

Prostatakrebscreening mittels PSA-Test

Berichtsplan

Auftrag: S19-01
Version: 1.0
Stand: 16.05.2019

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Prostatakrebscreening mittels PSA-Test

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

18.04.2019

Interne Auftragsnummer:

S19-01

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, Prostatatumoren, Nutzenbewertung, Systematische Übersicht

Keywords: Mass Screening, Prostatic 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 Kriterien für den Einschluss systematischer Übersichten	7
4.3 Informationsbeschaffung	7
4.3.1 Fokussierte Informationsbeschaffung nach systematischen Übersichten	8
4.3.2 Umfassende Informationsbeschaffung	8
4.3.3 Selektion relevanter Studien.....	8
4.4 Informationsbewertung und Synthese	9
4.4.1 Bewertung von systematischen Übersichten zu Screeningstudien.....	9
4.4.2 Darstellung der Einzelstudien.....	9
4.4.3 Bewertung des Verzerrungspotenzials der Ergebnisse.....	10
4.4.4 Metaanalysen	11
4.4.5 Sensitivitätsanalysen	11
4.4.6 Subgruppenmerkmale und andere Effektmodifikatoren	11
4.4.7 Aussagen zur Beleglage	12
5 Literatur	14

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss.....	6
Tabelle 2: Übersicht über die Kriterien für den Einschluss systematischer Übersichten	7
Tabelle 3: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit.....	13

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AHRQ	Agency for Healthcare Research and Quality
CAP	Comparison Arm for ProtecT
DRU	digital-rektale Untersuchung
G-BA	Gemeinsamer Bundesausschuss
HTA	Health Technology Assessment
IGeL	individuelle Gesundheitsleistung
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
NICE	National Institute for Clinical Excellence and Care
PSA	prostataspezifisches Antigen
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
USPSTF	U.S. Preventive Services Task Force

1 Hintergrund

Das Prostatakarzinom ist eine bösartige Veränderung der Prostata (Vorsteherdrüse), die mit Fortschreiten der Erkrankung das unmittelbar benachbarte Gewebe (Samenblasen, Harnblase, Dickdarm) infiltrieren und Fernmetastasen bilden kann. Gemessen an der Neuerkrankungsrate stellt das Prostatakarzinom in Deutschland mit 23,0 % aller Krebserkrankungen die häufigste Tumorerkrankung des Mannes dar. Nach einer Schätzung des Robert Koch-Instituts wurde im Jahr 2014 in Deutschland bei ca. 57 370 Männern die Diagnose erstmals gestellt [1]. Als wichtigster Risikofaktor für die Entstehung eines Prostatakarzinoms wird das Alter angesehen [1,2]. So tritt das Prostatakarzinom mit einem medianen Erkrankungsalter von 72 Jahren überwiegend im fortgeschrittenen Lebensalter in Erscheinung, wohingegen es vor dem 45. bis 50. Lebensjahr kaum auftritt [1].

Pro Jahr sterben in Deutschland knapp 14 000 Männer an den Folgen eines Prostatakarzinoms [1]. Die Prognose der Erkrankung hängt entscheidend vom Stadium seiner Ausbreitung sowie von der feingeweblichen Typisierung des Tumors hinsichtlich des Zelltyps, Entartungsgrads der Tumorzellen und der Veränderungen der Zellstruktur (häufig bewertet anhand des Gleason-Scores) ab [2]. Während Männer mit lokal begrenztem Prostatakarzinom und einem niedrigen Gleason-Score auch ohne sofortige invasive Behandlung eine günstige Prognose aufweisen, wird davon ausgegangen, dass Männer mit einem metastasierten Prostatakarzinom nicht kurativ behandelt werden können [2]. Vom Screening des Prostatakarzinoms verspricht man sich die Entdeckung von Prostatakarzinomen mit einem hohen Progressionsrisiko in einem heilbaren Stadium, um die Morbidität (zum Beispiel Schmerzen aufgrund von Knochenmetastasen) und Mortalität zu reduzieren, die mit metastasierten Prostatakarzinomen verbunden sind [2].

In Deutschland kommen derzeit 2 Screeningtests zum Einsatz: die digital-rektale Untersuchung (DRU) und der Test auf das prostataspezifische Antigen (PSA). Während die DRU im gesetzlichen Früherkennungsangebot für Männer ab 45 Jahren enthalten ist und somit von den gesetzlichen Krankenkassen erstattet wird, steht der PSA-Test ausschließlich als individuelle Gesundheitsleistung (IGeL) zur Verfügung.

Screeningmaßnahmen können nicht unerhebliche Schäden nach sich ziehen. Im Falle des PSA-Screenings sind insbesondere die hohe Anzahl falsch-positiver Testergebnisse, die mit nachfolgenden invasiven Biopsien verbunden sind, sowie der beträchtliche Anteil von Überdiagnosen, die häufig Übertherapien nach sich ziehen, zu nennen [4-6]. Bislang wird kontrovers diskutiert, ob der Nutzen des PSA-Screenings den Schaden überwiegt [4-7]. Dies spiegelt sich darin, dass sich aktuell weltweit nahezu alle der jeweils zuständigen nationalen Gesundheitsbehörden und auch Fachgesellschaften gegen ein organisiertes populationsbasiertes PSA-Screening aussprechen [8], gleichzeitig aber häufig eine individuelle Entscheidung der Betroffenen befürworten, die aus einem gemeinsamen Entscheidungsfindungsprozess mit einer Ärztin oder einem Arzt (Shared Decision Making) resultiert oder mithilfe einer Entscheidungshilfe getroffen wurde [2,9,10]. Kürzlich hat die U.S. Preventive Services Task Force (USPSTF) ihre Empfehlung von einer grundsätzlichen Ablehnung des

Screenings in eine Empfehlung für Shared Decision Making bei Männern in einem Alter von 55 bis 69 Jahren geändert [9] und es wurden die Ergebnisse der größten Studie zum PSA-Screening (CAP) veröffentlicht [11].

2 Fragestellung

Das Ziel der vorliegenden Untersuchung ist

- die Nutzenbewertung des Prostatakarzinomscreenings mittels PSA-Test

bei Teilnehmern ohne Verdacht auf ein Prostatakarzinom hinsichtlich patientenrelevanter Endpunkte.

3 Projektverlauf

Der Gemeinsame Bundesausschuss (G-BA) hat am 18.04.2019 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung des Prostatakrebscreenings mittels PSA-Test beauftragt.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

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 [12] erstellt.

4.1 Kriterien für den Einschluss von Studien in die Untersuchung

4.1.1 Population

In die Bewertung werden Studien mit Männern ohne Verdacht auf ein Prostatakarzinom aufgenommen.

4.1.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention bildet ein Screening mittels PSA-Test auf das Vorliegen eines Prostatakarzinoms. Die Behandlung eines Prostatakarzinoms muss in beiden Studienarmen vergleichbar sein, das heißt, vergleichbare Patienten (z. B. gleiches Stadium bzw. gleicher Tumorgad) müssen vergleichbare Behandlungen erhalten, unabhängig davon, ob das Prostatakarzinom beim betroffenen Mann über ein Screening mittels PSA-Test oder auf anderem Weg identifiziert wurde.

Als Vergleichsintervention gilt kein Screening mittels PSA-Test.

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Gesamtmortalität,
- prostatakarzinomspezifische Mortalität,
- Auftreten metastatischer Prostatakarzinome,
- Morbidität (zum Beispiel Schmerzen aufgrund von Knochenmetastasen),
- gesundheitsbezogene Lebensqualität,
- unerwünschte Ereignisse,
- Schäden, die sich direkt und indirekt aus dem Screening ergeben, einschließlich der Konsequenzen aus falschen Screeningbefunden und Überdiagnosen (zum Beispiel Folgen von Operationen und anderen Therapien).

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 verfügbar 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	Männer ohne Verdacht auf Prostatakrebs (siehe auch Abschnitt 4.1.1)
E2	Prüfintervention: Screening auf Prostatakrebs mittels PSA-Test (siehe auch Abschnitt 4.1.2)
E3	Vergleichsintervention: kein Screening auf Prostatakrebs mittels PSA-Test (siehe auch Abschnitt 4.1.2)
E4	patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert
E5	RCTs
E6	Publikationssprache: Englisch oder Deutsch
E7	Vollpublikation verfügbar ^a

a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [13] oder ein Bericht über die Studie, der den Kriterien des CONSORT-Statements [14] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.

CONSORT: Consolidated Standards of Reporting Trials; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; PSA: prostataspezifisches Antigen; RCT: randomisierte kontrollierte Studie

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 Probanden diese Kriterien erfüllt sind. Liegen für solche Studien Subgruppenanalysen für Probanden 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 Probanden vorliegen, die die Einschlusskriterien erfüllen.

4.2 Kriterien für den Einschluss systematischer Übersichten

Im Rahmen der Informationsbeschaffung werden in einem 1. Schritt hochwertige und aktuelle systematische Übersichten gesucht, die der Fragestellung des Berichts entsprechen (siehe Abschnitt 4.3). Die Kriterien für den Einschluss 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 Studienselktion verwendeten Einschlusskriterien.

In der folgenden Tabelle sind die Kriterien aufgelistet, die systematische Übersichten erfüllen müssen, um in die Bewertung eingeschlossen und hinsichtlich relevanter Studien beziehungsweise Dokumente gesichtet zu werden.

Tabelle 2: Übersicht über die Kriterien für den Einschluss systematischer Übersichten

Einschlusskriterien	
SR1	Als Zielpopulation wird in den Einschlusskriterien der SÜ definiert: erwachsene Männer (siehe auch Abschnitt 4.1.1)
SR2	Als Prüflintervention wird in den Einschlusskriterien der SÜ definiert: Screening auf Prostatakrebs mittels PSA-Test (siehe auch Abschnitt 4.1.2)
SR3	Als Vergleichsintervention wird in den Einschlusskriterien der SÜ definiert: kein Screening auf Prostatakrebs mittels PSA-Test (siehe auch Abschnitt 4.1.2)
SR4	Als patientenrelevante Endpunkte werden in den Einschlusskriterien der SÜ definiert: patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert
SR5	systematische Übersichten von RCTs
SR6	Publikationszeitraum: letzte 6 Jahre
SR7	Publikationssprache: Englisch oder Deutsch
SR8	Vollpublikation verfügbar ^a
<p>a: Als Vollpublikation gilt in diesem Zusammenhang ein Bericht, der den Kriterien des PRISMA-Statements [15] genügt. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PSA: prostataspezifisches Antigen; RCT: randomisierte kontrollierte Studie; SÜ: systematische Übersicht</p>	

4.3 Informationsbeschaffung

Der Pool aus abgeschlossenen und laufenden RCTs zum PSA-Screening ist seit Jahren bekannt. Mit dem Ziel einer möglichst effizienten Ergebniserhebung und Nutzung bereits vorliegender wissenschaftlicher Evidenz auf höchster Evidenzstufe wird in einem 1. Schritt eine fokussierte Informationsbeschaffung nach systematischen Übersichten durchgeführt. Ziel ist es, eine oder ggf. mehrere hochwertige und aktuelle systematische Übersichten auszuwählen, deren Primärstudien extrahiert und dann selektiert werden. Bei diesem Vorgehen wird von der verwendeten systematischen Übersicht 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 Übersicht oder die systematischen Übersichten abgedeckt wird.

4.3.1 Fokussierte Informationsbeschaffung nach systematischen Übersichten

Es erfolgt eine Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE sowie Cochrane Database of Systematic Reviews und Health Technology Assessment Database mit Einschränkung auf den Publikationszeitraum der letzten 6 Jahre. Zudem werden die Websites von HTA-Agenturen wie dem NICE oder der AHRQ nach systematischen Übersichten durchsucht.

4.3.2 Umfassende Informationsbeschaffung

Für die systematische Recherche nach relevanten Studien beziehungsweise Dokumenten für den Zeitraum, der nicht durch die systematische Übersicht oder die systematischen Übersichten abgedeckt wird, werden folgende Informationsquellen herangezogen:

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

- durch den G-BA übermittelte Dokumente
- Anhörungen zum Berichtsplan und Vorbericht
- Autorenanfragen

4.3.3 Selektion relevanter Studien

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

Die im Rahmen der fokussierten Informationsbeschaffung identifizierten systematischen Übersichten werden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der spezifischen Einschlusskriterien (siehe Tabelle 2) bewertet. Als potenziell relevant erachtete Dokumente werden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Aus den relevanten systematischen Übersichten

werden anschließend die jeweils eingeschlossenen Primärstudien extrahiert und auf ihre Relevanz bezüglich der Einschlusskriterien (siehe Tabelle 1) geprüft. Die gesamte oben beschriebene Selektion erfolgt durch 1 Reviewerin bzw. 1 Reviewer.

Selektion relevanter Studien bzw. Dokumente aus den Ergebnissen der bibliografischen Recherche

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.

Selektion relevanter Studien bzw. Dokumente aus weiteren Informationsquellen

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

- Studienregister,
- durch den G-BA übermittelte Dokumente.

Rechercheergebnisse aus den darüber hinaus berücksichtigten Informationsquellen werden von 1 Person 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.

4.4 Informationsbewertung und Synthese

4.4.1 Bewertung von systematischen Übersichten zu Screeningstudien

Es werden systematische Übersichten herangezogen, die die Einschlusskriterien des Berichtes erfüllen und die eine hohe Qualität bei der Informationsbeschaffung gewährleisten. Bei mehreren gleichwertigen Übersichtsarbeiten werden die aktuellsten herangezogen.

4.4.2 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 4.4.4 bis 4.4.6 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 Probanden basieren, das heißt, wenn der Anteil der Probanden, 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 Probanden zwischen den Gruppen größer als 15 Prozentpunkte ist.

4.4.3 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 des Probanden 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.

4.4.4 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 [16] 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-Mandel [17]. 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 4.4.5) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.4.6).

Abgesehen von den genannten Modellen können in bestimmten Situationen und mit besonderer Begründung Alternativen wie z. B. das Betabinomialmodell bei binären Daten [18] angewendet werden.

4.4.5 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 eingestuftes Ergebnis 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 4.4.7).

4.4.6 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 Probandengruppen und Behandlungsspezifika 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 einer der Subgruppen aufgetreten sind. Es ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Alter,
- Ausgestaltung des Screenings (z. B. PSA-Cut-off, Dauer des Screenings, Anzahl der Screeningrunden, Screeningintervall).

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 Probanden eingeschränkt werden (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt 4.4.7).

4.4.7 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 3 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 3: 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. Robert Koch-Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland (Ed). Krebs in Deutschland für 2013/2014. Berlin: RKI; 2017.
2. Leitlinienprogramm Onkologie der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften, Deutschen Krebsgesellschaft, Deutschen Krebshilfe. Interdisziplinäre Leitlinie der Qualität S3 zur Früherkennung, Diagnose und Therapie der verschiedenen Stadien des Prostatakarzinoms, Langversion 5.0 [online]. 04.2018 [Zugriff: 13.06.2018]. URL: https://www.leitlinienprogramm-onkologie.de/fileadmin/user_upload/Downloads/Leitlinien/Prostata_5_0/LL_Prostata_Langversion_5.0.pdf.
3. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. *BMJ* 2015; 350: g7773.
4. Fenton JJ, Weyrich MS, Durbin S, Liu Y, Bang H, Melnikow J. Prostate-specific antigen-based screening for prostate cancer: a systematic evidence review for the U.S. Preventive Services Task Force; AHRQ publication no. 17-05229-EF-1 [online]. 05.2018 [Zugriff: 23.04.2019]. (AHRQ Evidence Syntheses; Band 154). URL: https://www.ncbi.nlm.nih.gov/books/NBK518890/pdf/Bookshelf_NBK518890.pdf.
5. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. *Cochrane Database Syst Rev* 2013; (1): CD004720.
6. Ilic D, Djulbegovic M, Jung JH, Hwang EC, Zhou Q, Cleves A et al. Prostate cancer screening with prostate-specific antigen (PSA) test: a systematic review and meta-analysis. *BMJ* 2018; 362: k3519.
7. Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet* 2014; 384(9959): 2027-2035.
8. Ebell MH, Thai TN, Royalty KJ. Cancer screening recommendations: an international comparison of high income countries. *Public Health Rev* 2018; 39: 7.
9. Grossman DC, Curry SJ, Owens DK, Bibbins-Domingo K, Caughey AB, Davidson KW et al. Screening for prostate cancer: US Preventive Services Task Force recommendation statement. *JAMA* 2018; 319(18): 1901-1913.
10. Public Health England. Prostate cancer risk management programme (PCRMP): benefits and risks of PSA testing [online]. 29.03.2016 [Zugriff: 05.04.2019]. URL: <https://www.gov.uk/government/publications/prostate-cancer-risk-management-programme-psa-test-benefits-and-risks/prostate-cancer-risk-management-programme-pcrmp-benefits-and-risks-of-psa-testing#population-screening-for-prostate-cancer>.
11. Martin RM, Donovan JL, Turner EL, Metcalfe C, Young GJ, Walsh EI et al. Effect of a low-intensity PSA-based screening intervention on prostate cancer mortality: the CAP randomized clinical trial. *JAMA* 2018; 319(9): 883-895.

12. 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.
13. 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.
14. 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.
15. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; 339: b2535.
16. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for meta-analysis in medical research*. Chichester: Wiley; 2000.
17. 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.
18. 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.