



GRÜNE LISTE PRÄVENTION

Auswahl- und Bewertungskriterien für die
CTC Programm - Datenbank

Landespräventionsrat Niedersachsen
Frederick Groeger-Roth, Dr. Burkhard Hasenpusch
Fassung vom 1.11.2011



Die folgenden Auswahl- und Bewertungskriterien werden von den Autoren regelmäßig überprüft und ggf. neuen Entwicklungen angepasst.

A) Auswahlkriterien für die Aufnahme in die Datenbank

Aufgenommen in die Grüne Liste Prävention werden nur Aktivitäten, die der folgenden Definition für ein „evaluiertes Präventionsprogramm“ entsprechen.

- 1) Ein **Programm** soll eine Maßnahme oder eine Intervention sein, die
 - auf Replizierbarkeit an einem anderen Ort oder zu einem späteren Zeitpunkt angelegt ist (durch eine explizite Handlungsanleitung, z.B. ein geschriebenes Manual, Ablaufschema etc.)
 - auf ein oder mehrere (überprüf-, bzw. messbare) Ziele hin orientiert ist
 - für die jeweiligen Teilnehmer zeitlich begrenzt ist, aber als Maßnahme auf Dauer angeboten werden kann
 - zusätzlich zu einer vorhandenen Basisstruktur durchgeführt wird
- 2) „**Präventionsprogramme**“ sollen Programme sein, die starten, bevor das problematische Verhalten bei der Zielgruppe auftritt oder sich stabilisiert. Programme können grundsätzlich auf eine primäre, sekundäre oder tertiäre Präventionsebene ausgerichtet sein. Die Zielgruppe kann universell (an alle gerichtet), selektiv (mit erhöhtem Risiko) oder indiziert (mit ersten Anzeichen von Problemverhalten) sein. Im Rahmen der CTC-Datenbank werden nur primär- oder sekundärpräventive Programme, bzw. universell oder selektiv ausgerichtete Programme aufgenommen, da dies der Ausrichtung des CTC-Ansatzes entspricht. Das Bewertungsschema ist aber gleichermaßen für tertiärpräventive und indizierte Programme geeignet, dies gilt ebenso für reaktive (Behandlungs-)Ansätze.
- 3) Für das Programm liegt mindestens eine wissenschaftliche **Evaluationsstudie** aus dem deutschsprachigen Raum vor.
- 4) Die Datenbank für die **Umsetzung der CTC-Strategie** umfasst nur Programme, die darauf ausgerichtet sind, einen oder mehrere Risikofaktoren zu reduzieren, bzw. einen oder mehrere Schutzfaktoren zu stärken. Einbezogen werden nur diejenigen Faktoren, die innerhalb der CTC-Strategie als relevant für jugendliches Problemverhalten (Gewalt, Delinquenz, Substanzmissbrauch, Schulabbruch, frühe Schwangerschaften, Depressionen und Ängste) identifiziert wurden.
- 5) Das Programm ist in Deutschland, bzw. Niedersachsen implementierbar. Das beinhaltet die **Verfügbarkeit** von Materialien, Trainings oder technischer Unterstützung bei der Implementierung; es sei denn, das Programm ist ohne gesonderte Unterstützung umsetzbar.

B) Einstufung der Präventionsprogramme

Empfohlen werden in der Datenbank Programme der Bewertungsstufen 1 – 3, daneben existiert eine Kategorie „auf der Schwelle“ und die Kategorie „nicht aufgenommen“.

Stufe 1: Effektivität theoretisch gut begründet¹

1 a) Konzept:

- theoretisch klar begründetes Modell über die Wirksamkeit, die angenommenen Wirkfaktoren sind klar benannt (unter Bezug auf wissenschaftlich anerkannte theoretische Modelle)
- eingesetzte Methoden und Instrumente werden theoretisch gut begründet
- ein logischer Zusammenhang von: „Problemanalyse – beeinflussbaren Faktoren – Zielen – Zielgruppen – Methoden“ ist erkennbar
- Programm bezieht sich auf (aus der Forschung bekannte) Risiko- und Schutzfaktoren
- die Zielgruppe(n) sind präzise und umfassend beschrieben
- Handlungsanleitungen und Manuale sind nachvollziehbar aus dem Modell abgeleitet
- Ziele sind explizit formuliert und messbar/überprüfbar
- wenn nicht in Deutschland entwickelt, wurden der ursprüngliche Kontext und die erfolgten Anpassungen beschrieben

1b) Umsetzung

- eingesetzte Methoden und Instrumente sind didaktisch gut aufbereitet und nachvollziehbar beschrieben
- benötigte Materialien/ Manuale sind verfügbar und auf aktuellem Stand
- notwendige Trainings und Schulungen sind vorhanden
- entstehende Kosten sind transparent
- Support bei der Umsetzung ist verfügbar
- Instrumente für die Qualitätskontrolle bei der Umsetzung sind verfügbar

1c) Evaluation

- mindestens eine Evaluation mit dem Niveau 0 Sterne (s.u.) mit (überwiegend) positivem Ergebnis

Stufe 2: Effektivität wahrscheinlich

- 1 a) und b) wie oben
- mindestens eine Evaluationsstudie * - *** Sterne mit (überwiegend) positivem Ergebnis

Stufe 3: Effektivität nachgewiesen

- 1 a) und b) wie oben
- mindestens eine Evaluationsstudie **** oder ***** Sterne mit (überwiegend) positivem Ergebnis und mindestens hinreichender Beweiskraft

¹ Sobald ein Punkt nicht erfüllt ist, führt dies zu der Einordnung „Auf der Schwelle“. Sind die ersten drei Punkte der Konzeptqualität nicht erfüllt, erfolgt die Einordnung „keine Aufnahme“ (s.u.).

C) Bewertung der Evaluationsstudien

1) Bewertung des Evaluationsdesigns

Niveau	Evaluationstypen / Design
**** Fünf Sterne	<ul style="list-style-type: none"> RCT mit follow-up (i.d.R. 6 Monate)
**** Vier Sterne	<ul style="list-style-type: none"> quasi-experimentelle Studien mit follow-up (i.d.R. 6 Monate) Kohortendesign mit follow-up (i.d.R. 6 Monate)
<i>Schwellenwert Stufe 3: Effektivität nachgewiesen</i>	
*** Drei Sterne	<ul style="list-style-type: none"> RCT ohne follow-up Quasi-Experiment in der Praxis ohne follow-up Kohortendesign ohne follow-up
** Zwei Sterne	<ul style="list-style-type: none"> rein "klinisches" RCT oder klinisches Quasi-Experiment mit oder ohne follow-up Vorher-Nachher-Messung mit Kontrollgruppe(n) in der Praxis
* 1 Stern	<ul style="list-style-type: none"> Benchmark / Norm-Referenz-Studie Theory of Change – Studie
<i>Schwellenwert Stufe 2: Effektivität wahrscheinlich</i>	
0 kein Stern:	<ul style="list-style-type: none"> Teilnehmer-Zufriedenheits-Messung Vorher-Nachher-Messung ohne Kontrollgruppe Ziel-Erreichungs-Messung Qualitäts-Sicherungs-Studie
<i>Schwellenwert Stufe 1: Effektivität theoretisch gut begründet</i>	
Keine Aussage über Wirkung möglich	<ul style="list-style-type: none"> Literatur- und Dokumenten – Auswertungen, Prozess-Beschreibungen etc.

2) Bewertung der Beweiskraft der Evaluationsergebnisse

Evaluationsstudien werden berücksichtigt, wenn sie (entsprechend dem jeweiligen Untersuchungsdesign) eine nachvollziehbare Ergebnispräsentation nach allgemeinen wissenschaftlichen Standards beinhalten.

Für Studien mit Kontrollgruppen sollten Effektstärken (möglichst Cohen's d) berechnet sein. Für Effektstärken gilt, dass mindestens ein signifikanter Effekt in einer relevanten Verhaltensdimension und bei mindestens einer Zielgruppe vorliegen muss. Beim „follow-up“ (mindestens 6 Monate nach Abschluss der Intervention) muss mindestens ein positiver Effekt anhalten.

Niveau Evaluations-design	schwache Beweiskraft	hinreichende Beweiskraft	starke Beweiskraft	sehr starke Beweiskraft
4 – 5 Sterne	Interventions- und Kontrollgruppen n < 20 oder keine Studie im deutschsprachigen Raum mit 4 – 5 Sternen, aber eine oder mehrere ausländische Studien mit mindestens hinreichender Beweiskraft (#)	Interventions- und Kontrollgruppen n = 20 - 50 und mindestens eine Studie im deutschsprachigen Raum oder: 2 oder mehr Studien (4 – 5 Sterne) mit Interventions- und Kontrollgruppen n < 20 im deutschsprachigen Raum	Interventions- und Kontrollgruppen n > 50 oder 2 oder mehr Studien mit hinreichender Beweiskraft im deutschsprachigen Raum mit mehrdimensionalen Befragungen	Kriterien wie bei starker Beweiskraft plus mindestens ein groß angelegtes Feld-Experiment im deutschsprachigen Raum oder Kriterien wie bei starker Beweiskraft plus 2 oder mehrere ausländische Studien mit starker Beweiskraft
1 - 3 Sterne	vorläufige (* - **) oder schwache (***) Beweiskraft			
0 Sterne	keine Beweiskraft			

(#) Stufen der Übertragbarkeit bei im Ausland entwickelten Programmen:

- gegeben: mehrere Studien mit mindestens hinreichender Beweiskraft im deutschsprachigen Raum
- wahrscheinlich: mindestens 1 derartige Studie im deutschsprachigen Raum (Schwellenwert)
- möglich: keine Studie im deutschsprachigen Raum, aber ausreichende Begründung über vergleichbare Zielgruppen und Umstände ist vorhanden (führt zur Einordnung Stufe 1: „Effektivität theoretisch gut begründet“)
- unklar: führt zur Einordnung „auf der Schwelle“

3) Beurteilung der Evaluationsergebnisse:

- nur positiv: bei allen gemessenen Dimensionen ein signifikanter Effekt in die erwünschte Richtung
- überwiegend positiv: bei der Mehrzahl der gemessenen Dimensionen ein positiver signifikanter Effekt (Schwellenwert für Empfehlung)
- teils positiv und negativ: mehrere Dimensionen gehen in die unerwünschte Richtung
- kein Effekt
- überwiegend negativ / nur negativ

D) Gesamtübersicht Bewertungsschema

Konzept- und Umsetzungs-qualität	Evaluations-niveau und Beweiskraft	Evaluations-ergebnisse	Gesamtbewertung
1 a – c sind erfüllt	4 bis 5 Sterne: hinreichende Beweiskraft oder besser	(überwiegend) positiv	Effektivität nachgewiesen (Stufe 3)
1 a – c sind erfüllt	1 bis 3 Sterne 4 – 5 Sterne mit nur schwacher Beweiskraft	(überwiegend) positiv	Effektivität wahrscheinlich (Stufe 2)
1 a – c sind erfüllt	0 Sterne	(überwiegend) positiv	Effektivität theoretisch gut begründet (Stufe 1)
1 a – c sind (teilweise) erfüllt	keine Evaluation oder unterhalb 0 Sterne	teils positiv und negativ, oder unklar (0-5 Sterne)	auf der Schwelle
1 a – c sind (nicht) erfüllt	Alle	keine oder negative Effekte	Keine Aufnahme

Flussdiagramm
Einordnung Programme

Programm XY

Schritt 1: Prüfung der Konzept- und Umsetzungsqualität

Nicht erfüllt:
„keine Aufnahme“

teilweise erfüllt:
„Auf der Schwelle“

komplett erfüllt:
Weiter mit Schritt 2

Schritt 2: Prüfung des Evaluationsniveaus und der Beweiskraft der Evaluation

Unter 0 Sterne:
„Auf der Schwelle“

0 – 5 Sterne:
weiter mit Schritt 3

Schritt 3: Prüfung der Evaluationsergebnisse

Keine oder negative
Effekte:Keine Aufnahme

(überwiegend) positiv:
Empfehlung

teils - teils oder unklar:
„Auf der Schwelle“

Schritt 4: Gesamtbewertung

Stufe 1: Effektivität
theoretisch gut
begründet

Stufe 2:
Effektivität
wahrscheinlich

Stufe 3:
Effektivität
nachgewiesen

E) Das theoretische Modell

Die Bewertung von Programmen in der Datenbank wurde in Anlehnung an das theoretische Modell von Jan Veerman und Tom van Yperen (Veermann / van Yperen 2007) vorgenommen. Das von diesen Autoren vorgeschlagene „Entwicklungsmodell“ für effektive Interventionen geht davon aus, dass Interventionen (wie z.B. Präventionsprogramme) nicht nach dem simplen Raster unterschieden werden sollten, ob sie bewiesen wirksam („evidenzbasiert“) oder „nicht evidenzbasiert“ sind. Programme befinden sich vielmehr meist auf einem Entwicklungsweg, auf dem verschiedene Stufen (eine deskriptive, theoretische, indikative oder kausale Ebene) der Beweiskraft unterschieden werden können. Je niedriger die Entwicklungsstufe ist, desto freier können die Evaluationskonzepte gewählt werden. Für den Nachweis der Wirksamkeit auf der kausalen Ebene kommen aber nur die Evaluationsansätze in Frage, die auch kausale Aussagen erlauben. Die praktische Konsequenz dieses Modells lautet also nicht, es müssten mehr „randomisierte Kontrollstudien“ durchgeführt werden, sondern Interventionen wie z.B. Präventionsprogramme müssen ihre Annahmen über Wirkungszusammenhänge expliziter formulieren, damit sie entsprechenden Tests unterzogen werden können. Programmentwickler sollen somit durch Empfehlungslisten wie die vorliegende Datenbank motiviert werden, ihre Ansätze entsprechend des Entwicklungsmodells weiter zu verbessern. Im Unterschied zu vielen in den USA angewendeten Auswahlkriterien wird den theoretischen Erklärungsmodellen, **warum** und **auf welche Weise** etwas wirksam sein soll, bei Veermann und van Yperen mehr Bedeutung gegeben, ohne die Bedeutung von methodisch gut kontrollierten Evaluationsstudien damit zu verringern.

Dieses Modell liegt auch der niederländischen Datenbank effektiver Jugendinterventionen (<http://www.nji.nl/eCache/DEF/1/21/478.html>) zu Grunde.

Ebene der Beweiskraft	Typ der Forschung
4. Kausale Beweiskraft <i>Zeigen: die Wirkung wurde von der Intervention verursacht</i>	- Zufallsexperiment mit Kontrollgruppen (RCT)
3. Indikative Beweiskraft <i>Zeigen: die Ziele werden erreicht</i>	- Quasi-Experiment - “Theory of change” - Studien - Ziel-Erreichungs-Studien
2. Theoretische Beweiskraft <i>Begründen: warum sollte die Maßnahme wirken?</i>	- Reviews, Meta-Analysen - Expertenwissen (- Logische Modelle)
1. Deskriptive Beweiskraft <i>Beschreiben: was wird gemacht?</i>	- Prozess-Evaluation - Dokumenten-Analyse

Quellen:

Hawkins, J.D., Catalano, R.F. (2004): Communities That Care Prevention Strategies Guide. South Deerfield, MA: Channing Bete

Ince, D., Beumer, M., Jonkman, H. & Vergeer, M. (2007): Veelbelovend en effectief. Overzicht van preventieprojecten en programma's in de domeinen Gezin, School, Kinderen en jongeren, Wijk. Vierde druk, Netherland Youth Institut, Utrecht

Netherland Youth Institut NJI (2008): Database of Effective Youth Interventions, Utrecht

Veermann, J.W. / van Yperen, T.A. (2007): Degrees of freedom and degrees of certainty: A developmental model for the establishment of evidence-based youth care, Evaluation and Program Planning 30: 212 – 221

van Yperen, T.A. (2007): Integraal erkend. Naar een afstemming erkenning jeugdinterventies. Netherland Youth Institut, Utrecht

van Yperen, T.A. / van Bommel, M. (2009): Erkenning interventies: criteria 2009-2010, Erkenningcommissie (Jeugd)Interventies, Netherland Youth Institut, Utrecht



Herausgeber:

Landespräventionsrat Niedersachsen
Niedersächsisches Justizministerium
Am Waterlooplatz 5 A
30169 Hannover
www.lpr.niedersachsen.de

Sozialräumliche Prävention in Netzwerken – SPIN

Modellhafte Implementierung des Programms
Communities That Care – CTC in Niedersachsen
ist ein Kooperationsprojekt des
Landespräventionsrates Niedersachsen mit der
LAG Soziale Brennpunkte Nds. e.V.



SPIN wird gefördert durch:



With support from the Prevention of and Fight against
Crime Programme of the European Union
European Commission - Directorate-General Home
Affairs.

This publication reflects the views only of the author,
and the European Commission cannot be held
responsible for any use which may be made of the
information contained therein.

