



Lungenkrebs

**Kann Künstliche Intelligenz die Auswertung von
Röntgen- oder CT-Bildern der Lunge
unterstützen?**

VORLÄUFIGER THEMENCHECK-BERICHT

Projekt: T24-04

Version: 1.0

Stand: 24.03.2026

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

Thema

Lungenkrebs: Kann Künstliche Intelligenz die Auswertung von Röntgen- oder CT-Bildern der Lunge unterstützen?

Projektnummer

T24-04

Beginn der Bearbeitung

09.01.2025

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: themencheck@iqwig.de

Internet: www.iqwig.de

Schlagwörter

Künstliche Intelligenz, Computertomografie, thorakale Röntgenuntersuchung, Lungentumoren, Nutzenbewertung, Systematische Übersicht, Technikfolgen-Abschätzung – biomedizinische

Keywords

Artificial Intelligence, Tomography – X-Ray Computed, Radiography – Thoracic, Lung Neoplasms, Benefit Assessment, Systematic Review, Technology Assessment – Biomedical

Autorinnen und Autoren des IQWiG (Nutzenbewertung, Gesundheitsökonomische Bewertung und inhaltliche Gesamtverantwortung)

- Alexandra Korzeczek-Opitz
- Daniela Rüttgers
- Dominik Schierbaum
- Claudia Lenkewitz
- Moritz Felsch
- Stefan Sauerland
- Corinna Schaefer
- Mareike Störchel
- Sibylle Sturtz
- Sarah Thys

Wissenschaftliche Beraterinnen und Berater

- Daniel Pinto dos Santos, Klinik und Poliklinik für Diagnostische und Interventionelle Radiologie, Universitätsmedizin Mainz der Johannes Gutenberg-Universität Mainz, Mainz

Autorinnen und Autoren der Domänen Ethik, Soziales, Recht, Organisation und Umwelt- und Klimaaspekten

- Anne Rummer, DARUM. wissen. einfach. vermitteln. Marion Danner und Anne Rummer GbR, Köln
- Marion Danner, DARUM. wissen. einfach. vermitteln. Marion Danner und Anne Rummer GbR, Köln

Beteiligung von Betroffenen

Im Rahmen der Berichterstellung wurden Betroffene konsultiert.

An dem Betroffenenengespräch nahmen 5 Personen teil.

Ziel des Betroffenenengesprächs war es, Informationen zu folgenden Themenbereichen zu erhalten: Erfahrungen, Wünsche und Sorgen in Bezug auf die AI-Unterstützung bei der Auswertung von Röntgen- oder CT-Bildern.

Die Autorinnen und Autoren des Berichts danken den Betroffenen für ihre Beteiligung an dem persönlichen Austausch. Die Betroffenen waren nicht in die Berichterstellung eingebunden.

Dieser Bericht wurde durch das IQWiG in Zusammenarbeit mit externen Sachverständigen erstellt: Für die Bearbeitung der Fragestellung zu ethischen, sozialen, rechtlichen und organisatorischen Aspekten sowie den Klima- und Umweltaspekten wurden vom IQWiG externe Sachverständige beauftragt. Alle weiteren Inhalte wurden durch das IQWiG erstellt. Externe Sachverständige, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Satz 2 Sozialgesetzbuch Fünftes Buch – Gesetzliche Krankenversicherung „alle Beziehungen zu Interessenverbänden, Auftragsinstituten, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie, einschließlich Art und Höhe von Zuwendungen“ offenzulegen. Das Institut hat von jedem der Sachverständigen ein ausgefülltes Formular „Formblatt zur Offenlegung von Beziehungen“ erhalten. Die Angaben wurden von dem speziell für die Beurteilung der Interessenkonflikte eingerichteten Gremium des Instituts bewertet. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden. Die Selbstangaben der externen Sachverständigen und der externen Reviewerin bzw. des externen Reviewers zur Offenlegung von Beziehungen sind in Kapitel A10 zusammenfassend dargestellt.

Dieser vorläufige ThemenCheck-Bericht wird zur Anhörung gestellt und es können schriftliche Stellungnahmen eingereicht werden. Das Ende der Stellungnahmefrist wird auf der Website des IQWiG (www.iqwig.de) bekannt gegeben. Dort sind auch die notwendigen Formblätter zu finden. Stellungnahmen können alle interessierten Personen, Institutionen und Gesellschaften abgeben. Ggf. wird eine wissenschaftliche Erörterung zur Klärung unklarer Aspekte aus den schriftlichen Stellungnahmen durchgeführt. Die Anhörung kann zu Änderungen und / oder Ergänzungen des ThemenCheck-Berichts führen.

ThemenCheck-Kernaussagen

Fragestellungen des ThemenCheck-Berichts

Fragestellung 1: AI-unterstützte Befundung bei Verdacht auf Lungenkrebs

Die Ziele der vorliegenden Untersuchung sind

- die Nutzenbewertung einer Befundung von Röntgenbildern oder einer Computertomografie (CT) der Lunge durch Radiologinnen und Radiologen mit Unterstützung von künstlicher Intelligenz im Vergleich zu deren Befundung ausschließlich durch Radiologinnen und Radiologen
 - bei erwachsenen Patientinnen und Patienten mit Verdacht auf Lungenkrebs.
- die Bestimmung der Kosten (Interventionskosten) und die Bewertung der Kosteneffektivität sowie
- die Aufarbeitung ethischer, sozialer, rechtlicher und organisatorischer Aspekte sowie Umwelt- und Klimaaspekte, die mit der medizinischen Intervention verbunden sind.
- Ergänzende Betrachtung einer Befundung von Röntgenbildern oder einer Computertomografie der Lunge allein durch künstliche Intelligenz im Vergleich zu deren Befundung ausschließlich durch Radiologinnen und Radiologen

Fragestellung 2: AI-unterstützte Befundung in der Lungenkrebsfrüherkennung

Ziele der vorliegenden Untersuchung sind

- die Nutzenbewertung einer Befundung einer Niedrigdosis- Computertomografie (Low-Dose-CT) der Lunge durch Radiologinnen und Radiologen mit Unterstützung von künstlicher Intelligenz im Vergleich zur Befundung ausschließlich durch Radiologinnen und Radiologen
- bei Personen, die an einer Lungenkrebsfrüherkennung teilnehmen und gemäß den Kriterien der deutschen Lungenkrebs-Früherkennungs-Verordnung (LuKrFrühErkV) für eine Teilnahme berechtigt wären.
 - die Bestimmung der Kosten (Interventionskosten) und die Bewertung der Kosteneffektivität sowie
 - die Aufarbeitung ethischer, sozialer, rechtlicher und organisatorischer Aspekte sowie Umwelt- und Klimaaspekte, die mit der medizinischen Intervention verbunden sind

Ergänzende Betrachtung einer Befundung von einer Niedrigdosis-Computertomografie (Low-Dose-CT) der Lunge allein durch künstliche Intelligenz im Vergleich zu deren Befundung ausschließlich durch Radiologinnen und Radiologen

Schlussfolgerung des ThemenCheck-Berichts

Dieser ThemenCheck-Bericht untersuchte bei Patientinnen und Patienten mit Verdacht auf Lungenkrebs und bei Personen, die an einem Lungenkrebsfrüherkennungsscreening teilnehmen, die Frage eines Bürgers, ob Ärztinnen und Ärzte mit AI-Unterstützung Röntgen- oder CT-Bilder ebenso gut oder besser auswerten können als ohne AI und welche Folgen dies für die Versorgung haben könnte. Die Ergebnisse dieses Themen-Check-Berichts zu beiden Aspekten dieser Frage, werden im Folgenden unter

- Auswirkung der AI-Unterstützung auf Nutzen, Schaden und Kosteneffektivität der radiologischen Befundung und
- Folgen der Implementierung einer AI-unterstützten Diagnostik

zusammengefasst:

Auswirkung der AI-Unterstützung auf Nutzen, Schaden und Kosteneffektivität der radiologischen Befundung

Durch die umfassende Informationsbeschaffung wurde deutlich, dass keine Studien zur diagnostisch-therapeutischen-Kette vorliegen, die eine direkte Aussage zur Auswirkung einer AI-unterstützten Diagnostik auf patientenrelevante Endpunkte ermöglichen würde. Zudem zeigte sich, dass Studien die Anwendung einer AI-unterstützten Befundung häufig nur dahin gehend untersuchten, wie gut die AI-gestützte Diagnostik verdächtige Lungenrundherde erkennen konnte. Die Studien haben jedoch häufig nicht verifiziert, ob die Personen mit verdächtigen Lungenrundherden tatsächlich Lungenkrebs hatten. Es wurden 2 aussagekräftige Studien zur diagnostischen Güte identifiziert, die im Rahmen des Linked-Evidence-Ansatzes eine Untersuchung zu Nutzen und Schaden ermöglichen. In diesen Studien wurden nur Personen eingeschlossen, die an einer Lungenkrebsfrüherkennungsuntersuchung teilnahmen (Fragestellung 2). Patientinnen und Patienten mit Verdacht auf Lungenkrebs (Fragestellung 1) wurden in diesen Studien nicht untersucht. Für Fragestellung 1 liegt somit keine Evidenz vor. Die klinische Heterogenität der beiden Studien ließ keine metaanalytische Zusammenfassung der Ergebnisse zu. Aus den Ergebnissen der Einzelstudien zur diagnostischen Güte einer Screeningrunde bzw. einer einmaligen Diagnostik lässt sich aufgrund der geringen qualitativen Ergebnissicherheit kein Nutzen oder Schaden der AI-unterstützten Befundung im Vergleich zu einer alleinigen radiologischen Befundung ableiten. Insgesamt ist die Datenlage zur Nutzenbewertung der Prüfinerventionen für beide Fragestellungen nicht ausreichend.

Es wurde keine gesundheitsökonomische Evaluation zur AI-unterstützten Befundung von Röntgen-Bildern bei Patientinnen und Patienten mit Verdacht auf Lungenkrebs identifiziert (Fragestellung 1). Für die AI-unterstützte Befundung von CT-Bildern hingegen schloss dieser ThemenCheck-Bericht 1 gesundheitsökonomische Evaluation mit separat ausgewiesenen

Ergebnissen für beide Populationen ein. Zwar zeigte das Modell für sich genommen robuste Ergebnisse für eine gesundheitsökonomische Überlegenheit der AI-Unterstützung in der Screeningpopulation bzw. für eine gesundheitsökonomische Unterlegenheit der AI-Unterstützung in der Verdachtspopulation. Dabei sind jedoch die unzureichende Datenlage hinsichtlich der vorliegenden Nutzenparameter sowie Einschränkungen hinsichtlich der Übertragbarkeit zu bedenken. Daher reichen die Ergebnisse der eingeschlossenen Modellierungsstudie in diesem ThemenCheck-Bericht nicht für eine abschließende Bewertung der Kosteneffektivität aus.

Neben der fehlenden Evidenz kommt ergänzend dazu, dass unterschiedliche AI-Softwares aktuell nicht vergleichbar sind. Somit ist auf Basis der wenigen eingeschlossenen Studien zu einzelnen AI-Produkten eine allgemeingültige Aussage zum Nutzen und zur Kosteneffektivität der AI-unterstützten Befundung nicht ableitbar.

Zusammenfassend kann dieser ThemenCheck-Bericht aufgrund fehlender und unzureichender Evidenz die Frage, ob mit einer AI-unterstützten Befundung Betroffene von Lungenkrebs im Rahmen einer Verdachts- oder einer Früherkennungsuntersuchung genauso gut oder besser erkannt werden, nicht beantworten.

Voraussetzungen und Folgen der Implementierung einer AI-unterstützten Diagnostik

Der Einsatz einer AI-Unterstützung bei der Befundung kann sich in verschiedenster Weise auf eine mögliche Implementierung in die Versorgung, Gesellschaft und Umwelt auswirken. Wesentliche Aspekte, welche in diesem ThemenCheck-Bericht in mehreren Domänen betrachtet wurden, werden im Folgenden beschrieben.

Kosten

Die Bewertung der ökonomischen als auch organisatorischen Aspekte zeigt, dass in Deutschland vor allem für die Leistungserbringenden zusätzliche Kosten entstehen: für Leistungserbringende liegen die zusätzlichen Kosten für die AI-Nutzung bei der Befundung von Röntgen- beziehungsweise CT-Bildern schätzungsweise in einem niedrigen 1-stelligen Euro-Bereich pro Befundung. Darüber hinaus müssen Leistungserbringende mit weiterem Ressourcenaufwand für die Implementierung einer AI-unterstützten Befundung rechnen, die u. a. aufgrund der Integration der AI in bestehende Systeme und Prozesse sowie der Schulung von Personal entstehen. Für die gesetzlichen Krankenkassen sowie die gesetzlich versicherten Personen sind in der Regel aktuell keine zusätzlichen direkten Kosten durch die AI-Unterstützung zu erwarten, da bislang spezifische Abrechnungsmöglichkeiten in den aktuellen Vergütungskatalogen nicht enthalten sind.

Zeitliche Effizienz

Es wurden sowohl positive als auch negative Erwartungen zu möglichen Änderungen in der zeitlichen Effizienz identifiziert. In der Nutzenbewertung waren die Ergebnisse zum ergänzenden Endpunkt Zeitersparnis, welche auf einer Screeningpopulation basieren, nicht aussagekräftig, da in beiden Studien nur ein Teilaspekt der radiologischen Befundung betrachtet wurde. Im Rahmen der Nutzenbewertung lag keine Evidenz für Patientinnen und Patienten mit Verdacht auf Lungenkrebs vor. Auch in der gesundheitsökonomischen Bewertung lag keine allgemeingültige Evidenz bezogen auf die Zeitersparnis im Gesamtprozess vor. In Deutschland ist für die Lungenkrebsfrüherkennung durch die aktuelle gesetzliche Regelung keine Zeitersparnis möglich, da Radiologinnen und Radiologen, die Befundung erst allein und anschließend nochmals mit Software-Unterstützung durchführen sollen. Die Bewertung der organisatorischen Aspekte kommt damit übereinstimmend zu dem Schluss, dass im Rahmen der Früherkennung wegen der Vorgaben der Lungenkrebs-Früherkennungs-Verordnung kein personelles Einsparpotenzial zu erkennen ist. Ob das auch für die Befundung von Personen mit Verdacht auf Lungenkrebs gilt, für die derzeit keine Vorgaben für den Einsatz von AI-Systemen bestehen, ist unklar. Basierend auf den Ergebnissen dieses ThemenCheck-Berichts ist vor allem in der Implementierungsphase einer AI-unterstützten Befundung nicht von einer zeitlichen Effizienz für Leistungserbringende oder für Patientinnen und Patienten und zu screenende Personen auszugehen.

Weitere Aspekte

Aus ethischer Sicht ist insbesondere die Wahrung der Autonomie von Patientinnen und Patienten sowie von Ärztinnen und Ärzten von herausragender Bedeutung. Daneben bestehen hohe gerechtigkeitsethische Anforderungen, sei es in Hinblick auf das Training von AI-Systemen unter Berücksichtigung auch unterrepräsentierter sozialer Gruppen als auch in Hinblick auf einen gerechten Zugang – sofern der Nutzen von AI-Systemen hinreichend belegt ist.

Rechtlich ist die Anwendung von AI-Systemen zur Unterstützung bei der Befundung von Röntgen- und CT Bildern durch ein vielschichtiges Regelungsgefüge geprägt, das Fragen der Aufklärung, des Datenschutzes, des Marktzugangs und der Haftung umfasst. Von hervorzuhebender Bedeutung ist, dass die Verantwortung für Befundentscheidungen stets bei den menschlichen Akteuren verbleibt.

Auf Seiten der Betroffenen und der Radiologinnen und Radiologen werden in Befragungsstudien zwar viele positive Erwartungen hinsichtlich des Nutzens einer AI-Unterstützung und unterschiedliche Erfahrungen mit der Nutzung einer spezifischen AI zur unterstützten Befundung durch Radiologinnen und Radiologen berichtet. Gleichzeitig adressieren verschiedene Befragungsstudien, Stellungnahmen und Positionspapiere aber auch soziale und organisatorische Herausforderungen bei der Umsetzung einer AI-

unterstützten Befundung. Risiken werden insbesondere im Hinblick auf mögliche soziale Ungleichheiten im Zugang zu oder bei der Inanspruchnahme von AI-Leistungen gesehen. Organisatorisch besteht zudem das Risiko von De-Skilling-Effekten, die durch qualifikations-sichernde Maßnahmen und angepasste Arbeitsabläufe möglicherweise jedoch minimiert werden können. Ergänzend machen die erheblichen Umwelt- und Klimaauswirkungen der AI-Systeme – von Energie- und Wasserverbrauch bis hin zur Ressourcenabbau und Abfallmanagement – deutlich, dass auch ökologische Nachhaltigkeit in Entscheidungs- und Implementierungsprozessen systematisch berücksichtigt werden muss.

Schlussendlich ist ein Einsatz einer AI-unterstützten radiologischen Befundung erst dann vertretbar, wenn der patientenrelevante Nutzen, etwa die Verringerung von Mortalität und Morbidität sowie Verbesserung der gesundheitsbezogenen Lebensqualität, den Schaden, beispielsweise durch Fehl- oder Überdiagnosen für Patientinnen und Patienten oder für die untersuchten Personen überwiegt. Der vorliegende ThemenCheck-Bericht liefert allerdings mangels ausreichender Datenlage für keine der beiden Fragestellungen eine Nutzen- oder Schadenaussage. Unklar ist somit, ob die AI-Unterstützung bei der Auswertung von Röntgen- oder CT-Bildern mit einem patientenrelevanten Nutzen einhergeht. Die eingeschränkte Vergleichbarkeit zwischen AI-Anwendungen, die zu erwartenden Unterschiede in der diagnostischen Güte sowie die unterschiedlichen Anwendungsmöglichkeiten der AI verdeutlichen die Notwendigkeit weiterer Evidenzgenerierung für die Fragestellungen des Berichts: insbesondere sind Studien zur diagnostisch-therapeutischen Kette notwendig, um die Konsequenzen für die Therapie und die patientenrelevanten Endpunkte abzuschätzen. Des Weiteren sollten Studien zur diagnostischen Güte prospektiv geplant werden und eine ausreichende Teilnehmeranzahl umfassen, um zu aussagekräftigen Ergebnissen zu führen (siehe auch Einschlusskriterien der Fragestellung). Die aktuelle Attraktivität von AI-Systemen alleine, die in allen Lebensbereichen zu beobachten ist, kann vor diesem Hintergrund keine ausreichende Begründung für ihren routinemäßigen Einsatz bei der radiologischen Befundung von Lungenkrebs darstellen, da dies von der aktuellen Datenlage nicht gestützt wird.

Inhaltsverzeichnis

	Seite
ThemenCheck-Kernaussagen.....	6
Tabellenverzeichnis	18
Abbildungsverzeichnis	20
Abkürzungsverzeichnis.....	21
 ThemenCheck-Überblick	 23
1 Hintergrund	23
1.1 Gesundheitspolitischer Hintergrund und Auftrag.....	23
1.2 Medizinischer Hintergrund.....	23
1.3 Versorgungssituation	26
1.4 Anliegen des Themenvorschlagenden	26
2 Fragestellungen	27
2.1 Fragestellung 1: AI-unterstützte Befundung bei Verdacht auf Lungenkrebs.....	27
2.2 Fragestellung 2: AI-unterstützte Befundung in der Lungenkrebsfrüherkennung	27
Methoden	28
2.3 Methoden Nutzenbewertung.....	28
2.4 Methoden gesundheitsökonomische Bewertung.....	30
2.5 Methoden ethische Aspekte	31
2.6 Methoden soziale, rechtliche und organisatorische Aspekte sowie Umwelt- und Klimaaspekte	32
2.7 Austausch mit Betroffenen	33
3 Ergebnisse: Nutzenbewertung.....	34
3.1 Ergebnisse der umfassenden Informationsbeschaffung	34
3.2 Charakteristika der in die Bewertung eingeschlossenen Studien (Fragestellung 2)	34
3.3 Übersicht der patientenrelevanten Endpunkte (Fragestellung 2)	35
3.4 Bewertung des Verzerrungspotenzials der Ergebnisse aus Studien zur diagnostischen Güte (Fragestellung 2)	35

3.5	Bewertung der Übertragbarkeit der Ergebnisse aus Studien zur diagnostischen Güte (Fragestellung 2)	36
3.6	Ergebnisse der Studien zur diagnostischen Güte (Fragestellung 2)	36
3.6.1	AI-unterstützte Befundung vs. menschliche Befundung (Indextest 1 vs. Indextest 2)	37
3.6.2	Ergänzende Darstellung der alleinigen AI-Befundung.....	38
3.6.3	Ergebnisse zum ergänzenden dargestellten Endpunkt Zeitersparnis	39
3.7	Zusammenfassende Bewertung der Ergebnisse	39
4	Ergebnisse: Gesundheitsökonomische Bewertung	41
4.1	Interventionskosten.....	41
4.2	Systematische Übersicht gesundheitsökonomischer Evaluationen.....	43
4.2.1	Ergebnisse der Informationsbeschaffung.....	43
4.2.2	Charakteristika der in die Bewertung eingeschlossenen Studien	43
4.2.3	Ergebnisse gesundheitsökonomischer Evaluationen	45
4.2.4	Diskussion und Einordnung der gesundheitsökonomischen Ergebnisse für den vorliegenden Bericht	48
5	Ergebnisse: Ethische, soziale, rechtliche und organisatorische Aspekte sowie Umwelt- und Klimaaspekte	50
5.1	Gespräche mit Betroffenen	50
5.2	Ergebnisse zu ethischen Aspekten	50
5.2.1	Erwarteter Nutzen	51
5.2.2	Potenzieller Schaden und Belastungen	52
5.2.3	Auswirkungen auf die Autonomie	53
5.2.4	Auswirkungen auf die Gerechtigkeit	54
5.2.5	Erwartete Effizienz.....	55
5.2.6	Ausblick.....	55
5.3	Ergebnisse zu sozialen Aspekten	56
5.4	Ergebnisse zu rechtlichen Aspekten.....	61
5.4.1	Patientenautonomie I: informierte Einwilligung.....	62
5.4.2	Autonomie III: Datenschutz	63
5.4.3	Marktzugang I: Medizinprodukte	64
5.4.4	Klinische Studien.....	64
5.4.5	Kostenerstattung im öffentlichen Gesundheitssystem.....	64
5.4.6	Exkurs: Gibt es ein Recht auf eine AI-unterstützte Auswertung von Röntgen- oder CT-Aufnahmen der Lunge? Selbstbestimmungsrecht des Patienten vs. ärztliche Therapiefreiheit	65
5.4.7	Spezielle medizinische Bereiche	65

5.4.7.1	Arztvorbehalt	66
5.4.7.2	Pflicht zur Überprüfung von AI-Ergebnissen.....	66
5.4.7.3	Haftung.....	66
5.4.8	Ausblick.....	68
5.5	Ergebnisse zu organisatorischen Aspekten	69
5.6	Ergebnisse zu Umwelt- und Klimaaspekten	73
6	Domänenübergreifende Zusammenführung der Ergebnisse.....	77
7	Diskussion.....	78
7.1	Einordnung der Arbeitsergebnisse.....	78
7.2	ThemenCheck-Bericht im Vergleich zu anderen Publikationen.....	80
7.3	ThemenCheck-Bericht im Vergleich zu Leitlinien.....	81
7.4	Kritische Reflexion des Vorgehens.....	82
8	Schlussfolgerung	84
ThemenCheck-Details		88
A1	Projektverlauf.....	88
A1.1	Zeitlicher Verlauf des Projekts.....	88
A1.2	Spezifizierungen und Änderungen im Projektverlauf	89
A2	Details der Methoden – Methodik gemäß ThemenCheck-Berichtsprotokoll	90
A2.1	Nutzenbewertung.....	90
A2.1.1	Kriterien für den Einschluss von Studien zur diagnostisch-therapeutischen Behandlungskette.....	91
A2.1.1.1	Population	91
A2.1.1.2	Prüf- und Vergleichsintervention.....	91
A2.1.1.3	Patientenrelevante Endpunkte	92
A2.1.1.4	Studientypen	92
A2.1.1.5	Studiendauer.....	93
A2.1.1.6	Publikationssprache	93
A2.1.1.7	Tabellarische Darstellung der Kriterien für den Studieneinschluss.....	93
A2.1.1.8	Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	94
A2.1.2	Kriterien für den Einschluss von Studien zur diagnostischen Güte.....	94
A2.1.2.1	Population	94
A2.1.2.2	Index- und Referenztest.....	95
A2.1.2.3	Zielgröße	95

A2.1.2.4	Studientypen	96
A2.1.2.5	Studiendauer	96
A2.1.2.6	Publikationssprache	96
A2.1.2.7	Tabellarische Darstellung der Kriterien für den Studieneinschluss	97
A2.1.2.7.1	Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	98
A2.1.3	Informationsbeschaffung	98
A2.1.3.1	Fokussierte Informationsbeschaffung von systematischen Übersichten...	98
A2.1.3.2	Umfassende Informationsbeschaffung von Studien	99
A2.1.3.3	Anwendung von Limitierungen auf Datenbankebene	100
A2.1.3.4	Selektion relevanter Studien aus der umfassenden Informationsbeschaffung	101
A2.1.4	Informationsbewertung und -synthese	101
A2.1.4.1	Darstellung der Einzelstudien	101
A2.1.4.2	Bewertung des Verzerrungspotenzials der Ergebnisse	102
A2.1.4.2.1	Ergebnisse aus Studien zur diagnostisch-therapeutischen Behandlungskette	102
A2.1.4.2.2	Ergebnisse aus Studien zur diagnostischen Güte	103
A2.1.4.3	Metaanalysen	103
A2.1.4.3.1	Ergebnisse aus Studien zur diagnostisch-therapeutischen Behandlungskette	103
A2.1.4.3.2	Ergebnisse aus Studien zur diagnostischen Güte	104
A2.1.4.4	Sensitivitätsanalysen	105
A2.1.4.5	Subgruppenmerkmale und andere Effektmodifikatoren	105
A2.1.4.6	Aussagen zur Beleglage	106
A2.2	Gesundheitsökonomische Bewertung	107
A2.2.1	Interventionskosten	107
A2.2.2	Systematische Übersicht gesundheitsökonomischer Evaluationen	108
A2.2.2.1	Kriterien für den Einschluss von Studien in die systematische Übersicht	108
A2.2.2.1.1	Studientypen	108
A2.2.2.1.2	Publikationssprache	108
A2.2.2.1.3	Gesundheitssystem bzw. geografischer Bezug	108
A2.2.2.1.4	Tabellarische Darstellung der Kriterien	108
A2.2.2.2	Fokussierte Informationsbeschaffung	109
A2.2.2.2.1	Selektion relevanter Publikationen	109
A2.2.2.3	Informationsbewertung	109
A2.2.2.4	Informationsanalyse und -synthese	110

A2.3	Ethische Aspekte.....	110
A2.3.1	Berücksichtigung von Argumenten und Aspekten bei der Aufarbeitung ethischer Implikationen.....	110
A2.3.2	Informationsbeschaffung	110
A2.3.3	Informationsaufbereitung	111
A2.4	Soziale, rechtliche und organisatorische Aspekte sowie Umwelt- und Klimaaspekte	112
A2.4.1	Berücksichtigung von Argumenten und Aspekten bei der Aufarbeitung sozialer, rechtlicher und organisatorischer Implikationen sowie von Umwelt- und Klimaaspekten	112
A2.4.2	Informationsbeschaffung	112
A2.4.3	Informationsaufbereitung	113
A2.5	Domänenübergreifende Zusammenführung.....	114
A3	Details der Ergebnisse: Nutzenbewertung	115
A3.1	Informationsbeschaffung	115
A3.1.1	Fokussierte Informationsbeschaffung nach systematischen Übersichten	115
A3.1.2	Umfassende Informationsbeschaffung	115
A3.1.2.1	Primäre Informationsquellen.....	115
A3.1.2.1.1	Bibliografische Datenbanken	115
A3.1.2.1.2	Studienregister	116
A3.1.2.2	Weitere Informationsquellen und Suchtechniken.....	117
A3.1.2.2.1	Anwendung weiterer Suchtechniken	117
A3.1.3	Resultierender Studienpool.....	117
A3.1.4	Studien ohne berichtete Ergebnisse	118
A3.2	Charakteristika der in die Bewertung eingeschlossenen Studien	118
A3.2.1	Studiendesign und Studienpopulationen	118
A3.2.2	Bewertung des Verzerrungspotenzials der Studien zur diagnostischen Güte der AI-unterstützten Befundung im Vergleich zur ausschließlichen Befundung durch Radiologinnen und Radiologen.....	122
A3.3	Ergebnisse der Studien zur diagnostischen Güte	123
A3.3.1	Ergebnisse zur diagnostischen Güte der AI-unterstützten Befundung im Vergleich zur ausschließlichen Befundung durch Radiologinnen und Radiologen (Fragestellung 2).....	123
A3.3.2	Ergänzendes Ergebnis zur diagnostischen Güte der alleinigen AI-Befundung	125
A3.3.3	Ergänzend dargestellter Endpunkt Zeitersparnis	126
A4	Details der Ergebnisse: Gesundheitsökonomische Bewertung.....	127
A4.1	Bestimmung der Interventionskosten	127

A4.2	Systematische Übersicht gesundheitsökonomischer Evaluationen.....	130
A4.2.1	Fokussierte Informationsbeschaffung.....	130
A4.2.1.1	Primäre Informationsquellen.....	130
A4.2.1.2	Weitere Informationsquellen und Suchtechniken.....	131
A4.2.1.3	Resultierender Studienpool	132
A4.2.2	Charakteristika der in die Bewertung eingeschlossenen Studien	132
A4.2.2.1	Studiendesign.....	133
A4.2.2.2	Inputparameter.....	136
A4.2.3	Ergebnisse gesundheitsökonomischer Evaluationen	138
A4.2.4	Bewertung der Berichtsqualität und Übertragbarkeit	142
A4.2.4.1	Bewertung der Berichtsqualität.....	142
A4.2.4.2	Bewertung der Übertragbarkeit	145
A5	Details der Ergebnisse: Ethische, soziale, rechtliche und organisatorische Aspekte sowie Umwelt- und Klimaaspekte	148
A5.1	Ethische Aspekte.....	148
A5.1.1	Informationsbeschaffung zu ethischen Aspekten der zu bewertenden Technologie	148
A5.1.2	Identifizierte ethische Aspekte.....	148
A5.2	Soziale Aspekte.....	157
A5.2.1	Informationsbeschaffung zu sozialen Aspekten der zu bewertenden Technologie	157
A5.2.2	Identifizierte soziale Aspekte	158
A5.3	Rechtliche Aspekte	171
A5.3.1	Informationsbeschaffung zu rechtlichen Aspekten der zu bewertenden Technologie	171
A5.3.2	Identifizierte rechtliche Aspekte	171
A5.4	Organisatorische Aspekte	184
A5.4.1	Informationsbeschaffung zu organisatorischen Aspekten der zu bewertenden Technologie.....	184
A5.4.2	Identifizierte organisatorische Aspekte.....	184
A5.5	Umwelt- und Klimaaspekte	196
A5.5.1	Informationsbeschaffung zu Umwelt- und Klimaaspekten der zu bewertenden Technologie.....	196
A5.5.2	Identifizierte Umwelt- und Klimaaspekte.....	196
A6	Literatur.....	198
A7	Studienlisten.....	216
A7.1	Studienlisten Nutzenbewertung	216

A7.1.1	Liste der gesichteten systematischen Übersichten.....	216
A7.1.2	Liste der ausgeschlossenen Publikationen aus der bibliografischen Recherche	216
A7.2	Studienlisten gesundheitsökonomische Bewertung.....	234
A7.2.1	Liste der gesichteten systematischen Übersichten.....	234
A7.2.2	Liste der ausgeschlossenen Publikationen zum Thema	234
A7.3	Publikationslisten zu ethischen, sozialen, rechtlichen und organisatorischen Aspekten sowie Umwelt- und Klimaaspekten.....	235
A7.3.1	Liste der eingeschlossenen Publikationen zu ethischen Aspekten	235
A7.3.2	Liste der eingeschlossenen Publikationen zu sozialen Aspekten.....	237
A7.3.3	Liste der eingeschlossenen Publikationen zu rechtlichen Aspekten.....	240
A7.3.4	Liste der eingeschlossenen Publikationen zu organisatorischen Aspekten....	244
A7.3.5	Liste der eingeschlossenen Publikationen zu Umwelt- und Klimaaspekten ...	247
A8	Suchstrategien	248
A8.1	Suchstrategien zur Nutzenbewertung.....	248
A8.1.1	Bibliografische Datenbanken.....	248
A8.1.2	Studienregister	253
A8.2	Suchstrategien zur gesundheitsökonomischen Bewertung.....	254
A9	Zentrale Ergebnisse aus den Betroffeneninterviews	258
A10	Offenlegung von Beziehungen der externen Sachverständigen und des Reviewers bzw. der Reviewerin.....	263

Tabellenverzeichnis

	Seite
Tabelle 1: Übersicht über die Kriterien für den Einschluss von Studien zur diagnostisch-therapeutischen Behandlungskette in die Nutzenbewertung (Fragestellung 1)	93
Tabelle 2: Übersicht über die Kriterien für den Einschluss von Studien zur diagnostisch-therapeutischen Behandlungskette in die Nutzenbewertung (Fragestellung 2)	94
Tabelle 3: Übersicht über die Kriterien für den Einschluss von Studien zur diagnostischen Güte in die Nutzenbewertung (Fragestellung 1)	97
Tabelle 4: Übersicht über die Kriterien für den Einschluss von Studien zur diagnostischen Güte in die Nutzenbewertung (Fragestellung 2)	98
Tabelle 5: Regelmäßig abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit	107
Tabelle 6: Übersicht über die zusätzlichen Kriterien für den Einschluss von Studien in die ökonomische Bewertung	109
Tabelle 7: In Studienregistern identifizierte relevante Studien bzw. Dokumente	117
Tabelle 8: Resultierender Studienpool der Nutzenbewertung	117
Tabelle 9: In der Informationsbeschaffung identifizierte Studien ohne berichtete Ergebnisse	118
Tabelle 10: Charakterisierung der eingeschlossenen Studien zur diagnostischen Güte (Fragestellung 2)	118
Tabelle 11: Index- und Referenzstandard (Fragestellung 2)	119
Tabelle 12: Ein- / Ausschlusskriterien für Personen in den Studien (Fragestellung 2)	121
Tabelle 13: Charakterisierung der Studienpopulationen (Fragestellung 2)	121
Tabelle 14: Verzerrungspotenzial nach QUADAS 2 der Studien zur diagnostischen Güte (Fragestellung 2)	122
Tabelle 15: Bedenken bezüglich der Übertragbarkeit QUADAS 2 der Studien zur diagnostischen Güte (Fragestellung 2)	123
Tabelle 16: Ergebnisse der Studien zur diagnostischen Güte der AI-unterstützten Befundung im Vergleich zur Befundung ausschließlich durch Radiologinnen und Radiologen (Fragestellung 2)	124
Tabelle 17: Ergebnis zur diagnostischen Güte der alleinigen AI-Befundung (Fragestellung 2)	125
Tabelle 18: Ergebnisse – Ergänzender Endpunkt Zeitersparnis (Fragestellung 2)	126
Tabelle 19: Kosten für AI-Anwendungen aus Sicht von Selbstzahlenden	127
Tabelle 20: Kosten verschiedener AI-Systeme zur Unterstützung der Befundung von CT-Aufnahmen aus Sicht der Leistungserbringenden	128

Tabelle 21: Kosten verschiedener AI-Systeme zur Unterstützung der Befundung von Röntgenaufnahmen für Leistungserbringende	129
Tabelle 22: Durch Anwendung weiterer Suchtechniken identifizierte relevante Dokumente	132
Tabelle 23: Studienpool der gesundheitsökonomischen Bewertung	132
Tabelle 24: Studiencharakteristika	133
Tabelle 25: Modell	134
Tabelle 26: Daten zum Nutzen	136
Tabelle 27: Daten zu Nutzwerten	136
Tabelle 28: Daten zu Kosten	137
Tabelle 29: Ergebnisse Kosteneffektivität	139
Tabelle 30: Bewertung der Berichtsqualität	142
Tabelle 31: Bewertung der Übertragbarkeit	146
Tabelle 32: Informationsaufbereitung zu den identifizierten ethischen Aspekten	149
Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten	158
Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten	171
Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten ..	185
Tabelle 36: Informationsaufbereitung zu den identifizierten Umwelt- und Klimaaspekten ..	196

Abbildungsverzeichnis

	Seite
Abbildung 1: Zusammenfassung der Ergebnisse des vorliegenden ThemenCheck-Berichts als logisches Modell in Anlehnung an INTEGRATE-HTA	77
Abbildung 2: Ergebnis der umfassenden Informationsbeschaffung aus den bibliografischen Datenbanken und der Studienselektion – Nutzenbewertung.....	116
Abbildung 3: Ergebnis der bibliografischen Recherche und der Studienselektion – gesundheitsökonomische Evaluationen	131

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AI	Artificial Intelligence (künstliche Intelligenz)
AHRQ	Agency for Healthcare Research and Quality
BDSG	Bundesdatenschutzgesetz
BGB	Bürgerliches Gesetzbuch
BMG	Bundesministerium für Gesundheit
BTS	British Thoracic Society
CT	Computertomografie
CADe	Computer-Aided Detection (computer-unterstützte Erkennung)
CADx	Computer-Aided Diagnosis (computer-unterstützte Diagnose)
CHEERS-AI	Consolidated Health Economic Evaluation Reporting Standards for Interventions That Use Artificial Intelligence
CONSORT	Consolidated Standards of Reporting Trials
DigiG	Digitalgesetz
DRG	Diagnosis Related Groups (diagnosebezogene Fallgruppen)
DSGVO	Datenschutz-Grundverordnung
EBM	Einheitlicher Bewertungsmaßstab
EHDS	European Health Data Space (Europäischen Gesundheitsdatenraum)
ESTI	European Society of Thoracic Imaging
G-BA	Gemeinsamer Bundesausschuss
GDNG	Gesundheitsdatennutzungsgesetz
GKV	gesetzliche Krankenversicherung
GOÄ	Gebührenordnung für Ärzte
HTA	Health Technology Assessment (Gesundheitstechnologiebewertung)
ICER	Incremental Cost-Effectiveness Ratio (Kosteneffektivitätsverhältnis)
IFA	Informationsstelle für Arzneispezialitäten
ITT	Intention to treat
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KFE-RL	Krebsfrüherkennungsrichtlinie
K-LUCAS	Korean Lung-Cancer Screening Project
LuKrFrühErkV	Lungenkrebs-Früherkennungs-Verordnung
Lung-RADS	Lung Imaging Reporting and Data System

Abkürzung	Bedeutung
Low-Dose-CT	Niedrigdosis-Computergrafie
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NLST	National Lung Screening Trials
NPV	Negative Predictive Value (negativ prädiktiver Wert)
QALY	Qualitätsadjustiertes Lebensjahr
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
PPV	Positive Predictive Value (positiv prädiktiver Wert)
PPS	Personal Social Services
SGB	Sozialgesetzbuch
SÜ	systematische Übersicht
UICC	Union for International Cancer Control
VOPT	Verification of only positive Testers
ZB Med	deutsche Zentralbibliothek für Medizin

ThemenCheck-Überblick

1 Hintergrund

1.1 Gesundheitspolitischer Hintergrund und Auftrag

Der § 139b Absatz 5 des Sozialgesetzbuchs – Fünftes Buch – gesetzliche Krankenversicherung (SGB V) sieht vor, dass Versicherte und interessierte Einzelpersonen beim Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) Themen für die wissenschaftliche Bewertung von medizinischen Verfahren und Technologien vorschlagen können. Die Vorschläge für diese ThemenCheck-Berichte können auf der Website des IQWiG eingereicht werden.

Ziel des ThemenCheck Medizin ist es, die Beteiligung der Bürgerinnen und Bürger an der evidenzbasierten Medizin zu fördern und Antworten auf Fragestellungen zu finden, die für die Versorgung von Patientinnen und Patienten besonders bedeutsam sind.

1-mal pro Jahr wählt das IQWiG unter Beteiligung von Patientenvertreterinnen und -vertretern sowie Bürgerinnen und Bürgern bis zu 5 Themen für die Erstellung von ThemenCheck-Berichten aus. Die Bearbeitung der Fragestellung erfolgt durch das IQWiG in Zusammenarbeit mit externen Sachverständigen.

Die ThemenCheck-Berichte leitet das IQWiG an Institutionen in Deutschland weiter, die z. B. über die Leistungen und die Struktur des Gesundheitswesens entscheiden. Der Fachöffentlichkeit werden die ThemenCheck-Berichte über die Website des IQWiG (www.iqwig.de) zur Verfügung gestellt. Ebenso wird eine allgemein verständliche Darstellung der Ergebnisse mit dem Titel „ThemenCheck kompakt: Das Wichtigste verständlich erklärt“ veröffentlicht.

1.2 Medizinischer Hintergrund

Lungenkrebs, auch Lungen- oder Bronchialkarzinom genannt, ist ein bösartiger Tumor in der Lunge. Die Symptome der Erkrankung sind unspezifisch. Abhängig von der Größe und Lage des Lungenkrebses können Patientinnen und Patienten chronischen Husten, Atemnot (Dyspnoe), Schmerzen im Brustbereich, Gewichtsverlust oder Abhusten von bluthaltigem Sekret (Hämoptyse) aufweisen. Auch Nachtschweiß, Fieber und eine deutliche Abnahme der Leistungsfähigkeit (Leistungsknick) sind mögliche Symptome [1].

Lungenkrebs war in Deutschland im Jahr 2023 die zweithäufigste Krebsneuerkrankung bei Männern (nach Prostatakrebs) und bei Frauen nach Brustkrebs und Darmkrebs die

drithäufigste. Dabei war Lungenkrebs unter Männern die häufigste und unter Frauen die zweithäufigste krebssbedingte Todesursache (nach Brustkrebs) [2].

Hauptrisikofaktor für Lungenkrebs ist das Rauchen: etwa 9 von 10 Männern bzw. 8 von 10 Frauen mit Lungenkrebs erkrankten aufgrund von aktivem Rauchen. Weitere Risikofaktoren sind Passivrauchen, Luftverschmutzung, Exposition gegenüber Karzinogenen (z. B. Asbest oder Radon) [2].

Histologisch wird Lungenkrebs grob nach kleinzelligem Lungenkrebs (Small Cell Lung Cancer [SCLC]) oder nicht kleinzelligem Lungenkrebs (Non-small Cell Lung Cancer [NSCLC]) unterteilt. Der nicht kleinzellige Lungenkrebs wiederum wird histologisch weiter, unter anderem in das Adenokarzinom und dem Plattenepithelkarzinom differenziert [1]. Adenokarzinome machen 44 % aller Lungenkrebsfälle aus, etwa 21 % entfallen auf das Plattenepithelkarzinom und etwa 15 % auf das kleinzellige Bronchialkarzinom [2].

Neben der histologischen Einordnung wird der Schweregrad der Krebserkrankung durch die TNM-Klassifikation der Union for International Cancer Control (UICC) beschrieben [3]. Im Rahmen der TNM-Klassifikation wird die jeweilige Lungenkrebserkrankung hinsichtlich der Ausbreitung des Primärtumors, dem Lymphknotenbefall und dem Metastasierungsgrad eingestuft und kann dann einem Stadium zugeordnet werden. Die Stadieneinteilung des Lungenkrebs dient zusammen mit der histologischen Einteilung der Therapieplanung und der Prognoseabschätzung [4].

Die Prognose bei Lungenkrebs ist häufig ungünstig: so liegt die relative 5-Jahres-Überlebensrate für 2021 / 2022 / 2023 bei etwa 25 % für Frauen und bei 19 % für Männer. Der kleinzellige Lungenkrebs weist die schlechteste Prognose auf, da er zu einer frühen Metastasierung neigt: hier liegt die relative 5-Jahres-Überlebensrate verglichen mit der Allgemeinbevölkerung bei 12 % für Frauen und bei 7 % für Männer [2]. Die geringe 5-Jahres-Überlebensrate hängt laut S3-Leitlinie „Prävention, Diagnostik, Therapie und Nachsorge des Lungenkarzinoms“ unter anderem damit zusammen, dass in den Jahren 2021 bis 2023 unter allen Lungenkrebs-Erstdiagnosen insgesamt 54 % der Männer und 51 % der Frauen erst in den fortgeschrittenen UICC-Stadien III und IV diagnostiziert wurden. Die Prognose für UICC-Stadien I und II oder für asymptomatischer Patientinnen und Patienten ist günstiger [4].

Je nach Stadieneinteilung erfolgt die Behandlung des Lungenkrebses kurativ, d. h. mit dem Ziel der Heilung durch eine Operation und / oder einer Strahlentherapie mit oder ohne systemische Therapie (bspw. Chemotherapie), oder palliativ. Eine palliative Behandlung kann neben einer Operation auch eine Strahlentherapie und / oder eine systemische Therapie beinhalten. Bei einer palliativen Behandlung steht die Linderung der Begleitsymptomatik und eine Verbesserung der Lebensqualität im Vordergrund [4].

Um einen Lungenkrebsverdacht näher abzuklären und eine Lungenkrebsdiagnose zu stellen, sind mehrere Untersuchungen erforderlich. Wenn Patientinnen und Patienten mit Beschwerden eine Ärztin oder einen Arzt aufsuchen, erfolgt zunächst eine Anamnese und klinische Untersuchung sowie Laboruntersuchungen. Als weitere Basisdiagnostik des Lungenkarzinoms schließen sich bildgebende Verfahren zunächst in Form einer Röntgenuntersuchung und einer Computertomografie (CT) des Thorax (Brustkorbs) an. Die abschließende Diagnosesicherung bei einem auffälligen Befund in der Bildgebung erfolgt mittels Bronchoskopie. Im Rahmen der Bronchoskopie können Gewebeproben entnommen werden (sog. bronchoskopische transbronchiale Biopsie). Alternative Verfahren sind die CT-gesteuerte perkutane Biopsie oder die videoassistierte thorakoskopische Biopsie, letztere kann auch mit therapeutischer Zielsetzung durchgeführt werden. Die Gewebeproben werden dann histologisch untersucht. Nach der pathologischen Bestätigung der Malignität sind für die sich anschließende Therapie und die Prognose die Differenzierung des Tumortyps (Grading) und die Ausbreitung (Staging) entscheidend [4].

Die Lungenkrebsfrüherkennung mittels Niedrigdosis-Computertomografie (Low-Dose-CT) ist für (ehemalige) starke Raucherinnen und Raucher zwischen 50 und 76 Jahren ohne Verdacht auf Lungenkrebs zulässig [5]. Das Früherkennungsprogramm startet mit einer bildgebenden Untersuchung. Bei einem kontrollbedürftigen Befund schließt sich innerhalb einiger Monate (in jedem Fall vor der nächsten Früherkennung) eine weitere CT-Untersuchung an. Bei einem abklärungsbedürftigen Befund erfolgt die abschließende Diagnosesicherung mit den oben genannten invasiven Verfahren [5].

Bei den bildgebenden Verfahren werten Radiologinnen und Radiologen die Röntgen- oder CT-Bilder nach Veränderungen der Lunge aus, die auf bestimmte Erkrankungen hinweisen können. Durch Computerprogramme mit künstlicher Intelligenz (AI [artificial intelligence]), wie im AI Act [6] definiert, besteht die Möglichkeit, Radiologinnen und Radiologen bei der Auswertung / Befundung zu unterstützen. Die Unterstützung bei der Auswertung / Befundung kann auf unterschiedliche Art erfolgen: als computer-unterstützte Erkennung (CAdE, engl. „computer-aided detection“) oder als computer-unterstützte Diagnose (CAdx, engl. „computer-aided diagnosis“). Eine CAdE-Software erkennt abnormale Strukturen in den Schnittbildern und kennzeichnet diese zum Beispiel durch eine farbliche Umrandung. Eine CAdx-Software erkennt ebenfalls abnormale Strukturen und ordnet diese einer entsprechenden Diagnose zu [7]. Diese Computerprogramme sollen Radiologinnen und Radiologen bei der Befundung unterstützen, insbesondere bei der Bewertung des Wachstums von Lungenrundherden und der Erkennung von Auffälligkeiten sowie bei der Entscheidung über die weitere Diagnostik. Für diesen Bericht wird eine CAdE- als auch eine CAdx-Software als AI im Rahmen einer AI-unterstützten Befundung verstanden. Zudem ist eine potentiell denkbare Möglichkeit, dass zukünftig solche Computerprogramme die Befundung autonom übernehmen könnten. Für Patientinnen und Patienten mit Verdacht auf Lungenkrebs oder

Personen, die an einer Lungenkrebsfrüherkennung teilnehmen, soll die Versorgungsqualität mindestens gleichbleiben. Eine gute Versorgungsqualität bedeutet, dass die folgenden Qualitätsdimensionen in der Versorgung berücksichtigt werden: Patientensicherheit, Wirksamkeit, Ausrichtung der Versorgungsgestaltung an Patientinnen und Patienten, Rechtzeitigkeit und Verfügbarkeit, Angemessenheit sowie Koordination und Kontinuität [8].

1.3 Versorgungssituation

Für die Versorgung von Patientinnen und Patienten mit Verdacht auf Lungenkrebs arbeiten verschiedene Fachärztinnen und Fachärzte z. B. für Allgemeinmedizin, für Lungen- und Bronchialheilkunde (Pneumologie) oder für Radiologie mit Untersuchungslaboren zusammen. Die Behandlung von Lungenkrebs erfolgt zunächst meist stationär. Die Rehabilitation und Nachsorge kann in spezialisierten Rehabilitationskliniken oder, bei einer fortführenden Nachsorge, ambulant in pneumologischen, hausärztlichen oder onkologischen Praxen erfolgen [9].

Die Verordnung über die Zulässigkeit der Anwendung der low dose-CT zur Früherkennung von Lungenkrebs bei rauchenden Personen, die Lungenkrebs-Früherkennungs-Verordnung (LuKrFrühErkV), wurde 2024 vom Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz erlassen [10]. Lungenkrebsfrüherkennung mittels Low-Dose-CT, nicht aber mittels Röntgenaufnahmen, ist unter bestimmten Bedingungen (siehe Abschnitt 1.2) erlaubt [5]. Die Anforderungen an das medizinische Personal sowie an die Durchführung der Untersuchung und an die Befundung der Aufnahmen, einschließlich der Software zur computerassistenten Detektion sind ebenfalls in der LuKrFrühErkV festgelegt. Der Gemeinsame Bundesausschuss (G-BA) hat im Dezember 2025 die Einführung der Lungenkrebsfrüherkennung beschlossen. Ab April 2026 soll die Implementierung in die aktuelle Versorgung erfolgen [11].

1.4 Anliegen des Themenvorschlagenden

Der Themenvorschlagende berichtet, dass eine Software zugelassen worden sei, die mithilfe Künstlicher Intelligenz Röntgenbilder der Lunge auswerte. Er stellt sich die Frage, ob eine Künstliche Intelligenz Röntgenbilder ebenso gut oder sogar besser als Ärztinnen und Ärzte auswerten kann und welche Folgen dies für die Versorgung haben könnte.

2 Fragestellungen

2.1 Fragestellung 1: AI-unterstützte Befundung bei Verdacht auf Lungenkrebs

Die Ziele der vorliegenden Untersuchung sind

- die Nutzenbewertung einer Befundung von Röntgenbildern oder einer Computertomografie (CT) der Lunge durch Radiologinnen und Radiologen mit Unterstützung von künstlicher Intelligenz im Vergleich zu deren Befundung ausschließlich durch Radiologinnen und Radiologen
 - bei erwachsenen Patientinnen und Patienten mit Verdacht auf Lungenkrebs.
- die Bestimmung der Kosten (Interventionskosten) und die Bewertung der Kosteneffektivität sowie
- die Aufarbeitung ethischer, sozialer, rechtlicher und organisatorischer Aspekte sowie Umwelt- und Klimaaspekte, die mit der medizinischen Intervention verbunden sind.
- Ergänzende Betrachtung einer Befundung von Röntgenbildern oder einer Computertomografie der Lunge allein durch künstliche Intelligenz im Vergleich zu deren Befundung ausschließlich durch Radiologinnen und Radiologen

2.2 Fragestellung 2: AI-unterstützte Befundung in der Lungenkrebsfrüherkennung

Ziele der vorliegenden Untersuchung sind

- die Nutzenbewertung einer Befundung einer Niedrigdosis- Computertomografie (Low-Dose-CT) der Lunge durch Radiologinnen und Radiologen mit Unterstützung von künstlicher Intelligenz im Vergleich zur Befundung ausschließlich durch Radiologinnen und Radiologen
 - bei Personen, die an einer Lungenkrebsfrüherkennung teilnehmen und gemäß den Kriterien der deutschen Lungenkrebs-Früherkennungs-Verordnung (LuKrFrühErkV) für eine Teilnahme berechtigt wären.
- die Bestimmung der Kosten (Interventionskosten) und die Bewertung der Kosteneffektivität sowie
- die Aufarbeitung ethischer, sozialer, rechtlicher und organisatorischer Aspekte sowie Umwelt- und Klimaaspekte, die mit der medizinischen Intervention verbunden sind
- Ergänzende Betrachtung einer Befundung von einer Niedrigdosis-Computertomografie (Low-Dose-CT) der Lunge allein durch künstliche Intelligenz im Vergleich zu deren Befundung ausschließlich durch Radiologinnen und Radiologen

Methoden

2.3 Methoden Nutzenbewertung

Die Zielpopulation der Nutzenbewertung bildeten Patientinnen und Patienten mit Verdacht auf Lungenkrebs (Fragestellung 1), sowie Personen, die an einer Lungenkrebsfrüherkennung gemäß den Kriterien der LuKrFrühErkV teilnehmen (Fragestellung 2). Für beide Fragestellungen wurde sowohl nach Studien zur diagnostisch-therapeutischen Kette als auch nach Studien zur diagnostischen Güte gesucht.

Studien zur diagnostisch-therapeutischen Kette

Die Prüfintervention bildete die diagnostische Strategie unter Anwendung bildgebender Verfahren zur Befundung der Lunge durch Radiologinnen und Radiologen mit Unterstützung von AI. Als Vergleichsintervention galt die diagnostische Strategie unter Anwendung bildgebender Verfahren zur Befundung der Lunge ausschließlich durch Radiologinnen oder Radiologen.

Folgende patientenrelevante Endpunkte wurden als relevant für die diagnostisch-therapeutische Kette betrachtet:

- Mortalität (Gesamt mortalität, lungenkrebspezifische Mortalität)
- Morbidität, wie
 - krebsbedingte Symptomatik
- gesundheitsbezogene Lebensqualität
- Nebenwirkungen, wie
 - Schäden, die sich aus der Befundung oder aus nachfolgenden diagnostischen Untersuchungen (zum Beispiel aus invasiven Prozeduren wie Biopsien) ergeben, einschließlich der Konsequenzen aus falschen Befunden (falsch-positiv und falsch-negativ).
 - Zusätzliche bei Fragestellung 2: Überdiagnosen im Rahmen der Lungenkrebsfrüherkennung

Zur Bewertung der diagnostisch-therapeutischen Kette wurde nach randomisierten kontrollierten Studien (RCTs) sowie prospektiv vergleichenden Kohortenstudien gesucht. Hinsichtlich der Studiendauer bestand keine Einschränkung.

Parallel zur Erstellung des ThemenCheck-Berichtsprotokolls erfolgte eine Suche nach systematischen Übersichten in den Datenbanken MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews) und HTA Database sowie auf den Websites des National

Institute for Health and Care Excellence (NICE) und der Agency for Healthcare Research and Quality (AHRQ).

Es wurde geprüft, ob mindestens 1 hochwertige und aktuelle systematische Übersicht infrage kommt, deren Informationsbeschaffung als Grundlage für die Bewertung der diagnostischen-therapeutischen Kette verwendet werden kann (im Folgenden: Basis-SÜ).

Eine systematische Literaturrecherche nach Studien wurde in den Datenbanken MEDLINE, Embase und Cochrane Central Register of Controlled Trials durchgeführt. Darüber hinaus wurden folgende Informationsquellen und Suchtechniken berücksichtigt: Studienregister und die Sichtung von Referenzlisten identifizierter systematischer Übersichten.

Die Selektion relevanter Studien erfolgte von 2 Personen unabhängig voneinander. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst.

Studien zur diagnostischen Güte

Es konnte keine RCT zur diagnostisch-therapeutischen Behandlungskette identifiziert werden. Daher wurde die Nutzenbewertung basierend auf Studien zur diagnostischen Güte durchgeführt.

Indextest 1 bildete die Befundung der Lunge durch Radiologinnen und Radiologen mit Unterstützung von AI. Als Indextest 2 wurde die Befundung der Lunge ausschließlich durch Radiologinnen oder Radiologen festgelegt. Sofern relevante Studien zusätzlich auch die alleinigen Ergebnisse der AI als weiteren Indextest berichteten, wurden diese Ergebnisse ergänzend betrachtet. Als Referenztest für Studien zur diagnostischen Güte wurde die Biopsie, die klinische Nachbeobachtung oder Nachuntersuchung festgelegt.

Als Zielgröße wurden personenbezogene Vierfeldertafel-Daten zur Berechnung der diagnostischen Güte (z. B. Sensitivität, Spezifität) von Indextest 1 und Indextest 2 verwendet. Lagen keine Vierfeldertafel-Daten vor, wurden auch die Angaben der Sensitivität und Spezifität aus einer multicase-multireader Analyse [12] herangezogen, sofern die Anzahl der Referenztestpositiven und Referenztestnegativen sowie Angaben zu NPV und PPV aus dieser Analyse berichtet wurden.

Zur Bewertung der diagnostischen Güte wurden intraindividuelle Vergleiche in Form von prospektiv vergleichenden Kohortenstudien in die Nutzenbewertung eingeschlossen. Hinsichtlich der Studiendauer bestand keine Einschränkung.

Die Informationsbeschaffung von systematischen Übersichten und Studien zur diagnostischen Güte erfolgte unter derselben Methodik und gemeinsam mit den Studien zur diagnostisch-therapeutischen Kette.

Die Datenextraktion erfolgte in standardisierte Tabellen. Zur Einschätzung der qualitativen Ergebnissicherheit bei Studien zur diagnostischen Güte wurde das Verzerrungspotenzial der eingeschlossenen Studie zur Bewertung der diagnostischen Güte auf Basis des QUADAS-2-Instruments [13] bewertet. Das Verzerrungspotenzial von Primärstudien zur diagnostischen Güte wurde jeweils als niedrig oder hoch eingestuft. Zudem erfolgte eine Einschätzung zur Übertragbarkeit der Ergebnisse (geringe oder starke Bedenken). Die Ergebnisse der einzelnen Studien wurden hinsichtlich der Fragestellung beschrieben.

Aufgrund klinischer Heterogenität zwischen den eingeschlossenen Studien waren Metaanalysen nicht möglich. Ebenso waren Sensitivitätsanalysen und die Betrachtung von Effektmodifikatoren nicht möglich.

Die ermittelten Ergebnisse zu Sensitivität und Spezifität der verschiedenen Indextests wurden anschließend miteinander verglichen.

2.4 Methoden gesundheitsökonomische Bewertung

Zur Bestimmung der Interventionskosten wurden die anfallenden inkrementellen Kosten durch den Einsatz einer AI als Unterstützung bei der Befundung von Röntgenbildern oder CT-Aufnahmen über die Ressourcenverbräuche aus verschiedenen Perspektiven betrachtet. Zu diesem Zweck wurden sowohl Vergütungskataloge im Rahmen der gesetzlichen wie auch privatärztlichen Abrechnung nach regulierten oder verhandelten Preisen durchsucht. Für die Produktkosten der AI-Software wurden zusätzlich systematische Übersichtsarbeiten sowie die Treffer aus der Recherche nach gesundheitsökonomischen Evaluationen gesichtet.

Für die systematische Übersicht gesundheitsökonomischer Evaluationen wurde eine fokussierte Informationsbeschaffung in den bibliografischen Datenbanken MEDLINE, Embase und HTA Database durchgeführt. Zusätzlich wurden die Referenzlisten identifizierter systematischer Übersichten geprüft. Die Selektion relevanter Publikationen erfolgte durch 1 Person und wurde durch eine 2. Person qualitätsgesichert. Dabei wurden nur vergleichende Studien mit einer Aussage zur Kosteneffektivität einbezogen. Hinsichtlich der untersuchten Population, Prüf- und Vergleichsintervention bzw. Index- und Vergleichstest galten die gleichen Kriterien wie für die Nutzenbewertung.

Die für die Bewertung notwendigen Informationen wurden in standardisierte Tabellen extrahiert. Die Bewertung der Berichtsqualität orientierte sich an den Kriterien des Consolidated Health Economic Evaluation Reporting Standards for Interventions That Use Artificial Intelligence (CHEERS-AI) [14]. Der Bewertung der Übertragbarkeit lagen orientierend die Kriterien des European network for Health Technology Assessment (EUnetHTA) HTA adaptation toolkit [15] zugrunde.

2.5 Methoden ethische Aspekte

Die Informationsbeschaffung zu ethischen Aspekten erfolgte durch orientierende Recherchen nach einschlägigen Gesetzen, Verordnungen oder Richtlinien, nach interessenabhängigen Informationsquellen wie Internetseiten von Interessenvertretern sowie in den Datenbanken ETHMED, MEDLINE, LIVIVO (System der deutschen Zentralbibliothek für Medizin (ZB Med) in Köln) und BELIT. Außerdem wurden die in die Nutzenbewertung und die in die gesundheitsökonomische Bewertung eingeschlossenen Studien auf mögliche ethische Argumente geprüft. Schließlich wurde noch das Protokoll zur Dokumentation der Diskussion mit den befragten Betroffenen herangezogen.

Informationen aus allen Quellen der orientierenden Recherchen bzw. aus allen weiteren Dokumenten wurden von 1 Person auf Aussagen zu ethischen Argumenten und Aspekten der zu untersuchenden Technologien gesichtet und durch eine 2. Person qualitätsgesichert.

Die Suche nach Publikationen für die ethischen Aspekte orientierte sich nach Beauchamp und Childress an den medizinethischen Prinzipien Wohltun („beneficence“), Nichtschaden („nonmaleficence“), Achtung der Autonomie und Gerechtigkeit [16] in Bezug auf die vorliegende Fragestellung.

Darüber hinaus wurden bei der Recherche die ethischen Aspekte aus dem vereinfachten Fragenkatalog von Hofmann [17] und von Marckmann [18] berücksichtigt. Das Rahmengerüst nach Marckmann wurde für den Anwendungsbereich der vergleichenden Betrachtung zweier Fragestellungen erweitert.

Vor der vertieften Betrachtung der ethischen Aspekte zunächst 2 Anmerkungen:

1) Einsatz der AI

In der vorliegenden Fragestellung geht es um AI-Systeme zur Unterstützung der Auswertung von Röntgen- oder CT-Bildern; die finale Diagnoseentscheidung liegt bei Radiologinnen und Radiologen. Die folgenden Ausführungen beinhalten nicht den vollständigen Ersatz der ärztlichen Entscheidung durch AI.

2) Einheitliche Abhandlung der zwei Fragestellungen

Die ethische Bewertung erfolgt für die beiden Fragestellungen – bezogen auf Personen mit Verdacht auf Lungenkrebs einerseits und Personen, die eine Früherkennung in Anspruch nehmen andererseits – einheitlich, da die zu prüfenden Aspekte jeweils beide Zielgruppen betreffen.

2.6 Methoden soziale, rechtliche und organisatorische Aspekte sowie Umwelt- und Klimaaspekte

Für die Aufarbeitung sozialer, rechtlicher und organisatorischer Aspekte sowie von Klima- und Umweltaspekten wurden orientierende Recherchen in folgenden Informationsquellen durchgeführt: MEDLINE; LIVIVO (System der ZB Med in Köln); Google Scholar; Gesetze, Verordnungen oder Richtlinien (zum Beispiel Beck-Online oder juris); interessenabhängige Informationsquellen (z. B. Websites von Interessenvertretungen); nationale und regionale Register.

Zudem wurden die folgenden Dokumente auf mögliche soziale, rechtliche und / oder organisatorische sowie auf Klima- und Umwelt bezogene Argumente und Aspekte geprüft: in die Nutzenbewertung eingeschlossene systematische Übersichten, in die gesundheits-ökonomische Bewertung eingeschlossene Studien und das Protokoll zur Dokumentation des Interviews mit Betroffenen.

Informationen aus allen Quellen der orientierenden Recherchen bzw. aus allen weiteren Dokumenten wurden von 1 Person auf Aussagen zu sozialen, rechtlichen und organisatorischen, sowie auf Klima- und Umwelt bezogene Argumente und Aspekte der zu untersuchenden Technologie gesichtet. Die Ergebnisse wurden durch eine 2. Person qualitätsgesichert.

Zusätzlich wurden „reflective thoughts“ als Informationsquelle genutzt, also das reflektierte und auf das Wissen der Berichtautorinnen und -autoren bezogene Nachdenken über mögliche soziale, rechtliche oder organisatorische, sowie auf Klima- und Umwelt bezogene Argumente und Aspekte [19].

Die Informationsaufbereitung zu sozialen Aspekten orientiert sich an dem von Mozygemba 2016 [20] vorgeschlagenen umfassenden konzeptionellen Rahmen.

Die Informationsaufbereitung zu rechtlichen Aspekten orientiert sich an dem von Brönneke 2016 [21] entwickelten Leitfaden zur Identifikation von rechtlichen Aspekten.

Die Informationsaufbereitung zu organisatorischen Aspekten orientiert sich an dem von Perleth 2014 [22] vorgeschlagenen Raster zur Einschätzung der organisatorischen Folgen von Untersuchungs- bzw. Behandlungsmethoden.

Die Informationsaufbereitung zu Umwelt- und Klimaaspekten erfolgt im Rahmen einer deskriptiven Darstellung unter besonderer Berücksichtigung der Emission von Treibhausgasen, der Freisetzung toxischer Substanzen bzw. der Strahlenexposition, des Abfallmanagements sowie weiteren Aspekten der neuen Technologie.

Alle für die Informationsaufbereitung notwendigen Argumente und Aspekte wurden in Tabellen extrahiert.

2.7 Austausch mit Betroffenen

Während der Erstellung des HTA-Berichtsprotokolls wurden Betroffene zu verschiedenen Aspekten zum Thema „Kann Künstliche Intelligenz die Auswertung von Röntgen- oder CT-Bildern der Lunge unterstützen?“ befragt. Die Befragung der Betroffenen erfolgte innerhalb persönlicher Einzel-Gespräche entlang eines strukturierten Gesprächsleitfadens. Der Gesprächsleitfaden wurde basierend auf einer vorbereitenden Recherche nach und Sichtung von qualitativen und quantitativen Studien zum Thema erstellt. Es wurden Erfahrungen von Betroffenen mit der Röntgen- oder CT-Untersuchung der Lunge erhoben, sowie deren Einschätzungen und Präferenzen zu erwartetem Nutzen, Schaden und Implikationen der Umsetzung einer AI-Unterstützung bei der Auswertung von Röntgen- oder CT-Bildern. Explorativ wurden auch Einschätzungen und Präferenzen der Betroffenen hinsichtlich eines alleinigen Einsatzes einer AI zu Auswertung von Röntgen- oder CT-Bildern erfragt. Als Betroffene wurden Versicherte der GKV (gesetzliche Krankenversicherung) verstanden, die aus verschiedenen Gründen eine Röntgen-Untersuchung und / oder eine CT der Lunge bereits in Anspruch genommen oder eine Überweisung für eine Röntgen-Untersuchung oder ein CT der Lunge erhalten hatten (z. B. aktive oder vorherige Raucher aus unterschiedlichen Gründen, andere Anlässe für das Röntgen oder CT der Lunge, Vorliegen von Symptomen, die eine Lungenerkrankung vermuten ließen wie chronischer Husten oder Atemnot).

Die befragten Personen wurden frei rekrutiert. Im Sinne eines zielgerichteten Samplings wurde eine Mischung von Personen unterschiedlichen Geschlechts und Alters für die Teilnahme an einem Betroffenenengespräch gewonnen. Zu Beginn des Gesprächs wurden die Betroffenen darüber informiert, dass dieses aufgezeichnet wird. Im Nachgang verschriftlichte 1 Person das Gespräch im Wortlaut. Die Aufzeichnungen der von den Betroffenen berichteten Erfahrungen, Einschätzungen und Präferenzen wurden für die Erstellung des Berichts inhaltsanalytisch entlang der im Gesprächsleitfaden enthaltenen Themen (Kriterien) ausgewertet und zusammengefasst und für die Ausführungen zu ethischen, sozialen, rechtlichen und organisatorischen Aspekten herangezogen.

3 Ergebnisse: Nutzenbewertung

3.1 Ergebnisse der umfassenden Informationsbeschaffung

Es wurde keine systematische Übersicht als Basis-SÜ zum Zweck der Identifizierung von Primärstudien berücksichtigt.

Die Informationsbeschaffung ergab keine für die Fragestellung 1 zur AI-unterstützten Befundung bei Verdacht auf Lungenkrebs oder für die Fragestellung 2 zur AI-unterstützten Befundung in der Lungenkrebsfrüherkennung relevante randomisierte kontrollierte Studie zur diagnostisch-therapeutischen Behandlungskette. Es wurden keine geplanten sowie laufenden randomisierten kontrollierten Studien identifiziert. Hinsichtlich der Studien zur diagnostischen Güte ergab die Informationsbeschaffung keine relevante Studie für die Fragestellung 1. Für die Fragestellung 2 zwei relevante Studien mit verwertbaren Daten identifiziert. Die letzte Suche fand am 29.08.2025 statt.

Zusätzlich wurden für Fragestellung 2 drei abgeschlossene Studien zur diagnostischen Güte ohne berichtete Ergebnisse identifiziert.

3.2 Charakteristika der in die Bewertung eingeschlossenen Studien (Fragestellung 2)

In der südkoreanischen Studie Lee 2025 [23] wurden 366 aktuelle und ehemalige Raucherinnen und Raucher eingeschlossen, für die Low-Dose-CT-Bilder aus einem Studienzentrum des Korean Lung Cancer Screening Project (K-LUCAS) aus dem Jahr 2017 vorlagen. Personen, die an K-LUCAS teilnahmen, waren zwischen 55 und 74 Jahren alt und wiesen einen Zigarettenkonsum in Höhe von mindestens 30 Packungsjahren auf (1 Packungsjahr = Rauchen von ca. 20 Zigaretten täglich über 1 Jahr). Das Alter der eingeschlossenen Personen in Lee 2025 betrug im Mittel 64 Jahre (SD 5 Jahre) und die Anzahl der Packungsjahre betrug im Mittel 43 (SD 14). Mit einem Anteil von 97,8 % umfasste die Studie Lee 2025 ganz überwiegend Männer. In der Studie wurde die diagnostische Güte zweier Indextests gegenüber einem Referenzstandard verglichen. Indextest 1 war die AI-unterstützte Befundung durch Radiologinnen und Radiologen. Hierbei wurde der befundenden Radiologin oder dem befundenden Radiologen ein CT-Bild mit vorab durch die AI markierten Lungenrundherden sowie die Lung-RADS (Lung Imaging Reporting and Data System) Bewertung der AI zur Verfügung gestellt. Kriterien für die AI-Markierung eines Lungenrundherdes als auffällig werden nicht angegeben. Indextest 2 war die Befundung ausschließlich durch Radiologin und Radiologen ohne weitere AI-Unterstützung. Es gab 5 befundende Radiologinnen und Radiologen. Die CT-Aufnahmen wurden ebenso wie die Reihenfolge der Indextests randomisiert den Beurteilenden zugeteilt und der Abstand zwischen beiden Indextests betrug 1 Monat. Als positives Testergebnis galt für beide Indextests die Einschätzung eines Befundes nach Lung-RADS Kategorie 3 oder höher. Als Referenzstandard wurden die Angaben aus der elektronischen Patientenakte herangezogen,

ob ein Lungenkrebs innerhalb der letzten 5 Jahre nach der LDCT Untersuchung diagnostiziert wurde.

In die US-amerikanische Studie Lo 2018 [24] wurden 324 Personen eingeschlossen, deren Daten zufällig aus den Datenbanken des National Lung Screening Trials (NLST) (84 % der Personen) oder zweier Universitätskliniken ausgewählt wurden. Eingeschlossen wurden aktuelle und ehemalige asymptomatische Raucherinnen und Raucher im Alter zwischen 55 und 77 Jahren, für die low-dose CT-Bilder vorlagen. Symptomatische Patientinnen und Patienten mit Komorbiditäten wurden ausgeschlossen. Mit dem Ziel einer Prävalenzanreicherung wurden die untersuchten Personen so ausgewählt, dass bei einem Drittel auffällige Lungenrundherde vorlagen. Sofern ein primärer Lungenkrebs vorlag, sollte dieser durch Biopsie bestätigt worden sein. Bewertet wurde in der Studie die diagnostische Güte zweier Indextests im Vergleich zum Referenzstandard. Indextest 1 war hierbei die AI-unterstützte Befundung jeweils durch 12 Radiologinnen oder Radiologen unabhängig voneinander, wobei auf den CT-Bildern durch die AI sowohl die Darstellung von Gefäßen unterdrückt wurde als auch auffällige Lungenrundherde markiert wurden. Hierbei sollten Lungenrundherde mit einem Durchmesser von 5 mm oder mehr durch die AI markiert werden. Indextest 2 stellte die Befundung ohne AI jeweils durch dieselben 12 Radiologinnen und Radiologen dar. Der Abstand zwischen der Durchführung beider Indextests betrug mindestens 37 Tage. Referenztest war der histopathologische Befund beziehungsweise die Nachbeobachtung. Die Nachbeobachtung bei Personen mit auffälliger CT betrug im Mittel 6,5 Jahre. Berichtet wurden die Ergebnisse einer gemeinsamen Auswertung der Vierfeldertafeln aller 12 Radiologinnen und Radiologen anhand eines multicase-multireader Models [12] auf Basis einer ANOVA, diese sind daher interpretierbar als durchschnittliche Sensitivitäten, Spezifitäten, NPV und PPV.

3.3 Übersicht der patientenrelevanten Endpunkte (Fragestellung 2)

Aus 2 Studien konnten Daten zur diagnostischen Güte extrahiert werden. Hierbei wurden die Sensitivität und die Spezifität als Maß der diagnostischen Güte bewertet.

Zusätzlich berichteten die 2 Studien zur diagnostischen Güte Ergebnisse zur Befundungsdauer, die dem ergänzenden Endpunkt Zeitersparnis zugeordnet werden können. Der ergänzende Endpunkt Zeitersparnis sollte bei Studien zur diagnostisch-therapeutischen Kette zusätzlich betrachtet werden. Da keine Studien mit diesem Studiendesign vorliegen, werden die Ergebnisse aus den Studien zur diagnostischen Güte ergänzend dargestellt.

3.4 Bewertung des Verzerrungspotenzials der Ergebnisse aus Studien zur diagnostischen Güte (Fragestellung 2)

Für beide ausgewerteten Studien zur diagnostischen Güte wurde ein hohes Verzerrungspotenzial (VZP) festgestellt (siehe Tabelle 14). In beiden Studien lag das hohe VZP

(u. a.) am Patientenfluss und zeitlichen Ablauf: hier fehlten Angaben zur Durchführung der Nachbeobachtung als Referenztest während der Studiendauer zu allen Personen. Zudem erfolgte die Nachbeobachtung in beiden Studien über 5 oder mehr Jahre. Das Verzerrungspotenzial dieser langen Nachbeobachtungsdauer beruht darauf, dass diese über die geschätzte Lead-time von 1 bis 3 Jahren (LIT) hinausgeht: Personen mit einem negativen Befund im Indextest können im Referenztest nach 3 oder mehr Jahren Lungenrundherde aufweisen, welche sich erst nach Durchführung des Indextest entwickelt haben können und sich bei der Abklärung als maligne erweisen. Solche Personen würden dann als falsch-negatives Testergebnis des Indextests bewertet, obwohl diese Personen bei der Durchführung des Referenztests zu einem früheren Zeitpunkt als richtig-negativ eingestuft worden wären [25]. Zudem zeigte bei Lo 2018 die Patientenselektion zum Beispiel aufgrund des Fall-Kontroll-Designs der Studie ein hohes VZP. Auch war bei Lo 2018 für Indextest 1 und 2 unklar, ob die befundenden Personen gegenüber dem Referenztest verblindet waren und ob die Kriterien für ein testpositives Ergebnis prospektiv festgelegt worden waren.

3.5 Bewertung der Übertragbarkeit der Ergebnisse aus Studien zur diagnostischen Güte (Fragestellung 2)

In Lee 2025 gibt es keine Angaben zu den Tumortypen und Krankheitsursachen. Da es sich um eine asiatische Studie handelt, sind Unterschiede in der Häufigkeit der verschiedenen Tumortypen im Vergleich zum europäischen Raum möglich. Ob ein Unterschied vorliegt und ob dieser einen Einfluss auf die Auswertung der CT-Bilder hat, ist unklar. Die Bedenken hinsichtlich der Übertragbarkeit des Indextests 1 in Lee 2025 waren stark, da keine Versionsnummer der verwendeten AI-Software berichtet wurde. In der Gesamtschau wird die Übertragbarkeit der Ergebnisse von Lee 2025 als gegeben angesehen. In Lo 2018 waren die Bedenken bezüglich der Übertragbarkeit der Ergebnisse zur diagnostischen Güte bei der Patientenselektion und den Indextests stark, weil die Studienpopulation aufgrund von Prävalenzanreicherung sich deutlich von der Zielpopulation unterschied, Patientinnen und Patienten mit Komorbiditäten ausgeschlossen wurden und die herangezogenen Kriterien für ein testpositives Ergebnis nicht den Inhalten einer radiologischen Befundung in Deutschland entsprachen (siehe Tabelle 15). Domänenübergreifend lagen bei Lo 2018 somit starke Bedenken hinsichtlich der Übertragbarkeit der Ergebnisse zur diagnostischen Güte vor.

3.6 Ergebnisse der Studien zur diagnostischen Güte (Fragestellung 2)

Die Ergebnisse der beiden Studien werden getrennt dargestellt (siehe auch Tabelle 16), da die Indextests unterschiedlich sind: In der Studie Lee 2025 wurde die diagnostische Güte einer Screeningrunde untersucht, wohingegen Lo 2018 die diagnostische Güte einer einmaligen Diagnostik berichtet. Das bedeutet, dass Personen mit einem auffälligen Lungenrundherd, die eine zeitnahe Abklärung mittels erneuter Bildgebung benötigen (Lung-RADS = 3), in Lee 2025 als indextestpositiv betrachtet wurden. In Lo 2018 wurden nur die Personen als

indextestpositiv betrachtet, bei denen ein Lungenrundherd markiert wurde. Dies führte jedoch nicht zu einer zeitnahen Abklärung mittels erneuter Bildgebung, sondern unmittelbar zur Biopsie. Fälle mit als benigne eingeschätztem Befund ($n = 15$), die in Lo 2018 nachbeobachtet wurden aber keine Biopsie erhielten, wurden trotz Studieneinschluss aus der Analyse ausgeschlossen. Bei gleicher Anzahl an referenztestpositiven und – negativen Personen würde eine höhere Anzahl an indextestpositiven Personen die Wahrscheinlichkeit für eine hohe Sensitivität erhöhen, wohingegen eine höhere Anzahl an indextestnegativen Personen die Wahrscheinlichkeit für eine hohe Spezifität fördern würde.

3.6.1 AI-unterstützte Befundung vs. menschliche Befundung (Indextest 1 vs. Indextest 2)

Die Studie Lee 2025 berichtete für die AI-unterstützte Befundung für eine Screeningrunde (Indextest 1) eine Sensitivität (%) [95 %-KI] von 60,0 [23,1; 88,2], wohingegen für die ausschließliche Befundung durch eine Radiologin oder einen Radiologen für eine Screeningrunde (Indextest 2) eine Sensitivität von 80,0 [37,6; 96,4] vorlag. Zudem lag in der Studie Lee 2025 die Spezifität (%) [95 %-KI] der AI-unterstützten Befundung (Indextest 1) bei 70,1 [65,2; 74,6] wohingegen die Spezifität der Befundung ausschließlich durch Radiologinnen und Radiologen (Indextest 2) bei 78,7 [74,2; 82,6] lag. Die Punktschätzung der Spezifität von Indextest 2 liegt um 8,6 Prozentpunkte höher als Indextest 1. Zu beachten ist, dass die Schätzung der Sensitivität auf lediglich 5 Personen mit positivem Referenztest beruht, sodass diese Schätzung nicht belastbar ist.

Demgegenüber berichtete Lo 2018 für die einmalige Diagnostik eine Sensitivität von 80,0 [72,4; 87,6] für die AI-unterstützte Befundung und eine Sensitivität von 64,7 [57,1; 72,3] für die Befundung ausschließlich durch Radiologinnen und Radiologen. In Lo 2018 lag die Spezifität bei 84,4 [80,5; 88,3] für die AI-unterstützte Befundung und eine Spezifität von 89,9 [86,0; 93,8] für die Befundung ausschließlich durch Radiologinnen und Radiologen. Die Punktschätzung der Spezifität von Indextest 2 liegt um 5,5 Prozentpunkte höher als Indextest 1. Dabei ist zu beachten, dass die Population eine künstlich angereicherte Prävalenz aufweist und Fälle mit Komorbiditäten ausgeschlossen waren. In der Analyse in Lo 2018 wurden 95 maligne Befunde (von 93 Personen) berücksichtigt.

Berechnet man für beide Studien die Vierfeldertafeln basierend mit 1,2 % Prävalenz von Lungenkrebs, welche in Deutschland für eine Screeningrunde angenommen werden kann [26], ergibt sich bei 1200 erkrankten Personen von 100 000 gescreenten Personen folgendes Bild:

Für Lo 2018

- mit AI-Unterstützung haben von insgesamt 16 373 Personen mit positivem Testergebnis lediglich 960 Personen Lungenkrebs (5,9 % positiv prädiktiver Wert [PPV]) und von insgesamt 83 627 Personen mit negativem Testergebnis haben 83 387 Personen keinen Lungenkrebs (99,7 % negativ prädiktiver Wert [NPV])

- ohne AI-Unterstützung haben von insgesamt 10 755 Personen mit positivem Testergebnis 776 Personen Lungenkrebs (7,2 % PPV) und von insgesamt 89 245 Personen mit negativem Testergebnis haben 88 821 Personen keinen Lungenkrebs (99,5 % NPV)

Für Lee 2025

- mit AI-Unterstützung haben von insgesamt 30 261 Personen mit positivem Testergebnis lediglich 720 Personen Lungenkrebs (2,4 % PPV) und von insgesamt 69 739 Personen mit negativem Testergebnis haben 69 259 Personen keinen Lungenkrebs (99,3 % NPV)
- ohne AI-Unterstützung haben von insgesamt 22 004 Personen mit positivem Testergebnis 960 Personen Lungenkrebs (4,4 % PPV) und von insgesamt 77 996 Personen mit negativem Testergebnis haben 77 756 Personen keinen Lungenkrebs (99,5 % NPV)

Auch bei der höheren Prävalenz von 3,1 % für Lungenkrebs, welche über 4 Jahren Screening für Deutschland angenommen werden kann [27] ändern sich der PPV und NPV kaum, wie das Beispiel mit den Ergebnissen von Lo 2018 verdeutlicht:

- mit AI-Unterstützung haben von insgesamt 17 596 Personen mit positivem Testergebnis 2480 Personen Lungenkrebs (14,1 % PPV) und von insgesamt 82 404 Personen mit negativem Testergebnis haben 81 784 Personen keinen Lungenkrebs (99,2 % NPV)
- ohne AI-Unterstützung haben von insgesamt 11 793 Personen mit positivem Testergebnis 2006 Personen Lungenkrebs (17,0 % PPV) und von insgesamt 88 207 Personen mit negativem Testergebnis haben 87 113 Personen keinen Lungenkrebs (98,8 % NPV)

Die qualitative Ergebnissicherheit für beide Szenarien einer AI-unterstützten Befundung (Screeningrunde und einmalige Diagnostik) ist gering. Es liegt jeweils nur 1 Studie vor. Es kann keine Aussage für Lee 2025 und für Lo 2018 zu einem vergleichbaren Nutzen oder Anhaltspunkt hinsichtlich eines höheren oder geringeren Nutzens einer AI-unterstützten Befundung im Vergleich zu einer alleinigen Befundung durch Radiologinnen und Radiologen getroffen werden. Ergebnisse zur Konkordanz von Indextest 1 und Indextest 2 werden in beiden Studien nicht berichtet.

3.6.2 Ergänzende Darstellung der alleinigen AI-Befundung

Die Studie Lee 2025 berichtete zum Einsatz von AI allein eine Sensitivität von 80,0 [37,6; 96,4] und eine Spezifität von 66,2 [61,2; 70,9] (Tabelle 17). Da es sich um eine ergänzende Darstellung dieses Indextests handelt, werden keine Aussagen zu Nutzen und Schaden getroffen. Die Studie Lo 2018 berichtete zwar Ergebnisse zum Einsatz von AI allein, diese werden jedoch nicht berücksichtigt, da die Berechnung der Vierfeldertafel aus folgenden Gründen nicht

erfolgen kann: es ist zum einem unklar, wie viele Personen bei der Schätzung der berichteten Sensitivität berücksichtigt wurden und zum anderen wird die Spezifität in Lo 2018 nicht benannt.

3.6.3 Ergebnisse zum ergänzenden dargestellten Endpunkt Zeitersparnis

Beide Studien zur diagnostischen Güte berichteten zusätzlich die Interpretationszeit pro CT-Aufnahme für die befundende Radiologin oder Radiologen mit und ohne AI-Unterstützung. In beiden Studien waren die Radiologinnen und Radiologen in Bezug auf die jeweiligen Indextests nicht verblindet. Die in beiden Studien gemessene Zeit umfasste die Sichtung der CT-Aufnahme und Speicherung der Fallbewertung in der Software. In Lee 2025 waren die Radiologinnen und Radiologen bei der AI-unterstützten Befundung um 3 Sekunden schneller als bei der Befundung ohne AI-Unterstützung. Der Unterschied war nicht statistisch signifikant. In Lo waren die Radiologinnen und Radiologen bei der AI-unterstützten Befundung statistisch signifikant um 34,3 Sekunden schneller als zur Befundung ohne AI-Unterstützung. Zur Übertragbarkeit ist anzumerken, dass der Indextest, welcher in Lo 2018 untersucht wurde, nicht einer radiologischen Befundung in Deutschland entspricht (siehe auch Abschnitt 3.5). Das Ergebnis von Lo 2018 ist somit nicht aussagekräftig für die Dauer der radiologischen Befundung in Deutschland. Auch umfasst die erfasste Zeitspanne in den Studien mit im Mittel 2 bis 3 Minuten (siehe auch Tabelle 18) nur einen Teil der eigentlichen Zeit, welche für eine radiologische Befundung der Lunge angenommen wird [28].

Da es sich um eine ergänzende Darstellung handelt, werden keine Aussagen zu Nutzen und Schaden getroffen.

3.7 Zusammenfassende Bewertung der Ergebnisse

Auf eine tabellarische Darstellung der Landkarte der Beleglage in Bezug auf die Ergebnisse zur diagnostischen Güte wird verzichtet. Es lagen für keine Fragestellung RCTs zur diagnostisch-therapeutischen Behandlungskette vor. Für die diagnostische Güte einer AI-unterstützten Befundung in der Lungenkrebsfrüherkennung lagen 2 Studien mit geringer qualitativer Ergebnissicherheit vor, für die eine gemeinsame Betrachtung aufgrund klinischer Heterogenität nicht sinnvoll erschien.

Insgesamt kann aufgrund der verfügbaren Daten keine Aussage bezüglich eines Nutzens oder Schadens oder vergleichbaren Nutzens einer AI-unterstützten Befundung im Vergleich zu einer Befundung ausschließlich durch Radiologinnen und Radiologen getroffen werden.

Bewertung des Umfangs unpublizierter Daten

Es ist nicht verpflichtend, nicht randomisierte Studien bei Studienstart in einem Studienregister zu registrieren. Daher lässt sich ein Publikationsbias nur sehr eingeschränkt bewerten. Es wurden in der systematischen Recherche 3 Registereinträge zu abgeschlossenen

Studien zur diagnostischen Güte ohne berichtete Ergebnisse identifiziert (siehe Tabelle 9). Bei den Studien, die 2024 beziehungsweise 2025 abgeschlossen wurden, stehen die Publikationen womöglich aus. Die dritte Studie wurde 2019 abgeschlossen, bisher sind keine Ergebnisse zu dieser Studie veröffentlicht worden. Die 3 Registereinträge untersuchten 3 weitere unterschiedliche AI-Softwares und beschrieben Index- und Referenztest unzureichend. Somit ist unklar, ob diese 3 Studien mit einer der beiden berücksichtigten Studien vergleichbar wären. Unklar ist auch, ob die beiden abgeschlossenen Studien Konsequenzen auf die Nutzensaussage hätten. Da zudem keine hinreichenden Daten zur Verfügung standen, die eine Aussage zum Nutzen oder Schaden der Intervention ermöglicht hätten, ergibt sich aus einer möglichen Verzerrung durch unpublizierte Studienergebnisse keine Herabstufung einer bestehenden Nutzensaussage.

Nutzen-Schaden-Abwägung

Für Patientinnen und Patienten mit Verdacht auf Lungenkrebs wurden keine Studien identifiziert. Für Fragestellung 2 zur Lungenkrebsfrüherkennung lag Evidenz vor. Diese Evidenz war auf wenige Testgütestudien beschränkt, daher ließ sich kein vergleichbarer Nutzen oder Anhaltspunkt für einen Nutzen oder Schaden der AI-unterstützten radiologischen Befundung gegenüber einer allein radiologischen Befundung ableiten. Daher entfällt die Grundlage für eine Nutzen-Schaden-Abwägung.

4 Ergebnisse: Gesundheitsökonomische Bewertung

4.1 Interventionskosten

Der vorliegende Bericht vergleicht die Befundung von CT oder Röntgenbildern der Lunge durch Radiologinnen und Radiologen mit Unterstützung von AI mit einer Befundung durch Radiologinnen und Radiologen ohne Unterstützung durch AI. Dementsprechend fokussiert auch der Abschnitt zu den Interventionskosten auf das Kosteninkrement durch den zusätzlichen Einsatz von AI.

Derzeit existieren keine regulierten oder verhandelten Preise für die Unterstützung durch AI bei der Befundung von CT oder Röntgenbildern der Lunge bei Verdacht auf Lungenkarzinom bzw. bei einer Low-Dose-CT im Rahmen der Lungenkrebsfrüherkennung. Im Einheitlichen Bewertungsmaßstab (EBM) existieren auch keine weiteren speziellen Abrechnungsziffern für durch AI erbrachte Leistungen (Stand: 4. Quartal 2025) [29].

Demnach fallen aktuell für die gesetzliche Krankenversicherung durch den Einsatz von AI-Software zur Unterstützung der Befundung gemäß der Fragestellung des vorliegenden Berichtes keine zusätzlichen Kosten an.

Bei einer privatärztlichen Befundung besteht grundsätzlich die Möglichkeit, den Einsatz einer AI-Software analog z. B. über die Gebührenordnung für Ärzte (GOÄ)-Ziffer 5733 (Zuschlag für computergesteuerte Analyse [z. B. Kinetik, 3D-Rekonstruktion]) in Rechnung zu stellen. Die daraus resultierenden zusätzlichen Kosten in Höhe von 46,63 € (Einfachsatz) wären in diesem Fall von der privat versicherten Person zunächst selbst zu tragen und ggf. von der privaten Krankenversicherung zu erstatten. Eine entsprechende Abrechnung dieser Ziffer gegenüber einer gesetzlich versicherten Person im Sinne einer individuellen Gesundheitsleistung ist hingegen nicht möglich, wenn die durch die AI erbrachte Leistung als Bestandteil der vertragsärztlichen Versorgung angenommen werden kann [30,31].

Aufseiten der Leistungserbringenden lassen sich hingegen orientierend Kosten für die AI-Unterstützung bestimmen. Herstellerangaben, die im Auftrag des britischen National Institute for Health and Care Excellence (NICE) angefragt wurden, zeigen die Kosten je Befundung von verschiedenen AI-basierten Produkten zur Unterstützung der Befundung von CT- und Röntgenaufnahmen der Lunge [32-35]. Darüber hinaus geben sie Einblick in die verschiedenen Preisstrukturen der Hersteller:

Für die AI-Unterstützung bei der Befundung von Röntgenaufnahmen konnten in Abhängigkeit des konkreten Produkts über einen 5-Jahres-Zeitraum kaufkraftbereinigt in € umgerechnet [36] und auf das Jahr 2024 inflationiert [37] durchschnittliche Kosten von 0,63 € bis 3,12 € [33] je Befundung ermittelt werden (siehe Abschnitt A4.1, Tabelle 21). Es ist zu beachten, dass diese Kosten auf einem beispielhaften Jahresvolumen von 25 000 Aufnahmen basieren und

sich überwiegend auf die Kosten pro National Health Service (NHS) Trust¹ beziehen. Sie bilden daher lediglich eine orientierende Größe, in der die verschiedenen Kostenkomponenten (Installation, Integration, Schulungen, Lizenzgebühren, Wartung, Support) auf die Anzahl der Aufnahmen bzw. Befundungen umgelegt wurden, um trotz der variablen Preisstrukturen eine Vergleichbarkeit zu erreichen.

Für die AI-Unterstützung bei der Befundung von CT-Bildern ergeben sich anhand weiterer Angaben des NICE durchschnittlich zusätzliche Kosten von umgerechnet 2,52 € bis 9,45 € [32,34] pro Befundung (siehe Abschnitt A4.1, Tabelle 21). An dieser Stelle sind jedoch die zugrunde liegenden Annahmen, aus denen die Kosten pro Befundung ermittelt wurden, nur teilweise beschrieben. Zumindest für die untere Kostengrenze ist bekannt, dass keine Kosten für Implementierung und Wartung enthalten sind.

Darüber hinaus entstehen neben den dargestellten Kosten für die konkreten Produkte weitere Ressourcenverbräuche (beispielsweise die Zeit, in der das Personal Schulungen erhält oder Opportunitätskosten durch Umsatzausfälle, wenn während der Implementierung kein regulärer Betrieb möglich sein sollte). Des Weiteren könnte auch für die Aufklärung über den zusätzlichen Einsatz der AI ein erhöhter Zeitbedarf anfallen, der sowohl seitens der Leistungserbringenden als auch für die untersuchte Person zu zusätzlichen Kosten führt.

Insgesamt entstehen einerseits den Leistungserbringenden zunächst zusätzliche Kosten für die Implementierung bzw. die laufenden Kosten bei der Nutzung der AI, die vermutlich überwiegend in einem niedrigen 1-stelligen Bereich pro Befundung liegen. Diese Kosten sind jedoch aufgrund der vielfältigen Kostenstrukturen nicht einheitlich zu beziffern [32-34,38]. Andererseits ist unklar, inwieweit eine Zeitersparnis im Gesamtprozess erreicht werden kann, um Ressourcen (z. B. Personalkosten) einzusparen. Im Rahmen der Lungenkrebsfrüherkennung ist die Nutzung einer Software zur computerassistierten Detektion durch die Befunderin bzw. den Befunder erst nach einer Befundung ohne ein entsprechendes Hilfsmittel vorgesehen [5], sodass in diesem Setting keine Zeitersparnis im Vergleich zu einer Befundung ohne AI-Unterstützung zu erwarten ist. Ob ein kostendeckender Einsatz von AI-Unterstützung bei der Befundung parallel zur Radiologin und zum Radiologen in einem anderen Setting (Befundung bei Verdacht auf Lungenkrebs) möglich ist, ist hingegen abhängig von den konkreten Gegebenheiten und ob tatsächlich im Gesamtprozess (inkl. Aufklärung) eine Zeitersparnis erzielt werden kann [39].

Aktuell gibt es im EBM (und auch in der GOÄ) keine spezifische Abrechnungsmöglichkeiten für die AI-Unterstützung bei der Befundung. Somit sind für die GKV sowie für gesetzlich

¹ Ein NHS trust kann mehrere Krankenhäuser bzw. Einrichtungen allgemein betreiben.

Versicherte in der Regel derzeit keine zusätzlichen direkten Kosten durch die AI-Unterstützung zu erwarten.

4.2 Systematische Übersicht gesundheitsökonomischer Evaluationen

Nachfolgend werden die Ergebnisse zur Kosteneffektivität zusammenfassend dargestellt. Für detailliertere Informationen zur Informationsbeschaffung sowie zusätzliche detaillierte Extraktions- und Bewertungstabellen sei auf Abschnitt A4.2 verwiesen.

4.2.1 Ergebnisse der Informationsbeschaffung

Im Rahmen der fokussierten Informationsbeschaffung konnte 1 gesundheitsökonomische Evaluation für die diagnostisch-therapeutische Kette von Geppert 2025 [34] identifiziert werden. Dabei handelt es sich um einen Extended Research Article zu einem vom NICE beauftragten Diagnostics Assessment Report [40]. Ergänzend wurde das dazugehörige Studienprotokoll [41] sowie die seitens des NICE publizierten Dokumente zum zu untersuchenden Umfang („Final scope“) [35] und zur zusammenfassenden Empfehlung („Diagnostics guidance“) [42] herangezogen.

Gegenstand der identifizierten Bewertung [34] war (u. a.) die Bestimmung der Kosteneffektivität für den Einsatz von AI-Software als Unterstützung bei der Identifikation und Befundung von Lungenrundherden im Vergleich zur Identifikation und Befundung ohne AI-Unterstützung in CT-Aufnahmen des Thorax, die

- aufgrund von Symptomen eines vermuteten Lungenkarzinoms (Verdachtspopulation),
- im Rahmen eines Lungenkrebscreenings (Screeningpopulation),
- aufgrund anderer Indikationen (inzidentelle Population) oder
- als Maßnahme zur Kontrolle bereits zuvor identifizierter Lungenrundherde (Kontroll-Population)

angefertigt wurden.

Nachfolgend werden – in Übereinstimmung mit den Fragestellungen des vorliegenden Berichtes – die Ergebnisse für die Verdachts- und die Screeningpopulation dargestellt und diskutiert.

4.2.2 Charakteristika der in die Bewertung eingeschlossenen Studien

Da im Rahmen einer systematischen Informationsbeschaffung in Geppert 2025 [34] keine geeigneten Publikationen zur Kosteneffektivität der AI-Unterstützung identifiziert werden konnten, erfolgte eine de-novo- Modellierung in einem mehrstufigen Verfahren. Hinsichtlich der in der Modellierung verwendeten Input-Parameter zur diagnostischen Güte sei

angemerkt, dass diese nicht mittels eines adäquaten Referenztests ermittelt wurden und daher auch nicht im Rahmen der Nutzenbewertung berichtet werden.

Preliminary model

In einem vorläufigen Modell („preliminary model“) wurde im Rahmen einer Entscheidungsbaumanalyse das Kosteneffektivitätsverhältnis (incremental cost-effectiveness ratio [ICER]) pro korrekt identifizierter Person mit handlungsbedürftigem Lungenrundherd bestimmt. Als handlungsbedürftig galten dabei Lungenrundherde mit einer Größe ≥ 5 mm bei Abwesenheit von starken Anzeichen, die auf eine gutartige Struktur hinweisen sowie weitere Faktoren (beispielsweise Art des Knotens, Lage oder Morphologie). Da Prävalenzangaben zum Vorkommen von handlungsbedürftigen Lungenrundherden sowie geeignete Daten zur diagnostischen Güte nur für die Screeningpopulation vorlagen, wurden die weiteren Populationen (und somit auch die Verdachtspopulation) im preliminary model nicht untersucht. Die Angaben zur diagnostischen Güte von AI-Systemen als Unterstützung bei der CT-Befundung im Vergleich zur Befundung ohne AI-Unterstützung basierten dabei auf der ebenfalls in der Nutzendomäne identifizierten und bewerteten Publikation von Lo 2018 [24] (siehe Abschnitt 3.2). Bei der betrachteten Software, deren Kosten folglich für das Modell veranschlagt wurden, handelt es sich um die von Lo 2018 [24] untersuchte ClearReadCT (Riverain technologies).

Insbesondere aufgrund des beschränkten Zeithorizonts auf die initiale (Screening-)Untersuchung, sieht die Autorengruppe weiteren Bedarf für ein komplexeres Modell, das auch den aus der Erstuntersuchung resultierenden weiteren Patientenpfad untersucht (full model).

Full model

Das erweiterte Modell („full model“) folgt nach der initialen Untersuchung den Empfehlungen der British Thoracic Society (BTS) hinsichtlich weiterer diagnostischer oder kontrollierender Maßnahmen der identifizierten Lungenrundherde. Das Modell vergleicht dabei – wie auch schon im preliminary model – eine Befundung mit und ohne AI-Unterstützung u. a. für die Screening- und die Verdachtspopulation. Als primärer Endpunkt wurden die zusätzlichen Kosten pro gewonnenes qualitätsadjustiertes Lebensjahr (Quality-adjusted Life Year [QALY]) über einen lebenslangen Zeitraum ermittelt. Zu den weiteren Modifikationen gehören:

- Spezifizierung des Modells zur initialen Befundung anhand von zusätzlichen Angaben zum Typ der Lungenrundherde (solide oder subsolide) sowie der subtypspezifischen Größe eines detektierten Lungenrundherdes und
- Erweiterung des Modells um
 - die weiteren Pfade zur Nachverfolgung bzw. Abklärung der Lungenrundherde gemäß den Vorgaben der BTS,

- die natürliche Entwicklung von bösartigen Rundherden mittels eines separaten Modells je nach Subtyp (solide oder subsolide) und
- die Behandlung von Patientinnen und Patienten mit Lungenkarzinom

Die zugrundeliegenden Studien zur diagnostischen Güte enthalten oftmals keine Angaben zur Sensitivität und Spezifität für Subtypen und Größenkategorien der Rundherde. Um die Ergebnisse dieser Studien für die durch die BTS-Vorgaben festgelegten Kategorien der Lungenrundherde nutzen zu können, wurden daher separate Simulationen für die Verteilungen der wahren Größe detektierter Lungenrundherde sowie der geschätzten Größe durch Radiologinnen und Radiologen mit oder ohne AI-Unterstützung durchgeführt. Auf diese Weise sollte der Einfluss systematischer und zufälliger Fehler bei der Größenbestimmung und der daraus resultierenden falschen oder richtigen Zuordnung in die BTS-Kategorien im Modell berücksichtigt werden.

Die Angaben zur diagnostischen Güte von AI-Systemen als Unterstützung bei der CT-Befundung im Vergleich zur Befundung ohne AI-Unterstützung basierten für die Screeningpopulation auf Angaben von Hsu 2021 [43]. Dabei handelt es sich um eine taiwanesishe Studie, in der u. a. 57 Low-Dose-CT-Aufnahmen von erfahrenen (über 5 Jahre Berufserfahrung) bzw. 3 unerfahrenen (1 bis 2 Jahre Berufserfahrung) Radiologinnen und Radiologen sowohl mit als auch ohne AI-Unterstützung befundet wurden.

Für die symptomatische Population mit Verdacht auf Lungenkarzinom wurden Angaben zur diagnostischen Güte von Kozuka 2020 [44] herangezogen. In dieser japanischen Studie wurden retrospektiv 120 zufällig ausgewählte CT-Aufnahmen von Patientinnen und Patienten mit Verdacht auf Lungenkarzinom Radiologinnen und Radiologen mit 1 bis 5 Jahren Berufserfahrung sowohl mit als auch ohne AI-Unterstützung befundet.

Der Referenztest in Hsu 2021 und in Kozuka 2020 war die Markierung der Lungenrundherde von Expertenpanels bestehend aus 2 bzw. 3 langjährig erfahrenen Radiologinnen oder Radiologen [43,44].

Die im Modell betrachtete AI-Software entspricht den in den genannten Publikationen [43,44] untersuchten ClearReadCT (Riverain technologies) für die Screeningpopulation und InferRead CT (Infervision) für die Verdachtspopulation. Die Kostenangaben im Modell wurden entsprechend für diese Software berechnet.

4.2.3 Ergebnisse gesundheitsökonomischer Evaluationen

Preliminary model

Im Rahmen der Basisfallanalyse zeigte sich, dass die Befundung mit AI-Unterstützung der Befundung ohne AI-Unterstützung im gesundheitsökonomischen Sinne überlegen (Kosten und

Effektivität) aus der Perspektive des NHS bzw. Personal Social Services (PSS) ist. Je 39 CT-Befundungen würde ca. 1 Person mit handlungsbedürftigem Lungenrundherd zusätzlich identifiziert und dabei 113,73 £ eingespart (umgerechnet und inflationiert: 143,26 € im Jahr 2024). Die deterministischen Sensitivitätsanalysen zeigten dabei, dass das ICER insbesondere von der veranschlagten Zeit mit bzw. ohne AI-Unterstützung für die Befundung beeinflusst wird.

Full model

Screeningpopulation (Fragestellung 2)

Für die Screeningpopulation zeigte sich im Modell eine Überlegenheit der AI-Unterstützung gegenüber der Befundung ohne AI-Unterstützung, die sowohl in allen Endpunkten als auch in den deterministischen und probabilistischen Sensitivitätsanalysen durchgängig Bestand hatte. So führte im Basisfall die AI-Unterstützung bei der Befundung zu ca. 1 zusätzlichen QALY pro 125 CTs und dabei zu einer Einsparung in Höhe von 8777,50 £ (umgerechnet und inflationiert: 11056,75 € im Jahr 2024), während gemäß der probabilistischen Sensitivitätsanalyse die Befundung mit AI-Unterstützung in 100 % der Szenarien überlegen war [34].

Verdachtspopulation (Fragestellung 1)

Primärer Endpunkt im Modell

Für den primären Endpunkt (zusätzliche Kosten pro gewonnenem QALY) ist die Befundung ohne AI-Unterstützung der Befundung mit AI-Unterstützung überlegen: Bei der Befundung von symptomatischen Personen ergeben sich aus der Modellierung im Basisfall, bei Einsatz einer AI, pro 50 CTs ein Nutzwertverlust von ca. 1 QALY und zusätzliche Kosten von 5054,00 £ (umgerechnet und inflationiert: 6366,37 € im Jahr 2024). Diese Überlegenheit hatte auch in den meisten deterministischen Sensitivitätsanalysen Bestand. Ausgenommen hiervon ist der Fall, dass kein Nutzwertverlust durch eine reduzierte Lebensqualität für falsch-positive Befunde bzw. für gutartige Lungenrundherde während der CT-Kontrollen nach dem initialen Befund angenommen wurden (ICER für dieses Szenario: 12 709 £/QALY [umgerechnet und inflationiert: 16 009,14 € im Jahr 2024]). In der probabilistischen Sensitivitätsanalyse war in nahezu allen Szenarien die Befundung ohne AI-Unterstützung der Befundung mit AI-Einsatz überlegen [34].

Sekundäre Endpunkte im Modell

Bei Betrachtung weiterer Endpunkt ergab sich im Basisfall ein ICER von 38 316 £ pro korrekt identifizierten und folglich behandelten Fall mit Lungenkarzinom (umgerechnet und inflationiert: 48 265,51 € im Jahr 2024). Hinsichtlich der korrekten Identifikation von Personen mit handlungsbedürftigen Lungenrundherden ist die Befundung mit AI-Unterstützung der Befundung ohne AI-Unterstützung überlegen. Es werden in Geppert 2025 keine Ergebnisse der Sensitivitätsanalysen für die sekundären Endpunkte berichtet [34].

Zusammenfassung und Diskussion der Ergebnisse durch die Autorengruppe

Zusammenfassend zeigte sich im Modell für den primären Endpunkt (zusätzliche Kosten pro gewonnenem QALY) folgendes Ergebnis:

- für die Verdachtspopulation: Befundung ohne AI-Einsatz ist überlegen
- für die Screeningpopulation: Befundung mit AI-Einsatz ist überlegen

Bei der Verdachtspopulation beschreibt die Autorengruppe, dass die AI-Unterstützung in einem Zugewinn an QALYs auf Grundlage einer erhöhten Rate an detektierten (und frühzeitig behandelten) Karzinomen resultiert. Demgegenüber steht aber ein höherer Verlust an QALYs sowie zusätzliche Kosten durch die höhere Anzahl an falsch-positiven Befunden bzw. gutartigen Rundherden, die nachbeobachtet werden [34]. Bei der Screeningpopulation führte der unterstützende Einsatz von AI bei der Befundung hingegen zu weniger falsch-positiven Befunden, bedingt durch die unterschiedliche Angabe zur diagnostischen Güte im Modell (höhere Sensitivität bei AI-Unterstützung in beiden Populationen; geringere Spezifität durch AI-Unterstützung in der Verdachtspopulation und höhere Spezifität durch AI-Unterstützung in der Screeningpopulation).

Die Autorengruppe [34] schränkt die Ergebnisse dahin gehend ein, dass zwar für die Screeningpopulation die Datengrundlage grundsätzlich besser als für die Verdachtspopulation sei, insgesamt jedoch ein Mangel an Daten bzw. deren Qualität vorliege. So bewertet die Autorengruppe mittels QUADAS-2 das Verzerrungspotenzial für die eingeflossenen Studien zur diagnostischen Güte (Lo 2018 [24] für das preliminary model, Kozuka et al. (2020) [44] bzw. Hsu 2021 [43] für das full model) hinsichtlich der Domänen Patientenselektion und Indextest durchgängig als hoch. Zudem wurde neben Patientenselektion und Indextests auch für die Domäne Referenzstandard (Expertenpanel) starke Bedenken hinsichtlich der Übertragbarkeit der Ergebnisse auf die Fragestellung angegeben. Zudem weise die Studie von Hsu 2021 [43], aus der die Angaben zur diagnostischen Güte für das full model bei der Screeningpopulation entnommen wurden, abweichend zu anderen Studien für den Einsatz der AI-Unterstützung (neben einer höheren Sensitivität) auch eine höhere Spezifität im Vergleich zur Befundung ohne AI-Unterstützung aus.

Aufgrund der unklaren Vergleichbarkeit der AI-Produkte untereinander sind die Modellergebnisse zudem auf die AI-Produkte, für die entsprechende Daten zur diagnostischen Güte und zu den Kosten vorlagen (ClearRead CT für die Screeningpopulation und InferRead CT für die Verdachtspopulation), begrenzt. Direkt vergleichende Evidenz zu verschiedenen AI-Produkten konnte nicht identifiziert werden. Eine allgemeingültige Aussage zur Kosteneffektivität von AI-Unterstützung bei der Befundung lasse sich somit nicht ableiten.

Des Weiteren diskutiert die Autorengruppe methodische und strukturelle Limitationen des Modells, die ggf. nicht im Rahmen von Sensitivitäts- oder Szenarioanalysen ausreichend adressiert werden konnten (z. B. die ausschließliche Betrachtung des jeweils größten Lungenrundherdes oder die Annahme, dass keine neuen Rundherde bzw. Karzinome während der Beobachtungsdauer auftreten und gutartige Rundherde nicht fälschlicherweise als Karzinom befundet werden).

Insgesamt werden die Ergebnisse aus dem Modell seitens der Autorengruppe als hochgradig unsicher angesehen und sollten entsprechend kritisch interpretiert werden.

4.2.4 Diskussion und Einordnung der gesundheitsökonomischen Ergebnisse für den vorliegenden Bericht

Mit der Publikation von Geppert 2025 [34], die auf einem Auftrag des NICE basiert, konnte eine Evaluation mit einer hohen Berichtsqualität identifiziert werden (siehe hierzu die Bewertung einzelner Kriterien in Abschnitt A4.2.4.1, Tabelle 30).

Es ist jedoch zu bedenken, dass die Studie [34] Aussagen zum britischen Versorgungskontext trifft. Für das Lungenkrebsscreening in UK ist abweichend vom deutschen Versorgungskontext ein gestuftes Vorgehen vorgesehen. Dort erhalten aktive sowie ehemalige Raucherinnen und Raucher im Alter von 55 bis 74 Jahren eine Einladung zu einem Lungengesundheitscheck [45]. Eine Low-Dose-CT wird daran anschließend denjenigen angeboten, für die anhand verschiedener Kriterien (z. B. Packungsjahre, Familienanamnese, Vorerkrankungen), ein hohes Lungenkarzinomrisiko bestimmt wurde. In Deutschland werden hingegen Personen zu einer Low-Dose-CT eingeladen, die zwischen 50 und 75 sind und mindestens 25 Jahre (ohne lange Unterbrechungen) geraucht haben und dabei mindestens 15 Packungsjahre erreicht haben [5]. Die untersuchten Screeningpopulationen sind somit weitgehend ähnlich.

Demgegenüber ist für einige der im Full Modell zugrunde gelegten Annahmen bzw. Quellen die Übertragbarkeit auf den deutschen Versorgungskontext fraglich, insbesondere für

- die Prävalenz von Lungenrundherden bei der Verdachtspopulation, die anhand einer Kohorte mit 117 Personen aus 1 japanischen Krankenhaus ermittelt wurde und in der ca. 95 % der untersuchten Personen mindestens 1 Lungenrundherd aufwiesen [44],
- die diagnostische Güte (sowohl für die Verdachts- als auch die Screeningpopulation), die auf einzelnen vorselektierten Kohorten mit geringen Personenzahlen aus dem asiatischen Raum basiert und zudem nicht mittels eines adäquaten Referenztests (Biopsie, klinische Nachbeobachtung oder Nachuntersuchung; vgl. Abschnitt A2.1.2.2) ermittelt wurden, sondern anhand von Expertenpanels [43,44],
- die Modellstruktur, die sich bei der weiteren Nachverfolgung bzw. Abklärung von Lungenrundherden an den Vorgaben der BTS orientiert und von den Empfehlungen der

deutschen S3-Leitlinie zum Lungenkarzinom [4] hinsichtlich des Befundmanagements abweicht sowie

- für die zwischen den Ländern abweichenden Kosten- und Systemstrukturen (insbesondere durch NHS und PSS als Leistungserbringende und Kostenträger).

Es liegt eine zumindest in Teilen geeignete Modellstruktur vor. Das Modell berichtet auf Basis der getroffenen Annahmen, verwendeten Evidenz und adressierten Unsicherheitsaspekte eine gesundheitsökonomische Überlegenheit der Befundung mit AI-Unterstützung in der Screeningpopulation bzw. für die gesundheitsökonomische Überlegenheit der Befundung ohne AI-Unterstützung in der Verdachtspopulation. Jedoch unter Berücksichtigung der oben genannten Aspekte kann die Kosteneffektivität für die Befundung von CT durch Radiologinnen und Radiologen mit AI-Unterstützung im Vergleich zur Befundung ohne AI-Unterstützung nicht abschließend für den deutschen Versorgungskontext bewertet werden.

5 Ergebnisse: Ethische, soziale, rechtliche und organisatorische Aspekte sowie Umwelt- und Klimaaspekte

5.1 Gespräche mit Betroffenen

Es fanden insgesamt 5 Gespräche mit Betroffenen statt (3 Frauen und 2 Männer) (siehe Kapitel A9, Zentrale Ergebnisse aus den Betroffeneninterviews). Zu Beginn jedes Gesprächs wurden die Betroffenen über die Gesprächsziele und die Möglichkeit der AI-Unterstützung für Radiologinnen und Radiologen bei der Auswertung von Röntgen- / CT-Bildern unterrichtet. Jedes Gespräch erfolgte anhand eines strukturierten Gesprächsleitfadens und wurde durch 1 Person mit dem oder der Betroffenen geführt. Im Nachgang verschriftlichte der Interviewer die Gespräche, wertete diese inhaltsanalytisch aus und fasste die Ergebnisse zusammen. Die von den Betroffenen berichteten Erfahrungen, Einschätzungen und Präferenzen wurden insbesondere für die im Folgenden dargestellten ethischen, sozialen, rechtlichen und organisatorischen Aspekte des Themas herangezogen.

5.2 Ergebnisse zu ethischen Aspekten

Der vorliegende ThemenCheck-Bericht liefert mangels ausreichender Datenlage für keine der beiden Fragestellungen eine Nutzen- oder Schadenaussage. Unklar ist somit, ob die AI-Unterstützung bei der Auswertung von Röntgen- oder CT-Bildern mit einem patientenrelevanten Nutzen einhergeht, etwa Verringerung von Mortalität und Morbidität sowie Verbesserung der gesundheitsbezogenen Lebensqualität, weil mehr Personen mit Lungenkrebs früher entdeckt werden und dadurch ein besserer Therapieerfolg eintritt; und ob dieser Nutzen mögliche Schäden, zum Beispiel wegen Fehl- oder Überdiagnosen, überwiegt. Das – ein den Schaden überwiegender Nutzen – ist eine ethische Grundvoraussetzung für die Implementierung medizinischer Maßnahmen ist und außerdem Voraussetzung für die weitere ethische Bewertung:

„Only if there is a (sufficiently valid) actual or expected net-benefit does it make sense to continue the ethical evaluation.“ (Marckmann 2015 [18])

Im Einzelnen

Die ethische Betrachtung widmet sich der Frage, wie ein Einsatz von AI als Unterstützung bei der Auswertung von Röntgen- oder CT-Bildern ausgestaltet sein muss, wenn er mit den vier ethischen Grundprinzipien vereinbar sein soll – Wohltun („beneficence“), nicht schaden („nonmaleficence“), Achtung der Autonomie und Gerechtigkeit [16]. Wegen der multidimensionalen Natur der Fragestellung wurden für die Betrachtung der ethischen Aspekte vier miteinander verwobene Perspektiven berücksichtigt:

- die Perspektive derjenigen, bei denen ein Verdacht auf Lungenkrebs besteht oder die eine Früherkennung in Anspruch nehmen wollen (individuelle Ebene),

- die Perspektive derjenigen, die AI zur Unterstützung bei der Befundung anwenden,
- die Perspektive derjenigen, die AI entwickeln,
- die Perspektive derjenigen, die AI nutzen können sollen (Ebene bestimmter Bevölkerungsgruppen)

Vor diesem Hintergrund wurden die beiden Fragestellungen (AI-Unterstützung bei Personen mit Verdacht auf Lungenkrebs einerseits, bei Personen, die an einer Lungenkrebsfrüherkennung teilnehmen andererseits) in Hinblick auf „erwarteten Nutzen“, „potenziellen Schaden und Belastungen“, „Auswirkungen auf die Autonomie“, „gerechtigkeitsethische Auswirkungen“ und „erwartete Effizienz“ bewertet.

Insgesamt wurden 19 Publikationen eingeschlossen, siehe Abschnitt A7.3.1. Die in die Nutzenbewertung eingeschlossene Studien Lee 2025 [23] und Lo 2018 [24] sowie die in die gesundheitsökonomische Bewertung eingeschlossenen Studien lieferten keine zusätzlichen ethischen Argumente.

Für die Aspekte „erwarteter Nutzen“ und „potenzieller Schaden und Belastungen“ wurde zusätzlich auf die Ergebnisse aus der Nutzenbewertung zurückgegriffen, für die erwartete Effizienz auf die Ergebnisse der gesundheitsökonomischen Bewertung. Schließlich wurden die Synthesen aus den Gesprächen mit Betroffenen herangezogen.

Die tabellarische Zusammenfassung der ethischen Aspekte findet sich in Abschnitt A5.1.2 und Tabelle 32.

5.2.1 Erwarteter Nutzen

Ergebnis der Nutzenbewertung:

Die Domäne Nutzenbewertung liefert für keine der beiden Fragestellungen eine Nutzensaussage.

Erwartete positive Effekte:

Die ausgewertete Literatur lässt vielfältige Erwartungen hinsichtlich erwarteter positiver Wirkungen des Einsatzes von AI erkennen. Diese ersetzen nicht den fehlenden Nutznachweis. Sie werden gleichwohl vorgestellt, um zumindest einen Eindruck zu gewinnen, welche Hoffnungen in die Anwendung von AI-Systemen gesetzt werden.

Wichtigste Erwartung hinsichtlich eines positiven Effekts einer AI-Unterstützung bei der Auswertung von Röntgen und CT-Bildern wäre aus Sicht sowohl betroffener als auch behandelnder Personen eine Verbesserung klinischer Ergebnisse: eine frühere Erkennung von Krebserkrankungen, dadurch erfolgreichere und schonendere Therapien, bessere Behandlungsergebnisse sowie eine höhere Lebensqualität [42,46-49]. Weitere Hoffnungen beziehen

sich auf ein größeres Vertrauen in die Befundung, auf zeitliche Effizienz bei der Befundung, eine Entlastung des medizinischen Personals sowie organisatorische Vorteile im Gesundheitswesen [42,46,47,49].

Die Erwartungen von den für diesen Bericht befragten betroffenen Personen decken sich mit den genannten Aspekten. So gaben diese an, dass eine grundsätzlich bei einem auffälligen Befund vorgesehene „zweite Meinung“, wenn auch mit AI anstelle einer weiteren Ärztin oder eines weiteren Arztes, ein höheres Sicherheitsgefühl vermitteln und vor der Sorge schützen könnte, dass die Ärztin oder der Arzt etwas übersieht. Auch sahen sie Vorteile bei der Befundung wegen der Mischung der großen Datengrundlage der AI-Systeme und der Erfahrung der Ärztinnen und Ärzte (Abschnitt 5.1).

5.2.2 Potenzieller Schaden und Belastungen

Ergebnis der Nutzenbewertung:

Die Domäne Nutzenbewertung liefert für keine der beiden Fragestellungen eine Aussage zu einem potenziellen Schaden oder Belastungen.

Befürchtete negative Effekte:

Fast deckungsgleich zu den Hoffnungen, die sich an den Einsatz von AI-Systemen bei der Auswertung von Röntgen- und CT-Bildern knüpfen, lesen sich umgekehrt die Befürchtungen. Diese ersetzen nicht einen fehlenden Nachweis aus der Nutzenbewertung, werden aber für ein vollständigeres Bild von Hoffnungen und Sorgen gleichwohl vorgestellt. So besteht die Sorge vor Fehldiagnostik: zum einen, wenn Ergebnisse von AI-Systemen nicht von Radiologinnen und Radiologen geprüft werden; zum anderen, wenn wegen zunehmender technischer Unterstützung ärztliche Fähigkeiten abnehmen („De-Skilling“) und Sorgfaltspflichten im Umgang mit den Ergebnissen der AI-Systeme nicht im erforderlichen Maße wahrgenommen werden. Der „Automation Bias“, also das „blinde“ Folgen der AI-Empfehlung, würde begünstigt, wenn ärztliches Erfahrungswissen verloren geht. Daraus nährt sich die Befürchtung, dass fehlerhafte AI-Ergebnisse unnötige Folgediagnostik und Überbehandlung auslösen könnten und schlechtere Behandlungsergebnisse eintreten [49-52]. Auf Gruppenebene wird die Sorge geäußert, dass AI-Systeme zu ungünstigeren Gesundheitsergebnissen für bestimmte Gruppen beitragen können, wenn zum Beispiel bestimmte Gruppen bei der Entwicklung des AI-Systems unterrepräsentiert sind (Algorithmus-Bias) [53]. Neben Bedenken zum Datenschutz wird außerdem das Arzt-Patienten-Verhältnis in Gefahr gesehen, vor allem, wenn Ergebnisse von AI-Systemen nicht mit der ärztlichen Einschätzung übereinstimmen. Es besteht insbesondere in diesen Fällen die Sorge vor einem schwindenden Vertrauen in die (nur) ärztliche Befundung; damit einher geht die Sorge ärztlicher Fachkräfte vor rechtlichen Konsequenzen [46,47,51,53]. Aus ärztlicher Perspektive wird außerdem das Potenzial von AI-Systemen als bedrohlich wahrgenommen, mit ärztlicher Tätigkeit in Konkurrenz zu treten und menschliche Fachkräfte zu ersetzen [46,47,54].

Umgekehrt zu den Erwartungen positiver Auswirkungen werden außerdem Verschlechterungen der zeitlichen Effizienz bei der Befundung, eine ausbleibende Entlastung des medizinischen Personals sowie organisatorische Nachteile und Kostensteigerungen im Gesundheitswesen befürchtet [46-52].

Die Perspektive der Entwickler-Seite lenkt den Blick auf weitere mögliche Risiken: Es werden Bedenken erhoben in Hinblick auf die Einhaltung des Datenschutzes [46], die Manipulationsanfälligkeit von AI-Systemen, die vom Anwender unbemerkte Einflussnahme auf die Ergebnisausgabe von AI-Systemen erlaubt [55], oder die Gewährleistung der Cybersicherheit [48,56].

Die in der Literatur identifizierten Aspekte zu potenziellen Schäden und Belastungen tauchten auch in der Betroffenenbefragung für den vorliegenden Bericht auf. Auch hier wurde die Sorge geäußert, dass sich Ärztinnen und Ärzte zu sehr auf AI-Systeme verlassen könnten und vertraten die Ansicht, dass AI nur unterstützen, nicht aber die ärztliche Einschätzung ersetzen dürfe und AI-Ergebnisse von ärztlichen Fachkräften zu prüfen seien. Auch die Fehleranfälligkeit von AI-Systemen wurde benannt, wobei eine Präferenz für falsch-positive Ergebnisse in Abgrenzung zu falsch-negativen vorherrschte. Schließlich wurden auch Sorgen in Hinblick auf die Einhaltung des Datenschutzes geäußert (siehe Abschnitt 5.1).

5.2.3 Auswirkungen auf die Autonomie

Der Aspekt der Autonomie hat sowohl für Patientinnen und Patienten als auch für Ärztinnen und Ärzte Bedeutung.

Patientenautonomie

Für Patientinnen und Patienten steht die selbstbestimmte, autonome Entscheidung zur Inanspruchnahme von medizinischen Behandlungen, darunter fällt auch Diagnostik, im Vordergrund. Damit fordert die Wahrung der Patientenautonomie, dass Patientinnen und Patienten alle relevanten Informationen zu einer vorgeschlagenen Behandlung erhalten. Das beinhaltet auch eine Information über den Einsatz von AI-Systemen zur Unterstützung der Beurteilung von Röntgen- oder CT-Bildern. Schließlich müssen die Ergebnisse AI-unterstützter Befunderhebung für Patientinnen und Patienten nachvollziehbar sein, um hierauf weitere Behandlungsentscheidungen treffen zu können. Insofern kann Transparenz zu einer Mehrung von Patientenautonomie führen, wenn „AI-Entscheidungen“ in das Arzt-Patient-Gespräch zur Behandlungsentscheidung eingebunden werden. Im umgekehrten Fall kommt es zu einem Autonomieverlust, der als „Computer-Paternalismus“ beschrieben wird: Wenn lediglich ein Befund vorgelegt wird, ohne Informationen darüber, wie dieser zustande kam – ob AI eingebunden war, ob die Ergebnisse von einer Radiologin oder einem Radiologen überprüft wurden, ob ärztliche Einschätzung und AI-Ergebnis übereinstimmten oder nicht, und wenn nicht, warum die ärztliche Entscheidung anders ausgefallen ist – dann können Patientinnen

und Patienten den Befund nicht einordnen, und Folgeentscheidungen fallen nicht selbstbestimmt [47,52].

Zu Fragen der informationellen Selbstbestimmung im Zusammenhang mit einer möglichen Weiterverwendung von CT- oder Röntgenbildern siehe Abschnitt 5.4.2.

Therapiefreiheit

Für Ärztinnen und Ärzte als Anwender von AI-Systemen in der Diagnostik steht die ärztliche Therapiefreiheit im Fokus des Autonomie-Aspekts. Die ärztliche Therapiefreiheit umfasst die freie Wahl der Therapie, darunter fällt auch die Wahl der Diagnostik, wobei diese immer dem Wohl der Patientinnen und Patienten verpflichtet ist. Ähnlich wie auf Seite der Patientinnen und Patienten ist Voraussetzung für eine autonome Entscheidung die Transparenz: AI-Anwendende müssen wissen, ob das AI-System auf die einzelne Patientin, den einzelnen Patienten passt. Dafür muss bekannt sein, welcher Algorithmus der AI zugrunde liegt und mit welchen Patientendaten das System trainiert wurde. Es kann nicht verlangt werden, dass Ärztinnen und Ärzte die internen Prozesse eingesetzter AI-Systeme vollständig durchdringen. Aber sie sind doch zumindest auf diese Informationen angewiesen, um entscheiden zu können, ob sie die AI im konkreten Fall einsetzen können; und wenn sie sie einsetzen, sind sie auf diese Informationen angewiesen, um AI-generierte Ergebnisse interpretieren zu können [47,50,57,58]. Eine interessante Beobachtung ist in diesem Zusammenhang, dass mehr Transparenz, die es ärztlichen Fachkräften ermöglicht, Unstimmigkeiten in den AI-Vorhersagen zu beobachten, das Vertrauen von Radiologinnen und Radiologen in das AI-System verringern kann. Das ist ungünstig, wenn es das ärztliche Vertrauen in AI-Ergebnisse schwächt. Aber womöglich ist es angemessen, wenn die Leistung des AI-System nicht optimal und inkonsistent ist [58].

Die Bedeutung der Transparenz wurde auch bei den für diesen Bericht befragten Betroffenen deutlich: Sie wünschen sich einen offenen Umgang mit dem Einsatz von AI-Systemen, gerade dann, wenn AI-System und Radiologin oder Radiologe zu unterschiedlichen Ergebnissen kommen (siehe Abschnitt 5.1).

5.2.4 Auswirkungen auf die Gerechtigkeit

In Hinblick auf gerechtigkeitsethische Auswirkungen stehen Bedenken im Vordergrund, nicht dagegen Erwartungen an eine gerechtere Versorgungssituation. Bei der Anwendung von AI-Systemen steht die Sorge vor der Verstärkung von Gruppen-Ungleichheiten im Vordergrund, zum Beispiel bei fehlerhaften AI-Algorithmen oder bei Verwendung von AI-Trainingsdaten, die nicht divers genug sind, um auch in der Versorgung unterrepräsentierter Gruppe abzudecken oder ethnische Unterschiede zu berücksichtigen [50,53,59,60]. Von herausragender Bedeutung sind daher die Anforderungen an die Transparenz vor allem der Datengrundlage von AI-Systemen, und an die ausreichende Heterogenität der Datengrundlage selbst. Die

Beschaffung ausreichend heterogener Daten wird derzeit insbesondere als datenschutzrechtliches Problem gesehen. Um „gerechtere“ AI-Systeme zu schaffen, mit denen unterrepräsentierte Gruppen besser abgedeckt sind, werden AI-Systeme basierend auf föderativem Lernen entwickelt (federated learning). Beim föderativen Lernen werden AI-Systeme lokal in einzelnen Krankenhäusern, Zentren oder Praxen mit Patientendaten „gefüttert“. Das auf diese Weise weiterentwickelte System wird anschließend mit Daten aus anderen Häusern weiter „gefüttert“. Weil so individuelle Patientendaten nicht weitergegeben werden, entlastet das in datenschutzrechtlicher Hinsicht. Eine Kombination von Daten aus regionalen, nationalen oder internationalen Einrichtungen könnte größere Datendiversität schaffen; davon könnten sowohl unterrepräsentierte Patientengruppen profitieren als auch Krankenhäuser mit geringeren Ressourcen, die anders nicht an ausreichend Daten kommen. Das föderative Lernen befindet sich noch am Beginn der Entwicklung [61].

Sofern sich die Anwendung von AI-Systemen – zukünftig – in der Beurteilung von Röntgen- oder CT-Bildern der Lunge gegenüber der herkömmlichen Befundung überlegen erweisen sollte, fordert das Prinzip der Gerechtigkeit eine Sicherstellung des gerechten Zugangs für alle Patientengruppen, sowohl örtlich als auch gruppenzugehörig [47].

5.2.5 Erwartete Effizienz

Weil für die vorliegenden Fragestellungen keine belastbare Evidenz zu einem Nutzen oder Schaden der AI-unterstützten Befundung gegenüber der Befundung ohne AI-Unterstützung vorliegt und auch die Frage der Kosten-Wirksamkeit vorliegend nicht abschließend für den deutschen Versorgungskontext beantwortet werden kann (siehe Kapitel 3 und 4) bleibt die erwartete Effizienz unklar.

5.2.6 Ausblick

Vor dem Hintergrund der in der Literatur identifizierten, von Betroffenen berichteten und in den weiteren Domänen adressierten ethischen Aspekten wird deutlich, dass Evidenz zu Nutzen und Schaden unverzichtbar ist, um über die ethische Vertretbarkeit einer Implementierung von AI-Systemen zur Unterstützung bei der Befundung von Röntgen- oder CT-Bildern zu entscheiden. Außerdem wurde sichtbar, dass Erwartungen und Befürchtungen gegenläufig sind, dass die Sicherstellung der Autonomie von Patientinnen und Patienten wie auch von Ärztinnen und Ärzten eine zentrale Rolle spielt und dass gerechtigkeitsethische Anforderungen zu erfüllen sind. Die Aufgabe der Ethik wird es sein, den Prozess der Implementierung von AI-Systemen aufmerksam zu begleiten.

Der Deutsche Ethikrat hat sich in einer im März 2023 veröffentlichten Stellungnahme mit den Herausforderungen durch Künstliche Intelligenz befasst und zum Einsatz von AI-Systemen in der Medizin neun Empfehlungen formuliert [47], die sich mit den für diesen Bericht identifizierten Argumenten in Einklang bringen lassen (hier zusammenfassend formuliert):

- 1) Zur Qualitätssicherung ist eine enge Kooperation mit Zulassungsbehörden und medizinischen Fachgesellschaften bei der Entwicklung, Erprobung und Zertifizierung von AI-Produkten erforderlich, damit Schwachstellen früh erkannt werden.
- 2) Trainings-, Validierungs- und Testdatensätze müssen divers sein (z. B. Alter, Geschlecht, Vorerkrankungen), durch Monitoring und Dokumentationspflichten ist das sicherzustellen.
- 3) AI-Systeme sollen Ergebnisse so darstellen, dass Risiken wie Automation Bias erkennbar sind und die Anwender zu einer reflektierten Plausibilitätsprüfung anregen.
- 4) Bei Erhebung, Verarbeitung und Weitergabe von Gesundheitsdaten sind strenge Standards in Bezug auf Aufklärung, Datenschutz und Wahrung der Privatsphäre erforderlich.
- 5) Bei empirisch belegter Überlegenheit von AI-Anwendungen ist ein gerechter Zugang für alle betroffenen Patientengruppen sicherzustellen.
- 6) Bei empirisch belegter Überlegenheit von AI-Anwendungen ist die Anwendung von AI-Systemen in Aus-, Fort- und Weiterbildung für ärztliches Fachpersonal sowie andere Gesundheitsberufe zu übernehmen.
- 7) Bei routinemäßiger Anwendung von AI muss eine hohe methodischer Expertise zur Einordnung der Ergebnisse, Einhaltung von Sorgfaltspflichten bei Datenerhebung und -weitergabe sichergestellt sein und mit Fortbildungen gegen „De-Skilling“ vorgegangen werden.
- 8) Es muss sichergestellt sein, dass Patientinnen und Patienten über alle entscheidungsrelevanten Umstände informiert werden. Gegen ein Gefühl der Verobjektivierung ist gezielt vorzugehen; das Arzt-Patienten-Vertrauensverhältnis muss geschützt werden, ebenso die „sprechenden Medizin“ – und es sollte wegen des Einsatzes von AI-Systemen kein Personal abgebaut werden.
- 9) Ärztliche Fachkräfte dürfen nicht ersetzt werden; sie bleiben verantwortlich für Planung, Durchführung und Überwachung des Behandlungsprozesses.

Viele dieser Anforderungen werden in den Abschnitten zu sozialen und organisatorischen Aspekten adressiert. Im Abschnitt 5.4 „Rechtliche Aspekte“ werden sie nochmals explizit aufgegriffen, und es wird sich zeigen, dass Normgeber auf nationaler und europäischer Ebene diese Aspekte im Blick haben und suchen, ihnen rechtlich gerecht zu werden.

5.3 Ergebnisse zu sozialen Aspekten

Für die Auswertung der sozialen Aspekte wurden insgesamt 31 Publikationen eingeschlossen (vgl. Abschnitt A7.3.2). Die Informationsaufbereitung orientiert sich an dem von Mozygemba (2016) [20] vorgeschlagenen konzeptionellen Rahmen (vgl. Abschnitt A5.1.2, Tabelle 33).

Die Auswertung von sozialen Aspekten beruht auf den Ergebnissen der Betroffenenbefragungen, den eingeschlossenen Publikationen und „reflective thoughts“ [19]. Nach den bereits skizzierten ethischen Aspekten (vgl. Abschnitt 5.2) fokussiert dieser Abschnitt auf folgende soziale Aspekte: Verständnis des Gesundheitsproblems Lungenkrebs und der Befundung der Lunge bei Verdacht auf Lungenkrebs bzw. im Rahmen einer Lungenkrebsfrüherkennung; soziales Image und Verständnis des Einsatzes von AI bei der Befundung der Lunge bei Verdacht auf Lungenkrebs bzw. im Rahmen einer Lungenkrebsfrüherkennung; sowie soziokulturelle Aspekte der Implementierung der Technologie: soziokulturelle Eigenschaften der Zielgruppe, Organisation der Nutzung, soziale Ungleichheit und Beziehungen zwischen Betroffenen und Leistungserbringenden bzw. zwischen verschiedenen Leistungserbringenden.

Verständnis des Gesundheitsproblems

Im Folgenden wird zunächst das Verständnis des Gesundheitsproblems Lungenkrebs durch Betroffene hinsichtlich einer Befundung der Lunge bei Verdacht auf Lungenkrebs bzw. im Rahmen einer Lungenkrebs-Früherkennung ohne AI-Unterstützung beschrieben. Die LuKrFrühErkV [5], die im Mai 2024 verabschiedet wurde, ist vor dem Hintergrund einer hohen Lungenkrebsfatalität und der durch aussagekräftige Evidenz nachgewiesenen Möglichkeit einer lungenkrebspezifischen Mortalitätsreduktion bei bestimmten Personen durch eine frühe Krankheitserkennung entstanden [27,62-65]. In der Betroffenenbefragung als Teil der vorliegenden Bewertung zeigte sich jedoch, dass Personen, die bereits einmal eine Röntgen- oder eine CT-Untersuchung der Lunge mit oder ohne Verdacht auf eine Lungenkrebs-erkrankung durchlaufen hatten, überwiegend wenig Wissen darüber hatten, dass die frühe Entdeckung einer Lungenkrebserkrankung die Chancen auf Heilung erhöhen und das lungenkrebspezifische Mortalitätsrisiko senken kann. Allgemein nahmen die Betroffenen Lungenkrebs als eine Erkrankung wahr, die in kurzer Zeit zum Tode führt (vgl. Tabelle 33). Die gesichteten Studien berichten, dass Personen, die ihr persönliches Risiko für eine Lungenkrebserkrankung falsch oder zu hoch einschätzen, die aufgrund einer möglicherweise falschen Risikoeinschätzung Angst vor einer Lungenkrebsdiagnose haben oder die über die Früherkennung nicht oder nicht ausreichend von behandelnden Ärztinnen und Ärzte informiert werden, seltener eine Früherkennung in Anspruch nehmen [66-69]. Obwohl dies Personen sind, die von einer Früherkennung studienbasiert besonders profitieren könnten, denn es handelt sich hierbei vergleichsweise häufig um starke Raucher, ältere Menschen oder Frauen [66-71] (vgl. Abschnitt zu soziokulturellen Aspekten der Implementierung).

Soziales Image und Verständnis des Einsatzes von AI

Verständnis durch Betroffene

Keine der befragten Personen hat Erfahrung mit dem Einsatz von AI, weder in der Radiologie noch in anderen Bereichen der Gesundheitsversorgung. Dennoch beschreiben die befragten Personen mögliche Vorteile einer AI-unterstützten Untersuchung, wie zum Beispiel, dass

Radiologinnen und Radiologen zusammen mit einer AI möglicherweise weniger übersehen und so auch kleinste Veränderungen in der Lunge früh entdeckt werden können. Qualitative und quantitative Befragungs-Studien (vgl. Tabelle 33) berichten, dass Menschen trotz eines im Durchschnitt geringen Wissens über die AI eine AI-unterstützte Bildgebung überwiegend positiv bewerten und ähnliche Vorteile der AI-Unterstützung wie die Personen in der Betroffenenbefragung sehen [46,48,72]. Patientinnen und Patienten der Radiologie sehen über den Nutzen für sich persönlich hinaus die Möglichkeit, dass Gesundheitssysteme insgesamt und die Radiologinnen und Radiologen im Besonderen durch AI unterstützt und entlastet werden könnten, was gerade in Ländern mit bestehendem Fachkräftemangel als Vorteil der AI gewertet wird [48]. Die Studien berichten außerdem, dass, obwohl die AI insgesamt positiv gesehen wird, bestimmte Personengruppen einer AI-Unterstützung in der Gesundheitsversorgung eher kritisch oder skeptisch gegenüberstehen. Hierzu gehören eher Frauen, ältere Menschen, weniger gesunde, weniger gebildete und weniger Technik- bzw. AI-affine Personen [46,72] (vgl. Tabelle 33). Als besonders wichtig scheint den befragten Personen in der Betroffenenbefragung und in qualitativen Studien, dass bei Unterstützung durch eine AI die Radiologin oder der Radiologe die Deutungshoheit behält und die Ergebnisse einer AI-unterstützten Befundung mit dem bzw. der Betroffenen bespricht [46,48,72]. Eine Befundung durch die AI allein wurde von den befragten Personen in der Betroffenenbefragung und in Studien durchgehend abgelehnt. Als Risiken einer AI wurden von den befragten Personen in der Betroffenenbefragung und in den Studien besonders datenschutzrechtliche Aspekte genannt (vgl. hierzu auch Abschnitt 5.4 „Rechtliche Aspekte“) sowie die Möglichkeit von mehr falsch-positiven Befunden durch die AI als durch den Arzt allein [46,48]. Letztere wurden jedoch durchgehend als weniger schlimm oder zu bevorzugen eingeschätzt als falsch-negative Befunde, die bei Lungenkrebs in der Betroffenenbefragung als möglicherweise „tödlich“ wahrgenommen wurden.

Verständnis von Anwenderinnen und Anwendern

Qualitative Studien berichten, dass Leistungserbringende in der Radiologie, die noch nicht mit einer AI gearbeitet haben, eher skeptisch bezüglich deren Einsatz sind, wohingegen Leistungserbringende, die bereits Erfahrung mit dem Einsatz von AI in der Befundung von Lungenkrebs gemacht haben, viele Vorteile und wenige Nachteile bzw. „Dinge, die es zu beachten gibt“ beschreiben [48,73] (vgl. Tabelle 33). Als Vorteil wird die Nutzung der AI als zuverlässige zweite Meinung identifiziert, die die Arbeit schneller und effizienter machen kann. Als Nachteil wird die Möglichkeit beschrieben, dass durch eine unsachgemäße Nutzung von AI – insbesondere durch unerfahrene Radiologinnen und Radiologen allein – ein zu starkes Verlassen auf die AI zu Problemen bei der Befundung führen kann. Wichtig zu beachten sei hier, dass Erfahrungswissen durch die AI nicht abgebaut, sondern AI-spezifisch in gemischten Teams mit erfahrenen und unerfahrenen Radiologinnen oder Radiologen aufgebaut und erhalten werden müsse [48]. So könnten zum Beispiel Fehlerquellen oder Limitationen der AI (z. B. Darstellung von Blutgefäßen statt krankhafter Veränderungen) identifiziert und durch

die beteiligten Leistungserbringenden der Umgang hiermit erlernt werden. Eine gute Integration der AI in die bestehende personelle Infrastruktur sei wichtig [48]. Aber auch eine gute IT-Unterstützung (insbesondere in der Anfangsphase der AI-Anwendung) und die Integration in bestehende IT-Systeme wurde durch die befragten Leistungsbringer als Bedingung für die erfolgreiche Implementierung einer AI gewertet [48].

Soziokulturelle Aspekte der Implementierung der Technologie

Eine AI-unterstützte Befundung von Röntgen- oder CT-Bildern der Lunge kommt unter anderem für 2 Gruppen in Betracht: Personen mit Verdacht auf Lungenkrebs und Personen, die eine Früherkennungsuntersuchung nach der LuKrFrühErkV in Anspruch nehmen. Die Zielgruppe der LuKrFrühErkV sind Männer und Frauen über 50 und unter 76 Jahren, die einen hohen Zigarettenkonsum, wie in der LuKrFrühErkV spezifiziert, aufweisen [5,74]. Die Eigenschaften der Zielgruppe der LuKrFrühErkV kennzeichnen auch – wenn auch nur teilweise – die Personen mit Verdacht auf Lungenkrebs, da höheres Alter (> 50) und Rauchen die wichtigsten Risikofaktoren für eine Lungenkrebserkrankung sind [2,75]. Während die Inzidenz und Mortalität von Lungenkrebs in den letzten Jahrzehnten bei Männern gesunken ist, stieg sie bei Frauen an. Gleichlaufend stieg die Anzahl von Frauen, die rauchen, obwohl sie immer noch seltener rauchen und seltener an Lungenkrebs erkranken als Männer [2,76]. Frauen scheinen gemäß verschiedener Erhebungen genauso oder sogar mehr als Männer von einer Lungenkrebsfrüherkennung zu profitieren, erfüllten aber zum Beispiel seltener die Früherkennungs-Kriterien der amerikanischen United States Preventive Services Task Force (USPSTF) zum Zigarettenkonsum, werden seltener von Leistungserbringenden über die Möglichkeit der Lungenkrebsfrüherkennung informiert und nehmen diese seltener wahr [67,70,71,77-79]. Gleichzeitig zeigt sich, dass die Personengruppen, die seltener eine Lungenkrebsfrüherkennung in Anspruch nehmen – wie Frauen oder ältere Menschen [67-69,79] – auch zu den Gruppen gehören, die dem Einsatz einer AI-Unterstützung beim Röntgen oder CT eher kritisch gegenüberstehen [46]. Auch Personen mit geringerer Bildung oder geringer Technik bzw. IT(Informationstechnologie)-Affinität sehen AI-Technologien in der Gesundheitsversorgung eher kritisch [46,72]. Personen mit einer schlechteren Bildungs-, Arbeits- oder Einkommenssituation weisen zudem auch eine höhere Rauchprävalenz und Lungenkrebsinzidenz auf [80-82].

Zusammenfassend ist davon auszugehen, dass die Implementierung der Lungenkrebsfrüherkennung mit oder ohne AI-Unterstützung aufgrund der soziokulturellen Eigenschaften der Zielgruppe mit Herausforderungen verbunden sein kann und einige der Zielgruppen für eine Früherkennung schwer erreichbar sein werden. Sofern ein AI-System in der Befundung oder Früherkennung zum Einsatz kommen soll, ist zu prüfen, ob und wie Personen über die Vorgaben der LuKrFrühErkV [5] hinausgehend über den Einsatz von AI zu informieren sind (vgl. Abschnitte 5.2 und 5.4).

Soziokulturelle Aspekte der Implementierung der Technologie – soziale Ungleichheit bei Inanspruchnahme und Zugang

Soziale Ungleichheit bei der Inanspruchnahme einer AI-unterstützten Befundung kann zum einen bereits aus einer fehlerhaften Entwicklung einer AI-Technologie, d. h. beim Trainieren der Daten, resultieren (vgl. Tabelle 33). Es ist daher wesentlich, dass AI-Trainingsdaten ausgewogen die relevanten Zielgruppen abbilden und jede AI die genutzten Daten transparent und für Anwender nachvollziehbar dokumentiert [53,83-85].

Über die Entwicklung der AI hinaus kann soziale Ungleichheit bei der Inanspruchnahme aber auch daraus resultieren, dass bestimmte soziale Gruppen (wie Frauen, weniger gebildete oder informierte und ältere Menschen) aus den bereits genannten Gründen eine Lungenkrebsfrüherkennung nicht in dem Ausmaß wahrnehmen wie sie es könnten, zum Beispiel weil sie falsche Erwartungen an die Früherkennung haben, nicht oder unzureichend über die Vor- und Nachteile informiert sind oder ihnen diese seltener angeboten wird [65,67-69,71,86]. Die LuKrFrühErkV sieht vor, dass alle Anspruchsberechtigten über die Möglichkeit der Früherkennung von ihren behandelnden Ärztinnen und Ärzten (Allgemeinmediziner, Ärztinnen und Ärzte für innere Medizin oder für Arbeitsmedizin) mit Hilfe einer Versicherteninformation informiert [5,87]. Wichtig erscheint zur Vermeidung von Ungleichheiten zunächst eine Sensibilisierung der Leistungserbringenden für die Besonderheiten der genannten Zielgruppen. Eine zielgruppenspezifische und verständliche Ansprache dieser Zielgruppen sowie auf diese formal und inhaltlich abgestimmte und evidenzbasierte Informationen zur AI-unterstützten Befundung und deren Vor- und Nachteilen könnten deren Akzeptanz und Inanspruchnahme erhöhen [67,68,71,86].

Soziale Ungleichheit beim Zugang zu einer AI-unterstützten Befundung oder Früherkennung kann darüber hinaus dadurch entstehen, dass diese Leistungen nur in bestimmten, hochspezialisierten Lungenkrebszentren angeboten werden [79,84]. Da die Anschaffung, die Einführung und die dauerhafte Nutzung von AI-Systemen mit einem hohen finanziellen Aufwand (Anschaffungskosten und laufende Kosten, Training der Anwender, Anpassungen in der IT-Infrastruktur eines Zentrums) verbunden sind [88] (vgl. Abschnitte 4.2 und 5.5), kann möglicherweise nicht jedes Radiologiezentrum sich diese aufwändige Innovation leisten. Durch die regional dann möglicherweise sehr unterschiedliche Verteilung von Zentren, die eine AI-unterstützte Befundung anbieten, kann soziale Ungleichheit durch regionale bzw. kleinräumige Versorgungsunterschiede entstehen. Mobile Früherkennungszentren analog zu den Mammografie-Bussen werden als Lösungsansätze für diese Herausforderung besprochen [79,84]. Es ist jedoch fraglich, inwiefern eine AI-unterstützte Befundung oder Früherkennung auch in solchen Settings angeboten werden könnten.

5.4 Ergebnisse zu rechtlichen Aspekten

Die rechtlichen Aspekte der Frage: „Kann Künstliche Intelligenz die Auswertung von Röntgen- oder CT-Bildern der Lunge unterstützen?“ wurden anhand der Checkliste von Brönneke et al. 2016 [21] geprüft. Für die Auswertung der rechtlichen Aspekte wurden insgesamt 32 Publikationen eingeschlossen (vgl. Abschnitt A7.3.3).

Die Checkliste enthält 9 Punkte: 1. Patientenautonomie I: informierte Einwilligung; 2. Patientenautonomie II: Einwilligungsunfähigkeit; 3. Autonomie III: Datenschutz; 4. Marktzugang I: Medizinprodukte; 5. Marktzugang II: Arzneimittel; 6. Klinische Studien; 7. Intellektuelles Eigentum; 8. Kostenerstattung im öffentlichen Gesundheitssystem und 9. Spezielle medizinische Bereiche.

Die rechtlich wesentlichen Aspekte der Frage: „Kann Künstliche Intelligenz die Auswertung von Röntgen- oder CT-Bildern der Lunge unterstützen?“ wurden den Punkten 1. Patientenautonomie I: informierte Einwilligung, 3. Autonomie III: Datenschutz und 9. Spezielle medizinische Bereiche und daraus die Aspekte Haftung und Anforderungen an AI-Systeme insbesondere nach dem AI Act [6] zugeordnet. Diese Aspekte werden im Folgenden ausführlicher erörtert. Die Punkte 4. Marktzugang I: Medizinprodukte; 6. Klinische Studien und 8. Kostenerstattung im öffentlichen Gesundheitssystem und sind für die Fragestellung relevant, werden aber nur angerissen. Die übrigen Punkte aus der Checkliste von Brönneke et al. 2016 [21] 2. Patientenautonomie II: Einwilligungsunfähigkeit, 5. Marktzugang II: Arzneimittel und 7. Intellektuelles Eigentum sind vorliegend nicht relevant. Die vollständige Abbildung der Aspekte findet sich in Tabelle 34.

Die tabellarische Zusammenfassung der rechtlichen Aspekte findet sich in Abschnitt A5.1.2, Tabelle 34.

Vorab einige Anmerkungen:

1) Rechtliche Definition AI

Für die rechtlichen Betrachtungen wird die Definition für künstliche Intelligenz des AI Acts [6] zugrunde gelegt.

2) Einsatz der AI

In der vorliegenden Fragestellung geht es um AI-Systeme zur Unterstützung der radiologischen Befundung von Röntgen- oder CT-Bildern; die finale Diagnoseentscheidung liegt bei den Radiologinnen und Radiologen. Die folgenden Ausführungen beinhalten nicht den vollständigen Ersatz der ärztlichen Entscheidung durch AI. Aus verschiedenen Gründen entspricht das der rechtlichen Einordnung, nach der der vollständige Ersatz der ärztlichen Entscheidung nicht zulässig ist (siehe dazu im Folgenden).

3) Einheitliche Abhandlung der 2 Fragestellungen

Die rechtliche Bewertung erfolgte für die beiden Fragestellungen – bezogen auf Personen mit Verdacht auf Lungenkrebs einerseits und Personen, die eine Früherkennung in Anspruch nehmen andererseits – einheitlich; wo Unterschiede auftraten, wurde das kenntlich gemacht.

5.4.1 Patientenautonomie I: informierte Einwilligung

Der Aspekt der informierten Einwilligung ist in der vorliegenden Fragestellung 2-fach relevant. Erstens betrifft er die Einwilligung der Patientin oder des Patienten in die vorgeschlagene diagnostische Maßnahme (zum Beispiel Röntgen oder CT unter Einbindung eines AI-Systems). Zweitens umfasst er die Einwilligung in die Erhebung und Nutzung von Patientendaten sowohl für die Entwicklung eines AI-Systems als auch für die spätere Verwendung von Daten in einem AI-System. Unter die rechtliche Betrachtung der Patientenautonomie wird vorliegend der erste Fall gefasst. Fragen zur datenschutzrechtlichen Einwilligung fallen unter den Punkt Autonomie III: Datenschutz (siehe Abschnitt 5.4.2).

Ob eine Aufklärung speziell über den Einsatz eines AI-Systems als Unterstützung bei der Auswertung von Röntgen- oder CT-Bildern für eine rechtmäßige Einwilligung erforderlich ist, wird unterschiedlich beurteilt. Medizinische Maßnahmen als Eingriffe in die körperliche Unversehrtheit sind grundsätzlich nur dann gerechtfertigt, wenn die betroffene Patientin oder der betroffene Patient wirksam eingewilligt hat, § 630d BGB (Bürgerliches Gesetzbuch). Für die rechtswirksame Einwilligung hat eine Aufklärung nach den Vorgaben des § 630e BGB zu erfolgen. Danach ist über „sämtliche für die Einwilligung wesentlichen Umstände“ aufzuklären (§ 630e Abs. 1 S. 1 BGB), insbesondere über „Art, Umfang, Durchführung, zu erwartende Folgen und Risiken der Maßnahme sowie ihre Notwendigkeit, Dringlichkeit, Eignung und Erfolgsaussichten im Hinblick auf die Diagnose oder die Therapie“ (§ 630e Abs. 1 S. 2 BGB). Sonderregelungen gelten für Röntgenaufnahmen, die vorliegend aber nicht erörtert werden (siehe hierzu zum Beispiel Wigge 2021 [89]).

Befürworter einer umfassenden Aufklärungspflicht – einschließlich der Information über den Einsatz eines AI-Systems – verweisen vor allem auf die Einstufung des Einsatzes von AI-Systemen als „neue Methode“. In solchen Fällen müsse zusätzlich zur üblichen Aufklärung darauf hingewiesen werden, dass die Methode (noch) nicht allgemein anerkannt sei und unbekannte Risiken bestehen können. Dies folge dem Ziel der Aufklärung, Patientinnen und Patienten zu einer informierten Abwägung von Vor- und Nachteilen zu befähigen, was eine hinreichende Kenntnis über die beabsichtigte Behandlungsmethode voraussetze [55,90]. Darüber hinaus wird betont, dass gerade bei neuen Methoden häufig Fragen zu Sicherheit, Genauigkeit und zum Datenschutz auftreten; ebenso die Sorge, auf Daten reduziert statt als Person wahrgenommen zu werden [47]. Mit dieser Auffassung wären Patientinnen und Patienten darüber aufzuklären, wenn bei der Auswertung von Röntgen- oder CT-Bildern der

Lunge zur Früherkennung oder bei Verdacht auf Lungenkrebs ein unterstützendes AI-System eingesetzt wird.

Demgegenüber stützt sich die Auffassung, wonach nicht über den Einsatz eines AI-Systems aufzuklären ist, auf die Grundregel, dass (nur) über die wesentlichen Umstände der Behandlung aufzuklären sei. Der Einsatz von „Clinical Decision Support Systemen“ gehöre nur dann dazu, wenn diese als eigenständige Untersuchungs- oder Behandlungsmethode im Rahmen einer medizinischen Behandlung eingesetzt werden. Diene ein AI-System dagegen lediglich als unterstützende Neben- oder Begleitleistung, bestehe weder eine besondere Aufklärungs- noch eine Einwilligungspflicht [52]. Nach dieser Sichtweise ergibt sich keine Verpflichtung, über den Einsatz von AI-Systemen gesondert aufzuklären.

Welche dieser beiden Ansichten sich durchsetzt, bleibt abzuwarten. Wünschenswert ist, dass dabei der ethische Aspekt der Patientenautonomie berücksichtigt wird. Dieser spricht für eine umfassende Aufklärung von Patientinnen und Patienten (Abschnitt 5.4.1).

5.4.2 Autonomie III: Datenschutz

Für den Einsatz von AI-Systemen ist der Datenschutz von besonderer Bedeutung. Einschlägige Vorschriften finden sich in der Datenschutz-Grundverordnung (DSGVO) [91], im Bundesdatenschutzgesetz (BDSG) [92], im AI Act [6], im European Health Data Space Verordnung (EHDS-Verordnung) [83] sowie im Gesundheitsdatennutzungsgesetz (GDNG) und Digitalgesetz (DigiG) [93,94].

AI-Systeme werden in der Entwicklung mit personenbezogenen Daten gefüttert. Auch wenn es Ansätze gibt, durch Augmentation Daten künstlich zu erweitern, etwa um die Menge und Heterogenität der Daten zu erhöhen, sind doch zunächst Daten von „echten Menschen“ erforderlich. Bei den für solche Systeme erforderlichen Daten handelt es sich um besonders sensible und daher datenrechtlich besonders geschützte Gesundheitsdaten. Nach Art. 9 Abs. 1 DSGVO gehören sie zu den „besonderen Kategorien personenbezogener Daten“ und unterfallen damit grundsätzlich einem Verarbeitungsverbot. Dieses Verbot wird in Art. 9 Abs. 2 DSGVO durch definierte Ausnahmetatbestände aufgehoben, etwa, wenn die betroffene Person ausdrücklich in die Datenverarbeitung eingewilligt hat. Für die Entwicklung von AI-Systemen werden solche besonders geschützten Gesundheitsdaten in großer Menge benötigt. Das komplexe System des Datenschutzes wird daher oft als Hürde für den Fortschritt bemängelt (vgl. [54]). AI-Systeme, die auf einem föderativen Lernen (federated Learning) basieren, sollen Erleichterung schaffen, indem das System lokal „gefüttert“ und nur das trainierte Programm weitergegeben wird [61] (siehe auch Abschnitt 5.2.4). Personenbezogene Daten müssen so nicht geteilt werden, wodurch die datenschutzrechtlichen Anforderungen an die Weitergabe reduziert werden.

Ausführlicher zu einzelnen Datenschutz-Regelungen siehe Tabelle 34 in Abschnitt A5.3.2.

5.4.3 Marktzugang I: Medizinprodukte

Der Marktzugang ist in verschiedenen Regelwerken normiert. Grundlegend ist die Verordnung (EU) 2017/745 über Medizinprodukte (Medical Device Regulation, MDR) [95]; daneben tritt insbesondere der AI Act [6].

Diagnostische Software ist als Medizinprodukt einzustufen. Je nach Patientenrisiko fällt sie in der Regel in die Risikoklassen IIa-III (vergleiche Art. 2 Nrn. 1 und 12 MDR sowie die Leitlinien der Medical Device Coordination Group (MDCG) [96], Regel 11). AI-Systeme, die unter die MDR fallen, gelten nach dem AI Act automatisch als „Hochrisiko AI-Medizinprodukte“ und müssten ab dem 02.08.2027 sowohl die MDR-, als auch die AI Act-Anforderungen erfüllen. Ab Risikoklasse IIa ist eine benannte Stelle für die CE-Konformitätsbewertung erforderlich.

Adaptive, also weiterlernende AI-Medizinprodukte sind nach MDR schwer handhabbar. Zwar sind solche Systeme grundsätzlich zulassungsfähig, aber jede wesentliche Funktion zur Änderung durch Weiterlernen würde nach MDR eine erneute Bewertung erfordern – gegebenenfalls mehrmals täglich. Der Grund dafür ist, dass Software „wiederholbar zuverlässig und vorhersehbar“ funktionieren muss, Anhang I Kapitel II Nummer 17.1 MDR. Ab dem 02.08.2026 schafft der AI Act eine Ausnahme für adaptive AI-Medizinprodukte, wenn das Weiterlernen vorab definiert und bewertet wurde, Art. 43 Abs. 4 AI Act.

Ausführlicher zu einzelnen Marktzugangs-Regelungen siehe Tabelle 34 in Abschnitt A5.3.2.

5.4.4 Klinische Studien

Die in die Nutzenbewertung eingeschlossenen Studien entsprechen soweit ersichtlich den ethischen und rechtlichen Standards zum Schutz der Rechte von Patientinnen und Patienten als Studienteilnehmerinnen und Studienteilnehmern.

5.4.5 Kostenerstattung im öffentlichen Gesundheitssystem

GKV-Versicherte haben einen Anspruch auf eine bedarfsgerechte wissenschaftlich fundierte Behandlung, §§ 2, 12 SGB V [97]. Welche Leistungen übernommen werden, legt der GB-A in Richtlinien fest. Mit der Änderung der Krebsfrüherkennungsrichtlinie (KFE-RL) [74] am 05.09.2025 wird künftig (auch) die AI-unterstützte Auswertung von Röntgen- oder CT-Bildern im Rahmen der neuen Lungenkrebsfrüherkennung mittels Low-Dose-CT bei Rauchern zulasten der GKV erbracht werden, § 39 Abs. 5 KFE-RL [74] mit Verweis auf die LuKrFrühErkV [5]. Die LuKrFrühErkV legt einen Behandlungsablauf fest, nach dem Bilder zunächst von einer Radiologin oder einem Radiologen auszuwerten sind und im Anschluss mithilfe einer Software. Bei auffälligem Ergebnis hat derselbe Ablauf mit einer weiteren Radiologin oder einem weiteren Radiologen zu erfolgen. An die Qualifikation der beteiligten Ärztinnen und Ärzte werden explizit Aus-, Weiter- und Fortbildungsanforderungen gestellt.

Ausführlicher zu einzelnen Regelungen der Kostenerstattung siehe in Tabelle 34 in Abschnitt A5.3.2.

5.4.6 Exkurs: Gibt es ein Recht auf eine AI-unterstützte Auswertung von Röntgen- oder CT-Aufnahmen der Lunge? Selbstbestimmungsrecht des Patienten vs. ärztliche Therapiefreiheit

Unabhängig vom Leistungsanspruch gesetzlich versicherter Personen gegenüber der GKV (dazu siehe Abschnitt 5.4.5) bestimmt sich das Verhältnis zwischen Ärztinnen sowie Ärzten und Patientinnen sowie Patienten nach den allgemeinen Regeln des Behandlungsrechts. Einschlägig sind insbesondere die Vorschriften des BGB über den Behandlungsvertrag sowie die verfassungsrechtlich gewährleiteten Rechte der Beteiligten. Das grundrechtlich in Art. 2 Abs. 2 S. 1 GG und zivilrechtlich in § 630d Abs. 1 S. 1 BGB abgesicherte Selbstbestimmungsrecht von Patientinnen und Patienten überlässt es deren Willen zu entscheiden, ob sie eine medizinische Behandlung in Anspruch nehmen wollen und wenn ja welche; daraus folgt aber kein Anspruch auf jede mögliche medizinische Maßnahme [98]. Dem steht die ebenfalls grundrechtlich in Art. 12 Abs. 1 GG abgesicherte ärztliche Therapiefreiheit entgegen, die Ärztinnen und Ärzten einen Beurteilungsspielraum eröffnet und auch die Wahl der Diagnostik umfasst. Ärztinnen und Ärzte müssen jedoch nach den allgemein anerkannten fachlichen Standards behandeln, § 630a Abs. 2 BGB. Ärztinnen und Ärzte schulden nur dann neueste Therapiekonzepte wie zum Beispiel die Methode „AI-System“, wenn sie „risikoärmer sind und / oder bessere Heilungschancen versprechen, in der medizinischen Wissenschaft im Wesentlichen unumstritten sind und deshalb nur ihre Anwendung von einem sorgfältigen und auf Weiterbildung bedachten Arzt verantwortet werden kann“ [98].

5.4.7 Spezielle medizinische Bereiche

Zentrale Anforderungen an Hochrisiko-AI-Systeme nach dem AI Act (nicht abschließend)

- Nach dem AI Act müssen AI-Systeme so gestaltet sein, dass eine menschliche Aufsicht möglich ist und die finale Entscheidung bei Menschen liegt, Art. 14 AI Act.
- Anbieter sind verpflichtet, ein Risikomanagementsystem einzuführen, um Risiken für Sicherheit, Gesundheit, Grundrechte und Umwelt systematisch zu identifizieren, zu bewerten und zu mindern, Art. 9 AI Act.
- Die für Training, Validierung und Tests genutzten Datensätze müssen qualitativ hochwertig, repräsentativ, aktuell und frei von Verzerrungen sein, Art. 10 AI Act. Außerdem ist eine detaillierte technische Dokumentation erforderlich, die Aufbau, Funktion, Risiko, Management, Test und Ergebnisse beschreibt, Artikel 11 AI Act.
- Anwender müssen über Zweck, Funktionsweise und Grenzen das AI-System informiert werden, Art. 13 AI Act.

- Nach Inverkehrbringen müssen Anbieter ein Monitoringsystem einrichten, um Risiken durch Nutzung, Weiterlernen oder Updates zu erkennen und zu mindern, Art. 61 AI Act.
- Nach Art. 43 AI Act müssen Hochrisiko-AI Systeme eine Konformitätsbewertung durchlaufen.
- Es besteht eine Meldepflicht für Vorfälle und Sicherheitsprobleme und die Pflicht, erforderliche Korrekturmaßnahmen einzuleiten können, Art. 62 AI Act.

5.4.7.1 Arztvorbehalt

Nach Art. 14 AI Act dürfen AI-Systeme nur unterstützend eingesetzt werden, die endgültige Diagnoseentscheidung bleibt ärztliche Aufgabe („human in the loop“). Der Einsatz von AI-Systemen ist damit unproblematisch, solange ärztliches Handeln das System „beherrscht“, indem es „die durch das System ausgeführten Schritte vorgibt, plant, auslost, steuert und überwacht sowie die Ergebnisse kontrolliert“ [90]. Wenn ein pathologischer Befund übersehen oder nicht gemeldet wird, trägt damit die behandelnde Radiologin oder der behandelnde Radiologe die Verantwortung, sofern sie oder er hätte erkennen können, dass der Befund des AI-Systems falsch ist [90].

5.4.7.2 Pflicht zur Überprüfung von AI-Ergebnissen

Nach § 630a BGB sind Ärztinnen und Ärzte verpflichtet, Patientinnen und Patienten die bestmögliche Behandlung nach den allgemein anerkannten fachlichen Standards anzubieten. Dazu gehört auch die kritische Überprüfung der Ergebnisse der von ihnen eingesetzten technischen Instrumente [47]. Daraus folgt – wie auch bei der Verletzung des Arztvorbehalts – dass „Fehlentscheidungen“ der AI als Behandlungsfehler gelten können, wenn die ärztliche Überprüfungspflicht verletzt wurde.

5.4.7.3 Haftung

„Nichts auf der Welt ist perfekt, weder Menschen noch KI, und irgendwann wird das Modell einen Fehler machen. Was tun wir mit diesem Fehler? Wer trägt die Schuld und die letztendliche Verantwortung für diesen Fehler? Niemand macht einen Algorithmus dafür verantwortlich.“ (Bowers 2024 [57])

Die Bundesärztekammer fordert, dass klar geregelt sein müsse „bei welchen Fehlern der Hersteller und bei welchen Fehlern die Ärztin oder der Arzt beziehungsweise die medizinische Einrichtung haftbar ist“ [99]. Die Verteilung der Haftung dient im Allgemeinen dazu, Schäden aus rechtswidrigem oder fahrlässigem Handeln auszugleichen und Verantwortliche zur Rechenschaft zu ziehen. Die Arzthaftung soll sicherstellen, dass Ärztinnen und Ärzte sorgfältig nach anerkannten Standards behandeln. Die Produkthaftung dient dem Schutz von Verbrauchern vor Schäden durch fehlerhafte Produkte und zieht Hersteller für daraus

entstehende Schäden zur Verantwortung. Bislang gibt es keine speziellen Vorschriften zur Haftung für Schäden infolge des Einsatzes von AI-Systemen [55].

Für den vorliegenden Bereich werden 3 Haftungskonstellationen vorgestellt: 1. es entsteht ein Schaden beim Einsatz von AI; 2. es entsteht ein Schaden wegen Nichteinsatzes von AI und 3. es entsteht ein Schaden aufgrund fehlerhafter AI.

Zu 1. Es entsteht ein Schaden beim Einsatz von AI – Arzthaftung

Beim Einsatz von AI-Systemen haften die Ärztin oder der Arzt oder der (Krankenhaus-)Träger auf Grundlage des Behandlungsvertrages, §§ 630a, 280 BGB oder deliktisch nach § 823 ff. BGB. Arzthaftung ist stets Haftung für Pflichtverletzung und tritt ein, wenn Sorgfaltspflichten verletzt werden [90]. Ärztinnen und Ärzte unterliegen strengen Sorgfaltspflichten über den gesamten Behandlungsprozess. Beim Einsatz technischer Geräte, insbesondere von AI-Systemen, müssen Ärztinnen und Ärzte sicherstellen, dass diese wissenschaftlich anerkannt, regelmäßig gewartet, korrekt bedient und überwacht werden. Allerdings müssen sie keine Garantie für das fehlerfreie Funktionieren der eingesetzten medizinisch-technischen Apparate übernehmen: „Der Mediziner ist kein Techniker im Arztkittel“ [55]. Sehr wohl aber gehören zu den haftungsauslösenden Sorgfaltspflichtverletzungen zum Beispiel eine mangelhafte Aufklärung (dazu siehe oben unter „Patientenautonomie I“), eine fehlende Überprüfung von AI-Ergebnissen (siehe dazu oben unter „Arztvorbehalt“ und „Pflicht zur Überprüfung von AI-Ergebnissen“), der Einsatz nicht zertifizierter Medizinprodukte (dazu siehe oben unter „Marktzugang I“) oder Fehler bei Inbetriebnahme und Anwendung des AI-Systems [52].

Zu 2. Schaden wegen Nichteinsatzes von AI

Solange ein AI-System als neue Methode oder sogenannte Neulandmethode gilt, kann es die Ärztin oder der Arzt im Rahmen seiner Therapiefreiheit einsetzen, ist dazu aber nicht verpflichtet. Der Nichteinsatz ist dann nicht haftungsbegründend. Sobald sich eine Methode zur Standardtherapie entwickelt, bestimmt der Sorgfaltsstandard, ob sie eingesetzt werden muss. In diesem Fall ist zu prüfen, ob das Nicht-Verwenden der Methode ein haftungsbegründendes Zurückbleiben hinter dem geschuldeten Sorgfaltsstandard ist. Es besteht aber keine Pflicht, jede neue Therapie oder technische Ausstattung einzusetzen. Wenn sich eine Neulandmethode etabliert hat und so entscheidende Vorteile bietet, dass Patientinnen und Patienten darüber informiert werden müssen, um entscheiden zu können, ob sie die Behandlungen nach dem neuesten Standard wünschen, kann allerdings eine Aufklärungspflicht gegenüber Patientinnen und Patienten bestehen (vgl. [52,90]), die bei Nichtbeachtung haftungsbegründend wirken könnte.

Zu 3. Es entsteht ein Schaden beim Einsatz von AI – Produkthaftung

Beim Einsatz von AI-Systemen haftet der Hersteller nach dem Produkthaftungsgesetz (ProdHaftG) [100], wenn wegen eines Produktfehlers Personen- oder Sachschäden entstehen,

§ 1 ProdHaftG. Bislang war streitig, ob AI-Systeme als Produkt im Sinne des § 2 ProdHaftG gelten. Die EU-Produkthaftungsrichtlinie [101] von 2024 (ProdHaftRL) hat das nun klargestellt und definiert in Art. 4 Abs. 1 Nr. 1 S. 2 Software ausdrücklich als Produkt. Die Umsetzung in deutsches Recht soll bis Ende 2026 erfolgen. Der Referentenentwurf eines Gesetzes zur Modernisierung des Produkthaftungsrechts [102] nimmt die Einordnung von Software als Produkt in § 2 Abs. 1 Nr. 3 ProdHaftG (Referentenentwurf) auf. Ein Produktionsfehler liegt nach dem ProdHaftG vor, wenn das Produkt nicht die Sicherheit bietet, die berechtigterweise erwartet werden kann. Dazu gehören Fabrikationsfehler, Konstruktionsfehler oder Instruktionsfehler. Bei AI-Systemen können Fehler im Algorithmus, ein fehlerhaft designtes künstliches neuronales Netz, unvollständige Trainingsdaten oder mangelhafte Instruktionen / Dokumentation oder Sicherheitsmängel Produktmängel sein [55]. Nach Art. 7 Abs. 2 lit. c ProdHaftRL sind bei der Beurteilung der Fehlerhaftigkeit alle Umstände zu berücksichtigen, insbesondere auch die adaptive Lernfähigkeit von AI oder die Möglichkeit, dass das Produkt nach dem Inverkehrbringen neue Funktionen erwirbt.

5.4.8 Ausblick

Die relevanten Regelungen für die vorliegende Fragestellung sind komplex. Sie reichen vom Patientenrechtegesetz, §§ 630a ff BGB, über europäisches und nationales Datenschutzrecht, Marktzugangsrecht und spezifische europäische Vorgaben für AI-Systeme bis hin zu europäischem und nationalen Haftungsrecht. Vergleicht man die Anforderungen, die der Deutsche Ethikrat an die Implementierung von AI-Systemen stellt [47], mit den rechtlichen Vorgaben, kommt man zu diesem Ergebnis:

- 1) Qualitäts-Sicherung bei der Entwicklung: angelegt in MDR, MPBetreibV, AI Act
- 2) Datenrepräsentativität: angelegt in AI Act Art. 10, ProdHaftG
- 3) Automation-Bias-Vorsorge: angelegt in §§ 630a, 280 Abs. 1 BGB, Haftung wegen Sorgfaltspflichtverletzung
- 4) Datenschutz & Privatsphäre: angelegt in DGSVO, BDSG, EHDS, GDNG, DigiG
- 5) gerechter Zugang: für die Früherkennung angelegt in LuKrFrühErkV
- 6) Ausbildung & Weiterbildung: für die Früherkennung angelegt in LuKrFrühErkV
- 7) Sicherstellung hoher methodischer Expertise / Vermeidung von Kompetenzverlust: für die Früherkennung angelegt in LuKrFrühErkV
- 8) Kontextsensitivität: Aufklärungspflicht geregelt im Patientenrechtegesetz, Schutz vor Personalabbau mit Befundungsalgorithmus in der LuKrFrühErkV angelegt, Schutz des Arzt-Patienten-Vertrauensverhältnisses und der sprechenden Medizin rechtlich nicht geregelt

- 9) keine vollständige Ersetzung von ärztlichen Fachkräften: angelegt in AI Act Art. 14 (human in the loop), LuKrFrühErkV

Die ethischen Vorgaben sind somit weitgehend in rechtliche Vorgaben gefasst; wie die Umsetzung in der Praxis erfolgt, bleibt abzuwarten.

5.5 Ergebnisse zu organisatorischen Aspekten

Für die Auswertung der organisatorischen Aspekte wurden insgesamt 27 Publikationen eingeschlossen (vgl. Abschnitt A7.3.4). Die Informationsaufbereitung zu organisatorischen Aspekten erfolgte entlang des von Perleth (2014) [22] vorgeschlagenen Rasters zur Einschätzung der organisatorischen Folgen von Behandlungsmethoden (vgl. Tabelle 35).

Die Analyse zu organisatorischen Aspekten beruht auf den Ergebnissen der Betroffenenbefragungen, den eingeschlossenen Publikationen und „reflective thoughts“ [19]. Die folgende Zusammenfassung der in Tabelle 35 dargestellten Aspekte fokussiert auf den Einfluss der neuen Technologie einer AI-unterstützten Befundung der Lunge bei Verdacht auf Lungenkrebs bzw. im Rahmen einer Lungenkrebsfrüherkennung auf die Voraussetzungen und die Prozesse der Leistungserbringung.

Einfluss auf Voraussetzungen der Leistungserbringung

Änderung des Ortes der medizinischen Versorgung

Stand Dezember 2025 erfolgen weder die Befundung der Lunge bei Personen mit Verdacht auf Lungenkrebs noch die Lungenkrebsfrüherkennung regelhaft mit AI-Unterstützung. Die Befundung bei Personen mit Verdacht auf Lungenkrebs erfolgt durch Radiologinnen und Radiologen in einer ambulanten oder stationären radiologischen Einrichtung. Wenn wie geplant ab April 2026 die Lungenkrebsfrüherkennung gemäß LuKrFrühErkV Leistung der GKV wird [74,103], ist es wahrscheinlich, dass die Leistung, wie im Jahr 2021 vom Bundesamt für Strahlenschutz vorgeschlagen, primär in zertifizierten Lungenkrebszentren angeboten wird [63,104,105]. Dies scheint angesichts der fachlichen und technischen Anforderungen, die die LuKrFrühErkV an die Software-gestützte Früherkennung stellt, ein realistischer Ort der Versorgung. An dieser Stelle wird darauf hingewiesen, dass in der Begründung zur LuKrFrühErkV explizit davon ausgegangen wird, dass eine Software-gestützte CT-Untersuchung eingesetzt wird. Nicht beschrieben in der LuKrFrühErkV ist, welche Software oder auf welchen Algorithmen die Software beruhen sollte, wohingegen das Bundesministerium für Umwelt zwischen CAD- und AI-unterstützten Systemen unterscheidet [5].

Änderung bei den Qualifikationsanforderungen an Leistungserbringende

Die LuKrFrühErkV spezifiziert die Qualifikationsanforderungen an Leistungserbringende bei der Durchführung der Früherkennung [5]. Diese Anforderungen sind in Tabelle 35 zusammengefasst. Die LuKrFrühErkV benennt die Anforderungen übergreifend bezogen auf den Einsatz

CAD-basierter Systeme. Der Einsatz AI-unterstützter Systeme wird nicht explizit thematisiert. Sowohl nationale als auch internationale Fachgesellschaften, Gremien oder Experten betonen, dass der Einsatz von AI in der Lungenkrebsfrüherkennung in die Aus-, Weiter- und Fortbildungs-Curricula von Radiologinnen und Radiologen integriert werden muss [47,74,88,105-107]. In verschiedenen Stellungnahmen wird als wichtig herausgestellt, dass in der Ausbildung sowohl AI-spezifische als auch ausreichend AI-unabhängige Kenntnisse und Fertigkeiten vermittelt werden müssen, um so einem möglichen schleichenden Verlust von Erfahrungswissen, dem sogenannten De-Skilling, vorzubeugen [88,99]. Nur so kann gemäß verschiedener Veröffentlichungen unerfahrenen Leistungserbringenden der Aufbau von Erfahrungswissen auch weiterhin ermöglicht und vermieden werden, dass diese sich in, während und nach ihrer Ausbildung bei der Befundung möglicherweise zu sehr auf eine AI verlassen oder deren Ergebnisse aufgrund mangelnder Erfahrung falsch interpretieren [48,73,88,99,106]. Eine Standardisierung der Anforderungen an die Qualifizierung von Leistungserbringenden für die Früherkennung nimmt beispielsweise die European Society of Thoracic Imaging (ESTI) mit der „Lung Cancer Screening certification“ vor [86,108]. Inwiefern die AI-unterstützte Früherkennung hier integriert werden könnte, ist unklar. Qualitative Studien zeigen, dass neben der Einbindung von AI-Systemen zur Unterstützung der Befundung in die Aus-, Weiter- und Fortbildungscurricula von Radiologinnen und Radiologen auch die Einarbeitung und Schulung von Leistungserbringern auf eine spezifische AI gewährleistet sein muss, um die Akzeptanz und das Vertrauen der Leistungserbringenden in die AI zu fördern und diese auch dauerhaft erfolgreich in ein Radiologiesetting zu implementieren [48,73]. Während es sich hierbei folglich um einen wichtigen Teil der Implementierung neuer AI-Systeme handelt, wird die konkrete Ausgestaltung von Einführungsmodulen oder Schulungen bislang in nur etwa 40 % der Studien zu AI-Systemen überhaupt thematisiert [50,58].

Änderungen bei den Anforderungen an Personal, Material und Organisation der Leistungserbringung (Strukturqualität)

Personal

Bei der Befundung von Personen mit Verdacht auf Lungenkrebs sind durch den Einsatz einer AI grundsätzlich personelle Einsparungen möglich, wenn zum Beispiel Radiologinnen oder Radiologen die AI anstelle einer Kollegin oder eines Kollegen als „zweite Meinung“ hinzuziehen [48,73]. Die Befundung von Personen mit Verdacht auf Lungenkrebs muss in Deutschland durch eine Fachärztin oder einen Facharzt erfolgen. Eine Verifizierung des fachärztlichen Befundes ist nicht erforderlich. Ein personelles Einsparpotenzial durch das Ersetzen einer Zweitbefunderin oder eines Zweitbefunders durch eine AI ist daher in Deutschland bei Personen mit Verdacht auf Lungenkrebs nicht gegeben. Das Hinzuziehen einer AI zur Unterstützung der Befundung könnte zu falschen Diagnosen führen, wenn die AI zum Beispiel missverständliche Ergebnisse liefert, die die Ärztin oder der Arzt fehlinterpretiert. Eine weitere Befürchtung ist ein langfristiger Abbau von Erfahrungswissen (De-Skilling), da

dann möglicherweise der Austausch mit einem erfahreneren Kliniker oder einer Klinikerin wegfällt [48,88,99,107]. Da die LuKrFrühErkV klare Vorgaben an den Einsatz von Personal (Erst- und Zweitbefundung) stellt, ist auch hier kein personelles Einsparpotenzial durch eine AI-Unterstützung gegeben [5].

Material und Organisation der Leistungserbringung

Die Anforderungen an das Material zur Früherkennung, d. h. das jeweilige AI-System, und zur Organisation der Leistungserbringung bleiben in der LuKrFrühErkV zumindest in Teilen und besonders bezogen auf den Einsatz AI-unterstützter Systeme unklar [5]. Anforderungen an die Trainings- und Nutzungsdaten von AI-Systemen wurden bereits durch deutsche, europäische und internationale Vereinbarungen und rechtliche Vorgaben umfassend beschrieben bzw. definiert ([53], vgl. Abschnitt 5.4). Inwiefern die Transparenz und Nachvollziehbarkeit der Daten eines AI-Systems tatsächlich auch für die Leistungserbringenden (Radiologiezentren, Krankenhäuser oder einzelne Radiologinnen und Radiologen) gegeben sind, bleibt im Einzelfall zu prüfen. Zur Prüfung der Qualität von AI-Systemen wurden in Positionspapieren und wissenschaftlichen Publikationen Anforderungen definiert und Checklisten entwickelt [109-112]. Diese können Leistungserbringende, die den Einsatz eines AI-Systems in Erwägung ziehen, möglicherweise helfen, die Qualität einer AI und die Implikationen der Einführung eines solchen Systems strukturierter einzuschätzen.

Die Anforderungen an das Material und die Organisation der Erbringung einer AI-unterstützten Befundung oder Früherkennung in radiologischen Einrichtungen werden zudem beispielsweise in der Publikation der ESTI von Gleeson et al. 2023 [88] strukturiert aufbereitet (vgl. Tabelle 35). Gleeson et al. 2023 betonen, dass die Einführung einer AI in die bestehende Organisationsstruktur einer radiologischen Einrichtung mit erheblichem finanziellem, personellem und organisatorischem Aufwand einhergeht. Gleeson et al. 2023 heben die folgenden material- und organisationsbezogenen Herausforderungen besonders hervor, die auch verschiedene andere Publikationen als hinderliche Faktoren bei der Implementierung einer AI-unterstützten Befundung identifizieren [113,114]:

- Hohe Kosten der AI (Anschaffungs- und laufende Kosten, Kosten für Personal Aus- und Weiterbildung, Kosten für neue IT-Systeme etc.)
- Integration der AI in bestehende Prozesse, Arbeitsabläufe und besonders IT-Systeme (als besonders hinderlich wird eine unzureichende / mangelnde Kompatibilität zwischen AI und zentrumseigenem IT-System berichtet)
- Einfluss der AI auf Prozesse

Auch die mangelnde Kompatibilität zwischen zentrumseigenen AI-Systemen könnte in der AI-unterstützten Befundung bei einem Wechsel von Patientinnen und Patienten zwischen Zentren zu Problemen führen.

Qualitative Erhebungen bzw. Querschnittstudien weisen weitergehend darauf hin, dass die Integration einer AI in bestehende Prozesse, Arbeitsabläufe und IT-Systeme ohne zusätzliche Zeitverzögerungen und ohne Zusatz-Aufwand für die Radiologen und Radiologinnen eine Grundvoraussetzung dafür ist, dass AI von diesen längerfristig angenommen wird und sie Vertrauen in die AI aufbauen können [48,73]. Bestenfalls könnte eine AI zu zeitlichen und organisatorischen Einsparungen führen, wenn sie Abläufe beschleunigt oder beispielsweise die Dokumentationspflichten oder die Erstellung von Befundbriefen des Radiologen oder der Radiologin vorbereitet oder unterstützt. Im ungünstigen Fall kann eine AI zusätzlichen Aufwand durch einen erhöhten ärztlichen Kontrollbedarf oder durch eine nicht funktionierende Integration in die bestehende IT-Infrastruktur bewirken.

Befragungen legen zwar nahe, dass eine AI in verschiedenen internationalen Settings den Leistungserbringenden die Arbeit erleichtern und die Befundung möglicherweise schneller / effizienter gestalten und sogar zu Personaleinsparungen führen kann (AI als Zweitbefunderin) [48,73]. Belegt ist eine solche Effizienzsteigerung durch aussagekräftige Studien bislang jedoch nicht [84,113,115] (vgl. auch Abschnitt 3.6.3 zum ergänzenden Endpunkt Zeitersparnis). Bislang liegen weder für einen patientenrelevanten Nutzen oder Schaden einer AI-unterstützten Befundung bei Verdacht auf Lungenkrebs bzw. zur Früherkennung im Vergleich zur Befundung ohne AI aussagekräftige Studien vor (vgl. Abschnitt 3.7), noch konnten Kosten- oder Kosten-Wirksamkeitsstudien basierend auf aussagekräftigen Nutzen-Daten bisher nachweisen, dass AI im Lungenkrebscreening an einzelnen Stellen oder übergreifend zu Einsparungen führt bzw. kosten-effektiv ist (vgl. Kapitel 4). Es bleibt auch offen, ob die Einführung und die Nutzung einer AI-unterstützten Befundung bzw. Früherkennung die Prozesse der Leistungserbringung eher behindert bzw. verlangsamt oder längerfristig diese zu beschleunigen vermag.

Kommunikations- und Kooperationsformen

Zur Vermeidung insbesondere des beschriebenen De-Skilling-Problems wird vorgeschlagen, dass eine Befundung mit AI bestenfalls durch erfahrene und unerfahrene Kliniker gemeinsam innerhalb eines „mixed-expertise“-Ansatzes erfolgt. Studien haben diese Vorgehensweise eingesetzt bzw. untermauern deren Sinnhaftigkeit [116,117]. So können weniger erfahrene Radiologinnen und Radiologen in die AI-basierte Befundung eingebunden werden und Erfahrungswissen aufgebaut werden.

Für die Ansprache und die Kommunikation mit den Anspruchsberechtigten einer Lungenkrebsfrüherkennung gibt die LuKrFrühErkV vor, dass diese durch Fachärztinnen und Fachärzte der Inneren Medizin, der Allgemeinmedizin, oder der Arbeitsmedizin im Rahmen eines partizipativen Prozesses erfolgen soll. Zudem sollen den anspruchsberechtigten Personen evidenzbasierte Informationen zu Nutzen und Schaden der Lungenkrebsfrüherkennung in einer Versicherteninformation bereitgestellt werden [5,74,87,118]. Es ist unklar,

inwiefern die Aufklärungs- und Informationspflichten zum Lungenkrebscreening wie in der LuKrFrühErkV definiert sich auch auf die AI-unterstützte Befundung und auf die Befundung von Personen mit Verdacht auf Lungenkrebs erstrecken werden (vgl. Abschnitt 5.4).

Weitere Aspekte

Weitere organisatorische Aspekte beziehen sich auf die unterschiedlichen Interessenslagen der am Prozess der beteiligten Personen und Institutionen. Diese Interessenslagen können ausschlaggebend dafür sein, wie gut oder schlecht sich eine Innovation wie die AI-unterstützte Befundung bei Verdacht auf Lungenkrebs bzw. Früherkennung in einem Gesundheitssystem implementieren lassen. Bezüglich der Interessen der Betroffenen und der behandelnden Leistungserbringenden sei an dieser Stelle auf Tabelle 35 sowie die Ausführungen zu sozialen Aspekten in Abschnitt 5.3 verwiesen.

Aus Perspektive des Gesundheitssystems bleibt offen, ob und wo die Befundung bei Verdacht auf Lungenkrebs bzw. zur Lungenkrebsfrüherkennung mit oder ohne AI-Unterstützung angeboten wird [5]. Festzuhalten ist, dass die Früherkennung basierend auf den Angaben der LuKrFrühErkV und gemäß der Vorgaben des Gemeinsamen Bundesausschuss (G-BA) ab April 2026 den anspruchsberechtigten GKV-versicherten Personen als Kassenleistung zur Verfügung stehen soll [74,103,118].

Die Umsetzung der AI-unterstützten Befundung / Früherkennung wird angesichts der finanziellen, personellen und organisatorischen Herausforderungen wahrscheinlich am ehesten in zertifizierten Lungenkrebszentren erfolgen. Gleeson et al. 2023 fassen die Anforderungen an die Planung von Kapazitäten und Investitionen, die für die Einführung von AI-Systemen erforderlich sind, zusammen (vgl. Tabelle 35) [88]. Zusammenfassend bleibt festzuhalten, dass die Einführung einer AI-unterstützten Befundung mit hohen Kosten und großem Aufwand verbundenen sein wird; aussagekräftige Evidenz zu deren Nutzen oder Schaden im Vergleich zu einer ausschließlichen Befundung durch Radiologinnen und Radiologen fehlt jedoch bislang.

5.6 Ergebnisse zu Umwelt- und Klimaaspekten

Für die Auswertung der Umwelt- und Klimaaspekte wurden insgesamt 10 Publikationen eingeschlossen (vgl. A7.3.5). Zudem wurden „reflective thoughts“ der Berichtsautorinnen und -autoren [19] in die Analyse integriert. In der Befragung von Betroffenen wurden Umwelt- oder Klimaaspekte der AI-unterstützten Befundung der Lunge nicht adressiert.

Die Aufarbeitung der Umwelt- und Klimaaspekte befasst sich mit der Emission von Treibhausgasen, der Freisetzung toxischer Substanzen bzw. der Strahlenexposition, des Abfallmanagements sowie weiteren Aspekten der neuen Technologie (vgl. Tabelle 36).

Die für die Bewertung der Umwelt- und Klimaaspekte eingeschlossenen Studien befassen sich mit dem Einfluss von AI-Software in der Radiologie und hierbei besonders mit AI-Software die auf weiterentwickelten Algorithmen aufbaut: Deep-Learning- und neuronale Netzwerkmodelle verbrauchen dabei im Vergleich zu einfacheren Modellen deutlich mehr Ressourcen und Energie und verursachen einen größeren CO₂-Ausstoß, weil ihr Training größere Datenmengen und spezialisierte Rechenhardware erfordern [119].

Emission von Treibhausgasen

Nach einem Bericht der Organisation „Healthcare without Harm“ aus dem Jahr 2019 ist der Gesundheitssektor insgesamt für etwa 4,4 % der globalen Treibhausgas-Emissionen verantwortlich [120]. Für die Vereinigten Staaten liegen Schätzungen vor, dass etwa 8 bis 10 % der CO₂-Emissionen durch den Gesundheitssektor verursacht werden und dass etwa 1 % der weltweiten CO₂-Emissionen durch die medizinische Bildgebung entstehen [121]. Die Radiologie scheint folglich auch ohne die AI bereits für einen Teil von Treibhausgas-Emissionen verantwortlich zu sein, beispielsweise durch die Produktion und den Einsatz hochtechnologischer, energieintensiver Geräte. Mit dem Einsatz von AI in den verschiedenen radiologischen Anwendungen erhöht sich dieser Anteil [122]. Die Mehrheit der in der Gesundheitsversorgung zugelassenen und verwendeten medizinischen Geräte, die eine AI nutzen, sind radiologische Geräte wie die CT oder andere bildgebende Verfahren [123-125]. [84]

Zusätzliche CO₂-Emissionen beim Einsatz von AI-Systemen in der Radiologie entstehen durch stromintensive Datenserver einschließlich notwendiger Kühlsysteme, auf denen AI-generierte Daten cloudbasiert gelagert und verarbeitet werden [84,126,127]. Schätzungen gehen davon aus, dass solche Datenserver etwa 1 bis 2 % des weltweiten Stromverbrauchs ausmachen, mit steigender Tendenz. Und auch die Produktion von AI-Hardware sowie der Transport von Rohstoffen oder Komponenten für diese Hardware in die jeweiligen Herstellerländer geht mit Treibhausgas-Emissionen einher. Lösungsansätze, um die Treibhausgas-Emissionen in Zusammenhang mit AI zu reduzieren, sind zum Beispiel die Nutzung von überwiegend oder ausschließlich erneuerbaren und nicht mehr fossilen Energieträgern, die bislang noch vielfach eingesetzt werden. So könnten beispielsweise bei Planung und Bau neuer Rechenzentren oder Datenserver der gleichzeitige Bau neuer Solar- oder Windparks berücksichtigt werden, oder die Nähe zu solchen Parks bei der Standortwahl berücksichtigt werden [84].

Toxische Substanzen und Strahlenexposition

Für die Produktion der für AI-Systeme notwendigen Datenserver bzw. Hochleistungs-Hardware (wie Grafikprozessoren, GPUs, oder Central Processing Units, CPUs) werden toxische Stoffe wie Blei, Cadmium und Quecksilber benötigt [84,123]. Durch den schneller werdenden technologischen Fortschritt werden Hardware- bzw. Datenserver-Komponenten zudem immer schneller durch innovativere Produkte ersetzt, was die Nachfrage nach diesen

Substanzen erhöht und deren Umlauf und damit auch die Abfallproduktion beschleunigt. Außerdem erfordert die Herstellung von AI-Hardware den Abbau und die Veredelung verschiedener kritischer, d. h. seltener oder schwer zu beschaffender Rohstoffe, was wiederum mit erheblichen CO₂-Emissionen einhergeht [84,123]. Zu diesen kritischen Rohstoffen gehören seltene Erden wie Neodym oder Übergangs-Metalle wie Tantal oder Kobalt. Der Umgang mit den für AI-Systeme notwendigen toxischen Substanzen und die eingeschränkte Verfügbarkeit kritischer Substanzen impliziert Belastungen und Herausforderungen für Mensch und Umwelt.

Abfallmanagement

Der beschriebene Einsatz toxischer oder kritischer Substanzen für die Herstellung von AI-Hardware führt nach Ablauf von deren Betriebsdauer zu erheblichen Herausforderungen beim Abfallmanagement. Es entsteht toxischer Elektroschrott, sogenannter e-waste, der eine sachgemäße Entsorgung erfordert. Eine solche Entsorgung ist aufwändig und kostenintensiv [84,123]. Auf der anderen Seite gibt es zunehmende Bestrebungen, seltene Erden oder bestimmte Metalle zu recyceln.

Weitere

Eine weitere Implikation von AI-Systemen allgemein für Umwelt und Klima ist der hohe Wasserverbrauch für die Kühlung der für die AI benötigten Hochleistungsrechner und Datenserver. Schätzungen zufolge könnte der weltweite Einsatz von AI- Anwendungen bis 2027 einen Wasserverbrauch von 4,2 bis 6,6 Milliarden Kubikmetern verursachen. Dies entspricht dem 4- bis 6-fachen des jährlichen Wasserverbrauchs Dänemarks oder der Hälfte des Verbrauchs Großbritanniens [128]. Als Lösungsansatz für dieses Problem wird beispielsweise die Lokalisierung von großen Rechenzentren in kühlen Regionen oder in der Nähe großer Wasservorkommen besprochen [84], was jedoch wiederum beispielsweise zu einer unerwünschten Erwärmung von Wasserreserven führen könnte. Nicht zuletzt bleiben auch die Externalitäten eines Abbaus von seltenen Erden oder anderer Rohstoffe eine Herausforderung: der Abbau dieser Rohstoffe führt zur Zerstörung von Lebensräumen für Mensch und Tier, zum Verlust von biologischer Vielfalt und zu Wasser- und Umweltverschmutzung in den jeweiligen Abbau-Gebieten [84].

In Reaktion auf die zunehmend sichtbaren Auswirkungen von AI-Anwendungen in der Radiologie auf Klima und Umwelt werden Strategien für eine nachhaltige AI erforscht und diskutiert [84]. So können beispielsweise die AI-Modelle selbst effizienter werden und damit weniger CO₂ produzieren (z. B. durch Pruning, d. h. das Verkleinern von Modellen durch das Entfernen unwichtiger Komponenten, oder anderer Techniken wie einer Datenaugmentation), wenig effiziente Hardware kann durch „grüne“ Alternativen ersetzt werden (z. B. durch Tensor Processing Units, TPUs) oder die Versorgung aus fossilen Energieträgern kann durch erneuerbare Energien ersetzt werden. Auch beim Einsatz toxischer oder kritischer

Rohstoffe oder im Abfallmanagement werden alternative Vorgehensweisen erforscht, wie der Ersatz toxischer und kritischer durch weniger toxische und nicht kritische Rohstoffe oder das Recycling solcher Stoffe. Diese Ansätze zur Steigerung der Umweltverträglichkeit und Nachhaltigkeit von AI-Systemen sind wichtig und werden immer wichtiger, je mehr AI-Systeme in verschiedensten Bereichen, auch in der Radiologie, zum Einsatz kommen.

6 Domänenübergreifende Zusammenführung der Ergebnisse

Kontext

Initiales logisches Modell für T24-04: Lungenkrebs: Kann Künstliche Intelligenz (AI) die Auswertung von Röntgen- oder CT-Bildern der Lunge unterstützen?

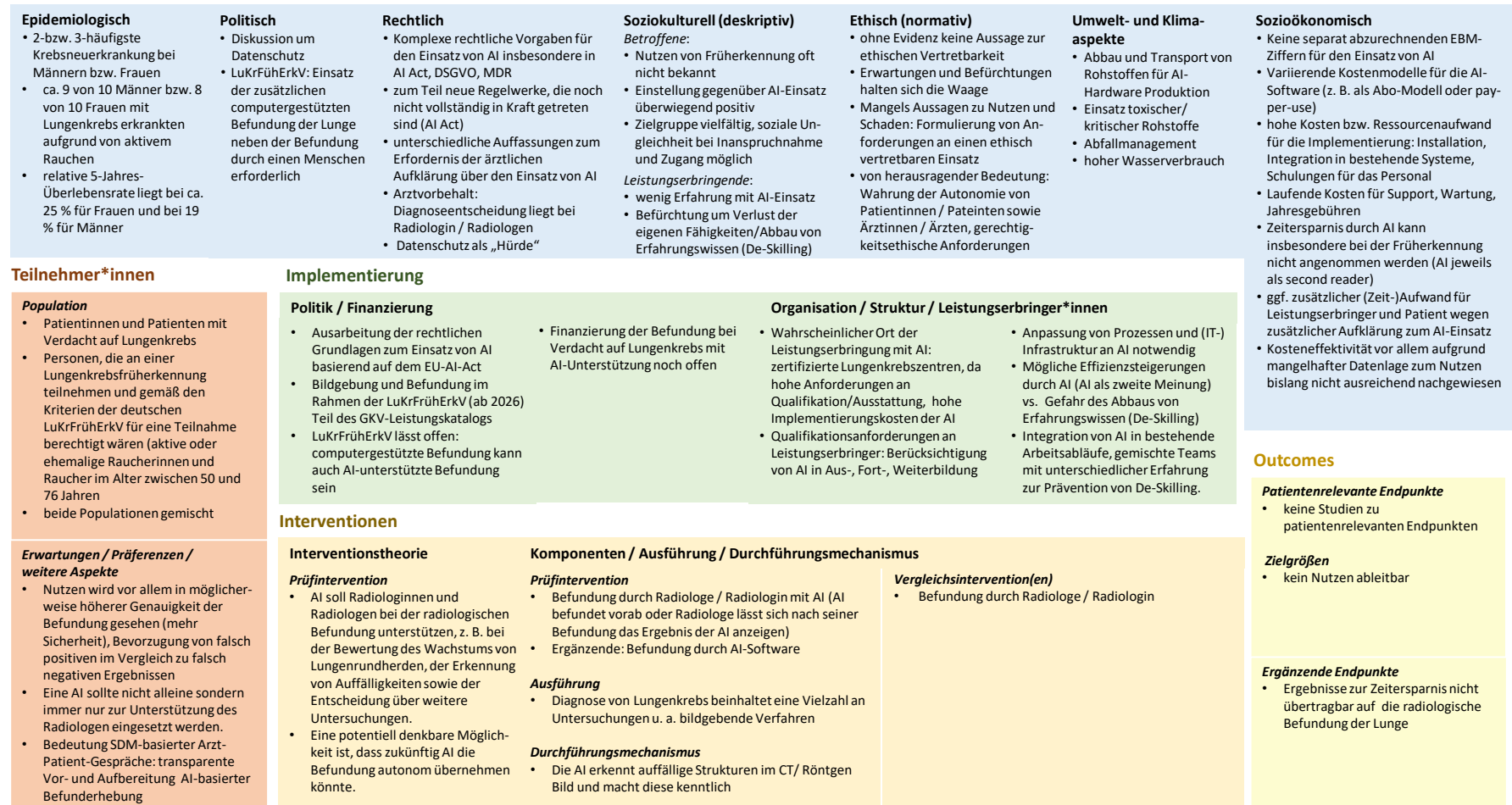


Abbildung 1: Zusammenfassung der Ergebnisse des vorliegenden ThemenCheck-Berichts als logisches Modell in Anlehnung an INTEGRATE-HTA [129]

7 Diskussion

7.1 Einordnung der Arbeitsergebnisse

Hinsichtlich der Frage nach dem Nutzen und Schaden einer AI-unterstützten Befundung in der Lungenkrebsfrüherkennung liegen 2 Testgütestudien mit geringer qualitativer Ergebnissicherheit vor. Grundsätzlich ist davon auszugehen, dass jedes Screening durch falsche Screeningbefunde und Überdiagnosen schadet. Ein Screening ist nur gerechtfertigt, wenn der Schaden durch den Nutzen mehr als aufgewogen wird. Die Konsequenzen eines falsch-negativen Screeningbefundes können dazu führen, dass eine weitere Diagnostik und eine notwendige Behandlung nicht oder nur verzögert stattfinden. Für ein Lungenkrebscreening wird dieses Risiko auf Patientenebene als ein bedeutsameres Risiko wahrgenommen als ein falsch-positiver Befund (siehe Abschnitt 5.2.2). Doch grundsätzlich stellt auch eine unnötige Abklärungsdiagnostik und potenziell damit einhergehende Komplikationen einen möglichen Schaden des Screenings dar. Wie in Abschnitt 3.6.1 aufgeführt, ergeben sich bei einer angenommenen Lungenkrebsprävalenz von 1,2 % bzw. 3,1 % zahlreiche falsch-positive Befunde, bei denen zur Abklärung eine erneute Untersuchung, wie eine CT, erfolgt oder, in selteneren Fällen, eine Biopsie. Dabei ist – bei geringer qualitativer Ergebnissicherheit – der errechnete PPV in der Gruppe mit dem AI-unterstützten Screening jeweils niedriger als in der Gruppe der alleinigen radiologischen Befundung und abhängig von der geschätzten Prävalenz (niedrigerer PPV bei niedrigerer Prävalenz). Das Screening soll in Deutschland ab April 2026 eingeführt werden [103]. Auch für Deutschland lässt sich eine Lungenkrebs-Prävalenz im Bereich von 1 % bis 3 % erwarten [130,131].

In einem Positionspapier zu Anforderungen an die Qualitätssicherung von AI-Modellen für die Lungenkrebsfrüherkennung wird gefordert, mit standardisierten Trainings- und Validierungsdaten zu arbeiten, die neben verschiedenen CT-Protokollen und Gerätetypen auch eine repräsentative Anzahl unterschiedlicher Läsionen und Varianten der Bildqualität abdecken [111]. Beim Vorliegen unterschiedlicher Trainingsdaten ist die diagnostische Güte verschiedener AI-Systeme nur bedingt vergleichbar. So zeigte eine Übersichtsarbeit für den Vergleich 7 verschiedener AI-Systeme bei der alleinigen Auswertung von Lungenrundherden durch die jeweilige AI-Software an Röntgenbildern Sensitivitäten zwischen 50 % und 89 % und Spezifitäten zwischen 80 % und 99 % [132]. Hinzu kommt, dass bei einigen Geräten die Sensitivität und Spezifität vom Anwender eingestellt werden können, indem sie Hinweise auf auffällige Befunde z. B. erst ab einer bestimmten Größe zulassen (z. B. [133]). Eine SÜ zu Studien, die AI in der Diagnose medizinischer Bilder einsetzen, berichtet, dass nur in 6 % der eingeschlossenen Studien eine externe Validierung des AI-Algorithmus erfolgte. Die betrachteten Studien stammten mehrheitlich (70,9 %) aus dem Bereich der Radiologie [134].

Neben der Überlegung, welche AI eine hohe diagnostische Güte aufweist und für das Lungenkrebscreening Anwendung finden könnte, ist auch der rechtliche und

organisatorische Aufwand von Relevanz. Die Ergebnisse für den Endpunkt Zeitersparnis wurden ergänzend für die beiden in der Nutzenbewertung eingeschlossenen Studien zur diagnostischen Güte für Fragestellung 2 (Screeningpopulation) dargestellt (siehe Abschnitt 3.6.3), die Übertragbarkeit der Ergebnisse ist jedoch als niedrig einzustufen. Die LuKrFrühErkV [10] gibt vor, dass im Rahmen der Früherkennung die CT-Aufnahme zunächst ohne und erst anschließend unter Zuhilfenahme einer Software befundet werden soll. Sofern der Erstbefunder oder die Erstbefunderin zu einem auffälligen Befund kommt, muss die CT-Aufnahme durch einen Zweitbefunder oder eine Zweitbefunderin ausgewertet werden, welche die Aufnahme wiederum zunächst allein und im Anschluss mit einer Software befundet. Aufgrund dieser Festlegung, in jedem Fall die AI-unterstützte Befundung zusätzlich durchzuführen, ist im Rahmen des Lungenkrebsfrüherkennung insgesamt ein höherer Zeitaufwand unvermeidlich, sodass die diagnostische Güte für die AI-unterstützte Befundung höher sein muss, um insgesamt von einem Vorteil ausgehen zu können.

Bei Patientinnen und Patienten mit Verdacht auf Lungenkrebs ist das Vorgehen in der Befundung aktuell nicht in einer Rechtsverordnung festgelegt, sodass eine Zeitersparnis bei der Befundung mit AI-Unterstützung denkbar wäre. Hierbei ist zu berücksichtigen, dass die hypothetischen Einsatzmöglichkeiten der AI in der Befundung vielfältig sind. So wurde in den eingeschlossenen Studien die AI als Assistent eingesetzt: die Befundung wurde mit AI-Unterstützung durchgeführt bzw. die zuerst nur durch den Menschen befundeten CT-Bilder wurden anschließend von derselben Person mit AI-Unterstützung befundet. Lee 2025 berichtet jedoch noch weitere mögliche Szenarien zur Nutzung von AI in der Befundung: Die AI als Vorabscreener, bei der im weiteren Verlauf nur die laut AI testpositiven Befunde durch Radiologinnen und Radiologen überprüft werden; die AI als Vorabscreener und Assistent, bei dem nur die durch die AI als testpositiv befundeten Aufnahmen anschließend mit AI-Unterstützung durch Radiologinnen und Radiologen geprüft werden; die AI als Back-up, bei der nur die diskrepanten Fälle zwischen menschlicher und AI-Befundung, die jeweils unabhängig voneinander erfolgt, durch denselben oder einen anderen Menschen erneut überprüft werden. Auch ein Einsatz einer AI zur alleinigen Befundung wird in Lee 2025 berichtet. In Studien, die den Einschlusskriterien der Nutzenbewertung nicht entsprachen, wurde zudem der Einsatz einer AI als zweiter Reader – die AI überprüft die Befundung der Radiologin oder des Radiologen und die markierten Diskrepanzen werden von einer weiteren Radiologin oder einem weiteren Radiologen überprüft – berichtet [135,136]. So sind die Anwendungsmöglichkeiten einer AI-Software grundsätzlich vielfältig, entsprechen jedoch nicht alle den aktuell geltenden rechtlichen Vorgaben innerhalb der Europäischen Union, nach denen z. B. die Befundentscheidung immer eine menschliche Entscheidung sein muss. (Abschnitt 5.4.7.1).

7.2 ThemenCheck-Bericht im Vergleich zu anderen Publikationen

Im Rahmen der systematischen Literaturrecherche konnten 6 Referenzen zu 4 SÜs identifiziert werden, die eine AI-unterstützte Befundung bei Röntgen- oder CT-Aufnahmen der Lunge betrachten (siehe Abschnitt A7.1.1). In keiner SÜ wurde eine RCT eingeschlossen, die die Einschlusskriterien der Nutzenbewertung dieses ThemenCheck-Berichts erfüllt.

Die Fragestellung der SÜ Colquitt 2024 [137] entspricht der der Fragestellung 1 der vorliegenden Nutzenbewertung, wobei sich die SÜ auf Studien, die die alleinige bzw. AI-unterstützte Befundung von Röntgenbildern betrachten, beschränkt. Für die Population der Patientinnen und Patienten mit Verdacht auf Lungenkrebs konnten in dieser SÜ ebenso wie in der vorliegenden Nutzenbewertung keine Studien identifiziert werden. In einer ergänzenden Darstellung wurden 6 retrospektive Kohortenstudien beschrieben, in denen eine unklare Patientenpopulation eingeschlossen wurde. Nur in 1 Studie, ohne den für die Nutzenbewertung geforderten Referenztest, wurde die diagnostische Güte anhand der identifizierten Lungenkrebsfälle dargestellt. Die Autoren berichten eine statistisch signifikant höhere Sensitivität und eine niedrigere Spezifität (nicht statistisch signifikant) für die AI-unterstützte Befundung im Vergleich zu einer alleinigen radiologischen Befundung. Die Signifikanzaussage ließ sich mit eigenen Berechnungen nicht bestätigen.

Die SÜ Geppert 2025 [34] betrachtet die diagnostische Güte und die Kosteneffektivität einer AI-unterstützten Befundung von CT-Aufnahmen im Vergleich zu einer nicht AI-unterstützten Befundung. Die Population in dieser SÜ war breiter gefasst, umfasste aber auch die beiden Populationen dieses ThemenCheck-Berichts. Zur Bewertung der diagnostischen Güte wurden in der SÜ 27 Studien eingeschlossen, welche mehrheitlich nicht der Fragestellung des ThemenCheck-Berichts entsprachen: Es wurden auch Studien, die die diagnostische Güte zur Erkennung von Lungenrundherden und nicht zur Erkennung von Lungenkrebs bewerteten, eingeschlossen sowie Studien mit alleiniger AI-Befundung. Auch wenn die Fragestellung der SÜ umfassender war, kamen die Studienautorinnen und -autoren ebenfalls zu dem Schluss, dass die aktuell bestehende Evidenz für die Beantwortung der Fragestellungen nicht ausreichend sei. Interessant ist hierbei, dass Geppert 2024 [138] abschätzen, dass die AI-gestützte Befundung aufgrund einer etwas höheren Sensitivität, bei aber etwas geringerer Spezifität im Vergleich zur Diagnostik ohne AI dazu führt, dass mehr als 100 Personen einen auffälligen Befund erhalten würden, um 1 zusätzlichen Fall von Lungenkrebs feststellen zu können. In der SÜ wurde keine geeignete gesundheitsökonomische Evaluation identifiziert, weshalb eine de-novo-Modellierung erfolgte, deren Ergebnisse zwar robust, aber aufgrund der verfügbaren Inputparameter als hochgradig unsicher einzustufen sind.

Der Studienpool in den SÜs Ewals 2023 [50] und Li 2021 [139] unterscheidet sich deutlich von dem der vorliegenden Nutzenbewertung und ist weder klar der Fragestellung 1 noch der Fragestellung 2 zuzuordnen. Es erfolgte keine Einschränkung auf eine Patientenpopulation, in

Li 2021 wurden neben Lungenrundherden bzw. Lungenkrebs unter anderem Studien an Patientinnen und Patienten mit auffälligem Befund (wie COVID-19, Tuberkulose oder interstitielle Lungenerkrankung) eingeschlossen. Auch entspricht der Referenzstandard in den eingeschlossenen Studien mehrheitlich nicht dem in der Nutzenbewertung geforderten Referenzstandard. Eine Gegenüberstellung der Ergebnisse aus den beiden SÜs im Vergleich zum vorliegenden Bericht erscheint daher nicht sinnvoll.

7.3 ThemenCheck-Bericht im Vergleich zu Leitlinien

Die aktuelle deutsche S3-Leitlinie „Prävention, Diagnostik, Therapie und Nachsorge des Lungenkarzinoms“ [4] umfasst Empfehlungen sowohl zur Versorgung von Patientinnen und Patienten mit Verdacht auf Lungenkrebs als zur Früherkennung von Lungenkrebs teilnehmen. Im Vereinigten Königreich gibt es 2 getrennte Leitlinien zur radiologischen Befundung mit AI-Unterstützung getrennt nach bildgebenden Verfahren: Röntgen und CT [42,49].

Patientinnen und Patienten mit Verdacht auf Lungenkrebs (Fragestellung 1)

Der Einsatz einer Software zur computerassistierten Detektion bei Patientinnen und Patienten mit Verdacht auf Lungenkrebs wird in der S3-Leitlinie nicht adressiert [4]. In der englischen Leitlinie wird die weitere Evidenzgenerierung für die AI-unterstützte Befundung von Röntgenbildern bei Patientinnen und Patienten mit Verdacht empfohlen [49]. Zusätzlich empfiehlt die Leitlinie zu Low-Dose-CT die AI-unterstützte Befundung für Patientinnen und Patienten mit Verdacht lediglich im Rahmen der Forschung zu nutzen [42]. In diesem ThemenCheck-Bericht war keine Nutzenbewertung für diese Population möglich, da keine Studien identifiziert wurden. Aufgrund der fehlenden Evidenz können auch die Kosteneffektivität und die ethischen Aspekte nicht abschließend bewertet werden.

Personen, die an einer Lungenkrebsfrüherkennung teilnehmen (Fragestellung 2)

Eines der Ergebnisse der organisatorischen Aspekte dieses ThemenCheck-Berichts ist, dass die Einführung einer AI-unterstützten Befundung oder Früherkennung mit erheblichem organisations- und prozessbezogenem Aufwand und Herausforderungen verbunden ist (s. Abschnitt 5.5). Die S3-Leitlinie [4] bietet mit ihren Empfehlungen einen ersten Rahmen für eine einheitliche deutschlandweite Umsetzung der verbindlichen Vorgaben der LuKrFrühErkV zum Einbezug einer Software zur computerassistierten Detektion: Zum Beispiel empfiehlt die S3-Leitlinie die Verwendung einer qualitätsgesicherten und für die Erst- und Zweitbefundung identischen Software zur computerassistierten Detektion, Volumetrie und Berechnung der Volumenverdopplungszeit. Zusätzlich soll für die Befundung die modifizierte Lung-RADS Klassifikation von 2022 mit integrierter Volumenverdopplungszeit (V-Lung-RADS 2022) verwendet werden. Nur eine in der Nutzenbewertung eingeschlossenen Studie, Lee 2025, verwendete Lung-RADS als Grundlage für die Befundung von Lungenrundherden.

Um die epidemiologische Qualitätssicherung zur gewährleisten empfiehlt die S3-Leitlinie u. a. die anonymisierte Erhebung zur Anzahl der diagnostizierten Tumortypen und -stadien als auch für alle Teilnehmenden Angaben zum Zeitintervall zwischen dem Aufnahmedatum der Low-Dose-CT und dem endgültigen Low-Dose-CT-Befund und bei abklärungsbedürftigem Befund bis zur der interdisziplinären Fallkonferenz [4]. Um Studien zur diagnostischen Güte für eine Nutzenbewertung heranziehen zu können, ist die Dokumentation solcher Charakteristika auch in klinischen Studien erforderlich um eine Aussage zum Nutzen der AI-unterstützten Befundung treffen zu können, vor allem wenn deutlich höhere Krebsdetektionsraten auftreten (siehe Abschnitt A2.1). Die in die Nutzenbewertung eingeschlossenen Studien berichten solche Charakteristika jedoch nicht.

Die Leitlinie aus dem Vereinigten Königreich zur radiologischen Befundung von CT-Aufnahmen beschreibt das gesundheitsökonomische Potenzial der AI-unterstützte Befundung für Personen, die an einer Lungenkrebsfrüherkennung teilnehmen. Gleichzeitig wird betont, dass weitere Evidenzgenerierung notwendig ist, um Aussagen bezüglich der kosteneffektivsten und klinisch relevantesten Software zu ermöglichen [42]. In diesem ThemenCheck-Bericht konnte auch für die Population der zweiten Fragestellung keine Aussage zum Nutzen und Schaden getroffen werden. Ebenfalls können aufgrund der fehlenden Evidenz bzw. eingeschränkten Übertragbarkeit auch die Kosteneffektivität und die ethischen Aspekte nicht abschließend bewertet werden.

7.4 Kritische Reflexion des Vorgehens

Dieser ThemenCheck-Bericht untersuchte als Prüfintervention bzw. als Indextest 1 die AI-unterstützte radiologische Befundung von Röntgen- oder CT-Aufnahme. Da keine verbindliche Definition von AI vorliegt, wurden für die Nutzenbewertung und für die gesundheitsökonomische Bewertung spezifische Anforderungen an die AI gestellt:

- Die AI unterstützt Radiologinnen und Radiologen, in dem sie abnormale Strukturen erkennt (CADE) und / oder abnormale Strukturen diagnostiziert (CADx).
- Es werden AI-Anwendungen eingeschlossen, bei denen es sich um CE-gekennzeichnete Medizinprodukte handelt.

Diese Anforderungen an die AI ermöglichten einen breiten Einschluss von Software jeglicher Art und sind im Einklang mit den Vorgaben der LuKrFrüErkV. Die eingeschlossenen Studien und SÜs zur diagnostischen Güte untersuchten jedoch ausschließlich AI-Systeme, die auf modernen Ansätzen des maschinellen Lernens aufbauen. Trotz der breiten Einschlusskriterien zu Softwarelösungen wurden keine klinischen Studien mit patientenrelevanten Endpunkten identifiziert. Es wurden zwar zahlreiche Studien zur diagnostischen Güte von AI im Rahmen der radiologischen Befundung durch die Recherche gefunden, allerdings untersuchten viele dieser Studien AI allein oder eine AI ohne CE-Kennzeichnung (siehe Abschnitt A7.1.2 nicht D2a

bzw. nicht D2b). Weiterhin verglich eine Vielzahl von Studien den Indextest 1 mit dem Expertenkonsensus, nicht aber zum definierten Referenztest dieses ThemenCheck-Berichts (siehe Abschnitt A7.1.2 nicht D4a bzw. nicht D4b). Im Rahmen des Linked-Evidence-Ansatzes ist jedoch für Aussagen zur diagnostischen Güte von Lungenkrebs ein Vergleich zu Biopsie oder Nachbeobachtung als sicherer diagnostischer Referenztest erforderlich. Für die Bewertung der ethischen und rechtlichen Aspekte wurde die Definition von AI-Systemen des AI Acts herangezogen. Diese Definition schließt, die in der Nutzenbewertung und in der gesundheitsökonomischen Bewertung berücksichtigten Systeme mit CE-Kennzeichnung und den spezifischen Anwendungen (CADe und CADx) ein. Bei der Betrachtung der Umwelt- und Klimaaspekte wurde eine engere Definition von AI in den eingeschlossenen Studien zugrunde gelegt, da komplexe Modelle wie maschinelles Lernen mehr Ressourcen verbrauchen als bisherige Softwarelösungen.

8 Schlussfolgerung

Dieser ThemenCheck-Bericht untersuchte bei Patientinnen und Patienten mit Verdacht auf Lungenkrebs und bei Personen, die an einem Lungenkrebsfrüherkennungsscreening teilnehmen, die Frage eines Bürgers, ob Ärztinnen und Ärzte mit AI-Unterstützung Röntgen- oder CT-Bilder ebenso gut oder besser auswerten können als ohne AI und welche Folgen dies für die Versorgung haben könnte. Die Ergebnisse dieses Themen-Check-Berichts zu beiden Aspekten dieser Frage, werden im Folgenden unter

- Auswirkung der AI-Unterstützung auf Nutzen, Schaden und Kosteneffektivität der radiologischen Befundung und
- Folgen der Implementierung einer AI-unterstützten Diagnostik

zusammengefasst:

Auswirkung der AI-Unterstützung auf Nutzen, Schaden und Kosteneffektivität der radiologischen Befundung

Durch die umfassende Informationsbeschaffung wurde deutlich, dass keine Studien zur diagnostisch-therapeutischen-Kette vorliegen, die eine direkte Aussage zur Auswirkung einer AI-unterstützten Diagnostik auf patientenrelevante Endpunkte ermöglichen würde. Zudem zeigte sich, dass Studien die Anwendung einer AI-unterstützten Befundung häufig nur dahingehend untersuchten, wie gut die AI-gestützte Diagnostik verdächtige Lungenrundherde erkennen konnte. Die Studien haben jedoch häufig nicht verifiziert, ob die Personen mit verdächtigen Lungenrundherden tatsächlich Lungenkrebs hatten. Es wurden 2 aussagekräftige Studien zur diagnostischen Güte identifiziert, die im Rahmen des Linked-Evidence-Ansatzes eine Untersuchung zu Nutzen und Schaden ermöglichen. In diesen Studien wurden nur Personen eingeschlossen, die an einer Lungenkrebsfrüherkennungsuntersuchung teilnahmen (Fragestellung 2). Patientinnen und Patienten mit Verdacht auf Lungenkrebs (Fragestellung 1) wurden in diesen Studien nicht untersucht. Für Fragestellung 1 liegt somit keine Evidenz vor. Die klinische Heterogenität der beiden Studien ließ keine metaanalytische Zusammenfassung der Ergebnisse zu. Aus den Ergebnissen der Einzelstudien zur diagnostischen Güte einer Screeningrunde bzw. einer einmaligen Diagnostik lässt sich aufgrund der geringen qualitativen Ergebnissicherheit kein Nutzen oder Schaden der AI-unterstützten Befundung im Vergleich zu einer alleinigen radiologischen Befundung ableiten. Insgesamt ist die Datenlage zur Nutzenbewertung der Prüfinerventionen für beide Fragestellungen nicht ausreichend.

Es wurde keine gesundheitsökonomische Evaluation zur AI-unterstützten Befundung von Röntgen-Bildern bei Patientinnen und Patienten mit Verdacht auf Lungenkrebs identifiziert (Fragestellung 1). Für die AI-unterstützte Befundung von CT-Bildern hingegen schloss dieser ThemenCheck-Bericht 1 gesundheitsökonomische Evaluation mit separat ausgewiesenen

Ergebnissen für beide Populationen ein. Zwar zeigte das Modell für sich genommen robuste Ergebnisse für eine gesundheitsökonomische Überlegenheit der AI-Unterstützung in der Screeningpopulation bzw. für eine gesundheitsökonomische Unterlegenheit der AI-Unterstützung in der Verdachtspopulation. Dabei sind jedoch die unzureichende Datenlage hinsichtlich der vorliegenden Nutzenparameter sowie Einschränkungen hinsichtlich der Übertragbarkeit zu bedenken. Daher reichen die Ergebnisse der eingeschlossenen Modellierungsstudie in diesem ThemenCheck-Bericht nicht für eine abschließende Bewertung der Kosteneffektivität aus.

Neben der fehlenden Evidenz kommt ergänzend dazu, dass unterschiedliche AI-Softwares aktuell nicht vergleichbar sind (Abschnitt 7.1). Somit ist auf Basis der wenigen eingeschlossenen Studien zu einzelnen AI-Produkten eine allgemeingültige Aussage zum Nutzen und zur Kosteneffektivität der AI-unterstützten Befundung nicht ableitbar.

Zusammenfassend kann dieser ThemenCheck-Bericht aufgrund fehlender und unzureichender Evidenz die Frage, ob mit einer AI-unterstützten Befundung Betroffene von Lungenkrebs im Rahmen einer Verdachts- oder einer Früherkennungsuntersuchung genauso gut oder besser erkannt werden, nicht beantworten.

Voraussetzungen und Folgen der Implementierung einer AI-unterstützten Diagnostik

Der Einsatz einer AI-Unterstützung bei der Befundung kann sich in verschiedenster Weise auf eine mögliche Implementierung in die Versorgung, Gesellschaft und Umwelt auswirken. Wesentliche Aspekte, welche in diesem ThemenCheck-Bericht in mehreren Domänen betrachtet wurden, werden im Folgenden beschrieben.

Kosten

Die Bewertung der ökonomischen als auch organisatorischen Aspekte zeigt, dass in Deutschland vor allem für die Leistungserbringenden zusätzliche Kosten entstehen: für Leistungserbringende liegen die zusätzlichen Kosten für die AI-Nutzung bei der Befundung von Röntgen- beziehungsweise CT-Bildern schätzungsweise in einem niedrigen 1-stelligen Euro-Bereich pro Befundung. Darüber hinaus müssen Leistungserbringende mit weiterem Ressourcenaufwand für die Implementierung einer AI-unterstützten Befundung rechnen, die u. a. aufgrund der Integration der AI in bestehende Systeme und Prozesse sowie der Schulung von Personal entstehen. Für die gesetzlichen Krankenkassen sowie die gesetzlich versicherten Personen sind in der Regel aktuell keine zusätzlichen direkten Kosten durch die AI-Unterstützung zu erwarten, da bislang spezifische Abrechnungsmöglichkeiten in den aktuellen Vergütungskatalogen nicht enthalten sind.

Zeitliche Effizienz

Es wurden sowohl positive als auch negative Erwartungen zu möglichen Änderungen in der zeitlichen Effizienz identifiziert. In der Nutzenbewertung waren die Ergebnisse zum ergänzenden Endpunkt Zeitersparnis, welche auf einer Screeningpopulation basieren, nicht aussagekräftig, da in beiden Studien nur ein Teilaspekt der radiologischen Befundung betrachtet wurde. Im Rahmen der Nutzenbewertung lag keine Evidenz für Patientinnen und Patienten mit Verdacht auf Lungenkrebs vor. Auch in der gesundheitsökonomischen Bewertung lag keine allgemeingültige Evidenz bezogen auf die Zeitersparnis im Gesamtprozess vor. In Deutschland ist für die Lungenkrebsfrüherkennung durch die aktuelle gesetzliche Regelung keine Zeitersparnis möglich, da Radiologinnen und Radiologen, die Befundung erst allein und anschließend nochmals mit Software-Unterstützung durchführen sollen. Die Bewertung der organisatorischen Aspekte kommt damit übereinstimmend zu dem Schluss, dass im Rahmen der Früherkennung wegen der Vorgaben der Lungenkrebs-Früherkennungs-Verordnung kein personelles Einsparpotenzial zu erkennen ist. Ob das auch für die Befundung von Personen mit Verdacht auf Lungenkrebs gilt, für die derzeit keine Vorgaben für den Einsatz von AI-Systemen bestehen, ist unklar. Basierend auf den Ergebnissen dieses ThemenCheck-Berichts ist vor allem in der Implementierungsphase einer AI-unterstützten Befundung nicht von einer zeitlichen Effizienz für Leistungserbringende oder für Patientinnen und Patienten und zu screenende Personen auszugehen.

Weitere Aspekte

Aus ethischer Sicht ist insbesondere die Wahrung der Autonomie von Patientinnen und Patienten sowie von Ärztinnen und Ärzten von herausragender Bedeutung. Daneben bestehen hohe gerechtigkeitsethische Anforderungen, sei es in Hinblick auf das Training von AI-Systemen unter Berücksichtigung auch unterrepräsentierter sozialer Gruppen als auch in Hinblick auf einen gerechten Zugang – sofern der Nutzen von AI-Systemen hinreichend belegt ist.

Rechtlich ist die Anwendung von AI-Systemen zur Unterstützung bei der Befundung von Röntgen- und CT Bildern durch ein vielschichtiges Regelungsgefüge geprägt, das Fragen der Aufklärung, des Datenschutzes, des Marktzugangs und der Haftung umfasst. Von hervorzuhebender Bedeutung ist, dass die Verantwortung für Befundentscheidungen stets bei den menschlichen Akteuren verbleibt.

Auf Seiten der Betroffenen und der Radiologinnen und Radiologen werden in Befragungsstudien zwar viele positive Erwartungen hinsichtlich des Nutzens einer AI-Unterstützung und unterschiedliche Erfahrungen mit der Nutzung einer spezifischen AI zur unterstützten Befundung durch Radiologinnen und Radiologen berichtet. Gleichzeitig adressieren verschiedene Befragungsstudien, Stellungnahmen und Positionspapiere aber auch soziale und organisatorische Herausforderungen bei der Umsetzung einer AI-

unterstützten Befundung. Risiken werden insbesondere im Hinblick auf mögliche soziale Ungleichheiten im Zugang zu oder bei der Inanspruchnahme von AI-Leistungen gesehen. Organisatorisch besteht zudem das Risiko von De-Skilling-Effekten, die durch qualifikations-sichernde Maßnahmen und angepasste Arbeitsabläufe möglicherweise jedoch minimiert werden können. Ergänzend machen die erheblichen Umwelt- und Klimaauswirkungen der AI-Systeme – von Energie- und Wasserverbrauch bis hin zur Ressourcenabbau und Abfallmanagement – deutlich, dass auch ökologische Nachhaltigkeit in Entscheidungs- und Implementierungsprozessen systematisch berücksichtigt werden muss.

Schlussendlich ist ein Einsatz einer AI-unterstützten radiologischen Befundung erst dann vertretbar, wenn der patientenrelevante Nutzen, etwa die Verringerung von Mortalität und Morbidität sowie Verbesserung der gesundheitsbezogenen Lebensqualität, den Schaden, beispielsweise durch Fehl- oder Überdiagnosen für Patientinnen und Patienten oder für die untersuchten Personen überwiegt. Der vorliegende ThemenCheck-Bericht liefert allerdings mangels ausreichender Datenlage für keine der beiden Fragestellungen eine Nutzen- oder Schadenaussage. Unklar ist somit, ob die AI-Unterstützung bei der Auswertung von Röntgen- oder CT-Bildern mit einem patientenrelevanten Nutzen einhergeht. Die eingeschränkte Vergleichbarkeit zwischen AI-Anwendungen, die zu erwartenden Unterschiede in der diagnostischen Güte sowie die unterschiedlichen Anwendungsmöglichkeiten der AI verdeutlichen die Notwendigkeit weiterer Evidenzgenerierung für die Fragestellungen des Berichts: insbesondere sind Studien zur diagnostisch-therapeutischen Kette notwendig, um die Konsequenzen für die Therapie und die patientenrelevanten Endpunkte abzuschätzen. Des Weiteren sollten Studien zur diagnostischen Güte prospektiv geplant werden und eine ausreichende Teilnehmeranzahl umfassen, um zu aussagekräftigen Ergebnissen zu führen (siehe auch Einschlusskriterien der Fragestellung). Die aktuelle Attraktivität von AI-Systemen alleine, die in allen Lebensbereichen zu beobachten ist, kann vor diesem Hintergrund keine ausreichende Begründung für ihren routinemäßigen Einsatz bei der radiologischen Befundung von Lungenkrebs darstellen, da dies von der aktuellen Datenlage nicht gestützt wird.

ThemenCheck-Details

A1 Projektverlauf

A1.1 Zeitlicher Verlauf des Projekts

Aus den im Vorschlagszeitraum von August 2022 bis Juli 2023 beim ThemenCheck Medizin eingereichten Themenvorschlägen wurde vom IQWiG im Jahr 2024 unter Beteiligung eines mit Patientinnen und Patienten sowie Bürgerinnen und Bürgern besetzten Auswahlbeirats das Thema „Lungenkrebs: Kann Künstliche Intelligenz die Auswertung von Röntgen- oder CT-Bildern der Lunge unterstützen?“ für die Erstellung eines ThemenCheck-Berichts mit der Projektnummer T24-04 ausgewählt.

Die Erstellung des ThemenCheck-Berichts gliedert sich in die folgenden Schritte:

Für die Bearbeitung der Fragestellung zu ethischen, sozialen, rechtlichen und organisatorischen Aspekten sowie den Klima- und Umweltaspekten wurden vom IQWiG externe Sachverständige beauftragt. Die weiteren Inhalte wurden durch das IQWiG erstellt.

Am 03.07.2025, 07.07.2025, 08.07.2025 und 14.07.2025 wurden im Rahmen der Projektbearbeitung Betroffene konsultiert.

Zunächst wurde ein ThemenCheck-Berichtsprotokoll erstellt, das in der Version 1.0 vom 04.08.2025 am 07.08.2025 auf der Website des IQWiG veröffentlicht wurde.

Auf Basis des Berichtsprotokolls wurde der vorliegende vorläufige ThemenCheck-Bericht durch das IQWiG zusammen mit den externen Sachverständigen unter Anwendung der Methodik des IQWiG erstellt. Vor der Veröffentlichung wurde ein Review des vorläufigen Berichts durch eine nicht projektbeteiligte Person durchgeführt. Der vorläufige ThemenCheck-Bericht wird zur Anhörung gestellt.

Im Anschluss an diese Anhörung erstellt das IQWiG zusammen mit den externen Sachverständigen den ThemenCheck-Bericht.

Der abschließende ThemenCheck-Bericht sowie eine allgemein verständliche Version (ThemenCheck kompakt: Das Wichtigste verständlich erklärt) werden an den Gemeinsamen Bundesausschuss (G-BA) und das Bundesministerium für Gesundheit (BMG) übermittelt und 2 Wochen später auf der Website des IQWiG veröffentlicht. An gleicher Stelle wird auch die Dokumentation der Anhörung zum vorläufigen ThemenCheck-Bericht veröffentlicht.

Dieser ThemenCheck-Bericht wurde in der International HTA Database [140] registriert.

A1.2 Spezifizierungen und Änderungen im Projektverlauf

Vorläufiger ThemenCheck-Bericht im Vergleich zum ThemenCheck-Berichtsprotokoll

Neben redaktionellen Änderungen ergaben sich folgende Spezifizierungen oder Änderungen im vorläufigen ThemenCheck-Bericht:

- Spezifizierung zu Abschnitt 1.2: Computerprogramme mit AI wurden spezifiziert.
- Spezifizierung zu Abschnitt A2.1.2.3: Die Studie Lo 2018 wurde eingeschlossen, auch wenn aus den Angaben der Studie keine personenbezogene Vierfeldertafel ableitbar war. Die Autorinnen und Autoren berichten die Anzahl der gescreenten Personen und die Anzahl der Referenztestpositiven und Referenztestnegativen sowie Sensitivität, Spezifität, PPV und NPV aus einer multireader-multicase Analyse nach Dorfman et al. [12].
- Spezifizierung zu Abschnitt A2.1.2.3: Ergebnisse zum ergänzenden Endpunkt Zeitersparnis wurde auch aus Studien zur diagnostischen Güte dargestellt.
- Spezifizierung zu Abschnitt A2.3.2: Die Informationsaufbereitung zu ethischen Aspekten erfolgte unter Berücksichtigung der übergeordneten Fragestellungen des vereinfachten Fragenkatalogs von Hofmann [17] und orientierte sich darüber hinaus an dem Rahmengerüst für die ethische Bewertung von Public-Health-Maßnahmen nach Marckmann [18].

A2 Details der Methoden – Methodik gemäß ThemenCheck-Berichtsprotokoll

Die folgenden Abschnitte geben den Wortlaut der Berichtsmethodik aus dem ThemenCheck-Berichtsprotokoll wieder. Über diese Methodik hinausgehende Spezifizierungen oder Änderungen der Methoden im Projektverlauf werden im Abschnitt A1.2 erläutert. Im folgenden Text wird an den entsprechenden Stellen auf diesen Abschnitt verwiesen.

Dieser ThemenCheck-Bericht wird auf Grundlage der Allgemeinen Methoden 7.0 [141] erstellt.

A2.1 Nutzenbewertung

Der Nutzen von diagnostischen Verfahren lässt sich anhand von randomisierten Interventionsstudien der gesamten diagnostisch-therapeutischen Behandlungskette bewerten (siehe Abschnitt A2.1.1.4). Daher wird, obwohl solche Studien insbesondere für die Fragestellung 2 (Lungenkrebsfrüherkennung) aufgrund der niedrigen Lungenkrebsprävalenz nur unter hohem Aufwand durchführbar sind, zunächst die Evidenz auf der Ebene relevanter randomisierter kontrollierter Studien (RCTs) ausgewertet.

Falls solche RCTs nicht oder in nicht ausreichender Quantität oder Qualität vorliegen, kann für diese Fragestellung ein Nutzen auch basierend auf der diagnostischen Güte abgeleitet werden (siehe Abschnitt A2.1.2.4).

Die diagnostische Güte der zu vergleichenden diagnostischen Verfahren kann z. B. anhand von Sensitivität und Spezifität für eine Nutzenableitung herangezogen werden. Dieses Vorgehen ist möglich, weil das zu prüfende diagnostische Verfahren, also die AI-gestützte Befundung alternativ zum bereits etablierten diagnostischen Verfahren eingesetzt werden soll und sich aufgrund der gleichbleibenden anschließenden Diagnosesicherung die therapeutischen Konsequenzen im Einzelfall nicht unterscheiden (siehe Abschnitt 3.5.1.2 Allgemeinen Methoden 7.0 [141]). Hierfür muss sichergestellt werden, insbesondere wenn die AI gestützte Diagnostik eine andere diagnostische Güte aufweist (z. B. deutlich höhere Krebsdetektionsrate), dass das mit dem AI-gestützten diagnostischen Verfahren identifizierbare Patientenspektrum tatsächlich gleich oder hinreichend ähnlich dem Patientenspektrum ist, für das der Nutzen in Therapiestudien nachgewiesen wurde. Die Gültigkeit dieser Linked-Evidence-Annahme wird überprüft, sofern die Studienergebnisse hierzu Anlass geben (z. B. deutlich höhere Krebsdetektionsrate in der AI-gestützten Befundung). Die Überprüfung der Verteilungen von histologischen Tumortypen als auch von Tumorstadien erfolgt mittels eines Vergleichs zum Referenzstandard jeweils für die AI-gestützten-Befundung und einer Befundung ausschließlich durch Radiologinnen und Radiologen.

Die im Folgenden dargestellten Kriterien für den Einschluss von Studien gelten, sofern nicht weiter spezifiziert, für beide Fragestellungen.

A2.1.1 Kriterien für den Einschluss von Studien zur diagnostisch-therapeutischen Behandlungskette

A2.1.1.1 Population

Fragestellung 1

In den ThemenCheck-Bericht werden Studien mit Patientinnen und Patienten ab 18 Jahren mit Verdacht auf Lungenkrebs aufgenommen. Ebenfalls können Studien mit Mischpopulationen berücksichtigt werden, die sowohl Patientinnen und Patienten mit Verdacht auf Lungenkrebs einschließen, als auch Personen, die an einer Lungenkrebsfrüherkennung (gemäß LuKrFrühErkV) teilnehmen und keiner der beiden Populationen eindeutig zuzuordnen sind.

Fragestellung 2

In den ThemenCheck-Bericht werden Studien mit Personen aufgenommen, die an einer Lungenkrebsfrüherkennung teilnehmen und gemäß der Kriterien der deutschen LuKrFrühErkV [5] für eine Teilnahme berechtigt wären. Hierbei handelt es sich um (ehemalige) starke Raucherinnen und Raucher im Alter zwischen 50 und 75 Jahren.

A2.1.1.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention stellt die diagnostische Strategie unter Anwendung bildgebender Verfahren zur Befundung der Lunge durch Radiologinnen und Radiologen mit Unterstützung von AI (Indextest 1) dar. Dabei gelten folgende Anforderungen an die Intervention:

- Die AI unterstützt Radiologinnen und Radiologen, in dem sie abnormale Strukturen erkennt (CADE) und / oder abnormale Strukturen diagnostiziert (CADx).
- Es werden AI-Anwendungen eingeschlossen, bei denen es sich um CE-gekennzeichnete Medizinprodukte handelt.
- **Fragestellung 1:** Bei Personen mit Verdacht auf Lungenkrebs erfolgt die Bildgebung primär mittels Röntgen. Eine Bildgebung mittels Low-Dose- oder High-Dose-CT ist bei dieser Population ebenfalls möglich.
Fragestellung 2: Bei Personen, die an einer Lungenkrebsfrüherkennung teilnehmen, erfolgt die Bildgebung mittels Low-Dose-CT.

Als Vergleichsintervention gilt die diagnostische Strategie unter Anwendung bildgebender Verfahren zur Befundung der Lunge ausschließlich durch Radiologinnen oder Radiologen. Für die Vergleichsintervention muss dasselbe bildgebende Verfahren angewendet werden, welches auch in der Prüfindervention durchgeführt wurde.

A2.1.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität (Gesamtmortalität, lungenkrebspezifische Mortalität)
- Morbidität, wie
 - krebbsbedingte Symptomatik
- gesundheitsbezogene Lebensqualität
- Nebenwirkungen, wie
 - Schäden, die sich aus der Befundung oder aus nachfolgenden diagnostischen Untersuchungen (zum Beispiel aus invasiven Prozeduren wie Biopsien) ergeben, einschließlich der Konsequenzen aus falschen Befunden (falsch-positiv und falsch-negativ).
 - Zusätzliche bei **Fragestellung 2**: Überdiagnosen im Rahmen der Lungenkrebsfrüherkennung

Ergänzend werden Angaben zur Zeitersparnis (z. B. Dauer der Befundung, Zeit bis zum Behandlungsbeginn, oder Zeit bis zur abschließenden Diagnose) sowie Biopsieraten bzw. Rate der invasiven Abklärungsdiagnostik betrachtet. Ein (höherer) Nutzen kann sich allein auf Basis dieser Endpunkte jedoch nicht ergeben.

A2.1.1.4 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 den zu erstellenden Bericht werden in erster Linie RCTs als relevante wissenschaftliche Informationsquelle in die Nutzenbewertung einfließen.

Falls keine RCTs in ausreichender Zahl und / oder Qualität vorliegen, werden schrittweise auch Studien einer niedrigeren Evidenzstufe eingeschlossen:

- 1) quasirandomisierte kontrollierte Studien,
- 2) prospektive vergleichende Kohortenstudien.

Dabei erfolgt der Einschluss von nicht randomisierten vergleichenden Studien ausschließlich bei adäquater Confounderkontrolle. Eine adäquate Confounderkontrolle liegt vor, wenn das Problem einer möglichen Strukturungleichheit bei der Planung und Auswertung der

entsprechenden Studien berücksichtigt wurde. Hierfür müssen Daten zu wesentlichen Basischarakteristika aller verglichenen Gruppen verfügbar sein, um den Einfluss wichtiger Confounder abschätzen zu können. Erforderlich sind hierfür Daten mindestens zu folgenden Confoundern: Alter, Geschlecht und Raucherstatus. Untersucht die Studie beide Populationen (wie in Abschnitt A2.1.1.1 zu Fragestellung 1 beschrieben) sind Daten zum Untersuchungsgrund (Abklärung eines Verdachts auf bzw. Früherkennung von Lungenkrebs) erforderlich.

A2.1.1.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

A2.1.1.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfügbar sein.

A2.1.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss

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

Tabelle 1: Übersicht über die Kriterien für den Einschluss von Studien zur diagnostisch-therapeutischen Behandlungskette in die Nutzenbewertung (Fragestellung 1)

Einschlusskriterien	
E1a	Patientinnen und Patienten mit Verdacht auf Lungenkrebs (siehe auch Abschnitt A2.1.1.1)
E2a	Prüfintervention: diagnostische Strategie unter Anwendung der Befundung der Lunge durch Radiologinnen und Radiologen mit Unterstützung von AI (siehe auch Abschnitt A2.1.1.2)
E3a	Vergleichsintervention: diagnostische Strategie unter Anwendung der Befundung der Lunge ausschließlich durch Radiologinnen und Radiologen (siehe auch Abschnitt A2.1.1.2)
E4a	patientenrelevante Endpunkte wie in Abschnitt A2.1.1.3 formuliert
E5a	Studientyp wie in Abschnitt A2.1.1.4 formuliert
E6a	Publikationssprache: Deutsch oder Englisch
E7a	Vollpublikation verfügbar ^a
a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Bericht über die Studie, der den Kriterien des CONSORT-(AI-) [142,143], TREND- [144], STARD- [145] oder STROBE-Statements [146] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und zu den Studienergebnissen nicht vertraulich sind. AI: Künstliche Intelligenz; CONSORT: Consolidated Standards of Reporting Trials; CT: Computertomografie; LuKrFrühErkV: Lungenkrebs-Früherkennungs-Verordnung; STARD: Standards for the Reporting of Diagnostic Accuracy Studies; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs	

Tabelle 2: Übersicht über die Kriterien für den Einschluss von Studien zur diagnostisch-therapeutischen Behandlungskette in die Nutzenbewertung (Fragestellung 2)

Einschlusskriterien	
E1b	Personen, die an einer Lungenkrebsfrüherkennung teilnehmen und gemäß den Kriterien der deutschen LuKrFrühErkV für eine Teilnahme berechtigt wären (siehe auch Abschnitt A2.1.1.1)
E2b	Prüfintervention: diagnostische Strategie unter Anwendung der Befundung der Lunge durch Radiologinnen und Radiologen mit Unterstützung von AI (siehe auch Abschnitt A2.1.1.2)
E3b	Vergleichsintervention: diagnostische Strategie unter Anwendung der Befundung der Lunge ausschließlich durch Radiologinnen und Radiologen (siehe auch Abschnitt A2.1.1.2)
E4b	patientenrelevante Endpunkte wie in Abschnitt A2.1.1.3 formuliert
E5b	Studientyp wie in Abschnitt A2.1.1.4 formuliert
E6b	Publikationssprache: Deutsch oder Englisch
E7b	Vollpublikation verfügbar ^a
a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Bericht über die Studie, der den Kriterien des CONSORT-(AI-) [142,143], TREND- [144], STARD- [145] oder STROBE-Statements [146] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und zu den Studienergebnissen nicht vertraulich sind. AI: Künstliche Intelligenz; CONSORT: Consolidated Standards of Reporting Trials; CT: Computertomografie; LuKrFrühErkV: Lungenkrebs-Früherkennungs-Verordnung; STARD: Standards for the Reporting of Diagnostic Accuracy Studies; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs	

A2.1.1.8 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für die Einschlusskriterien E1 (Population), E2 (Prüfintervention, bezogen auf die Interventionsgruppe der Studie) sowie E3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patientinnen und Patienten oder Personen diese Kriterien erfüllt sind. Liegen für solche Studien Subgruppenanalysen für Patientinnen und Patienten oder Personen 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 Patientinnen und Patienten oder Personen vorliegen, die die Einschlusskriterien erfüllen.

A2.1.2 Kriterien für den Einschluss von Studien zur diagnostischen Güte

A2.1.2.1 Population

Fragestellung 1: Analog zu den Einschlusskriterien für Studien zur diagnostisch-therapeutischen Behandlungskette (siehe Abschnitt A2.1.1.1), werden für Studien zur diagnostischen Güte solche mit Patientinnen und Patienten ab 18 Jahren mit Verdacht auf Lungenkrebs aufgenommen. Ebenfalls können Studien mit Mischpopulationen berücksichtigt werden, die sowohl Patientinnen und Patienten mit Verdacht auf Lungenkrebs einschließen,

als auch Personen, die an einer Lungenkrebsfrüherkennung (gemäß LuKrFrühErkV) teilnehmen und keiner der beiden Populationen eindeutig zuzuordnen sind.

Fragestellung 2: Analog zu den Einschlusskriterien für Studien zur diagnostisch-therapeutischen Behandlungskette (siehe Abschnitt A2.1.1.1), werden für Studien zur diagnostischen Güte solche mit Personen aufgenommen, die an einer Lungenkrebsfrüherkennung teilnehmen und gemäß der Kriterien der deutschen LuKrFrühErkV [5] für eine Teilnahme berechtigt wären.

A2.1.2.2 Index- und Referenztest

Der zu prüfende Indextest stellt die Befundung der Lunge durch Radiologinnen und Radiologen mit Unterstützung von AI (Indextest 1) dar. Die Befundung erfolgt auf Grundlage der Bildgebung der Lunge. Dabei gelten folgende Anforderungen an die Intervention:

- Die AI unterstützt Radiologinnen und Radiologen in dem sie abnormale Strukturen erkennt (CADE) und / oder abnormale Strukturen diagnostiziert (CADx).
- Es werden AI-Anwendungen eingeschlossen, bei denen es sich um CE-gekennzeichnete Medizinprodukte handelt.
- **Fragestellung 1:** Bei Personen mit Verdacht auf Lungenkrebs erfolgt die Bildgebung primär mittels Röntgen. Eine Bildgebung mittels Low-Dose- oder High-Dose-CT ist bei dieser Population ebenfalls möglich.

Fragestellung 2: Bei Personen, die an einer Lungenkrebsfrüherkennung teilnehmen, erfolgt die Bildgebung mittels Low-Dose-CT.

Als Vergleich gilt die Befundung der Lunge ausschließlich durch Radiologinnen oder Radiologen (Indextest 2). Die Befundung erfolgt auf Grundlage der Bildgebung der Lunge, hierbei muss dasselbe bildgebende Verfahren wie in Indextest 1 der Studie angewendet werden.

Den Referenztest bildet die Biopsie, die klinische Nachbeobachtung oder Nachuntersuchung.

Ergänzend werden auch die alleinigen Ergebnisse der AI betrachtet, sofern diese berichtet werden. Da es sich hierbei um eine ergänzende Darstellung handelt, werden keine Aussagen zu Nutzen und Schaden getroffen.

A2.1.2.3 Zielgröße

Eingeschlossen werden solche Studien, aus denen personenbezogene Vierfeldertafel-Daten zur Berechnung der diagnostischen Güte (z. B. Sensitivität, Spezifität) von Indextest 1 und Indextest 2 ableitbar sind. Zusätzlich werden auch Ergebnisse zur Konkordanz zwischen Indextest 1 und Indextest 2, sofern berichtet, berücksichtigt.

Eingeschlossen werden auch Studien im VOPT-Design (VOPT: Verification of only positive Testers), in welchem alle positiven Ergebnisse im Indextest mit dem Referenztest untersucht werden [147]. Eine Bewertung der testnegativen Fälle und damit eine Bestimmung der Sensitivität oder Spezifität des Tests ist mit solchen Studien nicht möglich.

Zu diesem Vorgehen gab es eine Spezifizierung im Projektverlauf, siehe Abschnitt A1.2

A2.1.2.4 Studientypen

Zur Bewertung der diagnostischen Güte werden sowohl intraindividuelle als auch interindividuelle Vergleichsstudien eingeschlossen. Aus prospektiv geplanten Kohortenstudien² können Ergebnisse zu intraindividuellen Vergleichen gewonnen werden. Hierbei erhält jede Person die beiden zu vergleichenden Indextests und den Referenztest.

Zu interindividuellen Vergleichen werden in erster Linie RCTs in die Nutzenbewertung einfließen, in denen je eine Patientengruppe zufällig einen der Indextests erhält und alle mittels Referenztest verifiziert werden. Sollten RCTs nicht in ausreichender Zahl und / oder Qualität vorliegen, werden auch quasirandomisierte Studien zu interindividuellen Vergleichen eingeschlossen. Der Einschluss von quasirandomisierten kontrollierten Studien erfolgt ausschließlich bei Berücksichtigung potenzieller Confounder. Hierfür müssen Daten zu wesentlichen Basischarakteristika aller verglichenen Gruppen verfügbar sein, um den Einfluss wichtiger Confounder abschätzen zu können. Erforderlich sind hierfür Daten mindestens zu folgenden Confoundern: Alter, Geschlecht und Raucherstatus.

Fragestellung 1: Untersucht die Studie beide Populationen (wie in Abschnitt A2.1.2.1 beschrieben) sind Daten zum Untersuchungsgrund (Abklärung eines Verdachts auf bzw. Früherkennung von Lungenkrebs) erforderlich.

A2.1.2.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

A2.1.2.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfügbar sein.

² Prospektive geplante Kohortenstudien umfassen auch Studien, in denen Patientinnen und Patienten bzw. Teilnehmerinnen und Teilnehmer retrospektiv eingeschlossen werden. Zentrale Aspekte der Ergebnissicherheit bei Studien zur diagnostischen Güte sind der konsekutive Einschluss der Patientinnen und Patienten bzw. Teilnehmerinnen und Teilnehmer sowie die Dokumentation fehlender Werte. Bei ausreichender Anzahl und Qualität werden nur die Studien eingeschlossen, die diese Voraussetzungen erfüllen.

A2.1.2.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss

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

Tabelle 3: Übersicht über die Kriterien für den Einschluss von Studien zur diagnostischen Güte in die Nutzenbewertung (Fragestellung 1)

Einschlusskriterien	
D1a	Patientinnen und Patienten mit Verdacht auf Lungenkrebs (siehe auch Abschnitt A2.1.2.1)
D2a	Indextest 1: Befundung der Lunge durch Radiologinnen und Radiologen mit Unterstützung von AI (siehe auch Abschnitt A2.1.2.2)
D3a	Vergleich: Indextest 2; Befundung der Lunge ausschließlich durch Radiologinnen und Radiologen (siehe auch Abschnitt A2.1.2.2)
D4a	Referenztest: Biopsie, klinische Nachbeobachtung oder Nachuntersuchung (siehe auch Abschnitt A2.1.2.2)
D5a	Zielgröße: auf die Beobachtungseinheit Person bezogene Vierfeldertafel-Daten zur diagnostischen Güte (siehe auch Abschnitt A2.1.2.3)
D6a	Studientyp wie in Abschnitt A2.1.2.4 formuliert
D7a	Publikationssprache: Deutsch oder Englisch
D8a	Vollpublikation verfügbar ^a
a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Bericht über die Studie, der den Kriterien des CONSORT(-AI)- [142,143], TREND- [144], STARD- [145] oder STROBE-Statements [146] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und zu den Studienergebnissen nicht vertraulich sind. AI: Künstliche Intelligenz; CONSORT: Consolidated Standards of Reporting Trials; CT: Computertomografie; LuKrFrühErkV: Lungenkrebs-Früherkennungs-Verordnung; STARD: Standards for the Reporting of Diagnostic Accuracy Studies; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs	

Tabelle 4: Übersicht über die Kriterien für den Einschluss von Studien zur diagnostischen Güte in die Nutzenbewertung (Fragestellung 2)

Einschlusskriterien	
D1b	Personen, die an einer Lungenkrebsfrüherkennung teilnehmen und gemäß den Kriterien der deutschen LuKrFrühErkV für eine Teilnahme berechtigt wären (siehe auch Abschnitt A2.1.2.1)
D2b	Indextest 1: Befundung der Lunge durch Radiologinnen und Radiologen mit Unterstützung von AI (siehe auch Abschnitt A2.1.2.2)
D3b	Vergleich: Indextest 2; Befundung der Lunge ausschließlich durch Radiologinnen und Radiologen (siehe auch Abschnitt A2.1.2.2)
D4b	Referenztest: Biopsie, klinische Nachbeobachtung oder Nachuntersuchung (siehe auch Abschnitt A2.1.2.2)
D5b	Zielgröße: auf die Beobachtungseinheit Person bezogene Vierfeldertafel-Daten zur diagnostischen Güte (siehe auch Abschnitt A2.1.2.3)
D6b	Studientyp wie in Abschnitt A2.1.2.4 formuliert
D7b	Publikationssprache: Deutsch oder Englisch
D8b	Vollpublikation verfügbar ^a
a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Bericht über die Studie, der den Kriterien des CONSORT(-AI)- [142,143], TREND- [144], STARD- [145] oder STROBE-Statements [146] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zu Studienmethodik und zu den Studienergebnissen nicht vertraulich sind. AI: Künstliche Intelligenz; CONSORT: Consolidated Standards of Reporting Trials; CT: Computertomografie; LuKrFrühErkV: Lungenkrebs-Früherkennungs-Verordnung; STARD: Standards for the Reporting of Diagnostic Accuracy Studies; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs	

A2.1.2.7.1 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für die Einschlusskriterien D1 (Population), D2 (Indextest 1 bezogen auf die Interventionsgruppe der Studie), D3 (Vergleich: Indextest 2, bezogen auf die Vergleichsgruppe der Studie) und D4 (Referenztest) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patientinnen und Patienten oder Personen diese Kriterien erfüllt sind. Liegen für solche Studien Subgruppenanalysen für Patientinnen und Patienten oder Personen vor, die die Einschlusskriterien erfüllen, wird auf diese Analysen zurückgegriffen. Studien, bei denen die Einschlusskriterien D1, D2, D3 und D4 bei weniger als 80 % erfüllt sind, werden nur dann eingeschlossen, wenn Subgruppenanalysen für Patientinnen und Patienten oder Personen vorliegen, die die Einschlusskriterien erfüllen.

A2.1.3 Informationsbeschaffung

A2.1.3.1 Fokussierte Informationsbeschaffung von systematischen Übersichten

Parallel zur Erstellung des ThemenCheck-Berichtsprotokolls 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 am 26.02.2025 statt. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Abschnitt A8.1. 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 eine 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) [148]. Kann mindestens eine diesbezüglich hochwertige und aktuelle Basis-SÜ identifiziert werden, werden die zugrunde liegenden Studien bzw. 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 ThemenCheck-Berichtsprotokolls anhand der darin festgelegten Kriterien entsprechend Tabelle 1 und Tabelle 2 sowie Tabelle 3 und Tabelle 4. In jedem Fall werden die Referenzlisten der identifizierten systematischen Übersichten hinsichtlich relevanter Primärstudien gesichtet (siehe Abschnitt A2.1.3.2).

A2.1.3.2 Umfassende Informationsbeschaffung von Studien

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

Für den Fall, dass mindestens eine systematische Übersicht als Basis-SÜ für die Informationsbeschaffung verwendet werden kann (siehe Abschnitt A2.1.3.1), 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 bzw. 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

Aufgrund der separaten Suche in ClinicalTrials.gov werden aus dem Suchergebnis des International Clinical Trials Registry Platform Search Portal Einträge dieses Registers entfernt.

Weitere Informationsquellen und Suchtechniken

- Anwendung weiterer Suchtechniken:
 - Sichten von Referenzlisten identifizierter systematischer Übersichten
- Anhörung zum vorläufigen ThemenCheck-Bericht
- Autorinnen- und Autorenanfragen

A2.1.3.3 Anwendung von Limitierungen auf Datenbankebene

Fokussierte Informationsbeschaffung von systematischen Übersichten

Die Suchen wurden auf kein Publikationsdatum eingeschränkt. Die MEDLINE Suchstrategie enthält Limitierungen auf deutsch- und englischsprachige Publikationen sowie auf Humanstudien.

Umfassende Informationsbeschaffung von Studien

Es ist keine zeitliche Einschränkung vorgesehen. Sollte die Informationsbeschaffung auf Grundlage einer Basis-SÜ erfolgen, wird eine entsprechende zeitliche Einschränkung in Betracht gezogen (siehe Abschnitt A2.1.3.2).

Mit den Suchstrategien werden folgende Publikationstypen ausgeschlossen: Kommentare (MEDLINE) und Editorials (MEDLINE, Embase), da diese i. d. R. keine Studien enthalten [149] 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 Suche Einträge aus Studienregistern ausgeschlossen.

A2.1.3.4 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 und Tabelle 2 sowie Tabelle 3 und Tabelle 4) 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.

A2.1.4 Informationsbewertung und -synthese

A2.1.4.1 Darstellung der Einzelstudien

Alle für die Bewertung notwendigen Informationen werden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Die Ergebnisse zu den in den Studien berichteten patientenrelevanten Endpunkten werden im Bericht 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 in den Abschnitten A2.1.4.3 bis A2.1.4.5 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 Patientinnen und Patienten oder Personen basieren, das heißt, wenn der Anteil der Patientinnen und Patienten oder 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 Patientinnen und Patienten oder Personen zwischen den Gruppen größer als 15 Prozentpunkte ist.

A2.1.4.2 Bewertung des Verzerrungspotenzials der Ergebnisse

A2.1.4.2.1 Ergebnisse aus Studien zur diagnostisch-therapeutischen Behandlungskette

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 (bei randomisierten Studien)
- Verdeckung der Gruppenzuteilung (bei randomisierten Studien)
- zeitliche Parallelität der Gruppen (bei nicht randomisierten kontrollierten Studien)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (bei nicht randomisierten kontrollierten Studien)
- Verblindung der Patientin oder des Patienten sowie der behandelnden Person (bei randomisierten Studien)
- 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 von Patientinnen und Patienten bzw. Personen bereits ein endpunktspezifisch hohes Verzerrungspotenzial.

Das Verzerrungspotenzial der Ergebnisse nicht randomisierter vergleichender Studien wird aufgrund der fehlenden Randomisierung zusammenfassend grundsätzlich als hoch bewertet.

A2.1.4.2.2 Ergebnisse aus Studien zur diagnostischen Güte

Die Bewertung des Verzerrungspotenzials und der Übertragbarkeit der eingeschlossenen Studie zur Bewertung der diagnostischen Güte erfolgt auf Basis des QUADAS-2-Instruments [13]. Das Verzerrungspotenzial von Primärstudien zur diagnostischen Güte wird als niedrig oder hoch eingestuft.

A2.1.4.3 Metaanalysen

A2.1.4.3.1 Ergebnisse aus Studien zur diagnostisch-therapeutischen Behandlungskette

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 [150] untersucht. Es wird außerdem untersucht, welche Faktoren eine vorhandene Heterogenität möglicherweise verursachen. Dazu zählen methodische Faktoren (siehe Abschnitt A2.1.4.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt A2.1.4.5). 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 folgendes Standardvorgehen gewählt sofern keine klaren Gründe dagegensprechen:

- 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 [151].
- 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. [152]. Zudem erfolgt ein Abgleich mit einer

qualitativen Zusammenfassung der Studienergebnisse. Für sonstige Effektmaße ist projektspezifisch zu entscheiden, ob die Methode nach Knapp-Hartung, eine qualitative Zusammenfassung der Studienergebnisse oder ein anderes 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 [153] 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 4 oder mehr Studien wird hierzu das Prädiktionsintervall im Forest Plot mit dargestellt.

A2.1.4.3.2 Ergebnisse aus Studien zur diagnostischen Güte

Neben den RCTs zur diagnostisch-therapeutischen Behandlungskette können für die Bewertung auch Studien zur diagnostischen Güte herangezogen werden. Bei diesen muss die diagnostische Güte des zu prüfenden diagnostischen Verfahrens, die AI-gestützte Befundung (Indextest 1) mit der diagnostischen Güte eines alternativen etablierten diagnostischen Verfahrens, Befundung ohne AI (Indextest 2) in Bezug auf den Referenztest verglichen werden.

Die Punktschätzungen und dazugehörigen univariaten 95 %-Konfidenzintervalle [154] aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Außerdem werden, sofern die dafür nötigen Anforderungen erfüllt sind, für die Maße der diagnostischen Güte uni- oder bivariate Metaanalysen durchgeführt [155]. Die Schätzung der Modellparameter erfolgt über ein generalisiertes lineares gemischtes Modell [156,157].

Falls die bivariate Metaanalyse präzise Schätzungen liefert, so werden bei diagnostischen Studien die beobachteten Paare aus Sensitivität und Spezifität zweidimensional grafisch dargestellt. Des Weiteren werden die aus der bivariaten Metaanalyse gewonnenen Schätzungen für die Erwartungswerte als gepoolte Paare der Sensitivität und der Spezifität mit den dazugehörigen 95 %-Konfidenzregionen dargestellt [158].

Der Algorithmus zum Schätzen der Parameter im bivariaten Modell kann zu unpräzisen Schätzungen führen, das heißt zu Schätzungen mit zu großen Standardfehlern und entsprechenden Konfidenzregionen. Auch kann der Algorithmus gegebenenfalls keine Schätzungen liefern, wenn das Maximum-Likelihood-Verfahren nicht konvergiert. In beiden Fällen fehlen brauchbare Schätzungen. Die Gründe hierfür können beispielsweise sein, dass zu wenige Studien vorliegen oder dass einzelne Studien extreme Werte aufweisen. Sind die resultierenden Schätzungen unpräzise, werden die Ergebnisse der bivariaten Metaanalysen in der Regel nicht dargestellt. In diesem Fall wird auf univariate Metaanalysen für die Sensitivität und Spezifität zurückgegriffen. Von der Berechnung einer gepoolten Schätzung wird abgesehen, falls sich die 95 %-Konfidenzintervalle der eingehenden Studien nur gering oder gar nicht überlappen und sich gleichzeitig die Schätzungen der Studien zu stark unterscheiden.

Das Vorliegen von Heterogenität wird anhand von Sensitivitäts- sowie Subgruppenanalysen untersucht.

A2.1.4.4 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 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 A2.1.4.6).

A2.1.4.5 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 Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- bzw. 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:

- Geschlecht,

- Alter,
- Art der AI,
- Art der Bildgebung (Verdachtsuntersuchung).

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

Bei Identifizierung möglicher Effektmodifikatoren erfolgt ggf. eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen. Bspw. kann der Beleg eines (höheren) Nutzens auf eine spezielle Subgruppe von Patientinnen und Patienten oder Personen bzw. von AI-programmen oder bildgebenden Verfahren eingeschränkt werden (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt A2.1.4.6).

A2.1.4.6 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 für Studien zur diagnostisch-therapeutischen Behandlungskette ist von den in Tabelle 5 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. Ergebnisse aus Studien zur diagnostischen Güte haben ebenfalls eine geringe qualitative Ergebnissicherheit.

Tabelle 5: 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		
				konkludente Effekte ^a		
			Metaanalyse statistisch signifikant	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 [141])						

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.

In der Bewertung wird jeweils ein höherer oder geringerer Nutzen oder Schaden der Befundung durch Radiologinnen und Radiologen mit Unterstützung von AI im Vergleich zum diagnostischen Verfahren ausschließlich durch Radiologinnen und Radiologen bewertet. Wenn endpunktübergreifend nicht mindestens ein Anhaltspunkt für einen höheren oder geringeren Nutzen oder Schaden der Befundung durch Radiologinnen und Radiologen mit Unterstützung von AI gezeigt werden kann, wird ein vergleichbarer Nutzen gegenüber der Befundung ausschließlich durch Radiologinnen und Radiologen überprüft.

Die Überprüfung des vergleichbaren Nutzens erfolgt sowohl für Studien zur diagnostisch-therapeutischen Behandlungskette als auch für Studien zur diagnostischen Güte. Im Rahmen der Linked-Evidence-Annahme (siehe Abschnitt A2.1) ist es für Studien zur diagnostischen Güte zur Ableitung eines vergleichbaren Nutzens ausreichend, dass eine mindestens vergleichbare diagnostische Güte vorliegt.

A2.2 Gesundheitsökonomische Bewertung

A2.2.1 Interventionskosten

Zur Bestimmung der Interventionskosten werden die durchschnittlichen Ressourcenverbräuche bestimmt, die jeweils direkt bei Anwendung der Prüf- und Vergleichsintervention erforderlich sind. Hierbei werden neben der Prüf- und Vergleichsintervention die zusätzlich

mit der Anwendung einhergehenden Leistungen berücksichtigt. Sofern die Prüf- oder Vergleichsintervention aus mehreren Leistungen besteht, werden alle Komponenten dargestellt. Für die anfallenden Leistungen werden soweit möglich die jeweils relevanten regulierten oder verhandelten Preise, z. B. aus der Datenbank der Informationsstelle für Arzneyspezialitäten (IFA), dem Einheitlichen Bewertungsmaßstab (EBM), dem Diagnosis-Related-Groups(DRG)-Katalog oder ähnlich geeignete Aufstellungen aus der Rentenversicherung oder des Statistischen Bundesamts angesetzt. Sofern notwendig, werden alternative Vorgehensweisen zur Bestimmung der Interventionskosten transparent dargestellt. Falls eine Therapie länger als ein Jahr dauert, werden die durchschnittlich pro Patientin bzw. Patient und Jahr anfallenden Kosten angegeben. Erstattungsfähige und nicht erstattungsfähige Kosten sowie Zuzahlungen werden getrennt ausgewiesen.

A2.2.2 Systematische Übersicht gesundheitsökonomischer Evaluationen

A2.2.2.1 Kriterien für den Einschluss von Studien in die systematische Übersicht

A2.2.2.1.1 Studientypen

In die systematische Übersicht gesundheitsökonomischer Studien werden vergleichende Studien mit einer Aussage zur Kosteneffektivität einbezogen [159], das heißt Kosten-Effektivitäts- / Wirksamkeits-Analysen, Kosten-Nutzwert-Analysen oder Kosten-Nutzen-Analysen (im engeren Sinne). Werden diese Studientypen im Rahmen der Recherche nicht identifiziert, erfolgt der Einschluss vergleichender gesundheitsökonomischer Studien mit einer Aussage zu den Kosten der Prüfintervention und Vergleichsintervention, das heißt Kosten-Kosten-Analysen.

A2.2.2.1.2 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfügbar sein.

A2.2.2.1.3 Gesundheitssystem bzw. geografischer Bezug

Für die systematische Übersicht gesundheitsökonomischer Studien erfolgt keine Einschränkung auf Studien aus einem bestimmten Gesundheitssystem oder Land.

A2.2.2.1.4 Tabellarische Darstellung der Kriterien

In der folgenden Tabelle sind die Kriterien aufgelistet, die die Studien zusätzlich zu den in Tabelle 1 oder Tabelle 2 beschriebenen Einschlusskriterien E1a bis E3a oder E1b bis E3b bzw. den in Tabelle 3 oder Tabelle 4 beschriebenen Einschlusskriterien D1a bis D3a oder D1b bis D3b erfüllen müssen, um für die ökonomische Bewertung eingeschlossen zu werden.

Ergänzend werden auch gesundheitsökonomische Studien berücksichtigt, in denen als Prüfintervention bzw. als Indextest die Befundung der Lunge allein durch AI untersucht wird. Ansonsten sind die in diesem Abschnitt benannten Kriterien zu berücksichtigen.

Tabelle 6: Übersicht über die zusätzlichen Kriterien für den Einschluss von Studien in die ökonomische Bewertung

Einschlusskriterien	
EÖ1	Studientyp: vergleichende gesundheitsökonomische Evaluation (siehe auch Abschnitt A2.2.2.1.1)
EÖ2	Publikationssprache: Deutsch oder Englisch
EÖ3	Vollpublikation verfügbar
EÖ: Einschlusskriterien ökonomische Bewertung	

A2.2.2.2 Fokussierte Informationsbeschaffung

Für die Bewertung gesundheitsökonomischer Aspekte wird eine systematische Recherche in Form einer fokussierten Informationsbeschaffung durchgeführt. Folgende primäre und weitere Informationsquellen sowie Suchtechniken werden dabei berücksichtigt:

Primäre Informationsquellen

- bibliografische Datenbanken
 - MEDLINE
 - Embase
 - HTA Database

Weitere Informationsquellen und Suchtechniken

- Anwendung weiterer Suchtechniken
 - Sichten von Referenzlisten identifizierter systematischer Übersichten
- Anhörung zum vorläufigen ThemenCheck-Bericht
- Autorinnen- und Autorenanfragen

A2.2.2.2.1 Selektion relevanter Publikationen

Die durch die Suche identifizierten Zitate werden durch 1 Person anhand der Einschlusskriterien (siehe Tabelle 6) selektiert. Das Ergebnis wird durch eine 2. Person qualitätsgesichert.

A2.2.2.3 Informationsbewertung

Datenextraktion

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

Bewertung der Berichtsqualität

Die Bewertung der Berichtsqualität der berücksichtigten gesundheitsökonomischen Studien orientiert sich an den Kriterien des Consolidated Health Economic Evaluation Reporting Standards for Interventions That Use Artificial Intelligence (CHEERS-AI-Statement) [14].

Bewertung der Übertragbarkeit

Die Bewertung der Übertragbarkeit der Ergebnisse orientiert sich an den Kriterien des European network for Health Technology Assessment (EUnetHTA) HTA adaptation toolkit [15].

A2.2.2.4 Informationsanalyse und -synthese

Es werden gesundheitsökonomische Studien betrachtet, die Aussagen zur Kosteneffektivität der Technologie gegenüber der Vergleichsintervention machen. Ergänzend können auch vergleichende gesundheitsökonomische Studien mit Aussagen ausschließlich zu den Kosten der Prüfindervention und Vergleichsintervention betrachtet werden.

Die Ergebnisse zu der in den Studien berichteten Kosteneffektivität bzw. der in den Studien berichteten Kosten und die Schlussfolgerungen der Autorinnen und Autoren werden im ThemenCheck-Bericht vergleichend beschrieben. Dabei sollen insbesondere auch Aspekte der Qualität der dargestellten Studien und deren Übertragbarkeit auf das deutsche Gesundheitssystem diskutiert werden. Zu diskutieren ist ebenfalls, welche Auswirkungen sich aus der Verwendung von Endpunkten ergeben, die von der Nutzenbewertung abweichen.

A2.3 Ethische Aspekte

A2.3.1 Berücksichtigung von Argumenten und Aspekten bei der Aufarbeitung ethischer Implikationen

Ethische Argumente und Aspekte finden sich in Publikationen, in Monografien, Projektberichten aber bspw. auch in Gesetzen und Verordnungen oder auf den Websites von Interessengruppen. Sie sind unabhängig vom Publikationstyp, -status und Studientyp.

Im ThemenCheck-Bericht werden Aussagen zu ethischen Aspekten und Argumenten der zu untersuchenden Technologie berücksichtigt.

A2.3.2 Informationsbeschaffung

Orientierende Recherche

Für die Informationsbeschaffung erfolgt in einem ersten Schritt eine Orientierung zu möglichen ethisch relevanten Argumenten und Aspekten durch eine Ausarbeitung des Hofmannschen Fragenkatalogs [17]. Anschließend wird eine orientierende Recherche in folgenden Informationsquellen durchgeführt:

- Gesetze, Verordnungen oder Richtlinien

- ETHMED (ETHik in der MEDizin / Literaturdatenbank der Universität Göttingen)
- interessenabhängige Informationsquellen, zum Beispiel Websites von Interessenvertreterinnen und -vertretern
- MEDLINE
- LIVIVO (System der ZB Med Köln)
- BELIT (Literaturdatenbank, die vom Deutschen Referenzzentrum für Ethik in den Biowissenschaften (DRZE, Bonn) entwickelt worden ist)

Informationen aus allen Informationsquellen der orientierenden Recherchen werden von 1 Person auf Aussagen zu ethischen Argumenten und Aspekten der zu untersuchenden Technologie gesichtet. Das Ergebnis wird durch eine 2. Person qualitätsgesichert.

Weitere Informationsquellen

Die folgenden Dokumente werden auf mögliche ethische Argumente geprüft:

- in die Nutzenbewertung eingeschlossene Studien
- in die gesundheitsökonomische Bewertung eingeschlossene Studien
- das Protokoll zur Dokumentation der Diskussion mit den befragten Betroffenen

Die Prüfung der Dokumente auf Aussagen zu ethischen Argumenten und Aspekten der zu untersuchenden Technologie erfolgt durch 1 Person. Das Ergebnis wird durch eine 2. Person qualitätsgesichert.

Sollten sich in den vorgenannten Informationsquellen nur unzureichende Informationen finden, können ergänzend auch weitere Stakeholder befragt werden.

Zusätzlich können „reflective thoughts“, also das reflektierte und auf das Wissen der Berichtsautorinnen und -autoren bezogene Nachdenken über mögliche ethische Argumente und Aspekte, als Informationsquelle genutzt werden [19].

A2.3.3 Informationsaufbereitung

Alle für die Informationsaufbereitung notwendigen Argumente und Aspekte werden in Tabellen extrahiert.

Die Informationsaufbereitung zu ethischen Aspekten erfolgt in Orientierung an den übergeordneten Fragestellungen des vereinfachten Fragenkatalogs von Hofmann [17] und orientiert sich darüber hinaus an einem um spezifische Aspekte von Künstlicher Intelligenz in der Diagnostik erweiterten Rahmengerüst für die ethische Bewertung von Public-Health-Maßnahmen von Marckmann [18]. Die Ergebnisse werden tabellarisch dargestellt.

Zu diesem Vorgehen gab es eine Spezifizierung im Projektverlauf, siehe Abschnitt A1.2

In der Informationsaufbereitung zu ethischen Aspekten erfolgt eine Auseinandersetzung mit sozialen und moralischen Normen und Werten, die in Beziehung zur Technologie des ThemenCheck-Berichts stehen.

A2.4 Soziale, rechtliche und organisatorische Aspekte sowie Umwelt- und Klimaaspekte

A2.4.1 Berücksichtigung von Argumenten und Aspekten bei der Aufarbeitung sozialer, rechtlicher und organisatorischer Implikationen sowie von Umwelt- und Klimaaspekten

Im ThemenCheck-Bericht werden Argumente bzw. Aspekte berücksichtigt, die Aussagen zu sozialen, rechtlichen und / oder organisatorischen Aspekten der zu untersuchenden Technologie beinhalten. Zusätzlich können auch Umwelt- und Klimaaspekte einer Technologie betrachtet werden.

Entsprechende Argumente und Aspekte finden sich in Publikationen, in Monografien, Projektberichten aber bspw. auch in Gesetzen und Verordnungen oder auf den Websites von Interessengruppen. Sie sind unabhängig vom Publikationstyp, -status und Studientyp.

A2.4.2 Informationsbeschaffung

Orientierende Recherche

Für die Aufarbeitung sozialer, rechtlicher und organisatorischer Aspekte sowie von Umwelt- und Klimaaspekten werden orientierende Recherchen durchgeführt.

Die orientierenden Recherchen werden in folgenden Informationsquellen durchgeführt:

- MEDLINE
- LIVIVO (System der ZB Med Köln)
- Google Scholar
- Gesetze, Verordnungen oder Richtlinien (zum Beispiel Beck-Online oder juris).
- interessenabhängige Informationsquellen, z. B. Websites von Interessenvertretungen
- nationale und regionale Register

Informationen aus allen Informationsquellen der orientierenden Recherchen werden von 1 Person auf Aussagen zu sozialen, rechtlichen und / oder organisatorischen Argumenten sowie Umwelt- und Klimaaspekten der zu untersuchenden Technologie gesichtet. Das Ergebnis wird durch eine 2. Person qualitätsgesichert.

Weitere Informationsquellen

Die folgenden Dokumente werden auf mögliche soziale, rechtliche und / oder organisatorische Argumente sowie Umwelt- und Klimaaspekte geprüft:

- in die Nutzenbewertung eingeschlossene Studien
- in die gesundheitsökonomische Bewertung eingeschlossene Studien
- das Protokoll zur Dokumentation der Diskussion mit den befragten Betroffenen

Die Prüfung der Dokumente auf Aussagen zu sozialen, rechtlichen und / oder organisatorischen Argumenten sowie Umwelt- und Klimaaspekten der zu untersuchenden Technologie erfolgt durch 1 Person. Das Ergebnis wird durch eine 2. Person qualitätsgesichert.

Sollten sich in den vorgenannten Informationsquellen nur unzureichende Informationen finden, können ergänzend auch weitere Stakeholder befragt werden.

Zusätzlich können „reflective thoughts“, also das reflektierte und auf das Wissen der Berichtautorinnen und -autoren bezogene Nachdenken über mögliche soziale, rechtliche oder organisatorische Argumente sowie Umwelt- und Klimaaspekte, als Informationsquelle genutzt werden [19].

A2.4.3 Informationsaufbereitung

Datenextraktion

Alle für die Informationsaufbereitung notwendigen Argumente und Aspekte werden in Tabellen extrahiert.

Soziale Aspekte

Soziale und soziokulturelle Aspekte im ThemenCheck-Bericht greifen die wechselseitigen Interaktionen zwischen Untersuchungs- bzw. Behandlungsmethode und sozialer Umwelt (z. B. Verteilung von Ressourcen in einer Gesellschaft, Zugang zu Technologien, Präferenzen von Patientinnen und Patienten, gesellschaftliche Normen und Wertvorstellungen) auf.

Die Informationsaufbereitung zu sozialen Aspekten orientiert sich an dem von Mozygemba 2016 [20] vorgeschlagenem umfassenden konzeptionellen Rahmen.

Rechtliche Aspekte

Rechtliche Aspekte im ThemenCheck-Bericht beziehen sich zum einen auf den rechtlichen Rahmen, in den die Untersuchungs- bzw. Behandlungsmethode und deren Bewertung eingebettet ist (z. B. Marktzulassung, Erstattungsstatus), und zum anderen auf die mit der Implementation und Nutzung der Gesundheitstechnologie verbundenen rechtlichen Aspekte

(z. B. Patientenautonomie). Unterschieden wird zwischen den technologie- und patientenbezogenen rechtlichen Aspekten.

Die Informationsaufbereitung zu rechtlichen Aspekten orientiert sich an dem von Brönneke 2016 [21] entwickelten Leitfaden zur Identifikation von rechtlichen Aspekten.

Organisatorische Aspekte

Organisatorische Aspekte umfassen die Wechselwirkungen, welche durch eine Untersuchungs- bzw. Behandlungsmethode auf die Organisation der Versorgung entstehen.

Die Informationsaufbereitung zu organisatorischen Aspekten orientiert sich an dem von Perleth 2014 [22] vorgeschlagenen Raster zur Einschätzung der organisatorischen Folgen von Untersuchungs- bzw. Behandlungsmethoden.

Umwelt- und Klimaaspekte

Umwelt- und Klimaaspekte umfassen die Auswirkungen, die Gesundheitstechnologien auf die Umwelt haben, etwa aufgrund von CO₂-Emissionen bei Herstellung und Transport, aufgrund von toxischen Substanzen (vor allem während des Herstellungsprozesses) oder aufgrund von Abfällen (z. B. bei der Entsorgung des Produkts oder Verpackungsmaterialien).

In der Informationsaufbereitung erfolgt, sofern zutreffend, eine deskriptive Darstellung möglicher Umwelt- und Klimaaspekte

- zu CO₂-Emissionen (z. B. während der Produktion, Nutzung und Entsorgung, während des Transports),
- zu toxischen Substanzen (z. B. hormonell wirksame Substanzen, Schwermetalle, umweltschädliche Gase oder umweltschädliche Strahlung),
- zum Abfallmanagement (z. B. Einmal- oder Mehrfachprodukt, Recyclingfähigkeit, Abfallmenge)

sowie ggf. zu weiteren Aspekten.

A2.5 Domänenübergreifende Zusammenführung

Die relevanten Argumentationsstränge und Ergebnisse aller Domänen (Nutzenbewertung, Gesundheitsökonomie, Ethik, Soziales, Recht und Organisation sowie Umwelt- und Klimaaspekte) werden abschließend qualitativ zusammengeführt. Dieser Schritt wird durch eine Diskussionsrunde mit allen Berichterstellerinnen und -erstellern, die sich für die Domänen verantwortlich zeichnen, unterstützt.

Für die Zusammenführung der Ergebnisse der Domänen des ThemenCheck-Berichts wird ein logisches Modell in Anlehnung an INTEGRATE-HTA erstellt [129].

A3 Details der Ergebnisse: Nutzenbewertung

A3.1 Informationsbeschaffung

A3.1.1 Fokussierte Informationsbeschaffung nach systematischen Übersichten

Von den 4 identifizierten systematischen Übersichten, denen 6 Publikationen zugeordnet werden konnten (siehe Abschnitt A7.1.1), wurde keine als Basis-SÜ herangezogen.

A3.1.2 Umfassende Informationsbeschaffung

A3.1.2.1 Primäre Informationsquellen

A3.1.2.1.1 Bibliografische Datenbanken

Abbildung 2 zeigt das Ergebnis der systematischen Literaturrecherche in den bibliografischen Datenbanken und der Studienselektion gemäß den Kriterien zum Studieneinschluss. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Abschnitt A8.1.1. Die letzte Suche fand am 29.08.2025 statt.

Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt A7.1.2.

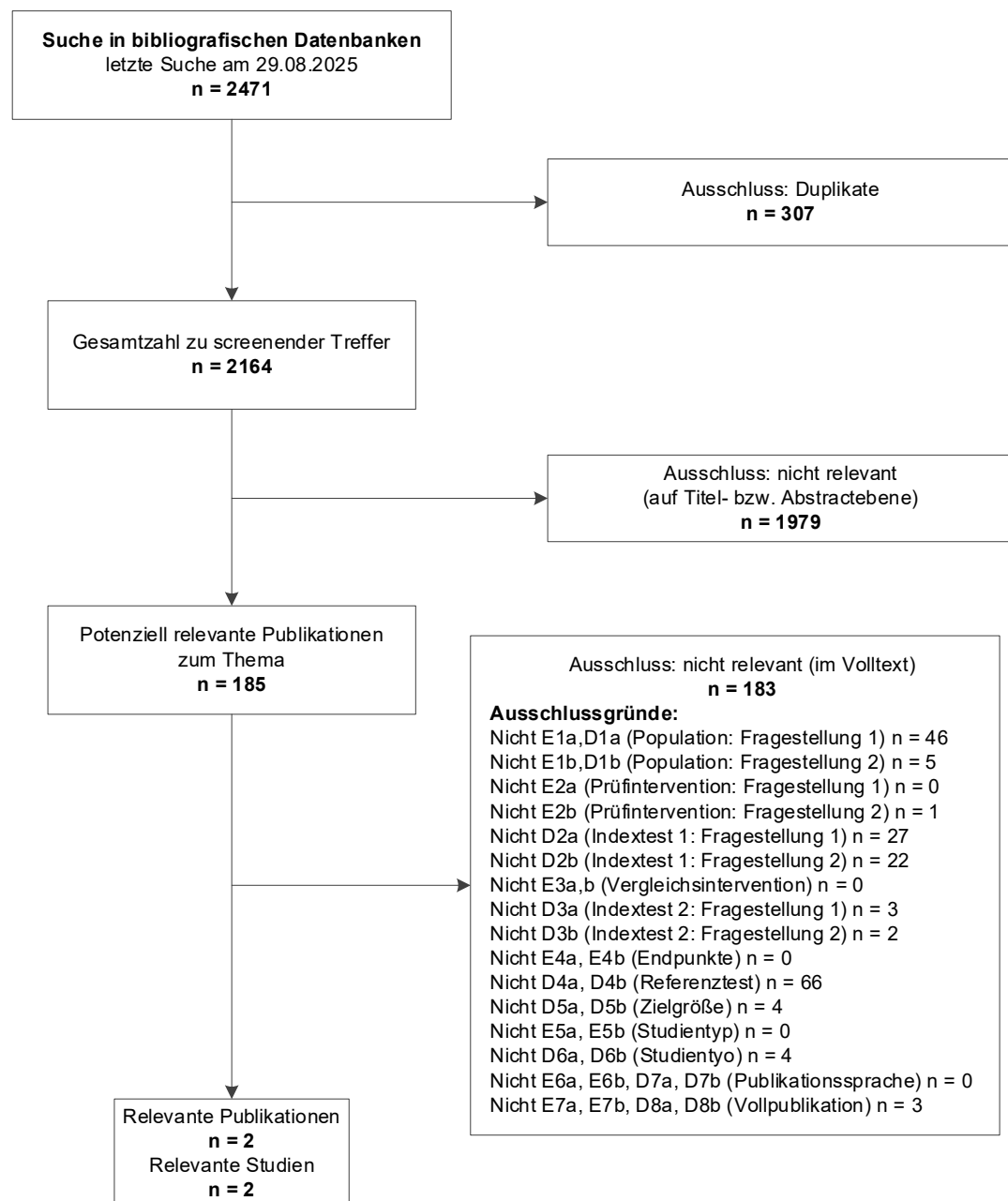


Abbildung 2: Ergebnis der umfassenden Informationsbeschaffung aus den bibliografischen Datenbanken und der Studienselktion – Nutzenbewertung

A3.1.2.1.2 Studienregister

Durch die Suche in Studienregistern wurde folgende relevante Studie bzw. Dokument identifiziert (Tabelle 7):

Tabelle 7: In Studienregistern identifizierte relevante Studien bzw. Dokumente

Studie	Studienregister ID	Studienregister	Ergebnisbericht in Studienregister vorhanden
Lo 2018	NCT02440139	ClinicalTrials.gov [160]	ja [160]

In den Studienregistern wurden darüber hinaus Studien ohne berichtete Ergebnisse identifiziert (siehe Abschnitt A3.1.4).

Die Suchstrategien für die Suche in Studienregistern finden sich in Abschnitt A8.1.2. Die letzte Suche in Studienregistern fand am 29.09.2025 statt.

A3.1.2.2 Weitere Informationsquellen und Suchtechniken

Über weitere Informationsquellen und Suchtechniken identifizierte relevante Studien bzw. Dokumente werden nachfolgend nur dargestellt, wenn sie nicht bereits über die primären Informationsquellen gefunden wurden.

A3.1.2.2.1 Anwendung weiterer Suchtechniken

Im Rahmen der Informationsbeschaffung wurden systematische Übersichten identifiziert – die entsprechenden Referenzen finden sich in Abschnitt A7.1.1. Die Referenzlisten dieser systematischen Übersichten wurden gesichtet.

Es fanden sich keine relevanten Studien bzw. Dokumente, die nicht über andere Rechenschritte identifiziert werden konnten.

Es wurde 1 abgeschlossene Studie ohne berichtete Ergebnisse identifiziert (siehe Abschnitt A3.1.4).

A3.1.3 Resultierender Studienpool

Durch die verschiedenen Rechenschritte konnten insgesamt 2 relevante Studien identifiziert werden (siehe auch Tabelle 8).

Tabelle 8: Resultierender Studienpool der Nutzenbewertung

Studie	Verfügbare Dokumente			
	Vollpublikation (in Fachzeitschriften)	Registereintrag / Ergebnisbericht aus Studienregistern	Studienbericht aus Herstellerunterlagen (nicht öffentlich zugänglich)	Sonstige Dokumente
Lee 2025	ja [23]	nein	nein	nein
Lo 2018	ja [24]	ja [160] / ja [160]	nein	nein

A3.1.4 Studien ohne berichtete Ergebnisse

In Tabelle 9 sind alle durch die Informationsbeschaffung identifizierten Studien ohne bisher berichtete Ergebnisse dargestellt.

Tabelle 9: In der Informationsbeschaffung identifizierte Studien ohne berichtete Ergebnisse

Studie	Dokumentart, Zitat	Studientyp	Geplante Fallzahl	Medizinprodukt, Hersteller, Land	Status (ggf. geplantes Studienende)
Fragestellung 1: Verdachtspopulation					
Es wurde keine Studie ohne berichtete Ergebnisse zur Fragestellung 1 identifiziert.					
Fragestellung 2: Screeningpopulation					
NCT06514248	Studienregistereintrag [161]	Studie zur diagnostischen Güte	350	V5med Lung AI, V5med, Taiwan ^a	abgeschlossen (05/2024)
NCT06751576	Studienregistereintrag [162]	Studie zur diagnostischen Güte	480	eyonis Lung Cancer Screening (LCS), Median Technologies, Frankreich ^b	abgeschlossen (03/2025)
NCT04119960	Studienregistereintrag [163]	Studie zur diagnostischen Güte	250	InferRead Lung CT AI, Infervision, China	abgeschlossen (10/2019)
a. Dies ist kein CE-gekennzeichnetes Medizinprodukt, es besteht jedoch eine FDA-Zulassung. b. Das Medizinprodukt befindet sich aktuell im Zulassungsverfahren der CE-Kennzeichnung [164] AI: künstliche Intelligenz; CT: Computertomografie					

A3.2 Charakteristika der in die Bewertung eingeschlossenen Studien

A3.2.1 Studiendesign und Studienpopulationen

Tabelle 10: Charakterisierung der eingeschlossenen Studien zur diagnostischen Güte (Fragestellung 2)

Studie	Studiendesign		Personenanzahl	Ort und Zeitraum der Durchführung	Dauer Nachbeobachtung
	Einschluss Studienpopulation ^a	Festlegung Indextests			
Lee 2025	retrospektiv, konsekutiv	unklar	366	Südkorea; k. A. ^b	5 Jahre
Lo 2018	retrospektiv, nicht konsekutiv	unklar	324	USA; 2015–2017	Mittelwert 6,5 Jahre
a. Bewertet wird, ob der Einschluss der Personen prospektiv bzw. retrospektiv und konsekutiv bzw. nicht konsekutiv erfolgte. b. Die Studienteilnehmer entsprachen denen aus einer vorausgegangenen Studie, die von 05/2017 bis 12/2018 durchgeführt wurde. k. A.: keine Angaben					

Tabelle 11: Index- und Referenzstandard (Fragestellung 2) (mehrseitige Tabelle)

Studie	Indextests	Referenztests
Lee 2025	Indextest I: AI-unterstützte Befundung durch Radiologin oder Radiologen <i>AI – Produkt:</i> LuCAS-Plus, Monitor Corporation <i>AI – Unterstützung:</i> Radiologe oder Radiologin sieht CT-Bild mit Markierungen der AI-detektierten Lungenrundherde sowie die Lung-RADS Bewertung und die Angaben zur Größe und Dichte des Lungenrundherdes von der AI <i>Anzahl Bewertende:</i> 5 <i>Expertise:</i> 3 Radiologinnen und Radiologen mit einer Spezialisierung auf den Thorax und 3–5 Jahre Erfahrung / 2 Radiologinnen oder Radiologen, die sich in der Weiterbildung zur Spezialisierung auf den Thorax befinden <i>Schwellenwert für AI-Markierung eines Lungenrundherdes:</i> k. A. Indextest II: Radiologin oder Radiologe <i>Anzahl Bewertende:</i> 5 <i>Expertise:</i> 3 Radiologinnen und Radiologen mit einer Spezialisierung auf den Thorax und 3–5 Jahre Erfahrung / 2 Radiologinnen oder Radiologen, die sich in der Weiterbildung zur Spezialisierung auf den Thorax befinden Beide Indextests: <i>Trennwert für positives Testergebnis pro Person:</i> Lung-RADS Kategorie ≥ 3 <i>Reihenfolge Indextests:</i> Die Reihenfolge der Indextests wurde pro Radiologin oder Radiologen randomisiert. Ein Teil der zu bewertenden CT-Aufnahmen werden in der ersten Runde nach Indextest 1 und der andere Teil nach Indextest 2 bewertet. In der 2. Bewertungsrunde werden die CT-Aufnahmen, die zuvor mit Indextest 1 bewertet wurden mit Indextest 2 bewertet und umgekehrt. <i>Abstand zwischen Indextest I und II:</i> 1 Monat <i>Bildgebendes Verfahren (Strahlendosis / Schichtdicke / Rekonstruktionskernel / Fensterung):</i> CT (Low-Dose / 1 mm / Hoch-Frequenz Kernel (Br59d, Somatom Force; YC, IQon Spectral CT) / k. A.)^a	Angaben aus elektronischer Patientenakte: Lungenkrebs innerhalb der letzten 5 Jahre nach der LDCT Untersuchung Angaben aus elektronischer Patientenakte: Lungenkrebs innerhalb der letzten 5 Jahre nach der Untersuchung
Lo 2018	Indextest I: AI-unterstützte Befundung durch Radiologin oder Radiologe <i>AI – Produkt:</i> ClearRead CT Vessel Suppression und ClearRead CT InSight, Riverain Technologies <i>AI – Unterstützung:</i> Original und AI verarbeitetes CT-Bild, in der die Gefäße unterdrückt werden, werden nebeneinander abgebildet. In beiden CT-Bildern sind AI-detektierte Lungenrundherde markiert <i>Anzahl Bewertende:</i> 12 <i>Expertise:</i> 6–26 Jahre <i>Schwellenwert für AI-Markierung eines Lungenrundherdes:</i> ≥ 5 mm (Durchmesser)^b	Histopathologischer Befund der Lungenrundherde bzw. Nachbeobachtung^c

Tabelle 11: Index- und Referenzstandard (Fragestellung 2) (mehrseitige Tabelle)

Studie	Indextests	Referenztests
	Indextest II: Radiologinnen und Radiologen <i>Anzahl Bewertende:</i> 12 <i>Expertise:</i> 6–26 Jahre Beide Indextests: <i>Trennwert für positives Testergebnis pro Person:</i> Markierung von Lungenrundherden in der CT Aufnahme. Die Markierung muss dabei innerhalb des durch ein Konsensuspanel festgelegten ellipsoiden Bereichs eines behandlungsbedürftigen Lungenrundherdes liegen <i>Reihenfolge Indextests:</i> Dieselbe Radiologin oder derselbe Radiologe führt die Befundung zunächst allein (Indextest 2) und dann erneut, diesmal mit AI Unterstützung durch (Indextest 1) <i>Befundergebnisse Indextest 1 und 2:</i> Radiologin oder Radiologe markiert die jeweilige Mitte der identifizierten Lungenrundherde und gibt eine Einschätzung des Verdachtgrades (Skala: 1–100) sowie Empfehlung für die weitere Diagnostik (z. B. CT-Nachuntersuchung, PET/CT oder Biopsie) ab. <i>Abstand zwischen Indextest I und II:</i> mindestens 37 Tage (MW 57 Tage) <i>Bildgebendes Verfahren (Strahlendosis / Schichtdicke / Rekonstruktionskernel / Fensterung):</i> CT (Low-Dose / ≤ 3 mm / Standard / Fensterung einstellbar für Lungen- oder Weichteilfenster)^d	Histopathologischer Befund bzw. Nachbeobachtung^c
a. 100 % der Personen stammen aus der K-LUCAS [165]. b. Lo 2018 beschrieb mit ≥ 4 mm ein zusätzliches Kriterium. Es bleibt unklar, welches Kriterium bei der AI-Markierung schlussendlich verwendet wurde. c. Nach Lo 2018 wurden „actionable nodules“ als Lungenrundherde identifiziert, bei denen 2 von 3 unabhängige Experten eine weitere Diagnostik als notwendig erachteten. Diese Personen wurden im Durchschnitt 6,5 Jahre nachbeobachtet. d. 85 % der Personen stammen aus der NLST [166]. AI: künstliche Intelligenz; CT: Computertomografie; k. A.: keine Angabe; K-LUCAS: Korean Lung-Cancer Screening Project; LDCT: Low-Dose-Computertomografie; Lung-RADS: Lung Imaging Reporting and Data System; MW: Mittelwert; NLST: National Lung Screening Trial		

Tabelle 12: Ein- / Ausschlusskriterien für Personen in den Studien (Fragestellung 2)

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Lee 2025	- Low-Dose-Thorax-CT-Aufnahmen aus einem Studienzentrum der K-LUCAS: - Alter: 55–74 Jahre ≥ 30 Packungsjahre - Rauchfrei: seit maximal 15 Jahre vor Studienbeginn	keine Ausschlusskriterien definiert
Lo 2018	- Alter: 55–77 Jahre - Primärer Lungenkrebs, durch Biopsie bestätigt - Asymptomatische Personen, die rauchen oder geraucht haben - Low-Dose-Thorax-CT-Aufnahmen aus der NLST^a (84 %), und von den Universitätskliniken Cleveland (14 %) und Maryland (4 %) - Randomisierte Auswahl der CT-Aufnahmen durch 3 Experten unter folgenden Kriterien: 2:1 Verteilung bez. CT-Aufnahmen ohne auffälligen Befund und CT-Aufnahmen mit Lungenrundherd - Max 5 Lungenrundherde / CT-Aufnahme - Größe der Lungenrundherde liegt zwischen 5–20mm	- Symptomatische Patientinnen und Patienten mit Komorbiditäten - Akuter Pneumothorax

a. Einschlusskriterien der NLST waren u. a. ein Alter zwischen 55–74 Jahren, > 30 Packungsjahre sowie auf Personen, die seit maximal 15 Jahren vor Studienbeginn rauchfrei sind.

CT: Computertomografie; K-LUCAS: Korean Lung-Cancer Screening Project; NLST: National Lung Screening Trial

Tabelle 13: Charakterisierung der Studienpopulationen (Fragestellung 2)

Studie	N	Alter [Jahre] MW (SD)	Geschlecht [w / m] %	Zigarettenkonsum [Packungsjahre] MW (SD)	Lungen- rundherde pro Person MW (SD)	Studienprävalenz von Lungenkrebs %
Lee 2025	366	64 (5)	2 / 98	43 (14)	1,9 (2,5)	1,4
Lo 2018	324	k. A.	k. A.	k. A.	k. A. ^a	28,7 ^b

a. 93 Personen haben 95 maligne Lungenrundherde. 36 Personen mit malignem Lungenrundherd haben zusätzlich 1 bis 4 benigne Lungenrundherde. 15 Personen haben nur benigne Lungenrundherde und 216 Personen haben keine Lungenrundherde oder einen normalen Befund.

b. eigene Berechnung

k. A.: keine Angaben; m: männlich; MW: Mittelwert; N: Anzahl eingeschlossener Personen; SD: Standardabweichung; w: weiblich

A3.2.2 Bewertung des Verzerrungspotenzials der Studien zur diagnostischen Güte der AI-unterstützten Befundung im Vergleich zur ausschließlichen Befundung durch Radiologinnen und Radiologen

Die Einschätzung des Verzerrungspotenzials nach QUADAS 2 ist in der folgenden Tabelle 14 dargestellt.

Tabelle 14: Verzerrungspotenzial nach QUADAS 2 der Studien zur diagnostischen Güte (Fragestellung 2)

Studie	Patienten-selektion (Domäne 1)	Indextest I (Domäne 2)	Indextest II (Domäne 2)	Referenz-standard (Domäne 3)	Patienten-fluss und zeitlicher Ablauf (Domäne 4)	Zusammen-fassende Einschätzung
Lee 2025	niedrig	unklar ^a	unklar ^a	niedrig	hoch ^b	hoch
Lo 2018	hoch ^c	hoch ^d	hoch ^d	niedrig	hoch ^e	hoch

a. Es ist unklar, ob der Trennwert prospektiv festgelegt wurde. Für die Studie liegt kein Registereintrag vor.
b. Der Referenztest umfasst eine Nachbeobachtungsdauer von 5 Jahren. Eine eindeutige Zuordnung der in der ersten Screeningrunde als testpositiv bewerteten Lungenrundherde zu den malignen Lungenrundherden liegt nicht vor. Standardisierte Nachbeobachtung im ersten Jahr des Follow-up nach der ersten Screeningrunde, für die weiteren 4 Jahre keine Angaben.
c. Keine konsekutive Stichprobe. Es lag ein Fall-Kontroll-Design und inadäquate Studienausschlüsse vor.
d. Unklar ob die Kriterien für das testpositive Ergebnis des Indextests prospektiv festgelegt wurden. Unklar ob die bewertende Person gegenüber dem Referenztest verblindet war.
e. Die Zeitspanne zwischen CT-Aufnahme und der Biopsie als Referenztest wurde nicht berichtet. Im Mittel betrug die Nachbeobachtungsdauer 6,5 Jahre. Es ist unklar, ob alle Personen ohne malignen Befund den gleichen Referenztest, d. h. gleiche Nachbeobachtungsdauer, erhalten haben. Es wurden nicht alle Personen und keine benignen Lungenrundherde in die Analyse einbezogen.

CT: Computertomografie

Die Einschätzung der Bedenken bezüglich der Übertragbarkeit der Ergebnisse nach QUADAS 2 ist in der folgenden Tabelle 15 dargestellt.

Tabelle 15: Bedenken bezüglich der Übertragbarkeit QUADAS 2 der Studien zur diagnostischen Güte (Fragestellung 2)

Studie	Patienten- selektion (Domäne 1)	Indextest I (Domäne 2)	Indextest II (Domäne 2)	Referenz- standard (Domäne 3)	Zusammen- fassende Einschätzung
Lee 2025	unklar ^a	stark ^c	gering	gering	gering
Lo 2018	stark ^b	stark ^{c, d}	stark ^d	gering	stark
a. Es werden keine histologischen Subtypen und Krankheitsursachen berichtet. Diese können sich jedoch zwischen Europa und Asien unterscheiden. b. Es ist eine selektive Fall-Kontroll-Studie mit erhöhter Prävalenz. Zudem wurden Patientinnen und Patienten mit Komorbiditäten ausgeschlossen. c. keine Angaben zur untersuchten AI Version. d. Der herangezogene Marker für ein testpositives Ergebnis des Indextests entspricht nicht der üblichen radiologischen Befundung. AI: künstliche Intelligenz					

A3.3 Ergebnisse der Studien zur diagnostischen Güte

A3.3.1 Ergebnisse zur diagnostischen Güte der AI-unterstützten Befundung im Vergleich zur ausschließlichen Befundung durch Radiologinnen und Radiologen (Fragestellung 2)

Tabelle 16: Ergebnisse der Studien zur diagnostischen Güte der AI-unterstützten Befundung im Vergleich zur Befundung ausschließlich durch Radiologinnen und Radiologen (Fragestellung 2)

Studie	N	n	Indextest	Trennwert	Referenztest	RP	FN	FP	RN	Sensitivität in %	[95 %-KI]	Spezifität in %	[95 %-KI]
Screeningrunde													
Lee 2025	366	366	AI-unterstützte Befundung	Lung-RADS ≥ 3	Lungenkrebsdiagnose vermerkt in der elektronische Patientenakte innerhalb von 5 Jahren nach dem Screening	3	2	108	253	60,0	[23,1; 88,2] ^a	70,1 ^a	[65,2; 74,6] ^a
			Radiologin oder Radiologe	Lung-RADS ≥ 3	Lungenkrebsdiagnose vermerkt in der elektronische Patientenakte innerhalb von 5 Jahren nach dem Screening	4	1	77	284	80,0	[37,6; 96,4] ^a	78,7 ^a	[74,2; 82,6] ^a
Einmalige Diagnostik													
Lo 2018	324	311 ^b	AI-unterstützte Befundung	Markierung Lungenrundherd im CT	Histopathologischer Befund bzw. Nachbeobachtung	k. A.	k. A.	k. A.	k. A.	80,0 ^c	[72,4; 87,6] ^{c, d}	84,4 ^c	[80,5; 88,3] ^{c, d}
			Radiologin oder Radiologe	Markierung Lungenrundherd im CT	Histopathologischer Befund bzw. Nachbeobachtung	k. A.	k. A.	k. A.	k. A.	64,7 ^{c, d}	[57,1; 72,3] ^{c, e}	89,9 ^c	[86,0; 93,8] ^{c, e}
a. eigene Berechnung; 95 %-KI nach Wilson b. 15 Personen mit benignen Tumoren, fließen nicht in die Analyse ein. Es wurden 95 maligne Befunde (93 CT-Aufnahmen) und 216 CT-Aufnahmen mit normalem Befund in der Analyse berücksichtigt. c. Multireader-multicase Analyse nach Dorfman-Berbaum-Metz [12] für 12 Reader und 311 Personen d. Die Sensitivität wurde aus Tabelle 2 in Lo 2018 extrahiert. Im Text werden 64,5 % als Sensitivität genannt. e. eigene Berechnung aus Sensitivität und zugehöriger Standardabweichung mittels Normalapproximation AI: künstliche Intelligenz; FN: falsch-negativ; FP: falsch-positiv; k. A.: keine Angaben; KI: Konfidenzintervall; n: Zahl ausgewerteter Personen; N: Zahl eingeschlossener Personen; RN: richtig-negativ; RP: richtig-positiv													

A3.3.2 Ergänzendes Ergebnis zur diagnostischen Güte der alleinigen AI-Befundung

Tabelle 17: Ergebnis zur diagnostischen Güte der alleinigen AI-Befundung (Fragestellung 2)

Studie	N	n	Indextest	Trennwert	Referenztest	RP	FN	FP	RN	Sensitivität in %	[95 %-KI]	Spezifität in %	[95 %-KI]
Lee 2025	366	366	Alleinige AI-Befundung	Lung-RADS ≥ 3	Lungenkrebsdiagnose vermerkt in der elektronische Patientenakte innerhalb von 5 Jahren nach dem Screening	4	1	122	239	80,0	[37,6; 96,4] ^a	66,2 ^a	[61,2; 70,9] ^a
a. eigene Berechnung; 95 %-KI nach Wilson AI: künstliche Intelligenz; FN: falsch-negativ; FP: falsch-positiv; KI: Konfidenzintervall; n: Zahl ausgewerteter Personen; N: Zahl eingeschlossener Personen; RN: richtig-negativ; RP: richtig-positiv													

A3.3.3 Ergänzend dargestellter Endpunkt Zeitersparnis

Tabelle 18: Ergebnisse – Ergänzend der Endpunkt Zeitersparnis (Fragestellung 2)

Studie Operationalisierung Indextests	N	Werte		Indextest 1 vs. Indextest 2		
		MW	[95 %-KI]	Differenz	[95 %-KI]	p-Wert
Interpretationszeit in Sekunden ^a						
Lee 2025						
AI-unterstützte Befundung	366	161	[149; 172]	3 ^b	k. A.	0,73
Radiologin oder Radiologe	366	164	[154; 175]			
AI allein ^c	366	k. A.	k. A.	k. A.	k. A.	k. A.
Lo 2018						
AI-unterstützte Befundung	3888 ^d	98,0	[88,0; 108,0]	34,3	[15,2; 53,5]	< 0,01
Radiologinnen und Radiologen	3888 ^d	132,3	[122,3; 142,3]			
a. Die Interpretationszeit wird in Lee 2025 durch die digitale Bewertungssoftware ab dem Zeitpunkt des Hochladens des Falles bis zur abschließenden Speicherung der Fallbewertung automatisch erfasst. In Lo 2018 umfasst die Interpretationszeit die Sichtung und Bewertung der jeweiligen CT-Aufnahme.						
b. eigene Berechnung						
c. Die alleinige Befundung durch AI wird ergänzend dargestellt.						
d. Es wurden CT-Aufnahmen von 324 Personen, die an einem Lungenkrebsfrüherkennungsscreening teilnahmen, von jeweils von 12 Radiologinnen und Radiologen ausgewertet.						
AI: künstliche Intelligenz; MW: Mittelwert; N: Anzahl der durchgeführten Befundungen von CT-Aufnahmen; KI: Konfidenzintervall						

A4 Details der Ergebnisse: Gesundheitsökonomische Bewertung

A4.1 Bestimmung der Interventionskosten

In Tabelle 19 werden die möglichen Kosten bei einer Abrechnung von AI-Anwendungen allgemein im privatärztlichen Bereich näherungsweise durch eine Analogabrechnung gemäß GOÄ dargestellt. In Tabelle 20 und Tabelle 21 werden Kosten verschiedener AI-Systeme aus Sicht der Leistungserbringenden dargestellt.

Tabelle 19: Kosten für AI-Anwendungen aus Sicht von Selbstzahlenden

Bezeichnung der Intervention	Bezeichnung und Kennzeichnung im relevanten Vergütungskatalog	Kosten pro Anwendung in €	Begründung (Quelle) Bezugsjahr	Erstattungsfähigkeit
AI-Unterstützung	Analog GOÄ-Ziffer 5733: Zuschlag für computergesteuerte Analyse (z. B. Kinetik, 3D-Rekonstruktion)	46,63 (Einfachsatz)	GOÄ [167] Wigge 2023 [168]	nein
AI: künstliche Intelligenz; GOÄ: Gebührenordnung für Ärzte				

Tabelle 20: Kosten verschiedener AI-Systeme zur Unterstützung der Befundung von CT-Aufnahmen aus Sicht der Leistungserbringenden

Produkt (Unternehmen)	Kosten je Befundung	Quelle	Kommentar
ClearRead CT (Riverain)	2,52 € ^{a, b}	Geppert 2025 [34]	Aufgrund verschiedener Kostenmodelle wurden die Herstellerangaben auf die Kosten pro Befundung umgerechnet.
InferRead CT Lung (Infervision)	4,21 € ^{a, b}		
Veye Chest ^c (Aidence)	6,30 €–9,45 € ^a	NICE MIB243 2021 [32]	Die Kosten pro Aufnahme variieren je nach gewählten Funktionen und Volumen.
Veolity (MeVis)	Anstelle von Kosten je Befundung lassen sich bei Veolity Kostenangaben finden ^d : ▪ unbefristete Lizenz: 55 425,47 € ^a ▪ Installation, Tests und Schulungen: 11 337,03 € ^a ▪ Jährliche Wartungskosten (ab 2. Jahr): 11 085,09 € ^a Jährliche Supportkosten: 7558,02 € ^a		-
a. kaufkraftbereinigt in € umgerechnet [36] und mittels Verbraucherpreisindex vom Basisjahr 2021 auf das Jahr 2024 inflationiert [37] b. ohne Kosten für Implementierung und Wartung c. mittlerweile durch Veye Lung Nodules ersetzt d. kein Kostenmodell mit pay-per-use als Komponente AI: künstliche Intelligenz; CT: Computertomografie; MIB: medtech innovation briefing; NICE: National Institute for Health and Care Excellence			

Tabelle 21: Kosten verschiedener AI-Systeme zur Unterstützung der Befundung von Röntgenaufnahmen für Leistungserbringende

Produkt (Unternehmen)	Implementierungskosten ^a	Jährliche Gebühren (bei 25 000 Aufnahmen)	Kosten je Befundung	Gesamtkosten im 1. Jahr	Gesamtkosten je Befundung im 5- Jahres-Durchschnitt	Quelle	Kommentar
Annalise Enterprise CXR (Annalise.ai)	6298,35 €– 31 491,75 € ^b	64 558,08 € ^b	-	83 453,13 € ^b	2,73 € ^b	NICE X-Ray 2023 [33]	Die Kosten beziehen sich – sofern nicht anders angegeben – auf Kosten pro NHS Trust. Den dargestellten Kosten liegt der beispielhafte Fall von 25 000 Aufnahmen im Jahr zugrunde. Abhängig vom Volumen können die Kosten unterschiedlich variieren.
red dot (Behold.ai)	12 596,70 € ^{b, c}	75 580,19 € ^{b, c}	-	88 176,89 € ^b	3,12 € ^b		
InferRead DR Chest (Infervision)	3779,01 € ^b	Lizenzgebühr: 20 154,72 € ^b Wartung: 7558,02 € ^b	-	31 491,75 € ^b	1,13 € ^b		
InferRead DR Chest (Infervision)	3779,01 € ^b	Wartung: 7558,02 € ^b	1,26 € ^b	42 828,78 € ^b	1,56 € ^b		
Lunit INSIGHT CXR (Lunit)	7558,02 € ^b	21 099,47 €– 62 983,49 € ^b	0,84 €– 2,52 € ^b	28 657,49 €– 70 541,51 € ^b	0,91 €–2,58 € ^b		
AI-Rad Companion Chest X-ray (Siemens Healthineers)	3023,21 € ^b	15 116,04 € ^b	-	18 139,25 € ^b	0,63 € ^b		

a. Die Implementierungskosten beinhalten Leistungen für die Installation, die Integration in bereits bestehende PACS und RIS sowie Schulungen für die Mitarbeiterinnen und Mitarbeiter.

b. kaufkraftbereinigt in € umgerechnet [36] und mittels Verbraucherpreisindex vom Basisjahr 2021 auf das Jahr 2024 inflationiert [37]

c. Kosten je Krankenhaus (1 NHS Trust kann mehrere Krankenhäuser bzw. Einrichtungen betreiben)

AI: künstliche Intelligenz; NHS: National Health Service; NICE: National Institute for Health and Care Excellence; PACS: Bildarchivierungs- und Kommunikationssystem; RIS: Radiologieinformationssystem

A4.2 Systematische Übersicht gesundheitsökonomischer Evaluationen

A4.2.1 Fokussierte Informationsbeschaffung

A4.2.1.1 Primäre Informationsquellen

Abbildung 3 zeigt das Ergebnis der fokussierten Literaturrecherche in den bibliografischen Datenbanken und der Studienselektion gemäß den Kriterien zum Studieneinschluss. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Abschnitt A8.2. Die letzte Suche fand am 07.08.2025 statt.

Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt A7.2.2.

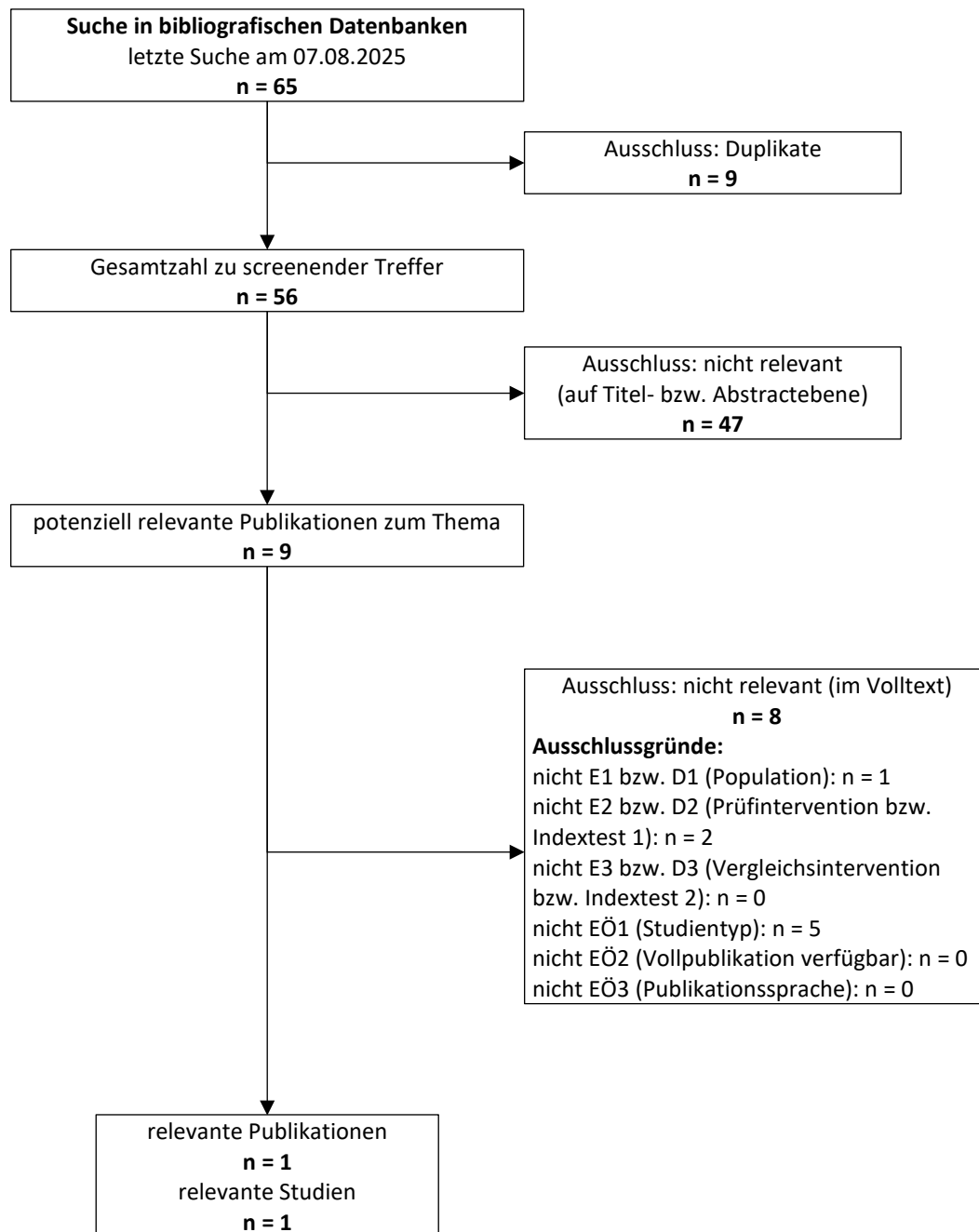


Abbildung 3: Ergebnis der bibliografischen Recherche und der Studienselektion – gesundheitsökonomische Evaluationen

A4.2.1.2 Weitere Informationsquellen und Suchtechniken

Über weitere Informationsquellen und Suchtechniken identifizierte relevante Studien bzw. Dokumente werden nachfolgend nur dargestellt, wenn sie nicht bereits über die primären Informationsquellen gefunden wurden.

Anwendung weiterer Suchtechniken

Im Rahmen der Informationsbeschaffung wurden systematische Übersichten identifiziert – die entsprechenden Referenzen finden sich in Abschnitt A7.2.1. Die Referenzlisten dieser systematischen Übersichten wurden gesichtet.

Es wurden folgende relevante Dokumente identifiziert, die nicht über andere Rechenschritte gefunden werden konnten:

Tabelle 22: Durch Anwendung weiterer Suchtechniken identifizierte relevante Dokumente

Studie	Verfügbare Dokumente ([Zitat])
Geppert 2025	NICE Final scope [35] Protokoll [41] Diagnostics Assessment Report [40] NICE Diagnostics Guidance [42]
NICE: National Institute for Health and Care Excellence	

A4.2.1.3 Resultierender Studienpool

Durch die verschiedenen Rechenschritte konnte insgesamt 1 relevante Studie identifiziert werden (siehe auch Tabelle 23).

Tabelle 23: Studienpool der gesundheitsökonomischen Bewertung

Studie	Verfügbare Dokumente [Zitat]
Geppert 2025	NICE Final scope [35] Protokoll [41] Diagnostics Assessment Report [40] NICE Diagnostics Guidance [42] Extended Research Article [34]
NICE: National Institute for Health and Care Excellence	

A4.2.2 Charakteristika der in die Bewertung eingeschlossenen Studien

A4.2.2.1 Studiendesign

In Tabelle 24 werden Studiencharakteristika der eingeschlossenen gesundheitsökonomischen Studien dargestellt.

Tabelle 24: Studiencharakteristika

Studie	Studiendesign und Vorgehensweise	Studienpopulation	Strategien		Land und Versorgungskontext	Endpunkt Kosteneffektivität	Studienfinanzierung
			Prüfintervention	Vergleichsintervention			
Geppert 2025	Kosten-Effektivitäts-Analyse / entscheidungsanalytische Modellierung (preliminary model)	Screening-population	CT-Auswertung durch Radiologin bzw. Radiologen + AI (ClearRead CT)	CT-Auswertung durch Radiologin bzw. Radiologen	UK	ICER: Kosten pro korrekt identifizierte Person mit handlungsbedürftigem Lungenrundherd	Beauftragt vom NICE und im Rahmen des Evidence Synthesis Programme finanziert (Fördernummer: NIHR135325)
	Kosten-Effektivitäts- bzw. Kosten-Nutzwert-Analyse / entscheidungsanalytische Modellierung (full model)	Screening- sowie Verdachts-population	CT-Auswertung durch Radiologin bzw. Radiologen + AI (ClearRead CT für Screening-population und InferRead CT für Verdachts-population)	CT-Auswertung ohne AI-Einsatz		ICERs: Kosten pro ▪ QALY (primärer Endpunkt) ▪ korrekt identifizierte Person mit handlungsbedürftigem Lungenrundherd (sekundärer Endpunkt) ▪ korrekt identifizierte und behandelte Person mit Lungenkarzinom (sekundärer Endpunkt)	
AI: künstliche Intelligenz; CT: Computertomografie; ICER: inkrementelles Kosteneffektivitätsverhältnis; NICE: National Institute for Health and Care Excellence; QALY: qualitätsadjustiertes Lebensjahr; UK: Vereinigtes Königreich							

In Tabelle 25 werden die Parameter der zugrunde liegenden Modellierung dargestellt.

Tabelle 25: Modell (mehrseitige Tabelle)

Studie	Modellierungs- technik	Anzahl und Art der Gesundheitszustände / Ereignisse	Zeithorizont	Wesentliche Annahmen	Umgang mit Unsicherheit / Aussagen zur Modellvalidierung
Geppert 2025	Entscheidungsbaum (preliminary model)	▪ handlungsbedürftig Lungenrundherd vorhanden: ja / nein ▪ Befund: Richtig oder falsch-positiv bzw. -negativ ▪ Größe des detektierten Lungenrundherdes: 5–8 mm / ≥ 8 mm oder ≥ 300 mm³	initiale Screening-Untersuchung	▪ Gehalt einer Röntgenassistenz als Proxy für das Gehalt eines Radiologen oder einer Radiologin ▪ Befundung mit AI-Unterstützung 2 Minuten schneller als Befundung ohne AI-Unterstützung	deterministische Sensitivitätsanalysen: ▪ univariate Analyse (Tornado-Diagramm) durch Variation der Input-Parameter anhand vorhandener Spannen bzw. durch arbiträre Anpassung von ± 10 % (Kosten für CT) bzw. ± 50 % (Zeit für Befundung), ▪ Szenarioanalysen durch Variation der Prävalenz handlungsbedürftiger Lungenrundherde (alternative Quelle) und der aufgewendeten Arbeitszeit für die Befundung (gleiche Dauer, Zeitersparnis bei Befundung ohne AI-Unterstützung)
	Entscheidungsbaum (full model)	▪ Lungenrundherd vorhanden: ja / nein ▪ Befund: Richtig oder falsch-positiv bzw. -negativ ▪ Art des Lungenrundherdes: solide / subsolide	Lebenszeit (Endpunkte: QALY und Lungenkarzinom) bzw. initiale Untersuchung (Endpunkt: handlungsbedürftiger Lungenrundherd)	▪ wenn mehrere Lungenrundherde bei einer Person identifiziert werden, wird der größte Lungenrundherd (gemäß BTS) als maßgeblich für das weitere Vorgehen betrachtet ▪ nicht detektierte Lungenrundherde konnten gutartig oder bösartig sein ▪ detektierte gutartige Lungenrundherde weisen kein (weiteres) Wachstum auf ▪ Kategorisierung als subsolider oder solider Lungenrundherd immer korrekt ▪ Prävalenz Lungenkarzinom bei detektierten Lungenrundherden bei Verdachtspopulation identisch zur Screeningpopulation	Deterministische Sensitivitätsanalysen für primären und sekundäre Endpunkte: ▪ univariate Analyse (Tornado-Diagramm) durch Variation der Input-Parameter anhand vorhandener Spannen bzw. durch arbiträre Anpassung von ± 10 % (Kosten für CT) bzw. ± 50 % (Zeit für Befundung) ▪ Szenarioanalysen: ▫ Variation der Prävalenz von Lungenrundherden im initialen CT (alternative Quelle für Screeningpopulation, arbiträre Anpassung bei Verdachtspopulation),

Tabelle 25: Modell (mehrseitige Tabelle)

Studie	Modellierungs- technik	Anzahl und Art der Gesundheitszustände / Ereignisse	Zeithorizont	Wesentliche Annahmen	Umgang mit Unsicherheit / Aussagen zur Modellvalidierung
		- Größe des detektierten soliden Lungenrundherdes: < 5 mm oder < 80 mm / 5–8 mm / ≥ 8 mm oder ≥ 300 mm³ - Größe des detektierten subsoliden Lungenrundherdes: < 5 mm / ≥ 5 mm - Behandlungspfad nach Befundung: kein Handlungsbedarf / Kontrolle (z. B. weitere CT) / weitere Diagnostik (z. B. Biopsie) / Behandlung - Malignität: gutartig / bösartig - Stadium: I–IV - Tod		- Gehalt einer Röntgenassistenz als Proxy für das Gehalt eines Radiologen oder einer Radiologin - Nutzwertverlust von 0,063 bei falsch-positiver Befundung bzw. bei CT-Überwachung trotz gutartigem Rundherd bis zum Ende der Überwachung - Personen ohne Lungenrundherd bzw. mit gutartigem Lungenrundherd mit Nutzwerten entsprechend der alters- und geschlechtsadjustierten Allgemeinbevölkerung - Nutzwertverlust von 0,2 bei Biopsie über einen Zeitraum von 3 Monaten - angenommene Gesamtmortalität unabhängig von Intervention und Population (Screening vs. Verdacht) - als initial falsch-negativ befundete Rundherde waren zu 0,04 % bösartig - es werden keine Karzinome durch die Strahlenbelastung verursacht - Befundung mit AI-Unterstützung 2 Minuten schneller als Befundung ohne AI-Unterstützung	- Variation der aufgewendeten Arbeitszeit für die Befundung (gleiche Dauer, Zeitersparnis bei Befundung ohne AI-Unterstützung), - Verwendung der diagnostischen Güte für Lungenrundherde ≥ 5 mm statt der diagnostischen Güte für Lungenrundherde allgemein und anschließendem Ausschluss von Lungenrundherden < 5 mm - Personen mit gutartigen Tumoren wurden nach CT-Kontrolle nach 2 (solide Rundherde) bzw. 4 (subsolide Rundherde) Kontrollen nicht weiter überwacht, - kein Nutzwertverlust bei falsch-positiver Befundung bzw. CT-Überwachung trotz gutartigem Rundherd Probabilistische Sensitivitätsanalyse für primären Endpunkt: Monte-Carlo-Simulation (10 000 Wiederholungen), Darstellung der Ergebnisse als CEAC
AI: künstliche Intelligenz; BTS: British Thoracic Society; CEAC: Kosten-Effektivitäts-Akzeptanz-Kurve; CT: Computertomografie					

A4.2.2.2 Inputparameter

In Tabelle 26 werden die in die Analyse eingehenden Daten zum Nutzen dargestellt.

Tabelle 26: Daten zum Nutzen

Studie	Klinische Endpunkte und deren Verwendung im Modell	Quellen [Zitat]
Geppert 2025 (preliminary modell)	Diagnostische Güte: Sensitivität und Spezifität zur Detektion handlungsbedürftiger Lungenrundherde (mit und ohne AI-Unterstützung)	Lo 2018 [24]
Geppert 2025 (full modell)	Diagnostische Güte: Sensitivität und Spezifität zur Detektion von Lungenrundherden (mit und ohne AI-Unterstützung)	Verdachtspopulation: Kozuka 2020 [44] Screeningpopulation: Hsu 2021 [43]
AI: künstliche Intelligenz		

In Tabelle 27 werden die in der Studie verwendeten Daten zu Nutzwerten dargestellt.

Tabelle 27: Daten zu Nutzwerten

Studie	Erhebungsinstrument (ggf. Tarif und Befragungskollektiv)	Quelle [Zitat]
Geppert 2025 (full modell)	Für Personen mit Lungenkarzinom: EQ-5D-3L (kanadischer Tarif basierend auf der Allgemeinbevölkerung)	Bajre 2017 [169] Naik 2015 [170]
	Für Personen ohne Lungenrundherde bzw. mit gutartigen Lungenrundherden: EQ-5D-3L (englischer Tarif basierend auf der Allgemeinbevölkerung)	Szende 2014 [171]
EQ-5D-3L: European Quality of Life – 5 Dimensions – 3 Levels		

In Tabelle 28 werden die in der Studie verwendeten Daten zu Kosten dargestellt.

Tabelle 28: Daten zu Kosten

Studie	Währung (Indexjahr)	Diskontrate	Perspektive	Kostenparameter <Quellen> (für Mengen und Preise)
Geppert 2025 (preliminary model)	GBP (2021/2022)	-	NHS and PSS	AI-Software [Information des Herstellers Riverain technologies] CT-Aufnahme mit / ohne Kontrastmittel [172] Personalkosten für Befundung [173]
Geppert 2025 (full model)	GBP (2021/2022)	3,5 %	NHS and PSS	AI-Software [Information des Herstellers Riverain technologies bzw. Infervision] CT-Aufnahme mit / ohne Kontrastmittel [172] Personalkosten für Befundung [173] Multidisziplinäre Fallkonferenz [172] Bildgesteuerte Nadelbiopsie [172] Positronen-Emissions-Tomografie [172] Behandlungskosten in Abhängigkeit vom Stadium [169]
AI: künstliche Intelligenz; CT: Computertomografie; GBP: Britische Pfund; NHS: National Health Service; PSS: Personal Social Services				

A4.2.3 Ergebnisse gesundheitsökonomischer Evaluationen

Die Ergebnisse der eingeschlossenen Studien werden in Tabelle 29 dargestellt. Die ermittelten Kosten pro Patientin bzw. Patient und / oder die inkrementellen Kosten-Nutzen-Verhältnisse werden jeweils entsprechend der in der Studie dargestellten Währung und dem entsprechenden Indexjahr angegeben. Die resultierenden Kostenangaben der eingeschlossenen Modelle werden zusätzlich kaufkraftbereinigt in Euro umgerechnet [36] und mittels des aktuellen Verbraucherpreisindex (VPI) des Statistischen Bundesamts auf das Jahr 2024 [37] inflationiert.

Tabelle 29: Ergebnisse Kosteneffektivität (mehrseitige Tabelle)

Studie	Kosten der Interventionen pro 1000 Personen mit CT-Untersuchung (Maß der Unsicherheit sofern angegeben)	Nutzen pro 1000 Personen mit CT-Untersuchung (Maß der Unsicherheit sofern angegeben)	inkrementelles Kosten-Nutzen-Verhältnis	Ergebnisse aus Sensitivitätsanalysen	Schlussfolgerungen der Autorinnen und Autoren
Geppert 2025 (preliminary model)	- AI-Unterstützung: 127 600 £ (160 733,87 €^a) - Ohne AI-Unterstützung: 130 500 £ (164 386,92 €^a)	- AI-Unterstützung: 149,3 korrekt identifizierte Personen mit handlungsbedürftigem Lungenrundherd - Ohne AI-Unterstützung: 123,8 korrekt identifizierte Personen mit handlungsbedürftigem Lungenrundherd	Befundung mit AI-Unterstützung ist überlegen	- Zeit(-ersparnis) bei Befundung als Haupttreiber des ICER - Auch bei den Szenarioanalysen bleibt das ICER in einem „akzeptablen“ Rahmen	Limitationen: - Diagnostische Güte würde anhand von Lungenrundherden bestimmt und dann auf Personen übertragen - Ausschließlich initiales Screening betrachtet - Ein komplexeres Modell ist notwendig (siehe full model)

Tabelle 29: Ergebnisse Kosteneffektivität (mehrsseitige Tabelle)

Studie	Kosten der Interventionen pro 1000 Personen mit CT-Untersuchung (Maß der Unsicherheit sofern angegeben)	Nutzen pro 1000 Personen mit CT-Untersuchung (Maß der Unsicherheit sofern angegeben)	inkrementelles Kosten-Nutzen-Verhältnis	Ergebnisse aus Sensitivitätsanalysen	Schlussfolgerungen der Autorinnen und Autoren
Geppert 2025 (full model, Screening-population)	- AI-Unterstützung: 400 410 £^b (504 384,41 €^a) bzw. 127 600 £^c (160 733,87 €^a) - Ohne AI-Unterstützung: 470 630 £^b (592 838,43 €^a) bzw. 130 500 £^c (164 386,92 €^a)	- AI-Unterstützung: - QALY: 6532,10 - korrekt identifizierte Person mit handlungsbedürftigem Lungenrundherd: 223,8 - korrekt identifizierte und behandelte Personen mit Lungenkarzinom: 8,342 - Ohne AI-Unterstützung: - QALY: 6524,10 - korrekt identifizierte Person mit handlungsbedürftigem Lungenrundherd: 178,7 - korrekt identifizierte und behandelte Personen mit Lungenkarzinom: 7,750	Befundung mit AI-Unterstützung ist in allen Endpunkten überlegen	- Deterministisch: - Zeit(ersparnis) bei Befundung als Haupttreiber des ICER, jedoch ohne Konsequenz für die Dominanz - Dominanz der Befundung mit AI-Unterstützung auch in allen Szenarioanalysen - Probabilistisch: sehr hohe Wahrscheinlichkeit, dass AI kosteneffektiv ist (100 %; unabhängig von der Zahlungsbereitschaft)	- AI wird als potenziell kosteneffektive Maßnahme in der Screeningpopulation angenommen. - Hauptgrund für die Überlegenheit sind weniger falsch-positive Befunde mit nachfolgender Überwachung durch AI-Unterstützung (die zugrunde liegende Quelle [43] wird jedoch kritisch gesehen, da andere Studien eher auf eine geringere Spezifität durch die AI-Unterstützung hindeuten) - Insgesamt hohes Maß an Unsicherheit durch eingeschränkte Datenverfügbarkeit, wodurch verschiedene Annahmen getroffen werden mussten

Tabelle 29: Ergebnisse Kosteneffektivität (mehreseitige Tabelle)

Studie	Kosten der Interventionen pro 1000 Personen mit CT-Untersuchung (Maß der Unsicherheit sofern angegeben)	Nutzen pro 1000 Personen mit CT-Untersuchung (Maß der Unsicherheit sofern angegeben)	inkrementelles Kosten-Nutzen-Verhältnis	Ergebnisse aus Sensitivitätsanalysen	Schlussfolgerungen der Autorinnen und Autoren
Geppert 2025 (full model, Verdachtspopulation)	- AI-Unterstützung: 816 520 £^b (1 028 545,64 €^a) bzw. 138 740 £^c (174 766,60 €^a) - Ohne AI-Unterstützung: 715 450 £^b (901 230,80 €^a) bzw. 142 750 £^c (179 817,87 €^a)	- AI-Unterstützung: - QALY: 6329,90 - korrekt identifizierte Person mit handlungsbedürftigem Lungenrundherd: 481,8 - korrekt identifizierte und behandelte Personen mit Lungenkarzinom: 15,51 - Ohne AI-Unterstützung: - QALY: 6349,89 - korrekt identifizierte Person mit handlungsbedürftigem Lungenrundherd: 333,4 - korrekt identifizierte und behandelte Personen mit Lungenkarzinom: 12,83	- QALY: Einsatz ohne AI ist überlegen - Person mit handlungsbedürftigem Lungenrundherd: Befundung mit AI-Unterstützung ist überlegen 38 316 £ (48 265,51 €^a) / Person mit korrekt identifiziertem und behandeltem Lungenkarzinom	- Deterministisch: - Sensitivität für Befundung ohne AI sowie Zeit(-ersparnis) für die Befundung als Haupttreiber des ICER - Dominanz der Befundung ohne AI auch in den meisten Szenarioanalysen (Ausnahme: kein Nutzwertverlust für falsch-positive Fälle bzw. für gutartige Rundherde während der CT-Kontrollen) - Probabilistisch: - sehr hohe Wahrscheinlichkeit, dass AI nicht kosteneffektiv ist (ca. 100 %)	- AI wird nicht als potenziell kosteneffektive Maßnahme in der Verdachtspopulation angenommen. - Grund dafür ist, dass der Nutzwertverlust bei der Vielzahl der Fälle mit gutartigen Rundherden die Nutzwertgewinne bei den zusätzlich identifizierten Fällen mit Lungenkarzinom überwiegt - Insgesamt hohes Maß an (insbesondere struktureller und Parameter-)Unsicherheit durch eingeschränkte Datenverfügbarkeit
a. kaufkraftbereinigt in € umgerechnet [36] und mittels Verbraucherpreisindex vom Basisjahr 2021 auf das Jahr 2024 inflationsiert [37] b. Kosten über den gesamten Zeithorizont bei Betrachtung von QALY bzw. identifizierter und behandelte Person mit Lungenkarzinom als Endpunkt c. Kosten im Rahmen der initialen Befundung bei Betrachtung von korrekt identifizierten Personen mit handlungsbedürftigem Lungenrundherd als Endpunkt AI: künstliche Intelligenz; CT: Computertomografie; ICER: inkrementelles Kosteneffektivitätsverhältnis; QALY: qualitätsadjustiertes Lebensjahr					

A4.2.4 Bewertung der Berichtsqualität und Übertragbarkeit

A4.2.4.1 Bewertung der Berichtsqualität

In Tabelle 30 werden die Ergebnisse der Bewertung der Berichtsqualität der eingeschlossenen Studie dargestellt. AI-spezifische Kriterien sind als solche gekennzeichnet.

Tabelle 30: Bewertung der Berichtsqualität (mehrseitige Tabelle)

Kriterium		Geppert 2025		Kommentare
		Berichtet	Begründet ^a	
Hintergrund				
1	Hintergrund der Studie und Präzisierung der Studienfrage	ja	ja	
Methoden				
2	Gesundheitsökonomischer Analyseplan	ja	nicht zutreffend	Es existiert ein Protokoll zum Diagnostic Assessment Report, in dem Spezifikationen für den (eingetretenen) Fall einer de-novo-Modellierung festgehalten wurden [41].
3	Charakterisierung der Studienpopulation und der Subgruppen	ja	ja	
4	Interventionsalternativen	ja	ja	
AI1	Anwender-Autonomie	ja	nein / nicht zutreffend	Auch wenn die AI-Software Rundherde eigenständig erkennen und vermessen kann, wird im Bericht davon ausgegangen, dass es sich lediglich um eine Unterstützung der Radiologinnen und Radiologen handelt. Keine klinische Entscheidung wird allein aufgrund der bzw. durch die AI getroffen.
5	Entscheidungs- / Versorgungskontext	ja	ja	
6	Wahl der Perspektive	ja	ja	
7	Zeithorizont	ja	ja	
8	Diskontierungsrate	ja	ja	

Tabelle 30: Bewertung der Berichtsqualität (mehrseitige Tabelle)

Kriterium		Geppert 2025		Kommentare
		Berichtet	Begründet ^a	
	Outcomeparameter			
9	Auswahl der in die Analyse eingehenden Outcomeparameter	ja	ja	
10	Erhebung der in die Analyse eingehenden Outcomeparameter	ja	ja	
AI2	Erhebung der AI-Outcomeparameter	ja	ja	
AI3	Erfassung des AI-Lernfortschritts über die Zeit	nicht zutreffend	nicht zutreffend	Es wurden keine adaptiven AI-Interventionen betrachtet, die im laufenden Betrieb lernen und ihre Algorithmen stetig anpassen. Regelmäßige Updates der AI-Algorithmen (und deren Auswirkungen) wurden nicht betrachtet.
AI4	Entwicklung der AI-Komponente	nein	nein	
AI5	Validierung der AI-Komponente	nein	nein	
AI6	Gesundheitsbezogener Outcome der AI-Komponente	ja	ja	
AI7	Populationsunterschiede	nein	nein	
11	Bewertung der in die Analyse eingehenden Outcomeparameter	ja	ja	
	Ressourcenverbrauch und Kosten			
12	Auswahl der in die Analyse eingehenden Ressourcenverbräuche und Kostenparameter	ja	Ja	
13	Erhebung der in die Analyse eingehenden Ressourcenverbräuche und Kostenparameter	ja	ja	
14	Bewertung der in die Analyse eingehenden Ressourcenverbräuche und Kostenparameter	ja	Ja	
15	Währung, Indexjahr und Umrechnung	ja	ja	
16	Modell			
16a	Wahl der Modellierungstechnik	ja	nein	

Tabelle 30: Bewertung der Berichtsqualität (mehrseitige Tabelle)

Kriterium		Geppert 2025		Kommentare
		Berichtet	Begründet ^a	
16b	Darstellung des Modells	ja	ja	
16c	Modellannahmen	ja	ja	
16d	Herleitung der Inputparameter	ja	ja	
16e	Modellvalidierung	nein	nicht zutreffend	
AI8	Modellierung des AI-Lernfortschritts	nicht zutreffend	nicht zutreffen	
Weitere analytische Verfahren				
17	Methoden zur Analyse von Heterogenität / Subgruppenergebnissen	teilweise	ja	Das Modell wird separat für verschiedene Populationen (u. a. Screening- und Verdachtspopulation) und – soweit verfügbar – mit jeweils für diese Populationen spezifischen Parametern (z. B. Prävalenzen, diagnostische Güte, Ressourcenverbräuche) ausgeführt.
18	Methoden zur Analyse der Unsicherheit (z. B. Sensitivitätsanalysen)	ja	ja	
19 Ergebnisse				
19a	absolute und inkrementelle Ergebnisse der Analyse / Modellierung für jede Intervention	ja	nicht zutreffend	
19b	Subgruppenergebnisse	teilweise	nicht zutreffend	Die Ergebnisse werden separat je Population (u. a. Screening- und Verdachtspopulation) dargestellt. Weitere Subgruppen innerhalb der einzelnen Populationen werden nicht untersucht.
19c	Einfluss durch Unsicherheiten (Ergebnisse von uni- / multivariaten, probabilistischen Sensitivitätsanalysen)	teilweise	nicht zutreffend	Im full model werden die Ergebnisse der Sensitivitätsanalysen nur für den primären Endpunkt (Kosten pro QALY) dargestellt.
AI9	Einfluss durch Unsicherheit aufgrund der AI	ja	ja	
20 Diskussion				
20a	Diskussion der Ergebnisse und Limitationen	ja	nicht zutreffend	

Tabelle 30: Bewertung der Berichtsqualität (mehrsseitige Tabelle)

Kriterium		Geppert 2025		Kommentare
		Berichtet	Begründet ^a	
20b	Diskussion der Übertragbarkeit / Generalisierbarkeit	ja	nicht zutreffend	
20c	Schlussfolgerungen	ja	nicht zutreffend	
AI10	Implementierung der AI-Komponente	ja	nicht zutreffend	Es werden u. a. Angaben zur möglichen Integration der AI-Software in bestehende Systeme sowie zu weiteren Ressourcenverbräuchen (z. B. Schulungen des Personals) gemacht.
21	Weiteres			
21a	Interessenkonflikte	ja	nicht zutreffend	
21b	Studienfinanzierung	ja	nicht zutreffend	
a. Die Studie enthält zum jeweiligen Kriterium eine kurze Erläuterung oder eine Quellenangabe. AI: künstliche Intelligenz; QALY: qualitätsadjustiertes Lebensjahr				

A4.2.4.2 Bewertung der Übertragbarkeit

In Tabelle 31 werden die Ergebnisse der Bewertung der Übertragbarkeit der eingeschlossenen Studie dargestellt.

Tabelle 31: Bewertung der Übertragbarkeit (mehrseitige Tabelle)

Übertragbarkeit auf die zu untersuchende Fragestellung und den Kontext	Geppert 2025	Kommentare
	Ja Nein Teilweise Unklar	
Bildet die Fragestellung der Studie exakt die Fragestellung des ThemenCheck-Berichts ab?	teilweise	Die Studie liefert Aussagen zur Kosteneffektivität von AI-Unterstützung bei CT-Untersuchungen. Röntgen-Untersuchungen waren nicht Bestandteil der Studie.
Entspricht die Studienpopulation in der Publikation der Population der Fragestellung des ThemenCheck-Berichts? (Bezugspunkte: Patientenselektion, Baselinecharakteristika)	teilweise	Es werden differenzierte Ergebnisse für die interessierenden Teilpopulationen (Screening- und Verdachtspopulation) berichtet Insbesondere für die Screeningpopulation gelten jedoch nicht die gleichen Selektionskriterien für ein CT wie in der Population gemäß LuKrFrüErkV [5].
Werden die für die Fragestellung des ThemenCheck-Berichts relevanten Endpunkte in der Studienpublikation berichtet?	ja	Die gesundheitsbezogene Lebensqualität (erhoben über den generischen EQ-5D) bildet die Grundlage für QALYs als primären Endpunkt.
Sind die in der Studienpublikation herangezogenen Daten für die Effekte übertragbar auf die Fragestellung des ThemenCheck-Berichts?	teilweise	Die Effekte der AI-Unterstützung im Vergleich zur Befundung ohne AI-Unterstützung fließen in Form der diagnostischen Güte (Sensitivität und Spezifität zur Detektion von Lungenrundherden) in das Modell ein. Die zugrunde liegenden Studien zeigen jedoch ein hohes Verzerrungspotenzial und sind nicht ohne Weiteres auf den deutschen Versorgungskontext übertragbar.
Sind die in der Studienpublikation herangezogenen Kosten übertragbar auf das deutsche Gesundheitssystem (Perspektive, Versorgungskontext, Empfehlungen in Leitlinien, zusätzlich erforderliche Leistungen etc.)?	unklar	Insbesondere hinsichtlich der Preisgestaltung der AI-Software ist unklar, inwieweit die auf Unternehmensangaben beruhenden Kosten aus UK auf Deutschland übertragen werden können. Zudem können insbesondere auch die weiteren Preise und Erstattungsmodalitäten zwischen UK und Deutschland abweichen.
Sind epidemiologische, demographische und sozioökonomische Parameter (Prävalenz und Inzidenz, genetische Varianten, Lebenserwartung, Patientenpräferenzen) übertragbar auf die Studienpopulation im ThemenCheck-Bericht?	teilweise	Die im Modell verwendeten Prävalenzen (z. B. für das Vorliegen von Lungenrundherden) wurden in Teilen anhand von vorselektierten Populationen bestimmt, die nicht dem deutschen Versorgungskontext entsprechen. Die Nutzwerte wurden anhand von britischen (für gutartige bzw. keine Lungenrundherde) bzw. kanadischen (für maligne Lungenrundherde) Erhebungen bestimmt sodass sie nur eingeschränkt auf Deutschland übertragbar sind.

Ist die Modellstruktur übertragbar auf die zu untersuchende Fragestellung des ThemenCheck-Berichts?	teilweise	Das full model basiert auf dem Versorgungspfad gemäß den Vorgaben der BTS. In der deutschen S3-Leitlinie [4] wird hingegen die Klassifikation gemäß V-Lung-RADS 2022 empfohlen, deren Einteilung und empfohlenen Maßnahmen sich von den Vorgaben der BTS teilweise unterscheiden.
Gibt es weitere relevante Aspekte, die einen Einfluss auf die Übertragbarkeit der Studienergebnisse auf die Fragestellung des ThemenCheck-Berichts haben?	ja	Es ist darauf hinzuweisen, dass die Nutzen- und Kostenparameter lediglich auf einzelnen AI-Produkten beruhen und somit nicht allgemeingültig für alle AI-Interventionen im Rahmen der Fragestellung sind
AI: künstliche Intelligenz; BTS: British Thoracic Society; CT: Computertomografie; EQ-5D: European Quality of Life – 5 Dimensions; LuKrFrüErkV: Lungenkrebs-Früherkennungs-Verordnung; QALY: qualitätsadjustiertes Lebensjahr; V-Lung-RADS: Lung CT Screening Reporting and Data System mit Volumenverdopplungszeit		

A5 Details der Ergebnisse: Ethische, soziale, rechtliche und organisatorische Aspekte sowie Umwelt- und Klimaaspekte

A5.1 Ethische Aspekte

A5.1.1 Informationsbeschaffung zu ethischen Aspekten der zu bewertenden Technologie

Insgesamt wurden durch die orientierende Recherche in bibliografischen Datenbanken und weiteren Informationsquellen 19 Publikationen eingeschlossen.

Die Zitate der relevanten Publikationen finden sich in Abschnitt A7.3.1.

A5.1.2 Identifizierte ethische Aspekte

In Tabelle 32 sind die identifizierten ethischen Aspekte dargestellt. Die Darstellung orientiert sich an dem Rahmengerüst für die ethische Bewertung von Public-Health-Maßnahmen von Marckmann [18], dass für den Anwendungsbereich der Betrachtung zweier Fragestellungen erweitert wurde. Für die Bearbeitung wurde vorausgesetzt, 1. dass unter AI „ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet [...]“ zu verstehen ist (Art. 3 Nr. 1 AI Act [6]); 2. dass es um AI-Systeme zur Unterstützung der Auswertung von Röntgen- oder CT-Bildern geht, bei der die finale Diagnoseentscheidung bei ärztlichen Fachkräften liegt und der vollständige Ersatz der ärztlichen Entscheidung durch AI nicht umfasst ist und 3. die ethische Bewertung erfolgte für die beiden Fragestellungen – bezogen auf Personen mit Verdacht auf Lungenkrebs einerseits und Personen, die eine Früherkennung in Anspruch nehmen andererseits – einheitlich erfolgte.

Tabelle 32: Informationsaufbereitung zu den identifizierten ethischen Aspekten (mehrseitige Tabelle)

Ethische Bewertungskriterien	Fragestellung 1: Verdacht auf Lungenkrebs	Fragestellung 2: Früherkennung nach LuKrFrühErkV	Quelle
1. erwarteter Nutzen	Ergebnisse aus der Nutzenbewertung:		
	keine Aussage zum Nutzen	keine Aussage zum Nutzen	Domäne Nutzenbewertung (Kapitel 3)
	Aspekte aus der gesichteten Literatur – mögliche positive Effekte der AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge		
	Bessere Therapierbarkeit durch frühere Erkennung - Erkennung von mehr Krebserkrankungen in einem früheren Stadium - höhere Sensitivität bei der Erkennung maligner Lungenrundherde und anderer Auffälligkeiten, die auf Krebs hindeuten können - damit Verbesserung der Möglichkeit einer erfolgreichen individuellen Therapie / möglichst schonenden Therapie und Heilung - Verbesserung der Behandlungsergebnisse - Verbesserung der Lebensqualität - Allgemein: Vorteile für Patienten durch den Einsatz von AI		Busch 2025 [46] Deutscher Ethikrat 2023 [47] (S. 202 / S. 204) Faric 2024 [48] NICE 2023 [42] NICE 2025 [49]
	Bessere Befundung - durch Unterstützung von Ärztinnen und Ärzten in Form einer Zweitmeinung - Unterstützung der Entscheidungsfindung durch Verbesserung der Dokumentation der Eigenschaften von Lungenrundherden, z. B. Volumen - Erleichterung der Beurteilung des Wachstums von Lungenrundherden unter CT-Überwachung		Busch 2025 [46] Li 2021 [139] NICE 2023 [42] NICE 2025 [49]
	Erhöhtes Vertrauen erhöhtes Vertrauen in die Befundung unauffälliger Röntgenbilder		NICE 2025 [49]
	Zeitersparnis - Beschleunigung der Auswertung digitaler Bilder - schnellere Übermittlung der Befunde an die hausärztliche Praxis - dadurch Zeitersparnis zwischen Röntgenaufnahme / Diagnose / Behandlung		Deutscher Ethikrat 2023 [47] (S. 202 / S. 204) NICE 2023 [42]

Tabelle 32: Informationsaufbereitung zu den identifizierten ethischen Aspekten (mehrseitige Tabelle)

Ethische Bewertungskriterien	Fragestellung 1: Verdacht auf Lungenkrebs	Fragestellung 2: Früherkennung nach LuKrFrühErkV	Quelle
	Besserer Umgang mit begrenzten Personalressourcen ▪ verbesserter Ressourceneinsatz im Gesundheitssystem ▪ bessere Bewältigung der hohen Nachfrage nach Bildauswertungen, ▪ Entlastung der Ärzteschaft von monotonen Routinearbeiten, dadurch mehr Zeit für die Vermittlung der Befunde im persönlichen Gespräch		Deutscher Ethikrat 2023 [47] (S. 202 / S. 204) NICE 2025 [49]
	Verbesserter Arbeitsablauf ▪ verbesserter Arbeitsablauf durch zuverlässige Identifikation unauffälliger Röntgenbilder und Priorisierung auffälliger Röntgenbilder		NICE 2025 [49]
	Einsparen von Ressourcen ▪ Einsparung von Ressourcen im Gesundheitssystem		NICE 2025 [49]
	Aspekte aus der Betroffenenbefragung – erwarteter Nutzen der AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge		
	▪ höheres Sicherheitsgefühl durch „zweite Meinung“, wenn auch mit AI ▪ Schutz vor der Sorge, dass die Ärztin oder der Arzt etwas übersieht ▪ Vorteile wegen einer guten Mischung von den vielen Fallbeispielen der AI-Systeme und der Erfahrung der Ärztinnen und Ärzte		Betroffenenbefragung (Kapitel A9)
	Soziale Aspekte – erwarteter Nutzen der AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge		
	▪ Arbeit wird schneller und effizienter		Domäne soziale Aspekte (Abschnitt 5.3)

Tabelle 32: Informationsaufbereitung zu den identifizierten ethischen Aspekten (mehrseitige Tabelle)

Ethische Bewertungskriterien	Fragestellung 1: Verdacht auf Lungenkrebs	Fragestellung 2: Früherkennung nach LuKrFrühErkV	Quelle
2. Potenzieller Schaden und Belastungen	Ergebnisse aus der Nutzenbewertung:		
	keine Aussage zum Schaden	keine Aussage zum Schaden	Domäne Nutzenbewertung (Kapitel 3)
	Aspekte aus der gesichteten Literatur – mögliche negative Effekte der AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge		
	Fehlerhafte AI / fehlende Korrektur fehlerhafter Befunde - geringe Spezifität bei der Erkennung von Krebs könnte dazu führen, dass mehr unnötige Folgeuntersuchungen ausgelöst werden - geringe Sensitivität könnte dazu führen, dass Lungenkrebs übersehen / erst in fortgeschrittenem Stadium erkannt wird, damit kostspieligere Behandlungen und schlechtere Behandlungsergebnisse - abnehmende oder inadäquate Reaktionen auf Warnsignale des Systems („alert fatigue“) - sich selbst erfüllende Vorhersage: Wenn ein System, das auf Ergebnisdaten z. B. von Krebspatienten trainiert ist, eine schlechte Prognose vorhersagt, wird palliativ statt kurativ behandelt und damit die Empfehlung des AI-Systems verstärkt		Ewals 2023 [50] Faric 2024 [48] Madanay 2025 [51] NICE 2025 [49] Zentrale Ethikkommission 2021 [52]
	Schwindende ärztliche Kompetenz - schleichender Kompetenzverlust infolge fortschreitender Delegation bestimmter Aufgaben an technische Systeme – „De-Skilling“ - Gefahr der Verletzung von Sorgfaltspflichten im Umgang mit AI durch „Automation-Bias“, begünstigt durch verlorengegangenes eigenes Erfahrungswissen - Sorge, dass weniger erfahrene Ärztinnen und Ärzte sich auf die AI-Beurteilung verlassen und wahrscheinlich fehlerhafte Beurteilungen nicht ausschließen		Deutscher Ethikrat 2023 [47] Faric 2024 [48] Zentrale Ethikkommission 2021 [52]
	Überdiagnostik und -behandlung / Fehlbehandlungen - unnötige Nachuntersuchen, Biopsien, diagnose- oder behandlungsbedingte Morbidität und Ängste bei Patientinnen und Patienten - Fehlbehandlungen aufgrund von Fehldiagnosen		Alderman 2025 [53] Ewals 2023 [50] Faric 2024 [48] NICE 2025 [49] Zentrale Ethikkommission 2021 [52]

Tabelle 32: Informationsaufbereitung zu den identifizierten ethischen Aspekten (mehrseitige Tabelle)

Ethische Bewertungskriterien	Fragestellung 1: Verdacht auf Lungenkrebs	Fragestellung 2: Früherkennung nach LuKrFrühErkV	Quelle
	Benachteiligung von Gruppen Sorge, dass AI-Systemen zu ungünstigeren Gesundheitsergebnissen für bestimmte Gruppen im Vergleich zu anderen beitragen		Alderman 2025 [53]
	Verschlechterung des Arzt-Patienten-Verhältnisses / schwindendes Vertrauen - Sorge, dass der Einsatz von AI den Kontakt zwischen Ärztinnen, Ärzten und Patientinnen, Patienten verringern könnte - Gefahr einer zunehmenden Verobjektivierung von Patientinnen und Patienten durch den Einsatz von AI-Systemen; dem ist durch den Einsatz aktiv entgegenzuwirken und das Vertrauensverhältnis zwischen den beteiligten Personen zu schützen - Schaden, wenn die Aussage einer Person als weniger zuverlässig angesehen wird als das Ergebnis eines Algorithmus, insbesondere, wenn AI-Empfehlung und ärztliche Empfehlung nicht übereinstimmen: Radiologinnen oder Radiologen, die weniger Folgeuntersuchungen empfehlen als die AI, müssen möglicherweise mit einem Vertrauensverlust der Patienten in ihre Fachkompetenz rechnen – aber nicht, wenn sie aggressivere Empfehlungen aussprechen als die AI - Sorge von Ärztinnen und Ärzten vor negativen Reaktionen von Patientinnen und Patienten - Sorge vor rechtlichen Konsequenzen, wenn die ärztliche Entscheidung von AI-Ergebnis abweicht.		Alderman 2025 [53] Busch 2025 [46] Deutscher Ethikrat 2023 [47] Madanay 2025 [51]
	Einsparung / Ersatz von Fachkräften - Potenzial der AI, mit ärztlicher Tätigkeit in Konkurrenz zu treten - Sorge, dass der Einsatz von AI menschliche Fachkräfte ersetzen könnte - Sorge vor Abwertung der sprechenden Medizin / Abbau von Personal durch verstärkte Nutzung von AI-Komponenten in der Versorgung		BÄK 2025 [54] Busch 2025 [46] Deutscher Ethikrat 2023 [47]
	fehlendes Vertrauen von einigen Patientinnen und Patienten in AI-Software		NICE 2025 [49]
	Keine Verbesserung der Arbeitsbelastung Sorge, dass die Anwendung von AI-Systemen keine Verringerung der Arbeitsbelastung von Ärztinnen und Ärzten		NICE 2025 [49]

Tabelle 32: Informationsaufbereitung zu den identifizierten ethischen Aspekten (mehrseitige Tabelle)

Ethische Bewertungskriterien	Fragestellung 1: Verdacht auf Lungenkrebs	Fragestellung 2: Früherkennung nach LuKrFrühErkV	Quelle
	Erhöhter Zeitaufwand ▪ Sorge, dass die Anwendung von AI-Systemen die Zeit für die Auswertung von Röntgenaufnahmen nicht verkürzt, sondern sogar verlängert ▪ keine Beschleunigung der Übermittlung des Befundberichts an die Hausarztpraxis		NICE 2025 [49]
	Kostensteigerung im Gesundheitswesen ▪ Sorge vor steigenden Kosten im Gesundheitswesen		Busch 2025 [46]
	Datenschutz ▪ Bedenken in Bezug auf den Datenschutz		Busch 2025 [46]
	Manipulation ▪ Gefahr von Machtmissbrauch und Manipulation aufgrund potenziell undurchsichtiger Entscheidungskriterien von AI-Systemen		Katzenmeier 2023 [55] (S. 75)
	Cybersicherheit ▪ Potenzielle Gefahr von Angriffen auf Deep-Learning-Algorithmen in der Radiologie ▪ Bedenken hinsichtlich der Cybersicherheit und der Anforderungen an die Datenspeicherung, einschließlich der damit verbundenen Kosten		Faric 2024 [48] Sorin 2023 [56]
	Aspekte aus der Betroffenenbefragung – negative Effekte der AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge		
	▪ Sorge, dass sich Ärztinnen und Ärzte zu sehr auf AI verlassen – AI darf ärztliche Beurteilung nicht ersetzen, nur unterstützen ▪ Erwartung, dass AI-Befunde von Ärztinnen und Ärzten geprüft werden, bevor sie weitergegeben werden ▪ Sorge vor Fehleranfälligkeit der AI-Systeme, aber: lieber falsch-positive als falsch-negative Befunde ▪ Sorge vor Nachteilen aufgrund Verletzung des Datenschutzes, z. B. bei Weitergabe von Befunden an die Krankenkasse		Betroffenenbefragung (Kapitel A9)

Tabelle 32: Informationsaufbereitung zu den identifizierten ethischen Aspekten (mehrseitige Tabelle)

Ethische Bewertungskriterien	Fragestellung 1: Verdacht auf Lungenkrebs	Fragestellung 2: Früherkennung nach LuKrFrühErkV	Quelle
3. Auswirkungen auf die Autonomie	Aspekte aus der gesichteten Literatur – Auswirkung der AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge“		
	Patientenautonomie Einbindung von „AI-Entscheidungen“ in SDM-Prozess führt zu Gewinn an Patientenautonomie; andernfalls führt es zu Autonomieverlust („Computer-Paternalismus“) ▪ Patientinnen und Patienten müssen darauf vertrauen können, dass sie als Person, ihr Wohl und selbstbestimmter Wille im Zentrum der Versorgung stehen und sie nicht den Eindruck bekommen, auf ihre Daten reduziert zu werden ▪ Patientinnen und Patienten sind darauf angewiesen, alle für sie relevanten Informationen für eine nicht nur rechtswirksame, sondern tatsächlich autonome Einwilligung in bestimmte Behandlungsvorschläge zu erhalten ▪ Transparenz: Erfordernis nachvollziehbarer Ergebnisse AI-unterstützter Befunderhebung für selbstbestimmte Behandlungsentscheidungen von Patientinnen und Patienten	Zentrale Ethikkommission [52]	

Tabelle 32: Informationsaufbereitung zu den identifizierten ethischen Aspekten (mehrseitige Tabelle)

Ethische Bewertungskriterien	Fragestellung 1: Verdacht auf Lungenkrebs	Fragestellung 2: Früherkennung nach LuKrFrühErkV	Quelle
	Ärztliche Autonomie - freie, allerdings den Patientinnen und Patienten verpflichtete Wahl der Untersuchungs- oder Behandlungsmethode - effektive Dokumentation von Datensätzen der AI-Systeme ermöglicht anhand dessen Inhalts, Kontext und Einschränkungen zu entscheiden, ob sie für den konkreten Zweck geeignet ist - Überschreiten der Grenze zwischen Entscheidungsassistenz und Entscheidungsübernahme beinhaltet die Gefahr der Unachtsamkeit oder sogar des Kontrollverlusts - Ärztinnen und Ärzte müssen weiterhin in der Lage sein, die Aufsicht über den Gesamtprozess von Diagnostik, Therapie, Prognostik und Prädiktion zu übernehmen, in den zunehmend auch Vorschläge von maschinellen Systemen eingebunden sind. Diese Aufgabe ist nicht delegierbar - Transparenz, Kontrolle über AI-Anwendungen, Erfordernis nachvollziehbarer Ergebnisse AI-unterstützter Befunderhebung für autonome Therapieentscheidungen von Ärztinnen und Ärzten und zur Gewährleistung der Interpretierbarkeit von AI-Entscheidungen („automation bias“, „Black Box-Entscheidungen“); Transparenz bedeutet Interpretierbarkeit		Alderman 2024 [53] Deutscher Ethikrat 2023 [47] Bowers 2024 [57] Busch 2025 [46] Eichelberger 2024 [90] Ewals 2024 [58] Wang 2024 [60]
	Aspekte aus der Betroffenenbefragung – Auswirkungen der AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge auf die Autonomie		
	- Wunsch nach offenem Umgang über den Einsatz von AI-Systemen - Wunsch nach Transparenz insbesondere, wenn AI-System und Radiologin oder Radiologe unterschiedlich befunden		Betroffenenbefragung (Kapitel A9)
4. Gerechtigkeitsethische Auswirkungen	Aspekte aus der gesichteten Literatur – Gerechtigkeitsethische Auswirkungen der AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge		
	Gefahren des Algorithmus-Bias Verstärkung von Gruppen-Ungleichheiten in Folge von AI-Anwendung z. B. bei fehlerhaften AI-Algorithmen oder bei Verwendung von AI-Trainingsdaten, die nicht divers genug sind, um auch in der Versorgung unterrepräsentierter Gruppe abzudecken / ethnische Unterschiede zu berücksichtigen		Alderman 2024 [53] Ewals 2023 [50] Thong 2023 [59] Wang 2024 [60]

Tabelle 32: Informationsaufbereitung zu den identifizierten ethischen Aspekten (mehrseitige Tabelle)

Ethische Bewertungskriterien	Fragestellung 1: Verdacht auf Lungenkrebs	Fragestellung 2: Früherkennung nach LuKrFrühErkV	Quelle
	Mehr Gerechtigkeit durch federated learning Die Kombination von Daten aus regionalen, nationalen oder internationalen Einrichtungen könnte Patientinnen und Patienten aus unterrepräsentierten Gruppen, mit seltenen Krankheiten und Krankenhäusern mit geringeren Ressourcen zugutekommen		Crowson 2016 [61]
	Sicherstellung der gerechten Zur-Verfügung-Stellung bei erwiesener Überlegenheit von AI-Systemen „Bei durch empirische Studien sorgfältig belegter Überlegenheit von AI-Anwendungen gegenüber herkömmlichen Behandlungsmethoden ist sicherzustellen, dass diese allen einschlägigen Patientengruppen zur Verfügung stehen.“ - dafür Integration in klinische Ausbildung des ärztlichen Fachpersonals, „um eine breitere Nutzung vorzubereiten und verantwortlich so gestalten zu können, dass möglichst alle Patientinnen und Patienten davon profitieren und bestehende Zugangsbarrieren zu den neuen Behandlungsformen abgebaut werden“		Deutscher Ethikrat 2023 [47] (Empfehlungen Medizin 5 und 6)
	Aspekte aus der Betroffenenbefragung – Auswirkungen der AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge auf die Gerechtigkeit		
	beim Trainieren der AI müssen auch Daten von Frauen und anderen Gruppen und anderen möglicherweise unterrepräsentierten Gruppen vertreten sein		Betroffenenbefragung (Kaitel A9)
	Soziale Aspekte – Auswirkungen der Auswirkungen der AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge auf die Gerechtigkeit auf die Gerechtigkeit		
	Es besteht die Gefahr, dass nicht in allen Regionen Deutschlands Zugang zur AI-unterstützten Auswertung von Röntgen- oder CT-Bildern der Lunge gewährleistet wird, wenn Praxen die erforderliche Technologie nicht vorhalten (können)		Domäne soziale Aspekte (Abschnitt 5.3)
5. Erwartete Effizienz	Gesundheitsökonomische Bewertung		
	keine abschließende Bewertung		Domäne Gesundheitsökonomie (Kapitel 4)

AI: künstliche Intelligenz; CT: Computertomografie; LuKrFrühErkV: Lungenkrebs-Früherkennungs-Verordnung; SDM: gemeinsame Entscheidungsfindung

A5.2 Soziale Aspekte

A5.2.1 Informationsbeschaffung zu sozialen Aspekten der zu bewertenden Technologie

Insgesamt wurden durch die orientierende Recherche in bibliografischen Datenbanken und weiteren Informationsquellen 31 Publikationen eingeschlossen.

Die Zitate der relevanten Publikationen finden sich in Abschnitt A7.3.2.

A5.2.2 Identifizierte soziale Aspekte

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehrseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
Soziales Konstrukt / Verständnis des Gesundheitsproblems	Verständnis durch Betroffene	Betroffenenbefragung Die 5 befragten Personen sehen bildgebende Untersuchungen zur Diagnose / Früherkennung von Lungenkrebs als wichtiges Mittel, eine schwerwiegende / tödliche Erkrankung zu erkennen. 3 der 5 befragten Personen war nicht klar, dass eine frühere Erkennung mit besseren Chancen auf Behandlung / Heilung und einer geringeren Mortalität einhergeht. Eine Patientin erhielt durch ein Lungen-CT aufgrund anderer Symptome einen Lungenkrebs-Befund – eine seltene Krebsart in frühem Stadium – der erfolgreich entfernt werden konnte, was sie als Beleg der Sinnhaftigkeit der Untersuchung betrachtete. Eine starke Raucherin betont, dass sie ein Lungenkrebscreening-Angebot annehmen würde, aber auch Angst vor den Ergebnissen / dem Wissen über eine mögliche Erkrankung hat und nicht sicher ist, ob in ihrem Alter (72 Jahre) Wissen oder Nichtwissen besser ist.	Betroffenenbefragung (Kapitel A9) [66-69]
		Literaturbasierte Ergebnisse Ein fehlerhaftes / mangelndes Wissen über die Lungenkrebsfrüherkennung und deren Vor- / Nachteile oder eine falsche Einschätzung des eigenen Risikos und hiermit verbundene Ängste können die Teilnahme – insbesondere von Risikopersonen – an einer Früherkennung erschweren. Auch persönliche Eigenschaften der Zielpersonen oder Kontextfaktoren einer Früherkennung können für die Akzeptanz und die Teilnahmebereitschaft relevant sein. <u>Salman 2025 (britische Studie):</u> Aktuelle systematische Übersichtsarbeit von qualitativen, quantitativen und Mixed-Methods-Studien zur Identifizierung von hinderlichen / förderlichen Faktoren für die Absicht und die tatsächliche Wahrnehmung einer Lungenkrebsfrüherkennung. 72 Studien wurden eingeschlossen. Häufigste hinderliche Faktoren: Ängste (entweder vor einer vermeintlich fatalen Diagnose oder vor den Nebenwirkungen der Behandlung, z. B. Belastung durch Röntgenstrahlung), Kosten / Unannehmlichkeiten der Früherkennung, mangelndes Wissen über die Früherkennung. Häufigste förderliche Faktoren: persönliche Empfehlung des Kliniklers, wahrgenommener persönlicher Nutzen einer Teilnahme an der Früherkennung. <u>Warner 2024 (USA):</u> Fokusgruppen-Befragungen mit insgesamt 28 möglichen Kandidatinnen für eine Lungenkrebsfrüherkennung. Mangelndes Wissen über die Möglichkeit / den Nutzen einer Früherkennung; wichtig für eine Teilnahme-Bereitschaft war auch, dass Frauen evidenzbasierte Informationen über die Vor- / Nachteile der Untersuchung und einer frühen Lungenkrebsentdeckung erhielten.	

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehreseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
		<u>Warner 2019 (USA)</u>: Erhebung basierend auf dem Health Information National Trends Survey: Frauen hatten eine 32 % geringere Wahrscheinlichkeit, über das Lungenkrebscreening von Leistungserbringenden informiert zu werden. <u>Ali 2015 (britischer RCT)</u>: Identifikation von hinderlichen Faktoren einer Teilnahme am UK Lung Cancer Screening-Programm bei Hoch-Risiko-Personen (50–75 Jahre) mittels Mixed-Methods-Ansatz. Weibliches Geschlecht, höheres Alter (> 71 Jahre), aktuell Raucher (im Vergleich zu ehemaligen Rauchern), niedrigerer sozioökonomischer Status oder die Einschätzung eines hohen persönlichen Risikos begründen in dieser Studie, dass ein Lungenkrebscreening eher <u>nicht</u> in Anspruch genommen wurde.	
Soziales Image / Verständnis des Einsatzes von AI bei der Befundung bzw. Früherkennung	Verständnis durch Betroffene Wahrgenommener Nutzen & Vorstellungen zum Nutzen	<u>Betroffenenbefragung</u> Die befragten Personen stehen positiv zum Einsatz einer AI zur Auswertung von Röntgen- oder CT-Bildern der Lunge. Lungenkrebs wird als tödliche Bedrohung wahrgenommen und der Einsatz von AI als Möglichkeit, die Erkrankung früher / sicherer zu erkennen und eine Art zweite Meinung zu bekommen (Zitate: „man sollte alle Vorteile der modernen Technik nutzen“; „als wenn so ein zweiter mit drauf geguckt hätte [...]. wenn es auch eine künstliche Intelligenz ist“; „wenn es hilft, auch kleinste Veränderungen entdecken, die der Arzt nicht gesehen hätte, ist das doch gut“; „erhöht das Sicherheitsgefühl“) <u>Literaturbasierte Ergebnisse</u> Befragte Personen haben mehrheitlich eine positive Einstellung zum Einsatz von AI in der Versorgung generell bzw. bei der Lungenkrebsfrüherkennung und der Befundung im Besonderen und sehen hierin einen möglichen Nutzen für sich persönlich. <u>Busch 2025 (internationale Erhebung)</u>: Querschnittserhebung an 74 Krankenhäusern in 43 Ländern bei 13 806 erwachsenen Patientinnen und Patienten der Radiologie (nicht lungenkrebspezifisch). Die befragten Personen äußerten überwiegend positive Einschätzungen zum Einsatz der AI in der Gesundheitsversorgung. Frauen, ältere Menschen, weniger gesunde Personen (d. h. Personen mit einem schlechteren Gesundheitszustand) und weniger IT- bzw. AI-affine Personen hatten weniger positive Einstellungen zur AI. Die Mehrheit bevorzugte eine nachvollziehbare und immer arztgesteuerte AI-Unterstützung.	Betroffenen-befragung (Kapitel A9) [46,48,72]

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehreseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
		<u>Fritsch 2022 (deutsche Studie)</u>: Von 452 befragten Personen äußerten mehr als die Hälfte positive Einstellungen zu AI in der Gesundheitsversorgung. Ältere Personen, Frauen, weniger gebildete oder weniger technik-affine Personen äußerten diesbezüglich skeptischere Einstellungen. <u>Faric 2024 (britische Studie)</u>: Semi-strukturierte Befragungen von 12 Patientinnen und Patienten im Alter zwischen 40 und 80 Jahren bezüglich des Einsatzes der AI-Software „Veye Lung Nodules (VLN)“ zur Entdeckung, Klassifizierung und Vermessung von Lungenherden in CT-Scans der Lunge, Hinweis: VLN ist eine CE-zertifizierte AI-Anwendung (siehe <u>Products – Health AI Register</u>). <u>Ergebnisse</u>: Die Befragten standen dem Einsatz der AI durchgehend positiv gegenüber und erwarten einen Nutzen durch die AI-Unterstützung bei der radiologischen Untersuchung	
	Wissen & Verständnis der Technologie	<u>Betroffenenbefragung</u> 5 Betroffene berichten, dass sie bislang keinen Kontakt mit AI in der Gesundheitsversorgung hatten – auch nicht bei ihrer Lungen-Untersuchung – und vor der Befragung kein gutes Verständnis darüber hatten, wie genau die AI bei der Befundung eingesetzt werden kann. <u>Literaturbasierte Ergebnisse</u> Das Wissen über den Einsatz von AI in der Gesundheitsversorgung ist bislang in der Bevölkerung nicht verbreitet. <u>Busch 2025</u>: Von den in Busch 2025 weltweit befragten Personen gaben etwa 72 % an, nichts oder wenig über AI in der Gesundheitsversorgung zu wissen, 26 % gaben an, ein gutes Wissen über AI zu haben oder gar Experten hierin zu sein, und 2 % beantworteten die Frage nicht. <u>Fritsch 2022 (deutsche Studie)</u>: Von 452 befragten Personen hatten nur weniger als 25 % ein gutes oder sehr gutes Wissen über den Einsatz von AI in der Gesundheitsversorgung. <u>Faric 2024</u>: Die befragten Personen hatten zuvor keine Erfahrungen mit dem eingesetzten AI-Tool. Die Antworten aus den strukturierten Befragungen waren daher allgemein. Die Befragten äußerten ein großes Vertrauen in die Klinikerinnen und Kliniker bzw. Radiologinnen und Radiologen, die dieses Tool zur Unterstützung ihrer Arbeit nutzen. Außerdem sahen die befragten Personen die AI als Möglichkeit das überausgelastete nationale Gesundheitssystem (National Health System, NHS) zu entlasten.	Betroffenen-befragung (Kapitel A9) [46,48,72]
	Einstellungen und Akzeptanz der Inanspruchnahme der Technologie	<u>Betroffenenbefragung</u> Die befragten Personen würden eine AI-unterstützte Lungenuntersuchung in Anspruch nehmen. Eine alleinige Auswertung durch AI wird abgelehnt. Die AI wird als eine Art zweite Meinung oder „datenbasierte Unterstützung“ betrachtet. Zitate: „...ich sehe eine AI als Unterstützung des Arztes, der nicht jeden Tag 100 000 Lungenbilder sieht“, „das, was der Arzt nicht kann – aufgrund von Erfahrung, Tagesform etc. – kann vielleicht die AI und vice versa.“	Betroffenen-befragung (Kapitel A9)

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehreseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
		Literaturbasierte Ergebnisse Insgesamt scheint eine breite Akzeptanz der Technologie vorzuliegen. <u>Busch 2025</u>: 83,3 % der Personen, die sich selbst als AI-Experten betrachteten (n = 211 Personen), hatten eine positive / sehr positive Einstellung zum Einsatz von AI in der Medizin / Gesundheitsversorgung, wohingegen das bei 73,5 % der Personen mit gutem Wissen (n = 3423), bei 54,4 % der Personen mit wenig Wissen (n = 8097) und bei 38 % der Personen ohne Wissen über AI (n = 1848) der Fall war. Fast 90 % aller Befragten hatten eine neutrale (28,4 %) oder eine positive (59,2 %) Einstellung zum Einsatz von AI bei der Interpretation von Röntgenbildern. Über 70 % der Befragten würden Einrichtungen wählen, die mit AI-Unterstützung arbeiten und über 65 % fanden es positiv, wenn Ärzte die AI zur Unterstützung nutzen. <u>Faric 2024</u>: Die Befragten sagten, dass sie ein Krankenhaus, dass AI-Unterstützung anbietet, einem nicht unterstützten Setting vorziehen würden. Angesichts der Überlastung des Gesundheitssystems gingen die befragten Personen nicht nur von einem Nutzen der AI für sich selbst aus, sondern auch von einer personellen und System-Entlastung durch die AI-Unterstützung.	[46,48]
	Wahrgenommenes Risiko der Intervention und Umgang hiermit	Betroffenenbefragung Alle Betroffenen sagten, dass sie einen falsch-positiven Befund einem falsch-negativen Befund bevorzugen würden. Zitat: „dann lieber Ängste und Sorgen umsonst, als nachher eine Erkrankung, die mich killt“, „Wenn man vernünftig aufklärt und die Situation vernünftig erklärt, kann man auch mit solchen Situationen (<i>gemeint sind hier falsch-positive Befunde</i>) besser umgehen als mit der Unsicherheit, dass der Arzt irgendwas übersieht, weil da gibt es ja keine Informationen dann drüber. Der Arzt übersieht und er übersieht. Fertig. Da wäre ich schon lieber auf der sicheren Seite“ Außerdem drücken Befragte ihre Sorge über Auswertungen durch die AI alleine aus, Zitate: „Nein, AI alleine, da hätte ich kein Vertrauen, das wär nicht gut.“, „dann verlässt sich der Radiologe nachher noch zu sehr oder nur noch auf die AI, das darf nicht sein.“ Auch bezüglich des Schutzes der eigenen Daten gab es Bedenken bei den Betroffenen.	Betroffenen-befragung (Kapitel A9)

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehreseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
		Literaturbasierte Ergebnisse <u>Busch 2025</u>: Die Mehrheit (46,5 %) würde eine AI wählen, die keine Diagnose übersieht, aber oft falsch-positiven Alarm gibt, im Vergleich zu 36,3 %, die eine AI wählen würden, die nie falschen Alarm gibt, aber manchmal einen Befund übersieht, und 17,2 %, die eine AI wählen würden, bei der falsch-positive und falsch-negative Befunde sich die Waage halten. (Zu beachten: befragte Personen waren Patientinnen und Patienten der Radiologie, jedoch nicht nur im Bereich Lungenkrebs). <u>Fritsch 2022 (deutsche Studie)</u>: Von 452 befragten Personen war die Mehrheit davon überzeugt, dass eine AI nur die Arbeit der Leistungserbringenden unterstützen kann, nicht jedoch diese ersetzen sollte.	[46,48]
	Verständnis durch Anwender Wahrgenommener Nutzen & Vorstellungen zum Nutzen	<u>Faric 2024</u>: semi-strukturierte Befragung von 22 Klinikerinnen und Kliniker bzw. Radiologinnen und Radiologen in Großbritannien zu den Themen „allgemeiner Hintergrund / Rolle des Kliniklers im Setting“, „Wissen / Anwendung von AI“, Erfahrungen / Meinung zu AI“. 20 der 22 Kliniker nutzten die Veye Lung Nodules (VLN)-Software regelmäßig beim Lungenkrebscreening. Als ein Thema der Befragungen wurde der Bereich „Wahrgenommene Vorteile / Nutzen“ identifiziert. Die Kliniker berichten zu diesem Thema, dass die AI ihre Arbeit „schneller und effizienter“ gemacht und die AI sich zu einer Art „second reader“ entwickelt habe. Zitat: „It’s like having a second eye for the radiologist. We all miss things, we’re human beings, but having sort of a second pair of eyes, a computer program scanning the scan and picking up a nodule that you may potentially have missed is definitely an extra reassurance for us, but obviously better for patients as well.“ (Consultant cardio-thoracic radiologist). Als Voraussetzung für den Nutzen der AI werden die AI-Fähigkeiten der Anwender und eine gute Integration in bestehende Arbeitsabläufe und IT-Systeme genannt.	[48]

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehreseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
	Wissen über die Technologie & Verständnis der Technologie	<u>Faric 2024:</u> Die meisten Kliniker und Radiologen bzw. Klinikerinnen und Radiologinnen hatten bereits Erfahrung mit der AI. Faric 2024 betonen, dass das Meinungsbild wahrscheinlich nicht die Einschätzungen von Radiologinnen oder Radiologen spiegelt, die noch nie mit einer AI gearbeitet haben. Hervorgehoben wird im Themenkomplex „Art der AI und Integration“, dass intensives Training der AI-Anwender und kontinuierliche IT-Unterstützung wichtig für eine effiziente Integration sind. Einschränkend stellen die Autoren klar, dass in der breiten Umsetzung wahrscheinlich das Training nicht in dem Ausmaß und die IT-Unterstützung nicht in der Intensität gegeben sein wird wie in diesem Studiensetting, was eine Barriere für die Umsetzung in anderen Settings sein könnte.	[48] Domäne organisatorische Aspekte (Abschnitt 5.5)
	Einstellungen und Akzeptanz der Inanspruchnahme der Technologie	<u>Faric 2024:</u> Der Erfolg einer AI wird im Themenkomplex „Übernahme der AI durch die Experten“ als entscheidend davon abhängig gesehen, ob Radiologinnen und Radiologen verstehen und erfahren, dass die AI den eigenen Arbeitsbereich und die eigene Rolle bedeutsamer und effizienter macht. Laut Aussage der Befragten kann die AI mögliche Schwächen von Klinikern und Radiologen bzw. Klinikerinnen und Radiologinnen kompensieren (z. B. bei weniger erfahrenen Klinikern), andererseits können die Stärken von Radiologen (erfahrene Kliniker) mögliche Schwächen der AI (Fehlalarme / Blind spots in bestimmten Lungenregionen) kompensieren. Die in der Studie befragten Personen berichten von überwiegend positiven Erfahrungen mit der Technologie, die die Arbeitsabläufe besser und effizienter mache. <u>Akudjedu 2023:</u> Online-Befragung von 314 Radiologen weltweit zeigt insgesamt positive Einstellungen zur AI, aber noch zu wenig Erfahrung / Wissen bezüglich AI. Bei wenig Erfahrung mit AI eher Unsicherheit bezüglich AI. Radiologen sehen Vorteile des AI-Einsatzes (Unterstützung der Ärztin oder des Arztes, Zeitersparnis, Aufwertung der Radiologin oder des Radiologen, bessere Versorgung der Patientinnen und Patienten), aber auch Herausforderungen (Problem des „De-Skillings“ durch Einsatz der AI, da Radiologen sich zu sehr auf AI verlassen).	[48,73] Domäne organisatorische Aspekte (Abschnitt 5.5)

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehreseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
	Wahrgenommenes Risiko der Intervention und Umgang hiermit	<u>Faric 2024:</u> Im Themenkomplex „Art der AI und Integration“ benennen einige Radiologen das Risiko, sich zu sehr auf die AI zu verlassen. Daher sei es wichtig, AI-Befunde immer wieder zu kontrollieren – was zwar auch Zeit koste, aber weniger als ohne die AI zu arbeiten. Auch hinsichtlich möglicher falsch-positiver Befunde durch die AI sei es wichtig, eine Sensibilität für Stellen, an denen solche Befunde entstehen können, zu entwickeln. Zitat: „Over time, users learned how to work around the limitations of the system, identifying which areas produced erroneous results and which parts were reliable. [...] sometimes the AI software would draw round a nodule, but it might also draw round something that wasn't a nodule, like a vessel, or like a benign pleural plaque or something like that. And we would sort of call those false positives. But we would just ignore that, it didn't sort of take up lots of our time. (Consultant radiologist, visiting professor, Belgium)“	[48] Domäne organisatorische Aspekte (Abschnitt 5.5)
Soziokulturelle Aspekte der Implementierung der Technologie & Organisation der Nutzung	Soziokulturelle Eigenschaften der Zielgruppe	Die Zielgruppe der Lungenkrebsfrüherkennung mit / ohne Software-Unterstützung gemäß LuKrFrühErkV sind aktuelle oder ehemalige starke Raucher zwischen 50 und 75 Jahren, die den Kriterien der Verordnung entsprechen. Auch die Personen mit Verdacht auf Lungenkrebs weisen häufig Eigenschaften der Zielgruppe der Früherkennung auf. Folgende Unterschiede / Besonderheiten lassen sich herausarbeiten: <u>Geschlechtsbezogene Unterschiede:</u> <u>Lungenkrebsinzidenz / Mortalität:</u> Männer > Frauen ▪ 2022: altersstandardisierte Inzidenz Frauen: 31,8/100 000, Inzidenz Männer: 48,9/100 000; Mortalität Frauen: 22,4/100 000, Mortalität Männer: 38,5/100 000; ▪ relative 5 Jahres-Überlebensrate: 25 % bei Frauen, 19 % bei Männern ▪ seit etwa 3 Jahrzehnten sinkt die Inzidenz / Mortalität bei Männern und steigt bei Frauen (besonders in jüngeren Alterskohorten) ▪ Frauen sind in Studien zum Lungenkrebscreening stark unterrepräsentiert, scheinen aber mehr als Männer zu profitieren (stärkere Mortalitätsreduktion) <u>Alter</u> Mit dem Alter steigen Lungenkrebsinzidenz / -mortalität (ab ca. 50 Jahren) kontinuierlich an, bei Männern stärker als bei Frauen	[5] [2,70,71,75,77,78]

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehreseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
		<u>Rauchprävalenz:</u> Männer (33 %) > Frauen (25 %) Rauchen als Hauptrisikofaktor für Lungenkrebs: 9 von 10 Erkrankungen bei Männern und 6 von 10 Erkrankungen bei Frauen durch aktives Rauchen. Auch Passivrauchen ist ein wichtiger Risikofaktor. <u>Rauchen und Anfälligkeit für Lungenkrebs:</u> Frauen > Männer US-Studien lassen vermuten, dass Frauen und bestimmte ethnische Gruppen, die eine ähnlich hohe Lungenkrebsinzidenz hatten wie die Zielpersonen der Früherkennung (gemäß US Preventive Services Task Force (USPSTF)-Empfehlung), von der Früherkennung profitiert hätten, obwohl sie offiziell nicht die rauchbezogenen Kriterien der Früherkennung erfüllten	[2,75-78,82]
		<u>Bildung / sozioökonomischer Status / soziale Deprivation:</u> Michalski 2022: Rauchprävalenz höher in niedrigen Bildungsgruppen (2022 / 2023: 37 %) vs. höheren Bildungsgruppen (mittlere Bildung: 30 % / hohe Bildung: 16,9 %) Niedriger sozioökonomischer Status / soziale Deprivation als Risikofaktor für erhöhte Lungenkrebsinzidenz; Deutschland: Männer im höchsten Deprivationsquintil haben um 46 % erhöhtes Risiko gegenüber dem niedrigsten Deprivationsquintil. Bei Frauen ist der Unterschied weniger ausgeprägt.“ (Deprivation gemessen über Bildungs-, Beschäftigungs- und Einkommenssituation) Die Ergebnisse bestätigen auch verschiedene internationale Studien.	[80-82]
	Organisation der Nutzung	Für die Lungenkrebsfrüherkennung sollte eine zielgerichtete Ansprache (siehe soziokulturelle Eigenschaften der Zielgruppe) und Informierung von möglichen Anspruchsberechtigten durch eine in der LuKrFrühErkV genannten leistungserbringende Person (Allgemeinmediziner oder Allgemeinmedizinerin, Facharzt oder Fachärztin der Inneren Medizin oder der Arbeitsmedizin) erfolgen. Wichtig erscheint, dass hierbei auf die unterschiedlichen Eigenschaften der Zielgruppe eingegangen wird. Die schriftlichen Informationen könnten z. B. zielgruppenspezifische und möglicherweise AI-bezogene Informationen für Personen enthalten (z. B. für Frauen, unterschiedliche Altersgruppen), die eine Früherkennung per se eher seltener in Anspruch nehmen und / oder einer AI-unterstützten Befundung möglicherweise kritisch gegenüberstehen.	[5] Domäne organisatorische Aspekte (Abschnitt 5.5)

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehreseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
	Soziale Ungleichheit bei Entwicklung und Nutzung der Technologie	<u>Entwicklung der Technologie (AI):</u> Betroffenenbefragung Eine Patientin spricht an, dass beim Trainieren einer AI auch Daten von Frauen und anderen Gruppen vertreten sein müssen. Literaturbasierte Ergebnisse <u>Alderman 2024 „Standing Together Recommendations“, Europäischer AI Act:</u> Bei der Entwicklung einer AI müssen bestimmte Vorgaben zur Minimierung von Verzerrungen berücksichtigt werden, die ansonsten zum Beispiel zur Ungleichbehandlung von Personengruppen durch die AI führen können. <u>Kocak 2025, Shah 2024, Europäischer AI Act:</u> Eine Verzerrung kann zum Beispiel durch die Unter- oder Fehl-Repräsentation bestimmter Gruppen oder Minderheiten (Frauen oder ethnische Gruppen) in den Trainingsdaten einer AI entstehen, analog zur Unter-Repräsentation von Frauen in klinischen Studien. Dies kann dazu führen, dass die AI für ebendiese Minderheiten möglicherweise keine richtigen Ergebnisse liefert. Bei der Entwicklung einer AI müssen daher der Trainingssatz und dessen mögliche Einschränkungen transparent beschrieben werden. Gemäß dem Europäischen AI Act muss sichergestellt sein, dass die Betreiber einer AI „in der Lage sind, zu verstehen, wie das AI-System funktioniert, seine Funktionalität zu bewerten und seine Stärken und Grenzen zu erfassen“ (vgl. Artikel 13 AI Act).	[53,83-85] Domäne rechtliche Aspekte (Abschnitt 5.4)
		<u>Nutzung der Technologie (AI):</u> <u>Soziale Ungleichheit durch unterschiedliche Inanspruchnahme:</u> Personen mit Verdacht auf Lungenkrebs erhalten vor einer Untersuchung mittels Röntgen-Untersuchung oder CT eine generelle Aufklärung (einschließlich über die Strahlenbelastung der Untersuchung). Es ist unklar, ob und wie diese Personen zukünftig auch über den Einsatz einer AI bei der Befundung informiert werden und ob dies Auswirkungen auf die Inanspruchnahme hat, insbesondere durch die genannten AI-kritischen Gruppen (z. B. Frauen oder weniger AI-affine Personen). Die Lungenkrebsfrüherkennung für aktuelle oder ehemalige starke Raucher zwischen 50 und 75 Jahren wird gemäß der LuKrFrühErkV voraussichtlich ab April 2026 den Zielpersonen in Deutschland angeboten. Unterstützt wird die Ansprache anspruchsberechtigter Personen durch die behandelnden Ärzte (Allgemeinmediziner, Ärztinnen und Ärzte für innere Medizin oder für Arbeitsmedizin) durch eine evidenzbasierte Versicherteninformation des Gemeinsamen Bundesausschusses. Sofern über den Einsatz der AI aufgeklärt / informiert wird, ist es möglich, dass besonders AI-kritische Personen die Früherkennung ablehnen.	Domäne rechtliche Aspekte (Abschnitt 5.4) [5,74,103,118]

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehrseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
		Studien legen nahe, dass bestimmte Eigenschaften der Zielgruppe der Früherkennung - weibliches Geschlecht, höheres Alter, Rauchen, oder niedriger sozioökonomischer Status - jeweils für sich und in Kombination miteinander mit einer höheren Wahrscheinlichkeit einhergehen, eine Früherkennung nicht wahrzunehmen und teilweise auch die AI-unterstützte Früherkennung eher abzulehnen. Studien weisen außerdem darauf hin, dass Menschen, die ihr persönliches Risiko für Lungenkrebs falsch einschätzen oder unzureichende Information hierzu haben, seltener eine Früherkennung in Anspruch nehmen. Es erscheint sinnvoll, insbesondere Frauen und ältere Menschen zielgerichtet und mit auf ihre Zielgruppe angepassten Informationen zur Früherkennung anzusprechen und zu informieren. <u>Kauczor 2024:</u> Im europäischen SOLACE (Strengthening the screening of Lung Cancer in Europe)-Projekt werden Frauen und bestimmte „schwer erreichbare“ Gruppen als Zielgruppen der Lungenkrebsfrüherkennung genannt, die gezielt angesprochen und mit Hilfe von auf diese abgestimmten Informationen für eine Teilnahme am Screening gewonnen werden sollten. Als „schwer erreichbare“ Gruppen identifiziert die SOLACE-Gruppe ethnische Minderheiten, Personen mit geringem sozioökonomischem Status oder Personen, die weit weg von einem Screening-Zentrum wohnen.	[46,66,68,69,71,80,105] Domäne organisatorische Aspekte (Abschnitt 5.5) [86]

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehreseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
		<u>Soziale Ungleichheit durch unterschiedlichen Zugang zur Früherkennung</u> Durch die hohen ressourcenbezogenen Anforderungen an die Umsetzung einer AI-unterstützten Befundung (hohe AI-Anschaffungskosten & laufende Kosten, Kosten für AI-Aus- und Weiterbildung, Integration der AI in bestehende IT-Systeme, Verfahren der Qualitätssicherung beim Einsatz der AI-unterstützten Früherkennung etc.) wird es nicht für alle Radiologiezentren in Deutschland realistisch / machbar sein, diese anzubieten. Es ist davon auszugehen, dass in den nächsten Jahren – analog zur Früherkennung – die AI-unterstützte Befundung für Personen mit Verdacht auf Lungenkrebs ebenfalls primär in einzelnen zertifizierten Lungenkrebszentren oder deren Partner-Zentren angeboten werden können. Dadurch könnte es zur regionalen / kleinräumigen Versorgungsunterschieden und damit zu Versorgungsungleichheiten kommen. <u>Kocak 2025:</u> Die hohen technischen / finanziellen / personellen Anforderungen an die Einrichtung AI-unterstützter Systeme können möglicherweise nur wenige, sehr gut ausgestattete Zentren erfüllen, die nicht überall verfügbar sind, wodurch soziale Ungleichheiten in der Versorgung entstehen können. <u>Wilder 2024:</u> US-amerikanische Erhebung zu den Faktoren, die die zu geringe Nutzung / Inanspruchnahme der Lungenkrebsfrüherkennung erklären. Neben demographischen (Frauen, bestimmte ethnische Minderheiten), sozioökonomischen (Bildung, Einkommen, Kosten des Screenings, Krankenversicherung ja / nein) wird auch die geographische Ungleichverteilung der Screening-Zentren als Grund für eine unterschiedliche Inanspruchnahme genannt. <u>Kauczor 2024:</u> Für Personen, die kein Früherkennungszentrum in ihrer Nähe haben, schlägt die europäische SOLACE-Gruppe zum Beispiel den Einsatz von mobilen CT-Trucks vor (ähnlich den Mammographie-Bussen in Deutschland), die die Früherkennung zu diesen Menschen bringen könnte [105]. Inwiefern solche Angebote auch eine AI-unterstützte CT-Befundung einbeziehen könnten, ist unklar.	Domäne gesundheitsökonomische Bewertung und organisatorische Aspekte (Abschnitte 4.2 und 5.5) [84] [79] [86]

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehreseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
	Beziehung zwischen Betroffenen und Leistungserbringern, die die Intervention einsetzen, Entscheidungsprozesse	Betroffenenbefragung Die befragten Personen wünschen sich einen offenen Umgang mit dem Einsatz der AI. 2 Personen betonen, dass Transparenz besonders dann wichtig ist, wenn es Abweichungen zwischen der AI-Einschätzung und dem Leistungserbringer / der Leistungserbringerin gibt und dass dann ggf. die Gegenprüfung durch eine 3. Person notwendig ist. Die meisten Befragten wollen eingebunden sein in den Prozess der Befundung und sich hierzu mit der Ärztin oder der Radiologin oder dem Arzt oder Radiologen austauschen. Außerdem ist es den Betroffenen wichtig, dass der Arzt die letzte Entscheidung mit ihnen zusammentrifft und die AI hierbei nur eine Unterstützung ist. <u>Zitate:</u> „Mir in verständlichen Worten erklären, wie die Diagnosestellung erfolgt und mich aufklärt, auch über den Einsatz der AI – damit man am Ende ein komplettes Bild hat.“, „Ich hoffe schon, dass da offen mit umgegangen wird. Ich würde schon gerne wissen, wie die AI da genutzt wird.“ Literaturbasierte Ergebnisse <u>Busch 2025:</u> Die Mehrheit der Befragten (72,9 %) befürworteten eine durch AI unterstützte Diagnose, bei der der Klinikerin oder Radiologin oder der Kliniker oder Radiologe die Entscheidungs- und Deutungshoheit behält. <u>Madanay 2025:</u> Untersuchung, wie Patientinnen und Patienten darauf reagieren, wenn Befunde durch die AI von Befunden der Radiologin oder des Radiologen abweichen. Die Studie stützt die Aussage aus der Betroffenenbefragung, dass besonders abweichende Befunde mit den Betroffenen besprochen werden sollten. <u>LuKrFrühErkV:</u> Durch die Bedeutung, die gemäß der LuKrFrühErkV der umfassenden, und verständlichen Informierung der Zielpersonen sowie der Aufklärung gemäß §630e BGB zukommt, scheint ein aktiver Austausch zwischen Klinikerinnen und Klinikern bzw. Radiologinnen und Radiologen und anspruchsberechtigten Personen im Sinne einer partizipativen Entscheidungsfindung unerlässlich. Zu klären ist die Frage, inwiefern die Aufklärung / Informierung der Betroffenen sich auf die Früherkennung mittels bildgebender Verfahren beschränkt, oder ob diese auch die AI-Unterstützung bei der Befundung einbeziehen wird. <u>Walsh 2023:</u> Diese aktuelle randomisiert kontrollierte Studie berichtet, dass ein Informationstool zum Lungenkrebscreening (Lung Cancer Assessment of Risk and Education, LungCare), welche Personen in der Interventionsgruppe erhalten, zu einer statistisch signifikant höheren Inanspruchnahme (32 %) des Screenings im Vergleich zur Kontrollgruppe (13 %) führte. Die Studie zeigt auf, dass eine gute und zielgerichtete Information über das Screening dessen Inanspruchnahme wirksam beeinflussen kann.	Betroffenenbefragung (Kapitel A9) [46] [51] [5] Domäne rechtliche Aspekte (Abschnitt 5.4) [174]

Tabelle 33: Informationsaufbereitung zu den identifizierten sozialen Aspekten (mehrseitige Tabelle)

Kategorie	Aspekt	Inhalt	Quelle(n)
	Beziehung zwischen Personen, die die Technologie nutzen	Die Zusammenarbeit im radiologischen Team muss einen neuen Player – die AI – integrieren, ohne hierdurch die Qualität der Befundung zu gefährden. Hierzu scheint es sinnvoll, dass Radiologinnen und Radiologen mit unterschiedlichem Erfahrungshintergrund zusammenarbeiten. Ein sogenannter „mixed-expertise approach“ bei einer AI-unterstützten Befundung kommt bereits an verschiedenen Stellen zum Einsatz. <u>Vogel-Claussen 2022, HANSE-Studie</u> Die HANSE-Studie testet derzeit ein umfassendes AI-unterstütztes Lungenkrebscreening bei (Ex-)Raucherinnen und (Ex-)Rauchern im norddeutschen Raum. Hier werden gemischte Teams mit unterschiedlichem Erfahrungshintergrund beim Screening eingesetzt. <u>LuKrFrühErkV</u>: Die Verordnung sieht ebenfalls einen Erfahrungsmix vor (1. / 2. Befunder mit unterschiedlichen Qualifikationen). Angesichts der Herausforderungen, die der Einsatz für AI für die Radiologinnen und Radiologen und die Qualität ihrer Befundung mit sich bringt, scheint dieser Ansatz sinnvoll und nachhaltig.	Domäne organisatorische Aspekte (Abschnitt 5.5) [5,116,117]
AI: künstliche Intelligenz; BGB: Bürgerliches Gesetzbuch; CAD: Computerassistierte Detektion; CT: Computertomografie; IT: Informationstechnologie; LuKrFrühErkV: Lungenkrebs-Früherkennungs-Verordnung; NHS: National Health Service; RCT: Randomized controlled Trial; USA: United States of America; USPSTF: US Preventive Services Task Force			

A5.3 Rechtliche Aspekte

A5.3.1 Informationsbeschaffung zu rechtlichen Aspekten der zu bewertenden Technologie

Insgesamt wurden durch die orientierende Recherche in bibliografischen Datenbanken und weiteren Informationsquellen 32 Publikationen eingeschlossen.

Die Zitate der relevanten Publikationen finden sich in Abschnitt A7.3.3.

A5.3.2 Identifizierte rechtliche Aspekte

In Tabelle 34 sind die identifizierten rechtlichen Aspekte dargestellt. Die Darstellung folgt dem Kriterienkatalog von Brönneke 2016 [21].

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten (mehrsseitige Tabelle)

Aspekt	Diskussion	Quelle
Patientenautonomie I: informierte Einwilligung	Einwilligung nach Aufklärung § 630c, 630e BGB ▪ Grundsätzliches Erfordernis der Aufklärung (§ 630e BGB) und Einwilligung (§ 630d BGB) des Patienten vor jeder medizinischen Maßnahme ▪ Aufzuklären ist über „sämtliche für die Einwilligung wesentlichen Umstände“ (§ 630e Abs. 1 S. 1 BGB), insbesondere über „Art, Umfang, Durchführung, zu erwartende Folgen und Risiken der Maßnahme sowie ihre Notwendigkeit, Dringlichkeit, Eignung und Erfolgsaussichten im Hinblick auf die Diagnose oder die Therapie“ (§ 630e Abs. 1 S. 2 BGB).	630a ff BGB (Patientenrechtegesetz, PatRG) [175]
	unterschiedliche Auffassungen zum Erfordernis der Aufklärung über den Einsatz von AI:	
	▪ Aufklärungsumfang abhängig von der Einordnung der AI-Unterstützung als „Standard-“ oder „neue“ Methode ▪ Einordnung als „neue Methode“: zusätzlich zur „normalen“ Aufklärung Information des Patienten darüber, dass es sich um keine allgemein anerkannte Standardmethode handelt und dass unbekannte Risiken nicht ausgeschlossen werden können ▪ Ziel der Aufklärung: Fähigkeit zur sorgfältigen Abwägung von Vor- und Nachteilen vor Einwilligung in die medizinische Maßnahme ▪ Einordnung als „Standardmethode“: Aufklärung nach den üblichen Grundsätzen („im Großen und Ganzen über Chancen und Risiken der Behandlung“); offen ist, ob darüber hinaus „eine Aufklärung geboten ist über (allein schon) den Umstand, dass ein [...] AI-System zum Einsatz kommen soll. Dafür könnte sprechen, dass (Langzeit-)Erfahrungen beim Einsatz solcher Systeme fehlen, dass möglicherweise derzeit	Eichelberger 2024, S. 102 ff. [90] Katzenmeier 2023 [55]

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehreseitige Tabelle)

Aspekt	Diskussion	Quelle
	eine allgemeine Erwartung besteht, „herkömmlich“ behandelt zu werden und dass möglicherweise eine allgemeine Skepsis gegen „AI“ besteht. ▪ Fazit: Einstweilen sollten beim Einsatz von AI-Systemen über diesen Umstand aufgeklärt werden, selbst wenn aus diesem Umstand für sich genommen keine gesteigerten Risiken folgen.“	
	Zur Aufklärungspflicht, wenn keine AI eingesetzt wird: ▪ möglich, wenn sich der Einsatz eines AI-Systems als neue Behandlungsmethode weitgehend durchgesetzt hat und „derart entscheidende Vorteile bietet, dass ein Patient davon erfahren muss, um für sich entscheiden zu können, ob er sich um die Behandlung nach dem neuesten Stand bemühen oder, sofern das möglich ist, mit der Behandlung abwarten will, bis auch der von ihm aufgesuchte Arzt oder seine Klinik über solche Therapiemöglichkeiten verfügen“.	Eichelberger 2024, S. 102 ff. [90]
	▪ Aufklärungspflicht umfasst die wesentlichen Umstände der Behandlung; dazu gehört der Einsatz von „Clinical Decision Support Systems“ (Anm.: umfasst AI-Systeme), die als eigenständige Untersuchungs- oder Behandlungsmethode im Rahmen einer medizinischen Behandlung eingesetzt werden und bereits zum medizinischen Standard zählen. ▪ Keine Aufklärungs- und Einwilligungspflicht, wenn AI lediglich eine unterstützende Neben- oder Begleitleistung ist (in Abgrenzung zu AI als eigenständige Untersuchungs- oder Behandlungsmethode, die die ärztliche Entscheidung faktisch ersetzt)	Zentrale Ethikkommission 2021 [52]
	Umfassende Aufklärung erforderlich, da neue Diagnostikinstrumente zu Fragen und Befürchtungen bei Patientinnen und Patienten führen: ▪ häufige Fragen beziehen sich auf Sicherheit, Genauigkeit und Datenschutz ▪ Befürchtungen z. B. Reduktion auf Daten statt Wahrnehmung als Person	Deutscher Ethikrat 2023 [47]
	▪ Information über und Einwilligung in Datenverarbeitung und -nutzung für AI-Entwicklung: s. Autonomie III: Datenschutz	Art. 12 Abs. 1 i. V. m. Art. 13 / 14 DSGVO [91]

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehrsseitige Tabelle)

Aspekt	Diskussion	Quelle
Patientenautonomie II: Einwilligungsunfähigkeit	Vorliegend nicht relevant, da sich der HTA ausschließlich auf einwilligungsfähige Personen bezieht	-
Autonomie III: Datenschutz	Vorgaben zum Datenschutz finden sich in verschiedenen Verordnungen und Gesetzen: Die DSGVO ist das grundlegende Regelwerk; alle Verarbeitung personenbezogener Gesundheitsdaten muss DSGVO-konform sein, es wird ergänzt durch das BDSG; daneben treten der AI ACT mit zusätzliche AI-spezifischen Anforderungen, die Gesundheitsdaten-Regulierung durch EHDS-Verordnung (EU) 2025/327), spezifiziert bezogen auf den Gesundheitssektor, wie Daten in Primär- und Sekundärnutzung genutzt, geteilt, und kontrolliert werden dürfen), sowie das GDNG und DigiG, Umsetzung der Vorgaben von EHDS, Regelung von Verfahren, Zuständigkeiten und Dateninfrastruktur. Art. 5 DSGVO Datenschutzgrundsätze: ▪ Rechtmäßigkeit: Verarbeitung personenbezogener Daten bedarf für die Rechtmäßigkeit einer Rechtsgrundlage, die die konkrete Datenverarbeitung erlaubt, z. B. Art. 6 Abs. 1 lit. a DSGVO die Datenverarbeitung, wenn die betroffene Person ihre Einwilligung gegeben hat ▪ Fairness: Gebot der fairen Behandlung; Verstoß liegt z. B. vor, wenn ein behandelnder Arzt den Umstand, dass ein Patient auf ihn angewiesen ist, dazu ausnutzt, um an Daten zu gelangen ▪ Transparenz: personenbezogene Daten müssen in einer für die betroffene Person nachvollziehbaren Weise verarbeitet werden; insbesondere ist die betroffene Person umfassend über die Verarbeitung ihrer Daten zu informieren ▪ Zweckbindung: bei der Erhebung von Daten ist der Zweck ihrer Verarbeitung festzulegen; die Daten dürfen grundsätzlich nur zu diesem Zweck verarbeitet werden ▪ Datenminimierung: Beschränkung der Datenverarbeitung und damit des Umfangs der gespeicherten und verarbeiteten personenbezogenen Daten auf das für die jeweils festgelegten Zwecke notwendige Maß. ▪ Richtigkeit: personenbezogene Daten müssen „sachlich richtig und erforderlichenfalls auf dem neuesten Stand sein“ ▪ Speicherbegrenzung: personenbezogene Daten dürfen nur so lange verarbeitet werden, wie es für die Zwecke, für die sie verarbeitet werden, erforderlich ist; gilt nicht bei Aufhebung des	DSGVO [91]; Herbst 2024, S. 78 ff. [176]

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehrsseitige Tabelle)

Aspekt	Diskussion	Quelle
	Personenbezugs durch Anonymisierung der Daten (Daten ohne Personenbezug fallen nicht in den Anwendungsbereich der DSGVO) ▪ Integrität, Vertraulichkeit: personenbezogene Daten sind so zu verarbeiten werden, dass eine angemessene Sicherheit dieser Daten gewährleistet ist; unter anderem Schutz vor Hackerangriffen.	
	Art 9 DSGVO: besonderer Schutz von Gesundheitsdaten, § 22 BDSG: Verarbeitung besonderer Kategorien personenbezogener Daten Verbot (Art. 9 Abs. 1 DSGVO) nebst Aufhebung des Verbots für bestimmte Fallgruppen unter bestimmten Voraussetzungen (Art. 9 Abs. 2 DSGVO); Ausnahmen zur Verbotsregelung (§ 22 BDSG) ▪ Ziel: Schutz personenbezogener Daten in der EU ▪ grundlegende Datenschutzanforderungen für die Verarbeitung personenbezogener Gesundheitsdaten, auch AI-Systeme ▪ die Verarbeitung sensibler Daten im Sinne des Art. 9 DSGVO ist zulässig, wenn die betroffene Person ausdrücklich einwilligt, Art. 9 Abs. 2 lit. a DSGVO ▪ für die medizinische Diagnostik (unter anderem) ist die Datenverarbeitung durch Ärztinnen und Ärzte oder durch sonstige Personen, die einer entsprechenden Geheimhaltungspflicht unterliegen, auch ohne ausdrückliche Einwilligung zulässig, Art. 9 Abs. 2 lit. h, Abs. 3 DSGVO; § 22 Abs. 1 Nr. 1 lit. b BDSG.; dann gelten besondere Anforderungen z. B. an die Pseudonymisierung	DSGVO [91]; Bundesdatenschutzgesetz (BDSG) [92]; Herbst 2024, S. 78 ff. [176]; Zentrale Ethikkommission [52]
	Art. 12 Abs. 1 i. V. m. Art. 13 oder 14 DSGVO: Informationspflicht bei Erhebung von personenbezogenen Daten bei den betroffenen Personen / bei einer anderen Person ▪ Regelung von Informationspflichten insbesondere zur Einhaltung der Grundsätze der Transparenz und Fairness (s. o. Art. 5 DSGVO); ▪ Information über die Umstände der Datenverarbeitung in „verständlicher und leicht zugänglicher Form in einer klaren und einfachen Sprache“	DSGVO [91]; Bundesärztekammer 2025a [99]
	Art. 17 DSGVO: Recht auf Löschung ("Recht auf Vergessenwerden") ▪ unklar, inwieweit die Datenverarbeitung zum Zweck der individuellen Diagnostik und Therapie zugleich systemimmanent dem Training und der Weiterentwicklung von Clinical Decision Support Systems“ (Anm.: umfasst AI-Systeme) dienen darf und ob eine ausdrückliche Einwilligung nötig ist.	DSGVO [91]; Zentrale Ethikkommission [52]

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehrsseitige Tabelle)

Aspekt	Diskussion	Quelle
	▪ Lösung über Anonymisierung, siehe oben Art. Art. 5 DSGVO, Grundsatz der Speicherbegrenzung	
	Art. 22 DSGVO: Automatisierte Entscheidungen im Einzelfall ▪ Ausschluss von ausschließlich auf automatisierter Verarbeitung beruhenden Entscheidungen, die „rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt“; das gilt insbesondere für die Verarbeitung von Gesundheitsdaten, Art. 22 Abs. 4 DSGVO. ▪ gilt daher nicht für AI-unterstützte Entscheidungen, bei denen eine menschliche Entscheidung erfolgt, z. B. ärztliche Plausibilitäts- und Richtigkeitskontrolle im Einzelfall ▪ Exkurs: anders für AI-Systeme, die ärztliche Entscheidungen ersetzen; solche dürfen grundsätzlich nur mit ausdrücklicher Einwilligung nach entsprechender Aufklärung eingesetzt werden	DSGVO [91]; Zentrale Ethikkommission [52]
	Art. 2 Abs. 7, Erwägungsgründe 103 f. AI Act (Verordnung (EU) 2024/1689 über künstliche Intelligenz) ▪ keine Veränderung der Rechte und Pflichten aus der DSGVO durch den AI Act ▪ keine Schaffung einer neuen Rechtsgrundlage für die Verarbeitung personenbezogener Daten ▪ Pflicht von Anbietern und Nutzern von AI-Systemen, die als Verantwortliche oder Auftragsverarbeiter im Sinne der Datenschutzvorschriften handeln, sicherzustellen, dass die Verarbeitung personenbezogener Daten im Einklang mit der DSGVO erfolgt	AI Act [6]
	Gesundheitsdaten-Regulierung durch EHDS-Verordnung ((EU) 2025/327), GDNG und DigiG: Digitalgesetz ▪ EHDS ergänzt die DSGVO mit Regelungen zur datenschutzkonformen Nutzung von Gesundheitsdaten, Art. 1 Abs. 4 EHDS, insbesondere ▪ Regelungen zur Verfügbarkeit, Kontrolle und Nutzung von Gesundheitsdaten für die Versorgung (z. B. medizinische Bildgebung und darauf gestützte Befunde, Primärnutzung, Art. 7-11 EHDS) ▪ Regelungen für die weitere Nutzung von Gesundheitsdaten (Sekundärnutzung; Forschung, Innovation, Politik, Art. 33-39 EHDS); innerhalb der EU können Gesundheitsdaten unter Wahrung des Datenschutzes für wichtige Zwecke wie z. B. der wissenschaftlichen Forschung genutzt werden – relevant im Zusammenhang mit der potenziellen	EHDS-Verordnung [83] BMG 2025 [93]; BfArM [177], Mühlhoff 2024 [178], GDNG [93] DigiG [94] Crowson 2022 [61]

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehreseitige Tabelle)

Aspekt	Diskussion	Quelle
	Gefahr der „missbräuchlichen Sekundärnutzung“, wenn z. B. trainierte Modelle für diskriminierende oder politische fragwürdige Zwecke genutzt werden ▪ Anwendung finden die Vorschriften der EHDS-VO gestaffelt 2, 4, 6 oder 10 Jahre nach Inkrafttreten (Veröffentlichung: 05.03.2025; Inkrafttreten: 26.03.2025; März 2027: Erlass von Durchführungsrechtsakten, um technische Details zu definieren; März 2029: Erste Gruppe von vorrangigen Datenkategorien (Patientenkurzakte, elektronische Verschreibungen) muss in allen Mitgliedstaaten für Primärnutzung austauschbar sein; Sekundärnutzungsregeln gelten für die meisten Datenkategorien wie z. B. elektronischen Patientenakten; März 2031: Austausch der zweiten Gruppe vorrangiger Daten (medizinische Bilder, Laborergebnisse, Krankenhausentlassungsberichte) für Primärnutzung; Sekundärnutzungsregeln auch für weitere Kategorien wie genomische Daten; März 2034: Drittstaaten und internationale Organisationen können an „HealthData@EU“ für Sekundärnutzung teilnehmen.) ▪ Erschließung von Gesundheitsdaten für die Forschung durch GDNG ▪ Regelungen zur Digitalisierung im Gesundheitswesen, z. B. elektronische Patientenakte (ePA), digitale Infrastruktur, digitale Anwendungen; Regelung der Erhebung und Nutzung von Daten durch das DigiG Wertung in der Literatur: enorme Herausforderung, die eine technische und semantische Interoperabilität zwischen den verschiedenen Infrastrukturen und IT-Systemen der Mitgliedstaaten erfordert.	
	„datenschutzrechtliche Hürden“ „In der öffentlichen Debatte sind zunehmend Stimmen zu vernehmen, die eine zu rigorose oder bürokratisierte Datenschutzpraxis in Deutschland kritisieren. Der Fokus [...] liege in Deutschland zu stark auf den Risiken statt auf den Chancen. Datenschutz dürfe in der Gesundheitsversorgung und -forschung nicht allein als Schutzfunktion verstanden werden, sondern als Ermöglichung einer gemeinwohlorientierten Datennutzung bei Wahrung der Persönlichkeitsrechte.“; dem sollen BDSG und DigiG Rechnung tragen	Bundesärztekammer 2025b (S. 13 f.) [54]
	Federated Learning Verschiedene Institutionen trainieren gemeinsam ein Modell; Patientendaten bleiben dabei lokal, nur	Crowson 2022 [61] Konečný 2016 [179]

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehreseitige Tabelle)

Aspekt	Diskussion	Quelle
	Modellparameter verlassen die Klinik, aus der die Daten stammen; nur das trainierte Programm wird wieder geteilt ▪ Potenzial vor allem hinsichtlich Datenschutz und Kooperation über Institutionen hinweg ▪ wachsender Bereich des maschinellen Lernens mit vielen vielversprechenden Anwendungsmöglichkeiten im Gesundheitswesen ▪ Ermöglichung größerer Datenvielfalt, dadurch keine „Überanpassung“ auf bestimmte Datensätze	
Marktzugang I: Medizinprodukte	Der Marktzugang ist verschiedenen Regelwerken normiert. Grundlegend ist die Verordnung (EU) 2017/745 über Medizinprodukte (MDR); daneben treten insbesondere der AI Act (siehe oben unter II Datenschutz) und die MPBetreibV.	

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehreseitige Tabelle)

Aspekt	Diskussion	Quelle
	Maßgeblich für die Anwendung der MDR ist die Einordnung eines Produktes als Medizinprodukt und die Feststellung der Risikoklasse; dafür gelten folgende Definitionen: ▪ MDR (EU) 2017/745 Art. 2 Nrn. 1 und 12: Definition „Medizinprodukt“ und „Zweckbestimmung“ ▪ Medical Device Coordination Group (MDCG) 2019-11: Abgrenzung Software / AI-System; Leitfaden zur Einordnung von Software als Medizinprodukt (Medical Device Software, MDSW) ▪ MDR Art. 51 + Anhang VIII Regel 9–13 (Regel 11 für Software): Risikoklassifizierung nach MDR; Software zur Diagnose fällt je nach Patientenrisiko in die Klassen IIa–III ▪ AI Act (EU) 2024/1689 Art. 6 Abs. 1 a. b; Anhänge I und III: AI-spezifische Einstufung: AI-Systeme, die unter MDR fallen, sind als „Hochrisiko-AI“ einzuordnen ▪ MDR Art. 52; AI Act Kapitel IV: Konformitätsbewertung; Beteiligung einer Benannten Stelle erforderlich (außer Klasse I bzw. A); CE-Kennzeichnung nach MDR und AI Act „Medizinprodukte, die ein AI-Hochrisiko-System enthalten, müssen ab 02.08.2027 die Anforderungen des AI Act der EU erfüllen. Medizinprodukte ab Risikoklasse IIa sind von einer sogenannten Benannten Stelle (notified body) auf ihre CE-Konformität zu prüfen. Medizinprodukte dieser Risikoklassen, die ein Hochrisiko-AI-System enthalten, müssen künftig neben den Anforderungen der MDR auch die des AI Act erfüllen. [...] Teile der MDR und des AI Acts stellen ähnliche Anforderungen zum Beispiel an das Risikomanagementsystem oder die technische Dokumentation der Produkte. Das könnte die gemeinsame Umsetzung beider Verordnungen zukünftig erleichtern. Doch gibt es auch Unterschiede. Mit dem Ziel, AI-bedingte Risiken zu mindern, fordert der AI Act die Möglichkeit einer menschlichen Aufsicht, während die Produkte verwendet werden. Die MDR sieht eine solche Aufsicht hingegen nicht zwingend vor.“ (Medizinischer Dienst Bund 2025)	MDR [95] AI Act [6] MDCG 2019 [96] Luckner 2025 [180] Bundesärztekammer 2025 [99] Medizinischer Dienst Bund 2025 [181]

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehrsseitige Tabelle)

Aspekt	Diskussion	Quelle
	Marktzugang für kontinuierlich / adaptiv lernende AI-Systeme: ▪ nach Anhang I Kapitel II Abschnitt 17.1 MDR ist Software so auszulegen, dass sie „Wiederholbarkeit, Zuverlässigkeit und Leistung entsprechend ihrer bestimmungsgemäßen Verwendung gewährleistet“; danach wäre bei jeder Änderung durch Weiterlernen eine Neuzulassung erforderlich ▪ Ab 02.08.2026 gilt nach Art. 43 Abs. 4 AI Act, dass AI-Systeme mit hohem Risiko, die bereits einem Konformitätsbewertungsverfahren unterzogen wurden, im Falle einer wesentlichen Änderung einem neuen Konformitätsbewertungsverfahren zu unterziehen sind. Das gilt nicht für weiterlernende AI-Systeme, die die Ausführung von Funktionen automatisch anpassen, sofern diese Änderungen vom Anbieter im Voraus festgelegt und zum Zeitpunkt der Konformitätsbewertung bewertet wurden (Erwägungsgrund 128 AI Act). ▪ MPBetreibV §§ 3, 4: Meldepflichten für Ärzte bei Medizinprodukten auftretende Vorkommnisse sind der zuständigen Behörde mitzuteilen; Fehlfunktionen oder Unzulänglichkeiten in den Herstellerinformationen, die eine erhebliche Gefährdung für Patienten, darstellen können, sind der zuständigen Behörde und dem Hersteller zu melden.	
	Aspekt Sicherheit und Qualität „Bei der Entwicklung, Erprobung und Zertifizierung medizinischer AI-Produkte bedarf es einer engen Zusammenarbeit mit den relevanten Zulassungsbehörden sowie insbesondere mit den jeweils zuständigen medizinischen Fachgesellschaften, um Schwachstellen der Produkte frühzeitig zu entdecken und hohe Qualitätsstandards zu etablieren.“	Deutscher Ethikrat 2023 (Empfehlung Medizin 1) [47] (S. 216)
Marktzugang II: Arzneimittel	Vorliegend nicht relevant, da die untersuchten Interventionen nicht unter den Arzneimittelbegriff des AMG fallen: § 2 AMG: Arzneimittel sind Stoffe oder Zubereitungen aus Stoffen, die zur Anwendung am oder im menschlichen Körper gedacht sind.	AMG [182]
Klinische Studien	Die in die Nutzenbewertung eingeschlossenen Studien entsprechen soweit ersichtlich den ethischen und rechtlichen Standards zum Schutz der Rechte von Patientinnen und Patienten als Studienteilnehmerinnen und Studienteilnehmern (unter anderem das Recht auf Einwilligung nach Aufklärung, Datenschutz, das Recht, keinen unnötigen	vgl. Brönneke et al 2016 [21], Lee 2025 [23], Lo 2018 [24]

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehreseitige Tabelle)

Aspekt	Diskussion	Quelle
	Tests ausgesetzt zu werden). Lo 2018 geben an, dass die Studie vom Institutional Review Board des Virginia Polytechnic Institute and State University genehmigt wurde. Lee 2025 gegen an, dass die Studie vom institutional review board genehmigt wurde und alle Personen ihre schriftliche Einwilligung zur Teilnahme an der Studie abgegeben haben.	
Intellektuelles Eigentum	Vorliegend nicht relevant, weil es um Interventionen allgemein geht, nicht um eine spezielle Intervention, für die die Rechte am intellektuellen Eigentum geprüft werden könnten.	-
Kostenerstattung im öffentlichen Gesundheitssystem	▪ Anspruch GKV-Versicherter auf bedarfsgerechte und dem Stand der Wissenschaft entsprechende medizinische Behandlung (§§ 2, 12 SGB V) nach SGB V ▪ Leistungskatalog der gesetzlichen Krankenversicherung legt der G-BA in Richtlinien fest ▪ Kostenerstattung für „CAD-basierte CT-Befundung“, damit auch AI-unterstützte Auswertung von Röntgen- oder CT-Bildern der Lunge mit Inkrafttreten der Änderung der KFE-RL am 05.09.2025 mit Einführung einer Lungenkrebsfrüherkennung mittels Low-Dose-Computertomografie bei Rauchern, § 39 Abs. 5 KFE-RL mit Verweis auf die LuKrFrühErkV ▪ Festlegung von Qualifikations- und Fortbildungsanforderungen	Sozialgesetzbuch Fünftes Buch SGB V [97] G-BA 2024 [74] LuKrFrühErkV [5]
	Recht auf AI-unterstützte Auswertung von Röntgen- oder CT-Bildern der Lunge? Selbstbestimmungsrecht des Patienten vs. ärztliche Therapiefreiheit: ▪ Patient: Recht auf Wahl der Behandlung, abgesichert durch Art. 2 Abs. 2 S. 1 GG sowie das Einwilligungserfordernis, § 630d Abs. 1 S. 1 BGB ▪ Arzt: Berufsausübungsfreiheit, abgesichert in Art. 12 Abs. 1 GG ▪ grundsätzlich kein Anspruch auf eine bestimmte Behandlungsmethode	Botta 2024 [98] Zentrale Ethikkommission 2021 [52]
Spezielle medizinische Bereiche: Anforderungen an AI-Systeme nach AI Act	Zu den zentralen Anforderungen an Hoch-Risiko-AI-Systeme gehören nach dem AI Act insbesondere: ▪ human in the loop / Aufsicht: siehe unten in dieser Tabelle und „Arztvorbehalt“ ▪ Anbieter müssen ein Risikomanagement einführen, um Risiken für Sicherheit, Gesundheit, Grundrechte und die Umwelt systematisch zu identifizieren, bewerten und zu mindern, Art. 9	

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehrsseitige Tabelle)

Aspekt	Diskussion	Quelle
	- Trainings-, Validierungs- und Testdatensätze müssen qualitativ hochwertig, repräsentativ, aktuell und frei von Verzerrungen sein, Art. 10 - Es muss eine detaillierte technische Dokumentation vorliegen, Artikel 11 - Anwender müssen über Zweck, Funktionsweise und Grenzen des AI-Systems informiert werden, Art. 13 - nach Inverkehrbringen müssen Anbieter ein Monitoringsystem einrichten, um Risiken nach Nutzung, Weiterlernen oder Updates zu erkennen und zu mindern, Art. 61	
Spezielle medizinische Bereiche: Arztvorbehalt	- AI Act Art. 14: AI-Systeme dürfen nur als Unterstützung dienen, die finale Diagnoseentscheidung liegt bei der behandelnden Radiologin oder dem behandelnden Radiologen („human in the loop“) - Festlegung der Befundungs-Reihenfolge beim Lungenkrebsfrüherkennungs-Screening gem. § 5 LuKrFrühErkV: zunächst nur ärztliche, danach AI-unterstützte Befundung, finale Entscheidung liegt in ärztlicher Hand - keine Delegation und damit Einsatz technischer Geräte unproblematisch, solange ärztliches Handeln das System „beherrscht“, indem es „die durch das System ausgeführten Schritte vorgibt, plant, auslost, steuert und überwacht sowie die Ergebnisse kontrolliert“	AI Act [6] Lungenkrebsfrüherkennungsverordnung (LuKrFrühErkV) [5] Eichelberger 2024 [90]
Spezielle medizinische Bereiche: Überprüfung von AI- Ergebnissen	Verpflichtung, Patientinnen und Patienten die aus ihrer Sicht beste Behandlung anzubieten, § 630a BGB („Die Behandlung hat nach den zum Zeitpunkt der Behandlung bestehenden, allgemein anerkannten fachlichen Standards zu erfolgen, soweit nicht etwas anderes vereinbart ist“; dazu gehört auch die kritische Überprüfung der Ergebnisse der von ihnen eingesetzten technischen Instrumente	Deutscher Ethikrat 2023 [47] (S. 202 f.)
Spezielle medizinische Bereiche: Ärztliche Qualifikation	Festlegung von Anforderungen an Weiterbildung, Fortbildung und Fallzahlen, § 6 LuKrFrühErkV	LuKrFrühErkV [5]
Spezielle medizinische Bereiche: Qualitätssicherung	Vorgaben zur Qualitätssicherung des Gesamtablaufs in der Früherkennungskette, auch der Befundung der CT-Aufnahme, § 7 LuKrFrühErkV	LuKrFrühErkV [5] Begründung zur LuKrFrühErkV [5,10]
Spezielle medizinische Bereiche: Haftung	Haftungs-Konstellationen - Schaden beim Einsatz von AI - Schaden wegen Nichteinsatzes von AI - Schaden aufgrund fehlerhafter AI	
	Arzthaftung vs. Hersteller- / Produkthaftung	Katzenmeier 2023 [55] (S. 77)

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehrsseitige Tabelle)

Aspekt	Diskussion	Quelle
	▪ Bislang keine speziellen Vorschriften zur Haftung für Schäden infolge des Einsatzes von AI ▪ Forderung Bundesärztekammer: „Es muss klar geregelt sein, bei welchen Fehlern der Hersteller und bei welchen Fehlern die Ärztin oder der Arzt bzw. die medizinische Einrichtung haftbar ist.“	Bundesärztekammer 2025 [99]
	Schaden beim Einsatz von AI: Arzthaftung (Arzt oder Krankenhausträger) ▪ Rechtsgrundlage: Behandlungsvertrag (§§ 630a, 280 Abs. 1 BGB) oder Delikt (§ 823 ff BGB) ▪ Haftung wegen Sorgfaltspflichtverletzung §§ 630a, 276 BGB z. B. bei Verletzung der Aufklärungspflicht; Verwendung eines nicht zertifizierten Medizinprodukts, Fehler bei der Inbetriebnahme und Anwendung des AI-Systems	Eichelberger 2024 [90] Zentrale Ethikkommission 2021 [52]
	Schaden wegen Nichteinsatzes von AI ▪ Pflicht zum Einsatz neuer Methoden erst dann, wenn die neue Methode ein geringeres Risiko oder bessere Heilungschancen verspricht als der bisherige Standard und das in der Wissenschaft im Wesentlichen unumstritten ist ▪ keine generelle Pflicht, stets neueste Therapien und Geräte einzusetzen, gegebenenfalls aber Informationspflicht gegenüber Patienten über neue, bereits etablierte Entwicklungen	BGH, Urteil vom 26.11.1991 (<u>Az. VI ZR 389/90</u>) BGH, Urteil vom 22.09.1987 (<u>Az. VI ZR 238/86</u>), Eichelberger 2024 [90] Zentrale Ethikkommission 2021 [52]
	Schaden beim Einsatz von AI: Produkthaftung ▪ der Hersteller eines Produktes ist zum Schadenersatz verpflichtet, wenn jemand durch den Fehler eines Produktes an seinen Rechtsgütern verletzt wird § 1 Abs. 1 S. 1 ProdHaftG; Verschulden ist nicht erforderlich ▪ AI-System als Produkt im Sinne des § 2 ProdHaftG bislang streitig ▪ explizite Einordnung von Software als Produkt gemäß Art. 4 Abs. 1 Nr. 1 S. 2 ProdHaftRL (RL (EU) 2024/2853, in Kraft seit 9.12.2024: AI-Systeme fallen als Software grundsätzlich unter die Produkthaftung; umgesetzt im Referentenentwurf § 2 Abs. 1 Nr. 3 ProdHaftG (Referentenentwurf), bis 12/2026 in nationales Recht umzusetzende Richtlinie; danach gilt: ▪ Schaden im Sinne des Art. 6 Abs. 1 lit. c ProdHaftRL sind in erster Linie Personenschäden und Sachschäden. ▪ Fehler im Sinne des § 3 ProdHaftG: wenn das Produkt nicht die Sicherheit bietet, die unter Berücksichtigung aller Umstände berechtigterweise erwartet werden kann; z. B. Fabrikationsfehler	ProdHaftG [100] Referentenentwurf zur Modernisierung des ProdHaftG [102] Katzenmeier 2023 [55] Zentrale Ethikkommission 2021 [52]

Tabelle 34: Informationsaufbereitung zu den identifizierten rechtlichen Aspekten
(mehreseitige Tabelle)

Aspekt	Diskussion	Quelle
	(mangelhafte Fertigung), Konstruktionsfehler (wie fehlerhafter Steuerungsalgorithmus, fehlerhaft designtes künstliches neuronales Netz, unvollständiger Trainingsdatensatz), Instruktionsfehler (fehlerhafte Gebrauchsinformation) ▪ Fehler im Sinne des Art. 7 Abs. 2 lit. c ProdHaftRL: „Bei der Beurteilung der Fehlerhaftigkeit eines Produktes sind alle Umstände zu berücksichtigen, einschließlich: [...] der Auswirkungen der Fähigkeit des Produkts, nach seinem Inverkehrbringen oder seiner Inbetriebnahme weiter zu lernen oder neue Funktionen zu erwerben, auf das Produkt“ ▪ Fehler im Sinne des § 7 Nr. 3 ProdHaftG Referentenentwurf: Ein Produkt ist fehlerhaft, wenn es nicht die Sicherheit bietet, die nach deutschem Recht oder nach dem Recht der Europäischen Union vorgeschrieben ist oder die erwartet werden darf. Bei der Beurteilung der Fehlerhaftigkeit sind alle Umstände zu berücksichtigen, einschließlich: [...] die Auswirkungen von Fähigkeiten des Produkts, nach dem Inverkehrbringen oder der Inbetriebnahme weiter zu lernen oder neue Funktionen zu erwerben, auf das Produkt.	
Spezielle medizinische Bereiche: Anforderungen an Betreiber von Hochrisiko-AI-Systemen	Kap. III, Art. 26 AI Act Anforderungen an Kliniken und Arztpraxen als Betreiber von AI-Systemen unter anderem bezüglich ▪ Sicherstellung des zweckbestimmungsgemäßen Gebrauchs der Systeme, auch hinsichtlich der menschlichen Aufsicht, ▪ Informationspflichten bezüglich schwerwiegender Vorfälle, ▪ die Aufbewahrung automatisch erzeugter Protokolle (z. B. Audit-Logs), ▪ die Information der Mitarbeitenden über den Einsatz der Systeme ▪ Einhaltung der Datenschutzbestimmungen	AI Act [6]

AI: künstliche Intelligenz; AMG: Arzneimittelgesetz; BDSG: Bundesdatenschutzgesetz; BfArM: Bundesinstitut für Arzneimittel und Medizinprodukte; BGB: Bürgerliches Gesetzbuch; CAD: Computerassistierte Detektion; CE: Conformité Européenne; CT: Computertomografie; DigiG: Digital- Gesetz; DSGVO: Datenschutz-Grundverordnung; EHDS: European Health Data Space; ePA: elektronische Patientenakte; EU: Europäische Union; G-BA: Gemeinsamer Bundesausschuss; GDNG: Gesundheitsdatennutzungsgesetz; GG: Grundgesetz; GKV: gesetzliche Krankenversicherung; HTA: Health Technology Assessment; IT: Informationstechnologie; KFE-RL: Krebsfrüherkennungs- Richtlinie; LuKrFrühErkV: Lungenkrebs-Früherkennungs-Verordnung; MDR: Medical Device Regulation; MDSW: Medical Device Software; MPBetreibV: Medizinprodukte-Betreiberverordnung; ProdHaftRL: Produkthaftungsrichtlinie; ProdHaftG: Produkthaftungsgesetz; SGB: Sozialgesetzbuch

A5.4 Organisatorische Aspekte

A5.4.1 Informationsbeschaffung zu organisatorischen Aspekten der zu bewertenden Technologie

Insgesamt wurden durch die orientierende Recherche in bibliografischen Datenbanken und weiteren Informationsquellen 27 Publikationen eingeschlossen.

Die Zitate der relevanten Publikationen finden sich in Abschnitt A7.3.4.

A5.4.2 Identifizierte organisatorische Aspekte

In Tabelle 35 sind die identifizierten organisatorischen Implikationen der Technologie entsprechend der Strukturierung nach Perleth 2014 [22] dargestellt.

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
Einfluss auf Voraussetzungen der Leistungserbringung	Änderung des Ortes der medizinischen Versorgung	<u>Zielgruppe 1: Personen mit Verdacht auf Lungenkrebs</u> Eine Auswertung von Röntgen- oder CT-Bildern der Lunge bei Personen mit Verdacht auf Lungenkrebs erfolgt derzeit in ambulanten oder stationären Radiologiezentren durch eine Radiologin oder einen Radiologen. Eine AI-unterstützte Befundung oder eine doppelte Befundung durch 2 Personen bei dieser Zielgruppe erfolgt nicht regelhaft. Eine AI-unterstützte Abklärung / Befundung wird innerhalb von Forschungs- oder Pilotprojekten in Zentren, die als Lungenkrebszentren zertifiziert sind, angeboten. Eine Aufstellung, wie viele der 81 im Jahr 2024 in Deutschland DKG-zertifizierten Lungenkrebszentren (mit 104 zertifizierten Standorten) bereits mit AI-unterstütztem Röntgen oder CT arbeiten, konnte nicht identifiziert werden.	[104]
		<u>Zielgruppe 2: Personen, die eine Lungenkrebsfrüherkennung in Anspruch nehmen können:</u> Derzeit erfolgt eine Lungenkrebsfrüherkennung in Deutschland nicht regelhaft. Im Prüfungsbericht des Bundesamtes für Strahlenschutz (BfS), der Grundlage der Lungenkrebs-Früherkennungsverordnung (LuKrFrühErkV) war, wurde vorgeschlagen, die Lungenkrebsfrüherkennung für anspruchsberechtigte Personen in von der Deutschen Krebsgesellschaft zertifizierten Lungenkrebszentren sowie assoziierten klinischen Einrichtungen durchzuführen. Dieser Vorschlag wird auch von den Deutschen Fachgesellschaften befürwortet. Das Screening in zertifizierten Lungenkrebszentren und ihren Partnern scheint angesichts der fachlichen und technischen Anforderungen, die die LuKrFrühErkV an die Software-gestützte Früherkennung stellt, ein realistischer Ort der Versorgung. An dieser Stelle wird darauf verwiesen, dass in der Begründung zur LuKrFrühErkV explizit davon ausgegangen wird, dass ein CAD-gestütztes Röntgen / CT eingesetzt wird, nicht notwendigerweise jedoch ein AI-unterstütztes System.	[5,63,105]

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
	Änderungen bei den Qualifikationsanforderungen für die Leistungserbringerinnen und -erbringer / zusätzliches oder reduziertes Personal	<u>Zielgruppe 1 und Zielgruppe 2</u> Es ist unklar, ob für beide Zielgruppen hinsichtlich der Qualifikationsanforderungen an Leistungserbringende für einen Einsatz AI-unterstützter Systeme zur Befundung vergleichbare Anforderungen gelten werden. Die Anforderungen sind in der LuKrFrühErkV übergreifend bezogen auf den Einsatz CAD-basierter Systeme formuliert. Der Einsatz AI-unterstützter Systeme wird nicht explizit thematisiert. <u>Qualifikationsanforderungen der LuKrFrühErkV an die Leistungserbringenden:</u> ▪ Für Ärztinnen und Ärzte, die die Zielpersonen gemäß § 2 der LuKrFrühErkV über die Früherkennung informieren, gilt: sie müssen approbierte bzw. zugelassene Fachärzte der Inneren Medizin, der Allgemeinmedizin, oder der Arbeitsmedizin sein, die zusätzliches Wissen in der Lungenkrebsfrüherkennung, z. B. durch Fortbildungen, erworben haben. Für Fachärztinnen und -ärzte der Radiologie (Erst- und Zweitbefundung), die die eigentliche Früherkennung mittels Thorax-CT durchführen: hier gelten über die im Strahlenschutzrecht vorgegebenen Voraussetzungen für die Durchführung eines Thorax CT hinausgehende Anforderungen gemäß § 6 der LuKrFrühErkV: Erfüllung bestimmter CT-Scan-Mindestmengen, die ein Arzt im Jahr vor Aufnahme der Ausübung der Früherkennung befundet und dokumentiert haben muss (200 Untersuchungen, gilt für Erstbefunderin oder Erstbefunder), sowie 100 (Erstbefunderin oder Erstbefunder) bzw. 200 (Zweitbefunderin oder Zweitbefunder) Untersuchungen im 1. Jahr und 200 (Erstbefunderin oder Erstbefunder) bzw. 400 (Zweitbefunderin oder Zweitbefunder) Untersuchungen ab dem 2. Jahr der Ausübung der Lungenkrebsfrüherkennung. Außerdem sind Fortbildungen in der Untersuchung zur Lungenkrebsfrüherkennung nachzuweisen. Die LuKrFrühErkV sieht vor, dass die Erstbefunderin oder der Erstbefunder zunächst ohne und dann mit Unterstützung einer CAD-Software befundet, die Zweitbefunderin oder der Zweitbefunder dies bei kontroll- oder abklärungsbedürftigen Befunden ebenfalls tut, und es im Abschluss zu einer gemeinsamen Befundbeurteilung kommt. Die LuKrFrühErkV formuliert keine expliziten Anforderungen an die Qualifikation der Leistungserbringende hinsichtlich AI-unterstützter Befundung, anders als der Bericht des BfS, der auf die Notwendigkeit der Qualifizierung / Weiterqualifizierung in CAD- und / oder AI-Software in der Lungenkrebsfrüherkennung explizit hinweist.	[5,63]

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
		<u>Qualifikationsanforderungen an Leistungserbringende - Einschätzungen in Deutschland</u> Positionspapiere und Stellungnahmen weisen darauf hin, dass politische und klinische Entscheidungsträger in Deutschland die Notwendigkeit einer (Weiter-) Qualifizierung der Leistungserbringenden hinsichtlich AI-unterstützter Systeme anerkennen (siehe zum Beispiel das GKV-Positionspapier aus 2024 oder die Stellungnahmen von Blum 2024 oder der Bundesärztekammer aus 2025). In letzterer wird auf die Bedeutung einer Qualifizierung über Erfahrungswissen und AI-Schulungen hingewiesen. Demgemäß sollte sichergestellt werden, dass in der ärztlichen Aus-, Weiter- und Fortbildung sowohl AI-spezifische als auch ausreichend AI-unabhängige Kenntnisse und Fertigkeiten vermittelt werden, um so einem schleichenden Verlust von Erfahrungswissen (dem sogenannten De-Skilling) vorzubeugen und erfahrene / unerfahrene Leistungserbringende für den Einsatz der AI umfassend auszubilden. Auch der deutsche Ethikrat weist darauf hin, dass „bei empirisch belegter Überlegenheit von AI-Anwendungen die Anwendung von AI-Systemen in Aus-, Fort- und Weiterbildung für ärztliches Fachpersonal sowie andere Gesundheitsberufe zu übernehmen“ sei.	[47,99,105-107]

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
		<u>Qualifikationsanforderungen- Internationale Stimmen / Studien</u> <u>Gleeson 2023</u>, Stellungnahme der European Society of thoracic imaging (ESTI): Analog zur Stellungnahme der Bundesärztekammer wird auf die möglichen negativen Effekte der Nutzung einer AI bezüglich der Ausbildung von Radiologinnen und Radiologen hingewiesen. <u>Akudjedu 2023</u>: Ergebnisse eine Online-Befragung von 400 in der Radiologie tätigen Ärztinnen und Ärzte (davon 314 Fragebögen ausgewertet, ca. 50 % Nordamerika, 20 % Asien und 10 % Europa): Mehr als 50 % der Befragten fühlten sich bereit, AI einzusetzen; mehr als 60 % der Befragten berichten über zu wenig verfügbares AI-Training. Fast 70 % waren dafür, AI-Inhalte in die ärztlichen Ausbildungs-Curricula für Radiologinnen und Radiologen aufzunehmen und sprachen in diesem Zusammenhang auch das Problem des De-Skillings bzw. die Bedeutung des Aufbaus von Erfahrungswissen ohne AI an (siehe hierzu auch <u>Faric 2024</u>). <u>Kauczor 2023 / Shaham 2025</u>: Das Projekt SOLACE (Strengthening the screening of Lung Cancer in Europe) startete im April 2023, endet im März 2026 und wird von der EU im Rahmen des Programms EU4Health gefördert. Ziel von SOLACE ist die Unterstützung der Entwicklung und Implementierung nationaler Lungenkrebscreening-Programme in Europa. Teil des SOLACE-Projektes ist die Entwicklung eines Trainings-Programms wie die "ESTI (European Society of Thoracic Imaging) Lung Cancer Screening Certification", ein europäisches Qualifizierungs-Programm für das Lungenkrebscreening. Ziel eines Arbeitspaket von SOLACE ist, die Anforderungen an die Ausbildung der am Lungenkrebs beteiligten Kliniker zu standardisieren. Unklar bleibt, ob und wie die AI-unterstützte Befundung hier integriert wird / werden soll. <u>Faric 2024</u>: Befragungen von Leistungserbringenden weisen darauf hin, dass eine gute Einführung und Schulung bezogen auf ein neu eingesetztes AI-System wichtig für die Akzeptanz und das Vertrauen der Anwender (Radiologinnen und Radiologen) in die AI sei, wodurch eine kontinuierliche / effiziente Nutzung der AI möglich sei. <u>Ewals 2024, Ewals 2023</u>: Obwohl die Einführung bzw. die AI-spezifische Schulung von Radiologinnen und Radiologen wichtig für die Akzeptanz und Nutzung durch Leistungserbringende zu sein scheint, berichten Studien zu neuen AI-Systemen eher selten (ca. 40 % der gesichteten Studien), wie die Radiologinnen und Radiologen auf eine spezifische AI geschult wurden.	[48,50,58,73,86,88,108,183]

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
	Änderungen bei den Anforderungen an Personal, Material und Organisation der Leistungserbringung (Strukturqualität)	<u>Personal für die Leistungserbringung</u> <u>Zielgruppe 1:</u> In verschiedenen Publikationen wird besprochen, dass durch die AI-unterstützte Befundung bei Personen mit Verdacht auf Lungenkrebs die AI als „Zweite Meinung“ (anstelle einer anderen Radiolog:in) genutzt und hierdurch das vorhandene Personal entlastet werden könnte. Hierdurch können möglicherweise personelle Ressourcen freigesetzt oder anders eingesetzt werden. Solche Einsparungen oder Verschiebungen menschlicher Ressourcen können aber auch zu Diagnose-Fehlern (beispielsweise, wenn die AI missverständliche Ergebnisse liefert) und längerfristig zu einem Abbau von Erfahrungswissen (De-Skilling) führen. <u>Zielgruppe 2:</u> Aufgrund der strikten Vorgaben der LuKrFrühErkV zur computergestützten Befundung durch Erst- und Zweitbefunder sowie zur Qualitätssicherung scheint ein personalbezogenes Entlastungspotenzial nicht gegeben.	[5,48,73,99,105,106]
		<u>Material und Organisation der Leistungserbringung</u> Die LuKrFrühErkV gibt klare Vorgaben hinsichtlich der Organisation der Früherkennung selbst (Ansprache der Anspruchsberechtigten, Durchführung der Untersuchung, Maßnahmen zur Qualitätssicherung, Anforderungen an den Computertomografen und die CAD-Software), aber nicht oder nur in geringem Ausmaß hinsichtlich der Organisation der Früherkennung in den Radiologiezentren selbst. Auch die Möglichkeit des Einsatzes einer AI-Software wird nicht konkretisiert. Es ist davon auszugehen, dass für beide Zielgruppen hinsichtlich Material und Organisation der Leistungserbringung in den Radiologiezentren vergleichbare Anforderungen gelten.	[5]

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
		<u>Material – Daten für Training und Validierung einer AI / Qualität einer AI</u> Grundsätzliche Anforderungen an die Daten, die für das Training einer AI genutzt und verarbeitet werden und die die Betreuung und Auswertung der Daten betrifft, wurden auf europäischer und deutscher Ebene definiert bzw. rechtlich festgehalten. Zur Prüfung der Qualität von AI-Systemen wurden darüber hinaus in Positionspapieren und wissenschaftlichen Publikationen Anforderungen definiert und Checklisten entwickelt. Diese können helfen, die Qualität einer AI und die Implikationen einer AI-Einführung strukturierter einzuschätzen und informierte Entscheidungen hierüber zu treffen. Sofern Daten für das Training einer AI (lernende bzw. weiterlernende AI-Systeme) in Deutschland aus bestehenden Patientendaten gezogen werden, müssen die Betroffenen hiermit einverstanden sein. Eine Studie zeigt auf, dass in Deutschland diesbezüglich noch Vorbehalte in der Bevölkerung bestehen. <u>Büthe 2024:</u> Die Bereitschaft zur Freigabe von Daten in anonymisierter Form muss gegeben sein. In einer repräsentativen Umfrage bei 4000 Personen in Deutschland Anfang 2024 wurde erhoben, dass 2/3 der Befragten grundsätzlich bereit waren, Daten aus ihrer elektronischen Patientenakte für Zwecke der Früherkennung und der medizinischen Forschung bereitzustellen, 1/3 der Personen war dagegen. Sorgen um die Sicherheit persönlicher Daten (Datenschutz) und den Verlust der Privatsphäre wurden geäußert.	[53,109-112,184] Domäne rechtliche Aspekte (Abschnitt 5.4)

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
		<u>Organisation der Leistungserbringung (Strukturqualität)</u> Die Stellungnahme der European Society of Thoracic Imaging fasst die Anforderungen an die Organisation bei der Implementierung einer AI in der klinischen Radiologie zusammen: ▪ Kritische Auswertung der für die AI genutzten Trainings- und Validierungsdaten (Anforderung an die AI: transparente Dokumentation aller relevanten Teile) ▪ Vorliegen einer CE-Kennzeichnung / Zulassung der AI ▪ Prüfung / Sicherstellung der Kompatibilität der AI mit der bestehenden IT-Infrastruktur ▪ Einschätzung der Anschaffungskosten und laufender Kosten der AI ▪ Sicherstellung der Einhaltung der Bestimmungen der europäischen Datenschutz-Grundverordnung (DSGVO, englisch: GDPR) ▪ Planung von Aus- / Weiterbildung für alle Nutzer der AI – neben AI-Wissen ist die Bildung von hohem Anwendungs-CT-Wissen / Erfahrung wichtig, um langfristige negative Effekte der AI auf die Ausbildung unerfahrener Radiologinnen und Radiologen zu vermeiden ▪ Nach Implementierung: Evaluation, Prüfung möglicher Zugangsbarrieren zur AI, System von User-Feedback und reguläre Prüfungen Einhaltung ethischer Anforderungen / Standards, Transparenz der AI, keine off-label Nutzung (keine AI zur Erst- / Allein-Befundung) etc.	[88]

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
		<u>Organisation der Leistungserbringung (Strukturqualität), hinderliche und förderliche Faktoren:</u> <u>Van Leeuwen 2024:</u> Basierend auf einer Umfrage bei Radiologieabteilungen in Krankenhäusern in den Niederlanden zwischen 2020 und 2022 wurde deutlich, dass AI immer häufiger und am häufigsten zur Unterstützung von CT-Untersuchungen der Lunge eingesetzt wird. Die befragten 110 Entscheidungsträger gaben an, dass die größten Hürden bei der Einführung von AI die hohen Kosten (Anschaffungs- und laufende Kosten bei mangelndem Budget) und die Integration in bestehende IT-Systeme seien (mangelnde Kompatibilität zwischen AI und IT-System als Herausforderung) <u>Jungmann 2024:</u> Jungmann 2024 weist darauf hin, dass aus Sicht der Nutzerinnen und Nutzer die Integration von AI in bestehende Prozesse, Arbeitsabläufe und IT-Systeme ohne zusätzliche Zeitverzögerungen, ohne Zusatz-Aufwand anwendbar sein müssen, damit AI von den Radiologinnen und Radiologen angenommen wird. Siehe hierzu auch die Erhebungen von <u>Faric 2024</u> und <u>Akudjedu 2023</u>: zusätzlich zur Betonung der Bedeutung der Integration in bestehende Prozesse, Arbeitsabläufe und IT-Systeme wird darauf hingewiesen, dass eine AI, wenn sie gut in bestehende IT-Systeme integriert ist, auch Dokumentationspflichten einer Befundung für die Radiologinnen und Radiologen vereinfachen / übernehmen kann.	[48,73,113,114]
Einfluss auf Prozesse	Alternative Technologien für dieselbe Fragestellung	Standard-Befundung durch Radiologinnen und Radiologen allein oder mit Hilfe einfacher CAD- Systeme. Einfache CAD-Systeme unterstützen die Radiologin oder den Radiologen durch regelbasierte oder algorithmische Verfahren, die auf bestimmten Voreinstellungen und Angaben beruhen. Sie sind abzugrenzen von AI-Systemen, die anhand großer Datensätze (> 1000) darauf trainiert werden, Muster in CT-Bildern zu erkennen (z. B. mittels machine learning, deep learning)	

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
	Nutzung von Gesundheitsleistungen / Ressourcen	Der Einsatz von AI in der Befundung wird neben den Anschaffungs- und laufenden Kosten für die AI selbst zu einem Mehreinsatz von Ressourcen beim Personal (für AI-Module für die Aus-, Fort- und Weiterbildung) und bei der Integration der AI in bestehende IT-Systeme führen. Befragungen legen zwar nahe, dass eine AI den Radiologinnen und Radiologen die Arbeit erleichtern und die Befundung möglicherweise schneller / effizienter gestalten und sogar zu Personaleinsparungen führen kann (AI als Zweitbefunder). Belegt ist eine solche Effizienzsteigerung durch aussagekräftige Studien bislang jedoch nicht. Betont wird, dass zwar inzwischen zahlreiche CE-zertifizierte AI-Systeme für die Radiologie auf dem Markt verfügbar sind, die stark beworben werden, es aber keine aussagekräftige Evidenz zum patientenrelevanten Nutzen und Schaden dieser AI-Systeme gibt. Diese Aussage unterstützen auch die fehlenden Ergebnisse zu Nutzen und Schaden der AI-unterstützten Befundung des vorliegenden Berichts. Auch Kosten- oder Kosten-Wirksamkeitsstudien konnten bisher nicht überzeugend zeigen, dass AI im Lungenkrebscreening an einzelnen Stellen oder übergreifend zu Einsparungen führt oder kosten-effektiv ist (z. B. durch effizientere Prozesse oder die Vermeidung später teurer Behandlungen)	(Kapitel 3 und 4) [84,88,114,115]
	Kommunikations- und Kooperationsformen	Die Zusammenarbeit im radiologischen Team muss einen neuen Player – die AI – integrieren, ohne hierdurch die Qualität der Befundung zu gefährden und z. B. Erfahrungswissen auf Dauer zu verringern. Hierzu scheint ein „mixed-expertise approach“ sinnvoll (d. h. erfahrene Kliniker arbeiten mit weniger erfahrenen in einem Team). Während gemäß LuKrFrühErkV klare Rollen und Kooperationsformen zwischen den Radiologinnen und Radiologen bei der Lungenkrebsfrüherkennung festgelegt sind, ist dies bei der Befundung von Patientinnen und Patienten mit Verdacht auf Lungenkrebs zumindest derzeit nicht gegeben.	[5,116,117]
		Die Kommunikation mit den Betroffenen wird für die Früherkennung von Lungenkrebs – entsprechend den Vorgaben der LuKrFrühErkV – im Rahmen eines partizipativen Entscheidungsprozesses erfolgen und auf evidenzbasierter Information (Versicherteninformation des G-BA) zum CAD / AI-unterstützten Lungen-CT beruhen. Es ist unklar, inwiefern die Aufklärungs- und Informationspflichten zum Lungenkrebscreening wie in der LuKrFrühErkV definiert sich auch auf die AI-unterstützte Befundung und auf die Befundung von Verdachtspersonen erstrecken werden.	[5,74] Domäne rechtliche Aspekte (Abschnitt 5.4)

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
Weitere Aspekte	Interessengruppen / Akzeptanz	Die Interessen von Betroffenen, Behandelnden und Gesellschaft können sehr unterschiedlich sein <u>Betroffene (bei Verdacht auf Lungenkrebs):</u> Wunsch nach schneller Abklärung: sichere Erkennung einer schweren Erkrankung und schnelle Behandlung versus Entwarnung <u>Betroffene (bei Früherkennung):</u> Wunsch nach Bestätigung (keine Erkrankung) versus Erkennen einer Erkrankung in frühem Erkrankungsstadium, schnelle Behandlung bzw. bessere Chancen auf Überleben / Heilung sowie Vermeidung von unerwünschten Nebenwirkungen. Wenn AI die frühere Erkennung verbessert, dann wird sie von den Betroffenen gewünscht. <u>Behandelnde:</u> Absicherung des eigenen Befunds, Vermeidung / Minimierung falsch-positiver bzw. falsch-negativer Befunde; Positive Einstellung zu AI-unterstützter Befundung, wenn damit Vorteile für die Befundung einhergehen (Zeitersparnis, bessere / effizientere Befundung) <u>Gesundheitswesen:</u> Früherkennung: Wunsch nach früher Erkennung von erkrankten Personen und mögliche Einsparpotentiale durch weniger Behandlungen in späten Stadien / Palliativbehandlung, Verringerung von Mortalität bei den Zielgruppen der Früherkennung, Vermeidung zu hoher Kosten (mögliche Abwägung zwischen hohen Kosten für AI und Einsparpotenzialen) – die Begründung zur LuKrFrühErkV lässt die Frage nach dem Einsatz AI-unterstützter Systeme und besonders deren Finanzierung offen.	Betroffenenbefragung (Abschnitt 5.3) [5]

Tabelle 35: Informationsaufbereitung zu den identifizierten organisatorischen Aspekten (mehrseitige Tabelle)

Themenbereich	Leitfrage	Erläuterungen	Quelle(n)
	Planung von Kapazitäten, Investitionen	<u>Gleeson 2023</u> fasst die Anforderungen an die Planung von Kapazitäten / Investitionen für die Einführung einer AI wie folgt zusammen. ▪ Erstellung einer Radiologie-spezifischen AI-Strategie mit klarer Definition der Rolle, des Typs und des Ziels einer AI-Anwendung im klinischen Workflow ▪ Planung einer kontinuierlichen Qualitätskontrolle ▪ Erstellung eines finanzierbaren Business-Plans ▪ CE-zertifizierte Software ▪ kritisches Review von Trainings- und Validierungsdaten zum Ausschluss von Bias etc. ▪ Integration der AI in klinischen Workflow und in IT-Infrastruktur ▪ Klärung / Vorbereitung von patientenbezogenen Aufklärungs- / Informationspflichten ▪ Sicherstellung des Datenschutzes gemäß DSGVO ▪ Planung der AI-spezifischen Aus- / Fortbildung der Radiologinnen und der Radiologen ▪ Berücksichtigung möglicher negativer Auswirkungen von AI auf die Weiterbildung unerfahrener Radiologinnen und Radiologen ▪ Klärung rechtlicher Fragen	[88]
AI: künstliche Intelligenz; BfAS: Bundesamt für Strahlenschutz; CAD: Computer-Assisted Detection; CE: Conformité Européenne; CT: Computertomografie; DKG: Deutsche Krebsgesellschaft; DSGVO: Datenschutz-Grundverordnung; ESTI: European Society of Thoracic Imaging; G-BA: Gemeinsamer Bundesausschuss; GKV: gesetzliche Krankenversicherung; IT: Informationstechnologie; LuKrFrühErkV: Lungenkrebs-Früherkennungs-Verordnung			

A5.5 Umwelt- und Klimaaspekte

A5.5.1 Informationsbeschaffung zu Umwelt- und Klimaaspekten der zu bewertenden Technologie

Insgesamt wurden durch die orientierende Recherche in bibliografischen Datenbanken und weiteren Informationsquellen 10 Publikationen eingeschlossen.

Die Zitate der relevanten Publikationen finden sich in Abschnitt A7.3.5.

A5.5.2 Identifizierte Umwelt- und Klimaaspekte

In Tabelle 36 sind die identifizierten Umwelt- und Klimaauswirkungen deskriptiv dargestellt.

Tabelle 36: Informationsaufbereitung zu den identifizierten Umwelt- und Klimaaspekten (mehreseitige Tabelle)

Umwelt- und Klimaaspekte	Quelle
CO²-Emissionen	
Der Gesundheitssektor ist derzeit für 4,4 % der globalen CO ₂ -Emissionen verantwortlich. USA: 8–10 % der CO ₂ -Emissionen durch den Gesundheitssektor; etwa 1 % der weltweiten Gesamt-CO ₂ -Emissionen entstehen durch die medizinische Bildgebung. Die Radiologie hat traditionell einen hohen Energieverbrauch / CO ₂ -Emissionen – die AI steigert das Problem. Die Mehrheit der medizinischen Geräte, die AI (Deep-Learning) nutzen und in den USA von der Food & Drug Administration zugelassen sind, gehören zur Radiologie – hierunter gehören die meisten in den Bereich der radiologischen Bildverarbeitung / Software und Computertomografie.	[120-125]
Kocak 2025 stellt verschiedene Schätzungen hinsichtlich des CO ₂ Ausstoßes für das Training von AI-Systemen in der Radiologie im Vergleich zu anderen AI-Anwendungen dar (verschiedene Quellen für diese Schätzungen werden in der Referenz herangezogen und sind dort benannt). Die Zahlen beziehen sich allein auf das Training einer AI. Zur Berechnung von CO ₂ -Ausstößen für Training <u>und</u> Inferenz der AI (d. h. Aussagen mit der AI basierend auf neuen Daten, z. B. Patienten-Daten) müssen die genannten Zahlen mit 1000 multipliziert werden); Ausstoß von CO ₂ -Äquivalenten (in Tonnen, CO ₂ -eq (t)): Mensch (pro Kopf und Jahr): 7–10 CO ₂ -eq (t) PKW (Lebenszeit+Benzin): 60 CO ₂ -eq (t) CT (5000 Scans): 40 CO ₂ -eq (t) CT (10 000 Scans): 80 CO ₂ -eq (t) MRT (10 000 Scans): 180 CO ₂ -eq (t) OpenAI's GPT-4: 21 000 CO ₂ -eq (t)	[84]
Cloud-basierte Lösungen zur Lagerung und Verarbeitung von Daten aus der Radiologie sind effizienter als die in-house Lagerung (da geteilte Infrastruktur bei cloud-basierter Lagerung). Cloud-basierte Lösungen mit großen Datenservern sind – einschließlich der notwendigen Kühlsysteme – sehr stromintensiv. Es wird geschätzt, dass diese Datenserver ungefähr 1–2 % des weltweiten Stromverbrauchs ausmachen, mit steigender Tendenz. AI-basierte Radiologiesysteme haben hieran einen signifikanten Anteil.	[84,126,127]

Tabelle 36: Informationsaufbereitung zu den identifizierten Umwelt- und Klimaaspekten (mehrsseitige Tabelle)

Umwelt- und Klimaaspekte	Quelle
Cloudbasierte / Server-Lösungen zur Datenlagerung- und Verarbeitung werden weltweit bislang häufig mit fossilen und nicht mit erneuerbaren Energieträgern betrieben. Der Ersatz fossiler durch erneuerbare Energien kann die CO2-Bilanz der Server verbessern.	[84,126,127]
Der Transport von Rohstoffen oder anderen Komponenten in die Länder, in der Hardware für die AI-Systeme produziert wird, geht ebenfalls mit CO2-Emissionen einher.	[84]
Toxische Substanzen / Strahlenexposition	
Hochleistungs-Hardware der AI-Systeme (wie Grafikprozessoren (GPUs), Central Processing Units (CPUs)) enthalten giftige Stoffe wie Blei, Cadmium und Quecksilber. Oft wird Hardware vor Ablauf ihrer Lebensdauer erneuert, da in großer Geschwindigkeit neuere, schnellere Prozessoren auf den Markt kommen, was somit auch den Umlauf toxischer Substanzen erhöht / beschleunigt. [84,123]	[84,123]
Herstellung von AI-Hardware erfordert den Abbau und die Veredelung von seltenen Metallen wie Neodym, Tantal und Kobalt, die endliche Ressourcen sind. Abbau und Veredelung dieser Metalle gehen wiederum mit signifikanten CO2-Emissionen einher.	[84,123]
Mögliche Optimierung der CT-Strahlendosis für die Befundung mit AI: In der Literatur werden verschiedene Eigenschaften von AI-Anwendungen genannt, die eine bessere Überwachung und Steuerung der CT-Strahlenexposition ermöglichen. Zum Beispiel: die AI analysiert historische Scandaten von Patienten und schlägt automatisch Protokolle für die Befundung mit geringerer Strahlendosis vor, ohne die Bildqualität zu beeinträchtigen. Adaptive AI-Modelle optimieren die Scanparameter auf der Grundlage patientenspezifischer Faktoren wie Alter und Gewicht, um die Strahlenbelastung zu minimieren. Die AI-unterstützte Qualitätssicherung reduziert die Wiederholungsrate von Scans, indem sie Bilder von schlechter Qualität in Echtzeit erkennt.	[84]
Abfallmanagement	
Problem der Entstehung von Elektronikschrott (e-waste): Elektronikschrott bezieht sich in der Radiologie / Krankenhäusern auf die Hardware, die für den Einsatz von AI erforderlich ist. Diese Geräte enthalten oft giftige Stoffe wie Blei, Cadmium und Quecksilber, die bei unsachgemäßer Entsorgung die Umwelt und die menschliche Gesundheit schädigen können. Eine sachgemäße Entsorgung geht mit Kosten einher. Lösungsmöglichkeit ist das Recycling von z. B. bei seltenen Erden oder bestimmten Metallen, sofern möglich.	[84,123]
Weitere	
Wasserverbrauch für die Kühlung der für die AI benötigten Hochleistungsrechnern / Datenservern: Schätzungen zufolge könnte der weltweite Bedarf an AI bis 2027 einen Wasserverbrauch von 4,2 bis 6,6 Milliarden Kubikmetern verursachen. Dies entspricht dem 4-6-fachen des jährlichen Wasserverbrauchs Dänemarks oder der Hälfte des Verbrauchs Großbritanniens.	[84,128]
Der Abbau von seltenen Erden und anderen Materialien, die für die Herstellung von AI-Hardware benötigt werden, kann Lebensräume zerstören, zum Verlust der biologischen Vielfalt sowie zur Wasser- / Umweltverschmutzung in den Abbau-Regionen / Ländern führen.	[84]
AI: künstliche Intelligenz; CO2: Kohlenstoffdioxid; CPUs: Central Processing Units; CT: Computertomografie; GPT: Generative Pre-trained Transformer; GPUs: Grafikprozessoren; MRT: Magnetresonanztomografie; PKW: Personenkraftwagen; USA: United States of America	

A6 Literatur

1. Lungenkarzinom [online]. 2021 [Zugriff: 10.04.2025]. URL: <https://www.pschyrembel.de/Lungenkrebs/K0461/doc/>.
2. Robert Koch-Institut. Krebs in Deutschland für 2021/2023 [online]. 2025 [Zugriff: 12.01.2026]. URL: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/krebs_in_deutschland_2025.pdf?blob=publicationFile.
3. Wittekind C. TNM Klassifikation maligner Tumoren. Weinheim: Wiley-VCH; 2020.
4. Leitlinienprogramm Onkologie. Prävention, Diagnostik, Therapie und Nachsorge des Lungenkarzinoms [online]. 2025 [Zugriff: 15.05.2025]. URL: https://www.leitlinienprogramm-onkologie.de/fileadmin/user_upload/Downloads/Leitlinien/Lungenkarzinom/Version_4/LL_Lungenkarzinom_Langversion_4.0.pdf.
5. Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz. Verordnung über die Zulässigkeit der Anwendung der Niedrigdosis-Computertomographie zur Früherkennung von Lungenkrebs bei rauchenden Personen (Lungenkrebs-Früherkennungs-Verordnung – LuKrFrühErkV). Bundesgesetzblatt Teil 1 2024; (162): 1-5.
6. Europäisches Parlament, Rat der Europäischen Union. Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz) [online]. 2024 [Zugriff: 07.01.2026]. URL: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj?eliuri=eli%3Areg%3A2024%3A1689%3Aoj&locale=de>.
7. National Institute for Health and Care Excellence. Artificial intelligence (AI)-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals: early value assessment [online]. 2025 [Zugriff: 10.04.2025]. URL: <https://www.nice.org.uk/guidance/htg696>.
8. Institut für Qualitätssicherung und Transparenz im Gesundheitswesen. Methodische Grundlagen; Version 2.1 [online]. 2024 [Zugriff: 01.07.2025]. URL: https://igtig.org/downloads/berichte/2024/IQTIG_Methodische-Grundlagen_Version-2.1_2024-11-27.pdf.
9. Deutsches Krebsforschungszentrum. Diagnose Lungenkrebs: Untersuchungen bei Krebsverdacht [online]. 2022 [Zugriff: 10.04.2025]. URL: <https://www.krebsinformationsdienst.de/lungenkrebs/diagnostik>.

10. Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz. Bekanntmachung der Begründung zur Lungenkrebs-Früherkennungs-Verordnung [online]. 2024 [Zugriff: 15.05.2025]. URL: <https://www.bundesanzeiger.de/pub/publication/RBUh7MbZCkbNPIUpceO/content/RBUh7MbZCkbNPIUpceO/BAnz%20AT%2027.06.2024%20B6.pdf?inline>.
11. Gemeinsamer Bundesausschuss. Bewertung der Lungenkrebsfrüherkennung mittels Niedrigdosis-Computertomographie bei Rauchern (§ 135 SGB V) [online]. 2024 [Zugriff: 05.01.2026]. URL: <https://www.g-ba.de/bewertungsverfahren/methodenbewertung/321/>.
12. Dorfman DD, Berbaum KS, Metz CE. Receiver operating characteristic rating analysis. Generalization to the population of readers and patients with the jackknife method. Invest Radiol 1992; 27(9): 723-731.
13. Whiting PF, Rutjes AW, Westwood ME et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155(8): 529-536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>.
14. Elvidge J, Hawksworth C, Avşar TS et al. Consolidated Health Economic Evaluation Reporting Standards for Interventions That Use Artificial Intelligence (CHEERS-AI). Value Health 2024; 27(9): 1196-1205. <https://doi.org/10.1016/j.jval.2024.05.006>.
15. European Network for Health Technology Assessment. EUnetHTA HTA Adaptation Toolkit & Glossary [online]. 2011 [Zugriff: 29.10.2024]. URL: https://www.eunetha.eu/wp-content/uploads/2011/01/EUnetHTA_adptation_toolkit_2011_version_5.pdf.
16. Beauchamp T, Childress J. Principles of Biomedical Ethics: Marking Its Fortieth Anniversary. Am J Bioeth 2019; 19(11): 9-12. <https://doi.org/10.1080/15265161.2019.1665402>.
17. Hofmann B, Droste S, Oortwijn W et al. Harmonization of ethics in health technology assessment: a revision of the Socratic approach. Int J Technol Assess Health Care 2014; 30(1): 3-9. <https://doi.org/10.1017/S0266462313000688>.
18. Marckmann G, Schmidt H, Sofaer N et al. Putting public health ethics into practice: a systematic framework. Front Public Health 2015; 3: 23. <https://doi.org/10.3389/fpubh.2015.00023>.
19. Lysdahl KB, Mozygemba K, Burns J et al. Guidance for assessing effectiveness, economic aspects, ethical aspects, socio-cultural aspects and legal aspects in complex technologies [online]. 2016 [Zugriff: 29.10.2024]. URL: <https://validatehta.eu/wp-content/uploads/2022/07/3.-Guidance-for-assessing-effectiveness-economic-aspects-ethical-aspects-socio-cultural-aspects-and-legal-aspects-in-complex-technologies.pdf>.

20. Mozygemba K, Hofmann B, Lysdahl KB et al. Guidance to assess socio-cultural aspects [online]. 2016 [Zugriff: 29.10.2024]. URL: <https://validatehta.eu/wp-content/uploads/2022/07/3.-Guidance-for-assessing-effectiveness-economic-aspects-ethical-aspects-socio-cultural-aspects-and-legal-aspects-in-complex-technologies.pdf>.
21. Brönneke JB, Hofmann B, Bakke Lysdal K et al. Guidance to assess legal aspects [online]. 2016 [Zugriff: 29.10.2024]. URL: <https://validatehta.eu/wp-content/uploads/2022/07/3.-Guidance-for-assessing-effectiveness-economic-aspects-ethical-aspects-socio-cultural-aspects-and-legal-aspects-in-complex-technologies.pdf>.
22. Perleth M, Gibis B, Velasco Garrido M et al. Organisationsstrukturen und Qualität. In: Perleth M, Busse R, Gerhardus A et al (Ed). Health Technology Assessment; Konzepte, Methoden, Praxis für Wissenschaft und Entscheidungsfindung. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2014. S. 265-279.
23. Lee M, Hwang EJ, Lee JH et al. Artificial Intelligence for Low-Dose CT Lung Cancer Screening: Comparison of Utilization Scenarios. AJR 2025. <https://doi.org/10.2214/ajr.25.32829>.
24. Lo SB, Freedman MT, Gillis LB et al. JOURNAL CLUB: Computer-Aided Detection of Lung Nodules on CT With a Computerized Pulmonary Vessel Suppressed Function. AJR Am J Roentgenol 2018; 210(3): 480–488. <https://doi.org/10.2214/ajr.17.18718>.
25. Benbassat J. Duration of lead time in screening for lung cancer. BMC Pulm Med 2021; 21(1): 4. <https://doi.org/10.1186/s12890-020-01385-3>.
26. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Erstellung einer Versicherteninformation zum Lungenkrebs-Screening mittels Niedrigdosis-Computertomografie; Rapid Report [online]. 2025. URL: <https://doi.org/10.60584/P24-02>.
27. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Lungenkrebscreening mittels Niedrigdosis Computertomografie; Abschlussbericht [online]. 2020 [Zugriff: 28.10.2025]. URL: https://www.iqwig.de/download/s19-02_lungenkrebscreening-mittels-low-dose-ct_abschlussbericht_v1-0.pdf.
28. Pitman A, Cowan IA, Floyd RA et al. Measuring radiologist workload: Progressing from RVUs to study ascribable times. J Med Imaging Radiat Oncol 2018; 62(5): 605-618. <https://doi.org/10.1111/1754-9485.12778>.
29. Kassenärztliche Bundesvereinigung. Der Einheitliche Bewertungsmaßstab; EBM 2025/Q4 [online]. 2025 [Zugriff: 30.10.2025]. URL: <https://ebm.kbv.de/>.
30. Wigge P. Abrechnung von KI-Leistungen in der Radiologie [online]. 2025 [Zugriff: 28.11.2025]. URL: <https://radiologiemagazin.de/it-ki/abrechnung-von-ki-leistungen-in-der-radiologie.html>.

31. Wigge P, Jentsch K. Der Einsatz von KI in der Radiologie – KI-Leistungen bei der Befundung als Bestandteil der vertragsärztlichen Versorgung? *Rofo* 2025; 197(8): 980-986. <https://doi.org/10.1055/a-2619-7824>.
32. National Institute for Health and Care Excellence. Artificial intelligence for analysing chest CT images [online]. 2021 [Zugriff: 07.01.2026]. URL: <https://www.nice.org.uk/advice/mib243>.
33. National Institute for Health and Care Research, National Institute for Health and Care Excellence. Artificial intelligence software for analysing chest X-ray images to identify suspected lung cancer; early value assessment report [online]. 2023 [Zugriff: 30.10.2025]. URL: <https://www.nice.org.uk/guidance/htg696/documents/diagnostics-assessment-report>.
34. Geppert J, Auguste P, Asgharzadeh A et al. Software with artificial intelligence-derived algorithms for detecting and analysing lung nodules in CT scans: systematic review and economic evaluation. *Health Technol Assess* 2025; 29(14): 1-234. <https://doi.org/10.3310/jytw8921>.
35. National Institute for Health and Care Excellence. Software with artificial intelligence derived algorithms for automated detection and analysis of lung nodules from CT scan images; Diagnostics Assessment Programme; Final scope [online]. 2021 [Zugriff: 24.09.2025]. URL: <https://www.nice.org.uk/guidance/htg687/documents/final-scope>.
36. Campbell & Cochrane Economics Methods Group, Evidence for Policy & Practice Centre. CCEMG - EPPI-Centre Cost Converter [online]. 2024 [Zugriff: 24.11.2025]. URL: <https://eppi.ioe.ac.uk/costconversion/>.
37. Statistisches Bundesamt. Verbraucherpreisindex: Deutschland, Jahre [online]. 2025 [Zugriff: 01.12.2025]. URL: <https://www-genesis.destatis.de/datenbank/online/statistic/61111/table/61111-0001>.
38. Tadavarthi Y, Vey B, Krupinski E et al. The State of Radiology AI: Considerations for Purchase Decisions and Current Market Offerings. *Radiology: Artificial Intelligence* 2020; 2(6): e200004. <https://doi.org/10.1148/ryai.2020200004>.
39. Du Y, Greuter MJW, Prokop MW et al. Pricing and cost-saving potential for deep-learning computer-aided lung nodule detection software in CT lung cancer screening. *Insights Into Imaging* 2023; 14(1): 208. <https://doi.org/10.1186/s13244-023-01561-z>.
40. Geppert J, Auguste P, Asgharzadeh A et al. Software with artificial intelligence derived algorithms for automated detection and analysis of lung nodules from CT scan images [DAP060]; Technology Assessment Report commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Care Excellence – Diagnostics Assessment Report [online]. 2022 [Zugriff: 25.09.2025]. URL: <https://www.nice.org.uk/guidance/htg687/documents/diagnostics-assessment-report>.

41. Geppert J, Auguste P, Asgharzadeh A et al. Software with artificial intelligence derived algorithms for automated detection and analysis of lung nodules from CT scan images [DAP60]; Diagnostic Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence – Final protocol [online]. 2021 [Zugriff: 24.09.2025]. URL: <https://www.nice.org.uk/guidance/htg687/documents/final-protocol>.
42. National Institute for Health and Care Excellence. AI-derived computer-aided detection (CAD) software for detecting and measuring lung nodules in CT scan images; NICE diagnostics guidance 55 [online]. 2023 [Zugriff: 18.11.2025]. URL: <https://www.nice.org.uk/guidance/htg687/chapter/1-Recommendations>.
43. Hsu HH, Ko KH, Chou YC et al. Performance and reading time of lung nodule identification on multidetector CT with or without an artificial intelligence-powered computer-aided detection system. Clin Radiol 2021; 76(8): 626.e23–626.e32. <https://doi.org/10.1016/j.crad.2021.04.006>.
44. Kozuka T, Matsukubo Y, Kadoba T et al. Efficiency of a computer-aided diagnosis (CAD) system with deep learning in detection of pulmonary nodules on 1-mm-thick images of computed tomography. Japanese Journal of Radiology 2020; 38(11): 1052-1061. <https://doi.org/10.1007/s11604-020-01009-0>.
45. National Health Services England. Lung cancer screening [online]. 2023 [Zugriff: 24.11.2025]. URL: <https://www.nhs.uk/tests-and-treatments/lung-cancer-screening/>.
46. Busch F, Hoffmann L, Xu L et al. Multinational Attitudes Toward AI in Health Care and Diagnostics Among Hospital Patients. JAMA Netw Open 2025; 8(6): e2514452. <https://doi.org/10.1001/jamanetworkopen.2025.14452>.
47. Deutscher Ethikrat. Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz; Stellungnahme [online]. 2023 [Zugriff: 07.01.2026]. URL: <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>.
48. Farič N, Hinder S, Williams R et al. Early experiences of integrating an artificial intelligence-based diagnostic decision support system into radiology settings: a qualitative study. J Am Med Inform Assoc 2023; 31(1): 24-34. <https://doi.org/10.1093/jamia/ocad191>.
49. National Institute for Health and Care Excellence. Artificial intelligence (AI)-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals: early value assessment; Recommendations [online]. 2025 [Zugriff: 18.11.2025]. URL: <https://www.nice.org.uk/guidance/htg696/chapter/1-Recommendations>.

50. Ewals LJS, van der Wulp K, van den Borne B et al. The Effects of Artificial Intelligence Assistance on the Radiologists' Assessment of Lung Nodules on CT Scans: A Systematic Review. *J Clin Med* 2023; 12(10). <https://doi.org/10.3390/jcm12103536>.
51. Madanay F, O'Donohue LS, Zikmund-Fisher BJ. Patient Reactions to Artificial Intelligence-Clinician Discrepancies: Web-Based Randomized Experiment. *J Med Internet Res* 2025; 27: e68823. <https://doi.org/10.2196/68823>.
52. Zentrale Kommission zur Wahrung ethischer Grundsätze in der Medizin und ihren Grenzgebieten (Zentrale Ethikkommission) bei der Bundesärztekammer. Entscheidungsunterstützung ärztlicher Tätigkeit durch Künstliche Intelligenz. *Dtsch Arztebl* 2021; 118: A1-A13.
53. Alderman JE, Palmer J, Laws E et al. Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together consensus recommendations. *Lancet Digit Health* 2025; 7(1): e64-e88. [https://doi.org/10.1016/s2589-7500\(24\)00224-3](https://doi.org/10.1016/s2589-7500(24)00224-3).
54. Bundesärztekammer. Künstliche Intelligenz in der Gesundheitsversorgung [online]. 2025 [Zugriff: 07.01.2026]. URL: https://www.bundesaerztekammer.de/fileadmin/user_upload/BAEK/Themen/Digitalisierung/Thesenpapier_KI_in_der_Gesundheitsversorgung_03.2025.pdf.
55. Katzenmeier C. Haftung für Schäden durch KI in der Medizin. In: Grundmann T, Hey J, Hobe S et al (Ed). *Die Macht der Algorithmen*. Baden-Baden: Nomos Verlagsgesellschaft; 2023. S. 73-102.
56. Sorin V, Soffer S, Glicksberg BS et al. Adversarial attacks in radiology - A systematic review. *Eur J Radiol* 2023; 167: 111085. <https://doi.org/10.1016/j.ejrad.2023.111085>.
57. Bowers PJT, Kirk FM. Insights into artificial intelligence and our intelligence-on the frontier of lung cancer screening. *J Thorac Dis* 2024; 16(11): 7905-7909. <https://doi.org/10.21037/jtd-24-1077>.
58. Ewals LJS, Heesterbeek LJJ, Yu B et al. The Impact of Expectation Management and Model Transparency on Radiologists' Trust and Utilization of AI Recommendations for Lung Nodule Assessment on Computed Tomography: Simulated Use Study. *JMIR AI* 2024; 3: e52211. <https://doi.org/10.2196/52211>.
59. Thong LT, Chou HS, Chew HSJ et al. Diagnostic test accuracy of artificial intelligence-based imaging for lung cancer screening: A systematic review and meta-analysis. *Lung Cancer* 2023; 176: 4-13. <https://doi.org/10.1016/j.lungcan.2022.12.002>.
60. Wang TW, Hong JS, Chiu HY et al. Standalone deep learning versus experts for diagnosis lung cancer on chest computed tomography: a systematic review. *Eur Radiol* 2024; 34(11): 7397-7407. <https://doi.org/10.1007/s00330-024-10804-6>.

61. Crowson MG, Moukheiber D, Arévalo AR et al. A systematic review of federated learning applications for biomedical data. PLOS Digit Health 2022; 1(5): e0000033. <https://doi.org/10.1371/journal.pdig.0000033>.
62. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Lungenkrebscreening mittels Niedrigdosis-Computertomografie; Rapid Report [online]. 2024 [Zugriff: 06.01.2026]. URL: https://doi.org/10.60584/S23-02_V1.1.
63. Bundesamt für Strahlenschutz. Lungenkrebsfrüherkennung mittels Niedrigdosis-Computertomographie; Wissenschaftliche Bewertung des Bundesamtes für Strahlenschutz gemäß § 84 Absatz 3 Strahlenschutzgesetz; Bericht [online]. 2023 [Zugriff: 22.10.2025]. URL: <https://doris.bfs.de/jspui/bitstream/urn:nbn:de:0221-2021082028027/5/BfS-35-21.pdf>.
64. Jonas DE, Reuland DS, Reddy SM et al. Screening for Lung Cancer With Low-Dose Computed Tomography: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA 2021; 325(10): 971-987. <https://doi.org/10.1001/jama.2021.0377>.
65. Adams SJ, Stone E, Baldwin DR et al. Lung cancer screening. Lancet 2023; 401(10374): 390-408. [https://doi.org/10.1016/s0140-6736\(22\)01694-4](https://doi.org/10.1016/s0140-6736(22)01694-4).
66. Salman M, Cotton A, Humphrey A et al. Patient barriers and facilitators to lung cancer screening uptake and intention to screen: a systematic review using the Theoretical Domains Framework. BMJ Open Respir Res 2025; 12(1). <https://doi.org/10.1136/bmjresp-2024-003127>.
67. Warner ET, Lathan CS. Race and sex differences in patient provider communication and awareness of lung cancer screening in the health information National Trends Survey, 2013–2017. Prev Med 2019; 124: 84-90. <https://doi.org/10.1016/j.ypmed.2019.05.001>.
68. Warner ET, Revette A, Restrepo E et al. Women's Information Needs and Educational Preferences Regarding Lung Cancer Screening. J Womens Health (Larchmt) 2024; 33(3): 318-327. <https://doi.org/10.1089/jwh.2023.0429>.
69. Ali N, Lifford KJ, Carter B et al. Barriers to uptake among high-risk individuals declining participation in lung cancer screening: a mixed methods analysis of the UK Lung Cancer Screening (UKLS) trial. BMJ Open 2015; 5(7): e008254. <https://doi.org/10.1136/bmjopen-2015-008254>.
70. Becker N, Motsch E, Trotter A et al. Lung cancer mortality reduction by LDCT screening- Results from the randomized German LUSI trial. Int J Cancer 2020; 146(6): 1503-1513. <https://doi.org/10.1002/ijc.32486>.
71. van der Aalst CM, Ten Haaf K, de Koning HJ. Implementation of lung cancer screening: what are the main issues? Transl Lung Cancer Res 2021; 10(2): 1050-1063. <https://doi.org/10.21037/tlcr-20-985>.

72. Fritsch SJ, Blankenheim A, Wahl A et al. Attitudes and perception of artificial intelligence in healthcare: A cross-sectional survey among patients. *Digit Health* 2022; 8: 20552076221116772. <https://doi.org/10.1177/20552076221116772>.
73. Akudjedu TN, Torre S, Khine R et al. Knowledge, perceptions, and expectations of Artificial intelligence in radiography practice: A global radiography workforce survey. *J Med Imaging Radiat Sci* 2023; 54(1): 104-116. <https://doi.org/10.1016/j.jmir.2022.11.016>.
74. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Krebsfrüherkennungs-Richtlinie (KFE-RL): Einführung der Lungenkrebsfrüherkennung mittels Niedrigdosis-Computertomographie bei Rauchern [online]. 2024 [Zugriff: 23.10.2025]. URL: https://www.g-ba.de/downloads/39-261-7301/2025-06-18_KFE-RL_Einf-Lungenkrebsfrueherkennung-Niedrigdosis-CT-Raucher_BAnz.pdf.
75. Zentrum für Krebsregisterdaten. Lungenkrebs (Bronchialkarzinom); ICD-10 C33–C34 [online]. 2024 [Zugriff: 28.10.2025]. URL: <https://www.krebsdaten.de/Krebs/DE/Content/Krebsarten/Lungenkrebs/lungenkrebs.html>.
76. Starker A, Schienkiewitz A, Damerow S et al. Verbreitung von Adipositas und Rauchen bei Erwachsenen in Deutschland – Entwicklung von 2003 bis 2023. *J Health Monit* 2025; 10(1): e12990.
77. Potter AL, Senthil P, Srinivasan D et al. Persistent race- and sex-based disparities in lung cancer screening eligibility. *J Thorac Cardiovasc Surg* 2024; 168(1): 248-260.e2. <https://doi.org/10.1016/j.jtcvs.2023.10.025>.
78. Potter AL, Xu NN, Senthil P et al. Pack-Year Smoking History: An Inadequate and Biased Measure to Determine Lung Cancer Screening Eligibility. *J Clin Oncol* 2024; 42(17): 2026-2037. <https://doi.org/10.1200/JCO.23.01780>.
79. Wilder FG, Cangut B, Jindani R et al. Lung cancer screening among minority groups: Identifying gaps in screening and opportunities for intervention. *JTCVS Open* 2024; 21: 341-348. <https://doi.org/10.1016/j.xjon.2024.07.001>.
80. Michalski N, Reis M, Tetzlaff F et al. German Index of Socioeconomic Deprivation (GISD): Revision, Aktualisierung und Anwendungsbeispiele. *Journal of Health Monitoring* 2022; 7. <https://doi.org/10.25646/10640>.
81. Pizzato M, Martinsen JI, Heikkinen S et al. Socioeconomic status and risk of lung cancer by histological subtype in the Nordic countries. *Cancer Med* 2022; 11(8): 1850-1859. <https://doi.org/10.1002/cam4.4548>.
82. Hovanec J, Kendzia B, Olsson A et al. Socioeconomic Status, Smoking, and Lung Cancer: Mediation and Bias Analysis in the SYNERGY Study. *Epidemiology* 2025; 36(2): 245-252. <https://doi.org/10.1097/ede.0000000000001807>.

83. Europäisches Parlament, Rat der Europäischen Union. Verordnung (EU) 2025/327 des Europäischen Parlaments und des Rates vom 11. Februar 2025 über den europäischen Gesundheitsdatenraum sowie zur Änderung der Richtlinie 2011/24/EU und der Verordnung (EU) 2024/2847 [online]. 2025 [Zugriff: 13.11.2025]. URL: <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32025R0327>.
84. Kocak B, Ponsiglione A, Romeo V et al. Radiology AI and sustainability paradox: environmental, economic, and social dimensions. *Insights Imaging* 2025; 16(1): 88. <https://doi.org/10.1186/s13244-025-01962-2>.
85. Shah SK, Krishnan V, Khan AA et al. Women are Underrepresented in Non-small Cell Lung Cancer Clinical Trials: A Systematic Review. *Ann Surg Oncol* 2024; 31(10): 6673-6679. <https://doi.org/10.1245/s10434-024-15720-z>.
86. Kauczor HU, von Stackelberg O, Nischwitz E et al. Strengthening lung cancer screening in Europe: fostering participation, improving outcomes, and addressing health inequalities through collaborative initiatives in the SOLACE consortium. *Insights Imaging* 2024; 15(1): 252. <https://doi.org/10.1186/s13244-024-01814-5>.
87. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Entscheidungshilfe: Früherkennung von Lungenkrebs bei starken Raucherinnen und Rauchern: Vor- und Nachteile der Niedrigdosis-Computer-Tomografie [online]. 2024 [Zugriff: 13.01.2026]. URL: <https://www.gesundheitsinformation.de/entscheidungshilfe-frueherkennung-von-lungenkrebs-bei-starken-raucherinnen-und-rauchern-vor-und-nachteile-der-niedrigdosis-computer-tomografie.html>.
88. Gleeson F, Revel MP, Biederer J et al. Implementation of artificial intelligence in thoracic imaging-a what, how, and why guide from the European Society of Thoracic Imaging (ESTI). *Eur Radiol* 2023; 33(7): 5077-5086. <https://doi.org/10.1007/s00330-023-09409-2>.
89. Wigge P. Informations- und Aufklärungspflichten des Radiologen vor Durchführung einer einzelnen Röntgenaufnahme. *Rofo* 2021; 193(05): 607-612. <https://doi.org/10.1055/a-1404-0411>.
90. Eichelberger J. Arzthaftung im Kontext von KI und Robotik. In: Ruschemeier H, Steinrötter B (Ed). *Der Einsatz von KI & Robotik in der Medizin; Interdisziplinäre Fragen*. Baden-Baden: Nomos Verlagsgesellschaft; 2024. S. 27-52.
91. Europäisches Parlament, Rat der Europäischen Union. Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung) [online]. 2016 [Zugriff: 07.01.2026]. URL: <http://data.europa.eu/eli/reg/2016/679/oj>.

92. Bundesministerium der Justiz und für Verbraucherschutz. Bundesdatenschutzgesetz vom 30. Juni 2017 (BGBl. I S. 2097), das zuletzt durch Artikel 7 des Gesetzes vom 6. Mai 2024 (BGBl. 2024 I Nr. 149) geändert worden ist [online]. 2024 [Zugriff: 08.01.2026]. URL: https://www.gesetze-im-internet.de/bdsg_2018/BDSG.pdf.
93. Bundesministerium der Justiz und für Verbraucherschutz. Gesetz zur verbesserten Nutzung von Gesundheitsdaten [online]. 2024 [Zugriff: 13.11.2025]. URL: <https://www.recht.bund.de/bgbl/1/2024/102/VO.html>.
94. Bundesministerium für Gesundheit. Gesetz zur Beschleunigung der Digitalisierung des Gesundheitswesens (Digital-Gesetz – DigiG) [online]. 2025 [Zugriff: 13.11.2025]. URL: <https://www.recht.bund.de/bgbl/1/2024/101/VO.html>.
95. Europäisches Parlament, Rat der Europäischen Union. Verordnung (EU) 2017/745 des Europäischen Parlaments und des Rates vom 5. April 2017 über Medizinprodukte, zur Änderung der Richtlinie 2001/83/EG, der Verordnung (EG) Nr. 178/2002 und der Verordnung (EG) Nr. 1223/2009 und zur Aufhebung der Richtlinien 90/385/EWG und 93/42/EWG des Rates [online]. 2017 [Zugriff: 07.01.2026]. URL: <http://data.europa.eu/eli/reg/2017/745/oj>.
96. Medical Device Coordination Group. Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR [online]. 2025 [Zugriff: 10.12.2025]. URL: https://health.ec.europa.eu/document/download/b45335c5-1679-4c71-a91c-fc7a4d37f12b_en?filename=mdcg_2019_11_en.pdf.
97. Bundesministerium der Justiz und für Verbraucherschutz. Das Fünfte Buch Sozialgesetzbuch – Gesetzliche Krankenversicherung – (Artikel 1 des Gesetzes vom 20. Dezember 1988, BGBl. I S. 2477, 2482), das zuletzt durch Artikel 3a des Gesetzes vom 22. Dezember 2025 (BGBl. 2025 I Nr. 371) geändert worden ist [online]. 2025 [Zugriff: 23.12.2025]. URL: https://www.gesetze-im-internet.de/sgb_5/SGB_5.pdf.
98. Botta J. (K)ein Recht auf Behandlung mit KI? Der Zugang zu intelligenten Medizinprodukten im Lichte des Medizin-, Sozial- und Verfassungsrechts. In: Ruschemeier H, Steinrötter B (Ed). Der Einsatz von KI & Robotik in der Medizin; Interdisziplinäre Fragen. Baden-Baden: Nomos Verlagsgesellschaft; 2024. S. 115-134.
99. Bundesärztekammer. Stellungnahme „Künstliche Intelligenz in der Medizin“. Dtsch Arztebl 2025; 122(4): 1-14.
https://doi.org/10.3238/arztebl.2025.Stellungnahme_KI_Medizin.
100. Bundesministerium der Justiz und für Verbraucherschutz. Produkthaftungsgesetz vom 15. Dezember 1989 (BGBl. I S. 2198), das zuletzt durch Artikel 5 des Gesetzes vom 17. Juli 2017 (BGBl. I S. 2421) geändert worden ist [online]. 2017 [Zugriff: 08.01.2026]. URL: <https://www.gesetze-im-internet.de/prodhaftg/ProdHaftG.pdf>.

101. Europäisches Parlament, Rat der Europäischen Union. Richtlinie (EU) 2024/2853 des Europäischen Parlaments und des Rates vom 23. Oktober 2024 über die Haftung für fehlerhafte Produkte und zur Aufhebung der Richtlinie 85/374/EWG des Rates [online]. 2024 [Zugriff: 08.01.2025]. URL: <https://eur-lex.europa.eu/eli/dir/2024/2853/oj>.
102. Bundesministerium der Justiz und für Verbraucherschutz. Entwurf eines Gesetzes zur Modernisierung des Produkthaftungsrechts; Referentenentwurf [online]. 2025 [Zugriff: 10.12.2025]. URL: https://bundestagszusammenfasser.de/wp-content/uploads/rewp21/1052_RefE_Produkthaftung.pdf.
103. Gemeinsamer Bundesausschuss. Lungenkrebs-Früherkennung bei starken Raucherinnen und Rauchern wird voraussichtlich ab April 2026 Kassenleistung [online]. 2025 [Zugriff: 23.10.2025]. URL: <https://www.g-ba.de/presse/pressemitteilungen-meldungen/1263/>.
104. Deutsche Krebsgesellschaft. Kennzahlenauswertung 2025; Jahresbericht der zertifizierten Lungenkrebszentren; Auditjahr 2024 / Kennzahlenjahr 2023 [online]. 2024 [Zugriff: 23.10.2025]. URL: https://www.krebsgesellschaft.de/unsere-themen/zertifizierung/jahresberichte-zertifizierter-zentren?file=files/content/unsere-themen/zertifizierung/jahresberichte/lungenkrebszentren/qualitaetsindikatoren_lungenkrebs_2025-a1_250804.pdf
105. Blum TG, Vogel-Claussen J, Andreas S et al. [Statement paper on the implementation of a national organized program in Germany for the early detection of lung cancer in risk populations using low-dose CT screening including management of screening findings]. Pneumologie 2024; 78(1): 15-34. <https://doi.org/10.1055/a-2175-4580>.
106. GKV-Spitzenverband. Verbesserung und Stärkung der ambulanten Versorgung; Positionspapier [online]. 2024 [Zugriff: 23.10.2025]. URL: https://www.gkv-spitzenverband.de/media/dokumente/service_1/publikationen/20240715_PoPa_Ambulante_Versorgung_barrierefrei.pdf.
107. Haserück A. Blickwinkel: Künstliche Intelligenz und Radiologie. Dtsch Arztebl International 2024; 121(21): 8626.
108. Shaham D, Kazerooni EA. Quality, Standards, and Optimal Training of Radiologists for Lung Cancer Screening. J Thorac Imaging 2025. <https://doi.org/10.1097/rti.0000000000000844>.
109. Tejani AS, Klontzas ME, Gatti AA et al. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update. Radiol Artif Intell 2024; 6(4): e240300. <https://doi.org/10.1148/ryai.240300>.
110. Fasterholdt I, Kjølhede T, Naghavi-Behzad M et al. Model for ASsessing the value of Artificial Intelligence in medical imaging (MAS-AI). Int J Technol Assess Health Care 2022; 38(1): e74. <https://doi.org/10.1017/s0266462322000551>.

111. Hahn HK, May MS, Dicken V et al. Anforderungen an die Qualitätssicherung von KI-Modellen für die Lungenkrebs-Früherkennung [online]. 2025 [Zugriff: 07.01.2026]. URL: <https://doi.org/10.48550/arXiv.2502.17639>.
112. Hadjiiski L, Cha K, Chan HP et al. AAPM task group report 273: Recommendations on best practices for AI and machine learning for computer-aided diagnosis in medical imaging. Med Phys 2023; 50(2): e1-e24. <https://doi.org/10.1002/mp.16188>.
113. van Leeuwen KG, de Rooij M, Schalekamp S et al. Clinical use of artificial intelligence products for radiology in the Netherlands between 2020 and 2022. Eur Radiol 2024; 34(1): 348-354. <https://doi.org/10.1007/s00330-023-09991-5>.
114. Jungmann F. Künstliche Intelligenz in der Radiologie. Akademie Aktuell 2024; 3: 26-31.
115. van Leeuwen KG, Schalekamp S, Rutten M et al. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. Eur Radiol 2021; 31(6): 3797-3804. <https://doi.org/10.1007/s00330-021-07892-z>.
116. Multusch MM, Hansen L, Heinrich MP et al. Impact of Radiologist Experience on AI Annotation Quality in Chest Radiographs: A Comparative Analysis. Diagnostics (Basel) 2025; 15(6). <https://doi.org/10.3390/diagnostics15060777>.
117. Vogel-Claussen J, Lasch F, Bollmann BA et al. Design and Rationale of the HANSE Study: A Holistic German Lung Cancer Screening Trial Using Low-Dose Computed Tomography. Rofo 2022; 194(12): 1333-1345. <https://doi.org/10.1055/a-1853-8291>.
118. Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Krebsfrüherkennungs-Richtlinie (KFE-RL); Einführung der Lungenkrebsfrüherkennung mittels Niedrigdosis-Computertomographie bei Rauchern [online]. 2024 [Zugriff: 10.12.2025]. URL: https://www.g-ba.de/downloads/40-268-11634/2025-06-18_KFE-RL_Einf-Lungenkrebsfrueherkennung-Niedrigdosis-CT-Raucher_TrG.pdf.
119. Strubell E, Ganesh A, McCallum A. Energy and Policy Considerations for Deep Learning in NLP [online]. 2019 [Zugriff: 07.01.2026]. URL: <https://doi.org/10.48550/arXiv.1906.02243>.
120. Karliner J, Slotterback S, Boyd R et al. Health care's climate footprint: the health sector contribution and opportunities for action. Eur J Public Health 2020; 30(Suppl 5). <https://doi.org/10.1093/eurpub/ckaa165.843>.
121. Doo FX, Vosschenrich J, Cook TS et al. Environmental Sustainability and AI in Radiology: A Double-Edged Sword. Radiology 2024; 310(2): e232030. <https://doi.org/10.1148/radiol.232030>.
122. Roletto A, Zanardo M, Bonfitto GR et al. The environmental impact of energy consumption and carbon emissions in radiology departments: a systematic review. Eur Radiol Exp 2024; 8(1): 35. <https://doi.org/10.1186/s41747-024-00424-6>.

123. Ueda D, Walston SL, Fujita S et al. Climate change and artificial intelligence in healthcare: Review and recommendations towards a sustainable future. *Diagnostic and Interventional Imaging* 2024; 105(11): 453-459. <https://doi.org/10.1016/j.diii.2024.06.002>.
124. Joshi G, Jain A, Araveeti SR et al. FDA-Approved Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices: An Updated Landscape. *Electronics* 2024; 13(3): 498. <https://doi.org/10.3390/electronics13030498>.
125. Muehlematter UJ, Bluethgen C, Vokinger KN. FDA-cleared artificial intelligence and machine learning-based medical devices and their 510(k) predicate networks. *Lancet Digit Health* 2023; 5(9): e618-e626. [https://doi.org/10.1016/S2589-7500\(23\)00126-7](https://doi.org/10.1016/S2589-7500(23)00126-7).
126. Thangam D, Muniraju H, Ramesh R et al. Impact of Data Centers on Power Consumption, Climate Change, and Sustainability. In: Kumar KD, Varadarajan V, Nasser N et al (Ed). *Computational Intelligence for Green Cloud Computing and Digital Waste Management*. Hershey: IGI Global Scientific Publishing; 2024. S. 60-83.
127. Koronen C, Åhman M, Nilsson LJ. Data centres in future European energy systems; energy efficiency, integration and policy. *Energy Efficiency* 2020; 13(1): 129-144. <https://doi.org/10.1007/s12053-019-09833-8>.
128. Ren S. How much water does AI consume? The public deserves to know [online]. 2023 [Zugriff: 30.10.2025]. URL: <https://oecd.ai/en/work/how-much-water-does-ai-consume>.
129. Wahlster P, Brereton L, Burns J et al. Guidance on the integrated assessment of complex health technologies – The INTEGRATE-HTA Model [online]. 2016 [Zugriff: 29.10.2024]. URL: <https://validatehta.eu/wp-content/uploads/2022/07/2.-Guidance-on-the-integrated-assessment-of-complex-health-technologies-the-INTEGRATE-HTA-model.pdf>.
130. Becker N, Motsch E, Gross ML et al. Randomized Study on Early Detection of Lung Cancer with MSCT in Germany: Results of the First 3 Years of Follow-up After Randomization. *J Thorac Oncol* 2015; 10(6): 890-896. <https://doi.org/10.1097/jto.0000000000000530>.
131. Vogel-Claussen J, Bollmann BA, May K et al. Effectiveness of NELSON versus PLCOm2012 lung cancer screening eligibility criteria in Germany (HANSE): a prospective cohort study. *Lancet Oncol* 2025; 26(12): 1541-1551. [https://doi.org/10.1016/s1470-2045\(25\)00490-5](https://doi.org/10.1016/s1470-2045(25)00490-5).
132. van Leeuwen KG, Schalekamp S, Rutten M et al. Comparison of Commercial AI Software Performance for Radiograph Lung Nodule Detection and Bone Age Prediction. *Radiology* 2024; 310(1): e230981. <https://doi.org/10.1148/radiol.230981>.
133. RayMedix. Produktmerkmale von Annalise Enterprise CXR [online]. [Zugriff: 30.12.2025]. URL: <https://www.raymedix.de/produkte/softwareloesungen/annaliseai.html>.

134. Kim DW, Jang HY, Kim KW et al. Design Characteristics of Studies Reporting the Performance of Artificial Intelligence Algorithms for Diagnostic Analysis of Medical Images: Results from Recently Published Papers. *Korean J Radiol* 2019; 20(3): 405-410. <https://doi.org/10.3348/kjr.2019.0025>.
135. Walstra ANH, Lancaster HL, Heuvelmans MA et al. Feasibility of AI as first reader in the 4-IN-THE-LUNG-RUN lung cancer screening trial: impact on negative-misclassifications and clinical referral rate. *Eur J Cancer* 2025; 216: 115214. <https://doi.org/10.1016/j.ejca.2024.115214>.
136. Bhatia BS, Morlese JF, Yusuf S et al. A real-world evaluation of the diagnostic accuracy of radiologists using positive predictive values verified from deep learning and natural language processing chest algorithms deployed retrospectively. *BJR Open* 2024; 6(1): tzad009. <https://doi.org/10.1093/bjro/tzad009>.
137. Colquitt J, Jordan M, Court R et al. Artificial intelligence software for analysing chest X-ray images to identify suspected lung cancer: an evidence synthesis early value assessment. *Health Technol Assess* 2024; 28(50): 1-75. <https://doi.org/10.3310/lkrt4721>.
138. Geppert J, Asgharzadeh A, Brown A et al. Software using artificial intelligence for nodule and cancer detection in CT lung cancer screening: systematic review of test accuracy studies. *Thorax* 2024; 79(11): 1040-1049. <https://doi.org/10.1136/thorax-2024-221662>.
139. Li D, Pehrson LM, Lauridsen CA et al. The Added Effect of Artificial Intelligence on Physicians' Performance in Detecting Thoracic Pathologies on CT and Chest X-ray: A Systematic Review. *Diagnostics (Basel)* 2021; 11(12). <https://doi.org/10.3390/diagnostics11122206>.
140. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. [X-ray image of the lung: Can artificial intelligence support the analysis of X-ray images?]; Original Title: Röntgenbild der Lunge: Kann Künstliche Intelligenz die Auswertung von Röntgenbildern unterstützen? [online]. [Zugriff: 06.01.2025]. URL: <https://database.inahta.org/article/30251>.
141. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Version 7.0 [online]. 2023 [Zugriff: 29.10.2024]. URL: https://www.iqwig.de/methoden/allgemeine-methoden_version-7-0.pdf.
142. 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>.
143. Liu X, Rivera SC, Moher D et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. *BMJ* 2020; 370: m3164. <https://doi.org/10.1136/bmj.m3164>.

144. Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004; 94(3): 361-366. <https://doi.org/10.2105/ajph.94.3.361>.
145. Bossuyt PM, Reitsma JB, Bruns DE et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med* 2003; 138(1): 40-44. <https://doi.org/10.7326/0003-4819-138-1-200301070-00010>.
146. 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>.
147. Pepe MS, Alonzo TA. Comparing disease screening tests when true disease status is ascertained only for screen positives. *Biostatistics* 2001; 2(3): 249-260. <https://doi.org/10.1093/biostatistics/2.3.249>.
148. 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>.
149. 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>.
150. Sutton AJ, Abrams KR, Jones DR et al. *Methods for Meta-Analysis in Medical Research*. Chichester: Wiley; 2000.
151. 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.
152. 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>.
153. 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>.
154. Leemis LM, Trivedi KS. A Comparison of Approximate Interval Estimators for the Bernoulli Parameter. *The American Statistician* 1996; 50(1): 63-68. <https://doi.org/10.1080/00031305.1996.10473544>.
155. Reitsma JB, Glas AS, Rutjes AW et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58(10): 982-990. <https://doi.org/10.1016/j.jclinepi.2005.02.022>.

156. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol* 2006; 59(12): 1331-1332. <https://doi.org/10.1016/j.jclinepi.2006.06.011>.
157. Menke J. Bivariate random-effects meta-analysis of sensitivity and specificity with SAS PROC GLIMMIX. *Methods Inf Med* 2010; 49(1): 54-64. <https://doi.org/10.3414/ME09-01-0001>.
158. Hotelling H. The generalization of student's ratio. *Ann Math Stat* 1931; 2(3): 360-378.
159. Drummond MF, Sculpher MJ, Claxton K et al. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press; 2015.
160. Virginia Polytechnic Institute, State University. Evaluation of Computer-Aided Lung Nodule Detection Software in Thoracic CT for Riverain Technologies LLC [online]. 2021 [Zugriff: 18.11.2025]. URL: <https://clinicaltrials.gov/study/NCT02440139>.
161. V5med. Clinical Evaluation of the V5-LU01 AI Software in Thoracic CT for V5med Inc [online]. 2024 [Zugriff: 18.11.2025]. URL: <https://clinicaltrials.gov/study/NCT06514248>.
162. Median Technologies. Retrospective Clinical Trial Comparing Radiologists' Diagnosis Accuracy in Lung Cancer Screening Population With and Without the Help of an AI/ML Tech-based SaMD (RELIVE) [online]. 2025 [Zugriff: 18.11.2025]. URL: <https://clinicaltrials.gov/study/NCT06751576>.
163. Infervision. Clinical Validation of InferRead Lung CT.A [online]. 2020 [Zugriff: 24.11.2025]. URL: <https://clinicaltrials.gov/study/NCT04119960>.
164. Median Technologies. Eyonis Lung Cancer Screening (LCS) meets primary endpoint in RELIVE clinical trial, the final pivotal study required for regulatory submissions [online]. 2025 [Zugriff: 24.10.2025]. URL: <https://mediantechnologies.com/investors/press-releases/eyonis-lung-cancer-screening-lcs-meets-primary-endpoint-in-relive-clinical-trial-the-final-pivotal-study-required-for-regulatory-submissions/>.
165. Lee J, Kim Y, Kim HY et al. Feasibility of implementing a national lung cancer screening program: Interim results from the Korean Lung Cancer Screening Project (K-LUCAS). *Transl Lung Cancer Res* 2021; 10(2): 723-736. <https://doi.org/10.21037/tlcr-20-700>.
166. Aberle DR, Adams AM, Berg CD et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011; 365(5): 395-409. <https://doi.org/10.1056/NEJMoa1102873>.
167. Verband der Privaten Krankenversicherungen. Gebührenordnung für Ärzte (GOÄ); mit verkürzten Leistungsbezeichnungen; Kurz-GOÄ [online]. 2013 [Zugriff: 12.11.2025]. URL: https://www.privat-patienten.de/fileadmin/Dateien/2_Beim_Arzt/PDF/GO%C3%84_Geb%C3%BChrenordnung_f%C3%BCr_%C3%84rzte.pdf.

168. Wigge P. KI-basierte Medizinprodukte in der Radiologie [online]. 2023 [Zugriff: 24.09.2025]. URL: <https://doi.org/10.13154/294-10190>.
169. Bajre MK, Pennington M, Woznitza N et al. Expanding the role of radiographers in reporting suspected lung cancer: A cost-effectiveness analysis using a decision tree model. *Radiography* 2017; 23(4): 273-278. <https://doi.org/10.1016/j.radi.2017.07.011>.
170. Naik H, Howell D, Su J et al. Stage specific health utility index scores of Canadian cancer patients. *J Clin Oncol* 2015; 33(15 Suppl): 6614. https://doi.org/10.1200/jco.2015.33.15_suppl.6614.
171. Szende A, Janssen B, Cabases J. Self-Reported Population Health; An International Perspective based on EQ-5D. Dordrecht: Springer; 2014.
172. National Health Services England. National Cost Collection for the NHS [online]. 2025 [Zugriff: 25.11.2025]. URL: <https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/>.
173. Jones K, Burns A. Costs of Health and Social Care 2021. Canterbury: Personal Social Services Research Unit; 2021.
174. Walsh JME, Karliner L, Smith A et al. LungCARE: Encouraging Shared Decision-Making in Lung Cancer Screening-a Randomized Trial. *J Gen Intern Med* 2023; 38(14): 3115-3122. <https://doi.org/10.1007/s11606-023-08189-1>.
175. Bundesministerium der Justiz und für Verbraucherschutz. Bürgerliches Gesetzbuch in der Fassung der Bekanntmachung vom 2. Januar 2002 (BGBl. I S. 42, 2909; 2003 I S. 738), das zuletzt durch Artikel 4 des Gesetzes vom 22. Dezember 2025 (BGBl. 2025 I Nr. 364) geändert worden ist" [online]. 2025 [Zugriff: 10.12.2025]. URL: <https://www.gesetze-im-internet.de/bgb/BGB.pdf>.
176. Herbst T. Herausforderungen an das Datenschutzrecht durch den Einsatz von Künstlicher Intelligenz in der Medizin. In: Ruschemeier H, Steinrötter B (Ed). *Der Einsatz von KI & Robotik in der Medizin; Interdisziplinäre Fragen*. Baden-Baden: Nomos Verlagsgesellschaft; 2024. S. 77-96.
177. Bundesinstitut für Arzneimittel und Medizinprodukte. Der EHDS erklärt; Eine Einführung in den Europäische Gesundheitsdatenraum (EHDS) [online]. [Zugriff: 13.11.2025]. URL: https://www.bfarm.de/DE/Das-BfArM/Aufgaben/Datenzugangs-und-Koordinierungsstelle/Europaeischer-Gesundheitsdatenraum/EHDS-erklart/_node.html.
178. Mühlhoff R. Das Risiko der Sekundärnutzung trainierter Modelle als zentrales Problem von Datenschutz und KI-Regulierung im Medizinbereich. In: Ruschemeier H, Steinrötter B (Ed). *Der Einsatz von KI & Robotik in der Medizin; Interdisziplinäre Fragen*. Baden-Baden: Nomos Verlagsgesellschaft; 2024. S. 27-52.

179. Konečný J, McMahan HB, Yu FX et al. Federated Learning; Strategies for Improving Communication Efficiency [online]. 2017 [Zugriff: 07.01.2026]. URL: <https://doi.org/10.48550/arXiv.1610.05492>.
180. Luckner S, Lauer W. [Regulatory classification of AI-enabled products for medical use on the basis of the EU AI Act and MDR/IVDR]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2025; 68(8): 854-861. <https://doi.org/10.1007/s00103-025-04091-9>.
181. Medizinischer Dienst Bund. Künstliche Intelligenz in Medizinprodukten: Neue Herausforderungen durch den „AI-Act“ der EU [online]. 2025 [Zugriff: 08.01.2026]. URL: https://md-bund.de/fileadmin/dokumente/Medizinprodukte_Artikel/2025-05-05_Kuenstliche_Intelligenz_in_Medizinprodukten.pdf.
182. Bundesministerium der Justiz und für Verbraucherschutz. Arzneimittelgesetz in der Fassung der Bekanntmachung vom 12. Dezember 2005 (BGBl. I S. 3394), das zuletzt durch Artikel 2 des Gesetzes vom 23. Oktober 2024 (BGBl. 2024 I Nr. 324) geändert worden ist [online]. 2024 [Zugriff: 07.01.2026]. URL: https://www.gesetze-im-internet.de/amg_1976/AMG.pdf.
183. European Society for Thoracic Imaging. ESTI Lung Cancer Screening Certification Project [online]. [Zugriff: 20.11.2025]. URL: <https://www.myesti.org/lungcancerscreeningcertificationproject/>.
184. Bütthe T, Steinert J, Sternberg H. Keine KI in der Medizin ohne die Patienten. Akademie Aktuell 2024; 3: 23-25.
185. Ayiku L, Finnegan A, Hudson T et al. Development and validation of the NICE artificial intelligence (AI) medical device intervention search filters for MEDLINE and Embase (Ovid). Int J Technol Assess Health Care 2025: 1-27. <https://doi.org/10.1017/S0266462324004823>.
186. 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.
187. Glanville J, Fleetwood K, Yellowlees A et al. Development and Testing of Search Filters to Identify Economic Evaluations in MEDLINE and EMBASE [online]. 2009 [Zugriff: 29.10.2024]. URL: https://www.cadth.ca/media/pdf/H0490_Search_Filters_for_Economic_Evaluations_mg_e.pdf.

A7 Studienlisten

A7.1 Studienlisten Nutzenbewertung

A7.1.1 Liste der gesichteten systematischen Übersichten

1. Colquitt J, Jordan M, Court R et al. Artificial intelligence software for analysing chest X-ray images to identify suspected lung cancer: an evidence synthesis early value assessment. *Health Technol Assess* 2024; 28(50): 1-75. <https://doi.org/10.3310/lkrt4721>.
2. Ewals LJS, van der Wulp K, van den Borne B et al. The Effects of Artificial Intelligence Assistance on the Radiologists' Assessment of Lung Nodules on CT Scans: A Systematic Review. *J Clin Med* 2023; 12(10). <https://doi.org/10.3390/jcm12103536>.
3. Geppert J, Asgharzadeh A, Brown A et al. Software using artificial intelligence for nodule and cancer detection in CT lung cancer screening: systematic review of test accuracy studies. *Thorax* 2024; 79(11): 1040-1049. <https://doi.org/10.1136/thorax-2024-221662>.
4. Li D, Pehrson LM, Lauridsen CA et al. The Added Effect of Artificial Intelligence on Physicians' Performance in Detecting Thoracic Pathologies on CT and Chest X-ray: A Systematic Review. *Diagnostics (Basel)* 2021; 11(12). <https://doi.org/10.3390/diagnostics11122206>.
5. National Institute for Health and Care Excellence. AI-derived computer-aided detection (CAD) software for detecting and measuring lung nodules in CT scan images; NICE diagnostics guidance 55 [online]. 2023 [Zugriff: 18.11.2025]. URL: <https://www.nice.org.uk/guidance/dg55>.
6. National Institute for Health and Care Excellence. Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals: early value assessment [online]. 2025 [Zugriff: 18.11.2025]. URL: <https://www.nice.org.uk/guidance/hte12>.

A7.1.2 Liste der ausgeschlossenen Publikationen aus der bibliografischen Recherche

Nicht E1a, D1a

1. Bae K, Oh DY, Yun ID et al. Bone Suppression on Chest Radiographs for Pulmonary Nodule Detection: Comparison between a Generative Adversarial Network and Dual-Energy Subtraction. *Korean J Radiol* 2022; 23(1): 139–149. <https://doi.org/10.3348/kjr.2021.0146>.
2. Bhure U, Cieciera M, Lehnick D et al. Incorporation of CAD (computer-aided detection) with thin-slice lung CT in routine 18F-FDG PET/CT imaging read-out protocol for detection of lung nodules. *European Journal Of Hybrid Imaging* 2023; 7(1): 17. <https://doi.org/10.1186/s41824-023-00177-2>.

3. Brown MS, Goldin JG, Suh RD et al. Lung micronodules: automated method for detection at thin-section CT--initial experience. *Radiology* 2003; 226(1): 256–262.
<https://doi.org/10.1148/radiol.2261011708>.
4. Chae KJ, Jin GY, Ko SB et al. Deep Learning for the Classification of Small (≤ 2 cm) Pulmonary Nodules on CT Imaging: A Preliminary Study. *Acad Radiol* 2020; 27(4): e55-e63.
<https://doi.org/10.1016/j.acra.2019.05.018>.
5. De Boo DW, Uffmann M, Weber M et al. Computer-aided detection of small pulmonary nodules in chest radiographs: an observer study. *Acad Radiol* 2011; 18(12): 1507–1514.
<https://doi.org/10.1016/j.acra.2011.08.008>.
6. Deng J, Zhao M, Li Q et al. Implementation of artificial intelligence in the histological assessment of pulmonary subsolid nodules. *Translational Lung Cancer Research* 2021; 10(12): 4574–4586. <https://doi.org/10.21037/tlcr-21-971>.
7. Freedman MT, Osicka T. Heat maps: an aid for data analysis and understanding of ROC CAD experiments. *Acad Radiol* 2008; 15(2): 249-259.
<https://doi.org/10.1016/j.acra.2007.07.010>.
8. Goo JM, Kim HY, Lee JW et al. Is the computer-aided detection scheme for lung nodule also useful in detecting lung cancer? *J Comput Assist Tomogr* 2008; 32(4): 570–575.
<https://doi.org/10.1097/RCT.0b013e318146261c>.
9. Hong EK, Roh B, Park B et al. Value of Using a Generative AI Model in Chest Radiography Reporting: A Reader Study. *Radiology* 2025; 314(3): e241646.
<https://doi.org/10.1148/radiol.241646>.
10. Huh JE, Lee JH, Hwang EJ et al. Effects of Expert-Determined Reference Standards in Evaluating the Diagnostic Performance of a Deep Learning Model: A Malignant Lung Nodule Detection Task on Chest Radiographs. *Korean J Radiol* 2023; 24(2): 155–165.
<https://doi.org/10.3348/kjr.2022.0548>.
11. Hunter B, Argyros C, Inglese M et al. Radiomics-based decision support tool assists radiologists in small lung nodule classification and improves lung cancer early diagnosis. *Br J Cancer* 2023; 129(12): 1949-1955. <https://doi.org/10.1038/s41416-023-02480-y>.
12. Jang S, Kim J, Lee JS et al. Real-World Diagnostic Performance and Clinical Utility of Artificial-Intelligence-Assisted Interpretation for Detection of Lung Metastasis on CT in Patients With Colorectal Cancer. *AJR* 2025. <https://doi.org/10.2214/ajr.25.33063>.
13. Jang S, Song H, Shin YJ et al. Deep Learning-based Automatic Detection Algorithm for Reducing Overlooked Lung Cancers on Chest Radiographs. *Radiology* 2020; 296(3): 652-661.
<https://doi.org/10.1148/radiol.2020200165>.

14. Kakeda S, Moriya J, Sato H et al. Improved detection of lung nodules on chest radiographs using a commercial computer-aided diagnosis system. *AJR Am J Roentgenol* 2004; 182(2): 505–510. <https://doi.org/10.2214/ajr.182.2.1820505>.
15. Kim RY, Oke JL, Pickup LC et al. Artificial Intelligence Tool for Assessment of Indeterminate Pulmonary Nodules Detected with CT. *Radiology* 2022; 304(3): 683–691. <https://doi.org/10.1148/radiol.212182>.
16. Kligerman S, Cai L, White CS. The effect of computer-aided detection on radiologist performance in the detection of lung cancers previously missed on a chest radiograph. *J Thorac Imaging* 2013; 28(4): 244–252. <https://doi.org/10.1097/RTI.0b013e31826c29ec>.
17. Koo YH, Shin KE, Park JS et al. Extravalidation and reproducibility results of a commercial deep learning-based automatic detection algorithm for pulmonary nodules on chest radiographs at tertiary hospital. *J Med Imaging Radiat Oncol* 2021; 65(1): 15–22. <https://doi.org/10.1111/1754-9485.13105>.
18. Kusano S, Nakagawa T, Aoki T et al. Efficacy of computer-aided diagnosis in lung cancer screening with low-dose spiral computed tomography: receiver operating characteristic analysis of radiologists' performance. *Japanese Journal of Radiology* 2010; 28(9): 649–655. <https://doi.org/10.1007/s11604-010-0486-1>.
19. Lee JH, Hong H, Nam G et al. Effect of Human-AI Interaction on Detection of Malignant Lung Nodules on Chest Radiographs. *Radiology* 2023; 307(5): e222976. <https://doi.org/10.1148/radiol.222976>.
20. Lee KH, Goo JM, Park CM et al. Computer-aided detection of malignant lung nodules on chest radiographs: effect on observers' performance. *Korean J Radiol* 2012; 13(5): 564–571. <https://doi.org/10.3348/kjr.2012.13.5.564>.
21. Liu J, Zhao L, Han X et al. Estimation of malignancy of pulmonary nodules at CT scans: Effect of computer-aided diagnosis on diagnostic performance of radiologists. *Asia Pac J Clin Oncol* 2021; 17(3): 216–221. <https://doi.org/10.1111/ajco.13362>.
22. Marten K, Engelke C, Seyfarth T et al. Computer-aided detection of pulmonary nodules: influence of nodule characteristics on detection performance. *Clin Radiol* 2005; 60(2): 196–206. <https://doi.org/10.1016/j.crad.2004.05.014>.
23. Matsuki Y, Nakamura K, Watanabe H et al. Usefulness of an artificial neural network for differentiating benign from malignant pulmonary nodules on high-resolution CT: evaluation with receiver operating characteristic analysis. *AJR Am J Roentgenol* 2002; 178(3): 657–663. <https://doi.org/10.2214/ajr.178.3.1780657>.
24. Matsumoto S, Ohno Y, Yamagata H et al. Potential contribution of multiplanar reconstruction (MPR) to computer-aided detection of lung nodules on MDCT. *Eur J Radiol* 2012; 81(2): 366–370. <https://doi.org/10.1016/j.ejrad.2010.12.020>.

25. Mezziane M, Obuchowski NA, Lababede O et al. A comparison of follow-up recommendations by chest radiologists, general radiologists, and pulmonologists using computer-aided detection to assess radiographs for actionable pulmonary nodules. *AJR Am J Roentgenol* 2011; 196(5): W542–549. <https://doi.org/10.2214/ajr.10.5048>.
26. Nam JG, Hwang EJ, Kim DS et al. Undetected Lung Cancer at Posteroanterior Chest Radiography: Potential Role of a Deep Learning-based Detection Algorithm. *Radiology Cardiothoracic Imaging* 2020; 2(6): e190222. <https://doi.org/10.1148/ryct.2020190222>.
27. Nam JG, Park S, Hwang EJ et al. Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. *Radiology* 2019; 290(1): 218–228. <https://doi.org/10.1148/radiol.2018180237>.
28. Sakai S, Soeda H, Takahashi N et al. Computer-aided nodule detection on digital chest radiography: validation test on consecutive T1 cases of resectable lung cancer. *J Digit Imaging* 2006; 19(4): 376–382. <https://doi.org/10.1007/s10278-006-0626-4>.
29. Schalekamp S, van Ginneken B, Koedam E et al. Computer-aided detection improves detection of pulmonary nodules in chest radiographs beyond the support by bone-suppressed images. *Radiology* 2014; 272(1): 252–261. <https://doi.org/10.1148/radiol.14131315>.
30. Shah SK, McNitt-Gray MF, De Zoysa KR et al. Solitary pulmonary nodule diagnosis on CT: Results of an observer study. *Acad Radiol* 2005; 12(4): 496–501. <https://doi.org/10.1016/j.acra.2004.12.017>.
31. Sim Y, Chung MJ, Kotter E et al. Deep Convolutional Neural Network-based Software Improves Radiologist Detection of Malignant Lung Nodules on Chest Radiographs. *Radiology* 2020; 294(1): 199–209. <https://doi.org/10.1148/radiol.2019182465>.
32. Suh YJ, Lee HJ, Sung P et al. A Novel Algorithm to Differentiate Between Multiple Primary Lung Cancers and Intrapulmonary Metastasis in Multiple Lung Cancers With Multiple Pulmonary Sites of Involvement. *J Thorac Oncol* 2020; 15(2): 203–215. <https://doi.org/10.1016/j.jtho.2019.09.221>.
33. Sun K, Chen S, Zhao J et al. Convolutional Neural Network-Based Diagnostic Model for a Solid, Indeterminate Solitary Pulmonary Nodule or Mass on Computed Tomography. *Front Oncol* 2021; 11: 792062. <https://doi.org/10.3389/fonc.2021.792062>.
34. Sung J, Park S, Lee SM et al. Added Value of Deep Learning-based Detection System for Multiple Major Findings on Chest Radiographs: A Randomized Crossover Study. *Radiology* 2021; 299(2): 450–459. <https://doi.org/10.1148/radiol.2021202818>.

35. Szucs-Farkas Z, Schick A, Cullmann JL et al. Comparison of dual-energy subtraction and electronic bone suppression combined with computer-aided detection on chest radiographs: effect on human observers' performance in nodule detection. *AJR Am J Roentgenol* 2013; 200(5): 1006-1013. <https://doi.org/10.2214/ajr.12.8877>.
36. Tang JSN, Lai JKC, Bui J et al. Impact of Different Artificial Intelligence User Interfaces on Lung Nodule and Mass Detection on Chest Radiographs. *Radiology Artificial intelligence* 2023; 5(3): e220079. <https://doi.org/10.1148/ryai.220079>.
37. Ueda D, Yamamoto A, Shimazaki A et al. Artificial intelligence-supported lung cancer detection by multi-institutional readers with multi-vendor chest radiographs: a retrospective clinical validation study. *BMC Cancer* 2021; 21(1): 1120. <https://doi.org/10.1186/s12885-021-08847-9>.
38. Wataya T, Yanagawa M, Tsubamoto M et al. Radiologists with and without deep learning-based computer-aided diagnosis: comparison of performance and interobserver agreement for characterizing and diagnosing pulmonary nodules/masses. *Eur Radiol* 2023; 33(1): 348-359. <https://doi.org/10.1007/s00330-022-08948-4>.
39. Xie J, He Y, Che S et al. Differential diagnosis of benign and lung adenocarcinoma presenting as larger solid nodules and masses based on multiscale CT radiomics. *PLoS One* 2024; 19(10): e0309033. <https://doi.org/10.1371/journal.pone.0309033>.
40. Yacoub B, Varga-Szemes A, Joseph Schoepf U et al. Impact of Artificial Intelligence Assistance on Chest CT Interpretation Times: A Prospective Randomized Study. *Am J Roentgenol* 2022; 219(5): 743-751. <https://doi.org/10.2214/ajr.22.27598>.
41. Yanagawa M, Nagatani Y, Hata A et al. Effect of spatial resolution on the diagnostic performance of machine-learning radiomics model in lung adenocarcinoma: comparisons between normal- and high-spatial-resolution imaging for predicting invasiveness. *Japanese Journal of Radiology* 2025. <https://doi.org/10.1007/s11604-025-01839-w>.
42. Yanagawa M, Niioka H, Kusumoto M et al. Diagnostic performance for pulmonary adenocarcinoma on CT: comparison of radiologists with and without three-dimensional convolutional neural network. *Eur Radiol* 2021; 31(4): 1978–1986. <https://doi.org/10.1007/s00330-020-07339-x>.
43. You S, Park JH, Park B et al. The diagnostic performance and clinical value of deep learning-based nodule detection system concerning influence of location of pulmonary nodule. *Insights Into Imaging* 2023; 14(1): 149. <https://doi.org/10.1186/s13244-023-01497-4>.

44. Yunaga H, Miyoshi H, Ochiai R et al. Image Quality and Lesion Detection of Multiplanar Reconstruction Images Using Deep Learning: Comparison with Hybrid Iterative Reconstruction. *Yonago Acta Med* 2024; 67(2): 100–107.

<https://doi.org/10.33160/yam.2024.05.001>.

45. Zeng JY, Ye HH, Yang SX et al. Clinical application of a novel computer-aided detection system based on three-dimensional CT images on pulmonary nodule. *Int J Clin Exp Med* 2015; 8(9): 16077–16082.

46. Zhang T, Wang Y, Sun Y et al. High-resolution CT image analysis based on 3D convolutional neural network can enhance the classification performance of radiologists in classifying pulmonary non-solid nodules. *Eur J Radiol* 2021; 141: 109810.

<https://doi.org/10.1016/j.ejrad.2021.109810>.

Nicht E1b, D1b

1. Nam JG, Hwang EJ, Kim J et al. AI Improves Nodule Detection on Chest Radiographs in a Health Screening Population: A Randomized Controlled Trial. *Radiology* 2023; 307(2): e221894. <https://doi.org/10.1148/radiol.221894>.

2. Nam JG, Kim HJ, Lee EH et al. Value of a deep learning-based algorithm for detecting Lung-RADS category 4 nodules on chest radiographs in a health checkup population: estimation of the sample size for a randomized controlled trial. *Eur Radiol* 2022; 32(1): 213–222.

<https://doi.org/10.1007/s00330-021-08162-8>.

3. Quanyang W, Lina Z, Yao H et al. Application of computer-aided detection for NCCN-based follow-up recommendation in subsolid nodules: Effect on inter-observer agreement. *Cancer Medicine* 2024; 13(2): e6967. <https://doi.org/10.1002/cam4.6967>.

4. Shu J, Wen D, Xu Z et al. Improved interobserver agreement on nodule type and Lung-RADS classification of subsolid nodules using computer-aided solid component measurement. *Eur J Radiol* 2022; 152: 110339. <https://doi.org/10.1016/j.ejrad.2022.110339>.

5. Wu J, Li R, Gan J et al. Application of artificial intelligence in lung cancer screening: A real-world study in a Chinese physical examination population. *Thoracic Cancer* 2024; 15(28): 2061–2072. <https://doi.org/10.1111/1759-7714.15428>.

Nicht D2a

1. Abadia AF, Yacoub B, Stringer N et al. Diagnostic Accuracy and Performance of Artificial Intelligence in Detecting Lung Nodules in Patients With Complex Lung Disease: A Noninferiority Study. *J Thorac Imaging* 2022; 37(3): 154–161.

<https://doi.org/10.1097/rti.0000000000000613>.

2. Aresta G, Ferreira C, Pedrosa J et al. Automatic Lung Nodule Detection Combined With Gaze Information Improves Radiologists' Screening Performance. *IEEE Journal of Biomedical & Health Informatics* 2020; 24(10): 2894–2901. <https://doi.org/10.1109/jbhi.2020.2976150>.

3. Bianconi F, Fravolini ML, Pizzoli S et al. Comparative evaluation of conventional and deep learning methods for semi-automated segmentation of pulmonary nodules on CT. *Quantitative Imaging in Medicine and Surgery* 2021; 11(7): 3286–3305. <https://doi.org/10.21037/qims-20-1356>.
4. Chassagnon G, Billet N, Rutten C et al. Learning from the machine: AI assistance is not an effective learning tool for resident education in chest x-ray interpretation. *Eur Radiol* 2023; 33(11): 8241-8250. <https://doi.org/10.1007/s00330-023-10043-1>.
5. Chen H, Wang XH, Ma DQ et al. Neural network-based computer-aided diagnosis in distinguishing malignant from benign solitary pulmonary nodules by computed tomography. *Chin Med J (Engl)* 2007; 120(14): 1211-1215.
6. Chiu HY, Peng RH, Lin YC et al. Artificial Intelligence for Early Detection of Chest Nodules in X-ray Images. *Biomedicines* 2022; 10(11). <https://doi.org/10.3390/biomedicines10112839>.
7. Do HM, Spear LG, Nikpanah M et al. Augmented Radiologist Workflow Improves Report Value and Saves Time: A Potential Model for Implementation of Artificial Intelligence. *Acad Radiol* 2020; 27(1): 96–105. <https://doi.org/10.1016/j.acra.2019.09.014>.
8. Du W, Luo X, Chen M. A Practical Deep Learning Model in Differentiating Pneumonia-Type Lung Carcinoma from Pneumonia on CT Images: ResNet Added with Attention Mechanism. *Journal of Oncology Print* 2022; 2022: 8906259. <https://doi.org/10.1155/2022/8906259>.
9. El Alam R, Jhala K, Hammer MM. Real-world Evaluation of Computer-aided Pulmonary Nodule Detection Software Sensitivity and False Positive Rate. *J Thorac Imaging* 2025. <https://doi.org/10.1097/rti.0000000000000835>.
10. Fan Z, Sun H, Ren C et al. Texture recognition of pulmonary nodules based on volume local direction ternary pattern. *Bioengineered* 2020; 11(1): 904-920. <https://doi.org/10.1080/21655979.2020.1807125>.
11. Han F, Qiu P, Feng K. Computed Tomography Imaging Features of Lung Cancer under Artificial Intelligence Algorithm and Its Correlation with Pathology. *Contrast Media Mol Imaging* 2023; 2023(no pagination): 9303688. <https://doi.org/10.1155/2023/9303688>.
12. Hwang SH, Shin HJ, Kim EK et al. Clinical outcomes and actual consequence of lung nodules incidentally detected on chest radiographs by artificial intelligence. *Sci Rep* 2023; 13(1): 19732. <https://doi.org/10.1038/s41598-023-47194-6>.
13. Ke X, Hu W, Su X et al. Potential of artificial intelligence based on chest computed tomography to predict the nature of part-solid nodules. *Clin Respir J* 2023; 17(4): 320-328. <https://doi.org/10.1111/crj.13597>.
14. Liu D, Zhao Y, Liu B. The effectiveness of deep learning model in differentiating benign and malignant pulmonary nodules on spiral CT. *Technol Health Care* 2024; 32(6): 5129–5140. <https://doi.org/10.3233/thc-241079>.

15. Liu X, Xu Y, Shu J et al. Preoperative CT and Radiomics Nomograms for Distinguishing Bronchiolar Adenoma and Early-Stage Lung Adenocarcinoma. *Acad Radiol* 2025; 32(2): 1054-1066. <https://doi.org/10.1016/j.acra.2024.08.047>.
16. Lv W, Wang Y, Zhou C et al. Development and validation of a clinically applicable deep learning strategy (HONORS) for pulmonary nodule classification at CT: A retrospective multicentre study. *Lung Cancer* 2021; 155: 78–86. <https://doi.org/10.1016/j.lungcan.2021.03.008>.
17. Mu J, Kuang K, Ao M et al. Deep learning predicts malignancy and metastasis of solid pulmonary nodules from CT scans. *Front Med (Lausanne)* 2023; 10: 1145846. <https://doi.org/10.3389/fmed.2023.1145846>.
18. Nishio M, Nishizawa M, Sugiyama O et al. Computer-aided diagnosis of lung nodule using gradient tree boosting and Bayesian optimization. *PLoS One* 2018; 13(4): e0195875. <https://doi.org/10.1371/journal.pone.0195875>.
19. Qi J, Deng Z, Sun G et al. One-step algorithm for fast-track localization and multi-category classification of histological subtypes in lung cancer. *Eur J Radiol* 2022; 154: 110443. <https://doi.org/10.1016/j.ejrad.2022.110443>.
20. Schultheiss M, Schober SA, Lodde M et al. A robust convolutional neural network for lung nodule detection in the presence of foreign bodies. *Sci Rep* 2020; 10(1): 12987. <https://doi.org/10.1038/s41598-020-69789-z>.
21. Selby HM, Mukherjee P, Parham C et al. Performance of alternative manual and automated deep learning segmentation techniques for the prediction of benign and malignant lung nodules. *Journal of Medical Imaging* 2023; 10(4): 044006. <https://doi.org/10.1117/1.Jmi.10.4.044006>.
22. Tartar A, Akan A, Kilic N. A novel approach to malignant-benign classification of pulmonary nodules by using ensemble learning classifiers. *Annual International Conference Of The IEEE Engineering In Medicine And Biology Society* 2014; 2014: 4651–4654. <https://doi.org/10.1109/embc.2014.6944661>.
23. Teng PH, Liang CH, Lin Y et al. Performance and educational training of radiographers in lung nodule or mass detection: Retrospective comparison with different deep learning algorithms. *Medicine (United States)* 2021; 100(23): E26270. <https://doi.org/10.1097/md.00000000000026270>.
24. Topff L, Steltenpool S, Ranschaert ER et al. Artificial intelligence-assisted double reading of chest radiographs to detect clinically relevant missed findings: a two-centre evaluation. *Eur Radiol* 2024; 34(9): 5876–5885. <https://doi.org/10.1007/s00330-024-10676-w>.

25. Tung YC, Su JH, Liao YW et al. Effective Invasiveness Recognition of Imbalanced Data by Semi-Automated Segmentations of Lung Nodules. *Biomedicines* 2023; 11(11). <https://doi.org/10.3390/biomedicines11112938>.

26. Wang YW, Wang JW, Yang SX et al. Proposing a deep learning-based method for improving the diagnostic certainty of pulmonary nodules in CT scan of chest. *Eur Radiol* 2021; 31(11): 8160-8167. <https://doi.org/10.1007/s00330-021-07919-5>.

27. Yu SN, Chiu MC, Chang YP et al. Improving Computer-Aided Thoracic Disease Diagnosis through Comparative Analysis Using Chest X-ray Images Taken at Different Times. *Sensors* 2024; 24(5). <https://doi.org/10.3390/s24051478>.

Nicht E2b

1. Madanay F, O'Donohue LS, Zikmund-Fisher BJ. Patient Reactions to Artificial Intelligence-Clinician Discrepancies: Web-Based Randomized Experiment. *J Med Internet Res* 2025; 27: e68823. <https://doi.org/10.2196/68823>.

Nicht D2b

1. Alves N, Bosma JS, Venkadesh KV et al. Prediction Variability to Identify Reduced AI Performance in Cancer Diagnosis at MRI and CT. *Radiology* 2023; 308(3): e230275. <https://doi.org/10.1148/radiol.230275>.

2. Ardila D, Kiraly AP, Bharadwaj S et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* 2019; 25(6): 954-961. <https://doi.org/10.1038/s41591-019-0447-x>.

3. Benzaquen J, Hofman P, Lopez S et al. Integrating artificial intelligence into lung cancer screening: a randomised controlled trial protocol. *BMJ Open* 2024; 14(2): e074680. <https://doi.org/10.1136/bmjopen-2023-074680>.

4. Chamberlin J, Kocher MR, Waltz J et al. Automated detection of lung nodules and coronary artery calcium using artificial intelligence on low-dose CT scans for lung cancer screening: accuracy and prognostic value. *BMC Med* 2021; 19(1): 55. <https://doi.org/10.1186/s12916-021-01928-3>.

5. Cho J, Kim J, Lee KJ et al. Incidence lung cancer after a negative ct screening in the national lung screening trial: deep learning-based detection of missed lung cancers. *J Clin Med* 2020; 9(12): 1-9. <https://doi.org/10.3390/jcm9123908>.

6. de Hoop B, De Boo DW, Gietema HA et al. Computer-aided detection of lung cancer on chest radiographs: effect on observer performance. *Radiology* 2010; 257(2): 532-540. <https://doi.org/10.1148/radiol.10092437>.

7. Gao S, Xu Z, Kang W et al. Artificial intelligence-driven computer aided diagnosis system provides similar diagnosis value compared with doctors' evaluation in lung cancer screening. *BMC Med Imaging* 2024; 24(1): 141. <https://doi.org/10.1186/s12880-024-01288-3>.

8. Garau N, Paganelli C, Summers P et al. A segmentation tool for pulmonary nodules in lung cancer screening: Testing and clinical usage. *Phys Med* 2021; 90: 23–29.
<https://doi.org/10.1016/j.ejimp.2021.08.011>.
9. Guo X, Liu X, Wang H et al. Enhanced CT images by the wavelet transform improving diagnostic accuracy of chest nodules. *J Digit Imaging* 2011; 24(1): 44–49.
<https://doi.org/10.1007/s10278-009-9248-y>.
10. Gupta S, Jacobson FL, Kong CY et al. Performance of Lung Nodule Management Algorithms for Lung-RADS Category 4 Lesions. *Acad Radiol* 2021; 28(8): 1037-1042.
<https://doi.org/10.1016/j.acra.2020.04.041>.
11. Hall H, Ruparel M, Quaife SL et al. The role of computer-assisted radiographer reporting in lung cancer screening programmes. *Eur Radiol* 2022; 32(10): 6891-6899.
<https://doi.org/10.1007/s00330-022-08824-1>.
12. Kim C, Yang Z, Park SH et al. Multicentre external validation of a commercial artificial intelligence software to analyse chest radiographs in health screening environments with low disease prevalence. *Eur Radiol* 2023; 33(5): 3501-3509. <https://doi.org/10.1007/s00330-022-09315-z>.
13. Kim EY, Kim YJ, Choi WJ et al. Concordance rate of radiologists and a commercialized deep-learning solution for chest X-ray: Real-world experience with a multicenter health screening cohort. *PLoS One* 2022; 17(2): e0264383.
<https://doi.org/10.1371/journal.pone.0264383>.
14. Lancaster HL, Jiang B, Davies MPA et al. Histological proven AI performance in the UKLS CT lung cancer screening study: Potential for workload reduction. *Eur J Cancer* 2025; 220: 115324. <https://doi.org/10.1016/j.ejca.2025.115324>.
15. Li F, Arimura H, Suzuki K et al. Computer-aided detection of peripheral lung cancers missed at CT: ROC analyses without and with localization. *Radiology* 2005; 237(2): 684-690.
<https://doi.org/10.1148/radiol.2372041555>.
16. Mazzone PJ, Obuchowski N, Phillips M et al. Lung cancer screening with computer aided detection chest radiography: design and results of a randomized, controlled trial. *PLoS One* 2013; 8(3): e59650. <https://doi.org/10.1371/journal.pone.0059650>.
17. Scholten ET, de Jong PA, Jacobs C et al. Interscan variation of semi-automated volumetry of subsolid pulmonary nodules. *Eur Radiol* 2015; 25(4): 1040-1047.
<https://doi.org/10.1007/s00330-014-3478-1>.
18. Walstra ANH, Lancaster HL, Heuvelmans MA et al. Feasibility of AI as first reader in the 4-IN-THE-LUNG-RUN lung cancer screening trial: impact on negative-misclassifications and clinical referral rate. *Eur J Cancer* 2025; 216: 115214.
<https://doi.org/10.1016/j.ejca.2024.115214>.

19. Wang J, Cai J, Tang W et al. A comparison of an integrated and image-only deep learning model for predicting the disappearance of indeterminate pulmonary nodules. *Comput Med Imaging Graph* 2025; 123: 102553. <https://doi.org/10.1016/j.compmedimag.2025.102553>.
20. Yoo H, Kim KH, Singh R et al. Validation of a Deep Learning Algorithm for the Detection of Malignant Pulmonary Nodules in Chest Radiographs. *JAMA Network Open* 2020; 3(9): e2017135. <https://doi.org/10.1001/jamanetworkopen.2020.17135>.
21. Yoo H, Lee SH, Arru CD et al. AI-based improvement in lung cancer detection on chest radiographs: results of a multi-reader study in NLST dataset. *Eur Radiol* 2021; 31(12): 9664-9674. <https://doi.org/10.1007/s00330-021-08074-7>.
22. Zhao J, Wang T, Wang B et al. Deep learning radiomics fusion model to predict visceral pleural invasion of clinical stage IA lung adenocarcinoma: a multicenter study. *J Cardiothorac Surg* 2025; 20(1): 246. <https://doi.org/10.1186/s13019-025-03488-6>.

Nicht D3a

1. Bhattacharjee A, Murugan R, Soni B et al. Ada-GridRF: A Fast and Automated Adaptive Boost Based Grid Search Optimized Random Forest Ensemble model for Lung Cancer Detection. *Physical and Engineering Sciences in Medicine* 2022; 45(3): 981-994. <https://doi.org/10.1007/s13246-022-01150-2>.
2. Han Y, Qi H, Wang L et al. Pulmonary nodules detection assistant platform: An effective computer aided system for early pulmonary nodules detection in physical examination. *Comput Methods Programs Biomed* 2022; 217: 106680. <https://doi.org/10.1016/j.cmpb.2022.106680>.
3. Wu W, Gao L, Duan H et al. Segmentation of pulmonary nodules in CT images based on 3D-UNET combined with three-dimensional conditional random field optimization. *Med Phys* 2020; 47(9): 4054-4063. <https://doi.org/10.1002/mp.14248>.

Nicht D3b

1. Hwang EJ, Goo JM, Kim HY et al. Variability in interpretation of low-dose chest CT using computerized assessment in a nationwide lung cancer screening program: comparison of prospective reading at individual institutions and retrospective central reading. *Eur Radiol* 2021; 31(5): 2845-2855. <https://doi.org/10.1007/s00330-020-07424-1>.
2. Tammemagi M, Ritchie AJ, Atkar-Khattra S et al. Predicting Malignancy Risk of Screen-Detected Lung Nodules-Mean Diameter or Volume. *J Thorac Oncol* 2019; 14(2): 203-211. <https://doi.org/10.1016/j.jtho.2018.10.006>.

Nicht D4a, D4b

1. Ahn JS, Ebrahimian S, McDermott S et al. Association of Artificial Intelligence-Aided Chest Radiograph Interpretation With Reader Performance and Efficiency. *JAMA Network Open* 2022; 5(8): e2229289. <https://doi.org/10.1001/jamanetworkopen.2022.29289>.

2. Beigelman-Aubry C, Raffy P, Yang W et al. Computer-aided detection of solid lung nodules on follow-up MDCT screening: evaluation of detection, tracking, and reading time. *AJR Am J Roentgenol* 2007; 189(4): 948–955. <https://doi.org/10.2214/ajr.07.2302>.
3. Bennani S, Regnard NE, Ventre J et al. Using AI to Improve Radiologist Performance in Detection of Abnormalities on Chest Radiographs. *Radiology* 2023; 309(3): e230860. <https://doi.org/10.1148/radiol.230860>.
4. Beyer F, Zierott L, Fallenberg EM et al. Comparison of sensitivity and reading time for the use of computer-aided detection (CAD) of pulmonary nodules at MDCT as concurrent or second reader. *Eur Radiol* 2007; 17(11): 2941–2947. <https://doi.org/10.1007/s00330-007-0667-1>.
5. Bhatia BS, Morlese JF, Yusuf S et al. A real-world evaluation of the diagnostic accuracy of radiologists using positive predictive values verified from deep learning and natural language processing chest algorithms deployed retrospectively. *BJR Open* 2024; 6(1): tzad009. <https://doi.org/10.1093/bjro/tzad009>.
6. Brown M, Browning P, Wahi-Anwar MW et al. Integration of Chest CT CAD into the Clinical Workflow and Impact on Radiologist Efficiency. *Acad Radiol* 2019; 26(5): 626–631. <https://doi.org/10.1016/j.acra.2018.07.006>.
7. Brown MS, Goldin JG, Rogers S et al. Computer-aided lung nodule detection in CT: results of large-scale observer test. *Acad Radiol* 2005; 12(6): 681–686. <https://doi.org/10.1016/j.acra.2005.02.041>.
8. Brown MS, Pais R, Qing P et al. An architecture for computer-aided detection and radiologic measurement of lung nodules in clinical trials. *Cancer Informatics [Electronic Resource]* 2007; 4: 25–31.
9. Chao HS, Tsai CY, Chou CW et al. Artificial Intelligence Assisted Computational Tomographic Detection of Lung Nodules for Prognostic Cancer Examination: A Large-Scale Clinical Trial. *Biomedicines* 2023; 11(1). <https://doi.org/10.3390/biomedicines11010147>.
10. Choi SY, Park S, Kim M et al. Evaluation of a deep learning-based computer-aided detection algorithm on chest radiographs: Case-control study. *Medicine* 2021; 100(16): e25663. <https://doi.org/10.1097/md.00000000000025663>.
11. Cui X, Zheng S, Heuvelmans MA et al. Performance of a deep learning-based lung nodule detection system as an alternative reader in a Chinese lung cancer screening program. *Eur J Radiol* 2022; 146: 110068. <https://doi.org/10.1016/j.ejrad.2021.110068>.
12. Das M, Muhlenbruch G, Mahnken AH et al. Small pulmonary nodules: effect of two computer-aided detection systems on radiologist performance. *Radiology* 2006; 241(2): 564–571. <https://doi.org/10.1148/radiol.2412051139>.

13. De Boo DW, van Hoorn F, van Schuppen J et al. Observer training for computer-aided detection of pulmonary nodules in chest radiography. *Eur Radiol* 2012; 22(8): 1659–1664. <https://doi.org/10.1007/s00330-012-2412-7>.
14. Diao K, Chen Y, Liu Y et al. Diagnostic study on clinical feasibility of an AI-based diagnostic system as a second reader on mobile CT images: a preliminary result. *Annals of Translational Medicine* 2022; 10(12): 668. <https://doi.org/10.21037/atm-22-2157>.
15. Fraioli F, Bertoletti L, Napoli A et al. Computer-aided detection (CAD) in lung cancer screening at chest MDCT: ROC analysis of CAD versus radiologist performance. *J Thorac Imaging* 2007; 22(3): 241–246. <https://doi.org/10.1097/RTI.0b013e318033aae8>.
16. Godoy MC, Kim TJ, White CS et al. Benefit of computer-aided detection analysis for the detection of subsolid and solid lung nodules on thin- and thick-section CT. *AJR Am J Roentgenol* 2013; 200(1): 74–83. <https://doi.org/10.2214/ajr.11.7532>.
17. Guo L, Zhou C, Xu J et al. Deep Learning for Chest X-ray Diagnosis: Competition Between Radiologists with or Without Artificial Intelligence Assistance. *Journal of Imaging Informatics in Medicine* 2024; 37(3): 922-934. <https://doi.org/10.1007/s10278-024-00990-6>.
18. Gurung J, Maataoui A, Khan M et al. Automated detection of lung nodules in multidetector CT: influence of different reconstruction protocols on performance of a software prototype. *Rofo* 2006; 178(1): 71–77. <https://doi.org/10.1055/s-2005-858831>.
19. Hata A, Yanagawa M, Yoshida Y et al. Combination of Deep Learning-Based Denoising and Iterative Reconstruction for Ultra-Low-Dose CT of the Chest: Image Quality and Lung-RADS Evaluation. *AJR Am J Roentgenol* 2020; 215(6): 1321–1328. <https://doi.org/10.2214/ajr.19.22680>.
20. Hempel HL, Engbersen MP, Wakkie J et al. Higher agreement between readers with deep learning CAD software for reporting pulmonary nodules on CT. *European Journal of Radiology Open* 2022; 9: 100435. <https://doi.org/10.1016/j.ejro.2022.100435>.
21. Hsu HH, Ko KH, Chou YC et al. Performance and reading time of lung nodule identification on multidetector CT with or without an artificial intelligence-powered computer-aided detection system. *Clin Radiol* 2021; 76(8): 626.e23–626.e32. <https://doi.org/10.1016/j.crad.2021.04.006>.
22. Iwasawa T, Matsumoto S, Aoki T et al. A comparison of axial versus coronal image viewing in computer-aided detection of lung nodules on CT. *Japanese Journal of Radiology* 2015; 33(2): 76–83. <https://doi.org/10.1007/s11604-014-0383-0>.
23. Jacobs C, van Rikxoort EM, Scholten ET et al. Solid, part-solid, or non-solid?: classification of pulmonary nodules in low-dose chest computed tomography by a computer-aided diagnosis system. *Invest Radiol* 2015; 50(3): 168-173. <https://doi.org/10.1097/rli.000000000000121>.

24. Jacobs C, van Rikxoort EM, Twellmann T et al. Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images. *Med Image Anal* 2014; 18(2): 374-384. <https://doi.org/10.1016/j.media.2013.12.001>.
25. Jankowski A, Martinelli T, Timsit JF et al. Pulmonary nodule detection on MDCT images: evaluation of diagnostic performance using thin axial images, maximum intensity projections, and computer-assisted detection. *Eur Radiol* 2007; 17(12): 3148-3156. <https://doi.org/10.1007/s00330-007-0727-6>.
26. Jeon KN, Goo JM, Lee CH et al. Computer-aided nodule detection and volumetry to reduce variability between radiologists in the interpretation of lung nodules at low-dose screening computed tomography. *Invest Radiol* 2012; 47(8): 457-461. <https://doi.org/10.1097/RLI.0b013e318250a5aa>.
27. Jin KN, Kim EY, Kim YJ et al. Diagnostic effect of artificial intelligence solution for referable thoracic abnormalities on chest radiography: a multicenter respiratory outpatient diagnostic cohort study. *Eur Radiol* 2022; 32(5): 3469-3479. <https://doi.org/10.1007/s00330-021-08397-5>.
28. Katase S, Ichinose A, Hayashi M et al. Development and performance evaluation of a deep learning lung nodule detection system. *BMC Med Imaging* 2022; 22(1): 203. <https://doi.org/10.1186/s12880-022-00938-8>.
29. Kaviani P, Kalra MK, Digumarthy SR et al. Frequency of Missed Findings on Chest Radiographs (CXR) in an International, Multicenter Study: Application of AI to Reduce Missed Findings. *Diagnostics* 2022; 12(10). <https://doi.org/10.3390/diagnostics12102382>.
30. Khan F, Das I, Kotnik M et al. AI-assisted detection for chest X-rays (AID-CXR): a multi-reader multi-case study protocol. *BMJ Open* 2024; 14(12): e080554. <https://doi.org/10.1136/bmjopen-2023-080554>.
31. Kim H, Lee KH, Han K et al. Development and Validation of a Deep Learning-Based Synthetic Bone-Suppressed Model for Pulmonary Nodule Detection in Chest Radiographs. *JAMA Network Open* 2023; 6(1): e2253820. <https://doi.org/10.1001/jamanetworkopen.2022.53820>.
32. Kozuka T, Matsukubo Y, Kadoba T et al. Efficiency of a computer-aided diagnosis (CAD) system with deep learning in detection of pulmonary nodules on 1-mm-thick images of computed tomography. *Japanese Journal of Radiology* 2020; 38(11): 1052-1061. <https://doi.org/10.1007/s11604-020-01009-0>.
33. Lan CC, Hsieh MS, Hsiao JK et al. Deep Learning-based Artificial Intelligence Improves Accuracy of Error-prone Lung Nodules. *Int J Med Sci* 2022; 19(3): 490–498. <https://doi.org/10.7150/ijms.69400>.

34. Lancaster HL, Zheng S, Aleshina OO et al. Outstanding negative prediction performance of solid pulmonary nodule volume AI for ultra-LDCT baseline lung cancer screening risk stratification. *Lung Cancer* 2022; 165: 133-140. <https://doi.org/10.1016/j.lungcan.2022.01.002>.
35. Lee HW, Jin KN, Oh S et al. Artificial Intelligence Solution for Chest Radiographs in Respiratory Outpatient Clinics: Multicenter Prospective Randomized Clinical Trial. *Annals of the American Thoracic Society* 2023; 20(5): 660-667. <https://doi.org/10.1513/AnnalsATS.202206-481OC>.
36. Lee IJ, Gamsu G, Czum J et al. Lung nodule detection on chest CT: evaluation of a computer-aided detection (CAD) system. *Korean J Radiol* 2005; 6(2): 89-93. <https://doi.org/10.3348/kjr.2005.6.2.89>.
37. Li L, Liu Z, Huang H et al. Evaluating the performance of a deep learning-based computer-aided diagnosis (DL-CAD) system for detecting and characterizing lung nodules: Comparison with the performance of double reading by radiologists. *Thoracic Cancer* 2019; 10(2): 183-192. <https://doi.org/10.1111/1759-7714.12931>.
38. Liu K, Li Q, Ma J et al. Evaluating a Fully Automated Pulmonary Nodule Detection Approach and Its Impact on Radiologist Performance. *Radiology Artificial intelligence* 2019; 1(3): e180084. <https://doi.org/10.1148/ryai.2019180084>.
39. Liu W, Wu Y, Zheng Z et al. Enhancing Diagnostic Accuracy of Lung Nodules in Chest Computed Tomography Using Artificial Intelligence: Retrospective Analysis. *J Med Internet Res* 2025; 27: e64649. <https://doi.org/10.2196/64649>.
40. Marten K, Seyfarth T, Auer F et al. Computer-assisted detection of pulmonary nodules: performance evaluation of an expert knowledge-based detection system in consensus reading with experienced and inexperienced chest radiologists. *Eur Radiol* 2004; 14(10): 1930-1938. <https://doi.org/10.1007/s00330-004-2389-y>.
41. Matsumoto S, Ohno Y, Yamagata H et al. Computer-aided detection of lung nodules on multidetector row computed tomography using three-dimensional analysis of nodule candidates and their surroundings. *Radiat Med* 2008; 26(9): 562-569. <https://doi.org/10.1007/s11604-008-0272-5>.
42. McCulloch CC, Kaucic RA, Mendonca PR et al. Model-based detection of lung nodules in computed tomography exams. *Thoracic computer-aided diagnosis. Acad Radiol* 2004; 11(3): 258-266. [https://doi.org/10.1016/s1076-6332\(03\)00729-3](https://doi.org/10.1016/s1076-6332(03)00729-3).
43. Miki S, Nomura Y, Hayashi N et al. Prospective Study of Spatial Distribution of Missed Lung Nodules by Readers in CT Lung Screening Using Computer-assisted Detection. *Acad Radiol* 2021; 28(5): 647-654. <https://doi.org/10.1016/j.acra.2020.03.015>.

44. Murchison JT, Ritchie G, Senyszak D et al. Validation of a deep learning computer aided system for CT based lung nodule detection, classification, and growth rate estimation in a routine clinical population. PLoS One 2022; 17(5): e0266799.
<https://doi.org/10.1371/journal.pone.0266799>.
45. Nam JG, Kim M, Park J et al. Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs. Eur Respir J 2021; 57(5).
<https://doi.org/10.1183/13993003.03061-2020>.
46. Peldschus K, Herzog P, Wood SA et al. Computer-aided diagnosis as a second reader: spectrum of findings in CT studies of the chest interpreted as normal. Chest 2005; 128(3): 1517–1523. <https://doi.org/10.1378/chest.128.3.1517>.
47. Peters AA, Wiescholek N, Muller M et al. Impact of artificial intelligence assistance on pulmonary nodule detection and localization in chest CT: a comparative study among radiologists of varying experience levels. Sci Rep 2024; 14(1): 22447.
<https://doi.org/10.1038/s41598-024-73435-3>.
48. Roos JE, Paik D, Olsen D et al. Computer-aided detection (CAD) of lung nodules in CT scans: radiologist performance and reading time with incremental CAD assistance. Eur Radiol 2010; 20(3): 549–557. <https://doi.org/10.1007/s00330-009-1596-y>.
49. Schalekamp S, van Ginneken B, Heggelman B et al. New methods for using computer-aided detection information for the detection of lung nodules on chest radiographs. Br J Radiol 2014; 87(1036): 20140015. <https://doi.org/10.1259/bjr.20140015>.
50. Selvam S, Peyrony O, Elezi A et al. Efficacy of a deep learning-based software for chest X-ray analysis in an emergency department. Diagn Interv Imaging 2025.
<https://doi.org/10.1016/j.diii.2025.03.007>.
51. Silva M, Schaefer Prokop CM, Jacobs C et al. Detection of Subsolid Nodules in Lung Cancer Screening: complementary Sensitivity of Visual Reading and Computer-Aided Diagnosis. Invest Radiol 2018; 53(8): 441–449.
<https://doi.org/10.1097/rli.0000000000000464>.
52. Singh R, Kalra MK, Homayounieh F et al. Artificial intelligence-based vessel suppression for detection of sub-solid nodules in lung cancer screening computed tomography. Quantitative Imaging in Medicine and Surgery 2021; 11(4): 1134–1143.
<https://doi.org/10.21037/qims-20-630>.
53. Smith D, Melville P, Fozzard N et al. Artificial intelligence software in pulmonary nodule assessment. J R Coll Physicians Edinb 2022; 52(3): 228–231.
<https://doi.org/10.1177/14782715221123856>.

54. Takahashi EA, Koo CW, White DB et al. Prospective Pilot Evaluation of Radiologists and Computer-aided Pulmonary Nodule Detection on Ultra-low-Dose CT With Tin Filtration. *J Thorac Imaging* 2018; 33(6): 396–401. <https://doi.org/10.1097/rti.0000000000000348>.
55. Takaishi T, Ozawa Y, Bando Y et al. Incorporation of a computer-aided vessel-suppression system to detect lung nodules in CT images: effect on sensitivity and reading time in routine clinical settings. *Japanese Journal of Radiology* 2021; 39(2): 159-164. <https://doi.org/10.1007/s11604-020-01043-y>.
56. Toda N, Hashimoto M, Iwabuchi Y et al. Validation of deep learning-based computer-aided detection software use for interpretation of pulmonary abnormalities on chest radiographs and examination of factors that influence readers' performance and final diagnosis. *Japanese Journal of Radiology* 2023; 41(1): 38-44. <https://doi.org/10.1007/s11604-022-01330-w>.
57. van den Berk IAH, Jacobs C, Kanglie M et al. An AI deep learning algorithm for detecting pulmonary nodules on ultra-low-dose CT in an emergency setting: a reader study. *European Radiology Experimental* 2024; 8(1): 132. <https://doi.org/10.1186/s41747-024-00518-1>.
58. van Leeuwen KG, Schalekamp S, Rutten M et al. Comparison of Commercial AI Software Performance for Radiograph Lung Nodule Detection and Bone Age Prediction. *Radiology* 2024; 310(1): e230981. <https://doi.org/10.1148/radiol.230981>.
59. von Berg J, Young S, Carolus H et al. A novel bone suppression method that improves lung nodule detection : Suppressing dedicated bone shadows in radiographs while preserving the remaining signal. *Int J Comput Assist Radiol Surg* 2016; 11(4): 641–655. <https://doi.org/10.1007/s11548-015-1278-y>.
60. White CS, Pugatch R, Koonce T et al. Lung nodule CAD software as a second reader: a multicenter study. *Acad Radiol* 2008; 15(3): 326-333. <https://doi.org/10.1016/j.acra.2007.09.027>.
61. Xu Y, Ma D, He W. Assessing the use of digital radiography and a real-time interactive pulmonary nodule analysis system for large population lung cancer screening. *Eur J Radiol* 2012; 81(4): e451–456. <https://doi.org/10.1016/j.ejrad.2011.04.031>.
62. Yanagawa M, Honda O, Yoshida S et al. Commercially available computer-aided detection system for pulmonary nodules on thin-section images using 64 detectors-row CT: preliminary study of 48 cases. *Acad Radiol* 2009; 16(8): 924–933. <https://doi.org/10.1016/j.acra.2009.01.030>.
63. Yoon SH, Park S, Jang S et al. Use of artificial intelligence in triaging of chest radiographs to reduce radiologists' workload. *Eur Radiol* 2024; 34(2): 1094-1103. <https://doi.org/10.1007/s00330-023-10124-1>.

64. Zhang Y, Jiang B, Zhang L et al. Lung Nodule Detectability of Artificial Intelligence-assisted CT Image Reading in Lung Cancer Screening. *Current Medical Imaging* 2022; 18(3): 327–334. <https://doi.org/10.2174/1573405617666210806125953>.

65. Zhao Y, de Bock GH, Vliegenthart R et al. Performance of computer-aided detection of pulmonary nodules in low-dose CT: comparison with double reading by nodule volume. *Eur Radiol* 2012; 22(10): 2076–2084. <https://doi.org/10.1007/s00330-012-2437-y>.

66. Zhu X, Zhu L, Song D et al. Comparison of single- and dual-energy CT combined with artificial intelligence for the diagnosis of pulmonary nodules. *Clin Radiol* 2023; 78(2): e99–e105. <https://doi.org/10.1016/j.crad.2022.09.114>.

Nicht D5a, D5b

1. Awai K, Murao K, Ozawa A et al. Pulmonary nodules: estimation of malignancy at thin-section helical CT--effect of computer-aided diagnosis on performance of radiologists. *Radiology* 2006; 239(1): 276–284. <https://doi.org/10.1148/radiol.2383050167>.

2. Kiraly AP, Cunningham CA, Najafi R et al. Assistive AI in Lung Cancer Screening: A Retrospective Multinational Study in the United States and Japan. *Radiology Artificial intelligence* 2024; 6(3): e230079. <https://doi.org/10.1148/ryai.230079>.

3. Way T, Chan HP, Hadjiiski L et al. Computer-aided diagnosis of lung nodules on CT scans: ROC study of its effect on radiologists' performance. *Acad Radiol* 2010; 17(3): 323–332. <https://doi.org/10.1016/j.acra.2009.10.016>.

4. Yoo H, Kim EY, Kim H et al. Artificial Intelligence-Based Identification of Normal Chest Radiographs: A Simulation Study in a Multicenter Health Screening Cohort. *Korean J Radiol* 2022; 23(10): 1009–1018. <https://doi.org/10.3348/kjr.2022.0189>.

Nicht D6a, D6b

1. Abe Y, Hanai K, Nakano M et al. A computer-aided diagnosis (CAD) system in lung cancer screening with computed tomography. *Anticancer Res* 2005; 25(1B): 483–488.

2. Gruden JF, Webb WR, Naidich DP et al. Multinodular disease: anatomic localization at thin-section CT--multireader evaluation of a simple algorithm. *Radiology* 1999; 210(3): 711–720. <https://doi.org/10.1148/radiology.210.3.r99mr21711>.

3. Hwang EJ, Goo JM, Kim HY et al. Implementation of the cloud-based computerized interpretation system in a nationwide lung cancer screening with low-dose CT: comparison with the conventional reading system. *Eur Radiol* 2021; 31(1): 475–485. <https://doi.org/10.1007/s00330-020-07151-7>.

4. Li J, Zhou L, Zhan Y et al. How does the artificial intelligence-based image-assisted technique help physicians in diagnosis of pulmonary adenocarcinoma? A randomized controlled experiment of multicenter physicians in China. *J Am Med Inform Assoc* 2022; 29(12): 2041–2049. <https://doi.org/10.1093/jamia/ocac179>.

Nicht E7a, E7b, D8a, D8b

1. Chen Y, Tian X, Fan K et al. The Value of Artificial Intelligence Film Reading System Based on Deep Learning in the Diagnosis of Non-Small-Cell Lung Cancer and the Significance of Efficacy Monitoring: A Retrospective, Clinical, Nonrandomized, Controlled Study. *Comput Math Methods Med* 2022; 2022: 2864170. <https://doi.org/10.1155/2022/2864170>.
2. Duncan SF, McConnachie A, Blackwood J et al. Radiograph accelerated detection and identification of cancer in the lung (RADICAL): a mixed methods study to assess the clinical effectiveness and acceptability of Qure.ai artificial intelligence software to prioritise chest X-ray (CXR) interpretation. *BMJ Open* 2024; 14(9): e081062. <https://doi.org/10.1136/bmjopen-2023-081062>.
3. Revel MP, Abdoul H, Chassagnon G et al. Lung CANcer SCreening in French women using low-dose CT and Artificial intelligence for DETection: the CASCADE study protocol. *BMJ Open* 2022; 12(12): e067263. <https://doi.org/10.1136/bmjopen-2022-067263>.

A7.2 Studienlisten gesundheitsökonomische Bewertung

A7.2.1 Liste der gesichteten systematischen Übersichten

1. National Institute for Health and Care Excellence. Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals: early value assessment [online]. 2025 [Zugriff: 18.11.2025]. URL: <https://www.nice.org.uk/guidance/hte12>.
2. National Institute for Health and Care Excellence. AI-derived computer-aided detection (CAD) software for detecting and measuring lung nodules in CT scan images; NICE diagnostics guidance 55 [online]. 2023 [Zugriff: 18.11.2025]. URL: <https://www.nice.org.uk/guidance/dg55>.
3. Behr CM, Oude Wolcherink MJ, Mj IJ et al. Population-Based Screening Using Low-Dose Chest Computed Tomography: A Systematic Review of Health Economic Evaluations. *Pharmacoeconomics* 2023; 41(4): 395-411. <https://doi.org/10.1007/s40273-022-01238-3>.
4. Colquitt J, Jordan M, Court R et al. Artificial intelligence software for analysing chest X-ray images to identify suspected lung cancer: an evidence synthesis early value assessment. *Health Technol Assess* 2024; 28(50): 1-75. <https://doi.org/10.3310/lkrt4721>.

A7.2.2 Liste der ausgeschlossenen Publikationen zum Thema

Nicht E1 bzw. D1

1. Du Y, Greuter MJW, Prokop MW et al. Pricing and cost-saving potential for deep-learning computer-aided lung nodule detection software in CT lung cancer screening. *Insights Into Imaging* 2023; 14(1): 208. <https://doi.org/10.1186/s13244-023-01561-z>.

Nicht E2 bzw. D2

1. Adams SJ, Mondal P, Penz E et al. Development and Cost Analysis of a Lung Nodule Management Strategy Combining Artificial Intelligence and Lung-RADS for Baseline Lung Cancer Screening. J Am Coll Radiol 2021; 18(5): 741–751. <https://doi.org/10.1016/j.jacr.2020.11.014>.
2. Ziegelmayr S, Graf M, Makowski M et al. Cost-Effectiveness of Artificial Intelligence Support in Computed Tomography-Based Lung Cancer Screening. Cancers (Basel) 2022; 14(7). <https://doi.org/10.3390/cancers14071729>.

Nicht EÖ1

1. Colquitt J, Jordan M, Court R et al. Artificial intelligence software for analysing chest X-ray images to identify suspected lung cancer: an evidence synthesis early value assessment. Health Technol Assess 2024; 28(50): 1–75. <https://doi.org/10.3310/lkrt4721>.
2. Duncan SF, McConnachie A, Blackwood J et al. Radiograph accelerated detection and identification of cancer in the lung (RADICAL): a mixed methods study to assess the clinical effectiveness and acceptability of Qure.ai artificial intelligence software to prioritise chest X-ray (CXR) interpretation. BMJ Open 2024; 14(9): e081062. <https://doi.org/10.1136/bmjopen-2023-081062>.
3. Lam S, Bai C, Baldwin DR et al. Current and Future Perspectives on Computed Tomography Screening for Lung Cancer: A Roadmap From 2023 to 2027 From the International Association for the Study of Lung Cancer. J Thorac Oncol 2024; 19(1): 36–51. <https://doi.org/10.1016/j.jtho.2023.07.019>.
4. Powell AC. Impact of the Artificial Nudge. Acad Radiol 2020; 27(1): 143–146. <https://doi.org/10.1016/j.acra.2019.09.010>.
5. Revel MP, Abdoul H, Chassagnon G et al. Lung Cancer SCreening in French women using low-dose CT and Artificial intelligence for DETection: The CASCADE study protocol. BMJ Open 2022; 12(12): e067263. <https://doi.org/10.1136/bmjopen-2022-067263>.

A7.3 Publikationslisten zu ethischen, sozialen, rechtlichen und organisatorischen Aspekten sowie Umwelt- und Klimaaspekten

A7.3.1 Liste der eingeschlossenen Publikationen zu ethischen Aspekten

1. Alderman JE, Palmer J, Laws E et al. Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together consensus recommendations. Lancet Digit Health 2025; 7(1): e64-e88. [https://doi.org/10.1016/s2589-7500\(24\)00224-3](https://doi.org/10.1016/s2589-7500(24)00224-3).
2. Bowers PJT, Kirk FM. Insights into artificial intelligence and our intelligence-on the frontier of lung cancer screening. J Thorac Dis 2024; 16(11): 7905-7909. <https://doi.org/10.21037/jtd-24-1077>.

3. Bundesärztekammer. Künstliche Intelligenz in der Gesundheitsversorgung [online]. 2025 [Zugriff: 07.01.2026]. URL: https://www.bundesaerztekammer.de/fileadmin/user_upload/BAEK/Themen/Digitalisierung/Thesenpapier_KI_in_der_Gesundheitsversorgung_03.2025.pdf.
4. Busch F, Hoffmann L, Xu L et al. Multinational Attitudes Toward AI in Health Care and Diagnostics Among Hospital Patients. JAMA Netw Open 2025; 8(6): e2514452. <https://doi.org/10.1001/jamanetworkopen.2025.14452>.
5. Crowson MG, Moukheiber D, Arévalo AR et al. A systematic review of federated learning applications for biomedical data. PLOS Digit Health 2022; 1(5): e0000033. <https://doi.org/10.1371/journal.pdig.0000033>.
6. Deutscher Ethikrat. Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz; Stellungnahme [online]. 2023 [Zugriff: 07.01.2026]. URL: <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>.
7. Eichelberger J. Arzthaftung im Kontext von KI und Robotik. In: Ruschemeier H, Steinrötter B (Ed). Der Einsatz von KI & Robotik in der Medizin; Interdisziplinäre Fragen. Baden-Baden: Nomos Verlagsgesellschaft; 2024. S. 27-52.
8. Ewals LJS, Heesterbeek LJJ, Yu B et al. The Impact of Expectation Management and Model Transparency on Radiologists' Trust and Utilization of AI Recommendations for Lung Nodule Assessment on Computed Tomography: Simulated Use Study. JMIR AI 2024; 3: e52211. <https://doi.org/10.2196/52211>.
9. Ewals LJS, van der Wulp K, van den Borne B et al. The Effects of Artificial Intelligence Assistance on the Radiologists' Assessment of Lung Nodules on CT Scans: A Systematic Review. J Clin Med 2023; 12(10). <https://doi.org/10.3390/jcm12103536>.
10. Farič N, Hinder S, Williams R et al. Early experiences of integrating an artificial intelligence-based diagnostic decision support system into radiology settings: a qualitative study. J Am Med Inform Assoc 2023; 31(1): 24-34. <https://doi.org/10.1093/jamia/ocad191>.
11. Katzenmeier C. Haftung für Schäden durch KI in der Medizin. In: Grundmann T, Hey J, Hobe S et al (Ed). Die Macht der Algorithmen. Baden-Baden: Nomos Verlagsgesellschaft; 2023. S. 73-102.
12. Li D, Pehrson LM, Lauridsen CA et al. The Added Effect of Artificial Intelligence on Physicians' Performance in Detecting Thoracic Pathologies on CT and Chest X-ray: A Systematic Review. Diagnostics (Basel) 2021; 11(12). <https://doi.org/10.3390/diagnostics11122206>.

13. Madanay F, O'Donohue LS, Zikmund-Fisher BJ. Patient Reactions to Artificial Intelligence-Clinician Discrepancies: Web-Based Randomized Experiment. J Med Internet Res 2025; 27: e68823. <https://doi.org/10.2196/68823>.
14. National Institute for Health and Care Excellence. AI-derived computer-aided detection (CAD) software for detecting and measuring lung nodules in CT scan images; NICE diagnostics guidance 55 [online]. 2023 [Zugriff: 18.11.2025]. URL: <https://www.nice.org.uk/guidance/htg687/chapter/1-Recommendations>.
15. National Institute for Health and Care Excellence. Artificial intelligence (AI)-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals: early value assessment; Recommendations [online]. 2025 [Zugriff: 18.11.2025]. URL: <https://www.nice.org.uk/guidance/htg696/chapter/1-Recommendations>.
16. Sorin V, Soffer S, Glicksberg BS et al. Adversarial attacks in radiology - A systematic review. Eur J Radiol 2023; 167: 111085. <https://doi.org/10.1016/j.ejrad.2023.111085>.
17. Thong LT, Chou HS, Chew HSJ et al. Diagnostic test accuracy of artificial intelligence-based imaging for lung cancer screening: A systematic review and meta-analysis. Lung Cancer 2023; 176: 4-13. <https://doi.org/10.1016/j.lungcan.2022.12.002>.
18. Wang TW, Hong JS, Chiu HY et al. Standalone deep learning versus experts for diagnosis lung cancer on chest computed tomography: a systematic review. Eur Radiol 2024; 34(11): 7397-7407. <https://doi.org/10.1007/s00330-024-10804-6>.
19. Zentrale Kommission zur Wahrung ethischer Grundsätze in der Medizin und ihren Grenzgebieten (Zentrale Ethikkommission) bei der Bundesärztekammer. Entscheidungsunterstützung ärztlicher Tätigkeit durch Künstliche Intelligenz. Dtsch Arztebl 2021; 118: A1-A13.

A7.3.2 Liste der eingeschlossenen Publikationen zu sozialen Aspekten

1. Akudjedu TN, Torre S, Khine R et al. Knowledge, perceptions, and expectations of Artificial intelligence in radiography practice: A global radiography workforce survey. J Med Imaging Radiat Sci 2023; 54(1): 104-116. <https://doi.org/10.1016/j.jmir.2022.11.016>.
2. Alderman JE, Palmer J, Laws E et al. Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together consensus recommendations. Lancet Digit Health 2025; 7(1): e64-e88. [https://doi.org/10.1016/s2589-7500\(24\)00224-3](https://doi.org/10.1016/s2589-7500(24)00224-3).
3. Ali N, Lifford KJ, Carter B et al. Barriers to uptake among high-risk individuals declining participation in lung cancer screening: a mixed methods analysis of the UK Lung Cancer Screening (UKLS) trial. BMJ Open 2015; 5(7): e008254. <https://doi.org/10.1136/bmjopen-2015-008254>.

4. Becker N, Motsch E, Trotter A et al. Lung cancer mortality reduction by LDCT screening- Results from the randomized German LUSI trial. *Int J Cancer* 2020; 146(6): 1503-1513. <https://dx.doi.org/10.1002/ijc.32486>.
5. Blum TG, Vogel-Claussen J, Andreas S et al. [Statement paper on the implementation of a national organized program in Germany for the early detection of lung cancer in risk populations using low-dose CT screening including management of screening findings]. *Pneumologie* 2024; 78(1): 15-34. <https://doi.org/10.1055/a-2175-4580>.
6. Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz. Verordnung über die Zulässigkeit der Anwendung der Niedrigdosis-Computertomographie zur Früherkennung von Lungenkrebs bei rauchenden Personen (Lungenkrebs-Früherkennungs-Verordnung – LuKrFrühErkV). *Bundesgesetzblatt Teil 1* 2024; (162): 1-5.
7. Busch F, Hoffmann L, Xu L et al. Multinational Attitudes Toward AI in Health Care and Diagnostics Among Hospital Patients. *JAMA Netw Open* 2025; 8(6): e2514452. <https://doi.org/10.1001/jamanetworkopen.2025.14452>.
8. Europäisches Parlament, Rat der Europäischen Union. Verordnung (EU) 2025/327 des Europäischen Parlaments und des Rates vom 11. Februar 2025 über den europäischen Gesundheitsdatenraum sowie zur Änderung der Richtlinie 2011/24/EU und der Verordnung (EU) 2024/2847 [online]. 2025 [Zugriff: 13.11.2025]. URL: <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32025R0327>.
9. Farič N, Hinder S, Williams R et al. Early experiences of integrating an artificial intelligence-based diagnostic decision support system into radiology settings: a qualitative study. *J Am Med Inform Assoc* 2023; 31(1): 24-34. <https://doi.org/10.1093/jamia/ocad191>.
10. Fritsch SJ, Blankenheim A, Wahl A et al. Attitudes and perception of artificial intelligence in healthcare: A cross-sectional survey among patients. *Digit Health* 2022; 8: 20552076221116772. <https://doi.org/10.1177/20552076221116772>.
11. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Krebsfrüherkennungs-Richtlinie (KFE-RL): Einführung der Lungenkrebsfrüherkennung mittels Niedrigdosis-Computertomographie bei Rauchern [online]. 2024 [Zugriff: 23.10.2025]. URL: https://www.g-ba.de/downloads/39-261-7301/2025-06-18_KFE-RL_Einf-Lungenkrebsfrueherkennung-Niedrigdosis-CT-Raucher_BAnz.pdf.
12. Gemeinsamer Bundesausschuss. Lungenkrebs-Früherkennung bei starken Raucherinnen und Rauchern wird voraussichtlich ab April 2026 Kassenleistung [online]. 2025 [Zugriff: 23.10.2025]. URL: <https://www.g-ba.de/presse/pressemitteilungen-meldungen/1263/>.

13. Hovanec J, Kendzia B, Olsson A et al. Socioeconomic Status, Smoking, and Lung Cancer: Mediation and Bias Analysis in the SYNERGY Study. *Epidemiology* 2025; 36(2): 245-252. <https://doi.org/10.1097/ede.0000000000001807>.
14. Kauczor HU, von Stackelberg O, Nischwitz E et al. Strengthening lung cancer screening in Europe: fostering participation, improving outcomes, and addressing health inequalities through collaborative initiatives in the SOLACE consortium. *Insights Imaging* 2024; 15(1): 252. <https://doi.org/10.1186/s13244-024-01814-5>.
15. Kocak B, Ponsiglione A, Romeo V et al. Radiology AI and sustainability paradox: environmental, economic, and social dimensions. *Insights Imaging* 2025; 16(1): 88. <https://doi.org/10.1186/s13244-025-01962-2>.
16. Madanay F, O'Donohue LS, Zikmund-Fisher BJ. Patient Reactions to Artificial Intelligence-Clinician Discrepancies: Web-Based Randomized Experiment. *J Med Internet Res* 2025; 27: e68823. <https://doi.org/10.2196/68823>.
17. Michalski N, Reis M, Tetzlaff F et al. German Index of Socioeconomic Deprivation (GISD): Revision, Aktualisierung und Anwendungsbeispiele. *Journal of Health Monitoring* 2022; 7. <https://doi.org/10.25646/10640>.
18. Multusch MM, Hansen L, Heinrich MP et al. Impact of Radiologist Experience on AI Annotation Quality in Chest Radiographs: A Comparative Analysis. *Diagnostics (Basel)* 2025; 15(6). <https://doi.org/10.3390/diagnostics15060777>.
19. Potter AL, Senthil P, Srinivasan D et al. Persistent race- and sex-based disparities in lung cancer screening eligibility. *J Thorac Cardiovasc Surg* 2024; 168(1): 248-260.e2. <https://doi.org/10.1016/j.jtcvs.2023.10.025>.
20. Potter AL, Xu NN, Senthil P et al. Pack-Year Smoking History: An Inadequate and Biased Measure to Determine Lung Cancer Screening Eligibility. *J Clin Oncol* 2024; 42(17): 2026-2037. <https://doi.org/10.1200/JCO.23.01780>.
21. Robert Koch-Institut. Krebs in Deutschland für 2021/2023 [online]. 2025 [Zugriff: 12.01.2026]. URL: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/krebs_in_deutschland_2025.pdf?blob=publicationFile.
22. Salman M, Cotton A, Humphrey A et al. Patient barriers and facilitators to lung cancer screening uptake and intention to screen: a systematic review using the Theoretical Domains Framework. *BMJ Open Respir Res* 2025; 12(1). <https://doi.org/10.1136/bmjresp-2024-003127>.
23. Shah SK, Krishnan V, Khan AA et al. Women are Underrepresented in Non-small Cell Lung Cancer Clinical Trials: A Systematic Review. *Ann Surg Oncol* 2024; 31(10): 6673-6679. <https://doi.org/10.1245/s10434-024-15720-z>.

24. Starker A, Schienkiewitz A, Damerow S et al. Verbreitung von Adipositas und Rauchen bei Erwachsenen in Deutschland – Entwicklung von 2003 bis 2023. J Health Monit 2025; 10(1): e12990.
25. van der Aalst CM, Ten Haaf K, de Koning HJ. Implementation of lung cancer screening: what are the main issues? Transl Lung Cancer Res 2021; 10(2): 1050-1063. <https://doi.org/10.21037/tlcr-20-985>.
26. Vogel-Claussen J, Lasch F, Bollmann BA et al. Design and Rationale of the HANSE Study: A Holistic German Lung Cancer Screening Trial Using Low-Dose Computed Tomography. Rofo 2022; 194(12): 1333-1345. <https://doi.org/10.1055/a-1853-8291>.
27. Walsh JME, Karliner L, Smith A et al. LungCARE: Encouraging Shared Decision-Making in Lung Cancer Screening-a Randomized Trial. J Gen Intern Med 2023; 38(14): 3115-3122. <https://doi.org/10.1007/s11606-023-08189-1>.
28. Warner ET, Lathan CS. Race and sex differences in patient provider communication and awareness of lung cancer screening in the health information National Trends Survey, 2013–2017. Prev Med 2019; 124: 84-90. <https://doi.org/10.1016/j.ypmed.2019.05.001>.
29. Warner ET, Revette A, Restrepo E et al. Women's Information Needs and Educational Preferences Regarding Lung Cancer Screening. J Womens Health (Larchmt) 2024; 33(3): 318-327. <https://doi.org/10.1089/jwh.2023.0429>.
30. Wilder FG, Cangut B, Jindani R et al. Lung cancer screening among minority groups: Identifying gaps in screening and opportunities for intervention. JTCVS Open 2024; 21: 341-348. <https://doi.org/10.1016/j.xjon.2024.07.001>.
31. Zentrum für Krebsregisterdaten. Lungenkrebs (Bronchialkarzinom); ICD-10 C33–C34 [online]. 2024 [Zugriff: 28.10.2025]. URL: <https://www.krebsdaten.de/Krebs/DE/Content/Krebsarten/Lungenkrebs/lungenkrebs.html>.

A7.3.3 Liste der eingeschlossenen Publikationen zu rechtlichen Aspekten

1. Botta J. (K)ein Recht auf Behandlung mit KI? Der Zugang zu intelligenten Medizinprodukten im Lichte des Medizin-, Sozial- und Verfassungsrechts. In: Ruschemeier H, Steinrötter B (Ed). Der Einsatz von KI & Robotik in der Medizin; Interdisziplinäre Fragen. Baden-Baden: Nomos Verlagsgesellschaft; 2024. S. 115-134.
2. Brönneke JB, Hofmann B, Bakke Lysdal K et al. Guidance to assess legal aspects [online]. 2016 [Zugriff: 29.10.2024]. URL: <https://validatehta.eu/wp-content/uploads/2022/07/3.-Guidance-for-assessing-effectiveness-economic-aspects-ethical-aspects-socio-cultural-aspects-and-legal-aspects-in-complex-technologies.pdf>.
3. Bundesärztekammer. Stellungnahme „Künstliche Intelligenz in der Medizin“. Dtsch Arztebl 2025; 122(4): 1-14. https://doi.org/10.3238/arztebl.2025.Stellungnahme_KI_Medizin.

4. Bundesärztekammer. Künstliche Intelligenz in der Gesundheitsversorgung [online]. 2025 [Zugriff: 07.01.2026]. URL: https://www.bundesaerztekammer.de/fileadmin/user_upload/BAEK/Themen/Digitalisierung/Thesenpapier_KI_in_der_Gesundheitsversorgung_03.2025.pdf.
5. Bundesinstitut für Arzneimittel und Medizinprodukte. Der EHDS erklärt; Eine Einführung in den Europäischen Gesundheitsdatenraum (EHDS) [online]. [Zugriff: 13.11.2025]. URL: https://www.bfarm.de/DE/Das-BfArM/Aufgaben/Datenzugangs-und-Koordinierungsstelle/Europaeischer-Gesundheitsdatenraum/EHDS-erklaert/_node.html.
6. Bundesministerium der Justiz und für Verbraucherschutz. Produkthaftungsgesetz vom 15. Dezember 1989 (BGBl. I S. 2198), das zuletzt durch Artikel 5 des Gesetzes vom 17. Juli 2017 (BGBl. I S. 2421) geändert worden ist [online]. 2017 [Zugriff: 08.01.2026]. URL: <https://www.gesetze-im-internet.de/prodhaftg/ProdHaftG.pdf>.
7. Bundesministerium der Justiz und für Verbraucherschutz. Arzneimittelgesetz in der Fassung der Bekanntmachung vom 12. Dezember 2005 (BGBl. I S. 3394), das zuletzt durch Artikel 2 des Gesetzes vom 23. Oktober 2024 (BGBl. 2024 I Nr. 324) geändert worden ist [online]. 2024 [Zugriff: 07.01.2026]. URL: https://www.gesetze-im-internet.de/amg_1976/AMG.pdf.
8. Bundesministerium der Justiz und für Verbraucherschutz. Bundesdatenschutzgesetz vom 30. Juni 2017 (BGBl. I S. 2097), das zuletzt durch Artikel 7 des Gesetzes vom 6. Mai 2024 (BGBl. 2024 I Nr. 149) geändert worden ist [online]. 2024 [Zugriff: 08.01.2026]. URL: https://www.gesetze-im-internet.de/bdsg_2018/BDSG.pdf.
9. Bundesministerium der Justiz und für Verbraucherschutz. Gesetz zur verbesserten Nutzung von Gesundheitsdaten [online]. 2024 [Zugriff: 13.11.2025]. URL: <https://www.recht.bund.de/bgbl/1/2024/102/VO.html>.
10. Bundesministerium der Justiz und für Verbraucherschutz. Bürgerliches Gesetzbuch in der Fassung der Bekanntmachung vom 2. Januar 2002 (BGBl. I S. 42, 2909; 2003 I S. 738), das zuletzt durch Artikel 4 des Gesetzes vom 22. Dezember 2025 (BGBl. 2025 I Nr. 364) geändert worden ist [online]. 2025 [Zugriff: 10.12.2025]. URL: <https://www.gesetze-im-internet.de/bgb/BGB.pdf>.
11. Bundesministerium der Justiz und für Verbraucherschutz. Das Fünfte Buch Sozialgesetzbuch – Gesetzliche Krankenversicherung – (Artikel 1 des Gesetzes vom 20. Dezember 1988, BGBl. I S. 2477, 2482), das zuletzt durch Artikel 3a des Gesetzes vom 22. Dezember 2025 (BGBl. 2025 I Nr. 371) geändert worden ist [online]. 2025 [Zugriff: 23.12.2025]. URL: https://www.gesetze-im-internet.de/sgb_5/SGB_5.pdf.

12. Bundesministerium der Justiz und für Verbraucherschutz. Entwurf eines Gesetzes zur Modernisierung des Produkthaftungsrechts; Referentenentwurf [online]. 2025 [Zugriff: 10.12.2025]. URL: https://bundestagszusammenfasser.de/wp-content/uploads/rewp21/1052_RefE_Produkthaftung.pdf.
13. Bundesministerium für Gesundheit. Gesetz zur Beschleunigung der Digitalisierung des Gesundheitswesens (Digital-Gesetz – DigiG) [online]. 2025 [Zugriff: 13.11.2025]. URL: <https://www.recht.bund.de/bgbl/1/2024/101/VO.html>.
14. Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz. Verordnung über die Zulässigkeit der Anwendung der Niedrigdosis-Computertomographie zur Früherkennung von Lungenkrebs bei rauchenden Personen (Lungenkrebs-Früherkennungs-Verordnung – LuKrFrühErkV). Bundesgesetzblatt Teil 1 2024; (162): 1-5.
15. Crowson MG, Moukheiber D, Arévalo AR et al. A systematic review of federated learning applications for biomedical data. PLOS Digit Health 2022; 1(5): e0000033. <https://doi.org/10.1371/journal.pdig.0000033>.
16. Deutscher Ethikrat. Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz; Stellungnahme [online]. 2023 [Zugriff: 07.01.2026]. URL: <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>.
17. Eichelberger J. Arzthaftung im Kontext von KI und Robotik. In: Ruschemeier H, Steinrötter B (Ed). Der Einsatz von KI & Robotik in der Medizin; Interdisziplinäre Fragen. Baden-Baden: Nomos Verlagsgesellschaft; 2024. S. 27-52.
18. Europäisches Parlament, Rat der Europäischen Union. Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung) [online]. 2016 [Zugriff: 07.01.2026]. URL: <http://data.europa.eu/eli/reg/2016/679/oj>.
19. Europäisches Parlament, Rat der Europäischen Union. Verordnung (EU) 2017/745 des Europäischen Parlaments und des Rates vom 5. April 2017 über Medizinprodukte, zur Änderung der Richtlinie 2001/83/EG, der Verordnung (EG) Nr. 178/2002 und der Verordnung (EG) Nr. 1223/2009 und zur Aufhebung der Richtlinien 90/385/EWG und 93/42/EWG des Rates [online]. 2017 [Zugriff: 07.01.2026]. URL: <http://data.europa.eu/eli/reg/2017/745/oj>.

20. Europäisches Parlament, Rat der Europäischen Union. Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz) [online]. 2024 [Zugriff: 07.01.2026]. URL: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj?eliuri=eli%3Areg%3A2024%3A1689%3Aoj&locale=de>.
21. Europäisches Parlament, Rat der Europäischen Union. Verordnung (EU) 2025/327 des Europäischen Parlaments und des Rates vom 11. Februar 2025 über den europäischen Gesundheitsdatenraum sowie zur Änderung der Richtlinie 2011/24/EU und der Verordnung (EU) 2024/2847 [online]. 2025 [Zugriff: 13.11.2025]. URL: <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32025R0327>.
22. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Krebsfrüherkennungs-Richtlinie (KFE-RL): Einführung der Lungenkrebsfrüherkennung mittels Niedrigdosis-Computertomographie bei Rauchern [online]. 2024 [Zugriff: 23.10.2025]. URL: https://www.g-ba.de/downloads/39-261-7301/2025-06-18_KFE-RL_Einf-Lungenkrebsfrueherkennung-Niedrigdosis-CT-Raucher_BAnz.pdf.
23. Herbst T. Herausforderungen an das Datenschutzrecht durch den Einsatz von Künstlicher Intelligenz in der Medizin. In: Ruschemeier H, Steinrötter B (Ed). Der Einsatz von KI & Robotik in der Medizin; Interdisziplinäre Fragen. Baden-Baden: Nomos Verlagsgesellschaft; 2024. S. 77-96.
24. Katzenmeier C. Haftung für Schäden durch KI in der Medizin. In: Grundmann T, Hey J, Hobe S et al (Ed). Die Macht der Algorithmen. Baden-Baden: Nomos Verlagsgesellschaft; 2023. S. 73-102.
25. Konečný J, McMahan HB, Yu FX et al. Federated Learning; Strategies for Improving Communication Efficiency [online]. 2017 [Zugriff: 07.01.2026]. URL: <https://doi.org/10.48550/arXiv.1610.05492>.
26. Lee M, Hwang EJ, Lee JH et al. Artificial Intelligence for Low-Dose CT Lung Cancer Screening: Comparison of Utilization Scenarios. AJR 2025. <https://doi.org/10.2214/ajr.25.32829>.
27. Lo SB, Freedman MT, Gillis LB et al. JOURNAL CLUB: Computer-Aided Detection of Lung Nodules on CT With a Computerized Pulmonary Vessel Suppressed Function. AJR Am J Roentgenol 2018; 210(3): 480–488. <https://doi.org/10.2214/ajr.17.18718>.

28. Luckner S, Lauer W. [Regulatory classification of AI-enabled products for medical use on the basis of the EU AI Act and MDR/IVDR]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2025; 68(8): 854-861. <https://doi.org/10.1007/s00103-025-04091-9>.
29. Medical Device Coordination Group. Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR [online]. 2025 [Zugriff: 10.12.2025]. URL: https://health.ec.europa.eu/document/download/b45335c5-1679-4c71-a91c-fc7a4d37f12b_en?filename=mdcg_2019_11_en.pdf.
30. Medizinischer Dienst Bund. Künstliche Intelligenz in Medizinprodukten: Neue Herausforderungen durch den „AI-Act“ der EU [online]. 2025 [Zugriff: 08.01.2026]. URL: https://md-bund.de/fileadmin/dokumente/Medizinprodukte_Artikel/2025-05-05_Kuenstliche_Intelligenz_in_Medizinprodukten.pdf.
31. Mühlhoff R. Das Risiko der Sekundärnutzung trainierter Modelle als zentrales Problem von Datenschutz und KI-Regulierung im Medizinbereich. In: Ruschemeier H, Steinrötter B (Ed). Der Einsatz von KI & Robotik in der Medizin; Interdisziplinäre Fragen. Baden-Baden: Nomos Verlagsgesellschaft; 2024. S. 27-52.
32. Zentrale Kommission zur Wahrung ethischer Grundsätze in der Medizin und ihren Grenzgebieten (Zentrale Ethikkommission) bei der Bundesärztekammer. Entscheidungsunterstützung ärztlicher Tätigkeit durch Künstliche Intelligenz. Dtsch Arztebl 2021; 118: A1-A13.

A7.3.4 Liste der eingeschlossenen Publikationen zu organisatorischen Aspekten

1. Akudjedu TN, Torre S, Khine R et al. Knowledge, perceptions, and expectations of Artificial intelligence in radiography practice: A global radiography workforce survey. J Med Imaging Radiat Sci 2023; 54(1): 104-116. <https://doi.org/10.1016/j.jmir.2022.11.016>.
2. Alderman JE, Palmer J, Laws E et al. Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together consensus recommendations. Lancet Digit Health 2025; 7(1): e64-e88. [https://doi.org/10.1016/s2589-7500\(24\)00224-3](https://doi.org/10.1016/s2589-7500(24)00224-3).
3. Blum TG, Vogel-Claussen J, Andreas S et al. [Statement paper on the implementation of a national organized program in Germany for the early detection of lung cancer in risk populations using low-dose CT screening including management of screening findings]. Pneumologie 2024; 78(1): 15-34. <https://doi.org/10.1055/a-2175-4580>.
4. Bundesamt für Strahlenschutz. Lungenkrebsfrüherkennung mittels Niedrigdosis-Computertomographie; Wissenschaftliche Bewertung des Bundesamtes für Strahlenschutz gemäß § 84 Absatz 3 Strahlenschutzgesetz; Bericht [online]. 2023 [Zugriff: 22.10.2025]. URL: <https://doris.bfs.de/jspui/bitstream/urn:nbn:de:0221-2021082028027/5/BfS-35-21.pdf>.

5. Bundesärztekammer. Stellungnahme „Künstliche Intelligenz in der Medizin“. Dtsch Arztebl 2025; 122(4): 1-14. https://doi.org/10.3238/arztebl.2025.Stellungnahme_KI_Medizin.
6. Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz. Verordnung über die Zulässigkeit der Anwendung der Niedrigdosis-Computertomographie zur Früherkennung von Lungenkrebs bei rauchenden Personen (Lungenkrebs-Früherkennungs-Verordnung – LuKrFrühErkV). Bundesgesetzblatt Teil 1 2024; (162): 1-5.
7. Bütke T, Steinert J, Sternberg H. Keine KI in der Medizin ohne die Patienten. Akademie Aktuell 2024; 3: 23-25.
8. Deutsche Krebsgesellschaft. Kennzahlenauswertung 2025; Jahresbericht der zertifizierten Lungenkrebszentren; Auditjahr 2024 / Kennzahlenjahr 2023 [online]. 2024 [Zugriff: 23.10.2025]. URL: https://www.krebsgesellschaft.de/unsere-themen/zertifizierung/jahresberichte-zertifizierter-zentren?file=files/content/unsere-themen/zertifizierung/jahresberichte/lungenkrebszentren/qualitaetsindikatoren_lungenkrebs_2025-a1_250804.pdf
9. Deutscher Ethikrat. Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz; Stellungnahme [online]. 2023 [Zugriff: 07.01.2026]. URL: <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>.
10. European Society for Thoracic Imaging. ESTI Lung Cancer Screening Certification Project [online]. [Zugriff: 20.11.2025]. URL: <https://www.myesti.org/lungcancerscreeningcertificationproject/>.
11. Ewals LJS, Heesterbeek LJJ, Yu B et al. The Impact of Expectation Management and Model Transparency on Radiologists' Trust and Utilization of AI Recommendations for Lung Nodule Assessment on Computed Tomography: Simulated Use Study. JMIR AI 2024; 3: e52211. <https://doi.org/10.2196/52211>.
12. Ewals LJS, van der Wulp K, van den Borne B et al. The Effects of Artificial Intelligence Assistance on the Radiologists' Assessment of Lung Nodules on CT Scans: A Systematic Review. J Clin Med 2023; 12(10). <https://doi.org/10.3390/jcm12103536>.
13. Farič N, Hinder S, Williams R et al. Early experiences of integrating an artificial intelligence-based diagnostic decision support system into radiology settings: a qualitative study. J Am Med Inform Assoc 2023; 31(1): 24-34. <https://doi.org/10.1093/jamia/ocad191>.
14. GKV-Spitzenverband. Verbesserung und Stärkung der ambulanten Versorgung; Positionspapier [online]. 2024 [Zugriff: 23.10.2025]. URL: https://www.gkv-spitzenverband.de/media/dokumente/service_1/publikationen/20240715_PoPa_Ambulante_Versorgung_barrierefrei.pdf.

15. Gleeson F, Revel MP, Biederer J et al. Implementation of artificial intelligence in thoracic imaging-a what, how, and why guide from the European Society of Thoracic Imaging (ESTI). *Eur Radiol* 2023; 33(7): 5077-5086. <https://doi.org/10.1007/s00330-023-09409-2>.
16. Hadjiiski L, Cha K, Chan HP et al. AAPM task group report 273: Recommendations on best practices for AI and machine learning for computer-aided diagnosis in medical imaging. *Med Phys* 2023; 50(2): e1-e24. <https://doi.org/10.1002/mp.16188>.
17. Hahn HK, May MS, Dicken V et al. Anforderungen an die Qualitätssicherung von KI-Modellen für die Lungenkrebs-Früherkennung [online]. 2025 [Zugriff: 07.01.2026]. URL: <https://doi.org/10.48550/arXiv.2502.17639>.
18. Haserück A. Blickwinkel: Künstliche Intelligenz und Radiologie. *Dtsch Arztebl International* 2024; 121(21): 8626.
19. Jungmann F. Künstliche Intelligenz in der Radiologie. *Akademie Aktuell* 2024; 3: 26-31.
20. Kauczor HU, von Stackelberg O, Nischwitz E et al. Strengthening lung cancer screening in Europe: fostering participation, improving outcomes, and addressing health inequalities through collaborative initiatives in the SOLACE consortium. *Insights Imaging* 2024; 15(1): 252. <https://doi.org/10.1186/s13244-024-01814-5>.
21. Kocak B, Ponsiglione A, Romeo V et al. Radiology AI and sustainability paradox: environmental, economic, and social dimensions. *Insights Imaging* 2025; 16(1): 88. <https://doi.org/10.1186/s13244-025-01962-2>.
22. Multusch MM, Hansen L, Heinrich MP et al. Impact of Radiologist Experience on AI Annotation Quality in Chest Radiographs: A Comparative Analysis. *Diagnostics (Basel)* 2025; 15(6). <https://doi.org/10.3390/diagnostics15060777>.
23. Tejani AS, Klontzas ME, Gatti AA et al. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update. *Radiol Artif Intell* 2024; 6(4): e240300. <https://doi.org/10.1148/ryai.240300>.
24. van Leeuwen KG, de Rooij M, Schalekamp S et al. Clinical use of artificial intelligence products for radiology in the Netherlands between 2020 and 2022. *Eur Radiol* 2024; 34(1): 348-354. <https://doi.org/10.1007/s00330-023-09991-5>.
25. van Leeuwen KG, Schalekamp S, Rutten M et al. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. *Eur Radiol* 2021; 31(6): 3797-3804. <https://doi.org/10.1007/s00330-021-07892-z>.
26. Vogel-Claussen J, Lasch F, Bollmann BA et al. Design and Rationale of the HANSE Study: A Holistic German Lung Cancer Screening Trial Using Low-Dose Computed Tomography. *Rofo* 2022; 194(12): 1333-1345. <https://doi.org/10.1055/a-1853-8291>.

27. Zentrum für Krebsregisterdaten. Lungenkrebs (Bronchialkarzinom); ICD-10 C33–C34 [online]. 2024 [Zugriff: 28.10.2025]. URL: <https://www.krebsdaten.de/Krebs/DE/Content/Krebsarten/Lungenkrebs/lungenkrebs.html>.

A7.3.5 Liste der eingeschlossenen Publikationen zu Umwelt- und Klimaaspekten

1. Doo FX, Vosshenrich J, Cook TS et al. Environmental Sustainability and AI in Radiology: A Double-Edged Sword. *Radiology* 2024; 310(2): e232030. <https://doi.org/10.1148/radiol.232030>.
2. Joshi G, Jain A, Araveeti SR et al. FDA-Approved Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices: An Updated Landscape. *Electronics* 2024; 13(3): 498. <https://doi.org/10.3390/electronics13030498>.
3. Karliner J, Slotterback S, Boyd R et al. Health care's climate footprint: the health sector contribution and opportunities for action. *Eur J Public Health* 2020; 30(Suppl 5). <https://doi.org/10.1093/eurpub/ckaa165.843>.
4. Kocak B, Ponsiglione A, Romeo V et al. Radiology AI and sustainability paradox: environmental, economic, and social dimensions. *Insights Imaging* 2025; 16(1): 88. <https://doi.org/10.1186/s13244-025-01962-2>.
5. Koronen C, Åhman M, Nilsson LJ. Data centres in future European energy systems; energy efficiency, integration and policy. *Energy Efficiency* 2020; 13(1): 129-144. <https://doi.org/10.1007/s12053-019-09833-8>.
6. Muehlematter UJ, Bluethgen C, Vokinger KN. FDA-cleared artificial intelligence and machine learning-based medical devices and their 510(k) predicate networks. *Lancet Digit Health* 2023; 5(9): e618-e626. [https://doi.org/10.1016/S2589-7500\(23\)00126-7](https://doi.org/10.1016/S2589-7500(23)00126-7).
7. Ren S. How much water does AI consume? The public deserves to know [online]. 2023 [Zugriff: 30.10.2025]. URL: <https://oecd.ai/en/wonk/how-much-water-does-ai-consume>.
8. Roletto A, Zanardo M, Bonfitto GR et al. The environmental impact of energy consumption and carbon emissions in radiology departments: a systematic review. *Eur Radiol Exp* 2024; 8(1): 35. <https://doi.org/10.1186/s41747-024-00424-6>.
9. Thangam D, Muniraju H, Ramesh R et al. Impact of Data Centers on Power Consumption, Climate Change, and Sustainability. In: Kumar KD, Varadarajan V, Nasser N et al (Ed). *Computational Intelligence for Green Cloud Computing and Digital Waste Management*. Hershey: IGI Global Scientific Publishing; 2024. S. 60-83.
10. Ueda D, Walston SL, Fujita S et al. Climate change and artificial intelligence in healthcare: Review and recommendations towards a sustainable future. *Diagnostic and Interventional Imaging* 2024; 105(11): 453-459. <https://doi.org/10.1016/j.diii.2024.06.002>.

A8 Suchstrategien

A8.1 Suchstrategien zur Nutzenbewertung

A8.1.1 Bibliografische Datenbanken

Suche nach systematischen Übersichten

1. PubMed

Suchoberfläche: NLM

#	Searches
#1	Search: lung OR pulmonary OR thorax OR chest
#2	Search: medical imaging OR diagnostic imaging OR x-ray OR CT OR computer tomograph* OR radiography OR radiographs
#3	Search: "Artificial Intelligence"[Mesh] OR (AI[tw] OR "artificial intelligence"[tw] OR AIVI[tw] OR "classification algorithm*" [tw] OR "computer heuristic*" [tw] OR "convolutional network*" [tw] OR DALL-E[tw] OR "decision support system*" [tw] OR "decision tree"[tw] OR DeepAI[tw] OR "deep learning"[tw] OR "data science"[tw] OR "feature detection"[tw] OR "generative pre-trained transformer"[tw] OR "generative pretrained transformer"[tw] OR Invideo[tw] OR "language learning model*" [tw] OR "large language model*" [tw] OR "learning algorithm*" [tw] OR "machine learning"[tw] OR (Markov[tw] AND model*[tw]) OR Midjourney[tw] OR ((multifactor*[tw] OR multicriteria[tw]) AND ("decision analysis"[tw] OR "decision making"[tw])) OR "natural language process*" [tw] OR "nearest neighbo*" [tw] OR "neural network*" [tw] OR "outlier detection"[tw] OR "pattern recognition"[tw] OR Perplexity[tw] OR "probability tree"[tw] OR "random forest"[tw] OR "representation learning"[tw] OR "Runway AI"[tw] OR "Runway Gen-1"[tw] OR "Stable Diffusion"[tw] OR "support vector machine*" [tw] OR "transfer learning"[tw] OR "Bing chat"[tw] OR ChatGPT*[tw] OR "Chat GPT"[tw] OR "Google* Bard"[tw] OR "Google* Gemini"[tw] OR "IBM Watson"[tw] OR "Microsoft* Bing"[tw] OR "Microsoft* Copilot"[tw] OR OpenAI[tw] OR "Open AI"[tw] OR PathAI[tw] OR "Path AI"[tw])
#4	Search: #1 AND #2 AND #3
#5	Search: #4 AND #2 AND #3 AND systematic[sb]

2. International HTA Database

Suchoberfläche: INAHTA

#	Searches
#1	lung OR pulmonary OR thorax OR chest
#2	medical imaging OR diagnostic imaging OR x-ray OR CT OR computer tomograph* OR radiography OR radiographs
#3	(AI OR "artificial intelligence" OR AIVI OR "classification algorithm*" OR "computer heuristic*" OR "convolutional network*" OR DALL-E OR "decision support system*" OR "decision tree" OR DeepAI OR "deep learning" OR "data science" OR "feature detection" OR "generative pre-trained transformer" OR "generative pretrained transformer" OR Invideo OR "language learning model*" OR "large language model*" OR "learning algorithm*" OR "machine learning" OR (Markov AND model*) OR Midjourney OR ((multifactor* OR multicriteria) AND ("decision analysis" OR "decision making")) OR "natural language process*" OR "nearest neighbo*" OR "neural network*" OR "outlier detection" OR "pattern recognition" OR Perplexity OR "probability tree" OR "random forest" OR "representation learning" OR "Runway AI" OR "Runway Gen-1" OR "Stable Diffusion" OR "support vector machine*" OR "transfer learning" OR "Bing chat" OR ChatGPT* OR "Chat GPT" OR "Google* Bard" OR "Google* Gemini" OR "IBM Watson" OR "Microsoft* Bing" OR "Microsoft* Copilot" OR OpenAI OR "Open AI" OR PathAI OR "Path AI")
#4	#3 AND #2 AND #1

Suche nach Primärstudien

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to August 28, 2025

Es wurde folgender Filter übernommen:

- AI-Interventionen: Ayiku 2025 [185]- NICE artificial intelligence (AI) medical device intervention search filters for MEDLINE and Embase (Ovid)

#	Searches
1	exp Lung Neoplasms/
2	(lung adj1 (cancer* or tumo?r* or neoplasm* or adenocarcinoma* or carcinoma*)).ab,ti.
3	((lung or pulmonary or chest) and (nodule* or lesion*)).ti,ab.
4	or/1-3
5	radiologist*.ti,ab.
6	(clinician* or doctor* or physician* or resident* or specialist* or observer* or reader*).ti,ab.
7	or/5-6
8	and/4,7
9	algorithm*.ti,kf.
10	(algorithm* adj2 (learn* or automate* or detect* or predict* or treatment* or therap* or radiolog* or AI or DL or data or dataset* or base* or classif*)).ab.

#	Searches
11	artificial intelligen*.ti,ab,kf.
12	AI.ti,kf.
13	(machine adj2 learn*).ti,ab,kf.
14	machinelearn*.ti,ab,kf.
15	(deep adj2 learn*).ti,ab,kf.
16	deeplearn*.ti,ab,kf.
17	neural network*.ti,ab,kf.
18	(convolutional adj1 network*).ti,ab,kf.
19	automate*.ti.
20	(automate* adj3 (system* or score* or software* or analysis* or analyse* or risk* or evaluat* or tool* or detect* or process*)).ab,kf.
21	(vector machine* or svm*).ti,ab,kf.
22	radiomic*.ti,ab,kf.
23	((supervised or unsupervised) adj3 (classifier* or prediction*)).ti,ab,kf.
24	or/9-23
25	((computer-assisted or computer-aided) adj3 (detect* or system*)).ti,ab.
26	or/24-25
27	and/8,26
28	(animals/ not humans/) or comment/ or editorial/ or exp review/ or meta analysis/ or consensus/ or exp guideline/
29	hi.fs. or case report.mp.
30	or/28-29
31	27 not 30
32	31 and (english or german or multilingual or undetermined).lg.

2. Embase

Suchoberfläche: Ovid

- Embase 1974 to 2025 May 30

Es wurden folgende Filter übernommen:

- RCT: Wong [186] – Strategy minimizing difference between sensitivity and specificity
- AI-Interventionen: Ayiku 2025 [185]- NICE artificial intelligence (AI) medical device intervention search filters for MEDLINE and Embase (Ovid)

#	Searches
1	exp Lung Tumor/ or Lung Nodule/
2	(lung adj1 (cancer* or tumo?r* or neoplasm* or adenocarcinoma* or carcinoma*)).ab,ti.
3	((lung or pulmonary or chest) and (nodule* or lesion*)).ti,ab.

#	Searches
4	or/1-3
5	radiologist*.ti,ab.
6	(clinician* or doctor* or physician* or resident* or specialist* or observer* or reader*).ti,ab.
7	or/5-6
8	and/4,7
9	algorithm*.ti,kf.
10	(algorithm* adj2 (learn* or automate* or detect* or predict* or treatment* or therap* or radiolog* or AI or DL or data or dataset* or base* or classif*)).ab.
11	artificial intelligen*.ti,ab,kf.
12	AI.ti,kf.
13	(machine adj2 learn*).ti,ab,kf.
14	machinelearn*.ti,ab,kf.
15	(deep adj2 learn*).ti,ab,kf.
16	deeplearn*.ti,ab,kf.
17	neural network*.ti,ab,kf.
18	(convolutional adj1 network*).ti,ab,kf.
19	automate*.ti.
20	(automate* adj3 (system* or score* or software* or analysis* or analyse* or risk* or evaluat* or tool* or detect* or process*)).ab,kf.
21	(vector machine* or svm*).ti,ab,kf.
22	radiomic*.ti,ab,kf.
23	((supervised or unsupervised) adj3 (classifier* or prediction*)).ti,ab,kf.
24	or/9-23
25	((computer-assisted or computer-aided) adj3 (detect* or system*)).ti,ab.
26	or/24-25
27	and/8,26
28	27 not medline.cr.
29	28 not (exp animal/ not exp human/)
30	29 not (Conference Abstract or Conference Review or Editorial).pt.
31	30 not ((afrikaans or albanian or arabic or armenian or azerbaijani or basque or belorussian or bosnian or bulgarian or catalan or chinese or croatian or czech or danish or dutch or english or esperanto or estonian or finnish or french or gallegan or georgian or german or greek or hebrew or hindi or hungarian or icelandic or indonesian or irish gaelic or italian or japanese or korean or latvian or lithuanian or macedonian or malay or norwegian or persian or polish or polyglot or portuguese or pushto or romanian or russian or scottish gaelic or serbian or slovak or slovene or spanish or swedish or thai or turkish or ukrainian or urdu or uzbek or vietnamese) not (english or german)).lg.
32	(random* or double-blind*).tw.
33	placebo*.mp.
34	or/32-33
35	and/31,34

3. The Cochrane Library

Suchoberfläche: Wiley

- Cochrane Central Register of Controlled Trials Issue 6 of 12, June 2025

Es wurde folgender Filter übernommen:

- AI-Interventionen (adaptiert für Central): Ayiku 2025 [185]- NICE artificial intelligence (AI) medical device intervention search filters for MEDLINE and Embase (Ovid)

#	Searches
#1	[mh "Lung Neoplasms"]
#2	(lung NEAR/1 (cancer* OR tumor* OR tumour OR neoplasm* OR adenocarcinoma* OR carcinoma*)):ti,ab
#3	((lung or pulmonary or chest) and (nodule* or lesion*)):ti,ab
#4	#1 OR #2 OR #3
#5	radiologist*:ti,ab
#6	(clinician* or doctor* or physician* or resident* or specialist* or observer* or reader*):ti,ab
#7	#5 OR #6
#8	#4 AND #7
#9	algorithm*
#10	(artificial NEXT intelligen*)
#11	AI
#12	(machine NEAR/2 learn*)
#13	machinelearn*
#14	(deep NEAR/2 learn*)
#15	deeplearn*
#16	(neural NEXT network*)
#17	(convolutional NEAR/1 network*)
#18	automate*
#19	(automate* NEAR/3 (system* or score* or software* or analysis* or analyse* or risk* or evaluat* or tool* or detect* or process*))
#20	(vector machine* or svm*)
#21	radiomic*
#22	((supervised or unsupervised) NEAR/3 (classifier* or prediction*))
#23	#9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22
#24	((computer-assisted or computer-aided) NEAR/3 (detect* or system*))
#25	#23 OR #24
#26	#8 AND #25
#27	#26 not (*clinicaltrial*gov* or *trialsearch*who* or *clinicaltrialsregister*eu* or *anzctr*org*au* or *trialregister*nl* or *irct*ir* or *isrctn* or *controlled*trials*com* or *drks*de*):so

#	Searches
#28	#27 not ((language next (afr or ara or aze or bos or bul or car or cat or chi or cze or dan or dut or es or est or fin or fre or gre or heb or hrv or hun or ice or ira or ita or jpn or ko or kor or lit or nor or peo or per or pol or por or pt or rom or rum or rus or slo or slv or spa or srp or swe or tha or tur or ukr or urd or uzb))) not (language near/2 (en or eng or english or ger or german or mul or unknown)))
#29	#28 in Trials

A8.1.2 Studienregister

1. ClinicalTrials.gov

Anbieter: U.S. National Institutes of Health

- URL: <http://www.clinicaltrials.gov>
- Eingabeoberfläche: Expert Search

Suchstrategie
AREA[ConditionSearch](lung cancer OR lung nodule) AND AREA[BasicSearch]((radiologist OR clinician OR doctor OR physician OR resident OR specialist OR observer OR reader) AND (artificial intelligence OR AI OR computer-assisted detection OR computer-aided detection OR automatic detection OR deep learning OR machine learning OR neural network OR classifier OR support vector machine OR SVM OR prediction OR radiomic OR algorithm))

2. International Clinical Trials Registry Platform Search Portal

Anbieter: World Health Organization

- URL: <https://trialsearch.who.int>
- Eingabeoberfläche: Basic Search

Suchstrategie
(lung OR ((pulmonary OR chest) AND nodule*)) AND (radiologist* OR clinician* OR doctor* OR physician* OR resident* OR specialist* OR observer* OR reader*) AND (artificial intelligence OR AI OR computer-assisted detection OR computer-aided detection OR automat* OR deep learning OR machine learning OR neural network* OR classifier* OR support vector machine OR SVM OR radiomic* OR algorithm*)

A8.2 Suchstrategien zur gesundheitsökonomischen Bewertung

1. MEDLINE

Suchoberfläche: Ovid

Ovid MEDLINE(R) ALL 1946 to August 06, 2025

Es wurden folgende Filter übernommen:

- Gesundheitsökonomische Studie: Glanville 2009 [187]- Emory University (Grady) filter
- AI-Interventionen: Ayiku 2025 [185]- NICE artificial intelligence (AI) medical device intervention search filters for MEDLINE and Embase (Ovid)

#	Searches
1	exp Lung Neoplasms/
2	(lung adj1 (cancer* or tumor* or neoplasm* or adenocarcinoma* or carcinoma*)).ab,ti.
3	((lung or pulmonary or chest) and (nodule* or lesion*)).ti,ab.
4	or/1-3
5	radiologist*.ti,ab.
6	(clinician* or doctor* or physician* or resident* or specialist* or observer* or reader*).ti,ab.
7	or/5-6
8	and/4,7
9	algorithm*.ti,kf.
10	(algorithm* adj2 (learn* or automate* or detect* or predict* or treatment* or therap* or radiolog* or AI or DL or data or dataset* or base* or classif*)).ab.
11	artificial intelligen*.ti,ab,kf.
12	AI.ti,kf.
13	(machine adj2 learn*).ti,ab,kf.
14	machinelearn*.ti,ab,kf.
15	(deep adj2 learn*).ti,ab,kf.
16	deeplearn*.ti,ab,kf.
17	neural network*.ti,ab,kf.
18	(convolutional adj1 network*).ti,ab,kf.
19	automate*.ti.
20	(automate* adj3 (system* or score* or software* or analysis* or analyse* or risk* or evaluat* or tool* or detect* or process*)).ab,kf.
21	(vector machine* or svm*).ti,ab,kf.
22	radiomic*.ti,ab,kf.
23	((supervised or unsupervised) adj3 (classifier* or prediction*)).ti,ab,kf.
24	or/9-23
25	((computer-assisted or computer-aided) adj3 (detect* or system*)).ti,ab.
26	or/24-25

#	Searches
27	and/8,26
28	(economic\$ or cost\$).ti.
29	cost benefit analysis/
30	treatment outcome/ and ec.fs.
31	or/28-30
32	31 not ((animals/ not humans/) or letter.pt.)
33	and/27,32
34	33 not (comment or editorial).pt.
35	34 and (english or german or multilingual or undetermined).lg.

2. Embase

Suchoberfläche: Ovid

Embase 1974 to 2025 August 05

Es wurden folgende Filter übernommen:

- Gesundheitsökonomische Studie: Glanville [187]- Embase G filter
- AI-Interventionen: Ayiku 2025 [185]- NICE artificial intelligence (AI) medical device intervention search filters for MEDLINE and Embase (Ovid)

#	Searches
1	exp Lung Tumor/ or Lung Nodule/
2	(lung adj1 (cancer* or tumor* or neoplasm* or adenocarcinoma* or carcinoma*)).ab,ti.
3	((lung or pulmonary or chest) and (nodule* or lesion*)).ti,ab.
4	or/1-3
5	radiologist*.ti,ab.
6	(clinician* or doctor* or physician* or resident* or specialist* or observer* or reader*).ti,ab.
7	or/5-6
8	and/4,7
9	algorithm*.ti,kf.
10	(algorithm* adj2 (learn* or automate* or detect* or predict* or treatment* or therap* or radiolog* or AI or DL or data or dataset* or base* or classif*)).ab.
11	artificial intelligen*.ti,ab,kf.
12	AI.ti,kf.
13	(machine adj2 learn*).ti,ab,kf.
14	machinelearn*.ti,ab,kf.
15	(deep adj2 learn*).ti,ab,kf.
16	deeplearn*.ti,ab,kf.

#	Searches
17	neural network*.ti,ab,kf.
18	(convolutional adj1 network*).ti,ab,kf.
19	automate*.ti.
20	(automate* adj3 (system* or score* or software* or analysis* or analyse* or risk* or evaluat* or tool* or detect* or process*)).ab,kf.
21	(vector machine* or svm*).ti,ab,kf.
22	radiomic*.ti,ab,kf.
23	((supervised or unsupervised) adj3 (classifier* or prediction*)).ti,ab,kf.
24	or/9-23
25	((computer-assisted or computer-aided) adj3 (detect* or system*)).ti,ab.
26	or/24-25
27	and/8,26
28	(Cost adj effectiveness).ab.
29	(Cost adj effectiveness).ti.
30	(Life adj years).ab.
31	(Life adj year).ab.
32	Qaly.ab.
33	(Cost or costs).ab. and Controlled Study/
34	(Cost and costs).ab.
35	or/28-34
36	and/27,35
37	36 not medline.cr.
38	37 not (exp animal/ not exp human/)
39	38 not ((Conference Abstract or Conference Review or Editorial).pt. or NCT*.ui.)
40	39 and (english or german).lg.
41	remove duplicates from 40

3. International HTA Database

Suchoberfläche: INAHTA

#	Searches
#1	lung OR pulmonary OR thorax OR chest
#2	medical imaging OR diagnostic imaging OR x-ray OR CT OR computer tomograph* OR radiography OR radiographs
#3	(AI OR "artificial intelligence" OR AIVI OR "classification algorithm*" OR "computer heuristic*" OR "convolutional network*" OR DALL-E OR "decision support system*" OR "decision tree" OR DeepAI OR "deep learning" OR "data science" OR "feature detection" OR "generative pre-trained transformer" OR "generative pretrained transformer" OR Invideo OR "language learning model*" OR "large language model*" OR "learning algorithm*" OR "machine learning" OR (Markov AND model*) OR Midjourney OR ((multifactor* OR multicriteria) AND ("decision analysis" OR "decision making")) OR "natural language process*" OR "nearest neighbo*" OR "neural network*" OR "outlier detection" OR "pattern recognition" OR Perplexity OR "probability tree" OR "random forest" OR "representation learning" OR "Runway AI" OR "Runway Gen-1" OR "Stable Diffusion" OR "support vector machine*" OR "transfer learning" OR "Bing chat" OR ChatGPT* OR "Chat GPT" OR "Google* Bard" OR "Google* Gemini" OR "IBM Watson" OR "Microsoft* Bing" OR "Microsoft* Copilot" OR OpenAI OR "Open AI" OR PathAI OR "Path AI")
#4	#3 AND #2 AND #1

A9 Zentrale Ergebnisse aus den Betroffeneninterviews

Die Betroffenenengespräche fanden am 3., 7., 8. und 14.07.2025 jeweils in Einzelgesprächen mit 2 Männern (87 und 66 Jahre alt) und 3 Frauen (55, 71 und 72 Jahre alt) anhand eines Befragungsleitfadens statt. 2 der befragten Personen waren ehemalige Raucher, 2 Nichtraucher und eine Person aktuelle Raucherin. Die Betroffenen waren Personen, die aus verschiedenen Gründen eine Röntgen- oder CT-Untersuchung bereits in Anspruch genommen oder eine Überweisung für eine solche Untersuchung der Lunge erhalten haben (z. B. aktive oder vorherige Raucher aus unterschiedlichen Gründen, andere Anlässe für die Röntgen / CT-Untersuchung der Lunge, Vorliegen von Symptomen, die eine Lungenerkrankung vermuten ließen wie chronischer Husten oder Atemnot). Da in Deutschland die AI zur Unterstützung der Befundung derzeit in der Regel nicht im niedergelassenen Bereich und auch nicht in Röntgen-Zentren eingesetzt werden, sondern auf bestimmte tertiäre Settings – insbesondere Universitätskrankenhäuser – beschränkt zu sein scheinen, hatten die Betroffenen durchgehend noch keine persönlichen Erfahrungen mit dem Einsatz einer AI bei ihren Röntgen / CT-Untersuchungen. Und auch wenn die Röntgen- oder CT-Bilder der Betroffenen mit Unterstützung einer AI ausgewertet worden wären, wäre es unwahrscheinlich, dass die Betroffenen hiervon Kenntnis gehabt hätten. Denn in der Regel werden Patientinnen und Patienten in Deutschland derzeit im Vorfeld einer CT-Untersuchung der Lunge zwar über die CT-Untersuchung selbst aufgeklärt (zum Beispiel zu möglichen Risiken durch die Bestrahlung), nicht jedoch über die Einbindung einer AI-basierten Auswertungs-Software, sofern eingesetzt (diese Information beruht auf der Aussage des ins Projekt eingebundenen Kliniklers).

Die Ergebnisse der Befragung der 5 betroffenen Personen sind im Folgenden entlang der Themen des Befragungsleitfadens zusammengefasst (Zitate im Wortlaut in Klammern):

Einschätzungen zu erwarteten Vorteilen durch die AI-Unterstützung (Zitate in Klammern)

- alle Vorteile der modernen Technik / Innovation nutzen („dass ich die ganze Technik, die es gibt, nutzen würde. Ich gehe davon aus, dass das geprüft ist und alles, ne, und dass das gut ist für mich, auf jeden Fall nicht schadet.“)
- höheres Sicherheitsgefühl, eine Art 2. Meinung erhalten, die Sorge nehmen, dass der Arzt etwas übersehen könnte („auch ein höheres Sicherheitsgefühl [...] als wenn so ein zweiter mit drauf geguckt hätte [...]. wenn es auch eine künstliche Intelligenz ist. Aber es würde einem vielleicht die Sorge nehmen, dass der Arzt etwas übersehen könnte.“)
- Die AI basiert auf vielen Daten und Fallbeispielen; Die Ärztin oder der Arzt bzw. die Radiologin oder der Radiologe hat ein breites Wissen und Erfahrung: das wurde als eine gute Mischung betrachtet („das, was der Arzt nicht kann – aufgrund von Erfahrung, Tagesform etc. – kann vielleicht die AI und vice versa [...]“, „weil es letztendlich doch eine unterstützende Maßnahme ist, die die KI halt mit einem großen Spektrum und viel

mehr Fallbeispielen besser handhaben kann als der Arzt“; „schon vielleicht einen kleinen Vorteil dem Arzt gegenüber hat... auch aufgrund der Summe der Daten“)

- auch kleinste Veränderungen finden (eine Patientin wurde mit Lungenkrebs in sehr frühem Stadium diagnostiziert, der Tumor wurde entfernt. „Es würde einem vielleicht die Sorge nehmen, dass der Arzt etwas übersehen könnte.“)
- In allen Gesprächen kam durch, dass Lungenkrebs als so schlimme / tödliche Erkrankung wahrgenommen wird, dass die frühe Entdeckung einer Erkrankung eine hohe Bedeutung hat. („[...] wenn ich dann einen negativen Befund kriege und später dann doch noch was kommt, das mich dann killt“; „bei Lungenkrebs besser 150 % als 100 %“). In den Gesprächen stellte sich aber heraus, dass 3 der 5 Personen vorher nicht wussten, dass bei früher Diagnose eine Lungenkrebserkrankung besser behandelbar und auch heilbar ist.
- Die einzige aktuelle Raucherin (72 Jahre) berichtet, dass sie eine Früherkennung in Anspruch nehmen würde („dann würde ich das schon in Anspruch nehmen. [...] ich rauche immerhin schon seit meinem 16. Lebensjahr“). Diese Patientin stellt aber gleichzeitig in Frage, ob das Wissen über eine Erkrankung immer dem Nichtwissen zu bevorzugen ist („mit 50 denkst du, du hast noch viel vor dir [...] mach was, aber mit über 70, Mann tot, hier ganz alleine, Verwandte weit weg, Tochter weit weg [...]. naja, es ist so ein Für und Wider [...] aber wenn es so ein Angebot für Raucher gibt, dann würde ich das auf jeden Fall machen, auch wenn ich Angst vor dem Ergebnis hätte. Ach, ich bin da so hin und her“).

Einschätzungen zu erwarteten / möglichen Nachteilen durch die AI-Unterstützung

- Eine gewisse Fehleranfälligkeit der Befundung wird toleriert, die Befragten betonen aber durchgehend, dass falsch-positive gegenüber falsch-negativen Befunden zu bevorzugen seien („Das würde ich sehen als normalen Fehler einer Technik, dass das vorkommen kann, nichts ist unfehlbar. Deshalb ist es ja auch besser, wenn der Arzt dann auch noch mal guckt, bevor es weitergeht.“, „dann lieber Ängste und Sorgen umsonst, als nachher eine Erkrankung, die mich killt“, „Ich glaube, ich würde mir lieber ein paar Sorgen und Ängste mehr machen und nachher stellt sich raus, war nix... ich würde dennoch lieber genau wissen, was da los ist“).
- Eine Person sieht keine wirklichen Nachteile der AI, solange der Datenschutz gewährleistet ist („[...] es muss gewährleistet sein, dass keine Daten der befundeten Personen zum Beispiel an eine Krankenkasse geleitet werden“)
- Wichtig erscheint aber allen, dass AI als Unterstützung und nicht als Ersatz der Radiologin oder des Radiologen bzw. der Ärztin oder des Arztes eingesetzt wird („das menschliche muss irgendwie dann schon dabei sein“; „Deshalb ist es ja auch besser, wenn der Arzt dann auch noch mal guckt, bevor es weitergeht [...]“). In diesem

Zusammenhang wurde die Sorge geäußert, dass die Leistungserbringer sich zu sehr auf die AI verlassen könnten („[...] Furcht, dass der Arzt nicht mehr so genau hinguckt, also sich auf die KI verlässt, zu sehr verlässt).

Einschätzung zur Rolle der Radiologin oder des Radiologen bzw. der Fachärztin oder des Facharztes beim Einsatz einer AI-Unterstützung

- Die Radiologin oder der Radiologe bzw. die Ärztin oder der Arzt wird als höhere Instanz gesehen, der die AI zur Unterstützung an die Seite gestellt bekommt. Außerdem wird er als „Vermittler“ zwischen Ergebnissen / Befund mit AI und betroffener Person gesehen („Der Radiologe / Arzt ist der Experte, der durch die AI unterstützt wird: das, was der Arzt nicht kann – aufgrund von Erfahrung, Tagesform etc. – kann vielleicht die AI und vice versa.“, „Ich verlasse mich am Ende immer auf die Expertise des Arztes, der muss das, was die AI macht, gegenprüfen“, „Der muss mir in einfachen, verständlichen Worten erklären, was er sieht oder vermutet, ob mit oder ohne AI“, „Ich werte die Einschätzung des Arztes immer noch höher als die der AI“, „[...] ich seh'ne AI als Unterstützung, als Unterstützung des Arztes, der nicht jeden Tag 100 000 Lungenbilder sieht, aber eben nicht die AI alleine, sondern immer in Verbindung mit ein oder 2 Ärzten“).

Einschätzung zum Austausch zwischen der Radiologin oder dem Radiologen bzw. der Fachärztin oder dem Facharzt und Betroffenen (Aufklärung / Information über AI-Einsatz)

- Die befragten Personen wünschen sich überwiegend einen offenen Umgang mit dem Einsatz der AI. 2 Personen betont, dass Transparenz besonders dann wichtig ist, wenn es Abweichungen zwischen der AI-Einschätzung und dem Leistungserbringer gibt und dass dann ggf. die Gegenprüfung durch eine 3. Person notwendig ist. Die meisten Befragten wollen eingebunden sein in den Prozess und sich hierzu mit Ärztin oder Arzt bzw. Radiologin oder Radiologen austauschen („Mir in verständlichen Worten erklären, wie die Diagnosestellung erfolgt und mich aufklärt, auch über den Einsatz der AI – damit man am Ende ein komplettes Bild hat.“, „dass er das in das Gespräch einbringt, die Ergebnisse von der AI und seine Gedanken dazu“, „Ich hoffe schon, dass da offen mit umgegangen wird. Ich würde schon gerne wissen, wie die AI da genutzt wird.“, „Kommt drauf an: wenn AI / Arzt voneinander in ihrer Einschätzung abweichen, dann will ich schon wissen, wie damit umgegangen wird, und ob dann noch ein 2. Arzt drauf schaut.“, „Wenn es Pille-Palle ist, kann der Arzt selbst entscheiden, aber wenn da doch größere Dinge sein könnten, will ich schon im Gespräch sein, es geht ja schließlich um mich“).
- zusätzliche Einschätzung der Betroffenen zur Aufklärung und Ergebnisdarstellung durch Ärztin oder Arzt bzw. dem Radiologin oder Radiologen („Ich denke mir, wenn man vernünftig aufklärt und die Situation vernünftig erklärt, kann man auch mit solchen Situationen besser umgehen als mit der Unsicherheit, dass der Arzt irgendwas übersieht,

weil da gibt es ja keine Informationen dann drüber. Der Arzt übersieht und er übersieht. Fertig. Da wäre ich schon lieber auf der sicheren Seite.“, „Ich will wissen was die AI als Diagnose hergibt, dass er das sehr hochwertet, dass er vielleicht voreingenommen ist von der Diagnose der AI [...]. Ich will wissen, wie er das macht, wie er die AI einschätzt und wie er die nutzt“)

Einschätzung zu Zeit / Dauer / Aufwand einer Diagnosestellung mit AI-Unterstützung

- Die meisten befragten Betroffenen gehen davon aus, dass die Zeit der Diagnosestellung gleich bleibt oder sogar kürzer ist („Kein Unterschied – der Arzt prüft schließlich immer noch gegen.“, „Ich gehe davon aus, dass das eher kürzer dauert und dann trotzdem gleich gut ist, vielleicht noch besser.“, „ich möchte nicht, dass es länger dauert“, „Das ist ja im Endeffekt nichts anderes, wie eine Datenbank die letztendlich Informationen sammelt und die dann verwendet, um eine vernünftige Auswertung vornehmen zu können“).
- Eine Person spricht an, dass der Einsatz einer AI wahrscheinlich keine Zeitersparnis bringt, da der Arzt ja dennoch sorgfältig prüfen muss, was die AI sagt. Hier zeigt sich bei einer Befragten auch wieder die Sorge, dass Ärzte sich zu sehr auf die AI verlassen („Also es wird da zeitlich keinen Unterschied geben, meines Erachtens. Es sei denn, der Arzt sagt, ich verlasse mich viel mehr auf AI und macht seinen Job halt nicht 100 %“).
- Einige Patienten sind unsicher bezüglich des Zeitaufwands der AI-Befundung („wenn ich sonst zum Radiologen gehe, der guckt sich die Bilder an und das Ergebnis kriegt man dann ja quasi sofort. Ich weiß nicht, wie lange so ‘ne AI braucht, um dann mein Bild mit dieser Masse an Inputs, die er hat, zu verarbeiten. Also, ich glaube, das dauert länger, als wenn der Arzt direkt mein Röntgenbild und an die Wand hängt [...] dann geht das eher schneller“).
- Zusätzliche Kosten für AI-Einsatz will keiner der Befragten übernehmen (1 Patient sagt, höchstens, wenn der Arzt ihm sagt, wie wichtig das ist).
- Eine Patientin bespricht die Möglichkeit, AI-Unterstützung bei der Früherkennung als IGeL in Anspruch zu nehmen („KI als IGeL? Ich sag mal so, solange es mir gut geht, würde ich sagen, och, muss das sein? Andere IGeL nutze ich schon... zum Beispiel das vaginale Ultraschall [...] weil das einfach besser ist, ne! (und bei der Lunge?) [...] also, ich sag mal so, wenn es mir nicht so gut ginge, ich Probleme hätte mit Husten oder so, würde ich das in Kauf nehmen, einfach so als Vorsorge, ich weiß nicht“).
- Zu den Kosten der AI gibt es eine abwägende Einschätzung („natürlich geh ich davon aus, dass die Anschaffungskosten so einer KI enorm sind. Aber die amortisieren sich mit der Zeit, weil der Arzt ja dadurch auch eine Zeitersparnis hat, ja, quasi eine zweite Meinung, ja, so ein Grundgerüst, auf das er sich stützen kann“).

Ängste / Sorgen bezüglich AI-Unterstützung / AI alleine

- Der Einsatz einer AI alleine zur Befundung wird durchgehend abgelehnt („Nein, AI alleine, da hätte ich kein Vertrauen, das wär gar nicht gut. Da macht der Arzt dann echt seinen Job nicht, sondern sourct einfach was aus, um es sich einfacher zu machen. Das wäre gar nicht gut.“, „ich weiß nicht, da hätte ich schon Sorgen, der Arzt oder der Radiologe sollte auf jeden Fall drüber schauen und mir dann die Information geben.“, „Nee, also, den Arzt brauch ich trotzdem noch, das Gespräch und diese Auswertung, also dieses Doppelspiel, das bräuchte ich schon [...] also nicht nur ein Computerprogramm, das sagt, da ist was oder nicht, schon dass da ein Arzt mit drin ist... und natürlich auch mit persönlichem Gespräch, die Hintergründe erfragt, alles einbezieht [...]“, „Nein, das finde ich gefährlich, sich total auf so eine AI zu verlassen. Da muss der Arzt schon auch eine Rolle spielen. Der ist sogar entscheidend, für mich jedenfalls.“, „Vielleicht hat die AI kleine Vorteile gegenüber dem Arzt, sollte aber nur zusammen mit dem Arzt agieren“).

Gedanken zur Entwicklung von AI-basierten Auswertungs-Systemen

- 3 Personen merken an, dass der Datenschutz (Schutz persönlicher Daten bei Entwicklung der AI und bei Nutzung) immer gewährleistet sein muss.
- 1 Patientin spricht an, dass beim Trainieren der AI mittels Daten auch Daten von Frauen und anderen Gruppen vertreten sein müssen und spricht in diesem Zusammenhang die mögliche Unterrepräsentation von Frauen in klinischen Studien an.
- 1 Patientin spricht an, dass Datenbanken nicht frei zugänglich sein dürfen / gut geschützt sein müssen, auch um nicht anfällig für Manipulation durch Dritte zu sein.
- kein Zugang zu den Daten für Krankenkassen (das sagt eine Patientin) – das darf nicht automatisch in die elektronische Patientenakte. Die Patientin betont, dass sie selbst darüber bestimmen will, wer was sieht und weiß.

Eigene Erfahrungen mit AI bei der Auswertung von Röntgen- / CT-Bildern

Es liegen keine Erfahrungen mit AI vor.

Eigene Erfahrungen mit AI in der Gesundheitsversorgung

Es liegen keine Erfahrungen mit AI oder Informationen aus Medien vor.

A10 Offenlegung von Beziehungen der externen Sachverständigen und des Reviewers bzw. der Reviewerin

Im Folgenden sind die Beziehungen der externen Sachverständigen und des Reviewers bzw. der Reviewerin zusammenfassend dargestellt. Alle Informationen beruhen auf Selbstangaben der einzelnen Personen anhand des „Formblatts zur Offenlegung von Beziehungen“ mit Stand 02/2023. Das aktuelle Formblatt ist unter www.iqwig.de abrufbar. Die in diesem Formblatt verwendeten Fragen finden sich im Anschluss an diese Zusammenfassung.

Externe Sachverständige

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7
Danner, Marion	ja	nein	nein	nein	nein	nein	nein
Rummer, Anne	ja	nein	nein	nein	nein	nein	nein
Pinto dos Santos, Daniel	ja	ja	ja	nein	nein	nein	nein

Im „Formblatt zur Offenlegung von Beziehungen“ (Version 02/2023) wurden folgende 7 Fragen gestellt:

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor bei einer Einrichtung des Gesundheitswesens (z. B. einer Klinik, einer Einrichtung der Selbstverwaltung, einer Fachgesellschaft, einem Auftragsforschungsinstitut), einem pharmazeutischen Unternehmen, einem Medizinproduktehersteller oder einem industriellen Interessenverband angestellt oder für diese / dieses / diesen selbstständig oder ehrenamtlich tätig bzw. sind oder waren Sie freiberuflich in eigener Praxis tätig?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor eine Einrichtung des Gesundheitswesens (z. B. eine Klinik, eine Einrichtung der Selbstverwaltung, eine Fachgesellschaft, ein Auftragsforschungsinstitut), ein pharmazeutisches Unternehmen, einen Medizinproduktehersteller oder einen industriellen Interessenverband beraten (z. B. als Gutachter/-in, Sachverständige/r, in Zusammenhang mit klinischen Studien als Mitglied eines sogenannten Advisory Boards / eines Data Safety Monitoring Boards [DSMB] oder Steering Committees)?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einer Einrichtung des Gesundheitswesens (z. B. einer Klinik, einer Einrichtung der Selbstverwaltung, einer Fachgesellschaft, einem Auftragsforschungsinstitut), einem pharmazeutischen Unternehmen, einem Medizinproduktehersteller oder einem industriellen Interessenverband Honorare erhalten (z. B. für Vorträge, Schulungstätigkeiten, Stellungnahmen oder Artikel)?

Frage 4: Haben Sie oder hat Ihr Arbeitgeber bzw. Ihre Praxis oder die Institution, für die Sie ehrenamtlich tätig sind, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einer Einrichtung des Gesundheitswesens (z. B. einer Klinik, einer Einrichtung der Selbstverwaltung, einer Fachgesellschaft, einem Auftragsforschungsinstitut), einem pharmazeutischen Unternehmen, einem Medizinproduktehersteller oder einem industriellen Interessenverband sogenannte Drittmittel erhalten (d. h. finanzielle Unterstützung z. B. für Forschungsaktivitäten, die Durchführung klinischer Studien, andere wissenschaftliche Leistungen oder Patentanmeldungen)? Sofern Sie in einer größeren Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, z. B. Klinikabteilung, Forschungsgruppe.

Frage 5: Haben Sie oder hat Ihr Arbeitgeber bzw. Ihre Praxis oder die Institution, für die Sie ehrenamtlich tätig sind, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen, z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren für Fortbildungen / Kongresse erhalten von einer Einrichtung des Gesundheitswesens (z. B. einer Klinik, einer Einrichtung der Selbstverwaltung, einer

Fachgesellschaft, einem Auftragsforschungsinstitut), einem pharmazeutischen Unternehmen, einem Medizinproduktehersteller oder einem industriellen Interessenverband? Sofern Sie in einer größeren Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, z. B. Klinikabteilung, Forschungsgruppe.

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile einer Einrichtung des Gesundheitswesens (z. B. einer Klinik, einem Auftragsforschungsinstitut), eines pharmazeutischen Unternehmens, eines Medizinprodukteherstellers oder eines industriellen Interessenverbands? Besitzen Sie Anteile eines sogenannten Branchenfonds, der auf pharmazeutische Unternehmen oder Medizinproduktehersteller ausgerichtet ist? Besitzen Sie Patente für ein pharmazeutisches Erzeugnis, ein Medizinprodukt, eine medizinische Methode oder Gebrauchsmuster für ein pharmazeutisches Erzeugnis oder ein Medizinprodukt?

Frage 7: Sind oder waren Sie jemals an der Erstellung einer medizinischen Leitlinie oder klinischen Studie beteiligt, die eine mit diesem Projekt vergleichbare Thematik behandelt/e? Gibt es sonstige Umstände, die aus Sicht von unvoreingenommenen Betrachtenden als Interessenkonflikt bewertet werden können, z. B. Aktivitäten in gesundheitsbezogenen Interessengruppierungen bzw. Selbsthilfegruppen, politische, akademische, wissenschaftliche oder persönliche Interessen?