

# **Beurteilung der Effektivität und der Qualität von Screening Programmen**

## **WHO, UK-NSC, USPSTF Systeme**

**1968 – 2002**

Schmid D, Piribauer F (Wien 2004)

# 1 WHO-Kriterien von Wilson&Jungner

Wilson und Jungner haben bereits in einem umfangreichen WHO-Bericht von 1968 Entscheidungskriterien über die Angemessenheit eines Screening-/Präventionsprogramms definiert:<sup>9</sup>

- Die Erkrankung soll ein bedeutsames Gesundheitsproblem sein,
- die Biologie – der natürliche Verlauf der Erkrankung von der latenten Phase bis zur manifesten Erkrankung – muss weitgehend verstanden sein,
- die Erkrankung muss eine identifizierbare Frühphase haben,
- die Behandlung im Frühstadium muss wirksamer sein als im Spätstadium,
- es muss einen geeigneten Test für die Entdeckung der Frühphase der Erkrankung geben,
- der Test muss für die Bevölkerung akzeptabel sein,
- die Untersuchungsintervalle des Screeningtests müssen von vornherein bekannt sein,
- Einrichtungen (Ressourcen) müssen bereits verfügbar sein, die den erhöhten Versorgungsbedarf, der durch bevölkerungsbasierte Screeningprogramme anfällt (wie z. B. definitive diagnostische Untersuchung, Folgebehandlungen), abdecken,
- das Risiko eines mit den Screeningmaßnahmen assoziierten physischen und psychischen Schadens muss bewiesenermaßen geringer sein als der Nutzen substantieller/moderater (Nettonutzen),
- die Kosten müssen in einem annehmbaren Verhältnis zum Gesundheitsnutzen des Programms stehen.

Diese ursprünglichen WHO-Kriterien wurden von dem 1996 etablierten UK National Screening Committee (NSC) aktualisiert.<sup>i,ii</sup> Dabei fand die internationale Arbeit auf dem Gebiet von Screening insbesondere in den USA<sup>iii</sup> und Kanada<sup>iv</sup> Berücksichtigung. Weiters wurde der neue strenge Standard erforderlicher wissenschaftlicher Methoden-Qualität (Evidenzgrad) für die Beurteilung von Wirksamkeit/Nutzen und Schaden von Versorgungsmaßnahmen miteinbezogen.

## 2 UK-NSC-Kriterien zur Beurteilung von Screeningprogrammen für Krankheiten

Die UK-National-Screening-Committee-Kriterien dienen der Beurteilung eines Screeningprogramms hinsichtlich Durchführbarkeit, Effektivität und Angemessenheit für bestimmte Krankheiten:

- Die Krankheit soll ein bedeutendes Gesundheitsproblem darstellen,
- die Epidemiologie und Biologie der Erkrankung einschließlich der Entwicklung vom latenten zum manifesten Erkrankungsstadium soll ausreichend verstanden sein,
- es sollen Risikofaktoren oder Erkrankungsmarker sowie Latenzperiode oder frühes symptomatisches Stadium bekannt und auffindbar sein,
- die kosteneffektiven primären präventiven bevölkerungsbasierten Interventionen sollten – sofern praktikabel – implementiert sein (z. B. Antirauchkampagne).

## 2.1 Test

- Der Screeningtest soll einfach, sicher, und genau sein,
- die Verteilung der Testmessergebnisse in der Zielpopulation soll bekannt und ein adäquater Cut-off-Wert definiert und festgesetzt sein,
- der Test soll für die Bevölkerung annehmbar sein,
- es soll eine allgemein akzeptierte Vorgangsweise (*agreed policy*) für weitere diagnostische Untersuchung(en) von Personen mit positivem Testergebnis geben.

## 2.2 Behandlung

- Es soll eine wirksame Behandlung oder Intervention für Patienten geben, die durch Früherkennung der Krankheit identifiziert wurden. Die Behandlung im Frühstadium der Krankheit muss erwiesenermaßen von höherem Nutzen sein als in der Spätphase,
- es soll Evidenz-basierte, allgemein akzeptierte Vorgangsweisen geben, die empfehlen, welchen Personen eine Behandlung – und in welcher Form – anzubieten ist,
- die Behandlung (*clinical management*) der Krankheit und deren Ergebnis (*outcome*) sollte durch alle Leistungserbringer des Gesundheitssystems vor Aufnahme in ein Screeningprogramm optimiert werden.

## 2.3 Screeningprogramm

- Es müssen gesicherte Ergebnisse von randomisiert kontrollierten klinischen Studien guter Qualität vorliegen, welche nach dem „*intention to treat*“-Prinzip analysiert wurden, die belegen, dass das vorgeschlagene Screeningprogramm Morbidität und/oder Mortalität der Zielkrankheit reduziert,

- wo Screening darauf ausgerichtet ist, eine informierte Entscheidungsfindung bei der Zielperson zu ermöglichen (z. B. Screening nach Down Syndrom oder genetischer Anlage für zystische Fibrose), muss Evidenz von qualitativ hochwertigen Studien vorliegen, welche die Genauigkeit des Tests belegen. Die Information über den Test und dessen Ergebnis muss für den Betroffenen wertvoll und leicht verständlich sein,
- es soll Evidenz dafür vorhanden sein, dass das komplette Screeningprogramm (Screeningtest, definitive diagnostische Untersuchung(en), Beratung, Behandlungseingriffe/Intervention) klinisch, gesellschaftlich und ethisch sowohl von allen Beteiligten des Gesundheitssystems als auch von der Öffentlichkeit akzeptiert wird,
- der Nutzen des Screening-/Präventionsprogramms soll den psychischen und physischen Schaden, verursacht durch Screeningtest bzw. dessen Ergebnisse, definitive diagnostische Untersuchung(en) und Behandlungseingriffe, überwiegen (Messzahlen für den Schaden: Anzahl der Personen, die durch das Screeningprogramm Schaden erleiden bezogen auf 1.000 Screenees oder auf ein gerettetes Leben),
- Opportunitätskosten des Screening-/Präventionsprogramms (inkl. Kosten für Screeningtest, diagnostische Untersuchung(en), Administration, Training und Qualitäts(sicherungs)management) sollen in einem ökonomisch annehmbaren Verhältnis zu den Ausgaben der gesamten medizinischen Versorgung (= Wert des Geldes!) stehen,
- das Vorliegen eines Plans für das Management wie auch für das Monitoring des Screeningprogramms ist unerlässlich. Vereinbarte qualitätssichernde Maßnahmen haben vorab beschlossen zu sein,
- Personalressourcen sowie Einrichtungen für Screeningtests, definitive Diagnose (Referenztests), Behandlung und Programmmanagement sind vor Beginn des Screeningprogramms in ausreichender Form zur Verfügung zu stellen,
- alle anderen Optionen zur Erkrankungsbehandlung sollten bereits berücksichtigt worden sein (z. B. Verbesserung der therapeutischen Maßnahmen, Alternativleistungen etc.), um zu gewährleisten, dass keine weiteren Kosten verursachenden Interventionen eingeführt oder laufende Interventionen innerhalb der verfügbaren Ressourcen erhöht werden,
- Evidenz-basierte Information mit Aufklärung über die Konsequenzen des Testens, der definitiven diagnostischen Untersuchung(en) und der Behandlung ist den potentiellen Teilnehmern zur Verfügung zu stellen, um sie bei ihrer informierten Entscheidungsfindung zu unterstützen.

- Ein Druck der Öffentlichkeit auf Verkürzung des Screeningintervalls und Erhöhung der Sensitivität des Testvorgangs ist zu erwarten. Entscheidungen über diese Parameter sind gegenüber der Öffentlichkeit wissenschaftlich zu rechtfertigen.

### **3 Methodik der U.S. Preventive Services Task Force**

Die U. S. American Preventive Services Task Force (USPSTF) wurde von dem U.S. Public Health Service (Gesundheitsverwaltung der U.S.-Bundesregierung) zur Beurteilung von Screening-/Präventionsprogrammen und anderen Leistungen für die öffentliche Gesundheit wie Impfprogramme und Chemoprophylaxe sowie Beratung einberufen. Mittlerweile ist die dritte Edition des Guide to Clinical Preventive Services, die periodische Updates von Empfehlungen für Screeningmaßnahmen bzw. Erweiterungen der bestehenden Empfehlungen beinhaltet, verfügbar.

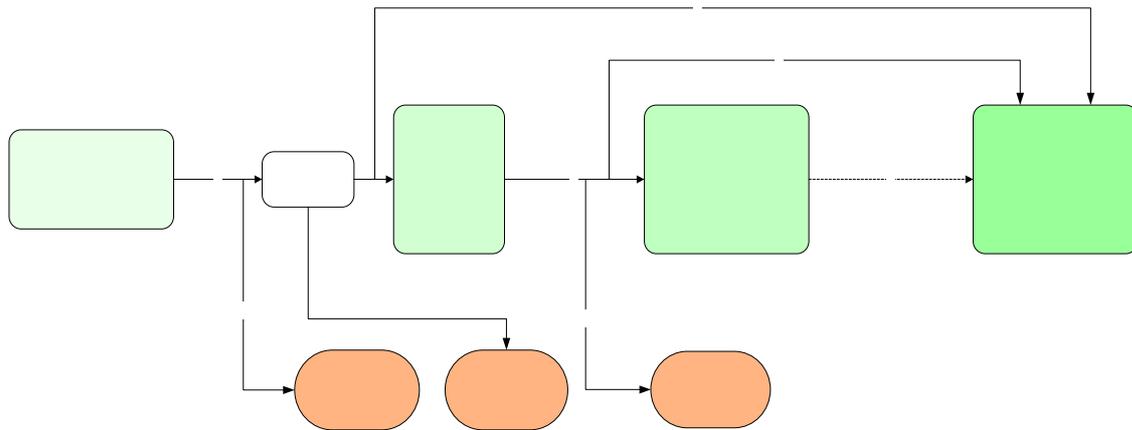
Diese dritte Ausgabe wendet die folgende aktualisierte Methodologie zur Beurteilung von Screeningprogrammen an:<sup>v</sup>

Das analytische Framework (siehe Graphik 1), ursprünglich der *Causal Pathway* in der zweiten Ausgabe<sup>vi</sup>, ermöglicht die Beurteilung der Effektivität eines Screeningprogramms trotz fehlender direkter Evidenz (1).

Das analytische Framework ist eine graphische Darstellung von Zielpopulation (2), Screeningtest, *Target Condition* (das heißt Risikofaktoren der Zielkrankheit, Zielkrankheit im asymptomatischen bzw. im Frühstadium), Diagnostik und Behandlung sowie zwischenzeitlichem Gesundheitsnutzen (*intermediate outcome*) und finalem Gesundheitsnutzen (*final outcome*; Reduktion von Morbidität, Mortalität der Zielkrankheit), die mittels Schlüsselfragen verknüpft sind.

Diese Schlüsselfragen beziehen sich auf Qualitäts-/Eignungskriterien des Screeningtests (3), den mit dem Screeningtest verbundenen Schaden (7), die Wirksamkeit der Behandlung (4 und 5) sowie den mit der Behandlung verbundenen Schaden (8) und den Zusammenhang von zwischenzeitlichem und finalem Gesundheitsnutzen (6).

### 3.1 Grafik 1: Das analytische Framework



Die Zahlen in der Graphik 1 beziehen sich auf folgende acht Schlüsselfragen:<sup>vii</sup>

1. Gibt es direkte Beweise, dass Screening Morbidität und/oder Mortalität der Krankheit reduziert?

2. Welche Prävalenz hat die Krankheit in der Zielgruppe? Kann eine Hochrisikogruppe verlässlich identifiziert werden?

3. Kann der Screeningtest die gesuchte Krankheit bzw. das Gesundheitsproblem exakt erkennen (Test-Validität)?

- Wie lautet die Sensitivität und Spezifität des Screeningtests?
- Variieren die Testergebnisse je nach durchführender Person signifikant (Reliabilität)?
- Können durch das Screeningprogramm Patienten früher identifiziert und behandelt werden?

4. Erbringt die Behandlung einen zwischenzeitlichen Gesundheitsnutzen?

- Ist die Behandlung unter idealen Bedingungen wirksam?
- Wie verhält sich dazu die Wirksamkeit der Behandlung unter Alltagsbedingungen?

5. Gibt es eine Behandlung für den klinisch diagnostizierten Patienten, welche bewiesenermaßen die Krankheitsprognose verbessert?

- Wie ähnlich ist der klinisch diagnostizierte Fall dem mittels Screening entdeckten Fall? Haben Personen, die beim Screening entdeckt werden, einen größeren Gesundheitsnutzen als jene, die klinisch diagnostiziert werden (Nutzensgewinn durch Frühtherapie)?

6. Ist der zwischenzeitliche Gesundheitsnutzen zuverlässig mit der Morbiditäts- und/oder Mortalitätsenkung der Zielkrankheit assoziiert (finaler Gesundheitsnutzen)?

7. Treten beim Screening unerwünschte Nebenwirkungen auf?

- Ist der Screeningtest für die Zielbevölkerung annehmbar?

- Was sind die möglichen Schäden des Screeningtests, und wie häufig treten sie auf?

8. Hat die Behandlung unerwünschte Nebenwirkungen?

Diese Schlüsselfragen werden mit Hilfe der nach umfangreichen systematischen Recherchen gefundenen Publikationen (= Evidenz) beantwortet. Die Qualität der gefundenen Publikationen (= Evidenz) für jede der Schlüsselfragen wird nach einem Drei-Kategorien-Schema bewertet (gut/befriedigend/schwach; für Details über Evaluierung der Evidenzqualität siehe Harris, 2001). Die logische kausal zusammenhängende Kette dieser bewerteten Einzelevidenzen ergibt die (indirekte) „Gesamt“evidenz. Die Gesamtevidenz ermöglicht nun Aussagen, ob ein bestimmtes Screening nachweislich zu einem Gesundheitsnutzen geführt hat, und wie verlässlich die Qualität der Datenlage derzeit ist.

Es mag auch sein, dass (derzeit noch) Evidenz fehlt oder von minderer, inakzeptabler Qualität oder widersprüchlich ist (= schwache Evidenz).

Die Qualität der Gesamtevidenz, wie sie im analytischen Framework ermittelt wird, fließt im nächsten Schritt in die Entscheidung für einen Empfehlungsgrad (*strenght of recommendation*) ein. Der Qualität der Evidenz wird zu diesem Zweck mit einem weiteren Faktor kombiniert. Dieser weitere Faktor bewertet das Ausmaß des Nettonutzens der Intervention. Es wird das Ausmaß des möglichen Nutzens mit dem Ausmaß des möglichen Schadens in Beziehung gebracht, und der daraus resultierende Nettonutzen in vier Stufen – substantiell, moderat, gering, null/negativ – diskutiert Die abschließende Gesamtempfehlung wird in einer der fünf Klassen A – D sowie I als Sonderklasse ausgedrückt (siehe Tabelle 1).

<b>Tabelle 1 Empfehlungsgrad für eine Intervention bei einem Gesundheitsproblem (<i>condition</i>)</b>				
<b>Stärke/Qualität der Evidenz</b>	<b>Stärke (Ausmaß) des Nettonutzens</b>			
	substantiell	moderat	gering	null/negativ
gut	A	B	C	D
befriedigend	B	B	C	D
schwach = I	I	I	I	I

Tabelle 1. Empfehlungsgrad: Aus der Evidenzstärke (vertikal) und dem Ausmaß des Nettonutzens (horizontal) ergibt sich der Grad der Gesamtempfehlung für ein bestimmtes Gesundheitsproblem.

## 3.2 Definition der Empfehlungsgrade

**A** Empfehlung: Die USPSTF spricht eine starke Empfehlung für die Einführung eines Screening-/Präventionsprogramms aus. Es wurde gute Evidenz für den Gesundheitsnutzen gefunden und der Nutzen überwiegt substantiell den Schaden (substantieller Nettonutzen).

**B** Empfehlung: Die USPSTF spricht eine Empfehlung für die Einführung eines Screening-/Präventionsprogramms aus. Es wurde mindestens befriedigende Evidenz für den Gesundheitsnutzen gefunden und der Nutzen überwiegt entweder substantiell oder moderat den Schaden (substantieller oder moderater Nettonutzen).

**C** Empfehlung: Die USPSTF gibt weder eine Empfehlung für noch gegen die Einführung eines Screening-/Präventionsprogramms ab. Es wurde mindestens befriedigende Evidenz für den Gesundheitsnutzen gefunden, aber die Nutzen-Schaden-Bilanz ist zu ungünstig (Nettonutzen gering), um eine allgemeine Empfehlung zu rechtfertigen.

**D** Empfehlung: Die USPSTF spricht eine Empfehlung gegen die Einführung eines Screening-/Präventionsprogramms aus. Es wurde mindestens befriedigende Evidenz für einen fehlenden Gesundheitsnutzen gefunden, oder der Schaden überwiegt den Nutzen (Null- oder Negativ-Nettonutzen).

**I** Empfehlung: Die USPSTF stellt fest, dass die Evidenz unzureichend ist, um eine Empfehlung für oder gegen die Einführung eines Screening-/Präventionsprogramms auszusprechen. Es fehlt die Evidenz für bzw. gegen den Gesundheitsnutzen, die Evidenz ist von minderer Qualität oder widersprüchlich und der Nettonutzen kann nicht bestimmt werden.

---

<sup>i</sup> First Report of the National Screening Committee, Second Report of the UK National Screening Committee; Health Departments of the United Kingdom, NSC Criteria. The criteria for appraising the viability, effectiveness and appropriateness of a screening programme, <http://www.cancerscreening.nhs.uk/>

<sup>ii</sup> Screening/ Case finding; National Service Frameworks: A practical aid to implementation in primary care August 2002

<sup>iii</sup> Harris RP et al: „Current methods of the U.S. Preventive Services Task Force, USPSTF: A review of the process“, Am J Prev Med 2001; 20 (3S): 21-35, <http://www.wahrq.gov/clinic/ajpmsuppl/harris1.htm>, <http://www.ahcpr.gov/clinic/uspstmeth.htm>

<sup>iv</sup> Canadian Task Force on Preventive Health Care: CTFPHC History/ Methodology, <http://www.ctfphc.org>

---

<sup>v</sup> Introducing the Third U.S. Preventive Services Task Force. Article originally in Am J Prev Med 2001; 20 (3S): 3-4. Agency for Healthcare Research and Quality, Rockville, MD. <http://www.ahrq.gov/clinic/ajpmsuppl/berg.htm>

<sup>vi</sup> U.S. Preventive Services Task Force. Guide to Clinical Preventive Services, 2nd ed. Washington, DC: Office of Disease Prevention and Health Promotion, U.S. Government Printing Office, 1996

<sup>vii</sup> Harris RP et al: „Current methods of the U.S. Preventive Services Task Force, USPSTF: A review of the process“ Am J Prev Med 2001; 20 (3S): 21-35, Chapter: Setting the Focus for Admissible Evidence Analytic Framework