



“Literatursuche und Evidenzsynthese: S3-Leitlinie maligne Nierentumore“

Teil 1: Forschungsfrage Diagnostik

Systematische Übersichtsarbeit

Department für Evidenzbasierte Medizin und
Klinische Epidemiologie

An die

Deutsche Gesellschaft für Urologie

Prof.ⁱⁿ Dr.ⁱⁿ Susanne Krege

Uerdingerstraße 64

40474 Düsseldorf

Von der

Donau-Universität Krems

Department für Evidenzbasierte Medizin und Klinische Epidemiologie

Dr.-Karl-Dorrek-Straße 30

3500 Krems an der Donau

Impressum

Donau-Universität Krems

Department für Evidenzbasierte Medizin und Klinische Epidemiologie

Dr.-Karl-Dorrek-Straße 30

3500 Krems

Tel. 02732 893 2911

Fax 02732 893 4910

gerald.gartlehner@donau-uni.ac.at

AutorInnen

Dr.ⁱⁿ Maria Flamm, MPH

Barbara Nußbaumer, Bakk. BSc. MSc.

Mag.^a Bita Mesgarpour

Megan G. Van Noord, MSIS

Univ.-Prof. Dr. Gerald Gartlehner, MPH

Interner Review

Michaela Strobelberger, MA

Dr. Peter Mahlknecht

Inhaltsverzeichnis

1	Einleitung.....	1
2	Methoden.....	2
2.1	Fragestellung	2
2.2	Literatursuche.....	3
2.3	Selektion der Literatur.....	3
2.4	Datenextraktion.....	4
2.5	Beurteilung der methodischen Validität	5
2.6	Auswahl relevanter Endpunkte	6
2.7	Beurteilung der Qualität der Evidenz	6
2.8	Synthese der Evidenz.....	7
3	Resultate Forschungsfrage 1	8
3.1	Primärdiagnostische Abklärung des Nierenzellkarzinoms.....	9
3.1.1	Primärdiagnostische Abklärung der Nieren mittels Sonographie vs. Computertomographie	10
3.1.2	Primärdiagnostische Abklärung der Nieren mittels Magnetresonanztomographie vs. Computertomographie	20
3.1.3	Primärdiagnostische Abklärung der Nieren mittels PET Scan vs. Computertomographie	25
3.1.4	Primärdiagnostische Abklärung der Nieren mittels Magnetresonanzangiographie vs. Computertomographie	30
4	Diskussion.....	31
5	Referenzen	36
6	Appendizes	37
6.1	Appendix 1: Suchstrategie 29. Jänner 2013.....	37
6.2	Appendix 2: Kontaktierte AutorInnen	41
6.3	Appendix 3: Ausgeschlossene Studien	42
6.3.1	Ausschluss wegen nicht zutreffenden Publikationstyps (6):	42
6.3.2	Ausschluss wegen nicht zutreffender Population (6):.....	42
6.3.3	Ausschluss wegen nicht zutreffender Intervention oder Kontrollintervention (5):	43
6.3.4	Ausschluss wegen nicht zutreffender Endpunkte (4):.....	43
6.3.5	Ausschluss wegen nicht zutreffenden Studiendesigns (7):	43
6.3.6	Ausschluss wegen nicht zutreffender Sprache (1):	44
6.3.7	Ausschluss da nur Abstract verfügbar (4):	44

Tabellenverzeichnis

Tabelle 1: Einschlusskriterien Forschungsfrage Diagnostik	4
Tabelle 2: Definition und Interpretation des Biasrisikos	5
Tabelle 3: <i>Darstellung der wesentlichsten Charakteristika und Resultate der eingeschlossenen Studien</i>	11
Tabelle 4: GRADE Evidenz Profil - Kontrastmittelverstärkter Ultraschall	17
Tabelle 5: GRADE Evidenz Profil - Nicht-verstärkter Ultraschall	18
Tabelle 6: GRADE Evidenz Profil - Computertomographie	19
Tabelle 7: <i>Darstellung der wesentlichsten Charakteristika und Resultate der eingeschlossenen Studie</i>	21
Tabelle 8: GRADE Evidenz Profil - Magnetresonanztomographie	23
Tabelle 9: GRADE Evidenz Profil – Multidetector Computertomographie	24
Tabelle 10: <i>Darstellung der wesentlichsten Charakteristika und Resultate der eingeschlossenen Studie</i>	26
Tabelle 11: GRADE Evidenz Profil – PET-Scan	28
Tabelle 12: GRADE Evidenz Profil –Computertomographie	29
Tabelle 13: Bestimmungsfaktoren der Stärke der Empfehlung	35

Abbildungsverzeichnis

Abbildung 1: Analytisches Rahmenwerk für Fragestellung 1 Diagnostik	2
Abbildung 2: Darstellung der Durchsicht der Literatur	8
Abbildung 3: Darstellung des Biasrisikos der eingeschlossenen Studien	10
Abbildung 4: Diagnostische Wertigkeit der einzelnen diagnostischen Tests	33

1 Einleitung

Das Nierenzellkarzinom macht rund 80-85% der malignen Nierentumore aus.¹ 2010 betrug die Inzidenz in Deutschland 17 pro 100.000 Männer und 8,4 pro 100.000 Frauen (altersstandardisierte Rate).² Es handelt sich um einen relativ symptomarmen Tumor, der häufig als Zufallsbefund entdeckt wird. Die häufigsten Therapieverfahren sind die operative Entfernung bzw. die systemische Chemotherapie.

Im Rahmen der derzeit laufenden Erstellung einer S3 Leitlinie zum Nierenzellkarzinom wurden zwei Teilbereiche zur externen Aufarbeitung der Evidenz an das Department für Evidenzbasierte Medizin und Klinische Epidemiologie an der Donau-Universität Krems ausgelagert.

Der vorliegende Bericht beschäftigt sich mit der Primärdiagnostik des Nierenzellkarzinoms.

2 Methoden

2.1 Fragestellung

Unter Einbeziehung des Auftraggebers, wurde folgende Fragestellung vor Durchführung der Literaturrecherchen definiert:

Gibt es bei PatientInnen, bei denen eine primärdiagnostische Abklärung wegen Verdachts auf Nierenzellkarzinom durchgeführt wird, Unterschiede in der Testgenauigkeit zwischen Computertomographie und Magnetresonanztomographie, Magnetresonanztomographie, Ultraschall oder PET (Positronen Emissions Tomographie)-Scan?

Gibt es Unterschiede hinsichtlich der Strahlenbelastung?

Abbildung 1 stellt die Fragestellung grafisch dar.

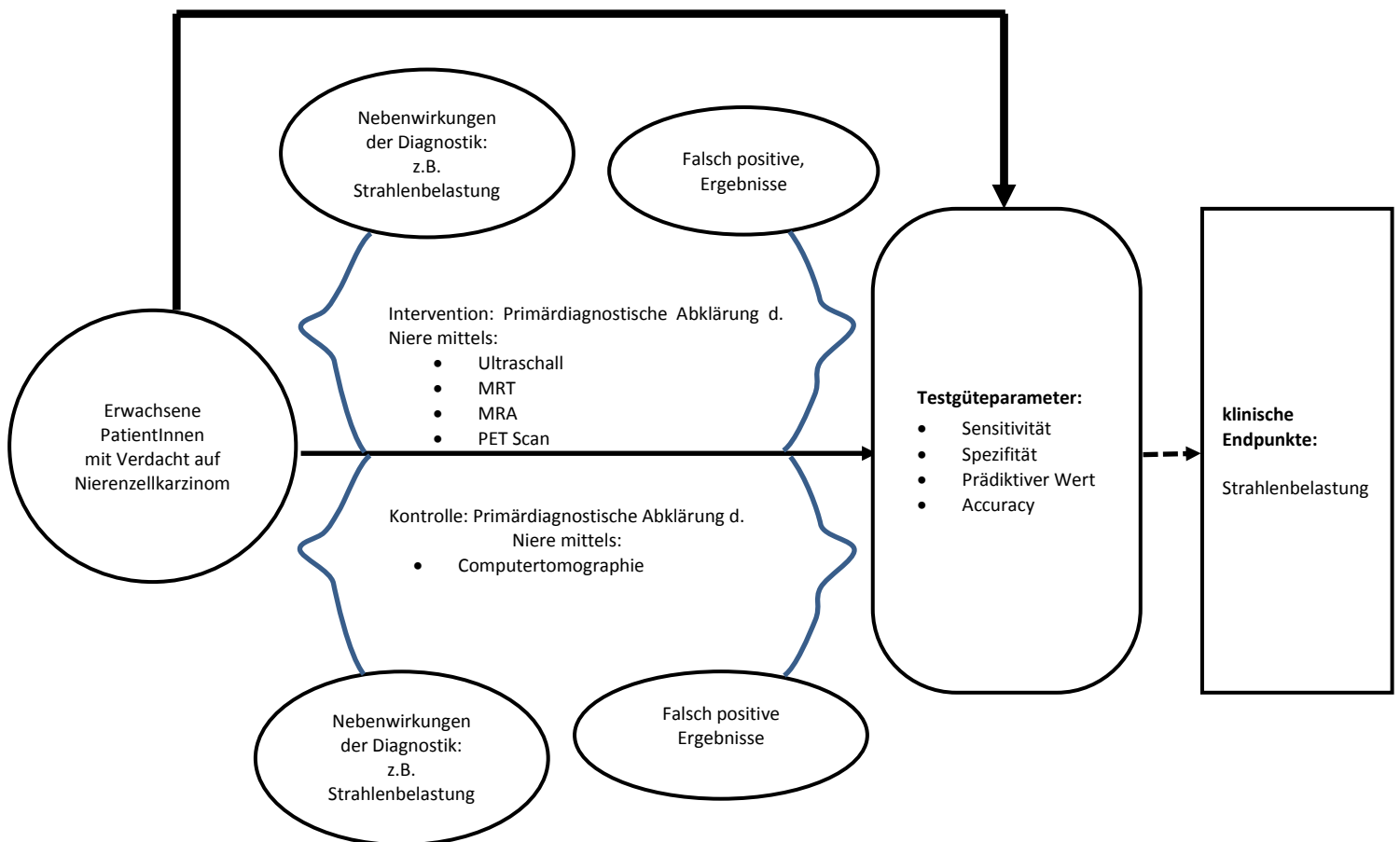


Abbildung 1: Analytisches Rahmenwerk für Fragestellung 1 Diagnostik

2.2 Literatursuche

Für die Identifikation relevanter Literatur zur Beantwortung der Fragestellung wurden folgende elektronische Literaturdatenbanken durchsucht: PubMed, die Cochrane Library, Web of Knowledge, CINAHL (Cumulative Index to Nursing and Allied Health Literature) und EMBASE (Excerpta Medica Database). Die Suche wurde mit Hilfe einer Kombination aus relevanten Schlüsselwörtern (Keywords) und, wenn anwendbar, mit MeSH (Medical Subject Headings)-Begriffen durchgeführt. Dabei wurden jeweils Schlüsselwörter oder MeSH-Begriffe zur diagnostischen Intervention und Kontrolle sowie zu möglichen Endpunkten kombiniert. Die Suche wurde auf Humanstudien im Suchzeitraum 2002 bis Januar 2013 limitiert. Die Suchstrategie umfasste beide Fragestellungen (d.h. Diagnostik und systemische Chemotherapie) und ist in Appendix 1 dargestellt.

Die systematische Suche in Datenbanken wurde mit einer umfassenden Handsuche in Referenzlisten von Übersichtsartikeln und eingeschlossenen Studien ergänzt.

Alle identifizierten Literaturstellen wurden in eine elektronische Datenbank importiert (EndNote X.4) und mit diesem Programm verwaltet.

2.3 Selektion der Literatur

Für die Auswahl der Studien entwickelte das Forschungsteam in Zusammenarbeit mit den AuftraggeberInnen á priori definierte Einschlusskriterien. Diese sind in Tabelle 1 zusammengefasst.

Die Literaturselektion erfolgte in zwei aufeinanderfolgenden Schritten (Erst- und Zweitselektion) und wurde jeweils durch zwei ProjektmitarbeiterInnen unabhängig voneinander durchgeführt. Bei der Erstselektion wurden alle Abstracts (bzw. die Titel, wenn kein Abstract vorhanden war) auf Basis der vorab definierten Ein- und Ausschlusskriterien gesichtet. Wenn sich beide BegutachterInnen einig waren, dass die Studie nicht den Auswahlkriterien entsprach, wurde sie ausgeschlossen. Bei Unsicherheiten über den Ein- oder Ausschluss eines Abstracts wurde das Abstract eingeschlossen und der Zweitselektion zugeführt.

Von den eingeschlossenen Abstracts wurden die Volltexte organisiert, die als Basis für die Zweitselektion dienten. Publikationen, wurden in diesem Schritt nur dann ausgeschlossen, wenn zwei BegutachterInnen der Meinung waren, dass sie nicht den zuvor definierten Einschlusskriterien entsprachen. Bei Uneinigkeiten wurde zur Entscheidungsfindung eine dritte Person herangezogen.

Tabelle 1: Einschlusskriterien Forschungsfrage Diagnostik

POPULATION:	<p>Erwachsene PatientInnen mit Verdacht auf Nierenzellkarzinom:</p> <ul style="list-style-type: none"> • Nierenzellkarzinom • Klarzelliges Karzinom • Hypernephrom • Papilläres Nierenzellkarzinom • Chromophobes Nierenzellkarzinom • Nicht-Klarzelliges Nierenzellkarzinom
INTERVENTION:	<p>Primärdiagnostische Abklärung der Nieren mittels</p> <ul style="list-style-type: none"> • Ultraschall • Magnetresonanztomographie (MRT) • Magnetresonanztomographie (MRA) • Positronenemissions Tomographie (PET) Scan
KONTROLLINTERVENTION:	Computertomographie
OUTCOMES:	<ul style="list-style-type: none"> • Testgüteparameter (Sensitivität, Spezifität, prädiktive Werte, Accuracy) • Strahlenbelastung
TIMING:	Suchzeitraum 2002-Jänner 2013
STUDIENDESIGNS:	<ul style="list-style-type: none"> • Randomisierte Kontrollierte Studien (RCTs) • Systematische Reviews und Meta-Analysen von RCTs oder prospektiven, kontrollierten Studien • Diagnostische Studien <p>Zusätzlich für Strahlenbelastung:</p> <ul style="list-style-type: none"> • Prospektive kontrollierte Beobachtungsstudien

2.4 Datenextraktion

Die Datenextraktion erfolgte anhand eines standardisierten Formulars, das in einer Pilotphase getestet und überarbeitet wurde. Die Datenextraktion wurde von einer Person durchgeführt und von einer weiteren Person kontrolliert. Unterschiedliche Beurteilungen und Diskrepanzen bei der Datenextraktion wurden mittels Diskussion und Konsens zwischen den zwei BegutachterInnen der Studie oder über die Beiziehung einer dritten, unabhängigen Person gelöst.

Für jede inkludierte Studie galt es folgende Charakteristika und Inhalte der Studien zu extrahieren: Beurteilung des Biasrisikos, detaillierte Informationen zur Studie (Autor, Jahr,

Finanzierung), Population, Dauer der Studie bzw. Beobachtungszeitraum, Studiendesign und Stichprobengröße, Test- und Referenzstandard sowie Endpunkte.

2.5 Beurteilung der methodischen Validität

Analog zur Vorgehensweise bei der Datenextraktion lagen auch für die Beurteilung des Biasrisikos der inkludierten Studien standardisierte Beurteilungsformulare vor. Das Instrument zur Beurteilung von diagnostischen Studien orientiert sich am QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2) Instrument.³ Das Instrument zur Beurteilung von RCTs (Randomisiert kontrollierten Studien) basiert auf dem Cochrane Risk of Bias Tool.⁴ Die Beurteilung der einzelnen Kriterien wurde von zwei BegutachterInnen unabhängig voneinander durchgeführt und resultierte in einer zusammenfassenden Bewertung des Biasrisikos jeder Studie in den Kategorien *niedriges, unklares und hohes Biasrisiko*⁴ (Tabelle 2). Unstimmigkeiten in der Beurteilung lösten die BegutachterInnen via Diskussion und Konsens.

Tabelle 2: Definition und Interpretation des Biasrisikos

Biasrisiko	Voraussetzung	Interpretation
Geringes Biasrisiko	Alle elementaren Komponenten zeigen ein geringes Biasrisiko	Es ist unwahrscheinlich , dass Bias das Ergebnis der Studie wesentlich verzerrt
Unklares Biasrisiko	Das Biasrisiko ist für eine oder mehrere Komponenten unklar.	Es ist unklar , inwieweit Bias die Ergebnisse der Studie verzerren könnte. Bias ist möglich und könnte die Korrektheit der Resultate in Frage stellen.
Hohes Biasrisiko	Das Biasrisiko ist für eine oder mehrere Komponenten hoch.	Es ist sehr wahrscheinlich , dass Bias das Ergebnis der Studie wesentlich verzerrt. Das Vertrauen in die Korrektheit der Resultate ist sehr gering.

2.6 Auswahl relevanter Endpunkte

In einem iterativen Prozess wurden gemeinsam mit den AuftraggeberInnen Endpunkte ausgewählt, die sowohl für die PatientInnen als auch für die Entscheidungsfindung im klinischen Alltag wesentlich sind. Die Auswahl fiel auf folgende Endpunkte:

- Sensitivität
- Spezifität
- Prädiktive Werte
- Strahlenbelastung

2.7 Beurteilung der Qualität der Evidenz

Die Beurteilung der Qualität der gesamten Evidenz basiert auf dem Ansatz der GRADE (Grading Recommendations Assessment, Development and Evaluation) Working Group.^{5,6}

Die Qualität der Evidenz wurde getrennt für wesentliche mit den Auftraggebern abgestimmte Endpunkte bewertet. Folgende Bereiche wurden für die Bewertung der Qualität der Evidenz berücksichtigt: Studiendesign, Biasrisiko, Indirektheit, Inkonsistenz, Impräzision, Publikationsbias.⁶

Die gesammelte Beurteilung der Qualität der Evidenz erfolgte in vier Kategorien:⁷

Hohe Qualität	Wir sind sehr sicher, dass der wahre Effekt nahe bei dem Effektschätzer liegt.
Moderate Qualität	Wir haben mäßig viel Vertrauen in den Effektschätzer: der wahre Effekt ist wahrscheinlich nahe bei dem Effektschätzer, aber es besteht die Möglichkeit, dass er relevant verschieden ist.
Niedrige Qualität	Unser Vertrauen in den Effektschätzer ist begrenzt: Der wahre Effekt kann durchaus relevant verschieden vom Effektschätzer sein.
Sehr niedrige Qualität	Wir haben nur sehr wenig Vertrauen in den Effektschätzer: Der wahre Effekt ist wahrscheinlich relevant verschieden vom Effektschätzer.

Die vorhandene Evidenz von Forschungsfrage 1 wurde nach GRADE für diagnostische Studien⁵ zusammengefasst und als Evidenzprofil dargestellt.

2.8 Synthese der Evidenz

Die vorhandene Evidenz wurde aufgrund der unzureichenden Studienlage deskriptiv zusammengefasst. In der Darstellung werden neben Informationen zu Studiengröße, Studiendauer und Biasrisiko auch Effektmaße wie Testgüteparameter, die Konsistenz der Resultate der verschiedenen Studien und Faktoren, die die externe Validität einschränken können, erläutert. Wenn mehrere Studien unterschiedlicher Qualität vorlagen, wurde der Fokus auf die Studie mit der höchsten Aussagekraft gelegt und genauer beschrieben.

3 Resultate Forschungsfrage 1

Insgesamt identifizierten unsere Literatursuchen 1553 relevante Abstracts. Von diesen wurden 256 als Volltexte eingeschlossen und einer genaueren Begutachtung unterzogen. Vier Studien entsprachen den a priori definierten Einschlusskriterien.

Abbildung 2 stellt den Verlauf der Literaturdurchsicht grafisch dar. Die Gründe für den Ausschluss von Studien auf Volltextebene sind in Appendix 3 dargestellt.

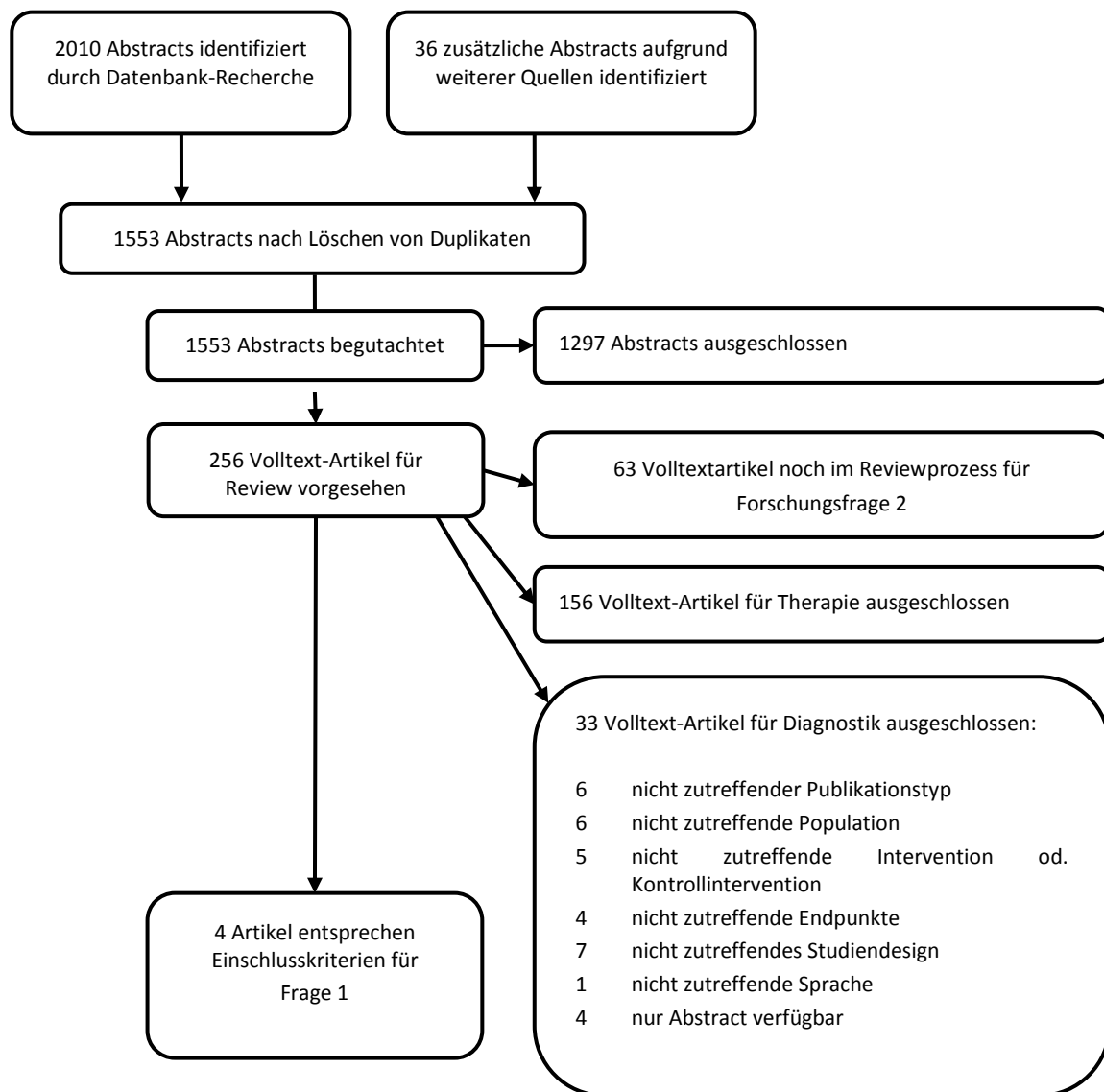


Abbildung 2: Darstellung der Durchsicht der Literatur

3.1 Primärdiagnostische Abklärung des Nierenzellkarzinoms

Im folgenden Kapitel fassen wir die vorhandene Evidenz aus prospektiven diagnostischen Studien zur Primärdiagnostik eines Tumors an der Niere zusammen. Wir stellen die Evidenz geordnet nach Untersuchungen dar. Am Ende des jeweiligen Unter-Kapitels wird die Stärke der Evidenz in Bezug auf die wesentlichsten Endpunkte in Form von GRADE Tabellen dargestellt.

Insgesamt wurden vier diagnostische Studien eingeschlossen.⁸⁻¹¹ Zwei Studien verglichen Sonographie mit Computertomographie hinsichtlich der gesuchten Testgüteparameter.^{8,11} Jeweils eine Studie verglich PET Scan¹⁰ bzw. Magnetresonanztomographie⁹ mit Computertomographie hinsichtlich der gesuchten Testgüteparameter. Wir konnten keine prospektive diagnostische Studie identifizieren, die Magnetresonanzangiographie mit Computertomographie verglich.

Eine detaillierte Darstellung des Biasrisikos der Studien erfolgt in Abbildung 3. Drei der eingeschlossenen diagnostischen Studien wurden mit einem hohen Biasrisiko beurteilt⁸⁻¹⁰, eine Studie mit unklarem Biasrisiko.¹¹

Keine der eingeschlossenen Studien machte Angaben über Sponsoren.

	Studie repräsentativ (Selection)	Auswahlkriterien klar definiert (Selection)	Referenztest ist "Goldstandard" (Vergleichbarkeit)	Indextest und Referenztest zum selben Zeitpunkt (Vergleichbarkeit)	Referenztest an der gesamten Studienpopulation durchgeführt (Vergleichbarkeit)	Referenztest unabhängig vom Resultat des Indextests (Vergleichbarkeit)	Indextest nicht Bestandteil des Referenztests (Vergleichbarkeit)	Beschreibung der Testdurchführung ermöglicht Replikation (Vergleichbarkeit)	Ergebnisse des Indextests unabhängig vom Referenztests interpretiert (Endpunkte)	Ergebnisse des Referenztests unabhängig vom Indextests interpretiert (Endpunkte)	Bedingungen bei Testinterpretation generalisierbar	Nicht interpretierbare Resultate berichtet (Endpunkte)	Drop-out Rate während der Tests berichtet
Aide 2003 ¹⁰	?	-	+	?	+	-	+	+	-	?	+	-	-
Beer 2006 ⁹	+	+	+	-	?	-	?	+	+	?	+	?	?
Quaia 2008 ⁸	+	+	+	?	-	-	?	+	+	?	+	?	?
Schmidt 2002 ¹¹	+	?	+	?	-	+	+	+	+	+	-	?	-

Abbildung 3: Darstellung des Biasrisikos der eingeschlossenen Studien.

3.1.1 Primärdiagnostische Abklärung der Nieren mittels Sonographie vs. Computertomographie

Für die Fragestellung über die primärdiagnostische Abklärung einer suspekten Raumforderung an der Niere mittels Sonographie vs. Computertomographie schlossen wir zwei diagnostische Studien, mit Daten über insgesamt 154 PatientInnen, ein.^{8,11} Tabelle 3 fasst die Charakteristika und Ergebnisse der eingeschlossenen Studien zusammen.

Tabelle 3: Darstellung der wesentlichsten Charakteristika und Resultate der eingeschlossenen Studien.

Studie; Studiendesign; Sponsor	Population (n); Alter (Jahre ± SA); % weiblich	Interventionen (Test) vs. Kontrolle (Referenztest, Goldstandard), technische Details	Relevante Endpunkte; Studiendauer	Resultate	Bias-Risiko
Quaia et al. (2008) ⁸ ; diagnostische Studie; kA	PatientInnen (40) mit komplexen Nieren-Zysten; durchschn. Alter: 62 ± 11; 55% weiblich	Nicht verstärkter Ultraschall vs. Kontrastmittel-verstärkter Ultraschall vs. CT oder MDCT Referenztest: Histologie (30) oder Follow-up (10) <i>Unverstärkter Ultraschall:</i> Sequoia, Acuson-Siemens oder EsaTune, Esaote, 2-5-MHz-Breitband Konvex-Array- Schallkopf <i>Kontrastmittelverstärkter Ultraschall:</i> Pulssequenzverfahren (Cadence Contrast-Pulse-Sequence, Acuson- Siemens Medical Solutions) mit Mikrobläschen oder Contrast Tuned Imaging (Esaote) mit kontrast- spezifischen Modi, Sendeleistung Beschallungswinkel (Mechanischer Index, 0,09-0,14), Dynamikumfang: 65dB, zeitliche Auflösung 75-100 Millisekunden (10-13 frames pro Sekunde), geringst mögliche frame-Rate (7-9Hz) <i>CT:</i> Single-detector Scanner TomoScan Avel, Philips Healthcare, Strahl- Kollimation: 5mm & Pitchfaktor 1,5 <i>MDCT Scanner:</i> Aquilion, Toshiba Medical Systems, Rotationszeit: 400 Millisekunden, Strahl-Kollimation: 64x0,5mm, normalisierter Pitch 1, Rekonstruktionsintervall: 3mm, Röhrenspannung: 120 kVp, Röhrenstrom :180-250 mAs	Sensitivität, Spezifität, Positiver prädiktiver Wert (PPW), Negativer prädiktiver Wert (NPW), Accuracy; kA	21 von 40 komplexen zystischen Läsionen waren Karzinome Nicht verstärkter Ultraschall: Sensitivität: 43-48%* Spezifität: 11-16%* PPW: 36-37%* NPW: 15-20%* Accuracy: 30-30%* Kontrastmittel verstärkter Ultraschall: Sensitivität: 86-95%* Spezifität: 63-79%* PPW: 74-82%* NPW: 83-92%* Accuracy: 80-83%* CT oder MDCT: Sensitivität: 81-95%* Spezifität: 42-68%* PPW: 61-74%* NPW: 67-89%* Accuracy: 63-75%* * Angabe als Range (minimaler Wert – maximaler Wert) von 3 Befundern	hoch

Nierenzellkarzinom Primärdiagnostik

Studie; Studiendesign; Sponsor	Population (n); Alter (Jahre ± SA); % weiblich	Interventionen (Test) vs. Kontrolle (Referenztest, Goldstandard), technische Details	Relevante Endpunkte; Studiendauer	Resultate	Bias-Risiko
Schmidt (2002) ¹¹ ; diagnostische Studie; kA	Erwachsene PatientInnen (114) mit fokalen Läsionen an der Niere; durchschn. Alter: 62 ± 13; 35% weiblich	Phase-Inversion Tissue Harmonic sonographic Imaging (Test) vs. B-Mode Ultraschall (Test) CT oder MR (wenn vorhanden auch Histologie) (Referenzstandard) Ultraschall: Sonoline Elegra (Software Version 6.0; Siemens) mit 3.5-MHz Konvex-Schallkopf Für B-Mode Ultraschall: 3,4 MHz Frequenz; Für Phase-inversion tissue harmonic sonographic imaging: Frequenzkombination von 2,0 und 4,0 MHz CT & MR: kA	Testgüteparameter: Sensitivität; Accuracy; hinsichtlich Läsionsdiagnostik; 10 Monate	88 Patienten mit 273 fokalen Läsionen an 127 Nieren (219 unkomplizierte Zysten, 15 komplizierte Zysten, 26 solide Tumoren und 13 andere Läsionen) Phase-Inversion Tissue Harmonic Imaging*: <ul style="list-style-type: none"> • Sensitivität (Läsion detektiert unabhängig von der Richtigkeit der Diagnose): 88 ± 8%[‡] (57 von 65 Läsionen ≥ 10mm detektiert) • Spezifität (Läsionsfreiheit richtig erkannt, auf Nierenebene): 97% • Accuracy (Diagnose konsistent mit Referenztest): 83 ± 9%[‡] B Mode Ultraschall*: <ul style="list-style-type: none"> • Sensitivität (Läsion detektiert unabhängig von der Richtigkeit der Diagnose): 66 ± 9%[‡] (73 von 111 Läsionen ≥ 10mm detektiert) • Spezifität (Läsionsfreiheit richtig erkannt, auf Nierenebene): 85% • Accuracy (Diagnose konsistent mit Referenztest): 64 ± 9%[‡] * Testgüteparameter, die sich auf die jeweilige Verwendung als „erste Untersuchungsart“ bei sequentieller Untersuchung bezogen. [‡] 95% Konfidenz Intervalle	unklar

Abkürzungen: CT, Computertomographie; dB, Dezibel; Hz, Hertz; kA, keine Angaben; kVp, Peak Kilovoltage; MDCT, Multidetector Computertomographie; mAs, Milliamperesekunden; mHz, Megahertz; mm, Millimeter; MRI, Magnetresonanztomographie; NPW, Negativer prädiktiver Wert; PPW, Positiver prädiktiver Wert; SA, Standardabweichung;

Eine Studie⁸ untersuchte die Testgenauigkeit von Ultraschall vor und nach intravenöser Injektion von Kontrastmittel (Schwefelhexafluorid Mikrobläschen) im Vergleich zur Testgenauigkeit der Computertomographie (technische Details siehe Tabelle 3) mit Kontrastmittel bei 40 PatientInnen mit komplexen Nierenzysten, die wegen möglicher Malignität weiter abgeklärt wurden. Die Untersuchungen wurden von jeweils drei UntersucherInnen unabhängig voneinander befundet. Die UntersucherInnen hatten zwei, sechs bzw. zehn Jahre Erfahrung in der radiologischen Diagnostik an der Niere. Als Referenztest wurde in 30 Fällen die histologische Diagnose und in zehn Fällen das radiologische „Follow-up“ herangezogen. Das Biasrisiko dieser Studie wurde als „hoch“ eingestuft, da verschiedene Referenztests angewandt wurden und Referenztest und Indextest nicht unabhängig voneinander durchgeführt wurden.

Die zweite Studie¹¹ untersuchte die Testgenauigkeit von zwei verschiedenen Ultraschall-Verfahren (Phase-Inversion Tissue Harmonic Imaging und B-Mode Ultraschall, technische Details siehe Tabelle 3) im Vergleich zur Testgenauigkeit der Computertomographie bei 114 PatientInnen mit fokalen Läsionen an der Niere. Die Reihenfolge der beiden Ultraschallverfahren erfolgte zufällig. In 70 Fällen wurde B-Mode Ultraschall und in 44 Fällen wurde Phase-Inversion Tissue Harmonic Imaging als erste Untersuchungsart durchgeführt. Es wurden jeweils nur die Testgüteparameter extrahiert, die sich auf die jeweilige Verwendung als „erste Untersuchungsart“ bei sequentieller Untersuchung bezogen. Denn bei der zweiten Untersuchung in einer Sequenz sind die beiden Untersuchungen nicht mehr unabhängig. Die Posttest-Wahrscheinlichkeit der ersten Untersuchung wird dann zur Prätest-Wahrscheinlichkeit der zweiten Untersuchung. Testgüteparameter sind somit nicht mehr direkt vergleichbar. Als Referenztest wurde entweder die Computertomographie, die Magnetresonanztomographie und bei Vorhandensein auch die Histologie verwendet. Die genaue Anzahl der Fälle pro jeweiliges Referenzverfahren wurde in der Studie nicht angegeben. Das Biasrisiko dieser Studie wurde als „unklar“ eingestuft, da verschiedene Referenztests angewandt wurden.

Die erhobenen Endpunkte in diesen zwei Arbeiten beschränkten sich auf Testgüteparameter. In der ersten Studie fanden sich keine Angaben zur Studiendauer. Die Studiendauer der zweiten Studie lag bei 10 Monaten.

Insgesamt zeigten die Ergebnisse der ersten Studie⁸, dass die diagnostische Testgüte (Accuracy) hinsichtlich der Diagnostik von Malignitäten bei komplexen zystischen Läsionen beim Kontrastmittel-verstärkten Ultraschall besser war als beim nicht verstärkten Ultraschall und bei der Computertomographie (Kontrastmittel-verstärkter Ultraschall vs. nicht verstärkter Ultraschall vs. Computertomographie: Untersucher 1, 83% vs. 30% vs. 75%; Untersucher 2, 83% vs. 30% vs. 63%; Untersucher 3, 80% vs. 30% vs. 70%; $p < 0,05$). Auch das

positive Wahrscheinlichkeitsverhältnis (Sensitivität/1-Spezifität) war bei allen drei UntersucherInnen beim kontrastverstärkten Ultraschall signifikant besser als bei den anderen beiden Untersuchungsarten ($p < 0,05$). Das positive Wahrscheinlichkeitsverhältnis (Sensitivität/1-Spezifität) des kontrastverstärkten Ultraschalls der drei UntersucherInnen war hinsichtlich der Malignitätserkennung bei komplexen Nierenzysten ähnlich der diagnostischen Vertrauenswürdigkeit der Computertomographie mit Kontrastmittel ($p > 0,05$).

Die Ergebnisse der zweiten Studie¹¹ zeigten, dass bei Läsionen eines Durchmessers von 10mm oder mehr, sowohl die diagnostische Testgüte (Accuracy) als auch die Sensitivität des Phase-Inversion Tissue Harmonic Imaging Ultraschallverfahrens statistisch signifikant besser war als die Treffsicherheit der Diagnose des B-Mode Ultraschallverfahrens (Accuracy: 83% vs. 64%; Sensitivität: 88% vs. 66%; p-Wert nicht berichtet). Bei Einschluss von Läsionen mit einem Durchmesser von < 10 mm war der Unterschied nicht mehr statistisch signifikant.

Ein Poolen der Ergebnisse der beiden Studien war aufgrund der Heterogenität der Untersuchungsarten, sowie der eingeschlossenen Populationen nicht möglich. Die Aussagekraft der Ergebnisse zu den einzelnen Testgüteparametern ist aufgrund der geringen Fallzahlen als eingeschränkt zu sehen.

Im Folgenden werden die einzelnen Endpunkte näher beschrieben. Eine Zusammenfassung der Stärke der Evidenz nach GRADE wird in Tabelle 4 und Tabelle 5 dargestellt.

3.1.1.1 Sensitivität

Der Anteil der richtig erkannten malignen Läsionen bei komplexen Nierenzysten (d.h. die Sensitivität) war beim Kontrastmittel-verstärkten Ultraschall 86%, 86% und 95% (Sensitivität Untersucher 1-3), bei der Computertomographie 81%, 81% und 95% (Sensitivität Untersucher 1-3) und beim nicht verstärkten Ultraschall 48%, 48% und 43% (Sensitivität Untersucher 1-3) (Studienpopulation $n=40$).⁸

Angaben zur Sensitivität des Phase-Inversion Tissue Harmonic Imaging Ultraschallverfahrens (88%) und des B-Mode Ultraschallverfahrens (66%) beziehen sich lediglich auf die Detektion einer fokalen Läsion, unabhängig von der Richtigkeit der Diagnose hinsichtlich Malignität (Studienpopulation $n=114$).¹¹

3.1.1.2 Spezifität

Der Anteil der richtig erkannten Benignität bei komplexen Nierenzysten (d.h. die Spezifität) war beim Kontrastmittel-verstärkten Ultraschall 79%, 79% und 63% (Spezifität UntersucherIn

1-3) und bei der Computertomographie 68%, 42% und 42% (Spezifität UntersucherIn 1-3). Die Spezifität des nicht verstärkten Ultraschalls war 11%, 11% und 16% (Spezifität UntersucherIn 1-3) (Studienpopulation n=40).⁸

Angaben zur Spezifität des Phase-Inversion Tissue Harmonic Imaging Ultraschallverfahrens (97%) und des B-Mode Ultraschallverfahrens (85%) beziehen sich auf das richtige Erkennen einer Läsionsfreiheit, nicht auf Benignität bzw. Malignität (Studienpopulation n=114).¹¹

3.1.1.3 Diagnostische Testgüte (Accuracy)

Die diagnostische Testgüte bei komplexen Nierenzysten (d.h. die Accuracy) war beim Kontrastmittel-verstärkten Ultraschall 83%, 83% und 80% (Accuracy UntersucherIn 1-3), bei der Computertomographie 75%, 63% und 70% (Accuracy UntersucherIn 1-3) und beim nicht verstärkten Ultraschall 30%, 30% und 30% (Accuracy UntersucherIn 1-3) (Studienpopulation n=40).⁸

Angaben zur Accuracy des Phase-Inversion Tissue Harmonic Imaging Ultraschallverfahrens (83%) und des B-Mode Ultraschallverfahrens (64%) beziehen sich auf die Konsistenz der Diagnose im Vergleich zum Referenztest (Studienpopulation n=114).¹¹

3.1.1.4 Positiver prädiktiver Wert

Die Wahrscheinlichkeit eines tatsächlich malignen Befundes bei positivem Untersuchungsbefund (d.h. positiver prädiktiver Wert [PPW]) war beim Kontrastmittel-verstärkten Ultraschall 82%, 82% und 74% (PPW UntersucherIn 1-3), bei der Computertomographie 74%, 61% und 65% (PPW UntersucherIn 1-3) und beim nicht verstärkten Ultraschall 37%, 37% und 36% (PPW UntersucherIn 1-3) (Studienpopulation n=40).⁸

3.1.1.5 Negativer prädiktiver Wert

Die Wahrscheinlichkeit eines tatsächlich benignen Befundes bei negativem Untersuchungsbefund (d.h. negativer prädiktiver Wert [NPW]) war beim Kontrastmittel-verstärkten Ultraschalls 83%, 83% und 92% (NPW UntersucherIn 1-3), bei der Computertomographie 76%, 67% und 89% (NPW UntersucherIn 1-3) und beim nicht verstärkten Ultraschall 15%, 15% und 20% (NPW UntersucherIn 1-3) (Studienpopulation n=40).⁸

3.1.1.6 Strahlenbelastung

Zu diesem Endpunkt, der sich in diesem Fall lediglich auf das Vergleichsverfahren Computertomographie beziehen würde, wurde keine Evidenz gefunden.

Tabelle 4: GRADE Evidenz Profil - Kontrastmittelverstärkter Ultraschall

Population: PatientInnen mit komplexen Nierenzysten

Test: Kontrastmittelverstärkter Ultraschall

Referenzstandard: Histologie oder Follow-up

Outcomes	Anzahl der Studien (PatientInnen)	Studiendesign	Faktoren die möglicherweise die Qualität der Evidenz einschränken					Stärke der Evidenz (GRADE)	Effekt (Effekt pro 1000)	Wichtigkeit
			Biasrisiko	Indirektheit	Inkonsistenz	Impräzision	Publikationsbias			
Sensitivität	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊖ Sehr niedrig	86%-95% (860-950 korrekt positive Testresultate pro 1000 Erkrankten)	Kritisch
Spezifität	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊖ Sehr niedrig	63%-79% (630-790 korrekt negative Testresultate pro 1000 Gesunden)	Kritisch
PPW	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊖ Sehr niedrig	74%-82% (740-820 korrekte Testresultate pro 1000 Personen mit einem positiven Test)	Kritisch
NPW	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊖ Sehr niedrig	83%-92% (830-920 korrekte Testresultate pro 1000 Personen mit einem negativen Test)	Kritisch
Accuracy	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊖ Sehr niedrig	80%-83% (800-830 korrekte Testresultate pro 1000 getesteten Personen) ³	Kritisch
Strahlenbelastung	Keine Evidenz								Wichtig	

Abkürzungen: n.a., nicht anwendbar; NPW, negativer prädiktiver Wert; PPW, positiver prädiktiver Wert;

¹ Indextest und Referenztest wurden nicht unabhängig voneinander durchgeführt; Verwendung unterschiedlicher Indextests

² Insgesamt nur 40 Personen eingeschlossen, um 2 Grade herabgestuft

³ Angaben als Range (minimaler Wert – maximaler Wert) von 3 UntersucherInnen

Legende Stärke der Evidenz (GRADE):

⊕⊕⊕⊕ Hoch

⊕⊕⊕⊖ Moderat

⊕⊕⊖⊖ Niedrig

⊕⊖⊖⊖ Sehr niedrig

Tabelle 5: GRADE Evidenz Profil - Nicht-verstärkter Ultraschall

Population: PatientInnen mit komplexen Nierenzysten

Test: Nicht-verstärkter Ultraschall

Referenzstandard: Histologie oder Follow-up

Outcomes	Anzahl der Studien (PatientInnen)	Studiendesign	Faktoren die möglicherweise die Qualität der Evidenz einschränken					Stärke der Evidenz (GRADE)	Effekt (Effekt pro 1000)	Wichtigkeit
			Biasrisiko	Indirektheit	Inkonsistenz	Impräzision	Publikationsbias			
Sensitivität	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊖ Sehr niedrig	43%-48% (430-480 korrekt positive Testresultate pro 1000 Erkrankten)	Kritisch
Spezifität	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊖ Sehr niedrig	11%-16% (110-160 korrekt negative Testresultate pro 1000 Gesunden)	Kritisch
PPW	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊖ Sehr niedrig	36%-37% (360-370 korrekte Testresultate pro 1000 Personen mit einem positiven Test)	Kritisch
NPW	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊖ Sehr niedrig	15%-20% (150-200 korrekte Testresultate pro 1000 Personen mit einem negativen Test)	Kritisch
Accuracy	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊖ Sehr niedrig	30% (300 korrekte Testresultate pro 1000 getesteten Personen) ³	Kritisch
Strahlenbelastung	Keine Evidenz								Wichtig	

Abkürzungen: n.a., nicht anwendbar; NPW, negativer prädiktiver Wert; PPW, positiver prädiktiver Wert;

¹ Indextest und Referenztest wurden nicht unabhängig voneinander durchgeführt; Verwendung unterschiedlicher Indextests

² Insgesamt nur 40 Personen eingeschlossen, um 2 Grade herabgestuft

³ Angaben als Range (minimaler Wert – maximaler Wert) von 3 UntersucherInnen

Legende Stärke der Evidenz (GRADE):

⊕⊕⊕⊕ Hoch

⊕⊕⊕⊖ Moderat

⊕⊕⊖⊖ Niedrig

⊕⊖⊖⊖ Sehr niedrig

Tabelle 6: GRADE Evidenz Profil - Computertomographie

Population: PatientInnen mit komplexen Nierenzysten

Test: Computertomographie

Referenzstandard: Histologie oder Follow-up

Outcomes	Anzahl der Studien (PatientInnen)	Studiendesign	Faktoren die möglicherweise die Qualität der Evidenz einschränken					Stärke der Evidenz (GRADE)	Effekt (Effekt pro 1000)	Wichtigkeit
			Biasrisiko	Indirektheit	Inkonsistenz	Impräzision	Publikationsbias			
Sensitivität	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	81%-95% (810-950 korrekt positive Testresultate pro 1000 Erkrankten)	Kritisch
Spezifität	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	42%-68% (420-680 korrekt negative Testresultate pro 1000 Gesunden)	Kritisch
PPW	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	61%-74% (610-740 korrekte Testresultate pro 1000 Personen mit einem positiven Test)	Kritisch
NPW	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	67%-89% (670-890 korrekte Testresultate pro 1000 Personen mit einem negativen Test)	Kritisch
Accuracy	1 Studie (40 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	63%-75% (630-750 korrekte Testresultate pro 1000 getesteten Personen) ³	Kritisch
Strahlenbelastung	Keine Evidenz								Wichtig	

Abkürzungen: n.a., nicht anwendbar; NPW, negativer prädiktiver Wert; PPW, positiver prädiktiver Wert;

¹ Indextest und Referenztest wurden nicht unabhängig voneinander durchgeführt; Verwendung unterschiedlicher Indextests

² Insgesamt nur 40 Personen eingeschlossen, um 2 Grade herabgestuft

³ Angaben als Range (minimaler Wert – maximaler Wert) von 3 UntersucherInnen

Legende Stärke der Evidenz (GRADE):

⊕⊕⊕⊕ Hoch

⊕⊕⊕⊖ Moderat

⊕⊕⊖⊖ Niedrig

⊕⊖⊖⊖ Sehr niedrig

3.1.2 Primärdiagnostische Abklärung der Nieren mittels Magnetresonanztomographie vs. Computertomographie

Wir schlossen für diese Fragestellung über die primärdiagnostische Abklärung einer suspekten Raumforderung mittels Magnetresonanztomographie an der Niere eine diagnostische Studie ein.⁹ Tabelle 7 fasst die Charakteristika und Ergebnisse der eingeschlossenen Studie zusammen.

Die Studie⁹ untersuchte die Testgenauigkeit der Magnetresonanztomographie im Vergleich zur Testgenauigkeit der Multidetector-Computertomographie (technische Details siehe Tabelle 7 mit Kontrastmittel bei 28 PatientInnen mit 56 renalen Läsionen, die im Ultraschall entdeckt wurden. Als Referenztest wurde bei 35 Läsionen die Histologie und bei 10 Läsionen das klinische bzw. radiologische „Follow-up“ herangezogen. 16 der renalen Läsionen waren Nierenzellkarzinome. Das Biasrisiko dieser Studie wurde als „hoch“ eingestuft, da verschiedene Referenztests angewandt wurden und Referenztest und Indextest nicht unabhängig voneinander durchgeführt wurden.

Die erhobenen Endpunkte dieser Studie beschränkten sich auf Testgüteparameter. Die Studiendauer lag bei 9 Monaten.

Zusammenfassend zeigt das Ergebnis sowohl für die Magnetresonanztomographie wie auch für die Computertomographie eine Sensitivität von 93% bei eingeschränkter Spezifität (71% und 68%).

Im Folgenden werden die einzelnen Endpunkte näher beschrieben. Eine Zusammenfassung der Stärke der Evidenz nach GRADE wird in Tabelle 8 und Tabelle 9 dargestellt.

Tabelle 7: Darstellung der wesentlichsten Charakteristika und Resultate der eingeschlossenen Studie.

Studie; Studiendesign; Sponsor	Population (N); Alter (Jahre ± SA); % weiblich	Interventionen (Test) vs. Kontrolle (Referenztest, Goldstandard), Technische Details	Relevante Endpunkte; Studiendauer	Resultate	Bias-Risiko
Beer et al. (2006) ⁹ ; diagnostische Studie; kA	PatientInnen (28) mit renalen Läsionen (56); durchschn. Alter: 65 ± 14; 29% weiblich	MDCT vs. MRI Goldstandard: Histologie (35 Läsionen); Follow-up (21 Läsionen) MDCT: 16-MDCT Scanner (Somatom Sensation 16, Siemens Medical Solutions), Kollimation: 16x0,75mm bei 120 kV und 200mAs mit CareDose Modus (Siemens Medical Solutions) MRI: 1,5-T supraleitendes bildgebendes Verfahren (Gyroscan Intera, Philips Medical System) Gradientensystem mit Amplitude: 23mT/m in 200 µ sek.	Sensitivität, Spezifität, Positiver prädiktiver Wert (PPW), Negativer prädiktiver Wert (NPW); Dauer: 9 Monate, Follow-up 6-12 Monate	16 der 56 Läsionen waren Nierenzellkarzinome MDCT: (auf Läsionen-Ebene) Sensitivität 93% (15 von 16 RCC) Spezifität: 68% (26 von 38 gutartigen Läsionen) PPW: 56% (15/27) NPW: 96% (26/27) MRI: (auf Läsionen-Ebene)* Sensitivität:93% (15 von 16 RCC) Spezifität: 71% (25 von 35 gutartigen Läsionen) PPW:56% (15/27) NPW: 96% (26/27) * Es geht aus der Arbeit nicht klar hervor, warum sich die Anzahl der benignen Läsionen beim MRI von der Anzahl der benignen Läsionen bei der CT unterscheidet. Diesbezüglich besteht eine gewisse Unsicherheit bezüglich der Werte PPW und NPW. (Autorenkontakt siehe Appendix 2)	hoch

Abkürzungen: CT, Computertomographie; kA, keine Angaben; MDCT, Multidetector Computertomographie; MRI, Magnetresonanztomographie; NPW, Negativer prädiktiver Wert; PPW, Positiver prädiktiver Wert; RCC, Nierenzellkarzinom; SA, Standardabweichung;

3.1.2.1 Sensitivität

Der Anteil der richtig erkannten malignen Läsionen (d.h. die Sensitivität) war sowohl bei der Computertomographie als auch bei der Magnetresonanztomographie 93%.

3.1.2.2 Spezifität

Der Anteil der richtig erkannten benignen Läsionen (d.h. die Spezifität) war 68% bei der Computertomographie und 71% bei der Magnetresonanztomographie.

3.1.2.3 Positiver prädiktiver Wert

Die Wahrscheinlichkeit eines tatsächlich malignen Befundes bei positivem Untersuchungsbefund (d.h. positiver prädiktiver Wert) wurden sowohl bei der Computertomographie als auch bei der Magnetresonanztomographie mit 56% angegeben. Diese Zahlen sind aufgrund der Unterschiede in Sensitivität und Spezifität allerdings nicht plausibel. Eine Kontaktaufnahme mit den Autoren zur Klarstellung war nicht erfolgreich.

3.1.2.4 Negativer prädiktiver Wert

Die Wahrscheinlichkeit eines tatsächlich benignen Befundes bei negativem Untersuchungsbefund (d.h. negativer prädiktiver Wert) wurde in der Studie sowohl bei der Computertomographie als auch bei der Magnetresonanztomographie mit 96% angegeben. Auch diese Zahlen sind aufgrund der Unterschiede in Sensitivität und Spezifität nicht plausibel. Eine Kontaktaufnahme mit den Autoren zur Klarstellung war nicht erfolgreich.

3.1.2.5 Strahlenbelastung

Zu diesem Endpunkt wurde keine Evidenz gefunden.

Tabelle 8: GRADE Evidenz Profil - Magnetresonanztomographie

Population: PatientInnen mit renalen Läsionen, die im Ultraschall entdeckt wurden

Test: Magnetresonanztomographie

Referenzstandard: Histologie oder Follow-up

Outcomes	Anzahl der Studien (PatientInnen)	Studiendesign	Faktoren die möglicherweise die Qualität der Evidenz einschränken					Stärke der Evidenz (GRADE)	Effekt (Effekt pro 1000)	Wichtigkeit
			Biasrisiko	Indirektheit	Inkonsistenz	Impräzision	Publikationsbias			
Sensitivität	1 Studie (28 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	93% (930 korrekt positive Testresultate pro 1000 Läsionen)	Kritisch
Spezifität	1 Studie (28 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	71% (710 korrekt negative Testresultate pro 1000 Gesunden)	Kritisch
PPW	1 Studie (28 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	56% (560 korrekte Testresultate pro 1000 Läsionen mit einem positiven Test) ³	Kritisch
NPW	1 Studie (28 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	96% (960 korrekte Testresultate pro 1000 Läsionen mit einem negativen Test) ³	Kritisch
Accuracy	Keine Evidenz								Kritisch	
Strahlenbelastung	Keine Evidenz								Wichtig	

Abkürzungen: n.a., nicht anwendbar; NPW, negativer prädiktiver Wert; PPW, positiver prädiktiver Wert;

¹ Indextest und Referenztest wurden nicht unabhängig voneinander durchgeführt; Verwendung unterschiedlicher Indextests

² Insgesamt nur 28 Personen eingeschlossen, um 2 Grade herabgestuft

³ Es geht aus der Arbeit nicht klar hervor, warum sich die Anzahl der benignen Läsionen beim MRI von der Anzahl der benignen Läsionen bei der CT unterscheidet. Diesbezüglich besteht eine gewisse Unsicherheit bezüglich der Werte PPW und NPW. (Autorenkontakt siehe Appendix 2)

Legende Stärke der Evidenz (GRADE):

⊕⊕⊕⊕ Hoch

⊕⊕⊕⊖ Moderat

⊕⊕⊖⊖ Niedrig

⊕⊖⊖⊖ Sehr niedrig

Tabelle 9: GRADE Evidenz Profil – Multidetector Computertomographie

Population: PatientInnen mit renalen Läsionen, die im Ultraschall entdeckt wurden

Test: Multidetector-Computertomographie

Referenzstandard: Histologie oder Follow-up

Outcomes	Anzahl der Studien (PatientInnen)	Studiendesign	Faktoren die möglicherweise die Qualität der Evidenz einschränken					Stärke der Evidenz (GRADE)	Effekt (Effekt pro 1000)	Wichtigkeit
			Biasrisiko	Indirektheit	Inkonsistenz	Impräzision	Publikationsbias			
Sensitivität	1 Studie (28 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	93% (930 korrekt positive Testresultate pro 1000 Läsionen)	Kritisch
Spezifität	1 Studie (28 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	68% (680 korrekt negative Testresultate pro 1000 Gesunden)	Kritisch
PPW	1 Studie (28 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	56% (560 korrekte Testresultate pro 1000 Läsionen mit einem positiven Test) ³	Kritisch
NPW	1 Studie (28 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	96% (960 korrekte Testresultate pro 1000 Läsionen mit einem negativen Test) ³	Kritisch
Accuracy	1 Studie (28 PatientInnen)	Nicht-konsequente Serie, verblindet.	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	Keine Angaben	Kritisch
Strahlenbelastung	Keine Evidenz									Wichtig

Abkürzungen: n.a., nicht anwendbar; NPW, negativer prädiktiver Wert; PPW, positiver prädiktiver Wert;

¹ Indextest und Referenztest wurden nicht unabhängig voneinander durchgeführt; Verwendung unterschiedlicher Indextests

² Insgesamt nur 28 Personen eingeschlossen, um 2 Grade herabgestuft

³ Es geht aus der Arbeit nicht klar hervor, warum sich die Anzahl der benignen Läsionen beim MRI von der Anzahl der benignen Läsionen bei der CT unterscheidet. Diesbezüglich besteht eine gewisse Unsicherheit bezüglich der Werte PPW und NPW. (Autorenkontakt siehe Appendix 2)

Legende Stärke der Evidenz (GRADE):

⊕⊕⊕⊕ Hoch

⊕⊕⊕⊖ Moderat

⊕⊕⊖⊖ Niedrig

⊕⊖⊖⊖ Sehr niedrig

3.1.3 Primärdiagnostische Abklärung der Nieren mittels PET Scan vs. Computertomographie

Wir schlossen für diese Fragestellung über die primärdiagnostische Abklärung einer suspekten Raumforderung an der Niere mittels PET Scaneine diagnostische Studie ein. Die Studie¹⁰ untersuchte die Testgenauigkeit der [¹⁸F] FDG PET (Fluorin-18 Fluoro-2-Desoxyglucose Positronen Emissions Tomographie) im Vergleich zur Testgenauigkeit der Computertomographie (technische Details siehe Tabelle 10) bei 35 PatientInnen mit Verdacht auf Nierenzellkarzinom. Als Referenztest wurde die Histologie und das klinische bzw. radiologische „Follow-up“ herangezogen. 30 der renalen Läsionen waren Nierenzellkarzinome. Tabelle 10 fasst die Charakteristika und Ergebnisse der eingeschlossenen Studie zusammen. Das Biasrisiko dieser Studie wurde als „hoch“ eingestuft, da verschiedene Referenztests angewandt wurden und der Index test unter Kenntnis des CT Befundes durchgeführt wurde.

Die erhobenen Endpunkte dieser Studie beschränkten sich auf Testgüteparameter. Die Studiendauer lag bei 29 Monaten.

Zusammenfassend zeigt das Ergebnis eine Sensitivität von 47% und eine Spezifität von 80% für den PET Scan. Die Sensitivität der Computertomographie war 97% bei einer Spezifität von 0%, bedingt durch das Fehlen von „richtig negativen“ Testergebnissen (vermutlich bedingt durch eine hohe Nierenzellkarzinom-Prävalenz von 85% in der hier vorliegenden Studienpopulation). Die Ergebnisse sind aufgrund der geringen Fallzahl und der prävalenzbedingten extrem hohen „Vortestwahrscheinlichkeit“ sicherlich nur sehr bedingt generalisierbar.

Im Folgenden werden die einzelnen Endpunkte näher beschrieben. Eine Zusammenfassung der Stärke der Evidenz nach GRADE wird in Tabelle 11 und Tabelle 12 dargestellt.

Tabelle 10: Darstellung der wesentlichsten Charakteristika und Resultate der eingeschlossenen Studie.

Studie; Studiendesign; Sponsor	Population (N); Alter (Jahre ± SA); % weiblich	Interventionen (Test) vs. Kontrolle (Referenztest, Goldstandard), Technische Details	Relevante Endpunkte; Studiendauer	Resultate	Bias-Risiko
Aide et al., 2003 ¹⁰ ;diagnostische Studie; kA	PatientInnen mit Verdacht auf Nieren- zellkarzinom (35); durchschnittliches Alter: 60 Jahre; ±14; 31% weiblich	FDG PET unter Kenntnis des CT Befundes vs. CT Histologie (Referenzstandard) FDG PET: HR+ Simens Ganzkörper PET, Schichtdicke: 5,1mm (full- width at half maximum 4,6mm, axial in der Mitte des Sichtfeldes) CT: Picker PQ 5000 CT, Schichtdicke: 8mm, kV:120, mAs:225, Pitch:1,5. Schichten, Knochenfenster (1.5000-300 HE)	Testgüteparameter: Sensitivität; Spezifität; Accuracy; hinsichtlich Tumordiagnostik; 29 Monate	30/35 Läsionen maligne FDG PET (Befundung unter Kenntnis des CT Befundes auf PatientInnen Ebene): <ul style="list-style-type: none"> • Sensitivität: 47% (95% KI 28%; 66%) • Spezifität: 80% (95% KI 28%; 99%) • PPW: 93% * • NPW: 20% * • Accuracy: 51% (95% KI 34%; 69%) CT: <ul style="list-style-type: none"> • Sensitivität: 97% (95% KI 83%; 100%) • Spezifität: 0% (95% KI 0%; 52%) • PPW: 85% * • NPW: 0% * • Accuracy: 83% (95% KI 66%; 93%) *Eigenberechnung aus den vorliegenden Zahlen	hoch

Abkürzungen: CT, Computertomographie; kA, keine Angaben; FDG PET, Fluorin-18 Fluoro-2-Desoxyglucose Positronen Emissions Tomographie; HE, Hounsfield Einheiten; KI, Konfidenzintervall; NPW, Negativer prädiktiver Wert; PPW, Positiver prädiktiver Wert; SA, Standardabweichung;

3.1.3.1 Sensitivität

Der Anteil der richtig erkannten malignen Läsionen (d.h. die Sensitivität) war beim PET Scan 47% [95% KI 28%; 66%] (die Befundung des PET Scans erfolgte unter Kenntnis des CT Befundes) und bei der Computertomographie 97% [95% KI 83%; 100%].

3.1.3.2 Spezifität

Der Anteil der richtig erkannten benignen Läsionen (d.h. die Spezifität) war 80% [95% KI 28%; 99%] beim PET Scan und 0% [95% KI 0%; 52%] bei der Computertomographie (bedingt durch das Fehlen „richtig negativer“ Ergebnisse in dieser Stichprobe).

3.1.3.3 Diagnostische Testgüte (Accuracy)

Die diagnostische Testgüte war 51% [95% KI 34%, 69%] beim PET Scan und 83% [95% KI 66%, 93%] bei der Computertomographie.

3.1.3.4 Positiver prädiktiver Wert

Die Wahrscheinlichkeit eines tatsächlich malignen Befundes bei positivem Untersuchungsbefund (d.h. positiver prädiktiver Wert) war beim PET Scan 93% (14/15) und bei der Computertomographie 85% (29/43).

3.1.3.5 Negativer prädiktiver Wert

Die Wahrscheinlichkeit eines tatsächlich benignen Befundes bei negativem Untersuchungsbefund (d.h. negativer prädiktiver Wert) war beim PET Scan 20% (4/20; 16 der negativen Testergebnisse waren „falsch negativ“) und bei der Computertomographie 0% (0/1; es gab keine „richtig negativen“ Ergebnisse in dieser Stichprobe).

3.1.3.6 Strahlenbelastung

Zu diesem Endpunkt wurde keine Evidenz gefunden.

Tabelle 11: GRADE Evidenz Profil – PET-Scan

Population: PatientInnen mit Verdacht auf Nierenzellkarzinom

Test: PET-Scan

Referenzstandard: Histologie oder Follow-up

Outcomes	Anzahl der Studien (PatientInnen)	Studiendesign	Faktoren die möglicherweise die Qualität der Evidenz einschränken					Stärke der Evidenz (GRADE)	Effekt (Effekt pro 1000)	Wichtigkeit
			Biasrisiko	Indirektheit	Inkonsistenz	Impräzision	Publikationsbias			
Sensitivität	1 Studie (35 PatientInnen)	Nicht-konsequente Serie, unverblindet	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊖⊖⊖ Sehr niedrig	47% (470 korrekt positive Testresultate pro 1000 Läsionen)	Kritisch
Spezifität	1 Studie (35 PatientInnen)	Nicht-konsequente Serie, unverblindet	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊖⊖⊖ Sehr niedrig	80% (800 korrekt negative Testresultate pro 1000 Gesunden)	Kritisch
PPW	1 Studie (35 PatientInnen)	Nicht-konsequente Serie, unverblindet	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊖⊖⊖ Sehr niedrig	93% (930 korrekte Testresultate pro 1000 Läsionen mit einem positiven Test) ³	Kritisch
NPW	1 Studie (35 PatientInnen)	Nicht-konsequente Serie, unverblindet	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊖⊖⊖ Sehr niedrig	20% (200 korrekte Testresultate pro 1000 Läsionen mit einem negativen Test) ³	Kritisch
Accuracy	1 Studie (35 PatientInnen)	Nicht-konsequente Serie, unverblindet	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊖⊖⊖ Sehr niedrig	51% (510 korrekte Testresultate pro 1000 getesteten Personen)	Kritisch
Strahlenbelastung	Keine Evidenz								Wichtig	

Abkürzungen: n.a., nicht anwendbar; NPW, negativer prädiktiver Wert; PET, Positronen Emissions Tomographie; PPW, positiver prädiktiver Wert;

¹ Indextest und Referenztest wurden nicht unabhängig voneinander durchgeführt; Verwendung unterschiedlicher Indextests, Ergebnis des Indextests war bei Durchführung des Referenztests bekannt; um 2 Grade herabgestuft

² Insgesamt nur 35 Personen eingeschlossen

³ Eigenberechnung aus den vorliegenden Zahlen

Legende Stärke der Evidenz (GRADE):

⊕⊕⊕⊕ Hoch

⊕⊕⊕⊖ Moderat

⊕⊕⊖⊖ Niedrig

⊕⊖⊖⊖ Sehr niedrig

Tabelle 12: GRADE Evidenz Profil –Computertomographie

Population: PatientInnen mit Verdacht auf Nierenzellkarzinom

Test: Computertomographie

Referenzstandard: Histologie oder Follow -up

Outcomes	Anzahl der Studien (PatientInnen)	Studiendesign	Faktoren die möglicherweise die Qualität der Evidenz einschränken					Stärke der Evidenz (GRADE)	Effekt (Effekt pro 1000)	Wichtigkeit
			Biasrisiko	Indirektheit	Inkonsistenz	Impräzision	Publikationsbias			
Sensitivität	1 Studie (35 PatientInnen)	Nicht-konsequente Serie, unverblindet	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	97% (970 korrekt positive Testresultate pro 1000 Läsionen)	Kritisch
Spezifität	1 Studie (35 PatientInnen)	Nicht-konsequente Serie, unverblindet	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	Keine Daten	Kritisch
PPW	1 Studie (35 PatientInnen)	Nicht-konsequente Serie, unverblindet	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	85% (850 korrekte Testresultate pro 1000 Läsionen mit einem positiven Test) ³	Kritisch
NPW	1 Studie (35 PatientInnen)	Nicht-konsequente Serie, unverblindet	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	Keine Daten	Kritisch
Accuracy	1 Studie (35 PatientInnen)	Nicht-konsequente Serie, unverblindet	Hoch ¹	Keine	n.a.	Impräzise ²	Keiner detektiert	⊕⊕⊕⊕ Sehr niedrig	83% (830 korrekte Testresultate pro 1000 getesteten Personen)	Kritisch
Strahlenbelastung	Keine Evidenz									Wichtig

Abkürzungen: n.a., nicht anwendbar; NPW, negativer prädiktiver Wert; PET, Positronen Emissions Tomographie; PPW, positiver prädiktiver Wert;

¹ Indextest und Referenztest wurden nicht unabhängig voneinander durchgeführt; Verwendung unterschiedlicher Indextests, Ergebnis des Indextests war bei Durchführung des Referenztests bekannt; um 2 Grade herabgestuft

² Insgesamt nur 35 Personen eingeschlossen

³ Eigenberechnung aus den vorliegenden Zahlen

Legende Stärke der Evidenz (GRADE):

⊕⊕⊕⊕ Hoch

⊕⊕⊕⊖ Moderat

⊕⊕⊖⊖ Niedrig

⊕⊖⊖⊖ Sehr niedrig

3.1.4 Primärdiagnostische Abklärung der Nieren mittels Magnetresonanztomographie vs. Computertomographie

Für diese Fragestellung konnte keine Evidenz gefunden werden.

4 Diskussion

Die Grundlage jeder evidenzbasierten Empfehlung muss ein genaues Abwägen des vorhandenen Nutzen- und Schadenspotenzials der bewerteten diagnostischen Maßnahmen sein. Wissenschaftliche Evidenz ist dabei eine wesentliche Säule. Für die Entscheidungsfindung beim Erstellen einer Empfehlung müssen aber auch andere Aspekte wie PatientInnenwerte oder klinische Erfahrung/Relevanz mit in Betracht gezogen werden, vor allem wenn adäquate wissenschaftliche Evidenz nur eingeschränkt vorhanden ist.

Bei der vorliegenden Fragestellung zur primärdiagnostischen Abklärung einer suspekten Raumforderung an der Niere konnten nur sehr wenige prospektive diagnostische Studien gefunden werden. Zu einer der gesuchten diagnostischen Maßnahmen (Magnetresonanzangiographie) konnte überhaupt keine Studie eingeschlossen werden. Zudem waren die identifizierten Studien sehr heterogen mit teilweise unterschiedlichen Untersuchungseinheiten (Individualebene vs. Läsionsebene) für die Berechnung der Testgüteparameter. Die Studien wurden zum Teil an sehr kleinen Stichproben durchgeführt, unter fallweiser Heranziehung verschiedener Referenztests innerhalb einer Studie. In drei der vier Studien wurde das Biasrisiko mit „hoch“ bewertet und in einer Studie mit „unklar“. Zusammenfassend ist somit die Qualität der Evidenz als deutlich eingeschränkt zu betrachten.

Bei der Bewertung der besten verfügbaren Evidenz zeigte sich zusammenfassend, dass sowohl die Computertomographie als auch die Magnetresonanztomographie eine Sensitivität von 93,3% hatten, allerdings bei eingeschränkter Spezifität (CT 68%; MR 71%).⁹ Die kontrastmittelverstärkte Sonographie zeigte bei der Diagnostik von Malignität bei komplexen zystischen Läsionen eine ähnliche Sensitivität wie die Computertomographie. Die nicht verstärkte Sonographie schnitt sowohl bei der Sensitivität, wie auch bei der Spezifität deutlich schlechter ab.⁸

Abbildung 4 stellt die diagnostischen Wertigkeiten der eingeschlossenen diagnostischen Tests in Form einer ROC (Receiver Operating Curve) grafisch dar. Ein idealer Test würde bei dieser Darstellung in der linken oberen Ecke liegen. Die Vergleichbarkeit der einzelnen Tests ist aufgrund heterogener Populationen, hohem Biasrisiko und unterschiedlichen Basisrisikos nur sehr eingeschränkt möglich. Es zeigt sich jedoch, dass nicht-Kontrastmittelverstärkter

Ultraschall und PET-Scan deutlich schlechtere diagnostische Wertigkeit haben, als andere diagnostische Tests.

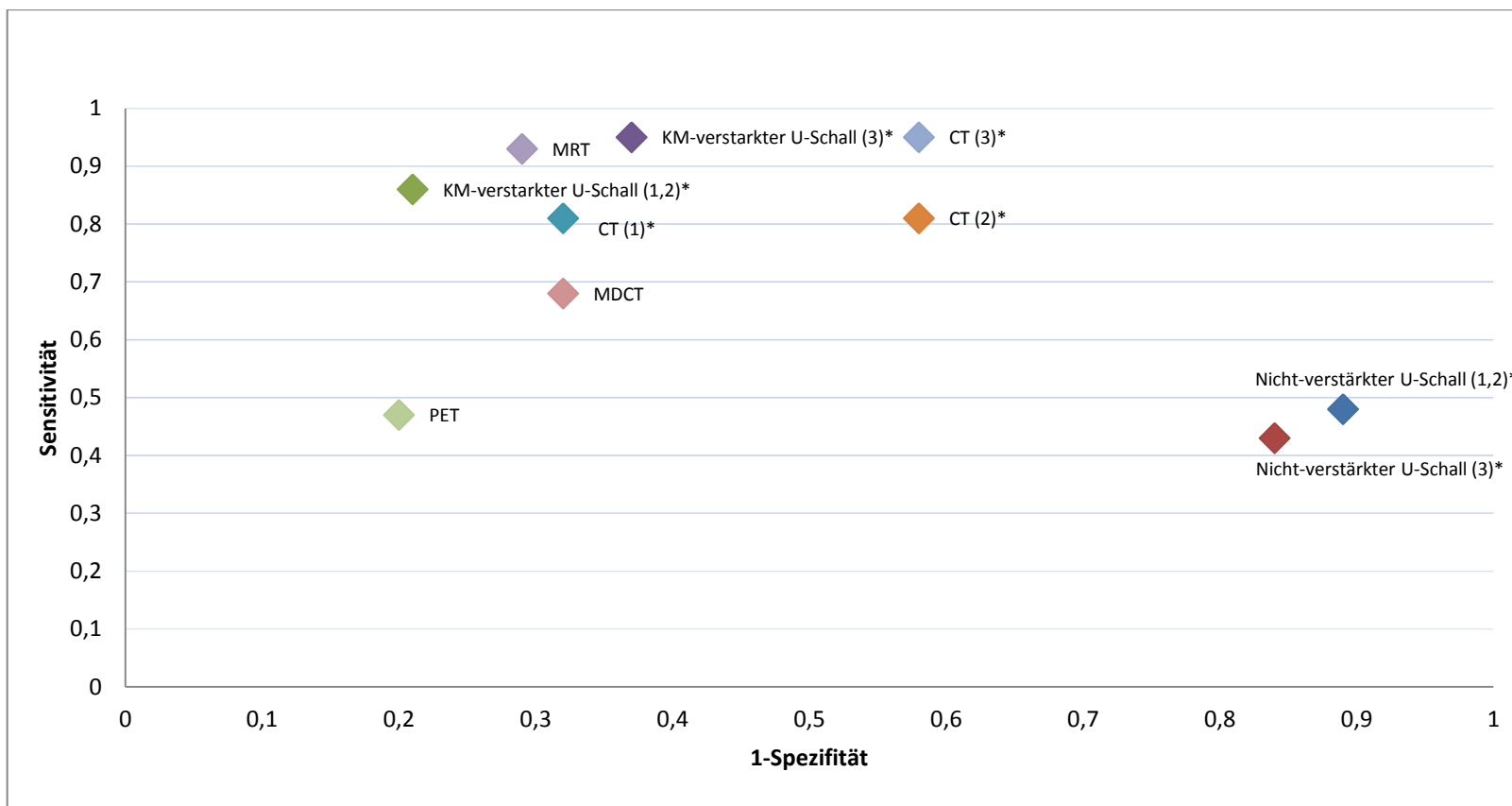


Abbildung 4: Diagnostische Wertigkeit der einzelnen diagnostischen Tests

Abkürzungen: CT, Computertomographie; KM, Kontrastmittel; MRT, Magnetresonanztomographie; MDCT, Multidetector Computertomographie; PET, Positronen Emissionstomographie; U-Schall, Ultraschall;

*Zahlen beziehen sich auf einzelne UntersucherInnen in der Studie

Im Rahmen der Literatursuche wurden eine Studie identifiziert, die wir an dieser Stelle in der Diskussion anführen möchten, obwohl die Arbeit aufgrund des Nicht-Entsprechens der Einschlusskriterien nicht in die Resultate dieses Berichtes inkludiert wurden.

Es handelt sich hierbei um einen systematischen Review mit Metaanalyse¹² hinsichtlich der diagnostischen Leistungsfähigkeit von [18F] FDG PET beim Nierenzellkarzinom. Zwei der inkludierten Studien zur Primärdiagnostik an der Niere waren zeitlich außerhalb des Suchzeitraumes. Eine Studie war retrospektiv angelegt und wurde daher aufgrund dieses Nichtentsprechens der Einschlusskriterien bereits im Abstract Review ausgeschlossen. Da es sich bei der Metaanalyse um ein gepooltes Ergebnis unter Einschluss einer retrospektiven Studie handelt, wurde die Metaanalyse nicht in unsere Arbeit eingeschlossen. Sowohl die globale Sensitivität (67%; 95% KI 59%-76%) wie auch die globale Spezifität (93%; 95% KI 86%-97%) waren in dieser Metaanalyse etwas höher, verglichen mit einer prospektiven Studie, die in unserem Report inkludiert wurde.¹⁰ Zusammenfassend kamen die AutorInnen der Metanalyse zu dem Schluss, dass die Leistung des PET Scans hinsichtlich Primärdiagnostik einer suspekten Raumforderung an der Niere limitiert ist, was auch unserem Ergebnis entspricht.

Die dargelegte Aufarbeitung der Evidenz dient als Grundlage für die Erstellung von Leitlinien-Empfehlungen. Bei der Erstellung von Empfehlungen unterscheidet GRADE zwischen „starken“ und „schwachen“ Empfehlungen (siehe Tabelle 13). In die Entscheidung über die Empfehlung fließen jedoch neben der Qualität der Evidenz noch weitere Faktoren, wie Gleichgewicht zwischen erwünschten und unerwünschten Effekten, PatientInnenwerte und Präferenzen sowie Kosten, ein. Die vorliegende Evidenz ist aufgrund bereits erwähnter Limitationen, wie geringe Fallzahlen und hohes Biasrisiko, für eine „starke Empfehlung“ wahrscheinlich nicht geeignet. Sowohl die Computertomographie als auch die Magnetresonanztomographie weisen in den vorliegenden Studien die höchsten Sensitivitätswerte bei suspekten renalen Läsionen auf. Keine der inkludierten Studien schloss patientenrelevante Endpunkte, wie etwa Strahlenbelastung, ein. Eine abschließende Bewertung sollte in der ExpertInnenrunde erfolgen.

Tabelle 13: Bestimmungsfaktoren der Stärke der Empfehlung

Faktor	Beschreibung
Gleichgewicht zwischen erwünschten und unerwünschten Effekten	Je größer der Unterschied zwischen erwünschten und unerwünschten Effekten, desto größer die Wahrscheinlichkeit einer stärkeren Empfehlung
Qualität der Evidenz	Je höher die Qualität der Evidenz, desto höher die Wahrscheinlichkeit einer stärkeren Empfehlung
Patientenwerte und Präferenzen	Je mehr Patientenwerte und Präferenzen variieren, desto größer ist die Unsicherheit bezüglich Patientenwerte, und desto wahrscheinlicher eine schwache Empfehlung
Kosten	Je höher die Kosten einer Empfehlung, desto größer die benötigten Ressourcen und desto weniger wahrscheinlich ist eine stärkere Empfehlung

5 Referenzen

1. Chow WH, Dong LM, Devesa SS. Epidemiology and risk factors for kidney cancer. *Nat Rev Urol.* May 2010;7(5):245-257.
2. Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID). GEKID-Atlas. 2012; <http://www.ekr.med.uni-erlangen.de/GEKID/>. Accessed 25.4.2013.
3. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Annals of Internal Medicine.* 2011;155(8):529-536.
4. Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration;2011.
5. Schünemann HJ, Mustafa R, Brozek J. Diagnostik und linked evidence – Wie robust muss die Kette sein? *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen.* 2012;106(3):153-160.
6. Schünemann H, Oxman AD, Brozek J, et al. GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ.* 2008;336(7654).
7. Meerpohl J, Langer G, Perleth M, Gartlehner G, Kaminski-Hartenthaler A, Schünemann H. GRADE-Leitlinien: 3. Bewertung der Qualität der Evidenz (Vertrauen in die Effektschätzer). *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen.* 2012;106(6):449-456.
8. Quaia E, Bertolotto M, Cioffi V, et al. Comparison of contrast-enhanced sonography with unenhanced sonography and contrast-enhanced CT in the diagnosis of malignancy in complex cystic renal masses. *AJR Am J Roentgenol.* Oct 2008;191(4):1239-1249.
9. Beer AJ, Dobritz M, Zantl N, Weirich G, Stofffuss J, Rummeny EJ. Comparison of 16-MDCT and MRI for characterization of kidney lesions. *AJR Am J Roentgenol.* Jun 2006;186(6):1639-1650.
10. Aide N, Cappele O, Bottet P, et al. Efficiency of [(18)F]FDG PET in characterising renal cancer and detecting distant metastases: a comparison with CT. *Eur J Nucl Med Mol Imaging.* Sep 2003;30(9):1236-1245.
11. Schmidt T, Hohl C, Haage P, et al. Diagnostic accuracy of phase-inversion tissue harmonic imaging versus fundamental B-mode sonography in the evaluation of focal lesions of the kidney. *AJR Am J Roentgenol.* Jun 2003;180(6):1639-1647.
12. Martinez de Llano SR, Delgado-Bolton RC, Jimenez-Vicioso A, et al. [Meta-analysis of the diagnostic performance of 18F-FDG PET in renal cell carcinoma]. *Rev Esp Med Nucl.* Jan-Feb 2007;26(1):19-29.

6 Appendizes

6.1 Appendix 1: Suchstrategie 29. Jänner 2013

MEDLINE (via PubMed)

ID	Query
#1	Search "Carcinoma, Renal Cell"[Mesh]
#2	Search Renal Cell Carcinoma*
#3	Search "Kidney Neoplasms"[Mesh]
#4	Search Kidney Cancer*
#5	Search Kidney Neoplasm*
#6	Search Kidney Tumor*
#7	Search Kidney Tumour*
#8	Search Renal Cancer*
#9	Search Renal Neoplasm*
#10	Search Renal Tumor*
#11	Search Renal Tumour*
#12	Search Clear-cell Carcinoma*
#13	Search Hypernephroma*
#14	Search Papillary Renal Cell Carcinoma*
#15	Search Chromophobe Renal Cell Carcinoma*
#16	Search Non Clear Cell Carcinoma*
#17	Search (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16)
#18	Search ("ultrasonography"[Subheading] OR "ultrasonography"[All Fields] OR "ultrasound"[All Fields] OR "ultrasonography"[MeSH Terms] OR "ultrasonics"[MeSH Terms] OR "ultrasonics"[All Fields])
#19	Search ("Magnetic Resonance Imaging"[Mesh] OR "MR Tomography"[tiab] OR "Magnetic Resonance Tomography"[tiab])
#20	Search ("Magnetic Resonance Angiography"[Mesh] OR "magnetic resonance angiography"[tiab])
#21	Search ("positron-emission tomography"[MeSH Terms] OR "positron-emission tomography"[tiab] OR "pet scan"[tiab] OR "pet scans"[tiab])
#22	Search (#18 OR #19 OR #20 OR #21)
#23	Search (#17 AND #22)
#24	Search "sunitinib"[Supplementary Concept] OR "sunitinib"[tiab] OR "sorafenib"[Supplementary Concept] OR "sorafenib"[tiab] OR "pazopanib"[Supplementary Concept] OR "pazopanib"[tiab] OR "axitinib"[Supplementary Concept] OR "axitinib"[tiab] OR "temsirolimus"[Supplementary Concept] OR "temsirolimus"[tiab] OR "everolimus"[Supplementary Concept] OR "everolimus"[tiab] OR "bevacizumab"[Supplementary Concept] OR "bevacizumab"[tiab] OR "tivozanib"[Supplementary Concept] OR "tivozanib"[tiab]
#25	Search (#17 AND #24)
#26	Search ("tomography, x-ray computed"[MeSH Terms] OR "x-ray computed tomography"[All Fields] OR "tomography"[All Fields] OR "tomography"[MeSH Terms] OR "computer tomography"[tiab])
#27	Search (#23 AND #26)
#28	Search (#25 OR #27)
#29	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh]))
#30	Search "randomized controlled trial"[Publication Type] OR "randomized controlled trials as topic"[MeSH Terms] OR "randomized controlled trial"[All Fields] OR "randomised controlled trial"[All Fields]
#31	Search (#29 AND #30)
#32	Search "Systematic Review" OR ("Review"[Publication Type] AND "systematic"[tiab])
#33	Search (#29 AND #32)
#34	Search "meta-analysis"[Publication Type] OR "meta-analysis as topic"[MeSH Terms] OR "meta-analysis"[All Fields]
#35	Search (#29 AND #34)
#36	Search "controlled clinical trial"[Publication Type] OR "Comparative Study"[Publication Type] OR "Random Allocation"[Mesh] OR "Single-Blind Method"[Mesh] OR "Double-Blind Method"[Mesh]

Nierenzellkarzinom Primäragnostik

ID	Query
#37	Search (#29 AND #36)
#38	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])) Filters: Clinical Trial
#39	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])) Filters: Clinical Trial; Clinical Trial, Phase I
#40	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])) Filters: Clinical Trial; Clinical Trial, Phase I; Clinical Trial, Phase II
#41	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])) Filters: Clinical Trial; Clinical Trial, Phase I; Clinical Trial, Phase II; Clinical Trial, Phase III
#42	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])) Filters: Clinical Trial; Clinical Trial, Phase I; Clinical Trial, Phase II; Clinical Trial, Phase III; Clinical Trial, Phase IV
#43	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])) Filters: Clinical Trial; Clinical Trial, Phase I; Clinical Trial, Phase II; Clinical Trial, Phase III; Clinical Trial, Phase IV; Comparative Study
#44	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])) Filters: Clinical Trial; Clinical Trial, Phase I; Clinical Trial, Phase II; Clinical Trial, Phase III; Clinical Trial, Phase IV; Comparative Study; Controlled Clinical Trial
#45	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])) Filters: Clinical Trial; Clinical Trial, Phase I; Clinical Trial, Phase II; Clinical Trial, Phase III; Clinical Trial, Phase IV; Comparative Study; Controlled Clinical Trial; Meta-Analysis
#46	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])) Filters: Clinical Trial; Clinical Trial, Phase I; Clinical Trial, Phase II; Clinical Trial, Phase III; Clinical Trial, Phase IV; Comparative Study; Controlled Clinical Trial; Meta-Analysis; Randomized Controlled Trial
#47	Search (#28 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])) Filters: Clinical Trial; Clinical Trial, Phase I; Clinical Trial, Phase II; Clinical Trial, Phase III; Clinical Trial, Phase IV; Comparative Study; Controlled Clinical Trial; Meta-Analysis; Randomized Controlled Trial; Systematic Reviews
#48	Search (#47 OR #37 OR #35 OR #33 OR #31) Sort by: PublicationDate

The Cochrane Library

ID	Query
#1	[mh "Carcinoma, Renal Cell"] or Renal Cell Carcinoma* or [mh "Kidney Neoplasms"] or Kidney Cancer* or Kidney Neoplasm* or Kidney Tumor* or Kidney Tumour* or Renal Cancer* or Renal Neoplasm* or Renal Tumor* or Renal Tumour* or Clear-cell Carcinoma* or Hypernephroma* or Papillary Renal Cell Carcinoma* or Chromophobe Renal Cell Carcinoma* or Non Clear Cell Carcinoma*
#2	ultrasonography or "ultrasound" or [mh ultrasonography] or "ultrasound" or [mh ultrasonics] or "ultrasonics" or [mh "Magnetic Resonance Imaging"] or "MR Tomography" or "Magnetic Resonance Tomography" or [mh "Magnetic Resonance Angiography"] or "magnetic resonance angiography" or [mh "positron-emission tomography"] or "positron-emission tomography" or "pet scan" or "pet scans"
#3	[mh "tomography, x-ray computed"] or "x-ray computed tomography" or "tomography" or [mh tomography] or "computer tomography"
#4	{and #1-#3}
#5	sunitinib or sorafenib or pazopanib or axitinib or temsirolimus or everolimus or bevacizumab or tivozanib
#6	#1 and #5
#7	#4 or #6

CINAHL

ID	Query
S30	S23 AND S29
S29	S24 OR S25 OR S26 OR S27 OR S28
S28	"randomized controlled trial" OR "randomised controlled trial" OR "Systematic Review" OR "meta-analysis" OR "controlled clinical trial" OR "Comparative Study" OR "Random Allocation" OR "Single-Blind Method" OR "Double-Blind Method"
S27	(MH "Comparative Studies")
S26	(MH "Meta Analysis")
S25	(MH "Systematic Review")

ID	Query
S24	(MH "Randomized Controlled Trials") OR "randomized controlled trial"
S23	S18 OR S22
S22	S17 NOT S21
S21	S19 NOT S20
S20	(MH "Human")
S19	(MH "Animals+")
S18	S17
S17	S14 OR S16
S16	S3 AND S15
S15	sunitinib OR sorafenib OR pazopanib OR axitinib OR temsirolimus OR everolimus OR bevacizumab OR tivozanib
S14	S10 AND S13
S13	S11 OR S12
S12	(MH "Tomography+")
S11	"computer tomography"
S10	S3 AND S9
S9	S4 OR S5 OR S6 OR S7 OR S8
S8	"ultrasonography" OR "ultrasound" OR "ultrasonics" OR "Magnetic Resonance Imaging" OR "MR Tomography" OR "Magnetic Resonance Tomography" OR "Magnetic Resonance Angiography" OR "positron-emission tomography" OR "pet scan" OR "pet scans"
S7	(MH "Tomography, Emission-Computed+")
S6	(MH "Magnetic Resonance Angiography")
S5	(MH "Magnetic Resonance Imaging+")
S4	(MH "Ultrasonography+")
S3	S1 OR S2
S2	Renal Cell Carcinoma* OR Kidney Cancer* OR Kidney Neoplasm* OR Kidney Tumor* OR Kidney Tumour* OR Renal Cancer* OR Renal Neoplasm* OR Renal Tumor* OR Renal Tumour* OR Clear-cell Carcinoma* OR Hypernephroma* OR Papillary Renal Cell Carcinoma* OR Chromophobe Renal Cell Carcinoma* OR Non Clear Cell Carcinoma*
S1	(MH "Carcinoma, Renal Cell+")

EMBASE

ID	Query
#1	'kidney carcinoma'/exp
#2	renal AND cell AND carcinoma* OR kidney AND cancer* OR kidney AND neoplasm* OR kidney AND tumor* OR kidney AND tumour* OR renal AND cancer* OR renal AND neoplasm* OR renal AND tumor* OR renal AND tumour* OR 'clear cell' AND carcinoma* OR hypernephroma* OR papillary AND renal AND cell AND carcinoma* OR chromophobe AND renal AND cell AND carcinoma* OR non AND clear AND cell AND carcinoma*
#3	#1 OR #2
#4	'ultrasound'/exp OR 'nuclear magnetic resonance imaging'/exp OR 'magnetic resonance angiography'/exp OR 'positron emission tomography'/exp OR 'echography'/exp
#5	'ultrasonography' OR 'ultrasound' OR 'ultrasonics' OR 'magnetic resonance imaging' OR 'mr tomography' OR 'magnetic resonance tomography' OR 'magnetic resonance angiography' OR 'positron-emission tomography' OR 'pet scan' OR 'pet scans'
#6	#4 OR #5
#7	#3 AND #6
#8	'computer assisted tomography'/exp OR 'computer tomography'
#9	#7 AND #8
#10	'sunitinib'/exp OR sunitinib OR 'sorafenib'/exp OR sorafenib OR 'pazopanib'/exp OR pazopanib OR 'axitinib'/exp OR axitinib OR 'temsirolimus'/exp OR temsirolimus OR 'everolimus'/exp OR everolimus OR 'bevacizumab'/exp OR bevacizumab OR 'tivozanib'/exp OR tivozanib
#11	#3 AND #10
#12	#9 OR #11
#13	#12 AND 'human'/de

ID	Query
#14	#12 AND 'human'/de AND ('article'/it OR 'review'/it)
#15	'randomized controlled trial'/exp OR 'randomized controlled trial' OR 'randomised controlled trial'/exp OR 'randomised controlled trial' OR 'systematic review'/exp OR 'systematic review' OR 'meta-analysis'/exp OR 'meta-analysis' OR 'controlled clinical trial'/exp OR 'controlled clinical trial' OR 'comparative study'/exp OR 'comparative study' OR 'random allocation'/exp OR 'random allocation' OR 'single-blind method'/exp OR 'single-blind method' OR 'double-blind method'/exp OR 'double-blind method'
#16	#14 AND #15

Web of Knowledge

ID	Query
# 9	#8 AND #4 Timespan=All Years Search language=English
# 8	#7 OR #6 Timespan=All Years Search language=English
# 7	#5 AND #3 Timespan=All Years Search language=English
# 6	#5 AND #2 AND #1 Timespan=All Years Search language=English
# 5	Topic=("Renal Cell Carcinoma" OR "Renal Cell Carcinomas" OR "Kidney Cancer" OR "Kidney Neoplasm" OR "Kidney Tumor" OR "Kidney Tumour" OR "Renal Cancer" OR "Renal Neoplasm" OR "Renal Tumor" OR "Renal Tumour" OR "Clear-cell Carcinoma" OR "Hypernephroma" OR "Papillary Renal Cell Carcinoma" OR "Chromophobe Renal Cell Carcinoma" OR "Non Clear Cell Carcinoma") Timespan=All Years Search language=English
# 4	Topic=("randomized controlled trial" OR "randomised controlled trial" OR "Systematic Review" OR "meta-analysis" OR "controlled clinical trial" OR "Comparative Study" OR "Random Allocation" OR "Single-Blind Method" OR "Double-Blind Method") Timespan=All Years Search language=English
# 3	Topic=(sunitinib OR sorafenib OR pazopanib OR axitinib OR temsirolimus OR everolimus OR bevacizumab OR tivozanib) Timespan=All Years Search language=English
# 2	Topic=("computer assisted tomography" OR "computer tomography" OR "x-ray computed tomography") Timespan=All Years Search language=English
# 1	Topic=("ultrasonography" OR "ultrasound" OR "ultrasonics" OR "Magnetic Resonance Imaging" OR "MR Tomography" OR "Magnetic Resonance Tomography" OR "Magnetic Resonance Angiography" OR "positron-emission tomography" OR "pet scan" OR "pet scans") Timespan=All Years Search language=English

6.2 Appendix 2: Kontaktierte AutorInnen

Autor	Frage	Antwort erhalten
Beer Ambros (Studie #549)	<p>Ich wäre Ihnen sehr dankbar, wenn Sie mir bei folgender Frage weiterhelfen könnten. Die Gesamtzahl der Läsionen war 56. Sie geben beim CT an, dass 15 von 16 RCC richtig diagnostiziert wurden (Sens 93,3%) und 26 von 38 benignen Läsionen (Spez. 68,4%). Beim MR sind es 15 von 16 RCC und 25 von 35 benignen Läsionen. Wieso ist die Gesamtzahl der benignen Läsionen beim MR geringer? Die Summe der benignen Läsionen und der RCCs ist in beiden Fällen nicht 56. Wie wurden die anderen Läsionen gewertet.</p> <p>Beim CT ist die PPW und NPW Berechnung mit richtig, falsch positiv und negativ klar. Bitte könnten Sie mir auch erklären wie die Berechnung des PPW und des NPW beim MRI erfolgt ist. Es sind die gleichen Werte wie beim CT, obwohl die Spezifität unterschiedlich ist.</p>	nein

6.3 Appendix 3: Ausgeschlossene Studien

6.3.1 Ausschluss wegen nicht zutreffenden Publikationstyps (6):

1. Tay KJ, Ho HS, Low AS, Cheng CW. Is contrast enhanced ultrasound a valid alternative diagnostic modality for renal cell carcinoma in patients with renal impairment? *Ann Acad Med Singapore*. 2012 Mar;41(3):127-8. PubMed PMID: 22538740. Pubmed Central PMCID: E1. Epub 2012/04/28. eng. <http://onlinelibrary.wiley.com/o/cochrane/clhta/articles/HTA-32011001077/frame.html>.
2. Gardner TA, Tirkes T, Mellon M, Koch MO. Imaging Techniques for the Patient With Renal Cell Carcinoma. *Seminars in Nephrology*. 2011;31(3):245-53. Pubmed Central PMCID: E1.
3. Sacco E, Pinto F, Totaro A, D'Addessi A, Racioppi M, Gulino G, et al. Imaging of renal cell carcinoma: State of the art and recent advances. *Urologia Internationalis*. 2011;86(2):125-39. Pubmed Central PMCID: E1.
4. Kang SK, Kim D, Chandarana H. Contemporary imaging of the renal mass. *Current Urology Reports*. 2011;12(1):11-7. Pubmed Central PMCID: E1.
5. Cohen D, Brown JJ. MR imaging of indeterminate renal lesions. *Applied Radiology*. 2008;37(11):19-32. Pubmed Central PMCID: E1.
6. Torabi M, Harisinghani MG, Gervais DA, Hahn PF, Arellano RS, Saksena MA, et al. Intermediate cystic renal masses: Radiology-pathology correlation. *Current Problems in Diagnostic Radiology*. 2005;34(3):116-25. Pubmed Central PMCID: E1.

6.3.2 Ausschluss wegen nicht zutreffender Population (6):

1. Hoeffel C, Pousset M, Timsit MO, Elie C, Mejean A, Merran S, et al. Radiofrequency ablation of renal tumours: diagnostic accuracy of contrast-enhanced ultrasound for early detection of residual tumour. *Eur Radiol*. 2010 Aug;20(8):1812-21. PubMed PMID: 20204642. Pubmed Central PMCID: E2. Epub 2010/03/06. eng.
2. Platzek I, Zastrow S, Deppe PE, Grimm MO, Wirth M, Laniado M, et al. Whole-body MRI in follow-up of patients with renal cell carcinoma. *Acta Radiol*. 2010 Jun;51(5):581-9. PubMed PMID: 20429759. Pubmed Central PMCID: E2. Epub 2010/05/01. eng.
3. Walter C, Kruessell M, Gindele A, Brochhagen HG, Gossmann A, Landwehr P. Imaging of renal lesions: evaluation of fast MRI and helical CT. *Br J Radiol*. 2003 Oct;76(910):696-703. PubMed PMID: 14512329. Pubmed Central PMCID: E2. Epub 2003/09/27. eng.
4. Desar IME, Stillebroer AB, Oosterwijk E, Leenders WPJ, Van Herpen CML, Van Der Graaf WTA, et al. 111In-bevacizumab imaging of renal cell cancer and evaluation of neoadjuvant treatment with the vascular endothelial growth factor receptor inhibitor sorafenib. *Journal of Nuclear Medicine*. 2010;51(11):1707-15. Pubmed Central PMCID: E2.
5. Hallscheidt PJ, Bock M, Riedasch G, Zuna I, Schoenberg SO, Autschbach F, et al. Diagnostic Accuracy of Staging Renal Cell Carcinomas Using Multidetector-Row Computed Tomography and Magnetic Resonance Imaging: A Prospective Study with Histopathologic Correlation. *Journal of Computer Assisted Tomography*. 2004;28(3):333-9. Pubmed Central PMCID: E2
6. Kitamura H, Fujimoto H, Tobisu KI, Mizuguchi Y, Maeda T, Matsuoka N, et al. Dynamic computed tomography and color Doppler ultrasound of renal parenchymal neoplasms: Correlations with histopathological findings. *Japanese journal of clinical oncology*. 2004;34(2):78-81. Pubmed Central PMCID: E2.

6.3.3 Ausschluss wegen nicht zutreffender Intervention oder Kontrollintervention (5):

1. Kim JH, Bae JH, Lee KW, Kim ME, Park SJ, Park JY. Predicting the histology of small renal masses using preoperative dynamic contrast-enhanced magnetic resonance imaging. *Urology*. 2012 Oct;80(4):872-6. PubMed PMID: 22854134. Pubmed Central PMCID: E3. Epub 2012/08/03. eng.
2. Krajewski KM, Guo M, Van den Abbeele AD, Yap J, Ramaiya N, Jagannathan J, et al. Comparison of four early posttherapy imaging changes (EPTIC; RECIST 1.0, tumor shrinkage, computed tomography tumor density, Choi criteria) in assessing outcome to vascular endothelial growth factor-targeted therapy in patients with advanced renal cell carcinoma. *Eur Urol*. 2011 May;59(5):856-62. PubMed PMID: 21306819. Pubmed Central PMCID: E3. Epub 2011/02/11. eng.
3. Divgi CR, Pandit-Taskar N, Jungbluth AA, Reuter VE, Gonen M, Ruan S, et al. Preoperative characterisation of clear-cell renal carcinoma using iodine-124-labelled antibody chimeric G250 (124I-cG250) and PET in patients with renal masses: a phase I trial. *Lancet Oncol*. 2007 Apr;8(4):304-10. PubMed PMID: 17395103. Pubmed Central PMCID: E3. Epub 2007/03/31. eng.
4. Lara PN, Jr., Quinn DI, Margolin K, Meyers FJ, Longmate J, Frankel P, et al. SU5416 plus interferon alpha in advanced renal cell carcinoma: a phase II California Cancer Consortium Study with biological and imaging correlates of angiogenesis inhibition. *Clin Cancer Res*. 2003 Oct 15;9(13):4772-81. PubMed PMID: 14581348. Pubmed Central PMCID: E3. Epub 2003/10/29. eng.
5. Graser A, Becker CR, Staehler M, Clevert DA, Macari M, Arndt N, et al. Single-phase dual-energy CT allows for characterization of renal masses as benign or malignant. *Invest Radiol*. 2010 Jul;45(7):399-405. PubMed PMID: 20498609. Pubmed Central PMCID: E3. Epub 2010/05/26. eng.

6.3.4 Ausschluss wegen nicht zutreffender Endpunkte (4):

1. Clevert DA, Minaifar N, Weckbach S, Jung EM, Stock K, Reiser M, et al. Multislice computed tomography versus contrast-enhanced ultrasound in evaluation of complex cystic renal masses using the Bosniak classification system. *Clin Hemorheol Microcirc*. 2008;39(1-4):171-8. PubMed PMID: 18503122. Pubmed Central PMCID: E4. Epub 2008/05/27. eng.
2. Ascenti G, Mazziotti S, Zimbaro G, Settineri N, Magno C, Melloni D, et al. Complex cystic renal masses: characterization with contrast-enhanced US. *Radiology*. 2007 Apr;243(1):158-65. PubMed PMID: 17392251. Pubmed Central PMCID: E4. Epub 2007/03/30. eng.
3. Tamai H, Takiguchi Y, Oka M, Shingaki N, Enomoto S, Shiraki T, et al. Contrast-enhanced ultrasonography in the diagnosis of solid renal tumors. *J Ultrasound Med*. 2005 Dec;24(12):1635-40. PubMed PMID: 16301719. Pubmed Central PMCID: E4. Epub 2005/11/23. eng.
4. Nathan PD, Vinayan A, Stott D, Juttla J, Goh V. CT response assessment combining reduction in both size and arterial phase density correlates with time to progression in metastatic renal cancer patients treated with targeted therapies. *Cancer Biology and Therapy*. 2010;9(1). Pubmed Central PMCID: E4.

6.3.5 Ausschluss wegen nicht zutreffenden Studiendesigns (7):

1. Mucksavage P, Ramchandani P, Malkowicz SB, Guzzo TJ. Is ultrasound imaging inferior to computed tomography or magnetic resonance imaging in evaluating renal mass size? *Urology*. 2012 Jan;79(1):28-31. PubMed PMID: 22202542. Pubmed Central PMCID: E5. Epub 2011/12/29. eng.

2. Mucksavage P, Kutikov A, Magerfleisch L, Van Arsdalen K, Wein AJ, Ramchandani P, et al. Comparison of radiographical imaging modalities for measuring the diameter of renal masses: is there a sizeable difference? *BJU Int.* 2011 Oct;108(8 Pt 2):E232-6. PubMed PMID: 21348913. Pubmed Central PMCID: E5. Epub 2011/02/26. eng.
3. Martinez de Llano SR, Delgado-Bolton RC, Jimenez-Vicioso A, Perez-Castejon MJ, Carreras Delgado JL, Ramos E, et al. [Meta-analysis of the diagnostic performance of 18F-FDG PET in renal cell carcinoma]. *Rev Esp Med Nucl.* 2007 Jan-Feb;26(1):19-29. PubMed PMID: 17286945. Pubmed Central PMCID: E5. Epub 2007/02/09. Meta-analisis sobre el rendimiento diagnostico de la tomografia por emision de positrones con 18F-FDG en el carcinoma de celulas renales. spa.
4. Ak I, Can C. F-18 FDG PET in detecting renal cell carcinoma. *Acta Radiol.* 2005 Dec;46(8):895-9. PubMed PMID: 16392617. Pubmed Central PMCID: E5. Epub 2006/01/06. eng.
5. Herts BR. Imaging for renal tumors. *Curr Opin Urol.* 2003 May;13(3):181-6. PubMed PMID: 12692438. Pubmed Central PMCID: E5. Epub 2003/04/15. eng.
6. Rankin SC, Webb JA, Reznek RH. Spiral computed tomography in the diagnosis of renal masses. *BJU Int.* 2000 Jul;86 Suppl 1:48-57. PubMed PMID: 10961275. Pubmed Central PMCID: E5. Epub 2000/08/29. eng.
7. Israel GM, Hindman N, Bosniak MA. Evaluation of cystic renal masses: comparison of CT and MR imaging by using the Bosniak classification system. *Radiology.* 2004 May;231(2):365-71. PubMed PMID: 15128983. Pubmed Central PMCID: E5. Epub 2004/05/07. eng.

6.3.6 Ausschluss wegen nicht zutreffender Sprache (1):

1. Ribeiro SM, Ajzen SA, Trindade JC. [Comparison of imaging methods for diagnosis of renal tumors and their calcifications]. *Rev Assoc Med Bras.* 2004 Oct-Dec;50(4):403-12. PubMed PMID: 15666022. Pubmed Central PMCID: E6. Epub 2005/01/25. Comparacao dos metodos de imagem no diagnostico dos tumores renais e calcificacoes nestas neoplasias. por.

6.3.7 Ausschluss da nur Abstract verfügbar (4):

1. Wang HY, Ding HJ, Chen JH, Chao CH, Lu YY, Lin WY, et al. Meta-analysis of the diagnostic performance of [18F]FDG-PET and PET/CT in renal cell carcinoma. *Cancer Imaging.* 2012;12:464-74. PubMed PMID: 23108238. Pubmed Central PMCID: E8. Epub 2012/10/31. eng.
2. Lei L, Fu NH, Yang B, Wei SP. Comparative analysis of contrast-enhanced ultrasound and contrast-enhanced CT in diagnosis of small renal cell carcinoma. *Chinese Journal of Medical Imaging Technology.* 2012;28(4):760-4. Pubmed Central PMCID: E8.
3. Yuan XC, Wang Y, Zhou AY, Chen L, Den LY, Liu J. Comparison on contrast-enhanced ultrasonography and CT in diagnosing renal cell carcinoma. *Chinese Journal of Medical Imaging Technology.* 2011;27(12):2502-5. Pubmed Central PMCID: E8.
4. Ko K, Kim B, Park BK. The use of contrastenhanced ultrasonography in the evaluation of cystic renal masses: comparison with CT. *Radiological Society of North America Scientific Assembly and Annual Meeting Program [Internet].* 2004 [cited 1 E8]; 152.