Ein ausgewogenes Ernährungsverhalten ist für die gesundheitliche Entwicklung vor allem von Kindern und Jugendlichen von Bedeutung (Mensink et al. 2007). Dies gilt besonders vor dem Hintergrund der zunehmenden Prävalenz von Adipositas im Kindes- und Jugendalter.

Die Daten des Kinder- und Jugendgesundheits surveys (KiGGS) in Deutschland zeigen, dass etwa 50% der befragten Jungen (N=7186) und Mädchen (N= 6919) täglich Brot und Milch konsumieren. 16% aller Befragten essen täglich Schokolade. 72% der Jungen und 75% der Mädchen essen ein bis drei Mal im Monat Fast Food (Bratwurst, Currywurst, Hamburger, Döner). 47% der Jungen und 55% der Mädchen gaben an, mindestens einmal am Tag frisches Obst zu sich zu nehmen (Mensink et al. 2007).

Die Prävalenz von Übergewicht und Adipositas ist von 1980 bis heute rasant gestiegen. Entsprechend internationalen Empfehlungen von Childhood Group der International Obesity Task Force (IOTF) und European Childhood Obesity Group (ECOG) werden Übergewicht und Adipositas mit Hilfe des Body Mass Index (BMI) definiert (Kromeyer-Hausschild et al. 2001).

Die Leitlinien der AGA definieren einen BMI oberhalb der 90. Perzentile für Alter und Geschlecht als „Übergewicht“, ein BMI größer der 97. Perzentile wird als „Adipositas“ definiert und ein Überschreiten der 99,5. Perzentile bezeichnet eine extreme Adipositas.

In Deutschland sind ca. 15% der Kinder und Jugendlichen im Alter von drei bis 17 Jahren übergewichtig und 6,3% leiden an Adipositas (Body Mass Index > 90. Perzentile) (Kurth und Schffrath Rosario 2007). Die Prävalenz ist in den neuen Bundesländern und in den sozial benachteiligten Gruppen noch höher (Benecke und Vogel 2003). Bislang ist auch nicht erkennbar, dass diese Entwicklung zum Stillstand kommt (Wabitsch et al. 2002).

Aber auch in anderen europäischen Ländern wie z. B. in den Niederlanden, in der Schweiz, in Italien und in Großbritannien, ist der Anteil der Übergewichtigen deutlich gestiegen. In ganz Europa sind 14 Millionen Kinder übergewichtig, davon 3 Millionen adipös. 1980 lag der Prozentsatz bei 24% übergewichtigen Kindern. Er wurde bis 2007 um 5% übertroffen (Kurth und Rosario 2007).

### **Die Folgen der Fehlernährung**

Die Aufnahme von zu fett- und zuckerreicher Nahrung, gekoppelt mit Bewegungsmangel und genetischer Disposition, kann zu chronischen Erkrankungen wie Malignomen, kardiovaskulären Erkrankungen, Diabetes mellitus und Adipositas führen. Diese Krankheitsbilder treten überwiegend in den westlichen Industrieländern und damit auch in Deutschland auf (Krocke 2003, S. 11). Neben den sozialen Auswirkungen wie z.B. Stigmatisierung in der Bevölkerung und Benachteiligung bei der Arbeitssuche (Schumann 1999) ist die Adipositas auch mit körperlichen Begleiterkrankungen (Komorbidität) verbunden.

Auf internistischem Gebiet kann Adipositas zu Bluthochdruckerkrankungen, koronarer Herzkrankheit, Diabetes mellitus, Hyperurikämie/Gicht, Dyslipoproteinämie etc. führen. Gelenkverschleißerscheinungen, Bandscheibenvorfälle und Schenkelhalsermüdungsfrakturen sind an dieser Stelle nur exemplarisch aus dem Fachgebiet der Orthopädie zu nennen (Märkl 2004, S. 5; Wabitsch 2000). Tabelle 22 zeigt noch einmal zusammenfassend, welche Krankheiten durch Adipositas begünstigt oder bedingt sein können.

Tabelle 22: Mit Adipositas assoziierte Krankheiten und Probleme (nach Wirth 1998)

Physischen Komorbiditäten	Psychosoziale Auswirkungen
Hypertonie, koronare	Vermindertes Selbstbewusstsein
Herzinsuffizienz	Stigmatisierung

Diabetes Mellitus Typ II	Soziale Isolation
Dyslipidämie	Diskriminierung
Hyperurikämie	Berufsprobleme
Fettleber, Hemien	Geringere Förderung in der Schule
.....	.....

Das Risiko der Persistenz der Adipositas im Alter ist für die Personen erhöht, die bereits als Kinder oder Jugendliche daran leiden (Wabitsch et al. 2002; Serdula et al. 1993). Führt man sich die Krankheitsfolgen eines gestörten Essverhaltens vor Augen, ist die Notwendigkeit von Präventionsprogrammen besonders bei Kindern und Jugendlichen gegeben.

### **Kosten von Fettleibigkeit und Adipositas**

Eine Vielzahl internationaler Studien betrachten die wirtschaftlichen Auswirkungen der Adipositas. Dabei ergab sich konsistent, dass sie für 2 bis 7% der gesamten Gesundheitsausgaben verantwortlich sind (WHO 2000).

In Frankreich belaufen sich die Kosten der mit Adipositas in Verbindung gebrachten Erkrankungen auf ungefähr 2% der gesamten Ausgaben der Gesundheitsvorsorge (Levy et al. 1995). In den Niederlanden liegen die Ausgaben bei 3 bis 4% gemessen am Aufwand der praktischen Ärzte, die Adipositas und Übergewicht behandeln (Seidell und Deerenberg 1994). Das National Health Service in Großbritannien gibt jährlich eine halbe Milliarde Pfund für die Behandlung von Adipositas aus.

In Großbritannien geht man davon aus, dass adipöse Personen 18 Millionen Krankheitstage pro Jahr verursachen (National Audit Office 2001). Das Robert-Koch-Institut stellte heraus, dass Adipositas mit einer erhöhten Inanspruchnahme von Allgemeinärzten in Deutschland einhergeht (von Lengerke und John 2005). 1993 veröffentlichte das Bundesministerium für Gesundheit einen Bericht, der die Kosten für ernährungsbedingte Krankheiten auf rund 83,5 Milliarden DM oder 30% aller Ausgaben im Gesundheitswesen beziffert (Hauner 2000; Kohlmeier et al. 1993).

### **Internationale Evidenz**

Es wird international empfohlen, wenigstens fünf Mal am Tag Obst und Gemüse zu konsumieren. Die Aufnahme von ungesättigten Fetten und die Verwendung von Salz sollen reduziert wer-

den. Gleichzeitig soll der Konsum von komplexen Kohlenhydraten gesteigert werden. Daneben wird ausreichende Bewegung empfohlen (Shepherd et al. 2006)

Im Bereich Ernährung gibt es eine Vielzahl von Reviews und Metaanalysen, die diese Empfehlung zu ihrem Untersuchungsschwerpunkt gemacht haben. An dieser Stelle kann nicht auf alle Reviews und Metaanalysen eingegangen werden. Deshalb wird exemplarisch anhand von zwei aktuellen Reviews und einer Metaanalyse die aktuelle Evidenz erläutert.

## **Reviews**

Blanchette und Brug (2005) haben 15 Studien zum Konsum von Obst und Gemüse bei sechs- bis zwölfjährigen Kindern untersucht. Dabei wurden sie den folgenden Bereichen zugeordnet: 1) Verfügbarkeit von Früchten und Gemüse 2) Verhalten von Eltern (Konsum der Eltern, child-feeding Erfahrungen) 3) Einflüsse durch Werbung im Fernsehen 4) Zugang zu Schulkantinen. Für diese Bereiche werden folgende Effekte angegeben: Portionen pro Tag, Portionen pro Mittagessen, Anzahl der Obststücke in der Schule pro Tag.

Die Ergebnisse der Interventionen zeigen einen Anstieg des Konsums von 0 bis 2,54 Portionen Obst und Gemüse pro Tag. Der Früchtekonsum ist dabei deutlicher gestiegen als der Gemüseverzehr.

Von 15 Interventionen waren zwölf schulbasiert. Dabei waren sieben als Multi-Komponenten-Intervention ausgelegt. Alle Multi-Komponenten-Interventionen hatten einen positiven Effekt (Anstieg von 0,2 bis 1,68 Portionen Obst und Gemüse pro Tag).

Das Review hat gezeigt, dass eine Kombination von Komponenten wie classroom curriculum, familienbasierter Unterstützung und ein schulischer Essensservice die höchste Effektivität hatte (French und Stables 2003).

Contento und Michela (1998) stellten heraus, dass der Versand von schriftlichem Infomaterial eine effektive Methode ist, um Eltern einzubinden.

Burchett (2003) analysierte in den USA fünf Interventionen im Setting Schule. Drei Studien hatten zum Ziel, den Früchte- und Obstkonsum in Grundschulen zu steigern. In den übrigen zwei Studien war der Konsum nur ein Aspekt von einer umfassenden kardiovaskulären Intervention (Perry et al. 1998). Als Effektmaß wurde auch hier die Anzahl der Früchte- und Obstportionen verwendet. Als evident kann hier festgehalten werden: Es besteht ein größerer Vorteil für Ge-

sundheitsprojekte, wenn der Fokus speziell auf den Früchte- und Obstkonsum gelegt wird. Generelle Gesundheitsratschläge führen allein nicht zum gewünschten Erfolg.

### **Meta-Analyse**

In der Meta-Analyse von Snethen et al. (2006) wurden aus 356 Studien sieben mit Interventionen im Bereich Gewichtsverlust bei Schulkindern auf ihre Effektivität hin untersucht. Als Effektgröße wurde Cohen's D benutzt. Als Effektmaß diente der Gewichtsverlust. Unter allen sieben Studien waren 14 Interventionen mit einer durchschnittlichen Effektgröße von  $d = 0,95$ . Das 95%-Konfidenzintervall für D erstreckt sich von 0,79 bis 1,11.

Als evidenzbasiert kann festgehalten werden: Eine strukturierte Diät und ein zusätzliches Fitnessprogramm können dazu beitragen, dass Kinder Gewicht verlieren.

## **8.2 Ergebnisse**

### **8.2.1 Allgemeine Daten zu den Interventionen im Bereich Ernährung und deren Bewertung**

In der Literaturrecherche unseres Forschungsprojekts wurden für den Bereich Ernährung 8.955 „Treffer“ gefunden. Für die vorliegende Untersuchung wurde entschieden, Interventionen zu evaluieren, die auf Kinder und Jugendliche fokussieren. Ein triftiger Grund liegt darin, dass die Prävalenz von Adipositas und Übergewicht besonders bei Kindern in den letzten Jahren gestiegen ist. Präventionsprogramme, die für Kinder angeboten werden, können Folgeerkrankungen der Adipositas im Alter vorbeugen.

Auf Grund der berücksichtigten Ein- und Ausschlusskriterien konnten 24 Publikationen evaluiert werden. Alle Interventionen fanden im Setting Schule statt. 72% der integrierten Interventionen wurden nach 2000 publiziert, alle sind englischsprachig. 64% der Untersuchungen sind randomisierte kontrollierte, 32% kontrollierten nicht randomisierte und 4% Eingruppenuntersuchungen.

Tabelle 23: Übersicht über die eingeschlossenen Studien im Handlungsfeld Ernährung

Kennnummer	Autor (Jahr)	Projektname (ggf.)	Land	Design
3004	Kaitosaari et al. (2003)	STRIP	Finnland	RCT
3005	Story et al. (2000)	5-a-Day Power Plus	USA	RCT
3006	Townsend et al. (2006)	EFNEP	USA	RCT
3007	Bere et al. (2006a)	FVMM	Norwegen	RCT
3008	Bere et al. (2006b)	FVMM	Norwegen	RCT
3009	Powers et al. (2005)		USA	CT
3010	Baranowski et al. (2000)	Gimme 5 Fruit and Vegetables	USA	RCT
3011	Perry et al. (2004)	Cafeteria Power Plus	USA	RCT
3013	Lowe et al. (2004)	Food Dudes	Großbritannien	Non-CT
3014	Nicklas et al. (1998)	Gimme 5 Fresh Nutrition	USA	RCT
3015	Resnicow et al. (1992)	Know you body	USA	CT
3016	Reynolds et al. (2000)	High 5 Project	USA	RCT
3017	Foerster et al. (1998)	5 a Day-Power Play	USA	CT
3020	Lytle et al. (1996)	Catch	USA	RCT
3021	Perry et al. (1998)	Catch Fruit and Vegetable Intake	USA	RCT
3022	Story et al. (2003)	GEMS	USA	RCT
1016	Belansky et al. (2006)		USA	CT
1017	Hoelscher et al. (2004)		USA	CT
1019	Trevino et al. (2004)		USA	RCT
1020	Kelder et al. (2005)	CATCH Kids Club	USA	CT
1021	Coleman et al. (2005)		USA	CT
1022	Stewart et al. (2003)		USA	RCT
1024	Gortmaker et al. (1999a)	Eat well and keep moving	USA	CT
1025	Gortmaker et al. (1999b)	Planet Health	USA	RCT

## 8.2.2 Effektmaße

Tabelle 24 zeigt eine Auswahl der am häufigsten verwendeten Effektmaße. Jede evaluierte Studie hatte zum Teil unterschiedliche Effektmaße. Ebenso wie beim Risikofaktor Bewegung sind die Effektmaße beim Risikofaktor Ernährung heterogen.

Tabelle 24: Ausgewählte Effektmaße im Handlungsfeld Ernährung

Kennung	%Kalorien-anteil aus Fett, aus Kohlenhydraten	Cholesterin	Verzehr Obst und Gemüse	Verzehr ausgewogener Nahrungsmittel	BMI/ Gewicht	Blutdruck	Anteil Zeit Bewegung	Weitere
3004	X	X						X
3005			X					X
3006				X				
3007			X					
3008			X					
3009			X					X
3010			X					
3011			X					
3013			X					
3014			X					X
3015		X			X	x		
3016	X		X					
3017			X					
3020	X							
3021			X					
3022			X		X		X	
1016								X
1017	X						X	
1019	X							
1020							X	X
1021	X				X		X	
1022	X				X		X	X
1024	X		X					X
1025							X	

In der folgenden Tabelle 25 wird die Evaluation der Studienqualität tabellarisch dargestellt. Die Qualität der Studien wird als Gesamtpunktzahl und anhand der ausgewählten Kriterien Randomi-

sation, Vergleichbarkeit der Gruppen, Validierung der Messinstrumente und der Powerberechnung wiedergegeben.

Tabelle 25: Qualität der Studiendurchführung im Handlungsfeld Ernährung

Kennnummer	Studien-design	Randomisation (16)	Vergleichbarkeit (19)	Messinstrumente (25)	Power (46)	Summe	Erzielbarer Maximal-score	Prozentualer Score
3004	RCT	0	1	1	1	5,5	7	78
3005	RCT	0	1	0	0	0	7	0
3006	RCT	0	0	1	0	3,5	7	50
3007	RCT	0	0	0	0	2	7	28
3008	RCT	0	0	0	0	3,5	7	50
3009	CT	X	0	0	0	1,5	6	25
3010	RCT	0	0	0	0	2,5	7	35
3011	RCT	0	1	0	0	3,5	7	50
3013	Non-CT	X	X	0	0	2	5	40
3014	RCT	1	0	0	0	4	7	57
3015	Non-CT	X	0	0,5	0	3	5	60
3016	RCT	1	1	0	0	4	7	57
3017	CT	X	0	0	0	2	6	33
3020	RCT	1	0	0	1	4	7	57
3021	RCT	0	0	0	0	2	7	28
3022	RCT	0	0	0	0	1	7	14
1016	CT	X	0	0	0	2	6	33
1017	CT	X	1	0	0	3	6	50
1019	RCT	1	1	0,5	0	4,5	7	64
1020	CT	X	0	0,5	0	1,5	6	25
1021	CT	X	0	1	0	3	6	50
1022	RCT	1	1	0	0	2	7	28
1024	CT	X	1	0,5	0	3,5	6	58
1025	RCT	1	1	0,5	0	4,5	7	64

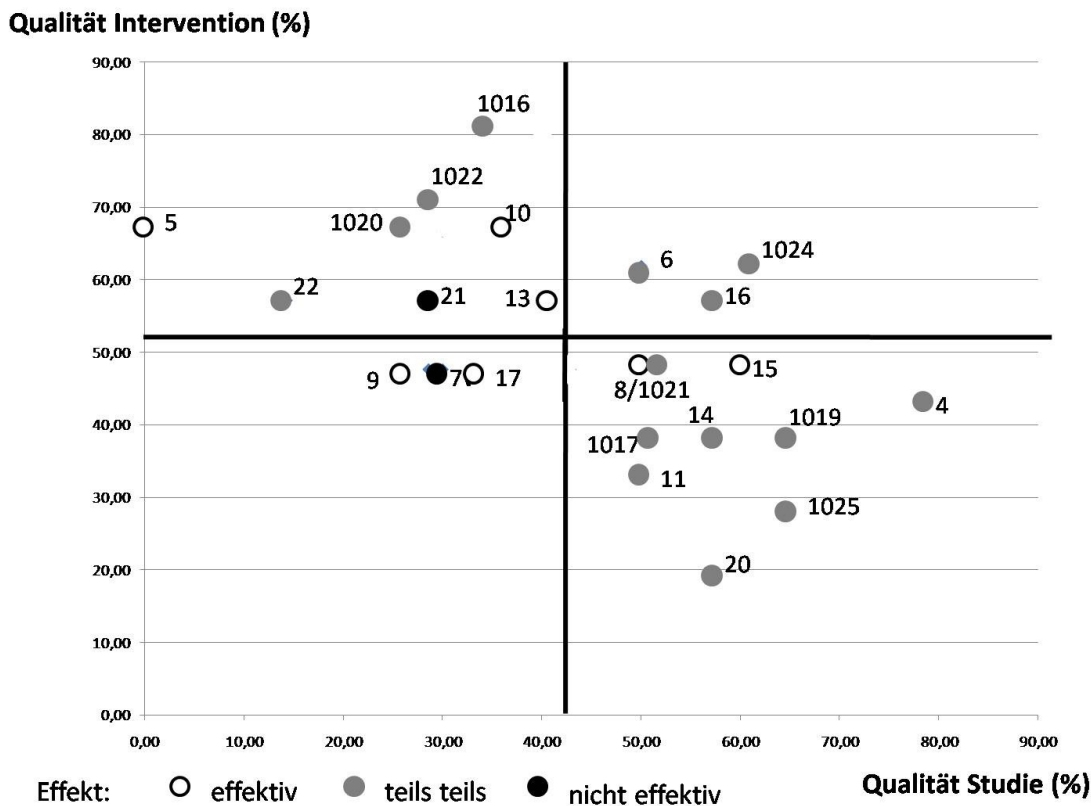
Nachdem gezeigt wurde, wie die evaluierten Studien in der Qualität der Studiendurchführung abgeschnitten haben, soll die folgende Tabelle 26 einen Überblick über die Qualität der Interventionen geben.

Tabelle 26: Qualität der Intervention im Handlungsfeld Ernährung

Kennnummer	Studien-design	Partizipation (54)	Empowerment (55)	Prozess dokumentiert (63)	Prozess evaluiert (64)	Summe	Erzielbarer Maximal-score	Prozentualer Score
3004	RCT	0	0	1	0	4,5	10,5	42
3005	RCT	0	1	1	1	7	10,5	66
3006	RCT	0	0	0	1	6,5	10,5	61
3007	RCT	0	1	0	1	5	10,5	47
3008	RCT	0	1	0	1	5	10,5	47
3009	CT	0	1	0	0	5	10,5	47
3010	RCT	0	1	1	1	7	10,5	66
3011	RCT	0	0	1	1	3,5	10,5	33
3013	Non-CT	0	1	0	1	6	10,5	57
3014	RCT	0	1	0	1	4	10,5	38
3015	CT	0	1	0	1	5	10,5	47
3016	RCT	0	0	1	1	6	10,5	57
3017	CT	0	1	0	1	5	10,5	47
3020	RCT	0	0	0	0	2	10,5	19
3021	RCT	0	1	1	0	6	10,5	57
3022	RCT	0	1	1	1	6	10,5	47
1016	CT	0	1	1	1	8,5	10,5	80
1017	CT	0	0	0	0	4	10,5	38
1019	RCT	0	0	0	0	4	10,5	38
1020	CT	0	0	1	1	7	10,5	66
1021	CT	0	0	0	0	5	10,5	47
1022	RCT	0	0	1	1	7,5	10,5	71
1024	CT	0	0	1	1	6,5	10,5	61
1025	RCT	0	0	1	0	3	10,5	28

Für den Bereich Ernährung werden im nachfolgenden Diagramm auf der x- Achse die Qualität der Studie und auf der y-Achse die Qualität der Intervention dargestellt. Die Farbe der einzelnen Punkte stellt die Effektivität der evaluierten Studie dar (schwarz= nicht effektiv, grau= Effekt teils teils vorhanden, weiß= effektiv).

Abbildung 15: Ergebnisse der Studienevaluation im Bereich Ernährung



Legende: In der Abbildung wurde auf die hohe Nummerierung verzichtet. Studie 4 ist demnach synonym zu Studie 3004, Studie 9= 3009, etc.

Der Durchschnittswert für die Qualität der Studie (x-Achse) liegt bei 43,3%, der Mittelwert für die Qualität der Intervention (y-Achse) liegt bei 51,2%. Die Abbildung zeigt die 24 evaluierten Studien. Von diesen zeigten sieben Interventionen einen Effekt, 15 Interventionen hatten nur einen Teileffekt und zwei Interventionen waren nicht effektiv. Eine Studie (1024) fällt in den als „stark empfehlenswert“ definierten Bereich (siehe dazu Methodenteil). Zwei Studien liegen im Bereich „eingeschränkt empfehlenswert“. Neun Studien haben zwar eine überdurchschnittliche Qualität der Studie, weisen aber eine unterdurchschnittliche Interventionsqualität auf. Umgekehrt wurden acht Interventionen mit überdurchschnittlicher Interventionsqualität bewertet, die eine unterdurchschnittliche Studienqualität erreichten.

In einigen Studien wurde eine Vielzahl an Effektmaßen analysiert. Für die vorliegende Analyse wurde aber nur eine begrenzte Anzahl dieser Effektmaße ausgewählt. Es wurden nur die Studien als effektiv bewertet, die eine signifikante Verbesserung aller ausgewählten Effektmaße zeigen

konnten. Wurde festgestellt, dass nur einzelne Parameter einen Effekt hatten, wurden die Studien als teils teils effektiv bewertet.

Zwischen Qualität und Effekt (s. Tabelle 27) wurden folgende Korrelationen nach Spearman erhoben.

Tabelle 27: Korrelation von Effekt, Qualität der Studiendurchführung und Qualität der Intervention im Handlungsfeld Ernährung

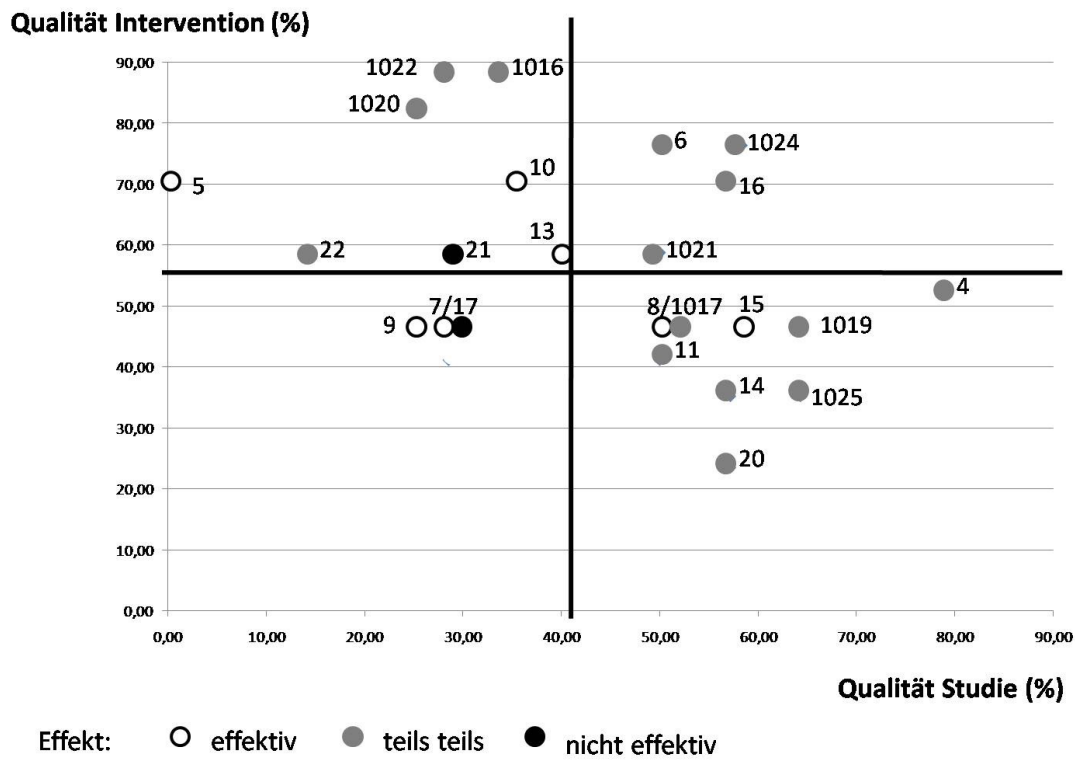
Zu korrelierende Größen	Spearmankorrelation
Qualität gesamt und Effekt	0,16
Qualität der Studie (QS) und Qualität der Intervention (QI)	-0,57
Qualität der Studie (QS) und Effekt	-0,03
Qualität der Intervention (QI) und Effekt	0,17

### **Darstellung der Sensitivitätsanalyse**

Auch für den Risikofaktor Ernährung wurde eine Sensitivitätsanalyse gemacht, indem auch hier die Kriterien Partizipation und Empowerment nicht in die Berechnung mit aufgenommen wurden.

Für die Sensitivitätsanalyse ist der Durchschnittswert für die Qualität der Intervention auf 57% gestiegen. Nach der neuen Berechnung liegen nun vier Studien im „eingeschränkt empfehlenswerten“ Bereich. Die Studie von Gortmaker et al. (1999a) fällt aus dem empfehlenswerten Bereich in den eingeschränkt empfehlenswerten Bereich zurück. Die Studie von Coleman et al. (2005) liegt in der Sensitivitätsanalyse nun sowohl über der durchschnittlichen Studienqualität als auch über dem Durchschnittswert der Interventionsqualität.

Abbildung 16: Ergebnisse der Sensitivitätsanalyse im Bereich Ernährung



Für die Sensitivitätsanalyse wurde folgende Korrelation zwischen Qualität und Effekt berechnet.

Tabelle 28: Korrelation von Effekt, Qualität der Studiendurchführung und Qualität der Intervention im Handlungsfeld Ernährung nach Sensitivitätsanalyse

Zu korrelierende Größen	Spearmankorrelation
Qualität gesamt und Effekt	0,28
Qualität der Studie (QS) und Qualität der Intervention (QI)	-0,35
Qualität der Studie (QS) und Effekt	0,07
Qualität der Intervention und Effekt	0,2

### 8.2.3 Darstellung einzelner Studien

Nachfolgend werden drei Studien vorgestellt, die aufgrund der Evaluation in die Qualitätsbereiche „stark empfehlenswert“ und „eingeschränkt empfehlenswert“ fallen. Liegt der Qualitätsscore über den beiden Durchschnittswerten, kann die jeweilige Studie (eingeschränkt) empfohlen werden. Im Bereich Ernährung befinden sich zwei Studien in diesem Bereich.

#### ***Gortmaker et al. (1999a) (1024)***

Die Studie *Eat Well and Keep Moving* wurde in den USA durchgeführt. Die Intervention wurde als eine quasi experimentelle Studie mit sechs Interventions- und acht Kontrollschulen angelegt. Die Anzahl der Kinder und Jugendlichen (6- bis 17jährige) lag in der Interventionsgruppe bei 173 und in der Kontrollgruppe bei 163 Personen.

Die Intervention umfasste spezielle, vom Lehrer durchgeführte Unterrichtseinheiten, zu Ernährung und Bewegung (13 Einheiten pro Schuljahr). Zudem wurde in vier der sechs Interventions- und acht Kontrollschulen zusätzliche Sportunterricht angeboten. 18 sogenannte Eat Well Karten wurden entwickelt, welche die Lehrer benutzten, um die Kinder über eine Vielzahl von Nahrungsmitteln aufzuklären. Ebenfalls fanden Kampagnen mit der Botschaft: Iss mehr Obst, schau weniger fern statt.

Die Studiendauer betrug 24 Monate. Es wurde nicht angegeben, ob es eine Nachbeobachtung gab.

Als Effektmaße dienten:

- Der Fettanteil und der Anteil gesättigten Fetts an der gesamten Kalorienaufnahme,
- Die Anzahl an konsumierten Obst und Gemüse pro Tag,
- Das Wissen um Ernährung und Bewegung.

Folgende Effekte wurden festgestellt: Durch die Intervention kam es zu einer signifikanten Abnahme des Anteils an Fett und gesättigten Fettsäuren ( $p= 0,04$  bzw.  $0,05$ ). Bei dem Konsum von Obst- und Gemüse gab es eine signifikante Zunahme ( $p= 0,01$ ) in der Interventionsgruppe gegenüber der Kontrollgruppe. Dagegen gab es keine signifikante Veränderung im Bereich Bewegung.

**Reynolds et al. (2000) (3016)**

Das High 5 Project wurde in den USA durchgeführt. Es ist eines von neun durchgeführten Interventionsprojekten mit der Botschaft 5 a Day. Ziel der Studie war es den Konsum von Früchten und Gemüse bei Kindern zu steigern. 28 Grundschulen wurden randomisiert nach Interventions- und Kontrollbedingungen. Die genaue Zahl der eingeschlossenen Personen wurde nicht angegeben. Die Intervention gliederte sich folgendermaßen:

- Spezielle Unterrichtseinheiten zum Thema Ernährung (insgesamt 14 Unterrichtseinheiten)
- Einbindung von Familienmitgliedern. Eltern erhielten das „Freggie Book“ (Fruit and Vegetable), das sieben Hausaufgaben wie zum Beispiel den häufigen Verzehr von Früchten und Gemüse enthielt. Eltern werden angehalten, mit ihren Kindern die Hausaufgaben durchzuführen.
- „Food Service“:

Die Köche der Interventionsschulen erhielten ein halbtägiges Training, um die Anforderungen der HIGH 5 Leitlinien umsetzen zu können.

Die Intervention wurde insgesamt 24 Monate durchgeführt. Im Anschluss fand eine Nachbeobachtung von einem Monat statt.

Als Effektmaße wurden zum einen der Früchte- und Gemüsekonsum (Anzahl der konsumierten Früchte bzw. Gemüse pro Tag) gewählt. Weitere verwendete Maße waren der prozentuale Kalorienanteil aus Fett, der prozentuale Kalorienanteil aus Kohlenhydraten und der prozentuale Kalorienanteil aus Eiweiß.

Die Intervention führte zu folgenden Effekten: Der durchschnittliche tägliche Konsum von Früchten und Gemüse hat signifikant zugenommen. Der Konsum von Früchten stieg um 1,21 Portionen in der Interventionsgruppe, und um 0,65 Portionen in der Kontrollgruppe ( $p < 0,001$ ). Der Gemüseverzehr stieg um 1,60 Portionen in der Interventionsgruppe, und um 1,25 in der Kontrollgruppe ( $p < 0,009$ ). Fasst man den Früchte- und Gemüsekonsum zusammen, so wurden 3,20 Portionen in der Interventionsgruppe und 2,21 Portionen in der Kontrollgruppe verzehrt ( $p < 0,001$ ).

Des Weiteren gab es eine signifikante Verbesserung im prozentualen Anteil von Kalorien aus Fett: Dieser lag in der Interventionsgruppe bei 31,6% und in der Kontrollgruppe bei 33,3% ( $p < 0,04$ ).

Es gab keine signifikante Verbesserung im prozentualen Anteil von Kalorien aus Kohlenhydraten: Der Anteil lag in der Interventionsgruppe bei 55,2% und in der Kontrollgruppe bei 53,2% ( $p < 0,0165$ ). Ebenfalls konnte keine Steigerung im Fruchtekonsum bei den Eltern festgestellt werden.

Die Intervention im Projekt HIGH 5 hat auf jeden Fall positive Effekte im Bereich Früchte- und Gemüsekonsum bei Kindern gezeigt. Bei dem Früchte- und Obstkonsum der Eltern waren konnten keinen positiven Effekt nachgewiesen werden.

Auch für diese Studie werden in folgender Tabelle die Qualitätsmerkmale dargestellt:

### ***Townsend et al. (2006) (3006)***

Die Studie *Expanded Food and Nutrition Education Program* ist ein Präventionsprogramm gemäß den Leitlinien des United States Department of Agriculture (USDA) und wird speziell für Familien mit geringem Einkommen angeboten. Es wurde in den USA durchgeführt. Ziel der Studie ist es, Eltern und Kindern ausführliches Wissen über Ernährung zu vermitteln, um somit das Wohlbefinden der Familien zu steigern. Als Studiendesign wurde eine randomisiert kontrollierte Studie gewählt. 229 Gruppen mit insgesamt 5.111 neun- bis elfjährige wurden randomisiert. Die Größe der Interventionsgruppe bei der Auswertung betrug 3.586 Kinder und Jugendliche. Die Größe der Kontrollgruppe umfasste bei der Auswertung 1.526 Kinder und Jugendliche.

Die Intervention hat den Namen *Eating Right is Basic* (ERIB) und wird für neun bis elf Jahre alte Kinder angeboten. Sie umfasst sieben Unterrichtseinheiten, die innerhalb von sechs bis acht Wochen durchgeführt werden. Der Fokus der Intervention liegt darin, Kindern und Jugendlichen den häufigen Verzehr von Früchten und Gemüse näher zu bringen.

Zur Intervention gehören auch experimentelle Aktivitäten wie Geschmacksproben, „food art“, „food puzzles“, Spiele und die Zubereitung von Früchten und Gemüse.

Insgesamt dauerte die Intervention zwölf Monate. Es gab in der Publikation keine konkrete Angabe, ob eine Nachbeobachtung stattgefunden hat.

Es wurden vier Effektmaße verwendet. Dazu zählte:

- Der Konsum von vielseitigen Nahrungsmitteln,
- Die Wissensverbesserung über Nahrung,
- Die Verbesserung der Fähigkeit gute Nahrungsmittel zu konsumieren und
- Die Verbesserung in der Nahrungsmittelzubereitung.

Der Konsum von vielseitigen Nahrungsmitteln hatte in der Interventionsgruppe mit einem Wert von 0,12 (change score) keinen signifikanten Effekt. Die Veränderung in der Kontrollgruppe lag bei 0,16 ( $p = 0,82$ ).

Es kam zu einer signifikanten Verbesserung im Wissen über Nahrung. Hier hat die Interventionsgruppe Werte von 0,65 und die Kontrollgruppe 0,25 ( $p < 0,0001$ ).

Auch in der Fähigkeit, ausgewogene Nahrungsmittel zu konsumieren, gab es eine signifikante Verbesserung: Interventionsgruppe 0,07; Kontrollgruppe 0,01 ( $p < 0,008$ ). Zu einer signifikanten Verbesserung kam es auch in der Nahrungsmittelzubereitung: Interventionsgruppe 0,97; Kontrollgruppe 0,40 ( $p < 0,0001$ ). Tabelle xx zeigt die diese Studie.

### **8.3 Diskussion**

Alle evaluierten Studien verfolgen ähnliche Ziele. Dennoch können viele Unterschiede im Studiendesign, in der Qualität (Studie und Intervention) und in der Durchführung der Interventionen festgestellt werden. Deshalb ist es schwer, aus einzelnen Evaluationsergebnissen ein Gesamtergebnis zu konzipieren. Die Studien wurden in 13 ausgewählten Ländern durchgeführt. Die Studienpopulation stammt teilweise aus unterschiedlicher ethnischer Herkunft. Die Unterschiede in der Ergebnissen sind möglicherweise auch auf kulturell unterschiedliche Essgewohnheiten zurückzuführen; daher kann eine Übertragbarkeit schwer bzw. unrealistisch sein (z.B. Story et al. 2003 (Studie 3022)).

#### **Interventionsformen**

Die Interventionen fanden alle im Setting Schule und teilweise zusätzlich auch zu Hause statt. Die Schule eignet sich gut, um die Studienteilnehmer durch Lehrkräfte zu beeinflussen und zur Mitarbeit zu bewegen.

Damit wurde in allen Studien eine klare Zielgruppe angesprochen, nämlich Kinder und Jugendliche. D.h. auch, dass die Interventionen speziell auf ihre Zielgruppe zugeschnitten wurden.

Am häufigsten wurden Unterrichtseinheiten zum Thema Ernährung konzipiert. Hier sind exemplarisch die Studien von Gortmaker SL et al. (1999a), Townsend MS et al. (2006) und Reynolds KD et al. (2000), zu nennen. Die Lehrer vermitteln den Schülern Informationen über gesunde Ernährung und die Auswirkungen von Fehlernährung. Des Weiteren wurden kleine Workshops angeboten, in denen die Schüler die Möglichkeit hatten, gesund zu kochen (u.a. Townsend et al.

2006). In manchen Interventionsschulen wurde das Mittagsmenu auf fettreduzierte Gerichte umgestellt. Zusätzlich wurde das Angebot an Obst und Gemüse erhöht (u.a. Story et al. 2000, Nicklas et al. 1998, Resnicow et al. 1992).

Einige Interventionen wurden multidimensional ausgelegt, d.h. es wurden Aktivitäten im Bereich Ernährung sowie auch im Bereich Bewegung angeboten (z.B. Belansky et al. 2006, Hoelscher et al. 2004, Trevino et al. 2004). Die Sportunterrichtseinheiten wurden erhöht und zusätzliche Sportstunden während des Nachmittags angeboten.

Einen abweichenden Interventionsansatz wird in der Studie von Lowe et al. 2004 beschrieben. In einem sechs minütigen Video kämpfen die „Food Dudes“ (eine Gruppe von 12-13 Jährigen) gegen den ‚Junk Punk‘, eine Figur, die versucht die Welt zu erobern, indem sie allen Menschen Früchte und Obst entzieht (für nähere Informationen siehe URL:[http://www.fooddudes.co.uk/prog\\_over.htm](http://www.fooddudes.co.uk/prog_over.htm)).

### **Wirksamkeit von Interventionen**

Die Mehrheit der Studien hatte zum Ziel, den Frucht- und Obstkonsum bei Kindern und Jugendlichen zu erhöhen. Einige Studien verfolgen als ein weiteres Ziel den Anteil von Fett und Kalorien an der täglichen Nahrung zu senken, um kardiovaskulären Erkrankungen vorzubeugen (z.B. Reynolds et al. 2000).

15 der evaluierten Studien haben einen Teileffekt. Das sie teilweise wirksam sind obwohl sie aus eigentlich sehr komplexen Maßnahmen zusammengesetzt sind, hat zwei verschiedenen Gründe:

- 1) Die Studien haben im Vergleich zu den anderen eine extrem hohe Anzahl an einbezogenen Effektmaßen, die aber nicht alle signifikante Verbesserungen verzeichnen konnten.
- 2) Es handelt sich um Studien mit komplexen Interventionen, in denen der Fokus auf zwei Risikofaktoren (Ernährung und Bewegung) lag.

Die Bestimmung mehrerer Effektparameter birgt die Gefahr, dass keine vollständige Effektivität erreicht werden kann. Unsere Ergebnisse zeigen ebenso, dass eine Kombination von Interventionen im Bereich Bewegung und Ernährung nur im Bereich Ernährung Wirksamkeit zeigte (u.a. Belansky et al. 2006, Hoelscher et al. 2004, Gortmaker et al. 1999b).

Es hat sich herausgestellt, dass der Erfolg einer Intervention gefördert wird, wenn sie über die Schule hinaus weitergeführt wird, also z.B im Elternhaus oder in der Freizeit (u.a. Belansky et al. 2006). Des Weiteren konnten positive Effekte festgestellt werden, wenn das Menu in den Cafete-

rias der Schulen auf fettreduzierte Produkte umgestellt worden ist (z.B. Story et al. 2000, Lowe et al. 2004).

Von Interesse waren auch die Länge der Intervention und die Nachbeobachtung. Die Untersuchung hat gezeigt, dass mit einigen Ausnahmen die evaluierten Interventionen in der Regel ein bis zwei Monate durchgeführt wurden. Ausnahmen bilden z.B. die STRIP Studie aus Finnland (Kaitosaari et al. 2003) mit 84 Monaten oder u.a. die Studie von Belansky et al. 2006 mit 24 Monate und einer Nachbeobachtung von mindestens einem Monat. Die Mehrzahl der evaluierten Interventionen wurde mit einer noch kürzeren Dauer durchgeführt. Hinzu kommt, dass in der Mehrheit der Studien keine nachbeobachtung stattgefunden hat.

Es ist schwierig von wissenschaftlicher Evidenz für langfristige Wirksamkeit zu sprechen, wenn Interventionen unter 24 Monate durchgeführt werden und keine Nachbeobachtungen stattfinden.

# 9 Rauchen

## 9.1 Stand der Forschung

### Epidemiologie des Rauchens

Rauchen ist weltweit die zahlenmäßig größte vermeidbare verhaltensbedingte Ursache von koronarer Herzkrankheit, zerebrovaskulärer Krankheiten, von Lungen- und Bronchialkrebs, chronisch obstruktiven Atemwegserkrankungen und frühkindlichen Entwicklungsstörungen infolge Rauchens in der Schwangerschaft (Fagerstrom 2002). Der Einfluss des Rauchens auf die vorzeitige Sterblichkeit ist in epidemiologischen Studien eindeutig nachgewiesen worden. Mehr als die Hälfte aller regelmäßigen Raucher sterben vorzeitig an den Folgen des Konsums. Die Hälfte verstirbt davon bereits im mittleren Lebensalter (Doll et al 2004). Der statistische Durchschnitt verlorener Lebenszeit liegt bei über 8 Jahren (Doll und Peto 1994).

Die Raucherprävalenz in der Gesamtbevölkerung liegt nach Surveydaten von 2003 in der über 18-jährigen Bevölkerung bei fast 33% (Männer 37,3%, Frauen 28%). 25,4% rauchen täglich und 7,1% gelegentlich. 26,9% sind ehemalige Raucher. Somit haben nur 40% (31% der Männer und fast 50% der Frauen) niemals geraucht (Lampert und Burger 2004). Die Raucherquote ist unter den 18 bis 29-jährigen Männern und Frauen nach Altersgruppen vergleichsweise am höchsten, aber die stärksten Raucher finden sich unter den 50 bis 59-Jährigen (Lampert und Burger 2005). Insgesamt rauchen 9,4% der über 18-Jährigen stark, das heißt nach WHO-Definition 20 Zigaretten und mehr täglich, 6,6 % rauchen mittelstark, das heißt 11 bis 19 Zigaretten. Der Anteil stark Rauchender unter den täglichen Konsumenten liegt bei 39,6% (Männer 47%, Frauen 31,2%), der Anteil mittelstark rauchender täglicher Konsumenten liegt bei 27,4% (RKI 2006 nach Lampert und Burger 2005).

Das Rauchen hat unter den Frauen in den letzten 10 Jahren zugenommen, während der Konsum bei den Männern leicht abgenommen hat.

Der sozioökonomische Einfluss auf das Rauchverhalten zeigt ab Mitte des 20. Jahrhunderts eine Verschiebung der Prävalenz von den sozial besser gestellten Bevölkerungsgruppen hin zu sozial schlechter gestellten. Heute ist der Sozialindikator mit dem größten und unabhängigen Erklärungswert zum Rauchverhalten die Schulbildung. Nach Adjustierung für Alter und Geschlecht

haben Personen mit Hauptschulabschluss im Vergleich zu Personen mit Abitur ein mehr als zweifach höheres Risiko für Tabakkonsum. Für Berufsgruppen und Lebenslagen gibt es ebenfalls differenzierte Analysen zum Rauchverhalten. Die Raucherprävalenz ist in sozioökonomisch benachteiligten Bevölkerungsgruppen vier- bis sechsfach höher als in besser gestellten Teilen der Bevölkerung (Helmert et al. 2001).

50% der Kinder sind durch Passivrauchen belastet und haben ein höheres Risiko für SIDS (Sudden Infant Death), Reifestörungen und Erkrankungen der Atemwege. (RKI und StBuA 2006)

### **Prävalenz des Rauchens bei Schwangeren**

In Deutschland rauchen 20,6% aller Schwangeren. Werdende Mütter der Unterschicht rauchen doppelt so häufig (40%) wie Schwangere der Oberschicht (fast 24%). 62% der ledigen oder geschiedenen werdenden Mütter rauchen im Vergleich zu 17% der verheirateten (Helmert et al. 1998). Der Anteil der Raucherinnen unter den Müttern mit Kindern unter einem Jahr liegt bei 28,4%, der Anteil der Mütter mit bis zu fünf Jahre alten Kindern bei 41,7%.

### **Folgen des Passivrauchen**

Die Begriffe „Passivrauchen“, „Passivrauchbelastung“ und „Environmental Tobacco Smoke“, kurz ETS, werden synonym verwendet. Sie bezeichnen die unfreiwillige Aufnahme von Tabakrauch über das Blut (Ungeborene) und das ungewollte Einatmen tabakrauchbelasteter Raumluft. In Deutschland sind 55% aller Nichtraucher im Alter zwischen 18 und 79 Jahren häufig Tabakrauchbelastung ausgesetzt (abhängig von Alter und anderen sozioökonomischen Faktoren) (DKFZ 2005). Passivrauchen erhöht das Risiko, an Herz-Kreislauf-Krankheiten, Schlaganfällen, Atemwegserkrankungen und Lungenkrebs zu erkranken. Dies bestätigen beispielsweise Studien der EPA (Environmental Protection Agency 1993), der IARC (International Agency for Research on Cancer 2002) und der Report des US Department of Health 2004 (Schulze 2006). Besonders gefährdet sind Berufs- und Wohngruppen, die nicht in rauchfreier Umgebung arbeiten und leben.

### **Rauchende Schwangere: Auswirkungen auf das ungeborene Kind**

Es ist empirisch nachgewiesen, dass die Konzentration an Kotinin, einem Abbauprodukt des Nikotins, im Blut der Fötus bei gleich bleibender Rauchbelastung gerade im letzten Abschnitt der Schwangerschaft zunimmt. Ungeborene Kinder rauchender Mütter haben unter anderem ein

zweifach erhöhtes Risiko einer Frühgeburt und ein dreifach erhöhtes Risiko einer Totgeburt. Sie leiden oft an einem verringerten Geburtsgewicht, haben einen kleineren Kopfumfang und ein vermindertes Längenwachstum. Das Ausmaß dieser Effekte korreliert positiv mit der Stärke des Rauchens der Schwangeren (Lindley et al. 2000, Bernstein et al. 2005).

Wird das Rauchen zu Beginn der Schwangerschaft eingestellt, ist die Möglichkeit ein Kind mit normalem Geburtsgewicht auf die Welt zu bringen, genau so groß wie bei einer nicht- rauchenden Schwangeren (Ahlsten et al. 1993).

Viele Schwangere, die während der Schwangerschaft nicht geraucht haben, fangen innerhalb der ersten zwölf Monate nach der Geburt wieder an zu rauchen (DKFZ 2003). Säuglinge rauchender Müttern sind durch Passivrauchen und durch die mit Tabakschadstoffen belastete Muttermilch doppelt belastet. Dabei ist der Nikotingehalt der Muttermilch abhängig von der Anzahl der täglich gerauchten Zigaretten und dem zeitlichen Abstand zwischen dem Stillen und der zuletzt gerauchten Zigarette.

Es besteht ein klarer Zusammenhang zwischen mütterlichem Rauchen und dem Risiko für plötzlichen Kindstod (Sudden infant death SID). Dies ist in den Industrieländern die häufigste Todesursache für Säuglinge unter einem Jahr (Dybing und Sanner 1999). Säuglinge rauchender Mütter haben ein mehr als doppelt so hohes Risiko, an SIDS zu sterben, als Kinder nicht rauchender Mütter. Das Risiko für SIDS steigt mit zunehmender Anzahl gerauchter Zigaretten der Mutter und mit der Anzahl der Raucher im Haushalt (Schlaud et al., 1996).

Die Atemwege kleiner Kinder sind empfindlich, weil sie sich in der Entwicklung befinden und ihr Abwehrsystem noch nicht ausreichend ausgebildet ist. Durch die Passivrauchbelastung wird sowohl das Wachstum der Lungen verzögert als auch die Lungenfunktion verringert. Nachgewiesenermaßen haben Kinder aus Raucherhaushalten ein erhöhtes Risiko für verschiedene akute und chronische Atemwegssymptome (Office of Research and Development, 1992).

## **Internationale Evidenz zu Raucherentwöhnung bei Schwangeren und Müttern**

### **Neugeborener**

National und international wird eine Vielzahl von Raucherentwöhnungsprogrammen angeboten, zu deren Effektivität es einige Übersichtsarbeiten gibt. Die in Deutschland angebotenen Entwöhnungsprogramme sind selten hinsichtlich ihrer Effektivität evaluiert (Kröger und Shaw 2000).

Drei Reviews und eine Meta-Analyse zeigen exemplarisch den Forschungsstand. In der Meta-Analyse von Mullen et al. (1994) wurden in elf randomisiert kontrollierten Studien aus den USA, Schweden und Kanada Raucherentwöhnungsmaßnahmen auf ihre Effektivität untersucht. In zehn der Studien lag das durchschnittliche relative Risiko (=RR) bei 1,5 (Min. 1,22 bis Max. 1,86) in einem 95%-Konfidenzintervall. Die Interventionsgruppen zeigten eine um 50% höhere Erfolgsrate als die Kontrollgruppen.

In der Übersichtsarbeit von Melvin et al. (2000) wurde die Evidenz der Raucherentwöhnungsberatung für Schwangere anhand der „Treating tobacco use and dependence: a clinical practice guideline“ der „Agency for Healthcare Research and Quality’s Publication“ untersucht. Als Grundlage dienten internationale Reviews und Meta-Analysen. Das durchschnittliche relative Risiko der erfolgreichen Entwöhnungsraten betrug 1,7 (95%-KI: 1,3; 2,2). Allerdings hatten nur fünf der 16 zu Grunde gelegten Arbeiten einen signifikanten Effekt.

In einem Cochrane Review von Lumley et al. aus dem Jahr 2004 wurden 64 randomisierte und quasi-randomisierte Studien auf ihre Effektivität bezüglich der Entwöhnungsraten, des Geburtsgewichts und weiterer Parameter bei Schwangeren untersucht. Dabei wurde als Haupteffekt das kontinuierliche Rauchen in der Endphase der Schwangerschaft gemessen. Es zeigte sich eine signifikante Reduktion der Raucherinnenprävalenz in 48 Studien, entsprechend einem relativen Risiko von 0,94 für das kontinuierliche Rauchen (95%-KI: 0,93; 0,95). 17 Interventionen, die durch eine hohe Interventionsintensität, objektive Messung der Effektparameter und hohe Qualität gekennzeichnet waren, kamen zu einem ähnlichen Ergebnis. Dort lag das relative Risiko bei 0,95 (95%-KI: 0,93; 0,97).

Im Review von Mullen (2004) wurde anhand von zwölf Studien die Rückfallprävention untersucht. Danach blieben lediglich 20 bis 30% der Frauen, die in der Schwangerschaft das Rauchen aufgehört hatten, länger als ein Jahr abstinent. Im Allgemeinen waren die Erfolge der Interventionen bis zu sechs Monate nach der Schwangerschaft nachweisbar. Nur zwei Studien zeigten noch Effekte nach einem Jahr. Einige der Studien hatten gar keinen Effekt nach der Geburt.

## 9.2 Ergebnisse

Auf Grund der hohen Trefferzahl und der Vielfalt der Interventionsansätze und Intervention im Handlungsfeld Rauchen wurde eine thematische Auswahl getroffen. Da zu Teilbereichen wie schulische Raucherprävention die Ergebnisse systematischer Reviews vorliegen, haben wir uns

entschieden, Präventionsmaßnahmen für Raucherinnen in der Schwangerschaft und in der nachgeburtlichen Phase (bis Kindesalter ein Jahr) auszuwerten. Aus den Treffern der Literaturrecherche wurden 34 Studien zur Raucherentwöhnung bei Schwangeren und Müttern neugeborener Kinder in die Analyse anhand des Kriterienkatalogs zur Interventions- und Studienqualität einbezogen. Tabelle 29 gibt eine Übersicht über die eingeschlossenen Studien.

Tabelle 29: Übersicht über die eingeschlossenen Studien im Handlungsfeld Rauchen

Kennung	Autor (Jahr)	Projektname	Land	Studiendesign
4002	Wright et al. (1996)		USA	Non-CT
4004	Hartmann et al. (1996)		USA	RCT
4005	Gielen et al. (1997)		USA	RCT
4007	McBride et al. (1999)		USA	RCT
4008	Emmons et al. (2000)		USA	CT
4009	Glasgow et al. (2000)		USA	RCT
4010	Fossum et al. (2004)		Schweden	CT
4011	Severson et al. (1997)		USA	RCT
4012	Neil-Urban et al. (2002)		USA	Non-CT
4013	Valanis et al. (2001)		USA	Non-CT
4014	Ershoff et al. (1999)		USA	RCT
4015	Kendrick et al. (1995)		USA	RCT
4017	de Vries et al. (2006)		Niederlande	RCT
4018	Aveyard et al. (2005)		Großbritannien	RCT
4019	Wiggins et al. (2005)		Großbritannien	RCT
4020	Lawrence et al. (2003)		Großbritannien	RCT

4021	Jaakkola et al. (2001)		Finnland	CT
4022	Campion et al. (1994)		Großbritannien	Non-CT
4023	Rigotti et al. (2006)		USA	RCT
4024	Gebauer et al. (1998)		USA	CT
4025	Dornelas et al. (2006)		USA	RCT
4026	Hegaard et al. (2003)		Dänemark	RCT
4027	Avidano et al. (2006)		USA	CT
4028	Stotts et al. (2004)		USA	RCT
4029	Tappin et al. (2005)		Großbritannien	RCT
4030	Walsh et al. (1997)		Australien	RCT
4031	Moore et al. (2002)		Großbritannien	RCT
4032	Lowe et al. (1998)		Australien	RCT
4033	Lowe et al. (1998)		Australien	RCT
4034	Hughes et al. (2000)		Kanada	RCT
4035	Pbert et al. (2004)		USA	RCT
4036	Stotts et al. (2002)		USA	RCT
4037	Secker-Walker et al. (1994)		USA	RCT
4038	Valbo, Nylander (1994)		Norwegen	RCT

Tabelle 30: Qualität der Studiendurchführung im Handlungsfeld Rauchen

Ken- nung	Studien- design	Rando- mi- sation (16)	Ver- gleich- barkeit (19)	Meßin- strumen- te (25)	Power (46)	Summe	erzielba- rer Ma- ximals- core	Prozen- tualer Score
4002	non-CT	X	X	0	0	0	5	0

4004	RCT	1	1	0	0	4	7	57
4005	RCT	0	0	0	0	2	7	28
4007	RCT	0	1	0	1	5	7	71
4008	CT	X	0	0	0	1	6	16
4009	RCT	1	1	0	0	3	7	42
4010	CT	X	1	0	0	3	6	50
4011	RCT	1	0	0,5	0	3,5	7	50
4012	non-CT	X	X	0	0	1	5	20
4013	non-CT	X	X	0	1	3	5	60
4014	RCT	C	1	0	1	2	7	28
4015	RCT	1	1	0	1	3	7	42
4017	RCT	1	1	0	1	6	7	85
4018	RCT	1	1	0	1	5	7	71
4019	RCT	1	1	0	1	5	7	71
4020	RCT	1	1	0	0	4	7	57
4021	CT	X	1	0	0	2	6	33
4022	non-CT	X	1	0	0	2	5	40
4023	RCT	1	1	0	1	6	7	85
4024	CT	X	1	0	1	4,5	6	75
4025	RCT	1	1	0	0	2	7	28
4026	RCT	1	0	0,5	0	2,5	7	35
4027	CT	X	1	0	0	2,5	6	41
4028	RCT	0	1	0	0	1	7	14
4029	RCT	1	1	0	1	5	7	71
4030	RCT	1	1	0	0	3	7	42
4031	RCT	1	1	0	1	4,5	7	64
4032	RCT	0	0	0	0	1	7	14
4033	RCT	0	1	0	0	2	7	28
4034	RCT	1	1	0	0	3	7	42
4035	RCT	0	1	1	0	4	7	57
4036	RCT	1	1	0	0	4	7	57
4037	RCT	0	1	0	1	4	7	57
4038	RCT	0	0	0	1	2,5	7	35

Der Durchschnitt der Studienqualität liegt bei 46,5% und wird von 16 Studien übertroffen. Sieben Studien erreichen mehr als 70% der erreichbaren Qualitätspunktzahl. Das Maximum liegt bei 85,7%. Ob validierte Meßinstrumente eingesetzt wurden, wird nur in einer Studie explizit und in zwei weiteren teilweise erwähnt.

Tabelle 31: Qualität der Intervention im Handlungsfeld Rauchen

Kenn- ung	Studien- design	Partizipa- tion (54)	Empo- werment (55)	Prozess doku- mentiert (63)	Prozess evaluiert (64)	Summe	Erziel- barer Maximal- score	Prozen- tualer Score
4002	non-CT	0	0	0	0	3	10,5	28
4004	RCT	0	1	1	0	5	10,5	47
4005	RCT	0	0	1	1	6	10,5	57
4007	RCT	0	1	1	0	5	10,5	47
4008	CT	0	1	0	0	5	10,5	47
4009	RCT	0	1	0	0	6	10,5	57
4010	CT	0	1	0	0	6	10,5	57
4011	RCT	0	1	1	0	5	10,5	47
4012	non-CT	0	1	0	0	4	10,5	38
4013	non-CT	0	1	1	0	6	10,5	57
4014	RCT	0	1	0	1	6	10,5	57
4015	RCT	0	0	0	1	5	10,5	47
4017	RCT	0	1	0	0	5	10,5	47
4018	RCT	1	1	0	0	6	10,5	57
4019	RCT	0	0	0	1	5	10,5	47
4020	RCT	0	1	0	0	5	10,5	47
4021	CT	0	0	0	1	4	10,5	38
4022	non-CT	0	0	0	0	3	10,5	28
4023	RCT	0	1	0	0	5	10,5	47
4024	CT	0	1	0	0	4	10,5	38
4025	RCT	0	1	0	0	5	10,5	47
4026	RCT	0	0	0	0	3	10,5	28
4027	CT	0	1	1	0	6	10,5	57
4028	RCT	1	1	0	0	5	10,5	47
4029	RCT	0	1	1	0	6	10,5	57
4030	RCT	0	1	0	0	3,5	10,5	33
4031	RCT	1	0	0	1	5,5	10,5	52
4032	RCT	0	1	0	1	4,5	10,5	42
4033	RCT	0	0	0	0	3,5	10,5	33
4034	RCT	0	1	0	0	3	10,5	28
4035	RCT	0	1	1	0	6	10,5	57
4036	RCT	0	1	0	0	4	10,5	38
4037	RCT	0	1	1	0	4	10,5	38
4038	RCT	0	1	0	0	3,5	10,5	33

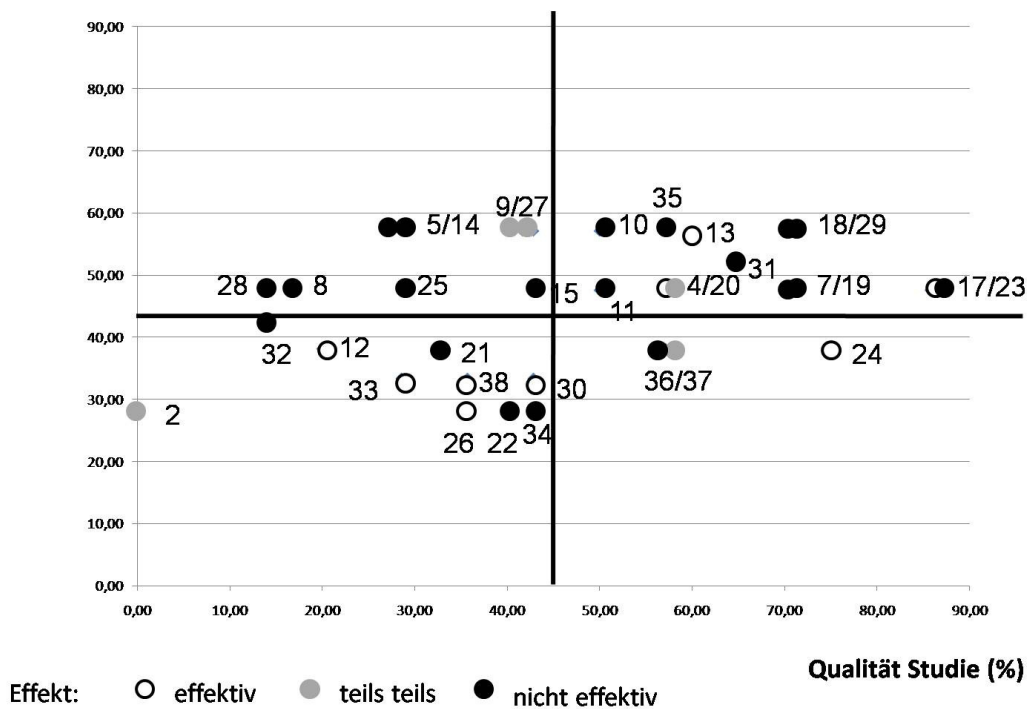
Die durchschnittliche Interventionsqualität liegt bei 45,2% der erreichbaren Punktzahl. 21 Studien liegen über dem Durchschnitt, das Maximum der erreichten Punktzahl liegt bei 71,4%. Nur

neun Studien geben an, den Interventionsprozess zu dokumentieren. Zum Teil gabe es auch in Studien ohne Angaben zur Dokumentation Hinweise auf eine Prozessevaluation.

Der Abbildung 17 lässt sich entnehmen, dass 21 der 34 Studien in Bezug auf die Interventionsqualität einen Score erreichen, der über dem Durchschnitt von 45,2% liegt, und 16 der 34 in Bezug auf die Studienqualität über dem Durchschnitt von 46,5% liegen. Die Schnittmenge der Studien mit dem höchsten Score in Studien- und in Interventionsqualität bilden dreizehn Studien, welche aufgrund der Gesamtqualitätsbewertung als eingeschränkt empfehlenswert eingestuft werden.

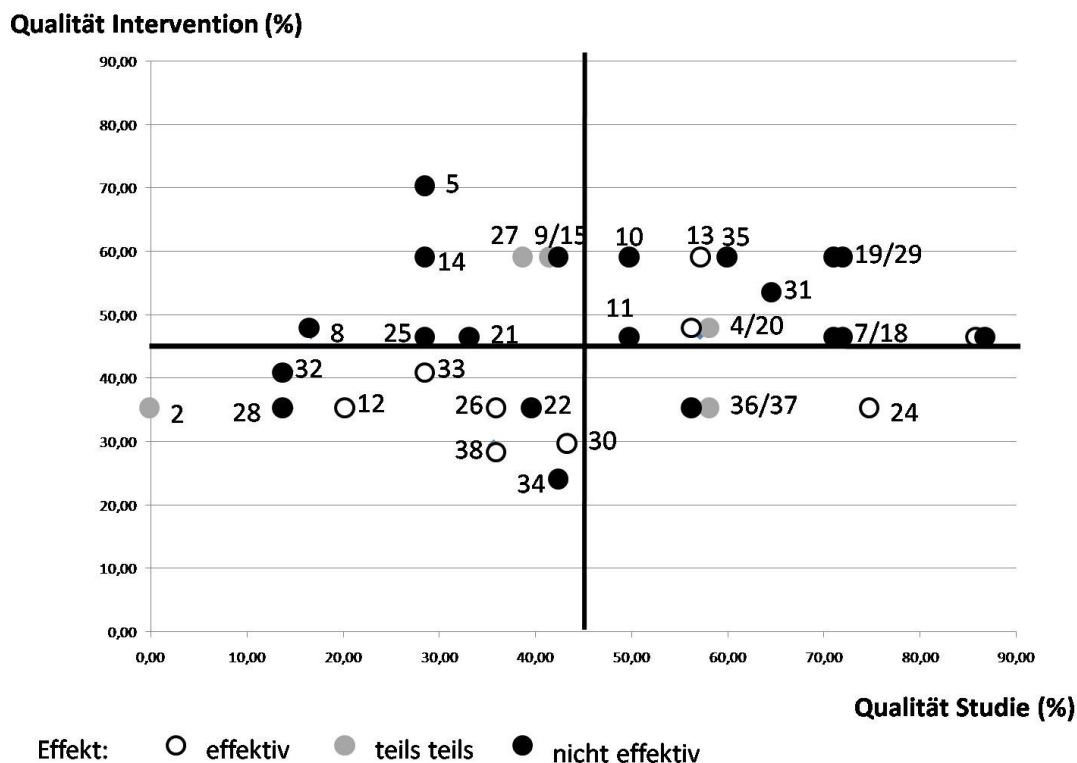
Abbildung 17: Vergleich Scores Studienqualität und Interventionsqualität von 34 Studien zur Raucherentwöhnung bei Schwangeren

**Qualität Intervention (%)**



Legende: 34 Studien sind in die Abbildung eingetragen. Es wird die Qualität der Studie gegen die Qualität der Intervention aufgetragen. Der Effekt wird als dritte Dimension durch unterschiedliche Punktfärbung eingebracht.

Abbildung 18: Sensitivitätsanalyse Risikofaktor Rauchen



Legende: Die Qualität der Intervention wurde neu berechnet ohne die beiden Qualitätskriterien Partizipation der Zielgruppe und Empowerment/ Selbstwirksamkeit. Auch wenn sich einige Studien in der Bewertung verschoben haben, so bleibt das Ergebnis bestehen: Dieselben 13 Studien kommen in den Bereich der Empfehlungen.

### 9.2.1 Darstellung einzelner Studien

Dreizehn Studien mit den höchsten Scores für Studien- und Interventionsqualität werden im Folgenden detaillierter dargestellt.

***Hartmann et al. (4004)***

Land: USA.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 107 vs. 100.

Kontrolle: Raucheranamnese.

Intervention: Nach Raucheranamnese und mündlicher Beratung der Schwangeren wird durch Pflegepersonal und Ärzte bei jedem Arztbesuch in einem Lehrkrankenhaus während der

Schwangerschaft Selbsthilfematerial zur Raucherentwöhnung verteilt (Windsor Self Directed 7 Day Smoking Cessation Guide). Dabei werden Zeitpunkte festgelegt, bis wann das Rauchen eingestellt werden soll ("Quitdates").

Studiendauer: 18 Monate.

Nachbeobachtung: keine.

Effektmaße: Raucherprävalenz, durch Kotinintest validiert; Anteil derjenigen, die das Rauchen um mindestens 50% reduziert haben; Anteil Personen, die unverändert rauchen, und Anteil der Raucher, die höchstens fünf Zigaretten am Tag rauchen.

Effekte und Signifikanzwert (p-Wert): Erfolgreiche Entwöhnung in der Intervention 20% vs. Kontrolle 10% (p= 0,052). Reduktion des Rauchens um 50 % in der Intervention bei 32% vs. Kontrolle bei 20% (p= 0,002). Unverändertes Rauchverhalten bei 31% vs. 40%. Verstärktes Rauchen bei der Intervention 18% vs. Kontrolle 30%. Höchstens 5 Zigaretten pro Tag: Intervention 50% vs. Kontrolle 31% (p= 0,007).

### ***McBride et al. (4007)***

Land: USA.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: Interventionsgruppe (1) 297, (2) 306, Kontrollgruppe 294.

Kontrolle: Informationsbroschüre per Post.

Intervention: Zwei Gruppen: (1) präpartale Beratung, (2) prä- und postpartale Beratung . Beide Interventionsgruppen erhielten auch persönliche Briefe sowie postalisch Material zur Rückfallprävention und telephonische Beratungsgespräche. Die zweite Interventionsgruppe erhielt auch postpartal telephonische Beratungsgespräche und postalisch einen Informationsbrief.

Studiendauer: Acht Monate.

Nachbeobachtung: Acht Monate.

Effektmaße: Sieben-Tage Abstinenz vor der Geburt; Rückfallrate nach der Geburt; Entwöhnung nach der Geburt; Sieben-Tage Abstinenz nach der Geburt.

Effekte und Signifikanzwert (p-Wert): Sieben-Tage Abstinenz vor der Geburt: beide Interventionsgruppen 50%, Kontrolle 47% (p< 0,17). Nach der Geburt: (signifikante Ergebnisse): Sieben-Tage Abstinenz über Follow-up von acht Wochen: Kontrolle 30%, präpartale Interventionsgrup-

pe IG (1) 35%, prä- und postpartale Interventionsgruppe IG (2) 39% ( $p=0,04$ ); Sieben-Tage Abstinenz über Follow-up von sechs Monaten: Kontrolle 26%, IG (1) 24%, IG (2) 33% ( $p=0,05$ ). Endpunkt: Sieben-Tage Abstinenz im Follow-up nach zwölf Monaten: Kontrolle 24%, IG (1) 23%, IG (2) 25% ( $p=0,96$ ).

***Fossum et al. (4010)***

Land: Schweden.

Studiendesign: CT.

Größe der Interventions- und Kontrollgruppe: 26 vs. 14.

Kontrolle: keine Angabe.

Intervention: Beratungsmethode für Mütter (fünf Teile): 1. Befragung der Mütter über das Bewusstsein der Effekte des Rauchens auf Kinder. 2. Bewusstseinsstärkung über Ausmaß der Tabakrauchbelastung in Nähe des Kindes. 3. Diskussion über Fragebogen zum Rauchverhalten und Befragung über aktuelles Rauchverhalten und Vorschläge zur Veränderung. 4. Unterstützung von Versuchen zur Änderung des Rauchverhaltens und Diskussion eventuell auftretender Probleme. 5. Unterstützung der Mütter, die während der Schwangerschaft aufgehört haben zu rauchen, um Rückfall nach Geburt vorzubeugen.

Studiendauer: Acht Monate.

Nachbeobachtung: keine.

Effektmaße: Aufhörrate, biochemisch validiert; Zigarettenzahl.

Effekte und Signifikanzwert (p-Wert):: keine signifikanten Effekte.

***Severson et al. (4011)***

Land: USA.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 1682 vs. 1219.

Kontrolle: Minimale Intervention: Besuch des Kinderarztes in Klinik nach Entbindung mit Bereitstellung von Informationsmaterial zur Raucherentwöhnung und Übergabe eines kurzen Briefs und einer Empfehlung zur Raucherentwöhnung.

Intervention: Erweiterte Intervention: Wie in "minimal intervention" und zusätzlich bei den ersten vier Besuchen beim Kinderarzt mit dem Neugeborenen (ca zwei - drei Wochen nach Geburt,

nach zwei Monaten, nach vier Monaten und nach sechs Monaten) kurze Beratung und Empfehlung zum Aufhören des Rauchens. Individualisierte Informationen in schriftlicher Form und Video bei einem der vier Besuche.

Studiendauer: Sechs Monate.

Nachbeobachtung: Sechs Monate.

Effektmaße: Aufhorrare, Rückfallrate, Aufhörbereitschaft.

Effekte und Signifikanzwert (p-Wert): keine signifikanten Ergebnisse für den Endpunkt zwölf Monate postpartum bei kontinuierlicher Abstinenz (sechs und zwölf Monate). Effekte bei Müttern, die während der Intervention immer noch rauchten: Anzahl Zigaretten pro Tag: Minimal 12,88 versus erweiterte Intervention 11,65 ( $p < 0,01$ ). Bereitschaft, das Rauchen aufzuhören: Minimal 5,95 vs. erweiterte Intervention 6,49 ( $p < 0,001$ ).

### ***Valanis et al. (4013)***

Land: USA.

Studiendesign: non-CT, da historischer Vergleich dreier Studiengruppen.

Größe nach der Intervention: 2055 Teilnehmende.

Intervention: Abschätzung und Dokumentation des Raucherstatus und Bereitschaft zur Raucherentwöhnung. Kurze individuelle Beratung zur Raucherentwöhnung und Besprechung mit der Schwangeren über Maßnahmen (Videos, Infobroschüren, Briefe für rauchende und nichtrauchende Partner der Schwangeren, Selbsthilfematerial), das Rauchen aufzugeben. Dokumentation, welche Maßnahmen in Anspruch genommen wurden.

Studiendauer: unklar.

Nachbeobachtung: zwölf Monate.

Effektmaße: Rauchstopp sechs bis neun Monate postpartum.

Effekte und Signifikanzwert (p-Wert): Raucherstatus während der Schwangerschaft als kontinuierliche Nichtraucherin: Historischen Vergleichsgruppe 28,9% versus Interventionsgruppe 38,8% ( $p = 0,001$ ). Status aufgehört, aber rückfällig: Historische Vergleichsgruppe 13,6% versus Interventionsgruppe 13,2% ( $p = 0,001$ ). Raucherstatus ein Jahr postpartum als Nichtraucherin in letzten 6 Monaten: Historische Vergleichsgruppe 14,9% versus Interventionsgruppe 18,4% ( $p = 0,005$ ). Raucherin bei einem-Jahr Follow up 78,3% versus Interventionsgruppe 76,3 ( $p = 0,005$ ).

Aktion, um Kind vor Tabakrauch zu schützen: Kontrollgruppe 73,4% versus Interventionsgruppe 88%, signifikant.

***De Vries et al. (4017)***

Land: Niederlande.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 136 vs. 175.

Kontrolle: Standardbehandlung, Faltblatt zur Raucherentwöhnung, keine spezifische Beratung.

Intervention: Raucherentwöhnung und Rückfallprävention bei Schwangeren durch die betreuende Hebamme. Die Intervention wurde zur bei den routinemäßigen Besuchen der Hebammen bei der Schwangeren durchgeführt und umfasste Videomaterial, Selbsthilfematerial und Broschüren für den Partner und Beratung und Unterstützung zur Raucherentwöhnung der Schwangeren.

Studiendauer: Fünf Monate.

Nachbeobachtung: 1,5 Monate.

Effektmaße: Sieben Tage Abstinenz an zwei Messzeitpunkten; kontinuierliche Abstinenz (Zigarettenabstinenz für vergangene sieben Tage sowohl für T1 als auch für T2); Raucherstatus des Partners durch Befragung der Schwangeren.

Effekte und Signifikanzwert (p-Wert): Der letzte Messzeitpunkt war sechs Wochen nach der Geburt. Kontinuierliche Abstinenz (Zigarettenabstinenz für vergangene sieben Tage sowohl für T1 als auch für T2). Interventionsgruppe 12% versus Kontrollgruppe 3%; signifikant.

Empfehlung und Übertragung: Im Hinblick auf die Raucherentwöhnung bei Schwangeren ist die Intervention in ihren zeitlichen Grenzen effektiv. Die im Vergleich hohe Qualität sowohl der Studie als auch der Intervention implizieren, dass eine derartige Maßnahme auch in Deutschland umsetzbar ist und erfolgreich in der sekundären Raucherprävention eingesetzt werden kann.

***Aveyard et al. (4018)***

Land: Großbritannien.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 629 vs. 289.

Kontrolle: Standardraucherentwöhnung durch Hebammen, kurze Ratschläge und Empfehlung zur Entwöhnung sowie eine Informationsbroschüre.

Intervention: Interventionsgruppe IG (1) wird von qualifizierten Hebammen mit ausführlichem schriftlichem Selbsthilfematerial versorgt, mündlich beraten und weiter betreut. Gespräch dauert ca. 15 Minuten. Interventionsgruppe IG (2): Die Hebammen sind wie in der ersten Interventionsgruppe qualifiziert. Zusätzlich arbeiten die Schwangeren mit einem Computerprogramm zur Anwendung des Selbsthilfeprogramms mit Auswertungs- und Feedbackmöglichkeit.

Studiendauer: unklar.

Nachbeobachtung: 1,8 Monate.

Effektmaße: Raucherpunktprävalenz des Partners (30. Schwangerschaftswoche (SSW) und 10 Tage nach der Geburt).

Effekte und Signifikanzwert (p-Wert): Aufhören des Partners zum Termin 30 SSW: KG 3,3% versus IG (1) 4,1% versus IG (2) 5,2% ( $p=0,77$ ). Rate 10 Tage nach der Geburt: KG 8% versus IG (1) 4,7% versus IG (2) 7,9% ( $p=0,4$ ). Weiterhin wurde das "Inventory of socially supportive behaviours" eingesetzt, wobei die Veränderung von Baseline bis zur 30. SSW beim sog. ISSB combined score gemessen wurde: KG -0,13 versus IG (1) -0,09 versus IG (2) -0,08 ( $p=0,68$ ). Zehn Tage nach der Geburt: Sog. ISSB combined score in KG 0,16 versus IG (1) 0,10 versus IG (2) 0,16 ( $p=0,53$ ).

### ***Wiggins et al. (4019)***

Land: Großbritannien.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: Interventionsgruppe 1: 158, Interventionsgruppe 2: 145, Kontrollgruppe 298.

Kontrolle: Routinemäßige NHS-Versorgung. Ein Besuch nach der Geburt zur Standardbetreuung. Weitere Beratung in der Klinik möglich.

Intervention: Intervention eine Unterstützung durch Gesundheitsberater: Angebot für einjährige, monatliche Betreuung zu Hause mit mündlicher Beratung zu allgemeinen Fragen der Mutter und Hinweisen. Beginn: ca. zehn Wochen nach Geburt. Intervention zwei Gemeindegruppe: Mitgliedschaft in einer von acht Gemeindegruppen für ein Jahr. Betreuung für Mütter mit Kindern bis zu fünf Jahren. Verschiedene Dienste wie Telefonberatung, Hausbesuche, Gruppensitzungen. Kontrollgruppe NHS Standardbetreuung für Mütter, wenn das Baby 10-15 Tage alt ist (ein Hausbe-

such). Danach Möglichkeit zur Betreuung in der Klinik. Keine weiteren planmäßigen Hausbesuche.

Studiendauer: Zwölf Monate.

Nachbeobachtung: Sechs Monate.

Effektmaße: Raucherprävalenz der Mutter, nach zwölf Monaten und Nachbeobachtung nach 18 Monaten.

Effekte und Signifikanzwert (p-Wert): Mutter raucht gemessen nach dem relativen Risiko (95%-Konfidenzintervall): zwölf-Monate: Gesundheitsberatung versus Kontrolle 0,86 (0,62 - 1,19) und Gemeindegruppe versus Kontrolle 0,97 (0,97 - 1,33). 18-Monate: Gesundheitsberatung versus Kontrolle 0,98 (0,69 - 1,39); Gemeindegruppe versus Kontrolle 1,06 (0,76 - 1,47). Nicht signifikant.

### ***Lawrence et al. (4020)***

Land: Großbritannien.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: Intervention (1) 243, Intervention (2) 251, Kontrolle 217.

Kontrolle: Standardversorgung: Raucherberatung durch Hebamme mit Informationsbroschüre "Thinking about stopping".

Intervention: Intervention auf der Basis des "Transtheoretischen Modells" durch Hebammen in zwei Interventionsgruppen: IG (1) TTM-basiertes Selbsthilfemanual (n=243) - "Prochange programme for a healthy pregnancy", sechs Manuale zu je 30 Seiten, ein Manual pro Phasenwechsel. Unterstützung und Betreuung durch Hebamme. IG (2) TTM-basiertes Selbsthilfemanual und zusätzlich Sitzungen mit Computerprogramm zur individuellen Raucherentwöhnungsberatung.

Studiendauer: unklar.

Nachbeobachtung: 1,8 Monate.

Effektmaße: Aufhörrate biochemisch validiert, Punktprävalenz 28.-30. Schwangerschaftswoche (SSW) und zehn Tage postpartum. Zigarettenzahl pro Tag.

Effekte und Signifikanzwert (p-Wert): 30. SSW kombinierte Ergebnisse von IG (1) und (2) im Vergleich zur Kontrolle gemessen nach adjustiertem Odds ratio (OR) (95%-Konfidenzintervall).

IG (1): Anhaltende Abstinenz seit 20. SSW: OR 1,41 (0,65 - 3,06). Punktprävalenz: 2,11 (1,14 -

3,88). IG (2): Anhaltende Abstinenz seit 20. SSW: OR 2,16 (0,88 - 5,31). Punktprävalenz: 2,66 (1,23 - 5,75). Punktprävalenz für Rauchabstinenz 18 Monate Follow up: OR 1,12 (0,65 - 2,22). Prävalenz für anhaltende Abstinenz seit 10 Tage postpartum: OR 1,15 (0,22 - 5,97). Unterschiede im täglichen Zigarettenkonsum (Selbsteinschätzung) und der Kotininkonzentration nach Interventionsgruppe. Vergleich Baseline-Prävalenz zu Prävalenz 30. SSW. Gemessen als Änderung der Zigarettenzahl/Tag: Kontrolle 5,3 versus IG(1) 4,4 versus IG(2) 1,4 ( $p < 0,001$ ). Änderung der Kotininkonzentration: Kontrolle 0,98 versus IG (1) 0,67 versus IG (2) -0,6 ( $p = 0,033$ ). Vergleich Baseline-Prävalenz mit Prävalenz 10 Tage postpartum gemessen als Änderung der Zigarettenzahl/Tag: Kontrolle 6,1 versus IG (1) 5,7 versus IG(2) 1,1 ( $p < 0,001$ ). Änderung der Kotininkonzentration: Kontrolle 0,62 versus IG (1) versus 1,52 IG(2) 0,36 ( $p = 0,194$ ). Vergleich Baseline-Prävalenz zu Prävalenz 18 Monate postpartum gemessen als Änderung der Zigarettenzahl/Tag: Kontrolle 8,0 versus IG (1) 6,3 versus IG (2) 4,5 ( $p = 0,001$ ). Empfehlung und Übertragung: teilweise effektive Intervention im eingeschränkt empfehlenswerten Bereich. Überprüfung der Machbarkeit innerhalb der strukturellen Bedingungen in Deutschland.

***Rigotti et al. (4023)***

Land: USA.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 220 vs. 222.

Kontrolle: „Best-Practice“ Kurzberatung, fünf-minütiges Beratungstelefonat durch einen qualifizierten Berater.

Intervention: Telefonische Raucherberatung für Schwangere durch qualifiziertes Personal während der Schwangerschaft und zwei Monate postpartum. Bis zu fünf Telefonate mit durchschnittlich insgesamt 68 Min. Dauer. Ergänzend Feedback per e-mail sowie Informationsmaterial. Individuell auf den Einzelfall zugeschnittene Beratung und Materialbereitstellung. Im Vergleich Kontrollgruppe mit einem Telefonat (fünf Minuten im Schnitt) und kurzer Broschüre, die per e-mail zugeht.

Studiendauer: 3,5 Monate.

Nachbeobachtung: Einen Monat.

Effektmaße: Sieben-Tage-Abstinenz, biochemisch validiert, am Ende der Schwangerschaft, Punktprävalenz drei Monate postpartum, kontinuierliche Abstinenz zu beiden Zeitpunkten, Reduktion der Zigarettenzahl.

Effekte und Signifikanzwert (p-Wert): Sieben-Tage-Abstinenz, biochemisch validiert am Ende der Schwangerschaft in Interventionsgruppe 10% versus Kontrollgruppe 7,5% (p= 0,39). Drei Monate postpartum: IG 6,7% versus KG 7,1% (p= 1,0). Kontinuierliche Abstinenz IG 4,8% versus KG 3,3% (p= 0,47). Signifikante Reduktion ( $\geq 50\%$ ) in IG 29,2% versus KG 21,7% (p= 0,09). Drei Monate postpartum IG 17,7% versus KG 16,2% (p= 0,69). Keine signifikanten Effekte.

### ***Tappin et al. (4029)***

Land: Großbritannien.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 351 vs. 411.

Kontrolle: Standardgesundheitsberatung mit Informationen bezüglich der Raucherentwöhnung bei Schwangeren anhand eines Gesundheits-Aufklärungsbuchs für alle Schwangeren in Schottland.

Intervention: Raucherentwöhnung bei Schwangeren. Zusätzlich zu Standardentwöhnungsempfehlungen bieten speziell qualifizierte Hebammen zwei bis fünf Besuche bei den zu betreuenden Schwangeren an. Durchführung von "motivational interviewing" während dieser Sitzungen.

Studiendauer: Drei Monate.

Nachbeobachtung: Keine.

Effektmaße: Effektstärke wurde anhand von Aufhörraten unmittelbar nach der Intervention erhoben und durch Kotininmessungen überprüft.

Effekte und Signifikanzwert (p-Wert): Aufhörrate in IG 4,8% versus KG 4,6%, damit Relatives Risiko 1,05 (95%-KI: 0,55; 1,98). Reduktion der Zigarettenzahl IG 4,3% versus KG 6,3%, damit Relatives Risiko 0,68 (95%-KI: 0,36; 1,25). Unveränderte Zigarettenzahl: IG 85,8% versus KG 78,3%, damit Relatives Risiko 1,10 (95%-KI: 1,02; 1,17). Zunahme der Zigarettenzahl: IG 5,1% versus KG 10,7%, damit Relatives Risiko 0,48 (95%-KI 0,28; 0,81). Keine signifikanten Ergebnisse.

***Moore et al. (4031)***

Land: Großbritannien.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 600 vs. 695.

Kontrolle: Standardversorgung ("usual care").

Intervention: Auf der Veränderungsbereitschaft basierendes Selbsthilfemanual „Stop For Good“ zusätzlich zur Standardversorgung. Einführung in die Benutzung und Beratung durch betreuende Hebamme beim ersten Routinebesuch. Weitere Broschüren (individuell zusammengestellt und angepasst) werden im Wochenrhythmus zugesandt.

Studiendauer: 1,8 Monate.

Nachbeobachtung: keine.

Effektmaße: Sieben-Tage Abstinenz, biochemisch validiert, Anzahl Zigaretten pro Tag.

Effekte und Signifikanzwert (p-Wert): Raucherstatus Ende des 2. Schwangerschaftstrimesters: Anteil der Nichtraucherinnen IG 18,8% versus KG 20,7%, nicht signifikant (95%- KI -3,5%; 7,3%). Differenz 1,9%, keine signifikant unterschiedlichen Raucherentwöhnungsraten am Ende des 2. Schwangerschaftstrimesters.

***Pbert et al. (4035)***

Land: USA.

Studiendesign: RCT.

Größe der Interventions- und Kontrollgruppe: 120 zu 161.

Kontrolle: Standardbehandlung

Intervention: Beratung und Information der Schwangeren in drei Gesundheitszentren in Boston durch speziell qualifizierte Ärzte und Schwestern während der routinemäßigen Sitzungen während und nach der Schwangerschaft. Zusätzlich Verteilung von Informationsmaterial in den jeweiligen Einrichtungen. ("Quit-together-study").

Studiendauer: Zwei Monate.

Nachbeobachtung: Sechs Monate.

Effektmaße: Sieben-Tage Abstinenz, biochemisch validiert, Zigarettenzahl pro Tag.

Effekte und Signifikanzwert (p-Wert): kein statistischer signifikanter Effekt unabhängig vom Zeitpunkt der Messung beim Vergleich zwischen Kontroll- und Interventionsgruppe.

### 9.3 Diskussion

Von dreizehn aufgrund der Studien- und Interventionsqualität – entsprechend der Indikatoren des Kriterienkatalogs eingeschränkt empfehlenswerten Studien – weisen neun keine signifikanten Effekte auf, eine ist teilweise effektiv je nach Effektmaß (4020) und drei sind effektiv (4004, 4013, 4017; s. Abb. 17). Diese drei Studien weisen alle ein multimodales Vorgehen auf, in welchem sowohl persönliche Beratung und Motivation der Schwangeren, als auch die Bereitstellung von Selbsthilfematerial und Informationen angewendet wurde. Allerdings sind auch andere Studien bzw. Programme multimodal angelegt, die in dieser Kategorie eingeschränkter Empfehlung aber keinen positiven Effekt hatten.

Exemplarisch greifen wir im Folgenden aus den dreizehn Studien einzelne heraus. Studie 4017 z. B. hat eine vergleichsweise hohe Qualität und ist signifikant effektiv, wie die Interventionsergebnisse in den Niederlanden zeigen. Eine derartige Maßnahme sollte versuchsweise auf Deutschland übertragen werden, da vergleichbare kulturelle Bedingungen für die Intervention bestehen und die Machbarkeit innerhalb der Strukturen in Deutschland leicht prüfbar ist.

Studie 4029 (Tappin et al. 2005) war trotz der vergleichsweise hohen Qualität ineffektiv. Hohe Qualität ist somit kein Garant für Erfolg. Ob unter anderen strukturellen Bedingungen erwünschte Effekte erzielt werden können, ist nicht abschließend zu bewerten.

Wie schon gesagt, finden sich im Handlungsfeld Rauchen auch unter den nicht effektiven Studien multimodale Interventionen. In weiteren Programmen, die bei hoher Qualität ein negatives Ergebnis hatten, wurde allerdings nur eine Beratung durchgeführt, und in zweien wurde neben der telefonischen Beratung Material bereitgestellt.

Vier der 34 untersuchten Studien hatten ihren Fokus auf sozioökonomisch benachteiligte Schwangere (4005, 4019, 4022, 4033) gerichtet. Diese speziellen Programme zeigten keine signifikanten Effekte. Die Anwendung des auf der Veränderungsbereitschaft beruhenden („Stage-of-Change“) Ansatzes bewirkte in den vorliegenden Studien nicht die gewünschten Effekte. Die Studie 27, eine der fünf Studien (4015, 4023, 4027, 4031, 4034), die diesen Ansatz explizit zur Grundlage der Beratung der Schwangeren oder Gestaltung von Material machten, war teilweise signifikant effektiv.

Bei den effektiven Studien, die unterhalb der Grenze der Mittelwerte bei der Interventions- und Studienqualität liegen (4012, 4024, 4026, 4030, 4033, 4038), lässt sich nicht beurteilen, ob derartige Interventionen unter höherer Studienqualität ihre Effekte beibehalten würden. Die Studie

4022 (Campion et al. 1994) liegt ebenfalls in diesem Bereich, ist aber nicht effektiv. Es ist denkbar, dass derartige Programme bei höherer Gesamtqualität der Intervention und der Studie an Wirksamkeit gewinnen können.

Drei Studien konnten identifiziert werden (4004, 4013, 4017), deren Übertragbarkeit auf deutsche Verhältnisse sich zu diskutieren lohnt, da sie eine vergleichsweise hohe Studien- und Interventionsqualität sowie einen positiven Effekt haben. Die Heterogenität der Methoden in den Studien und Interventionen erlaubt dennoch keine klare Empfehlung. Es muss im Einzelfall geklärt werden, ob die strukturellen Voraussetzungen der effektiven Programme eine Übertragung zulassen. Zusammenfassend ist festzuhalten, dass auch die internationale Studienlage keine eindeutigen Empfehlungen erlaubt. In der Tendenz sind effektive Interventionen mit hoher Qualität in der internationalen Studienlage multimodal angelegt und wurden durch Hebammen, Ärzte oder Pflegepersonal während der routinemäßigen Schwangerschaftsuntersuchungen durchgeführt.

Die strukturellen Verhältnisse in Deutschland sind daraufhin zu überprüfen, ob sie die Übernahme von Präventionsmaßnahmen für Schwangere und Mütter neugeborener Kinder flächendeckend zulassen und wie diese finanziert werden können.

## **10 Gesamtdiskussion und Fazit**

### **10.1 Zur Situation der Prävention**

Evaluation in Prävention und Gesundheitsförderung ist nach wie vor ein stark emotional besetztes Thema in einem methodisch und gesundheitspolitisch kontrovers geführten Diskurs. In Teilen hinken Positionen in Deutschland zu Evaluation von Prävention und Gesundheitsförderung sowohl der Kuration als auch dem Stand der Evaluation in den führenden Ländern wie z. B. Finnland hinterher. In den 80er Jahren des letzten Jahrhunderts war es extrem schwierig und umstritten, eine umfassende und übergreifende Datenlage zu einer kurativen Fragestellung zusammenzutragen. Es wurde behauptet, das medizinische Vorgehen sei Kunst und damit nur in Grenzen mess- und bewertbar. Auch wenn dies so nicht mehr behauptet wird, war und ist zum Teil die Praxis bei der Auswahl von zu fördernden Projekten in Prävention und Gesundheitsförderung von dieser Einstellung geprägt. Dies wird insbesondere deutlich, wenn der Vergleich zu den Ländern gezogen wird, die in der Priorisierung von Interventionen nach klar standardisierten Verfahren führend sind. So sind Priorisierung und Evaluation in Finnland seit Jahren selbstverständlich von der Mess- und Bewertbarkeit von Interventionen in Prävention und Gesundheitsförderung geprägt.

### **10.2 Katalog zur Priorisierung von Präventionsvorhaben**

Mit dem hier vorliegenden Forschungsvorhaben ist unsere Arbeitsgruppe in Deutschland keineswegs die einzige Forschergruppe, die sich dem Problem der Priorisierung von Interventionen in Prävention und Gesundheitsförderung stellt und widmet. Mit dieser Arbeit legen wir aber ein (neues) Instrument vor, das zur Priorisierung von Interventionen herangezogen werden kann.

Unser methodisches Vorgehen versteht sich als Synthese aus insgesamt drei Forschungsbereichen: Evidenzbasierte Medizin, Evaluation in der Prävention sowie Gesundheitsförderung und schließlich Qualitätssicherung. Damit werden die Ansätze best practice und systematisches Review bzw. Metanalyse im Sinne eines Health Technology Assessment für Interventionen der Prävention und Gesundheitsförderung zusammen geführt.

Es wurde ein Screeninginstrument entwickelt, das es ermöglicht, Interventionen, deren Ergebnis (positiver oder negativer oder in sich widersprüchlicher Effekte) hinlänglich abgesichert ist, von denen zu trennen, bei denen Unsicherheit über ihr Ergebnis angebracht ist. Konkret heißt das, dass die Interventionen nach den drei Dimensionen Effekt, Qualität der Studiendurchführung und Qualität der Intervention evaluiert worden sind.

Dabei wurde zunächst aus der evidenzbasierten Medizin sowie aus bestehenden Instrumenten zur Qualitätssicherung von Interventionen der Prävention und Gesundheitsförderung ein Instrument (Kriterienkatalog) entwickelt. Dieser Kriterienkatalog vereint neben rein deskriptiven Kriterien wie Dauer der Studie etc. auch Kriterien, die Indikatoren für Qualität sind bzw. sein können. Für die Studiendurchführung werden Kriterien wie Darstellung der Randomisierung, Berechnung von statistischer Power und Validierung der Auswertungsinstrumente herangezogen. Für die Qualität der Intervention werden Kriterien wie Dokumentation und Evaluation des Prozesses sowie Einbeziehen und Schulung der Multiplikatoren einbezogen. Zwei klassische Kriterien der Prävention und Gesundheitsförderung, nämlich Partizipation der Zielgruppe und Empowerment/ Stärkung der Selbstwirksamkeit, werden ebenfalls berücksichtigt. Da sie im Gegensatz zu den anderen aber wesentlich stärker unter dem Vorzeichen normativer Setzung diskutiert werden bzw. dem Streit unterliegen, ob sie zunächst als wirksam für das Ergebnis einer Intervention nachgewiesen werden müssen, wurde eine Sensitivitätsanalyse angeschlossen, bei der diese beiden Kriterien aus der Wertung genommen wurden. Dabei haben sich jedoch die Ergebnisse nur minimal verändert; im Grunde blieben dieselben Studien in den Bereichen empfehlenswerter Interventionen. Dies hat vor allem den Grund, dass nur wenige Studien überhaupt die Kriterien Empowerment und Partizipation erfüllten, so dass sich im Gefüge der Studien zueinander wenig änderte. Gleichmaßen wollen wir damit nicht sagen, dass man auf die Qualitätskriterien Partizipation und Empowerment verzichten kann, wie in der Diskussion (s. oben 3.4) deutlich wurde.

### **10.3 Zur Übertragbarkeit und Wiederholbarkeit von Studienergebnissen**

Generell geht die Medizin davon aus, dass Wirksamkeiten von Therapien mit einer hohen Wahrscheinlichkeit wiederholt werden können. Nahezu die gesamte Pharmakotherapie basiert auf dieser Annahme. In Bereichen, wo (auch) Verhalten und sozialwissenschaftliche Messinstrumente

zum Einsatz kommen, kann der Zusammenhang zwischen Intervention und Effekt sehr viel schwächer ausfallen und muss nicht unbedingt reproduzierbar sein. Mehrere Aspekte aus den von uns evaluierten Studien sollen erwähnt werden:

- 1) Die Bewertung der Effekte von Studien mit widersprüchlichen Effekten kann unterschiedliche Ursachen haben. Zum einen können sich Subgruppen- und Gesamtanalyse unterscheiden. Zum anderen gibt es viele Studien, die über unterschiedliche Effektparameter in sich widersprüchliche Ergebnisse haben. Auch wenn diese oft in den Darstellungen der Publikationen als Erfolg dargestellt werden, so scheint es aus unserer Sicht als berechtigt, die Bewertung nicht als uneingeschränkt positiv auszusprechen. Beispiele finden sich etwa insbesondere auch im Bereich der Studien zu Treppensteigen im Betrieb (u. a. Adams und White 2002, Marshall et al. 2002, Titze et al. 2001). Widersprüche zwischen verschiedenen Messmethoden des Erfolges, z. B. objektive Messung der Treppennutzung durch Infrarotkameras versus subjektive Messung durch Befragung (bei Marshall et al. 2002), sollten dokumentiert werden (s. aktuell Engbers et al. 2007a). Problematisch bleibt weiterhin, dass eine Wertung erforderlich wird, welche vor Beginn der Studie festgelegten Erfolgsparameter wichtiger sind als andere. Ansonsten kann ein widersprüchliches Ergebnis nicht bewertet werden. Anzumerken ist, dass überhaupt vorab Erfolgskriterien formuliert werden sollten.
- 2) Studien, die widersprüchliche Erfolge zeigen, sind nach unserer Kategorisierung nicht per se als Misserfolg zu bewerten. Die genaue Betrachtung der Subgruppen oder einzelner Parameter muss angeschlossen werden und kann dazu führen, diese Interventionen zu kopieren bzw. abzulehnen oder mit weiteren Forschungsfragen erneut durchzuführen. So zeigte beispielsweise die Studie von Pate et al. (2005) eine Zunahme an Bewegung, jedoch keine Gewichtsveränderung. Dies kann einerseits ein Fehlschlagen im Hinblick darauf bedeuten, dass die Übergewichtigen sich nicht mehr bewegen als zuvor und sich daher keine Änderung ergibt. Andererseits können aber Muskeln auf- und Fett abgebaut werden, so dass sich dadurch im Durchschnitt keine Gewichtsveränderung ergeben hat. Eine erneute Durchführung könnte hier eine Subgruppenanalyse anschließen, eine exaktere Messmethode der Adipositas (Trizepsfalten etc.) einsetzen oder eine qualitative Befragung integrieren, um herauszufinden, welche der beiden Ursachen vorliegt und ob damit tendenziell eher ein Erfolg oder ein Misserfolg der Intervention in Bezug auf Ge-

wichtsveränderung vorliegt. Im Einzelnen muss also jede Intervention, deren Ergebnis als widersprüchlich darzustellen ist, nach Gewichtung von Parametern oder nach Zielgruppen von den Experten im Blick auf die Geldgeber und die Übertragung beurteilt werden. Unser Instrument legt hier also ein erstes Auswählen von Studien vor. Daran anschließend müssen Interventionen von hoher Qualität aufgegriffen werden und nach ihrem Effekt differenziert interpretiert werden.

- 3) Des Weiteren wurden alle Interventionen nur auf Grund der Publikationen der Evaluationsstudie bewertet. Möglicherweise haben die Studiendurchführenden in den Publikationen Informationen nicht gegeben, die zu einer schwächeren Bewertung anhand der Qualitätskriterien geführt haben, obwohl sie bei der Durchführung beachtet worden sind. Diese Probleme sind im Übrigen auch aus der „klassischen“ EbM bekannt. Die Vorenthaltung von Information führt in unserem Katalog in einigen Bewertungen zum Abzug in der Qualität der Studiendurchführung. Im späteren praktischen Umgang gäbe es verschiedene Vorgehensweisen:
  - a. Die Studie wird schlechter beurteilt und fällt ggf. aus dem systematischen Review oder aus der Metaanalyse.
  - b. Die Studienleiter werden kontaktiert, Informationen werden nachgefordert. Die Studie kann dann auf Grundlage der zusätzlichen Information in das systematische Review eingezogen werden.

Wir haben uns gegen das zweite Vorgehen entschieden. Die Zeit des Forschungsprojekts war zu kurz, um umfangreiche Korrespondenz zu vielen Studien zu führen. Zu einigen Studien, die herausragenden Charakter vermuten ließen, wurden jedoch über die Studienleiter nochmals Informationen angefordert. Drei verschiedene Szenarien schlossen sich an:

- c. Die Studienleiter antworteten selbst auf wiederholte Kontaktierung nicht.
- d. Die Studienleiter antworteten, aber konnten die Information nicht geben.
- e. Die Studienleiter gaben die Information.

Selbst wenn man alle Studienleiter anschreiben würde, um Informationen zu den Studien zu erhalten, die in den Bereich fallen, zu dem Aussagen nicht gemacht werden können, ergäbe sich eine Verzerrung ergeben, wenn nur einige antworten würden. Wir haben uns

daher dafür entschieden, die Studien nur auf Grundlage des schriftlichen Befundes zu bewerten.

Generell gehen wir davon aus, dass Interventionen mit hoher Qualität auch eine Darstellung in der Literatur gefunden haben. Eine Literatursuche wird somit immer die besten Interventionen finden. Es handelt sich um einen positiven Bias (Verzerrung). Ebenso verhält es sich bei den Effekten der Intervention. Solche mit positiv nachgewiesenen Effekten werden eher in Studien in der Literatur dargestellt als solche mit negativem oder fehlendem Effekt. Auch dies ist eher positiv zu sehen, da Interventionen mit Effekt stärker verbreitet werden sollen. Kritisch gesehen werden muss natürlich, dass bei einem Bias zu Lasten nicht wirksamer Studien diese eventuell nochmals in anderen Regionen ohne Kenntnis der fehlenden Effekte wiederholt werden. Dieser Bias ist vornehmlich ein Problem der Journale, welcher eher erfolgreiche, als nicht erfolgreiche Interventionen publizieren.

## **10.4 Verhaltensprävention versus Verhältnisprävention**

Die meisten der Studien unserer Recherche sind dem Bereich Verhaltensprävention zuzuordnen, nur wenige fallen in die Verhältnisprävention. Weiterhin fallen nur wenig Studien in den Bereich Gesundheitsförderung. Dies kann zum einen unserer Suchstrategie geschuldet sein. Jedoch haben wir auch Datenbanken einbezogen, die nicht nur aus klassisch medizinisch dominierten Journalen gespeist werden und somit auch für soziale Programme etc. offen sind. Möglicherweise sind aber Interventionen bzw. Studien, die nicht als Verhaltensprävention ausgelegt sind, zunächst gar nicht an der Erhebung der von uns gewählten Outcomeparameter, sprich Verhalten und/ oder Gewicht/ Ausdauer etc., interessiert, sondern gehen zunächst nur auf die Ebene des Wissens oder der Einstellungen.

Auch die Schlagwörter wurden zwar auf Grundlage von Risikofaktoren gebildet, aber lassen in ihrer Undeterminiertheit, z. B. Bewegung, auch die Aufnahme nicht mit Risikofaktoren verschlagworteter Studien zu, denn es wurde ja nicht nach koronare Herzkrankheit oder Diabetes gesucht.

Das Phänomen einer Engführung auf Interventionen der Verhaltensprävention kann aber auch den gegenwärtigen Stand der Dinge abbilden. Interventionen der Gesundheitsförderung sind

schwieriger zu evaluieren und damit auch zu publizieren. Allerdings wäre unser Instrument auch auf Interventionen der Gesundheitsförderung im weiteren Sinne anwendbar.

## **10.5 Empfehlung auf Basis der Evaluation**

Im Folgenden werden die Ergebnisse der Evaluation ausgewählter Studien in den vier Handlungsfeldern zusammenfassend in Form von Empfehlungen dargestellt:

### **10.5.1 Bewegung**

Im Handlungsfeld Frauen und Mädchen und Bewegung kann man den Trend erkennen, dass die zwei effektiven bzw. teilweise effektiven Studien, die in den Bereich stark empfehlenswert kamen, multistrategisch ausgerichtet waren, eine vorherige Befragung der Zielgruppe einschlossen und möglichst alle Beteiligten wie Schule, Gemeinde etc. integrierten (Wen et al. 2002, Pate et al. 2005). Unter den Studien im Bereich eingeschränkter Empfehlung zeigte eine Studie mit einem sehr engen verhaltenspräventiven Ansatz keinen Effekt (Robbins et al. 2006). Zwei weitere Studien mit einem Effekt teils-teils haben ebenfalls einen sehr engen Ansatz und sind von der Anlage her eher dem Feld efficacy (Schofield et al. 2005) bzw. einem Pilotcharakter (Williams et al. 2005) zuzuordnen.

Für den Bereich Betrieb zeigt sich, dass alle Aktionen, die mit Plakaten etc. das Treppensteigen isoliert steigern wollen, nicht mehr durchgeführt werden sollten, da sie scheitern (Adams and White 2002, Marshall et al. 2002). Nur zwei der acht Studien zum Treppensteigen im betrieblichen Setting, die von uns evaluiert wurden, erreichen unserer Ansicht nach das nötige Qualitätsniveau. Die einzige Studie, die nach Qualität in den Bereich empfehlenswert kam, und einen positiven Effekt hat, entspricht zwar eher einem von oben verordneten Programm (Harrell et al. 1996). Sie weist damit aber in der Tendenz auf, dass Bewegung im Betrieb, hier zwar das spezielle Klientel von Polizeiangehörigen, „einfach“ durchgeführt werden muss, so dass man im Betrieb alle Mitarbeiter erreicht. Weiterhin zeigen die weiteren gescheiterten Studien Faktoren, die man als Misserfolgskriterien beurteilen könnte: a) Arbeitsbefreiung wird nicht gewährt (Gerdle et al. 1992). b) Die Intervention ist isoliert auf Beratung oder Fitnesstestung als Anreiz ausgelegt (Aittasalo et al. 2004).

## 10.5.2 Depression

Die zur Prävention der Depression analysierten Studien konzentrieren sich auf wenige Programme. Am häufigsten wurden die Programme Leichtigkeit im sozialen Alltag (LISA), das Programm FRIENDS sowie Resourceful Adolescent Programm (RAP) evaluiert, die auch alle unter den Studien mit überdurchschnittlichen Qualitätsscores vertreten sind. Für alle drei Programme ist mindestens eine der Studien mit überdurchschnittlichen Qualitätscore teilweise effektiv. Besonders hervorzuheben sind die beiden mit überdurchschnittlichen Qualitätsscores durchgeführten Studien zu dem Programm RAP, die beide als teilweise effektiv bewertet sind. Weitere Studien für dieses Programm mit unterdurchschnittlicher Studienqualität liegen nicht vor. (Merry et al. 2004; Shochet et al. 2001) Die Studie von Merry et al. 2004 zu dem Programm RAP erreicht als einzige Studie den Qualitätslevel stark empfehlenswert.

Für das Programm LISA (Pössel et al. 2004; Groen et al. 2003; Pössel et al. 2006a) und für das Programm FRIENDS (Lowry-Webster et al. 2003) liegen auch Studien mit unterdurchschnittlicher Studienqualität vor. Die Studien für das Programm FRIENDS sind aber alle zumindest teilweise effektiv und weisen damit einheitliche Tendenzen auf. Für das Programm LISA liegen dagegen sowohl ineffektive als auch teilweise effektive Studien vor.

Angesichts von Anhaltspunkten, dass die Ausgestaltung der Interventionen über die in den Qualitätsscores erfassten Aspekten hinaus großen Einfluss auf die Ergebnisse haben (Pössel et al. 2006 b), ist im Bereich der Prävention von Depressionen und Angststörungen noch weiterer Forschungsbedarf gegeben. Trotz der positiven Tendenzen der kognitiv-verhaltenstherapeutischen Ansätze sind gegenwärtig keine klaren Empfehlungen möglich. Sie scheinen jedoch dem rein edukativen Ansatz überlegen, für den die zwei analysierten Studien – weder die mit unterdurchschnittlichen noch die mit überdurchschnittlichen Qualitätsscores – keine Wirksamkeit zeigen konnten. (Clark et al. 1993) Eine flächendeckende Umsetzung kognitiv verhaltenstherapeutischer Interventionen zur universalen Prävention der Depression erscheint gegenwärtig aufgrund der offenen Fragen zur optimalen Ausgestaltung verfrüht.

## 10.5.3 Ernährung

Für die 24 evaluierten Studien über alle Qualitätsstufen ergab sich, dass bei den acht Interventionen, die kombiniert Ernährung und Bewegung in den Blick nahmen, nur der Ernährungsbereich

einen positiven Effekt hatte, der Bewegungsbereich einen negativen. Alle drei Interventionen, die in den Bereich empfehlenswert fielen, hatten entweder einen Ansatz, der sowohl Ernährung und Bewegung einbezog (Gortmaker et al. 1999a) oder einen Ansatz, in dem die Eltern ganz eng eingebunden waren (Reynolds et al. 2000, Townsend et al. 2006).

Es hat sich herausgestellt, dass der Erfolg einer Intervention gefördert wird, wenn sie über die Schule hinaus weitergeführt wird, also z.B. im Elternhaus oder in der Freizeit (u.a. Belansky et al. 2006). Des Weiteren konnten positive Effekte festgestellt werden, wenn das Menu in den Cafeterias der Schulen auf fettreduzierte Produkte umgestellt wurde (z.B. Story et al. 2000, Lowe et al. 2004).

#### **10.5.4 Rauchen**

Zusammenfassend ist festzuhalten, dass auch die internationale Studienlage keine eindeutigen Empfehlungen erlaubt. In der Tendenz sind effektive Interventionen mit hoher Qualität in der internationalen Studienlage multimodal angelegt – auch wenn man sagen muss, dass wiederum multimodale Programme auch gescheitert sind – und wurden durch Hebammen, Ärzte oder Pflegepersonal während der routinemäßigen Schwangerschaftsuntersuchungen durchgeführt.

### **10.6 Übertragung der Ergebnisse auf Deutschland**

Was kann für eine zukünftige Priorisierung aus unseren Ergebnissen abgeleitet werden? Lassen sich aus den anhand des Kriterienkatalogs gescreenten und als von hoher Qualität bewerteter Interventionen weitere Schlüsse ziehen auf Gemeinsamkeiten dieser Interventionen, die als Erfolgsgaranten gelten können?

- 1) Insgesamt muss festgehalten werden, dass es international wenige Studien gibt, die sowohl hohe Studienqualität als auch hohe Interventionsqualität aufweisen. Zwar haben wir nur Ausschnitte aus vier Handlungsfeldern betrachtet, aber in keinem Feld konnten Studien über 80% in beiden Qualitätsdimensionen erreichen. Dies zeigt, dass selbst in den dreizehn im Bereich Prävention führenden Nationen vermehrt Anstrengungen zu hoher Qualität in der Studiendurchführung als auch in der Interventionsdurchführung zu fordern sind.

- 2) Weiterhin muss angesichts der kurzen Studiendauern, dem Ausbleiben von Nachbeobachtung und der geringen Zahl von eingeschlossenen Personen gefordert werden, dass Fragestellungen in größeren Studien, beispielsweise auf der Ebene der Europäischen Union, bearbeitet werden. Damit wäre auch zugleich die Frage der Übertragbarkeit als inhärentes Merkmal für Studien integriert.
- 3) Es ist oft festzustellen, dass viele kleine und wenig koordinierte Projekte und Studien mit einer Vielzahl von verschiedenen Erhebungsinstrumenten und Endpunkten nicht zielführend sind. Es ist nicht immer nachvollziehbar, warum welche Endpunkte gewählt wurden und wann man sich für die Endpunkte entschieden hat. In diesem Sinne muss es eine Konzentration geben, damit in wenigeren, aber größer angelegten Studien klare Fragestellungen beantwortet werden. Dies birgt die Gefahr, dass möglicherweise viele kleinere Projektinitiatoren zunächst enttäuscht werden, aber die Einsicht, dass weniger Studien mit hohem Anspruch mehr Ergebnis bedeutet, muss sich durchsetzen.
- 4) Es zeigte sich, dass weder die Qualität der Intervention noch die Qualität der Studie positiv oder negativ korrelieren. Dieses Ergebnis beruht zwar nur auf den von uns ausgewählten Kriterien, die aber die international derzeit anerkannte Summe abbilden. Auch wurden von uns nur fünf Interventionsbereiche aus den vier Handlungsfeldern Bewegung, Depression, Ernährung und Rauchen untersucht. Allerdings war für den Bereich Depression sogar festzustellen, dass eine in der Tendenz höhere Qualität in der Studien- und Interventionsdurchführung eher zu einem negativen Erfolg führt. Dies ist zunächst erkenntnistheoretisch nicht erstaunlich, auch wenn oft vermutet wird, gute Qualität in der Studiendurchführung und Gestaltung der Intervention würden einen positiven Erfolg ermöglichen bzw. sichern. Hier gilt es gerade anzumerken, wie die EbM zu berichten weiß, dass schlechte Studienqualität eher zu positiven Effekten führt. Es kann aus unserer Sicht nicht argumentiert werden, dass es zunächst reicht, sich auf die Einhaltung von Qualitätskriterien zu verlassen. Im Gegenteil, auch für Interventionen der Gesundheitsförderung bzw. verhältnispräventive Interventionen kann es nicht reichen, nur Qualität einzuhalten. Es müssen auch Effekte gemessen werden.
- 5) Auf Grundlage der geringen Zahl an empfehlenswerten Studien können keine Erfolgsfaktoren abgeleitet werden. Dieses Ergebnis war für uns zunächst enttäuschend, da eine Empfehlung über das Vorgehen bei Interventionen hohen praktischen Wert gehabt hätte.

Zumindest konnten wir jedoch ein Kondensat derjenigen Interventionen aufstellen, welche sowohl hohe Qualität der Intervention und eine ausreichende Evaluation über Studien aufweisen. Gesetzmäßigkeiten ergeben sich jedoch nicht. Dies führt im Umkehrschluss jedoch nicht nur zu einer fehlenden Empfehlung für gutes Design, sondern auch zu fehlendem Ausschluss schlechten Designs. Mit anderen Worten ist ein Qualitätswettbewerb notwendig, um nach einer genügenden Zahl an Interventionen und begleitenden Studien nochmals nach ausschlaggebenden Kriterien zu suchen.

- 6) Ob Interventionen, die entweder in der Interventions- oder Studienqualität „durchgefallen“ sind, ggf. doch erneut durchgeführt werden können, muss eine Einzelanalyse zeigen, die im Rahmen dieses Forschungsprojekts nicht vorgelegt werden konnte. In diese Einzelanalyse müsste neben einer genauen Beurteilung der Probleme, z. B. schlechte oder mangelnde Qualifikation der Multiplikatoren – wobei dies nur selten erhoben, zugegeben und dann auch publiziert wird – auch die Frage gestellt werden, ob die Studie in einem anderen Setting, für eine andere Zielgruppe nicht doch erfolgversprechend sein könnte.
- 7) Insbesondere für die Evaluation von sozial übergreifenden Interventionen sollten die Ansätze aus Skandinavien stärker in Deutschland Berücksichtigung finden. Dies sollte wie auch die Pflicht zur Einordnung in die bisherige Forschung und die konzertierte Durchführung von Studien zu neuen Fragen in der Intervention und Prävention sollte der Gesetzgeber aufnehmen.

## **10.7 Ertrag für die Priorisierung von Interventionen**

Insgesamt kann unser Kriterienkatalog als ein Screeninginstrument angesehen werden, das sich auf alle Bereiche von Prävention und Gesundheitsförderung anwenden lässt. Die Kriterien zur Qualität der Studiendurchführung wie der Intervention sind adaptierbar und veränderbar innerhalb des Katalogs von 58 Fragen. Das Instrument ist somit anwendbar im Sinne eines Health Technology Assessment und darüber hinaus hilfreich im Rahmen von Priorisierungsprozessen, wenn Ressourcen für Interventionen verteilt werden müssen. Des Weiteren sollte an Hand dieser Screeninginstruments und der dargestellten Erfahrungen aus Interventionen folgendermaßen vorgegangen werden:

- Alle Anträge auf Förderung müssen einen Nachweis haben, dass die Literatur zu dieser Intervention durchsucht und bewertet worden ist.

- Die Bewertung der Studien zeigt dann, ob die Intervention vielversprechend ist, bzw. ob sie umgekehrt innovativ ist.
- Dies kann mittels unseres Screeningsinstruments mit wenigen Fragen festgestellt werden.
- Sollte eine Intervention bei bisher schlechter Studienqualität durchgeführt worden sein, müsste dann begründet werden, in wie fern die neue Durchführung positiv sein kann.

Die Einhaltung auch der formalen Qualitätskriterien wie Prozessdokumentation oder Schulungen der Multiplikatoren ist streng zu fordern. Selbst wenn dies keinen positiven Effekt garantieren kann, so wird damit die Verlässlichkeit des Ergebnisses garantiert, so dass daraus Positiv- wie Negativlisten zu fördernder bzw. nicht mehr zu fördernder Interventionen entwickelt werden können.

# 11 Literatur

## 11.1 Literatur Evaluation

Antonovsky, A. 1987: Unraveling the mystery of health. How people manage stress and stay well. In: Social and Behavioral Science Series. Jossey, Bass, Wiley, San Francisco

Beck, D. 2003: Sense of Coherence (SOC) – ein brauchbares Analysekonzept in der betrieblichen Gesundheitsförderung? Gütebeurteilung des Sense of Coherence (SOC) – Messkonzepte von Aaron Antonovsky anhand einer Querschnittsuntersuchung bei Angestellten eines Berliner Dienstleistungsunternehmens. In: Unveröffentlichte Diplomarbeit. Freie Universität, Berlin

Bloch, E. 1985: Geist der Utopie, 1918. In: Werkausgabe Bd. 16, Suhrkamp Verlag, Frankfurt am Main

Bödeker, W. 2006: Evidenzbasierung in Gesundheitsförderung und Prävention – Der Wunsch nach Legitimation und das Problem der Nachweisstrenge. In: Bödeker, W., Kreis, J. (Hg.) Evidenzbasierung in Gesundheitsförderung und Prävention. BKK Gesundheitsförderung und Selbsthilfe, Bd. 17, Bremerhaven, S. 1-12

Canadian Task Force on the Periodic Health Examination 1979: The periodic health examination. In: Canadian Medical Association Journal, Jg. 121 (1979), S. 1193-1254

Edwards, A., Elwyn, G. 2001: Evidence-based Patient Choice. Inevitable or impossible. In: Oxford University Press, Oxford, New York, Toronto

Egger, M., Jüni, P., Bartlett, C. et al. 2003: How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. In: Health Technology Assessment, Jg. 7 (2003), S. 1-76

Elkeles, T. 2006: Evaluation von Gesundheitsförderung und Evidenzbasierung? In: Böderker, W., Kreis, J. (Hrsg.) Evidenzbasierung in Gesundheitsförderung und Prävention. BKK Gesundheitsförderung und Selbsthilfe, Bd. 17, Bremerhaven, S. 111-154

Engbers, L.H., Poppel, M.N.M. van, W. van 2007a: Measuring stair use in two office buildings. A comparison between an objective and a self-reported method. *Scandinavian Journal of Medical Sciences and Sports*, Jg. 17 (2007), S. 165-71

Engbers, L.H., Poppel, M.N.M. van, Mechelen W. van 2007b: Modest effects of a controlled worksite environmental intervention on cardiovascular risk in office workers. *Preventive Medicine*, Jg. 44 (2007), S. 356-62

Gadamer, H.G. 1993: Über die Verborgenheit der Gesundheit. Suhrkamp Verlag, Frankfurt am Main

Glasziou, P., Vandenbroucke, J., Chalmers, I. 2004: Assessing the quality of research. In: *British Medical Journal*, Jg. 328 (2004), S. 39-41

Harden, A., Garcia, J., Oliver, S. et al. 2004: Applying systematic review methods to studies of people's views: an example from public health research. In: *Journal of Epidemiology and Community Health*, Jg. 58 (2004), S. 794-800

Hurrelmann, K. 2000: Gesundheitssoziologie. Eine Einführung in sozialwissenschaftliche Theorien von Krankheitsprävention und Gesundheitsförderung, 4. Auflage, Juventa, Weinheim/München

Hurrelmann, K., Klotz, T., Haisch, J. 2004: Lehrbuch Prävention und Gesundheit. Hans Huber Verlag, Bern, Göttingen, Toronto, Seattle

Institute of Medicine (2001) Health and Behavior. The Interplay of Biological, Behavioral, and Societal Influences. National Academy Press, Washington, DC.(zitiert als IoM 2001)

International Union for Health Promotion and Education (IUHPE) 1999: The Evidence of Health Promotion Effectiveness: a report for the European Commission by the IUHPE. ECSC-EC-EAEC, Brussels-Luxemburg (zitiert als IUHPE 1999)

Jackson, N., Waters, E. for the Guidelines for Systematic Reviews in Health Promotion and Public Health Taskforce 2005: Criteria for the systematic review of health promotion and public health interventions. In: Health Promotion International, Jg. 20 (2005), S.367-374

Khushf, G. 2007: An agenda for future debate on concepts of health and disease. In: Medicine, Health Care and Philosophy, Jg. 10 (2007), S.19-27

Kickbusch, I. 2003: Gesundheitsförderung. In: Schwartz, F.W., Badura, B., Busse, R. et al. (Hrsg.) Public Health. Gesundheit und Gesundheitswesen. Urban & Fischer, München, S. 181-189

Kreis, J. 2006: Anwendung der Methodik der Cochrane Reviews auf Interventionen der Gesundheitsförderung und Prävention. In: Bödeker, W., Kreis, J. (Hrsg.) Evidenzbasierung in Gesundheitsförderung und Prävention. BKK Gesundheitsförderung und Selbsthilfe, Bd. 17, Bremerhaven, S. 155-184

McQueen, D.V. 2002: The evidence debate. Evaluating Evidence for Public Health Intervention. In: Journal of Epidemiology and Community Health, Jg. 56 (2002), S. 83f

McQueen, D.V. 2001: Strengthening the evidence base for health promotion. In: Health Promotion International, Jg. 16 (2001), S. 261-268

Moyer, V.A., Elliott, E.J. 2004: Evidence-Based Pediatrics and Child Health. In: British Medical Journal Books, London, 2<sup>nd</sup> edition

Nutbeam, D. 1998: Evaluating Health Promotion. In: Health Promotion International, Jg. 13 (1998), S. 27-44

Ogilvie, D., Egan, M., Hamilton, V. et al. 2005: Systematic reviews of health effects of social interventions: 2. Best available evidence: how low should you go? In: Journal of Epidemiology and Community Health, Jg. 59 (2005), S. 886-892

Oliver, S., Harden, A., Rees, R. et al. 2005: An emerging framework for including different types of evidence in systematic reviews for public policy. In: Evaluation, Jg. 11 (2005), S. 428-446

Ottawa-Charta 1986: abgedruckt in: Paulus P. (Hrsg.) (1992) Prävention und Gesundheitsförderung. Perspektiven für die Psychosoziale Praxis, GwG-Verlag, Köln 17-22.

Proper, K., Koning, M., Beek, A.J. van der, 2003: The effectiveness of worksite physical activity programs on physical activity, physical fitness, and health. In: Clinical Journal of Sport Medicine, Jg. 13 (2003), S. 106-117

Proper, K.I., Staal, B.J., Hildebrandt, V.H. et al. 2002: Effectiveness of physical activity programs at worksite with respect to work-related outcomes. In: Scandinavian Journal of Work, Environment and Health, Jg. 28 (2002), H. 2, S. 75-84

Richardson, D.P. 2005: The scientific substantiation of health claims with particular reference to the grading of evidence. In: European Journal of Nutrition, Jg. 44 (2005), S. 319-324

Rosenberg, W., Donald, A. 1995: Evidence based medicine: An appraisal to clinical problem-solving. In: British Medical Journal, Jg. 310 (1995), S. 1122-1126

Rosenbrock, R. 2004: Evidenzbaisierung und Qualitätssicherungin der gesundheitsbezogenen Primärprävention. In: Zeitschrift für Evaluation, Jg. 1 (2004), S. 71-80

Rychetnik, L., Frommer, M. 2002: A Schema for Evaluation Evidence on Public Health Interventions; Version 4. National Public Health Partnership, Melbourne, <http://www.nphp.gov.au/publications/phpractice/schemaV4.pdf>, vormals erreichbar über: <http://www.dhs.vic.gov.au/nphp/ppi/evidence/isspaper/index.htm> (am 25.10.2005)

Rychetnik, L., Wise, M. 2004: Advocating evidence-based health promotion: reflections and a way forward. In: Health Promotion International, Jg. 19 (2004), S. 247-57

Sackett, D., Rosenberg, W., Gray, J. et al. 1996: Evidence-based medicine: What it is and what it isn't. In: British Medical Journal, Jg. 312 (1996), S. 71f

Schwartz, F.W. 2000: Erwartungen an ein Qualitätsmanagement der Prävention durch die Gesetzliche Krankenversicherung. Referat am 17.8.2000, <http://www.google.de/search?q=cache:nmUFpC7QonoJ:www.epi.mh-hannover.de/eng/staff/schwartz/docs/mmm170800.pdf> (am 25.10.2005)

Syme, L. 1986: Strategies for Health Promotion. In: Preventive Medicine, Jg. 15 (1986), S. 492-507

Tang, K.C., Ehsani, J.P., McQueen, D.V. 2003: Evidence based health promotion: recollections, reflections, and reconsiderations. In: Journal of Epidemiology and Community Health, Jg. 57 (2003), S. 841-843

Timmer, A., Antes, G. 2006: Evidenzsynthese und Evidenzdisseminierung: Aufgabe und Struktur der Cochrane Collaboration. In: Bödeker, W., Kreis, J. (Hrsg.) Evidenzbasierung in Gesundheitsförderung und Prävention. BKK Gesundheitsförderung und Selbsthilfe, Bd. 17, Bremerhaven, S. 13-32

Trojan, A. 2006a: „Das Bohren dicker Bretter“: Gesundheit fördern durch komplexe soziale Interventionen. In: Sozial- und Präventivmedizin, Jg. 51 (2006), S. 249-251

Trojan, A. 2006b: Zu Chancen und Grenzen der Evidenzbasierung sozialer komplexer Interventionen. In: Bödeker, W., Kreis, J. (Hrsg.) Evidenzbasierung in Gesundheitsförderung und Prävention. BKK Gesundheitsförderung und Selbsthilfe, Bd. 17, Bremerhaven, S. 73-110

Wallerstein, N. 2006: What is the evidence on effectiveness of empowerment to improve health? Copenhagen, WHO Regional Office for Europe (Health Evidence Network report; <http://www.euro.who.int/Document/E88086.pdf> (am 16.12.2006)

Walter, U., Schwartz, F.W. 2003: Prävention, in: Schwartz F.W., Badura B., Busse R. et al. (Hrsg.). Public Health. Gesundheit und Gesundheitswesen, 2. erweiterte Auflage, Urban & Fischer, München/Jena, S. 189-214

Wong, M.L. 2002: Evidence-based Health Promotion: Applying it in Practice. In: Annals Academy of Medicine Singapore, Jg. 31 (2002); S. 656-662

## **11.2 Literatur Bewegung**

Adams, J., White, M. 2002: A systematic approach of the development and evaluation of an intervention promoting stair use. In: Health Education Journal, Jg. 61 (2002), S. 272-286

Aittasalo, M., Miilunpalo, S., Suni, J. 2004: The effectiveness of physical activity counseling in a work-site setting. A randomized, controlled trial. In: Patient Education and Counseling, Jg. 55 (2004), S. 193-202

Aldana, S.G., Barlow, M., Smith, R. et al. 2005: The Diabetes Prevention Program: A worksite experience. In: American Association of Occupational Health Nurses Journal, Jg. 53 (2005), S. 599-505

Allender, S., Foster, C., Scarborough, P. et al. 2007: The burden of physical activity-related ill health in the UK. In: Journal of Epidemiology and Community Health, Jg. 61 (2007), S. 344-348

Altantis, E., Chow, C.M., Kriby, A. et al. 2006: Worksite intervention effects on physical health: a randomized controlled trial. In: Health Promotion International, Jg. 21 (2006), S. 191-200

American College of Sports 1998: The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. In: Medicine and Science in Sports and Exercise, Jg. 30 (1998), S. 975-91

American Heart Association 2002: Biostatistical fact sheet – Populations. Women and coronary heart disease fact sheet, <http://www.americanheart.org> (am 14.10.2006)

American Heart Association 2003: Heart disease and stroke statistics 2003 update. Dallas, TX: American Heart Association 2002, S. 4

Andersen, L.B., Schnohr, P., Schroll, M. et al. 2000: All-cause mortality associated with physical activity during leisure time, work, sports, and cycling to work. In: Archives of Internal Medicine, Jg. 160 (2000), S. 1621-1628

Anderson, L.H., Martinson, B.C., Crain, A.L. et al. 2005: Health care charges associated with physical inactivity, overweight and obesity. In: Preventing Chronic Disease, Jg. 2 (2005), A09

Badland, H.M., Schofield, G.M. 2005: Posters in a sample of professional worksites have no effect on objectively measured physical activity. In: Health Promotion Journal of Australia, Jg. 16 (2005), S. 78-81

Barnes, P.M., Schoenborn, C.A. 2003: Physical activity among adults: United States, 2000. Hyattsville, MD: National Center for Health Statistics

Blamey, A., Mutrey, N. 2004: Changing the individual to promote health-enhancing physical activity: the difficulties of producing evidence and translating it into practice. In: Journal of Sports Sciences, Jg. 22 (2004), S. 741-754

Boutelle, K.N., Jeffery, R.W., Murray, D.M. et al. 2001: Using signs, artwork, and music to promote stair use in a public building. In: American Journal of Public Health, Jg. 91 (2001), S. 2004-2006

British Heart Foundation 2003: Coronary heart disease statistics, <http://www.dphpc.ox.ac/bhfhprg/stats/2000/2002/physicalactivity.html> (am 11.10.2006)

Buchner, D., Miles, R. 2002: Seeking a contemporary understanding of factors that influence physical activity. In: American Journal of Preventive Medicine, Jg. 23 (2002), S. 3-4

Bucksch, J., Schlicht, W. 2006: Health-enhancing physical activity and the prevention of chronic diseases – An epidemiological review. In: Social Preventive Medicine, Jg. 51 (2006), H. 5, S. 281-301

Bundesregierung 2007: Gesunde Ernährung und Bewegung – Schlüssel für mehr Lebensqualität. Eckpunktepapier der Bundesregierung, ([http://www.besser-essen-mehr-bewegen.de/fileadmin/grafiken/Aktionsplan\\_Eckpunktepapier\\_0507.pdf](http://www.besser-essen-mehr-bewegen.de/fileadmin/grafiken/Aktionsplan_Eckpunktepapier_0507.pdf)) (am 03.07.2007)

Calfas, K., Taylor, C. 1994: Effects of physical activity on psychological variables in adolescents. In: Pediatric Exercise Science, Jg. 6 (1994), S. 406-423

Carels, R.A., Darby, L.A., Cacciapaglia, H.M. et al. 2004: Reducing cardiovascular risk factors in postmenopausal women through a lifestyle change intervention. In: Journal of Women's Health, Jg. 13 (2004), S. 412-426

Casperson, C.J. 1995: Prevalence of physical inactivity in the United States. NIH Consensus Development Conference: Physical Activity and Cardiovascular Health. In: Prepson and Abstracts Bethesda, MD: Office of the Director, National Institute of Health

Coleman, K.J., Gonzalez, E.C. 2001: Promoting stair use in a US-Mexican border community. In: American Journal of Public Health, Jg. 91 (2001), S. 2007-2009

Dishman, R.K., Motl, R.W., Saunders, R., Felton, G., Ward, D.S., Dowda, M., Pate, R.R. 2004: Self-efficacy partially mediates the effect of a school-based physical-activity intervention among adolescent girls. *Preventive Medicine*, Jg. 38, (2004), S. 628-36

Eden, K.B., Orleans, C.T., Mulrow, C.D. et al. 2002: Does counseling by clinicians improve physical activity? A summary of the evidence for the US Preventive Services Task Force. In: *Annals of Internal Medicine*, Jg. 137 (2002), S. 208-215

Emmons, K.M., Linnan, L.A., Shadel, W.G. et al. 1999: The Working Healthy Project: a worksite health-promotion trial targeting physical activity, diet, and smoking. In: *Journal of Occupational and Environmental Medicine*, Jg. 41 (1999), S. 545-555

Eves, F.F., Webb, O.J. 2006: Worksite interventions to increase stair climbing: reasons for caution. In: *Preventive Medicine*, Jg. 43 (2006), H. 1, S. 4-7

Ewart, C.K., Young, D.R., Hagberg, J.M. 1998: Effects of school-based aerobic exercise on blood pressure in adolescent girls at risk for hypertension. In: *American Journal of Public Health*, Jg. 88 (1998), S. 949-951

Fahrenwald, N.L., Atwood, J.R., Walker, S.N. et al. 2004: A randomized pilot test of „Moms on the Move“: a physical activity intervention for WIC mothers. In: *Annals of Behavioral Medicine*, Jg. 27 (2004), S. 82-90

Felton, G., Saunders, R.P., Ward, D.S. et al. 2005: Promoting physical activity in girls: a case study on one school's success. In: *Journal of School Health*, Jg. 75 (2005), S. 57-62

Fisher, S.P., Fisher, M.M. 1995: Development, implementation, and evaluation of a health promotion program in a college setting. In: *Journal of American College Health*, Jg. 44 (1995), S. 81-84

Foster, C., Hillsdon, M., Cavill, N. et al. 2006: Interventions that use the environment to encourage physical activity. Evidence review. In: National Institute for Health and Clinical Excellence, <http://www.nice.org.uk> (am 4.2.2007)

Foster, C., Hillsdon, M. 2004: Changing the environment to promote health-enhancing physical activity. In: *Journal of Sports Sciences*, Jg. 22 (2004), S. 755-769

Gamble, R.P., Boreham, A.G., Stevens, A.B. 1993: Effects of a 10-week exercise intervention programme on exercise and work capacities in Belfast's ambulancemen. *Occupational Medicine* Jg. 43 (1993), S. 85-89

Gerdle, B., Brulin, C., Elert, J. et al. 1992: Effect of a general fitness program on musculoskeletal symptoms, clinical status, physiological capacity, and perceived work environment among home care service personnel. *Journal of Occupational Rehabilitation* Jg. 5(1993), S. 1-16

Glasgow, R.E., Terbourg, J.R., Strycker, L.A. et al. 1997: Take Heart II: Replication of a worksite health promotion trial. *Journal of Behavioral Medicine* Jg. 20 (1997), S. 143-61

Grønningsæter, H., Hytten, K., Skauli, G. et al. 1992: Improved health and coping by physical exercise or cognitive behavioral stress management training in a work environment. *Psychology and Health* Jg. 7 (1992), S. 147-63.

Grosch, J.W., Alterman, T., Petersen, M.R. et al. 1998: Worksite health promotion programs in the U.S.: factors associated with availability and participation. In: *American Journal of Health Promotion*, Jg. 13 (1998), H. 1, S. 36-45

Gruber, J. 1986: Physical activity and self-esteem development in children: A meta-analysis. In: Stull, G., Eckert, H. (Hrsg.) *Effects of Physical Activity on Children*. Champaign, IL: Human Kinetics

Grunbaum, J.A., Kann, L., Kinchen, S.A. 2002: Youth Risk Behavior Surveillance-United States 2001. In: Morbidity and Mortality Weekly Report, Jg. 51 (2002), H. 4, S. 1-62

Harrell, J.S., Johnston, L.F., Griggs, T.R. et al. 1996: An occupation based physical activity intervention program. In: American Association of Occupational Health Nurses Journal, Jg. 44 (1996), S. 377-384

Hillsdon, M., Thorogood, M., White, I. et al. 2002: Advising people to take more exercise is ineffective: a randomized controlled trial of physical activity promotion in primary care. In: International Journal of Epidemiology, Jg. 31 (2002), S. 808-815

Hu, F.B., Sigal, R.J., Rich-Edwards, J.W. et al. 1999: Walking compared with vigorous physical activity and risk of type 2 diabetes in women: a prospectiv study. In: Journal of the American Medical Association, Jg. 282 (1999), S. 1433-1439

Johnson, J.M., Ballin, S.D. 1996: Surgeon General's report on physical activity and health is hailed as historic step toward a healthier nation. In: Circulation, Jg. 94 (1996), H. 9, S. 2045

Jones, D.A., Ainsworth, B.E., Croft, J.B. et al. 1998: Moderate leisure-time physical activity: Who is meeting the public health recommendations? A national cross-sectional study. In: Archives of Family Medicine, Jg. 7 (1998), S. 285-289

Kahn, E.B., Ramsey, L.T., Brownson, R.C. 2002: The effectiveness of interventions to increase physical activity – a systematic review. In: American Journal of Preventive Medicine, Jg. 22 (2002), S. 73-108

Kerr, J., Evelt, F., Carroll, D. 2001: Can posters prompt stair use in a worksite environment? In: Journal of Occupational Health, Jg. 43 (2001), S. 205-7

Kerr, N.A., Yore, M.M., Ham, S.A. et al. 2004: Increasing stair use in a worksite through environmental changes. In: American Journal of Health Promotion, Jg. 18 (2004), S. 312-15

Kimm, S.Y.S., Glynn, N.W., Kriska, A.M. et al. 2002: Decline in physical activity in black girls and white girls during adolescence. In: New England Journal of Medicine, Jg. 347 (2002), S. 709-715

Klocke, A. 2006: Gesundheitszustand und –verhalten von Kindern und Jugendlichen im internationalen Vergleich. In: Gerber, A., Lauterbach, K.W. (Hrsg.): Gesundheitsökonomie und Pädiatrie. Stuttgart, S. 56-63

Knowler, W.C., Barrett-Connor, E., Folwer, S.E. et al. Diabetes Prevention Program Research Group 2002: Reduction in the incidence of Type 2 diabetes with lifestyle intervention or metformin. In: New England Journal of Medicine, Jg. 346 (2002), H. 6, S. 393-403

Kohl, H.W. 3rd, Powell, K.E., Gordon, N.F., Blair, S.N., Pfaffenbarger, R.S. Jr. 1992: Physical activity, physical fitness, and sudden cardiac death. Epidemiologic Review Jg. 14 (1992), S. 37-58

Kuller, L.H., Simkin-Silverman, L.R., Wing, R.R. et al. 2001: Women's Health Lifestyle Project: a randomized clinical trial. In: Circulation, Jg. 103 (2001), S. 32-37

Kurth, B.M., Schaffrath Rosario, A. 2007: Die Verbreitung von Übergewicht und Adipositas bei Kindern und Jugendlichen in Deutschland. Ergebnisse des bundesweiten Kinder- und Jugendgesundheits surveys (KiGGS). In: Bundesgesundheitsblatt, Jg. 50 (2007), S. 736-43

Lampert, T., Starker, A., Mensink, G.B.M. 2006: Sport und Bewegung. Folienpräsentation KiGGS-Symposium, Berlin 25.09.2006, Robert Koch-Institut, [http://www.kiggs.de/experten/downloads/dokumente/ppt\\_SportBewegung.pdf](http://www.kiggs.de/experten/downloads/dokumente/ppt_SportBewegung.pdf) (am 03.03.2007)

Lawlor, D.A., Hanratty, B. 2002: The effects of physical activity advice given in routine primary consultations: a systematic review. In: Journal of Public Health Medicine, Jg. 23 (2002), S. 219-226

- Lee, C., White, S.W. 1997: Controlled trial of a minimal-intervention exercise program for middle-aged working women. In: *Psychology and Health*, Jg. 12 (1997), S. 361-74
- Malina, R.M. 1996: Tracking of physical activity and physical fitness in children and adolescents. In: *Critical Reviews in Food Science and Nutrition*, Jg. 33 (1996), S. 403-408
- Manson, J.E., Lee, I.M. 1996: Exercise for women: How much pain for optimal gain? In: *New England Journal of Medicine*, Jg. 334 (1996), S.1325-1326
- Marshall, A.L., Bauman, A.E., Patch, C. et al. 2002: Can motivational signs prompt increases in incidental physical activity in an Australian health-care facility? In: *Health Education Research*, Jg. 17 (2002), S. 743-749
- Martin, B.W., Jimmy, G., Marti, B. 2001: Bewegungsförderung bei Inaktiven: Eine Herausforderung auch in der Schweiz. In: *Therapeutische Umschau*, Jg. 58 (2001), S.196-201
- Masse, C., Anderson, C. 2003: Ethnic differences among correlates of physical activity in women. In: *American Journal of Health Promotion*, Jg. 17 (2003), H. 6, S. 357-360
- Mensink, G. 2003: Bundes-Gesundheitssurvey: Körperliche Aktivität. Beiträge zur Gesundheitsberichterstattung des Bundes. Berlin: Robert-Koch Institut
- Mutrie, N., Carney, C., Blamey, A. et al. 2002: Walk in to work out: a randomized controlled trial of a self help intervention to promote active commuting. In: *Journal of Epidemiology and Community Health*, Jg. 56 (2002), S. 407-412
- Mutrie, N., Parfitt, G. 1998: Physical activity and its link with mental, social and moral health in young people. In: Biddle, S., Sallis, J., Cavill, N. (Hrsg.): *Young and Active? Young People and Health-enhancing Physical Activity: Evidence and Implications*. In: London: Health Education Authority, S. 49-68

Oguma, Y., Shinoda-Tagawa, T. 2004: Physical activity decreases cardiovascular risk in women: review and meta-analysis. In: American Journal of Preventive Medicine, Jg. 26 (2004), S. 407-418

Pate, R.R., Pratt, M., Blair, S.N. 1995: Physical activity and public health: A recommendation from the Center for Disease Control and Prevention and the American College of Sports Medicine. In: Journal of the American Medical Association, Jg. 273 (1995), S. 402-407

Pate, R.R., Ward, D.S., Saunders, R.P. 2005: Promotion of physical activity among high-school girls: a randomized controlled trial. In: American Journal of Public Health, Jg. 95 (2005), S. 1582-1587

Pegus, C., Bazzarre, T.L., Brown, J.S. et al. 2002: Effect of the Heart at work program on awareness of risk factors, self-efficacy, and health behaviors. In: Journal of Occupational and Environmental Medicine, Jg. 44 (2002), S. 228-36

Pohjonen, T., Ranta, R. 2001: Effects of worksite physical exercise intervention on physical fitness, perceived health status, and work ability among home care workers: five-year follow-up. In: Preventive Medicine, Jg. 32 (2001), S. 465-475

Pritchard, J.E., Nowson, C.A., Wark, J.D. 1997: A worksite program for overweight middle-aged men achieves lesser weight loss with exercise than with dietary change. In: Journal of the American Dietary Association, Jg. 97 (1997), S. 37-42

Proper, K., Koning, M., Beek, A.J. et al. 2003: The effectiveness of worksite physical activity programs on physical activity, physical fitness, and health. In: Clinical Journal of Sports Medicine, Jg. 13 (2003), S. 106-117

Proper, K.I., Staal, B.J., Hildebrandt, V.H. et al. 2002: Effectiveness of physical activity programs at worksite with respect to work-related outcomes. In: Scandinavian Journal of Work, Environment and Health, Jg. 28 (2002), H. 2, S. 75-84

- Ransdell, L.B., Robertson, L.A., Ornes, L. et al. 2004: Generations exercising together to improve fitness (GET FIT): a pilot study designed to increase physical activity and improve health related fitness in three generations of women. In: *Women and Health*, Jg. 40 (2004), S. 77-94
- Riddoch, C. 1998: Relationships between physical activity and health among young people. In: Biddell, S., Sallis, S., Cavill, N. (Hrsg.): *Young and Active? Young People and Health-enhancing Physical Activity: Evidence and Implications*. In: London: Health Education Authority, S. 17-49
- Robbins, L.B., Gretebeck, K.A., Kazanis, A.S. et al. 2006: Girls on the Move program of increased participation. In: *Nurs Research*, Jg. 55 (2006), S. 206-216
- Russell, W.D., Dzewaltowski, D.A., Ryan, G.J. 1999: The effectiveness of a point-of-decision prompt in deterring sedentary behavior. In: *American Journal of Health Promotion*, Jg. 13 (1999), S. 257-259
- Rutten, A. 2006: Beyond sport and physical activity: the concept of movement in health promotion. In: *Sozial- und Präventivmedizin*, Jg. 51 (2006), S. 252-253
- Rutten, A., Abu-Omar, K. 2004: Prevalence of physical activity in the European Union. In: *Sozial und Präventivmedizin*, Jg. 49 (2004), H. 4, S. 481-490
- Rutten, A., Ziemainz, H., Schena, F. et al. 2003: Using different physical activity measurements in eight European countries. Results of the European Physical Activity Surveillance System (EU-PASS) time series survey. In: *Public Health Nutrition*, Jg. 6 (2003), H. 4, S. 371-376
- Scharff, D.P., Homan, S., Kreuter, M. et al. 1999: Factors associated with physical activity in women across the life span: Implications for program development. In: *Women and Health*, Jg. 29 (1999), S. 115-134

Schofield, L., Mummery, K., Schofield, G. 2005: Effects of a controlled pedometer-intervention trial for low-active adolescent girls. In: *Medicine and Science in Sports and Exercise*, Jg. 37 (2005), S. 1414-1420

Sorensen, G., Barbeau, E., Stoddard, A.M. et al. 2005: Promoting behavior change among working-class multiethnic workers: results of the healthy directions- small business study. In: *American Journal of Public Health*, Jg. 95 (2005), S. 1389-1395

Titze, S., Martin, B.W., Seiler, R. et al. 2001: A worksite intervention module encouraging the use of stairs: results and evaluation issues. In: *Sozial- und Praeventivmedizin*, Jg. 46 (2001), S. 13-19

Tuomilehto, J., Lindstrom, J., Eriksson, J.G. et al. Finnish Diabetes Prevention Study Group 2001: Prevention of Type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. In: *New England Journal of Medicine*, Jg. 344 (2001), H. 18, S. 1343-1350

US Department of Health and Human Services 1996: *Physical Activity and Health: A Report of the Surgeon General*. Atlanta, Ga: US Department of Health and Human Services, Centers for Disease Control and Prevention, and National Center for Chronic Disease Prevention and Health Promotion

US Department of Health and Human Services 2000: *Healthy people 2010: Understanding and improving health*. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion

Vuori, I.M., Oja, P., Paronen, O. 1994: Physically active commuting to work – testing its potential for exercise promotion. In: *Medicine and Science in Sports and Exercise*, Jg. 26 (1994), S. 844-850

Wen, L.M., Thomas, M., Jones, H. et al. 2002: Promoting physical activity in women: evaluation of a 2-year community based intervention. In: *Health Promotion International*, Jg. 17 (2002), S. 127-137

Williams, B.R., Bezner, J.U., Chesbro, S.B. et al. 2005: The effect of a behavioral contract on adherence to a walking program. In: *Topics in Geriatric Rehabilitation*, Jg. 21 (2005), S. 332-342

World Health Organisation 2002: *European Health Report*. WHO Regional Publications, European Series no. 97. Copenhagen: WHO

Zizzi, S., Vitullo, E., Rye, J. et al. 2006: Impact of a three-week pedometer intervention on high school students' daily step counts and perceptions on physical activity. In: *American Journal of Health Education*, Jg. 37 (2006), H. 1, S. 35-40

### **11.3 Literatur Depression**

Andlin-Sobocki, P., Wittchen, H.U. 2005: Cost of affective disorders in Europe. In: *European Journal of Neurology*, Jg. 12 (2005), Suppl. 1, S. 34-38

Andrews, G., Szabo, M., Burns, J. 2002: Preventing major depression in young people. In: *British Journal of Psychiatry*, Jg. 181 (2002), S. 460-462

Barrett, P., Turner, C. 2001: Prevention of anxiety symptoms in primary school children: Preliminary results from a universal school-based trial. In: *British Journal of Clinical Psychology*, Jg. 40 (2001), S. 399-410

Barrett, P.M., Farrell, L.J., Ollendick, T.H. et al. 2006: Long-term outcomes of an Australian universal prevention trial of anxiety and depression symptoms in children and youth: an evaluation of the Friends program. In: *Journal of Clinical Child and Adolescent Psychology*, Jg. 35 (2006), H. 3, S. 403-411

Barth, J., Schumacher, M., Hermann-Lingen, C. 2004: Depression as a risk factor for mortality in patients with coronary heart disease: a meta-analysis. In: *Psychosomatic Medicine*, Jg. 66 (2004), H. 6, S. 802-813

Birmaher, B., Ryan, N.D., Williamson, D.E. et al. 1996: Childhood and adolescent depression: A review of the past 10 years. Part I. In: *Journal of the American Academy of Child and Adolescent Psychiatry*, Jg. 35 (1996), S. 1427-1439

Booth, B.M., Zhang, M., Rost, K.M. et al. 1997: Measuring outcomes and costs for major depression. In: *Psychopharmacology Bulletin*, Jg. 33 (1997), H.4, S. 653-658

Bramesfeld, A., Schwartz, F.W. 2007: Volkskrankheit Depression. Nur ein Problem der Gesundheitsversorgung? In: *Gesellschaftspolitische Kommentare, Sonderausgabe 1* (2007), S. 3-8

Clarke, G.N., Hawkins, W., Murphy, M. et al. 1993: School-based primary prevention of depressive symptomatology in adolescents. Findings from two studies. In: *Journal of Adolescents Research*, Jg. 8 (1993), H. 2, S. 183-204

Clark, G.N., Hawkins, W., Murphy, M., et al. 1995: Targeted Prevention of unipolar depressive disorder in an at risk sample of high school adolescents: a randomized trial of a group cognitive intervention. In: *Journal of the American Academy of Child and Adolescent Psychiatry*, Jg. 34 (1995), S. 312-321

Costello, E., Angold, A., Burns, B.J. et al. 1996: The great smoky Mountains Study of Youth. Goals, design, methods, and the prevalence of DSM-III-R disorders. In: *Archives of General Psychiatry*, Jg. 53 (1996), S. 1129-1136

Cuijpers, P., Van Straten, A., Smit, F. 2005: Preventing the incidence of new cases of mental disorders. A Meta-analytic review. In: *Journal of Nervous and Mental Disease*, Jg. 193 (2005), H. 2, S. 119-125

Dubow, E.F., Schmidt, D., McBride J, et al. 1993: Teaching children to cope with stressful experiences: Initial implementation and evaluation of a primary prevention program. In: Journal of Clinical Child Psychology 1003, Jg. 22 (1993), H. 4, S. 428-440

Durlak, J.A., Wells, A.M. 1997: Primary prevention mental health programmes for children and adolescents: a meta-analytic review. In: American Journal of Community Psychology, Jg. 25 (1997), S. 115-152

Ferketich, A.K., Schwartzbaum, J.A., Frid, D.J. et al. 2000: Depression as an antecedent to Heart Disease among Woman and Men in the NHANES I Study. In: Archives of Internal Medicine, Jg. 160 (2000), S. 1261-1268

Friemel, S., Bernert, S., Angermeyer, M.C. et al. 2005: Die direkten Ksoten von depressiven Erkrankungen in Deutschland. Ergebnisse aus dem European Study of the Epidemiology of Mental Disorders (ESEMeD) Projekt. In: Psychiatrische Praxis, Jg. 32 (2005), S. 113-121

Greenberg, P.E., Kessler, R.C., Birnbaum, H.C. et al. 2003: The economic burden of depression in the united states: How did it change between 1990 and 2000? In: Journal of Clinical Psychiatry, Jg. 64 (2003), S. 1465-1475

Grobe, T.G., Dörning, H. 2003: Schwerpunkt: Depressive Erkrankungen. Gesundheitsreport der Techniker Krankenkasse mit Daten und Fakten zur Arbeitsunfähigkeit und Arzneimittelverordnung 2002. In: Techniker Krankenkasse 2003, S. 52-88

Groen, G., Pössel, P., Al-Wiswasi S, et al. 2003: Universelle, schulbaiserte Prävention der Depression im Jugendalter: Ergebnisse einer Follow-up-Studie. In: Kindheit und Entwicklung, Jg. 12 (2003), H. 3, S. 164-174

GVG 2006: Bericht 6. nationales Gesundheitsziel. Depressive Erkrankungen: verhindern, früh erkennen, nachhaltig behandeln. In: BMG (Hrsg.) Gesundheitsziele.de, Berlin

Harrington, R., Clark, A. 1998: Prevention and early intervention for depression in adolescence and early adult life. In: *European Archives of Psychiatry and Clinical Neuroscience*, Jg. 248 (1998), S. 32-45

Hinz, A., Schwarz, R. 2001: Angst und Depression in der Allgemeinbevölkerung. Eine Normierungsstudie zur Hospital Anxiety and Depression Scale. In: *Psychotherapie, Psychosomatik, Medizinische Psychologie (PPmP)*, Jg. 51 (2001), S.193-200

Horowitz, J.L., Garber, J. 2006: The prevention of depressive Symptoms in children and adolescents: a meta-analytic review. In: *Journal of Consulting and Clinical Psychology*, Jg. 74 (2006), H. 3, S. 401-415

Jané-Llopis, E. 2002: What makes the ounce of prevention effective? A meta-analysis of mental health promotion and mental disorder prevention programmes. Nijmegen: Nijmegen University

Jané-Llopis, E., Barry, M., Hosman, C., Patel, V. 2005: Mental health promotion works: a review. In: *International Union for Health Promotion and Education (IUHPE)*, Supplement 2, S. 9-25

Joukaama, M., Heliövaara, M., Knekt, P. et al. 2001: Mental disorders and cause-specific mortality. In: *The British Journal of Psychiatry*, Jg. 179 (2001), S. 498-502

Kapfhammer, H.P. 2007: Depressive Störungen. Eine diagnostische und therapeutische Herausforderung auch in der Primärversorgung. In: *Internist*, Jg. 48 (2007), S. 173-188

Kenardy, J., Mc Caffery, K., Rosa, V. 2003: Internet-delivered indicated prevention for anxiety disorders: a randomized controlled trial. In: *Behavioural and Cognitive Psychotherapy*, Jg. 31 (2003), S. 279-289

Kessler, R.C., McGonagle, K.A., Zhao, S. et al. 1994: Lifetime and 12-month prevalence of DSM-III-psychiatric disorders in the United States: results from the National Comorbidity Survey. In: Archives of General Psychiatry, Jg. 51 (1994), S. 8-19

Lock, S., Barrett, P.M. 2003: A longitudinal Study of developmental differences in universal preventive intervention for child anxiety. In: Behaviour Change, Jg. 20 (2003), H. 4, S. 183-199

Lowry-Webster, H.M., Barrett, M., Lock, S. 2003: A universal prevention trial of anxiety symptomology during childhood: Results at 1-year follow-up. In: Behavior Change, Jg. 20 (2003), H. 1, S. 25-43

Manz, R., Junge, J., Margraf, J. 2001: Prävention von Angst und Depression bei Jugendlichen. Ergebnisse einer Follow-up Untersuchung nach 6 Monaten. In: Zeitschrift für Gesundheitspsychologie, Jg. 9 (2001), H. 4, S. 168-179

Merry, S., McDowell, H., Hetrick, S. et al. 2003: Psychological and /or educational interventions for the prevention of depression in children and adolescents. In: Oxford: Cochrane Library

Merry, S., McDowell, H., Wild, C.J. et al. 2004: A randomized placebo-controlled trial of a school-based depression prevention program. In: Journal of the American Academy of Child and Adolescent Psychiatry, Jg. 43 (2004), S. 538-547

Möller, H.J. 2003: Suicide, Suicidality and suicide prevention in affective disorders. In: Acta Psychiatrica Scandinavica, Jg. 108 (2003), S. 73-80

Murray, C.J.L., Lopez, A.D. 1997: Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. In: The Lancet, Jg. 349 (1997), S. 1498-1504

Oria, J., Cureton, V.Y., Canham, D. 2001: Evaluation of the effectiveness of a youth leadership class in the prevention of depression in adolescents. In: The Journal of school Nursing, Jg. 17 (2001), H. 4, S. 204-209

Pössel, P., Baldus, C., Horn, A.B. et al. 2005: Influence of general self-efficacy on the effects of a school-based universal primary prevention program of depressive symptoms in adolescents: a randomized and controlled follow-up study. In: *Journal of Child Psychology and Psychiatry*, Jg. 46 (2005), H. 9, S. 982-994

Pössel, P., Horn, A.B., Groen et al. 2004: School-based prevention of depressive symptoms in adolescents: A 6 month follow up. In: *Journal of the American Academy of Child and Adolescent Psychiatry*, Jg. 43 (2004), S. 1003-1010

Pössel, P., Horn, A., Hautzinger, M. 2003: Erste Ergebnisse eines Programms zur schulbasierten Prävention von depressiven Symptomen bei Jugendlichen. In: *Zeitschrift für Gesundheitspsychologie*, Jg. 11 (2003), H. 1, S. 10-20

Pössel, P., Horn, A., Hautzinger, M. 2006a: Vergleich zweier schulbasierter Programme zur Prävention depressiver Symptome bei Jugendlichen. In: *Zeitschrift für Klinische Psychologie und Psychotherapie*, Jg. 35 (2006), H. 2, S. 109-116

Pössel, P., Schneider, S., Seemann, S. 2006b: Effekte und Kosten universaler Prävention von Internalisierungsstörungen bei Kindern und Jugendlichen. In: *Verhaltenstherapie*, Jg. 16 (2006), S. 201-210

Salize, H.J., Stamm, K., Schubert, M. et al. 2004: Behandlungskosten von Patienten mit Depressionsdiagnosen in haus- und fachärztlicher Versorgung. In: *Psychiatrische Praxis*, Jg. 31 (2004), S. 147-156

Sheffield, J.K., Spence, S.H., Papee, R.M. et al. 2006: Evaluation of universal, indicated, and combined cognitive-behavioral approaches to the prevention of depression among adolescents. In: *Journal of Consulting and Clinical Psychology*, Jg. 74 (2006), H. 1, S. 66-79

Shochet, I.M., Dadds, M.R., Holland, D. et al. 2001: The efficacy of a universal school-based program to prevent adolescent depression. In: *Journal of Clinical Child Psychology*, Jg. 30 (2001), H. 3, S. 303-315

Spence, S.H., Sheffield, J.K., Donovan, C.L. 2003: Preventing adolescent depression: An evaluation of the problem solving for life program. In: *Journal of Consulting and Clinical Psychology*, Jg. 71 (2003), H. 1, S. 3-13

Statistisches Bundesamt 1998: Depressionen. Gesundheitsberichterstattung des Bundes. In: *Gesundheitsbericht für Deutschland*, Metzler-Poeschel, Stuttgart, S. 219-222

Üstun, T.B., AYUSO-Mateos, J.L., Chatterji, S. et al. 2004: Global burden of depressive disorders in the year 2000. In: *British Journal of Psychiatry*, Jg. 184 (2004), S. 386-392

Wade, T.J., Cairney, J., Pevalin, D.J. 2002: Emergence of gender differences in depression during adolescents: National panel results from three countries. In: *Journal of the American Academy of Child and Adolescent Psychiatry*, Jg. 41 (2002), S. 190-198

Weber, A., Hörmann, G., Köllner, V. 2006: Die Epidemie des 21. Jahrhunderts? Der Erhalt oder die Wiederherstellung der seelischen Gesundheit müssen als gesamtgesellschaftliche Aufgabe begriffen werden. In: *Deutsches Ärzteblatt*, Jg. 103 (2006), S. 834-841

Wittchen, H.U., Jacobi, F. 2005: Psychische Störungen in Deutschland und in der EU. Größenordnung und Belastung, <http://www.tu-dresden.de/presse/psyche/pdf> (am 03.07.2007)

Wittchen, J. 2006: Epidemiologie. In: Stoppe, G., Bramesfeld, A. und Schwartz, F.W. (Hrsg.) *Volkskrankheit Depression? Bestandaufnahme und Perspektiven*, Springer Verlag, Berlin/Heidelberg, S. 15-37

Wulsin, L.R., Vaillant, G.E., Wells, V.E. 1999: A systematic Review of the Mortality of Depression. In: *Psychosomatic Medicine*, Jg. 17 (1999), S. 61-66

## 11.4 Literatur Ernährung

Baranowski, T., Davis, M., Resnicow, K. et al. 2000: Gimme 5 Fruit, and Vegetables for Fun and Health: Outcome Evaluation. In: Health Education and Behavior, Jg. 27 (2000), S. 96-111

Belansky, E.S., Romaniello, C., Morin, C. et al. 2006: Adapting and Implementing a Long-Term Nutrition and Physical Activity Curriculum to a rural, Low-Income, Biethnic Community. In: Journal of Nutrition Education and Behavior, Jg. 38 (2006), S. 106-113

Benecke, A., Vogel, H. 2003: Übergewicht und Adipositas. In: Gesundheitsberichterstattung des Bundes 16, Robert Koch Institut (Hrsg.)

Bere, E., Veierod, M.B., Bjelland, M. et al. 2006a: Outcome and process evaluation of a Norwegian school-randomized fruit and vegetable intervention: Fruits and Vegetables Make the Marks (FVMM). In: Health Education Research, Jg. 21 (2006), H. 2, S. 258-267

Bere, E., Veierod, M.B., Bjelland, M. et al. 2006b: Free school fruit- sustained effect 1 year later. In: Health Education Research, Jg. 21 (2006), H. 2, S. 268-275

Blanchette, L., Brug, J. 2005: Determinants of fruit and vegetable consumption among 6-12-year-old children and effective intervention to increase consumption. In: Journal of Human Nutrition and Dietetics, Jg. 18 (2005), S. 431-443

Bönnhoff, N. 2005: Der Einfluss des Ernährungsverhaltens und der kritischen Nährstoffe. In: Bjarnason-Wehrens, B., Dorel, S. (Hrsg.): Übergewicht und Apipositas im Kindes- und Jugendalter, Sankt Augustin

Burchett, H. 2003: Increasing fruit and vegetable consumption among British primary schoolchildren: a review. In: Health Education, Jg. 2 (2003), S. 99-109

Coleman, K.J., Tiller, C.T., Sanchez, J. et al. 2005: Prevention of the Epidemic Increase in Child Risk of Overweight in Low-Income Schools. In: Archives of Pediatrics and Adolescent Medicine, Jg. 159 (2005), S. 217-224

Contento, I.R., Michela, J.L. 1998: Nutrition and food choice behavior among children and adolescents. In: Goreczny, A.J., Hersen, M. (Hrsg.): Handbook of Pediatric and Adolescent Health Psychology, S. 249-273

Foerster, S.B., Gregson, J., Beall, D.L. et al. 1998: The California Children's 5 a Day-Power Play! Campaign: Evaluation of a Large-Scale Social Marketing Initiative. In: Family and Community Health, Jg. 21 (1998), H. 1, S. 46-64

French, S.A., Stables, G. 2003: Environmental intervention to promote vegetable and fruit consumption among youth in school settings. In: Preventive Medicine, Jg. 37 (2003), S. 593-610

Gortmaker, S.L., Cheung, W.Y., Peterson, K.E. et al. 1999a: Impact of a School-Based Interdisciplinary Intervention on Diet and Physical Activity Among Urban Primary School Children. In: Archives of Pediatrics and Adolescent Medicine, Jg. 153 (1999), S. 975-983

Gortmaker, S.L., Peterson, K., Wiecha, J. et al. 1999b: Reducing Obesity via a School-Based Interdisciplinary Intervention among Youth. In: Archives of Pediatrics and Adolescent Medicine, Jg. 153 (1999), S. 409-418

Hauner, H. 2000: Brauchen wir mehr Ernährungsmedizin? In: Deutsche Medizinische Wochenschrift, Jg. 125 (2000), S. 255

Hoelscher, D.M., Feldman, H.A., Johnson, C.C. et al. 2004: School-based health education programs can be maintained over time: results from the CATCH Institutionalization study. In: Preventive Medicine, Jg. 38 (2004), H. 5, S.594-606

Kaitosaari, T., Rönnemaa, T., Raitakari, O. et al. 2003: Effect of 7-Year Infancy-Onset Dietary Intervention on Serum Lipoproteins and Lipoprotein Subclasses in Healthy Children in the Prospective, Randomized Special Turku Coronary Risk Factor Intervention Project for Children (STRIP) Study. In: Journal of the American Heart Association, Jg. 108 (2003), S. 672-677

Kelder, S., Hoelscher, D.M., Barroso, C.S. et al. 2005: The CATCH Kids Club: a pilot after-school study for improving elementary students' nutrition and physical activity. In: Public Health Nutrition, Jg. 8 (2005), H. 2, S. 133-140

Kohlmeier, L., Kroke, A., Pötzsch, J. et al. 1993: Ernährungsabhängige Krankheiten und ihre Kosten. In: Band 27 der Schriftenreihe des Bundesministers für Gesundheit. Nomos Verlagsgesellschaft. Baden-Baden

Krocke, A. 2003: Primärprävention chronischer Erkrankungen durch Ernährung – von der epidemiologischen Grundlagenforschung zur Formulierung von Empfehlungen. Habilitationsschrift, Berlin

Krocke, A., Boeing, H. 2003: Die Rolle der Ernährung bei der Entstehung und Prävention chronischer Erkrankungen, [http://www.sanavita-gesundheitsforum.de/index.php?option=com\\_content&task=view&id=42&Itemid=57](http://www.sanavita-gesundheitsforum.de/index.php?option=com_content&task=view&id=42&Itemid=57) (am 19.06.2007)

Kromeyer-Hauschild, K. et al. 2001: Perzentile für den Body Mass Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. In: Monatschrift Kinderheilkunde, Jg. 149 (2001), S. 807-818

Kurth, B.-M., Schaffrath, R. 2007: Die Verbreitung von Übergewicht und Adipositas bei Kindern und Jugendlichen in Deutschland. Ergebnisse des bundesweiten Kinder- und Jugendgesundheits-surveys (KiGGS) Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz 5/6, S. 736-743

Levy, E. et al. 1995: The economic costs of obesity: the French situation. In: International Journal of Obesity, Jg. 19 (1995), S. 788-792

Lowe, C.F., Horne, P.F., Tapper, K. et al. 2004: Effects of a peer modeling and rewards-based intervention to increase fruit and vegetable consumption in children. In: European Journal of Clinical Nutrition, Jg. 58 (2004), S. 510-522

Lytle, L.A., Elaine, J.S., Nichaman, M.Z. et al. 1996: Changes in Nutrient Intakes of Elementary School Children Following a school-based Intervention: Results from the Catch Study. In: Preventive Medicine, Jg. 25 (1996), S. 465-477

Märkl, I. 2004: Gesundheitsbezogene Lebensqualität bei Kindern mit Adipositas Frem- versus Selbsteinschätzung, Wie ist die Fremdeinschätzung aus Sicht der Eltern? Dissertation Hamburg

Mensing, G.B.M., Kleiser, C., Richter, A. 2007: Lebensmittelverzehr bei Kindern und Jugendlichen in Deutschland. Ergebnisse des Kinder- und Jugendgesundheits surveys (KiGGS). Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz 5/6, S. 607-623

National Audit Office 2001: Tackling Obesity in England. London, [https://www.nao.gov.uk/publications/nao\\_reports/00-01/0001220.pdf](https://www.nao.gov.uk/publications/nao_reports/00-01/0001220.pdf) (am 25.06.2007)

Nickals, T.A., Johnson, C.C., Myers, L. et al. 1998: Outcomes of a High School Program to Increase Fruit and Vegetable Consumption: Gimme 5 – A Fresh Nutrition Concept for Students. In: Journal of School Health, Jg. 68 (1998), H. 6, S. 248-253

Perry, C.L., Bishop, D.B., Taylor, G.L. et al. 2004: A Randomized School Trial of Environmental Strategies to Encourage Fruit and Vegetable Consumption among Children. In: Health Education and Behavior, Jg. 31 (2004), S. 65-76

Perry, C.L., Lytle, L.A., Feldman, H. et al. 1998: Effects of the child and adolescent trial for cardiovascular health (CATCH) on fruit and vegetable intake. In: *Journal of Nutrition Education*, Jg. 30 (1998), S. 354-360

Powers, A.R., Struempfer, B.J., Guarino, A. et al. 2005: Effects of a Nutrition Education Program on the Dietary Behavior and Nutrition Knowledge of Second-Grade and Third-Grade Students. In: *Journal of School Health*, Jg. 75 (2005), H. 4, S. 129-133

Resnicow, K., Cohn, L., Reinhardt, J. et al. 1992: A Three- Year Evaluation of the Know Your Body Program in Inner-City Schoolchildren. In: *Health Education Quarterly*, Jg. 19 (1992), H. 4, S. 463-480

Reynolds, K.D., Franklin, F.A., Binkley, D. et al. 2000: Increasing the Fruit and Vegetable Consumption of Fourth-Graders: Results from the High 5 Project. In: *Preventive Medicine*, Jg. 30 (2000), S. 309-319

Schumann, J. 1999: Psychologische Aspekte der Adipositas. In: *Ernährungs-Umschau*, Jg. 46 (1999), S. 12-14

Seidell, J.C., Deerenberg, I. 1994: Obesity in Europe – prevalence and consequences for the use of medical care. In: *PharmacoEconomics*, Jg. 5 (1994), S. 38-44

Serdula, M.K., Ivery, D., Coates, R.J. et al. 1993: Do obese children become obese adults? A review of the Literature. In: *Preventive Medicine*, Jg. 22 (1993), S. 167-177

Shepherd, J. et al. 2006: Young people and healthy eating: a systematic review of research on barriers and facilitators. In: *Health Education Research*, Jg. 21 (2006), H. 2, S. 239-257

Snethen, J.A. 2006: Effective Weight Loss for Overweight Children: A Meta-Analysis of Intervention Studies. In: *Journal of Pediatric Nursing*, Jg. 21 (2006), S. 45-56

Stewart, G.J., Steckler, A., Martin, C. J. et al. 2003: Pathways process evaluation results: a school-based prevention trial to promote healthful diet and physical activity in American Indian third, fourth, and fifth grade students. In: Preventive Medicine, Jg. 37 (2003), H. 6, S. 80-90

Story, M., Sherwood, N.E., Himes, J.H. et al. 2003: An After-School Obesity Prevention Program for African-American Girls: The Minnesota GEMS Pilot Study. In: Ethnicity and Disease, Jg. 13 (2003), S. 54-64

Story, M., Warren Mays, R., Bishop, D.B. et al. 2000: 5-a-Day Power Plus: Process Evaluation of a Multicomponent Elementary School Program to Increase Fruit and Vegetable Consumption. In: Health Education and Behavior, Jg. 27 (2000), H. 2, S. 187-200

Treviño, R.P., Zenong, Y., Hernandez, A. et al. 2004: Impact of the Bienestar School-Based Diabetes Mellitus Prevention Program on Fasting Capillary Glucose Levels. In: Archives of Pediatrics and Adolescent Medicine, Jg. 158 (2004), H. 9, S. 911-917

Townsend, M.S., Johns, M., Shilts, M.K. et al. 2006: Evaluation of a USDA Nutrition Education Program for Low-income Youth. In: Journal of Nutrition Education and Behavior, Jg. 38 (2006), S. 30-41

Von Lengerke, T., John, J. 2005: Gesundheitsökonomische Aspekte der Adipositas-Bisherige Ergebnisse der Kooperativen Gesundheitsforschung in der Region Augsburg (KORA). In: Pressemitteilung, [www.dgk.org/kongresse/ft/ft2005/presstexte/39\\_Lengerke.pdf](http://www.dgk.org/kongresse/ft/ft2005/presstexte/39_Lengerke.pdf) (am 22.06.2007)

Wabitsch, M., Kunze, D., Keller, E. et al. 2002: Adipositas bei Kindern und Jugendlichen in Deutschland. Deutliche und anhaltende Zunahme der Prävalenz- Aufruf zum Handeln. In: Fortschritte der Medizin, Jg. 120 (2002), Originalien 4, S. 99-106

Wabitsch, M. 2000: Overweight and obesity in European children: definition and diagnostic procedures risk factors and consequences for later health outcome. In: European Journal of Pediatrics, Jg. 159 (2000), Suppl 1, S. 8-13

WHO 2000: Obesity: preventing and managing the global epidemic. In: WHO Technical Report Series 894, Geneva

Wirth, A. 1998: Gesundheitsrisiken der Adipositas. In: Aktuelle Ernährungsmedizin, Jg. 23 (1998), S. 172-177

## 11.5 Literatur Rauchen

Ahlsten, G., Cnattingius, S., Lindmark, G. 1993: Cessation of smoking during pregnancy improves foetal growth and reduces infant morbidity in the neonatal period. A population-based prospective study. In: Acta Paediatrica, Jg. 82 (1993), S. 177–181

Aveyard, P., Lawrence, T., Evans, O. et al. 2005: The influence of in-pregnancy smoking cessation programmes on partner quitting and women's social support mobilization: a randomized controlled trial. In: BioMed Central Public Health, Jg. 29 (2005), H. 5 (80)

Avidano Britton, G.R., Brinthaup, J., Stehle, J.M. et al. 2006: The effectiveness of a nurse-managed perinatal smoking cessation program implemented in a rural county. In: Nicotine & Tobacco Research, Jg. 8 (2006), H. 1, S. 13-28

Bernstein, I.M., Mongeon, J.A., Badger, G.J. et al. 2005: Maternal smoking and its association with birth weight. In: Obstetrics & Gynecology, Jg. 106 (2005), S. 986–991

Bundeszentrale für gesundheitliche Aufklärung (2001, 2004, 2005) Die Drogenaffinität Jugendlicher in der Bundesrepublik Deutschland 2001, 2004. Neue Ergebnisse zur Entwicklung des Rauchverhaltens 2005. BZgA, Köln

Campion, P., Owen, L., McNeill, A. et al. 1994: Evaluation of a mass media campaign on smoking and pregnancy. In: Addiction, Jg. 89 (1994), H. 10, S. 1245-1254

De Vries, H., Bakker, M.J., Mullen, P.D. et al. 2006: The effects of smoking cessation counseling by midwives on Dutch pregnant women and their partners. In: Patient Education and Counseling, Jg. 63 (2006), S. 177-187

Deutsches Krebsforschungszentrum (Hrsg.) 2003: Passivrauchende Kinder in Deutschland – Frühe Schädigung für ein ganzes Leben, Heidelberg

Deutsches Krebsforschungszentrum (Hrsg.) 2005: Passivrauchen – ein unterschätztes Gesundheitsrisiko, Heidelberg: komplus

Doll, R., Peto, R., Wheatley, K. et al. 1994: Mortality in relation to smoking: 40 years' observations on male British doctors. In: British Medical Journal, Jg. 309 (1994), S. 901-911

Dornelas, E.A., Magnavita, J., Beazoglou, T. et al. 2006: Efficacy and cost-effectiveness of a clinic-based counseling intervention tested in an ethnically diverse sample of pregnant smokers. In: Patient Education and Counseling, Jg. 64 (2006), S. 342-349

Dybing, E., Sanner, T. 1999: Passive smoking, sudden infant death syndrome (SIDS) and childhood infections. In: Human and Experimental Toxicology, Jg. 18 (1999), S. 202–205

Emmons, K.M., Sorensen, G., Klar, N. et al. 2000: Healthy Baby Second Hand Smoke Study: project brief. In: Tobacco Control, Jg. 9 (2000), Suppl. III, S. iii58-iii60

Ershoff, D.H., Quinn, V.P., Boyd, N.R. et al. 1999: The Kaiser Permanente prenatal smoking-cessation trial. When more isn't better, what is enough? In: American Journal of Preventive Medicine, Jg. 17 (1999), H. 3, S. 161-168

Fagerstrom, K. 2002: The Epidemiology of smoking: health consequences and benefits of cessation. In: Drugs, Jg. 62 (2002), Suppl. 2, S. 1-9

Fossum, B., Arborelius, E., Bremberg, S. 2004: Evaluation of a counseling method for the prevention of child exposure to tobacco smoke: an example of client-centered communication. In: Preventive Medicine, Jg. 38 (2004), S. 295-301

Gebauer, C., Kwo, C.-Y., Haynes, E.F. et al. 1998: A nurse-managed smoking cessation intervention during pregnancy. In: Journal of Obstetric, Gynecologic and Neonatal Nursing, Jg. 27 (1998), H. 1, S. 47-53

Gielen, A.C., Windsor, R., Faden, R.R. et al. 1997: Evaluation of a smoking cessation intervention for pregnant women in an urban prenatal clinic. In: Health Education Research, Jg. 12 (1997), H. 2, S. 247-254

Glasgow, R.E., Whitlock, E.P., Eakin, E.G. et al. 2000: A brief smoking cessation intervention for women in low-income Planned Parenthood clinics. In: American Journal of Public Health, Jg. 90 (2000), H. 5, S. 786-789

Hartmann, K.E., Thorp, J.M., Pahel-Short, L. et al. 1996: Graduate education - A randomized controlled trial of smoking cessation intervention in pregnancy in an academic clinic. In: Obstetrics & Gynecology, Jg. 87 (1996), H. 4, S. 621-626

Hegaard, H.K., Kjaergaard, H., Moller, L.F. et al. 2003: Multimodal intervention raises smoking cessation rate during pregnancy. In: Acta Obstetrica et Gynecologica Scandinavica, Jg. 82 (2003), S. 813-819

Helmert, U., Lang, P., Cuelenaere, B. 1998: Rauchverhalten von Schwangeren und Müttern mit Kleinkindern. In: Sozial- und Präventivmedizin, Jg. 43 (1998), S. 51-58

Helmert, U., Borgers, D., Bammann, K. 2001: Soziale Determinanten des Rauchverhaltens in Deutschland: Ergebnisse des Mikrozensus 1995. In: Sozial- und Präventivmedizin, Jg. 46 (2001), S. 172-181

Hughes, E.G., Lamont, D.A., Beercroft, M.L. et al. 2000: Randomized trial of a "stage-of-change" oriented smoking cessation intervention in infertile and pregnant women. In: *Fertility and Sterility*, Jg. 74 (2000), H. 3, S. 498-503

Jaakkola, N., Zahlsen, K., Jaakkola, J.J.K. 2001: Effects of a population-based smoking cessation programme on smoking in pregnancy. In: *European Journal of Public Health*, Jg. 11 (2001), S. 446-449

Kendrick, J.S., Zahniser, S.C., Miller, N. et al. 1995: Integrating smoking cessation into routine public prenatal care: The smoking cessation in pregnancy project. In: *American Journal of Public Health*, Jg. 85 (1995), H. 2, S. 217-222

Kröger, C., Shaw, R. 2000: Raucherentwöhnung in Deutschland: Grundlagen und kommentierte Übersicht. (Gesundheitsförderung Konkret Bd. 2). Köln: Bundeszentrale für gesundheitliche Aufklärung

Lampert, T., Burger, M. 2004: Rauchgewohnheiten in Deutschland – Ergebnisse des telefonischen Bundes-Gesundheitssurveys 2003. In: *Gesundheitswesen*, Jg. 66 (2004), S. 511-517

Lampert, T., Burger, M. 2005: Verbreitung und Strukturen des Tabakkonsums in Deutschland. In: *Bundesgesundheitsblatt - Gesundheitsforschung – Gesundheitsschutz*, Jg. 48 (2005), S. 1231-1241

Lawrence, T., Aveyard, P., Evans, O. et al. 2003: A cluster randomized controlled trial of smoking cessation in pregnant women comparing interventions based on the transtheoretical (stages of change) model to standard care. In: *Tobacco Control*, Jg. 12 (2003), S. 168-177

Lindley, A.A., Becker, S., Gray, R.H. et al. 2000: Effect of continuing or stopping smoking during pregnancy on infant birth weight, crown-heel length, head circumference, ponderal index and brain: body weight ratio. In: *American Journal of Epidemiology*, Jg. 152 (2000), H. 3, S. 219–225

Lowe, J.B., Balanda, K.P., Clare, G. 1998: Evaluation of antenatal smoking cessation programs for pregnant women (TRIAL 1+2). In: Australian and New Zealand Journal of Public Health, Jg. 22 (1998), H. 1, S. 55-59

Lumley, J., Oliver, S.S., Chamberlain, C. et al. 2004: Interventions for promoting smoking cessation during pregnancy (Cochrane Review). In: Cochrane Database of Systematic Reviews 4, Art.-No.: CD001055, DOI: 10.1002/14651858.CD001055.pub2

McBride, C.M., Curry, S.J., Lando, H.A. et al. 1999: Prevention of relapse in women who quit smoking during pregnancy. In: American Journal of Public Health, Jg. 89 (1999), H. 5, S. 706-711

Melvin, C.L., Mullen, P.D., Windsor, R.A. et al. 2000: Recommended cessation counselling for pregnant women who smoke: a review of the evidence. In: Tobacco Control, 9 suppl., S. iii80–iii84

Moore, L., Campbell, R., Whelan, A. et al. 2002: Self help smoking cessation in pregnancy: cluster randomised controlled trial. In: British Medical Journal, Jg. 325 (2002), S. 1383-1389

Mullen, P.D. 2004: How can more smoking suspension during pregnancy become lifelong abstinence? Lessons learned about predictors, interventions and gaps in our accumulated knowledge. In: Nicotine & Tobacco research, Jg. 6 (2004), Suppl. 2, S. S217–S238

Mullen, P.D., Ramirez, G., Groff, J.Y. 1994: A meta-analysis of randomized trials of prenatal smoking cessation interventions. In: American Journal of Obstetrics and Gynecology, Jg. 171 (1994), S. 1328–1334

Neil-Urban, S., LaSala, K., Todd, S.J. 2002: Community Collaboration: Using nursing students in a smoking cessation program for pregnant women. In: Journal of Nursing Education, Jg. 41 (2002), H. 2, S. 76-79

Office of Research and Development, U.S. Environmental Protection Agency 1992: Respiratory health effects of passive smoking: lung cancer and other disorders. In: Office of Health and Environmental Assessment, Washington, D.C.

Pbert, L., Ockone, J.K., Zapka, J. et al. 2004: A community health center smoking-cessation intervention for pregnant and postpartum women. In: American Journal of Preventive Medicine, Jg. 26 (2004), H. 5, S. 377-385

Rigotti, N.A., Park, E.R., Regan, S. et al. 2006: Efficacy of telephone counseling for pregnant smokers - a randomized controlled trial. In: Obstetrics & Gynecology, Jg. 108 (2006), H. 1, S. 83-92

Robert-Koch-Institut und Statistisches Bundesamt (Hrsg.) 2006: Gesundheit in Deutschland. In: Gesundheitsberichterstattung des Bundes, Juli 2006

Schlaud, M., Kleemann, W.J., Poets, C.F. et al. 1996: Smoking during pregnancy and poor antenatal care: two major preventable risk factors for sudden infant death syndrome. In: International Journal of Epidemiology, Jg. 25 (1996), H. 5, S. 959-965

Secker-Walker, R.H., Solomon, L.J., Flynn, B.S. et al. 1994: Individualized smoking cessation counseling during prenatal and early postnatal care. In: American Journal of Obstetrics and Gynecology, Jg. 171 (1994), H. 5, S. 1347-1355

Severson, H.H., Andrews, J.A., Lichtenstein, E. et al. 1997: Reducing maternal smoking and relapse: Long-term evaluation of a pediatric intervention. In: Preventive Medicine, Jg. 26 (1997), S. 120-130

Stotts, A.L., DeLaune, K.A., Schmitz, J.M. et al. 2004: Impact of a motivational intervention on mechanisms of change in low-income pregnant smokers. In: Addictive Behaviors, Jg. 29 (2004), S. 1649-1657

Stotts, A.L., DiClemente, C.C., Dolan-Mullan, P. 2002: Ono-to-One. A motivational intervention for resistant pregnant smokers. In: *Addictive Behaviors*, Jg. 27 (2002), S. 275-292

Tappin, D.M., Lumsden, M.A., Gilmour, W.H. et al. 2005: Randomised controlled trial of home based motivational interviewing by midwives to help pregnant smokers quit or cut down. In: *British Medical Journal*, Jg. 331 (2005), S. 373-377

Valanis, B., Lichtenstein, E., Mullooly, J.P. et al. 2001: Maternal smoking cessation and relapse prevention during health care visits. In: *American Journal of Preventive Medicine*, Jg. 20 (2001), H. 1, S. 1-8

Valbo, A., Nylander, G. 1994: Smoking cessation in pregnancy. Intervention among heavy smokers. In: *Acta Obstetrica et Gynecologica Scandinavica*, Jg. 73 (1994), S. 215-219

Walsh, R.A., Redman, S., Brinsmead, M.W. et al. 1997: A smoking cessation program at a public antenatal clinic. In: *American Journal of Public Health*, Jg. 87 (1997), H. 7, S. 1201-1204

Wiggins, M., Oakley, A., Roberts, I. et al. 2005: Postnatal support for mothers living in disadvantaged inner city areas: a randomized controlled trial. In: *Journal of Epidemiology and Community Health*, Jg. 59 (2005), S. 288-295

Wright, L.N., Pahel-Short, L., Hartmann, K. et al. 1996: Statewide assessment of a behavioral intervention to reduce cigarette smoking by pregnant women. In: *American Journal of Obstetrics and Gynecology*, Jg. 175 (1996), H. 2, S. 283-288