

Was bringt ein evidenzbasierter Ansatz in Prävention und Gesundheitsförderung?

Norbert Schmacke

1 Einleitung

Medizin, Primärprävention und Gesundheitsförderung stellen – idealtypisch – getrennte Welten dar, was ihr Verständnis von der Erhaltung oder Wiedererlangung von Gesundheit anbelangt, und doch verbindet sie ein entscheidender Punkt: Sie offerieren eine Alternative zu Nichts-Tun bzw. Abwarten von Spontanverläufen des Lebens einzelner oder ganzer Gruppen. Anders formuliert: Auch Prävention und Gesundheitsförderung stellen Interventionen dar, die einer Legitimation bedürfen. Dies ist die erste Antwort auf die Frage in der Überschrift, und die Antwort ist nur scheinbar trivial. Nicht alle Akteure in der Gesellschaft sind wohl prinzipiell davon überzeugt, dass es immer nötig ist, die Frage nach Nutzen und Risiken von gesundheitsrelevanten Interventionen zu stellen, da sie manches für selbstverständlich halten. Die klinisch tätigen Ärzte machen sich gern über die RCT-Ära mit dem Beispiel vom Fallschirm lustig, dessen Nutzen nicht in doppelt verblindeten randomisierten Studien (weder Patient noch Therapeut wissen, welche Patientinnen/Patienten die neue Behandlung, z.B. ein Arzneimittel, erhalten, und welche mit einem Placeboansatz behandelt werden, z.B. ein Leerpräparat ohne Wirkstoff) untersucht werden müsse: zu wahr, aber vom Problem ablenkend. Und so gibt es mit Sicherheit auch eine Reihe bedeutender Rahmeneinflüsse auf die Gesundheit, die prinzipiell keiner weiteren wissenschaftlichen Unterfütterung bedürfen. Das betrifft ganz im Sinne von „Old Public Health“ die Erkenntnis, dass die Verbesserung der allgemeinen Lebensbedingungen, so eine hygienische Umwelt, ausreichende Ernährung und Zivilisierung der Arbeitswelt unabdingbare Voraussetzungen für die Entfaltung der gesundheitlichen Wohlfahrt von Gesellschaften sind. Bedenklich ist demgegenüber eine ganz andere Variante einer nach wie vor stark eingeschränkten Akzeptanz des Evidenzgedankens, die sich vielleicht besonders gut an der Debatte um die Früherkennung von Erkrankungen erläutern lässt. Hier ist offenkundig die über Jahrhunderte verankerte Überzeugung „Vorbeugen ist besser als Heilen“ so mächtig, dass es eine gelegentlich aussichtslos erscheinende Anstrengung darstellt, den Nutzen von

Früherkennungsuntersuchungen in Frage zu stellen, sogar dann, wenn es tatsächlich keinen wissenschaftlichen Nachweis für einen Vorteil dieser Vorgehensweise gibt (siehe Weymayr & Koch, 2003; Welch, 2004 sowie Mühlhauser, 2007). Im Falle mancher Krebserkrankungen kommt dem Screening daher am Ende vielleicht vor allem die Funktion symbolischen Handelns zu, demzufolge doch nicht tatenlos abgewartet werden dürfe, bis die Erkrankung sich eingestellt habe. Dies kann mit gewisser Berechtigung sogar für Screeningverfahren behauptet werden, die hoch plausibel erscheinen (zur aktuellen Studienlage siehe Ransohoff, 2009). Wenn ich mich für eine Koloskopie zur Früherkennung von Dickdarmkrebs entscheide, habe ich gezeigt, dass ich mich der neuen kulturellen Norm gegenüber konform verhalte und mich und die Gesellschaft von „unnötiger“ Last befreie. Dabei spielt der, meist überschätzte, tatsächliche Nutzen nur eine partielle Rolle. Wichtiger scheint das Gefühl, im Falle einer Erkrankung „unschuldig“ zu sein. Im sicheren Wissen „alles getan“ zu haben, hat man sich im Falle einer Erkrankung nichts vorzuwerfen. Von derartiger aus Angst genährter Legitimation ist hier freilich nicht die Rede, auch wenn als sicher gelten darf, dass individuelle wie kollektive Entscheidungen über Gesundheitsfragen sehr häufig komplexer Natur sind und die biometrische Nutzen-Risiko-Betrachtung allein oft nicht trägt. Bleibt einleitend noch einmal zu betonen, dass es nach allen Erfahrungen im Umgang mit Gesundheitsthemen wichtig ist, die Frage nach dem Nutznachweis nie aus dem Auge zu verlieren. Mit Blick auf die häufig leidvolle Geschichte der medizinischen Prävention hat Skrabanek in diesem Kontext gemahnt: „We should not confuse „prevention“ with „hopes of prevention““ (Skrabanek, 1990, S. 188). Und noch scheint es nötig zu sein, auch in der Gesundheitsförderung für ein größeres Problembewusstsein zu werben. Helen Roberts bringt das Problem wie folgt auf den Punkt:

„A few leaflets here, telling parents how to do their jobs better, a bit of social engineering there, trying to iron out a little local difficulty with housing or transport. What could be the harm in that? So long as people’s hearts were in the right place, brains were not thought to need to be quite so fully engaged in changing communities as in changing lipid lowering medication.“ (Roberts, 2004, S. 729).

Ganz so unschuldig sind eben viele Interventionen in die Lebenswelten von Menschen nicht, und im Kern richten sich die Botschaften überwiegend an Menschen aus sozial benachteiligten Schichten:

„Many public health interventions work with vulnerable or disadvantaged groups, such as children, those experiencing poverty, unemployment or other forms of marginalization. All of these groups would benefit from the increased protection from potential harm caused by poorly conducted intervention studies which would be gained from increased transparency and accountability“ (Waters et al., 2007, S. 324).

Anders formuliert: Gerade die hohe Verantwortlichkeit der Gesundheitsförderung für sozial benachteiligte Bevölkerungsgruppen verdeutlicht die Notwendigkeit, sich ein angemessenes Urteil über den Nutzen von Interventionen zu verschaffen.

Der vorliegende Beitrag konzentriert sich nun auf die aus wissenschaftlicher Sicht unverzichtbare Frage, welche Anforderungen an wissenschaftliche Belege für Nutzen und Risiken von Primärprävention und Gesundheitsförderung gestellt werden können und sollen. Dabei soll vor allem dargelegt werden, wie es um die Möglichkeit methodisch hochwertiger kontrollierter Vergleiche bestellt ist und welche Schlussfolgerungen aus den deshalb vorgestellten RCTs für die Hauptfrage des Artikels gezogen werden können (es sei auf ergänzende grundlegende Ausführungen in den Beiträgen von Veronika Müller und Gabriele Meyer in diesem Buch hingewiesen). Diese Schwerpunktsetzung erscheint wichtig, weil häufig zu hören ist, RCTs seien im Bereich der Gesundheitsförderung wegen der Komplexität der Intervention oder wegen unlösbarer Umsetzungshürden (so vor allem Rekrutierungs- und Finanzierungsprobleme) nicht durchführbar. Es ist, so die nachfolgend diskutierte Kernaussage, völlig unnötig, die RCT im Falle der Gesundheitsförderung als utopische oder unmögliche Studienform ad acta zu legen und gewissermaßen eine eigene Methodik für die Nutzenbewertung von Gesundheitsförderung einzufordern. Es ist eigentlich unnötig zu sagen, dass damit kein Monopolanspruch der RCT in diesem Bereich der Evaluationsforschung erhoben wird: es geht „nur“ um die Frage, welche Studiendesigns für welche Fragestellungen angemessen sind.

2 Die Perry Preschool Study

Zur Erläuterung der Kernaussage wird ein wissenschaftliches Experiment aus dem Erziehungsbereich etwas ausführlicher beschrieben. Von 1962 bis 1967 untersuchten David Weikart und sein Team in Ypsilanti (Michigan/USA) die Kurz- und Langzeiteffekte einer qualitativ hochwertigen Vorschulerziehung in einem sozialen Brennpunkt. Es wurden in einem Randomisierungsverfahren 58 Kinder afroamerikanischer Herkunft dem Förderprogramm zugeteilt, 65 Kinder ohne Fördermaßnahme bildeten den zweiten Arm dieser RCT, deren Endpunkte in den Lebensaltern 3 bis 11 und erneut im Alter von jeweils 14, 15, 19, 27 und 40 Jahren erhoben und publiziert wurden (Schweinhart et al., 2005). Einige herausragende Ergebnisse zugunsten der Interventionsgruppe in Zahlen (jeweils Interventions- vs. Kontrollgruppe):

- Highschool-Abschluss: gesamt 65% vs. 45% (bei den Schülerinnen 84% vs. 32%)

- in einem Beschäftigungsverhältnis im Alter von 40 Jahren: 76% vs. 62% (wobei in diesem Alter mehr Männer als Frauen einer Erwerbstätigkeit nachgingen, im Alter von 27 Jahren war es noch umgekehrt)
- Jahreseinkommen im Alter von 40 Jahren: 20.800 US\$ vs. 15.300 US\$
- Inhaftierungen (mehr als fünf) im Alter von 40 Jahren: 36% vs. 55%
- Beteiligung von Vätern an der Kindererziehung: 57% vs. 30 %
- Gebrauch von Schlafmitteln und Tranquillizern: 17% vs. 43%
- Gebrauch von Cannabis: 48% vs. 71%
- Gebrauch von Heroin: 0% vs. 9%.

Bezogen auf US\$ im Jahr 2000 und bei einer Diskontierung von 3% wird für die Studie eine Investition von 15.166 US\$ pro Kind und ein ökonomischer Return von 244.812 US\$ angegeben, v.a. resultierend aus eingesparten Inhaftierungskosten und höheren Steuereinnahmen. Als größter Einzelposten fiel die Einsparung von Inhaftierungskosten der männlichen Studienteilnehmer ins Gewicht. Akzeptiert man diesen ökonomischen Ansatz, erscheinen die Investitionen im Interventionsarm unter dem Aspekt der Übertragbarkeit des kontrollierten Experiments in den gesellschaftlichen Alltag nicht so utopisch wie sie bei isolierter Schilderung wirken könnten. Die betreffenden Kinder wurden über zwei Jahre von Lehrerinnen und Lehrern mit dem Abschluss Bachelor in Erziehungswissenschaften in Kleingruppen von fünf bis sechs Kindern täglich für 2,5 Stunden betreut. Zugrunde lag ein theoretisch begründetes, ausgefeiltes, auf Aktivierung der Kinder abzielendes Curriculum, zudem erfolgten regelmäßige Besuche in den Familien. Die Autoren sprechen selber an, dass es auf den ersten Blick fast unmöglich wirkt, die in der Studie angelegten Qualitätskriterien für verbesserte Erziehung zum Maßstab für die Anwendung in der Praxis zu nehmen, einschließlich der Tatsache, dass entsprechend qualifizierte Erzieherinnen und Erzieher besser bezahlt werden müssten. Nur in einem US-Staat, in Rhode Island, wird aktuell ein BA-Abschluss für den Einsatz in der Vorschulerziehung verlangt. Aus methodischer Sicht ist besonders interessant, dass die relativ kleine „Fallzahl“ in der Studie – entgegen weit verbreiteter Annahmen – angesichts der großen Unterschiede zwischen Interventions- und Kontrollgruppe kein Problem dargestellt hat. Für eine ausführlichere Methodendebatte muss hier auf die Monographien der Studie verwiesen werden (Schweinhart et al., 2005). Es ist aber wohl nicht übertrieben, die Perry Preschool Study als ein Highlight der pädagogisch motivierten Interventionsforschung mit direktem Bezug zu Public Health zu bezeichnen. Diese Bewertung findet sich auch in dem Cochrane Review aus dem Jahr 2000 (Zoritch et al., 2000) zur Tagesbetreuung von Kindern im Vorschulalter, der insgesamt acht RCTs mit insgesamt 2.203 Kindern ausfindig machen konnte, sämtlich aus den USA mit dem Fokus auf sozial benachteiligte Familien. Methodische Schwächen

dieser RCTs werden benannt, ohne die übereinstimmenden Kernergebnisse im Grundsatz in Frage zu stellen: Derartige Interventionen verbessern die schulischen Ergebnisse, fördern die persönliche Entwicklung, verringern den Umfang an Delinquenz und zeigen positive Ergebnisse für den (leider nicht immer wünschenswert breit erhobenen) Gesundheitszustand. Es ist leider festzuhalten, dass in diesen Studien die Daten zur Gesundheitsentwicklung umfassender hätten erhoben werden können: eine verpasste Chance dieser insgesamt kleinen Zahl relevanter Studien zur Förderung durch strukturierte Vorschulerziehung. Oakley und Fullerton sprechen an anderer Stelle bezüglich derartiger Studien von der Nadel im Heuhaufen (Oakley & Fullerton, 1996). Die Perry Preschool Study ist unter einem speziellen Aspekt für die Gesamtdenkte um die Evaluation von Ansätzen der Gesundheitsförderung interessant. Sie zeigt, dass relevante Ziele mittels einer komplexen Intervention in einem klassischen kontrollierten Design mit wenigstens akzeptablem, wenn nicht gar vorbildlichem Aufwand beforshbar sind. RCTs von komplexen Interventionen sind, so lässt sich schlussfolgern, auch in der Gesundheitsförderung keine Hexerei. Die Perry Preschool Study könnte zudem ein guter Grund sein, den in epidemiologischen Studien immer wieder gezeigten Zusammenhang von formaler Bildung und Gesundheitschancen in prospektiven Studien mit pädagogisch, sozialwissenschaftlich und medizinisch relevanten Endpunkten neu ins Blickfeld zu nehmen. Nachdem die Debatte um die Pisa-Studien verstärkt die Aufmerksamkeit auf das Thema der sozialen Benachteiligung gelenkt hat, wäre ein idealer Zeitpunkt, derartige Forschung mit einigen wenigen gut begründeten Interventionsstudien zu beginnen.

3 Experimentelle Designs in der Sozialforschung

Die Verteilung von Maßnahmen zur Verbesserung der gesundheitlichen, sozialen und ökonomischen Situation von Individuen und Bevölkerungsgruppen nach dem Zufallsprinzip steht immer auch unter dem Verdacht mangelnder Fairness oder wird gar mit dem Vorwurf ethischer Unverantwortlichkeit verbunden. Silverman und Chalmers (2001) haben in einem wenig zitierten, kurzen historischen Rückblick auf das Problem der „Chancenlotterie“ thematisiert, dass seit dem 17. Jahrhundert Bemühungen belegt sind, vermeintlich heilbringende Interventionen wie den Aderlass durch zufällige Zuteilung von Patienten auf ihren Nutzen und Schaden hin zu untersuchen. Den Beginn der „modernen“ RCT-Ära markiert dann die Studie zum Nutzen von Streptomycin bei Tuberkulosekranken. Und sie zeigen, dass dieselbe methodische Philosophie dem zitierten Ansatz der Förderung durch Vorschulerziehung zugrunde lag. Die Autoren reflektieren, dass das Bauchgefühl vielen Menschen sagt, dass die Lotterie in der Medizin und auf dem Weg zu mehr Chancengleichheit nichts verloren habe, und dass dabei nicht bedacht wird, mit welchen anderen Methoden

denn Unsicherheit aus dem Weg geräumt werden kann, ob Medizin oder soziale Interventionen dem Status Quo tatsächlich überlegen sind. Thomson et al. (2004) argumentieren darüber hinausgehend, dass gerade zur Klärung des Nutzens sozialpolitischer Maßnahmen die Zufallsverteilung unter den Bedingungen knapper Mittel auch aus ethischer Sicht die fairste Möglichkeit der Schaffung von mehr Klarheit darstellt: eine Gedankenführung, die im Vereinigten Königreich vermutlich auf deutlich größere Akzeptanz stößt als in Deutschland, weil dort eine längere Erfahrung im Umgang mit Ressourcenknappheit nach dem zweiten Weltkrieg zu verzeichnen ist.

Oakley (1998) hat ebenfalls gezeigt, wie fruchtbar der historische Rückblick für die aktuellen Methodendebatten ist. Randomisierung in Gruppen wird in den Erziehungswissenschaften seit Anfang des 20. Jahrhunderts diskutiert und in Studien erprobt. Im selben Zeitraum unternahmen Forschergruppen um Greenwood an der Columbia University und Chapin an der University of Minnesota sozialwissenschaftliche Experimente zum Nachweis des Nutzens etwa von Gesundheitserziehung in ländlichen Regionen, von Anstrengungen sozialen Wohnungsbaus oder von Erziehungsmaßnahmen für delinquente Jugendliche. Chapin publizierte 1947 sein Buch „Experimental Designs in Sociological Research“, Greenwood definierte in „Experimental Sociology“ (1945, S. 72) derartige Experimente als „the proof of a causal hypothesis through the study of two controlled contrasting situations“. Ein weiterer Meilenstein war die Publikation „Experimental and Quasi-experimental Designs for Research“ (Campbell & Stanley, 1966). Boruch veröffentlichte 1978 eine Bibliographie, die 245 randomisierte „field experiments“ umfasste, bezogen u.a. auf Maßnahmen der sozialen Wohlfahrt, Erziehung, Massenmedien und der psychologisch-psychiatrischen Hilfen (Boruch, 1978). Oakley spricht vom goldenen Zeitalter der Evaluation in den Sozialwissenschaften (Oakley, 1998, S. 1240). In dieser Phase US-amerikanischer Sozialpolitik wurden neue Programme rechtlich daran gekoppelt, dass 1% des Budgets für Evaluation verwendet wurde. In diese Phase fällt auch die Forschung der Rand Cooperation zu Reformansätzen in der Krankenversicherung, die randomisierte Studien einschloss (Manning et al., 1984). Unerwartete negative Ergebnisse bezüglich des Nutzens sozialpolitischer Maßnahmen und die Länge der Laufzeit derartiger Studien waren nach Oakley Hauptgründe, warum die Politik sich von dieser Ära ab den 1980er Jahren zunächst verabschiedete. Roberts weist mit anderen Worten auf dieses Problems hin: Die Sponsoren von Kampagnen der Gesundheitsförderung wollen in der Regel raschen Erfolg und eine gute Story, nicht aber Hinweise auf begrenzten oder fehlenden Nutzen (Roberts, 2004, S. 730). Mehrere Autorinnen und Autoren, so Connor et al. (1999) bedauern im Rückblick auf die Zeit der florierenden amerikanischen Sozialforschung, dass die Chance verpasst wurde, neben den verwendeten sozialökonomischen vermehrt gesundheitliche Outcome-Parameter einzuschließen.

Vor diesem Hintergrund wird zum einen verständlicher, warum Gesundheitsförderung sich nach wie vor schwer tut, mit Forschungsergebnissen aufzuwarten, welche die Endpunkte Lebenserwartung, Krankheitslast und Wohlbefinden überzeugend abbilden. Zum anderen bleibt aber unverständlich und kontraproduktiv, dass ausgerechnet der „Hort“ des modernen Verständnisses von Gesundheitsförderung, die Weltgesundheitsorganisation (WHO), sich gegen Ende des 20. Jahrhunderts relativ dogmatisch gegen die Einbindung von RCTs in die Evaluationsforschung ausgesprochen hat. Plädierte die WHO 1998 zunächst noch gegenüber der Politik, man solle das gesamte Spektrum an quantitativen und qualitativen Forschungsmethoden einsetzen, wurde im selben Jahr verkündet, der Einsatz von RCTs sei in der Regel unangemessen, irreführend und unnötig kostspielig (WHO, 1998a, b). Diese Linie findet dann einen Niederschlag in einem Theorieverständnis, das mit der immanenten Aversion gegen Kontrolldesigns in nichts dem Unverständnis vieler Ärztinnen und Ärzte gegenüber RCTs in der klinischen Forschung nachsteht. So sagt Jackie Green, Hochschullehrerin für Gesundheitsförderung an der Leeds Metropolitan University: „(T)here is a very real danger of ending up with little more than a menu of *proven* interventions from which to select and without a rational base to guide that selection“, und sie fordert: „Theory needs to be developed and tested not only in controlled situations, but also in the real world where inductive insights can shape the development of theory and its relevance in specific contexts“ (Green, 2000, S. 129). Hier findet, jenseits auch nur eines einzigen konkreten Beispiels, dieselbe Ex-Cathedra-Verurteilung randomisierter, kontrollierter Studien statt wie es in der klinischen Forschung bis vor kurzem auch weithin noch üblich war. Es gab, anders formuliert, immer nur vereinzelte Stimmen in der Evaluationsforschung, welche darauf insistierten, das Ziel nicht aufzugeben, nach den best möglichen Studiendesigns für den Nachweis kausaler Zusammenhänge zwischen Interventionen und angestrebten Zielen zu suchen: „Health promotion researchers are advised to insist on randomized studies more often than they do. They could educate instructors and administrators about the need for randomization rather than accept objections such as trials as justification for not performing them“, so der Verhaltensforscher Brian Flay von der University of South California (Flay, 1986, S. 463).

4 Systematische Unterstützung bei der Literatursuche

Interventionen der Gesundheitsförderung und der primären Prävention sind häufig der Kategorie „komplexe Interventionen“ zuzuordnen. Die damit einhergehenden Probleme der Standardisierung der Intervention und der Durchführung aussagefähi-

ger Evaluationen wurden erstmals im Jahr 2000 vom Medical Research Council (MRC) im Sinne einer Leitlinie für die Forschung dargestellt; Craig und andere haben diese Empfehlungen dann 2008 noch einmal deutlich erweitert publiziert (Craig et al., 2008). Diese MRC-Empfehlungen können als Referenzquelle gelten. Dem entspricht eine stattliche Sammlung von Homepages, welche jüngere Ansätze von Gesundheitsförderung und Prävention in kontrollierten Studiendesigns vorstellen; sie sollen hier knapp aufgelistet werden:

Guide to Community Preventive Services (Centers for Disease Control and Prevention <http://www.thecommunityguide.org/index.html>): Auf dieser Website finden sich frei zugängliche Reviews zu Interventionen in verschiedenen gesundheitlichen Bereichen (Jugendgesundheit, Alkoholmissbrauch, Asthma, Übergewicht usw.). Ein Blick in die jeweiligen Reviews erlaubt eine rasche Orientierung, für welche Interventionsansätze Studien welcher Güte vorhanden sind. So finden sich z.B. Studien zur Gewichtsreduktion in Schule oder Betrieb oder körperlicher Aktivität und die Bandbreite reicht von schlichten Informationsprogrammen bis zu hoch komplexen Interventionen. Auch wenn sich die Ergebnisse auf englischsprachige Publikationen beschränken, bieten sie eine Fundgrube an Orientierungshilfen.

Informativ sind weiter die von den Centers for Disease Control (CDC) erstellten „**Morbidity and Mortality Weekly Reports**“ (MMWR). Die CDCs verstehen die drei auf der Website <http://www.cdc.gov/mmwr/w> angebotenen Formate (MMWR Weekly, MMWR Recommendations and Reports und MMWR Surveillance Summaries) als ihre primären Verbreitungsmedien für wissenschaftliche Informationen aus dem Bereich Public Health. In den Recommendations und Reports findet sich z.B. ein Überblick über die Erfahrungen mit unterschiedlichen Ansätzen der Gewaltprävention an Schulen sowie viele andere Themen höchster gesellschaftlicher Relevanz (www.cdc.gov/mmwr/preview/mmwrhtml/rr5607a1.htm).

Neben der einschlägig bekannten Cochrane Collaboration liefert die **Cochrane Public Health Review Group** einen weiteren Einblick in Studien zur Evaluation von Gesundheitsförderung (www.vichealth.vic.gov.au/cochrane).

Auch das **NHS Centre for Reviews and Dissemination** (www.york.ac.uk/inst/crd/wph.htm), die **International Union for Health Promotion and Education** (www.iuhpe.nyu.edu/pbs) sowie last not least die **Campbell Collaboration** (www.campbellcollaboration.org). bieten einen guten Überblick über vorhandene Studien mit methodisch hochwertigen Designs.

5 Die Chancen des kommenden (?) Präventionsgesetzes

Wenn man die ideologische Grundsatzdebatte im positiven Sinne überwunden hat, kann man sich den Herausforderungen und Problemen für kontrollierte Vergleiche in der Evaluationsforschung in Ruhe zuwenden. Sie gleichen bei näherer Betrachtung den Problemen bei der Durchsetzung des RCT-Designs als Standard für klinische Studien mehr als man denken mag. Wenn Sanson-Fisher et al. etwa schreiben: „A major problem for randomized community intervention studies ist the availability and accessibility of an appropriate number of eligible and widely dispersed communities“ (Sanson-Fisher et al., 2007, S. 157), so ist dem fraglos zuzustimmen; diese Hürde muss aber auch jede anspruchsvolle Multicenter-Studie in der klinischen Forschung überwinden, und viele Ansätze scheitern auch hier dann oft nicht an prinzipiellen, sondern an logistischen und ökonomischen Fragen. Was das Geld anbelangt, so beziffern dieselben Autoren die Kosten einiger interessanter RCTs im Präventionsbereich auf 2,5 bis 6 Millionen US\$: Das sind Summen, die in der klinischen Forschung niemanden in Unruhe versetzen würden (Sanson-Fisher et al., 2007, S. 158).

In einer systematischen Auswertung der methodischen Qualität und Belastbarkeit internationaler Studien zu Prävention und Gesundheitsförderung mit dem Fokus auf Bewegung, Depression, Ernährung und Rauchen kommen Klever-Deichert et al. zu folgender Schlussfolgerung: „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“ (Klever-Deichert, 2007, S. 12). Dabei stellt sich als Hauptproblem nicht heraus, dass es einen grundsätzlichen Mangel an Studien mit kontrollierten Designs gäbe: Das Hauptproblem ist neben einer Vielzahl an methodisch inadäquaten Studien die mangelnde Qualität vieler RCTs. Mit dieser mit Blick auf das geplante Präventionsgesetz durchgeführten Studie wird ein weiteres Mal deutlich, dass die prinzipiellen Bedenken gegen Kontrolldesigns, wie sie Elkeles (2006) beispielsweise geäußert hat, von dem Problem ablenken, dass es dringend erforderlich ist, die Studienkultur in der Prävention und Gesundheitsförderung zu verbessern und dass hierfür die notwendigen Mittel eingefordert und eingeworben werden müssen. Aus dem offenkundigen Forschungsdefizit zu schließen, die Forderung nach hochwertigen methodischen Studien schade der Entwicklung der Gesundheitsförderung generell, wirkt wie eine Vogel-Strauß-Politik – so irritierend bis ärgerlich allemal die

Schiefverteilung der Ressourcenverteilung in der Gesundheitsforschung zu Gunsten der klinischen Forschung, insbesondere der Arzneimittelstudien auch ist.

So lautet die abschließende Beantwortung der Fragestellung: Auch in der Primärprävention und Gesundheitsförderung müssen methodisch angemessene Studien die notwendige Legitimation für die Implementierung von Interventionen herstellen. Auf RCTs muss innerhalb des methodischen Arsenal nicht deshalb verzichtet werden, weil sie angeblich prinzipiell unangemessen oder undurchführbar sind. Insofern wird dafür plädiert, künftig wie in der klinischen Forschung bei der Entwicklung von Studiendesigns zur Testung neuer Ansätze der Prävention und Gesundheitsförderung darzulegen, wie das jeweils gewählte Studiendesign begründet worden ist. Dies könnte in einer Kettenreaktion zu zweierlei führen: 1. dass über Second-Best-Designs ehrlicher geredet werden kann, wenn z.B. Rekrutierungs- oder Finanzierungsprobleme zur Realisierung von RCTs unübersehbar sind und dass 2. gleichzeitig die Schar derer wächst, die ihre Stimme in Richtung der Forschungsförderung und der Politik lauter erheben, die Weiterentwicklung von Standards in Prävention und Gesundheitsförderung nicht wie bisher sträflich zu vernachlässigen. Die Perry Preschool Study hat vor über 40 Jahren gezeigt, dass Pionierforschung von höchster gesellschaftlicher Relevanz möglich ist, mit deren Rückendeckung Implementierung und nachfolgende Evaluationen eine ganz andere Durchsetzungskraft aufweisen, als wenn sie sich auf noch so plausible Annahmen stützen müssen. Insofern leitet sich daraus mit Blick auf das viel diskutierte Präventionsgesetz im Grunde die Forderung ab, einen Teil der Gelder in eine beratende Institution zu investieren, die Studiendesigns vor Aufnahme von Interventionen bewertet und eine Homepage vorbildlicher Ansätze in Primärprävention und Gesundheitsförderung für den deutschsprachigen Raum aufbaut. Und sicher wäre es auch ratsam, wie in der klinischen Forschung zu verlangen, dass Studiendesigns *vor* Aufnahme von Evaluationen publiziert werden. Im großen Stil hat dies übrigens 2007 auch die WHO gefordert, indem sie für den Aufbau eines Registers für Interventionsstudien plädierte (Waters et al., 2007). So sympathisch diese Forderung auch sein mag: die Bindung einer derartigen Idee an einen gesetzlichen Rahmen, hier das nach wie vor ausstehende Präventionsgesetz, erscheint für die Implementierung einer Überfliegeridee allemal besser. Prävention und Gesundheitsförderung sind – darüber sollte sicher kein Streit aufkommen – auf jeden Fall zu wichtig, um sie allein dem gesunden Menschenverstand zu überlassen.

Literatur

- Boruch, R. F. (1978). Randomized field experiments for program planning, development and evaluation: an illustrative bibliography. *Evaluation Review*, 2, 655-695.
- CDC (2000). Internet site for the Community Guide: Verfügbar unter: www.health.gov/communityguide.
- Connor, J., Rodgers, A. & Priest, P. (1999). Randomised studies of income supplementation: a lost opportunity to assess health outcomes. *Journal of Epidemiology and Community Health*, 53, 725-730.
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I. & Petticrew, M. (2008). Developing and evaluating complex interventions: new guidance. Verfügbar unter: <http://www.mrc.ac.uk/complexinterventionsguidance>
- Elkeles, T. (2006). Evaluation von Gesundheitsförderung und Evidenzbasierung? In W. Bödeker & J. Kreis (Hg.), *Evidenzbasierte Gesundheitsförderung und Prävention* (S. 111-153). Bremerhaven: Wirtschaftsverlag NW.
- Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine*, 15, 451-474.
- Green, J. (2000). The role of theory in evidence-based health promotion practice. *Health Education and Research*, 22, 125-129.
- Greenwood, E. (1945). *Experimental sociology*. New York: Octagon Books.
- Klever-Deichert, G., Gerber, A., Schröer, M. A. & Plamer, E. (2006). Internationale erfolgreiche Interventionen der Prävention und Gesundheitsförderung und ihre Übertragbarkeit auf Deutschland. *Studien zu Gesundheit, Medizin und Gesellschaft*; Köln: Ausgabe 09/2007 vom 12.11.2007. Verfügbar unter: www.uk-koeln.de/kai/igmg/sgmg/2007-09_praeventionspolitik_international.pdf.
- Manning, W. G., Leibowitz, A., Goldberg, G. A., Rogers, W. H., Newhouse, H. P. (1984). A controlled trial of the effect of a prepaid group practice on use of services. *New England Journal of Medicine*, 310, 105-1010.
- McQueen, D. V. (2001). Strengthening the evidence base for health promotion. *Health Promotion International*, 16, 261-268.
- Medical Research Council (2000). *A framework for development and evaluation of RCTs for complex interventions to improve health*. London: MRC.
- Mühlhauser, I. (2007) Ist Vorbeugen besser als Heilen? *Zeitschrift für Ärztliche Fortbildung und Qualitätssicherung*, 101, 293-299.
- Oakley, A. (1998). Experimentation and social interventions: a forgotten but important history. *British Medical Journal*, 317, 1239-1242.
- Oakley, A. & Fullerton, D. (1996). The lamp-post of research: support or illumination? In A. Oakley & I. Roberts (Eds.), *Evaluating social interventions* (pp. 4-38). Ilford, Essex: Barnados.
- Ransohoff, D. F. (2009). How much does colonoscopy reduce colon cancer mortality? *Annals of Internal Medicine*, 150, 50-52.
- Roberts, H. (2004). Intervening in communities: challenges for public health. *Journal of Epidemiology and Community Health*, 58, 729-730.
- Sanson-Fisher, R. W., Bonevski B., Lawrence W. G. & D'Este C. (2007). Limitations of the randomized controlled trial in evaluating population-based health interventions. *American Journal of Preventive Medicine*, 33, 155-161.

- Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R. & Nores, M. (2005). *Lifetime effects: The High/Scope Perry Preschool Study Through Age 40* (Monographs of the High/Scope Educational Research Foundation, 14). Ypsilanti, MI: High/Scope Press.
- Silverman, W. A. & Chalmers, I. (2001). Casting and drawing lots: a time honoured way of dealing with uncertainty and ensuring fairness. *British Medical Journal*, 323, 1467-1468.
- Skrabanek, P. (1990). Why is preventive medicine exempted from ethical constraints? *Journal of Medical Ethics*, 16, 187-190.
- Thomson, H., Hoskins, R., Petticrew, M., Ogilvie, D., Craig, N., & Quinn, T. (2004). Evaluating the health effects of social interventions. *British Medical Journal*, 328, 282-285.
- Waters, E., Priest, N., Armstrong, R., Oliver, S., Baker, P., McQueen, D. V., Summberbell, C., Kelly, M.R. & Swinburn, B. (2007). The role of a prospective public health intervention study register in building public health evidence: proposal for content and use. *Journal of Public Health*, 29, 322-327.
- Welsh, H. G. (2004). *Should I be tested for cancer? Maybe not and here's why*. California Press.
- Weymayr, C. & Koch, K. (2003). *Mythos Krebsvorsorge. Schaden und Nutzen der Früherkennung*. Frankfurt/M.: Eichborn.
- WHO (1998a). *Fifty-First World Health Assembly WHA51.12: Health Promotion*. Geneva: WHO.
- WHO (1998b). *Health promotion evaluation: recommendations to policymakers*. Copenhagen: WHO.
- Zoritch, B., Roberts, I. & Oakley, A. (2000). Day care for pre-school children. *Cochrane Database of Systematic Review* 2000, Issue 3. Art. No.:CD000564. DOI: 10.1002/14651858.CD000564