



Lungenkrebs

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

**DOKUMENTATION DER ANHÖRUNG ZUM
VORLÄUFIGEN THEMENCHECK-BERICHT**

Projekt: T24-04

Version: 1.0

Stand: 06.08.2026

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

Inhaltsverzeichnis

Abkürzungsverzeichnis.....	4
1 Dokumentation der Anhörung.....	5
Anhang A Dokumentation der Stellungnahmen	6

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AI	Artificial Intelligence (künstliche Intelligenz)
CT	Computertomografie
HTA	Health Technology Assessment (Gesundheitstechnologiebewertung)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
LuKrFrüErkV	Lungenkrebs-Früherkennungs-Verordnung
UICC	Union for International Cancer Control

1 Dokumentation der Anhörung

Am 24.03.2026 wurde der vorläufige ThemenCheck-Bericht in der Version 1.0 vom 24.03.2026 veröffentlicht und zur Anhörung gestellt. Bis zum 21.04.2026 konnten schriftliche Stellungnahmen eingereicht werden. Insgesamt wurden 2 Stellungnahmen form- und fristgerecht abgegeben. Diese Stellungnahmen sind im Anhang abgebildet.

Da sich aus den schriftlichen Stellungnahmen keine Unklarheiten ergaben, war die Durchführung einer wissenschaftlichen Erörterung der Stellungnahmen nicht erforderlich.

Die im Rahmen der Anhörung vorgebrachten Aspekte wurden hinsichtlich valider wissenschaftlicher Argumente für eine Änderung des vorläufigen ThemenCheck-Berichts überprüft. Eine Würdigung der in der Anhörung vorgebrachten wesentlichen Aspekte befindet sich im Kapitel „Würdigung der Anhörung zum vorläufigen ThemenCheck-Bericht“. Im ThemenCheck-Bericht sind darüber hinaus Änderungen, die sich durch die Anhörung ergeben haben, zusammenfassend dargestellt. Der ThemenCheck-Bericht ist auf der Website des IQWiG unter www.iqwig.de veröffentlicht.

Anhang A Dokumentation der Stellungnahmen

Inhaltsverzeichnis

	Seite
A.1 Stellungnahmen von Organisationen, Institutionen und Firmen.....	A 2
A.1.1 Deutsche Röntgengesellschaft (DRG), Gesellschaft für medizinische Radiologie und bildgeführte Therapie e. V. und Deutsche Gesellschaft für Pneumologie und Beatmungsmedizin e. V. (DGP)	A 2
A.1.2 MeVis Medical Solutions AG	A 14

A.1 Stellungnahmen von Organisationen, Institutionen und Firmen

A.1.1 Deutsche Röntgengesellschaft (DRG), Gesellschaft für medizinische Radiologie und bildgeführte Therapie e. V. und Deutsche Gesellschaft für Pneumologie und Beatmungsmedizin e. V. (DGP)

Autorinnen und Autoren

- Barkhausen, Jörg
- Bayerl, Nadine
- Blum, Torsten Gerriet
- Hahn, Horst
- Hamer, Okka
- Kuhl, Christiane
- May, Matthias
- Wielpütz, Mark

Stellungnahme zum vorläufigen ThemenCheck-Bericht

Berichtnr: T24-04

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

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigelegten Anlagen werden in dem Dokument „Dokumentation der Anhörung zum vorläufigen ThemenCheck-Bericht“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>	
Barkhausen, Jörg; Prof. Dr.	
Hahn, Horst; Prof. Dr.	
Bayerl, Nadine; Dr.	
May, Matthias; Prof. Dr.	
Wielpütz, Mark; Prof. Dr.	
Hamer, Okka; Prof. Dr.	
Kuhl, Christiane; Prof. Dr.	
Blum, Torsten Gerriet; Prof. Dr.	
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)	
☒	im Namen folgender Institution / Organisation: Deutsche Röntgengesellschaft (DRG) und Deutsche Gesellschaft für Pneumologie und Beatmungsmedizin (DGP)
☐	als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Dokument haben, tragen Sie diese bitte vor der Tabelle ein. Spezifische Aspekte adressieren Sie bitte in der Tabelle.

Die beteiligten Fachgesellschaften begrüßen ausdrücklich, dass sich das IQWiG im ThemenCheck mit dem wichtigen und hochkomplexen Thema KI in der radiologischen Lungendiagnostik beschäftigt. Der grundsätzlichen Aussage, dass ausreichend valide Daten fehlen, um den Nutzen einer KI-Unterstützung bei der Auswertung von Röntgen- oder CT-Bildern der Lunge abschließend zu beurteilen, stimmen wir zu. **Die Schlussfolgerung, dass der Einsatz von KI-unterstützter radiologischer Befundung erst dann vertretbar ist, wenn der patientenrelevante Nutzen nachgewiesen ist, sehen wir jedoch aus mehreren Gründen kritisch.** Sie steht zudem in direktem Widerspruch zur im gleichen Absatz getätigten Aussage, wonach der vorliegende ThemenCheck-Bericht mangels ausreichender Datenlage für keine der beiden Fragestellungen eine Nutzen- oder Schadenaussage liefert.

Der Einsatz von KI liefert mehr Funktionen als in der Literatur evaluiert und im ThemenCheck untersucht wurden: Verbesserung der Bildqualität, Dosisreduktion, Läsionsdetektion, Läsionsklassifikation, Läsionsvolumenmessung, Verlaufsbeurteilung und Lungenstrukturanalyse. Für jeden dieser Aspekte in der radiologischen Aufarbeitung von Lungenbefunden liegen unterschiedliche und separat zu bewertende Studien vor und die Beurteilung von Künstlicher Intelligenz sollte daher nicht pauschal, sondern stets nach der konkreten Aufgabe differenziert und für Röntgen und Computertomographie getrennt erfolgen.

In Bezug auf die besonders relevante Läsionsvolumenmessung ist zu berücksichtigen, dass die vorliegende Fassung des ThemenCheck die Variabilität der menschlichen Befundung als Vergleichsmaßstab nicht adressiert. Bereits Erasmus et al. (JCO 2003 [1]) zeigten, dass fünf erfahrene Thoraxradiologen bei der Messung von NSCLC-Tumoren in 30–43% der Fälle allein durch Messvariabilität zu einer abweichenden Einschätzung der Tumorprogression kommen. Die LuKrFrühErkV und die V-Lung-RADS-2022-Klassifikation fordern eine Volumetrie und die Berechnung der Volumenverdopplungszeit zur Verlaufsbeurteilung von Lungenrundherden. Diese Messwerte sind ohne Softwareunterstützung nicht mit vertretbarem Aufwand erhebbar oder reproduzierbar – die dedizierte KI- bzw. Software-basierte Volumetrie ist daher keine optionale Ergänzung, sondern eine messtechnische Notwendigkeit im Alltag der Lungenkrebsfrüherkennung.

Wir werden den Nutzen von KI nur dann beurteilen können, wenn wir jetzt beginnen, KI in Deutschland unter wissenschaftlicher Begleitung gezielt einzusetzen und die generierten

Daten strukturiert auszuwerten. Durch die existierenden, aktuellen Vorarbeiten, die Vorgaben der Lungenkrebsfrüherkennungsverordnung (LuKrFrühErkV) und die Neueinführung der Methode in Deutschland ist jetzt der optimale Zeitpunkt für einen Paradigmenwechsel: Wenn wir in Zukunft das Potenzial von KI nutzen wollen, ist der Einsatz auf Basis der vorhandenen Evidenz indiziert, und es gilt, KI in großen Kohorten zu nutzen, um den Mehrwert im Lungenkrebscreening und in der Patientenversorgung abschließend und umfassend nachzuweisen.

Die dafür erforderliche, grundsätzlich positive Bewertung von KI insbesondere im Kontext der Lungenkrebsfrüherkennung wird durch ein Health Technology Assessment des National Institute for Health and Care Research aus Großbritannien (Geppert 2025 [3]) unterstützt, das 27 Studien einschloss und alle vom IQWiG adressierten Populationen (Screening, Tumorverdacht, inzidentelle Befunde) abdeckt. Geppert 2025 kommt zu folgenden Kernaussagen bezüglich der Beurteilung von Rundherden [3]:

- (1) Artificial intelligence assistance generally improved the detection of any nodules compared with unaided reading.
- (2) Radiologist reading time generally decreased with artificial intelligence assistance in research settings.
- (3) AI-assisted analysis may reduce variability of and improve consistency in the measurement and clinical management of lung nodules.
- (4) For screening, artificial intelligence-assisted computed tomography image analysis was cost-effective in the base case and all sensitivity and scenario analyses.

Darüber hinaus unterstützt die Europäische Kommission den Einsatz von Künstlicher Intelligenz und hat gerade den Aufbau von European AI-powered advanced screening centres initiiert ([European AI-powered advanced screening centres | Shaping Europe's digital future](#) [2]). Diese Zentren sollen die klinische Validierung unter realen Bedingungen sowie die lokale Leistungsbewertung von KI-Lösungen ermöglichen und damit Erkenntnisse liefern, die den großflächigen Einsatz im gesamten Gesundheitssystem prüfen.

Kapitel / Abschnitt (Seite) im Dokument	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Alle Volltexte der Literaturstellen, die zitiert werden, müssen als PDF-Dateien eindeutig benannt, im Online-Formular hochgeladen werden.</i>
Seite 7	<u>Anmerkung:</u> Der Begriff „Lungenkrebsfrüherkennungsscreening“ ist ein Pleonasmus und sollte ersetzt werden. <u>Vorgeschlagene Änderung</u> Dieser ThemenCheck-Bericht untersuchte bei Patientinnen und Patienten mit Verdacht auf Lungenkrebs sowie bei Personen, die an der Lungenkrebsfrüherkennung mittels Niedrig-Dosis-Computertomographie teilnehmen, ...

Kapitel / Abschnitt (Seite) im Dokument	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Alle Volltexte der Literaturstellen, die zitiert werden, müssen als PDF-Dateien eindeutig benannt, im Online-Formular hochgeladen werden.</i>
Seite 9	<u>Anmerkung:</u> Die Aussage, dass durch die „Vorgaben der Lungenkrebs-Früherkennungs-Verordnung kein personelles Einsparpotenzial durch KI zu erkennen ist“, sollte aus mehreren Gründen überprüft werden. Erstens kommt das Health Technology Assessment des National Institute for Health and Care Research (Geppert 2025 [3]) zu einem anderen Ergebnis. Zweitens sollen, analog zum Mammographie-Screening, auch bei der Lungenkrebs-Früherkennung alle CT-Untersuchungen doppelt befundet werden. Gemäß LuKreFrühErkV wird diese Vorgabe bei allen Lung-RADS-1- und 2-Befunden durch die Kombination aus einer Radiologin/einem Radiologen und einer KI-basierten Software erfüllt. Eine ärztliche Zweitbefundung ist nur bei höhergradigen Lung-RADS-Befunden erforderlich. Die aktuellen Daten der HANSE-Studie (Vogel-Claussen 2025 [7]) zeigen, dass bei mehr als 90% der Untersuchungen, die in der HANSE-Studie als Lung-RADS 1 und 2 bewertet wurden, durch den Einsatz einer KI-basierten Software keine ärztliche Zweitbefundung erforderlich ist, was ein erhebliches personelles Einsparpotenzial darstellt. Dieses wurde durch die Verordnung also bereits umgesetzt. Drittens wird hier nur die Detektion von Lungenrundherden betrachtet. Die vom G-BA empfohlene Lung-RADS-Klassifikation erfordert jedoch eine Segmentierung und Volumetrie der Lungenrundherde. Eine manuelle 3D-Segmentierung ist extrem zeitaufwendig, klinisch nicht validiert und ärztlichen Kolleg:innen im Rahmen der Lungenkrebsfrüherkennung nicht zumutbar. Der Einsatz einer KI-basierten Software ist bei der Segmentierung alternativlos und im Vergleich zur manuellen Bearbeitung mit einer erheblichen Zeitersparnis verbunden. Auch dies ist in den Vorgaben der LuKreFrühErkV bereits berücksichtigt worden. <u>Vorgeschlagene Änderung</u> Nach unserer Einschätzung ist eine grundlegende Überarbeitung der Aussagen zum personellen Einsparpotential erforderlich.

Kapitel / Abschnitt (Seite) im Dokument	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Alle Volltexte der Literaturstellen, die zitiert werden, müssen als PDF-Dateien eindeutig benannt, im Online-Formular hochgeladen werden.</i>
Seite 10	<u>Anmerkung:</u> Neben einer Reduktion der Mortalität und Morbidität oder einer Verbesserung der gesundheitsbezogenen Lebensqualität gibt es zahlreiche weitere Gründe, die bereits einzeln den Einsatz von KI-unterstützten Softwares rechtfertigen. Erstens: Der Vergleichsstandard (menschliche Befundung) ist selbst hochvariabel – Erasmus et al. (JCO 2003 [1]) belegten Abweichungen in der Tumorklassifikation von 30–43% allein durch Interobserver-Variabilität bei manuellen Tumormessungen. Die Reduktion dieser Variabilität durch standardisierte Software hat somit einen eigenständigen patientenrelevanten Nutzen in der Entscheidungsfindung bei der Rundherdbeurteilung. Zweitens: KI bietet mit der softwarebasierten Volumetrie und Volumenverdopplungszeit (VDT) eine Funktionalität, die manuell nicht reproduzierbar erbracht werden kann (Geppert 2025, Seite 39–48 [3]). Drittens: Da die Mortalität einer Erkrankung von vielen Faktoren abhängt, ist es fraglich, ob ein einzelner diagnostischer Test diese Parameter allein beeinflussen kann. Für diesen Nachweis sind Studien mit extrem großen Fallzahlen erforderlich, die nur durch den flächendeckenden Einsatz und bestenfalls Randomisierung der Methoden erreicht werden können. Dies gilt gleichermaßen für die CT-basierte Früherkennung wie für den im Bericht ebenfalls adressierten, aber evidenzmäßig nicht bearbeiteten Einsatz von KI bei der Befundung von Röntgen-Thorax-Aufnahmen bei Verdacht auf Lungenkrebs (Fragestellung 1), für die ebenfalls keine Studien identifiziert werden konnten. <u>Vorgeschlagene Änderung</u> Die Aussagen zum Einsatz von KI in der Lungenkrebsfrüherkennung sollten überprüft werden.
Seite 24	<u>Anmerkung:</u> Im Jahr 2024 war das Lungenkarzinom bei Frauen erstmals die häufigste krebsbedingte Todesursache in Deutschland (Pressemitteilung des Deutschen Krebsregisters 2025 [8]) <u>Vorgeschlagene Änderung:</u> Aktuelle Daten aus dem Jahr 2024 zeigen, dass Lungenkrebs inzwischen bei Männern und Frauen die häufigste krebsbedingte Todesursache ist.

Kapitel / Abschnitt (Seite) im Dokument	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Alle Volltexte der Literaturstellen, die zitiert werden, müssen als PDF-Dateien eindeutig benannt, im Online-Formular hochgeladen werden.</i>
Seite 24	<u>Anmerkung:</u> Die Aussage, 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, gibt die klinische Realität verzerrt wieder, da die Auflistung 22 bzw. 23% fehlende Stadienangaben aus klinischen Krebsregistern bzw. 9% unzureichende Stadienangaben aus Totenscheinen berücksichtigt (Abb. 3.12.4, RKI-Bericht „Krebs in Deutschland“ für 2021-2023 [6]). Wir wissen aus den Daten der Arbeitsgemeinschaft Deutscher Tumorzentren, dass die Gesamtüberlebenskurven bei Patienten mit unzureichenden UICC-Stadienangaben zwischen den Stadien III und IV verlaufen. Bei 9 % der Fälle mit Erstdiagnose eines Lungenkarzinoms anhand von Totenscheinen ist keine Überrepräsentation früher Stadien zu erwarten. <u>Vorgeschlagene Änderung:</u> Bei der Verteilung der UICC-Stadien sollten nur die Fälle berücksichtigt werden, in denen gültige Werte vorliegen. In diesem Fall werden ohne Screening bei Frauen 25% und bei Männern 21% der Lungenkarzinome in den Stadien I und II diagnostiziert. Dagegen zeigen die Daten der HANSE-Studie (Vogel-Claussen 2025 [7]) und die aus dem NHS England Lung Cancer Screening Programme (Lee 2026 [5]) übereinstimmend, dass bei einem Screening mit Softwareunterstützung $\frac{3}{4}$ der Lungenkarzinome in den Stadien I und II diagnostiziert werden.
Seite 25	<u>Anmerkung:</u> Weder eine CAdE- noch eine CAdx-Software ermöglicht die Bewertung des Tumorwachstums. Die Berechnung der Tumorverdopplungszeit erfordert die Segmentierung und Volumetrie der Rundherde. <u>Vorgeschlagene Änderung:</u> Im gesamten Dokument sollte bei der Bewertung der KI-Unterstützung zwischen der Detektion und der Segmentierung/Volumetrie von Lungenrundherden differenziert werden. Eine ausführliche Diskussion dieser beiden Aspekte findet sich im Health Technology Assessment des National Institute for Health and Care Research (Geppert 2025, Seiten 25–45 [3]). Die entsprechenden Abschnitte im vorläufigen Bericht sollten umfassend überarbeitet und ergänzt werden.

Kapitel / Abschnitt (Seite) im Dokument	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Alle Volltexte der Literaturstellen, die zitiert werden, müssen als PDF-Dateien eindeutig benannt, im Online-Formular hochgeladen werden.</i>
Seite 34	<u>Anmerkung:</u> Radiologin <u>Vorgeschlagene Änderung:</u> Radiologinnen
Seite 37	<u>Anmerkung:</u> Die Berechnungen in Kapitel 3.6.1 sind nicht einfach nachvollziehbar und in wesentlichen Punkten methodisch fragwürdig. Erstens beruht die Sensitivitätsschätzung in Lee 2025 [4] auf nur 5 referenztestpositiven Personen (3 vs. 4 korrekt erkannt), was Konfidenzintervalle von [23,1; 88,2] vs. [37,6; 96,4] ergibt – statistisch kein nachweisbarer Unterschied. Dennoch wird aus diesen nicht-belastbaren Punktschätzern eine Hochrechnung auf 100.000 Personen abgeleitet, die die KI systematisch benachteiligt (PPV 2,4% vs. 4,4%). Zweitens ist der Anteil der positiven Testergebnisse mit 16% bzw. 30% in den Studien auffällig hoch im Vergleich zu den Ergebnissen der HANSE-Studie und anderer aktueller Screening-Programme. Drittens zeigt Geppert 2025 (Seite 88–93, full model [3]), dass KI-unterstützte Befundung für die Screeningpopulation gesundheitsökonomisch durchgehend überlegen war – ein Ergebnis, das der Bericht nicht angemessen würdigt. <u>Vorgeschlagene Änderung:</u> Die Berechnungen sowie deren Interpretation sind entsprechend zu ergänzen bzw. differenzierter darzustellen.
Seite 60	<u>Anmerkung:</u> Da Untersuchung und Auswertung räumlich und zeitlich getrennt durchgeführt werden können, ist es problemlos möglich, eine AI-unterstützte Befundung auch in mobilen Früherkennungszentren anzubieten. <u>Vorgeschlagene Änderung:</u> Der letzte Satz zum Einsatz von KI in mobilen Untersuchungseinheiten sollte gestrichen werden.

Kapitel / Abschnitt (Seite) im Dokument	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Alle Volltexte der Literaturstellen, die zitiert werden, müssen als PDF-Dateien eindeutig benannt, im Online-Formular hochgeladen werden.</i>
Seite 69	<u>Anmerkung:</u> Die Aussage, dass die Befundung im Rahmen der Lungenkrebsfrüherkennung zum Stand von Dezember 2025 nicht routinemäßig mithilfe von KI erfolgt, gibt die Situation in Deutschland verzerrt wieder. Da die Vergütung erst im März 2026 abschließend geregelt wurde und das Verfahren erst seit dem 01.04.2026 als Leistung der GKV angeboten wird, konnten 2025 routinemäßige Lungenkrebsfrüherkennungsuntersuchungen weder mit noch ohne KI durchgeführt werden. Auf der anderen Seite nutzen alle aktuell in Deutschland laufenden Studien zur Lungenkrebsfrüherkennung KI-Unterstützung zur Detektion und zur Volumetrie von Lungenrundherden. <u>Vorgeschlagene Änderung:</u> Dieser Abschnitt sollte korrigiert werden.
Seite 69	<u>Anmerkung:</u> Die Aussage, dass es wahrscheinlich ist, dass die Leistung „primär in zertifizierten Lungenkrebszentren angeboten wird“, ist nicht korrekt. Die Durchführung der Untersuchung und die Erstbefundung setzen neben der Genehmigung der Kassenärztlichen Vereinigung die Befugnis zur Teilnahme an der vertragsärztlichen Versorgung voraus. Die Untersuchungen werden daher weit überwiegend in radiologischen Praxen durchgeführt. Voraussetzung für die Zweitbefundung ist dagegen die Tätigkeit an einer Einrichtung, die auf die Untersuchung und Behandlung von Lungenkrebs gemäß § 43 Absatz 4 der Krebsfrüherkennungs-Richtlinie spezialisiert ist. <u>Vorgeschlagene Änderung:</u> Dieser Abschnitt sollte korrigiert werden.

Kapitel / Abschnitt (Seite) im Dokument	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Alle Volltexte der Literaturstellen, die zitiert werden, müssen als PDF-Dateien eindeutig benannt, im Online-Formular hochgeladen werden.</i>
Seite 71	<u>Anmerkung:</u> Die Aussage, dass durch klare Vorgaben der LuKrFrühErkV zum Einsatz von Personal (Erst- und Zweitbefundung) auch hier kein personelles Einsparpotenzial durch eine AI-Unterstützung gegeben ist, trifft nicht zu. Siehe ausführlichen Kommentar zu Seite 9. <u>Vorgeschlagene Änderung:</u> Nach unserer Einschätzung ist eine grundlegende Überarbeitung der Aussagen zum personellen Einsparpotential, insbesondere im Zusammenhang mit der Lungenkrebsfrüherkennung, erforderlich.
Seite 72	<u>Anmerkung:</u> Die finale Version der Versicherteninformation enthält den Hinweis auf den Einsatz spezieller Computerprogramme im Rahmen der Befundung. <u>Vorgeschlagene Änderung:</u> Der letzte Satz des Abschnitts ‚Kommunikations- und Kooperationsformen‘ sollte gestrichen werden.
Seite 73	<u>Anmerkung:</u> Die Aussage, dass die Umsetzung der AI-unterstützten Befundung/Früherkennung angesichts der finanziellen, personellen und organisatorischen Herausforderungen voraussichtlich am ehesten in zertifizierten Lungenkrebszentren erfolgen wird, ist nicht zutreffend (siehe Kommentar zu Seite 69). <u>Vorgeschlagene Änderung:</u> Dieser Abschnitt sollte korrigiert werden.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

[1] Erasmus, J. J., Gladish, G. W., Broemeling, L., et al. (2003) Interobserver and intraobserver variability in measurement of non–small-cell carcinoma lung lesions: Implications for assessment of tumor response. JCO 21(13), 2574–2582.

<https://doi.org/10.1200/JCO.2003.01.144>

[2] Europäische Kommission. (2025) European AI-powered advanced screening centres. Shaping Europe’s Digital Future. <https://digital-strategy.ec.europa.eu/en/policies/ai-screening-centres>

[3] Geppert, J., Auguste, P., Asgharzadeh, A., et al. (2025) Software with artificial intelligence-derived algorithms for detecting and analysing lung nodules in CT scans: Systematic review and economic evaluation. HTA 29(14), 1–234.

<https://doi.org/10.3310/JYTW8921>

[4] Lee, M., Hwang, E. J., Lee, J. H., et al. (2025) Artificial intelligence for low-dose CT lung cancer screening: Comparison of utilization scenarios. AJR 225(1), e2532829.

<https://doi.org/10.2214/AJR.25.32829>

[5] Lee, R. W., Nair, A., Balata, H., et al. (2026) Implementation of the NHS England Lung Cancer Screening Programme over 5 years. Nature Medicine.

<https://doi.org/10.1038/s41591-026-04292-y>

[6] Robert Koch-Institut. (2025) Krebs in Deutschland für 2021–2023. Zentrum für Krebsregisterdaten. <https://edoc.rki.de/handle/176904/13155>

[7] Vogel-Claussen, J., Bollmann, B.-A., May, K., et al. (2025) Effectiveness of NELSON versus PLCOm2012 lung cancer screening eligibility criteria in Germany (HANSE): A prospective cohort study. The Lancet Oncology 26(12), 1541–1551. [https://doi.org/10.1016/S1473-2045\(25\)00490-5](https://doi.org/10.1016/S1473-2045(25)00490-5)

[8] Pressemitteilung des Deutschen Krebsregisters. <https://www.dkr.de/lungenkrebs-jetzt-haeufigster-krebstodesursache-bei-frauenkrebs> jetzt häufigste Krebstodesursache bei Frauen
[DKR - Homepage](#)

A.1.2 MeVis Medical Solutions AG

Autorinnen und Autoren

- Schörner, Stephanie
- Woetzel, Jan

Stellungnahme zum vorläufigen ThemenCheck-Bericht

Berichtnr: T24-04

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

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation der Anhörung zum vorläufigen ThemenCheck-Bericht“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>	
Schörner, Stephanie Marie; Dr.	
Woetzel, Jan	
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)	
☒	im Namen folgender Institution / Organisation: MeVis Medical Solutions AG
☐	als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Dokument haben, tragen Sie diese bitte vor der Tabelle ein. Spezifische Aspekte adressieren Sie bitte in der Tabelle.

Allgemein erscheint uns zur Nutzenbewertung der AI-Unterstützung bei der Befundung von a) Hochrisikogruppen in der Früherkennung und b) konkretem Verdacht auf Lungenkrebs die Wahl der finalen Endpunkte Mortalität, Morbidität und Lebensqualität als alleinige Quelle der Evidenzgewinnung methodisch nicht vollständig.

Bereits die Bewertung von Teilaspekten von AI als Werkzeug für die Verbesserung von Teilaspekten, insbesondere der Detektionsqualität, Objektivierung und Reproduzierbarkeit von Messungen wie der Volumetrie in der Verlaufskontrolle oder allgemein Erhöhung der Qualität von entscheidungsunterstützenden Zwischenergebnissen erscheint uns ausreichend, um auch einen systemischen Effekt hochwahrscheinlich erscheinen zu lassen.

Kapitel / Abschnitt (Seite) im Dokument	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Alle Volltexte der Literaturstellen, die zitiert werden, müssen als PDF-Dateien eindeutig benannt, im Online-Formular hochgeladen werden.</i>
Kosten (S. 8)	<u>Anmerkung</u> Neben der Qualitätserhöhung ist auch das ökonomische Einsparpotential erkennbar, insbesondere durch Entfall des menschlichen 2nd read, Workflowverbesserung und Strukturierung sowie der Sicherstellung der Versorgung durch zeiteffektive Befundung [geppert2025, jacobs2021, riel2019, vogel-claussen2025] <u>Vorgeschlagene Änderung</u> Berücksichtigung der vorliegenden, positiven Effekte auch zur Verringerung der systemischen Kosten zur Erreichung der Versorgungssicherheit.
(S. 25f.)	<u>Anmerkung</u> Für eine valide Studienauswahl, die methodische Nachvollziehbarkeit sowie die Bewertung einer potenziellen Zeitersparnis erfordert ist eine präzise Beschreibung der zu vergleichenden Befundungsprozesse. Diese scheint zu fehlen. Sowohl der Befundungsworkflow ohne KI als auch der Workflow mit KI-Unterstützung sind nicht hinreichend operationalisiert, sodass unklar bleibt, was genau miteinander verglichen wird. <u>Vorgeschlagene Änderung</u> Eine entsprechende Konkretisierung des Vergleichs sollte auf S. 25/26 erfolgen.

2.3ff (S 28.)	<u>Anmerkung</u> Allgemein erscheint uns zur Nutzenbewertung der AI-Unterstützung bei der Befundung von a) Hochrisikogruppen in der Früherkennung und b) konkretem Verdacht auf Lungenkrebs die Wahl der Endpunkte Mortalität, Morbidität und Lebensqualität als alleinige Quelle der Evidenzgewinnung methodisch nicht vollständig zur Bewertung des potentiellen Nutzens von AI. Auch die Bewertung von Teilaspekten wie der Befundungsqualität und Fehlerrate lässt wesentliche früher Rückschlüsse auf den zu erwartenden Impact wahrscheinlich erscheinen. Die AI fungiert nicht als eigenständige diagnostische Instanz, sondern als Teil der Diagnostik der Leitlinien für Diagnostik und Therapie. Der Effekt von AI erscheint auch anhand von Teilaspekten bewertbar. Hier exemplarisch durch Bewertung der Befundungsqualität durch Konsens in der Diagnostik oder Vergleich der diagnostischen Qualität der Früherkennung durch a) einen menschlichen Reader, b) zwei menschliche Reader oder c) einen menschlichen Reader und den zweiten Read durch die AI mittels CAD-e zur Detektion übersehener Lungenrundherde. Im Anwendungskontext ist der Nutzen von AI auch daran messbar, ob das System mit AI eine mit einem zweiten menschlichen Radiologen vergleichbare oder bessere Performance in Teilaspekten erreicht. Ein Expertenkonsens als Referenzstandard ist methodisch zulässig und angemessen zur Bewertung der Teilaspekte. Patientenzentrierte Endpunkte wie Mortalität oder Morbidität dagegen werden primär durch das organisierte Screeningprogramm, nachgelagerte diagnostische Entscheidungen und Therapien beeinflusst und nicht isoliert durch den Einsatz eines unterstützenden AI Detektionssystems. Zudem definiert nicht die AI die klinische Entscheidung, sondern verringert die Fehlerrate in der Detektion und erhöht die Messgenauigkeit. Die in den S3-Leitlinien referenzierte Lung-RADS Klassifikation legt die Kriterien zur Abklärungsbedürftigkeit fest. Die Aufgabe der KI beschränkt sich auf Detektion und ggf. Quantifizierung; die finale klinische Entscheidungslogik bleibt leitlinienbasiert. <u>Vorgeschlagene Änderung</u> Berücksichtigung und Diskussion der vorliegenden wissenschaftlichen Ergebnisse für isolierte Teilaspekte der Diagnostik (Detektion, Messgenauigkeit, Grad des Konsens) und der Beeinflussung von Änderungen im Patientenmanagement. Vgl. [geppert2025, jacobs2021, jiang2025, riel2019, cheo2025].
--------------------------	---

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

[geppert2025] Geppert, J., Auguste, P., Asgharzadeh, A., Ghiasvand, H., Patel, M., Brown, A., ... & Chen, Y. F. (2025). Software with artificial intelligence-derived algorithms for detecting and analysing lung nodules in CT scans: systematic review and economic evaluation. *Health Technology Assessment (Winchester, England)*, 29(14), 1.

[jacobs2021] Jacobs, Colin, et al. "Assisted versus manual interpretation of low-dose CT scans for lung cancer screening: impact on lung-RADS agreement." *Radiology: Imaging Cancer* 3.5 (2021): e200160.

[jiang2025] Jiang, Beibei, et al. "AI performance for nodule volume doubling time in the follow-up of the UKLS lung cancer screening study compared to expert consensus and histological validation." *European Journal of Cancer* (2025): 116137.

[riel2019] van Riel, Sarah J., et al. "Observer variability for Lung-RADS categorisation of lung cancer screening CTs: impact on patient management." *European radiology* 29.2 (2019): 924-931.

[cheo2025] Cheo, Hao Min, Chern Yue Glen Ong, and Yonghan Ting. "A systematic review of AI performance in lung cancer detection on CT thorax." *Healthcare*. Vol. 13. No. 13. MDPI, 2025.

[vogel-claussen2025] Vogel-Claussen, J., Bollmann, B.-A., May, K., et al. (2025) Effectiveness of NELSON versus PLCOm2012 lung cancer screening eligibility criteria in Germany (HANSE): A prospective cohort study. *The Lancet Oncology* 26(12), 1541–1551. [https://doi.org/10.1016/S1470-2045\(25\)00490-5](https://doi.org/10.1016/S1470-2045(25)00490-5)