

Risikoadaptiertes PSA- und MRT-Screening zur Früherkennung eines Prostatakarzinoms

A horizontal bar composed of 18 squares of varying shades of blue and grey, spanning the width of the page. The word 'BERICHTSPLAN' is centered in white text on a dark blue segment of the bar.

BERICHTSPLAN

Projekt: S25-02

Version: 1.0

Stand: 05.02.2026

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Risikoadaptiertes PSA- und MRT-Screening zur Früherkennung eines Prostatakarzinoms

Auftraggeber

Gemeinsamer Bundesausschuss

Datum des Auftrags

16.10.2025

Interne Projektnummer

S25-02

Anschrift des Herausgebers

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Siegburger Str. 237
50679 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: info@iqwig.de

Internet: www.iqwig.de

Schlagwörter

Reihenuntersuchung, Prostatatumoren, Nutzenbewertung, Systematische Übersicht

Keywords

Mass Screening, Prostatic Neoplasms, Benefit Assessment, Systematic Review

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	v
Abkürzungsverzeichnis.....	vi
1 Hintergrund.....	1
2 Fragestellung.....	3
3 Projektverlauf.....	4
4 Methoden	5
4.1 Kriterien für den Einschluss von vergleichenden Interventionsstudien zur Untersuchung von Screeningalgorithmen.....	6
4.1.1 Population	6
4.1.2 Prüf- und Vergleichsintervention	6
4.1.3 Patientenrelevante Endpunkte	7
4.1.4 Studientypen	7
4.1.5 Studiendauer	7
4.1.6 Publikationssprache	7
4.1.7 Setting.....	8
4.1.8 Tabellarische Darstellung der Kriterien für den Studieneinschluss	8
4.2 Empirische Prüfung der Annahmen des PSA- / MRT-Screenings bei fehlender Evidenz aus Interventionsstudien zu Screeningalgorithmen	8
4.2.1 Kriterien für den Einschluss von Studien zu Hypothese 1.....	10
4.2.1.1 Population.....	10
4.2.1.2 Prüf- und Vergleichsintervention	10
4.2.1.3 Patientenrelevante Endpunkte.....	10
4.2.1.4 Studientypen.....	10
4.2.1.5 Studiendauer.....	11
4.2.1.6 Studiengröße.....	11
4.2.1.7 Publikationssprache	11
4.2.1.8 Setting	11
4.2.1.9 Tabellarische Darstellung der Kriterien für den Studieneinschluss.....	11
4.2.2 Kriterien für den Einschluss von Studien zu Hypothese 2.....	12
4.2.2.1 Population.....	12
4.2.2.2 Prüf- und Vergleichsintervention	12
4.2.2.3 Zielgrößen	13

4.2.2.4	Studientypen.....	13
4.2.2.5	Studiendauer.....	13
4.2.2.6	Publikationssprache.....	13
4.2.2.7	Setting.....	13
4.2.2.8	Tabellarische Darstellung der Kriterien für den Studieneinschluss.....	13
4.3	Informationsbeschaffung.....	14
4.3.1	Fokussierte Informationsbeschaffung von systematischen Übersichten zu vergleichenden Interventionsstudien zur Untersuchung von Screeningalgorithmen	14
4.3.2	Umfassende Informationsbeschaffung von vergleichenden Interventionsstudien zur Untersuchung von Screeningalgorithmen.....	15
4.3.3	Informationsbeschaffung zur empirischen Prüfung der Annahmen des PSA- / MRT-Screenings bei fehlender Evidenz aus Interventionsstudien zu Screeningalgorithmen	16
4.3.4	Anwendung von Limitierungen auf Datenbankebene	17
4.3.5	Selektion relevanter Studien aus der umfassenden Informationsbeschaffung...	17
4.3.6	Selektion relevanter Studien aus der fokussierten Informationsbeschaffung	18
4.4	Informationsbewertung und -synthese	18
4.4.1	Darstellung der Einzelstudien.....	18
4.4.1.1	Vergleichende Interventionsstudien zur Untersuchung von Screeningalgorithmen bzw. zur Therapie eines Prostatakarzinoms	18
4.4.1.2	Prospektiv geplante Kohortenstudien (Prognosestudien)	19
4.4.2	Bewertung des Verzerrungspotenzials der Ergebnisse.....	19
4.4.2.1	Vergleichende Interventionsstudien zur Untersuchung von Screeningalgorithmen bzw. zur Therapie eines Prostatakarzinoms	19
4.4.2.2	Prospektiv geplante Kohortenstudien (Prognosestudien)	20
4.4.3	Metaanalysen	20
4.4.4	Sensitivitätsanalysen	21
4.4.5	Subgruppenmerkmale und andere Effektmodifikatoren.....	22
4.4.6	Aussagen zur Beleglage	22
5	Literatur	24
Anhang A	Suchstrategien in bibliografischen Datenbanken.....	26

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Studieneinschluss.....	8
Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss zur Überprüfung der Hypothese 1.....	12
Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss zur Überprüfung der Hypothese 2.....	14
Tabelle 4: Regelmäßig abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit	23

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AHRQ	Agency for Healthcare Research and Quality
DRU	digital-rektale Untersuchung
EAU	European Association of Urology
EBRT	External Beam Radiotherapy (externe Strahlentherapie)
G-BA	Gemeinsamer Bundesausschuss
HTA	Health Technology Assessment
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ISUP	International Society of Urological Pathology
ITT	Intention to treat
MRT	Magnetresonanztomografie
NICE	National Institute for Health and Care Excellence
PI-RADS	Prostate Imaging-Reporting and Data System
PSA	prostataspezifisches Antigen
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RKI	Robert Koch-Institut
SÜ	systematische Übersicht

1 Hintergrund

Das Prostatakarzinom ist eine bösartige Erkrankung der Vorsteherdrüse (Prostata), die durch ein unkontrolliertes Wachstum der Epithelzellen in den Drüsenläppchen entsteht. Mit fortschreitender Erkrankung kann das Karzinom in benachbartes Gewebe – etwa die Samenblasen, Harnblase oder das Rektum – einwachsen und Fernmetastasen bilden.

Prostatakrebs kann bei allen Personen auftreten, die (funktionales) Prostatagewebe besitzen. Dies betrifft in erster Linie (cisgender) Männer, kann jedoch auch Transfrauen, nichtbinäre oder intergeschlechtliche Personen einschließen, sofern eine Prostata vorhanden ist. Im Bericht wird daher die neutrale Bezeichnung „Personen“ verwendet.

Nach Angaben des Robert Koch-Instituts (RKI) wurden im Jahr 2022 etwa 75 000 Neuerkrankungen eines Prostatakarzinoms diagnostiziert sowie im Jahr 2023 etwa 15 000 Todesfälle registriert [1]. Ein wichtiger Risikofaktor für die Entstehung eines Prostatakarzinoms ist das Alter: Unter 50 Jahren tritt die Erkrankung selten auf, die meisten Diagnosen werden bei über 70-Jährigen gestellt [1]. Auch familiäre bzw. genetische Faktoren (z. B. BRCA2-Mutationen) gelten als Risikofaktoren [2,3].

In der Literatur wird davon ausgegangen, dass die Prognose vom Tumorstadium sowie von der feingeweblichen Charakterisierung abhängt, insbesondere vom Gleason-Score bzw. ISUP(International Society of Urological Pathology)-Grad, welche den Entartungsgrad der Tumorzellen beschreiben [4]. Personen mit lokal begrenztem Prostatakarzinom und niedrigen Gleason- bzw. ISUP-Werten haben häufig auch mit aktiver Überwachung („active surveillance“) eine gute Prognose [5].

Ziel eines Screenings auf Prostatakarzinom ist die frühzeitige Entdeckung klinisch relevanter, aggressiver Tumoren, die ohne Behandlung zu Metastasen, symptomatischer Erkrankung oder Tod führen könnten. Eine frühzeitige Diagnose soll eine rechtzeitige Therapie ermöglichen und dadurch Krankheitsprogression und Mortalität senken.

Dem möglichen Nutzen stehen jedoch erhebliche Risiken und Unsicherheiten gegenüber. Dazu gehören insbesondere Überdiagnosen (Erkennung von indolenten, klinisch irrelevanten Tumoren, die nie symptomatisch geworden wären), Übertherapien (unnötige Behandlungen mit potenziellen Nebenwirkungen wie Inkontinenz oder erektile Dysfunktion) und falsch-positive Testergebnisse, die zu psychischer Belastung und invasiven Folgeeingriffen (Biopsien) führen können. Eine Nutzenbewertung des IQWiG zum Prostatakarzinomscreening mittels Bestimmung des prostataspezifischen Antigens (PSA) (und ggf. anschließender Biopsie) zeigte zwar einen Nutzen hinsichtlich der prostatakarzinomspezifischen Mortalität, jedoch auch einen Schaden durch Überdiagnosen. Insgesamt kam das Institut zu dem Schluss, dass der Schaden einer solchen Screeningstrategie den Nutzen überwiegt [6]. Im Rahmen der

damaligen Bewertung wurden bereits verschiedene mögliche Maßnahmen zur Verringerung der Screeningschäden diskutiert, insbesondere eine Anpassung der Screeningstrategie.

In Deutschland können Personen ab 45 Jahren jährlich eine gesetzliche Früherkennungsuntersuchung auf Prostatakrebs in Anspruch nehmen. Diese umfasst derzeit ein ärztliches Beratungsgespräch und eine digital-rektale Untersuchung (DRU).

Die im Sommer 2025 aktualisierte S3-Leitlinie empfiehlt die DRU nicht mehr als Screeningmethode, da ihre Sensitivität und Spezifität für die Früherkennung unzureichend sei [4]. Stattdessen wird ein risikoadaptierter Ansatz vorgeschlagen, der auf der Bestimmung des PSA basiert: Ab einem PSA-Wert ≥ 3 ng/ml soll eine weiterführende Diagnostik erfolgen, bestehend aus einer urologischen Konsultation und Risikobewertung (z. B. mithilfe der Risk Calculators 3 und 4 der Prostate Cancer Research Foundation) sowie regelhaft einer Magnetresonanztomografie (MRT) der Prostata. Abhängig vom MRT-Befund (insbesondere dem PI-RADS[Prostate Imaging-Reporting and Data System]-Scores) soll eine gezielte und eine zusätzliche systematische Prostatabiopsie durchgeführt werden, um eine histologische Sicherung der Diagnose zu ermöglichen.

Ziel dieses Screeningalgorithmus ist es, die mit dem PSA-basierten Screening verbundenen Schäden (insbesondere durch Überdiagnosen und daraus folgenden Übertherapien) zu reduzieren ohne den mit dem PSA-basierten Screening verbundenen Mortalitätsvorteil zu beeinträchtigen. Der Ansatz beruht auf der Annahme, dass die MRT eine ausreichend verlässliche Differenzierung zwischen Personen ermöglicht, die von einer nachfolgenden Biopsie sowie einer daraus möglichen Tumordiagnose und -therapie profitieren, und solchen Personen, die davon nicht profitieren.

Mehrere internationale Leitlinien beschreiben ebenfalls einen risikoadaptierten Screeningansatz, bei dem PSA-Tests mit individueller Risikostratifizierung kombiniert werden und bei auffälligen PSA-Werten eine MRT-Untersuchung als Triage-Test vor einer möglichen Biopsie eingesetzt wird (bspw. die Leitlinie der European Association of Urology [EAU]; [7]). In mehreren Ländern wird derzeit die Einführung eines organisierten, risikoadaptierten Screenings auf Prostatakarzinom diskutiert oder bereits pilotiert (bspw. Finnland [8] oder Schweden [9]). Diese Programme sollen die Machbarkeit, Akzeptanz und diagnostische Effizienz eines risikoadaptierten, stufenweisen Vorgehens im realen Versorgungskontext prüfen.

2 Fragestellung

Das Ziel der vorliegenden Untersuchung ist

- die Nutzenbewertung eines risikoadaptierten PSA- und MRT-Screenings zur Früherkennung eines Prostatakarzinoms im Vergleich zu keiner spezifischen Prostatakarzinom-Früherkennung

bei Personen im Alter von 50 bis 70 Jahren ohne Verdacht auf ein Prostatakarzinom.

3 Projektverlauf

Der Gemeinsame Bundesausschuss (G-BA) hat am 16.10.2025 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung des risikoadaptierten PSA- und MRT-Screenings zur Früherkennung eines Prostatakarzinoms beauftragt.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Im Rahmen der Projektbearbeitung werden Betroffene eingebunden, um einen Eindruck von ihren Erwartungen an ein Prostatakrebscreening zu gewinnen.

Auf Basis des Berichtsplans wird die vorläufige Bewertung vorgenommen. Diese wird in einem Vorbericht veröffentlicht, der zur Anhörung gestellt wird. Im Anschluss an die Anhörung zum Vorbericht erstellt das IQWiG einen Abschlussbericht. Dieser Bericht wird an den G-BA übermittelt und grundsätzlich 4 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 8.0 [10] erstellt.

Die Bewertung eines risikoadaptierten PSA- und MRT-Screenings zur Früherkennung eines Prostatakarzinoms (im Folgenden: PSA- / MRT-Screening) erfolgt folgendermaßen:

Ansatz (1)

Primär soll die Bewertung des Nutzens und Schadens mithilfe vergleichender Interventionsstudien erfolgen, die den Effekt des gesamten Screeningalgorithmus – bestehend aus PSA-Test und MRT-Diagnostik (ggf. ergänzt durch Risikoscores) – gegenüber keiner spezifischen Prostatakarzinom-Früherkennung (ggf. ergänzt durch eine DRU) auf patientenrelevante Endpunkte direkt untersuchen (siehe Abschnitt 4.1).

Ansatz (2)

Liegen derartige Studien zu einem direkten Vergleich nicht in ausreichender Zahl oder Qualität vor, werden folgende Studien herangezogen, die eine qualitative Abschätzung des unter Ansatz (1) beschriebenen Vergleichs ermöglichen:

- 2a) Studien zum Vergleich eines PSA-basierten Screenings gegenüber keiner spezifischen Früherkennung (ggf. ergänzt durch DRU) sowie
- 2b) Studien zum Vergleich eines Screeningalgorithmus aus PSA-Test und MRT-Diagnostik (ggf. unter Einbeziehung eines Risikoscores) gegenüber einem alleinigen PSA-Screening.

Das Ziel des Vorgehens gemäß Ansatz (2) ist, die Teilergebnisse der einzelnen Screeningvergleiche qualitativ zu verknüpfen, um den Effekt des zu bewertenden Screeningalgorithmus auf patientenrelevante Endpunkte im Vergleich zu keinem Screening abschätzen zu können.

Konkret liefern die Studien unter (2a) Schätzungen des Effekts eines PSA-basierten Screenings gegenüber keiner spezifischen Früherkennung (ggf. ergänzt durch DRU), während Studien unter (2b) Schätzungen des Effekts einer Hinzunahme einer MRT-Untersuchung zu einem PSA-basierten Screening (ggf. einschließlich Risikoscore) gegenüber einem reinen PSA-Screening ermöglichen. Das PSA-basierte Screening fungiert damit als gemeinsamer Brückenkompator, um Erkenntnisse zum Vergleich des kombinierten Screeningalgorithmus aus PSA + MRT (und ggf. Risikoscore) gegenüber keiner spezifischen Früherkennung (ggf. ergänzt durch DRU) zu gewinnen.

Für eine solche qualitative Abschätzung ist es erforderlich, dass die Studien in zentralen Aspekten – etwa hinsichtlich Population, diagnostischer Verfahren oder Studienzeitraum – ausreichend vergleichbar sind. Weichen diese Merkmale zu stark voneinander ab, steigt das

Risiko systematischer Verzerrungen. In solchen Fällen können die betreffenden Studien nicht für die qualitative Abschätzung herangezogen werden.

Ansatz (3)

Wenn auch diese qualitative Abschätzung nicht möglich ist z. B. weil keine ausreichenden Langzeitdaten zu patientenrelevanten Endpunkten – insbesondere zur Mortalität – vorliegen, werden anhand von Studien die zentralen Annahmen überprüft, die erfüllt sein müssen, damit ein risikoadaptiertes PSA- / MRT-Screening einen Nutzen bieten kann (siehe Abschnitt 4.2). Diese Annahmen lauten,

- dass es unter Personen mit erhöhtem PSA-Wert solche gibt, die von einer nachfolgenden Biopsie und ggf. Tumordiagnose und -therapie nicht profitieren (Annahme A1), und
- dass der ergänzende Einsatz der MRT zusätzlich zur PSA-Bestimmung eine ausreichend verlässliche Differenzierung ermöglicht zwischen Personen, die von einer nachfolgenden Biopsie und ggf. Tumordiagnose und -therapie profitieren und solchen, die nicht profitieren (Annahme A2; siehe auch Kapitel 1).

4.1 Kriterien für den Einschluss von vergleichenden Interventionsstudien zur Untersuchung von Screeningalgorithmen

4.1.1 Population

In die Bewertung gemäß Ansatz (1) und (2) des o. g. Vorgehens werden Studien mit Personen ohne Verdacht auf ein Prostatakarzinom aufgenommen.

4.1.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention gemäß Ansatz (1) ist eine sequenziell geordnete, risikoadaptierte Früherkennung des Prostatakarzinoms, die auf einem PSA-Test und einer sich anschließenden MRT-Diagnostik (ggf. unter Einbeziehung von Risikoscores) basiert. Entsprechend werden solche Studien berücksichtigt, welche den gesamten Screeningalgorithmus aus PSA-Test und MRT-Diagnostik (ggf. ergänzt durch Risikoscores) gegenüber keiner spezifischen Prostatakarzinom-Früherkennung (ggf. ergänzt durch DRU) untersuchen. Für die Bewertung des Nutzens ist sicherzustellen, dass die Behandlung des Prostatakarzinoms in den Studienarmen vergleichbar erfolgt. Das bedeutet, dass Patienten mit einem vergleichbaren Tumorstadium oder -grad (z. B. nach Gleason-Score oder ISUP-Grad) eine vergleichbare Therapie erhalten, unabhängig davon, wie das Karzinom diagnostiziert wurde.

Liegen derartige Studien zu einem direkten Vergleich nicht in ausreichender Zahl oder Qualität vor, erfolgt gemäß Ansatz (2) eine qualitative Abschätzung des beschriebenen Vergleichs. Hierfür werden Studien berücksichtigt, welche die folgenden Screeningalgorithmen untersuchen:

- 2a) Studien zum Vergleich eines PSA-basierten Screenings gegenüber keiner spezifischen Früherkennung (ggf. ergänzt durch DRU) sowie
- 2b) Studien zum Vergleich eines Screeningalgorithmus aus PSA-Test und MRT-Diagnostik (ggf. unter Einbeziehung eines Risikoscores) gegenüber einem alleinigen PSA-basierten Screening. Dies umfasst auch Studien, die Personen mit positiven PSA-Test auf eine MRT-Diagnostik randomisieren (Anreicherungsdesign).

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden patientenrelevante Zielgrößen der folgenden Endpunktkategorien betrachtet:

- Mortalität (z. B. Gesamtmortalität, prostatakarzinomspezifische Mortalität)
- Morbidität (z. B. Auftreten eines metastasierten Prostatakarzinoms, Schmerzen aufgrund von Knochenmetastasen)
- gesundheitsbezogene Lebensqualität
- unerwünschte Ereignisse (u. a. 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 in Abschnitt 4.1.2 genannten Interventionen und alle in Abschnitt 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 Setting

Es werden nur solche Studien eingeschlossen, die in Staaten durchgeführt wurden, bei denen von einer Übertragbarkeit auf das deutsche Gesundheitssystem ausgegangen werden kann (insbesondere westliche Industrieländer).

4.1.8 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle 1 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	Population: Personen ohne Verdacht auf ein Prostatakarzinom
E2-1	Prüfintervention (siehe auch Abschnitt 4.1.2): sequenziell geordnete, risikoadaptierte Früherkennung des Prostatakarzinoms mittels PSA-Test und MRT-Diagnostik (ggf. unter Einbeziehung von Risikoscores)
E2-2a	Früherkennung des Prostatakarzinoms mittels PSA-Test
E2-2b	sequenziell geordnete, risikoadaptierte Früherkennung des Prostatakarzinoms mittels PSA-Test und MRT-Diagnostik (ggf. unter Einbeziehung von Risikoscores)
E3-1	Vergleichsintervention (siehe auch Abschnitt 4.1.2): keine spezifische Prostatakarzinom-Früherkennung (ggf. ergänzt durch DRU)
E3-2a	keine spezifische Prostatakarzinom-Früherkennung (ggf. ergänzt durch DRU)
E3-2b	Früherkennung des Prostatakarzinoms mittels PSA-Test
E4	Zielgrößen der Endpunktkategorien wie in Abschnitt 4.1.3 formuliert
E5	Studientyp: RCTs
E6	Publikationssprache: Deutsch oder Englisch
E7	Setting: Staaten, bei denen von einer Übertragbarkeit auf das deutsche Gesundheitssystem ausgegangen werden kann (siehe auch Abschnitt 4.1.7)
E8	Vollpublikation verfügbar ^a
a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [11] oder ein Bericht über die Studie, der den Kriterien des CONSORT-Statements [12] 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; DRU: digital-rektale Untersuchung; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; MRT: Magnetresonanztomografie; PSA: prostataspezifisches Antigen; RCT: randomisierte kontrollierte Studie	

4.2 Empirische Prüfung der Annahmen des PSA- / MRT-Screenings bei fehlender Evidenz aus Interventionsstudien zu Screeningalgorithmen

Sollten die identifizierten Studien weder eine direkte Bewertung des Effekts des PSA- / MRT-Screenings (siehe Ansatz (1) in Kapitel 4) noch eine qualitative Abschätzung des Effekts (siehe Ansatz (2) in Kapitel 4) ermöglichen, werden die Annahmen überprüft, die mindestens erfüllt

sein müssen, damit ein solches Screeningverfahren einen patientenrelevanten Nutzen bieten kann. Die beiden zentralen Annahmen lauten:

- A1) Unter Personen mit erhöhtem PSA-Wert gibt es solche, die von einer nachfolgenden Biopsie sowie einer möglichen Tumordiagnose und -therapie nicht profitieren.
- A2) Die zusätzliche MRT-Untersuchung ermöglicht eine ausreichend verlässliche Differenzierung zwischen diesen Personen und solchen, die von Tumordiagnose und -therapie profitieren.

Um diese Annahmen empirisch zu prüfen, wird der in Screening- und Diagnosestudien regelhaft erhobene Tumorgrad bei Erstdiagnose (z. B. nach Gleason / ISUP) herangezogen. Der Tumorgrad beschreibt die mikroskopische Morphologie des Drüsengewebes und gilt als ein klinisch relevanter Parameter für die Prognose des Prostatakarzinoms [4]. Er bildet zusammen mit PSA-Wert und DRU die Grundlage etablierter Risikostratifizierungen und leitet damit maßgeblich therapeutische Entscheidungen ab.

Die Verwendung des Tumorgrads zur Überprüfung der Annahmen beruht auf folgender Logik: Wenn (A1) der Tumorgrad valide prognostiziert, welche Personen einer sofortigen kurativen Therapie bedürfen und bei welchen dies nicht unmittelbar notwendig ist, und (A2) ein PSA- / MRT-basiertes Screening Personen mit solchermaßen hoch- bzw. niedriggradigen Progressionsrisiko zuverlässig erkennt, dann kann ein solcher Screeningalgorithmus im Vergleich zu einer Früherkennung mittels alleinigem PSA-Test unnötige Diagnostik und Übertherapien reduzieren, ohne das Mortalitätsrisiko zu erhöhen.

Hieraus ergeben sich die folgenden überprüfbaren Hypothesen:

H1) Hypothese zu Annahme A1:

Personen mit niedrigem Progressionsrisiko weisen unter Abwarten oder aktiver Überwachung kein höheres Mortalitätsrisiko auf als Personen, die unmittelbar kurativ behandelt werden. Ein niedriges Progressionsrisiko ergibt sich insbesondere aus einem niedrigen Tumorgrad (d. h. ISUP-Grad 1 bzw. Gleason Score ≤ 6) und einem PSA-Wert < 10 (siehe hierzu die Risikoklassifikation der EAU [13]).

Ein solches Ergebnis würde darauf hindeuten, dass das Weglassen von weiterer Diagnostik und Therapie bei diesen Personen angemessen ist.

H2) Hypothese zu Annahme A2:

Ein PSA- / MRT-basiertes Screening führt im Vergleich zu einem allein PSA-basierten Screening zu weniger Diagnosen niedriggradiger Tumoren bei mindestens gleicher Anzahl höhergradiger Tumoren, sodass beim PSA- / MRT-basierten Screening der Anteil höhergradiger an allen diagnostizierten Tumoren größer ist.

Dies würde darauf hinweisen, dass ein PSA- / MRT-basiertes Screening mindestens

ebenso viele Personen mit behandlungsbedürftigem Prostatakarzinom einer Biopsie zuführt wie ein alleiniges PSA-basiertes Screening, während gleichzeitig bei Personen mit niedrigem Progressionsrisiko eine unnötige weiterführende Diagnostik und Therapie vermieden wird.

Zeigt die empirische Prüfung der Annahme A1, dass diese nicht haltbar ist, wird die Annahme A2 nicht geprüft.

Im Folgenden werden die Einschlusskriterien der Studien zur Prüfung dieser beiden Hypothesen dargestellt.

4.2.1 Kriterien für den Einschluss von Studien zu Hypothese 1

4.2.1.1 Population

Eingeschlossen werden Studien mit Personen im Alter von 50 bis 70 Jahren mit lokalisiertem Prostatakarzinom und niedrigem Tumorgrad bzw. niedrigem Progressionsrisiko.

4.2.1.2 Prüf- und Vergleichsintervention

Als Prüfinerventionen gelten kurative Behandlungsstrategien für lokalisierendes Prostatakarzinom mit niedrigem Tumorgrad, insbesondere

- radikale Prostatektomie
- Strahlentherapie (Brachytherapie oder externe Strahlentherapie [EBRT])

Die Behandlungsformen und -zeitpunkte müssen auf in Deutschland etablierte therapeutische Strategien übertragbar sein. Eingesetzte Arzneimittel müssen im Rahmen ihres in Deutschland gültigen Zulassungsstatus verwendet werden.

Als Vergleichsintervention gilt aktive Überwachung / active surveillance oder Abwarten.

4.2.1.3 Patientenrelevante Endpunkte

Für die Überprüfung werden die folgenden patientenrelevanten Endpunkte betrachtet:

- Gesamtmortalität
- Prostatakarzinomspezifische Mortalität

4.2.1.4 Studientypen

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 hier zu untersuchende Hypothese 1, dass Personen mit

niedrigem Tumorgrad unter Abwarten oder aktiver Überwachung kein höheres Mortalitätsrisiko aufweisen als Personen, die unmittelbar kurativ behandelt werden.

Falls keine ausreichend belastbaren RCTs vorliegen, werden zusätzlich prospektiv geplante, vergleichende Kohortenstudien einbezogen. Falls auch keine ausreichend belastbaren prospektiv geplanten, vergleichenden Kohortenstudien vorliegen, werden zusätzlich prospektiv geplante, 1-armige Studien mit unbehandelten Personen mit nachvollziehbarem Patientenfluss einbezogen, sofern sie Rückschlüsse auf die genannten Endpunkte zulassen (Prognosestudien).

4.2.1.5 Studiendauer

Eingeschlossen werden Studien mit einer Mindestdauer der Nachbeobachtung von 10 Jahren.

4.2.1.6 Studiengröße

Für prospektiv geplante vergleichende Kohortenstudien bzw. 1-armige Studien gilt eine Mindeststudiengröße von 1000 eingeschlossenen Personen.

4.2.1.7 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfasst sein.

4.2.1.8 Setting

Es werden nur solche Studien eingeschlossen, die in Staaten durchgeführt wurden, bei denen von einer Übertragbarkeit auf das deutsche Gesundheitssystem ausgegangen werden kann (insbesondere westliche Industrieländer).

4.2.1.9 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle 2 sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss zur Überprüfung der Hypothese 1

Einschlusskriterien	
H1-1	Population: Personen im Alter von 50 bis 70 Jahren mit lokalisiertem Prostatakarzinom und niedrigem Tumorgrad
H1-2	Prüfintervention: kurative Behandlungsoptionen bei lokalisiertem Prostatakarzinom mit niedrigem Tumorgrad, insbesondere: ▪ radikale Prostatektomie ▪ Strahlentherapie (Brachytherapie oder EBRT) (siehe auch Abschnitt 4.2.1.2)
H1-3	Vergleichsintervention: aktive Überwachung / active surveillance oder Abwarten (siehe auch Abschnitt 4.2.1.2)
H1-4	Endpunkte: Gesamtmortalität, prostatakarzinomspezifische Mortalität
H1-5	Studientyp: RCTs, sofern vorhanden; andernfalls prospektive vergleichende Kohortenstudien, sofern vorhanden; andernfalls prospektiv geplante, 1-armige Studien mit unbehandelten Personen (siehe Abschnitt 4.2.1.4)
H1-6	Studiendauer: mindestens 10 Jahre (siehe Abschnitt 4.2.1.5)
H1-7	Studiengröße: min. 1000 eingeschlossene Personen für prospektiv geplante vergleichende Kohortenstudien bzw. 1-armige Studien (siehe auch Abschnitt 4.2.1.6)
H1-8	Publikationssprache: Deutsch oder Englisch
H1-9	Setting: Staaten, bei denen von einer Übertragbarkeit auf das deutsche Gesundheitssystem ausgegangen werden kann (siehe auch Abschnitt 4.2.1.7)
H1-10	Vollpublikation verfügbar ^a
a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [11] oder ein Bericht über die Studie, der den Kriterien des CONSORT- [12] oder STROBE-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; EBRT: External Beam Radiotherapy; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; RCT: randomisierte kontrollierte Studie	

4.2.2 Kriterien für den Einschluss von Studien zu Hypothese 2

4.2.2.1 Population

Eingeschlossen werden Studien mit Personen ohne Verdacht auf ein Prostatakarzinom.

4.2.2.2 Prüf- und Vergleichsintervention

Die Prüfintervention ist ein Screeningalgorithmus aus PSA-Test und MRT-Diagnostik (ggf. unter Einbeziehung von Risikoscores). Ebenfalls eingeschlossen werden Studien mit Anreicherungsdesign, in denen Personen nach positivem PSA-Test randomisiert entweder einer MRT-Diagnostik oder keiner weiteren Diagnostik vor Biopsie zugeteilt werden.

Die Vergleichsintervention besteht aus einem alleinigen PSA-basiertem Screening.

4.2.2.3 Zielgrößen

Als relevante Zielgrößen werden die Biopsie-Rate und die hierbei festgestellten Tumorgrade betrachtet.

4.2.2.4 Studientypen

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 Überprüfung der Hypothese 2.

Für alle in Abschnitt 4.2.2.2 genannten Interventionen ist eine Evaluation im Rahmen von RCTs möglich und praktisch durchführbar. Für die Überprüfung der Hypothese 2 werden daher RCTs als relevante wissenschaftliche Literatur in die Bewertung einfließen.

4.2.2.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

4.2.2.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfasst sein.

4.2.2.7 Setting

Es werden nur solche Studien eingeschlossen, die in Staaten durchgeführt wurden, bei denen von einer Übertragbarkeit auf das deutsche Gesundheitssystem ausgegangen werden kann (insbesondere westliche Industrieländer).

4.2.2.8 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle 3 sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss zur Überprüfung der Hypothese 2

Einschlusskriterien	
H2-1	Population: Personen ohne Verdacht auf ein Prostatakarzinom
H2-2	Prüfintervention: sequenziell geordnete Früherkennung des Prostatakarzinoms mittels PSA-Test und MRT-Diagnostik (ggf. unter Einbeziehung von Risikoscores)
H2-3	Vergleichsintervention: Früherkennung des Prostatakarzinoms mittels PSA-Test
H2-4	Zielgröße: Biopsierate und Tumorgrad
H2-5	Studiendesign: RCTs
H2-6	Publikationssprache: Deutsch oder Englisch
H2-7	Setting: Staaten, bei denen von einer Übertragbarkeit auf das deutsche Gesundheitssystem ausgegangen werden kann (siehe auch Abschnitt 4.1.7)
H2-8	Vollpublikation verfügbar ^a
a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [11] oder ein Bericht über die Studie, der den Kriterien des CONSORT-Statements [12] 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; MRT: Magnetresonanztomografie; PSA: prostataspezifisches Antigen; RCT: randomisierte kontrollierte Studie	

4.3 Informationsbeschaffung

4.3.1 Fokussierte Informationsbeschaffung von systematischen Übersichten zu vergleichenden Interventionsstudien zur Untersuchung von Screeningalgorithmen

Parallel zur Erstellung des Berichtsplans erfolgte eine Recherche nach systematischen Übersichten in MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews), der International Health Technology Assessment (HTA) Database, sowie auf den Websites des National Institute for Health and Care Excellence (NICE) und der Agency for Healthcare Research and Quality (AHRQ).

Die Suche fand in Erwartung des Auftrags am 29.09.2025 statt. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Anhang A. Die Selektion erfolgte durch 1 Person und wurde anschließend von einer 2. Person überprüft. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst.

Es wird geprüft, ob mindestens 1 hochwertige und aktuelle systematische Übersicht infrage kommt, deren Informationsbeschaffung als Grundlage verwendet werden kann (im Folgenden: Basis-SÜ). Dafür erfolgt eine Bewertung der Qualität der Informationsbeschaffung dieser systematischen Übersicht(en) mit den entsprechenden Items aus A Measurement Tool to Assess Systematic Reviews 2 (AMSTAR 2) [15]. Kann mindestens eine diesbezüglich hochwertige und aktuelle Basis-SÜ identifiziert werden, werden die zugrunde liegenden

Studien beziehungsweise Dokumente von 1 Person auf ihre Relevanz für die vorliegende Bewertung geprüft und das Ergebnis von einer 2. Person überprüft. Bewertungen der eingeschlossenen Studien oder die Datenextraktion werden nicht übernommen.

Die finale Entscheidung, ob und wenn ja welche systematische(n) Übersicht(en) als Basis-SÜ herangezogen werden, erfolgt nach Fertigstellung des Berichtsplans anhand der darin festgelegten Kriterien entsprechend Tabelle 1. In jedem Fall werden die Referenzlisten der identifizierten systematischen Übersichten hinsichtlich relevanter Primärstudien gesichtet (siehe Abschnitt 4.3.2).

4.3.2 Umfassende Informationsbeschaffung von vergleichenden Interventionsstudien zur Untersuchung von Screeningalgorithmen

Für die umfassende Informationsbeschaffung wird eine systematische Recherche nach relevanten Studien beziehungsweise Dokumenten durchgeführt.

Für den Fall, dass mindestens eine systematische Übersicht als Basis-SÜ für die Informationsbeschaffung verwendet werden kann (siehe Abschnitt 4.3.1), wird diese für die Informationsbeschaffung von Studien für den von der Übersicht abgedeckten Zeitraum herangezogen. Dieser Teil der Informationsbeschaffung wird ergänzt um eine systematische Recherche nach relevanten Studien beziehungsweise Dokumenten für den nicht von der Übersicht abgedeckten Zeitraum.

Für den Fall, dass keine Basis-SÜ identifiziert werden kann, findet eine systematische Recherche für den gesamten relevanten Zeitraum statt.

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

Die Einträge von ClinicalTrials.gov werden aus dem Suchergebnis des International Clinical Trials Registry Platform Search Portal entfernt.

Weitere Informationsquellen und Suchtechniken

- durch den G-BA übermittelte Dokumente
- G-BA-Website und IQWiG-Website
- Anwendung weiterer Suchtechniken
 - Sichten von Referenzlisten identifizierter systematischer Übersichten
- Anhörung zum Vorbericht
- Autorenanfragen

4.3.3 Informationsbeschaffung zur empirischen Prüfung der Annahmen des PSA- / MRT-Screenings bei fehlender Evidenz aus Interventionsstudien zu Screeningalgorithmen

Die Überprüfung der Hypothese 1 erfolgt über eine fokussierte Informationsbeschaffung nach vergleichenden Interventionsstudien zur Behandlung eines Prostatakarzinoms. Für diese Rechenschritte werden folgende primäre und weitere Informationsquellen sowie Suchtechniken berücksichtigt:

- bibliografische Datenbanken
 - MEDLINE
 - Cochrane Central Register of Controlled Trials
- Weitere Informationsquellen und Suchtechniken
 - durch den G-BA übermittelte Dokumente
 - Websites von G-BA und IQWiG
 - Anhörung zum Vorbericht

Sollten diese beiden Schritte keine ausreichend belastbaren Studien identifizieren, erfolgen weitergehende Recherchen nach prospektiv geplanten vergleichenden Kohortenstudien bzw. ggf. nach prospektiv geplanten 1-armigen Studien mit unbehandelten Personen (Prognosestudien) in den o. g. Quellen.

Zur Überprüfung der Hypothese 2 wird auf die Informationsbeschaffung der vergleichenden Interventionsstudien zur Untersuchung von Screeningalgorithmen zurückgegriffen (zur Informationsbeschaffung dieser Studien siehe Abschnitt 4.3.1 und 4.3.2).

4.3.4 Anwendung von Limitierungen auf Datenbankebene

Fokussierte Informationsbeschaffung von systematischen Übersichten

Die Suchen wurden auf das Publikationsdatum ab 2022 eingeschränkt. Die MEDLINE-Suchstrategien enthalten Limitierungen auf deutsch- und englischsprachige Publikationen sowie auf Humanstudien.

Informationsbeschaffung von vergleichenden Interventionsstudien

Es ist keine zeitliche Einschränkung vorgesehen für die umfassende Informationsbeschaffung zur Untersuchung von Screeningalgorithmen. Sollte die Informationsbeschaffung auf Grundlage einer Basis-SÜ erfolgen, wird eine entsprechende zeitliche Einschränkung in Betracht gezogen (siehe Abschnitt 4.3.2).

Mit den Suchstrategien werden folgende Publikationstypen ausgeschlossen: Kommentare (MEDLINE) und Editorials (MEDLINE, Embase), da diese i. d. R. keine Studien enthalten [16], sowie Conference Abstract und Conference Review (Embase). Außerdem enthalten die Suchstrategien Limitierungen auf deutsch- und englischsprachige Publikationen sowie auf Humanstudien (MEDLINE, Embase). In der Embase-Suche werden MEDLINE-Datensätze und in der Suche im Cochrane Central Register of Controlled Trials Einträge aus Studienregistern ausgeschlossen.

4.3.5 Selektion relevanter Studien aus der umfassenden Informationsbeschaffung

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

Duplikate werden mit Hilfe des Literaturverwaltungsprogrammes EndNote entfernt. 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, Tabelle 2 und Tabelle 3) 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 den beiden aufgelöst.

Selektion relevanter Studien bzw. Dokumente aus weiteren Informationsquellen

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

- Studienregister.

Die 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. Diskrepanzen werden durch Diskussion zwischen den beiden aufgelöst.

4.3.6 Selektion relevanter Studien aus der fokussierten Informationsbeschaffung

Die Ergebnisse der fokussierten Informationsbeschaffung werden von einer Person auf relevante Studien gesichtet. Anschließend prüft eine 2. Person den gesamten Prozess. Etwaige Diskrepanzen werden zwischen beiden diskutiert und geklärt.

4.4 Informationsbewertung und -synthese

4.4.1 Darstellung der Einzelstudien

4.4.1.1 Vergleichende Interventionsstudien zur Untersuchung von Screeningalgorithmen bzw. zur Therapie eines Prostatakarzinoms

Alle für die Nutzenbewertung notwendigen Informationen werden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Die Ergebnisse zu den in den Studien berichteten (patienten-)relevanten Endpunkten werden im Bericht beschrieben.

Der Anteil der Überdiagnosen wird in 2 unterschiedlichen Bezugsgrößen berechnet: (1) bezogen auf die im Screeningzeitraum diagnostizierten Personen und (2) bezogen auf alle zum Screening eingeladenen Personen.

- 1) Der Anteil der Überdiagnosen unter den diagnostizierten Personen wird folgendermaßen berechnet: Von der Anzahl der Diagnosen der Interventionsgruppe (aus dem gesamten Zeitraum von Screening- und Nachbeobachtungsphase) wird die Anzahl der Diagnosen der Kontrollgruppe (aus demselben genannten Zeitraum) subtrahiert. Für die Nachbeobachtungsphase wird der längste Beobachtungszeitraum zugrunde gelegt. Die Differenz wird durch die Anzahl der Diagnosen in der Interventionsgruppe während des Screeningzeitraums dividiert. Das Ergebnis ist der Anteil der Überdiagnosen (Überdiagnoserisiko) unter den Diagnosen des Screenings. In Fällen, in denen die Behandlungsgruppen nicht gleich groß sind, wird bei der Berechnung die Zahl der Diagnosen in der Kontrollgruppe entsprechend gewichtet.
- 2) Der Anteil der Überdiagnosen unter den zum Screening eingeladenen Personen wird folgendermaßen berechnet: Von der Anzahl der Diagnosen der Interventionsgruppe (aus dem gesamten Zeitraum von Screening- und Nachbeobachtungsphase) wird die Anzahl der Diagnosen der Kontrollgruppe (aus demselben genannten Zeitraum) subtrahiert. Für die Nachbeobachtungsphase wird der längste Beobachtungszeitraum zugrunde gelegt. Die Differenz wird durch die Anzahl der Personen der Interventionsgruppe dividiert. Das Ergebnis ist der Anteil der Überdiagnosen (Überdiagnoserisiko) unter den zum Screening eingeladenen Teilnehmerinnen und Teilnehmern. In Fällen, in denen die

Behandlungsgruppen nicht gleich groß sind, wird bei der Berechnung die Zahl der Diagnosen in der Kontrollgruppe entsprechend gewichtet.

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 in den Abschnitten 4.4.3 bis 4.4.5 beschriebenen Verfahren eingesetzt. Für Ergebnisse, die nicht in Metaanalysen einfließen, werden nur dann eigene Berechnungen durchgeführt, wenn dies zur Einschätzung der statistischen Signifikanz notwendig ist.

Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Personen basieren, das heißt, wenn der Anteil der Personen, 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 Personen zwischen den Gruppen größer als 15 Prozentpunkte ist.

4.4.1.2 Prospektiv geplante Kohortenstudien (Prognosestudien)

Alle Ergebnisse aus der fokussierten Informationsbeschaffung zu Prognosestudien werden in verkürzter Form dargestellt. Die Darstellung umfasst die Charakteristika des Studiendesigns, der Studienpopulation sowie die Ergebnisse zu den Endpunkten. Wenn in Studien Auswertungen zu relevanten Endpunkten für mehrere Zeitpunkte vorliegen, erfolgt eine Darstellung der Ergebnisse aus der längsten Nachbeobachtungsdauer.

4.4.2 Bewertung des Verzerrungspotenzials der Ergebnisse

4.4.2.1 Vergleichende Interventionsstudien zur Untersuchung von Screeningalgorithmen bzw. zur Therapie eines Prostatakarzinoms

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 für die endpunktübergreifende Bewertung des Verzerrungspotenzials der Ergebnisse

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung der an der Studie teilnehmenden Person sowie der behandelnden Personen
- ergebnisunabhängige Berichterstattung

B: Kriterien für die endpunktspezifische 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.

Im Falle subjektiv erhobener Endpunkte ergibt sich aus einer fehlenden Verblindung der an der Studie teilnehmenden Person bereits ein endpunktspezifisch hohes Verzerrungspotenzial.

4.4.2.2 Prospektiv geplante Kohortenstudien (Prognosestudien)

Die methodische Bewertung der Qualität der Prognosestudien erfolgt mittels des PROBAST-Instruments (Prediction model Risk Of Bias Assessment Tool; [17]).

4.4.3 Metaanalysen

Die geschätzten Effekte und Konfidenzintervalle aus den Studien werden mittels Forest Plots dargestellt. Die Heterogenität zwischen den Studien wird mithilfe des statistischen Tests auf Vorliegen von Heterogenität [18] untersucht. Es wird außerdem untersucht, welche Faktoren eine vorhandene Heterogenität möglicherweise verursachen. Dazu zählen methodische Faktoren (siehe Abschnitt 4.4.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.4.5). Falls vorhandene Heterogenität durch solche Faktoren zumindest zum Teil erklärt werden kann, so wird der Studienpool nach diesen Faktoren aufgespaltet und die weiteren Berechnungen erfolgen in den getrennten Studienpools. Bei statistisch nachgewiesener (unerklärter) Heterogenität ist eine gemeinsame Effektschätzung nicht sinnvoll und es erfolgt eine qualitative Zusammenfassung der Studienergebnisse.

In Abhängigkeit von der Anzahl der Studien wird zur Durchführung von Metaanalysen regelhaft folgendes Standardvorgehen gewählt:

- 2 Studien: Anwendung des Modells mit festem Effekt, und zwar mithilfe der inversen Varianzmethode bei stetigen Daten bzw. der Mantel-Haenszel-Methode bei binären Daten [19].

- 3 bis 4 Studien: Anwendung des Modells mit zufälligen Effekten, und zwar – für die Effektmaße SMD, Odds Ratio, relatives Risiko und Hazard Ratio – mithilfe einer bayesschen Metaanalyse mit nicht informativen A-priori-Verteilungen für den Behandlungseffekt und informativen A-priori-Verteilungen für den Heterogenitätsparameter τ gemäß Lilienthal et al. [20]. Zudem sollte ein Abgleich mit einer qualitativen Zusammenfassung der Studienergebnisse mithilfe des Konzepts der konkludenten Effekte erfolgen. Das bedeutet, dass in den Fällen, in denen aus der bayesschen Metaanalyse kein Effekt abgeleitet wird, zur Kontrolle eine qualitative Zusammenfassung der Studienergebnisse erfolgt. Sollten hierbei konkludente Effekte gefunden werden, so wird insgesamt ein Effekt abgeleitet, der jedoch nicht quantifizierbar ist. Für sonstige Effektmaße ist projektspezifisch zu entscheiden, ob die Methode nach Knapp-Hartung, eine qualitative Zusammenfassung der Studienergebnisse oder ein anderes metaanalytisches Verfahren anzuwenden ist.
- 5 Studien und mehr: Anwendung des Modells mit zufälligen Effekten, und zwar mithilfe der Knapp-Hartung-Methode. Zunächst werden gepoolte Effekte nach der Methode von Knapp-Hartung – mit und ohne Ad-hoc-Varianzkorrektur – sowie der Paule-Mandel-Methode zur Schätzung des Heterogenitätsparameters τ [21] und gepoolte Effekte nach der Methode von DerSimonian-Laird berechnet. Es wird geprüft, ob das Konfidenzintervall nach Knapp-Hartung (ohne Ad-hoc-Varianzkorrektur) schmäler ist als das nach DerSimonian-Laird. In diesem Fall wird die Effektschätzung nach Knapp-Hartung mit Ad-hoc-Varianzkorrektur, ansonsten ohne Ad-hoc-Varianzkorrektur weiterverwendet. Im Anschluss wird geprüft, ob diese Effektschätzung informativ ist. Als informativ wird die Schätzung dann bezeichnet, falls das Konfidenzintervall (des gemeinsamen Effekts) in der Vereinigung der Konfidenzintervalle der Einzelstudien enthalten ist. In diesem Fall wird diese Effektschätzung (nach Knapp-Hartung) zur finalen Bewertung herangezogen. Ansonsten wird eine gemeinsame Effektschätzung als nicht sinnvoll erachtet und es erfolgt eine qualitative Zusammenfassung der Studienergebnisse.

Bei 5 oder mehr Studien wird das Prädiktionsintervall im Forest Plot mit dargestellt.

4.4.4 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 eingestuftter 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 4.4.6).

4.4.5 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 Personengruppen 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 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 Werte, Sequenztechnik des MRT (bi- vs. multiparametrisch), Risikoscore.

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 eingeschränkt werden (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt 4.4.6).

4.4.6 Aussagen zur Beleglage

Für jeden Endpunkt der vergleichenden Interventionsstudien zur Untersuchung von Screeningalgorithmen (Ansatz (1) und (2); siehe Kapitel 4) wird eine Aussage zur Beleglage des (höheren) Nutzens oder (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 4 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 4: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit

		Anzahl Studien				
		1 (mit statistisch signifikantem Effekt)	≥ 2			
			gemeinsame Effektschätzung sinnvoll	gemeinsame Effektschätzung nicht sinnvoll		
				Metaanalyse statistisch signifikant	konkludente Effekte ^a	
			deutlich		mäßig	nein
Qualitative Ergebnis- sicherheit	hoch	Hinweis	Beleg	Beleg	Hinweis	–
	mäßig	Anhaltspunkt	Hinweis	Hinweis	Anhaltspunkt	–
	gering	–	Anhaltspunkt	Anhaltspunkt	–	–
a. Unter konkludenten Effekten wird eine Datensituation verstanden, in der es möglich ist, einen Effekt im Sinne der Fragestellung abzuleiten, obwohl eine gemeinsame Effektschätzung nicht sinnvoll möglich ist (siehe Abschnitt 3.1.4 der Allgemeinen Methoden 8.0 [10]).						

Abschließend erfolgt eine endpunktübergreifende Bewertung des Nutzens. In diese übergreifende Bewertung wird auch die Datenvollständigkeit und die sich daraus möglicherweise ergebende Verzerrung aufgrund von Publication Bias oder Outcome Reporting Bias einbezogen.

Eine qualitative Abschätzung (siehe Abschnitt 4.1.2) bzw. eine empirische Prüfung der Annahmen des PSA- / MRT-Screenings (siehe Abschnitt 4.2) dient als mögliche zusätzliche Informationsquelle für das Fazit. Ein Nutzen kann sich allein auf Basis einer solchen Abschätzung aufgrund einer zu geringen qualitativen Ergebnissicherheit jedoch regelhaft nicht ergeben.

Falls kein Anhaltspunkt für einen (höheren) Nutzen oder (höheren) Schaden ableitbar ist, wird eine Aussage zum Potenzial der Methode getroffen und es werden ggf. Eckpunkte einer Erprobungsstudie formuliert.

5 Literatur

1. Robert Koch-Institut. Prostatakrebs (Prostatakarzinom) [online]. 2024 [Zugriff: 10.10.2025]. URL: https://www.krebsdaten.de/Krebs/DE/Content/Krebsarten/Prostatakrebs/prostatakrebs_no_de.html.
2. Castro E, Goh C, Olmos D et al. Germline BRCA Mutations Are Associated With Higher Risk of Nodal Involvement, Distant Metastasis, and Poor Survival Outcomes in Prostate Cancer. J Clin Oncol 2013; 31(14): 1748-1757. <https://doi.org/10.1200/JCO.2012.43.1882>.
3. Robert Koch-Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland. Krebs in Deutschland für 2019/2020 [online]. 2023 [Zugriff: 06.11.2025]. URL: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/krebs_in_deutschland_2023.pdf?blob=publicationFile.
4. Leitlinienprogramm Onkologie. S3-Leitlinie Prostatakarzinom Version 8.1 [online]. 2025 [Zugriff: 10.10.2025]. URL: <https://register.awmf.org/de/leitlinien/detail/043-022OL>.
5. Raychaudhuri R, Lin DW, Montgomery RB. Prostate Cancer: A Review. JAMA 2025; 333(16): 1433-1446. <https://doi.org/10.1001/jama.2025.0228>.
6. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Prostatakarzinom-Screening mittels PSA-Test; Abschlussbericht [online]. 2020 [Zugriff: 10.10.2025]. URL: https://www.iqwig.de/download/s19-01_psa-screening_abschlussbericht_v1-1.pdf.
7. Cornford P, van den Bergh RCN, Briers E et al. EAU-EANM-ESTRO-ESUR-ISUP-SIOG Guidelines on Prostate Cancer - 2024 Update. Part I: Screening, Diagnosis, and Local Treatment with Curative Intent. Eur Urol 2024; 86(2): 148-163. <https://doi.org/10.1016/j.eururo.2024.03.027>.
8. Auvinen A, Tammela TLJ, Mirtti T et al. Prostate Cancer Screening With PSA, Kallikrein Panel, and MRI: The ProScreen Randomized Trial. JAMA 2024; 331(17): 1452-1459. <https://doi.org/10.1001/jama.2024.3841>.
9. Bratt O, Godtman RA, Jiborn T et al. Population-based Organised Prostate Cancer Testing: Results from the First Invitation of 50-year-old Men. Eur Urol 2024; 85(3): 207-214. <https://doi.org/10.1016/j.eururo.2023.11.013>.
10. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Version 8.0 [online]. 2025. URL: https://doi.org/10.60584/Allgemeine-Methoden_V8.0.
11. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3 [online]. 1995 [Zugriff: 02.09.2024]. URL: https://database.ich.org/sites/default/files/E3_Guideline.pdf.

12. Moher D, Hopewell S, Schulz KF et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869. <https://doi.org/10.1136/bmj.c869>.
13. European Association of Urology. Prostate Cancer [online]. 2025 [Zugriff: 17.12.2025]. URL: <https://uroweb.org/guidelines/prostate-cancer>.
14. Von Elm E, Altman DG, Egger M et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577. <https://doi.org/10.7326/0003-4819-147-8-200710160-00010>.
15. Shea BJ, Reeves BC, Wells G et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017; 358: j4008. <https://doi.org/10.1136/bmj.j4008>.
16. Waffenschmidt S, Navarro-Ruan T, Hobson N et al. Development and validation of study filters for identifying controlled non-randomized studies in PubMed and Ovid MEDLINE. Res Synth Methods 2020; 11(5): 617-626. <https://doi.org/10.1002/jrsm.1425>.
17. Moons KGM, Wolff RF, Riley RD et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. Ann Intern Med 2019; 170(1): W1-W33. <https://doi.org/10.7326/M18-1377>.
18. Sutton AJ, Abrams KR, Jones DR et al. Methods for meta-analysis in medical research. Chichester: Wiley; 2000.
19. Schulz A, Schürmann C, Skipka G et al. Performing Meta-analyses with Very Few Studies. In: Evangelou V, Veroniki AA (Ed). Meta-Research; Methods and Protocols. New York: Humana; 2022. S. 91-102.
20. Lilienthal J, Sturtz S, Schürmann C et al. Bayesian random-effects meta-analysis with empirical heterogeneity priors for application in health technology assessment with very few studies. Res Synth Methods 2024; 15(2): 275-287. <https://doi.org/10.1002/jrsm.1685>.
21. Veroniki AA, Jackson D, Bender R et al. Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis. Res Synth Methods 2019; 10(1): 23-43. <https://doi.org/10.1002/jrsm.1319>.
22. Wong SS, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

Anhang A Suchstrategien in bibliografischen Datenbanken

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to September 26, 2025

Es wurde folgender Filter übernommen:

- Systematische Übersicht: Wong [22] – High specificity strategy (adaptiert)

#	Searches
1	exp Prostatic Neoplasms/
2	(prostat* adj5 cancer*).ti,ab.
3	or/1-2
4	Mass Screening/
5	Early Detection of Cancer/
6	(prostat* adj2 screening*).ti,ab.
7	or/4-6
8	and/3,7
9	Cochrane database of systematic reviews.jn.
10	(search or MEDLINE or systematic review).tw.
11	(meta analysis or network meta analysis or systematic review).pt.
12	or/9-11
13	12 not (exp animals/ not humans.sh.)
14	and/8,13
15	14 and (english or german or multilingual or undetermined).lg.
16	limit 15 to yr="2022 -Current"

2. International HTA Database

Suchoberfläche: INAHTA

#	Searches
1	Prostatic Neoplasms[mhe]
2	prostat* AND cancer*
3	#2 OR #1
4	Mass Screening[mh]
5	Early Detection of Cancer[mh]
6	screening* OR screened*
7	#6 OR #5 OR #4
8	#7 AND #3
9	(*) FROM 2022 TO 2025
10	#9 AND #8