

**Bewertung des Nutzens einer  
Früherkennungs-  
untersuchung für Personen  
unter 55 Jahren mit  
familiärem Darmkrebsrisiko**

**Vorläufiger Berichtsplan**

Auftrag: S11-01  
Version: 1.0  
Stand: 10.10.2011

# Impressum

**Herausgeber:**

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

**Thema:**

Bewertung des Nutzens einer Früherkennungsuntersuchung für Personen unter 55 Jahren mit familiärem Darmkrebsrisiko

**Auftraggeber:**

Gemeinsamer Bundesausschuss

**Datum des Auftrags:**

17.03.2011

**Interne Auftragsnummer:**

S11-01

**Anschrift des Herausgebers:**

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen  
Dillenburger Str. 27  
51105 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

Berichte@iqwig.de

[www.iqwig.de](http://www.iqwig.de)

Bei dem vorliegenden Berichtsplan handelt es sich um eine vorläufige Version. Zu diesem Berichtsplan können Stellungnahmen abgegeben werden, die zu einer Ergänzung und / oder Überarbeitung des Berichtsplans führen können. Die Frist für den Eingang der Stellungnahmen befindet sich auf der Website des IQWiG ([www.iqwig.de](http://www.iqwig.de)), ebenso wie die dafür notwendigen Formblätter und ein Leitfaden.

**Schlagwörter:** kolorektale Tumoren, familiäres Darmkrebsrisiko, Anamneseerhebung, Screening, systematische Übersicht

**Keywords:** Colorectal Neoplasms, Family History, Medical History Taking, Mass Screening, Systematic Review

# Inhaltsverzeichnis

|  | Seite      |
|--|------------|
| <b>Tabellenverzeichnis</b> .....   | <b>v</b>   |
| <b>Abbildungsverzeichnis</b> .....   | <b>vi</b>  |
| <b>Abkürzungsverzeichnis</b> .....   | <b>vii</b> |
| <b>1 Hintergrund</b> .....   | <b>1</b>   |
| <b>2 Ziele der Untersuchung</b> .....  | <b>9</b>   |
| <b>3 Projektbearbeitung</b> .....  | <b>14</b>  |
| <b>3.1 Zeitlicher Verlauf des Projekts</b> .....   | <b>14</b>  |
| <b>4 Methoden</b> .....  | <b>15</b>  |
| <b>4.1 Kriterien für den Einschluss von Studien in die Untersuchung</b> .....  | <b>15</b>  |
| 4.1.1 Teilziel A: Bestimmung des Erkrankungsrisikos bei positiver<br>Familienanamnese (Prognose).....  | 15         |
| 4.1.1.1 Population .....   | 15         |
| 4.1.1.2 Prognostische Merkmale.....  | 15         |
| 4.1.1.3 Patientenrelevante Endpunkte.....  | 15         |
| 4.1.1.4 Studientypen .....   | 15         |
| 4.1.1.5 Sonstige Studiencharakteristika.....   | 15         |
| 4.1.2 Teilziel B: Bewertung der diagnostischen Güte anamnestischer Instrumente .....   | 16         |
| 4.1.2.1 Population .....   | 16         |
| 4.1.2.2 Prüfinstrument- und Referenztest .....   | 16         |
| 4.1.2.3 Gütemaße .....   | 16         |
| 4.1.2.4 Studientypen .....   | 16         |
| 4.1.2.5 Sonstige Studiencharakteristika.....   | 16         |
| 4.1.3 Teilziel C: Vergleichende Nutzenbewertung umfassender Screeningstrategien<br>unter Einbeziehung anamnestischer Instrumente .....   | 17         |
| 4.1.3.1 Population .....   | 17         |
| 4.1.3.2 Prüf- und Vergleichsintervention.....  | 17         |
| 4.1.3.3 Patientenrelevante Endpunkte.....  | 17         |
| 4.1.3.4 Studientypen .....   | 18         |
| 4.1.4 Teilziel D: Vergleichende Nutzenbewertung verschiedener<br>Screeningmaßnahmen im Rahmen einer Screeningstrategie bei Personen mit<br>erhöhtem Risiko aufgrund einer positiven Familienanamnese ..... | 19         |
| 4.1.4.1 Population .....   | 19         |
| 4.1.4.2 Prüf- und Vergleichsintervention.....  | 19         |

|            |  |           |
|------------|--|-----------|
| 4.1.4.3    | Patientenrelevante Endpunkte.....  | 19        |
| 4.1.4.4    | Studientypen .....   | 19        |
| 4.1.5      | Tabellarische Übersicht über die Kriterien für den Studieneinschluss .....             | 21        |
| 4.1.6      | Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen..... | 22        |
| <b>4.2</b> | <b>Informationsbeschaffung.....</b>  | <b>22</b> |
| 4.2.1      | Bibliografische Literaturrecherche .....   | 22        |
| 4.2.2      | Suche nach weiteren publizierten und nicht publizierten Studien.....                   | 23        |
| 4.2.3      | Selektion relevanter Studien.....  | 23        |
| <b>4.3</b> | <b>Informationsbewertung .....</b>   | <b>23</b> |
| 4.3.1      | Bewertung von Studien zum Erkrankungsrisiko (Teilziel A) .....                         | 24        |
| 4.3.2      | Bewertung von Studien zur diagnostischen Güte (Teilziel B) .....                       | 24        |
| 4.3.3      | Bewertung von Studien zum patientenrelevanten Nutzen (Teilziel C und D) .....          | 24        |
| <b>4.4</b> | <b>Informationssynthese und -analyse .....</b>   | <b>25</b> |
| 4.4.1      | Gegenüberstellung der Ergebnisse der Einzelstudien .....                               | 26        |
| 4.4.2      | Meta-Analysen .....  | 26        |
| 4.4.3      | Sensitivitätsanalyse.....  | 28        |
| 4.4.4      | Subgruppenmerkmale und andere Effektmodifikatoren .....                                | 28        |
| <b>5</b>   | <b>Literaturverzeichnis .....</b>  | <b>30</b> |

**Tabellenverzeichnis**

|  | <b>Seite</b> |
|--|--------------|
| Tabelle 1: Erkrankungs- und Sterberisiko in Deutschland nach Alter und Geschlecht,<br>ICD-10 C18–C21 ..... | 2            |
| Tabelle 2: Klassifikationsschema TNM .....   | 4            |
| Tabelle 3: Stadiengruppierung .....  | 5            |
| Tabelle 4: Prüf- und Vergleichsintervention für Teilziel C .....   | 17           |
| Tabelle 5: Übersicht über die Kriterien für den Studieneinschluss.....                                     | 21           |

**Abbildungsverzeichnis**

|  | <b>Seite</b> |
|--|--------------|
| Abbildung 1: Allgemeine Effekte und Nutzen von Krebsfrüherkennungsprogrammen.....                    | 6            |
| Abbildung 2: Studiendesign zur Erfassung einer Wechselwirkung.....                                   | 10           |
| Abbildung 3: Kontrollierte Interventionsstudie zur Evaluation eines anamnestischen Tests..           | 11           |
| Abbildung 4: Kontrollierte Interventionsstudie zur Evaluation einer anamnestischen<br>Strategie..... | 12           |
| Abbildung 5: Studiendesign mit Testergebnis als Einschlusskriterium.....                             | 13           |

**Abkürzungsverzeichnis**

| <b>Abkürzung</b> | <b>Bedeutung</b>  |
|------------------|---|
| BMI              | Body-Mass-Index   |
| CCT              | Controlled Clinical Trial   |
| CONSORT          | Consolidated Standards of Reporting Trials  |
| DNS              | Desoxyribonukleinsäure  |
| FAP              | Familiäre adenomatöse Polyposis   |
| FOBT             | Fäkaler Okkultbluttest  |
| G-BA             | Gemeinsamer Bundesausschuss   |
| HNPCC            | Hereditäres nicht polypöses kolorektales Karzinom   |
| ICD-10           | International Statistical Classification of Diseases and Related Health Problems, 10th Revision |
| IQWiG            | Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen                                |
| ITT              | Intention-to-Treat  |
| HSROC            | Hierarchical Summary Receiver Operating Curve   |
| KRK              | Kolorektales Karzinom   |
| MID              | Minimal Important Difference  |
| QUADAS           | Quality Assessment of Diagnostic Accuracy Studies   |
| RCT              | Randomized Controlled Trial   |
| RKI              | Robert Koch-Institut  |
| SMD              | Standardised Mean Difference  |
| STARD            | Standards for the Reporting of Diagnostic Accuracy Studies                                      |
| STROBE           | Strengthening the Reporting of Observational Studies in Epidemiology                            |
| TREND            | Transparent Reporting of Evaluations with Nonrandomized Designs                                 |
| UICC             | International Union Against Cancer  |

## 1 Hintergrund

### Definition des Krankheitsbildes

Unter Darmkrebs werden Krebserkrankungen des Dickdarms (Kolon), des Mastdarms (Rektum) und des Darmausgangs (Anus) zusammengefasst. In den meisten Fällen handelt es sich um Neubildungen, die sich aus der Schleimhaut von Dick- oder Mastdarm entwickeln (Adenokarzinome) [1,2]. In der 10. Ausgabe der Internationalen Statistischen Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (ICD-10) wird Darmkrebs als bösartige Neubildung folgender Organe eingeteilt: Kolon (C18), rektosigmoidaler Übergang (C19), Rektum (C20) sowie Anus und Analkanal (C21) [3]. Im engeren Sinn umfasst das kolorektale Karzinom (KRK) nur die Einträge C18 bis C20. Bei der Erfassung von Inzidenz- und Mortalitätsraten wird jedoch vielfach der Eintrag C21 mit eingeschlossen [4].

### Epidemiologie und Krankheitslast

In Deutschland ist das KRK bei beiden Geschlechtern sowohl die zweithäufigste Krebserkrankung als auch die zweithäufigste krebsbedingte Todesursache [1]. Nach Angaben des Robert Koch-Instituts (RKI) lag die Zahl der Neuerkrankungen in Deutschland 2006 für Männer bei 36 300 und für Frauen bei 32 440 [1]; für das Jahr 2010 schätzte das RKI 39 410 bzw. 33 620 Erkrankungsfälle. Im Jahr 2006 verursachte das KRK bei Männern 13 756 und bei Frauen 13 469 Todesfälle [1].

Wie auch weltweit [5] so liegt in Deutschland die jährliche Inzidenz bei Männern höher als bei Frauen (90,0 vs. 77,1 pro 100 000 Einwohner) und das Lebenszeitrisiko, an einem KRK zu erkranken, ist ebenfalls geringfügig höher (Männer: 7,7 % vs. Frauen: 6,5 %) [1]. Männer erkranken im Mittel mit 69, Frauen mit 75 Jahren [6]. Die geschätzten altersstandardisierten Inzidenzraten von Frauen und Männern sind bis etwa Mitte bzw. Ende der 1990er-Jahre im Vergleich zu 1980 um 34 % bzw. 40 % gestiegen und seitdem (bis 2006) leicht gesunken. Die altersstandardisierten Sterberaten für Darmkrebs waren bis zum Beginn der 1990er-Jahre weitgehend konstant und sind seitdem bis 2006 um 24 % bei Männern und um 33 % bei Frauen gesunken. Die 5-Jahres-Überlebensrate (Daten aus dem Jahr 2006) nach Erkrankung wird in Deutschland auf 53 bis 63 % beziffert [1].

Es gibt große internationale Unterschiede im Auftreten von Darmkrebs [5,7]: Die Inzidenzen variieren um das 10-Fache; die höchsten Inzidenzraten werden für westliche Industrieregionen (Europa, Nordamerika, Australien / Neuseeland) beschrieben, die niedrigsten für viele Länder Afrikas und Asiens [5]. Während die Inzidenzraten in den meisten Ländern der westlichen Industrieregionen seit Langem vergleichsweise stabil sind, sind sie insbesondere in einigen asiatischen Ländern, z. B. Japan und Singapur, in den letzten Jahren stark angestiegen [4]. Unterschiede in der wirtschaftlichen Entwicklung, Urbanisierung und Industrialisierung und damit verbundene Änderungen der Ernährungsweise und des Lebensstils werden als Gründe für die geografischen Unterschiede und die jüngeren Veränderungen angeführt [4,7,8].

## Ursache der Erkrankung

Für die Mehrzahl der betroffenen KRK-Patienten ist die genaue Ursache ihrer Erkrankung unklar, es lassen sich jedoch bestimmte Risikofaktoren und Risikogruppen beschreiben, von denen im Folgenden eine Auswahl dargestellt wird.

Der bedeutendste Risikofaktor für das KRK ist das Alter. Mit zunehmendem Alter nimmt das Risiko, an einem KRK zu erkranken, zu. So steigt in Deutschland das Risiko, in den nächsten 10 Jahren an einem KRK zu erkranken, bei Männern zwischen 40 und 70 Jahren von 0,2 auf 3,7 %, bei Frauen von 0,2 auf 2,4 % [1] (siehe auch Tabelle 1).

Tabelle 1: Erkrankungs- und Sterberisiko in Deutschland nach Alter und Geschlecht, ICD-10 C18–C21 (entnommen aus „Krebs in Deutschland 2005 / 2006“ des RKI [1])

| <b>Männer</b>    |                                  | <b>Erkrankungsrisiko</b> |               |            |                                  | <b>Sterberisiko</b> |               |            |            |
|------------------|----------------------------------|--------------------------|---------------|------------|----------------------------------|---------------------|---------------|------------|------------|
| im Alter von     | <b>in den nächsten 10 Jahren</b> |                          | <b>jemals</b> |            | <b>in den nächsten 10 Jahren</b> |                     | <b>jemals</b> |            |            |
|                  | 40 Jahren                        | 0,2%                     | (1 von 480)   | 7,8%       | (1 von 13)                       | 0,1%                | (1 von 1800)  | 3,3%       | (1 von 30) |
| 50 Jahren        | 0,9%                             | (1 von 120)              | 7,8%          | (1 von 13) | 0,3%                             | (1 von 390)         | 3,4%          | (1 von 30) |            |
| 60 Jahren        | 2,3%                             | (1 von 44)               | 7,5%          | (1 von 13) | 0,7%                             | (1 von 140)         | 3,3%          | (1 von 30) |            |
| 70 Jahren        | 3,7%                             | (1 von 27)               | 6,3%          | (1 von 16) | 1,4%                             | (1 von 70)          | 3,1%          | (1 von 32) |            |
| Lebenszeitrisiko |                                  |                          | 7,7%          | (1 von 13) |                                  |                     |               | 3,3%       | (1 von 31) |
| <b>Frauen</b>    |                                  | <b>Erkrankungsrisiko</b> |               |            |                                  | <b>Sterberisiko</b> |               |            |            |
| im Alter von     | <b>in den nächsten 10 Jahren</b> |                          | <b>jemals</b> |            | <b>in den nächsten 10 Jahren</b> |                     | <b>jemals</b> |            |            |
|                  | 40 Jahren                        | 0,2%                     | (1 von 560)   | 6,5%       | (1 von 15)                       | < 0,1%              | (1 von 2300)  | 2,9%       | (1 von 34) |
| 50 Jahren        | 0,6%                             | (1 von 160)              | 6,4%          | (1 von 16) | 0,2%                             | (1 von 650)         | 2,9%          | (1 von 34) |            |
| 60 Jahren        | 1,5%                             | (1 von 67)               | 6,0%          | (1 von 17) | 0,4%                             | (1 von 260)         | 2,9%          | (1 von 35) |            |
| 70 Jahren        | 2,4%                             | (1 von 42)               | 5,0%          | (1 von 20) | 0,9%                             | (1 von 110)         | 2,7%          | (1 von 37) |            |
| Lebenszeitrisiko |                                  |                          | 6,5%          | (1 von 15) |                                  |                     |               | 2,9%       | (1 von 34) |

In ca. 70 % der Fälle tritt das KRK sporadisch auf, d. h. ohne dass eine familiäre Häufung oder hereditäre Vorbelastung festzustellen ist. Die restlichen ca. 30 % der KRK-Erkrankungen lassen sich mit einer hereditären oder familiären Vorbelastung in Verbindung bringen [9]. Bei bis zu 5 % handelt es sich um monogen erbliche („hereditäre“) Formen, bei denen Keimbahnmutationen gut charakterisierte genetische Syndrome wie z. B. das hereditäre kolorektale Karzinom ohne Polyposis (HNPCC, Hereditary Non-Polyposis Colorectal Cancer, Lynch-Syndrom) oder die familiäre adenomatöse Polyposis (FAP) hervorrufen. Bis zu 25 % der KRK-Fälle treten „familiär gehäuft“ auf, d. h. es kann ein gehäuftes Vorkommen der Krebserkrankung innerhalb einer Familie beobachtet werden, ohne dass jedoch konkrete genetische Ursachen identifiziert sind [9]. Generell geht man davon aus, dass das Erkrankungsrisiko weiterer Familienangehöriger vom Grad der Verwandtschaft sowie von der Anzahl der betroffenen Verwandten und deren Alter zum Zeitpunkt der Erkrankung abhängt [9-11].

Es wird angenommen, dass Ernährung und Lebensstil eine bedeutsame Rolle bei der Verursachung von Darmkrebserkrankungen spielen [7,8]. Allerdings sind viele Zusammenhänge nicht abschließend geklärt [7]. Insbesondere wird davon ausgegangen, dass der häufige Verzehr von rotem und verarbeitetem Fleisch, Rauchen, übermäßiger Alkoholkonsum und starkes Übergewicht das Risiko für KRK erhöhen, wohingegen körperliche Aktivität das Risiko senken soll [7,8]. Darüber hinaus wird ein möglicher protektiver Effekt durch den gesteigerten Verzehr von Ballaststoffen, Obst oder Gemüse diskutiert [7,8].

Die chronisch entzündlichen Darmerkrankungen Colitis ulcerosa und Morbus Crohn sind Risikoerkrankungen bezüglich der Entwicklung eines KRK [12]. Bei Colitis ulcerosa korreliert das Risiko der Karzinomentwicklung eng mit der Dauer, der Ausdehnung und der Intensität der Entzündung [12,13]. Auch bei Patienten mit Morbus Crohn muss von einem erhöhten Risiko ausgegangen werden, allerdings ist hier die Datenlage noch nicht so eindeutig [12,13].

### **Verlauf der Erkrankung (natürlicher Verlauf)**

Die meisten kolorektalen Karzinome sind Adenokarzinome und entstehen aus gutartigen Neubildungen (Adenomen) der Dickdarmschleimhaut (sogenannte Adenom-Karzinom-Sequenz). Allerdings entwickelt sich nur ein kleiner Prozentsatz dieser gutartigen Neubildungen zu einem invasiven Karzinom [14,15]. Das größte Risiko für eine Entartung haben fortgeschrittene Adenome [16]. Hierunter versteht man alle Adenome  $\geq 1$  cm Größe sowie auch Adenome  $< 1$  cm, wenn sie hochgradige Dysplasien oder eine villöse / tubulovillöse Architektur aufweisen (z. B. [16]).

Die Symptome des KRK sind oft unspezifisch und abhängig von der anatomischen Lage des Tumors [17,18]; bei einigen Patienten treten Symptome erst im weit fortgeschrittenen Krankheitsstadium auf [18]. Nur ein Teil aller Erkrankungen manifestiert sich in einem frühen, chirurgisch gut heilbaren Stadium [19]. Die häufigsten Symptome des KRK oder von großen Polypen sind rektale Blutungen, anhaltende Änderungen der Stuhlgewohnheiten und Anämie. Bei fortgeschrittenen Tumoren kommt es häufig zu Gewichtsverlust, Übelkeit und Appetitlosigkeit sowie abdominellen Schmerzen [18]. Diese Schmerzen können durch einen Darmverschluss, eine Peritonealkarzinose oder eine Darmperforation, die zu einer Peritonitis führt, verursacht werden [20]. Das KRK kann sich im Verlauf der Erkrankung auf andere Organe ausbreiten. Metastasen bilden sich häufig in den lokoregionären Lymphknoten, im Peritoneum, in der Leber und in der Lunge [20].

### **Klassifikationsschemata**

Das KRK wird, wie andere maligne Tumoren, nach der TNM-Klassifikation eingeteilt (Tabelle 2). Erfasst werden von der Klassifikation die Größe und Ausdehnung des Primärtumors (T), das Vorhandensein von regionären Lymphknotenmetastasen (N von Nodi lymphatici) und das Vorhandensein von Fernmetastasen (M) [18]:

Tabelle 2: Klassifikationsschema TNM nach [21]

|          |   |
|----------|---|
| <b>T</b> | <b>Primärtumor (bezieht sich hier auf die Lokalisation des KRK):</b>  |
| TX       | Primärtumor kann nicht beurteilt werden   |
| T0       | Kein Anhalt für Primärtumor   |
| Tis      | Carcinoma in situ: intraepithelial oder Infiltration der Lamina propria   |
| T1       | Tumor infiltriert Submukosa   |
| T2       | Tumor infiltriert Muscularis propria  |
| T3       | Tumor infiltriert durch die Muscularis propria in die Subserosa oder in nicht peritonealisiertes perikolisches oder perirektales Gewebe |
| T4       | Tumor infiltriert direkt in andere Organe oder Strukturen und / oder perforiert das viszerale Peritoneum                                |
| <b>N</b> | <b>Regionäre Lymphknoten:</b>   |
| NX       | Regionäre Lymphknoten können nicht beurteilt werden   |
| N0       | Keine regionären Lymphknotenmetastasen  |
| N1       | Metastase / n in 1 bis 3 regionären Lymphknoten   |
| N2       | Metastasen in 4 oder mehr regionären Lymphknoten  |
| <b>M</b> | <b>Fernmetastasen:</b>  |
| M0       | Keine Fernmetastasen  |
| M1       | Fernmetastasen  |

Es existieren verschiedene Systeme zur Einteilung der Stadien; die Stadiengruppierung nach der UICC und deren Entsprechungen mit der Klassifikation nach Dukes sind in Tabelle 3 zusammengefasst. Die entsprechende Einstufung einer Tumorerkrankung erlaubt prognostische Aussagen und bestimmt häufig auch die weitere Therapie.

Tabelle 3: Stadiengruppierung nach [22]

| Stadium | T       | N        | M   | Dukes |
|---------|---------|----------|-----|-------|
| 0       | Tis     | N0       | M0  | -     |
| I       | T1      | N0       | M0  | A     |
|         | T2      | N0       | M0  | A     |
| IIA     | T3      | N0       | M0  | B     |
| IIB     | T4a     | N0       | M0  | B     |
| IIC     | T4b     | N0       | M0  | B     |
| IIIA    | T1–T2   | N1 / N1c | M0  | C     |
|         | T1      | N2a      | M0  | C     |
| IIIB    | T3–T4a  | N1 / N1c | M0  | C     |
|         | T2–T3   | N2a      | M0  | C     |
|         | T1–T2   | N2b      | M0  | C     |
| IIIC    | T4a     | N2a      | M0  | C     |
|         | T3–T4a  | N2b      | M0  | C     |
|         | T4b     | N1–N2    | M0  | C     |
| IVA     | Jedes T | Jedes N  | M1a | -     |
| IVB     | Jedes T | Jedes N  | M1b | -     |

### Allgemeine Effekte und Nutzen von Krebsfrüherkennungsprogrammen

Üblicherweise wird für die Bewertung von Methoden zur Krebsfrüherkennung der Nachweis von Effekten auf die krankheitsspezifische Mortalität, idealerweise auch auf die Gesamtsterblichkeit, gefordert [23]. Wie in Abbildung 1 vereinfacht dargestellt, resultiert eine frühe Identifikation und Therapie einer Tumorerkrankung und / oder ggf. ihrer Vorstufen nicht zwangsläufig in einer Reduktion der betreffenden Morbidität und / oder krankheitsspezifischen Mortalität. Im Rahmen von Krebsfrüherkennungsprogrammen bei symptomfreien Personen werden auch solche frühen Krankheitsstadien identifiziert und therapiert, die sich ohne ein Screening entweder zurückgebildet hätten oder klinisch niemals auffällig geworden wären [24]. Eine Intervention ist in solchen Fällen also überflüssig, hat potenzielle Nebenwirkungen und führt weder zu einer Reduktion der Inzidenz von fortgeschrittenen und damit therapieintensiveren Krankheitsstadien mit i. d. R. schlechterer Prognose noch zu einer Reduktion der krankheitsspezifischen Mortalität.

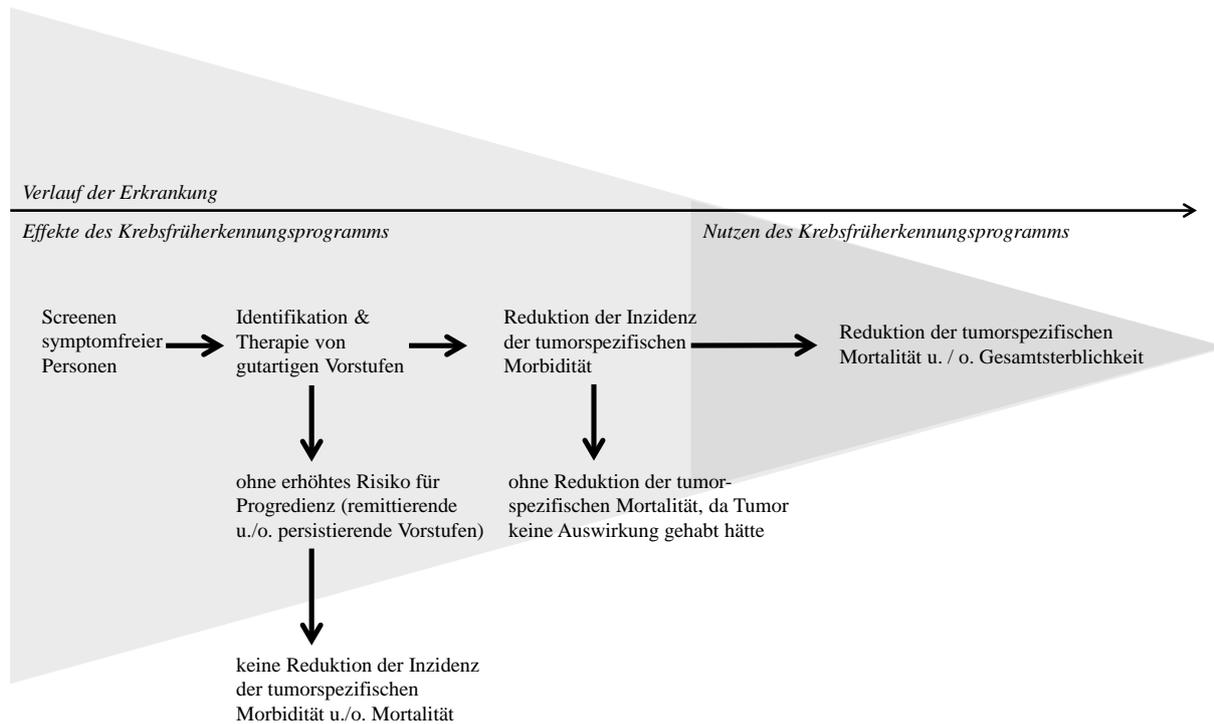


Abbildung 1: Allgemeine Effekte und Nutzen von Krebsfrüherkennungsprogrammen

### Untersuchungsverfahren zur Früherkennung des kolorektalen Karzinoms

Zur Früherkennung kolorektaler Karzinome bei Personen mit durchschnittlichem Risiko werden derzeit nach [12] keine anderen Verfahren als Tests auf verborgenes Blut im Stuhl (Okkultbluttests), Koloskopien („große Darmspiegelung“) und Sigmoidoskopien („kleine Darmspiegelung“) empfohlen.

Bei den in Deutschland als FOBT eingesetzten Guaiak-Tests wird eine Stuhlprobe mittels mit Guaiak imprägnierten Filterpapiers daraufhin untersucht, ob Blut im Stuhl vorhanden ist. Dieser Test beruht auf dem Nachweis der Pseudoperoxidaseaktivität des Hämoglobins [25], stellt somit einen indirekten Test dar. Ein positiver Test kann auf ein KRK hinweisen [12]. Die Genauigkeit des Tests hängt von verschiedenen Faktoren wie z. B. der Anzahl der Proben ab. Im Falle eines positiven Befundes wird Patienten eine endoskopische Untersuchung des gesamten Dickdarms (Koloskopie) empfohlen [12].

Sowohl die Koloskopie als auch die Sigmoidoskopie ist ein endoskopisches Verfahren. Während der Untersuchung können jeweils Gewebeentnahmen zur histopathologischen Abklärung verdächtiger Adenome vorgenommen und Polypen entfernt werden (Polypektomie). Bei der Koloskopie wird der gesamte Dickdarm endoskopisch untersucht. Als Vorbereitung ist eine komplette Darmreinigung nötig [26,27]. In Deutschland erfolgen Koloskopien in der Regel in Analgosedierung [28]. Bei der Sigmoidoskopie werden lediglich etwa die letzten 50 bis 60 Zentimeter des Dickdarms mit einem Endoskop untersucht [27]. Auch hier ist eine Darmreinigung notwendig; sowohl die Vorbereitungszeit als auch die Dauer der

Untersuchung ist kürzer als bei der Koloskopie [27]. Bei positivem Befund wird eine Koloskopie zur weiteren Abklärung der nicht erreichten Dickdarmabschnitte empfohlen [12]. Die möglichen Komplikationen der Koloskopie und Sigmoidoskopie umfassen insbesondere Perforationen, Blutungen und kardiorespiratorische Ereignisse. Insgesamt sind schwerwiegende Komplikationen und solche mit Todesfolge sehr selten [29]. Blutungen und Perforationen im Rahmen von Endoskopien ereignen sich insbesondere im Zusammenhang mit Polypektomien [30].

Weitere derzeit diskutierte Verfahren zur Früherkennung von Darmkrebs sind bspw. immunologische Stuhltestverfahren, bei denen ebenfalls der Stuhl auf verborgenes Blut untersucht wird, molekulare Screeningverfahren zur Identifikation von vom Tumor stammenden DNS-Veränderungen sowie bildgebende Verfahren wie die CT- bzw. MRT-Kolonografie (siehe dazu auch die S3-Leitlinie zum kolorektalen Karzinom) [12]. Im Rahmen der Krebsfrüherkennungsrichtlinien der Gesetzlichen Krankenversicherung werden ausschließlich der FOBT und die Koloskopie erstattet (siehe Abschnitt „Gegenwärtiges Darmkrebscreening im Rahmen des Krebsfrüherkennungsprogramms in Deutschland“).

### **Management kolorektaler Adenome / Polypen**

Im Rahmen der Koloskopie entdeckte Adenome werden noch während der Untersuchung unter Dokumentation der Lokalisation entfernt und für eine histologische Aufbereitung geborgen [12]: In Abhängigkeit von verschiedenen Faktoren kommt dafür üblicherweise entweder die Polypektomie mit der Schlinge oder eine endoskopische Mukosaresektion zum Einsatz. Dabei sollten Polypen immer vollständig entfernt werden. Die entfernten Polypen müssen geborgen und einer histologischen Untersuchung unterzogen werden. Nach Abtragung nicht neoplastischer Polypen, z. B. einzelner hyperplastischer Polypen, ist eine Nachsorge nicht notwendig [12]. Im Gegensatz dazu ist nach der Entfernung von Adenomen immer eine Kontrolluntersuchung nach in der Regel 3 bis 5 Jahren notwendig. Der Zeitpunkt der Kontrolluntersuchung richtet sich dabei nach der Anzahl, Größe und Histologie der entfernten Adenome [12]. Da die Koloskopie keine 100%ige Sensitivität für Polypen und insbesondere flache Läsionen hat [31], besteht auch nach primär unauffälliger Koloskopie und erfolgreicher, kompletter Polypektomie kein 100%iger Schutz vor Darmkrebs. Auch sogenannte Intervallkarzinome sind möglich: Das sind Karzinome, die zwischen 2 Screeninguntersuchungen entdeckt werden, entweder durch andere Untersuchungen oder weil sie durch Symptome auffallen. Intervallkarzinome setzen sich aus zuvor übersehenen oder nicht vollständig entfernten Tumoren zusammen sowie aus Tumoren, die besonders schnell wachsen und sich deshalb innerhalb des Untersuchungsintervalls zum Karzinom entwickeln konnten [31].

### **Therapie des invasiven Kolonkarzinoms**

Das therapeutische Vorgehen bei der Behandlung einer Darmkrebserkrankung hängt insbesondere vom Stadium und von der Lage des Tumors, vom Alter und Zustand des Patienten und, daraus abgeleitet, von der Zielsetzung (kurativ oder palliativ) ab. Die

chirurgische Entfernung des Tumors und ggf. seiner Metastasen ist das regelmäßig angestrebte Ziel in der Therapie des KRK bei der Mehrheit der Patienten. Das operative Vorgehen ist je nach Lage des Tumors unterschiedlich, in jedem Fall schließt die Resektion die Entfernung regionärer Lymphabflussgebiete ein [12]. Die chirurgische Entfernung kann je nach primärem Stadium des Karzinoms, Resektionsergebnis und Tumorlokalisation von adjuvanter Chemotherapie, Radiochemotherapie oder neoadjuvanter Radio- und / oder Radiochemotherapie begleitet werden [12].

### **Gegenwärtiges Darmkrebsscreening im Rahmen des Krebsfrüherkennungsprogramms in Deutschland**

Über die übliche Diagnostik bei einem Tumorverdacht (z. B. aufgrund von Symptomen) hinaus haben gesetzlich Versicherte in Deutschland gemäß den Krebsfrüherkennungsrichtlinien [32] Anspruch auf Früherkennungsuntersuchungen auf Darmkrebs: Diese umfassen für 50- bis 54-Jährige jährlich einen Test auf okkultes Blut im Stuhl. Seit dem Jahr 2002 haben Versicherte beiderlei Geschlechts ab dem Alter von 55 Jahren Anspruch auf insgesamt 2 Koloskopien, wobei die zweite frühestens 10 Jahre nach der ersten durchgeführt wird. Alternativ können sich Versicherte ab 55 Jahren dafür entscheiden, alle 2 Jahre ihren Stuhl auf okkultes Blut testen zu lassen. Ist das Ergebnis positiv, so besteht ein Anspruch auf die Durchführung einer Koloskopie zur Abklärung [32].

Vor diesem Hintergrund soll der vorliegende Bericht dem Gemeinsamen Bundesausschuss Informationen dazu liefern, inwiefern Personen, die einen oder mehrere an Darmkrebs erkrankte Verwandte haben, ein erhöhtes Darmkrebsrisiko aufweisen, wie verlässlich sie durch anamnestische Instrumente wie Fragebögen oder Interviews identifiziert werden können, und ob und in welcher Form Personen mit einem aufgrund einer positiven Familienanamnese erhöhten Darmkrebsrisiko durch die Teilnahme an einem modifizierten Früherkennungsangebot einen patientenrelevanten Nutzen erfahren. Ausdrücklich ausgenommen sind im Folgenden Personen mit hereditären Formen von Darmkrebs.

## 2 Ziele der Untersuchung

Ziel der vorliegenden Untersuchung ist die Bewertung des patientenrelevanten Nutzens einer Darmkrebsfrüherkennungsuntersuchung für Personen unter 55 Jahren mit familiärem Darmkrebsrisiko (ausgenommen hereditäre Erkrankungen). Die Fragestellung des Berichts wurde in mehrere Teilziele aufgegliedert, die im Folgenden dargestellt werden.

### **Teilziel A: Bestimmung des Erkrankungsrisikos bei positiver Familienanamnese**

Die Einführung eines modifizierten Screenings für Personen mit positiver Familienanamnese ist nur dann sinnvoll, wenn diese Personengruppe im Vergleich zur Normalbevölkerung tatsächlich ein erhöhtes Risiko aufweist, an Darmkrebs zu erkranken.

Im Rahmen von Teilziel A wird deshalb für Personen mit positiver Familienanamnese das Risiko, an Darmkrebs zu erkranken oder zu sterben, bestimmt. Dabei soll untersucht werden, welche unterschiedlichen Definitionen von Risikogruppen es gibt (z. B. in Abhängigkeit von der Anzahl der erkrankten Verwandten oder dem Alter der Verwandten zum Zeitpunkt der Erkrankung) und wie hoch ggf. die jeweilige Risikoerhöhung im Vergleich zur Normalbevölkerung ist.

Sollten sich keine Hinweise darauf finden, dass eine positive Familienanamnese mit einem erhöhten Erkrankungsrisiko einhergeht, fehlt die Grundlage für die weitere Bearbeitung der im Folgenden dargestellten Teilziele.

### **Teilziel B: Bewertung der diagnostischen Güte anamnestischer Instrumente**

Voraussetzung dafür, Personen mit erhöhtem Risiko aufgrund einer positiven Familienanamnese ein modifiziertes Screening anbieten zu können, ist, dass diese Personengruppe mit angemessener Zuverlässigkeit innerhalb der Normalbevölkerung identifiziert werden kann. Dafür können zum Beispiel anamnestische Instrumente wie Fragebögen oder Interviews zum Einsatz kommen.

Im Rahmen von Teilziel B wird deshalb untersucht, wie zuverlässig (d. h. mit welcher diagnostischen Güte) anamnestische Instrumente, die im Rahmen eines Screenings eingesetzt werden könnten, Personen mit erhöhtem Risiko aufgrund einer positiven Familienanamnese innerhalb der Normalbevölkerung identifizieren.

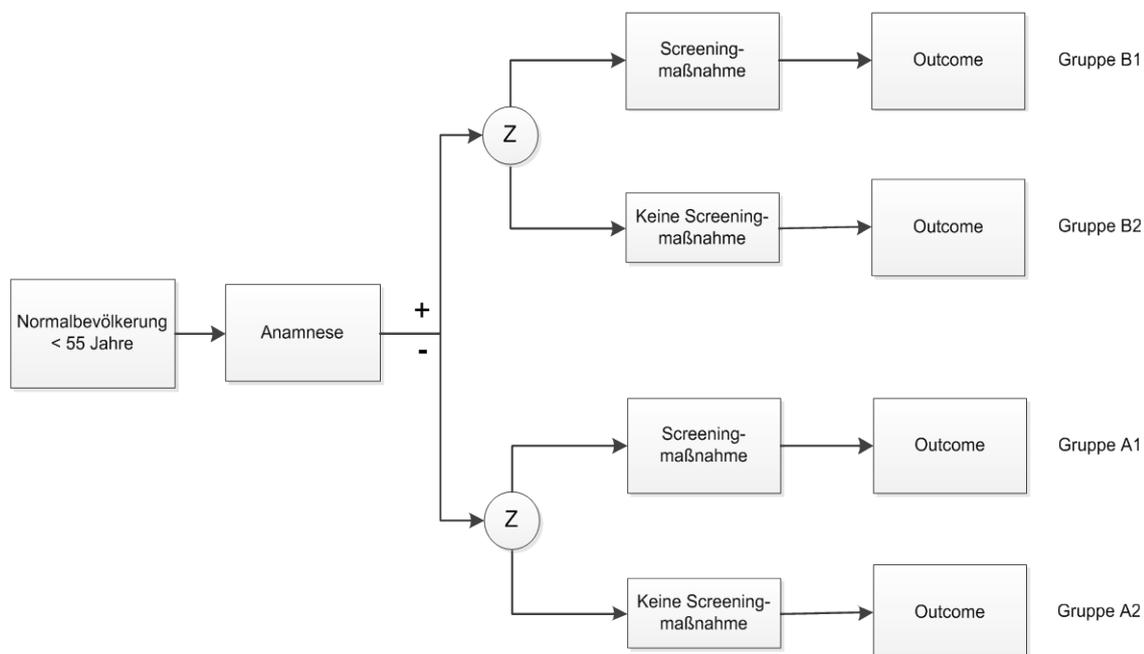
### **Teilziel C: Vergleichende Nutzenbewertung umfassender Screeningstrategien unter Einbeziehung anamnestischer Instrumente**

Im Rahmen von Teilziel C soll die Frage beantwortet werden, welchen Nutzen eine Screeningstrategie hat, bei der in der Normalbevölkerung unter 55 Jahren zunächst ein Anamneseinstrument eingesetzt wird, um Personen mit einem erhöhten Risiko aufgrund einer positiven Familienanamnese zu identifizieren, und anschließend Personen mit einem positiven Testergebnis eine Screeningmaßnahme angeboten wird. Dabei interessiert auch, ob sich

unterschiedliche Screeningstrategien (z. B. durch den Einsatz unterschiedlicher Anamneseinstrumente) in ihrem Nutzen unterscheiden.

Prinzipiell sind verschiedene Studiendesigns denkbar, mit denen diese Frage adressiert werden kann [33]. Im Folgenden wird der Begriff „Screeningstrategie“ verwendet, um eine vollständige diagnostische Kette einschließlich eines anamnestischen Instruments und (bei positiver Anamnese) eines diagnostischen Untersuchungsverfahrens zur Früherkennung von Darmkrebs zu beschreiben. Als „Screeningmaßnahme“ werden im Folgenden die eigentlichen Maßnahmen zur Früherkennung (z. B. Koloskopie, FOBT) bezeichnet, die primär diagnostische Verfahren darstellen, aber im Falle der Koloskopie und Sigmoidoskopie ggf. auch eine therapeutische Komponente (Polypektomie) beinhalten. Eine Screeningstrategie besteht also aus einer oder mehreren Screeningmaßnahmen in Kombination.

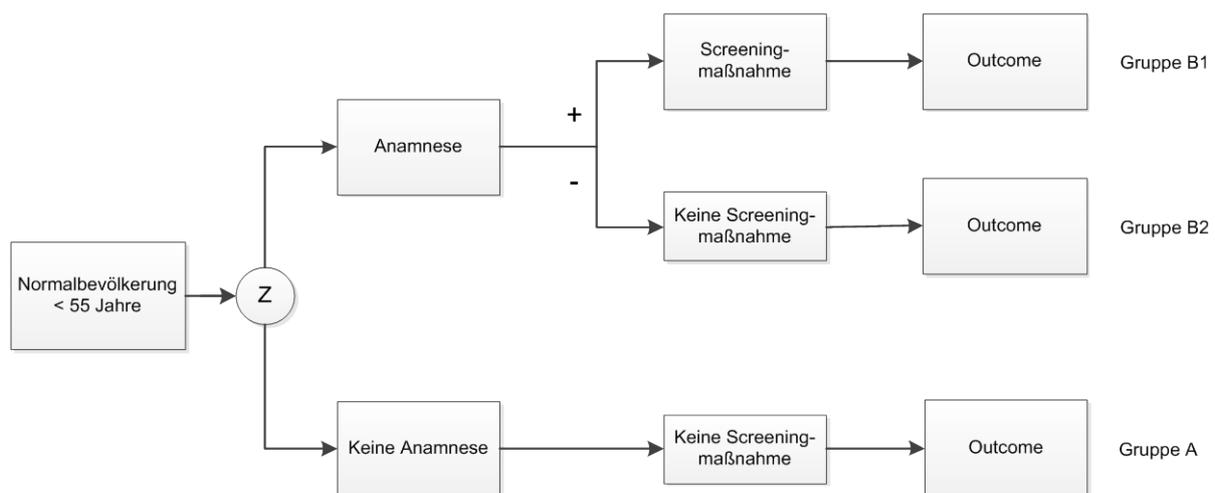
Eine Möglichkeit besteht in einem Studiendesign, bei dem an allen Teilnehmern (Normalbevölkerung) zu Studienbeginn die Familienanamnese erhoben wird und unabhängig vom Testergebnis eine Zuteilung (idealerweise zufällig) zu den Bedingungen „Screeningmaßnahme“ / „keine Screeningmaßnahme“ erfolgt (Abbildung 2). Eine Wechselwirkung zwischen dem Testergebnis und dem Effekt des Screenings läge vor, wenn der Effekt des Screenings bei den Personen mit positiver Familienanamnese (d. h. Vergleich Gruppe B1 vs. B2) größer wäre als bei Personen ohne Familienanamnese (d. h. Vergleich Gruppe A1 vs. A2). Dies würde bedeuten, dass ein positives Testergebnis (d. h. eine positive Familienanamnese) Personen identifiziert, die von der Screeningmaßnahme besonders profitieren.



Z = Zuteilung zu den Interventionen

Abbildung 2: Studiendesign zur Erfassung einer Wechselwirkung (adaptiert nach [33])

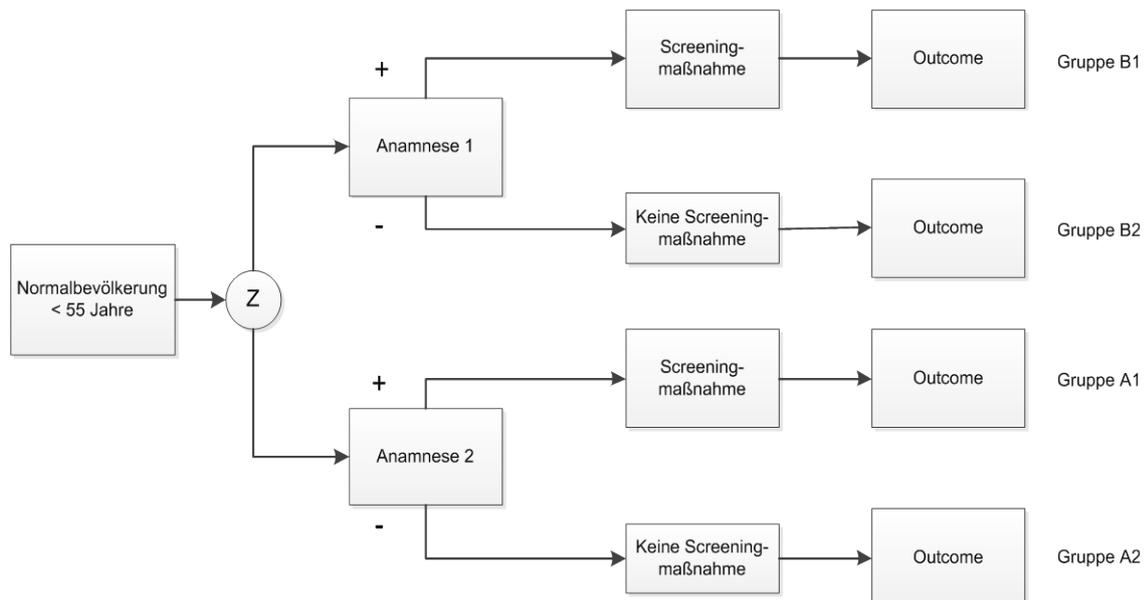
Ebenfalls möglich wäre ein Studiendesign wie in Abbildung 3 dargestellt: Hier würden Personen der Normalbevölkerung zunächst idealerweise zufällig zu den Bedingungen „Anamnese“ und „keine Anamnese“ (Gruppe A) zugeteilt. Die Personen, die an einer Anamnese teilnehmen, würden dann bei positivem Testergebnis einer Screeningmaßnahme (z. B. Koloskopie) unterzogen (Gruppe B1), diejenigen mit negativem Testergebnis nicht (Gruppe B2). Ein Vergleich der Gruppen B (B1+B2) und A evaluiert somit den Effekt der Screeningstrategie, bestehend aus der Anamnese und der darauffolgenden Screeningmaßnahme. Falls sich kein Unterschied zwischen den beiden Gruppen zeigt, ist allerdings unklar, ob dies daran liegt, dass der Test unzureichend war, oder daran, dass die Screeningmaßnahme in dieser Gruppe keinen Effekt hatte. Dieses Studiendesign würde auch die Beurteilung eventueller Schäden durch die Erhebung der Familienanamnese erlauben.



Z = Zuteilung zu den Interventionen

Abbildung 3: Kontrollierte Interventionsstudie zur Evaluation eines anamnestischen Tests (adaptiert nach [33])

Erweiterungen dieser Studiendesigns sind möglich, zum Beispiel um den Nutzen und Schaden verschiedener anamnestischer Instrumente miteinander zu vergleichen, wie in Abbildung 4 veranschaulicht:



Z = Zuteilung zu den Interventionen

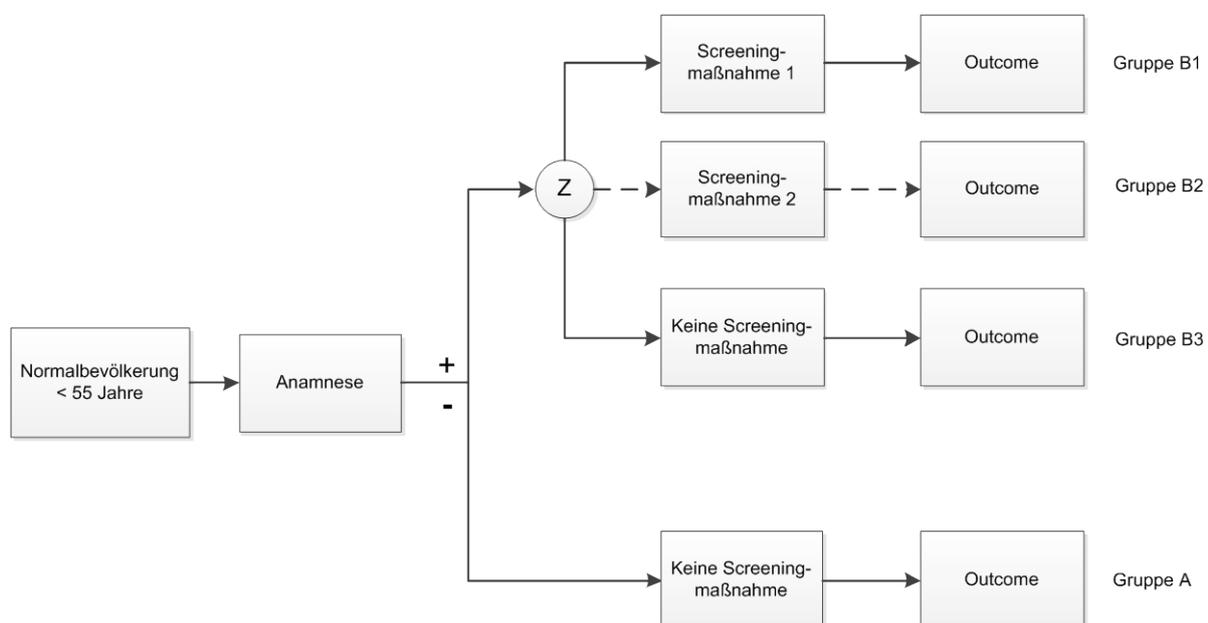
Abbildung 4: Kontrollierte Interventionsstudie zur Evaluation einer anamnestischen Strategie (adaptiert nach [33])

#### **Teilziel D: Vergleichende Nutzenbewertung verschiedener Screeningmaßnahmen im Rahmen einer Screeningstrategie bei Personen mit erhöhtem Risiko aufgrund einer positiven Familienanamnese**

Sollten sich zu Teilziel C keine Studien finden, so soll untersucht werden, welchen Nutzen verschiedene Maßnahmen im Rahmen eines modifizierten Screenings für Personen haben, bei denen bereits ein erhöhtes Risiko aufgrund einer positiven Familienanamnese festgestellt worden ist. Dies bezieht sich auf unterschiedliche Screeningmaßnahmen (insbesondere Koloskopie und FOBT), aber auch auf andere mögliche Unterschiede in den Screeningstrategien (z. B. verschiedene Zeitabstände zwischen den Untersuchungen).

Ziel D ist deshalb die vergleichende Nutzenbewertung verschiedener Maßnahmen im Rahmen einer Screeningstrategie im Vergleich zu keiner oder anderen Maßnahmen im Rahmen einer Screeningstrategie bei Personen mit erhöhtem Risiko aufgrund einer positiven Familienanamnese unter 55 Jahren hinsichtlich patientenrelevanter Endpunkte. Abbildung 5 veranschaulicht, wie eine mögliche Studie zur Beantwortung dieser Fragestellung angelegt sein könnte: Dabei würde in der Normalbevölkerung in einem früheren Alter als dem, in dem bisher das Darmkrebscreening angeboten wird (also z. B. mit 45 Jahren), ein Instrument zur Feststellung einer positiven Familienanamnese eingesetzt. Den Personen, für die das Testergebnis negativ ist (also keine positive Familienanamnese vorliegt), werden keine weiteren Maßnahmen angeboten (Gruppe A). Die Personen mit einem positiven Testergebnis

werden idealerweise zufällig zu 2 Studienarmen zugeteilt. Der einen Gruppe wird eine Screeningmaßnahme (z. B. Koloskopie) (Gruppe B1) angeboten, der anderen nicht (Gruppe B3). Ein Vergleich der Ergebnisse in Gruppe B1 versus Gruppe B3 evaluiert den Nutzen der Screeningmaßnahme für Personen mit einem positiven Testergebnis. Alternativ (oder zusätzlich) ist ein Studienarm mit einer alternativen Screeningmaßnahme denkbar, der einen Vergleich verschiedener Screeningmaßnahmen bei Personen mit positivem Testergebnis erlaubt (Gruppe B1 versus Gruppe B2). Darüber hinaus kann ein solches Studiendesign Informationen zur prognostischen Güte des Anamneseinstruments liefern, wenn der Darmkrebsstatus von Personen in Gruppe B3 mit dem bei Personen in Gruppe A verglichen wird – unter der Voraussetzung, dass die beiden Gruppen ansonsten vergleichbar sind. Der Nachteil dieses Designs ist, dass evtl. Schäden durch die Erhebung der Familienanamnese nicht beurteilt werden können.



Z = Zuteilung zu den Interventionen

Abbildung 5: Studiendesign mit Testergebnis als Einschlusskriterium (adaptiert nach [33])

Im Unterschied zu den Studien, die für Teilziel C beschrieben werden, in denen die gesamte Screeningkette abgebildet wird, können die für Teilziel D zu betrachtenden Studien lediglich indirekt Evidenz für den Nutzen der hier interessierenden Screeningstrategie liefern. Voraussetzung dafür ist, dass 1. entsprechende Screeningmaßnahmen bei Personen mit erhöhtem Risiko aufgrund einer positiven Familienanamnese einen Effekt hinsichtlich patientenrelevanter Endpunkte erzielen und 2. ein anamnestisches Instrument identifiziert werden konnte, mit dem zuverlässig Personen identifiziert werden, die den Einschlusskriterien dieser Studien entsprechen (sogenanntes Linked-Evidence-Konzept [34,35]).

### **3 Projektbearbeitung**

#### **3.1 Zeitlicher Verlauf des Projekts**

Der Gemeinsame Bundesausschuss (G-BA) hat in seiner Sitzung vom 17.03.2011 beschlossen, das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung des Nutzens einer Früherkennungsuntersuchung für Personen unter 55 Jahren mit familiärem Darmkrebsrisiko zu beauftragen.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Während der Erstellung des Berichtsplans wurde am 07.07.2011 eine Vertreterin der Deutschen ILCO e. V. zur Festlegung patientenrelevanter Zielgrößen konsultiert.

Der vorliegende vorläufige Berichtsplan (Version 1.0) wird zur Anhörung gestellt. Hierzu können schriftlich Stellungnahmen eingereicht werden. Das Ende der Stellungnahmefrist wird auf der Website des IQWiG ([www.iqwig.de](http://www.iqwig.de)) bekannt gegeben. Stellungnahmen können von allen interessierten Personen, Institutionen und Gesellschaften abgegeben werden. Die Stellungnahmen müssen bestimmten formalen Anforderungen genügen, die ebenfalls auf der Website des IQWiG in einem entsprechenden Leitfaden dargelegt sind. Gegebenenfalls wird eine wissenschaftliche Erörterung zur Klärung unklarer Aspekte aus den schriftlichen Stellungnahmen durchgeführt. Die Anhörung kann zu Änderungen und / oder Ergänzungen des Berichtsplans führen. Im Anschluss an diese Anhörung wird der dann gültige Berichtsplan publiziert.

## **4 Methoden**

### **4.1 Kriterien für den Einschluss von Studien in die Untersuchung**

Da der Auftrag auf Personen unter 55 Jahren abzielt, werden zu allen Teilzielen in erster Linie Studien oder Auswertungen zu dieser Altersgruppe betrachtet. Wenn die Datenlage für Personen unter 55 Jahren nicht ausreichend ist, werden die Ergebnisse zu anderen Altersgruppen als ergänzende Information dargestellt. In diesem Fall wird eine mögliche Effektmodifikation durch das Alter evaluiert.

#### **4.1.1 Teilziel A: Bestimmung des Erkrankungsrisikos bei positiver Familienanamnese (Prognose)**

##### **4.1.1.1 Population**

- Normalbevölkerung

##### **4.1.1.2 Prognostische Merkmale**

- Positive Familienanamnese bzw. keine positive Familienanamnese bezüglich Darmkrebs entsprechend der jeweiligen Definition in den Studien

##### **4.1.1.3 Patientenrelevante Endpunkte**

Für die Untersuchung werden folgende patientenrelevante Endpunkte verwendet:

- Auftreten des kolorektalen Karzinoms
- Krankheitsspezifisches (tumorspezifisches) Überleben

##### **4.1.1.4 Studientypen**

- Kohortenstudien
- Eingebettete Fall-Kontroll-Studien
- Fall-Kontroll-Studien

In den Studien werden relevante Störgrößen erkennbar berücksichtigt (z. B. durch adäquate Confounderadjustierung). Insbesondere sollten die folgenden Faktoren berücksichtigt sein: Ernährung, körperliche Aktivität, Gewicht / BMI, Alkoholkonsum, Rauchen.

##### **4.1.1.5 Sonstige Studiencharakteristika**

- Durchführung der Studie in westlichen Industrienationen (Europa, USA, Kanada, Neuseeland, Australien)

## **4.1.2 Teilziel B: Bewertung der diagnostischen Güte anamnestischer Instrumente**

### **4.1.2.1 Population**

- Normalbevölkerung

### **4.1.2.2 Prüfinstrument- und Referenztest**

- Das Prüfinstrument ist ein anamnestisches Instrument (z. B. Fragebogen, Interview) zur Feststellung einer positiven Familienanamnese für Darmkrebs. Das anamnestische Instrument muss vordefinierte Fragen beinhalten, davon mindestens 1 zur Feststellung einer positiven Familienanamnese für Darmkrebs.
- Der Referenztest muss eine Klärung der Verwandtschaftsverhältnisse und eine Feststellung des Darmkrebsstatus aller relevanten Verwandten anhand einer unabhängigen Datenquelle (z. B. Register, Krankenakten) beinhalten.

### **4.1.2.3 Gütemaße**

- Diagnostische Genauigkeit mittels Sensitivität / Spezifität oder positiver / negativer Likelihood-Quotienten, daraus abgeleitet ggf. prädiktive Aussagen mittels z. B. positivem / negativem prädiktiven Wert. Dabei müssen die Vierfeldertafel-Daten zur Berechnung der diagnostischen Güte der Primärstudien ableitbar sein.

### **4.1.2.4 Studientypen**

- Studien (vergleichend oder nicht vergleichend), aus denen sich die unter 4.1.2.3 genannten Gütemaße für die untersuchten anamnestischen Instrumente, ggf. auch zum Vergleich mehrerer untereinander, ableiten lassen.

### **4.1.2.5 Sonstige Studiencharakteristika**

- Durchführung der Studie in westlichen Industrienationen (Europa, USA, Kanada, Neuseeland, Australien)

### 4.1.3 Teilziel C: Vergleichende Nutzenbewertung umfassender Screeningstrategien unter Einbeziehung anamnestischer Instrumente

#### 4.1.3.1 Population

- Normalbevölkerung

#### 4.1.3.2 Prüf- und Vergleichsintervention

Die verschiedenen Möglichkeiten, den Nutzen einer Screeningstrategie unter Einbeziehung einer vorgeschalteten Anamnese zur Identifizierung von Personen mit erhöhtem Risiko aufgrund einer positiven Familienanamnese für Darmkrebs mittels unterschiedlicher Studiendesigns (vergleichend) zu bewerten, sind in Kapitel 2 beschrieben. Die folgende Tabelle 4 gibt einen Überblick über die jeweilige Prüf- und Vergleichsintervention, wie sie sich in den jeweiligen Studiendesigns darstellen.

Tabelle 4: Prüf- und Vergleichsintervention für Teilziel C

| Studiendesign /<br>Abbildung  | Prüfintervention   | Vergleichsintervention   |
|---|--|--|
| Studiendesign zur Erfassung einer Wechselwirkung <sup>a</sup> /<br>Abbildung 2  | Screeningmaßnahme (Gruppen B1, A1)<br><br>Erfassung einer möglichen Wechselwirkung durch den Vergleich des Effekts des Screenings in Abhängigkeit vom Anamneseergebnis | Keine Screeningmaßnahme (Gruppen B2, A2)   |
| Prospektive kontrollierte Interventionsstudie zur Evaluation eines anamnestischen Tests <sup>a</sup> /<br>Abbildung 3     | Anamnese mit nachgeschalteter Screeningmaßnahme in Abhängigkeit vom Anamneseergebnis (Gruppen B1+B2)   | Keine Anamnese und keine Screeningmaßnahme (Gruppe A)  |
| Prospektive kontrollierte Interventionsstudie zur Evaluation einer anamnestischen Strategie <sup>a</sup> /<br>Abbildung 4 | Anamnese 1 mit nachgeschalteter Screeningmaßnahme in Abhängigkeit vom Anamneseergebnis (Gruppen B1+B2)   | Anamnese 2 mit nachgeschalteter Screeningmaßnahme in Abhängigkeit vom Anamneseergebnis (Gruppen A1+A2) |
| a: Für eine genaue Definition des erforderlichen Studiendesigns siehe 4.1.3.4.  |  |  |

#### 4.1.3.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte verwendet:

- Gesamtüberleben
- Krankheitsspezifisches (tumorspezifisches) Überleben
- Auftreten des kolorektalen Karzinoms
- Auftreten fortgeschrittener Adenome
- Gesundheitsbezogene Lebensqualität sowie psychosoziale Aspekte

- Schäden, die sich direkt und indirekt aus der Befragung ergeben, einschließlich der Konsequenzen aus falschen anamnestischen Befunden und Überdiagnosen (z. B. Stigmatisierung / Etikettierung / Beunruhigung, unnötige Koloskopien)
- Schäden, die sich direkt und indirekt aus nachfolgenden diagnostischen Untersuchungen (z. B. Koloskopien) ergeben, einschließlich der Konsequenzen aus falschen Screeningbefunden und Überdiagnosen

Subjektive Endpunkte (zum Beispiel gesundheitsbezogene Lebensqualität) werden nur dann berücksichtigt, wenn sie mit validen Messinstrumenten (zum Beispiel Skalen) erfasst worden sind.

#### **4.1.3.4 Studientypen**

Randomisierte kontrollierte Studien (RCTs) sind, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt wurden, mit der geringsten Ergebnisunsicherheit behaftet. Sie liefern daher die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention.

Für alle unter 4.1.3.2 genannten Interventionen und alle unter 4.1.3.3 genannten Endpunkte ist eine Evaluation im Rahmen von randomisierten kontrollierten Studien möglich und praktisch durchführbar. Für den Vergleich zweier verschiedener Screeningstrategien werden daher RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen.

Ist die auf randomisierten kontrollierten Studien basierende Datenlage nicht hinreichend, um den patientenrelevanten Nutzen und Schaden mit ausreichender Ergebnissicherheit schätzen zu können (z. B. kleine Fallzahl, wenige Ereignisse [geringe Präzision], hohes Verzerrungspotenzial), werden gegebenenfalls auch verfügbare eindeutig prospektive, jedoch nicht randomisierte Interventionsstudien mit zeitlich parallelen Kontrollgruppen herangezogen (sogenannte Controlled Clinical Trials: CCTs), sofern das Problem einer möglichen Strukturungleichheit (unfairer Vergleich) der Beobachtungsgruppen adäquat in der Planung und Auswertung der entsprechenden Studien berücksichtigt wurde. Solche Studien können zwar die Aussage von aggregierten Ergebnissen aus qualitativ belastbaren RCTs in der Regel nicht qualitativ ändern, diese aber ggf. bestärken.

#### **4.1.4 Teilziel D: Vergleichende Nutzenbewertung verschiedener Screeningmaßnahmen im Rahmen einer Screeningstrategie bei Personen mit erhöhtem Risiko aufgrund einer positiven Familienanamnese**

##### **4.1.4.1 Population**

- Personen mit positiver Familienanamnese für Darmkrebs

##### **4.1.4.2 Prüf- und Vergleichsintervention**

- Screeningmaßnahme zur Früherkennung von Darmkrebs versus keine oder andere Screeningmaßnahme (z. B. Koloskopie, Sigmoidoskopie, FOBT; unterschiedliche Zeitabstände)

##### **4.1.4.3 Patientenrelevante Endpunkte**

Für die Untersuchung werden folgende patientenrelevante Endpunkte verwendet:

- Gesamtüberleben
- Krankheitsspezifisches (tumorspezifisches) Überleben
- Auftreten des kolorektalen Karzinoms
- Auftreten fortgeschrittener Adenome
- Gesundheitsbezogene Lebensqualität sowie psychosoziale Aspekte
- Schäden, die sich direkt und indirekt aus nachfolgenden diagnostischen Untersuchungen (z. B. Koloskopien) ergeben, einschließlich der Konsequenzen aus falschen Screeningbefunden und Überdiagnosen

Subjektive Endpunkte (zum Beispiel gesundheitsbezogene Lebensqualität) werden nur dann berücksichtigt, wenn sie mit validen Messinstrumenten (zum Beispiel Skalen) erfasst worden sind.

##### **4.1.4.4 Studientypen**

Randomisierte kontrollierte Studien (RCTs) sind, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt wurden, mit der geringsten Ergebnisunsicherheit behaftet. Sie liefern daher die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention.

Für alle unter 4.1.4.2 genannten Interventionen und alle unter 4.1.4.3 genannten Endpunkte ist eine Evaluation im Rahmen von randomisierten kontrollierten Studien möglich und praktisch durchführbar. Für den Vergleich zweier verschiedener Screeningmaßnahmen werden daher RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen.

Ist die auf randomisierten kontrollierten Studien basierende Datenlage nicht hinreichend, um den patientenrelevanten Nutzen und Schaden mit ausreichender Ergebnissicherheit schätzen zu können (z. B. kleine Fallzahl, wenige Ereignisse [geringe Präzision], hohes Verzerrungspotenzial), werden gegebenenfalls auch verfügbare eindeutig prospektive, jedoch nicht randomisierte Interventionsstudien mit zeitlich parallelen Kontrollgruppen herangezogen (sogenannte Controlled Clinical Trials: CCTs), sofern das Problem einer möglichen Strukturungleichheit (unfairer Vergleich) der Beobachtungsgruppen adäquat in der Planung und Auswertung der entsprechenden Studien berücksichtigt wurde. Solche Studien können zwar die Aussage von aggregierten Ergebnissen aus qualitativ belastbaren RCTs in der Regel nicht qualitativ ändern, diese aber ggf. bestärken.

#### 4.1.5 Tabellarische Übersicht über die Kriterien für den Studieneinschluss

Die folgende Tabelle zeigt die Kriterien für den Einschluss von Studien in die Bewertung.

Tabelle 5: Übersicht über die Kriterien für den Studieneinschluss

| <b>Einschlusskriterien Teilziel A (Prognose)</b> |  |
|--|--|
| EA1  | Normalbevölkerung (siehe auch Abschnitt 4.1.1.1)   |
| EA2  | positive Familienanamnese bzw. keine positive Familienanamnese bezüglich Darmkrebs (siehe auch Abschnitt 4.1.1.2)  |
| EA3  | patientenrelevante Endpunkte wie in Abschnitt 4.1.1.3 formuliert   |
| EA4  | Kohortenstudien, eingebettete Fall-Kontroll-Studien, Fall-Kontroll-Studien jeweils mit erkennbarer Berücksichtigung relevanter Störgrößen (siehe auch Abschnitt 4.1.1.4) |
| EA5  | Vollpublikation verfügbar <sup>a</sup>   |
| EA6  | Durchführung der Studie in westlichen Industrienationen (Europa, USA, Kanada, Neuseeland, Australien)  |
| <b>Einschlusskriterien Teilziel B (Diagnose)</b> |  |
| EB1  | Normalbevölkerung (siehe auch Abschnitt 4.1.2.1)   |
| EB2  | anamnestisches Instrument zur Feststellung einer positiven Familienanamnese für Darmkrebs als Prüfinstrument und umfassender Referenztest (siehe auch Abschnitt 4.1.2.2) |
| EB3  | Angaben zur diagnostischen Genauigkeit wie in Abschnitt 4.1.2.3 formuliert   |
| EB4  | Studien, aus denen sich Gütemaße zur diagnostischen Genauigkeit wie unter 4.1.2.3 definiert ableiten lassen (siehe auch Abschnitt 4.1.2.4)                               |
| EB5  | Vollpublikation verfügbar <sup>a</sup>   |
| EB6  | Durchführung der Studie in westlichen Industrienationen (Europa, USA, Kanada, Neuseeland, Australien)  |

(Fortsetzung)

Tabelle 5: Übersicht über die Kriterien für den Studieneinschluss (Fortsetzung)

| <b>Einschlusskriterien Teilziel C (Screeningstrategie)</b>   |  |
|--|--|
| EC1  | Normalbevölkerung (siehe auch Abschnitt 4.1.3.1)   |
| EC2  | Prüf- und Vergleichsintervention entsprechend Abschnitt 4.1.3.2  |
| EC3  | patientenrelevante Endpunkte wie in Abschnitt 4.1.3.3 formuliert   |
| EC4  | randomisierte kontrollierte Studien sowie prospektive kontrollierte Interventionsstudien mit geeigneten Maßnahmen zur Sicherstellung der Strukturgleichheit (siehe auch Abschnitt 4.1.3.4) |
| EC5  | Vollpublikation verfügbar <sup>a</sup>   |
| <b>Einschlusskriterien Teilziel D (Screeningmaßnahme)</b>  |  |
| ED1  | Personen mit positiver Familienanamnese für Darmkrebs (siehe auch Abschnitt 4.1.4.1)   |
| ED2  | Screeningmaßnahme zur Früherkennung von Darmkrebs (siehe auch Abschnitt 4.1.4.2)   |
| ED3  | keine oder andere Screeningmaßnahme (siehe auch Abschnitt 4.1.4.2)   |
| ED4  | patientenrelevante Endpunkte wie in Abschnitt 4.1.4.3 formuliert   |
| ED5  | randomisierte kontrollierte Studien sowie prospektive kontrollierte Interventionsstudien mit geeigneten Maßnahmen zur Sicherstellung der Strukturgleichheit (siehe auch Abschnitt 4.1.4.4) |
| ED6  | Vollpublikation verfügbar <sup>a</sup>   |
| <b>Ausschlusskriterium (für alle Teilziele)</b>  |  |
| A1   | Mehrfachpublikationen ohne relevante Zusatzinformationen   |
| a: Als Vollpublikation gilt in diesem Zusammenhang auch die nicht vertrauliche Weitergabe eines Studienberichts an das Institut oder die nicht vertrauliche Bereitstellung eines Berichts über die Studien, der den Kriterien des CONSORT- [36], TREND- [37], STARD- [38] oder STROBE-Statements [39] genügt und eine Bewertung der Studie ermöglicht. |  |

#### 4.1.6 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für das Einschlusskriterium E1 (Population) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patienten dieses Kriterium erfüllt ist. Liegen für solche Studien entsprechende Subgruppenanalysen vor, wird auf diese Analysen zurückgegriffen. Studien, bei denen das Einschlusskriterium E1 bei weniger als 80 % erfüllt ist, werden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorliegen.

## 4.2 Informationsbeschaffung

### 4.2.1 Bibliografische Literaturrecherche

Die systematische Literaturrecherche nach relevanten Studien soll in folgenden Quellen durchgeführt werden:

- Suche nach Primärstudien in den bibliografischen Datenbanken MEDLINE, EMBASE, NHS Economic Evaluation Database und Cochrane Central Register of Controlled Trials (Clinical Trials) (letztgenannte nur für Teilziele C und D)
- Suche nach systematischen Übersichten: Identifizierung mittels Suche in den Datenbanken MEDLINE und EMBASE parallel zur Suche nach relevanter Primärliteratur sowie mittels Suche in den Datenbanken Cochrane Database of Systematic Reviews (Cochrane Reviews), Database of Abstracts of Reviews of Effects (Other Reviews) und Health Technology Assessment Database (Technology Assessments)

#### **4.2.2 Suche nach weiteren publizierten und nicht publizierten Studien**

Zusätzlich zur Suche in bibliografischen Datenbanken sollen folgende Quellen zur Identifizierung publizierter und nicht publizierter Studien herangezogen werden:

- systematische Übersichten
- durch den G-BA übermittelte Unterlagen
- öffentlich zugängliche Studienregister
- im Rahmen der Anhörung zum vorläufigen Berichtsplan und zum Vorbericht eingereichte Informationen

Gegebenenfalls werden Informationen von Autoren einzelner Publikationen bzw. Herausgebern spezifischer Instrumente, zum Beispiel zur Frage nach nicht publizierten Teilaspekten, eingeholt.

#### **4.2.3 Selektion relevanter Studien**

Die Selektion relevanter Studien erfolgt durch 2 Reviewer unabhängig voneinander. Dazu wird das Ergebnis der Recherche in den oben genannten Quellen herangezogen.

### **4.3 Informationsbewertung**

Die Bewertung der Informationen der eingeschlossenen Studien hängt stark von den verfügbaren Angaben und der Qualität der jeweiligen Publikationen und weiterer Informationsquellen ab. Alle für die Nutzenbewertung relevanten Ergebnisse werden hinsichtlich ihrer Ergebnissicherheit, bestehend aus dem Verzerrungspotenzial und der Präzision der Ergebnisse, überprüft.

#### **Datenextraktion**

Alle für die Nutzenbewertung notwendigen Informationen werden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert.

#### **4.3.1 Bewertung von Studien zum Erkrankungsrisiko (Teilziel A)**

Die eingeschlossenen Studien zum Erkrankungsrisiko werden hinsichtlich ihres Verzerrungspotenzials bewertet. Bezüglich möglicher Verzerrungen sind bei Kohortenstudien, eingebetteten Fall-Kontroll-Studien und Fall-Kontroll-Studien insbesondere folgende Punkte bedeutsam [40,41]:

- Auswahl der Studienpopulation
- Angemessenheit der Nachbeobachtung (für Kohortenstudien)
- Adäquate Erfassung der prognostischen Faktoren
- Adäquate Erfassung der Endpunkte
- Adäquater Umgang mit Störgrößen
- Ergebnisunabhängige Berichterstattung

#### **4.3.2 Bewertung von Studien zur diagnostischen Güte (Teilziel B)**

Die Bewertung des Verzerrungspotenzials der Studien zur diagnostischen Güte wird mithilfe des QUADAS-2-Instruments durchgeführt [42,43]. Das Verzerrungspotenzial der Studien zur diagnostischen Güte wird als „niedrig“ oder „hoch“ eingestuft.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führt nicht zum Ausschluss aus der Bewertung der diagnostischen Güte. Die Klassifizierung dient vielmehr der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

#### **4.3.3 Bewertung von Studien zum patientenrelevanten Nutzen (Teilziel C und D)**

Das Verzerrungspotenzial der Ergebnisse wird für jede in die Nutzenbewertung eingeschlossene Studie bewertet, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu werden insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte, die das Verzerrungspotenzial beeinflussen, systematisch extrahiert und bewertet:

##### **A: Aspekte des Verzerrungspotenzials der Ergebnisse auf Studienebene**

- Erzeugung der Randomisierungssequenz (bei randomisierten Studien)
- Verdeckung der Gruppenzuteilung (bei randomisierten Studien)
- zeitliche Parallelität der Gruppen (bei nicht randomisierten kontrollierten Studien)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (bei nicht randomisierten kontrollierten Studien)
- Verblindung des Patienten sowie des Behandlers (bei randomisierten Studien)
- ergebnisgesteuerte Berichterstattung

**B: Aspekte des Verzerrungspotenzials der Ergebnisse auf Endpunktebene**

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung

Das Verzerrungspotenzial wird als „niedrig“ oder „hoch“ eingestuft. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Für die Bewertung eines Endpunkts wird zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft. Falls diese Einstufung als „hoch“ erfolgt, wird das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet. Ansonsten finden die unter B genannten endpunktspezifischen Aspekte Berücksichtigung.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führt nicht zum Ausschluss aus der Nutzenbewertung. Die Klassifizierung dient vielmehr der Diskussion heterogener Studienergebnisse und beeinflusst die Sicherheit der Aussage.

**4.4 Informationssynthese und -analyse**

Die Informationen werden einer Informationssynthese und -analyse unterzogen. Wenn möglich, werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die unten beschriebenen Werkzeuge eingesetzt. Eine abschließende zusammenfassende Bewertung der Informationen erfolgt darüber hinaus in jedem Fall.

Bei der Betrachtung patientenrelevanter Endpunkte, die mithilfe von (komplexen) Skalen operationalisiert werden, ist es in besonderer Weise notwendig, neben der statistischen Signifikanz der Effekte die Relevanz der beobachteten Wirkungen der untersuchten Interventionen zu bewerten, da die Komplexität der Skalen häufig eine sinnvolle Interpretation geringfügiger Unterschiede erschwert. Die Relevanzbewertung kann auf Basis von Mittelwertdifferenzen und Responderanalysen vorgenommen werden. Welches dieser Verfahren gewählt werden kann, hängt insbesondere von der Verfügbarkeit der Daten aus den Primärstudien ab.

Um skalen- und indikationsspezifischen Besonderheiten gerecht zu werden, soll die Relevanzbewertung primär auf Basis validierter (bzw. etablierter) Relevanzkriterien für die jeweilige Skala durchgeführt werden. Für die Betrachtung von Mittelwertdifferenzen ist dies z. B. eine validierte bzw. etablierte Irrelevanzschwelle für den Gruppenunterschied: Liegt das zum beobachteten Effekt korrespondierende Konfidenzintervall vollständig oberhalb dieser Irrelevanzschwelle, wird davon ausgegangen, dass die Effektstärke nicht in einem sicher

irrelevanten Bereich liegt. Bei der Betrachtung von Responderanalysen ist es erforderlich, dass bei diesen Analysen ein validiertes bzw. etabliertes Responsekriterium angewendet wurde (im Sinne einer individuellen Minimal Important Difference [MID]). Liegt bei einer solchen Auswertung ein statistisch signifikanter Unterschied der Anteile der Responder zwischen den Gruppen vor, wird dies als relevanter Effekt angesehen, da die Responserdefinition bereits eine Schwelle der Irrelevanz (nämlich die MID) beinhaltet.

Falls skalenspezifisch validierte bzw. etablierte Relevanzkriterien nicht vorliegen, muss auf ein allgemeines statistisches Maß zur Relevanzbewertung zurückgegriffen werden. In diesem Fall werden standardisierte Mittelwertdifferenzen (SMDs in Form von Hedges'  $g$ ) betrachtet. Als Irrelevanzschwelle wird dann 0,2 verwendet: Liegt das zum beobachteten Effekt korrespondierende Konfidenzintervall vollständig oberhalb dieser Irrelevanzschwelle, wird davon ausgegangen, dass die Effektstärke nicht in einem sicher irrelevanten Bereich liegt. Dies soll gewährleisten, dass der beobachtete Effekt hinreichend sicher mindestens als „klein“ angesehen werden kann [44,45].

#### **4.4.1 Gegenüberstellung der Ergebnisse der Einzelstudien**

Die Ergebnisse zu den in den Studien berichteten patientenrelevanten Endpunkten werden im Bericht vergleichend beschrieben.

In bestimmten Fällen werden einzelne Ergebnisse aus den Studien zu einem Endpunkt nicht dargestellt bzw. nicht in die Nutzenbewertung, Bewertung des Risikos bzw. Bewertung der diagnostischen Güte einbezogen. Dies trifft insbesondere zu, wenn viele Patienten nicht in der Auswertung enthalten sind. Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patienten basieren, das heißt wenn der Anteil der Patienten ohne jegliche Berücksichtigung in der Auswertung (Nichtberücksichtigungsanteil) größer als 30 % ist. In der Literatur werden zum Teil bereits Nichtberücksichtigungsanteile größer als 20 Prozentpunkte als nicht mehr aussagekräftig betrachtet [46].

Ausnahmen von dieser Regel können zum Beispiel dann gemacht werden, wenn aus logistischen Gründen für ganze Zentren (ganze Randomisierungsblöcke) keine Daten erhoben wurden und dies bereits bei der Studienplanung vorgesehen war [47].

Die Ergebnisse werden auch dann nicht in die Nutzenbewertung, Bewertung des Risikos bzw. Bewertung der diagnostischen Güte einbezogen, wenn der Unterschied der Nichtberücksichtigungsanteile zwischen den Gruppen größer als 15 Prozentpunkte ist.

#### **4.4.2 Meta-Analysen**

Sofern die Studien hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbar sind, werden die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst. Für die statistische Auswertung werden primär die Ergebnisse aus Intention-to-Treat-

Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet. Die Meta-Analysen erfolgen in der Regel auf Basis von Modellen mit zufälligen Effekten [48]. In begründeten Ausnahmefällen werden Modelle mit festen Effekten eingesetzt. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, werden diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt.

### **Studien zum patientenrelevanten Nutzen**

Für kontinuierliche Variablen wird die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges'  $g$ , als Effektmaß eingesetzt. Bei binären Variablen werden Meta-Analysen primär anhand des Odds Ratio und beim Vergleich von Überlebenszeiten anhand des Hazard Ratio durchgeführt. Bei kategorialen Variablen wird ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und von den verfügbaren Daten verwendet [49].

Die Effektschätzer und Konfidenzintervalle aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Anschließend erfolgt die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes  $I^2$  und des statistischen Tests auf Vorliegen von Heterogenität [50]. Ist die Heterogenität der Studienergebnisse nicht bedeutsam ( $p > 0,2$  für Heterogenitätstest), wird der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt. Bei bedeutsamer Heterogenität werden die Ergebnisse nur in begründeten Ausnahmefällen gepoolt. Außerdem wird untersucht, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.4.3) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.4.4).

### **Diagnostestudien**

Die Effektschätzer und Konfidenzintervalle aus den Studien werden ebenfalls mittels Forest Plots zusammenfassend dargestellt. Außerdem wird, sofern die dafür nötigen Anforderungen als erfüllt gelten können, für die Testgütekriterien eine Meta-Analyse anhand der Sensitivität und Spezifität in einer bivariaten Analyse durchgeführt [51]. Die Schätzung der Modellparameter (Sensitivität, Spezifität und Parameter für die hierarchical summary receiver operating curve, HSROC) erfolgt über einen Maximum-Likelihood-Ansatz in einem nicht linearen gemischten Modell. In begründeten Ausnahmefällen kommen auch andere Effektmaße zum Einsatz.

Zusätzlich werden bei diagnostischen Studien die beobachteten Paare aus Sensitivität und Spezifität grafisch zweidimensional dargestellt. Studien, in denen verschiedene diagnostische Verfahren direkt verglichen wurden, werden gekennzeichnet. Des Weiteren werden die aus bivariaten Meta-Analysen gewonnenen Schätzer für die Erwartungswerte als gepoolte Paare der Sensitivität und der Spezifität mit den dazugehörigen 95%-Konfidenzellipsen grafisch dargestellt. Schließlich werden die geschätzten HSROC-Kurven abgebildet. Die grafischen Darstellungen der Sensitivität und Spezifität und der HSROC-Kurve können Tendenzen zu einer besseren Testgüte in den vorliegenden Studien aufzeigen. Schneiden sich die Konfidenzellipsen zweier Technologien nicht, so wird davon ausgegangen, dass in den

vorliegenden Studien ein Unterschied in der diagnostischen Güte besteht. Die Richtung des Unterschieds wird ggf. grafisch bestimmt. Überlappen sich 2 Konfidenzellipsen beträchtlich oder schneiden sich 2 HSROC-Kurven im Bereich der Grafik, in dem auch Paare aus Sensitivität und Spezifität beobachtet wurden, wird davon ausgegangen, dass in den vorliegenden Studien kein Unterschied in der diagnostischen Güte besteht.

Das Vorliegen von Heterogenität wird anhand von Sensitivitätsanalysen und Subgruppenanalysen untersucht.

#### **4.4.3 Sensitivitätsanalyse**

Zur Einschätzung der Robustheit der Ergebnisse sind Sensitivitätsanalysen hinsichtlich methodischer Faktoren geplant. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ wird für Sensitivitätsanalysen verwendet.

Der Algorithmus zum Schätzen der Parameter im bivariaten Modell kann ggf. keinen Schätzer liefern, da das Maximum-Likelihood-Verfahren nicht konvergiert, oder zu unrealistischen Schätzern führen, d. h. zu Schätzern mit zu großen Standardfehlern und entsprechenden Konfidenzintervallen. Die Gründe hierfür können beispielsweise sein, dass zu wenige Studien vorliegen, als dass die Parameter geschätzt werden könnten. Aber auch einzelne Studien mit extremen Werten können dazu führen. Hierzu werden Sensitivitätsanalysen durchgeführt, um diese Studien zu identifizieren.

Das Ergebnis der Sensitivitätsanalysen kann die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestufteffekt kann zum Beispiel dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen Nutzen attestiert wird.

#### **4.4.4 Subgruppenmerkmale und andere Effektmodifikatoren**

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Dies können direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen sein. Im Gegensatz zu den in Abschnitt 4.4.3 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- bzw. Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung von Effektmodifikatoren werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Es

ist vorgesehen, teilzielübergreifend folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Geschlecht
- Alter

Sollten sich aus den verfügbaren Informationen Anhaltspunkte für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden.

Bei Identifizierung möglicher Effektmodifikatoren erfolgt gegebenenfalls eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen. Beispielsweise kann der Beleg eines Zusatznutzens auf eine spezielle Subgruppe von Patienten eingeschränkt werden.

## 5 Literaturverzeichnis

1. Husmann G, Kaatsch P, Katalinic A, Bertz J, Haberland J, Kraywinkel K et al. Krebs in Deutschland 2005/2006: Häufigkeiten und Trends. Berlin: Robert Koch-Institut; 2010.  
URL: [http://edoc.rki.de/documents/rki\\_fv/re2vZ21t28Ir8Y/PDF/22aJOdYnmXV0I.pdf](http://edoc.rki.de/documents/rki_fv/re2vZ21t28Ir8Y/PDF/22aJOdYnmXV0I.pdf).
2. Bertz J, Dahm S, Haberland J, Kraywinkel K, Kurth BM, Wolf U. Verbreitung von Krebserkrankungen in Deutschland: Entwicklung der Prävalenzen zwischen 1990 und 2010. Berlin: Robert Koch-Institut; 2010.  
URL: [http://edoc.rki.de/documents/rki\\_fv/re2vZ2t28Ir8Y/PDF/23GSS31yB0GKUUhU.pdf](http://edoc.rki.de/documents/rki_fv/re2vZ2t28Ir8Y/PDF/23GSS31yB0GKUUhU.pdf).
3. Deutsches Institut für Medizinische Dokumentation und Information. ICD-10-GM: internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision; German modification. Köln: DIMDI; 2011.  
URL: <http://www.dimdi.de/dynamic/de/klassi/downloadcenter/icd-10-gm/version2011/systematik>.
4. Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. CA Cancer J Clin 2009; 59(6): 366-378.
5. International Agency for Research on Cancer, World Health Organization. Colorectal cancer incidence and mortality worldwide in 2008: summary [online]. In: GLOBOCAN. 2010 [Zugriff: 18.07.2011]. URL: <http://globocan.iarc.fr/factsheets/cancers/colorectal.asp>.
6. Batzler WU, Giersiepen K, Hentschel S, Husemann G, Kaatsch P, Katalinic A et al. Krebs in Deutschland 2003-2004: Häufigkeiten und Trends. Berlin: Robert Koch-Institut; 2008.  
URL: <http://www.ekr.med.uni-erlangen.de/GEKID/Doc/kid2008.pdf>.
7. World Cancer Research Fund, American Institute for Cancer Research. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Washington: AICR 2007.  
URL: <http://eprints.ucl.ac.uk/4841/1/4841.pdf>.
8. Chan AT, Giovannucci EL. Primary prevention of colorectal cancer. Gastroenterology 2010; 138(6): 2029-2043.e10.
9. Jaspersion KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. Gastroenterology 2010; 138(6): 2044-2058.
10. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. Am J Gastroenterol 2001; 96(10): 2992-3003.
11. Butterworth AS, Higgins JP, Pharoah P. Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis. Eur J Cancer 2006; 42(2): 216-227.

12. Schmiegel W, Reinacher-Schick A, Arnold D, Graeven U, Heinemann V, Porschen R et al. S3-Leitlinie "Kolorektales Karzinom": Aktualisierung 2008. *Z Gastroenterol* 2008; 46(8): 799-840.
13. Basseri RJ, Basseri B, Papadakis KA. Dysplasia and cancer in inflammatory bowel disease. *Expert Rev Gastroenterol Hepatol* 2011; 5(1): 59-66.
14. Winawer SJ. A quarter century of colorectal cancer screening: progress and prospects. *J Clin Oncol* 2001; 19(18): 6S-12S.
15. Leslie A, Carey FA, Pratt NR, Steele RJ. The colorectal adenoma-carcinoma sequence. *Br J Surg* 2002; 89(7): 845-860.
16. Brenner H, Hoffmeister M, Stegmaier C, Brenner G, Altenhofen L, Haug U. Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies. *Gut* 2007; 56(11): 1585-1589.
17. Braunwald E, Fauci AS, Kasper DL, Hauser SL, Longo DL, Jameson JL (Ed). *Harrison's principles of internal medicine*. New York: McGraw-Hill Professional; 2001.
18. National Institute of Clinical Excellence. Improving outcomes in colorectal cancers: guidance on cancer services; manual update [online]. 2004.  
URL: <http://www.nice.org.uk/nicemedia/pdf/CSGCCfullguidance.pdf>.
19. Krebsinzidenz 2001 und 2002 im Erfassungsgebiet des Gemeinsamen Krebsregisters: Jahresbericht ; Gemeinsames Krebsregister der Länder Berlin, Brandenburg, Mecklenburg-Vorpommern, Sachsen-Anhalt und der Freistaaten Sachsen und Thüringen (GKR) [online]. [Zugriff: 04.10.2011]. (Schriftenreihe des GKR 1/2005).  
URL: [http://www.berlin.de/imperia/md/content/gkr/publikationen/jahresberichte/jb2001\\_02.pdf?start&ts=1182861148&file=jb2001\\_02.pdf](http://www.berlin.de/imperia/md/content/gkr/publikationen/jahresberichte/jb2001_02.pdf?start&ts=1182861148&file=jb2001_02.pdf)
20. Ahnen DJ, Macrae FA. Clinical manifestations, diagnosis, and staging of colorectal cancer [online]. In: UpToDate. 13.12.2010 [Zugriff: 28.07.2011]. URL: <http://www.uptodate.com>.
21. Wittekind C, Meyer HJ (Ed). *TNM: Klassifikation maligner Tumoren*. Weinheim: Wiley-Blackwell; 2010.
22. American Joint Committee on Cancer. Cancer colon and rectum cancer staging [online]. 2009 [Zugriff: 04.08.2011].  
URL: <http://www.cancerstaging.org/staging/posters/colon8.5x11.pdf>.
23. Giersiepen K, Hense HW, Klug SJ, Antes G, Zeeb H. Entwicklung, Durchführung und Evaluation von Programmen zur Krebsfrüherkennung: ein Positionspapier. *Z Arztl Fortbild Qualitätssich* 2007; 101(1): 43-49.

24. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010; 102(9): 605-613.
25. Lock G. Okkultes Blut im Stuhl In: Guder WG, Nolte J (Ed). *Das Laborbuch für Klinik und Praxis*. München: Elsevier; 2005. S. 848-849.
26. Ell C, Friedrich-Rust M, Schmitt W. Positionspapier der Sektion Endoskopie der DGVS zur Darmreinigung vor Koloskopie. *Z Gastroenterol* 2007; 45(11): 1191-1198.
27. Steckelberg A, Mühlhauser I. Darmkrebs Screening: Früherkennung [online]. 02.2011 [Zugriff: 15.08.2011]. URL: <http://www.gesundheit.uni-hamburg.de/upload/NeueDarmkrebsbroschuere2011.pdf>.
28. Riphaut A, Rabofski M, Wehrmann T. Sedierung in der gastrointestinalen Endoskopie in Deutschland - eine bundesweite Evaluation. *Z Gastroenterol* 2007; 45(8): P131.
29. Whitlock EP, Lin JS, Liles E, Beil TL, Fu R. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2008; 149(9): 638-658.
30. Levin B, Lieberman DA, McFarland B, Andrews KS, Brooks D, Bond J et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology* 2008; 134(5): 1570-1595.
31. Bechtler M, Eickhoff A, Riemann JF. Das Intervallkarzinom und mögliche Ursachen. *Dtsch Med Wochenschr* 2008; 133(47): 2458-2462.
32. Gemeinsamer Bundesausschuss. Richtlinie des Gemeinsamen Bundesausschusses über die Früherkennung von Krebserkrankungen (Krebsfrüherkennungs-Richtlinie / KFE-RL) [online]. 16.12.2010 [Zugriff: 05.08.2011]. URL: [http://www.g-ba.de/downloads/62-492-510/RL\\_KFU\\_2010-12-16.pdf](http://www.g-ba.de/downloads/62-492-510/RL_KFU_2010-12-16.pdf).
33. Lijmer JG, Bossuyt PM. Various randomized designs can be used to evaluate medical tests. *J Clin Epidemiol* 2009; 62(4): 364-373.
34. Medical Services Advisory Committee. Guidelines for the assessment of diagnostic technologies [online]. 2005. [Zugriff: 08.03.2011]. URL: [http://www.health.gov.au/internet/msac/publishing.nsf/Content/D81BE529B98B3DB6CA2575AD0082FD1B/\\$File/Diag%20Guidelines%20Sept%202005%20updated%2021%20may%202007.pdf](http://www.health.gov.au/internet/msac/publishing.nsf/Content/D81BE529B98B3DB6CA2575AD0082FD1B/$File/Diag%20Guidelines%20Sept%202005%20updated%2021%20may%202007.pdf).

35. Donegan E, Gajraj E, Garner S, Garrett Z, Lord J. Briefing paper for methods review workshop on diagnostic technologies [online]. 30.11.2007 [Zugriff: 11.03.2008].  
URL: <http://www.nice.nhs.uk/media/8AD/5C/FinalBriefingPaperDiagnosticsAMENDED120707.pdf>.
36. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ et al. CONSORT 2010: explanation and elaboration; updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c869.
37. Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004; 94(3): 361-366.
38. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. *Ann Intern Med* 2003; 138(1): 40-44.
39. Von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.
40. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med* 2006; 144(6): 427-437.
41. Altman DG. Systematic reviews of evaluations of prognostic variables. *BMJ* 2001; 323(7306): 224-228.
42. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; 3: 25.
43. University of Bristol. QUADAS 2 [online]. 19.07.2011 [Zugriff: 29.07.2011].  
URL: <http://www.bris.ac.uk/quadas/quadas-2>.
44. Fayers P, Hays R. Assessing quality of life in clinical trials: methods and practice. Oxford: Oxford University Press; 2005.
45. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale: Erlbaum; 1988.
46. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; 359(9308): 781-785.
47. Lange S. The all randomized/full analysis set (ICH E9): may patients be excluded from the analysis? *Drug Inf J* 2001; 35(3): 881-891.

48. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7(3): 177-188.

49. Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

50. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414): 557-560.

51. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58(10): 982-990.