



Lungenkrebscreening mittels Niedrigdosis- Computertomografie

Berichtsplan

Auftrag: S19-02
Version: 1.0
Stand: 22.07.2019

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Lungenkrebscreening mittels Niedrigdosis-Computertomografie

Auftraggeber

Gemeinsamer Bundesausschuss

Datum des Auftrags

18.07.2019

Interne Auftragsnummer

S19-02

Anschrift des Herausgebers

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

Schlagwörter: Reihenuntersuchung, Computertomografie, Lungentumoren, Nutzenbewertung, Systematische Übersicht

Keywords: Mass Screening, Tomography – X-Ray Computed, Lung Neoplasms, Benefit Assessment, Systematic Review

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	iv
Abkürzungsverzeichnis	v
1 Hintergrund	1
2 Fragestellung	3
3 Projektverlauf	4
4 Methoden	5
4.1 Kriterien für den Einschluss von Studien in die Untersuchung	5
4.1.1 Population.....	5
4.1.2 Prüf- und Vergleichsintervention	5
4.1.3 Patientenrelevante Endpunkte	5
4.1.4 Studientypen	5
4.1.5 Studiendauer	6
4.1.6 Publikationssprache.....	6
4.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss	6
4.1.8 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen.....	6
4.2 Informationsbeschaffung	7
4.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten.....	7
4.2.2 Umfassende Informationsbeschaffung	7
4.2.3 Selektion relevanter Studien.....	8
4.3 Informationsbewertung und -synthese	8
4.3.1 Bewertung von systematischen Übersichten zu Screeningstudien.....	8
4.3.2 Darstellung der Einzelstudien.....	9
4.3.3 Bewertung des Verzerrungspotenzials der Ergebnisse.....	9
4.3.4 Metaanalysen	10
4.3.5 Sensitivitätsanalysen	11
4.3.6 Subgruppenmerkmale und andere Effektmodifikatoren	11
4.3.7 Aussagen zur Beleglage	12
5 Literatur	13

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss.....	6
Tabelle 2: Regelmäßig abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit.....	12

Abkürzungsverzeichnis

Abkürzung	Bedeutung
G-BA	Gemeinsamer Bundesausschuss
CT	Computertomografie
IARC	Internationale Agentur für Krebsforschung
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RKI	Robert Koch-Institut

1 Hintergrund

Ein Lungenkarzinom bezeichnet einen epithelialen malignen Tumor im Lungen- oder Bronchialsystem [1]. Laut Robert Koch-Institut (RKI) wurde im Jahr 2014 in Deutschland bei Männern Lungenkrebs nach Prostatakrebs am zweithäufigsten diagnostiziert und war die häufigste krebsbedingte Todesursache. Bei Frauen war Lungenkrebs nach Brustkrebs und Darmkrebs eine der am häufigsten diagnostizierten Krebserkrankungen und die zweithäufigste Todesursache (nach Brustkrebs). Männer erkranken im Median mit 69, Frauen mit 68 Jahren [2]. Die Internationale Agentur für Krebsforschung (IARC) schätzte, dass im Jahr 2018 weltweit unter allen Krebsneuerkrankungen, bei einer gemeinsamen Betrachtung beider Geschlechter, Lungenkrebs am häufigsten auftrat und die häufigste Krebstodesursache war [3].

Der bedeutendste Risikofaktor für Lungenkrebs ist das Rauchen. Bei etwa 9 von 10 erkrankten Männern und bei mindestens 6 von 10 erkrankten Frauen wurde der Lungenkrebs auf aktives Rauchen zurückgeführt. Auch Passivrauchen erhöht das Risiko [2]. Umweltexpositionen wie Radon oder Feinstaub sowie Asbest sind weitere Risikofaktoren für Lungenkrebs [4]. Die Symptome der Erkrankung sind unspezifisch. Zu den häufigsten Symptomen gehören anhaltender Husten und Dyspnoe. Bei fortgeschrittener Erkrankung treten auch Abgeschlagenheit, Gewichtsverlust, Thoraxschmerzen, Knochenschmerzen und Hämoptysen auf [1].

Bei Lungenkrebs wird grob zwischen den nicht kleinzelligen Lungenkarzinomen (Non-small Cell Lung Cancer [NSCLC]) und den kleinzelligen Lungenkarzinomen (Small Cell Lung Cancer [SCLC]) unterschieden. Diese Aufteilung beruht auf Unterschieden in Therapie und Prognose, denn Letztere sind rasch progredient und führen früh zu einer systemischen Streuung. Die wichtigsten histologischen Haupttypen zur weiteren Unterteilung der nicht kleinzelligen Karzinome sind das Adenokarzinom, das Plattenepithelkarzinom und das großzellige Karzinom [1]. Anhand weiterer Kriterien beispielsweise auf immunhistochemischer Basis lassen sich weitere Subtypen bestimmen, die für die Therapiekonzepte entscheidend sein können [5].

Darüber hinaus wird Lungenkrebs nach der TNM-Klassifikation (8. Auflage) der Union for International Cancer Control (UICC) eingeteilt [6], welche die Erkrankung entsprechend der Ausbreitung des Primärtumors, dem Lymphknotenbefall sowie dem Metastasierungsgrad einstuft und verschiedenen Stadien zuordnet [7]. In den Jahren 2013 und 2014 wurden unter allen Lungenkrebs-Erstdiagnosen insgesamt 53 % der Männer und 51 % der Frauen in den Stadien T3 und T4 diagnostiziert. Damit einher geht eine geringe 5-Jahres-Überlebensrate. Diese wurde auf 15 % (Männer) beziehungsweise 20 % (Frauen) beziffert [2]. Eine Früherkennung von Lungenkrebs könnte daher die Prognose verbessern. Bereits 1993 startete die groß angelegte amerikanische „Prostate, Lung, Colorectal, and Ovarian“ (PLCO)-Studie und untersuchte, ob durch ein jährliches Screening mittels Röntgen-Thorax die Lungenkrebsmortalität gesenkt werden kann. Über 150 000 Personen nahmen an der Studie teil. Dabei zeigte sich nach der 4-jährigen Screeningphase und insgesamt 13 Jahren Studienlaufzeit für das Screening mittels Röntgen-Thorax kein statistisch signifikanter Unterschied für die Lungenkrebsmortalität im Vergleich zur Standardversorgung. Die

Stadienverteilung bei Diagnosestellung war vergleichbar zwischen den Gruppen. Das Screening mittels Röntgen-Thorax führte nicht zur Reduzierung von Lungenkrebstumoren in späten Stadien. Eine Stadienverschiebung zu früheren Krankheitsstadien ist ausgeblieben [8].

Gegenwärtig stellt die Niedrigdosis-CT (Low-Dose-CT) ein Verfahren dar, welches für die Früherkennung von Lungenkrebs geeignet sein könnte. Die Bildqualität der CT wurde seit ihrer Einführung ständig weiterentwickelt. Mittels Niedrigdosis-Protokollen kann durch Veränderung verschiedener CT-Parameter wie Röhrenspannung und -strom eine Reduzierung der Strahlendosis erreicht werden bei ausreichender Bildqualität für die Befundung [9]. Die S3-Leitlinie des Leitlinienprogramms Onkologie enthält eine „kann“-Empfehlung für das Screening auf Lungenkrebs mittels Low-Dose-CT für Personen im Alter zwischen 55 und 74 Jahren, die mehr als 30 Packungsjahre konsumiert haben und weniger als 15 Jahre nikotinabstinent waren, oder für Personen ab 50 Jahren, die mehr als 20 Packungsjahre konsumiert haben und einen zusätzlichen Risikofaktor aufweisen, wie z. B. Asbestexposition oder positive Familienanamnese für ein Lungenkarzinom. Die Autoren der S3-Leitlinie weisen darauf hin, dass Daten aus laufenden Studien vor weitergehenden Empfehlungen abgewartet werden sollten [6].

Ein bekanntes Problem des Lungenkrebscreenings mittels Low-Dose-CT ist die hohe Rate falsch-positiver Befunde [6]. Lungenrundherde sind ein häufiger CT-Befund [10]. Für die Malignitätswahrscheinlichkeit eines Rundherds sind die Größe und die Morphologie wesentliche Faktoren. Aber auch andere Faktoren wie zum Beispiel die Lokalisation, die Wachstumsrate sowie Alter und Geschlecht der Person sind zur Abschätzung der Malignitätswahrscheinlichkeit zu berücksichtigen [11]. Bei einem auffälligen Screeningbefund erfolgen weitere Untersuchungen zur Abklärung. Als Methoden stehen die bronchoskopische transbronchiale Biopsie, die CT-gesteuerte perkutane Biopsie und die videoassistierte thoroskopische Biopsie zur Verfügung. Nach der pathologischen Bestätigung der Malignität sind für die sich anschließende Therapie und die Prognose die Differenzierung des Tumortyps (Grading) und die Ausbreitung (Staging) entscheidend [6].

Derzeit existiert in Deutschland kein systematisches Screening auf Lungenkrebs. In diesem Versorgungskontext sind für diese Bewertung daher Studien zum Vergleich Low-Dose-CT-Screening versus kein Screening relevant. Aktuelle systematische Übersichtsarbeiten [12-14], die das Lungenkrebscreening mittels Low-Dose-CT untersuchen, betrachten sowohl Studien, in denen als Vergleichsintervention kein Screening durchgeführt wurde, als auch Studien, in denen als Vergleichsintervention ein Screening mit einer anderen Diagnostik, insbesondere Röntgen-Thorax, durchgeführt wurde. Stützen lässt sich dieses Vorgehen durch die Ergebnisse der PLCO-Studie [8], die Anlass zu der Annahme geben, dass kein Screening und ein Screening mittels Röntgen-Thorax zumindest im Hinblick auf ihren Effekt auf die lungenkrebspezifische Mortalität vergleichbar sind.

2 Fragestellung

Das Ziel der vorliegenden Untersuchung ist

- die Nutzenbewertung eines Screenings auf Lungenkrebs mittels Niedrigdosis-Computertomografie (Low-Dose-CT) im Vergleich zu keinem (bzw. keinem systematischen) Screening

bei Personen mit erhöhtem Lungenkrebsrisiko aufgrund von derzeitigem oder ehemaligem starken Tabakkonsum hinsichtlich patientenrelevanter Endpunkte.

3 Projektverlauf

Der Gemeinsame Bundesausschuss (G-BA) hat am 18.07.2019 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung des Lungenkrebscreenings mittels Niedrigdosis-Computertomografie beauftragt.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Während der Erstellung des Berichtsplans war eine Konsultation von Betroffenen unter anderem zur Diskussion von patientenrelevanten Endpunkten und relevanten Subgruppen vorgesehen. Trotz Anfragen bei verschiedenen Patientenorganisationen kam eine solche Konsultation nicht zustande.

Der vorliegende 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) bekannt gegeben. Stellungnahmen können alle interessierten Personen, Institutionen und Gesellschaften abgeben. Die Stellungnahmen müssen bestimmten formalen Anforderungen genügen. Diese sind ebenfalls auf der Website des IQWiG dargelegt. Kommt es durch die Anhörung zu einer Änderung der Berichtsmethodik, wird ein überarbeiteter Berichtsplan erstellt und veröffentlicht. Die Dokumentation der Anhörung zum Berichtsplan wird auf der Website des IQWiG veröffentlicht.

Auf Basis des Berichtsplans wird die vorläufige Bewertung vorgenommen. Diese wird in einem Vorbericht veröffentlicht, der zur Anhörung gestellt wird. Der Vorbericht wird gegebenenfalls zusätzlich einem externen Review unterzogen. Im Anschluss an die Anhörung zum Vorbericht erstellt das IQWiG einen Abschlussbericht. Dieser Bericht wird an den G-BA übermittelt und grundsätzlich 8 Wochen später auf der Website des IQWiG veröffentlicht. An selber Stelle wird auch die Dokumentation der Anhörung zum Vorbericht veröffentlicht.

Der Zeitplan für alle Arbeitsschritte der Berichterstellung für dieses Projekt ist auf der Website des IQWiG dargelegt.

4 Methoden

Diese Bewertung wird auf Grundlage der Allgemeinen Methoden 5.0 [15] erstellt.

4.1 Kriterien für den Einschluss von Studien in die Untersuchung

4.1.1 Population

In die Bewertung werden Studien mit aktiven oder ehemaligen Raucherinnen und Rauchern ohne Verdacht auf Lungenkrebs aufgenommen.

4.1.2 Prüf- und Vergleichsintervention

Die Prüfindervention ist das Screening auf Lungenkrebs mittels Low-Dose-CT. Als Vergleichsintervention gilt kein (bzw. kein systematisches) Screening. Auch ein Screening auf Lungenkrebs mittels Röntgen-Thorax wird im Sinne einer Sensitivitätsanalyse als Vergleichsintervention berücksichtigt.

Für ein Lungenkrebscreening mittels Röntgen-Thorax gibt es keinen Nutznachweis. Die Ergebnisse der PLCO-Studie geben Anlass zur Annahme der Vergleichbarkeit keines Screenings und eines Screenings mittels Röntgen-Thorax [8]. Daher werden entsprechende Sensitivitätsanalysen geplant (siehe Abschnitt 4.3.5).

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität (Gesamtmortalität, lungenkrebspezifische Mortalität),
- Morbidität (zum Beispiel Auftreten des Lungenkarzinoms),
- gesundheitsbezogene Lebensqualität,
- unerwünschte Ereignisse,
- Schäden, die sich aus der Screeningmaßnahme oder aus nachfolgenden diagnostischen Untersuchungen (zum Beispiel invasive Prozeduren wie Biopsien) ergeben, einschließlich der Konsequenzen aus falschen Screeningbefunden (falsch-positiv und falsch-negativ) und Überdiagnosen.

4.1.4 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 4.1.2 genannten Interventionen und alle unter 4.1.3 genannten Endpunkte ist eine Evaluation im Rahmen von RCTs möglich und praktisch durchführbar.

Für den zu erstellenden Bericht werden daher RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen.

4.1.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

4.1.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfasst sein.

4.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss

Einschlusskriterien	
E1	aktive oder ehemalige Raucherinnen und Raucher ohne Verdacht auf Lungenkrebs (siehe auch Abschnitt 4.1.1)
E2	Prüfintervention: Screening auf Lungenkrebs mittels Low-Dose-CT (siehe auch Abschnitt 4.1.2)
E3	Vergleichsintervention: kein (bzw. kein systematisches) Screening auf Lungenkrebs oder ein Screening auf Lungenkrebs mittels Röntgen-Thorax (siehe auch Abschnitt 4.1.2)
E4	patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert
E5	Studientyp: RCT
E6	Publikationssprache: Deutsch oder Englisch
E7	Vollpublikation verfügbar ^a
<p>a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [16] oder ein Bericht über die Studie, der den Kriterien des CONSORT-Statements [17] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; CT: Computertomografie; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; RCT: randomisierte kontrollierte Studie</p>	

4.1.8 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für die Einschlusskriterien E1 (Population), E2 (Prüfintervention, bezogen auf die Interventionsgruppe der Studie) und E3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Teilnehmerinnen und Teilnehmern diese Kriterien erfüllt sind. Liegen für solche Studien Subgruppenanalysen für Teilnehmerinnen und Teilnehmer vor, die die Einschlusskriterien erfüllen, wird auf diese Analysen zurückgegriffen. Studien, bei denen die Einschlusskriterien E1, E2 und E3 bei weniger als 80 % erfüllt sind, werden nur dann eingeschlossen, wenn Subgruppenanalysen für Teilnehmerinnen und Teilnehmer vorliegen, die die Einschlusskriterien erfüllen.

4.2 Informationsbeschaffung

Mit dem Ziel einer möglichst effizienten Identifizierung relevanter Screeningstudien und durch Nutzung bereits vorliegender wissenschaftlicher Evidenz auf höchster Evidenzstufe wird in einem 1. Schritt eine fokussierte Informationsbeschaffung von systematischen Übersichten durchgeführt. Ziel ist es, 1 oder mehrere hochwertige und aktuelle systematische Übersichten auszuwählen, aus der oder denen Primärstudien identifiziert und anschließend bezüglich der spezifischen Einschlusskriterien des Berichtes selektiert werden. Bei diesem Vorgehen wird von der verwendeten systematischen Übersicht oder den verwendeten systematischen Übersichten allein das Rechercheergebnis, nicht aber die Bewertung der eingeschlossenen Primärstudien oder die Datenextraktion übernommen. In einem 2. Schritt erfolgt dann eine Aktualisierung der Informationsbeschaffung für den Zeitraum, der nicht durch die systematische(n) Übersicht(en) abgedeckt wird.

4.2.1 Fokussierte Informationsbeschaffung von systematischen Übersichten

Die Suche nach relevanten systematischen Übersichten erfolgt in den Datenbanken MEDLINE sowie Cochrane Database of Systematic Reviews und Health Technology Assessment Database mit Einschränkung auf den Publikationszeitraum der letzten 6 Jahre. Es werden hochwertige und aktuelle systematische Übersichten gesucht, die der Fragestellung des Berichts entsprechen. Die Kriterien für die Berücksichtigung dieser systematischen Übersichten orientieren sich dabei an den Kriterien für den Einschluss von Primärstudien (siehe Abschnitt 4.1) und beziehen sich auf die in den systematischen Übersichten für die Studienselektion verwendeten Einschlusskriterien.

4.2.2 Umfassende Informationsbeschaffung

Für die umfassende Informationsbeschaffung wird eine systematische Recherche nach relevanten Studien beziehungsweise Dokumenten durchgeführt. Diese erfolgt für den Zeitraum, der nicht durch die systematische Übersicht abgedeckt wird. Folgende primäre und weitere Informationsquellen sowie Suchtechniken werden dabei berücksichtigt:

Primäre Informationsquellen

- bibliografische Datenbanken
 - MEDLINE
 - Embase
 - Cochrane Central Register of Controlled Trials
- Studienregister
 - U.S. National Institutes of Health. ClinicalTrials.gov
 - World Health Organization. International Clinical Trials Registry Platform Search Portal

Weitere Informationsquellen und Suchtechniken

- Sichten von Referenzlisten identifizierter systematischer Übersichten
- Anhörungen zum Berichtsplan und Vorbericht
- Autorenanfragen

4.2.3 Selektion relevanter Studien

Selektion relevanter Studien bzw. Dokumente aus den Ergebnissen der fokussierten Informationsbeschaffung

Die im Rahmen der fokussierten Informationsbeschaffung identifizierten systematischen Übersichten werden von einer Reviewerin oder einem Reviewer gesichtet und dann hinsichtlich ihrer Relevanz bewertet. Aus der relevanten systematischen Übersicht werden anschließend die jeweils eingeschlossenen Primärstudien extrahiert und auf ihre Relevanz bezüglich der Einschlusskriterien (siehe Tabelle 1) geprüft. Die Selektion der eingeschlossenen Primärstudien aus der systematischen Übersicht erfolgt durch 2 Personen.

Selektion relevanter Studien bzw. Dokumente aus den Ergebnissen der umfassenden Informationsbeschaffung

Die in bibliografischen Datenbanken identifizierten Treffer werden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien (siehe Tabelle 1) bewertet. Als potenziell relevant erachtete Dokumente werden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgen durch 2 Personen unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen beiden aufgelöst.

Rechercheergebnisse aus den folgenden Informationsquellen werden von 2 Personen unabhängig voneinander in Bezug auf ihre Relevanz bewertet:

- Studienregister

Rechercheergebnisse aus den darüber hinaus berücksichtigten Informationsquellen werden von 1 Person in Bezug auf Studien gesichtet. Die identifizierten Studien werden dann auf ihre Relevanz geprüft. Der gesamte Prozess wird anschließend von einer 2. Person überprüft. Sofern in einem der genannten Selektionsschritte Diskrepanzen auftreten, werden diese jeweils durch Diskussion zwischen den beiden aufgelöst.

4.3 Informationsbewertung und -synthese

4.3.1 Bewertung von systematischen Übersichten zu Screeningstudien

Es wird 1 oder werden mehrere aktuelle systematische Übersichten herangezogen, die die Einschlusskriterien des Berichtes erfüllen und die eine hohe Qualität bei der Informationsbeschaffung gewährleisten. Die Qualität der Informationsbeschaffung wird anhand von Punkt 3 der AMSTAR-Checkliste geprüft [18].

4.3.2 Darstellung der Einzelstudien

Alle für die Nutzenbewertung notwendigen Informationen werden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Ergeben sich im Abgleich der Informationen aus unterschiedlichen Dokumenten zu einer Studie (aber auch aus multiplen Angaben zu einem Aspekt innerhalb eines Dokumentes selbst) Diskrepanzen, die auf die Interpretation der Ergebnisse erheblichen Einfluss haben könnten, wird dies an den entsprechenden Stellen im Ergebnisteil des Berichts dargestellt.

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

Die relevanten Ergebnisse werden endpunktspezifisch pro Studie auf ihr jeweiliges Verzerrungspotenzial überprüft. Anschließend werden die Informationen zusammengeführt und analysiert. Wenn möglich werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die unter 4.3.4 bis 4.3.6 beschriebenen Verfahren eingesetzt. Eine abschließende zusammenfassende Bewertung der Informationen erfolgt in jedem Fall.

Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Teilnehmerinnen und Teilnehmer basieren, das heißt, wenn der Anteil der Teilnehmerinnen und Teilnehmer, die nicht in der Auswertung berücksichtigt werden, größer als 30 % ist.

Die Ergebnisse werden auch dann nicht in die Nutzenbewertung einbezogen, wenn der Unterschied der Anteile nicht berücksichtigter Teilnehmerinnen und Teilnehmer zwischen den Gruppen größer als 15 Prozentpunkte ist.

4.3.3 Bewertung des Verzerrungspotenzials der Ergebnisse

Das Verzerrungspotenzial der Ergebnisse wird endpunktspezifisch für jede in die Nutzenbewertung eingeschlossene Studie bewertet. Dazu werden insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Kriterien systematisch extrahiert und bewertet:

A: Kriterien zur endpunktübergreifenden Bewertung des Verzerrungspotenzials der Ergebnisse

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung der Teilnehmerin oder des Teilnehmers sowie der behandelnden Personen
- ergebnisunabhängige Berichterstattung

B: Kriterien zur endpunktspezifischen Bewertung des Verzerrungspotenzials der Ergebnisse

- Verblindung der Endpunkterheber
- Umsetzung des Intention-to-treat(ITT)-Prinzips
- ergebnisunabhängige Berichterstattung

Für die Ergebnisse randomisierter Studien wird das Verzerrungspotenzial zusammenfassend als niedrig oder hoch eingestuft. Wird bereits hinsichtlich der unter (A) aufgeführten Kriterien ein endpunktübergreifend hohes Verzerrungspotenzial festgestellt, gilt dieses damit für alle Ergebnisse aller Endpunkte als hoch, unabhängig von der Bewertung endpunktspezifischer Aspekte. Andernfalls finden anschließend die unter (B) genannten Kriterien pro Endpunkt Berücksichtigung.

4.3.4 Metaanalysen

Die geschätzten Effekte und Konfidenzintervalle aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Anschließend wird die Heterogenität des Studienpools anhand des statistischen Tests auf Vorliegen von Heterogenität [19] untersucht. Ergibt der Heterogenitätstest ein statistisch nicht signifikantes Ergebnis ($p \geq 0,05$), wird davon ausgegangen, dass die Schätzung eines gemeinsamen (gepoolten) Effekts sinnvoll ist. Im Fall von mindestens 5 Studien erfolgt die Metaanalyse mithilfe des Modells mit zufälligen Effekten nach der Methode von Knapp und Hartung unter Verwendung des Heterogenitätsschätzers nach Paule und Mandel [20]. Als Ergebnis wird der gemeinsame Effekt inklusive Konfidenzintervall dargestellt. Weil die Heterogenität im Fall weniger Studien nicht verlässlich geschätzt werden kann, werden bei 4 oder weniger Studien gegebenenfalls Modelle mit festem Effekt verwendet. Dazu müssen die Studien ausreichend ähnlich sein und es darf keine Gründe geben, die gegen die Anwendung eines Modells mit festem Effekt sprechen. Ist ein Modell mit festem Effekt nicht vertretbar, kann eine qualitative Zusammenfassung erfolgen.

Ergibt der Heterogenitätstest ein statistisch signifikantes Ergebnis ($p < 0,05$), wird im Fall von mindestens 5 Studien nur das Prädiktionsintervall dargestellt. Bei 4 oder weniger Studien erfolgt eine qualitative Zusammenfassung. In beiden Fällen wird außerdem untersucht, welche Faktoren diese Heterogenität möglicherweise verursachen. Dazu zählen methodische Faktoren (siehe Abschnitt 4.3.5) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.3.6).

Abgesehen von den genannten Modellen können in bestimmten Situationen und mit besonderer Begründung Alternativen wie z. B. das Betabinomialmodell bei binären Daten [21] angewendet werden.

4.3.5 Sensitivitätsanalysen

Bestehen Zweifel an der Robustheit von Ergebnissen wegen methodischer Faktoren, die beispielsweise durch die Wahl bestimmter Cut-off-Werte, Ersetzungsstrategien für fehlende Werte, Erhebungszeitpunkte oder Effektmaße begründet sein können, ist geplant, den Einfluss solcher Faktoren in Sensitivitätsanalysen zu untersuchen. Das Ergebnis solcher Sensitivitätsanalysen kann die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestuft Effekt kann zum Beispiel dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen (höheren) Nutzen attestiert wird (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt 4.3.7).

Für die Fragestellung dieser Bewertung ist der Vergleich Low-Dose-CT-Screening versus kein Screening relevant. Zusätzlich werden Studien betrachtet, in denen als Vergleichsintervention ein Screening mittels Röntgen-Thorax durchgeführt wurde. Die metaanalytische Zusammenfassung beider Vergleichsinterventionen gemeinsam im Vergleich zum Low-Dose-CT-Screening wird im Sinne einer Sensitivitätsanalyse durchgeführt, um zu prüfen, ob diese Studien zusätzliche Informationen zur Fragestellung des Berichts liefern können.

4.3.6 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Ziel ist es, mögliche Effektunterschiede zwischen Teilnehmergruppen und Screeningspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- beziehungsweise Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Metaregressionen oder Metaanalysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Subgruppenanalysen werden nur durchgeführt, falls jede Subgruppe mindestens 10 Personen umfasst und bei binären Daten mindestens 10 Ereignisse in 1 der Subgruppen aufgetreten sind. Es ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Geschlecht,
- Alter,
- Raucherstatus,
- Tabakkonsum (z. B. Packungsjahre),
- Ausgestaltung des Screenings (zum Beispiel Screeningintervall, Anzahl der Screeningrunden).

Sollten sich aus den verfügbaren Informationen 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 (höheren) Nutzens auf eine spezielle Subgruppe von Teilnehmerinnen und Teilnehmern eingeschränkt werden (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt 4.3.7).

4.3.7 Aussagen zur Beleglage

Für jeden Endpunkt wird eine Aussage zur Beleglage des (höheren) Nutzens und (höheren) Schadens getroffen. Dabei sind 4 Abstufungen der Aussagesicherheit möglich: Es liegt entweder ein Beleg (höchste Aussagesicherheit), ein Hinweis (mittlere Aussagesicherheit), ein Anhaltspunkt (schwächste Aussagesicherheit) oder keine dieser 3 Situationen vor. Der letzte Fall tritt ein, wenn keine Daten vorliegen oder die vorliegenden Daten keine der 3 übrigen Aussagen zulassen. In diesem Fall wird die Aussage „Es liegt kein Anhaltspunkt für einen (höheren) Nutzen oder (höheren) Schaden vor“ getroffen.

Die regelhaft abzuleitende Aussagesicherheit ist von den in Tabelle 2 dargestellten Kriterien abhängig. Die qualitative Ergebnissicherheit ist abhängig vom Design der Studie. Ergebnisse randomisierter Studien mit niedrigem Verzerrungspotenzial haben eine hohe, Ergebnisse randomisierter Studien mit hohem Verzerrungspotenzial eine mäßige qualitative Ergebnissicherheit. Ergebnisse nicht randomisierter vergleichender Studien haben eine geringe qualitative Ergebnissicherheit.

Tabelle 2: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit

		Anzahl Studien				
		1 (mit statistisch signifikantem Effekt)	≥ 2			
			homogen	heterogen		
			Metaanalyse statistisch signifikant	gleichgerichtete Effekte ^a		
				deutlich	mäßig	nein
qualitative Ergebnis- sicherheit	hoch	Hinweis	Beleg	Beleg	Hinweis	–
	mäßig	Anhaltspunkt	Hinweis	Hinweis	Anhaltspunkt	–
	gering	–	Anhaltspunkt	Anhaltspunkt	–	–
a: Gleichgerichtete Effekte liegen vor, wenn trotz Heterogenität eine deutliche oder mäßige Richtung der Effekte erkennbar ist.						

5 Literatur

1. Bösch D. Lunge und Atemwege. Berlin: Springer; 2014.
2. Robert Koch-Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland (Ed). Krebs in Deutschland für 2013/2014. Berlin: RKI; 2017. URL: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/kid_2017/krebs_in_deutschland_2017.pdf?_blob=publicationFile.
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68(6): 394-424.
4. Robert Koch-Institut (Ed). Bericht zum Krebsgeschehen in Deutschland 2016. Berlin: RKI; 2016. URL: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebsgeschehen/Krebsgeschehen_download.pdf?_blob=publicationFile.
5. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB et al. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J Thorac Oncol* 2015; 10(9): 1243-1260.
6. Leitlinienprogramm Onkologie der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften, Deutsche Krebsgesellschaft, Deutsche Krebshilfe. S3-Leitlinie Prävention, Diagnostik, Therapie und Nachsorge des Lungenkarzinoms: Langversion 1.0; AWMF-Registernummer 020/007OL [online]. 02.2018 [Zugriff: 09.07.2019]. URL: https://www.awmf.org/uploads/tx_szleitlinien/020-007OL_1_S3_Lungenkarzinom_2018-03.pdf.
7. Wittekind C (Ed). TNM-Klassifikation maligner Tumore. Weinheim: Wiley-VCH; 2017.
8. Oken MM, Hocking WG, Kvale PA, Andriole GL, Buys SS, Church TR et al. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. *JAMA* 2011; 306(17): 1865-1873.
9. Diederich S, Lenzen Z, Puskas A, Koch AT, Yelbuz TM, Eameri M et al. Niedrigdosiscomputertomographie des Thorax: experimentelle und klinische Untersuchungen. *Radiologe* 1996; 36(6): 475-482
10. MacMahon H, Austin JH, Gamsu G, Herold CJ, Jett JR, Naidich DP et al. Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society. *Radiology* 2005; 237(2): 395-400.
11. MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR et al. Guidelines for management of incidental pulmonary nodules detected on ct images: from the Fleischner Society 2017. *Radiology* 2017; 284(1): 228-243.
12. Mazzone PJ, Silvestri GA, Patel S, Kanne JP, Kinsinger LS, Wiener RS et al. Screening for lung cancer: CHEST guideline and expert panel report. *Chest* 2018; 153(4): 954-985.

13. Snowsill T, Yang H, Griffin E, Long L, Varley-Campbell J, Coelho H et al. Low-dose computed tomography for lung cancer screening in high-risk populations: a systematic review and economic evaluation. *Health Technol Assess* 2018; 22(69): 1-276.
14. Wang X, Liu H, Shen Y, Li W, Chen Y, Wang H. Low-dose computed tomography (LDCT) versus other cancer screenings in early diagnosis of lung cancer: a meta-analysis. *Medicine (Baltimore)* 2018; 97(27): e11233.
15. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 5.0. Köln: IQWiG; 2017. URL: https://www.iqwig.de/download/Allgemeine-Methoden_Version-5-0.pdf.
16. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3; current step 4 version [online]. 30.11.1995 [Zugriff: 05.09.2018]. URL: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf.
17. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ et al. CONSORT 2010: explanation and elaboration; updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c869.
18. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007; 7: 10.
19. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for meta-analysis in medical research*. Chichester: Wiley; 2000.
20. Veroniki AA, Jackson D, Viechtbauer W, Bender R, Knapp G, Kuss O et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. *Cochrane Database Syst Rev* 2015; (Suppl 1): 25-27.
21. Kuss O. Statistical methods for meta-analyses including information from studies without any events: add nothing to nothing and succeed nevertheless. *Stat Med* 2015; 34(7): 1097-1116.