

Positronenemissionstomographie (PET) und PET/CT bei Ösophaguskarzinom

Vorbericht (vorläufige Nutzenbewertung)

Auftrag: D06-01H
Version: 1.0
Stand: 15.05.2012

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Positronenemissionstomographie (PET) und PET/CT bei Ösophaguskarzinom

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

21.12.2006

Interne Auftragsnummer:

D06-01H

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

Berichte@iqwig.de

www.iqwig.de

Dieser Bericht wurde unter Beteiligung externer Sachverständiger erstellt. Externe Sachverständige, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Nr. 2 Sozialgesetzbuch – Fünftes Buch – Gesetzliche Krankenversicherung „alle Beziehungen zu Interessenverbänden, Auftragsinstituten, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie, einschließlich Art und Höhe von Zuwendungen“ offenzulegen. Das Institut hat von jedem der Sachverständigen ein ausgefülltes Formular „Offenlegung potenzieller Interessenkonflikte“ erhalten. Die Angaben wurden durch das speziell für die Beurteilung der Interessenkonflikte eingerichtete Gremium des Instituts bewertet. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden.

Bei dem vorliegenden Vorbericht handelt es sich um eine vorläufige Nutzenbewertung. Zu diesem Vorbericht können Stellungnahmen abgegeben werden, die zu einer Ergänzung und / oder Überarbeitung des Berichts führen können. Die Frist für den Eingang der Stellungnahmen befindet sich auf der Website des IQWiG (www.iqwig.de), ebenso wie die dafür notwendigen Formblätter und ein Leitfaden.

Kurzfassung

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 21.12.2006 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissensstandes zur Positronen-emissionstomographie (PET), auch als Integrationsgerät von PET und Computertomographie (PET/CT), bei 14 verschiedenen Krankheitsbildern beauftragt. Der vorliegende Vorbericht bezieht sich auf den Teilauftrag zu Ösophaguskarzinomen.

Fragestellung

Der vorliegende Bericht verfolgte 2 Ziele:

1) Ermittlung des patientenrelevanten Nutzens der PET bzw. PET/CT

Das primäre Ziel des Berichts war die Beschreibung des patientenrelevanten Nutzens, den Ärzte und Patienten beim Primärstaging, beim Restaging und bei der Rezidivdiagnostik von Ösophaguskarzinomen von einer PET bzw. PET/CT erwarten können. Unter „Nutzen“ wurden die Veränderungen verstanden, die kausal auf den Einsatz der PET bzw. PET/CT zurückzuführen sind und für den Patienten fassbare Konsequenzen haben.

2) Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT

Sollten zu wenige aussagekräftige Studien zur Ermittlung des patientenrelevanten Nutzens identifiziert werden (Ziel 1), so sollte zusätzlich eine Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT vorgenommen werden (Ziel 2). Dabei sollte geprüft werden, inwieweit die PET bzw. PET/CT den diagnostischen Standardverfahren ohne PET überlegen ist. Das bedeutet: Verbessert sich durch den Einsatz der PET bzw. PET/CT das Primärstaging, das Restaging oder die Erkennung von Rezidiven? Ebenso sollte geprüft werden, ob sich mittels PET bzw. PET/CT zuverlässigere prognostische Aussagen im Rahmen der genannten Indikationen treffen lassen, als es mit den bisherigen diagnostischen Standardverfahren möglich ist.

Methoden

Für die Nutzenbewertung sollten im Rahmen einer systematischen Übersicht (randomisierte) kontrollierte Studien – z. B. Strategie mit vs. ohne PET bzw. PET/CT – mit patientenrelevanten Endpunkten (z. B. Mortalität / Morbidität) berücksichtigt werden.

Für die Beurteilung der diagnostischen und prognostischen Güte sollte ein „Review of Reviews“, also eine Bewertung auf der Basis publizierter Evidenzsynthesen, dienen. Für den Zeitraum und die Fragestellungen, die die aktuellste Evidenzsynthese mit ihrer Literatursuche nicht abdeckte, sollte die relevante Primärliteratur (prospektive Kohorten- und Querschnittsstudien) durch eigene Ergänzungsrecherchen zusätzlich identifiziert werden.

Im Rahmen der Nutzenbewertung wurde eine systematische Literaturrecherche nach RCTs und im Rahmen der Ergänzungsrecherche nach Studien zur diagnostischen und

prognostischen Güte in den folgenden Datenbanken durchgeführt: EMBASE, MEDLINE und Cochrane Central Register of Controlled Trials (Clinical Trials). Zur Ermittlung von Evidenzsynthesen wurde zusätzlich in folgenden Datenbanken recherchiert: Cochrane Database of Systematic Reviews (Cochrane Reviews), Database of Abstracts of Reviews of Effects (Other Reviews) und Health Technology Assessment Database (Technology Assessments). Die Literaturrecherche umfasste den Zeitraum bis zum 16.02.2010 (Primärstudien) bzw. 17.02.2011 (Evidenzsynthesen). Eine Aktualisierung der Recherche ist für den Zeitraum nach Publikation des Vorberichts geplant. Darüber hinaus wurden die vom G-BA übermittelten Unterlagen, öffentlich zugängliche Studienregister, die im Rahmen der Anhörung zum vorläufigen Berichtsplan übermittelten Unterlagen sowie die Datenbanken von Leitlinienerstellern durchsucht. Daneben wurde in 4 Kongressbänden recherchiert. Die Literaturverzeichnisse potenziell relevanter Evidenzsynthesen wurden ebenfalls durchsucht.

Das Literaturscreening wurde von 2 Reviewern unabhängig durchgeführt. Nach der Bewertung der Studienqualität wurden die Ergebnisse der einzelnen Studien nach Fragestellungen geordnet und beschrieben. Darüber hinaus wurden die Studien, die im Rahmen der Ergänzungsrecherche eingeschlossen wurden, im Hinblick auf ihre Übertragbarkeit auf den deutschen Versorgungskontext bewertet.

Ergebnisse

Patientenrelevanter Nutzen

Durch die systematische Recherche nach publizierter Literatur konnte für keine Fragestellung des vorliegenden Berichts eine vergleichende Studie identifiziert werden, die eine Aussage zum patientenrelevanten (Zusatz-)Nutzen der PET bzw. PET/CT für die Fragestellungen des Primärstaging, des Restaging nach Abschluss der Therapie und der Rezidivdiagnostik ermöglicht hätte. Auch die Suche in Kongressbänden und Studienregistern lieferte keine Hinweise auf noch laufende vergleichende Studien.

Diagnostische und prognostische Güte

3 Evidenzsynthesen und 30 Primärstudien aus der Ergänzungsrecherche entsprachen den Einschlusskriterien des vorliegenden Berichts. Aus den Evidenzsynthesen, die ausschließlich die Indikation Restaging untersuchten, entsprachen 11 Primärstudien den Einschlusskriterien des vorliegenden Berichts. Damit basiert der Bericht auf insgesamt 41 Primärstudien: Primärstaging (n = 23), Restaging (n = 14) und Rezidivdiagnostik (n = 4).

Primärstaging

Für das N-Staging wurden direkte Vergleiche zwischen PET und CT in 12 Studien untersucht. In den direkten bivariaten Meta-Analysen für das N-Staging konnte kein statistisch signifikanter Unterschied zwischen der diagnostischen Güte von PET und CT festgestellt werden.

Für das M-Staging wurden direkte Vergleiche zwischen PET und CT in 5 Studien untersucht. In den direkten bivariaten Meta-Analysen konnte kein statistisch signifikanter Unterschied zwischen der diagnostischen Güte von PET und CT festgestellt werden. Die Robustheit dieser gepoolten Ergebnisse wurde in mehreren Sensitivitätsanalysen bestätigt.

In einer prognostischen Studie wurde ein direkter Vergleich von FDG-PET und TI-SPECT durchgeführt. Keiner der PET- sowie SPECT relevanten Merkmale zeigte nach der multifaktoriellen Modellierung einen statistisch signifikanten Zusammenhang.

Restaging

Von den insgesamt 14 eingeschlossenen Studien untersuchten 2 Studien neben der diagnostischen auch die prognostische Güte der PET bzw. PET/CT. Eine Studie untersuchte ausschließlich die prognostische Güte der PET.

Sechs der aus den Evidenzsynthesen stammenden Studien untersuchten die PET bzw. PET/CT in der Indikation des Interimsstaging. Drei dieser Studien untersuchten den Einsatz der PET während einer neoadjuvanten Chemotherapie. Hier zeigte sich für die PET eine Spannweite der Punktschätzer der Sensitivitäten von 80 % (bei einer zugehörigen Spezifität von 78 % bzw. 100 %) bis 88 % (zugehörige Spezifität 75 %). Für die Punktschätzer der Spezifitäten ergab sich entsprechend eine Spannweite von 75 % (zugehörige Sensitivität 88 %) bis 100 % (zugehörige Sensitivität 80 %). Für alle Punktschätzer waren weite Konfidenzintervalle zu beobachten. In den weiteren 3 Studien wurde der Einsatz der PET bzw. PET/CT während einer neoadjuvanten Chemoradiotherapie untersucht. Allerdings konnten für 2 Studien die Ergebnisse nicht dargestellt werden, weil der Anteil fehlender Werte größer als 30 % war. In einer Studie ergab sich eine Sensitivität von 44 % und eine Spezifität von 52 %. Auch hier waren die Konfidenzintervalle weit.

Für das Restaging nach Abschluss einer neoadjuvanten Therapie wurde in 3 Studien ein direkter Vergleich diagnostischer Technologien untersucht. Der Vergleich PET bzw. PET/CT und Endosonografie (EUS) wurde in 2 Studien untersucht, wovon eine Studie jedoch nur 10 Patienten mit der PET untersuchen konnte. In der zweiten Studie zeigte sich für die PET/CT eine Sensitivität von 87 % (95 %-KI [60; 98]) und eine Spezifität von 88 % (95 %-KI [72; 97]). Im Gegensatz dazu hatte die EUS eine Sensitivität von 20 % (95 %-KI [4; 48]) und eine Spezifität von 94 % (95 %-KI [80; 99]).

Rezidivdiagnostik

Für die Rezidivdiagnostik konnte keine Evidenzsynthese identifiziert werden. Im Rahmen der Ergänzungsrecherche wurden 4 Primärstudien eingeschlossen. Alle 4 Studien untersuchten die diagnostische Güte der PET bzw. PET/CT. Eine Studie untersuchte zusätzlich noch die prognostische Güte.

Zum direkten Vergleich zwischen PET bzw. PET/CT und konventioneller Diagnostik lagen keine Ergebnisse vor.

Fazit

Der patientenrelevante Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen ist nicht belegt. Es konnten weder laufende noch abgeschlossene vergleichende Studien zum patientenrelevanten Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen identifiziert werden.

Zu der 2. Fragestellung des Berichts, der diagnostischen und prognostischen Güte, konnten aus 3 eingeschlossenen Evidenzsynthesen und der Ergänzungsrecherche insgesamt 41 Primärstudien eingeschlossen werden.

Die meisten Studien wurden in der Indikation Primärstaging durchgeführt. Zwölf Studien untersuchten die PET direkt mit der CT beim N-Staging und 5 Studien untersuchten den direkten Vergleich von PET mit CT beim M-Staging. Es konnte keine Aussage zugunsten eines Verfahrens getroffen werden, da sich die Konfidenzregionen in den bivariaten Meta-Analysen stark überlappten.

Zu den Indikationen Restaging und Rezidivdiagnostik liegen bisher zu wenige hochwertige Studien vor, um eine verlässliche Aussage zur diagnostischen und prognostischen Güte der PET oder PET/CT treffen zu können. Unklar bleibt insbesondere der Vorteil dieses metabolischen Verfahrens gegenüber den morphologischen bildgebenden Verfahren. Ob der Einsatz der PET bzw. PET/CT auch zu einer Verbesserung patientenrelevanter Endpunkte führt, ist bisher in keiner der 3 Indikationen untersucht worden.

Weitere methodisch hochwertige Studien sind (insbesondere bezüglich der Fragestellung des Behandlungsansprechens neoadjuvanter Therapien) dringend erforderlich, um den patientenrelevanten Nutzen oder Schaden der PET bzw. PET/CT von Ösophaguskarzinomen verlässlich bewerten zu können. Bisher konnten weder laufende noch abgeschlossene parallel vergleichende Studien zum patientenrelevanten Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen identifiziert werden.

Schlagwörter: Positronenemissionstomographie, Computertomographie, Staging, Rezidiv, Ösophagustumoren, systematische Übersicht

Keywords: Positron-Emission Tomography, Tomography, X-Ray Computed, Staging, Recurrence, Esophageal Neoplasms, Systematic Review

Inhaltsverzeichnis

	Seite
Kurzfassung	iv
Tabellenverzeichnis	xiv
Abbildungsverzeichnis	xviii
Abkürzungsverzeichnis.....	xix
1 Hintergrund	1
1.1 Definition des Krankheitsbildes.....	1
1.2 Epidemiologie und individuelle Krankheitslast	1
1.3 Ursache der Erkrankung.....	1
1.4 Verlauf der Erkrankung	2
1.5 Klassifikation.....	3
1.6 Diagnostische Verfahren	4
1.7 Derzeit übliche therapeutische Verfahren.....	5
1.8 PET und PET/CT.....	6
1.9 Empfehlungen wissenschaftlicher Fachgesellschaften zum Einsatz der PET bei Ösophaguskarzinomen	8
2 Ziele der Untersuchung.....	9
3 Projektbearbeitung.....	11
3.1 Zeitlicher Verlauf des Projekts.....	11
3.2 Dokumentation der Änderungen im Projektverlauf	12
4 Methoden.....	13
4.1 Ermittlung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT	15
4.1.1 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT anhand von Evidenzsynthesen	16
4.1.2 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT anhand von Primärstudien.....	16
4.1.2.1 Population	16
4.1.2.2 Prüf- und Vergleichsinterventionen.....	16
4.1.2.3 Zielgrößen.....	16
4.1.2.4 Studientypen	17
4.1.2.5 Sonstige Studiencharakteristika.....	18
4.1.3 Tabellarische Darstellung der Einschlusskriterien („Nutzenstudien“).....	18

4.2	Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT	18
4.2.1	Einschlusskriterien zur Ermittlung der diagnostischen und prognostischen Güte anhand von Evidenzsynthesen.....	19
4.2.1.1	Populationen	19
4.2.1.2	Prüftechnologien, Vergleichstechnologien („Komparatoren“) und Referenztest	19
4.2.1.3	Zielgrößen.....	19
4.2.1.4	Den Evidenzsynthesen zugrunde liegende Studientypen	20
4.2.1.5	Sonstige Charakteristika der Evidenzsynthesen	21
4.2.2	Einschlusskriterien zur Ermittlung der diagnostischen und prognostischen Güte anhand von Primärstudien („Ergänzungsrecherche“).....	22
4.2.2.1	Populationen	22
4.2.2.2	Prüftechnologien (Indextest I), Vergleichstechnologien (Indextest II) und Referenztest	22
4.2.2.3	Zielgrößen.....	22
4.2.2.4	Studientypen	22
4.2.2.5	Sonstige Studiencharakteristika.....	22
4.2.2.6	Tabellarische Darstellung der Einschlusskriterien („Ergänzungsrecherche“ nach Primärstudien zur diagnostischen und prognostischen Güte)	23
4.2.3	Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen.....	23
4.2.4	Suche nach zusätzlichen Informationen zu relevanten Studien	23
4.3	Informationsbeschaffung.....	23
4.3.1	Bibliografische Literaturrecherche.....	23
4.3.2	Suche nach weiteren publizierten und nicht publizierten Studien und Evidenzsynthesen	24
4.3.2.1	Suche in potenziell relevanten Evidenzsynthesen	24
4.3.2.2	Suche in öffentlich zugänglichen Studienregistern	24
4.3.2.3	Suche in durch den G-BA übermittelten Unterlagen.....	25
4.3.2.4	Suche in Kongressbänden.....	25
4.3.2.5	Suche in den Datenbanken der Leitlinienanbieter	25
4.3.3	Selektion relevanter Studien und Evidenzsynthesen.....	25
4.3.4	Suche nach zusätzlichen Informationen zu relevanten Studien	26
4.3.5	Nutzung von Informationen aus der Anhörung.....	26
4.4	Informationsbewertung	26
4.4.1	Bewertung von Studien zum patientenrelevanten Nutzen und Schaden	26
4.4.2	Bewertung von Evidenzsynthesen zur diagnostischen und prognostischen Güte	27

4.4.3	Bewertung von Primärstudien zur diagnostischen und prognostischen Güte	28
4.5	Informationssynthese und -analyse	28
4.5.1	Gegenüberstellung der Ergebnisse der Einzelstudien	28
4.5.1.1	Gegenüberstellung der Ergebnisse von (Primär-)Studien zur Bewertung des patientenrelevanten Nutzens und Schadens.....	29
4.5.1.2	Gegenüberstellung der Ergebnisse von Evidenzsynthesen zur diagnostischen und prognostischen Güte.....	30
4.5.1.3	Gegenüberstellung der Ergebnisse von Primärstudien zur diagnostischen und prognostischen Güte („Ergänzungsrecherche“).....	30
4.5.1.4	Gegenüberstellung der Ergebnisse aus Evidenzsynthesen und der Ergänzungsrecherche („Robustheitsprüfung“)	30
4.5.2	Meta-Analysen	31
4.5.2.1	Meta-Analysen für vergleichende Studien	31
4.5.2.2	Meta-Analysen für Diagnosestudien	31
4.5.3	Sensitivitätsanalysen	32
4.5.4	Subgruppenmerkmale und andere Effektmodifikatoren	32
4.6	Änderungen der Methodik im Vergleich zum Berichtsplan.....	33
5	Ergebnisse.....	35
5.1	Ergebnisse der Informationsbeschaffung.....	35
5.1.1	Ergebnis der bibliografischen Literaturrecherche	35
5.1.2	Weitere publizierte und nicht publizierte Studien.....	37
5.1.2.1	Evidenzsynthesen	37
5.1.2.2	Studienregister	37
5.1.2.3	Suche in den Datenbanken der Leitlinienanbieter	38
5.1.2.4	Unterlagen des G-BA	38
5.1.2.5	Kongressbände.....	38
5.1.2.6	Informationen aus der Anhörung.....	39
5.1.2.7	Zusätzliche Informationen zu relevanten Studien	39
5.1.3	Resultierender Studienpool	39
5.1.3.1	Studien zur Bewertung des patientenrelevanten Nutzens.....	39
5.1.3.2	Evidenzsynthesen zur diagnostischen und prognostischen Güte.....	39
5.1.3.3	Primärstudien zur diagnostischen und prognostischen Güte (Ergänzungsrecherche)	40
5.2	Teilziel 1: patientenrelevanter Nutzen	44
5.3	Teilziel 2a: diagnostische und prognostische Güte beim Primärstaging	45
5.3.1	Diagnostische und prognostische Güte auf Grundlage in die Bewertung eingeflossener Evidenzsynthesen beim Primärstaging.....	45
5.3.2	Diagnostische Güte der eingeschlossenen Primärstudien beim Primärstaging....	45

5.3.2.1	Studiendesigns und Studienpopulationen der eingeschlossenen Primärstudien zum Primärstaging.....	48
5.3.2.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien beim Primärstaging	61
5.3.2.3	Einschätzung der Übertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien auf den deutschen Versorgungskontext beim Primärstaging ..	62
5.3.2.4	Ergebnisse zur diagnostischen Güte der eingeschlossenen Studien beim Primärstaging	64
5.3.3	Prognostische Güte auf Grundlage von Primärstudien beim Primärstaging (Ergänzungsrecherche)	83
5.3.3.1	Studiendesigns und Studienpopulationen der eingeschlossenen Prognosestudien beim Primärstaging	85
5.3.3.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Prognosestudien beim Primärstaging	90
5.3.3.3	Einschätzung der Übertragbarkeit der Ergebnisse der eingeschlossenen Prognosestudien beim Primärstaging auf den deutschen Versorgungskontext.....	92
5.3.3.4	Ergebnisse zur prognostischen Güte der eingeschlossenen Primärstudien beim Primärstaging	92
5.4	Teilziel 2b: diagnostische und prognostische Güte beim Restaging.....	98
5.4.1	Diagnostische und prognostische Güte beim Restaging auf Grundlage in die Bewertung eingeflossener Evidenzsynthesen.....	98
5.4.1.1	Studiendesigns und Studienpopulationen der eingeschlossenen Evidenzsynthesen beim Restaging	98
5.4.1.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Evidenzsynthesen	101
5.4.1.3	Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien aus den Evidenzsynthesen.....	101
5.4.1.4	Ergebnisse zur diagnostischen Güte auf Grundlage der eingeschlossenen Evidenzsynthesen	102
5.4.2	Diagnostische Güte beim Restaging auf Grundlage von Primärstudien (Ergänzungsrecherche)	106
5.4.2.1	Studiendesigns und Studienpopulationen der eingeschlossenen Primärstudien beim Restaging.....	106
5.4.2.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien beim Restaging (Ergänzungsrecherche).....	110
5.4.2.3	Einschätzung der Übertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien auf den deutschen Versorgungskontext beim Restaging (Ergänzungsrecherche)	110
5.4.2.4	Ergebnisse zur diagnostischen Güte der eingeschlossenen Primärstudien beim Restaging (Ergänzungsrecherche)	111

5.4.3	Prognostische Güte der eingeschlossenen Primärstudien beim Restaging (Evidenzsynthesen und Ergänzungsrecherche)	116
5.4.3.1	Studiendesigns und Studienpopulationen der eingeschlossenen Prognosestudien beim Restaging (Evidenzsynthesen und Ergänzungsrecherche).....	117
5.4.3.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Prognosestudien beim Restaging (Evidenzsynthesen und Ergänzungsrecherche).....	120
5.4.3.3	Einschätzung der Übertragbarkeit der Ergebnisse der Prognosestudien auf den deutschen Versorgungskontext beim Restaging (Evidenzsynthesen und Ergänzungsrecherche).....	122
5.4.3.4	Ergebnisse der Primärstudien zur prognostischen Güte beim Restaging (Evidenzsynthesen und Ergänzungsrecherche)	122
5.4.3.5	Studiendesign und Studienpopulation der Studie zur „prognostic accuracy“ beim Restaging (Ergänzungsrecherche)	125
5.4.3.6	Einschätzung des Verzerrungspotenzials der Studie zur „prognostic accuracy“ beim Restaging	127
5.4.3.7	Einschätzung der Übertragbarkeit der Studie zur „prognostic accuracy“ beim Restaging	127
5.4.3.8	Ergebnisse der Studie zur „prognostic accuracy“ beim Restaging.....	127
5.5	Teilziel 2c: diagnostische und prognostische Güte bei der Rezidivdiagnostik..	129
5.5.1	Diagnostische und prognostische Güte auf Grundlage in die Bewertungen eingeflossener Evidenzsynthesen bei der Rezidivdiagnostik.....	129
5.5.2	Diagnostische Güte auf Grundlage der eingeschlossenen Primärstudien bei der Rezidivdiagnostik.....	129
5.5.3	Studiendesigns und Studienpopulationen der eingeschlossenen Primärstudien bei der Rezidivdiagnostik.....	130
5.5.3.1	Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien bei der Rezidivdiagnostik.....	133
5.5.3.2	Einschätzung der Übertragbarkeit der Ergebnisse der Primärstudien auf den deutschen Versorgungskontext	134
5.5.3.3	Ergebnisse der Primärstudien zur diagnostischen Güte bei der Rezidivdiagnostik	135
5.5.4	Prognostische Güte auf Grundlage der Primärstudien bei der Rezidivdiagnostik.....	138
5.5.4.1	Studiendesign und Studienpopulation der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	138
5.5.4.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	141
5.5.4.3	Einschätzung der Übertragbarkeit der Ergebnisse der Prognosestudie auf den deutschen Versorgungskontext bei der Rezidivdiagnostik.....	143

5.5.4.4	Ergebnisse der Primärstudien zur prognostischen Güte bei der Rezidivdiagnostik	143
5.6	Zusammenfassung der Ergebnisse	145
6	Diskussion.....	147
6.1	Gesamtbewertung der Evidenzlage.....	147
6.1.1	Studien zur Untersuchung des patientenrelevanten Nutzens.....	147
6.1.2	Diagnostische und prognostische Güte der PET und PET/CT beim Primärstaging.....	147
6.1.3	Diagnostische und prognostische Güte der PET und PET/CT beim Restaging.	149
6.1.4	Diagnostische und prognostische Güte der PET und PET/CT bei der Rezidivdiagnostik	156
6.2	Vergleich mit den Empfehlungen wissenschaftlicher Fachgesellschaften zum Einsatz der PET und PET/CT beim Ösophaguskarzinom	157
6.3	Strahlenschutzüberlegungen.....	159
6.4	Allgemeine methodische Probleme.....	160
7	Fazit.....	161
8	Liste der eingeschlossenen Studien	162
9	Literatur	170
Anhang A	– Suchstrategien	183
Anhang B	– Liste der im Volltext überprüften, aber ausgeschlossenen Evidenzsynthesen mit Ausschlussgründen	196
Anhang C	– Liste der ausgeschlossenen Primärstudien mit Ausschlussgründen	201
Anhang D	– Beurteilung der Qualität von Evidenzsynthesen.....	212
Anhang D.1	– Qualitätsbeurteilung von Evidenzsynthesen nach Oxman und Guyatt-Kriterien	212
Anhang D.2	– Ergebnisse der Qualitätsbewertung der Evidenzsynthesen, die nach Oxman und Guyatt bewertet wurden, aber nicht in den Bericht eingeschlossen wurden.....	213
Anhang D.3	– Bewertung des Verzerrungspotenzials von Primärstudien auf Basis von QUADAS-Kriterien	214
Anhang D.4	– Einschätzung der Übertragbarkeit der Studienergebnisse von diagnostischen Primärstudien.....	215
Anhang D.5	– Bewertung des Verzerrungspotenzials von prognostischen Primärstudien.....	216
Anhang D.6	– Einschätzung der Übertragbarkeit Studienergebnisse der prognostischen Primärstudien	217
Anhang E	– Ergebnisse aus einfaktorierter und multifaktorieller Analyse der eingeschlossenen Prognosestudien	218

Tabellenverzeichnis

	Seite
Tabelle 1: Inzidenz und Mortalität nach Altersgruppen in Deutschland 2004 (Fälle pro 100 000).....	2
Tabelle 2: Stadieneinteilung nach dem TNM-System	3
Tabelle 3: Stadieneinteilung nach AJCC / UICC.....	4
Tabelle 4: Anwendung der FDG-PET bei Ösophaguskarzinomen. Bewertung durch die 3. Interdisziplinäre Konsensuskonferenz Onko-PET III.....	8
Tabelle 5: Indikationen für den Einsatz der PET im Rahmen der Diagnostik von Ösophaguskarzinomen	10
Tabelle 6: Klassifikationsschemata für Evaluierungsstudien zu diagnostischen Testverfahren.....	14
Tabelle 7: Evidenzklassifizierung des G-BA für Unterlagen zu diagnostischen Methoden....	15
Tabelle 8: Tabellarische Darstellung der Einschlusskriterien für Primärstudien zur Nutzenbewertung der PET bzw. PET/CT	18
Tabelle 9: Tabellarische Darstellung der Einschlusskriterien für Evidenzsynthesen zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT	22
Tabelle 10: Tabellarische Darstellung der Einschlusskriterien für Primärstudien zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT (Ergänzungsrecherche).....	23
Tabelle 11: Suche in Kongressbänden	38
Tabelle 12: Studienpool zur Beurteilung der diagnostischen und prognostischen Güte der PET bzw. PET/CT zum Restaging	40
Tabelle 13: Überblick über die in den Evidenzsynthesen eingeschlossenen Primärstudien (n = 11) zur diagnostischen Güte beim Restaging (alphabetisch geordnet).....	40
Tabelle 14: Studienpool zur diagnostischen und prognostischen Güte der eingeschlossenen Primärstudien (Ergänzungsrecherche).....	42
Tabelle 15: Designs, primäre Studienziele und untersuchte Zielgrößen der eingeschlossenen Primärstudien beim Primärstaging	48
Tabelle 16: Index- und Referenztest der eingeschlossenen Primärstudien beim Primärstaging	52
Tabelle 17: Eingeschlossene Patienten der Primärstudien beim Primärstaging (Ergänzungsrecherche).....	58
Tabelle 18: Verzerrungspotenzial der eingeschlossenen Primärstudien beim Primärstaging..	62

Tabelle 19: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien beim Primärstaging	64
Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging	66
Tabelle 21: Ergebnisse der bivariaten Meta-Analyse für das N-Staging (gepoolte Schätzer, direkte Vergleiche)	79
Tabelle 22: Ergebnisse der bivariaten Meta-Analyse für das M-Staging (gepoolte Schätzer, direkte Vergleiche)	81
Tabelle 23: Ergänzend betrachtete Ergebnisse zu Änderungen im klinischen Management im Vergleich zur konventionellen Diagnostik (CT und EUS) beim Primärstaging.....	83
Tabelle 24: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudien beim Primärstaging	85
Tabelle 25: Index- und Referenztest der eingeschlossenen Prognosestudien beim Primärstaging	87
Tabelle 26: Eingeschlossene Patienten der eingeschlossenen Prognosestudien beim Primärstaging	88
Tabelle 27: Verzerrungspotenzial der eingeschlossenen Prognosestudien beim Primärstaging	91
Tabelle 28: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien zum Primärstaging	92
Tabelle 29: Ergebnisse der Primärstudien zur prognostischen Güte beim Primärstaging	94
Tabelle 30: Fragestellung und Methodik der Evidenzsynthesen zur diagnostischen und prognostischen Güte der PET bzw. PET/CT beim Restaging.....	99
Tabelle 31: Ergebnisse der Qualitätsbewertung der eingeschlossenen Evidenzsynthesen nach Oxman und Guyatt.....	101
Tabelle 32: Ergebnisse der Studien aus den Evidenzsynthesen zum Restaging	104
Tabelle 33: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche).....	107
Tabelle 34: Index- und Referenztest der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche).....	108
Tabelle 35: Eingeschlossene Patienten der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche).....	109
Tabelle 36: Verzerrungspotenzial der eingeschlossenen Studien beim Restaging (Ergänzungsrecherche).....	110

Tabelle 37: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Studien beim Restaging (Ergänzungsrecherche).....	111
Tabelle 38: Ergebnisse der Studien zur diagnostischen Güte der PET und anderer Indextests beim Restaging von Ösophaguskarzinomen (Ergänzungsrecherche).....	112
Tabelle 39: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudien beim Restaging	118
Tabelle 40: Index- und Referenztest der eingeschlossenen Prognosestudien beim Restaging.....	118
Tabelle 41: Eingeschlossene Patienten der eingeschlossenen Prognosestudien beim Restaging.....	119
Tabelle 42: Verzerrungspotenzial der eingeschlossenen Prognosestudien beim Restaging ..	121
Tabelle 43: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Prognosestudien bei der Restaging.....	122
Tabelle 44: Ergebnisse der Studien zur prognostischen Güte der PET beim Restaging.....	124
Tabelle 45: Design, primäres Studienziel und untersuchte Zielgröße der Studie zur „prognostic accuracy“ beim Restaging	125
Tabelle 46: Index- und Referenztest der Studie zur „prognostic accuracy“ beim Restaging	125
Tabelle 47: Eingeschlossene Patienten der Studie zur „prognostic accuracy“ beim Restaging.....	126
Tabelle 48: Verzerrungspotenzial der Studie zur „prognostic accuracy“ bei der Rezidivdiagnostik.....	127
Tabelle 49: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der Studie zur „prognostic accuracy“ beim Restaging	127
Tabelle 50: Ergebnisse der Studie zur „prognostic accuracy“ bei der Restaging	128
Tabelle 51: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Studien bei der Rezidivdiagnostik.....	130
Tabelle 52: Index- und Referenztest der eingeschlossenen Studien bei der Rezidivdiagnostik.....	131
Tabelle 53: Eingeschlossene Patienten der eingeschlossenen Studien bei der Rezidivdiagnostik.....	132
Tabelle 54: Verzerrungspotenzial der eingeschlossenen Studien bei der Rezidivdiagnostik	133
Tabelle 55: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Studien bei der Rezidivdiagnostik.....	134
Tabelle 56: Ergebnisse der Studien zur diagnostischen Güte der PET bzw PET/CT bei der Rezidivdiagnostik.....	136

Tabelle 57: Ergänzend betrachtete Ergebnisse zu Änderungen im klinischen Management im Vergleich zur CT bei der Rezidivdiagnostik.....	138
Tabelle 58: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik	139
Tabelle 59: Index- und Referenztest der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	139
Tabelle 60: Eingeschlossene Patienten der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	140
Tabelle 61: Verzerrungspotenzial der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	142
Tabelle 62: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	143
Tabelle 63: Ergebnisse der Studien zur prognostischen Güte der PET bei der Rezidivdiagnostik.....	144
Tabelle 64: Ergebnisse der Evidenzsynthesen Van Westreenen 2004, Van Vliet 2008 und des IQWiG-Vorberichts beim Primärstaging bei Ösophaguskarzinomen	149
Tabelle 65: Schlussfolgerungen der eingeschlossenen Evidenzsynthesen beim Restaging von Ösophaguskarzinomen	151

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnisse der bibliografischen Literaturrecherche und des Literaturscreening.....	35
Abbildung 2: Forest Plot der univariaten Ergebnisse für die diagnostische Güte der PET und CT beim N-Staging von Ösophaguskarzinomen.....	77
Abbildung 3: Ergebnisse der bivariaten Meta-Analyse PET versus CT beim N-Staging von Ösophaguskarzinomen (direkte Vergleiche).....	79
Abbildung 4: Forest Plot der Ergebnisse für die diagnostische Güte der PET beim M-Staging von Ösophaguskarzinomen.....	80
Abbildung 5: Ergebnisse der bivariaten Meta-Analyse PET versus CT beim M-Staging von Ösophaguskarzinomen (direkte Vergleiche).....	81
Abbildung 6: Forest Plot der univariaten Ergebnisse für die diagnostische Güte der PET bzw. PET/CT und weiterer bildgebender Verfahren beim Restaging von Ösophaguskarzinomen nach Abschluss der Therapie.....	114
Abbildung 7: Forest Plot der univariaten Ergebnisse für die diagnostische Güte der PET bzw. PET/CT und weiterer bildgebender Verfahren beim Interimsstaging von Ösophaguskarzinomen.....	115
Abbildung 8: Forest Plot der Ergebnisse für die diagnostische Güte der PET bei der Rezidiverkennung von Ösophaguskarzinomen.....	137

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AJCC	American Joint Committee on Cancer
ASA	American Society of Anesthesiologists
ASCO	American Society of Clinical Oncology
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
CCT	Controlled Clinical Trials
CEA	Tumormarker Carcinoembryonales Antigen
CT	Computertomographie
EANM	European Association of Nuclear Medicine
ECCO	European Cancer Organisation
ESMO	European Society of Medical Oncology
EUS	Endosonografie
FDG	2-[¹⁸ F]Fluor-Deoxy-D-Glukose
G-BA	Gemeinsamer Bundesausschuss
G-I-N	Guidelines International Network
HR	Hazard Ratio
HSROC	Hierarchical Summary Receiver Operating Curve
HTA	Health Technology Assessment
ITT	Intention to treat
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KI	Konfidenzintervall
NAC	Neoadjuvante Chemotherapie
NAC(R)	Neoadjuvante Chemotherapie (Radiotherapie)
NCCN	National Comprehensive Cancer Network
NGC	National Guideline Clearinghouse
NPV	Negativer Prädiktiver Wert
MRT	Magnetresonanztomographie
OP	Operation
OR	Odds Ratio
PEG	Perkutane Endoskopische Gastrotomie
PET	Positronenemissionstomographie
PET/CT	Integration von PET und CT in einem Untersuchungsgerät
PICO	Patient, intervention, comparison (Vergleichsintervention), outcome (patientenrelevanter Endpunkt)

Abkürzung	Bedeutung
POSSUM	physiological and operative severity score for the enumeration of mortality and morbidity
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RCT	randomised controlled trial (randomisierte kontrollierte Studie)
RKI	Robert Koch-Institut
ROC	Receiver operating characteristic
SIGN	Scottish Intercollegiate Guidelines Network
SUV	standardized uptake value
TI	Thaliumchlorid
SPECT	Single Photon Emission Computed Tomography
TI-SPECT	Thaliumchlorid Single Photon Emission Computed Tomography
TNM-Klassifikation	Tumorstadieneinteilung nach Tumorausdehnung (T = Tumor); Lymphknotenbefall (N = Nodi lymphatici) und Metastasierung (M = Metastase)
UICC	Union internationale contre le cancer
VOPT	Verification of Only Positive Testers

1 Hintergrund

1.1 Definition des Krankheitsbildes

Als Speiseröhrenkrebs bezeichnet man eine Gruppe von Tumoren im Bereich der Speiseröhre (Ösophagus).

Histologisch sind die beiden wichtigsten Vertreter das Plattenepithelkarzinom (Häufigkeit 60 % bis 70 %) und das Adenokarzinom (30 % bis 40 %). Andere Histologien wie anaplastische bzw. kleinzellige Karzinome, Zylindrome, Karzinoide und Leiomyosarkome machen unter 5 % der Fälle aus [1].

Während Plattenepithelkarzinome überwiegend im mittleren und unteren Drittel der Speiseröhre entstehen, finden sich Adenokarzinome am häufigsten im unteren Drittel [2]. Zudem unterscheiden sie sich auch in der Entstehung und der Therapie. Dies wird nachfolgend dargestellt.

1.2 Epidemiologie und individuelle Krankheitslast

Weltweit gilt der Speiseröhrenkrebs als die sechsthäufigste Krebserkrankung, wobei 80 % bis 85 % der Fälle in Entwicklungsländern, insbesondere im sogenannten „Asian oesophageal cancer belt“ (Türkei, Iran, Afghanistan, Russland und China), auftreten [3,4]. Nach Angaben der World Health Organization lag die globale Inzidenz im Jahr 2008 bei 326 245 Fällen [5]. Die höchste Inzidenz tritt in der Altersgruppe der 50- bis 70-Jährigen auf, wobei Männer 3- bis 5-mal häufiger betroffen sind als Frauen [6,7]. Die Inzidenz von Speiseröhrenkrebs ist global betrachtet sehr unterschiedlich. Die höchsten Inzidenzraten finden sich in Asien sowie im Osten und Süden Afrikas [6].

In Deutschland erkranken pro Jahr nach den Daten der deutschen bevölkerungsbezogenen Krebsregister etwa 3900 Männer und etwa 1050 Frauen an Speiseröhrenkrebs. Bei Frauen entspricht dies einem Anteil von 0,5 % an allen bösartigen Neubildungen, bei Männern einem Anteil von 1,7 %. Das mittlere Erkrankungsalter liegt bei Männern bei 65 Jahren, das der Frauen bei 70 Jahren. Laut Angaben des RKI erkranken in Deutschland Männer etwa 4-mal häufiger als Frauen. Der Anteil der Todesfälle durch Speiseröhrenkrebs ist für Männer mit 3,1 % deutlich höher als für Frauen mit 1,1 % [2].

1.3 Ursache der Erkrankung

Verschiedene Faktoren begünstigen die Entwicklung von Tumoren der Speiseröhre.

Alkohol- und Tabakkonsum gelten als die wichtigsten Risikofaktoren für die Entstehung von Speiseröhrenkrebs [2,6,8]. Eine familiäre Häufung von Erkrankungsfällen ist bekannt [2].

Die häufige Einnahme von heißen Speisen und Getränken ist mit einem erhöhten Krebsrisiko verbunden [7]. Als weiterer Risikofaktor in der Entstehung von Ösophaguskarzinomen gilt

das Essen von nitrosamin-haltigen Speisen, wie z. B. gepökeltem Fleisch [9]. Ebenso ist eine Ernährung, die wenig frisches Obst und Gemüse beinhaltet, mit einem erhöhten Risiko der Entstehung von Plattenepithel- und Adenokarzinomen assoziiert [6].

Tumoren der Speiseröhre treten gehäuft im Alter mit einem Inzidenzspitzen im achten Lebensjahrzehnt und häufiger bei Männern als bei Frauen auf [6,7,10].

Ob Umwelteinflüsse wie vorherige Strahlenbelastung oder bestimmte Tätigkeiten in der Industrie das Risiko für das Auftreten von Plattenepithelkarzinomen ebenfalls erhöhen, wird diskutiert [6]. Übergewicht und regelmäßiger Verzehr von Fleisch gelten als weitere Risikofaktoren [6,10].

Ein Barrett-Ulkus, eine Schleimhautveränderung infolge von häufigem Rückfluss von Mageninhalt in die Speiseröhre (Refluxösophagitis), wird als Präkanzerose für Adenokarzinome angesehen [2,6].

1.4 Verlauf der Erkrankung

Die Mortalitätsraten für das Adeno- und Plattenepithelkarzinom sind vergleichbar, da beide meist erst in einem späten Stadium diagnostiziert werden [2,6,11,12]. Die 5-Jahres-Überlebensrate beträgt in westlichen Industrienationen etwa 10 % [6]. Die Empfehlungen des Tumorzentrums Freiburg gehen von einer Rate zwischen 5 % und 10 % aus [1]. Daten des Robert Koch-Institutes geben die relative 5-Jahres-Überlebensrate in Deutschland für Männer mit etwa 22 % und für Frauen mit 20 % an [2].

Tabelle 1: Inzidenz und Mortalität nach Altersgruppen in Deutschland 2004 (Fälle pro 100 000) [2]

Alter in Jahren	Männer		Frauen	
	Inzidenz	Mortalität	Inzidenz	Mortalität
bis unter 15	0,0	0,0	0,0	0,0
15 bis unter 35	0,1	0,1	0,0	0,0
35 bis unter 40	0,8	0,6	0,0	0,0
55 bis unter 60	21,7	19,1	5,0	3,8
60 bis unter 65	28,1	24,0	5,3	4,6
65 bis unter 70	34,8	29,0	5,7	5,6
70 bis unter 75	33,5	30,3	7,2	5,6
75 bis unter 80	29,5	32,4	7,5	8,8
80 bis unter 85	26,3	33,3	9,2	11,1
85 und älter	23,2	33,7	9,3	16,2

1.5 Klassifikation

Klinisch-pathologisch lassen sich Tumoren der Speiseröhre nach dem TNM-Schema einteilen. Dabei beschreibt das „T“ (Tumor) die Ausdehnung des Primärtumors. Das „N“ (Knoten, von lateinisch nodus) gibt an, ob nahe liegende Lymphknoten befallen sind, und das „M“ (Metastasen) definiert, ob und wo Absiedlungen des Primärtumors vorhanden sind.

Tabelle 2: Stadieneinteilung nach dem TNM-System [13]

T Primärtumor	
TX	Primärtumor nicht beurteilbar
T0	kein Primärtumor
Tis	Carcinoma in situ
T1	Tumor begrenzt auf Lamina propria und Lamina submucosa
T2	Tumor infiltriert die Muscularis propria
T3	Tumor infiltriert die Adventitia
T4	Tumor infiltriert angrenzende extraösophageale Strukturen
N Lymphknotenbefall	
NX	regionäre Lymphknoten nicht beurteilbar
N0	Lymphknoten tumorfrei
N1	Befall regionaler Lymphknoten
M Fernmetastasen	
Mx	Fernmetastasen nicht beurteilbar
M0	keine Fernmetastasen
Tumoren des oberen Drittels	
M1a	Metastasen in Lymphknoten im Bereich des Halses
M1b	andere Fernmetastasen
Tumoren des mittleren Drittels	
M1a	nicht anwendbar
M1b	Befall nicht regionaler Lymphknoten oder / und andere Fernmetastasen
Tumoren des unteren Drittels	
M1a	Metastasen nahe dem Truncus coeliacus
M1b	andere Fernmetastasen

Entsprechend den Ergebnissen der TNM-Klassifikation werden Tumoren der Speiseröhre in verschiedene Grade eingeteilt.

Tabelle 3: Stadieneinteilung nach AJCC / UICC [13]

Stadium	T	N	M
0	Tis	N0	M0
I	T1	N0	M0
II A	T2, T3	N0	M0
II B	T1, T2	N1	M0
III	T3	N1	M0
	T4	jedes N	M0
IV A	jedes T	jedes N	M1a
IV B	jedes T	jedes N	M1b

1.6 Diagnostische Verfahren

Eine Reihe von klinischen Symptomen kann auf einen Tumor im Bereich der Speiseröhre hindeuten. Häufig treten Schluckstörungen, Erbrechen, Appetitlosigkeit, Gewichtsverlust und Minderung der körperlichen Leistungsfähigkeit auf [1,10]. Außerdem können Schmerzen hinter dem Brustbein, Heiserkeit, Husten oder Atemnot während der Nahrungsaufnahme Symptome eines Tumors im Bereich der Speiseröhre sein [1]:

Im Zuge der Recherchen wurde eine Leitlinie der Deutschen Krebsgesellschaft und der Deutschen Gesellschaft für Chirurgie (AWMF Nr. 032 / 008, Stand: 10 / 2003) identifiziert. Diese interdisziplinäre Leitlinie „Ösophaguskarzinom“ ist zum Zeitpunkt der Entstehung des vorliegenden Berichts nicht mehr aktuell und somit nicht mehr verfügbar.

Nach Empfehlungen des Tumorzentrums Freiburg können folgende Schritte in der Diagnostik von Tumoren der Speiseröhre zum Einsatz kommen [1]:

- Anamnese, inklusive Abklärung eventuell vorliegender Risikofaktoren
- Klinische Untersuchung
- Labordiagnostik (Blutbild, Gerinnungswerte, Leber- und Nierenwerte)
- Endoskopie der Speiseröhre mit Biopsie zur Histologiegewinnung und Erfassung der Tumorausdehnung
- Röntgenübersichtsaufnahme des Thorax
- Ösophagusbreischluck (mit wasserlöslichem Kontrastmittel)
- Abdomensonographie
- Computertomographie von Thorax, Hals und Abdomen (zum Staging bzw. Ausschluss intraabdomineller Metastasen)
- Bronchoskopie zum Ausschluss einer Infiltration der Luftröhre

- Laryngoskopie zum Ausschluss einer Beteiligung des rückläufigen Kehlkopfnerve (Nervus laryngeus recurrens) und von Zweitkarzinomen im Bereich des Kehlkopfes (Larynx) bzw. Rachens (Pharynx)
- Ggf. endosonographische Untersuchungen (zur weiteren Abklärung der Tumorausdehnung)
- Skelettszintigraphie (bei Verdacht auf Knochenmetastasen)
- Ggf. Bestimmung des Tumormarkers CEA (Carcinoembryonales Antigen, zur Verlaufsbestimmung)

Aufgrund der therapieassoziierten Mortalität wird eine genaue Erfassung des Gesamtzustandes des Patienten empfohlen [1,10]. Dabei können verschiedene Scoringssysteme wie POSSUM (physiological and operative severity score for the enumeration of mortality and morbidity) und ASA (American Society of Anesthesiologists) eingesetzt werden [10].

Die Empfehlungen des Tumorzentrums Freiburg entsprechen dabei im Wesentlichen den Empfehlungen des Scottish Intercollegiate Guidelines Network (SIGN) [10]. Nach den schottischen Empfehlungen könnte die Endoskopie bei Patienten mit hohem Risiko für therapieassoziierte Mortalität ggf. noch um eine Chromoendoskopie ergänzt werden. Dabei können verschiedene Färbemittel eingesetzt werden, die die Diagnose von kleineren Läsionen verbessern sollen. Der Nutzen dieser Diagnostik ist allerdings noch nicht abschließend geklärt.

1.7 Derzeit übliche therapeutische Verfahren

Die Therapie von Tumoren der Speiseröhre richtet sich nach der Tumorausbreitung und dem allgemeinen Gesundheitszustand des Patienten [1,14]. In über der Hälfte der Patienten wird eine Operation aufgrund des unzureichenden Gesundheitszustandes abgelehnt [10]. Es lässt sich die kurative Therapie von palliativer Therapie unterscheiden.

Wichtigste kurative Therapie ist die Resektion der Speiseröhre mit gleichzeitiger Entfernung der benachbarten Lymphknoten. Dazu kommen Patienten infrage, bei denen keine Fernmetastasen vorliegen und keine lokale Inoperabilität besteht (z. B. bei Infiltration der Luftröhre [Trachea]) [1,12,14]. Dieser Eingriff sollte in Einrichtungen, die bereits ausreichende Erfahrungen mit ähnlichen Operationen haben, durch erfahrene Chirurgen durchgeführt werden [10]. Bezüglich einer ablativen endoskopischen Therapie für Tumoren mit geringer Ausbreitung (T1 / N0 / M0) gibt es widersprüchliche Empfehlungen. Zwei Publikationen sprechen keine Empfehlung aus, da diese Therapie noch Gegenstand aktueller Forschung sei [1,14]. Dagegen wird ein endoskopisches Verfahren durch das SIGN empfohlen [10]. Für lokalisierte Tumoren kann unter Umständen Radiotherapie als Primärtherapie zum Einsatz kommen [10,11]. Ein multimodaler Therapieansatz unter Einbeziehung von Operation, Chemotherapie und Strahlentherapie ist einfachen Ansätzen überlegen [1,10,12,14,15]. Eine kombinierte Radiochemotherapie könnte zum Einsatz

kommen, um einen Tumor vor einer Operation zu verkleinern, oder als Teil der Nachsorge nach einer Resektion. Kontrollen im Abstand von 3 Monaten werden in der Anfangszeit nach der Therapie empfohlen [1].

In der palliativen Therapie stehen supportive Maßnahmen wie Bougierung (Aufdehnung der tumorbedingten Engstelle), Tubus- und Stentimplantation und Anlage eines Zugangs zum Magen durch die Bauchwand (sogenannte perkutane endoskopische Gastrotomie [PEG-Sonde]) im Vordergrund. So lässt sich oftmals eine Verbesserung der Lebensqualität erzielen [1,10]. Die Radiotherapie kann zum Einsatz kommen, um bei Patienten in schlechtem Allgemeinzustand Schluckbeschwerden zu vermindern [10]. Eine Chemotherapie (ggf. in Kombination mit Radiotherapie) kann palliativ bei Patienten mit Fernmetastasen verwendet werden [1,10,16]. In der weiteren Behandlung wird ein symptomorientiertes Vorgehen empfohlen [1,10].

1.8 PET und PET/CT

Die Positronenemissionstomographie (PET) ist ein nichtinvasives diagnostisches Verfahren, das freigesetzte Photonen registriert. Dazu werden radioaktive Substanzen (sogenannte Tracer) über eine Vene appliziert. Durch den Zerfall des Tracers werden Positronen freigesetzt. Diese kollidieren mit Elektronen und setzen bei ihrer Annihilation 2 Photonen frei, die in entgegengesetzter Richtung (180 Grad) fortfliegen. Mittels ringförmig um die Patienten angeordneter Photonendetektoren werden diese erfasst. Aus der räumlichen Verteilung der registrierten Photonen wird auf die räumliche Verteilung der radioaktiven Substanz im Körperinneren geschlossen.

In der Anfangszeit der PET wurden sogenannte Nichtvollringssysteme hergestellt. Die Empfindlichkeit der später entwickelten Vollring-PET-Systeme war derjenigen der ursprünglichen so stark überlegen, dass ausschließlich diese neueren Systeme eine breite Anwendung fanden. Die ersten Vollring-PET-Systeme hatten nur einen Detektorring; bei modernen Systemen werden heute mehrere Ringe nebeneinander angeordnet, wodurch die Sensitivität des Gesamtsystems noch weiter gesteigert wurde.

Im Gegensatz zu den klassischen bildgebenden Verfahren, bei denen Informationen über die anatomische Struktur und damit die Lokalisation von Läsionen erhoben werden (sogenannte morphologische Bildgebung), kann die PET dazu Funktion und Metabolismus der Gewebe charakterisieren (sogenannte metabolische Bildgebung). Dadurch können biochemische und physiologische Prozesse sichtbar gemacht werden, weswegen man die PET zu den Verfahren der funktionellen Bildgebung zählt.

Als Tracer finden unterschiedliche radioaktiv markierte Substanzen Verwendung, die bei den verschiedensten Stoffwechselfvorgängen eine Rolle spielen. Je nach Fragestellung variiert die Auswahl des Tracers. Als häufigster Tracer für die PET wird derzeit die Fluorodeoxyglukose ($[^{18}\text{F}]$ -FDG / FDG) eingesetzt.

Zur Befundung wird unter anderem die Stoffwechselaktivität eines bestimmten Areals mit der des umliegenden Gewebes verglichen. Aufgrund von Unterschieden in der physiologischen Aufnahme des Tracers in verschiedenen Körperregionen kann die Zuordnung einer bestimmten Region als potenziell pathologischer sogenannter „erhöhter Uptake“ schwierig sein und erfordert eine erhebliche Expertise des Untersuchers. Um bei der Interpretation eine gewisse Standardisierung zu erreichen, kann der Metabolismus als Standardized Uptake Value (SUV) wiedergegeben werden. Allerdings ist ein erhöhter Umsatz nicht malignom-spezifisch, vielmehr können sich auch andere Prozesse mit erhöhtem Stoffwechsel dahinter verbergen (z. B. Entzündungen).

Die Aktivitätsanreicherungen bei der PET lassen sich anatomisch jedoch nicht immer präzise lokalisieren, da in PET-Bildern in erster Linie Stoffwechselprozesse abgebildet werden; hinzu kommt die im Vergleich zur CT geringere Ortsauflösung von etwa 4 bis 6 mm. Eine technische Weiterentwicklung stellt eine Kombination der PET und der Computertomographie (CT) in einem Untersuchungsgerät (PET/CT) dar. Hier wird der Patient in einem Untersuchungsang durch die beiden Detektorringssysteme von CT und PET gefahren. Die entstehenden Bilder werden im Computer fusioniert. Üblicherweise wird die CT-Information in Graustufen abgebildet und von der PET-Information farbig überlagert. Mit dieser Methode wird die hohe Auflösung der CT (Ortsauflösung < 1 mm) mit den hochauflösenden Stoffwechsellinformationen der PET verknüpft. Die kombinierten PET/CT-Geräte sind seit dem Jahr 2001 auf dem Markt und verdrängen seitdem reine PET-Scanner bei den Neuanschaffungen in deutschen Krankenhäusern [17].

Mittels moderner Software lässt sich eine Überlagerung (sogenannte Soft-Fusion bzw. Koregistrierung) von CT-Bildern und PET-Daten auch aus separaten Untersuchungsschritten berechnen. Dies geschieht mithilfe gemeinsamer Referenzpunkte wie etwa Knochenstrukturen oder anhand externer Positionsdaten. Die Bildqualität eines durch eine solche Rekonstruktion von PET- und CT-Daten entstehenden Bildes ist allerdings i. d. R. schlechter als die einer kombinierten PET/CT.

Zu berücksichtigen sind des Weiteren die Weiterentwicklungen der bildlichen Auflösung sowohl der PET als auch der CT. In der PET-Technologie werden heute High-Definition-, High-Resolution- und Time-of-Flight-Systeme eingesetzt. Im Vergleich zu den vor 5 bis 10 Jahren produzierten PET-Systemen, die eine Ortsauflösung von 5 bis 6 mm erzielten, liegt die Ortsauflösung heute bei einer Größenordnung von 3 bis 4 mm. Durch die Time-of-Flight-Systeme kann das Rauschen reduziert und die Ortsauflösung gesteigert werden. Dadurch können zunehmend kleinere Tumormanifestationen nachgewiesen werden. Eine vergleichbare Entwicklung gab es auch bei der CT [18].

1.9 Empfehlungen wissenschaftlicher Fachgesellschaften zum Einsatz der PET bei Ösophaguskarzinomen

Wie in Abschnitt 1.6 dargestellt, sind evidenzbasierte deutsche Leitlinien nicht mehr verfügbar (AWMF Nr. 032 / 008) oder erwähnen die PET bzw. PET/CT nicht [1]. Die PET wird durch die Empfehlungen des Freiburger Tumorzentrums nicht als Standard empfohlen, da die Rolle der PET in der Diagnostik von Speiseröhrenkrebs „derzeit evaluiert“ wird [1].

Relativ wenige internationale Leitlinien liegen zur Diagnostik von Ösophaguskarzinomen vor, was u. a. aus einer Suche nach dem Begriff „oesophageal cancer“ auf der Homepage des National Guideline Clearinghouse hervorgeht [19]. Oftmals sind diese älter und die PET bzw. PET/CT findet darin keine Erwähnung [8]. Durch das SIGN wird der routinemäßige Einsatz der PET im Staging von Speiseröhrentumoren nicht empfohlen [10].

Dagegen sieht die ESMO die PET als „hilfreich“ in der Identifikation von bisher nicht beschriebenen Metastasen und in der Diagnose bei Verdacht auf ein Rezidiv an [14]. Zudem könnte die PET in der Kontrolle des Therapieansprechens eine Rolle spielen [14].

In einer Richtlinie der niederländischen Vereinigung für Gastroenterologie (Nederlands Genootschap van Maag-Darm-Leverartsen), die 2010 veröffentlicht wurde, wird die FDG-PET für das TNM-Staging nur eingeschränkt empfohlen [9]. Durch die begrenzte Spatialauflösung sei die FDG-PET zur Abgrenzung des Primärtumors nur schlecht einsetzbar. Zur Bestimmung von Metastasen im lokalen Lymphsystem wird Ultraschall als Methode der Wahl beschrieben. In der Diagnostik von Fernmetastasen scheint die FDG-PET der Computertomographie überlegen zu sein. Daher sollte FDG-PET auch zur Indikationsstellung und Planung operativer Eingriffe eingesetzt werden. Dieses wird ebenfalls in der internationalen Leitlinie des NCCN empfohlen [20].

Von der Deutschen Gesellschaft für Nuklearmedizin e. V., der Deutschen Röntgengesellschaft e. V. und den Vertretern onkologischer Fachgesellschaften wurde im Rahmen der 3. Interdisziplinären Konsensuskonferenz Onko-PET III im Jahr 2000 der klinische Nutzen der PET mit FDG bei verschiedenen Indikationen untersucht und bewertet [21]. Tabelle 4 fasst die Empfehlungen der 3. Interdisziplinären Konsensuskonferenz Onko-PET III zusammen:

Tabelle 4: Anwendung der FDG-PET bei Ösophaguskarzinomen. Bewertung durch die 3. Interdisziplinäre Konsensuskonferenz Onko-PET III [21]

Indikation	Anwendung der FDG-PET
Differenzialdiagnose zur Dignität der Tumoren	Aufgrund unzureichender Daten noch nicht endgültig beurteilbar
Staging	Regelmäßige klinische Anwendung
Therapiekontrolle	Aufgrund unzureichender Daten noch nicht endgültig beurteilbar
Rezidivdiagnostik	Aufgrund unzureichender Daten noch nicht endgültig beurteilbar

2 Ziele der Untersuchung

Die vorliegende Untersuchung sollte als Erstes prüfen, welcher patientenrelevante Nutzen oder Schaden mit der bildgebenden Diagnostik mittels PET bzw. PET/CT im Vergleich oder in Ergänzung zur konventionellen Diagnostik beim Krankheitsbild Ösophaguskarzinom verbunden ist.

Falls dies aufgrund der vorliegenden Datenlage nicht möglich ist, weil keine oder zu wenige aussagekräftige Studien vorliegen, sollte als Zweites untersucht werden, ob sich durch den Einsatz der PET bzw. PET/CT die prognostische bzw. diagnostische Güte der Diagnosestrategie verbessert.

Alle Fragestellungen sollten auch im Rahmen der deutschen Strahlenschutzgesetzgebung erörtert werden.

Folgende Technologievergleiche sollten betrachtet werden:

- Diagnostische Strategie mit PET versus diagnostische Strategie ohne PET (einzelne Verfahren, ihre Kombination) oder keine weitere Diagnostik
- Diagnostische Strategie mit PET/CT versus diagnostische Strategie ohne PET/CT (einzelne Verfahren, ihre Kombination) oder keine weitere Diagnostik
- Diagnostische Strategie mit PET/CT versus diagnostische Strategie mit PET

Unter konventioneller diagnostischer Strategie wurde im vorliegenden Bericht die CT, Endosonographie (EUS) oder die Histologie verstanden. Es sollen aber darüber hinaus alle weiteren in den gefundenen Studien angestellten diagnostischen Vergleiche dargestellt werden.

Es erfolgte keine Eingrenzung der Fragestellungen auf bestimmte Tracer.

Allgemeine Voraussetzungen für Studien zur Nutzenbewertung („Nutzenstudien“)

Unter „patientenrelevanter Nutzen und Schaden“ (im Weiteren kurz als „Nutzen und Schaden“ bezeichnet) wurden Veränderungen verstanden, die für den Patienten fassbare Konsequenzen haben, wie z. B. die Auswirkung auf die Mortalität, Morbidität (darunter auch das Auftreten erneuter Rezidive) und gesundheitsbezogene Lebensqualität. Mit dem Begriff „Nutzen“ wurden kausal begründete positive Effekte, mit dem Begriff „Schaden“ kausal begründete negative Effekte einer medizinischen Intervention auf patientenrelevante Endpunkte (Abschnitt 4.1.2.3) bezeichnet. Der Begriff „Nutzenbewertung“ bezeichnet den gesamten Prozess der Evaluation medizinischer Interventionen hinsichtlich ihrer kausal begründeten positiven und negativen Effekte im Vergleich mit einer klar definierten anderen Intervention, einem Placebo oder keiner Intervention [22].

Allgemeine Voraussetzungen für Studien zur Bewertung der diagnostischen und prognostischen Güte

Wenn ausreichend Evidenz zum patientenrelevanten Nutzen und Schaden vorgelegen hätte, wäre eine Nutzenbewertung alleine anhand der Beantwortung dieser Fragestellung möglich gewesen. In diesem Fall wäre die Nutzenbewertung abgeschlossen gewesen. Da jedoch keine ausreichende Evidenz zur Bewertung des patientenrelevanten Nutzens und Schadens vorlag, wurde eine Bewertung der diagnostischen und der prognostischen Güte der PET bzw. der PET/CT durchgeführt.

Unter „diagnostische Güte“ wird eine korrekte Zuordnung der Patienten zum jeweiligen Gesundheitszustand verstanden, unter „prognostische Güte“ eine korrekte Vorhersage der zukünftigen klinischen Ereignisse.

Indikationen für den Einsatz der PET bzw. PET/CT laut Konkretisierungen des G-BA

In Tabelle 5 sind die laut Konkretisierung des G-BA festgelegten Indikationen für den Einsatz der PET und PET/CT dargestellt. Im Rahmen des vorliegenden Berichts soll sowohl der patientenrelevante Nutzen (Fragestellung 1) als auch die diagnostische und prognostische Güte (Fragestellung 2) für die in Tabelle 5 angegebenen Indikationen untersucht werden. Dabei wird die sogenannte Interims-PET (= Diagnostik bei noch laufender Therapie zur Ermittlung des Behandlungsansprechens) unter die Kategorie Restaging subsumiert.

Tabelle 5: Indikationen für den Einsatz der PET im Rahmen der Diagnostik von Ösophaguskarzinomen

Erkrankung	Indikation			
	Primärdiagnostik	Primärstaging	Restaging ^a	Rezidivdiagnostik
Ösophaguskarzinom	-	X	X	X

a: Das Restaging (Therapieansprechen) beinhaltet auch die Interims-PET.

3 Projektbearbeitung

3.1 Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 21.12.2006 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissensstandes zur Positronenemissionstomographie (PET), auch als Integrationsgerät von PET und Computertomographie (PET/CT), bei 14 verschiedenen Krankheitsbildern beauftragt.

Der dem vorliegenden Vorbericht zugrunde liegende Berichtsplan bezog sich auf den Teilauftrag zu 7 onkologischen Krankheitsbildern (23 Indikationen bei Adenokarzinom des Pankreas, malignem Melanom, Schilddrüsenkarzinom, Ösophaguskarzinom, Ovarialkarzinom, Mammakarzinom und Knochen- und Weichteiltumoren) [23]. Der vorliegende Vorbericht bezieht sich auf den Teilauftrag zum Ösophaguskarzinom. Die restlichen 6 Krankheitsbilder werden in separaten Berichten bearbeitet. Spezifische Passagen für diese 6 Krankheitsbilder wurden im vorliegenden Vorbericht nicht berücksichtigt.

Die endgültigen Konkretisierungsunterlagen wurden dem Institut mit dem Schreiben des G-BA vom 11.03.2008 übersandt.

Zur Erstellung des (für die o. g. 7 Krankheitsbilder gemeinsamen) Berichtsplans wurden am 04.09.2009 Patientenvertreter konsultiert. Für das Krankheitsbild Ösophaguskarzinom konnten jedoch keine Patientenvertreter zur Konsultation im Hinblick auf patientenrelevante Endpunkte gewonnen werden.

Der vorläufige (gemeinsame) Berichtsplan in der Version 1.0 vom 16.02.2010 wurde am 24.02.2010 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 07.04.2010 konnten schriftliche Stellungnahmen eingereicht werden. Unklare Aspekte aus den schriftlichen Stellungnahmen zum vorläufigen Berichtsplan wurden am 11.05.2010 in einer wissenschaftlichen Erörterung mit den Stellungnehmenden diskutiert. Die Dokumentation und Würdigung der Anhörung zum Berichtsplan ist auf der Website des IQWiG veröffentlicht. Der endgültige Berichtsplan in der Version 1.0 vom 05.07.2010 wurde am 12.07.2010 auf der Website des IQWiG veröffentlicht.

Der vorliegende Vorbericht wird zur Anhörung gestellt. Hierzu können schriftlich Stellungnahmen eingereicht werden. Das Ende der Stellungnahmefrist wird auf der Website des IQWiG (www.iqwig.de) bekannt gegeben. Stellungnahmen können von allen interessierten Personen, Institutionen und Gesellschaften abgegeben werden. Die Stellungnahmen müssen bestimmten formalen Anforderungen genügen, die ebenfalls auf der Website des IQWiG in einem entsprechenden Leitfaden dargelegt sind. Gegebenenfalls wird eine wissenschaftliche Erörterung zur Klärung unklarer Aspekte aus den schriftlichen Stellungnahmen durchgeführt. Die Anhörung kann zu Änderungen und / oder Ergänzungen des Berichts führen. Im Anschluss an diese Anhörung wird der Abschlussbericht erstellt.

Dieser Bericht wird an den G-BA übermittelt und 8 Wochen später auf der Website des IQWiG veröffentlicht.

3.2 Dokumentation der Änderungen im Projektverlauf

In der Tabelle 10 (Darstellung der Einschlusskriterien für die Primärliteratur zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT) wurde der Einschlussgrund E7c „Publikationszeitraum ist durch Literatursuche der eingeschlossenen Evidenzsynthesen nicht abgedeckt“ hinzugefügt.

Die Einschlusskriterien für vergleichende Studien zur Bewertung des Nutzens und des potenziellen Schadens wurden vereinfacht (vgl. Abschnitt 4.1.2.3).

Der Abschnitt 4.2.3 „Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen“ wurde ergänzt.

In der Tabelle 9 und der Tabelle 10 wurde unter den Einschlusskriterien E2b und E2c der Verweis auf die Beschreibung der Evaluierung der Prüftechnologie korrigiert.

4 Methoden

In der Vergangenheit verband man mit dem Denken über die Bewertung von diagnostischen Tests primär die Beschreibung der technischen Merkmale des Tests und die Evaluation seiner diagnostischen Güte, d. h. seiner Fähigkeit, kranke von gesunden Personen zu unterscheiden [24]. Damit verband man die Erwartung, dass sich eine bessere diagnostische Güte automatisch in eine bessere Patientenversorgung umsetzen würde. Seitdem hat sich das Denken über den Stellenwert diagnostischer Tests im therapeutischen Management weiterentwickelt, wobei neben der verbesserten diagnostischen Güte die Konsequenzen, die aus dem diagnostischen Test für den Arzt, den Patienten und die Gesellschaft resultieren, in das Zentrum der Betrachtung rücken [25,26]. Fryback und Thornbury haben dazu in ihrer Publikation „The efficacy of diagnostic imaging“ von 1991 [25] 6 Stufen identifiziert, die von den technischen Auswirkungen („technical efficacy“) als unterste Stufe über die diagnostische Genauigkeit („diagnostic accuracy efficacy“) als 2. Stufe bis hin zu den therapeutischen Effekten („therapeutic efficacy“) als 4. Stufe und den patientenrelevanten Effekten („patient outcome efficacy“) als 5. Stufe reichen. Als oberste Stufe betrachten sie schließlich die gesellschaftlichen Konsequenzen („societal efficacy“) der diagnostischen Technologie (s. auch Tabelle 6).

Tabelle 6: Klassifikationsschemata für Evaluierungsstudien zu diagnostischen Testverfahren

	Köbberling et al. [26]	Fryback & Thornbury [25]	Evidenzklassifizierung G-BA [27]
Patientenrelevante Zielgrößen	Phase 4: Wirksamkeit	6. Stufe: Auswirkung auf Systemebene 5. Stufe: Auswirkung im Hinblick auf patientenrelevante Endpunkte	Evidenzstufe I
Surrogatzielgröße (Management-änderung)		4. Stufe: Auswirkung im Hinblick auf das therapeutische Denken des Behandelnden 3. Stufe: Auswirkung im Hinblick auf das (differenzial-) diagnostische („Nachtstwahrscheinlichkeit“)	Evidenzstufe II
Sensitivität (SN), Spezifität (SP), Likelihood Ratio (LR), prädiktive Werte	Phase 3: Diagnostische Genauigkeit bei nicht bekanntem Krankheitsstatus	2. Stufe: Auswirkung im Hinblick auf Diskriminationsfähigkeit	Evidenzstufe III
Sensitivität (SN), Spezifität (SP), Likelihood Ratio (LR)	Phase 2: Diagnostische Genauigkeit bei bekanntem Krankheitsstatus		Evidenzstufe III
Analytische Sensitivität, Spezifität, Reproduzierbarkeit	Phase 1: Technische Voruntersuchungen	1. Stufe: Technische Auswirkungen	Evidenzstufe IV

Diese erweiterte Betrachtungsweise findet zunehmend ihren Niederschlag in der klinischen Forschung, wo vermehrt die Auswirkungen eines neuen Tests auf die Entscheidungsfindung des Arztes, auf eine verbesserte Behandlung für den Patienten und auf die Konsequenzen für die Gesellschaft empirisch untersucht werden. Die Mehrzahl der westlichen Gesundheitssysteme zieht Überlegungen zu den Konsequenzen und dem Nutzen und Schaden diagnostischer Tests für Arzt, Patient und Gesellschaft in Betracht, wenn sie über den Wert eines diagnostischen Verfahrens nachdenkt. Auch der G-BA hat sich in seiner Verfahrensordnung bei der Bewertung diagnostischer Tests (§ 18 Abs. 7) dieser Sichtweise angeschlossen und diese Methodik auch in der aktuellen Verfahrensordnung fortgeschrieben [27]. In Tabelle 6 sind verschiedene Klassifikationsschemata für Evaluierungsstudien zu diagnostischen Testverfahren einander gegenübergestellt. Neben der Einteilung von Fryback und Thornbury und der Evidenzklassifizierung des G-BA (s. Tabelle 7) wurde auch das Phasenschema nach Köbberling [26] mit aufgenommen.

Tabelle 7: Evidenzklassifizierung des G-BA für Unterlagen zu diagnostischen Methoden

Evidenzklasse nach G-BA [27]	Kriterien
I a	Systematische Übersichtsarbeiten von Studien der Evidenzstufe I b
I b	Randomisierte kontrollierte Studien
I c	Andere Interventionsstudien
II a	Systematische Übersichtsarbeiten von Studien zur diagnostischen Testgenauigkeit der Evidenzstufe II b
II b	Querschnitt- und Kohortenstudien, aus denen sich alle diagnostischen Kenngrößen zur Testgenauigkeit (Sensitivität und Spezifität, Wahrscheinlichkeitsverhältnisse, positiver und negativer prädiktiver Wert) berechnen lassen
III	Andere Studien, aus denen sich die diagnostischen Kenngrößen zur Testgenauigkeit (Sensitivität und Spezifität, Wahrscheinlichkeitsverhältnisse) berechnen lassen
IV	Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. Ä.; nicht mit Studien belegte Meinungen anerkannter Experten, Berichte von Expertenkomitees und Konsensuskonferenzen

4.1 Ermittlung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT

Um den patientenrelevanten Nutzen und Schaden der bildgebenden diagnostischen Technologien PET bzw. PET/CT für die Behandlung von Patienten mit Ösophaguskarzinom zu bewerten, genügt es nicht festzustellen, inwieweit die PET bzw. die PET/CT den gängigen diagnostischen Methoden hinsichtlich einer korrekten Identifizierung des Tumors oder des Rezidivs (bzw. seiner Ausbreitung im Körper) überlegen ist. Vielmehr geht es bei einer erweiterten Bewertung darum festzustellen, inwieweit sich eine präzisere Diagnosestellung auf die Behandlung und das weitere Management des Patienten auswirkt und inwieweit in letzter Konsequenz, das veränderte Management patientenrelevante Endpunkte wie z. B. Mortalität, Morbidität oder die Lebensqualität beeinflusst.

Diese Fragen lassen sich beispielsweise durch Studien untersuchen, in denen die Patienten je nach PET-Befund randomisiert entweder der bisherigen, von der PET unabhängigen Standardtherapie oder einer aufgrund des PET-Befundes adaptierten Therapie zugeteilt werden (sogenannte „Managementstudien“). Aber auch andere Studiendesigns kommen infrage [28]. Wie sich die Behandlungsoptionen (z. B. aggressiveres oder weniger aggressives Vorgehen, Chemo- oder Strahlentherapie oder beides) verändern und welche Auswirkungen die unterschiedlichen Strategien schließlich auf patientenrelevante Endpunkte haben, lässt sich am besten durch Studien mit ausreichender Nachbeobachtungszeit beantworten (s. auch Abschnitt 4.1.2.3).

4.1.1 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT anhand von Evidenzsynthesen

Im Folgenden werden die Begriffe Sekundärliteratur, systematische Übersichten, HTA-Berichte oder Leitlinien, die auf systematischen Übersichten basieren, unter dem Begriff „Evidenzsynthese“ zusammengefasst.

Vorrecherchen hatten ergeben, dass vermutlich keine qualitativ ausreichende Evidenzsynthese zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bei den genannten Krankheitsbildern (siehe IQWiG-Berichte D0601-E-K) vorlag. Daher sollte der patientenrelevante Nutzen im vorliegenden Bericht nicht auf Basis von Evidenzsynthesen, sondern anhand von Primärstudien untersucht werden.

4.1.2 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT anhand von Primärstudien

4.1.2.1 Population

Studien zu Patienten mit vermutetem oder nachgewiesenem, primärem oder rezidiertem Ösophaguskarzinom wurden eingeschlossen.

4.1.2.2 Prüf- und Vergleichsinterventionen

Als Prüfintervention galt eine diagnostische Strategie mit Anwendung einer Vollring-PET bzw. einer Vollring-PET/CT (Indextest I). Als Vergleichsintervention (Indextest II) wurde eine diagnostische Strategie ohne Anwendung der PET bzw. der PET/CT betrachtet. Ein weiterer Vergleich beinhaltete eine diagnostische Strategie mit Anwendung der Vollring-PET/CT als Prüfintervention und eine diagnostische Strategie mit Vollring-PET (ohne CT) als Indextest II. Folgende Vergleiche sollten betrachtet werden:

- Diagnostische Strategie mit PET versus diagnostische Strategie ohne PET (einzelne Verfahren, ihre Kombination) oder keine weitere Diagnostik
- Diagnostische Strategie mit PET/CT versus diagnostische Strategie ohne PET/CT (einzelne Verfahren, ihre Kombination) oder keine weitere Diagnostik
- Diagnostische Strategie mit PET/CT versus diagnostische Strategie mit PET

Verschiedene (randomisierte) Vergleiche nach vorheriger Diagnostik mit der PET bzw. PET/CT (z. B. Randomisierung der diskordanten Ergebnisse der PET und CT) waren ebenfalls möglich [28].

4.1.2.3 Zielgrößen

Als Zielgrößen wurden die nachfolgend angeführten Endpunkte verwendet, die eine Beurteilung patientenrelevanter Interventionsziele ermöglichten:

- Mortalität
- Krankheitsfreies Überleben (oder ein vergleichbarer Endpunkt)
- Morbidität, z. B. unerwünschte Ereignisse durch aggressivere / weniger aggressive Behandlung (behandlungsbedingte unerwünschte Ereignisse)
- Gesundheitsbezogene Lebensqualität (z. B. SF-36)

Eine Änderung im Patientenmanagement alleine (ohne Verknüpfung zu den oben genannten Zielgrößen) stellte keine patientenrelevante Zielgröße im Sinne dieses Berichts dar und wurde daher nur ergänzend betrachtet.

4.1.2.4 Studientypen

Randomisierte kontrollierte Studien (RCTs) liefern für die Bewertung des Nutzens und Schadens einer medizinischen Intervention die zuverlässigsten Ergebnisse, weil sie, sofern methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt, mit der geringsten Ergebnisunsicherheit behaftet sind. Es wurden deshalb primär RCTs betrachtet.

Im Fall, dass die auf randomisierten kontrollierten Studien basierende Datenlage nicht hinreichend war, um den patientenrelevanten Nutzen und Schaden der PET bzw. PET/CT mit ausreichender Ergebnissicherheit schätzen zu können (z. B. kleine Fallzahl, wenige Ereignisse [geringe Präzision], hohes Verzerrungspotenzial), sollten gegebenenfalls auch verfügbare eindeutig prospektiv geplante, jedoch nicht randomisierte Interventionsstudien mit zeitlich parallelen Kontrollgruppen herangezogen werden (sogenannte Controlled Clinical Trials: CCTs), sofern das Problem einer möglichen Strukturungleichheit (unfairer Vergleich) der Beobachtungsgruppen adäquat in der Planung und Auswertung der entsprechenden Studien berücksichtigt wurde. Solche Studien können zwar die Aussage von aggregierten Ergebnissen aus qualitativ belastbaren RCTs in der Regel nicht qualitativ ändern, diese aber ggf. bestärken.

Sollten sich auch in der 2. Kategorie (CCTs) keine ausreichenden Daten finden, so sollten gegebenenfalls auch verfügbare eindeutig prospektiv geplante, jedoch nicht randomisierte Studien ohne Zuteilung zu den Diagnosegruppen durch die Untersuchungsleiter (Beobachtungsstudien) mit zeitlich parallelen Kontrollgruppen herangezogen werden, sofern das Problem einer möglichen Strukturungleichheit (unfairer Vergleich) der Beobachtungsgruppen adäquat in der Planung und Auswertung der entsprechenden Studien berücksichtigt wurde. Auch diese Studien können die Aussage von aggregierten Ergebnissen aus qualitativ belastbaren RCTs in der Regel nicht qualitativ ändern, diese aber ggf. bestärken.

Dieses Vorgehen schien für die zugrunde liegende Fragestellung gerechtfertigt, da sich erst sehr langsam der Paradigmenwechsel in der Bewertung diagnostischer Verfahren hin zu vergleichenden Studien mit patientenrelevanten Endpunkten vollzieht. Zudem erlauben

möglicherweise nicht randomisierte kontrollierte Interventionsstudien – wie oben skizziert – Hinweise auf die Ausgestaltung zukünftiger aussagekräftigerer Studien.

4.1.2.5 Sonstige Studiencharakteristika

Es erfolgte keine weitere Einschränkung.

4.1.3 Tabellarische Darstellung der Einschlusskriterien („Nutzenstudien“)

In die Nutzenbewertung wurden alle Studien einbezogen, die alle nachfolgenden Einschlusskriterien erfüllten.

Tabelle 8: Tabellarische Darstellung der Einschlusskriterien für Primärstudien zur Nutzenbewertung der PET bzw. PET/CT

Einschlusskriterien für Primärstudien zur Evaluierung des patientenrelevanten Nutzens und Schadens	
E1a	Patienten wie unter 4.1.2.1 beschrieben
E2a	Vergleich von 2 diagnostischen Strategien wie unter 4.1.2.2 beschrieben
E3a	Erhebung von patientenrelevanten Endpunkten wie unter 4.1.2.3 beschrieben
E4a	Randomisierte kontrollierte Studien (Bei nicht hinreichend auf randomisierten kontrollierten Studien basierender Datenlage sollten ggf. prospektive nicht randomisierte kontrollierte Studien [mit und ohne experimentelle Gruppenzuteilung], wie in Abschnitt 4.1.2.4 beschrieben, eingeschlossen werden.)
E5a	Keine Mehrfachpublikationen ohne relevante Zusatzinformationen
E6a	Vollpublikation verfügbar ^a
a: Als Vollpublikation galt in diesem Zusammenhang auch die nicht vertrauliche Weitergabe eines Studienberichts an das Institut oder die nicht vertrauliche Bereitstellung eines Berichts über die Studien, der den Kriterien des CONSORT- [29] oder TREND-Statements [30] genügte und eine Bewertung der Studie ermöglichte.	

4.2 Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT

In den orientierenden Vorrecherchen fanden sich 5 Evidenzsynthesen, die Teilaspekte der Fragestellungen zur diagnostischen Güte der PET oder PET/CT beim Krankheitsbild des Ösophaguskarzinoms untersuchten. Mit dem Ziel einer möglichst raschen Ergebnisgewinnung und Nutzung bereits vorliegender wissenschaftlicher Evidenz auf höchster Evidenzstufe sollte die Bewertung der diagnostischen Güte der PET und PET/CT primär auf Grundlage von Evidenzsynthesen erfolgen, sofern sichergestellt war, dass sie bestimmte Mindestanforderungen erfüllten. Sollte sich keine geeignete Evidenzsynthese zu einer Fragestellung finden, so sollten die diagnostische und die prognostische Güte anhand von Primärstudien beurteilt werden. Waren die identifizierten Arbeiten nicht mehr aktuell (Suchzeitraum vor Februar 2009), wurde für den Zeitraum, den sie nicht abdeckten, relevante Primärliteratur recherchiert und im Bericht berücksichtigt (Ergänzungsrecherche).

Nachfolgend werden zunächst die Kriterien für einen Einschluss von Evidenzsynthesen in die geplante Bewertung aufgeführt (Abschnitt 4.2.1). Anschließend erfolgt die Darstellung für einzuschließende Primärstudien („Ergänzungsrecherche“; Abschnitt 4.2.2).

4.2.1 Einschlusskriterien zur Ermittlung der diagnostischen und prognostischen Güte anhand von Evidenzsynthesen

4.2.1.1 Populationen

Die Evidenzsynthesen mussten Angaben zu Patienten enthalten wie in Abschnitt 4.1.2.1 beschrieben.

4.2.1.2 Prüftechnologien, Vergleichstechnologien („Komparatoren“) und Referenztest

Die den Evidenzsynthesen zugrunde liegenden Primärstudien mussten die Angaben zur diagnostischen Güte anhand folgender Vergleiche ermittelt haben:

Prüftechnologie („Indextest I“) war die Vollring-PET bzw. PET/CT. Als Vergleichstechnologie („Indextest II“) diente jegliche diagnostische Strategie ohne PET bzw. PET/CT („konventionelle Diagnostik“). Außerdem wurde der Vergleich PET vs. PET/CT zugelassen (s. Abschnitt 4.1.2.2).

Indextest I und Indextest II mussten zu einem Referenztest (= dem besten zur Verfügung stehenden Test) in Beziehung gebracht worden sein, der für eine adäquate Interpretation der Ergebnisse unabhängig vom Index- und Vergleichstest erhoben worden war, da es anderenfalls zu Verzerrungen der Güte- und/oder Variabilitätsschätzer kommen konnte (sogenannter Verifikationsbias oder Inkorporationsbias) [31-35].

Als Referenztest (= bester zur Verfügung stehender Test) galt bei der Bewertung der diagnostischen Güte der histopathologische Befund ggf. in Kombination mit dem individuellen Krankheitsverlauf und bei der Bewertung der prognostischen Güte der tatsächlich beobachtete individuelle Krankheitsverlauf im Rahmen der Nachbeobachtung. Obwohl in vielen Fällen der Bewertung der diagnostischen Güte lediglich der individuelle Krankheitsverlauf als Referenzstandard vorliegt, muss konstatiert werden, dass dieser häufig mit systematischen Verzerrungen einhergeht (Disease progression bias, Differential verification bias). In vielen Fällen ist beispielsweise schwer zu beurteilen, was ein adäquater Zeitraum zwischen Index- und Referenztest ist.

4.2.1.3 Zielgrößen

Es wurden Evidenzsynthesen eingeschlossen, deren Ergebnisse auf Primärstudien beruhten, die relevante Angaben zu folgenden Zielgrößen enthielten:

- Beschreibung der diagnostischen Genauigkeit mittels Sensitivität / Spezifität (ggf. ROC-Kurven) [Receiver Operating Characteristic] oder positiven / negativen Likelihood-

Quotienten, daraus abgeleitet ggf. prädiktive Aussagen mittels z. B. positiven / negativen prädiktiven Wertes. Dabei mussten die patientenbezogenen Vierfeldertafel-Daten zur Berechnung der diagnostischen Güte der Primärstudien ableitbar sein.

- Prognostische Aussagen (Wahrscheinlichkeit) hinsichtlich zukünftiger klinischer Ereignisse (z. B. Überleben, Rezidiv) anhand geeigneter statistischer Maßzahlen (z. B. aus Kaplan-Meier-Schätzungen). Es sollten dazu adäquate Maßnahmen zur Berücksichtigung von Störgrößen durchgeführt worden sein.

4.2.1.4 Den Evidenzsynthesen zugrunde liegende Studientypen

Berücksichtigt wurden Ergebnisse von Evidenzsynthesen, die auf Daten beruhten, die anhand folgender Primärstudientypen erhoben worden waren:

- Diagnostische Güte

Prospektiv geplante Querschnittstudien, ggf. mit klinischer Nachbeobachtung, durchgeführt in der Anwendungssituation (gemäß Phase 3 nach Köbberling et al. [26])

- Prognostische Güte

Prospektiv geplante Kohortenstudien: Um die prognostische Güte der PET möglichst unverzerrt bestimmen zu können, sollte eine Gruppe von Patienten, die zu einem bestimmten Zeitpunkt rekrutiert und mit der PET untersucht wurde, im zeitlichen Verlauf untersucht werden (Longitudinalstudie). Dabei waren zum einen ein konsekutiver Einschluss der Patienten und die Dokumentation aller fehlenden Werte notwendig. Zum Zweiten sollten die PET-Positiven und die PET-Negativen zeitlich parallel, prospektiv und im direkten Vergleich beobachtet worden sein (historische Vergleiche oder der Vergleich positiver und negativer Ergebnisse aus unterschiedlichen Patientenkollektiven liefern keine verlässlichen Daten). Im Idealfall sollte die weitere Behandlung dieser Patienten ohne das Wissen um die Ergebnisse der PET erfolgen (Verblindung).

Für die Interpretation des Überlebens war ein Follow-up von mindestens 6 Monaten entscheidend. Evidenzsynthesen, die sowohl prospektive als auch retrospektive Studien enthielten, wurden nur dann berücksichtigt, wenn eine separate Auswertung der in ihnen enthaltenen prospektiv geplanten Studien möglich war.

- Diskordanzstudien und VOPT-Design

Im Fall, dass die Datenlage aus Studien, welche die diagnostischen Vorgehensweisen mit dem Referenztest direkt verglichen, unzureichend war, konnten Evidenzsynthesen der o. g. Studientypen herangezogen werden, die lediglich die diskordanten Befunde zwischen den alternativen diagnostischen Vorgehensweisen untersuchen. Diskordanzstudien mussten so konzipiert sein, dass alle diskordanten Befunde analog zu den Studien zur diagnostischen und prognostischen Güte mit dem gleichen, validen Referenztest untersucht wurden (s. Abschnitt 4.2.1.2). Eine Abwandlung der Diskordanzstudien stellt

das VOPT-Design (Verification of Only Positive Testers) dar, in welchem alle Ergebnisse, die in mindestens einem der beiden Indextests positiv sind, mit dem Referenztest untersucht werden [36].

4.2.1.5 Sonstige Charakteristika der Evidenzsynthesen

Neben diesen genannten Charakteristika mussten die Evidenzsynthesen folgende Kriterien erfüllen:

- Es musste sich um Evidenzsynthesen handeln, die auf systematische, reproduzierbare und transparente Weise Ergebnisse von Primärstudien zusammenfassend dargestellt und bewertet hatten. Daher mussten die Publikationen auf dem Qualitätsindex von Oxman und Guyatt mindestens 5 von 7 möglichen Punkten erreichen [37-39].
- Die jeweilige Forschungsfrage musste im Format einer sog. PICO-Frage formuliert sein, die Population, Indextest I, ggf. Indextest II, Referenztest und Endpunkte explizit benannte, oder die PICO-Fragestellung musste sich der Publikation eindeutig entnehmen lassen.
- Ein möglicherweise in den Primärstudien zu konstatierender Verifikations- bzw. Inkorporationsbias (vgl. Abschnitt 4.2.1.2) sollte adressiert worden sein.
- Die Ergebnisse der Qualitätsbewertung, die Kerninformationen über die Population und die Ergebnisse einer jeden Studie sollten transparent und vergleichbar abgebildet worden sein und die Daten sollten eindeutig mit den zugehörigen Studien verknüpft werden können.

Tabellarische Darstellung der Einschlusskriterien (Evidenzsynthesen zur diagnostischen und prognostischen Güte)

Es wurden alle Evidenzsynthesen einbezogen, die alle nachfolgenden Einschlusskriterien erfüllten.

Tabelle 9: Tabellarische Darstellung der Einschlusskriterien für Evidenzsynthesen zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT

Einschlusskriterien zur Evaluierung der diagnostischen und prognostischen Güte anhand von Evidenzsynthesen	
E1b	Patienten wie in Abschnitt 4.1.2.1 beschrieben
E2b	Prüftechnologie: Vollring-PET bzw. PET/CT, Evaluierung wie in Abschnitt 4.2.1.2 beschrieben
E3b	Zielgrößen wie in Abschnitt 4.2.1.3 beschrieben
E4b	Evidenzsynthesen aus prospektiv geplanten Querschnitt- oder Kohortenstudien wie in den Abschnitten 4.2.1.4 und 4.2.1.5 beschrieben
E5b	Keine Mehrfachpublikationen ohne relevante Zusatzinformation
E6b	Vollpublikation verfügbar ^a
a: Als Vollpublikation galt in diesem Zusammenhang auch die nicht vertrauliche Weitergabe eines Studienberichts an das Institut oder die nicht vertrauliche Bereitstellung eines Berichts über die Studien, der den Kriterien des PRISMA-Statements [40] oder des MOOSE-Statemats [41] genügte und eine Bewertung der Studie ermöglichte.	

4.2.2 Einschlusskriterien zur Ermittlung der diagnostischen und prognostischen Güte anhand von Primärstudien („Ergänzungsrecherche“)

Im Rahmen einer Ergänzungsrecherche wurde in Ergänzung der Evidenzsynthesen Primärliteratur identifiziert und anhand folgender Kriterien eingeschlossen:

4.2.2.1 Populationen

Patienten wie in Abschnitt 4.1.2.1 beschrieben.

4.2.2.2 Prüftechnologien (Indextest I), Vergleichstechnologien (Indextest II) und Referenztest

Prüftechnologie war die Vollring-PET bzw. -PET/CT. Vergleichstechnologie, Referenztest und Evaluierung wie in Abschnitt 4.2.1.2 beschrieben.

4.2.2.3 Zielgrößen

Für die Zielgrößen der Primärstudien aus der Ergänzungsrecherche galten dieselben Einschlusskriterien wie für die Zielgrößen der Primärstudien in den Evidenzsynthesen (vgl. Abschnitt 4.2.1.3).

4.2.2.4 Studientypen

Für die Primärstudien der Ergänzungsrecherche galten dieselben Einschlusskriterien wie für die in den Evidenzsynthesen enthaltenen Primärstudien (vgl. Abschnitt 4.2.1.4).

4.2.2.5 Sonstige Studiencharakteristika

Es erfolgte keine weitere Einschränkung.

4.2.2.6 Tabellarische Darstellung der Einschlusskriterien („Ergänzungsrecherche“ nach Primärstudien zur diagnostischen und prognostischen Güte)

Es wurden alle Studien einbezogen, die alle nachfolgenden Einschlusskriterien erfüllten.

Tabelle 10: Tabellarische Darstellung der Einschlusskriterien für Primärstudien zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT (Ergänzungsrecherche)

Einschlusskriterien für Primärstudien zur Evaluierung der diagnostischen und prognostischen Güte	
E1c	Patienten wie in Abschnitt 4.1.2.1 beschrieben
E2c	Prüftechnologie: Vollring-PET bzw. PET/CT, Evaluierung wie in Abschnitt 4.2.1.2 beschrieben
E3c	Zielgrößen wie in Abschnitt 4.2.1.3 beschrieben
E4c	Prospektiv geplante Querschnitt- oder Kohortenstudien wie in Abschnitt 4.2.1.4 beschrieben
E5c	Keine Mehrfachpublikationen ohne relevante Zusatzinformation
E6c	Vollpublikation verfügbar ^a
E7c	Durch Literatursuche der eingeschlossenen Evidenzsynthesen nicht abgedeckt
a: Als Vollpublikation galt in diesem Zusammenhang auch die nicht vertrauliche Weitergabe eines Studienberichts an das Institut oder die nicht vertrauliche Bereitstellung eines Berichts über die Studien, der den Kriterien des STARD-Statements [42] oder STROBE-Statements [43] genügte und eine Bewertung der Studie ermöglichte.	

4.2.3 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für das Einschlusskriterium E1 (Population) reichte es aus, wenn bei mindestens 80 % der eingeschlossenen Patienten dieses Kriterium erfüllt war. Lagen für solche Studien entsprechende Subgruppenanalysen vor, wurde auf diese Analysen zurückgegriffen. Studien, bei denen das Einschlusskriterium E1 bei weniger als 80 % erfüllt war, wurden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorlagen.

Ebenfalls eingeschlossen wurden Studien, die zu mindestens 80 % das Einschlusskriterium E2 erfüllten. Dies galt allerdings ausschließlich für die Indextests.

4.2.4 Suche nach zusätzlichen Informationen zu relevanten Studien

Es war geplant, Anfragen an die Autoren der Publikationen zu stellen, falls Informationen, die einen relevanten Einfluss auf die Bewertung erwarten ließen, den vorliegenden Studiendokumenten nicht oder nur ungenau zu entnehmen waren.

4.3 Informationsbeschaffung

4.3.1 Bibliografische Literaturrecherche

Die systematische Suche nach relevanten Publikationen erfolgte in folgenden Quellen:

- Suche nach Primärstudien in den bibliografischen Datenbanken MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials (Clinical Trials)
- Suche nach relevanten Evidenzsynthesen: Identifizierung mittels Suche in den Datenbanken MEDLINE und EMBASE parallel zur Suche nach relevanter Primärliteratur sowie mittels Suche in den Datenbanken Cochrane Database of Systematic Reviews (Cochrane Reviews), Database of Abstracts of Reviews of Effects (Other Reviews) und Health Technology Assessment Database (Technology Assessments)

Sollten relevante HTA-Berichte gefunden werden, welche die Einschlusskriterien für Evidenzsynthesen erfüllten, sollten die jeweiligen HTA-Organisationen zu eventuellen Aktualisierungen befragt werden.

Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Anhang A. Die Suche nach Evidenzsynthesen erfolgte am 17.02.2010, die Suche nach Primärstudien wurde am 16.02.2010 durchgeführt.

4.3.2 Suche nach weiteren publizierten und nicht publizierten Studien und Evidenzsynthesen

Mit dem Ziel, weitere veröffentlichte und unveröffentlichte Studien zu ermitteln, wurden weitere Quellen berücksichtigt. Die Rechercheergebnisse wurden anschließend auf zusätzliche relevante Studien und Studienunterlagen untersucht.

4.3.2.1 Suche in potenziell relevanten Evidenzsynthesen

Potenziell relevante Evidenzsynthesen wurden hinsichtlich weiterer relevanter Publikationen bzw. Studien gesichtet.

4.3.2.2 Suche in öffentlich zugänglichen Studienregistern

Die folgenden öffentlich zugänglichen Studienregister wurden durchsucht:

- U.S. National Institutes of Health. ClinicalTrials.gov [online].
URL: <http://www.clinicaltrials.gov>
- World Health Organization. International Clinical Trials Registry Platform Search Portal [online].
URL: <http://apps.who.int/trialsearch>
- UK Clinical Research Network Study Portfolio [online].
URL: <http://public.ukcrn.org.uk/search/>

Die letzte Suche in öffentlich zugänglichen Studienregistern fand am 09.03.2011 statt.

4.3.2.3 Suche in durch den G-BA übermittelten Unterlagen

Im Rahmen des Stellungnahmeverfahrens, das zu Beginn der Auftragsbearbeitung durch den G-BA durchgeführt wurde, wurden 290 Publikationen benannt. Alle diese Publikationen wurden im Volltext beschafft und separat auf ihre Relevanz für die jeweiligen Berichtsfragestellungen geprüft.

4.3.2.4 Suche in Kongressbänden

Es wurde eine Suche in Abstractbänden bzw. auf den Websites folgender internationaler Kongresse durchgeführt.

- American Society of Clinical Oncology (ASCO)
- European Association of Nuclear Medicine (EANM)
- European Cancer Organisation (ECCO)
- European Society of Medical Oncology (ESMO)

Die Suche fand für die Jahre 2006 bis 2010 statt.

4.3.2.5 Suche in den Datenbanken der Leitlinienanbieter

Um Evidenzsynthesen zu identifizieren, die im Rahmen einer Leitlinienerstellung angefertigt wurden und die möglicherweise nicht durch die oben beschriebene Recherchestrategie identifizierbar waren, wurde eine Recherche nach evidenzbasierten Leitlinien in folgenden Datenbanken durchgeführt:

- Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF)
- Guidelines International Network (G-I-N)
- National Guideline Clearinghouse (NGC)

Die letzte Suche nach Leitlinien fand am 10.03.2010 statt.

4.3.3 Selektion relevanter Studien und Evidenzsynthesen

Die Selektion relevanter Primärstudien und Evidenzsynthesen erfolgte durch 2 Reviewer unabhängig voneinander. Dazu wurde das Ergebnis der Recherche in den oben genannten Quellen herangezogen.

Die Fragestellung 1 sollte, wie in Abschnitt 4.1 beschrieben, auf Basis der Primärliteratur bearbeitet werden.

Die Fragestellung 2 zur diagnostischen und prognostischen Güte wurde primär auf Basis von Evidenzsynthesen beantwortet. Zunächst wurde daher systematisch nach Evidenzsynthesen gesucht und die gefundenen Arbeiten mit dem Qualitätsindex von Oxman und Guyatt

bewertet (vgl. Abschnitt 4.2.1.5). Anschließend wurde für den Zeitraum und für Teilfragestellungen, die durch den Recherchezeitraum der Evidenzsynthesen nicht abgedeckt waren, eine zusätzliche Recherche nach Primärliteratur durchgeführt (Ergänzungsrecherche; vgl. Abschnitt 4.2.2). Sollten für eine Fragestellung nicht ausreichend belastbare Daten aus Evidenzsynthesen vorliegen, so sollte sie ausschließlich auf Basis der Primärliteratur beantwortet werden.

4.3.4 Suche nach zusätzlichen Informationen zu relevanten Studien

Es war geplant, Anfragen an die Autoren der Publikationen zu stellen, falls Informationen, die einen relevanten Einfluss auf die Bewertung erwarten ließen, den vorliegenden Studiendokumenten nicht oder nur ungenau zu entnehmen waren.

4.3.5 Nutzung von Informationen aus der Anhörung

Im Anschluss an die Veröffentlichung des vorläufigen Berichtsplans erfolgte eine Anhörung, die sich unter anderem auch auf in die Nutzenbewertung einzubeziehende Informationen beziehen konnte. Relevante Informationen aus dieser Anhörung konnten in die Nutzenbewertung einfließen.

4.4 Informationsbewertung

Die Dokumentation und Bewertung der Studiencharakteristika und -ergebnisse der relevanten Publikationen erfolgte anhand von standardisierten Datenextraktionsbögen und Evidenztabellen. Ein Reviewer führte die Datenextraktion unter Verwendung des Extraktionsbogens durch. Ein zweiter Reviewer überprüfte die Extraktion. Etwaige Diskrepanzen in der Bewertung wurden durch Diskussion zwischen den Gutachtern, ggf. unter Hinzuziehung einer 3. Person, aufgelöst.

4.4.1 Bewertung von Studien zum patientenrelevanten Nutzen und Schaden

Die Bewertung der Informationen der eingeschlossenen Studien hing stark von den verfügbaren Angaben und der Qualität der jeweiligen Publikationen und weiterer Informationsquellen ab. Alle für die Nutzenbewertung relevanten, mit validierten Messinstrumenten erhobene Ergebnisse sollten hinsichtlich ihrer Ergebnissicherheit, bestehend aus dem Verzerrungspotenzial und der Präzision, überprüft werden.

Bewertung des Verzerrungspotenzials der Ergebnisse

Das Verzerrungspotenzial der Ergebnisse sollte für jede in die Nutzenbewertung eingeschlossene Studie bewertet werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollten insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte, die das Verzerrungspotenzial beeinflussen, systematisch extrahiert und bewertet werden:

A: Aspekte des Verzerrungspotenzials der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (*bei randomisierten Studien*)
- Verdeckung der Gruppenzuteilung (*bei randomisierten Studien*)
- Zeitliche Parallelität der Gruppen (*bei nichtrandomisierten kontrollierten Studien*)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (*bei nicht randomisierten kontrollierten Studien*)
- Verblindung des Patienten sowie des Behandlers
- Ergebnigesteuerte Berichterstattung

B: Aspekte des Verzerrungspotenzials der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- Ergebnigesteuerte Berichterstattung

Das Verzerrungspotenzial sollte als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Für die Bewertung eines Endpunkts sollte zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Bei einer Einstufung als „hoch“ sollte das Verzerrungspotenzial für den Endpunkt i. d. R. auch als „hoch“ bewertet werden. Ansonsten sollten die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ sollte nicht zum Ausschluss aus der Nutzenbewertung führen. Die Klassifizierung diene vielmehr der Diskussion heterogener Studienergebnisse und beeinflusste die Sicherheit der Aussage.

4.4.2 Bewertung von Evidenzsynthesen zur diagnostischen und prognostischen Güte

Es wurde eine Übersicht der in den Evidenzsynthesen jeweils eingeschlossenen Studien mit den wesentlichen Studiencharakteristika in tabellarischer Form erstellt. Die Ergebnisse der zugrunde liegenden Primärstudien wurden dabei nach den Technologien Indextest I, ggf. Indextest II, und Referenztest und den Studientypen aufgeteilt.

Die Bewertung des Verzerrungspotenzials einer Evidenzsynthese erfolgte anhand der Kriterien von Oxman und Guyatt [38,39]. Diese Kriterien wurden um die Angabe von Interessenkonflikten ergänzt (vgl. Anhänge Anhang D.1 und Anhang D.2).

4.4.3 Bewertung von Primärstudien zur diagnostischen und prognostischen Güte

Die Studien wurden zunächst nach Indikationen, Technologievergleichen und Studientypen aufgeteilt und separat dargestellt. Die Angaben zu Studienpopulationen, verglichenen Technologien, methodischen Aspekten und Ergebnissen einzelner Studien wurden tabellarisch dargestellt.

Die Bewertung des Verzerrungspotenzials der Primärstudien zur diagnostischen Güte wurde mithilfe eines adaptierten QUADAS-Instruments durchgeführt [34]. Dieses wurde um ein Item zur Interdependenz von Index- und Referenztest sowie um ein generelles Item zur ergebnisgesteuerten Berichterstattung ergänzt. Außerdem wurde ein Fragenblock zur Übertragbarkeit hinzugefügt (vgl. Anhang D.3 und Anhang D.4).

Für die Bewertung des Verzerrungspotenzials von prognostischen Studien war zusätzlich zu QUADAS die adäquate Berücksichtigung von möglichen Confoundern als ein weiteres Qualitätskriterium vorgesehen [44]. Im Zuge der Berichtsbearbeitung erwies sich dieses Instrument für die Bewertung von Prognosestudien jedoch nicht als geeignet. Auf Grundlage der Arbeiten von Hayden et al. [44] und Altman et al. [45] wurde daher ein neues Instrument entwickelt (vgl. Anhang D.5 und Anhang D.6).

Das Verzerrungspotenzial der Primärstudien zur diagnostischen und prognostischen Güte wurde als „niedrig“ oder „hoch“ eingestuft.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ führte nicht zum Ausschluss aus der Bewertung der diagnostischen und prognostischen Güte. Die Klassifizierung diente vielmehr der Diskussion heterogener Studienergebnisse und beeinflusste die Sicherheit der Aussage.

4.5 Informationssynthese und -analyse

Die Informationen wurden einer Informationssynthese und -analyse unterzogen. Wenn möglich, wurden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die unten beschriebenen Werkzeuge (Meta-Analyse, Sensitivitäts- und Subgruppenanalyse) eingesetzt. Darüber hinaus erfolgte eine abschließende zusammenfassende Bewertung der Informationen. Die Ergebnisse der Evidenzsynthesen wurden entsprechend den Darstellungen in den Publikationen berichtet. Auch hier erfolgte eine Darstellung der Ergebnisse jeder darin enthaltenen Primärstudie. Zusammenfassende Ergebnisse der Evidenzsynthesen sollten ebenfalls berichtet werden.

4.5.1 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse zu den in den Studien berichteten Endpunkten wurden im Bericht vergleichend beschrieben. Weiterhin erfolgte eine Abschätzung der Heterogenität der Ergebnisse innerhalb der jeweiligen Fragestellung.

In bestimmten Fällen sollten einzelne Ergebnisse aus den Studien zu einem Endpunkt nicht dargestellt bzw. nicht in die Nutzenbewertung oder Bewertung der diagnostischen und prognostischen Güte einbezogen werden. Dies sollte insbesondere zutreffen, wenn Daten zu vielen Patienten nicht in der Auswertung berücksichtigt waren. Ergebnisse sollten i. d. R. nicht in die Bewertung einfließen, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patienten basierten, das heißt, wenn der Anteil der fehlenden Werte größer als 30 % war. In der Literatur werden zum Teil bereits Ausfallraten größer als 20 % als nicht mehr aussagekräftig betrachtet [46]. Die Festlegung der Grenze auf 30 % kann daher als vergleichsweise liberal angesehen werden. Das beschriebene Vorgehen wurde auch angewendet, wenn in vergleichenden Studien der Unterschied der fehlenden Werte zwischen den Gruppen größer als 15 Prozentpunkte war.

Ausnahmen von dieser Regel sollten z. B. dann gemacht werden, wenn aus logistischen Gründen für ganze Zentren (ganze Randomisierungsblöcke) keine Daten erhoben wurden, obwohl dies bereits bei der Studienplanung vorgesehen war [47].

Die Ergebnisse sollten auch dann nicht in die Nutzenbewertung oder Bewertung der diagnostischen und prognostischen Güte einbezogen werden, wenn der Unterschied der Nichtberücksichtigungsanteile zwischen den Gruppen größer als 15 Prozentpunkte war.

4.5.1.1 Gegenüberstellung der Ergebnisse von (Primär-)Studien zur Bewertung des patientenrelevanten Nutzens und Schadens

Im Bericht sollten die Studien anhand des Designs (Studiendesign, Studiendauer, Zahl der randomisierten / zugeteilten Patienten, Zuteilungsmechanismus, Ort und Zeitraum der Studiendurchführung, relevante Zielkriterien) beschrieben werden. Darüber hinaus sollten die in der jeweiligen Studie verglichenen diagnostischen Strategien (z. B. eingesetzte Technologien, eventuelle Effekte der Vorselektion der Patienten etc.) dargestellt werden. Die Studienpopulationen sollten weiterhin durch demografische Daten (Alter, Geschlecht), durch Charakteristika der betrachteten Krankheitsentitäten und durch prognostische Angaben beschrieben werden. Zusätzlich sollte die Zahl der Studienabbrecher dargestellt werden.

Die Ergebnisse der jeweiligen Behandlungsgruppen zu den in den Studien berichteten patientenrelevanten Endpunkten sollten im Bericht vergleichend beschrieben werden. Dazu sollten zunächst die in den Studien berichteten jeweiligen Endpunkte den im Berichtsplan definierten patientenrelevanten Zielgrößen (s. Abschnitt 4.1.2.3) zugeordnet werden. Endpunkte aus Studien, die keiner vordefinierten Zielgröße gemäß Abschnitt 4.1.2.3 zugeordnet werden konnten, sollten als nicht patientenrelevant angesehen werden und sollten daher in der Nutzenbewertung nicht berücksichtigt werden.

4.5.1.2 Gegenüberstellung der Ergebnisse von Evidenzsynthesen zur diagnostischen und prognostischen Güte

Bei der Beurteilung der diagnostischen und prognostischen Güte wurden zunächst die zu der Fragestellung vorhandenen Ergebnisse der Evidenzsynthesen abgebildet. Die Ergebnisse aus den zugrunde liegenden Primärstudien wurden dabei – soweit notwendig und möglich – zunächst nach den Fragestellungen und Ordnungskriterien „Indextest I“ (z. B. PET), „Indextest II“ (z. B. PET/CT) und „Referenztest“ klassifiziert und qualitativ zusammengefasst. Zudem erfolgte eine Abschätzung der Heterogenität der Ergebnisse innerhalb dieser Klassen. Beim Vorliegen inkonsistenter Ergebnisse zu einer Fragestellung wurden mögliche Erklärungen für die beobachtete Heterogenität (z. B. Unterschiede bei der Berücksichtigung eines möglichen Verifikations- oder Inkorporationsbias) beschrieben [48].

Sollten durch eine Ergänzungsrecherche für den Zeitraum nach Abschluss des Recherchezeitraums der eingeschlossenen Evidenzsynthesen noch weitere relevante Primärstudien identifiziert werden, so sollten diese separat und analog zu den (Primär-)Studien zur Bewertung des patientenrelevanten Nutzens und Schadens präsentiert werden (s. Abschnitt 4.5.1.3).

4.5.1.3 Gegenüberstellung der Ergebnisse von Primärstudien zur diagnostischen und prognostischen Güte („Ergänzungsrecherche“)

Im Bericht wurden die Studien anhand der zugrunde liegenden Methodik (Studiendesign, Studiendauer, Ort und Zeitraum der Studiendurchführung, relevante Endpunkte) beschrieben. Die Studienpopulationen wurden weiterhin durch demografische Daten (Alter, Geschlecht), durch Charakteristika der betrachteten Krankheitsentitäten und durch prognostische Angaben beschrieben. Zusätzlich sollte die Zahl der Studienabbrecher dargestellt werden.

Bei der Beurteilung der diagnostischen und prognostischen Güte wurden die zu der Fragestellung vorhandenen Ergebnisse der Primärstudien abgebildet. Die Ergebnisse wurden dabei – soweit notwendig und möglich – zunächst nach den Fragestellungen und Ordnungskriterien „Indextest I“, „Indextest II“ und „Referenztest“ klassifiziert und qualitativ zusammengefasst. Bei prognostischen Primärstudien wurden – soweit untersucht – multifaktorielle Analysen und der Einfluss von möglichen Confoundern auf die prognostische Güte präsentiert.

4.5.1.4 Gegenüberstellung der Ergebnisse aus Evidenzsynthesen und der Ergänzungsrecherche („Robustheitsprüfung“)

Fanden sich sowohl relevante Evidenzsynthesen als auch relevante Primärstudien in der Ergänzungsrecherche, wurde geprüft, ob die (quantitativ zusammengefassten) Ergebnisse der aktuelleren Primärstudien die Ergebnisse der Evidenzsynthesen infrage stellen konnten („Robustheitsprüfung“). Nur wenn dies zu erwarten war, wurden zusammenfassende

Analysen über alle Studien – sowohl jene, die in den systematischen Übersichten vorhanden waren, als auch jene aus der Ergänzungsrecherche – in Erwägung gezogen.

4.5.2 Meta-Analysen

4.5.2.1 Meta-Analysen für vergleichende Studien

Sofern die Primärstudien hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbar waren, wurden die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst. Für die statistische Auswertung vergleichender Studien wurden primär die Ergebnisse aus Intention-to-Treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben waren, verwendet.

Die Meta-Analysen erfolgten i. d. R. auf Basis von Modellen mit zufälligen Effekten [49]. In begründeten Ausnahmefällen sollten Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorlagen, sollten diese nach Möglichkeit anhand der vorhandenen Informationen eigenständig berechnet bzw. näherungsweise bestimmt werden.

Bei Studien zum patientenrelevanten Nutzen und Schaden sollte für stetige Variablen die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges g , als Effektmaß eingesetzt werden. Bei binären Variablen sollten Meta-Analysen primär anhand des Odds Ratios und beim Vergleich von Überlebenszeiten anhand des Hazard Ratio durchgeführt werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollten mittels Forest Plots zusammenfassend dargestellt werden. Anschließend sollte die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität erfolgen [50]. In den Fällen, in denen die Heterogenität der Studienergebnisse nicht bedeutsam war ($p \geq 0,2$ für Heterogenitätstest), sollte der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollten die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem sollte untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe 4.5.3) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.5.4).

4.5.2.2 Meta-Analysen für Diagnosestudien

Die Schätzer und Konfidenzintervalle aus den Studien werden ebenfalls mittels Forest Plots zusammenfassend dargestellt. Außerdem wird, sofern die dafür nötigen Anforderungen als erfüllt gelten können, für die Testgütekriterien eine Meta-Analyse anhand der Sensitivität und Spezifität in einer bivariaten Analyse durchgeführt [51,52]. Die Schätzung der Modellparameter (Sensitivität, Spezifität und Parameter für die „hierarchical summary receiver operating curve“, HSROC) erfolgt über einen Maximum-Likelihood-Ansatz in einem nicht linearen gemischten Modell. In begründeten Ausnahmefällen kommen auch andere

Effektmaße zum Einsatz. Zusätzlich werden bei diagnostischen Studien die beobachteten Paare aus Sensitivität und Spezifität zweidimensional grafisch dargestellt. Studien, in denen Technologien wie PET und MRT direkt verglichen wurden, wurden gekennzeichnet. Des Weiteren wurden die aus der bivariaten Meta-Analyse gewonnenen Schätzer für die Erwartungswerte als gepoolte Paare der Sensitivität und der Spezifität mit den dazugehörigen 95 %-Konfidenzregionen grafisch dargestellt. Schließlich wurden die geschätzten HSROC-Kurven abgebildet. Die grafischen Darstellungen von Sensitivität und Spezifität und der HSROC-Kurve können Tendenzen zu einer besseren Testgüte in den vorliegenden Studien aufzeigen.

Schneiden sich die Konfidenzregionen zweier Technologien nicht, so wird auf einen statistisch signifikanten Unterschied der diagnostischen Güte in den vorliegenden Studien geschlossen. Die Richtung des Unterschieds wird ggf. grafisch bestimmt. Überlappen sich 2 Konfidenzregionen oder schneiden sich 2 HSROC-Kurven im Bereich der Grafik, in dem auch Paare aus Sensitivität und Spezifität beobachtet wurden, kann in den vorliegenden Studien keine Aussage zu einem eventuell vorliegenden Unterschied in der diagnostischen Güte getroffen werden.

4.5.3 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der gepoolten Ergebnisse waren Sensitivitätsanalysen hinsichtlich methodischer Faktoren geplant. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ sollte hierfür verwendet werden. Dies sollte sowohl Primärstudien zum patientenrelevanten Nutzen und Schaden als auch jene zur diagnostischen und prognostischen Güte betreffen. Bei Einschluss von mehreren Evidenzsynthesen sollten wenn möglich Sensitivitätsanalysen anhand der Beurteilung mit dem Index von Oxman und Guyatt [38,39] durchgeführt werden (Stufe 5 versus 6 versus 7; alle < 5 wurden ausgeschlossen). Weitere methodische Faktoren konnten aus den im Rahmen der Informationsbewertung aufgefallenen Merkmalen abgeleitet werden, z. B. die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes.

Das Ergebnis der Sensitivitätsanalysen kann die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestufteffekt kann z. B. dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen Nutzen oder Schaden attestiert wird.

4.5.4 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse wurden hinsichtlich potenzieller Effektmodifikatoren, d. h. klinischer Faktoren, die die Effekte beeinflussen, untersucht. Dies konnten direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen sein. Im Gegensatz zu den in Abschnitt 4.5.3 beschriebenen methodischen Faktoren für Sensitivitätsanalysen bestand hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und

Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte war die auf einem Homogenitäts- bzw. Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung von Effektmodifikatoren wurden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalteten, und aus Subgruppenanalysen einbezogen. Außerdem sollten eigene Analysen in Form von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren erfolgen. Es war vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Geschlecht
- Tumorentitäten
- Vorangegangene Therapie: Die diagnostische Güte und die Präzision der PET oder PET/CT nach einer Strahlentherapie sollen mit der diagnostischen Güte nach einer Chemotherapie verglichen werden, da es z. B. denkbar ist, dass eine vorherige Strahlentherapie durch eine stärkere Entzündungsreaktion vermehrt zu falsch positiven Befunden führt etc.
- Zeitpunkt der PET-Diagnostik nach Abschluss der Primärtherapie

Im Falle, dass in den verfügbaren Informationen Anhaltspunkte für weitere mögliche Effektmodifikatoren vorlagen, konnten diese ebenfalls begründet einbezogen werden.

Bei Identifizierung möglicher Effektmodifikatoren sollte ggf. eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen erfolgen. Beispielsweise hätte ggf. der Beleg eines Zusatznutzens auf eine spezielle Subgruppe von Patienten eingeschränkt werden können.

4.6 Änderungen der Methodik im Vergleich zum Berichtsplan

In Tabelle 10 (Darstellung der Einschlusskriterien für die Primärliteratur zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT) wurde der Einschlussgrund E7c „Durch Literatursuche der eingeschlossenen Evidenzsynthesen nicht abgedeckt“ hinzugefügt.

Die Einschlusskriterien für vergleichende Studien zur Bewertung des Nutzens und des potenziellen Schadens wurden vereinfacht (vgl. Abschnitt 4.1.2.3). Allerdings sollten Studien, die alleine die Managementänderungen (ohne gleichzeitige Dokumentation der patientenrelevanten Endpunkte) berichteten, nicht eingeschlossen werden. Da sich in diesem Bericht keine vergleichende Studie fand, hat diese Änderung keine Auswirkung auf die Berichtsergebnisse und damit auf das Fazit.

Nach Erstellung des Berichtsplans wurde in Anlehnung an andere Berichte des Instituts der Abschnitt 4.2.3 „Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen“ ergänzt und prospektiv für alle weiteren PET-Berichte festgelegt. Die Grenze von

80 % ist willkürlich festgelegt, erscheint jedoch vergleichsweise liberal, da bei einer noch größeren Durchmischung der Populationen bzw. Interventionen eine externe Übertragbarkeit der Ergebnisse nicht mehr gewährleistet wäre.

Im Absatz 4.5.2 wurde das Vorgehen der bivariaten Meta-Analyse erläutert, das für die Testgütekriterien anhand der Sensitivität und Spezifität durchgeführt wurde.

5 Ergebnisse

5.1 Ergebnisse der Informationsbeschaffung

5.1.1 Ergebnis der bibliografischen Literaturrecherche

Abbildung 1 zeigt das Ergebnis der systematischen Literaturrecherche nach Primärstudien und Evidenzsynthesen in den bibliografischen Datenbanken und das Literaturscreening gemäß Ein- / Ausschlusskriterien. In dieser Abbildung ist sowohl die Recherche nach Evidenzsynthesen als auch jene nach Primärliteratur dargestellt.

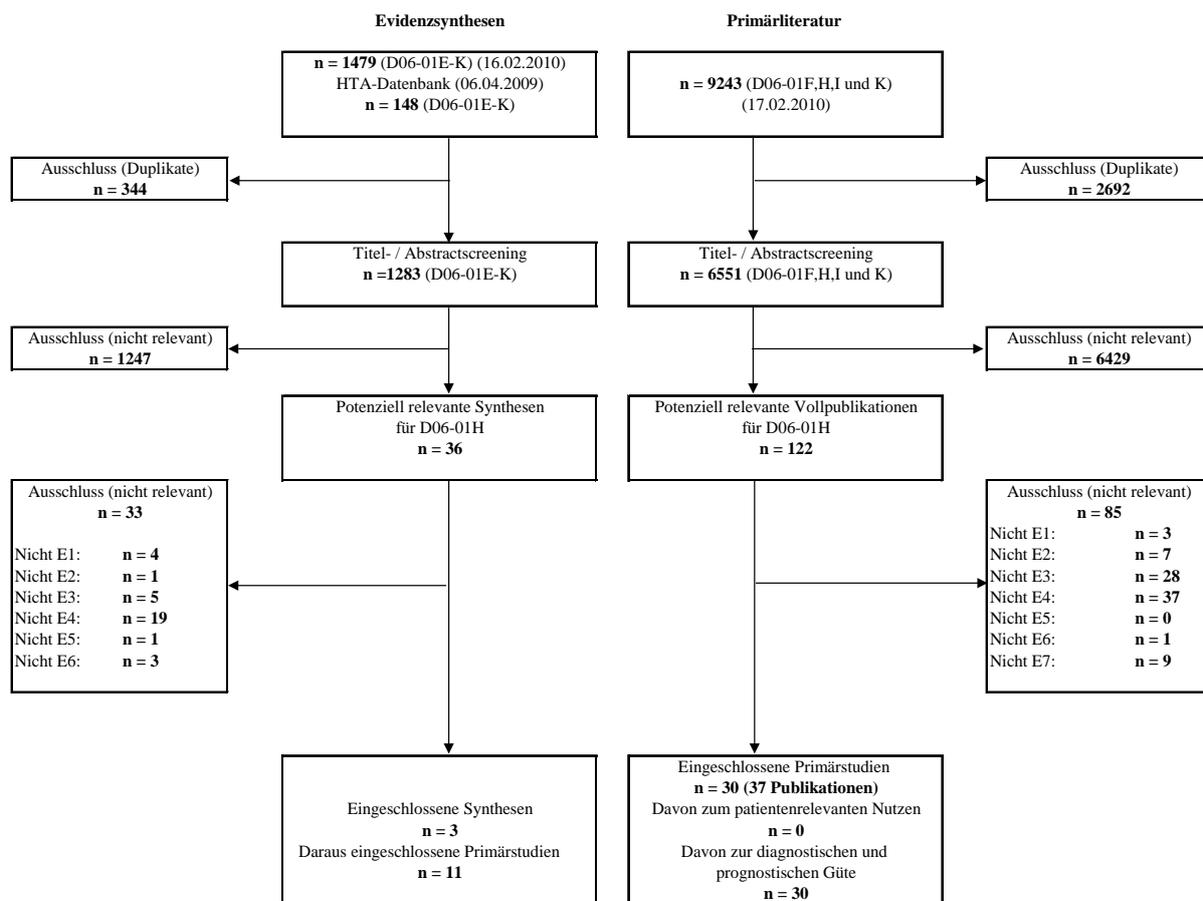


Abbildung 1: Ergebnisse der bibliografischen Literaturrecherche und des Literaturscreening

Suche nach Studien zur Nutzenbewertung

Die Primärliteraturrecherche wurde für 4 Projekte (D06-01F, H, I und K) gleichzeitig durchgeführt, da diese Projekte von einer Sachverständigengruppe gemeinsam bearbeitet wurden. In der Primärrcherche wurden für die Krankheitsbilder malignes Melanom (D06-01F), Ösophaguskarzinom (D06-01H), Ovarialkarzinom (D06-01I) und Knochen- und Weichteiltumoren (D06-01K) insgesamt 9243 Referenzen identifiziert.

Nach Ausschluss von 2692 Duplikaten ergab sich eine Gesamtzahl von 6551 zu screenenden Treffern.

6429 Treffer wurden von 2 Reviewern nach Konsentierung zunächst diskrepanter Einschätzungen übereinstimmend im Rahmen des Titel- und Abstractscreenings als nicht relevant ausgeschlossen. Ein häufiger Grund für den Ausschluss war, dass sich aufgrund der oben beschriebenen Recherche viele Arbeiten nicht mit dem Krankheitsbild der Ösophaguskarzinome befassten. Ebenfalls ausgeschlossen wurden Referenzen, die thematisch oder methodisch nicht den Fragestellungen des Berichts entsprachen (z. B. andere diagnostische Verfahren oder gesundheitsökonomische Analysen). Aus der bibliografischen Literaturrecherche verblieben damit 122 potenziell relevante Treffer, die im Volltext gesichtet wurden.

Zur Fragestellung des Nutzens (Fragestellung 1) konnte keine vergleichende Primärstudie identifiziert werden.

Suche nach Evidenzsynthesen zur diagnostischen und prognostischen Güte

Durch die Suche nach Evidenzsynthesen wurden für die Projekte D06-01E-K insgesamt 1479 Treffer identifiziert. Eine indikationsunabhängige Suche nach PET in der Health Technology Assessment Database (Technology Assessment) identifizierte weitere 148 Treffer. Nach Ausschluss von 344 Duplikaten ergab sich eine Gesamtzahl von 1283 zu screenenden Treffern.

1247 Treffer wurden von beiden Reviewern nach Konsentierung zunächst diskrepanter Einschätzungen übereinstimmend im Rahmen des Titel- und Abstractscreenings als nicht relevant ausgeschlossen. Aus der bibliografischen Literaturrecherche verblieben damit 36 potenziell relevante Treffer, die im Volltext gesichtet wurden. Hiervon wurden 33 aufgrund fehlender Relevanz ausgeschlossen. Die Zitate der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang B.

Für die Indikation Restaging konnten 3 Evidenzsynthesen in den Bericht eingeschlossen werden. Da keine HTA-Berichte als Evidenzsynthesen in den Bericht eingeschlossen wurden, wurden keine Anfragen an Agenturen von HTA-Berichten gestellt.

Ergänzungsrecherche zur diagnostischen und prognostischen Güte

Für die Indikation Restaging verfügte die Arbeit von Ngamruengphong 2010 [53] über die aktuellste Recherche der eingeschlossenen Evidenzsynthesen (bis Februar 2008). Im Rahmen der Ergänzungsrecherche wurden daher nur Arbeiten eingeschlossen, die mit einer Überlappung von 6 Monaten, also ab Oktober 2007, publiziert wurden. Mit der Überlappung von 6 Monaten wurde sichergestellt, dass auch Studien im vorliegenden Bericht berücksichtigt wurden, die zwar zwischen Oktober 2007 und März 2008 publiziert wurden,

jedoch noch nicht in den Literaturdatenbanken gelistet waren, als die Autoren der Evidenzsynthesen ihre Recherchen durchführten.

Die Ergänzungsrecherche zur diagnostischen und prognostischen Güte (Fragestellung 2) für das Restaging im Krankheitsbild der Ösophaguskarzinome wurde zusammen mit der Suche nach Studien zur Nutzenbewertung (Fragestellung 1) für die Indikationen Primärstaging und Rezidivdiagnostik durchgeführt.

Aus der bibliografischen Literaturrecherche verblieben 122 potenziell relevante Treffer, die im Volltext gesichtet wurden. Hiervon wurden 85 aufgrund fehlender Relevanz ausgeschlossen. Die Zitate der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang C.

Alle Vollpublikationen, die sich mit dem Restaging von Ösophaguskarzinomen beschäftigten, aber vor Oktober 2007 in den Literaturdatenbanken gelistet waren, wurden mit der Begründung, dass das Einschlusskriterium E7c („Publikationszeitraum ist durch die Literatursuche der eingeschlossenen Evidenzsynthesen nicht abgedeckt“) verletzt war, ausgeschlossen. Die 85 im Volltext gescreenten, aber im Rahmen der Ergänzungsrecherche ausgeschlossenen Referenzen sind mit den jeweiligen Ausschlussgründen in Anhang C gelistet.

Für die Indikationen Primärstaging und Rezidivdiagnostik wurden keine Evidenzsynthesen, die den Einschlusskriterien und Qualitätsansprüchen des vorliegenden Berichts entsprachen, eingeschlossen. Die Bewertung der diagnostischen und prognostischen Güte dieser beiden Indikationen basierte ausschließlich auf Primärstudien, die im Rahmen der Ergänzungsrecherche identifiziert wurden.

Die Ergänzungsrecherche lieferte 30 Primärstudien (37 Publikationen) zur diagnostischen und prognostischen Güte für das Primärstaging, das Restaging und die Rezidivdiagnostik.

5.1.2 Weitere publizierte und nicht publizierte Studien

5.1.2.1 Evidenzsynthesen

Im Rahmen der Ergänzungsrecherche wurde eine Suche in der Literaturliste der ausgeschlossenen Evidenzsynthesen durchgeführt. Hier fand sich jedoch keine zusätzliche potenziell relevante Primärstudie, die nicht im Rahmen der Ergänzungsrecherche ohnehin identifiziert worden wäre.

5.1.2.2 Studienregister

Die Suche in den Studienregistern ergab keine Hinweise auf laufende oder abgeschlossene Studien, die für die Fragestellung des vorliegenden Berichts relevant wären.

5.1.2.3 Suche in den Datenbanken der Leitlinienanbieter

Eine Recherche in den Datenbanken unterschiedlicher Leitlinienanbieter (AWMF, G-I-N und NGC) ergab keine Leitlinie, die auf Basis einer Evidenzsynthese der vorhandenen klinischen Studien erstellt wurde bzw. deren Methodik und Ergebnisse entsprechend den Einschlusskriterien des vorliegenden Berichts dokumentiert worden wären.

5.1.2.4 Unterlagen des G-BA

Im Rahmen des Stellungnahmeverfahrens, das zu Beginn der Auftragsbearbeitung durch den G-BA durchgeführt wurde, wurden 290 Publikationen benannt. Alle diese Publikationen wurden im Volltext beschafft und separat auf ihre Relevanz für die jeweiligen Berichtsfragestellungen geprüft. Von diesen befasste sich eine Studie, Brücher 2001 [54], mit dem Krankheitsbild der Ösophaguskarzinome. Diese Studie wurde in den vorliegenden Bericht für die Indikation Restaging eingeschlossen.

5.1.2.5 Kongressbände

Für alle Indikationen des Berichts erfolgte eine Suche in den Kongressbänden der folgenden Kongresse: American Society of Clinical Oncology (ASCO), European Association of Nuclear Medicine (EANM), European Cancer Organisation (ECCO) und European Society of Medical Oncology (ESMO). Der Zeitraum der Suche beschränkte sich auf die Jahre 2006 bis 2010. Es wurde ausschließlich nach RCTs gesucht. Dabei wurden folgende Suchbegriffe ggf. mit entsprechenden Trunkierungen und Boole'schen Verknüpfungen verwendet: ("random" oder "RCT"), ("PET" oder "positron"). Es wurden keine entsprechenden Studien für das Ösophaguskarzinom durch die Suche identifiziert, die genauen Ergebnisse sind in Tabelle 11 dargestellt.

Tabelle 11: Suche in Kongressbänden

Kongress	Überprüfter Zeitraum	Anzahl der Treffer nach Eingabe der Suchbegriffe in die digitale Suchmaske	Anzahl der Treffer, die einen relevanten Hinweis auf eine potenziell relevante Studie geben, nach Sichtung des Titels und / oder des gesamten Abstracts	Anzahl der Vollpublikationen, die aufgrund des Abstracts gefunden wurden
ECCO	2007 ^a	1573	0	0
	2009 ^{a,b}	Ergebnisse bei ESMO		
ASCO	2006-2010	743	0	0
ESMO	2006-2010	436	0	0
EANM	2006-2010	117	0	0
a: Kongress findet nur in ungeraden Jahren statt; b: 2009 gemeinsame Konferenz mit ESMO. Die Treffer der ECCO-Konferenz 2009 sind in den Suchergebnissen für ESMO enthalten.				

5.1.2.6 Informationen aus der Anhörung

Aus der Anhörung zum Berichtsplan ergaben sich keine Hinweise auf zusätzliche relevante Studien.

5.1.2.7 Zusätzliche Informationen zu relevanten Studien

Für die vorliegende Bewertung wurden für 5 Studien Autorenanfragen versendet, um konkrete Informationen in Bezug auf unklare Ergebnisdarstellung oder fehlende Daten (Choi 2000 und 2004[55,56], Duong 2006a [57], Duong 2006b [58], Omluo 2008 [59], Ott 2006 [60] und Shenfine 2009 [61]) zu erhalten. Zwei Autorenanfragen wurden beantwortet und enthielten ausreichende Informationen.

5.1.3 Resultierender Studienpool

5.1.3.1 Studien zur Bewertung des patientenrelevanten Nutzens

Die systematische Literaturrecherche ergab keine prospektive, vergleichende Studie für die Bewertung des patientenrelevanten Nutzens der PET bzw. PET/CT im Krankheitsbild der Ösophaguskarzinome, wie sie in den Einschlusskriterien des vorliegenden Berichts gefordert wird.

5.1.3.2 Evidenzsynthesen zur diagnostischen und prognostischen Güte

Zur diagnostischen und prognostischen Güte der PET bzw. PET/CT fanden sich 3 Evidenzsynthesen (Ngamruengphong 2010, Rebollo Aquirre 2009 und Westerterp 2005) in der Indikation Restaging, die den Einschlusskriterien des vorliegenden Berichts entsprachen und im Detail ausgewertet wurden (vgl. Tabelle 12) [53,62,63]. Aus diesen Arbeiten konnten 11 Primärstudien von insgesamt 21 Studien, die die PET bzw. PET/CT untersuchten, in den vorliegenden Bericht eingeschlossen werden (vgl. Tabelle 13).

Für die Fragestellungen des Primärstaging entsprach zunächst auch die Evidenzsynthese von van Westreenen 2004 [64] den Einschlusskriterien und erreichte 5 von 7 Punkten mithilfe des Index von Oxman und Guyatt. Nach eingehender Analyse musste diese allerdings von der Bewertung ausgeschlossen werden, da keine der eingeschlossenen Primärstudien den Einschlusskriterien des vorliegenden Berichts entsprach. Für die Rezidivdiagnostik fanden sich keine Evidenzsynthesen, die den Einschlusskriterien des vorliegenden Berichts entsprachen. Evidenzsynthesen, die aufgrund der Bewertung nach Oxman und Guyatt ausgeschlossen wurden, sind im Anhang D.2 tabellarisch aufgeführt.

Tabelle 12: Studienpool zur Beurteilung der diagnostischen und prognostischen Güte der PET bzw. PET/CT zum Restaging

Evidenzsynthesen	Zugeordnete Publikation	Ref.
Ngamruengphong 2010	Ngamruengphong S et al. Dis Esophagus 2010; 23(3): 216-231	[53]
Rebollo Aquirre 2009	Rebollo Aquirre AC et al. Ann Surg 2009; 250 (2): 247-254.	[62]
Westerterp 2005	Westerterp M et al. Radiology 2005; 236(3): 841-851.	[63]

Ein Studienspiegel der 3 eingeschlossenen Evidenzsynthesen wird in Tabelle 13 dargestellt. Aufgrund unterschiedlicher Recherchezeiträume und Quellen für die Recherchen ergaben sich nur teilweise Überlappungen bezüglich einzelner Studien in den eingeschlossenen Evidenzsynthesen. Brücher 2001, Flamen 2002 und Kroep 2003 z. B. waren in Ngamruengphong 2010, aber nicht in Westerterp 2005 enthalten. Da alle 3 Evidenzsynthesen in den gleichen Datenbanken recherchierten, konnte nicht eindeutig festgestellt werden, wie es zu diesen unterschiedlichen Einschlüssen kam.

Tabelle 13: Überblick über die in den Evidenzsynthesen eingeschlossenen Primärstudien (n = 11) zur diagnostischen Güte beim Restaging (alphabetisch geordnet)

Publikation	Jahr	Referenz	Einschluss in Ngamruengphong 2010	Einschluss in Rebollo Aquirre 2009	Einschluss in Westerterp 2005
Brücher	2001	[54]	X	-	X
Cerfolio	2005	[65]	X	X	-
Flamen	2002	[66]	X	-	X
Gillham	2006	[67]	X	-	-
Kroep	2003	[68]	X	-	X
Levine	2006	[69]	X	X	-
Ott	2006	[70]	X	-	-
Song	2005	[71]	X	X	-
Weber	2001	[72]	-	-	X
Westerterp	2006	[73]	X	X	-
Wieder	2004	[74]	X	X	-

5.1.3.3 Primärstudien zur diagnostischen und prognostischen Güte (Ergänzungsrecherche)

Zur diagnostischen und prognostischen Güte der PET bzw. PET/CT fanden sich im Rahmen der Ergänzungsrecherche 30 Primärstudien (37 Publikationen).

Für das Primärstaging konnten 23 Studien (30 Publikationen) eingeschlossen werden. Von diesen 23 Studien untersuchten 18 ausschließlich die diagnostische Güte der PET bzw. PET/CT. Eine Studie, Shenfine 2009 [61], untersuchte sowohl die diagnostische als auch prognostische Güte und 4 Studien, Cheze-Le Rest 2008 [75], Choi 2004 [55], Chung 2008

[76] und Omloo 2008 [59], untersuchten alleine die prognostische Güte der PET bzw. PET/CT beim Primärstaging.

In der Indikation Restaging konnten im Rahmen der Ergänzungsrecherche zusätzlich zu den durch die Evidenzsynthese identifizierten 3 weitere Studien eingeschlossen werden. Eine dieser Studien (Kim 2007) [77] untersuchte dabei ausschließlich die prognostische Güte der PET bzw. PET/CT.

Vier Studien wurden für die Rezidivdiagnostik eingeschlossen. Eine dieser Studie (Guo 2007 [78]) untersuchte die diagnostische und die prognostische Güte. Die 3 weiteren Studien untersuchten die diagnostische Güte der PET bzw. PET/CT.

Tabelle 14: Studienpool zur diagnostischen und prognostischen Güte der eingeschlossenen Primärstudien (Ergänzungsrecherche)

Studie	Zugeordnete Publikation(en)	Referenz
Primärstaging (diagnostische Kohortenstudien)		
Buchmann 2006	Buchmann et al. Nucl Med (Stuttg) 2006; 45(6): 235-241.	[79]
Choi 2000 ^a	Choi et al. J Nucl Med 2000; 41(5): 808-815.	[56]
Choi 2010	Choi et al. Surg Endosc 2010; 24(6): 1380-1386.	[80]
Flamen 2000	Flamen et al. J Clin Oncol 2000; 18(18): 3202-3210.	[81]
	Lerut et al. Ann Surg 2000; 232(6): 743-752.	[82]
Heeren 2004	Heeren et al. J Nucl Med 2004; 45(6): 980-987.	[83]
Itoh 2008	Itoh et al. Japanese Journal of Clinical Radiology 2008; 53(4): 501-507.	[84]
Jager 2001	Jager et al. Eur J Nucl Med 2001; 28(12): 1845-1849.	[85]
Kato 2005	Kato et al. Anticancer Res 2005; 25(6C): 4439-4444.	[86]
Kato 2008	Kato et al. Oncol Rep 2008; 20(4): 857-862.	[87]
Kneist 2004	Kneist et al. Arch Surg 2004; 139(10): 1043-1049.	[88]
	Kneist et al. World J Surg 2003; 27(10): 1105-1112.	[89]
	Kneist et al. Viszeralchirurgie 2003; 38(1): 32-38.	[90]
	Junginger et al. Dtsch Med Wochenschr 2002; 127(38): 1935-1941.	[91]
	Kneist et al. Chirurg 2003; 74(10): 922-930.	[92]
Kole 1998	Kole et al. Br J Cancer 1998; 78(4): 521-527.	[93]
Liberale 2004	Liberale et al. Eur J Surg Oncol 2004; 30(9): 942-947.	[94]
Lowe 2005	Lowe et al. Molecular Imaging & Biology 2005; 7(6): 422-430.	[95]
Noble 2009	Noble et al. Clin Radiol 2009; 64(7): 699-705.	[96]
Rankin 1998	Rankin et al. Clin Radiol 1998; 53(9): 659-665.	[97]
Schreurs 2008	Schreurs et al. Eur J Radiol 2008; 67(1): 105-111.	[98]
Shenfine 2009	Shenfine et al. Dis Esophagus 2009; 22(8): 668-675.	[61]
Sihvo 2004	Sihvo et al. J Gastrointest Surg 2004; 8(8): 988-996.	[99]
	Rasanen et al. Ann Surg Oncol 2003; 10(8): 954-960.	[100]
Yoon 2003	Yoon et al. Radiology 2003; 227(3): 764-770.	[101]
Primärstaging (prognostische Kohortenstudien)		
Cheze-Le Rest 2008	Cheze-Le Rest et al. Nucl Med Commun 2008; 29(7): 628-635.	[75]
	Teyton et al. Medecine Nucleaire 2008; 32(6): 323-331.	[102]
Choi 2004 ^a	Choi et al. J Nucl Med 2004; 45(11): 1843-1850.	[55]
Chung 2008	Chung et al. World J Surg 2008; 32(1): 69-75.	[76]
Omloo 2008	Omloo et al. Endoscopy 2008; 40(6): 464-471.	[59]
Shenfine 2009	Shenfine et al. Dis Esophagus 2009; 22(8): 668-675.	[61]

(Fortsetzung)

Tabelle 14: Studienpool zur diagnostischen und prognostischen Güte der eingeschlossenen Primärstudien (Ergänzungsrecherche) (Fortsetzung)

Restaging (diagnostische Kohortenstudien)		
Higuchi 2008	Higuchi et al. J Thorac Cardiovasc Surg 2008; 136(1): 205-212.	[103]
Roedl 2009a	Roedl et al. Mol Imaging Biol 2009; 11(1): 54-60.	[104]
Restaging (prognostische Kohortenstudien)		
Kim 2007	Kim et al. Eur J Cancer 2007; 43(9): 1385-1391.	[77]
Higuchi 2008	Higuchi et al. J Thorac Cardiovasc Surg 2008; 136(1): 205-212.	[103]
Rezidivdiagnostik (diagnostische Kohortenstudien)		
Guo 2007	Guo et al. J Nucl Med 2007; 48(8): 1251-1258.	[78]
Roedl 2008a	Roedl et al. Ann Thorac Surg 2008; 86(4): 1131-1138.	[105]
Teyton 2009	Teyton et al. J Gastrointest Surg 2009; 13(3): 451-458.	[106]
Zhang 2006	Zhang et al. Zhonghua Wei Chang Wai Ke Za Zhi 2006; 9(1): 56-58.	[107]
Rezidivdiagnostik (prognostische Kohortenstudie)		
Guo 2007	Guo et al. J Nucl Med 2007; 48(8): 1251-1258.	[78]
a: Aus den Studien wird nicht eindeutig ersichtlich, ob nicht auch Patienten aus Choi 2000 in Choi 2004 enthalten sind. Diesbezüglich wurde eine Autorenanfrage gestellt, die bis zur Veröffentlichung nicht beantwortet wurde. Weil die Studien unterschiedliche Fragestellungen untersuchten, wurden sie jeweils als einzelne Studien eingeschlossen.		

5.2 Teilziel 1: patientenrelevanter Nutzen

Es konnte keine vergleichende Studie identifiziert werden, aus der Ergebnisse zum patientenrelevanten Nutzen der PET oder PET/CT zur Diagnostik von Ösophaguskarzinomen abgeleitet werden konnten. Auch die Suche in den Studienregistern und Kongressbänden lieferte keine Hinweise auf aktuell laufende, kontrollierte Studien in diesem Krankheitsbild.

5.3 Teilziel 2a: diagnostische und prognostische Güte beim Primärstaging

5.3.1 Diagnostische und prognostische Güte auf Grundlage in die Bewertung eingeflossener Evidenzsynthesen beim Primärstaging

Für diese Fragestellung konnte keine Evidenzsynthese gefunden werden, die den Einschlusskriterien des vorliegenden Berichts entsprach. Die Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT wurde auf Basis der Primärstudien durchgeführt.

5.3.2 Diagnostische Güte der eingeschlossenen Primärstudien beim Primärstaging

In der Recherche wurden 19 Primärstudien zur diagnostischen Güte der PET und PET/CT beim Primärstaging identifiziert (vgl. Tabelle 14). In den 19 für das Primärstaging eingeschlossenen Diagnosestudien wurden insgesamt 1336 Patienten untersucht. In 15 Studien wurden Angaben zum Geschlecht der eingeschlossenen Patienten gemacht, die in der Summe 957 Männer und 164 Frauen ergeben. In 13 von 19 Studien wurden Patienten mit einem Plattenepithelkarzinom oder einem Adenokarzinom der Speiseröhre untersucht. In 4 Studien wurden ausschließlich Patienten mit einem Plattenepithelkarzinom [56,79,87,101] und in 2 Studien nur Patienten mit einem Adenokarzinom der Speiseröhre eingeschlossen [61,99]. In einer Studie fehlten die Angaben bzgl. der Primärdiagnose [97]. 58 % der eingeschlossenen Studien wurden in Europa (2 Studien in Deutschland), 32 % im asiatischen Raum und 11 % der Studien in den USA oder Australien durchgeführt. Der Einsatz der nichtintegrierten FDG-PET wurde in 17 von 19 Studien, die integrierte FDG-PET/CT in 3 Studien und der Einsatz beider Technologien in einer Studie untersucht. Als Tracer wurde in allen Studien Fluorodeoxyglukose (FDG) benutzt. In der Studie von Jager 2001 wurde neben FDG auch der Tracer ^{11}C -Cholin eingesetzt [85]. Beim Vergleich der Ziele der Studien fiel auf, dass sie unterschiedliche Fragestellungen untersuchten. Neun der eingeschlossenen Studien analysierten ausschließlich das N-Staging, 2 Studien (Kato 2005 und Noble 2005 [86,96]) das M-Staging und 8 Studien beide Bereiche. Als Referenztest wurde Histologie ggf. in Verbindung mit einer Nachbeobachtung eingesetzt.

Im Folgenden werden besondere Auffälligkeiten bezüglich der Fragestellungen und wesentlichen methodischen Charakteristika, die in einigen Studien auftraten, erläutert. Alle weiteren ausführlichen Daten sind den Tabellen der folgenden Abschnitte zu entnehmen.

In der Studie von **Buchmann 2006** [79] wurde bei 17 Patienten die Pathologie nach OP in Kombination mit den Ergebnissen der CT als Referenztest angewendet. Bei 3 Patienten, die nicht operabel waren oder eine neoadjuvante Radiochemotherapie erhielten, wurde nicht die PET zum N- und M-Staging eingesetzt, sondern die CT. Diese 3 Patienten konnten auf Basis der Tabelle 2 der Publikation identifiziert werden und somit aus der Analyse der Vierfeldertafel des vorliegenden Berichts herausgenommen werden.

In der Studie von **Flamen 2000** [81] wurden 74 Patienten mit neu nachgewiesenem Ösophaguskarzinom eingeschlossen. Die Studie befasste sich mit dem M-Staging sowie auch mit dem N-Staging. Die Ergebnisse der Studie wurden in 2 Arbeiten berichtet [81,82], wobei sich die Publikation von Lerut 2000 [82] ausschließlich mit der präoperativen Erkennung von Lymphknotenmetastasen (N-Staging) bei Patienten mit Tumoren der Speiseröhre oder des gastroösophagealen Übergangs beschäftigt. In der Studie wurde die diagnostische Güte der FDG-PET im Vergleich zu den Technologien CT und EUS untersucht. Von den eingeschlossenen 74 Patienten konnten bei 34 Patienten Fernmetastasen festgestellt werden. Die in der Arbeit von Flamen 2000 dargestellten Ergebnisse zum M-Staging konnten allerdings nicht verwendet werden, da folgende Referenztests für die Bestätigung von Organmetastasen und / oder Fernmetastasen von den Autoren angegeben wurden: Histologie, radiologische Verfahren oder Nachbeobachtung. Nach Definition des Vorberichts entsprechen radiologische Verfahren nicht den Anforderungen für adäquate Referenztests. Die Ergebnisse der diagnostischen Güte der FDG-PET, CT und EUS bezüglich der Indikation N-Staging wurden mithilfe der Histologie als Referenzstandard abgeklärt und konnten dementsprechend verwendet werden. Das N-Staging wurde bei 39 Patienten durchgeführt. Die Daten der Vierfeldertafel konnten mithilfe der Tabelle 1 der Publikation von Lerut 2000 rekonstruiert und im vorliegenden Bericht dargestellt werden.

Zu **Itoh 2008** [84] ist zu erwähnen, dass die Ganzkörper CT und die Ganzkörper PET separat durchgeführt wurden. Die Bilder wurden am Ende mithilfe einer Software fusioniert.

Die Studie von **Jager 2001** [85] untersuchte als einzige Studie neben dem Tracer FDG auch den Einsatz des Tracers ¹¹C-Cholin. Bei allen eingeschlossenen Patienten wurde der regionale Lymphknotenstatus histologisch gesichert (N-Staging). Bei einigen Patienten (die Anzahl ist unklar) wurde die klinische Nachbeobachtung als Referenztest zur Bestätigung von Fernmetastasen (M-Staging) verwendet.

Für die Patienten in der Studie von **Kato 2008** [87], die zwischen 1999 und 2005 rekrutiert wurden (n = 117), lagen getrennte Ergebnisse für PET und CT vor. Die Patienten, die 2006 oder 2007 rekrutiert wurden (n = 50), wurden mittels PET/CT untersucht. Die Studie berichtete die Ergebnisse für PET, CT und PET/CT, wobei die Ergebnisse der PET/CT im vorliegenden Bericht nicht verwendet wurden, weil der Anteil fehlender Werte (Patientenzahl) größer als 30 % war.

Die Ergebnisse der Arbeit von **Kneist 2004** [88] wurden in 5 verschiedenen Publikationen berichtet [88-92]. In diesen Arbeiten wurden verschiedene Indikationen betrachtet, sodass die Zahl der jeweils berichteten Patienten von der jeweiligen Fragestellung abhängt. Der Studie konnten Angaben zur diagnostischen Güte von CT und FDG-PET zum Primärstaging von Lymphknoten- und Fernmetastasen nach Ösophaguskarzinom entnommen werden.

Bei **Liberale 2004** [94] wurden 58 Patienten mit nachgewiesenen Karzinomen der Speiseröhre und des gastroösophagealen Übergangs eingeschlossen. Ziel der Arbeit war die

Bewertung der diagnostischen Güte von FDG-PET im präoperativen Staging sowie des Einflusses auf Behandlungsentscheidungen. Für den vorliegenden Bericht konnten keine relevanten Daten zu Änderungen der Behandlung nach FDG-PET eingeschlossen werden. Zwar geben die Autoren dieses als Ziel der Arbeit an, berichten aber keine Daten dazu. Ergebnisse für die Erkennung bösartiger Lymphknoten wurden sowohl für PET als auch für eine Kombination aus CT und EUS berichtet, wobei positive Befunde in einem der Tests als positiver Indextest gewertet wurden. Als Referenztest wurde die Histologie verwendet. Dabei wurden nur Patienten eingeschlossen, die keine neoadjuvante Therapie erhielten ($n = 27$). Zudem wurden 3 Patienten mit Stenosen der Speiseröhre, in denen ein EUS nur eingeschränkt möglich war, ausgeschlossen, sodass insgesamt die Ergebnisse von 24 Patienten vorlagen.

Die Arbeit von **Shenfine 2009** [61] betrachtete sowohl eine diagnostische als auch eine prognostische Fragestellung, sodass im vorliegenden Bericht beide Ergebnisse wiedergegeben werden, auch wenn sie sich auf dieselben Patienten beziehen (siehe hierzu Abschnitt 5.3.3). Als diagnostische Fragestellung wird untersucht, ob die präoperative FDG-PET ein fortgeschrittenes Krankheitsstadium (definiert als $T \geq 3$ oder N1) feststellen kann.

Zur Studie von **Shivo 2004** [99] gibt es eine weitere Arbeit von Rasanen 2003, in der Zwischenergebnisse präsentiert wurden [100]. Shivo 2004 wurde als Hauptpublikation der Studie angesehen. Die im vorliegenden Bericht dargestellten Daten stammen aus Shivo 2004. Ziel der Arbeit war die postoperative Bewertung der diagnostischen Güte von EUS, CT und FDG-PET. Für alle Patienten lagen histologische Ergebnisse vor, allerdings wurden 12 Patienten (10 in der früheren Publikation) von der Bewertung der Lymphknotenmetastasen ausgeschlossen, da bei ihnen keine Lymphadenektomie durchgeführt wurde. In einem weiteren Fall wurde ein Befund der Bildgebung durch eine klinische Nachbeobachtung über 4 Jahre kontrolliert. Es wurde kein Grund angegeben, warum bei dieser Person keine Histologie gemacht wurde.

In der Studie von **Yoon 2003** [101] wurden konsekutiv 136 Patienten mit nachgewiesenen Plattenepithelkarzinomen der Speiseröhre und des gastroösophagealen Übergangs eingeschlossen. 55 Patienten wurden nicht in die Analyse eingeschlossen (36 keine OP, 12 präoperative Chemo- oder Chemoradiotherapie und 7 mit Ösophagektomie ohne Lymphknotendissektion).

5.3.2.1 Studiendesigns und Studienpopulationen der eingeschlossenen Primärstudien zum Primärstaging

Tabelle 15: Designs, primäre Studienziele und untersuchte Zielgrößen der eingeschlossenen Primärstudien beim Primärstaging

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenzstufe ^b	Verzerrungspotenzial
Buchmann 2006	Diagnostische Kohortenstudie Untersuchungszeitraum „vorwiegend“ zwischen 2003 und 2004. Deutschland	Bewertung der diagnostischen Güte von FDG-PET für N- und M-Staging von Ösophaguskarzinomen und des Einflusses auf die Behandlungsentscheidung.	Diagnostische Güte, Managementänderung	I Ib	hoch
Choi 2000	Diagnostische Kohortenstudie Untersuchungszeitraum zwischen Februar 1997 und Dezember 1998. Südkorea	Bewertung der diagnostischen Güte von FDG-PET zur Erkennung von Lymphknotenmetastasen bei Patienten mit Ösophaguskarzinomen vor chirurgischer Entfernung im Vergleich zu CT und EUS.	Diagnostische Güte	I Ib	niedrig
Choi 2010	Diagnostische Kohortenstudie Untersuchungszeitraum zwischen September 2005 und Juli 2008. Südkorea	Bewertung der diagnostischen Güte von EUS zum Primärstaging bei Patienten mit Ösophaguskarzinomen vor chirurgischer Entfernung im Vergleich zu CT und PET.	Diagnostische Güte	I Ib	niedrig
Flamen 2000	Diagnostische Kohortenstudie Rekrutierungszeitraum zwischen Oktober 1997 und Dezember 1998. Belgien	Bewertung der diagnostischen Güte von FDG-PET, CT, EUS und der Kombination aus CT und EUS zur Erkennung von Fernmetastasen und Lymphknotenmetastasen.	Diagnostische Güte	I Ib	hoch
Heeren 2004	Diagnostische Kohortenstudie Rekrutierungszeitraum zwischen Januar 1996 und Januar 2002. Niederlande	Bewertung der diagnostischen Güte von FDG-PET im Vergleich zur konventionellen Bildgebung bei Patienten mit Karzinomen der Speiseröhre oder der gastroösophagealen Übergangs.	Diagnostische Güte	I Ib	hoch
Itoh 2008	Diagnostische Kohortenstudie Rekrutierung der Patienten zwischen Januar und Dezember 2006. Japan	Bewertung der diagnostischen Güte von FDG-PET/CT (Fusion) bei Patienten mit bereits operierten Karzinomen der Speiseröhre bezüglich Lymphknotenmetastasen.	Diagnostische Güte	I Ib	niedrig

(Fortsetzung)

Tabelle 15: Designs, primäre Studienziele und untersuchte Zielgrößen der eingeschlossenen Primärstudien beim Primärstaging
(Fortsetzung)

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenzstufe ^b	Verzerrungspotenzial
Jager 2001	Diagnostische Kohortenstudie Keine Angaben zum Rekrutierungs- und Untersuchungszeitraum. Niederlande	Bewertung von ¹¹ C-Cholin- und FDG-PET im Vergleich zu Histopathologie als Goldstandard zum Staging von Ösophaguskarzinomen.	Diagnostische Güte	I Ib	hoch
Kato 2005	Diagnostische Kohortenstudie Rekrutierungszeitraum zwischen Januar 2000 und September 2003. Japan	Untersuchung von FDG-PET im Vergleich zu Knochenszinigraphie zur Erkennung von Knochenmetastasen von Patienten mit Ösophaguskarzinomen.	Diagnostische Güte	I Ib	hoch
Kato 2008	Diagnostische Kohortenstudie Rekrutierungszeitraum zwischen Januar 1999 und April 2007. PET: 1999 bis 2005 PET/CT: 2006 bis 2007 Japan	Bewertung von FDG-PET/CT im Vergleich mit FDG-PET alleine in der Erkennung des Lymphknotenstatus von Patienten mit Ösophaguskarzinomen.	Diagnostische Güte	I Ib	niedrig
Kneist 2004	Diagnostische Kohortenstudie Untersuchungszeitraum zwischen Januar 2001 und November 2002. Deutschland	Bewertung der diagnostischen Güte von PET zur Erkennung von Metastasen von Ösophaguskarzinomen und des Einflusses auf die Entscheidung des operativen Vorgehens.	Diagnostische Güte, Managementänderungen	I Ib	hoch
Kole 1998	Diagnostische Kohortenstudie Keine Angaben zum Rekrutierungs- und Untersuchungszeitraum. Niederlande	Bewertung der diagnostischen Güte von FDG-PET zur Erkennung von Lymphknotenmetastasen im Vergleich zu CT und zu chirurgischen und histologischen Ergebnissen als Referenzstandard bei Patienten mit Tumoren der Speiseröhre.	Diagnostische Güte	I Ib	hoch
Liberale 2004	Diagnostische Kohortenstudie Rekrutierungszeitraum zwischen November 1998 und Dezember 2002. Belgien	Bewertung von FDG-PET in der präoperativen Bewertung von Patienten mit Tumoren der Speiseröhre oder gastroösophagealen Übergangs sowie den Einfluss auf Behandlungsentscheidungen.	Diagnostische Güte	I Ib	hoch

(Fortsetzung)

Tabelle 15: Designs, primäre Studienziele und untersuchte Zielgrößen der eingeschlossenen Primärstudien beim Primärstaging
(Fortsetzung)

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenz- stufe ^b	Verzerrungs- potenzial
Lowe 2005	Diagnostische Kohortenstudie Rekrutierungszeitraum zwischen November 2000 und Juli 2002. USA	Bewertung der diagnostischen Güte von CT, EUS und PET zum Staging von Ösophaguskarzinomen sowie des Einflusses auf Behandlungsentscheidungen.	Diagnostische Güte	I Ib	hoch
Noble 2009	Diagnostische Kohortenstudie Rekrutierungszeitraum zwischen Oktober 2006 und September 2007. Vereinigtes Königreich	Untersuchung des Einflusses von PET/CT auf die Behandlungsentscheidungen bei Patienten mit neu diagnostizierten Ösophaguskarzinomen.	Diagnostische Güte, Management- änderungen	I Ib	hoch
Rankin 1998	Diagnostische Kohortenstudie Keine Angaben zum Rekrutierungs- und Untersuchungszeitraum. Vereinigtes Königreich	Vergleich der Ergebnisse von CT und FDG-PET zum präoperativen Staging von Ösophaguskarzinomen.	Diagnostische Güte	I Ib	hoch
Schreurs 2008	Diagnostische Kohortenstudie Rekrutierungszeitraum Januar 2001 und März 2006. Niederlande	Bewertung der Bedeutung von US zum Staging von Ösophaguskarzinomen	Diagnostische Güte	I Ib	niedrig
Shenfine 2009	Multizentrische diagnostische Kohortenstudie Rekrutierungszeitraum zwischen November 2002 und November 2005. Australien	Stellenwert einer präoperativen FDG-PET, um ein fortgeschrittenes Stadium eines Adenokarzinoms der Speiseröhre feststellen zu können.	Diagnostische Güte	I Ib	hoch
Shivo 2004	Diagnostische Kohortenstudie Rekrutierungszeitraum zwischen Dezember 1998 und Oktober 2003. Finnland	Staging von Adenokarzinomen von Speiseröhre und gastroösophagealen Übergangs mit CT, EUS und PET und Vergleich der Ergebnisse mit Histologie nach OP. Zudem Bestimmung der prognostischen Genauigkeit der Vorhersage des Überlebens der Patienten.	Diagnostische Güte	I Ib	hoch

(Fortsetzung)

Tabelle 15: Designs, primäre Studienziele und untersuchte Zielgrößen der eingeschlossenen Primärstudien beim Primärstaging
(Fortsetzung)

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenz- stufe ^b	Verzerrungs- potenzial
Yoon 2003	Diagnostische Kohortenstudie Rekrutierungszeitraum zwischen Februar 2000 und Juli 2001. Südkorea	Bestimmung der diagnostischen Güte von CT und FDG- PET zur Bestimmung des Primärtumors und zur Erkennung von Lymphknotenmetastasen.	Diagnostische Güte	IIb	hoch
<p>a: Studienziel gemäß Formulierung der Autoren b: entspricht der Evidenzstufe des G-BA CT = Computertomographie; EUS = Endoskopischer Ultraschall; FDG = ¹⁸Fluordeoxyglukose; M = Metastase(n); md-CT = Multidetector computed tomography; N = Nodus ((Lymph-) Knoten)); OP = Operation; PET = Positronenemissionstomographie; US = Ultraschall; USA = Vereinigte Staaten von Amerika</p>					

Tabelle 16: Index- und Referenztest der eingeschlossenen Primärstudien beim Primärstaging

Studie	Indextests	Referenztest
Buchmann 2006	PET (ECAT Exact, CTI / Siemens, Deutschland / USA) 60 Minuten nach i.v.-Injektion von 326 (\pm 37) MBq. Bilderstellung von Schädel bis oberer Oberschenkel. Schichtdicke 3,375 mm. Befundung durch 2 erfahrene Nuklearmediziner. Im Falle von größerer Diskrepanz Hinzuziehen eines 3. Befunders und Bewertung im Konsens. Keine Angaben zur Verblindung.	Histologie nach OP und CT (17 Patienten) oder klinisch basierend auf CT-Ergebnissen (3 Patienten, die inoperabel waren oder neoadjuvante Radiochemotherapie erhielten) ^a
Choi 2000	PET (Advance PET scanner, GE Medical Systems, USA) 45 Minuten nach i.v.-Injektion von 370 MBq FDG. Regionale Lymphknoten wurden als bösartig gewertet, wenn FDG-Aufnahme im Vergleich zum normalen Lungenparenchym in 2 oder mehr transversalen Schnitten erhöht war. Befundung im Konsens durch 2 Nuklearmediziner, die gegenüber CT, EUS und Histologie verblindet waren. CT (HiSpeed Advantage Spiral-CT, GE Medical Systems, USA) nach i.v.-Injektion von 100 ml KM (Iopamidol, Bracco Diagnostics, Italien). Bei Lokalisation des Primärtumors im oberen Drittel der Speiseröhre zusätzlich Ultraschall des Halses. Regionale Lymphknoten > 10 mm (kürzeste Achse) wurden als bösartig gewertet. Befundung durch einen Radiologen, der gegenüber EUS, PET und Histologie verblindet war. EUS (GF-UM 20; Olympus Optical Co., Japan) mit 7,5 oder 12 MHz Schallkopf. Beurteilung der Bösartigkeit regionaler Lymphknoten (klare Abgrenzung, runde Erscheinung, Hypoechogenität, Größe > 10 mm). Befundung durch einen Gastroenterologen, der gegenüber CT ^b , PET und Histologie verblindet war.	Histologie nach OP
Choi 2010	PET (Philips Gemini, DA best, Niederlande) 60 Minuten nach i.v.-Injektion von 555-740 MBq (15-20 mCi; 0,22 mg/kg KG) FDG. Aufnahme von Schädel bis mittlerer Oberschenkel. Läsionen wurden als bösartig gewertet, wenn FDG-Aufnahme im Vergleich zum umgebenden Gewebe erhöht war. Keine Angabe zu den Befundern und zur Konsensfindung. CT (Genesis High-Speed CT, GE Medical Systems, USA) nach i.v.-Injektion von 1,4 ml/kg KG KM (Ultravist 370, Schering, Deutschland) mit 120 kVp und 230 mAs. Aufnahme von Hals bis unterer Nierenpol. Lymphknoten > 8 mm (kürzeste Achse) wurden als bösartig gewertet. Keine Angabe zu den Befundern und zur Konsensfindung. EUS (GF-UM 2000; Olympus Optical Co., Japan) mit 5 MHz Schallkopf. Zusätzlich Verwendung von Hochfrequenz-Ultraschall (UM-2R, 20 MHz, Olympus Optical Co., Japan) für oberflächige oder bösartige Strukturen nach Instillation von Wasser in das Lumen und Entfernung von Luft aus dem Lumen der Speiseröhre. Beurteilung der Bösartigkeit der Lymphknoten (klare Abgrenzung, runde Erscheinung, Hypoechogenität, Größe > 10 mm). Lymphknoten innerhalb von 2 cm um den Truncus coeliacus wurden als Fernmetastasen gewertet. Befundung durch einen erfahrenen Endoskopiker. Keine Angaben zur Verblindung.	Histologie nach OP. TNM-Stadieneinteilung nach Klassifikation der AJCC. Als regionale Lymphknoten galten paratracheale, im vorderen und hinteren Mediastinum gelegene, unterhalb der Luftröhre gelegene, perikardiale, nahe dem Ligamentum pulmonale gelegene sowie am rückläufigen Kehlkopfnerv gelegene Lymphknoten. Bösartige Lymphknoten im Hals und nahe dem Truncus coeliacus wurden wie Organmetastasen als Fernmetastasen gewertet. Keine Angaben zum Befunder.

(Fortsetzung)

Tabelle 16: Index- und Referenztest der eingeschlossenen Primärstudien beim Primärstaging (Fortsetzung)

Studie	Indextests	Referenztest
Flamen 2000 ^c	<p>PET (931 / 08 / 12 scanner, CTI / Siemens, USA) 60 Minuten nach i.v.-Injektion von 6,5 MBq/kg KG FDG (maximal 555 MBq). Bilderstellung von Oberkiefer bis mittleres Abdomen. Befundung verblindet gegenüber allen Patientendaten, keine Angaben zum Befunder.</p> <p>CT (keine Angaben zum Gerät) nach i.v.-Injektion von jodhaltigem KM (120 ml für Aufnahmen von Thorax und Abdomen, sonst 80 ml, 350 mg Jod/ml). Lymphknoten > 10 mm (für runde LK) bzw. > 15 mm (ovale LK) wurden als bösartig bewertet. Auswertung durch Radiologen, der gegenüber den Ergebnissen der anderen bildgebenden Verfahren verblindet war.</p> <p>EUS (keine Angaben zum Gerät) mit Konvex- oder Linearsonde. Bei Patienten mit Stenosen wurde üblicherweise keine Dilatation durchgeführt und so nur der sichtbare Teil (oberhalb der Stenose) betrachtet. Bewertung für Bösartigkeit basierend auf Größe, Form, Begrenzung, Echogenität. Befundung durch 1 von 3 Bewertern mit 4-12 Jahren Erfahrung.</p>	Histologie nach OP. Keine Angaben zur Befundung.
Heeren 2004	<p>PET (ECAT 951 / 31 oder ECAT HR+, Siemens / CTI, Deutschland/USA) 90 Minuten nach i.v.-Injektion von 400-580 MBq (6 MBq/kg KG) FDG. Bewertung nach UICC-Klassifikation. Befundung durch 2 erfahrene Nuklearmediziner, die gegenüber den Ergebnissen von CT und EUS verblindet waren.</p> <p>CT (SR 7000, Philips Medical Systems, Niederlande oder Somatron Plus 4 Spiral-CT, Siemens, Deutschland) nach i.v.-Injektion oder oraler Aufnahme von KM (keine weiteren Angaben). Befundung erfolgte verblindet gegenüber EUS, PET und Histologie.</p> <p>EUS (GF-UM 20; 7,5–12 MHz oder MH-908; 7,5 MHz, Olympus Optical Co., Japan). Befundung erfolgte verblindet gegenüber CT, PET und Histologie.</p>	Histologie nach OP oder ultraschallgeführte (Benelux, FGUX-36, 5-7,5 MHz, Pentax, Japan) FNAB mit 22 Gauge und 8 cm Nadel. Klinische oder radiologische Nachbeobachtung Keine Angabe zum Befunder.
Itoh 2008	PET/CT (Fusion) (Philips GEMINI-GXL, Philips, Niederlande) 60 Minuten nach i.v.-Injektion von durchschnittlich 4,0 MBq/kg KG FDG (180 MBq/Ampulle, Nihon Medi + Physics, Japan). Quantitative (Messung des SUV _{max}) und visuelle Befundung. Keine Angaben zur Expertise des Befunders.	Histologie nach Biopsie. Der LK-Status wurde in N0-N4 klassifiziert. Keine weiteren Angaben.
Jager 2001	<p>PET (ECAT Exact 951 / 31 oder HCAT HR+ positron camera, Siemens, Deutschland) 5 Minuten nach i.v.-Injektion von 370 MBq (Spannweite 232-403) ¹¹C-Cholin Herstellung nach Hara und Yuasa^d) bzw. 90 Minuten nach i.v.-Injektion von 370 MBq (Spannweite 170-405) FDG.</p> <p>Befundung unabhängig durch 2 erfahrene Nuklearmediziner, die gegenüber den anderen Staging-Ergebnissen verblindet waren.</p>	Histologie oder klinische Nachbeobachtung. Keine weiteren Angaben.

(Fortsetzung)

Tabelle 16: Index- und Referenztest der eingeschlossenen Primärstudien beim Primärstaging (Fortsetzung)

Studie	Indextests	Referenztest
Kato 2005	<p>PET (SET 2400 W, Shimadzu Corporation, Japan) 40 Minuten nach i.v.-Injektion von 275–370 MBq FDG. Aufnahme von Kopf bis Oberschenkel. Qualitative Befundung durch mindestens 2 erfahrene Nuklearmediziner.</p> <p>Knochenszintigraphie (E.CAM Gammakamera, Toshiba Corporation, Japan) 3–4 Stunden nach i.v.-Injektion von 740 MBq ^{99m}Tc-Technetium-Methyldiphosphonat. Keine Angaben zur Befundung.</p>	<p>Histologie nach OP oder klinisch-radiologische Nachbeobachtung, inkl. CT, MRT, Knochenszintigraphie, FDG-PET und spezifische konventionelle Röntgenuntersuchungen über mindestens 6 Monate.</p> <p>Keine Angaben zur Befundung.</p>
Kato 2008	<p>PET (SET 2400 W, Shimadzu Corporation, Japan) 40 Minuten nach i.v.-Injektion von 275–370 MBq FDG. Aufnahme von Kopf bis Oberschenkel. Qualitative Befundung durch mindestens 2 erfahrene Nuklearmediziner.</p> <p>PET/CT (GE Discovery ST8, GE, USA). Zunächst CT-Untersuchung, dann PET-Untersuchung desselben Volumens.</p> <p>CT-Aufnahme von Hals bis Unterrand der Leber vor und nach i.v.-Injektion von KM (keine weiteren Angaben), keine orale Gabe von KM. Befundung durch 2 Radiologen, die gegenüber den Ergebnissen der PET verblindet waren.</p>	<p>Histologie nach OP.</p> <p>Keine Angaben zur Befundung.</p>
Kneist 2004	<p>PET (ECAT Exact PET scanner, CTI / Siemens, Deutschland / USA) 60 Minuten nach i.v.-Injektion von FDG (durchschnittlich 337 MBq, Spannweite 278–369 MBq). Bilderstellung von Schädel bis oberer Oberschenkel. Schichtdicke 3,375 mm. Rekonstruktion mit OSEM-Algorithmus zur Erstellung einer 128x128-Matrix. Befundung durch mind. 2 erfahrene Nuklearmediziner, die gegenüber den Ergebnissen der CT verblindet waren.</p> <p>CT (Siemens, keine weiteren Angaben zum Gerät bzw. Gerät in einweisender Klinik) von Hals, Thorax und Abdomen. Schichtdicke 5 mm. Lymphknoten (LK) > 6 mm oder kleinere LK, von denen mehr als 3 als Klumpen dargestellt wurden, wurden als positiv gewertet. Für Lymphknoten nahe der Vena azygos und unterhalb der Luftröhrengabelung > 10 mm.</p> <p>Keine Angaben zur Expertise der Untersucher oder zur Konsensusfindung.</p>	<p>Histologie nach OP, klinische Nachbeobachtung</p>

(Fortsetzung)

Tabelle 16: Index- und Referenztest der eingeschlossenen Primärstudien beim Primärstaging (Fortsetzung)

Studie	Indextests	Referenztest
Kole 1998	<p>PET (951 / 31 ECAT Positronenkamera, Siemens / CTI, USA) 30 Minuten nach i.v.-Injektion von 10 mCi FDG. Aufnahme von Hals bis Becken. Befundung erfolgte verblindet gegenüber den Ergebnissen von CT und EUS. Keine Angaben zum Bewerter.</p> <p>CT (SR 7000, Philips Medical Systems, Niederlande oder Somatron Plus 4 spiral CT, Siemens Medical Systems, Deutschland) nach oraler und intravenöser Gabe von KM (keine näheren Angaben). Aufnahme von überlappenden 10-mm Schichten. Befundung erfolgte verblindet gegenüber den Ergebnissen der PET, aber mit Kenntnis des EUS-Befundes.</p>	Histologie nach OP. Keine Angaben zur Befundung.
Libérale 2004	<p>PET (962 HR + PET scanner, CTI, USA) 60 Minuten nach i.v.-Injektion von 6,5 MBq/kg KG FDG. Visuelle Befundung durch 1 Nuklearmediziner, der gegenüber Patientendaten nicht verblindet war.</p> <p>CT (Somatom Plus 4 oder Somatom Volume Zoom; Siemens, Deutschland) nach i.v.-Injektion von 20 mg Butylscopolamin (Buscopan, Boehringer Ingelheim, Deutschland) und 120–150 ml jodhaltigem KM (Ultravist 370, Schering, Deutschland). Aufnahme von Thorax und Abdomen sowie ggf. zervikales CT für Läsionen der oberen Speiseröhre. Befundung durch Radiologen, die gegenüber klinischen Informationen nicht verblindet waren.</p> <p>EUS (keine Angaben zum Gerät) durch 1 Endoskopiker basierend auf cTNM-Klassifikation. Lymphknoten > 10 mm und mit zum Primärtumor vergleichbarer Echogenität wurden als bösartig betrachtet.</p>	Histologie nach Biopsie oder OP. Befundung durch 2 erfahrene Pathologen nach pTNM-Klassifikation (UICC 1997).
Lowe 2005	<p>PET (GE Advance PET Tomograph, General Electric Medical Systems, Inc., USA) 60 Minuten nach i.v.-Injektion von 555–740 MBq FDG. Bilderstellung von Hals bis Becken. Zwei Befunder bewerteten die Bilder visuell. Diskrepanzen wurden im Konsens entschieden. FDG-Aufnahme in der Läsion höher als Hintergrund wurde als bösartig gewertet.</p> <p>CT (General Electric Light Speed Plus, CTI oder High-Speed CT Scanner, General Electric Medical Systems, USA) nach oraler Gabe von 60 ml Esopho-Cat (3,0 % w / w; E-A-EM Inc., USA) und nach i.v.-Injektion von nicht- jodhaltigen jodhaltigem Kontrastmittel (Omnipaque, 300 mg/ml; Amersham Health Inc., USA). Schichtdicke 5 mm. Befundung durch 2 Bewerter, die gegenüber den Ergebnissen anderer bildgebender Verfahren verblindet waren.</p> <p>EUS Zunächst endoskopische Betrachtung von Speiseröhre, Magen und Duodenum zur Bestimmung der Tumorausdehnung. Ggf. Biopsieentnahme und Dilatation als Vorbereitung der EUS. EUS (Olympus GF-UM30; Olympus America Corporation Inc., USA) mit 7,5–12 MHz zur Bestimmung der Tumortiefe (T-Stadium) sowie zur Erkennung von perigastrischen und mediastinalen Lymphknoten (N-Stadium) und Fernmetastasen (M-Stadium). Befundung erfolgte unabhängig durch 2 erfahrene Ärzte.</p>	Histologie nach Biopsie oder OP. Keine Angaben zur Befundung. Alle Patienten wurden für 2 Jahre nach erstem Staging unter Verwendung der Krankenakten nachbeobachtet.

(Fortsetzung)

Tabelle 16: Index- und Referenztest der eingeschlossenen Primärstudien beim Primärstaging (Fortsetzung)

Studie	Indextests	Referenztest
Noble 2009	PET/CT (keine Angaben zum Gerät) 45 Minuten nach i.v.-Injektion von 400 MBq FDG. Keine Gabe von Kontrastmittel für CT. Unphysiologische FDG-Aufnahme außerhalb des Primärtumors wurde als bösartig gewertet. Die Ergebnisse wurden in interdisziplinären Tumorkonferenzen (Viszeralchirurgen, Thoraxchirurgen, Onkologen, Pathologen, Strahlenmediziner und Radiologen) besprochen.	Histologie nach Biopsie oder OP. Klinische Nachbeobachtung inkl. weiterer Bildgebung (z. B. MRT), Kolonoskopie oder Laparoskopie. Keine Angaben zur Befundung.
Rankin 1998	PET (ECAT 951R PET Scanner, Siemens / CTI, USA) 60 Minuten nach i.v.-Injektion von 350 MBq FDG. Befundung im Konsens durch 2 Nuklearmediziner, die gegenüber den Ergebnissen von CT und OP verblindet waren. CT (keine Angaben zum Gerät) nach oraler Gabe von 400 ml KM (keine weiteren Angaben). Die „Mehrheit der Patienten“ zudem mit i.v.-KM (keine weiteren Angaben). Mediastinale Lymphknoten wurden bösartig gewertet, wenn kurze Achse > 1 cm und LK der kleinen Magenkurvatur > 8 mm. Ausdrücke wurden durch 2 Radiologen vor OP befundet. Diese waren gegenüber PET verblindet.	Histologie nach OP. Keine Angaben zur Befundung
Schreurs 2008	FDG-PET wurde mit einem ECAT 951 / 31 oder einem ECAT HR + Positron-Kamera (Siemens / CTI, Knoxville, TN, USA) durchgeführt. Alle Patienten fasteten über 4 Stunden bevor eine mittlere Dosis von 400–580 MBq FDG (je nach Körpergewicht) intravenös verabreicht wurde. FDG-PET-Scans wurden von 1 oder 2 Nuklearmediziner interpretiert. CT Gerät (Somatom Sensation 16 Siemens, Medical Systems, Erlangen, Deutschland). MD-CT Gerät: 3 und 5 mm Kollimation. Scans mit intravenösen und oralen Kontrastmitteln wurden von der Basis des Schädels bis zum oberen Abdomen, einschließlich der gesamten Leber, durchgeführt. EUS 7,5-13,5 MHz Linear-Array-Schallkopf (Siemens Antaris, Siemens Medical Systems, Erlangen, Deutschland)	Histologie nach OP und / oder Nachbeobachtung über 6 Monate
Shenfine 2009	PET (Phillips Allegro GSO PET Scanner, Phillips Medical Systems, Niederlande) 45 Minuten nach i.v.-Injektion von 210-427 MBq FDG. Läsionen mit einem $SUV_{max} > 5$ wurden als bösartig gewertet. Befundung durch Nuklearmediziner.	Histologie nach OP. Keine Angaben zur Befundung.
Shivo 2004	PET (GE Advance scanner, GE Medical Systems, USA) 50 Minuten nach i.v.-Injektion von 370 MBq FDG (mediane Dosis). Aufnahme von Oberkiefer bis zum mittleren Abdomen. Keine Angaben zu CT- oder EUS-Methoden oder zur Befundung der bildgebenden Verfahren.	Histologie nach OP, klinische Nachbeobachtung Keine Angaben zur Befundung

(Fortsetzung)

Tabelle 16: Index- und Referenztest der eingeschlossenen Primärstudien beim Primärstaging (Fortsetzung)

Studie	Indextests	Referenztest
Yoon 2003	<p>PET (Advanced PET Scanner, GE Medical Systems, USA) 45 Minuten nach i.v.-Injektion von 370 MBq (10 mCi) FDG. Aufnahme von Kopf bis Oberschenkel. Erhöhte FDG-Aufnahme im Vergleich zum Hintergrund wurde als bösartig gewertet. Befundung durch erfahrenen Nuklearmediziner, der gegenüber den Ergebnissen der CT verblindet war.</p> <p>CT (HiSpeed Advantage scanner, GE Medical Systems, USA) nach i.v.-Injektion von 100 ml jodhaltigem KM (Iopamidol). Aufnahme mit 120 kVp und 250 mA von Hals bis Nieren. LK > 10 mm wurden als bösartig gewertet. Befundung durch erfahrenen Radiologen, der gegenüber den PET-Ergebnissen verblindet war.</p>	<p>Histologie nach OP, klinische Nachbeobachtung (Spanne 3 bis 19 Monate. Keine Angaben zur Befundung.</p>
<p>a: Es blieb unklar, ob PET-negative Patienten nachbeobachtet wurden. b: Angabe im Text („unaware of the PET, ES and histologic findings“) wurde als Fehler gedeutet und entsprechend korrigiert. c: Die Angaben stammen aus der Publikation von Lerut 2000 [82]. c: siehe [108]</p> <p>AJCC = American Joint Committee on Cancer; CT = Computertomographie; EUS = Endoskopischer Ultraschall; FDG = ¹⁸Fluordeoxyglukose; FNAB = Feinnadelaspirationsbiopsie; GE = General Electric; i.v. = intravenös; kg = Kilogramm; KG = Körpergewicht; KM = Kontrastmittel; kVP = Peak kilovoltage; M = Metastase(n); mAs = Milliamperesekunde(n); MBq = Megabecquerel; mCi = Millicurie; MHz = Megahertz; MRT = Magnetresonanztomographie; N = nodus ((Lymph-) Knoten); OP = Operation; PET = Positronenemissionstomographie; s = Sekunde(n); SUV = Standardized Uptake Value; T = Tumor; US = Ultraschall; USA = Vereinigte Staaten von Amerika</p>		

Tabelle 17: Eingeschlossene Patienten der Primärstudien beim Primärstaging (Ergänzungsrecherche)

Studie	Patienten [n]	Geschlecht [n (m/w)]	Alter (Mittelwert; Median, Spannweite)	Diagnose	Vorbehandlung
Buchmann 2006	20	20 (14 / 6)	59; 59; 35–77	Histologie: Plattenepithelkarzinom (20)	k. A.
Choi 2000	48 ^a	48 (45 / 3)	k. A.; k. A.; 46–77	Histologie: Plattenepithelkarzinom (48) Differenzierungsgrad: guter (15), mittlerer (20), niedriger (6) sarkomatoides Karzinom (4), Karzinom in situ (3). Lokalisation: k. A.	k. A.
Choi 2010	109	109 (103 / 6)	62,7; k. A.; k. A.	Histologie: Plattenepithelkarzinom (101), Adenokarzinom (5), adenoid-zystischer Tumor (2), Karzinosarkom (1). Differenzierungsgrad: guter (23), mittlerer (76), niedriger (7), unbestimmt (3). Lokalisation: oberes Drittel (5), mittleres Drittel (40), unteres Drittel der Speiseröhre (53); gastroösophagealer Übergang (11).	k. A.
Flamen 2000	74 ^b	k. A.	k. A.; k. A.; k. A.	Histologie: Plattenepithelkarzinom (21), Adenokarzinom (53). Lokalisation: mittleres Drittel (13), unteres Drittel der Speiseröhre (30); gastroösophagealer Übergang (31).	Keine Vorbehandlung.
Heeren 2009	74	74 (60 / 14)	62; k. A.; 21–78	Histologie: Plattenepithelkarzinom (12), Adenokarzinom (62). Lokalisation: oberes Drittel (3), mittleres Drittel (5), unteres Drittel der Speiseröhre (32); gastroösophagealer Übergang (34).	k. A.
Itoh 2008	96	96 (75 / 21)	k. A.; 64; 38–84	Histologie: Plattenepithelkarzinom (92); Adenokarzinom (3), gemischtes Karzinom (1). Differenzierungsgrad: guter (23), mittlerer (59), niedriger (14). Ausdehnung des Primärtumors: T0 (1), T1a (10), T1b (32), T2 (11), T3 (41), T4 (1).	k. A.
Jager 2001	18	18 (15 / 3)	k. A.; 59; 22–75	Histologie: Plattenepithelkarzinom (4), Adenokarzinom (11), gastrointestinaler Stromatumor (1), Karzinom in situ (2). Lokalisation: oberes Drittel (2), mittleres Drittel (2), unteres Drittel der Speiseröhre (14).	k. A.

(Fortsetzung)

Tabelle 17: Eingeschlossene Patienten der Primärstudien beim Primärstaging (Ergänzungsrecherche) (Fortsetzung)

Studie	Patienten [n]	Geschlecht [n (m/w)]	Alter (Mittelwert; Median, Spannweite)	Diagnose	Vorbehandlung
Kato 2005	44	k. A.	k. A.; 63,4; 43–76	k. A.	k. A.
Kato 2008	167	167 (154 / 13)	64,1; k. A.; 41–79	Histologie: Plattenepithelkarzinom (167). Lokalisation: oberes Drittel (22), mittleres Drittel (81), unteres Drittel der Speiseröhre (64). Klassifikation: T1 (76), T2 (13), T3 (69), T4 (9). Tumorstadium (UICC 2002): I (52), II (49), III (36), IV (30).	Keine Vorbehandlung
Kneist 2004	81 ^c	81 (65 / 16)	k. A.; 61; 44–79	Histologie: Plattenepithelkarzinom (40), Adenokarzinom (41). Lokalisation: oberes Drittel (8), mittleres Drittel (16), unteres Drittel der Speiseröhre (57). UICC-Stadium: I (9), II (18), III (16), IV (11)	k. A.
Kole 1998	26 ^d	26 (22 / 4)	60; k. A.; 41–76	Histologie: Plattenepithelkarzinom (4), Adenokarzinom (21), bösartiger Stromtumor (1). Lokalisation: unteres Drittel der Speiseröhre (13), gastroösophagealer Übergang (13).	k. A.
Liberale 2004	58	58 (47 / 11)	60; k. A.; 37–80	Histologie: Plattenepithelkarzinom (31), Adenokarzinom (26), Mischform (1). Lokalisation: oberes Drittel (6), mittleres Drittel (22), unteres Drittel der Speiseröhre oder gastroösophagealer Übergang (30). P-Stadium: 0 (2), I (7), IIA (10), IIB (7), III (11), 4A (1), 4B (15), unbestimmt (5), kein Resttumor (3), keine Angaben (2).	k. A.
Lowe 2005	69 ^e	k. A.	Patienten > 18 Jahre (keine weiteren Angaben)	Histologie: Plattenepithelkarzinom (7), Adenokarzinom (62).	k. A.
Noble 2009	191	191 (149 / 42)	k. A.; 66; k. A.	Histologie: Plattenepithelkarzinom (49), Adenokarzinom (133), andere (9). Lokalisation: oberes Drittel (3), mittleres Drittel (26), mittleres/unteres (11), unteres Drittel der Speiseröhre (105), gastroösophagealer Übergang (46).	k. A.

(Fortsetzung)

Tabelle 17: Eingeschlossene Patienten der Primärstudien beim Primärstaging (Ergänzungsrecherche) (Fortsetzung)

Studie	Patienten [n]	Geschlecht [n (m/w)]	Alter (Mittelwert; Median, Spannweite)	Diagnose	Vorbehandlung
Rankin 1998	25	k. A.	k. A.	Histologie: Plattenepithelkarzinom (6), Adenokarzinom (19).	k. A.
Schreurs 2008	55	55 (51 / 4)	64,0; k. A.; 44–84	Histologie: Adenokarzinom (41), Plattenepithelkarzinom (10), andere (3), unbekannt (1). Lokalisation: oberes Drittel (4), mittleres Drittel (39), gastro-ösophagealer Übergang (12).	k. A.
Shenfine 2009	45	45 (37 / 8)	k. A.; 72; 38–82	Histologie: Adenokarzinom (45). Differenzierungsgrad: guter oder mittlerer (21), niedriger (24). Lokalisation: unteres Drittel der Speiseröhre (24); gastroösophagealer Übergang (21).	Keine Vorbehandlung
Shivo 2004	55	55 (42 / 13)	60,5 ^f ; k. A.; k. A.	Histologie: Adenokarzinom (20) oder gastroösophagealen Übergang (35). Keine weiteren Angaben.	k. A.
Yoon 2003	81 ^g	81 (78 / 3)	63; k. A.; 31–90	Histologie: Plattenepithelkarzinomen (81) Lokalisation: oberes Drittel (7), mittleres Drittel der Speiseröhre (42), unteres Drittel der Speiseröhre oder gastroösophagealer Übergang (32).	k. A.
<p>a: In die Studie wurden konsekutiv 61 Patienten eingeschlossen, von denen 5 die OP ablehnten und 8 nicht-operationsfähig waren.</p> <p>b: 74 Patienten eingeschlossen. Für das N-Staging lagen Daten für 36 Patienten vor. Die Daten für das M-Staging wurden nicht verwendet, weil kein adäquater Referenztest verwendet wurde.</p> <p>c: 90 Patienten konsekutiv eingeschlossen. 9 Patienten (2 neoadjuvante Chemotherapie, 2 mit Zeit zwischen Test und OP > 45 Tage, 5 intra- oder postoperativer Nachweis von Karzinomen der Cardia) wurden ausgeschlossen.</p> <p>d: Die Studie schloss 26 Patienten ein. Vier Patienten wurden nicht operiert und daher von der Analyse ausgeschlossen. Ein weiterer Patient wurde von der Bewertung der PET ausgeschlossen.</p> <p>e: Es wurden zunächst 75 Patienten eingeschlossen. Sechs Patienten wurden wieder ausgeschlossen wegen anderer Tumoren.</p> <p>f: eigene Berechnung</p> <p>g: Die Studie schloss 136 Patienten ein. 55 Patienten wurden nicht in die Analyse eingeschlossen (36 keine OP, 12 präoperative Chemo- oder Chemoradiotherapie und 7 mit Ösophagektomie ohne Lymphknotendissektion).</p> <p>m = männlich; k. A. = keine Angabe(n); n = Anzahl; OP = Operation; UICC = Union internationale contre le cancer (dt. Internationale Vereinigung gegen Krebs); w = weiblich</p>					

5.3.2.2 Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien beim Primärstaging

Vierzehn der eingeschlossenen Studien wurden mit einem hohen Verzerrungspotenzial und 5 Studien mit einem niedrigen Verzerrungspotenzial bewertet. Das Verzerrungspotenzial der Diagnosestudien wird in Tabelle 18 dargestellt. Ein „Plus“ bedeutet, dass das jeweilige Kriterium für die entsprechende Studie mit einem niedrigen Verzerrungspotenzial bewertet wurde. Ein „Minus“ bedeutet, dass das genannte Kriterium mit einem hohen Verzerrungspotenzial bewertet wurde. Ein Fragezeichen bedeutet, dass in der Vollpublikation Informationen fehlten, um das Verzerrungspotenzial einschätzen zu können. Das Item „F10: sonstige Aspekte“ wurde in 18 von 19 Studien mit unklar bewertet, weil in den Studien nicht berichtet wurde, ob zwischen dem Index- und dem Referenztest eine Therapie lag. Für das Item „F8: Intention To Diagnose-/Intention To Screen-Analyse“ wurde ein „Plus“ nur dann vergeben, wenn von einem „konsekutiven Einschluss“ berichtet wurde oder angegeben war, dass alle ursprünglich eingeschlossenen Patienten am Ende auch in die statistische Analyse gingen. Eine differenzielle Verifikation (F5) lag in den meisten Studien vor, da Patienten ohne Hinweise aus der PET bzw. PET/CT auf eine Metastasierung üblicherweise nicht biopsiert, sondern nachbeobachtet wurden. Das heißt, dass dieses Item meistens mit „Plus“ oder „Fragezeichen“ bewertet wurde. Für die Gesamtbewertung fällt dieses aber nicht ins Gewicht, da die Nachbeobachtung – sofern sie adäquat lang ist – ebenfalls einen ausreichend exakten Referenztest darstellt.

Generell wurde festgestellt, dass bei 15 Studien (Ausnahmen: Buchmann 2006, Kneist 2004, Shenfine 2009 und Yoon 2003) keine genauen Angaben zur Dauer zwischen Index- und Referenztest gemacht wurden. Bei 14 Studien außer bei Choi 2000, Heeren 2004, Jager 2001 und Lowe 2005 war es unklar, ob die Befunder bei der Interpretation der Ergebnisbilder der Indextests und / oder Referenztests verblindet waren. Die Ergebnisse bei Yoon 2003 wurden nicht verblindet ausgewertet.

Tabelle 18: Verzerrungspotenzial der eingeschlossenen Primärstudien beim Primärstaging

Studie	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Verzerrungs- potenzial der Ergebnisse
Buchmann 2006	+	+	+	+	+	+	?	-	+	?	hoch
Choi 2000	+	?	+	+	+	+	+	+	+	?	niedrig
Choi 2010	+	?	+	+	+	+	?	+	+	?	niedrig
Flamen 2000	+	?	+	+	-	?	?	-	+	?	hoch
Heeren 2004	+	?	+	-	+	+	+	-	+	?	hoch
Itoh 2008	+	?	+	+	+	+	?	+	+	?	niedrig
Jager 2001	+	?	+	+	?	+	+	-	+	?	hoch
Kato 2005	+	?	+	+	-	-	?	?	+	?	hoch
Kato 2008	+	?	+	+	+	+	?	+	+	?	niedrig
Kneist 2004	+	+	+	-	+	+	?	+	+	?	hoch
Kole 1998	+	?	+	-	+	+	?	-	+	?	hoch
Liberale 2004	+	?	+	?	?	?	?	-	+	+	hoch
Lowe 2005	+	?	+	+	?	+	+	-	+	?	hoch
Noble 2009	+	?	+	+	-	?	?	?	+	?	hoch
Rankin 1998	+	?	+	-	+	+	?	-	+	?	hoch
Schreurs 2008	+	?	+	+	-	+	?	+	+	?	niedrig
Shenfine 2009	+	+	+	+	+	+	?	+	-	?	hoch
Shivo 2004	+	?	+	+	+	+	?	-	+	?	hoch
Yoon 2003	+	+	+	-	+	+	-	-	+	?	hoch
+: Kriterium erfüllt; -: Kriterium nicht erfüllt; ?: unklar											

5.3.2.3 Einschätzung der Übertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien auf den deutschen Versorgungskontext beim Primärstaging

In 7 Studien konnten keine Anhaltspunkte für eine Nichtübertragbarkeit gefunden werden. Eine generelle Übertragbarkeit dieser Studien auf den deutschen Versorgungskontext wird daher angenommen. Bei 13 Studien (Buchmann 2006 [79], Choi 2000 [56], Flamen 2000 [81], Heeren 2004 [83], Jager 2001 [85], Kato 2008 [87], Kneist 2004 [88], Liberale 2004 [94], Lowe 2005 [95], Noble 2009 [96], Rankin 1998 [97] und Schreurs 2008 [98], Yoon 2003 [101]) konnten Anhaltspunkte gefunden werden, die eine Übertragbarkeit der Ergebnisse für das Primärstaging auf den deutschen Versorgungskontext einschränken könnten:

- Die Bewertung der Indextests in den Studien von Buchmann 2006 [79], Heeren 2004 [83], Jager 2001 [85], Kato 2008 [87], Noble 2009 [96], Rankin 1998 [97] und Schreurs 2008 [98] wurde gemeinsam bzw. auch unabhängig voneinander durch mehrere Bewerter gemacht.

- In den Studien von Choi 2000 [56], Flamen 2000 [81], Jager 2001 [85], Kneist 2004 [88] und Lowe 2005 [95] wurden die Ergebnisse der Indextests ohne Kenntnis vorhergehender klinischer Untersuchungen und / oder ohne Kenntnis der Ergebnisse der jeweiligen anderen Indextests beurteilt.
- Die Studie von Liberale 2004 [94] schloss ursprünglich 58 Patienten ein. Es wurden aber schließlich nur 24 Patienten ausgewertet. In der Studie von Yoon 2003 [101] wurden ursprünglich 136 Patienten eingeschlossen. 55 Patienten wurden nicht in die Analyse eingeschlossen. Die Ergebnisse der Studien wurden im vorliegenden Bericht dennoch präsentiert, obwohl mehr als 30 % der Daten fehlten, weil die nicht ausgewerteten Patienten nicht der diagnostischen Fragestellung entsprachen (dieses wurde klinisch begründet; genauere Erläuterung unter 5.3.2). Allerdings könnte dadurch die Generalisierbarkeit der Ergebnisse stark eingeschränkt sein.

Die beiden erstgenannten entsprechen möglicherweise nicht dem klinischen Alltag, da meistens den Beurteilern der PET bzw. aller weiteren Tests die Ergebnisse vorangegangener Diagnostik bekannt sind. Die Bewertung der PET bzw. PET/CT erfolgte in 7 Studien gemeinsam durch 2 oder mehrere Bewerter, sodass es plausible Gründe dafür gibt, dass die Ergebnisse möglicherweise nur für einen Teil der Anwendungssituationen in Deutschland gelten.

Tabelle 19: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien beim Primärstaging

Studie	F1	F2	F3	F4	F5	F6	F7	Generelle Einschätzung der Übertragbarkeit
Buchmann 2006	-	-	+	-	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Choi 2000	-	-	-	+	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Choi 2010	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Flamen 2000	-	-	-	+	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Heeren 2004	-	-	+	-	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Itoh 2008	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Jager 2001	-	-	+	+	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Jager 2001	-	-	+	+	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Kato 2005	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Kato 2008	-	-	+	-	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Kneist 2004	-	-	-	+	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Kole 1998	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Liberale 2004	+	-	-	-	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Lowe 2005	-	-	+	+	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Noble 2009	-	-	+	-	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Rankin 1998	-	-	+	-	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Schreurs 2008	-	-	+	-	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
Shenfine 2009	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Shivo 2004	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Yoon 2003	+	-	-	-	-	-	-	Anhaltspunkt für Nichtübertragbarkeit
+: konkrete Anhaltspunkte für eine Nichtübertragbarkeit liegen vor								
-: keine Anhaltspunkte für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden)								

5.3.2.4 Ergebnisse zur diagnostischen Güte der eingeschlossenen Studien beim Primärstaging

In diesem Abschnitt werden die Ergebnisse der Primärstudien zur diagnostischen Güte der PET bzw. der PET/CT und verschiedener anderer Indextests dargestellt. Insgesamt konnten 19 Studien zur diagnostischen Güte beim Primärstaging von Ösophaguskarzinomen identifiziert werden. Wesentliche Charakteristika und die Ergebnisse der eingeschlossenen Studien sind in Tabelle 20 zusammengefasst. Neben den jeweiligen Schätzern wurden auch die Konfidenzintervalle berichtet. Sofern diese nicht in der Studie präsentiert wurden, wurden sie eigenständig nach der Methode von Clopper und Pearson berechnet [109]. Die Daten der Vierfeldertafeln wurden nicht dargestellt, wenn der Anteil der in die Auswertung eingeschlossenen Patienten weniger als 70 % betrug. In allen Studien wurde FDG als Tracer eingesetzt. Die Studie von Jager 2001 untersuchte neben der FDG-PET auch die ¹¹C-Cholin-PET.

Insgesamt wurde in 16 dieser Studien ein direkter Vergleich angestellt. FDG-PET bzw. FDG-PET/CT wurde mit den Technologien CT, EUS, Knochenszintigrafie, ¹¹C-Cholin-PET oder einer Kombination aus CT und EUS direkt verglichen. In 12 Studien wurde PET direkt mit CT verglichen. Diese hohe Anzahl an direkten Vergleichen wird im vorliegenden Bericht metaanalytisch untersucht.

Die Daten der Vierfeldertafeln wurden nicht dargestellt, wenn der Anteil der in die Auswertung eingeschlossenen Patienten weniger als 70 % betrug. In allen Studien wurde FDG als Tracer eingesetzt. Die Studie von Jager 2001 untersuchte neben der FDG-PET auch die ¹¹C-Cholin-PET.

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging

Studie	n (N)	Indikation	Indextest	Trennwerte (Definition Indextest positiv)	Referenztest	Trennwerte	RP	FN	FP	RN	Sensitivität % (95 % KI)	Spezifität % (95 % KI)
Buchmann 2006	16 ^a (20)	N-Staging	FDG-PET	FDG-Aufnahme im Vergleich zum umgebenden Gewebe erhöht	Histologie (keine Angaben zur Nachbeobachtung)	k. A.	2	8	0	6	20 ^c (2,5 bis 55,6) ^c	100 ^c (54,1 bis 100) ^c
	17 ^b (20)	M-Staging	FDG-PET				3	1	1	12	75 ^c (19,4 bis 99,4) ^c	92,3 ^c (64,0 bis 99,8) ^c
	17 ^b (20)						2	2	1	12	50 ^c (6,8 bis 93,2) ^c	92,3 ^c (64,0 bis 99,8) ^c
Choi 2000	48 (48)	N-Staging	FDG-PET	FDG-Aufnahme in LK in 2 Schnitten > FDG-Aufnahme im Lungenparenchym	Histologie	k. A.	26	6	2	14	81 (63,6 bis 92,8) ^{c,d}	87,5 (61,7 bis 98,4) ^{c,d}
			CT	Kurze Achse des LK > 10 mm			13	19	0	16	41 (23,7 bis 59,4) ^{c,d}	100 (79,4 bis 100) ^{c,d}
	45 (48)		EUS	LK scharf begrenzt, rund, hypoechogen und > 10 mm.			15	15	4	11	50 (31,3 bis 68,7) ^{c,d}	73,3 (44,9 bis 92,2) ^{c,d}

(Fortsetzung)

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Trennwerte (Definition Indextest positiv)	Referenztest	Trennwerte	RP	FN	FP	RN	Sensitivität % (95 % KI)	Spezifität % (95 % KI)
Choi 2010	109 (109)	N-Staging	FDG-PET	Aufnahme im Vergleich zum umgebenden Gewebe erhöht.	Histologie (keine Angaben zur Nachbeobachtung)	N1 (+); N0 (-)	27	28	7	47	49 (35,4 bis 62,9 ^{c,d})	87(75,1 bis 94,6 ^{c,d})
		M-Staging	FDG-PET	FDG-Aufnahme im Vergleich zum umgebenden Gewebe erhöht		M1 (+); M0 (-)	2	3	1	103	40 (5,3 bis 85,3 ^{c,d})	99 (94,8 bis 100 ^{c,d})
		N-Staging	CT	LK > 8 mm.		N1 (+); N0 (-)	19	36	4	50	35 (22,2 bis 48,6 ^{c,d})	93 (82,1 bis 97,9 ^{c,d})
		M-Staging	CT	k. A.		M1 (+); M0 (-)	1	4	1	103	20 (0,5 bis 71,6 ^{c,d})	99 (94,8 bis 100 ^{c,d})
		N-Staging	EUS	LK scharf begrenzt, rund, hypoechogen und > 10 mm.		N1 (+); N0 (-)	23	32	5	49	42 (28,7 bis 55,9 ^{c,d})	91 (79,7 bis 96,9 ^{c,d})
		N-Staging	Kombination von FDG-PET, CT und EUS ^e	k. A.		N1 (+); N0 (-)	36	19	11	43	65 (51,4 bis 77,8 ^{c,d})	80 (66,5 bis 89,4 ^{c,d})
		M-Staging	Kombination von FDG-PET und CT ^e	k. A.		M1 (+); M0 (-)	2	3	2	102	40 (5,3 bis 85,3 ^{c,d})	98 (93,2 bis 99,8 ^{c,d})
Flamen 2000	39 (42) ^f	N-Staging	FDG-PET	N1-2 (+); N0 (-)	Histologie nach OP	N1-2 (+); N0 (-)	12	15	1	11	44,4 (25,5 bis 64,7)	91,7 (61,5 bis 99,8)
			CT				13	14	1	11	48,1 (28,7 bis 68,1)	91,7 (61,5 bis 99,8)
			EUS				19	3	7	5	86,4 (65,1 bis 97,1)	41,7 (15,2 bis 72,3)
	34 (42) ^f											

(Fortsetzung)

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Trennwerte (Definition Indextest positiv)	Referenztest	Trennwerte	RP	FN	FP	RN	Sensitivität % (95 % KI)	Spezifität % (95 % KI)	
Heeren 2004	61 (74)	N-Staging	FDG-PET	N1 (+), N0 (-) nach TNM-Klassifikation	FNAB, Histologie, Klinische oder radiologische Nachbeobachtung	N1 (+); N0 (-)	22	18	6	15	55 (38,5 bis 70,7 ^{c,d})	71 (47,8 bis 88,7 ^{c,d})	
	72 (74)	M-Staging	FDG-PET	M1a (+), M0 (-) nach TNM-Klassifikation		M1a (+); M0 (-)	17	7	1	47	71 (48,9 bis 87,4 ^{c,d})	98 (88,9 bis 99,9 ^{c,d})	
	60 (74)	N-Staging	CT	N1 (+), N0 (-) nach TNM-Klassifikation		N1 (+); N0 (-)	17	22	2	19	44 (27,8 bis 60,4 ^{c,d})	90 (69,6 bis 98,8 ^{c,d})	
	72 (74)	M-Staging	CT	M1a (+), M0 (-) nach TNM-Klassifikation		M1a (+); M0 (-)	5	19	1	47	21 (7,1 bis 42,2 ^{c,d})	98 (88,9 bis 99,9 ^{c,d})	
	43 ⁿ (74)	N-Staging	EUS	N1 (+), N0 (-)		N1 (+), N0 (-)	- ⁿ	- ⁿ	- ⁿ				
	52 (74)	M-Staging	EUS	M1a (+), M0 (-) nach TNM-Klassifikation		M1a (+); M0 (-)	2	12	1	37	14 (1,8 bis 42,8 ^{c,d})	97 (86,2 bis 99,9 ^{c,d})	
	72 ^s (74)	M-Staging	CT oder EUS	M1a in CT und/ oder EUS (+); M0 (-) nach TNM-Klassifikation		M1a (+); M0 (-)	7	17	2	46	29 (12,6 bis 51,1 ^{c,d})	96 (85,7 bis 99,5 ^{c,d})	
Itoh 2008	96 (96)	N-Staging	FDG-PET/CT (Fusion)	k. A.	Histologie nach Biopsie	N1 (+), N0 (-)	24	35	4 ^h	33	41 (28,1 bis 54,3) ^{a,b}	89,2 ^a (74,6 bis 97,0) ^{a,b}	

(Fortsetzung)

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Trennwerte (Definition Indextest positiv)	Referenztest	Trennwerte	RP	FN	FP	RN	Sensitivität % (95 % KI)	Spezifität % (95 % KI)	
Jager 2001	17 (18)	M-Staging	FDG-PET	M1b (+)	Laparotomie, OP oder klinische Nachbeobachtung	Als positiv definiert: Stadium M1b	4	0	0	13	100 ^c (39,8 bis 100)	100 (75,3 bis 100)	
	16 (18)		¹¹ C-Cholin-PET				3	1	0	12	75,0 ^c (19,4 bis 99,4)	100 ^c (73,5 bis 100)	
	17 (18)	M-Staging	FDG-PET	M1a (+)		Als positiv definiert: Stadium M1a	4	2	0	11	66,7 (22,3 bis 95,7)	100 ^c (71,5 bis 100)	
	16 (18)		¹¹ C-Cholin-PET				3	2	0	11	60,0 (14,7 bis 94,7)	100 ^c (71,5 bis 100)	
	17 (18)	N- und M-Staging	FDG-PET	Jede Metastase (N oder M; +) ⁱ ; NOM0 (-)		Laparotomie, OP oder klinische Nachbeobachtung	N1 oder M1 (+); NOM0 (-)	4	8	0	5	33,3 (9,9 bis 65,1 ^{c,d})	100 ^c (47,8 bis 100 ^{c,d})
	16 (18)		¹¹ C-Cholin-PET					3	8	0	5	27,3 ^c (6,0 bis 61,0 ^{c,d})	100 ^c (47,8 bis 100 ^{c,d})
Kato 2005	44 (44)	M-Staging	FDG-PET	Nachweis von Knochenmetastasen (+)	Histologie oder klinisch-radiologische Nachbeobachtung	Nachweis von Knochenmetastasen (+)	12	1	2	29	92 (64,0 bis 99,8 ^{c,d})	94 (78,6 bis 99,2 ^{c,d})	
			Knochen-szintigraphie				10	3	5	26	77 (46,2 bis 95,0 ^{c,d})	84 (66,3 bis 94,5 ^{c,d})	

(Fortsetzung)

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Trennwerte (Definition Indextest positiv)	Referenztest	Trennwerte	RP	FN	FP	RN	Sensitivität % (95 % KI)	Spezifität % (95 % KI)
Kato 2008	117 (167)	N-Staging	FDG-PET	N1 (+), N0 (-)	Histologie nach OP	k. A.	33	27	8	49	55,0 (41,6 bis 67,9 ^{c,d})	86,0 (74,2 bis 93,7 ^{c,d})
			CT				29	31	15	42	48,3 (35,2 bis 61,6 ^{c,d})	73,7 (60,3 bis 84,5 ^{c,d})
	50 ^a (167)	PET/CT	- ⁿ	- ⁿ	- ⁿ	- ⁿ	- ⁿ	- ⁿ				
Kneist 2004	81 ^j (81)	M-Staging	FDG-PET	FDG-Aufnahme > umgebendes Gewebe	Histologie nach OP, Biopsie, klinische Nachbeobachtung	k. A.	16	4	8	53	80,0 ^c (56,3 bis 94,3 ^{c,d})	86,9 ^c (75,8 bis 94,2 ^{c,d})
		M-Staging	CT	Für Fern-LK > 6 mm ^k bzw. Verklumpung			16	4	8	53	80,0 ^c (56,3 bis 94,3 ^{c,d})	86,9 ^c (75,8 bis 94,2 ^{c,d})
	41 ^j (81)	M-Staging	FDG-PET (AJCC-Stadium IV)	FDG-Aufnahme > umgebendes Gewebe	7	10	0	24	41 (24 bis 68)	100 (87 bis 100)		
	42 ^j (81)	M-Staging	CT (AJCC-Stadium IV)	Für Fern-LK > 6 mm bzw. Verklumpung	12	4	6	20	75 (50 bis 92)	77 (58 bis 89)		
	42 (81)	N-Staging	FDG-PET (Thorakale LK)	FDG-Aufnahme > umgebendes Gewebe	11	14	1 ^c	16 ^c	42 (23 bis 63)	94 (72 bis 100)		
	41 ^j (81)	N-Staging	CT (Thorakale LK)	LK > 6 mm ^k bzw. Verklumpung (mehr als 3 kleinere Lymphknoten in einer Gruppe)	18	6	5 ^c	12 ^c	75 (55 bis 89)	71 (46 bis 88)		

(Fortsetzung)

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Trennwerte (Definition Indextest positiv)	Referenztest	Trennwerte	RP	FN	FP	RN	Sensitivität % (95 % KI)	Spezifität % (95 % KI)
Kneist 2004 (Fortsetzung)	39 ⁿ (81)	N-Staging	FDG-PET (LK kleine Magenkurvatur)	FDG-Aufnahme > umgebendes Gewebe			– ⁿ	– ⁿ				
	39 ⁿ (81)	N-Staging	CT (LK kleine Magenkurvatur)	LK > 6 mm ^k bzw. Verklumpung			– ⁿ	– ⁿ				
	39 ⁿ (81)	N-Staging	FDG-PET (LK Truncus coeliacus)	FDG-Aufnahme > umgebendes Gewebe			– ⁿ	– ⁿ				
	39 ⁿ (81)	N-Staging	CT (LK Truncus coeliacus)	LK > 6 mm ^k bzw. Verklumpung			– ⁿ	– ⁿ				
	4 ⁿ (81)	N-Staging	CT (Zervikale LK)	LK > 6 mm bzw. Verklumpung			– ⁿ	– ⁿ				
Kole 1998	21 (26)	N-Staging	FDG-PET	FDG-Aufnahme > Hintergrund	Histologie nach OP	k. A.	12	1	1	7	92 (64,0 bis 99,8 ^{c,d})	88 (47,3 bis 99,7 ^{c,d})
	22 (26)		CT	Kurze Achse des LK > 10 mm (+)			5	8	0	9	38 (13,9 bis 68,4 ^{c,d})	100 (66,4 bis 100 ^{c,d})
Liberale 2004	24 (58)	N-Staging	FDG-PET	N1-2 (+); N0 (-)	Histologie nach Biopsie oder OP	N1-2 (+); N0 (-)	3	5	3	13	37,5 ^c (8,5 bis 75,5 ^{c,d})	81,3 ^c (54,4 bis 96,0 ^{c,d})
			CT+EUS	Bösartiger Befund in min. 1 Untersuchung.			2	6	8	8	25,0 ^c (3,2 bis 65,1 ^{c,d})	50,0 ^c (24,7 bis 75,3 ^{c,d})

(Fortsetzung)

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Trennwerte (Definition Indextest positiv)	Referenztest	Trennwerte	RP	FN	FP	RN	Sensitivität % (95 % KI)	Spezifität % (95 % KI)
Lowe 2005	59 (69)	N-Staging	FDG-PET	FDG-Aufnahme > Hintergrund	Histologie nach Biopsie oder OP, klinische Nachbeobachtung	N1 (+); N0 (-)	36	8	6	9	82 (67 bis 92)	60 (32 bis 84)
	48 (69) ⁿ	M-Staging	FDG-PET	FDG-Aufnahme > Hintergrund		M1 (+); M0 (-)	- ⁿ	- ⁿ				
	59 (69)	N-Staging	CT	k. A.		N1 (+); N0 (-)	37	7	5	10	84 (69 bis 93)	67 (38 bis 88)
	48 (69) ⁿ	M-Staging	CT			M1 (+); M0 (-)	- ⁿ	- ⁿ				
	59 (69)	N-Staging	EUS	k. A.		N1 (+); N0 (-)	38	6	5	10	86 (73 bis 95)	67 (38 bis 88)
	48 (69) ⁿ	M-Staging	EUS			M1 (+); M0 (-)	- ⁿ	- ⁿ				
Noble 2009	191 (191)	M-Staging	FDG-PET/CT	Unphysiologisch erhöhte FDG-Aufnahme	Histologie nach Biopsie oder OP. Klinische Nachbeobachtung (inkl. MRT, Koloskopie oder Laparoskopie)	M1 (+); M0 (-)	21	2	10	158	91,3 ^c (72,0 bis 98,9 ^{c,d})	94,0 ^c (89,3 bis 97,1 ^{c,d})

(Fortsetzung)

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Trennwerte (Definition Indextest positiv)	Referenztest	Trennwerte	RP	FN	FP	RN	Sensitivität % (95 % KI)	Spezifität % (95 % KI)
Rankin 1998	18 (25)	N-Staging	FDG-PET	Kurze Achse von periösophagealen LK > 10 mm (+)	Histologie nach OP	N1 (+); N0 (-) (periösophageal)	3	5	1	9	37,5 ^c (8,5 bis 75,5 ^{c,d})	90,0 ^c (55,5 bis 99,7 ^{c,d})
	19 (25)					N1 (+); N0 (-) (kleine Magenkurvatur)	1	8	1	9	11,1 ^c (0,3 bis 48,2 ^{c,d})	90,0 ^c (55,5 bis 99,7 ^{c,d})
	18 (25)		CT	k. A.		N1 (+); N0 (-) (periösophageal)	4	4	2	8	50,0 ^c (15,7 bis 84,3 ^{c,d})	80,0 ^c (44,4 bis 97,59 ^{c,d})
	19 (25)					N1 (+); N0 (-) (kleine Magenkurvatur)	5	4	0	10	55,6 ^c (21,2 bis 86,3 ^{c,d})	100 ^c (69,2 bis 100 ^{c,d})
Schreurs 2008	55 (55)	N-Staging	CT	Axialer LK Durchmesser > 11 mm für jugulodigastric Knoten und > 10 mm für alle anderen Halslymphknoten	Histologie und / oder klinische Anzeichen einer Krankheit während der ersten 6 Monate Follow-up		4	1	3	47	80 (28,4 bis 99,5) ^a	94 (83,5 bis 98,7) ^a
	44 (55)		FDG-PET				4	0	2	38	100 (39,8 bis 100) ^a	95 (83,1 bis 99,4) ^a
	44 (55)		CT + FDG-PET				4	0	1	39	100 (39,8 bis 100) ^a	98 (86,8 bis 99,9) ^a

(Fortsetzung)

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Trennwerte (Definition Indextest positiv)	Referenztest	Trennwerte	RP	FN	FP	RN	Sensitivität % (95 % KI)	Spezifität % (95 % KI)
Shenfine 2009	45 (45)	N-Staging	FDG-PET	Fortgeschrittene Erkrankung (T ≥ 3 oder N1), SUV ≥ 5 (+), frühes Krankheitsstadium (T ≤ 2 und N0), SUV < 5 (-)	Histologie nach OP	identisch zu Spalte Indextest	22	9	4	10	71 (52,0 bis 85,8 ^{c,d})	71,4 (41,9 bis 91,6 ^{c,d})
Shivo 2004	43 (55) ¹	N-Staging	FDG-PET	Unphysiologisch erhöhte FDG-Aufnahme	Histologie nach OP, klinische Nachbeobachtung	N1 (+); N0 (-)	9	17	0	17	35 (17,2 bis 55,7 ^{c,d})	100 (80,5 bis 100 ^{c,d})
	55 (55)	M-Staging	FDG-PET			M1 (+); M0 (-)	10	9	4	32	53 (28,9 bis 75,6 ^{c,d})	89 (73,9 bis 96,9 ^{c,d})
	42 (55)						7	8	3	24	47 (21,3 bis 73,4 ^{c,d})	89 (70,8 bis 97,6 ^{c,d})
	43 (55) ¹	N-Staging	CT	k. A.		N1 (+); N0 (-)	11	15	3	14	42 (23,4 bis 63,1 ^{c,d})	82 (56,6 bis 96,2 ^{c,d})
	55 (55)	M-Staging	CT			M1 (+); M0 (-)	6	13	1	35	32 (12,6 bis 56,6 ^{c,d})	97 (85,5 bis 99,9 ^{c,d})
	42 (55)					5	10	1	26	33 (11,8 bis 61,6 ^{c,d})	96 (81,0 bis 99,9 ^{c,d})	
	43 (55) ¹	N-Staging	EUS			N1 (+); N0 (-)	22	4	8	9	85 (65,1 bis 95,6 ^{c,d})	53 (27,8 bis 77,0 ^{c,d})
	43 (55) ¹	N-Staging	FDG-PET und CT	k. A., wie Ergebnisse kombiniert wurden		N1 (+); N0 (-)	13	13	0	17	50 (29,9 bis 70,1 ^{c,d})	100 (80,5 bis 100 ^{c,d})
55 (55)	M-Staging	FDG-PET und CT	M1 (+); M0 (-)			12	7	0	36	64 ^m (38,4 bis 83,7 ^{c,d})	100 (90,3 bis 100 ^{c,d})	

(Fortsetzung)

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Trennwerte (Definition Indextest positiv)	Referenztest	Trennwerte	RP	FN	FP	RN	Sensitivität % (95 % KI)	Spezifität % (95 % KI)
Shivo 2004 (Fortsetzung)	43 (55) ¹	N-Staging	EUS and CT	k. A., wie Ergebnisse kombiniert wurden		N1 (+); N0 (-)	22	4	3	14	85 (65,1 bis 95,6 ^{c,d})	82 (56,6 bis 96,2 ^{c,d})
	55 (55)	M-Staging	EUS und CT			M1 (+); M0 (-)	8	11	0	36	42 (20,3 bis 66,5 ^{c,d})	100 (90,3 bis 100 ^{c,d})
	43 (55) ¹	N-Staging	FDG-PET, CT und EUS	k. A., wie Ergebnisse kombiniert wurden		N1 (+); N0 (-)	22	4	0	17	85 (65,1 bis 95,6 ^{c,d})	100 (80,5 bis 100 ^{c,d})
	55 (55)	M-Staging	FDG-PET, CT und EUS			M1 (+); M0 (-)	14	5	0	36	74 (48,8 bis 90,9 ^{c,d})	100 (90,3 bis 100 ^{c,d})
Yoon 2003	81 (136)	N-Staging	FDG-PET	FDG-Aufnahme > Hintergrund (+)	Histologie nach OP, klinische Nachbeobachtung	N1 (+); N0Mx (-)	25	14	13	29	64 (47,2 bis 78,8 ^{c,d})	69 (52,9 bis 82,4 ^{c,d})
		M-Staging	FDG-PET	FDG-Aufnahme > Hintergrund (+)		M1 (+); M0 (-)	3	4	1	73	43 (9,9 bis 81,6 ^{c,d})	99 (92,7 bis 100 ^{c,d})
		N-Staging	CT	LK > 10 mm (+)		N1 (+); N0Mx (-)	12	27	6	36	31 (17,0 bis 47,6 ^{c,d})	86 (71,5 bis 94,6 ^{c,d})
		M-Staging	CT	k. A.		M1 (+); M0 (-)	1	6	4	70	14 (0,4 bis 57,9 ^{c,d})	95 (86,7 bis 98,5 ^{c,d})

(Fortsetzung)

Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging (Fortsetzung)

- a: Die Daten stammen aus Tabelle 2 der Publikation. Drei Patienten wurden ausgeschlossen, weil diese eine Therapie zwischen Index- und Referenztest erhielten. 17 Patienten waren damit für den vorliegenden Bericht relevant. Für das N-Staging wurde ein Patient aus der Bewertung ausgeschlossen, da der Referenztest uneinheitlich war. CT zeigte verdächtige Läsion, während Pathologie unauffällig blieb.
- b: Nach Pathologie, falsch- positiver Befund (Metastase im Lungenapex) und richtig- positiver Befund (bronchohilar) bei einem Patienten. Hier sind die 2 Möglichkeiten dargestellt, die sich ergaben, wenn dieser Patient nur in jeweils einer der Kategorien gezählt wurde: In der oberen Reihe als RP, unten als FN gewertet. Es ergaben sich zur Publikation abweichende Sensitivitäten. In Tabelle 3 der Publikation wurde der Schätzer der Sensitivität mit 60 % angegeben. Möglicherweise wurde der Grenzfall jeweils zur Hälfte als RP und FN gewertet.
- c: Wert(e) wurde(n) selbst berechnet.
- d: Keine oder unzureichende Angabe des 95 %-KI in der Publikation.
- e: Keine Angabe, wie diese Tests kombiniert wurden.
- f: Die Daten stammen aus der Publikation Lereut 2000 (zweite Publikation zur Studie).
- g: Von den 72 in der Studie eingeschlossenen Patienten erhielten 52 EUS.
- h: Im Text der Publikation wurde der Teiler (RN+FP) zur Berechnung der Spezifität mit 47 angegeben, was zu einer falschen Anzahl ausgewerteter Patienten führen würde. Tabelle 2 gab die Zahl mit 37 an, was zur richtigen Anzahl ausgewerteter Patienten führte. Hier wurden diese Daten verwendet. Somit hatten 4 Patienten einen falsch positiven Befund.
- i: Außerdem Angaben zur Sensitivität für N1, M1a und M1b. Jedoch wurden die entsprechenden Schätzer der Spezifität nicht berichtet.
- j: Anzahl der Patienten, die für die jeweilige Indikation (N- oder M-Staging) berichtet wurden.
- k: Für Lymphknoten nahe der Vena azygos und unterhalb der Luftröhrengabelung > 10 mm.
- l: Räsänen et al. 2003 war eine Zwischenauswertung von Shivo 2004 [100]. Räsänen enthält eine Teilmenge der Patienten von Shivo und untersucht das N-Staging.
- m: Rechnerisch ergab sich hier eine leicht abweichende Sensitivität von 63,2 %.
- n: Die Ergebnisse wurden nicht dargestellt, weil der Anteil fehlender Werte größer als 30 % war.

AJCC = American Joint Committee on Cancer; CT = Computertomographie; FDG = ¹⁸Fluorodeoxyglukose; FN = falsch negativ; FNAB = Feinnadelaspirationsbiopsie; FP = falsch positiv; k. A. = keine Angabe(n); KI = Konfidenzintervall; LK = Lymphknoten; M = Metastase(n); MRT = Magnetresonanztomographie; N = nodus ((Lymph-) Knoten); PET = Positronenemissionstomographie; RN = richtig negativ; RP = richtig positiv; SUV = standardized uptake value; UICC = Union internationale contre le cancer (dt. Internationale Vereinigung gegen Krebs)

Direkte Vergleiche zwischen PET und CT für das N-Staging

In der Abbildung 2 wird der Forest Plot der Sensitivität und der Spezifität für alle Studien, die einen direkten Vergleich zwischen PET und CT für das N-Staging durchführten und in die bivariate Meta-Analyse in den vorliegenden Bericht eingingen, dargestellt.

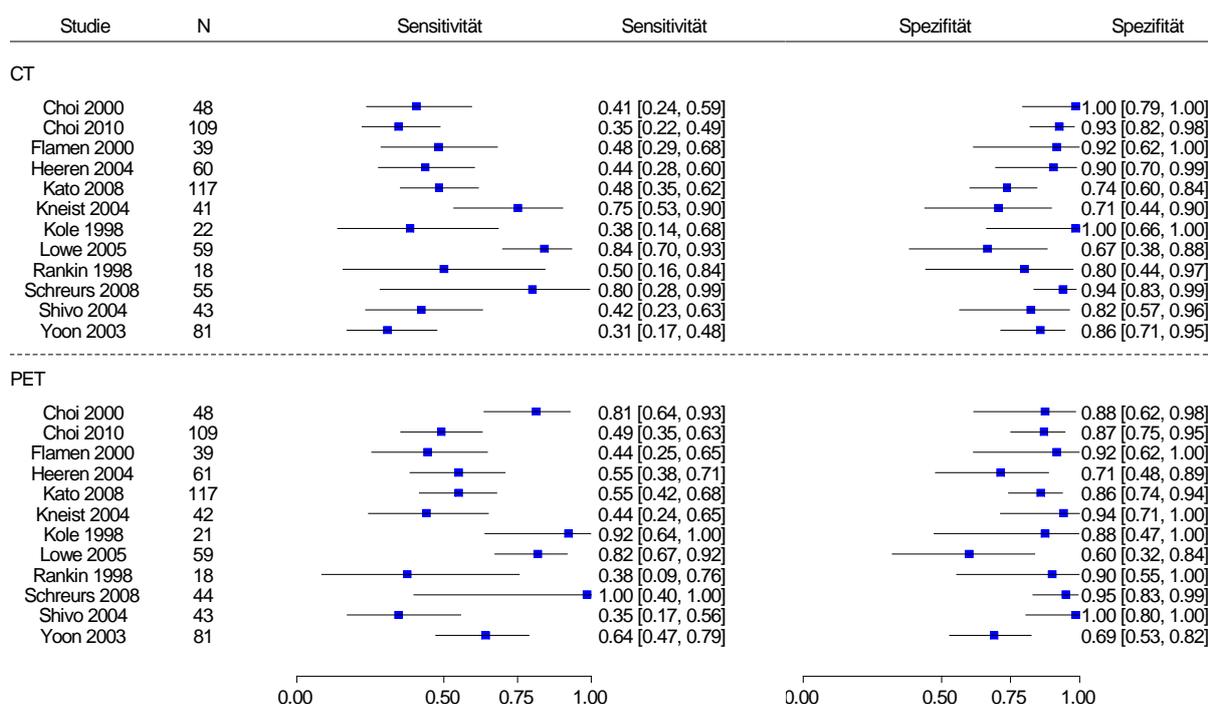


Abbildung 2: Forest Plot der univariaten Ergebnisse für die diagnostische Güte der PET und CT beim N-Staging von Ösophaguskarzinomen

Die Abbildung 3 zeigt die direkten bivariaten Vergleiche zwischen PET und CT für das N-Staging basierend auf 12 Studien (Darstellung der direkten Vergleiche (oben links), bivariate Sensitivitäts- / Spezifitäts-Punktschätzer mit Konfidenzellipsen (oben rechts) und HSROC-Kurven (unten links)).

In der Abbildung oben links werden die direkten Vergleiche zwischen PET und CT dargestellt. Die grauen Linien zwischen je einem blauen und einem schwarzen Kreis verbinden die PET- und CT-Ergebnisse aus derselben Studie. In 7 von 12 Vergleichen ergab sich ein höherer Punktschätzer der PET für Sensitivität im Vergleich zur CT. In 2 von diesen 7 direkten Vergleichen fand sich für die PET ein höherer Punktschätzer für die Spezifität im Vergleich zur CT. In 5 von 12 direkten Vergleichen ergab sich ein höherer Punktschätzer der PET für die Spezifität im Vergleich zur CT. In 2 von diesen 5 direkten Vergleichen fand sich für die PET ein höherer Punktschätzer für die Sensitivität im Vergleich zur CT. Diese direkten Vergleiche zeigen eine große Unstrukturiertheit: Die diagnostische Güte der PET

und CT sind ohne erkennbares Muster vermischt und grenzen sich nicht räumlich voneinander ab.

Die Kreise in der obigen rechten Abbildung zeigen die jeweiligen Schätzer der eingeschlossenen Primärstudien, wobei die Größe der Kreise die Anzahl der eingeschlossenen Patienten widerspiegelt. In der Abbildung unten links wurde ein HSROC-Modell angepasst, in dem die 2 Diagnoseverfahren als Kovariable eingegangen sind. Dabei wurde angenommen, dass die Varianzen der zufälligen Effekte für alle Diagnoseverfahren gleich sind. Die Punktschätzer der mittleren Sensitivität und Spezifität wurden aus den geschätzten HSROC-Kurven ermittelt. Die HSROC-Kurven liegen sehr nahe beieinander und schneiden sich, woraus nicht auf einen Unterschied der beiden diagnostischen Verfahren geschlossen werden kann. Der ausgefüllte Punkt bzw. das Dreieck stellen die gepoolten Punktschätzer dar. Sie entsprechen jeweils dem mittleren Paar von Sensitivität und Spezifität der PET bzw. CT. Die 95 %-Konfidenzregionen überlappen sich, sodass keine Aussage darüber getroffen werden kann, ob sich die Verfahren in der erwarteten Sensitivität und Spezifität unterscheiden. Die gepoolten Schätzer der bivariaten Meta-Analyse sind Tabelle 21 zu entnehmen.

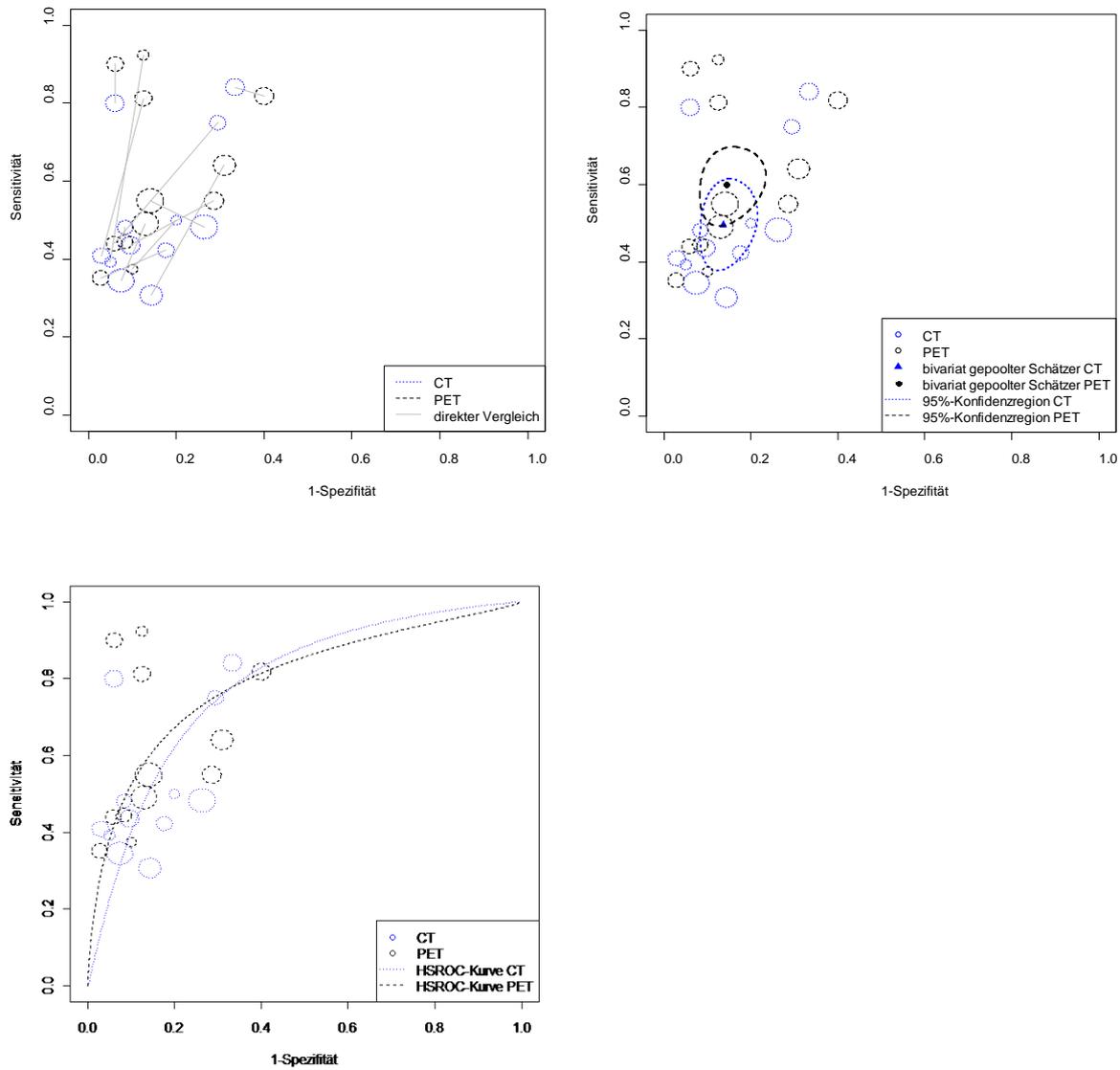


Abbildung 3: Ergebnisse der bivariaten Meta-Analyse PET versus CT beim N-Staging von Ösophaguskarzinomen (direkte Vergleiche)

Tabelle 21: Ergebnisse der bivariaten Meta-Analyse für das N-Staging (gepoolte Schätzer, direkte Vergleiche)

Diagnosemethode	Sensitivität in % (95 %-KI) ^a	Spezifität in % (95 %-KI) ^a
PET	59,9 (51–68)	85,8 (79–91)
CT	49,6 (40–59)	86,4 (80–91)
a: Das Konfidenzintervall stammt aus der bivariaten Modellierung		

Direkte Vergleiche zwischen PET und CT für das M-Staging

In Abbildung 4 wird der Forest Plot der Sensitivität und der Spezifität für alle Studien, die einen direkten Vergleich zwischen PET und CT für das M-Staging durchführten und in die bivariate Meta-Analyse in den vorliegenden Bericht eingingen, dargestellt.

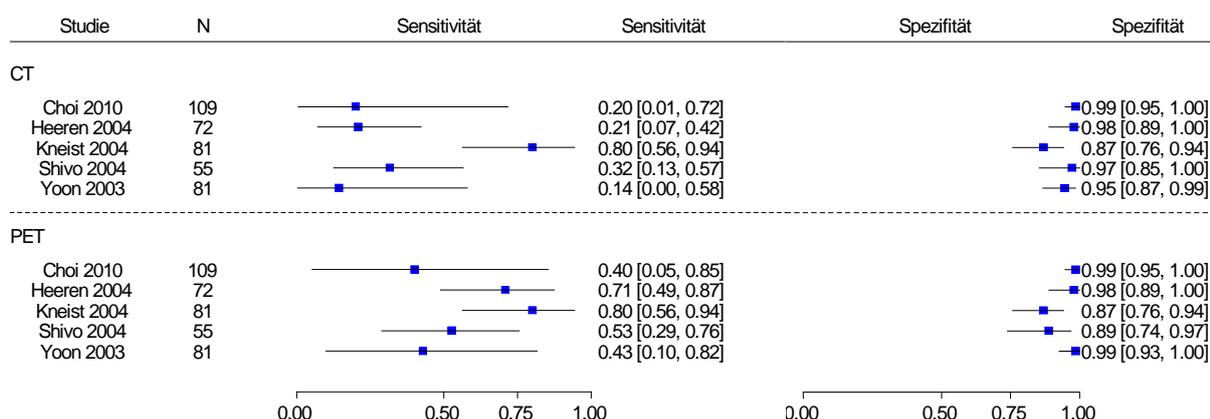


Abbildung 4: Forest Plot der Ergebnisse für die diagnostische Güte der PET beim M-Staging von Ösophaguskarzinomen

Die Abbildung 5 stellt die direkten bivariaten Vergleiche zwischen PET und CT für das M-Staging dar (Darstellung der direkten Vergleiche (oben links), bivariate Sensitivitäts- / Spezifitäts-Punktschätzer mit Konfidenzellipsen (oben rechts) und HSROC-Kurven (unten links)). Ein direkter Vergleich beider Technologien für das M-Staging wurde in 5 Studien untersucht [80,83,88,99,101]. Mangels fehlender Daten konnten keine weiteren Vergleichstechnologien in die Meta-Analyse integriert werden.

In der Abbildung oben links werden die direkten Vergleiche zwischen PET und CT dargestellt. In 4 von 5 Vergleichen ergab sich ein höherer Punktschätzer der PET für Sensitivität im Vergleich zur CT. In einem von diesen 4 direkten Vergleichen fand sich für die PET ein höherer Punktschätzer für die Spezifität im Vergleich zur CT. In einem direkten Vergleich gab es keinen Unterschied zwischen dem Punktschätzer der PET für Spezifität und dem für die CT. In 4 von 5 Vergleichen waren die Punktschätzer der PET für Spezifität identisch mit denen für die CT. In 3 von diesen 4 Vergleichen erzielte die PET eine höhere Sensitivität als die CT.

Die Abbildung oben rechts stellt dar, dass sich die 95 %-Konfidenzregionen überlappen, sodass keine Aussage darüber getroffen werden kann, ob sich die Verfahren in der erwarteten Sensitivität und Spezifität unterscheiden. Darüber hinaus schneiden sich die geschätzten HSROC-Kurven von PET und CT (vgl. Abbildung unten links), sodass auch hieraus nicht auf

einen Unterschied der beiden Verfahren geschlossen werden kann. Die gepoolten Schätzer der bivariaten Meta-Analyse sind Tabelle 22 zu entnehmen.

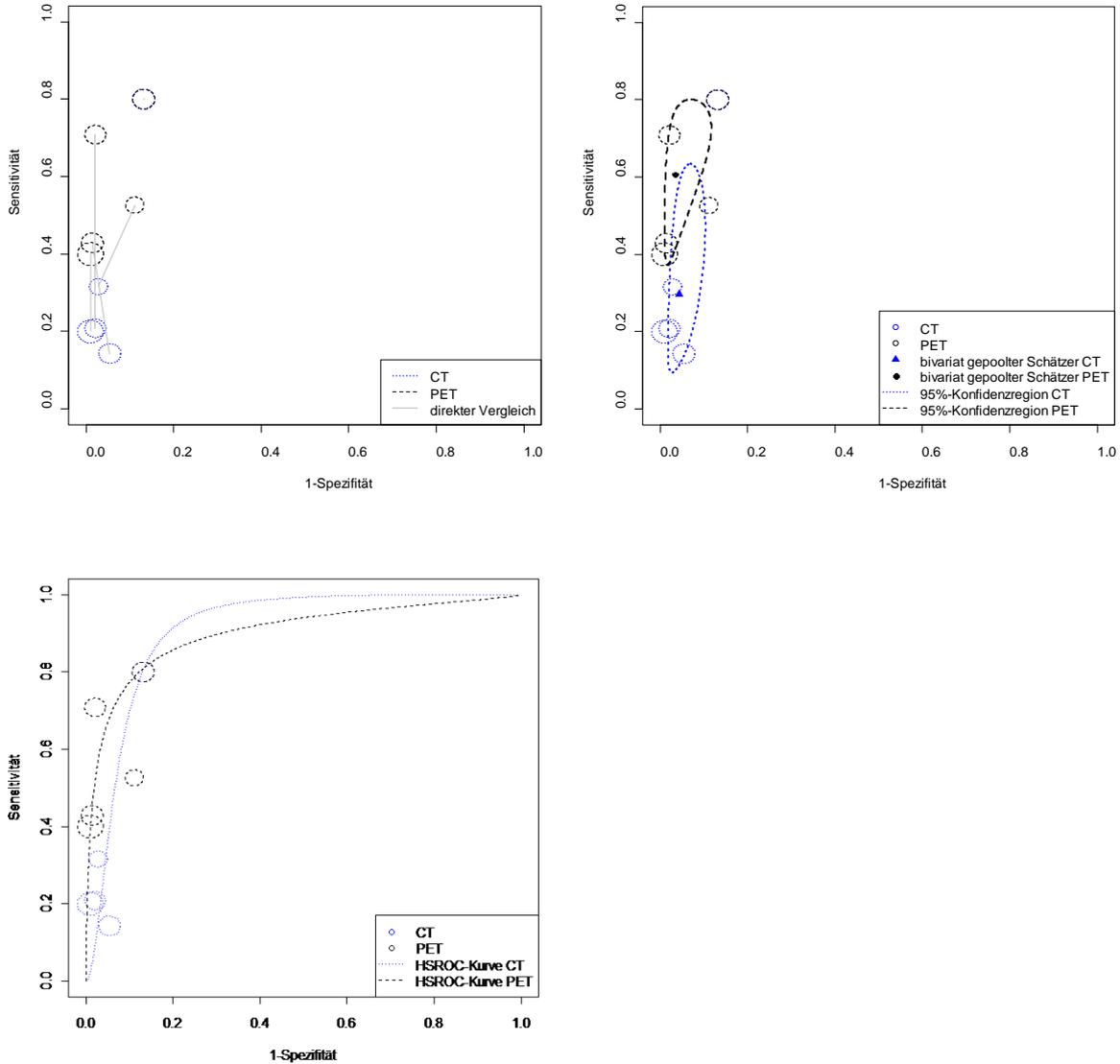


Abbildung 5: Ergebnisse der bivariaten Meta-Analyse PET versus CT beim M-Staging von Ösophaguskarzinomen (direkte Vergleiche)

Tabelle 22: Ergebnisse der bivariaten Meta-Analyse für das M-Staging (gepoolte Schätzer, direkte Vergleiche)

Diagnosemethode	Sensitivität in % (95 %-KI) ^a	Spezifität in % (95 %-KI) ^a
PET	60,7 (42–77)	96,6 (91–99)
CT	29,8 (12–57)	95,7 (91–98)

a: Das Konfidenzintervall stammt aus der bivariaten Modellierung

Subgruppenmerkmale und andere Effektmodifikatoren

In der Studie von Kneist 2004 [88] wurden verschiedene Subgruppenanalysen durchgeführt. Für das M-Staging wurde für die PET und die CT neben den AJCC-Stadien I bis III das Stadium IV explizit noch in einer Subgruppenanalyse untersucht. Daneben wurden in der Studie weitere Subgruppenanalysen für die PET und CT bezüglich verschiedener Lymphknoten durchgeführt. Die Daten der letztgenannten Subgruppenanalysen wurden allerdings nicht im Bericht dargestellt, weil der Anteil fehlender Werte größer als 30 % war.

Sensitivitätsanalysen

Sensitivitätsanalysen bezüglich des Verzerrungspotenzials wurden nicht durchgeführt, da sich die Sensitivitäten und Spezifitäten der hoch und niedrig verzerrten Studien vor dem Hintergrund des unterstellten Modells nicht systematisch unterschieden. Aufgrund der durchgeführten Sensitivitätsanalysen ist davon auszugehen, dass die hier präsentierten Ergebnisse robust sind.

Ergänzend betrachtete Ergebnisse zu Änderungen im klinischen Management

In der Studie von Buchmann 2006 [79] wurde eine Änderung im klinischen Management basiert auf dem Einsatz der PET (siehe Tabelle 23). Bei einem Patienten wurden nur durch den Einsatz der PET Metastasen der supraklavikulären Lymphknoten entdeckt. Der Status des Patienten wurde von M0 auf M1 hinaufgesetzt. Der Patient befand sich damit im AJCC-Stadium IV. Aufgrund dessen kam es zur Therapieveränderung.

In der Studie von Kneist 2004 [88] wurden Angaben zur Änderung im klinischen Management basierend auf den Einsatz der PET gemacht. Durch die Ergebnisse der PET wurde bei insgesamt 3 von 81 in der Studie eingeschlossenen Patienten eine Änderung im klinischen Management herbeigeführt (siehe Tabelle 23).

Auch in der Studie von Noble 2009 [96] wurde berichtet, dass von 191 eingeschlossenen Patienten sich die Behandlung bei 50 Patienten durch den Einsatz der PET/CT veränderte. Die PET/CT wurde zusätzlich zur konventionellen Diagnostik CT und EUS eingesetzt (Add-on). Nähere Angaben zur Managementänderung werden nicht dargestellt (siehe Tabelle 23).

Tabelle 23: Ergänzend betrachtete Ergebnisse zu Änderungen im klinischen Management im Vergleich zur konventionellen Diagnostik (CT und EUS) beim Primärstaging

Studie	Indextest	Anzahl Patienten in der Studie	Davon Patienten mit Managementänderung: n (%)	Angaben zur Managementänderung
Buchmann 2006	FDG-PET	20	1 (5)	Nachweis Metastasen der supraclavikulären Lymphknoten durch PET. Nach Bestätigung in der Pathologie; Änderung des Status (M0→M1) und Therapieänderung (Resektion).
Kneist 2004	FDG-PET	81	3 (4)	1 Patient: Durch das PET-Bild wurde ein Adenom entdeckt. Dieses Adenom wurde nach transanaler endoskopischer Mikrooperation diagnostiziert. 1 Patient: Kolon Karzinom wurde diagnostiziert und reseziert. 1 Patient: Karzinom der Epiglottis, das durch lokale Resektion entfernt wurde.
Noble 2009	FDG-PET/CT	191	50 (26)	k. A.
CT = Computertomographie; FDG = ¹⁸ Fluorodeoxyglukose; k. A. = keine Angabe(n); n = Anzahl; OP = Operation; PET = Positronenemissionstomographie				

5.3.3 Prognostische Güte auf Grundlage von Primärstudien beim Primärstaging (Ergänzungsrecherche)

Zur prognostischen Güte fanden sich 5 Arbeiten (Cheze-Le Rest 2008 [75], Choi 2004 [55], Chung 2008 [76], Omloo 2008 [59] und Shenfine 2009 [61,75]). In 3 der 5 Studien wurden Behandlungen, die während der Nachbeobachtungszeit durchgeführt wurden, nicht standardisiert oder randomisiert durchgeführt, in 2 Studien blieb dieses unklar.

Eine der prognostischen Kohortenstudien wurde in 2 Aufsätzen **Cheze-Le Rest 2008** und **Teyton 2009** ([75,106]) publiziert. Es wurden konsekutiv 47 Patienten mit neu diagnostiziertem Ösophaguskarzinom eingeschlossen, bei denen ein Primärstaging vor Aufnahme der Behandlung durchgeführt wurde. Ziel der Studie war die Bewertung, ob eine FDG-PET-Untersuchung im Rahmen des Routinestagings von Patienten mit Ösophaguskarzinomen das Gesamtüberleben vorhersagen kann. 25 der eingeschlossenen Patienten verstarben zum Zeitpunkt der letzten Nachbeobachtung (Median der Nachbeobachtungszeit 32 Monate; Spanne 24 bis 52 Monate).

Die Studie von **Choi 2004** [55] schloss konsekutiv 69 Patienten mit neu diagnostiziertem Ösophaguskarzinom ein, bei denen ein Primärstaging vor Aufnahme der Behandlung durchgeführt wurde. Ziel der Studie war die Bewertung, ob eine FDG-PET-Untersuchung im Rahmen des Routinestagings von Patienten mit Ösophaguskarzinomen das Gesamtüberleben und das krankheitsfreie Überleben vorhersagen kann.

In **Chung 2008** [76] wurden konsekutiv 100 Patienten mit neu diagnostiziertem Ösophaguskarzinom eingeschlossen, bei denen ein Primärstaging vor Aufnahme der Behandlung durchgeführt wurde. Ziel der Studie war die Bewertung der prognostischen Güte von TISPECT und FDG-PET. Dabei wurden verschiedene Faktoren hinsichtlich ihrer Wertigkeit in der Vorhersage des krankheitsfreien Überlebens untersucht. Am Ende der klinischen Nachbeobachtung (durchschnittlich 20,3 Monate, Spannweite 0 bis 62,7 Monate) wiesen 38 der 100 Patienten keine Krankheitssymptome auf.

Die Studie von **Omloo 2008** [59] hatte das Ziel, die prognostische Güte der PET bei Patienten mit Ösophaguskarzinomen zu untersuchen, welche hauptsächlich durch OP behandelt wurden. Weiter wurde untersucht, welchen zusätzlichen Wert die PET im Vergleich zur EUS hatte. Es wurden 125 Personen mit histologisch nachgewiesenen primären Ösophaguskarzinomen eingeschlossen. Die klinische Nachbeobachtung betrug 25 Monate. Es wurden verschiedene prognostische Merkmale hinsichtlich ihrer Wertigkeit in der Vorhersage des Gesamtüberlebens untersucht. Das mediane Überleben der Patienten mit einem Rezidiv betrug 13 Monate (Interquartilsabstand 8 bis 21 Monate) und bei Patienten ohne Rezidiv am Ende der Nachbeobachtung 35 Monate (Interquartilsabstand 28 bis 42 Monate). Sechs Patienten starben aus anderen Gründen.

Shenfine 2009 [61] untersuchte die diagnostische Güte und prognostische Güte der PET bzw. PET/CT beim Primärstaging. An dieser Stelle werden die Ergebnisse der prognostischen Güte dargestellt. Die in der Studie berichteten Ergebnisse zur diagnostischen Güte wurden bereits unter 5.3.2 erläutert. Es wurden 45 Patienten mit neu nachgewiesenen Adenokarzinomen der Speiseröhre und des gastroösophagealen Übergangs eingeschlossen. Die Arbeit untersuchte, ob eine präoperative quantitative Bewertung von FDG-PET ein unabhängiger Faktor zur Vorhersage des krankheitsfreien Überlebens und des Gesamtüberlebens von Patienten mit Adenokarzinomen der Speiseröhre ist. Als Referenztest wurde die Nachbeobachtung mit einer Länge von 18 bis 61 Monaten (Median 44 Monate) verwendet.

5.3.3.1 Studiendesigns und Studienpopulationen der eingeschlossenen Prognosestudien beim Primärstaging

Tabelle 24: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudien beim Primärstaging

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenzstufe ^b	Verzerrungs- potenzial
Cheze-Le Rest 2008	Prognostische Kohortenstudie Rekrutierung der Patienten konsekutiv zwischen Januar 2003 und Mai 2005. Frankreich	Bewertung, ob eine FDG-PET-Untersuchung im Rahmen des Routinestagings von Patienten mit Ösophaguskarzinomen das Gesamtüberleben vorhersagen kann.	Prognostische Güte von FDG-PET zur Vorhersage des Gesamtüberlebens	II ^c	hoch
Choi 2004	Prognostische Kohortenstudie Rekrutierung der Patienten konsekutiv zwischen Februar 1997 und Mai 2000. Südkorea	Bewertung, ob PET-Parameter (SUV des Primärtumors, Tumurlänge, Anzahl der Lymphknoten mit erhöhter FDG-Aufnahme und PET-Stadium) eine Vorhersage des Krankheitsverlaufes von Patienten mit Ösophaguskarzinomen ermöglichen.	Prognostische Güte von FDG-PET zur Vorhersage des krankheitsfreien Überlebens und des Gesamtüberlebens.	II ^c	hoch
Chung 2008	Prognostische Kohortenstudie Rekrutierung der Patienten konsekutiv zwischen Oktober 2000 und Juni 2004. Südkorea	Untersuchung der prognostischen Güte von TI-SPECT und FDG-PET.	Prognostische Güte von TI-SPECT und FDG-PET zur Vorhersage des krankheitsfreien Überlebens.	II ^c	hoch
Omloo 2008	Prognostische Kohortenstudie Rekrutierung der Patienten zwischen Oktober 2002 und August 2004. Niederlande	Beurteilung der prognostischen Güte der PET bei Patienten mit Ösophaguskarzinom, die primär operativ behandelt wurden. Darüber hinaus sollte der zusätzliche Einsatz der PET verglichen mit dem derzeit angewandten EUS untersucht werden.	Prognostische Güte von EUS und FDG-PET zur Vorhersage des krankheitsfreien Überlebens	II ^c	hoch

(Fortsetzung)

Tabelle 24: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudien beim Primärstaging (Fortsetzung)

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenzstufe ^b	Verzerrungs- potenzial
Shenfine 2009	Multizentrische prognostische Kohortenstudie Rekrutierungszeitraum zwischen November 2002 und November 2005. Australien	Bestimmung ob eine präoperative quantitative Bewertung von FDG-PET ein unabhängiger Faktor zur Vorhersage des Überlebens von Patienten mit Adeno- karzinomen der Speiseröhre ist.	Prognostische Güte von FDG- PET zur Vorhersage des krankheitsfreien Überlebens und des Gesamtüberlebens.	II ^c	hoch
<p>a: Studienziel gemäß Formulierung der Autoren b: entspricht der Evidenzstufe des G-BA c: Die Verfahrensordnung des G-BA sieht derzeit keine Evidenzstufe für prognostische Studien vor. Sie werden im Rahmen dieses Berichts daher mit Stufe 2 gekennzeichnet.</p> <p>EUS = endoskopischer Ultraschall ; FDG = ¹⁸Fluorodeoxyglukose; PET = Positronenemissionstomographie; SPECT = Single Photon Emission Computed Tomography (Einzelphotonen-Emissions-Tomographie)</p>					

Tabelle 25: Index- und Referenztest der eingeschlossenen Prognosestudien beim Primärstaging

Studie	Indextest	Referenztest
Cheze-Le Rest 2008	PET (Allegro-dedicated PET Scanner, Philips Medical System, USA) 60 Minuten nach i.v.-Injektion von 5 MBq/ kg KG FDG (Mittelwert: 355 MBq). Aufnahme von Kopf bis Oberschenkel. Befundung durch erfahrene Nuklearmediziner, die gegenüber Ergebnissen der CT, EUS und Histologie verblindet waren. Visuelle und semiquantitative Auswertung (SUV_{max}).	Gesamtüberleben, Nachbeobachtung 24 bis 52 Monate, Median: 32 Monate
Choi 2004	PET (Advance PET Scanner, GE Medical Systems, USA) 45 Minuten nach i.v.-Injektion von 370 MBq FDG. Aufnahme von Kopf bis Oberschenkel. Befundung durch 2 Nuklearmediziner, die gegenüber Ergebnissen der CT, EUS und Histologie verblindet waren. Visuelle und semiquantitative Auswertung (SUV_{max}).	Krankheitsfreies Überleben und Gesamtüberleben. Keine Angaben zur Dauer der Nachbeobachtung.
Chung 2008	PET (Advance PET Scanner oder Discovery LS PET/CT Scanner, GE Medical Systems, USA) 45 Minuten nach i.v.-Injektion von 370 MBq FDG. Befundung unabhängig durch 2 Nuklearmediziner. Für SPECT: Quantitative Analyse. Berechnung des Verhältnisses von TI-Aufnahme der „region of interest“ zu TI-Aufnahme des Mediastinums. Für PET: quantitative und visuelle Auswertung.	Krankheitsfreies Überleben. Durchschnittliche Nachbeobachtungszeit 20,3 Monate (Spannweite 0–62,7 Monate)
Omloo 2008	PET (Ecat Exact HR + PET scanner (Siemens / CTI, Knoxville, Tennessee, USA) Blutzuckerwerte wurden vor und nach PET-Untersuchung mithilfe der Hexokinase-Methode (Hitachi 747, Boehringer-Mannheim, Mannheim, Deutschland) gemessen. Scans wurden in einem 2-dimensionalen (2-D) acquisition Acquisition-Modus mit einem Emission-Übertragungs Übertragungs-Scan-Protokoll durchgeführt. Emission-Scans (ab Mitte Schädel bis Mitte Oberschenkel) von 5-minütiger Länge pro Position wurden 90 Minuten nach intravenöser Injektion von einer mittleren Dosis von 398 MBq (Bereich von 130 ± 810 MBq, je nach Körpergewicht) durchgeführt. EUS Radialscanner (GF-UM130 oder GF-UM160, 5 ± 20 MHz, Olympus Medical Systems, Tokio, Japan). EUS wurde bei Patienten in einer linken Dekubitusposition und unter Sedierung mit 2,5 ± 10 mg Midazolam intravenös durchgeführt.	Gesamtüberleben, Follow-up mindestens 25 Monate
Shenfine 2009	PET (Phillips Allegro GSO PET scanner, Phillips Medical Systems, Niederlande) 45 Minuten nach i.v.-Injektion von 210–427 MBq FDG. Läsionen mit einem $SUV_{max} > 5$ wurden als bösartig gewertet. Befundung durch Nuklearmediziner.	Nachbeobachtung 18 bis 61 Monate, Median 44 Monate
EUS = endoskopischer Ultraschall; FDG = ¹⁸ Fluordeoxyglukose; GE = General Electric; i.v. = intravenös; kg = Kilogramm; KG = Körpergewicht; MBq = Mega Becquerel; TI =Thallium		

Tabelle 26: Eingeschlossene Patienten der eingeschlossenen Prognosestudien beim Primärstaging

Studie	Patienten [n]	Geschlecht [n (m/w)]	Alter (MW, Median, Spannweite)	Diagnose	Behandlung
Cheze-Le Rest 2008	47 ^a	47 (42 / 5)	61,1; 58; 41–89	Histologie: Plattenepithelkarzinom (36), Adenokarzinom (11). Differenzierungsgrad: guter (22), mittlerer (13), niedriger (12) Lokalisation: obere (9), mittlere (18), untere Speiseröhre (20) AJCC Stadium: I (2), II (18), III (15), IV (12)	Alle Patienten mit neu diagnostiziertem Ösophaguskarzinom zum Zeitpunkt des Indextests. Nach Primärstaging OP (11) ggf. in Kombination mit Chemo- und Strahlentherapie (12) oder Chemo- und Strahlentherapie ohne OP (24).
Choi 2004	69 ^b	69 (64 / 5)	63; k. A.; 41–77	Histologie: Plattenepithelkarzinom (69). Differenzierungsgrad: guter (21), mittlerer (29), niedriger (19) Lokalisation: obere (9), mittlere (50), untere Speiseröhre (10) Stadium: Klinisch: I (10), II (26), III (23), IV (10). Pathologie: 0-I (11), II (17), III (12), IV (29)	Alle Patienten mit neu diagnostiziertem Ösophaguskarzinom zum Zeitpunkt des Indextests. Nach Primärstaging OP (36) ggf. in Kombination mit Chemo- und/ oder Strahlentherapie (13) oder OP mit adjuvanter Chemotherapie (20)
Chung 2008	100	100 (94 / 6)	65; k. A.; 42–85	Histologie: Plattenepithelkarzinom (94), Adenokarzinom (5), undifferenziert mit lymphoidem Stroma (1). Lokalisation: obere (12), mittlere (62), untere Speiseröhre (26) Stadium (AJCC 2002): 0 (2), I (18), II (27), III (30), IV (23)	Alle Patienten mit neu diagnostiziertem Ösophaguskarzinom zum Zeitpunkt des Indextests. Nach Primärstaging OP (77), neoadjuvante Chemotherapie oder keine Chirurgie wegen Fernmetastasierung (23).
Omloo 2008	125	125 (104 / 21)	63; k. A.; 37–82	Histologie: Adenokarzinom (106), Plattenepithelkarzinom (19). Lokalisation: Type I (90), Type II (35). Klinisches Stadium: I (12), IIa (38), IIb (8), III (67).	Kein Patient erhielt eine neoadjuvante oder adjuvante Therapie.

(Fortsetzung)

Tabelle 26: Eingeschlossene Patienten der eingeschlossenen Prognosestudien beim Primärstaging (Fortsetzung)

Studie	Patienten [n]	Geschlecht [n (m/w)]	Alter (MW, Median, Spannweite)	Diagnose	Behandlung
Shenfine 2009	45	45 (37 / 8)	k. A.; 72; 38–82	Histologie: Adenokarzinom (45). Differenzierungsgrad: guter oder mittlerer (21), niedriger (24). Lokalisation: unteres Drittel der Speiseröhre (24); gastroösophagealer Übergang (21).	Keine Vorbehandlung
<p>a: 52 Patienten eingeschlossen. Fünf Patienten, die innerhalb von 30 Tagen nach OP nicht krebsbedingt verstarben, wurden nicht in die Analyse eingeschlossen.</p> <p>b: 89 Patienten eingeschlossen. Ein Patient mit unvollständiger Resektion 9 Patienten, die eine OP ablehnten, und 10 nicht OP-fähige Patienten wurden ausgeschlossen.</p> <p>m = männlich; MW = Mittelwert; w = weiblich</p>					

5.3.3.2 Einschätzung des Verzerrungspotenzials der eingeschlossenen Prognosestudien beim Primärstaging

Einige Aspekte, die zur Einschätzung des Verzerrungspotenzials der Studie von **Cheze-Le Rest 2008** [75] erforderlich waren, wurden nicht berichtet, sodass eine Bewertung erschwert war. Konkrete Angaben zur Länge der Nachbeobachtungszeit wurden nicht gemacht. Die multivariate Analyse schloss neben 3 Parametern der PET nur eine weitere Variable (alleinige OP) ein, während andere Faktoren (wie z. B. Alter und Geschlecht), die möglicherweise einen Einfluss haben könnten, nicht berücksichtigt wurden. Außerdem wurde der Faktor „Tumordifferenzierung“, der in der univariaten Analyse statistisch signifikant war, nicht in die multifaktorielle Analyse eingeschlossen. Die Anzahl der in der multifaktoriellen Analyse eingeschlossenen Faktoren scheint hinsichtlich der Anzahl der Todesfälle während der Nachbeobachtung zu hoch zu sein. Zudem ist unklar, wie die Schwellenwerte für die Analyse der PET-Parameter ermittelt wurden und ob diese im Vorfeld definiert waren. Zu Studienbeginn waren die Patienten in unterschiedlichen Krankheitsstadien und wurden im Studienverlauf je nach klinischer Situation unterschiedlich behandelt. Es gab keine separaten Auswertungen nach Krankheitsstadien. Insgesamt ist von einem hohen Verzerrungspotenzial auszugehen.

In der Studie **Choi 2004** [55] fehlten Angaben zur Nachbeobachtungszeit. Die Schwellenwerte der Analysen wurden nicht angegeben. Die eingeschlossenen Patienten waren zu Studienbeginn in unterschiedlichen Krankheitsstadien. Es gab keine separaten Auswertungen nach Krankheitsstadien.

In der uni- und multifaktoriellen Analyse bei **Chung 2008** [76] wurde die Behandlung, die möglicherweise einen Einfluss haben könnte, nicht berücksichtigt, sodass ein hohes Verzerrungspotential besteht. Die Anzahl der in der multifaktoriellen Analyse eingeschlossenen Faktoren (7) scheint hinsichtlich der Gesamtzahl der eingeschlossenen Patienten (100) zu hoch zu sein. Die eingeschlossenen Patienten waren darüber hinaus zu Studienbeginn in unterschiedlichen Krankheitsstadien. Es gab keine separaten Auswertungen nach Krankheitsstadien.

In der Studie von **Omloo 2008** [59] wurde der Schwellenwert für SUV im Rahmen der Studie definiert. Die eingeschlossenen Patienten waren zu Studienbeginn in unterschiedlichen Krankheitsstadien. Es gab keine separaten Auswertungen nach Krankheitsstadien.

In der Studie von **Shenfine 2009** [61] blieb unklar, ob die Schwellenwerte für die verschiedenen eingeschlossenen Faktoren im Vorfeld definiert wurden. Zudem finden sich keine Angaben zur Verblindung der Patienten und die postoperative Behandlung erfolgte nicht standardisiert. Außerdem basiert das Model auf einer zu geringen Anzahl von Ereignissen.

Tabelle 27: Verzerrungspotenzial der eingeschlossenen Prognosestudien beim Primärstaging

Studie	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	Verzerrungspotenzial der Ergebnisse
Cheze-Le Rest 2008	+	-	-	?	+	+	+	?	+	?	+	+	+	+	-	-	-	hoch
Choi 2004	+	-	+	?	+	+	+	?	+	?	+	+	+	+	-	-	-	hoch
Chung 2008	+	-	+	+	+	+	+	-	+	?	+	+	+	-	-	+	-	hoch
Omloo 2008	+	-	+	+	+	+	+	-	+	?	+	+	?	-	?	+	+	hoch
Shenfine 2009	+	-	+	+	+	+	?	?	+	?	+	+	?	-	?	+	-	hoch

+: Kriterium erfüllt; -: Kriterium nicht erfüllt; ?: unklar

5.3.3.3 Einschätzung der Übertragbarkeit der Ergebnisse der eingeschlossenen Prognosestudien beim Primärstaging auf den deutschen Versorgungskontext

In 2 Studien konnten keine Anhaltspunkte für eine Nichtübertragbarkeit gefunden werden. Eine generelle Übertragbarkeit dieser Studien auf den deutschen Versorgungskontext wird angenommen. Bei 3 von 5 eingeschlossenen Prognosestudien konnten Anhaltspunkte gefunden werden, die eine Übertragbarkeit der Ergebnisse der Prognosestudien auf den deutschen Versorgungskontext einschränken:

- Die Bewertung der Ergebnisse der Technologien erfolgte in den Studien von Cheze-Le Rest 2008 [75], Choi 2004 [55] und Chung 2008 [76] gemeinsam bzw. auch unabhängig voneinander durch mehrere erfahrene Bewerter.
- In den Studien von Cheze-Le Rest 2008 [75] und Choi 2004 wurden die Ergebnisse der Indextests ohne Kenntnis vorhergehender klinischer Untersuchungen beurteilt. Dies führt zwar zu einer hohen internen Validität der Studienergebnisse, ist aber in der klinischen Routine eher unwahrscheinlich, was zu einer Reduktion der Übertragbarkeit führen könnte.

Genannte Aspekte entsprechen möglicherweise nicht dem klinischen Alltag, da meistens den Beurteilern der PET bzw. aller weiteren Tests die Ergebnisse vorangegangener Diagnosen bekannt sind. Darüber hinaus wurden in der Studie die Bilder von 2 Klinikern unabhängig voneinander interpretiert, was im Versorgungsalltag auch eher selten vorkommen dürfte. Aus diesen Gründen ergeben sich Anhaltspunkte dafür, dass die Ergebnisse möglicherweise nur für einen Teil der Anwendungssituationen in Deutschland gelten.

Tabelle 28: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien zum Primärstaging

Studie	F1	F2	F3	F4	F5	F6	F7	Generelle Einschätzung der Übertragbarkeit
Cheze-Le Rest 2008	-	-	+	+	-	-	-	Anhaltspunkt für eine Nichtübertragbarkeit
Choi 2004	-	-	+	+	-	-	-	Anhaltspunkt für eine Nichtübertragbarkeit
Chung 2008	-	-	+	-	-	-	-	Anhaltspunkt für eine Nichtübertragbarkeit
Omloo 2008	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Shenfine 2009	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
+: konkrete Anhaltspunkte für eine Nichtübertragbarkeit liegen vor								
-: keine Anhaltspunkte für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden)								

5.3.3.4 Ergebnisse zur prognostischen Güte der eingeschlossenen Primärstudien beim Primärstaging

In Tabelle 29 werden die Ergebnisse der einfaktoriellen und multifaktoriellen Analyse der eingeschlossenen Prognosestudien dargestellt.

Von 5 eingeschlossenen Prognosestudien hatten 4 Studien das Ziel, das Gesamtüberleben zu untersuchen. Zwei dieser 4 Studien untersuchten zusätzlich noch das krankheitsfreie Überleben. Eine Studie hatte ausschließlich das Ziel, das krankheitsfreie Überleben zu untersuchen. In einer Studie wurde ein direkter Vergleich zwischen FDG-PET und Ti-SPECT durchgeführt. Es wurde eine Vielzahl an prognostischen Merkmalen untersucht, wobei in der folgenden Tabelle nur die PET-relevanten Merkmale dargestellt werden. Alle weiteren Merkmale werden im Anhang E präsentiert. Von allen prognostischen Merkmalen der 5 eingeschlossenen Studien blieben 3 PET-relevante Merkmale nach multiplem Testen statistisch signifikant.

Tabelle 29: Ergebnisse der Primärstudien zur prognostischen Güte beim Primärstaging

Studie	Zielgröße	Anzahl ausgewerteter Patienten	Prognostisches Merkmal	Cut-off zum prognostischen Merkmal	Patienten mit prognostischen Merkmal / alle Patienten	Ergebnisse aus einfaktorierter Analyse		Ergebnisse aus multifaktorieller Analyse					
						p-Wert	HR	p-Wert	HR				
Chez-Le Rest 2008	Gesamtüberleben	47	SUVmax im Primärtumor	> 9 ^a	k. A.	0,02	k. A.	0,01	3,16				
			Lymphknoten mit erhöhter FDG-Aufnahme	≥ 1	k. A.					0,004	k. A.	0,05	2,43
				≥ 2	k. A.					k. A.	k. A.	0,01	3,16
			Anzahl abnormaler FDG foci	1	22/47					0,006	k. A.	k. A.	k. A.
				mehr als 1	25/47								
	FDG positiv bzgl. nodal involvement	k. A.	k. A.	0,03	k. A.	k. A.	k. A.						
Choi 2004	Krankheitsfreies Überleben	69	Anzahl Lymphknoten mit erhöhter FDG-Aufnahme	0	26/69	< 0,001	k. A.	< 0,001	1,87				
				1	16/69								
				2	16/69								
				≥ 3	11/69								
			Tumurlänge in der PET	< 3 cm	14/69	0,018	k. A.	k. A.	k. A.				
				3 bis 5 cm	26/69								
				≥ 5 cm	29/69								
			PET Stadium	N0 M0	26/69,	< 0,001	k. A.	k. A.	k. A.				
				N1 M0	28/69								
				M1	15/69								

(Fortsetzung)

Tabelle 29: Ergebnisse der Primärstudien zur prognostischen Güte beim Primärstaging (Fortsetzung)

Studie	Zielgröße	Anzahl ausgewerteter Patienten	Prognostisches Merkmal	Cut-off zum prognostischen Merkmal	Patienten mit prognostischen Merkmal / alle Patienten	Ergebnisse aus einfaktorierter Analyse		Ergebnisse aus multifaktorieller Analyse	
						p-Wert	HR	p-Wert	HR
Choi 2004 (Fortsetzung)	Gesamtüberleben		Anzahl Lymphknoten mit erhöhter FDG-Aufnahme	0	26/69	< 0,001	k. A.	0,024	1,71
				1	16/69				
				2	16/69				
				≥ 3	11/69				
			Tumurlänge in der PET	< 3 cm	14/69,	0,001	k. A.	0,010	2,74
				3 bis 5 cm	26/69				
				≥ 5 cm	29/69				
			Maximum SUV des Primärtumors	< 6,3	14/69	0,035	k. A.	k. A.	k. A.
				6,3 bis 13,7	27/69				
				≥ 13,7	28/69				
			PET-Stadium	N0 M0	26/69	0,001	k. A.	k. A.	k. A.
				N1 M0	28/69				
M1	15/69								
Chung 2008	Krankheitsfreies Überleben	100	Visualisierung durch FDG-PET	Visualisierung	91/100	0,01	k. A.	0,97	k. A.
			Höchster SUV der FDG-PET	≥ 15	55/100	0,003	k. A.	0,78	2,648 ^b
			Visualisierung durch TI-SPECT	Visualisierung	85/100	0,01	k. A.	0,72	4,88 ⁱ
			Verhältnis der TI-Aufnahme von Tumor zu Mediastinum in der SPECT nach 4 Stunden	≥ 2,3	58/100	0,04	k. A.	0,66	1,98 ⁱ

(Fortsetzung)

Tabelle 29: Ergebnisse der Primärstudien zur prognostischen Güte beim Primärstaging (Fortsetzung)

Studie	Zielgröße	Anzahl ausgewerteter Patienten	Prognostisches Merkmal	Cut-off zum prognostischen Merkmal	Patienten mit prognostischen Merkmal / alle Patienten	Ergebnisse aus einfaktorierter Analyse		Ergebnisse aus multifaktorieller Analyse	
						p-Wert	HR	p-Wert	HR
Omloo 2008	Gesamtüberleben	125	SUV niedrig versus hoch, cut-off: Median	k.A.	k.A.	0,04	Unklar ^c	0,6	Unklar ^c
Shenfine 2009	Gesamtüberleben	45	SUV _{max}	≤ 5	19/45	0,006	k. A.	Unklar ^c	Unklar ^c
				> 5	26/45	0,007	k. A.	Unklar ^c	Unklar ^c
	Krankheitsfreies Überleben	45	FDG PET qualitatives Ergebnis	Negative Lymphknoten	37/45	0,02	k. A.	> 0,05	k. A.
				Positive Lymphknoten	8/45				
Krankheitsfreies Überleben	45	SUV _{max}	≤ 5	19/45	0,013	k.A.	Unklar ^c	Unklar ^c	
			> 5	26/45	0,015	k.A.	Unklar ^c	Unklar ^c	

a: In Tabelle 3 „< 9“, in Tabelle 5 „> 9“.
 b: Korrespondenz mit den Autoren der Studie
 c: Bei den Angaben zum multifaktoriellen Modell bleibt unklar, auf welche Zielgröße sich die Autoren beziehen. Sie geben für SUVmax einen HR von 0,37 und einen p-Wert von 0,08 an.
 k. A.=keine Angaben

Direkte Vergleiche zwischen PET bzw. PET/CT und anderen bildgebenden Verfahren

Ein direkter Vergleich zwischen FDG-PET und ^{201}Tl SPECT wurde in der Studie von Chung 2008 [76] durchgeführt. Alle PET- sowie auch die SPECT-relevanten Merkmale waren nach multiplem Testen nicht mehr statistisch signifikant. Die Autoren der Studien wiesen selbst darauf hin, dass die Zunahme der ^{201}Tl mit dem FDG SUV korreliert, weil sich beide bildgebenden Verfahren ähneln. Damit tritt das Problem der Multikollinearität auf.

Subgruppenmerkmale und andere Effektmodifikatoren

Es wurden keine Subgruppenmerkmale in den eingeschlossenen Studien untersucht.

5.4 Teilziel 2b: diagnostische und prognostische Güte beim Restaging

5.4.1 Diagnostische und prognostische Güte beim Restaging auf Grundlage in die Bewertung eingeflossener Evidenzsynthesen

Für die Fragestellung konnten 3 Evidenzsynthesen eingeschlossen werden (Ngamruengphong 2010 [53], Rebollo Aquirre 2009 [62] und Westerterp 2005 [63]).

5.4.1.1 Studiendesigns und Studienpopulationen der eingeschlossenen Evidenzsynthesen beim Restaging

Die Evidenzsynthese von Ngamruengphong 2010 [53] hatte das Ziel, die FDG-PET und die EUS-Technologie im Hinblick auf das Ansprechen einer neoadjuvanten Therapie bei Patienten mit Ösophaguskarzinomen zu untersuchen. Sie beschränkte sich dabei auf die diagnostische und prognostische Güte der Technologien. Der Einschluss umfasste die Zeit von Januar 1980 bis Februar 2008. Als Indextest wurde die PET oder PET/CT mit dem Tracer FDG und der EUS betrachtet. Die pathologische Untersuchung wurde als Referenztest herangezogen.

Die Evidenzsynthese von Rebollo Aquirre 2009 [62] hatte das Ziel, das Behandlungsansprechen durch den Einsatz der PET bzw. PET/CT bei Patienten mit Ösophaguskarzinomen zu untersuchen. Die Recherche beschränkte sich auf den Zeitraum von Januar 1999 bis August 2006. Eine pathologische Untersuchung, andere bildgebende Verfahren und eine klinische Nachbeobachtung von über einem Jahr wurden als Referenztests verwendet. Allerdings wurde nur die Studien in dem vorliegenden Bericht eingeschlossen, in denen die Histologie und / oder ggf. Nachbeobachtung als Referenztests verwendet wurden. Bildgebende Verfahren, die außerhalb der Nachbeobachtung angewendet wurden, wurden im Vorbericht nicht als Referenztest akzeptiert.

Die Evidenzsynthese von Westerterp 2005 [63] verfolgte das Ziel, die FDG-PET, die CT und die EUS im Hinblick auf Behandlungsansprechen bei Patienten mit Ösophaguskarzinomen zu untersuchen. Es wurde von Januar 1980 bis Januar 2004 recherchiert. Eingeschlossene Studien mussten 10 oder mehr Patienten umfassen. Eine pathologische Untersuchung wurde als Referenztest angesehen.

Die Fragestellungen von Ngamruengphong 2010 und Westerterp 2005 gehen über die des vorliegenden Berichts hinaus, da sie systematisch auch nach Studien suchten, die nur die CT oder den EUS untersuchten, ohne einen direkten Vergleich zur PET oder PET/CT anzustellen. Darüber hinaus waren die Einschlusskriterien der 3 Evidenzsynthesen weiter gefasst als im vorliegenden Bericht, da beispielsweise auch retrospektive Studien eingeschlossen wurden. Die Fragestellungen und Methodik der einzelnen Evidenzsynthesen werden in Tabelle 30 zusammengefasst.

Tabelle 30: Fragestellung und Methodik der Evidenzsynthesen zur diagnostischen und prognostischen Güte der PET bzw. PET/CT beim Restaging

Evidenzsynthese / Fragestellung	Literaturrecherche	Ein- / Ausschlusskriterien, Studienselektion	Studienbewertung	Informationssynthese
<p>Ngamruengphong 2010 PET und PET/CT bei Patienten mit Ösophaguskarzinom</p>	<p><u>Quellen:</u> MEDLINE, EMBASE, Cochrane Library, <u>Suchstrategie:</u> selbst entwickelte Suchstrategie, <u>Zeitraum:</u> Jan 1980 bis Feb 2008</p>	<p><u>Einschlusskriterien:</u> auf Englisch publizierte Primärstudien, PET oder EUS zum Therapieansprechen vor und nach Therapie, histologisch nachgewiesenes Ösophaguskarzinom, Histologie als Referenztest, min. 10 Patienten. <u>Ausschlusskriterien:</u> ungenügende Daten um eine Vierfeldertafel zu berechnen, Daten wurden nicht auf Patientenbasis erhoben, Publikationen, die sich auf identische Patienten beziehen. <u>Selektion:</u> k. A. zum Screeningprozess, 2 Personen reviewten unabhängig voneinander, ob relevante Studien den Einschlusskriterien entsprachen.</p>	<p>Nach den Kriterien des QUADAS-Instruments</p>	<p>-</p>
<p>Rebollo Aquirre 2009 PET bzw. PET/CT bei Patienten mit Ösophaguskarzinomen</p>	<p><u>Quellen:</u> MEDLINE, EMBASE, Cochrane Library, ClinicalTrial.gov, the National Research Register of the National Health Service, the Centre for Reviews and Dissemination databases (International Network of Agencies for Health Technology Assessment, Database of Abstracts of Reviews of Effects and NHS Economic Evaluation Database). Handsuche <u>Suchstrategie:</u> selbst entwickelte Suchstrategie, <u>Zeitraum:</u> Jan 1999 bis Aug 2006</p>	<p><u>Einschlusskriterien:</u> Patienten mit nachgewiesenem Ösophaguskarzinom, Durchführung einer neoadjuvanten Therapie möglich, PET oder PET/CT, als Referenztests Histologie, andere bildgebende Verfahren und klinische Nachbeobachtung > 1 Jahr. <u>Ausschlusskriterien:</u> Extraktion der Daten der einzelnen Tests nicht möglich, kein Originalartikel, Kongressabstracts, Editorials, Comments, Leserbriefe, Review Artikel und Fallberichte, < 10 Patienten, Publikationen mit unzureichenden Informationen, um den genauen Einschluss zu bestimmen. <u>Selektion:</u> k. A.</p>	<p>Nach den Kriterien des QUADAS-Instruments</p>	<p>-</p>

(Fortsetzung)

Tabelle 30: Fragestellung und Methodik der Evidenzsynthesen zur diagnostischen und prognostischen Güte der PET bzw. PET/CT beim Restaging (Fortsetzung)

Evidenzsynthese / Fragestellung	Literaturrecherche	Ein- / Ausschlusskriterien, Studienselektion	Studienbewertung	Informations- synthese
Westerterp 2005 FDG-PET, CT und EUS bei Ösophagus- karzinomen	<u>Quellen:</u> MEDLINE, EMBASE, Cochrane Library <u>Suchstrategie:</u> selbst entwickelte Suchstrategie <u>Zeitraum:</u> Jan 1980 bis Jan 2004	<u>Einschlusskriterien:</u> > 10 Patienten, Histologie als Referenztest, Untersuchung von PET, CT und EUS zum Therapieansprechen vor und nach Therapie. <u>Ausschlusskriterien:</u> Publikationen, die sich auf identische Patienten beziehen. <u>Selektion:</u> Titel- und Abstractscreening und Volltextscreening wurden von 2 Personen unabhängig durchgeführt.	Nach den Kriterien der Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests. Die Kriterien wurden teilweise modifiziert.	-

5.4.1.2 Einschätzung des Verzerrungspotenzials der eingeschlossenen Evidenzsynthesen

Zwei der 3 eingeschlossenen Evidenzsynthesen erreichten einen Punktwert von 7 auf dem Qualitätsindex von Oxman und Guyatt. Die Evidenzsynthese von Rebollo Aquirre 2009 erreichte einen Punktwert von 5. Die methodische Schwachstelle der Evidenzsynthese von Rebollo Aquirre 2009 lag in der Selektion relevanter Studien aus den Recherchen. Der Prozess der Selektion wurde in der Publikation nicht beschrieben. Dies weist auf einen potenziellen Selektionsbias hin. Für die Indikation Restaging lagen 2 weitere qualitativ hochwertige Evidenzsynthesen (Ngamruengphong 2010 und Westerterp 2005) vor, in welchen das Literaturscreening von 2 unabhängigen Reviewern durchgeführt wurde und den Zeitraum bis 2008 abdeckte. Es gab daher keine Anhaltspunkte für einen Selektionsbias. Da die Evidenzsynthese von Rebollo Aguirre 2009 ansonsten alle Kriterien des Qualitätsindex erfüllte, wurde sie in die Bewertung des vorliegenden Berichts eingeschlossen.

Tabelle 31: Ergebnisse der Qualitätsbewertung der eingeschlossenen Evidenzsynthesen nach Oxman und Guyatt [38,39]

Evidenzsynthesen	OG1	OG2	OG3	OG4	OG5	OG6	OG7	OG8	OG9	Total
Ngamruengphong 2010	+	+	+	+	+	+	+	+	+	7
Westerterp 2005	+	+	+	+	+	+	+	+	+	7
Rebollo Aguirre 2009	+	+	+	-	+	+	+	+	+	5
OG1 bis OG9: Qualitätskriterien nach Oxman und Guyatt; +: Kriterium erfüllt; -: Kriterium nicht erfüllt										

In den Evidenzsynthesen von Rebollo Aquirre 2009 und Ngamruengphong 2010 wurde angegeben, dass kein Interessenkonflikt vorhanden war. In Westerterp 2005 fanden sich keine Angaben zu den Interessenkonflikten der Autoren.

5.4.1.3 Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien aus den Evidenzsynthesen

Aus den 3 Evidenzsynthesen entsprachen 11 Primärstudien zur diagnostischen und prognostischen Güte der PET bzw. PET/CT den Einschlusskriterien des vorliegenden Berichts in der Indikation des Restagings [54,65-74].

Die Qualitätsbewertung wurde in den Evidenzsynthesen von Rebollo Aquirre 2009 und Ngamruengphong 2010 anhand des QUADAS-Instruments durchgeführt. Beide Evidenzsynthesen stellten ihre Bewertung pro Item in Form einer Tabelle und eines Scores dar. Die Autoren vergaben für jedes erfüllte Item einen Punkt, sodass eine Studie am Ende 14 Punkte erreichen konnte. Die Qualitätsbewertung in der Evidenzsynthese von Westerterp 2005 wurde mithilfe der Kriterien der Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests durchgeführt. Das Verzerrungspotenzial wurde

anhand 7 Items analysiert. Für jedes erfüllte Item wurde ein Punkt vergeben, sodass eine Studie maximal 7 Punkte erhalten konnte.

Ngamruengphong 2010 definieren nicht, ab wann für sie Studien mit einem hohen Verzerrungspotenzial bewertet wurden. Sie gaben an, dass keine der Studien weniger als 10 Punkte erreichte. Die 11 Studien, die im vorliegenden Bericht eingeschlossen wurden, erzielten bei Ngamruengphong 2010 10 bis 13 Punkte [54,65-71,73,74]. Die Schwächen der Studien lagen darin, dass a) partielle Verifikation nicht vermieden wurde, b) die Ergebnisse des Referenztests nicht ohne Kenntnisse der Resultate des Indextests interpretiert wurden und c) nichtinterpretierbare Werte nicht genannt wurden.

Ab wann eine Studie als verzerrt eingestuft wurde, wurde in der systematischen Übersicht von Rebollo Aquirre 2009 nicht definiert. Fünf Studien aus der Evidenzsynthese von Rebollo Aquirre 2009 wurden in den vorliegenden Bericht eingeschlossen [65,69,71,73,74]. Alle 5 Studien waren auch in Ngamruengphong 2010 enthalten. Die 5 Studien aus Rebollo Aquirre 2009 erreichten Scores von 8 bis 11 Punkten. Die Schwächen der Studien lagen darin, dass a) die Selektionskriterien der Patienten nicht beschrieben wurden, b) partielle Verifikation nicht vermieden wurde, d) die Ergebnisse des Referenztests nicht ohne Kenntnisse der Resultate des Indextests interpretiert wurden, e) die Ergebnisse des Indextests nicht ohne Kenntnisse der Resultate des Referenztests interpretiert wurden und c) nichtinterpretierbare Werte nicht genannt wurden.

Vier Studien aus Westerterp 2005 entsprachen den Einschlusskriterien des vorliegenden Berichts. Drei der Studien waren auch in den Evidenzsynthesen von Rebollo Aquirre 2009 und Ngamruengphong 2010 [54,66,68] eingeschlossen. Die Studie von Weber 2001 war nur in Westerterp 2005 enthalten. Drei der untersuchten Studien [54,66,72] erhielten 6 von 7 möglichen Punkten, die Studie von Kroep 2003 [68] wurde mit der Maximalpunktzahl bewertet. Die Schwächen der 3 Studien lagen darin, dass die Ergebnisse des Referenztests nicht ohne Kenntnisse der Resultate des Indextests interpretiert wurden, und dass die Interpretation der Ergebnisse nicht ohne klinische Hintergrundinformation erfolgte.

Die Qualitätsbewertungen fielen in den 3 unterschiedlichen Evidenzsynthesen vergleichbar aus, wobei die Bewertungen von Rebollo Aquirre 2009 etwas kritischer waren als die der anderen beiden Autoren.

5.4.1.4 Ergebnisse zur diagnostischen Güte auf Grundlage der eingeschlossenen Evidenzsynthesen

In Tabelle 32 werden die Ergebnisse zur diagnostischen Güte auf Basis der extrahierten Evidenzsynthesen dargestellt. Bezüglich der Studie von Song 2005 [71] fanden sich in den Evidenzsynthesen von Rebollo Aquirre 2009 und Ngamruengphong 2010 diskrepante Angaben zu den Daten der Vierfeldertafel, sodass in diesem Fall die Primärstudie herangezogen wurde. Neben den jeweiligen Schätzern für Sensitivität und Spezifität werden in Tabelle 32 auch die Konfidenzintervalle dargestellt. Sofern diese nicht in der

Evidenzsynthese berichtet wurde, wurden sie im Rahmen des vorliegenden Berichts nach der Methode von Clopper und Pearson 1934 [109] berechnet.

6 der aus den Evidenzsynthesen stammenden Studien untersuchten die PET bzw. PET/CT in der Indikation des Interimsstaging. Die Studien von Kroep 2003, Ott 2006 und Weber 2001 [60,68,72] untersuchten den Einsatz der PET während einer neoadjuvanten Chemotherapie. Hier zeigte sich für die PET eine Spannweite der Punktschätzer der Sensitivitäten von 80 % (bei einer zugehörigen Spezifität von 78 % bzw. 100 %) bis 88 % (zugehörige Spezifität 75 %). Für die Punktschätzer der Spezifitäten ergab sich eine Spannweite von 75 % (zugehörige Sensitivität 88 %) bis 100 % (zugehörige Sensitivität 80 %). Für alle Punktschätzer waren weite Konfidenzintervalle zu beobachten. Die Studien von Gillham 2006, Westerterp 2006 und Wieder 2004 [67,73,74] untersuchten den Einsatz der PET bzw. PET/CT während einer neoadjuvanten Chemoradiotherapie. Allerdings konnten für die Studien von Westerterp 2006 und Wieder 2004 die Ergebnisse nicht dargestellt werden, weil der Anteil fehlender Werte größer als 30 % war. In der Studie von Gillham 2006 ergab sich eine Sensitivität von 44 % und eine Spezifität von 52 %. Auch hier waren die Konfidenzintervalle weit.

In 6 Studien wurde die diagnostische Güte der PET beim Restaging (nach Abschluss der Therapie) untersucht. In einer Studie wurden die Ergebnisse nicht dargestellt, weil der Anteil fehlender Werte größer als 30 % war. Auf Basis von 5 Studien zeigte die PET eine Spannweite der Punktschätzer der Sensitivitäten von 27 % (zugehörige Spezifität 95 %) bis 100 % (zugehörige Spezifität 55 % bzw. 100 %). Für die Punktschätzer der Spezifität ergab sich entsprechend eine Spannweite von 55 % (zugehörige Sensitivität 100 %) bis 100 % (zugehörige Sensitivität 100 %). In 2 Studien hatte die PET/CT eine geschätzte Sensitivität von 44 % und 86,7 % und eine geschätzte Spezifität von 52 % und 87,9 %. Für beide Technologien sind zum Teil sehr breite Konfidenzintervalle zu beobachten.

Tabelle 32: Ergebnisse der Studien aus den Evidenzsynthesen zum Restaging

Studie	Anzahl Patienten	Index-test	Popula-tion	Cut-Off	Zeitpunkt Response-Monitoring	RP	FN	FP	RN	Sensitivität % (95 %- KI)	Spezifität % (95 %- KI)
Brücher 2001	24 ^a	PET	AJCC II und III	Reduction des FDG-PET uptake 52 %	Nach Therapie	13	0	5	6	100 (75,3 bis 100) ^b	55 (23,4 bis 83,3) ^b
Cerfolio 2005	48 ^c	PET/CT	AJCC II und IV	SUV \geq 2,5	Nach Therapie	13	2	4	29	86,7 (59,5 bis 98,3) ^b	87,9 (71,8 bis 96,6) ^b
	48 ^c	EUS	AJCC II und IV	k. A.	Nach Therapie	3	12	2	31	20 (4,3 bis 48,1) ^b	94 (79,8 bis 99,3) ^b
Flamen 2002	36	PET	AJCC III	Reduktion der Tumor-to-liver uptake ratio > 80 %	Nach Therapie	10	4	4	18	71 (41,9 bis 91,6) ^b	82 (59,7 bis 94,8) ^b
Gillham 2006	32	PET und integrierte PET/CT ^e	AJCC II und III	Reduktion der SUV > 20 %	Eine Woche nach Beginn der Therapie (Interimsstaging)	4	5	11	12	44 (13,7 bis 78,8) ^b	52 (30,6 bis 73,2) ^b
Kroep 2003	11 ^f	PET	AJCC III	SUV > 40 %	Während Therapie (Interimsstaging)	4	1	0	6	80 ^{bg} (28,4 bis 99,5) ^b	100 ^g (54,1 bis 100) ^b
	10 ^f			SUV > 60 %	Nach Therapie	4	0	0	6	100 (39,8 bis 100) ^b	100 (54,1 bis 100) ^b
	12 ^f	EUS		k. A.	Beide Zeitpunkte	5	0	0	7	100 (47,8 bis 100) ^b	100 (59,0 bis 100) ^b
	11 ^f	CT		k. A.	k. A.	2	2	2	5	50 (6,8 bis 93,2) ^b	71 (29,0 bis 96,3) ^b
Levine 2006 ^h	41	PET	AJCC I und IV	SUV: 4, Post-ChT + RT	Nach Therapie	- ^h	- ^h				

(Fortsetzung)

Tabelle 32: Ergebnisse der Studien aus den Evidenzsynthesen zum Restaging (Fortsetzung)

Studie	Anzahl Patienten	Index-test	Population	Cut-Off	Zeitpunkt Response-Monitoring	RP	FN	FP	RN	Sensitivität % (95 %- KI)	Spezifität % (95 %- KI)
Ott 2006	56 ⁱ	PET	AJCC II und III	≥ 35 % nach nach 2-wöchiger Chemotherapie	2 Wochen nach Beginn der Therapie (Interimsstaging)	8	2	10	36	80 (44,4 bis 97,5) ^b	78 (63,6 bis 89,1) ^b
Song 2005	32	PET	AJCC II und III	Maximale SUV < 2,5	Nach Therapie	3 ^a	8 ^a	1 ^a	20 ^a	27 ^d (6,0 bis 61,0) ^b	95 ^d (76,2 bis 99,9) ^b
Weber 2001	37 ^j	PET	AJCC III	k. A.	2 Wochen nach Beginn der Therapie (Interimsstaging)	8	1	7	21	88 ^b (51,8 bis 99,7) ^{a,b}	75 ^b (55,1 bis 89,3) ^b
Westerterp 2006 ^h	26	PET	AJCC II und III	SUV: ≥ 31 %, 2 Wochen Behandlung	Während Therapie (Interimsstaging)	_{-h}	_{-h}	_{-h}	_{-h}	_{-h}	_{-h}
Wieder 2004	23 ^h	PET	AJCC II und III	SUV: ≥ 30 %, 2 Wochen Behandlung	Während Therapie (Interimsstaging)	_{-h}	_{-h}	_{-h}	_{-h}	_{-h}	_{-h}
	33	PET	AJCC II und III	SUV: 52 %, präoperativer PET scan	Nach Therapie	17	2	6	8	89 (66,9 bis 98,7) ^b	57 (28,9 bis 82,3) ^b
<p>a: Es wurden ursprünglich 27 Patienten in die Studie eingeschlossen. Daten liegen nur von 24 Patienten vor.</p> <p>b: Eigene Berechnung.</p> <p>c: Cerfolio 2005 schlossen insgesamt 67 Patienten in ihre Studien ein. Für die EUS und für die PET liegen allerdings nur Daten von 48 Patienten vor.</p> <p>d: Daten kommen aus der Primärstudie.</p> <p>e: Die Studie von Gillham 2006 untersuchte die PET-Technologie alleine und die integrierte PET/CT. Die Ergebnisse für die einzelnen Gruppen wurden nicht separat dargestellt.</p> <p>f: Es wurden ursprünglich 13 Patienten in die Studien eingeschlossen.</p> <p>g: Die Angaben für Sensitivität und Spezifität wurden von den Autoren Ngamruengphong et al. 2009 vertauscht.</p> <p>h: Die Ergebnisse wurden nicht dargestellt, weil der Anteil fehlender Werte größer als 30 % war.</p> <p>i: Ott 2006 schlossen insgesamt 65 Patienten in ihre Studien ein. Es liegen allerdings nur Daten von 56 Patienten vor.</p> <p>j: Weber 2001 schlossen insgesamt 40 Patienten in ihre Studien ein. Es liegen allerdings nur Daten von 37 Patienten vor.</p>											

5.4.2 Diagnostische Güte beim Restaging auf Grundlage von Primärstudien (Ergänzungsrecherche)

5.4.2.1 Studiendesigns und Studienpopulationen der eingeschlossenen Primärstudien beim Restaging

Im Rahmen der Ergänzungsrecherche wurden 2 weitere Diagnosestudien für die Indikation des Restagings identifiziert (Higuchi 2008 [103] und Roedl 2009a [104]).

Higuchi 2008 [103]

Die Studie untersuchte die diagnostische Güte der PET/CT beim Restaging und bei der Rezidivdiagnostik. Die Ergebnisse der Studie konnten im vorliegenden Bericht allerdings nur für die Indikation Restaging verwendet werden, da für die Ergebnisse der Rezidivdiagnostik keine validen Angaben zum Referenztest gemacht wurden. Die Studie untersuchte zum einen die diagnostische Güte der FDG-PET für das Behandlungsansprechen und zum anderen die diagnostische Güte der FDG-PET zur Vorhersage des Überlebens nach OP. Damit handelte es sich bei der Arbeit zum einen um eine Diagnosestudie und zum anderen auch um eine Prognosestudie, die als Assoziationsmaß die Sensitivität und Spezifität verwendete („prognostic accuracy study“). An dieser Stelle werden die Ergebnisse der diagnostischen Güte der PET bezüglich des Behandlungsansprechens dargestellt. Es wurden 50 Patienten mit lokal fortgeschrittenen Plattenepithelkarzinomen der Speiseröhre eingeschlossen. Bei allen Patienten schloss sich eine chirurgische Resektion an eine neoadjuvante Chemo- oder Chemoradiotherapie an.

Roedl 2009a [104]

In die Studie von Roedl 2009a wurden 49 Patienten mit lokal fortgeschrittenen Plattenepithelkarzinomen der Speiseröhre eingeschlossen. Ziel der Arbeit war die Bewertung der diagnostischen Güte der PET und PET/CT zur Vorhersage des Ansprechens auf neoadjuvante Chemoradiotherapie und der Vorhersage des krankheitsfreien Überlebens bei Patienten mit Speiseröhrenkarzinomen. Die Autoren der Studie haben zwar ein prognostisches Modell berechnet, es fehlten aber Angaben darüber, ob eine adäquate Adjustierung für Störgrößen durchgeführt wurde. Deshalb werden die Ergebnisse für das krankheitsfreie Überleben nicht im Bericht dargestellt.

Die Patienten wurden vor Aufnahme und nach Abschluss der Behandlung mit FDG-PET bzw. PET/CT untersucht, wobei verschiedene Parameter erhoben wurden. Bei allen Patienten wurde das Behandlungsansprechen histologisch untersucht.

Tabelle 33: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche)

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenzstufe ^b	Verzerrungspotenzial
Higuchi 2008	Diagnostische Kohortenstudie Rekrutierung der Patienten zwischen März 2000 und August 2004. Japan	Bewertung der diagnostischen Genauigkeit der FDG-PET bei der Bestimmung des Behandlungsansprechens von Patienten mit Plattenepithelkarzinomen.	Diagnostische Genauigkeit von FDG-PET bei der Bestimmung des Behandlungsansprechens.	I Ib	hoch
Roedl 2009a	Diagnostische Kohortenstudie Keine Angaben zum Rekrutierungs- und Untersuchungszeitraum. USA	Untersuchung von metabolischen Markern zur Bestimmung der Tumorausdehnung zur Vorhersage des Ansprechens auf Chemoradiotherapie und des krankheits- freien Überlebens von Patienten mit Ösophaguskarzinomen.	Diagnostische Genauigkeit von FDG-PET bzw. PET/CT zur Bestimmung des Ansprechens auf Chemoradiotherapie. Angabe der Ergebnisse verschiedener Parameter.	I Ib	hoch
<p>a: Studienziel gemäß Formulierung der Autoren b: entspricht der Evidenzstufe des G-BA FDG = ¹⁸Fluorodeoxyglukose; PET = Positronenemissionstomographie; USA = Vereinigte Staaten von Amerika</p>					

Tabelle 34: Index- und Referenztest der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche)

Studie	Indextest	Referenztest
Higuchi 2008	PET (HEADTOME/SET 2400 W, Shimadzu Co, Japan) 60 Minuten nach i.v.-Injektion von ca. 370 MBq FDG. Läsionen mit $SUV_{max} \geq 2,5$ nach Behandlung wurden als bösartig gewertet. Keine Angaben zur Befundung.	Histologie Positives Ansprechen \geq Grad 2 (Grad 0: kein, Grad 1: weniger als 1/3, Grad 2: mehr als 2/3, Grad 3: komplettes Verschwinden der Tumorzellen) ^a . Keine weiteren Angaben.
Roedl 2009a	PET/CT (Biograph Sensation 16 scanner, Siemens, Erlangen, Germany) ca. 60 Minuten nach i.v.-Injektion von 15 mCi (555 MBq) vor Beginn und nach Ende der Chemoradiotherapie. Tumoransprechen durch Abnahme von Tumorgröße, Tumordurchmesser, SUV des Tumors, Index (Tumordurchmesser x SUV des Tumors) und SUV_{max} . Beurteilung unabhängig durch 2 Untersucher (jeweils 4 Jahre Erfahrung in PET/CT-Befundung). Die Befunder waren gegenüber Diagnose und klinischen Angaben der Patienten verblindet.	Histologie Tumoransprechen < 10 % wachstumsfähige Tumorzellen. Keine weiteren Angaben.
a: Vermutlich wurde die Klassifikation durch die Autoren nicht richtig dargestellt, da unklar ist, welchem Grad eine Reduktion von mehr als 1/3, aber weniger als 2/3 entspricht. FDG = ¹⁸ Fluorodeoxyglukose; MBq = Megabecquerel; mCi = Millicurie; PET = Positronenemissionstomographie; SUV = standardized uptake value		

Tabelle 35: Eingeschlossene Patienten der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche)

Studie	Patienten [n]	Geschlecht [n (m/w)]	Alter	Diagnose	Behandlung
Higuchi 2008	50	50 (41/9)	k. A.; 62,2; 44–77	Histologie: Plattenepithelkarzinom (50). Lokalisation: zervikale (11), obere (5), mittlere (17), untere Speiseröhre (17). Klinisches Stadium (vor Behandlung): I (0), IIA (0), IIB (0), III (24), IVA (11), IVB (15) Resektabilität: R0 (48), R1 (0), R2 (2)	Nach Studienbeginn: 35 Chemo- und 15 Chemoradiotherapie. 11 zervikale, 39 subtotale Ösophagektomie.
Roedl 2009a	49	49 (39/10)	68,9; k. A.; k. A.	Histologie: Plattenepithelkarzinom (49). Lokalisation: obere (23), mittlere (22), untere Speiseröhre (4). Tumorstadium (vor Behandlung): T2N0M0 (8), T2N1M0 (8), T3N0M0 (10), T3N1M0 (23).	Alle Patienten nach Studienbeginn: neoadjuvante Chemoradiotherapie (50,4 Gy, Cisplatin und 5-Fluoruracil nach 1 und 5 Woche(n) nach Beginn).
Gy = Gray; OP = Operation; LK = Lymphknoten; m = männlich; T = Tumor; w = weiblich					

5.4.2.2 Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien beim Restaging (Ergänzungsrecherche)

In der folgenden Tabelle werden die Ergebnisse der Einschätzung des Verzerrungspotenzials der eingeschlossenen Studien beim Restaging dargestellt. Beide Studien hatten ein hohes Verzerrungspotenzial.

Tabelle 36: Verzerrungspotenzial der eingeschlossenen Studien beim Restaging (Ergänzungsrecherche)

Studie	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Verzerrungspotenzial der Ergebnisse
Higuchi 2008	+	+	+	+	?	+	?	-	+	?	hoch
Roedl 2009a	+	?	+	+	?	+	?	+	-	?	hoch
+: Kriterium erfüllt; -: Kriterium nicht erfüllt; ?: unklar											

In der Studie von **Higuchi 2008** [103] wurden keine genauen Angaben zur Verblindung bei der Bewertung von Index- und Referenztest gemacht und es war unklar, wie viel Zeit zwischen Bildgebung und Referenztest lag. Darüber hinaus wurden keine genauen Angaben zur Verblindung bei Bewertung des Index- und Referenztests gemacht. In der Studie von **Roedl 2009a** [104] wurde eine Reihe verschiedener Referenztests verwendet. Es wurden keine genauen Angaben zur Verblindung bei Bewertung des Index- und Referenztests gemacht. Angaben fehlten, ob Patienten zwischen Indextest und Referenztest oder während der Nachbeobachtung behandelt wurden. Darüber hinaus lag in der Studie eine ergebnisgesteuerte Berichterstattung vor, weil der optimale Trennwert und die Kombination der PET und PET/CT-Parameter in Kenntnis der Daten abgeleitet wurden.

5.4.2.3 Einschätzung der Übertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien auf den deutschen Versorgungskontext beim Restaging (Ergänzungsrecherche)

Neben der Bewertung des Verzerrungspotenzials wurde auch die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext evaluiert. Es fand sich eine Studie (Roedl 2009a [104]), die Anhaltspunkte für eine Nichtübertragbarkeit auf den deutschen Versorgungskontext nachwies:

Zum einen wurde die Bewertung der Ergebnisse der Diagnose und der klinischen Hintergrundinformation verblindet ausgewertet. Zum anderen erfolgte die Interpretation der Bilder von 2 Bewertern, die mindestens 4 Jahre Erfahrung im Bereich PET/CT nachweisen konnten (siehe Tabelle 37).

Genannte Aspekte entsprechen möglicherweise nicht dem klinischen Alltag, da meistens den Beurteilern der PET bzw. aller weiteren Tests die Ergebnisse vorangegangener Diagnosen

bekannt sind. Darüber hinaus wird davon ausgegangen, dass die Interpretation der Bilder in der Regel nicht von 2 Klinikern voneinander unabhängig erfolgt, sodass es plausible Gründe dafür gibt, dass die Ergebnisse möglicherweise nur für einen Teil der Anwendungssituationen in Deutschland gelten.

Tabelle 37: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Studien beim Restaging (Ergänzungsrecherche)

Studie	F1	F2	F3	F4	F5	F6	F7	Generelle Einschätzung der Übertragbarkeit
Higuchi 2008	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Roedl 2009a	-	-	+	+	-	+	-	Anhaltspunkte für Nichtübertragbarkeit
+: konkrete Anhaltspunkte für eine Nichtübertragbarkeit liegen vor								
-: keine Anhaltspunkte für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden)								

5.4.2.4 Ergebnisse zur diagnostischen Güte der eingeschlossenen Primärstudien beim Restaging (Ergänzungsrecherche)

In diesem Abschnitt werden die Ergebnisse der Primärstudien zur diagnostischen Güte der PET und PET/CT (siehe Tabelle 38) aus der Ergänzungsrecherche dargestellt. Insgesamt wurden noch 2 weitere Primärstudien (Higuchi 2008 und Roedl 2009a), die nicht durch die Evidenzsynthesen abgedeckt wurden, eingeschlossen.

Tabelle 38: Ergebnisse der Studien zur diagnostischen Güte der PET und anderer Indextests beim Restaging von Ösophaguskarzinomen (Ergänzungsrecherche)

Studie	n (N)	Indikation	Indextest	Cut-off	Referenztest	Cut-off	RP	FN	FP	RN	Sensitivität % (95 %-KI)	Spezifität % (95 %-KI)
Higuchi 2008	50 (50)	Bewertung des Tumorsprechens nach neoadjuvanter Chemo- oder Chemoradiotherapie	FDG-PET (nach Behandlung)	SUV _{max} < 2,5	Histologie	> 2/3 Verschwinden von Tumorzellen	27 ^a	3 ^a	2 ^a	18 ^a	90,0 ^a (73,5 bis 97,9) ^{a,b,c}	90,0 ^a (68,3 bis 98,8) ^{a,b,c}
Roedl 2009a	49 (49)	Bewertung des Tumoransprechens nach neoadjuvanter Chemoradiotherapie	FDG-PET (Abnahme der Werte zwischen Prä- und Post-Behandlung)	59 % (Reduktion des Bereiches)	Histologie lebende Tumorzellen	< 10 %	17 ^a	5 ^a	4 ^a	23 ^a	77 (54,6 bis 92,2) ^{a,b}	85 (66,3 bis 95,8) ^a
				32 % (Reduktion des Durchmessers)		18 ^a	4 ^a	3 ^a	24 ^a	82 (59,7 bis 94,8) ^{a,b}	89 (70,8 bis 97,6) ^{a,b}	
				43 % (Reduktion des SUV des Bereichs)		13 ^a	9 ^a	0 ^a	27 ^a	59 (36,4 bis 79,3) ^{a,b}	100 (87,2 bis 100) ^{a,b}	
				56 % (Index = Durchmesser – SUV)		20 ^a	2 ^a	2 ^a	25 ^a	91 (70,8 bis 98,9) ^{a,b}	93 (75,7 bis 99,1) ^{a,b}	
			FDG-PET/CT (Abnahme der Werte Prä- und Post-Behandlung)	54 % (Bereich)	15 ^a	7 ^a	2 ^a	25 ^a	68 (45,1 bis 86,1) ^{a,b}	93 (75,7 bis 99,1) ^{a,b}		
			23 % (Durchmesser)	20 ^a	2 ^a	3 ^a	24 ^a	91 (70,8 bis 98,9) ^{a,b}	89 (70,8 bis 97,6) ^{a,b}			
			42 % (SUV des Bereichs)	15 ^a	7 ^a	0 ^a	27 ^a	68 (45,1 bis 86,1) ^{a,b}	100 (87,2 bis 100) ^{a,b}			
			55 % (Index = Durchmesser – SUV)	20 ^a	2 ^a	2 ^a	25 ^a	91 (70,8 bis 98,9) ^{a,b}	93 (75,7 bis 99,1) ^{a,b}			
		PET und / oder PET/CT ^d	42 % (SUV _{max}) ^d			18 ^a	4 ^a	8 ^a	19 ^a	82 (59,7 bis 94,8) ^{a,b}	70 (49,8 bis 86,2) ^{a,b}	

(Fortsetzung)

Tabelle 38: Ergebnisse der Studien zur diagnostischen Güte der PET und anderer Indextests beim Restaging von Ösophaguskarzinomen (Ergänzungsrecherche) (Fortsetzung)

a: Wert(e) wurde(n) selbst berechnet.

b: Keine oder unzureichende Angabe des 95 %-KI in der Publikation.

c: Selbst berechnete Werte für Sensitivität und Spezifität weichen von den angegebenen Ergebnissen in der Publikation ab, da die Autoren einen positiven Befund in der PET als Nichtansprechen und einen negativen Befund als Ansprechen definiert haben. Hier vereinheitlicht.

d: Es wurde nicht beschrieben, auf welchen Indextest oder eventuell auf eine Kombination beider Tests sich dieser Cut-off bezieht.

FDG = ¹⁸Fluordeoxyglukose; FN = falsch negativ; FP = falsch positiv; KI = Konfidenzintervall; PET = Positronenemissionstomographie; RN = richtig negativ; RP = richtig positiv; SUV = Standardized Uptake Value

Univariate Ergebnisse für die diagnostische Güte der PET zw. PET/CT und anderer bildgebenden Verfahren auf Basis der Studien aus Evidenzsynthesen und Ergänzungsrecherche

In der folgenden Abbildung 6 werden die Ergebnisse der PET, PET/CT und der EUS beim Restaging von Ösophaguskarzinomen nach Beendigung der Therapie univariat dargestellt. In Abbildung 7 werden die Ergebnisse für die diagnostische Güte der PET bzw. PET/CT und weitere bildgebende Verfahren beim Interimsstaging (die PET wurde während der Therapie durchgeführt) univariat dargestellt. In beiden Abbildungen werden nur die Studien dargestellt, aus denen eindeutig hervorging, dass die PET bzw. PET/CT nach Beendigung der Therapie oder während der Therapie erfolgte.

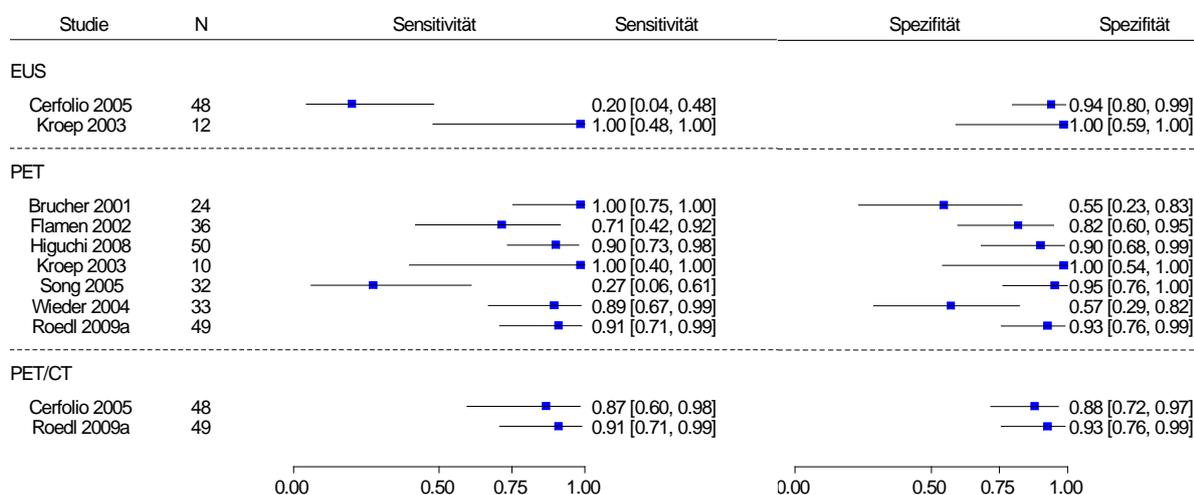


Abbildung 6: Forest Plot der univariaten Ergebnisse für die diagnostische Güte der PET bzw. PET/CT und weiterer bildgebender Verfahren beim Restaging von Ösophaguskarzinomen nach Abschluss der Therapie

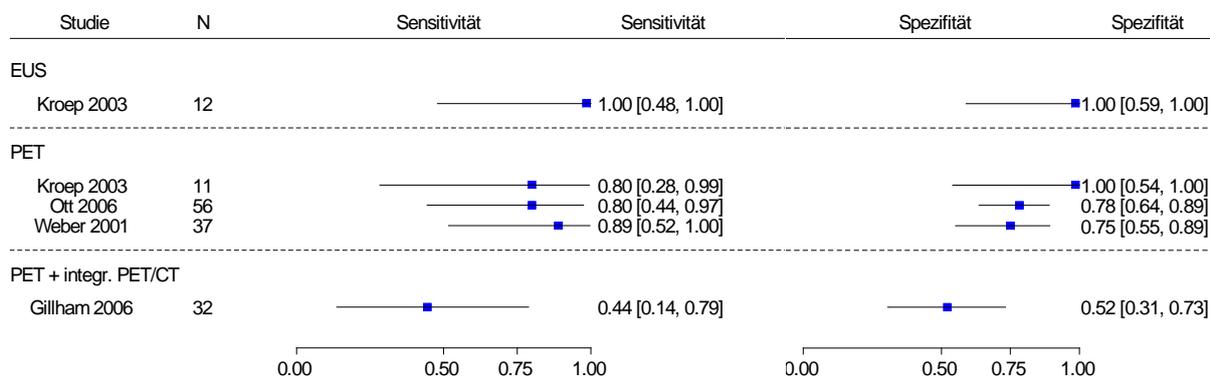


Abbildung 7: Forest Plot der univariaten Ergebnisse für die diagnostische Güte der PET bzw. PET/CT und weiterer bildgebender Verfahren beim Interimsstaging von Ösophaguskarzinomen

Direkte Vergleiche zwischen PET bzw. PET/CT und anderen bildgebenden Verfahren auf Basis der Studien aus Evidenzsynthesen und Ergänzungsrecherche

In den 11 eingeschlossenen Studien für das Restaging wurden 3 direkte Vergleiche zwischen PET und anderen bildgebenden Verfahren berichtet. Für die Indikation Restaging konnte keine bivariate Meta-Analyse berechnet werden, da Vergleichstechnologien nicht in ausreichendem Maß vorhanden waren.

Für das Restaging (nach Abschluss der Therapie) wurde in 2 Studien ein direkter Vergleich der Technologien untersucht.

Die Studie von Cerfolio 2005 analysierte den direkten Vergleich von PET/CT und EUS bei 48 eingeschlossenen Patienten. In der Spezifität zeigen die beiden Technologien ähnliche Punktschätzer, in der Sensitivität erreicht die PET/CT 87 % (KI 60 % bis 98 %) und die EUS eine Sensitivität von 20 % (KI 4 % bis 48 %).

In der zweiten Studie (Roedl 2009a) wurde die PET im Vergleich zur PET/CT an 49 Patienten untersucht. Je nach Grenzwert erreichten die Technologien unterschiedliche Werte für Sensitivität und Spezifität.

Für das Interimsstaging (PET wird während einer noch nicht abgeschlossenen Therapie durchgeführt) wurden in der Studie von Kroep 2003 [68] die Technologien PET, EUS und CT auf Basis einer sehr geringen Patientenzahl (je nach Technologie 10 bis 12 Patienten) miteinander verglichen. Die Punktschätzer der Sensitivitäten und Spezifitäten ließen sich nicht direkt miteinander vergleichen, weil zum einen die Angaben zum Zeitpunkt der Durchführung (während oder nach der Therapie) bei der CT fehlten. Zum anderen wurde die

EUS zwar während der Therapie und auch nach der Therapie durchgeführt, die Werte für Sensitivität und Spezifität wurden aber nicht getrennt berichtet.

Subgruppenmerkmale und andere Effektmodifikatoren auf Basis der Studien aus Evidenzsynthesen und Ergänzungsrecherche

Es wurden keine Subgruppenmerkmale in den eingeschlossenen Studien untersucht.

Ergänzend betrachtete Ergebnisse zu Änderungen im klinischen Management auf Basis der Studien aus Evidenzsynthesen und Ergänzungsrecherche

Es wurden keine Änderungen im klinischen Management berichtet.

5.4.3 Prognostische Güte der eingeschlossenen Primärstudien beim Restaging (Evidenzsynthesen und Ergänzungsrecherche)

In der Evidenzsynthese von Ngamruengphong 2010 wurden neben der diagnostischen Güte der PET bzw. PET/CT auch Angaben über die prognostische Güte gemacht, soweit die Studien beides untersuchten. Aus den berichteten Daten konnte nicht abgeleitet werden, ob die Studien ein multifaktorielles Modell berechnet hatten. Deshalb wurden die Primärstudien angesehen, über die Angaben zur prognostischen Güte gemacht wurden und die bereits im Rahmen der diagnostischen Fragestellung eingeschlossen worden waren [54,66,70,74,110].

In 2 dieser Studien wurde kein multifaktorielles Modell berechnet, sodass die Einschlusskriterien nicht erfüllt waren [54,74].

In der Studie von Flamen 2002 [66] wurde zwar ein multifaktorielles Modell berechnet, es fehlten aber Angaben, ob für Störgrößen adjustiert wurde. Daher konnten die Daten aus Flamen 2002 nicht weiter verwendet werden.

In der Studie von **Ott 2006** [70] wurde ein multifaktorielles Modell berechnet, sodass die Studie den Einschlusskriterien des vorliegenden Berichts entsprach. Die Daten zur prognostischen Güte wurden der Primärstudie entnommen.

Weil die eingeschlossenen Evidenzsynthesen nicht explizit nach Prognosestudien recherchiert hatten, wurde im Rahmen der Ergänzungsrecherche nach weiteren Prognosestudien gesucht. Es konnten 2 weitere Prognosestudien **Higuchi 2008** [103] und **Kim 2007** [77] für die Indikation Restaging, die den Einschlusskriterien des vorliegenden Berichts entsprachen, eingeschlossen werden. Bei der Studie von Higuchi 2008 handelte es sich um eine Prognosestudie, die als Assoziationsmaß die Sensitivität und Spezifität verwendet. In der Studie wurde ein einfaktorielles, aber kein multifaktorielles Modell berechnet, sodass die Studie als „prognostic accuracy study“ eingeschlossen wurde.

In den folgenden Abschnitten werden zunächst die Ergebnisse der beiden Prognosestudien (**Ott 2006** [70] und **Kim 2007** [77]) beim Restaging präsentiert. Im Anschluss erfolgt die Darstellung der „prognostic accuracy study“ **Higuchi 2008** [103].

5.4.3.1 Studiendesigns und Studienpopulationen der eingeschlossenen Prognosestudien beim Restaging (Evidenzsynthesen und Ergänzungsrecherche)

Die Studie von **Kim 2007** [77] untersuchte die diagnostische und auch die prognostische Güte der PET und schloss 62 Patienten vor neoadjuvanter Chemoradiotherapie ein. Die Arbeit untersuchte zum einen die diagnostische Güte der FDG-PET zur Beurteilung des Behandlungsansprechens von Patienten mit operablen Ösophaguskarzinomen, die eine neoadjuvante Chemotherapie erhielten. Zum anderen untersuchte sie die Beurteilung der prognostischen Güte der FDG-PET zur Vorhersage des Überlebens von Patienten nach präoperativer Chemotherapie und Operation. Dabei wurden verschiedene Faktoren unifaktoriell und multifaktoriell hinsichtlich ihrer Erklärungskraft in der Vorhersage des krankheitsfreien Überlebens und des Gesamtüberlebens untersucht. Am Ende der klinischen Nachbeobachtung (Median 19,3 Monate, Spannweite 3,9 bis 57,1 Monate) waren 20 der 62 ausgewerteten Patienten verstorben. Vier Patienten verstarben an postoperativen Komplikationen und wurden nicht in die Analyse eingeschlossen.

In der Studie von **Ott 2006** [70] wurden 56 Patienten mit Nachweis eines Adenokarzinoms im distalen Ösophagus oder gastroösophagealen Übergang eingeschlossen, bei denen PET-Untersuchungen vor und nach Behandlung (neoadjuvante Chemotherapie mit nachfolgender Resektion) durchgeführt wurden. Ziel der Studie war die Bewertung des metabolischen Ansprechens in der FDG-PET zur Vorhersage des Überlebens (rezidivfreies Überleben (ab Datum der Operation) und Gesamtüberleben (ab ersten Tag der neoadjuvanten Chemotherapie)). Als primärer Endpunkt wurde das Gesamtüberleben gewählt, weil er auf alle eingeschlossenen Patienten und nicht nur auf eine Subgruppe angewandt werden konnte. Dabei wurden verschiedene Faktoren hinsichtlich ihres Einflusses auf das Überleben univariat und multivariat untersucht.

Tabelle 39: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudien beim Restaging

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenzstufe ^b	Verzerrungspotenzial
Kim 2007	Diagnostische und prognostische Kohortenstudie Rekrutierung der Patienten ab März 1999. Südkorea	Bewertung der prognostischen Güte von FDG-PET zur Vorhersage des Überlebens von Patienten nach präoperativer Chemotherapie und vor Operation.	Prognostische Güte von FDG-PET zur Vorhersage des krankheitsfreien Überlebens und des Gesamtüberlebens.	II ^c	hoch
Ott 2006	Prognostische Kohortenstudie Rekrutierung der Patienten zwischen Januar 1999 und Juli 2002. Deutschland	Bewertung des metabolischen Ansprechens in der FDG-PET zur Vorhersage des Überlebens.	Prognostische Güte von FDG-PET zur Vorhersage des rezidivfreien Überlebens und des Gesamtüberlebens..	II ^c	hoch
<p>a: Studienziel gemäß Formulierung der Autoren b: entspricht der Evidenzstufe des G-BA c: Die Verfahrensordnung des G-BA sieht derzeit keine Evidenzstufe für prognostische Studien vor. Sie werden im Rahmen dieses Berichts daher mit Stufe 2 gekennzeichnet. FDG = ¹⁸Fluorodeoxyglukose; PET = Positronenemissionstomographie</p>					

Tabelle 40: Index- und Referenztest der eingeschlossenen Prognosestudien beim Restaging

Studie	Indextest	Referenztest
Kim 2007	PET (ECAT HR ⁺ oder Biograph Sensation 16, Siemens, USA) 60 Minuten nach i.v.-Injektion von ca. 550 MBq FDG. Aufnahme von Schädelbasis bis oberer Oberschenkel. Visuelle und semiquantitative Analyse (SUV _{max}) durch erfahrenen Nuklearmediziner. Bewertung der metabolischen Antwort nach EORTC-Kriterien.	Überleben (Datum der Operation bis Tod)
Ott 2006	PET (keine Angaben zum Gerät) 40 Minuten nach i.v.-Injektion von 250-370 MBq FDG. Quantitative Bewertung. Keine Angaben zum Bewerter und zur evtl. Entscheidungsfindung.	Rezidivfreies Überleben (ab Datum der Operation) und Gesamtüberleben (ab ersten Tag der neoadjuvanten Chemotherapie). Nachbeobachtung 26 bis 67 Monate, Median: 42 Monate
<p>CT = Computertomographie; FDG = ¹⁸Fluorodeoxyglukose; i.v. = intravenös; MBq = Mega Becquerel; PET = Positronenemissionstomographie; SUV = standardized uptake value; TI = ²⁰¹Thaliumchlorid; USA = Vereinigte Staaten von Amerika</p>		

Tabelle 41: Eingeschlossene Patienten der eingeschlossenen Prognosestudien beim Restaging

Studie	n (N)	Geschlecht [n (m/w)]	Alter (MW, Median, Spannweite)	Diagnose	Vorbehandlung
Kim 2007	62 (89)	62 (54/8)	k. A.; 63; 45–74	Differenzierungsgrad: guter (6), mittlerer (45), niedriger (11) Lokalisation: obere (3), mittlere (25), untere Speiseröhre (34) Klinisches Stadium: IIA (15), IIB (18), III (29)	Nachbeobachtung 3,9 bis 57,1 Monate, Median: 19,3 Nach Studienbeginn neoadjuvante Chemoradiotherapie (9: 5-FU, Cisplatin und Bestrahlung; 53: Capecitabin, Cisplatin und Bestrahlung)
Ott 2006	56 (65)	56 (50/6)	58; k. A., k. A.	Histologie: Adenokarzinome im distalen Ösophagus oder gastroösophagealen Übergang (56). Lokalisation: intestinal (40), nicht intestinal (16). Siewert-Einteilung der Karzinome des gastroösophagealen Übergangs ^a : I (35), II (21).	Keine Angaben zur Vorbehandlung. Nach Studienbeginn: Cisplatin, Leucovorin, Fluoruracil. Zusätzlich Paclitaxel für Patienten mit AEG I. OP 3-4 Wochen nach Chemotherapie.
<p>a: Siewert JR, Stein HJ. Classification of adenocarcinoma of the oesophagogastric junction. Br J Surg 1998; 85(11): 1457-1459. AEG = adenocarcinomas of the esophagogastric junction; m = männlich; MW = Mittelwert; n = Anzahl; OP = Operation; PET = Positronenemissionstomographie; w = weiblich</p>					

5.4.3.2 Einschätzung des Verzerrungspotenzials der eingeschlossenen Prognosestudien beim Restaging (Evidenzsynthesen und Ergänzungsrecherche)

In der Studie von **Kim 2007** [77] wurde nicht beschrieben, ob und ggf. wie die Patienten während der klinischen Nachbeobachtung behandelt wurden. Zudem fehlen in der multifaktoriellen Analyse weitere Faktoren, die das Überleben der Patienten beeinflussen könnten, wie Alter und klinisches Stadium. Die eingeschlossenen Patienten waren zu Studienbeginn in unterschiedlichen Krankheitsstadien.

In **Ott 2006** [70] wurden potenziell relevante Faktoren, wie Alter und Behandlung, in der multifaktoriellen Analyse nicht berücksichtigt. Die eingeschlossenen Patienten waren zu Studienbeginn in unterschiedlichen Krankheitsstadien.

Tabelle 42: Verzerrungspotenzial der eingeschlossenen Prognosestudien beim Restaging

Studie	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	Verzerrungspotenzial der Ergebnisse
Kim 2007	+	-	+	+	+	+	+	+	+	?	+	+	+	-	?	+	-	hoch
Ott 2006	+	-	+	+	+	+	+	+	+	?	+	+	+	-	?	+	-	hoch
+: Kriterium erfüllt; -: Kriterium nicht erfüllt; ?: unklar																		

5.4.3.3 Einschätzung der Übertragbarkeit der Ergebnisse der Prognosestudien auf den deutschen Versorgungskontext beim Restaging (Evidenzsynthesen und Ergänzungsrecherche)

Neben der Bewertung des Verzerrungspotenzials wurde auch für die beiden Prognosestudien die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext evaluiert. In der Studie von Ott 2006 fand sich ein Anhaltspunkt für eine Nichtübertragbarkeit auf den deutschen Versorgungskontext, da im Rahmen der Studien die Ergebnisse der Tests ohne Kenntnis vorhergehender klinischer Untersuchungen und ohne Kenntnis der Ergebnisse weiterer Tests beurteilt wurden (siehe Tabelle 62).

Genannter Aspekt entsprechen möglicherweise nicht dem klinischen Alltag, da meistens den Beurteilern der PET bzw. aller weiteren Tests die Ergebnisse vorangegangener Diagnosen bekannt sind. Eine generelle Übertragbarkeit ist damit eingeschränkt.

Tabelle 43: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Prognosestudien bei der Restaging

Studie	F1	F2	F3	F4	F5	F6	F7	Generelle Einschätzung der Übertragbarkeit
Kim 2007	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Ott 2006	-	-	-	+	-	-	-	Anhaltspunkt für eine Nichtübertragbarkeit
+: konkrete Anhaltspunkte für eine Nichtübertragbarkeit liegen vor								
-: keine Anhaltspunkte für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden)								

5.4.3.4 Ergebnisse der Primärstudien zur prognostischen Güte beim Restaging (Evidenzsynthesen und Ergänzungsrecherche)

In der folgenden Tabelle werden die Ergebnisse der ein- und multifaktoriellen Analysen für die eingeschlossenen Prognosestudien beim Restaging dargestellt. Es wurde eine Vielzahl an prognostischen Merkmalen untersucht, wobei in der folgenden Tabelle nur die PET-relevanten Merkmale dargestellt werden. Alle weiteren Merkmale werden im Anhang E präsentiert.

In der Studie von **Kim 2007** [77] wurden in der einfaktoriellen Analyse Alter, Geschlecht, Lokalisation des Primärtumors, histologischer Grad, klinisches Stadium, verschiedene PET-Parameter (SUV_{max} des Primärtumors, SUV-Änderung nach Behandlung im Vergleich zu vor Behandlung) sowie Verhältnis zwischen komplettem und partiellem Ansprechen in CT bzw. PET betrachtet. Statistisch signifikante Faktoren ($p < 0,1$) wurden in die multifaktorielle Analyse eingeschlossen, wobei beide eingeschlossenen Faktoren (SUV-Abnahme $\geq 80\%$, Vorliegen eines kompletten Ansprechens in der PET) hinsichtlich des krankheitsfreien Überlebens statistisch signifikant blieben. Beim Gesamtüberleben blieb der Parameter „Ansprechen“ signifikant, während die SUV-Abnahme nicht mehr statistisch signifikant war. An dieser Stelle muss erwähnt werden, dass, wenn 2 PET-Parameter in einem Modell integriert werden, das Problem der Kollinearität eintritt.

Für das Gesamtüberleben bei **Ott 2006** [70] blieb in der multifaktoriellen Analyse, welche die FDG-Abnahme nach Behandlung, das Tumorstadium nach Behandlung, das histopathologische Ansprechen sowie das Vorliegen von Lymphknoten in der Histologie nach Behandlung einschloss, lediglich das Nichtvorliegen von Lymphknoten statistisch signifikant. In der multifaktoriellen Analyse für das rezidivfreie Überleben war lediglich das metabolische Ansprechen statistisch signifikant. Für das rezidivfreie Überleben unterscheidet sich das Ergebnis für die uni- und multifaktorielle Analyse nicht.

Tabelle 44: Ergebnisse der Studien zur prognostischen Güte der PET beim Restaging

Studie	Zielgröße	n	Prognostisches Merkmal	Cut-off	Patienten mit prog. Merkmal	Ergebnisse aus einfaktorieller Analyse		Ergebnisse aus multifaktorieller Analyse	
						p-Wert	HR	p-Wert	HR
Kim 2007	Krankheitsfreies Überleben	62 ^a	SUV-Aufnahme	Referenz: < 80 %	k. A. ^b	0,016	k. A.	0,025	0,25
			Komplettes metabolisches Ansprechen	Referenz: Ja		k. A.	k. A.	0,006	3,58
	Gesamtüberleben	62 ^a	SUV-Aufnahme	Referenz: < 80 %		0,023	k. A.	0,64	0,52
			Komplettes metabolisches Ansprechen	Referenz: Ja		k. A.	k. A.	0,033	3,09
Ott 2006	Gesamtüberleben ^e	50 ^c	Metabolisches Ansprechen (Reduktion der FDG-Aufnahme zwischen prä- und postoperativer PET)	35 %	k. A.	0,039	0,34	> 0,05	k. A.
		41	Metabolisches Ansprechen (Reduktion der FDG-Aufnahme zwischen prä- und postoperativer PET)	35 %	k. A.	0,05	k. A.	> 0,05	k. A.
	Rezidivfreies Überleben	41 ^d	Metabolisches Ansprechen (Reduktion der FDG-Aufnahme zwischen prä- und postoperativer PET)	35 %	18/56	0,018	0,22	0,018	0,22

a: Vier Patienten verstarben an postoperativen Komplikationen und wurden nicht in die Analyse eingeschlossen. Es wird nicht berichtet, warum für „komplettes metabolisches Ansprechen“ lediglich 61 Patienten berichtet werden.

b: Durchschnittliche SUV-Abnahme des Primärtumors 83,3 %

c: Außerdem einfaktoriell modelliert, aber p-Wert ≥ 0,1: baseline FDG uptake, FDG uptake zum Tag 14, R0-Kategorie.

d: Außerdem einfaktoriell modelliert: baseline FDG uptake, FDG uptake day 14, ypT-stage less than T3, histopathologic response alle p-Wert > 0,1.

FDG = ¹⁸Fluorodeoxyglukose; HR = Hazard Ratio; KI = Konfidenzintervall; PET = Positronenemissionstomographie; SUV = standardized uptake value

5.4.3.5 Studiendesign und Studienpopulation der Studie zur „prognostic accuracy“ beim Restaging (Ergänzungsrecherche)

Die diagnostische als auch prognostische Kohortenstudie von **Higuchi 2008** [103] schloss 50 Patienten mit histologisch nachgewiesenen Plattenepithelkarzinomen des Ösophagus ein, die eine neoadjuvante Chemo- bzw. Chemoradiotherapie vor OP erhielten. Ziel der Arbeit war es zu bewerten, ob FDG-PET das histologische Ansprechen und das postoperative Überleben vorhersagen kann.

Die Patienten wurden über einen Median von 26,5 Monaten (Spannweite 5,4 bis 84,2 Monate) nachbeobachtet.

Tabelle 45: Design, primäres Studienziel und untersuchte Zielgröße der Studie zur „prognostic accuracy“ beim Restaging

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenzstufe ^b	Verzerrungspotenzial
Higuchi 2008	„Prognostic accuracy study“ Rekrutierung der Patienten zwischen März 2000 und August 2004. Japan	Bewertung, ob FDG-PET das histologische Ansprechen und das Überleben nach Operation von Patienten mit Plattenepithelkarzinomen vorhersagen kann.	Diagnostische Güte der FDG-PET zur Vorhersage des fall- und krankheits-spezifischen Überlebens.	II ^c	hoch
<p>a: Studienziel gemäß Formulierung der Autoren b: entspricht der Evidenzstufe des G-BA c: Die Verfahrensordnung des G-BA sieht derzeit keine Evidenzstufe für prognostische Studien vor. Sie werden im Rahmen dieses Berichts daher mit Stufe 2 gekennzeichnet. FDG = ¹⁸Fluordeoxyglukose; PET = Positronenemissionstomographie</p>					

Tabelle 46: Index- und Referenztest der Studie zur „prognostic accuracy“ beim Restaging

Studie	Indextest	Referenztest
Higuchi 2008	PET (HEADTOME/SET 2400 W, Shimadzu Co, Japan) 60 Minuten nach i.v.-Injektion von ca. 370 MBq FDG. Läsionen mit SUVmax $\geq 2,5$ wurden als bösartig gewertet. Keine Angaben zur Befundung.	Nachbeobachtung (Median 26,5 Monate, Spannweite 5,4 bis 84,2 Monate), keine weiteren Angaben. Bestimmung des fall-spezifischen und krankheits-freien Überlebens nach 1, 3 und 5 Jahren.
<p>FDG = ¹⁸Fluordeoxyglukose; GE = General Electric; i.v. = intravenös; MBq = Mega Becquerel; PET = Positronenemissionstomographie; SUV = standardized uptake value</p>		

Tabelle 47: Eingeschlossene Patienten der Studie zur „prognostic accuracy“ beim Restaging

Studie	Patienten [n]	Geschlecht [n (m/w)]	Alter (MW, Median, Spannweite)	Diagnose	Vorbehandlung
Higuchi 2008	50	50 (41/9)	k. A.; 62,2; 44–77	Histologie: Plattenepithelkarzinom (50). Lokalisation: zervikale (11), obere (5), mittlere (17), untere Speiseröhre (17). Klinisches Stadium (vor Behandlung): I (0), IIA (0), IIB (0), III (24), IVA (11), IVB (15) Resektabilität: R0 (48), R1 (0), R2 (2)	Nach Studienbeginn: 35 Chemo- und 35 Chemoradiotherapie. 11 zervikale, 39 subtotale Ösophagektomie.
K. A = keine Angabe(n); m = männlich n = Anzahl; OP = Operation; T = Tumor; w = weiblich					

5.4.3.6 Einschätzung des Verzerrungspotenzials der Studie zur „prognostic accuracy“ beim Restaging

Es blieb in der Studie von Higuchi 2008 [103] unklar, ob die Befunder des Referenztests verblindet waren und ob die Patienten zwischen Index- und Referenztest behandelt wurden. Darüber hinaus wurde in der Studie die Nachbeobachtung als Referenztest durchgeführt, bei der unklar blieb, wie das Rezidiv bestimmt wurde. In der Studie wurde keine Intention-to-Diagnose-Analyse durchgeführt. Es konnte nicht ausgeschlossen werden, dass die fehlenden Werte nicht zufällig und in nicht akzeptabler Häufigkeit auftraten.

Tabelle 48: Verzerrungspotenzial der Studie zur „prognostic accuracy“ bei der Rezidivdiagnostik

Studie	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Verzerrungspotenzial der Ergebnisse
Higuchi 2008	?	+	+	+	?	+	?	-	+	?	hoch
+: Kriterium erfüllt; -: Kriterium nicht erfüllt; ?: unklar											

5.4.3.7 Einschätzung der Übertragbarkeit der Studie zur „prognostic accuracy“ beim Restaging

In der Studie von Higuchi 2008 konnten keine Anhaltspunkte für eine Nichtübertragbarkeit identifiziert werden. Die Ergebnisse wurden daher als auf den deutschen Versorgungskontext übertragbar angesehen (vgl. Tabelle 49).

Tabelle 49: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der Studie zur „prognostic accuracy“ beim Restaging

Studie	F1	F2	F3	F4	F5	F6	F7	Generelle Einschätzung der Übertragbarkeit
Higuchi 2008	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
-: keine Anhaltspunkte für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden)								

5.4.3.8 Ergebnisse der Studie zur „prognostic accuracy“ beim Restaging

In der folgenden Tabelle werden die Ergebnisse der „prognostic accuracy study“ von Higuchi 2008 dargestellt. Es bestand eine Diskrepanz zwischen den im Text und den in den Tabellen berichteten Ergebnissen. Die Sensitivität wurde niedrig geschätzt (47,6 bis 59,4 %), während die Schätzer der Spezifität höher (70,8 % bis 87,5 %) lagen. Die Daten wurden kontrolliert und noch einmal selbst berechnet.

Tabelle 50: Ergebnisse der Studie zur „prognostic accuracy“ bei der Restaging

Studie	n (N)	Indextest	Cut-off	Referenztest	Cut-off	RP	FN	FP	RN	Sensitivität % (95 %-KI)	Spezifität % (95 %-KI)
Higuchi 2008	50 (50)	FDG-PET (nach Behandlung)	SUVmax ≤ 2,5 (Hinweis auf Behandlungs- ansprechen)	Klinische Nachbeobach- tung	Überleben nach 1 Jahr	7 ^a	1 ^a	22 ^a	20 ^a	87,5 ^a (47,3 bis 99,7) ^a	47,6 ^a (32,0 bis 63,6) ^a
	50 (50)				Überleben nach 5 Jahren	18 ^a	7 ^a	11 ^a	14 ^a	72,0 ^a (50,6 bis 87,9) ^a	56,0 ^a (34,9 bis 75,6) ^a
	50 (50)				Überleben nach 3 Jahren	17 ^a	5 ^a	12 ^a	16	77,3 ^a (54,6 bis 92,2) ^a	57,1 ^a (37,2 bis 75,5) ^a
	48 ^b (50)				Krankheits- freies Überleben nach 1 Jahr	14 ^a	2 ^a	13 ^a	19 ^a	87,5 ^a (61,7 bis 98,4) ^a	59,4 ^a (40,6 bis 76,3) ^a
	48 (50)				Krankheits- freies Überleben nach 3 Jahren	17 ^a	7 ^a	10 ^a	14 ^a	70,8 ^a (48,9 bis 87,4) ^a	58,3 ^a (36,6 bis 77,9) ^a
	48 (50)				Krankheits- freies Überleben nach 5 Jahren	17 ^a	7 ^a	10 ^a	14 ^a	70,8 ^a (48,9 bis 87,4) ^a	58,3 ^a (36,6 bis 77,9) ^a

a: Wert(e) wurde(n) berechnet.
 b: 50 Patienten wurden eingeschlossen. Zwei Patienten wurden von der Analyse ausgeschlossen, da sie eine R2-Resektion bei fortgeschrittener metastasierter Erkrankung hatten.
 FDG = ¹⁸Fluorodeoxyglukose; PET = Positronenemissionstomographie; SUV = standardized uptake value

5.5 Teilziel 2c: diagnostische und prognostische Güte bei der Rezidivdiagnostik

5.5.1 Diagnostische und prognostische Güte auf Grundlage in die Bewertungen eingeflossener Evidenzsynthesen bei der Rezidivdiagnostik

Für diese Fragestellung konnte keine Evidenzsynthese gefunden werden, die den Einschluss- und Qualitätskriterien des vorliegenden Berichts entsprach. Die Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT wurde auf Basis der Primärstudien untersucht.

5.5.2 Diagnostische Güte auf Grundlage der eingeschlossenen Primärstudien bei der Rezidivdiagnostik

In der Recherche wurden 4 Primärstudien zur diagnostischen Güte der PET und PET/CT bei der Rezidivdiagnostik identifiziert (vgl. Tabelle 13). In den 4 für die Rezidivdiagnostik eingeschlossenen Diagnosestudien wurden insgesamt 175 Patienten untersucht. In allen Studien wurden Angaben zum Geschlecht der eingeschlossenen Patienten gemacht: Insgesamt handelt es sich um 139 Männer und 36 Frauen. In 3 von 4 Studien wurden sowohl Patienten mit einem Plattenepithelkarzinom als auch einem Adenokarzinom der Speiseröhre untersucht. In einer Studie wurden ausschließlich Patienten mit einem Plattenepithelkarzinom der Speiseröhre eingeschlossen. Eine Studie wurde in Frankreich, eine in den USA und 2 Studien in China durchgeführt.

Der Einsatz der nichtintegrierten FDG-PET wurde in einer Studie, die integrierte FDG-PET/CT in 2 Studien und der Einsatz beider Technologien in einer Studie untersucht. Als Tracer wurde in allen Studien Fluorodeoxyglukose (FDG) benutzt. Als Referenztests wurden Histologie, Biopsie ggf. in Verbindung mit einer Nachbeobachtung eingesetzt.

Im Folgenden werden Besonderheiten bezüglich Fragestellungen und wesentlicher methodischer Charakteristika, die in einigen Studien auftraten, erläutert. Alle weiteren ausführlichen Daten sind den Tabellen der folgenden Abschnitte zu entnehmen.

Die Arbeit von **Guo 2007** [78] betrachtete sowohl eine diagnostische als auch eine prognostische Fragestellung, sodass im vorliegenden Bericht beide Ergebnisse wiedergegeben werden. An dieser Stelle wird die Bewertung der diagnostischen Güte von PET/CT erläutert. Die Ergebnisse zur prognostischen Güte werden in Abschnitt 5.5.4. dargestellt. Die Studie von **Teyton 2009** [106] verfolgte das Ziel, FDG-PET und CT zur Rezidiverkennung bei asymptomatischen Patienten zu vergleichen. Ergebnisse wurden für die Erkennung jeglicher Rezidive (lokal und Fernmetastasen) sowie jeweils für lokale Rezidive und Fernmetastasen berichtet. Im vorliegenden Bericht wurden nur die Ergebnisse der Erkennung jeglicher Rezidive dargestellt, da die getrennte Darstellung der Ergebnisse unvollständig war. Darüber hinaus wurden nur die Daten der PET dargestellt, da die fehlenden Werte der CT die 30 %-Schwelle überschritt.

5.5.3 Studiendesigns und Studienpopulationen der eingeschlossenen Primärstudien bei der Rezidivdiagnostik

Tabelle 51: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Studien bei der Rezidivdiagnostik

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenzstufe ^b	Verzerrungspotenzial
Guo 2007	Diagnostische Kohortenstudie Rekrutierung der Patienten zwischen Juni 2003 und Juni 2005. China	Bewertung der diagnostischen Güte von PET/CT bei Patienten mit Rezidivverdacht nach Abschluss der Behandlung eines Plattenepithelkarzinoms.	Diagnostische Güte von FDG-PET/CT zur Erkennung von Rezidiven.	I Ib	hoch
Roedl 2008a	Diagnostische Kohortenstudie Keine Angaben zum Rekrutierungs- und Untersuchungszeitraum, USA	Bewertung der diagnostischen Güte von PET/CT im Vergleich zu PET zur Rezidiverkennung.	Diagnostische Güte von FDG-PET bzw. PET/CT zur Erkennung von jeglichen Rezidiven.	I Ib	hoch
Teyton 2009	Diagnostische Kohortenstudie Rekrutierung der Patienten zwischen Oktober 2003 und September 2006. Frankreich	Vergleich von FDG-PET und CT zur Rezidiverkennung.	Diagnostische Güte von FDG-PET zur Erkennung von jeglichen Rezidiven.	I Ib	niedrig
Zhang 2006	Diagnostische Kohortenstudie Rekrutierung der Patienten zwischen Juni 2003 und Juni 2004. China	Bewertung und Vergleich der FDG-PET/CT und CT zur Erkennung von lokalen Rezidiven und Metastasen nach Resektion von Karzinomen der Speiseröhre.	Diagnostische Güte von FDG-PET/CT und CT zur Erkennung von Lokalrezidiven und Fernmetastasen.	I Ib	hoch
<p>a: Studienziel gemäß Formulierung der Autoren b: entspricht der Evidenzstufe des G-BA CT = Computertomographie; FDG = ¹⁸Fluorodeoxyglukose; M = Metastase(n); N = Nodus (Knoten); PET = Positronenemissionstomographie; T = Tumor</p>					

Tabelle 52: Index- und Referenztest der eingeschlossenen Studien bei der Rezidivdiagnostik

Studie	Indextest	Referenztest
Guo 2007	<p>PET/CT (Discovery LS PET/CT system; GE Healthcare, USA) 60 Minuten nach i.v.-Injektion von 370 MBq FDG. Aufnahme von Kopf bis oberer Oberschenkel. Schichtdicke 4,25 mm. CT-Parameter: 120 kV, 90 mA, 6:1 pitch, Schichtdicke 4,25 mm, 0,8 Sekunden pro Rotation.</p> <p>Befundung im Konsensus durch 3 erfahrene Nuklearmediziner auf Xeleris Workstation (GE Healthcare, USA). Keine Verblindung gegenüber klinischen Daten oder Befunden anderer bildgebender Verfahren. Abschließende Bewertung (positiv, grenzwertig, negativ) vorwiegend visuell (erhöhte FDG-Aufnahme im Vergleich zum umgebenden Gewebe) unter Hinzuziehung semiquantitativer Ergebnisse ($SUV \geq 2,5$).</p>	Histologie oder klinische Nachbeobachtung zur Erkennung von Rezidiven (lokal, regional und Fernmetastasen)
Roedl 2008a	<p>PET/CT (Biograph 16, Siemens Medical Solutions, Deutschland) 60 Minuten nach i.v.-Injektion von 555 MBq (15 mCi) FDG. CT nach i.v.-Injektion von 100 ml jodhaltigem KM. Pitch 1,5, Tubenspannung 140 kV.</p> <p>Befundung des PET-Anteils durch ein Team erfahrener Nuklearmediziner. Dann Bewertung der PET/CT durch ein Team aus Nuklearmedizinern und Radiologen. Zwei Befunder werteten unabhängig voneinander aus und der Mittelwert wurde in der Auswertung verwendet.</p>	Auffällige Befunde wurden mit Biopsie abgeklärt. Negative Befunde wurden nachbeobachtet und während dessen mit EUS untersucht. Keine weiteren Angaben.
Teyton 2009	<p>PET (Allegro dedicated PET scanner, Philips Medical Systems, Niederlande) 60 Minuten nach i.v.-Injektion von durchschnittlich 355 MBq (5-6 MBq/kg KG) FDG. Aufnahme von Kopf bis Oberschenkel. Läsionen wurden als bösartig gewertet, wenn die FDG-Aufnahme im Vergleich zum Mediastinum in 2 Schichten erhöht war. Die PET-Bilder wurden von 2 erfahrenen Nuklearmediziner in Unkenntnis der Ergebnisse der CT, der endoskopischen Ultraschall-Befunde und der histologischen Ergebnisse ausgewertet.</p> <p>PET-Untersuchungen alle 6 Monate während der ersten 2 Jahre der Nachbeobachtung und ab dann jährlich.</p>	Biopsie oder regelmäßige klinische Nachbeobachtung alle 4 bis 6 Monate über einen Zeitraum von 2 Jahren (klinische Untersuchung, CT von Thorax und Abdomen, Ultraschall des Bauches, Endoskopie mindestens alle 6 Monate). PET wurde als routinemäßige Prozedur (alle 6 Monate) während der Nachbeobachtung durchgeführt.
Zhang 2006	<p>PET/CT (GE Discovery LS, GE Medical Systems, USA) 60 Minuten nach i.v.-Injektion von 3,7-5,5 MBq/kg KG. Läsionen mit $SUV > 2,5$ wurde als bösartig befundet. CT: Spiral CT thin-layer scanning vom Hals bis zum oberen Brustkorb. Intravenöse Injektion eines bildgebenden Mediums. Schichten 3-5 mm. Keine Angaben zu den Befundern oder zum Bewertungsprozess.</p>	Histologie (keine weiteren Angaben) oder klinische Nachbeobachtung über mindestens 6 Monate (keine weiteren Angaben).
<p>CT = Computertomographie; FDG = ¹⁸Fluorodeoxyglukose; GE = General Electric; i.v. = intravenös; kg = Kilogramm; KG = Körpergewicht; KM = Kontrastmittel; kV = Kilovolt; mA = Milliampere; MBq = Mega Becquerel; mCi = Millicurie; PET = Positronenemissionstomographie; SUV = standardized uptake value; USA = Vereinigte Staaten von Amerika</p>		

Tabelle 53: Eingeschlossene Patienten der eingeschlossenen Studien bei der Rezidivdiagnostik

Studie	Patienten [n]	Geschlecht [n (m/w)]	Alter	Diagnose	Vorbehandlung
Guo 2007	56 ^a	56 (47/9)	k. A.; k. A.; 38–77	Histologie: Plattenepithelkarzinomen (56) Differenzierungsgrad: guter (11), mittlerer (21), niedriger (15), unklar (9). Lokalisation: obere (6), mittlere (34), untere Speiseröhre (16). Stadium (TNM): IIb (5), III (38), IV (13).	Abschluss der Behandlung (26 operativ, 30 Strahlentherapie) vor Aufnahme in die Studie. Wenn erforderlich, Gabe von 5- Fluoruracil als Erhaltungs- Chemotherapie.
Roedl 2008a	47	47 (35/12)	66,4; k. A.	Histologie: Plattenepithelkarzinom (11), Adenokarzinom (36). Lokalisation: obere (5), mittlere (10), untere Speiseröhre (11), gastroösophagealer Übergang (21). Ausdehnung des Primärtumors: T2 (16), T3 (31).	Nach Studienbeginn: Alle Patienten mit neoadjuvanter Chemoradiotherapie (50,4 Gy + Cisplatin + 5-Fluoruracil) und OP (thorakoabdominale Ösopha- gektomie für tiefliegende Läsionen und Ösophagektomie nach Ivor-Lewis mit ausgedehnter Lymphadenektomie für höher gelegene Läsionen).
Teyton 2009	41	41 (38/3)	60,7; 59; 43–83	Histologie: Plattenepithelkarzinom (31), Adenokarzinom (10). Differenzierungsgrad: guter (22), mittlerer (15), niedriger (4). Lokalisation: obere (3), mittlere (20), untere Speiseröhre (18). Stadium: I (6), IIa (15), IIb (5), III (15)	Nach Studienbeginn: 25 nur OP, 7 OP mit adjuvanter Chemotherapie und ggf. Strahlentherapie, 9 OP mit neoadjuvanter Chemotherapie und Strahlentherapie
Zhang 2006	31	31 (19/12)	60; k. A.; 38–78	Histologie: Plattenepithelkarzinom (26), Adenokarzinom (5).	7 ohne Nachbehandlung, 11 Bestrahlung (45–50 Gy) und Chemotherapie (keine weiteren Angaben), 4 nur Bestrahlung, 9 nur Chemotherapie
<p>a: 112 Patienten eingeschlossen. 56 wurden ausgeschlossen: 24 Patienten hatten unvollständige medizinische Unterlagen, 22 Patienten erhielten PET/CT zur Kontrolle des Behandlungsansprechens, 8 Patienten hatten Infektionen zum Zeitpunkt der PET/CT-Untersuchung und bei 2 Patienten wurde histologisch ein Adenokarzinom nachgewiesen.</p> <p>Gy = Gray; m = männlich; k. A. = keine Angabe(n); LK = Lymphknoten; M = Metastase; N = Nodus ((Lymph-) Knoten); n = Anzahl; OP = Operation; T = Tumor; w = weiblich</p>					

5.5.3.1 Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien bei der Rezidivdiagnostik

Drei der eingeschlossenen Studien zur Rezidivdiagnostik wurden mit einem hohen Verzerrungspotenzial bewertet. Die Studie von **Teyton 2009** [106] wurde mit einem niedrigen Verzerrungspotenzial bewertet. Die methodischen Charakteristika der Diagnosestudien werden in Tabelle 54 dargestellt. Im Anschluss an die Tabelle erfolgt eine kurze Zusammenfassung im Hinblick auf das Verzerrungspotenzial der jeweiligen Studie.

Generell wurde festgestellt, dass bei allen Studien keine genauen Angaben über die Zeitspanne zwischen Index- und Referenztest gemacht wurden. Bei 2 Studien blieb unklar, ob die Befunder des Referenztests verblindet waren. Bei 3 Studien fehlten Angaben, ob die Patienten zwischen Index- und Referenztest behandelt wurden.

In der Studie von **Guo 2007** [78] waren die Befunder der PET nicht verblindet und es war unklar, ob die Patienten zwischen Index- und Referenztest behandelt wurden. Bei dieser Studie bestand das Risiko der differenziellen Verifikation, weil 2 unterschiedliche Referenztests (Histologie und klinische Nachbeobachtung) zum Einsatz kamen. In der Studie von **Roedl 2008a** [105] blieb unklar, ob nach einer 18-monatigen Nachbeobachtung nur die PET-negativen Patienten oder auch PET-positive Patienten die EUS anstelle der Biopsie als Referenztest erhielten. Bei **Teyton 2009** [106] blieb unklar, ob der Indextest (FDG-PET) auch Teil des Referenztests (klinisch-radiologische Nachbeobachtung) war. Da der Indextest Ergebnisse aller FDG-PET während der Nachbeobachtung einschloss, war die Zeitdauer zwischen Index- und Referenztests variabel und wurde nicht berichtet. Die Ergebnisse von PET wurden mit denen der CT verglichen, allerdings fehlten Angaben, wie die CT durchgeführt wurde. In der Studie von **Zhang 2006** [107] blieb unklar, ob und ggf. wie Patienten nach dem Indextest behandelt wurden. Es wurde nicht berichtet, ob die Befundung der Tests verblindet erfolgte. Der Studie konnten nur unzureichende Angaben zum Referenztest (Histologie oder klinische Nachbeobachtung über mindestens 6 Monate) entnommen werden. Dadurch konnten eine differenzielle Verifikation oder ein Inkorporationsbias nicht ausgeschlossen werden.

Tabelle 54: Verzerrungspotenzial der eingeschlossenen Studien bei der Rezidivdiagnostik

Studie	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Verzerrungspotenzial der Ergebnisse
Guo 2007	+	?	+	+	-	+	-	+	+	?	hoch
Roedl 2008a	+	?	+	+	-	+	?	+	+	?	hoch
Teyton 2009	+	?	+	+	-	?	+	+	+	?	niedrig
Zhang 2006	?	?	+	+	?	?	?	?	+	?	hoch
+: Kriterium erfüllt; -: Kriterium nicht erfüllt; ?: unklar											

5.5.3.2 Einschätzung der Übertragbarkeit der Ergebnisse der Primärstudien auf den deutschen Versorgungskontext

In allen 4 eingeschlossenen Studien konnten Anhaltspunkte gefunden werden, die eine Übertragbarkeit der Ergebnisse für die Rezidivdiagnostik auf den deutschen Versorgungskontext möglicherweise einschränken:

- In Zhang 2006 [107] konnten nicht ausreichend Details über die Durchführung des Referenztests und die Länge der Nachbeobachtung entnommen werden. Damit ist eine Replikation des Tests erschwert.
- Die Bewertung der Ergebnisse der Indextests in den Studien von Guo 2007 [78], Roedl 2008a [105] und Teyton 2009 [106] wurde gemeinsam bzw. auch unabhängig voneinander durch mehrere Bewerter gemacht.

Genannte Aspekte entsprechen möglicherweise nicht dem klinischen Alltag, da die Interpretation der Bilder in der Regel nicht von 2 Klinikern unabhängig voneinander erfolgt. Die nachfolgende Tabelle 55 stellt die Ergebnisse der Einschätzung der Übertragbarkeit der Ergebnisse der Primärstudien auf den deutschen Versorgungskontext bei der Rezidivdiagnostik dar.

Tabelle 55: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Studien bei der Rezidivdiagnostik

Studie	F1	F2	F3	F4	F5	F6	F7	Generelle Einschätzung der Übertragbarkeit
Guo 2007	-	-	+	-	-	-	-	Anhaltspunkt für eine Nichtübertragbarkeit
Roedl 2008a	-	-	+	-	-	-	-	Anhaltspunkt für eine Nichtübertragbarkeit
Teyton 2009	-	-	+	-	-	-	-	Anhaltspunkt für eine Nichtübertragbarkeit
Zhang 2006	-	+	-	-	-	-	-	Anhaltspunkt für eine Nichtübertragbarkeit
+: konkrete Anhaltspunkte für eine Nichtübertragbarkeit liegen vor								
-: keine Anhaltspunkte für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden)								

5.5.3.3 Ergebnisse der Primärstudien zur diagnostischen Güte bei der Rezidivdiagnostik

In Tabelle 56 werden die Ergebnisse zur diagnostischen Güte der Primärstudien bei der Rezidivdiagnostik dargestellt. Neben den jeweiligen Schätzern wurden auch die Konfidenzintervalle berichtet. Sofern diese nicht in der Studie präsentiert wurden, wurden sie eigenständig nach der Methode von Clopper und Pearson 1934 [109] berechnet.

In den 4 Studien wurden 5 Analysen zur diagnostischen Güte der PET bzw. PET/CT durchgeführt. Der Einsatz der PET wurde in 2 Studien (Roedl 2008a und Teyton 2009) untersucht, wohingegen die PET/CT in 3 Studien untersucht wurde (Guo 2007, Roedl 2008a und Zhang 2006). Die PET hatte eine Sensitivität von 89 % (Spezifität 55 %) und 100 % (Spezifität 88,9 %). Die Präzisionen aller Schätzungen waren sehr gering und die Konfidenzintervalle entsprechend weit.

Die Sensitivität der PET/CT lag im Bereich von 89 % (bei einer Spezifität von 75 %) bis 100 % (bei einer Spezifität von 85,7 %). Die Spezifität bewegte sich in einer Spanne von 54,5 % (bei einer Sensitivität von 95,6 %) bis 85,7 (bei einer Sensitivität von 100 %). Die Ergebnisse der CT konnten nicht dargestellt werden. In der Studie wurden für die CT eine Sensitivität von 61,7 % und eine Spezifität von 78,6 % angegeben, die sich auf Basis der eingeschlossenen Patientenzahl von $n = 31$ nicht rekonstruieren ließen. Die Präzisionen der Schätzer der Sensitivität waren gering und die Konfidenzintervalle entsprechend weit. Außerdem waren die Konfidenzintervalle der Schätzer der Spezifität sehr gering und die Konfidenzintervalle entsprechend sehr weit.

Tabelle 56: Ergebnisse der Studien zur diagnostischen Güte der PET bzw PET/CT bei der Rezidivdiagnostik

Studie	n (N)	Indextest	Cut-off	Referenztest	Cut-off	RP	FN	FP	RN	Sensitivität % (95 %-KI)	Spezifität % (95 %-KI)
Guo 2007	56 (56)	FDG-PET/CT	Rezidiv pro Patient	Histologie oder klinische Nachbeobachtung (min. 6 Monate)	Jedes Rezidiv	43	2	5	6	95,6 (84,9 bis 99,5) ^{a,b}	54,5 (23,4 bis 83,3) ^{a,b}
Roedl 2008a	47 (47)	FDG-PET	Rezidiv pro Patient	Biopsie, klinische Nachbeobachtung (Median 25 Monate), nach 18 Monaten: EUS	Jedes Rezidiv	24	3	9	11	89 (70,8 bis 97,6) ^{a,b}	55 (31,5 bis 76,9) ^b
		FDG-PET/CT				24	3	5	15	89 (70,8 bis 97,6) ^{a,b}	75 (50,9 bis 91,3) ^{a,b}
Teyton 2009	41 (41)	FDG-PET	Rezidiv pro Patient	Biopsie oder klinische Nachbeobachtung	Jedes Rezidiv	23	0	2 ^a	16 ^a	100 (85,2 bis 100) ^{a,b}	88,9 ^a (65,3 bis 98,6) ^{a,b}
Zhang 2006	31 ^c (31)	FDG-PET/CT	Läsionen pro Patient mit SUV > 2,5	Histologie oder klinische Nachbeobachtung > 6 Monate	Jedes Rezidiv	23 ^a	0 ^a	1 ^a	6 ^a	100 (85,2 bis 100) ^{a,b}	85,7 (42,1 bis 99,6) ^{a,b}
	31 (31)	CT	k. A.	Histologie oder klinische Nachbeobachtung > 6 Monate	Jedes Rezidiv	_d	_d	_d	_d	_d	_d

a: Wert(e) wurde(n) selbst berechnet.
 b: Keine Angabe des 95 %-KI in der Publikation.
 c: 6 Patienten starben während der Nachbeobachtung. Einer davon war PET/CT negativ. Da Sensitivität mit 100 % angegeben wird, gehen wir davon aus, dass dieser nicht krebbsbedingt verstarb und daher nicht in der Vierfeldertafel eingeschlossen wurde.
 d: Im Text wurden eine Sensitivität von 61,7 % und eine Spezifität von 78,6 % angegeben. Mit einer Patientenzahl von n = 31 lassen sich diese Werte nicht rekonstruieren.
 CT = Computertomographie; FDG = ¹⁸Fluorodeoxyglukose; FN = falsch negativ; FP = falsch positiv; k. A. = keine Angabe(n); KI = Konfidenzintervall; PET = Positronenemissionstomographie; RN = richtig negativ; RP = richtig positiv; TBR = tumor to background ratio (Verhältnis von FDG-Aufnahme im Tumor zu Hintergrund)

Univariate Ergebnisse für die diagnostische Güte der PET

In der nachfolgenden Abbildung 8 werden die Ergebnisse der diagnostischen Güte der PET bzw. PET/CT bei der Rezidivdiagnostik von Ösophaguskarzinomen univariat dargestellt. ImForestPlot werden die Ergebnisse der 4 Studien getrennt nach Indextest dargestellt.

Ein indirekter Vergleich zwischen PET und PET/CT zeigt, dass beide Technologien hohe Werte für Sensitivität und mittlere Werte für die Spezifität erreichten.

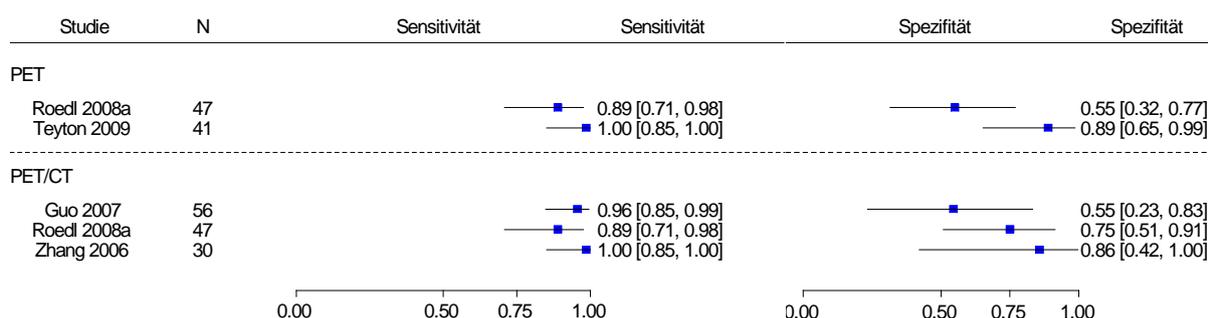


Abbildung 8: Forest Plot der Ergebnisse für die diagnostische Güte der PET bei der Rezidiverkennung von Ösophaguskarzinomen

Direkte Vergleiche zwischen PET bzw. PET/CT und anderen bildgebenden Verfahren

Ein direkter Vergleich zwischen PET und PET/CT wurde nur in der Studie Roedl 2008a [105] durchgeführt. Dieser zeigte, dass die PET und die PET/CT die gleichen Werte für Sensitivität erzielten (89 %), die Konfidenzintervalle für die Schätzer der Spezifität schließen jeweils den Schätzer der anderen Technologie ein. Zhang 2006 untersuchte den direkten Vergleich zwischen der PET/CT und CT. Aus der Studie konnten allerdings nur die Punktschätzer für Sensitivität und Spezifität der PET/CT abgeleitet werden (100 %, 95 %-KI [85; 100] und 85,7 %, 95 %-KI [42; 100]). In der Studie wurden für die CT-Werte für Sensitivität eine Spezifität angegeben, die sich auf Basis der eingeschlossenen Patientenzahl von $n = 31$ nicht rekonstruieren ließen. Auf dieser Basis kann keine Meta-Analyse gerechnet werden.

Subgruppenmerkmale und andere Effektmodifikatoren

Es wurden keine Subgruppenmerkmale in den eingeschlossenen Studien berichtet.

Sensitivitätsanalysen

Es wurden keine Sensitivitätsanalysen durchgeführt.

Ergänzend betrachtete Ergebnisse zu Änderungen im klinischen Management

In der Studie von Teyton 2009 [106] wurde eine Änderung im klinischen Management basierend auf dem Einsatz der PET (siehe Tabelle 57) beschrieben. Durch den Einsatz der

PET wurden Fernmetastasen in Lunge, Leber oder Knochen entdeckt, die durch den Einsatz der CT nicht erkannt wurden. Dies führte bei 6 von 41 untersuchten Patienten (15 %) zu einer Änderung im klinischen Management. In der Studie wurde berichtet, dass basierend auf dem Einsatz der PET 5 Patienten mit einem Rezidiv noch eine zusätzliche Chemotherapie bekamen.

Tabelle 57: Ergänzend betrachtete Ergebnisse zu Änderungen im klinischen Management im Vergleich zur CT bei der Rezidivdiagnostik

Studie	Indextest	Anzahl der Patienten	Patienten mit Managementänderung: n (%)	Angaben zur Managementänderung
Teyton 2009	FDG-PET	41	6 (15)	bei 5 Patienten mit Rezidiv zusätzlich Chemotherapie
FDG = ¹⁸ Fluorodeoxyglukose; PET = Positronenemissionstomographie				

5.5.4 Prognostische Güte auf Grundlage der Primärstudien bei der Rezidivdiagnostik

Im Rahmen der Ergänzungsrecherche wurde eine Primärstudie zur prognostischen Güte (Guo 2007 [78]) für die Indikation Rezidivdiagnostik identifiziert.

5.5.4.1 Studiendesign und Studienpopulation der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik

Wie erwähnt, betrachtete die Arbeit von **Guo 2007** [78] sowohl eine diagnostische als auch eine prognostische Fragestellung, sodass im vorliegenden Bericht beide Ergebnisse wiedergegeben werden. In diesem Abschnitt werden die Ergebnisse der prognostischen Güte der PET/CT dargestellt. Es wurden 56 Patienten mit behandelten Plattenepithelkarzinomen des Ösophagus eingeschlossen, bei denen der Verdacht auf ein Rezidiv bestand. In diesem Abschnitt wird die Bewertung der prognostischen Güte von FDG-PET zur Vorhersage des krankheitsfreien Überlebens und des Gesamtüberlebens erläutert. Dabei wurden verschiedene Faktoren hinsichtlich ihres Einflusses in der Vorhersage des Gesamtüberlebens untersucht. Eine klinische Nachbeobachtung mit einer Länge von 4 bis 34,5 Monaten (Median 24,5 Monate) wurde als Referenztest eingesetzt.

Tabelle 58: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik

Studie	Studiendesign	Studienziel ^a	Zielkriterium	Evidenzstufe ^b	Verzerrungspotenzial
Guo 2007	Diagnostische und prognostische Kohortenstudie Rekrutierung der Patienten zwischen Juni 2003 und Juni 2005. China	Bewertung der prognostischen Güte von PET/CT bei Patienten nach Abschluss der Behandlung eines Plattenepithelkarzinoms.	Prognostische Güte von FDG-PET/CT zur Vorhersage des Gesamtüberlebens.	II ^c	hoch

a: Studienziel gemäß Formulierung der Autoren
b: entspricht der Evidenzstufe des G-BA
c: Die Verfahrensordnung des G-BA sieht derzeit keine Evidenzstufe für prognostische Studien vor. Sie werden im Rahmen dieses Berichts daher mit Stufe 2 gekennzeichnet.
CT = Computertomographie; FDG = ¹⁸Fluorodeoxyglukose; PET = Positronenemissionstomographie;

Tabelle 59: Index- und Referenztest der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik

Studie	Indextest	Referenztest
Guo 2007	PET/CT (Discovery LS PET/CT system; GE Healthcare, USA) 60 Minuten nach i.v.-Injektion von 370 MBq FDG. Aufnahme von Kopf bis oberer Oberschenkel. Schichtdicke 4,25 mm. CT-Parameter: 120 kV, 90 mA, 6:1 pitch, Schichtdicke 4,25 mm, 0,8 Sekunden pro Rotation. Befundung im Konsens durch 3 erfahrene Nuklearmediziner auf Xeleris Workstation (GE Healthcare, USA). Keine Verblindung gegenüber klinischen Daten oder Befunden anderer bildgebender Verfahren. Abschließende Bewertung (positiv, grenzwertig, negativ) vorwiegend visuell (erhöhte FDG-Aufnahme im Vergleich zum umgebenden Gewebe) unter Hinzuziehung semiquantitativer Ergebnisse (SUV \geq 2,5).	Gesamtüberleben (Zeit von Behandlungsende bis Tod jeglicher Ursache). Mediane Nachbeobachtung 24,5 Monate (Spannweite 4 bis 34,5 Monate).

CT = Computertomographie; FDG = ¹⁸Fluorodeoxyglukose; GE = General Electric; i.v. = intravenös; mA = Milliampere; mAs = Milliamperesekunde; MBq = Mega Becquerel; PET = Positronenemissionstomographie; SUV = standardized uptake value; USA = Vereinigte Staaten von Amerika

Tabelle 60: Eingeschlossene Patienten der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik

Studie	Patienten [n]	Geschlecht [n (m/w)]	Alter (MW, Median, Spannweite)	Diagnose	Vorbehandlung
Guo 2007	56	56 (47/9)	k. A.; k. A.; 38- 77	Histologie: Plattenepithelkarzinom (56) Differenzierungsgrad: guter (11), mittlerer (21), niedriger (15), unklar (9). Lokalisation: obere (6), mittlere (34), untere Speiseröhre (16) Stadium (TNM): IIb (5), III (38), IV (13)	Abschluss der Behandlung (26 operativ, 30 Strahlentherapie) vor Aufnahme in die Studie. Wenn erforderlich, Gabe von 5- Fluoruracil als Erhaltungs- Chemotherapie.
K. A.= keine Angabe(n); m= männlich; M= Metastase; N= Nodus ([Lymph-] Knoten); n= Anzahl; T= Tumor; w= weiblich					

5.5.4.2 Einschätzung des Verzerrungspotenzials der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik

Der Schwellenwert für SUV wurde in der Studie von **Guo 2007** [78] im Rahmen der Studie (in Kenntnis der Ergebnisse) definiert. Außerdem war unklar, ob und ggf. wie die Patienten während der klinischen Nachbeobachtung behandelt wurden. Die eingeschlossenen Patienten befanden sich zu Studienbeginn in unterschiedlichen Krankheitsstadien.

Tabelle 61: Verzerrungspotenzial der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik

Studie	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	Verzerrungspotenzial der Ergebnisse
Guo 2007	+	-	+	+	+	+	+	-	+	?	+	+	+	-	?	+	+	hoch
+: Kriterium erfüllt; -: Kriterium nicht erfüllt; ?: unklar																		

5.5.4.3 Einschätzung der Übertragbarkeit der Ergebnisse der Prognosestudie auf den deutschen Versorgungskontext bei der Rezidivdiagnostik

Neben der Bewertung des Verzerrungspotenzials wurde auch für die Prognosestudie bei der Rezidivdiagnostik die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext evaluiert (siehe Tabelle 62).

Die Bewertung der Ergebnisse der Indextests von **Guo 2007** [78] wurde gemeinsam bzw. auch unabhängig voneinander durch mehrere Bewerter gemacht. Damit konnte ein Anhaltspunkt gefunden werden, der eine Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext einschränken könnte.

Tabelle 62: Anhaltspunkte für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik

Studie	F1	F2	F3	F4	F5	F6	F7	Generelle Einschätzung der Übertragbarkeit
Guo 2007	-	-	+	-	-	-	-	Anhaltspunkt für eine Nichtübertragbarkeit
+: konkrete Anhaltspunkte für eine Nichtübertragbarkeit liegen vor								
-: keine Anhaltspunkte für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden)								

5.5.4.4 Ergebnisse der Primärstudien zur prognostischen Güte bei der Rezidivdiagnostik

In der folgenden Tabelle werden die Ergebnisse der ein- und multifaktoriellen Analysen für die eingeschlossene Prognosestudie bei der Rezidivdiagnostik dargestellt. Es wurde eine Vielzahl an prognostischen Merkmalen untersucht, wobei in der folgenden Tabelle nur die PET-relevanten Merkmale dargestellt werden. Alle weiteren Merkmale werden im Anhang E präsentiert.

In der einfaktoriellen Analyse bei Guo 2007 waren Primärtumore im oberen sowie im mittleren Ösophagus-Drittel, primär operative Behandlung, $SUV \geq 9,26$ und eine systemische Ausbreitung der Krankheit statistisch signifikante Prädiktoren des Gesamtüberlebens. Diese Faktoren wurden in die multifaktorielle Analyse eingeschlossen, wobei außer Primärtumoren im mittleren Drittel alle eingeschlossenen Faktoren statistisch signifikant blieben.

Tabelle 63: Ergebnisse der Studien zur prognostischen Güte der PET bei der Rezidivdiagnostik

Studie	Zielgröße	n (N)	Prognostisches Merkmal	Cut-off	Patienten mit prog. Merkmal	Ergebnisse aus einfaktorieller Analyse		Ergebnisse aus multifaktorieller Analyse	
						p-Wert	HR	p-Wert	HR
Guo 2007	Gesamtüberleben	56 (56)	SUV	$\geq 9,26$	26 / 56	0,002	1,105	0,029	1,071
HR= hazard ratio; SUV= standardized uptake value									

5.6 Zusammenfassung der Ergebnisse

Patientenrelevanter Nutzen

Durch die systematische Recherche nach publizierter Literatur konnte für keine Fragestellung des vorliegenden Berichts eine vergleichende Studie identifiziert werden, die eine Aussage zum patientenrelevanten (Zusatz-)Nutzen der PET bzw. PET/CT für die Fragestellungen des Primärstaging, des Restagings nach Abschluss der Therapie und der Rezidivdiagnostik ermöglicht hätte. Auch die Suche in Kongressbänden und Studienregistern lieferte keine Hinweise auf noch laufende vergleichende Studien.

Diagnostische und prognostische Güte

3 Evidenzsynthesen und 30 Primärstudien aus der Ergänzungsrecherche entsprachen den Einschlusskriterien des vorliegenden Berichts. Aus den Evidenzsynthesen, die ausschließlich die Indikation Restaging untersuchten, entsprachen 11 Primärstudien den Einschlusskriterien des vorliegenden Berichts. Damit basiert der Bericht auf insgesamt 41 Primärstudien: Primärstaging (n = 23), Restaging (n = 14) und Rezidivdiagnostik (n = 4).

Primärstaging

Für das N-Staging wurden direkte Vergleiche zwischen PET und CT in 12 Studien untersucht. In den direkten bivariaten Meta-Analysen für das N-Staging konnte kein statistisch signifikanter Unterschied zwischen der diagnostischen Güte von PET und CT festgestellt werden.

Für das M-Staging wurden direkte Vergleiche zwischen PET und CT in 5 Studien untersucht. In den direkten bivariaten Meta-Analysen konnte kein statistisch signifikanter Unterschied zwischen der diagnostischen Güte von PET und CT festgestellt werden. Die Robustheit dieser gepoolten Ergebnisse wurde in mehreren Sensitivitätsanalysen bestätigt.

In einer prognostischen Studie wurde ein direkter Vergleich von FDG-PET und TI-SPECT durchgeführt. Keiner der PET- sowie SPECT relevanten Merkmale zeigte nach der multifaktoriellen Modellierung einen statistisch signifikanten Zusammenhang.

Restaging

Von den insgesamt 14 eingeschlossenen Studien untersuchten 2 Studien neben der diagnostischen auch die prognostische Güte der PET bzw. PET/CT. Eine Studie untersuchte ausschließlich die prognostische Güte der PET.

Sechs der aus den Evidenzsynthesen stammenden Studien untersuchten die PET bzw. PET/CT in der Indikation des Interimsstaging. Drei dieser Studien untersuchten den Einsatz der PET während einer neoadjuvanten Chemotherapie. Hier zeigte sich für die PET eine Spannweite der Punktschätzer der Sensitivitäten von 80 % (bei einer zugehörigen Spezifität von 78 % bzw. 100 %) bis 88 % (zugehörige Spezifität 75 %). Für die Punktschätzer der Spezifitäten ergab sich entsprechend eine Spannweite von 75 % (zugehörige Sensitivität

88 %) bis 100 % (zugehörige Sensitivität 80 %). Für alle Punktschätzer waren weite Konfidenzintervalle zu beobachten. In den weiteren 3 Studien wurde der Einsatz der PET bzw. PET/CT während einer neoadjuvanten Chemoradiotherapie untersucht. Allerdings konnten für 2 Studien die Ergebnisse nicht dargestellt werden, weil der Anteil fehlender Werte größer als 30 % war. In einer Studie ergab sich eine Sensitivität von 44 % und eine Spezifität von 52 %. Auch hier waren die Konfidenzintervalle weit.

Für das Restaging nach Abschluss einer neoadjuvanten Therapie wurde in 3 Studien ein direkter Vergleich diagnostischer Technologien untersucht. Der Vergleich PET bzw. PET/CT und Endosonografie (EUS) wurde in 2 Studien untersucht, wovon eine Studie jedoch nur 10 Patienten mit der PET untersuchen konnte. In der zweiten Studie zeigte sich für die PET/CT eine Sensitivität von 87 % (95 %-KI [60; 98]) und eine Spezifität von 88 % (95 %-KI [72; 97]). Im Gegensatz dazu hatte die EUS eine Sensitivität von 20 % (95 %-KI [4; 48]) und eine Spezifität von 94 % (95 %-KI [80; 99]).

Rezidivdiagnostik

Für die Rezidivdiagnostik konnte keine Evidenzsynthese identifiziert werden. Im Rahmen der Ergänzungsrecherche wurden 4 Primärstudien eingeschlossen. Alle 4 Studien untersuchten die diagnostische Güte der PET bzw. PET/CT. Eine Studie untersuchte zusätzlich noch die prognostische Güte.

Zum direkten Vergleich zwischen PET bzw. PET/CT und konventioneller Diagnostik lagen keine Ergebnisse vor.

6 Diskussion

6.1 Gesamtbewertung der Evidenzlage

6.1.1 Studien zur Untersuchung des patientenrelevanten Nutzens

Im Krankheitsbild der Ösophaguskarzinome fand sich keine vergleichende Studie, aus der eine Aussage zum direkt kausal bedingten patientenrelevanten (Zusatz-)Nutzen der PET oder PET/CT abgeleitet werden konnte. Auch die Suche in Studienregistern und Kongressbänden lieferte keine Hinweise auf aktuell laufende vergleichende Studien. Diese Tatsache ist vor allem vor dem Hintergrund der relativ hohen Inzidenz dieser Erkrankung bedauerlich.

Bisher gibt es jedoch keine sicheren Belege dafür, dass durch die Bildgebung mittels PET bzw. PET/CT diese Ziele tatsächlich erreicht werden. Die wenigen identifizierten Studien, die über Managementänderungen aufgrund der PET bzw. PET/CT berichteten, führten nicht aus, ob die PET-bedingten Behandlungsänderungen Effekte auf patientenrelevante Endpunkte hatten [79,88,96,106]. Diese Problematik bestand beispielsweise auch in den MUNICON-I- und II-Studien [110,111], die im Abschnitt 6.1.3 näher beschrieben werden.

Für das Krankheitsbild der Ösophaguskarzinome ist es deshalb dringend erforderlich, den patientenrelevanten Nutzen der Diagnostik mittels PET oder PET/CT in prospektiv geplanten, randomisierten kontrollierten Studien zu untersuchen [112].

6.1.2 Diagnostische und prognostische Güte der PET und PET/CT beim Primärstaging

Für die Indikation Primärstaging wurden 23 Primärstudien in den vorliegenden Bericht eingeschlossen. Die Bewertung der PET bzw. PET/CT erfolgte auf Basis der Ergänzungsrecherche. Es konnten zwar 2 Evidenzsynthesen von Westreenen 2004 [64] und van Vliet 2008 [113] für die Indikation identifiziert werden, sie wurden aber nicht eingeschlossen, da beide Mängel auf dem Qualitätsindex von Oxman und Guyatt 1991 [38,39] aufwiesen. Allein anhand der Anzahl klinischer Studien lässt sich vermuten, dass die Indikation Primärstaging von Ösophaguskarzinomen offenbar mit den größten Hoffnungen der Anwender von PET- und PET/CT-Geräten verbunden ist, hier einen (zusätzlichen) Nutzen für die Patienten erzielen zu können.

Im Gegensatz zu van Westreenen 2004 [64], der wie die vorliegende Arbeit nur nach Studien recherchierte, die den Einsatz der PET bzw. PET/CT untersuchten, suchte van Vliet 2008 [113] zusätzlich nach Studien, die den Einsatz der EUS und CT bei Patienten mit einem Ösophaguskarzinom analysierten. Im Kontrast zum vorliegenden Bericht schlossen beide Evidenzsynthesen sowohl retrospektive als auch prospektive Studien ein und berichteten über indirekte Vergleiche.

Van Vliet et al. 2008 [113] poolten 9 Studien, die den Einsatz der FDG-PET für das M-Staging untersuchten. Von diesen entsprachen 6 Studien den Einschlusskriterien des vorliegenden Berichts. Allerdings wurden nur 4 Studien [81,83,99,101] in die bivariate Meta-

Analyse des vorliegenden Vorberichts eingeschlossen, weil in diesen Studien ein direkter Vergleich zwischen der PET und einer anderen Technologie durchgeführt wurde. Für das N-Staging poolten van Vliet et al. 2008 10 Studien, von denen ebenfalls 6 den Einschlusskriterien des vorliegenden Berichts entsprachen. Fünf [56,83,95,99,101] dieser 6 Studien wurden aufgrund der Durchführung eines direkten Vergleichs in die bivariaten Meta-Analyse für das N-Staging im vorliegenden Bericht eingeschlossen.

Van Westreenen et al. 2004 [64] poolten 12 Studien, die N-Staging und ebenfalls M-Staging untersuchten. Von diesen wurden 7 für N- und M-Staging in den vorliegenden Bericht eingeschlossen, wobei 4 [56,81,88,101] dieser 7 Studien in der bivariaten Meta-Analyse für das N-Staging und M-Staging berücksichtigt wurden, weil in diesen ein direkter Vergleich der Technologien durchgeführt wurde.

Vergleicht man die gepoolten Ergebnisse der PET des vorliegenden Vorberichts mit den gepoolten Ergebnissen der 2 Evidenzsynthesen, so kann festgestellt werden, dass die gepoolten Schätzer für PET vor allem angesichts der breiten Konfidenzintervalle ähnlich sind und sich die Ergebnisse nicht widersprechen.

Im vorliegenden Bericht wurden bivariate Meta-Analysen nur für direkte Vergleiche durchgeführt. Direkte Vergleiche zwischen der PET, der PET/CT und anderen bildgebenden Verfahren geben Auskunft darüber, welche diagnostische Güte mit der PET bzw. PET/CT als Substitut anderer Verfahren erzielt werden könnte. Für das N- und M-Staging lagen genügend Studien vor, die sowohl PET (Indextest I) als auch CT als Indextest II untersuchten. Ein Vergleich der gepoolten Schätzer für die CT aus dem vorliegenden Bericht mit denen aus van Vliet 2008 [113] sollte vor dem Hintergrund erfolgen, dass die Evidenzsynthese indirekte Vergleiche durchführte und der vorliegende Bericht nur direkte Vergleiche untersuchte. Dennoch, trotz der unterschiedlichen Methodik (direkter versus indirekter Vergleich) wurden ähnliche Ergebnisse erzielt.

Die gepoolten Ergebnisse der 2 Evidenzsynthesen sowie die Ergebnisse des vorliegenden Berichts werden in Tabelle 64 dargestellt. Die 95 %-Konfidenzregionen der gepoolten Ergebnisse aus dem vorliegenden Bericht überlappen sich, sodass keine Aussage darüber getroffen werden kann, ob sich die Verfahren in der erwarteten Sensitivität und Spezifität unterscheiden.

Tabelle 64: Ergebnisse der Evidenzsynthesen Van Westreenen 2004, Van Vliet 2008 und des IQWiG-Vorberichts beim Primärstaging bei Ösophaguskarzinomen

		Sensitivität in % (95 %-KI)	Spezifität in % (95 %-KI)
<i>N-Staging</i>			
IQWiG-Vorbericht^a	PET	59,9 (51 bis 68)	85,8 (79 bis 91)
	CT	49,6 (40 bis 59)	86,4 (80 bis 91)
Van Westreenen 2004	PET	51 (34 bis 69)	84 (76 bis 91)
Van Vliet 2008	PET	57 (43 bis 70)	85 (76 bis 95)
	CT	50 (41 bis 60)	83 (77 bis 89)
	EUS	80 (75 bis 84)	70 (65 bis 75)
<i>M-Staging</i>			
IQWiG-Vorbericht^a	PET	60,7 (42 bis 77)	96,5 (91 bis 99)
	CT	29,8 (12 bis 57)	95,7 (91 bis 98)
Van Westreenen 2004	PET	67 (58 bis 76)	97 (90 bis 100)
Van Vliet 2008	PET	71 (62 bis 79)	93 (89 bis 97)
	CT	52 (33 bis 71)	91 (86 bis 96)
a: Die gepoolten Ergebnisse stammen aus den direkten Vergleichen zum N- und M-Staging.			

Van Vliet 2008 wiesen darüber hinaus in ihrer Diskussion darauf hin, dass der PET besondere Bedeutung im Bereich der Erkennung von Fernmetastasen bei Patienten mit Ösophaguskarzinomen zukommt. Durch den Einsatz der PET sollten bis zu 20 % mehr Fernmetastasen erkannt werden als durch den alleinigen Einsatz der konventionellen Diagnostik. Diese Aussage konnte durch die bivariate Meta-Analyse des vorliegenden Berichts nicht bestätigt werden, da kein statistisch signifikanter Unterschied zwischen den beiden Technologien festgestellt werden konnte.

Van Westreenen 2004, van Vliet 2008 und der vorliegende Bericht fanden Ergebnisse zum klinischen Management. Es wurde diskutiert, dass die Ergebnisse von einarmigen, diagnostischen oder prognostischen Studien nahelegen, dass sich durch den Einsatz der PET möglicherweise eine bessere dem jeweiligen Stadium angepasste Therapie realisieren lässt. Bisher gibt es jedoch keine belastbaren Belege dafür, dass durch die Bildgebung mittels PET bzw. PET/CT dieses Ziel tatsächlich erreicht wird.

6.1.3 Diagnostische und prognostische Güte der PET und PET/CT beim Restaging

Für die Indikation Restaging konnten insgesamt 14 Studien eingeschlossen werden. Von diesen stammen 11 Primärstudien aus den eingeschlossenen Evidenzsynthesen [53,62,63]. Drei Studien wurden im Rahmen der Ergänzungsrecherche gesichtet und eingeschlossen

[77,103,104]. Die Primärstudien, die im Rahmen der Evidenzsynthese eingeschlossen wurden, untersuchten mehrheitlich die nichtintegrierte FDG-PET. In einer Studie wurde auch die PET/CT untersucht.

Aus den 3 berücksichtigten Evidenzsynthesen konnten nicht alle Primärstudien eingeschlossen werden, da sich eine Reihe von Primärstudien fand, die ein retrospektives Design aufwies. Darüber hinaus konnten nicht von allen Studien die Daten der Vierfeldertafel rekonstruiert werden.

Alle 3 eingeschlossenen Evidenzsynthesen hatten eine hohe Qualität nach Oxman und Guyatt 1991 [38,39]. Zwei der eingeschlossenen Evidenzsynthesen [53,63] verwendeten das QUADAS-Instrument, eine Evidenzsynthese [62] das Cochrane-Instrument. Ein Vergleich der Bewertungen ist damit erschwert, weil die einzelnen Items der 2 Instrumente nicht identisch sind. Alle Evidenzsynthesen gaben einen Score an, definierten aber nicht, ab wann sie Studien als verzerrt einstufen.

Direkte Vergleiche mit weiteren Technologien innerhalb der Studien wurden nur in 3 Studien berichtet. Die Vergleiche beinhalteten unterschiedliche Technologien, sodass keine bivariate Meta-Analyse für diese Indikation berechnet werden konnte. Vergleichbar zur Evidenzsynthese von Ngamruengphong 2010 waren auch die Ergebnisse beim Restaging des vorliegenden Berichts sehr heterogen: Die Studien schlossen verschiedene Patientenspektren ein und wählten unterschiedliche Referenztests (z. B. unterschiedliche klinische Nachbeobachtungszeiten). Darüber hinaus verwendeten die Studien unterschiedliche Trennwerte, welches eine Vergleichbarkeit der Studien untereinander erschwerte.

Die Evidenzsynthese von Rebollo Aquirre 2009 [62] untersuchte nur die diagnostische Güte der FDG-PET beim Restaging. Im Gegensatz dazu verfolgte die Evidenzsynthese von Westerterp 2005 [63] das Ziel, den Einsatz der CT, EUS und FDG-PET beim Restaging von Ösophaguskarzinomen zu analysieren. Ngamruengphong 2010 [53] untersuchte die EUS und die FDG-PET. Beide letztgenannten Evidenzsynthesen führten dementsprechend eine umfangreiche Recherche nach Studien durch, die PET und/oder CT und/oder EUS untersuchten. Die Ergebnisse wurden in beiden Evidenzsynthesen nach Technologie getrennt dargestellt. Die Ergebnisse wurden in den Diskussionsabschnitten der Evidenzsynthesen indirekt miteinander verglichen und erläutert. Indirekte Vergleiche bergen jedoch ein sehr hohes Verzerrungspotenzial, insbesondere bei Studien zur diagnostischen Güte (keine Randomisierung) und Studien mit geringen Fallzahlen, sodass sie im vorliegenden Bericht nicht angestellt wurden. Daher sollten diese Vergleiche aus den Evidenzsynthesen äußerst kritisch betrachtet werden.

In der folgenden Tabelle 65 werden die Schlussfolgerungen der Autoren der 3 eingeschlossenen Evidenzsynthesen zusammengefasst.

Tabelle 65: Schlussfolgerungen der eingeschlossenen Evidenzsynthesen beim Restaging von Ösophaguskarzinomen

Evidenzsynthese	Schlussfolgerungen zur Indikation des Restagings bei Ösophaguskarzinomen ^a
Rebollo Aquirre 2009	Die FDG-PET scheint das bestverfügbare bildgebende Verfahren für das Behandlungsansprechen einer neoadjuvanten Therapie für Ösophaguskarzinome zu sein. Prospektive Studien mit höheren Patientenzahlen zur diagnostischen und prognostischen Güte werden gefordert, um die Ergebnisse der Technologie zu bestätigen und um Responder und Nichtresponder zu bestimmen.
Ngamruengphong 2010	Die Untersuchung zeigte, dass die diagnostische Güte der EUS und FDG-PET beim Behandlungsansprechen bei Patienten mit Ösophaguskarzinomen ähnlich war. Jede Technologie hat Vor- und Nachteile. Sie sollten als sich ergänzende Technologien und nicht als konkurrierende Technologien betrachtet werden. Die diagnostische Güte der PET beim Interimsstaging und die der nach der neoadjuvanten Therapie sind vergleichbar. Das Interimsstaging sollte aber bevorzugt werden, weil Nichtresponder von Respondern unterschieden werden können und Patienten dadurch eine Chemoradiotherapie erspart bleiben könnte.
Westerterp 2005	Die CT hat eine niedrige diagnostische Genauigkeit beim Behandlungsansprechen bei Patienten mit einem Ösophaguskarzinom. EUS und FDG-PET haben eine ähnlich hohe diagnostische Güte. Allerdings ist eine EUS nach einer Chemo- oder Radiotherapie nicht immer realisierbar. Studien mit hohen Patientenzahlen und ausreichender Power, welche die Detektionsraten der unterschiedlichen Technologien direkt vergleichen, werden gefordert.
a: Schlussfolgerungen gemäß Formulierung der Autoren	

Insgesamt äußerten sich die Autoren der eingeschlossenen Evidenzsynthesen eher zurückhaltend. Keine Autorengruppe gab eine ausdrückliche Empfehlung zum Einsatz der PET bei Ösophaguskarzinomen ab. In 2 Evidenzsynthesen wurden weitere qualitativ hochwertige Studien mit hohen Patientenzahlen gefordert. Keine der Evidenzsynthesen fordert prospektiv vergleichende Studien mit patientenrelevanten Endpunkten.

Therapieansprechen mittels PET bzw. PET/CT während einer neoadjuvanten Therapie (Interimsstaging)

In der Indikation des Interimsstagings bei neoadjuvanter Chemotherapie (NAC) und neoadjuvanter Chemoradiotherapie (NACR)¹ ergab sich aus der Gesamtschau der identifizierten Studien eine besondere Konstellation:

Unter der Annahme, dass das Ziel einer NAC(R) (einzig) darin besteht, die Tumorlast zu reduzieren, um eine effektive Operation (mit R0-Resektabilität) überhaupt möglich zu machen oder die Invasivität der Operation zu reduzieren, kann das Ergebnis der Histologie nach Resektion als Goldstandard für die Differenzierung zwischen Patienten, die von einer NAC(R) nicht profitieren bzw. profitieren können, betrachtet werden. Nach der Feststellung des Nichtansprechens in einer Interims-PET bzw. PET/CT-Untersuchung könnte also

¹ NAC oder NACR wird im Folgenden mit NAC(R) abgekürzt

unmittelbar eine Vermeidung der Fortführung einer nicht effektiven, aber toxischen Therapie erreicht werden.

Bei einem negativ prädiktiven Wert (NPV) der PET bzw. PET/CT von 100 % hätten alle Patienten, die nach 2 Zyklen NAC(R) kein Ansprechen zeigen würden, auch nach Abschluss der gesamten Therapie keine relevante Reduktion des Tumors und damit keinen Vorteil, sondern nur einen potenziellen Schaden durch die NAC(R). Bei einem NPV < 100 % muss es Patienten geben, die trotz negativer PET bzw. PET/CT am Ende der NAC(R) eine relevante Tumorreduktion erzielen würden; d. h., man würde diesen Patienten eine potenziell nützliche NAC(R) vorenthalten. Es wird daher eine Schwelle für den NPV (< 100 %) geben, ab der der Vorteil einer Diagnose mit PET bzw. PET/CT (Reduktion ineffektiver Chemotherapien) ihre Nachteile (fälschlicherweise nicht Behandelte) überwiegt. Die A-priori-Festlegung einer solchen Schwelle wäre notwendig, um eine ergebnisgesteuerte Berichterstattung zu vermeiden.

Insbesondere die aufeinander aufbauenden Studien von Weber 2001 [72] und Ott 2006 [60] untersuchten mit einem sehr homogenen Patienten- und Therapiespektrum den Einsatz der PET während einer NAC an Patienten mit Adenokarzinomen des ösophagogastralen Übergangs (AEG I und II). Daher wurden sie an dieser Stelle gesondert gewürdigt. Diese Studien entsprachen den Einschlusskriterien des vorliegenden Berichts und wurden als „Prognostic accuracy“-Studien für die Indikation Restaging eingeschlossen.

Die Studie von Weber 2001 schloss 40 Patienten ein, wobei nur 37 Patienten in die Analyse eingingen (vgl.5.4.1.4). Ziel der Arbeit war die Ermittlung eines geeigneten Grenzwerts für einen PET-Parameter zur Vorhersage des Therapieansprechens 2 Wochen nach Beginn der NAC. Es konnte gezeigt werden, dass bei den metabolischen Respondern (Reduktion der Tumoraktivität von mehr als 35 %) eine histopathologische Responserate von 54 % (95 %-KI [28,1; 78,6]) erzielt werden kann. Im Vergleich dazu lag die (histopathologische) Responserate bei den (metabolischen) Nonrespondern bei nur 5 % (95 %-KI [0; 13,2]). Die Resultate dieser Studie sind potenziell hoch verzerrt, da der Grenzwert von > 35 % erst im Rahmen der Studie festgelegt wurde, um die Differenzierung zwischen klinischen Respondern und Nichtrespondern zu optimieren.

In die Studie von Ott 2006 [70] wurden 65 Patienten eingeschlossen, von denen 56 in die Analyse eingingen. Die PET-Untersuchungen wurden vor und während der NAC und unmittelbar vor der Operation durchgeführt. Ziel der Studie war die Bestätigung der hohen Vorhersagegenauigkeit der PET unter Anwendung des von Weber 2001 ermittelten Grenzwerts von > 35 %. Diejenigen Patienten, die nach 14 Tagen NAC ein metabolisches Ansprechen zeigten, hatten nach Abschluss der gesamten NAC und nach erfolgter Operation (etwa 13 bis 14 Wochen nach Beginn der NAC) eine histopathologische Ansprechrate von 44 % (KI [21,5; 67,4]). Die Nonresponder erzielten dagegen eine histologische Responserate nach Abschluss der NAC von 5 % (KI [0; 12,4]).

In der Studie von Weber 2001 hatten 28 von 37 (75 %, 95 %-KI [61,9; 89,5]) Patienten kein histopathologisches Ansprechen nach Abschluss der NAC. In der Studie von Ott 2006 hatten nach Abschluss der NAC 46 der 56 Patienten (82 %, 95 %-KI [72,1; 92,2]) kein relevantes Ansprechen in der Histologie. In diesen Studien lag der NPV bei etwa 95 %.

Es wurden weitere Studien in diesen Bericht eingeschlossen, die die PET im Interimsstaging untersuchten. Nicht alle dieser Studien beschränkten sich auf AEG I+II. Neben Weber 2001 und Ott 2006 untersuchte auch die Studie von Kroep 2003 [68] den Einsatz der PET während einer NAC. Die Studien von Gillham 2006 [67], Westerterp 2006 [73] und Wieder 2004 [74] untersuchten den Einsatz der PET während einer NACR. In allen 4 Studien war für jeden ausgewerteten Patienten die SUV-Reduktion vom ersten zum zweiten PET-Bild tabellarisch angegeben, sodass eine nachträgliche Auswertung zu einem Grenzwert von 35 % möglich war, auch wenn in den Studien andere Grenzwerte für die Analysen gewählt worden waren. Keine der Studien war randomisiert. Alle Patienten wurden unabhängig vom PET-Ergebnis weiter mit einer NAC(R) behandelt.

Die Studie von Kroep 2003 [68] schloss insgesamt 13 Patienten mit einem Adenokarzinom oder Plattenepithelkarzinom des gastroösophagealen Übergangs oder des intrathorakalen Ösophagus ein. Ein Patient erhielt nach der Interims-PET keine NAC mehr, ein anderer Patient verstarb. Ein weiterer Patient hatte keinen ausreichenden Kontrast in der Baseline-PET, war aber histologischer Responder. In der Studie hatten nach Abschluss der NAC 6 von 10 Patienten kein relevantes Ansprechen in der Histologie. In dieser Studie lag der NPV bezogen auf den Grenzwert von > 35 % bei 100 % (KI [0,40; 1,00]).

Die Studie von Gillham 2006 [67] berichtete das Ergebnis von 32 eingeschlossenen Patienten mit einem Adenokarzinom oder Plattenepithelkarzinom des gastroösophagealen Übergangs oder des Ösophagus.. Nach Abschluss der NAC(R) hatten 23 von 32 Patienten kein relevantes Ansprechen in der Histologie. Der NPV (bei > 35 %) lag bei 76 % (KI [0,53; 0,92]).

In der Studie von Westerterp 2006 [73] wurde angegeben, dass 26 Patienten mit einem Adenokarzinom oder Plattenepithelkarzinom des Ösophagus mit ≤ 2 cm Involvierung des Magens konsekutiv eingeschlossen wurden. Nach Abschluss einer 5-wöchigen NACR hatten 9 von 26 Patienten kein relevantes histopathologisches Ansprechen. Der NPV (bei > 35 %) lag bei 64 % [0,31; 0,89].

In der Studie von Wieder 2004 [74] wurden die Daten von 38 Patienten mit einem Plattenepithelkarzinom des intrathorakalen Ösophagus berichtet. Bei 6 Patienten wurde die Interims-PET-Untersuchung nicht wie in den anderen Studien nach 2 Wochen, sondern erst nach 4 Wochen durchgeführt. Von den verbleibenden 32 Patienten lag bei 5 Patienten keine Angabe zum histologischen Ansprechen vor, bei 5 Patienten wurde keine Interims-PET durchgeführt, wobei einem Patienten beides fehlte. Von den verbleibenden 23 Patienten hatten nach Beendigung der NACR 8 Patienten kein Ansprechen in der Histologie. Der NPV (bei > 35 %) lag in dieser Studie bei 78 % (KI [0,40; 0,97]).

Zusammenfassend kann also festgestellt werden, dass in den 2 Studien zur diagnostischen Güte der PET beim Interimsstaging einer NAC die Punktschätzer der NPV bei 95 % (38 Patienten mit negativem PET-Ergebnis) und 100 % (4 Patienten mit negativem PET-Ergebnis) und jene der 3 Studien zur NACR zwischen 64 % (11 Patienten mit negativem PET-Ergebnis) und 78 % (9 Patienten mit negativem PET-Ergebnis) lagen. Die Studien waren allesamt sehr klein, sodass diese Ergebnisse mit einer erheblichen Unsicherheit behaftet sind.

Zusammenfassung

Methodisch ergab sich für das Interimsstaging eine besondere Konstellation, weil

- zwei aufeinander aufbauende Studien (Weber 1999 und Weber 2001) [72,114] mit einem sehr homogenen Patienten- und Therapiespektrum untersucht wurden, aus denen ein Grenzwert von $> 35\%$ hervorging, der in Ott 2006 [60] validiert wurde,
- weitere Studien zu dieser Fragestellung Daten auf Patientenebene zur Verfügung stellten, die eine nachträgliche Auswertung zum Cut-off von $> 35\%$ erlaubten (Kroep 2003 [60,68] für die NAC bzw. Gillham 2006, Westerterp 2006 und Wieder 2004 für die NACR [67,73,74]),
- der Anteil ineffektiver neoadjuvanter Therapien in dieser Indikation relativ hoch war, sodass mithilfe der PET relativ vielen Patienten eine unwirksame neoadjuvante Therapie hätte erspart werden können, und
- die prognostische Güte der PET in Hinblick auf das Nichtansprechen des Tumors in den Studien hoch ausfiel.

Alle dargestellten NPVs bezogen sich ausschließlich auf den Grenzwert 35 %. Für andere Grenzwerte werden andere Effekte zu erwarten sein. Für eine Nutzenbewertung der PET blieb die Frage offen, ob ein potenzieller Nutzen bei Patienten den potenziellen Schaden bei Patienten überwiegt. Dazu müsste ein Schwellenwert für den NPV prospektiv festgelegt werden. Außerdem müsste gezeigt werden, dass eine hohe „Prognostic accuracy“ auch tatsächlich zu Änderungen in der klinischen Anwendung der Chemotherapie führt (Auswirkung im Hinblick auf das therapeutische Denken des Behandelnden). Daher sind hier prospektive vergleichende Interventionsstudien mit patientenrelevanten Endpunkten, besonders zur Mortalität, notwendig, um den Nutzen, aber auch den möglichen Schaden der PET-gesteuerten NAC(R) mit größerer Sicherheit und in seinem Ausmaß richtig beurteilen zu können. Auch die Autoren der MUNICON-II-Publikation wiesen darauf hin, dass sie einen internationalen multizentrischen RCT planen [115].

Eine Limitation der in diesem Abschnitt präsentierten Analyse der NPV-Werte bestand darin, dass sich die Berechnungen auf den ausschließlich für die NAC validierten Grenzwert von $> 35\%$ beziehen. Dieser Grenzwert wurde hier auch für die NACR angewendet, wobei er für die NAC(R) in den Studien nicht explizit vorgeschlagen wurde. Darüber hinaus wurde der validierte Grenzwert von $> 35\%$ in den Studien von Weber 2001 und Ott 2006 ausschließlich bei Patienten mit AEG I und II ermittelt. Dieser Grenzwert wurde im vorliegenden Bericht

auf Studien übertragen, die Patienten mit anderen Karzinomen und -anderen Lokalisationen des Ösophagus einschlossen. Es bleibt jedoch fraglich, ob der hier angewandte Grenzwert tatsächlich auf die anderen Studien übertragbar ist.

Nach der Leitlinie der ESMO wurde neben der NAC bei Patienten mit einem Adenokarzinom oder einem Plattenepithelkarzinom des Ösophagus auch die NACR empfohlen [14]. Als mögliche Störgröße für die Beurteilung des Behandlungsansprechens mithilfe einer PET wurde eine durch Radiotherapie induzierte Ösophagitis, welche einen Einfluss auf das Ergebnis der PET hätte, angesehen [116]. Die Ergebnisse einer aktuellen Meta-Analyse von Chen 2011 zeigten allerdings, dass es keinen statistisch signifikanten Unterschied zwischen dem Einsatz der PET bei einer NAC und dem Einsatz bei einer NACR gab [116]. Daher könnte vermutet werden, dass die Ergebnisse der PET bzw. PET/CT bei NAC auf die NACR übertragbar sind.

Zusammenhang zu Mortalität

Dass ein Nichtansprechen in der PET tatsächlich zu Managementänderungen führen kann, wurde in den MUNICON-I- und -II-Studien gezeigt. Für einen Einschluss der MUNICON I und II als Nutzenstudien fehlte in beiden Studien eine zeitlich parallele Kontrollgruppe, die einen randomisierten Vergleich mit einer PET-gesteuerten Therapie erlaubte.

Wenn es das Standardvorgehen ist, dass alle Patienten eine komplette NAC(R) bekommen, so ändert sich die Mortalität der metabolischen Responder nicht, denn sie bekommen weiterhin (auch nach der PET) eine NAC(R). Nur die Mortalität der metabolischen Nonresponder darf unter der Strategie, diesen Patienten keine NAC(R) zu geben, nicht relevant schlechter sein als bei Standardvorgehen. Es lag keine Studie vor, die die Mortalität von metabolischen Nonrespondern, die aufgrund des PET-Befundes keine NAC bzw. NACR bekommen hatten, mit der Mortalität von metabolischen Nonrespondern, die trotzdem eine NAC bzw. NAC(R) bekommen hatten, direkt verglich.

Für die NAC war es jedoch auf Basis der Studien von Weber 1999 und 2001 [72,114], Ott 2006 [60] und MUNICON 2007 [110] möglich, einen indirekten Vergleich nach Bucher 1997 [117] von Hazard Ratios (HR) der PET-Ansprecher bzw. PET-Nichtansprecher durchzuführen. Wenn das HR in der MUNICON-Studie nicht größer würde als in den Studien von Weber 2001 und Ott 2006 [60,72], könnte davon ausgegangen werden, dass der Endpunkt Überleben vergleichbar blieb, trotz der Vermeidung von NACs bei einer großen Anzahl von Patienten. Um diese Hypothese zu testen, bedarf es (streng genommen) eines Nichtunterlegenheitsdesigns.

Das HR verkleinerte sich tatsächlich in der MUNICON-Studie im Vergleich zur Studie von Ott 2006. Aufgrund der geringen Studiengrößen ist das aus dem indirekten Vergleich resultierende Konfidenzintervall für das HR jedoch so breit, dass keine Aussage zur Nichtunterlegenheit eines PET-gesteuerten Verzichts auf die NAC im Vergleich zu einer Fortführung der NAC beim metabolischen Nonresponder gemacht werden kann. Bei dieser

Auswertung ist zusätzlich zu berücksichtigen, dass es sich um einen indirekten Vergleich von Daten aus Beobachtungsstudien handelt, der per se mit einer sehr hohen Unsicherheit behaftet ist. Um eine ausreichende Teststärke für eine Nichtunterlegenheitsfragestellung bzgl. der Mortalität zu erreichen, hätte in diesen Studien ein Vielfaches an Patienten eingeschlossen werden müssen.

Ausblick

Der durchschnittliche Effekt einer neoadjuvanten Therapie im Vergleich zu einer unmittelbaren Operation ist eher schwach. Vorhandene Evidenzsynthesen zeigen, dass diejenigen Patienten, bei denen die NAC anspricht, durchaus deutlich profitieren, dass jedoch der Großteil der Patienten auf die NAC nicht anspricht und daher nur die Nebenwirkungen erleiden muss [118,119]. In den gesichteten Studien scheint die PET in der Lage zu sein, Ansprecher von Nichtansprechen zu einem frühen Zeitpunkt zu unterscheiden (Weber 2001, Ott 2006, MUNICON I [60,72,110]). Direkt vergleichende Studien zum patientenrelevanten Nutzen und Schaden einer PET gesteuerten NAC fehlen jedoch bislang.

In zukünftigen Studien sollten daher idealerweise randomisierte und kontrollierte Designs mit patientenrelevanten Endpunkten angewendet werden. Dazu stehen verschiedene Designvarianten zur Verfügung ([28,120]).

6.1.4 Diagnostische und prognostische Güte der PET und PET/CT bei der Rezidivdiagnostik

Für die Rezidivdiagnostik konnte keine Evidenzsynthese identifiziert werden. Im Rahmen der Ergänzungsrecherche wurden allerdings 4 Studien eingeschlossen. Eine Studie untersuchte dabei neben der diagnostischen Güte auch die prognostische Güte der PET bzw. PET/CT.

Im Bereich der Diagnostik von Rezidiven des Ösophaguskarzinoms argumentieren die Autoren Law et al. 1996, dass über zwei Drittel der Rezidive zwischen einem Jahr oder spätestens 2 Jahren nach Beendigung der ersten Therapie wieder auftreten. [121]. Es bedarf also diagnostischer Verfahren mit hohen Werten für die Sensitivität. Zurzeit werden die CT, die EUS und die MRI als konventionelle bildgebende Verfahren im Bereich der Rezidiverkennung eingesetzt. Diese Verfahren haben allerdings einige Nachteile bei der Rezidivdiagnostik von Ösophaguskarzinomen: Durch die Therapie des Primärtumors entstehen im Zielgewebe Ödeme, Fibrosen oder Vernarbungen. Diese können eine Diagnose mithilfe der genannten Diagnostiken erschweren bzw. verzerren [122,123]. Durch den Einsatz der PET erhofft man sich, diese Hindernisse zu überwinden [124].

Allerdings wurde in den eingeschlossenen Studien nur ein direkter Vergleich zur konventionellen Diagnostik mittels CT angestellt, von dem allerdings die Werte der Sensitivität und Spezifität aus den Angaben in der Studie nicht rekonstruiert werden konnten. So lässt sich also keine Aussage bzgl. einer diagnostischen Überlegenheit der PET bzw. PET/CT im Vergleich zur konventionellen Diagnostik ableiten.

6.2 Vergleich mit den Empfehlungen wissenschaftlicher Fachgesellschaften zum Einsatz der PET und PET/CT beim Ösophaguskarzinom

Eine Recherche auf der Homepage des National Guideline Clearinghouse der Agency for Healthcare Research and Quality machte deutlich, dass relativ wenige internationale Leitlinien zur Diagnostik von Ösophaguskarzinomen vorliegen [19]. Leitlinien, die gefunden wurden, waren oftmals älter und die PET bzw. PET/CT fand darin keine Erwähnung [8]. In der internationalen Leitlinie des NCCN wurde der Einsatz der PET und der PET/CT für das M-Staging empfohlen. Es wurde darauf hingewiesen, dass die diagnostische Güte der PET/CT besser sei als die der PET. Diese Empfehlung basierte auf der Meta-Analyse von van Westreenen 2006 [64], welche nicht in dem vorliegenden Bericht eingeschlossen wurde, weil die Primärstudien nicht den Einschlusskriterien des Vorberichts entsprachen. Besonders für das N-Staging wurde der Einsatz der PET basierend auf der Studie von Flamen 2000 [81] empfohlen. Die genannte Studie wurde im Rahmen des Berichts für die Indikation Primärstaging eingeschlossen. Wie bereits erwähnt, gab es zur Studie von Flamen 2000, in welcher das N- und M-Staging untersucht wurde, eine zweite Publikation von Lerut 2000 [82]. Das N-Staging wurde bei 39 Patienten durchgeführt. Die Daten der Vierfeldertafel konnten mithilfe der Tabelle 1 der Publikation von Lerut 2000 rekonstruiert und im vorliegenden Bericht dargestellt werden. Für das N-Staging wurde ein direkter Vergleich der PET mit der CT und EUS durchgeführt. Die PET und die CT hatten eher niedrige Sensitivitäten (PET: 44,4 %, 95 %-KI [26; 65], CT: 48,1 %, 95 %-KI [29; 68]) und eher hohe Spezifitäten (PET und CT: 91,7 %, 95 %-KI [62; 100]). Die EUS hatte eine eher niedrige Spezifität (41,7 %, 95 %-KI [15; 72]) und eine eher hohe Sensitivität (86,4 %, 95 %-KI [65; 97]). Bei der Interpretation der Ergebnisse sollte jedoch das hohe Verzerrungspotenzial der Studie beachtet werden. In der aktuellen AWMF-S3-Leitlinie zu Adenokarzinomen des Magens und des ösophagogastralen Übergangs wurde der Einsatz der PET/CT als Add-on zur konventionellen Diagnostik für das N- und M-Staging bei Patienten mit Adenokarzinomen des ösophagogastralen Übergangs als eher mäßig empfohlen [125]. Die Empfehlung basierte auf einer nichtsystematischen Leitlinie [126].

Darüber hinaus wies die Leitlinie darauf hin, dass die PET als Technologie für die Prognose bzgl. des Gesamtüberlebens bei Patienten ohne Metastasen Verwendung finden sollte. Die Empfehlung basierte auf einem narrativen Review [127]. Ebenfalls wurde der Einsatz der PET auch für das Behandlungsansprechen einer Chemotherapie vor einer Operation empfohlen. Diese Empfehlung wurde auf Basis von 4 Studien gemacht: Swisher 2004 [128], Levine 2006 [69], Weber 2001 [72] und Flamen 2001 [66]. Von diesen 4 Studien wurden 3 Studien in den vorliegenden Bericht eingeschlossen [66,69,72]. Die Studie von Swisher 2004 [128] entsprach nicht den Einschlusskriterien des Vorberichts, weil es sich um eine retrospektive Studie handelte. Die Leitlinie hob weiter hervor, dass die integrierte PET/CT Vorteile gegenüber der PET nachweist. Sie wies dabei auf die Untersuchung von Cerfolio 2005 [65] hin, in der die PET/CT eine höhere diagnostische Güte als der Ultraschall und die CT hatte. Genannte Studie wurde ebenfalls in den Vorbericht für die Indikation Restaging

eingeschlossen. Die Leitlinie fordert jedoch für die Zukunft, weitere Studien mit der integrierten PET/CT durchzuführen.

Die im Jahr 2006 publizierte Leitlinie des SIGN gab keine Empfehlung zum routinemäßigen Einsatz der PET beim Staging von Speiseröhrentumoren [10]. Die Empfehlung basierte auf der Evidenzsynthese von van Westreenen 2004, die wie erwähnt nicht in den vorliegenden Bericht eingeschlossen werden konnte [64].

Auch in der Leitlinie der niederländischen Vereinigung für Gastroenterologie (Nederlands Genootschap van Maag-Darm-Leverartsen), die zuletzt 2010 veröffentlicht wurde, wurde die FDG-PET für das TNM-Staging nur eingeschränkt empfohlen [9]. Durch die begrenzte Spatialauflösung sei die FDG-PET zur Abgrenzung des Primärtumors nur schlecht einsetzbar. Zur Bestimmung von Metastasen im lokalen Lymphsystem wurde Ultraschall als Methode der Wahl beschrieben. In der Diagnostik von Fernmetastasen schien die FDG-PET der CT überlegen zu sein. Die Empfehlungen basierten auf den Evidenzsynthesen von van Vliet 2008 [113] und Facey 2007 [129]. Die Evidenzsynthese von van Vliet 2008 wurde nicht im Vorbericht eingeschlossen, weil diese nur 3 von 7 Punkten auf dem Qualitätsindex von Oxman und Guyatt 1991 [38,39] erreichte. Die Ergebnisse der Untersuchung wurden allerdings im Abschnitt 6.1.2 diskutiert. Die Arbeit von Facey 2007 [129] wurde im Rahmen des Vorberichts ebenfalls gesichtet und mit E4 (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation) ausgeschlossen.

Die European Society for Medical Oncology (ESMO) hingegen stufte die PET als „hilfreich“ in der Identifikation von bisher nicht beschriebenen Metastasen und in der Diagnose bei Verdacht auf ein Rezidiv [14]. Darüber hinaus könnte die PET in der Kontrolle des Therapieansprechens eine Rolle spielen [14]. Worauf diese Empfehlung basierte, wurde in der Leitlinie nicht deutlich.

Von der Deutschen Gesellschaft für Nuklearmedizin e. V., der Deutschen Röntgengesellschaft e. V. und den Vertretern onkologischer Fachgesellschaften wurde im Rahmen der 3. Interdisziplinären Konsensuskonferenz Onko-PET III im Jahr 2000 der klinische Nutzen der PET mit FDG bei verschiedenen Indikationen untersucht und bewertet [21]. Aus der Leitlinie ging die Empfehlung hervor, die PET bzw. PET/CT beim N- und M-Staging anzuwenden. Die Empfehlung basierte auf 6 Studien, die zwischen 1997 und 2001 publiziert wurden [54,81,82,93,130,131]. Vier dieser Arbeiten [54,81,82,93] wurden auch in den vorliegenden Vorbericht eingeschlossen. Allerdings handelte es sich bei den Arbeiten von Flamen 2000 und Lerut 2000 um dieselbe Studie. Im Vorbericht wurden nur die Ergebnisse für das N-Staging (aus der Publikation von Lerut 2000) eingeschlossen, weil die Autoren für das M-Staging keinen nach Definition des Vorberichts validen Referenztest verwendeten. Die Studie von Brücher 2001 [54] wurde in den Bericht zwar eingeschlossen, jedoch untersuchte sie nicht das Primärstaging, sondern das Therapieansprechen (Restaging). Die Studien von Block 1997 und Flanagan 1997 [130,131] wurden nicht eingeschlossen, weil es sich um retrospektive Studien handelte.

Derzeit gibt es nur wenige Empfehlungen für einen Einsatz der PET bzw. PET/CT im Bereich der Rezidivdiagnostik. Wie erwähnt, empfiehlt die ESMO den Einsatz der PET bei der Rezidivdiagnostik (Evidenzlevel II, B) [14]. Darüber hinaus ließen sich weder internationale systematische Reviews zur Rezidivdiagnostik finden, noch gab es weitere Empfehlungen von internationalen Fachgesellschaften: Die Deutsche Gesellschaft für Nuklearmedizin e. V., die Deutsche Röntgengesellschaft e. V. und die Vertreter onkologischer Fachgesellschaften erläuterten z. B., dass derzeit unzureichende Daten vorliegen, um den Einsatz der PET in der Rezidivdiagnostik zu empfehlen [21]. Die Cancer-Care-Ontario-Gesellschaft unterstützte diese Ansicht, indem sie erläuterten, dass sie bislang nur 2 Studien [78,132] zum Einsatz der PET bei der Rezidivdiagnostik identifizieren konnten und hieraus keine Empfehlung ableiten konnten [133]. Eine dieser Studien (Guo 2007 [78]) wurde auch im Rahmen des vorliegenden Berichts eingeschlossen. Die zweite Publikation von Jingu 2010 lag außerhalb des Recherchezeitraumes des vorliegenden Berichts und wird im Rahmen des Abschlussberichts bearbeitet werden.

Zusammenfassend lässt sich festhalten, dass zur zentralen Fragestellung des vorliegenden Berichts, nämlich zum patientenrelevanten Nutzen der PET bzw. PET/CT beim Primärstaging und / oder Restaging und / oder Rezidivdiagnostik, keine Aussagen gemacht werden können. Es fehlen aktuell methodisch verlässliche und ausreichend große Studien.

6.3 Strahlenschutzüberlegungen

Ein wichtiger Aspekt der PET-Diagnostik ist die Reduktion der Strahlenbelastung im Vergleich zu konventionellen bildgebenden Verfahren. Eine Reduktion der Strahlendosis wird jedoch nur dann erzielt, wenn die PET bzw. PET/CT andere diagnostische Verfahren ersetzt, deren Strahlenbelastung höher ist. In den meisten hier identifizierten Studien wird die PET bzw. PET/CT jedoch als zusätzliche Diagnostik („add-on“) eingesetzt. Die Strahlenbelastung der Diagnostik kann in diesen Fällen daher insgesamt nur zunehmen.

Eine Reduktion der Strahlenbelastung ist darüber hinaus nur in den Fällen für die Patienten relevant, in denen sie nicht mit Verschlechterungen anderer patientenrelevanter Endpunkte einhergeht. Die Effekte der Diagnostik auf patientenrelevante Endpunkte sind jedoch nur in vergleichenden Studien messbar.

Es fand sich keine Studie, die den patientenrelevanten Nutzen unterschiedlicher diagnostischer Alternativen gegen ihren strahlenbedingten Schaden abwägt. Hier besteht noch Forschungsbedarf. Solange derartige Studien nicht vorliegen, kann aus der alleinigen Aussage, dass die PET oder die (Low-dose)PET/CT mit weniger Strahlenbelastung verbunden ist als beispielsweise die CT, nicht unmittelbar auf einen größeren Nutzen dieser Technologie geschlossen werden.

6.4 Allgemeine methodische Probleme

Die häufig geringe Berichtsqualität vorhandener Studien erwies sich als generelles Problem im Rahmen der Fragestellung des vorliegenden Berichts. Ebenso wenig konnte vielen Arbeiten entnommen werden, welche Grenzwerte verwendet wurden, inwieweit die Index- und Referenztests untereinander verblindet befundet wurden oder zu welchen Zeitpunkten im Therapieverlauf die PET bzw. die PET/CT, andere Indextests und die Referenzdiagnostik durchgeführt wurden. Insbesondere bei schnell wachsenden Tumoren können nämlich durch inadäquate Abstände zwischen Erstbehandlung, Indextest 1 / Indextest 2 und Referenztest erhebliche Verzerrungen hervorgerufen werden (disease progression bias) [134]. Insgesamt wäre daher eine präzisere Beschreibung des Studiendesigns für das Verständnis, aber auch die weitere quantitative Zusammenfassung zukünftiger Studien sehr hilfreich.

7 Fazit

Der patientenrelevante Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen ist nicht belegt. Es konnten weder laufende noch abgeschlossene vergleichende Studien zum patientenrelevanten Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen identifiziert werden.

Zu der 2. Fragestellung des Berichts, der diagnostischen und prognostischen Güte, konnten aus 3 eingeschlossenen Evidenzsynthesen und der Ergänzungsrecherche insgesamt 41 Primärstudien eingeschlossen werden.

Die meisten Studien wurden in der Indikation Primärstaging durchgeführt. Zwölf Studien untersuchten die PET direkt mit der CT beim N-Staging und 5 Studien untersuchten den direkten Vergleich von PET mit CT beim M-Staging. Es konnte keine Aussage zugunsten eines Verfahrens getroffen werden, da sich die Konfidenzregionen in den bivariaten Meta-Analysen stark überlappten.

Zu den Indikationen Restaging und Rezidivdiagnostik liegen bisher zu wenige hochwertige Studien vor, um eine verlässliche Aussage zur diagnostischen und prognostischen Güte der PET oder PET/CT treffen zu können. Unklar bleibt insbesondere der Vorteil dieses metabolischen Verfahrens gegenüber den morphologischen bildgebenden Verfahren. Ob der Einsatz der PET bzw. PET/CT auch zu einer Verbesserung patientenrelevanter Endpunkte führt, ist bisher in keiner der 3 Indikationen untersucht worden.

Weitere methodisch hochwertige Studien sind (insbesondere bezüglich der Fragestellung des Behandlungsansprechens neoadjuvanter Therapien) dringend erforderlich, um den patientenrelevanten Nutzen oder Schaden der PET bzw. PET/CT von Ösophaguskarzinomen verlässlich bewerten zu können. Bisher konnten weder laufende noch abgeschlossene parallel vergleichende Studien zum patientenrelevanten Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen identifiziert werden.

8 Liste der eingeschlossenen Studien

Eingeschlossene Evidenzsynthesen

Ngamruengphong S, Sharma VK, Nguyen B, Das A. Assessment of response to neoadjuvant therapy in esophageal cancer: an updated systematic review of diagnostic accuracy of endoscopic ultrasonography and fluorodeoxyglucose positron emission tomography. *Dis Esophagus* 2010; 23(3): 216-231.

Rebollo Aguirre AC, Ramos-Font C, Villegas Portero R, Cook GJR, Llamas Elvira JM, Tabares AR. 18F-fluorodeoxyglucose positron emission tomography for the evaluation of neoadjuvant therapy response in esophageal cancer: systematic review of the literature. *Ann Surg* 2009; 250(2): 247-254.

Westerterp M, Van Westreenen HL, Reitsma JB, Hoekstra OS, Stoker J, Fockens P et al. Esophageal cancer: CT, endoscopic US, and FDG PET for assessment of response to neoadjuvant therapy; systematic review. *Radiology* 2005; 236(3): 841-851.

Eingeschlossene Primärstudien (aus den Evidenzsynthesen)

Brücher BL, Weber W, Bauer M, Fink U, Avril N, Stein HJ et al. Neoadjuvant therapy of esophageal squamous cell carcinoma: response evaluation by positron emission tomography. *Ann Surg* 2001; 233(3): 300-309.

Cerfolio RJ, Bryant AS, Ohja B, Bartolucci AA, Eloubeidi MA. The accuracy of endoscopic ultrasonography with fine-needle aspiration, integrated positron emission tomography with computed tomography, and computed tomography in restaging patients with esophageal cancer after neoadjuvant chemoradiotherapy. *J Thorac Cardiovasc Surg* 2005; 129(6): 1232-1241.

Flamen P, Van Cutsem E, Lerut A, Cambier JP, Haustermans K, Bormans G et al. Positron emission tomography for assessment of the response to induction radiochemotherapy in locally advanced oesophageal cancer. *Ann Oncol* 2002; 13(3): 361-368.

Gillham CM, Lucey JA, Keogan M, Duffy GJ, Malik V, Raouf AA et al. ¹⁸F-FDG uptake during induction chemoradiation for oesophageal cancer fails to predict histomorphological tumour response. *Br J Cancer* 2006; 95(9): 1174-1179.

Kroep JR, Van Groeningen CJ, Cuesta MA, Craanen ME, Hoekstra OS, Comans EF et al. Positron emission tomography using 2-deoxy-2-[¹⁸F]-fluoro-D-glucose for response monitoring in locally advanced gastroesophageal cancer: a comparison of different analytical methods. *Mol Imaging Biol* 2003; 5(5): 337-346.

Levine EA, Farmer MR, Clark P, Mishra G, Ho C, Geisinger KR et al. Predictive value of ¹⁸F-fluoro-deoxy-glucose-positron emission tomography (¹⁸F-FDG-PET) in the identification of responders to chemoradiation therapy for the treatment of locally advanced esophageal cancer. *Ann Surg* 2006; 243(4): 472-478.

Ott K, Weber WA, Lordick F, Becker K, Busch R, Herrmann K et al. Metabolic imaging predicts response, survival, and recurrence in adenocarcinomas of the esophagogastric junction. *J Clin Oncol* 2006; 24(29): 4692-4698.

Song SY, Kim JH, Ryu JS, Lee GH, Kim SB, Park SI et al. FDG-PET in the prediction of pathologic response after neoadjuvant chemoradiotherapy in locally advanced, resectable esophageal cancer. *Int J Radiat Oncol Biol Phys* 2005; 63(4): 1053-1059.

Weber WA, Ott K, Becker K, Dittler HJ, Helmberger H, Avril NE et al. Prediction of response to preoperative chemotherapy in adenocarcinomas of the esophagogastric junction by metabolic imaging. *J Clin Oncol* 2001; 19(12): 3058-3065.

Westerterp M, Omloo JM, Sloof GW, Hulshof MC, Hoekstra OS, Crezee H et al. Monitoring of response to pre-operative chemoradiation in combination with hyperthermia in oesophageal cancer by FDG-PET. *Int J Hyperthermia* 2006; 22(2): 149-160.

Wieder HA, Brücher BL, Zimmermann F, Becker K, Lordick F, Beer A et al. Time course of tumor metabolic activity during chemoradiotherapy of esophageal squamous cell carcinoma and response to treatment. *J Clin Oncol* 2004; 22(5): 900-908.

Eingeschlossene Primärstudien (Ergänzungsrecherche)**Buchmann 2006**

Buchmann I, Hansen T, Brochhausen C, Kneist W, Oberholzer K, Junginger T et al. FDG-PET in the initial staging of squamous cell oesophageal carcinoma. Nucl Med (Stuttg) 2006; 45(6): 235-241.

Cheze-Le Rest 2008

Cheze-Le Rest C, Metges J-P, Teyton P, Jestin-Le Tallec V, Lozac'h P, Volant A et al. Prognostic value of initial fluorodeoxyglucose-PET in esophageal cancer: a prospective study. Nucl Med Commun 2008; 29(7): 628-635.

Teyton P, Metges JP, Jestin-Le Tallec V, Lozac'h P, Volant A, Visviki D et al. Prognostic value of initial FDG PET in esophageal cancer [Französisch]. Medecine Nucleaire 2008; 32(6): 323-331.

Choi 2010

Choi J, Kim SG, Kim JS, Jung HC, Song IS. Comparison of endoscopic ultrasonography (EUS), positron emission tomography (PET), and computed tomography (CT) in the preoperative locoregional staging of resectable esophageal cancer. Surg Endosc 2010; 24(6): 1380-1386.

Choi 2004

Choi JY, Jang HJ, Shim YM, Kim K, Lee KS, Lee KH et al. ¹⁸F-FDG PET in patients with esophageal squamous cell carcinoma undergoing curative surgery: prognostic implications. J Nucl Med 2004; 45(11): 1843-1850.

Choi 2000

Choi JY, Lee KH, Shim YM, Lee KS, Kim JJ, Kim SE et al. Improved detection of individual nodal involvement in squamous cell carcinoma of the esophagus by FDG PET. J Nucl Med 2000; 41(5): 808-815.

Chung 2008

Chung HW, Lee KH, Lee EJ, Lee SJ, Cho YS, Choi JY et al. Comparison of uptake characteristics and prognostic value of ²⁰¹Tl and ¹⁸F-FDG in esophageal cancer. World J Surg 2008; 32(1): 69-75.

Flamen 2000

Lerut T, Flamen P, Ectors N, Van Cutsem E, Peeters M, Hiele M et al. Histopathologic validation of lymph node staging with FDG-PET scan in cancer of the esophagus and gastroesophageal junction: a prospective study based on primary surgery with extensive lymphadenectomy. *Ann Surg* 2000; 232(6): 743-752.

Flamen P, Lerut A, Van Cutsem E, De Wever W, Peeters M, Stroobants S et al. Utility of positron emission tomography for the staging of patients with potentially operable esophageal carcinoma. *J Clin Oncol* 2000; 18(18): 3202-3210.

Guo 2007

Guo H, Zhu H, Xi Y, Zhang B, Li L, Huang Y et al. Diagnostic and prognostic value of ¹⁸F-FDG PET/CT for patients with suspected recurrence from squamous cell carcinoma of the esophagus. *J Nucl Med* 2007; 48(8): 1251-1258.

Heeren 2004

Heeren PAM, Jager PL, Bongaerts F, Van Dullemen H, Sluiter W, Plukker JTM. Detection of distant metastases in esophageal cancer with ¹⁸F-FDG PET. *J Nucl Med* 2004; 45(6): 980-987.

Higuchi 2008

Higuchi I, Yasuda T, Yano M, Doki Y, Miyata H, Tatsumi M et al. Lack of fludeoxyglucose F 18 uptake in posttreatment positron emission tomography as a significant predictor of survival after subsequent surgery in multimodality treatment for patients with locally advanced esophageal squamous cell carcinoma. *J Thorac Cardiovasc Surg* 2008; 136(1): 205-212.

Itho 2008

Itoh K, Hosokawa M, Fujita M. ¹⁸F-FDG-PET/CT in esophageal cancer: comparison between PET and histopathological evaluation on metastasis to locoregional lymph node [Japanisch]. *Japanese Journal of Clinical Radiology* 2008; 53(4): 501-507.

Jager 2001

Jager PL, Que TH, Vaalburg W, Pruijm J, Elsinga P, Plukker JT. Carbon-11 choline or FDG-PET for staging of oesophageal cancer? *Eur J Nucl Med* 2001; 28(12): 1845-1849.

Kato 2008

Kato H, Kimura H, Nakajima M, Sakai M, Sano A, Tanaka N et al. The additional value of integrated PET/CT over PET in initial lymph node staging of esophageal cancer. *Oncol Rep* 2008; 20(4): 857-862.

Kato 2005

Kato H, Miyazaki T, Nakajima M, Takita J, Kimura H, Faried A et al. Comparison between whole-body positron emission tomography and bone scintigraphy in evaluating bony metastases of esophageal carcinomas. *Anticancer Res* 2005; 25(6C): 4439-4444.

Kim 2007

Kim MK, Ryu JS, Kim S-B, Ahn JH, Kim SY, Park SL et al. Value of complete metabolic response by ¹⁸F-fluorodeoxyglucose-positron emission tomography in oesophageal cancer for prediction of pathologic response and survival after preoperative chemoradiotherapy. *Eur J Cancer* 2007; 43(9): 1385-1391

Kneist 2004

Junginger T, Kneist W, Schreckenberger M, Menzel C, Oberholzer K, Bartenstein P. Positronen-Emissions-Tomographie zum präoperativen Staging des Ösophaguskarzinoms. *Dtsch Med Wochenschr* 2002; 127(38): 1935-1941.

Kneist W, Schreckenberger M, Bartenstein P, Menzel C, Oberholzer K, Junginger T. Prospective evaluation of positron emission tomography in the preoperative staging of esophageal carcinoma. *Arch Surg* 2004; 139(10): 1043-1049.

Kneist W, Schreckenberger M, Bartenstein P, Grünwald F, Oberholzer K, Junginger T. Positron emission tomography for staging esophageal cancer: does it lead to a different therapeutic approach? *World J Surg* 2003; 27(10): 1105-1112.

Kneist W, Schreckenberger M, Bartenstein P, Grünwald F, Oberholzer K, Junginger T. Positronenemissionstomographie oder Computertomographie zur präoperativen Diagnostik von Fernmetastasen bei Ösophaguskarzinom. *Viszeralchirurgie* 2003; 38(1): 32-38.

Kneist W, Schreckenberger M, Bartenstein P, Grünwald F, Oberholzer K, Junginger T. Die Positronenemissionstomographie zur präoperativen Lymphknotendiagnostik bei Ösophaguskarzinom. *Chirurg* 2003; 74(10): 922-930.

Kole 1998

Kole AC, Plukker JT, Nieweg OE, Vaalburg W. Positron emission tomography for staging of oesophageal and gastroesophageal malignancy. *Br J Cancer* 1998; 78(4): 521-527.

Liberale 2004

Liberale G, Van Laethem JL, Gay F, Goldman S, Nagy N, Coppens E et al. The role of PET scan in the preoperative management of oesophageal cancer. *Eur J Surg Oncol* 2004; 30(9): 942-947.

Lowe 2005

Lowe VJ, Booya F, Fletcher JG, Nathan M, Jensen E, Mullan B et al. Comparison of positron emission tomography, computed tomography, and endoscopic ultrasound in the initial staging of patients with esophageal cancer. *Molecular Imaging & Biology* 2005; 7(6): 422-430.

Noble 2009

Noble F, Bailey D, Tung K, Byrne JP. Impact of integrated PET/CT in the staging of oesophageal cancer: a UK population-based cohort study. *Clin Radiol* 2009; 64(7): 699-705.

Omloo 2008

Omloo JMT, Sloof GW, Boellaard R, Hoekstra OS, Jager PL, Van Dullemen HM et al. Importance of fluorodeoxyglucose-positron emission tomography (FDG-PET) and endoscopic ultrasonography parameters in predicting survival following surgery for esophageal cancer. *Endoscopy* 2008; 40(6): 464-471.

Rankin 1998

Rankin SC, Taylor H, Cook GJ, Mason R. Computed tomography and positron emission tomography in the pre-operative staging of oesophageal carcinoma. *Clin Radiol* 1998; 53(9): 659-665.

Roedl 2009

Roedl JB, Halpern EF, Colen RR, Sahani DV, Fischman AJ, Blake MA. Metabolic tumor width parameters as determined on PET/CT predict disease-free survival and treatment response in squamous cell carcinoma of the esophagus. *Mol Imaging Biol* 2009; 11(1): 54-60.

Roedl 2008

Roedl JB, Harisinghani MG, Colen RR, Fischman AJ, Blake MA, Mathisen DJ et al. Assessment of treatment response and recurrence in esophageal carcinoma based on tumor length and standardized uptake value on positron emission tomography-computed tomography. *Ann Thorac Surg* 2008; 86(4): 1131-1138.

Schreurs 2008

Schreurs LM, Verhoef CC, Van der Jagt EJ, Van Dam GM, Groen H, Plukker JT. Current relevance of cervical ultrasonography in staging cancer of the esophagus and gastroesophageal junction. *Eur J Radiol* 2008; 67(1): 105-111.

Shenfine 2009

Shenfine J, Barbour AP, Wong D, Thomas J, Martin I, Gotley DC et al. Prognostic value of maximum standardized uptake values from preoperative positron emission tomography in resectable adenocarcinoma of the esophagus treated by surgery alone. *Dis Esophagus* 2009; 22(8): 668-675.

Shivo 2004

Rasanen JV, Sihvo EIT, Knuuti MJ, Minn HRI, Luostarinen MES, Laippala P et al. Prospective analysis of accuracy of positron emission tomography, computed tomography, and endoscopic ultrasonography in staging of adenocarcinoma of the esophagus and the esophagogastric junction. *Ann Surg Oncol* 2003; 10(8): 954-960.

Sihvo EIT, Rasanen JV, Knuuti MJ, Minn HRI, Luostarinen MES, Viljanen T et al. Adenocarcinoma of the esophagus and the esophagogastric junction: positron emission tomography improves staging and prediction of survival in distant but not in locoregional disease. *J Gastrointest Surg* 2004; 8(8): 988-996.

Teyton 2009

Teyton P, Metges JP, Atmani A, Jestin-Le Tallec V, Volant A, Visvikis D et al. Use of positron emission tomography in surgery follow-up of esophageal cancer. *J Gastrointest Surg* 2009; 13(3): 451-458.

Yoon 2003

Yoon YC, Lee KS, Shim YM, Kim BT, Kim K, Kim TS. Metastasis to regional lymph nodes in patients with esophageal squamous cell carcinoma: CT versus FDG PET for presurgical detection prospective study. *Radiology* 2003; 227(3): 764-770.

Zhang 2006

Zhang JD, Yu JM, Guo HB, Sun XR, Xing LG, Li G. Clinical value of positron emission tomography-CT for the diagnosis of postoperative recurrence and metastasis in the patients with esophageal cancer [Chinesisch]. *Zhonghua Wei Chang Wai Ke Za Zhi* 2006; 9(1): 56-58.

9 Literatur

1. Ruf G, Farthmann EH, Engelhardt R, Frommhold H, Henke M, Schaefer HE. Empfehlungen zur standardisierten Diagnostik, Therapie und Nachsorge des Ösophaguskarzinoms [online]. 06.2001 [Zugriff: 17.09.2009]. URL: http://www.tumorzentrum-freiburg.de/tumorzentrum/live/Medizin-Info/Leitlinien/oesophagus_karzinom.pdf.
2. Robert Koch-Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland (Ed). Krebs in Deutschland 2003-2004: Häufigkeiten und Trends. Berlin: RKI; 2008.
3. Cancer Research UK. Oesophageal cancer: UK incidence statistics [online]. 24.07.2009 [Zugriff: 20.01.2010]. URL: <http://info.cancerresearchuk.org/cancerstats/types/oesophagus/incidence/index.htm>.
4. Hongo M, Nagasaki Y, Shoji T. Epidemiology of esophageal cancer: Orient to Occident; effects of chronology, geography and ethnicity. J Gastroenterol Hepatol 2009; 24(5): 729-735.
5. World Health Organization. Fact sheet: world [online]. In: GLOBOCAN 2008: Cancer Incidence and Mortality Worldwide in 2008; Fast Stats. 2010 [Zugriff: 22.09.2011]. URL: <http://globocan.iarc.fr/factsheets/populations/factsheet.asp?uno=900>.
6. Holmes RS, Vaughan TL. Epidemiology and pathogenesis of esophageal cancer. Semin Radiat Oncol 2007; 17(1): 2-9.
7. Kollarova H, Machova L, Horakova D, Janoutova G, Janout V. Epidemiology of esophageal cancer: an overview article. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub 2007; 151(1): 17-20.
8. Allum WH, Griffin SM, Watson A, Colin-Jones D. Guidelines for the management of oesophageal and gastric cancer. Gut 2002; 50(Suppl V): 1-23.
9. Nederlands Genootschap van Maag-Darm-Leverartsen. Oesofaguscarcinoom: landelijke richtlijn; versie 3.0 [online]. 01.12.2010 [Zugriff: 27.09.2011]. URL: http://www.mdl.nl/uploads/240/809/Richtlijn_Oesofaguscarcinoom_definitief_december_2010.pdf.
10. Scottish Intercollegiate Guidelines Network. Management of oesophageal and gastric cancer: a national clinical guideline. Edinburgh: SIGN; 2006. (SIGN publication; Band 87). URL: <http://www.sign.ac.uk/pdf/sign87.pdf>.
11. Wong-Rebecca K, Malthaner R. Combined chemotherapy and radiotherapy (without surgery) compared with radiotherapy alone in localized carcinoma of the esophagus. Cochrane Database Syst Rev 2006; (1): CD002092.

12. Malthaner R, Collin S, Fenlon D, Rhodes S. Preoperative chemotherapy for resectable thoracic esophageal cancer. *Cochrane Database Syst Rev* 2006; (3): CD001556.
13. Sobin LH, Gospodarowicz MK, Wittekind C. TNM classification of malignant tumours. New York: Wiley-Blackwell; 2009.
14. Stahl M, Oliveira J. Esophageal cancer: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann Oncol* 2009; 20(Suppl 4): iv32-iv33.
15. Arnott SJ, Duncan W, Gignoux M, Girling D, Hansen H, Launois B et al. Preoperative radiotherapy for esophageal carcinoma. *Cochrane Database Syst Rev* 2005; (4): CD001799.
16. Homs-Marjolein YV, Gaast A, Siersema PD, Steyerberg EW, Kuipers EJ. Chemotherapy for metastatic carcinoma of the esophagus and gastro-esophageal junction. *Cochrane Database Syst Rev* 2006; (4): CD004063.
17. Deutsche Forschungsgemeinschaft. PET-CT [online]. 09.02.2005 [Zugriff: 20.10.2009]. URL: http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/wgi/download/pet_ct_stellungnahme_050209.pdf.
18. Kadrmas DJ, Casey ME, Conti M, Jakoby BW, Lois C, Townsend DW. Impact of time-of-flight on PET tumor detection. *J Nucl Med* 2009; 50(8): 1315-1323.
19. Agency for Healthcare Research and Quality. National Guideline Clearinghouse [online]. [Zugriff: 22.09.2011]. URL: <http://www.guideline.gov>.
20. Gross E, Busse B, Riemenschneider M, Neubauer S, Seck K, Klein HG et al. Strong association of a common dihydropyrimidine dehydrogenase gene polymorphism with fluoropyrimidine-related toxicity in cancer patients. *PLoS ONE* 2008; 3(12): e4003.
21. Reske SN, Kotzerke J. FDG-PET for clinical use: results of the 3rd German Interdisciplinary Consensus Conference, "Onko-PET III", 21 July and 19 September 2000. *Eur J Nucl Med* 2001; 28(11): 1707-1723.
22. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 4.0 [online]. 23.09.2011 [Zugriff: 14.05.2012]. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf.
23. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Positronenemissionstomographie (PET) und PET/CT bei Adenokarzinom des Pankreas, malignem Melanom, Schilddrüsenkarzinom, Ösophaguskarzinom, Ovarialkarzinom, Mammakarzinom und Knochen- und Weichteiltumoren; Berichtsplan; Aufträge: D06-01E-K [online]. 05.07.2010 [Zugriff: 12.01.2012]. URL: https://www.iqwig.de/download/D06-01_E-K_Berichtsplan_PET_bei_onkologischen_Indikationen_.pdf.

24. How to read clinical journals; II: to learn about a diagnostic test. *Can Med Assoc J* 1981; 124(6): 703-710.
25. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991; 11(2): 88-94.
26. Köbberling J, Trampisch HJ, Windeler J. Memorandum for the evaluation of diagnostic measures. *J Clin Chem Clin Biochem* 1990; 28(12): 873-879.
27. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. 19.01.2012 [Zugriff: 28.03.2012]. URL: http://www.g-ba.de/downloads/62-492-598/VerfO_2012-01-19.pdf.
28. Lijmer JG, Bossuyt PM. Various randomized designs can be used to evaluate medical tests. *J Clin Epidemiol* 2009; 62(4): 364-373.
29. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med* 2001; 134(8): 657-662.
30. Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004; 94(3): 361-366.
31. Cheran SK, Herndon JE, Patz EF. Comparison of whole-body FDG-PET to bone scan for detection of bone metastases in patients with a new diagnosis of lung cancer. *Lung Cancer* 2004; 44(3): 317-325.
32. Kirby AM, George MN. The role of FDG PET in the management of lymphoma: practical guidelines. *Nucl Med Commun* 2007; 28(5): 355-357.
33. Whiting P, Rutjes AW, Dinnes J, Reitsma JB, Bossuyt PM, Kleijnen J. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 2004; 8(25): 1-234.
34. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; 3: 25.
35. Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med* 2004; 140(3): 189-202.
36. Pepe MS, Alonzo TA. Comparing disease screening tests when true disease status is ascertained only for screen positives. *Biostatistics* 2001; 2(3): 249-260.

37. Jadad AR, Murray WE. Randomized controlled trials: questions, answers and musings. Malden: Blackwell; 2007.
38. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991; 44(11): 1271-1278.
39. Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA et al. Agreement among reviewers of review articles. *J Clin Epidemiol* 1991; 44(1): 91-98.
40. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; 151(4): 264-269, W64.
41. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA* 2000; 283(15): 2008-2012.
42. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 2003; 138(1): 1-12.
43. Von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.
44. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med* 2006; 144(6): 427-437.
45. Altman DG. Systematic reviews of evaluations of prognostic variables. *BMJ* 2001; 323(7306): 224-228.
46. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; 359(9308): 781-785.
47. Lange S. The all randomized/full analysis set (ICH E9): may patients be excluded from the analysis? *Drug Inf J* 2001; 35(3): 881-891.
48. Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *CMAJ* 1997; 156(10): 1411-1416.
49. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7(3): 177-188.
50. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414): 557-560.

51. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58(10): 982-990.
52. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: analysing and presenting results [online]. In: Deeks JJ, Bossuyt PM, Gatsonis C. *Cochrane handbook for systematic reviews of diagnostic test accuracy: version 1.0*. 23.12.2010 [Zugriff: 05.12.2011]. URL: <http://srdta.cochrane.org/sites/srdta.cochrane.org/files/uploads/Chapter%2010%20-%20Version%201.0.pdf>.
53. Ngamruengphong S, Sharma VK, Nguyen B, Das A. Assessment of response to neoadjuvant therapy in esophageal cancer: an updated systematic review of diagnostic accuracy of endoscopic ultrasonography and fluorodeoxyglucose positron emission tomography. *Dis Esophagus* 2010; 23(3): 216-231.
54. Brücher BL, Weber W, Bauer M, Fink U, Avril N, Stein HJ et al. Neoadjuvant therapy of esophageal squamous cell carcinoma: response evaluation by positron emission tomography. *Ann Surg* 2001; 233(3): 300-309.
55. Choi JY, Jang HJ, Shim YM, Kim K, Lee KS, Lee KH et al. ¹⁸F-FDG PET in patients with esophageal squamous cell carcinoma undergoing curative surgery: prognostic implications. *J Nucl Med* 2004; 45(11): 1843-1850.
56. Choi JY, Lee KH, Shim YM, Lee KS, Kim JJ, Kim SE et al. Improved detection of individual nodal involvement in squamous cell carcinoma of the esophagus by FDG PET. *J Nucl Med* 2000; 41(5): 808-815.
57. Duong CP, Demitriou H, Weih L, Thompson A, Williams D, Thomas RJS et al. Significant clinical impact and prognostic stratification provided by FDG-PET in the staging of oesophageal cancer. *Eur J Nucl Med Mol Imaging* 2006; 33(7): 759-769.
58. Duong CP, Hicks RJ, Weih L, Drummond E, Leong T, Michael M et al. FDG-PET status following chemoradiotherapy provides high management impact and powerful prognostic stratification in oesophageal cancer. *Eur J Nucl Med Mol Imaging* 2006; 33(7): 770-778.
59. Omloo JMT, Sloof GW, Boellaard R, Hoekstra OS, Jager PL, Van Dullemen HM et al. Importance of fluorodeoxyglucose-positron emission tomography (FDG-PET) and endoscopic ultrasonography parameters in predicting survival following surgery for esophageal cancer. *Endoscopy* 2008; 40(6): 464-471.
60. Ott K, Weber WA, Lordick F, Becker K, Busch R, Herrmann K et al. Metabolic imaging predicts response, survival, and recurrence in adenocarcinomas of the esophagogastric junction. *J Clin Oncol* 2006; 24(29): 4692-4698.

61. Shenfine J, Barbour AP, Wong D, Thomas J, Martin I, Gotley DC et al. Prognostic value of maximum standardized uptake values from preoperative positron emission tomography in resectable adenocarcinoma of the esophagus treated by surgery alone. *Dis Esophagus* 2009; 22(8): 668-675.
62. Rebollo Aguirre AC, Ramos-Font C, Villegas Portero R, Cook GJR, Llamas Elvira JM, Tabares AR. ¹⁸F-fluorodeoxyglucose positron emission tomography for the evaluation of neoadjuvant therapy response in esophageal cancer: systematic review of the literature. *Ann Surg* 2009; 250(2): 247-254.
63. Westerterp M, Van Westreenen HL, Reitsma JB, Hoekstra OS, Stoker J, Fockens P et al. Esophageal cancer: CT, endoscopic US, and FDG PET for assessment of response to neoadjuvant therapy; systematic review. *Radiology* 2005; 236(3): 841-851.
64. Van Westreenen HL, Westerterp M, Bossuyt PM, Pruijm J, Sloof GW, Van Lanschot JJ et al. Systematic review of the staging performance of ¹⁸F-fluorodeoxyglucose positron emission tomography in esophageal cancer. *J Clin Oncol* 2004; 22(18): 3805-3812.
65. Cerfolio RJ, Bryant AS, Ohja B, Bartolucci AA, Eloubeidi MA. The accuracy of endoscopic ultrasonography with fine-needle aspiration, integrated positron emission tomography with computed tomography, and computed tomography in restaging patients with esophageal cancer after neoadjuvant chemoradiotherapy. *J Thorac Cardiovasc Surg* 2005; 129(6): 1232-1241.
66. Flamen P, Van Cutsem E, Lerut A, Cambier JP, Haustermans K, Bormans G et al. Positron emission tomography for assessment of the response to induction radiochemotherapy in locally advanced oesophageal cancer. *Ann Oncol* 2002; 13(3): 361-368.
67. Gillham CM, Lucey JA, Keogan M, Duffy GJ, Malik V, Raouf AA et al. ¹⁸FDG uptake during induction chemoradiation for oesophageal cancer fails to predict histomorphological tumour response. *Br J Cancer* 2006; 95(9): 1174-1179.
68. Kroep JR, Van Groeningen CJ, Cuesta MA, Craanen ME, Hoekstra OS, Comans EF et al. Positron emission tomography using 2-deoxy-2-[¹⁸F]-fluoro-D-glucose for response monitoring in locally advanced gastroesophageal cancer: a comparison of different analytical methods. *Mol Imaging Biol* 2003; 5(5): 337-346.
69. Levine EA, Farmer MR, Clark P, Mishra G, Ho C, Geisinger KR et al. Predictive value of 18-fluoro-deoxy-glucose-positron emission tomography (¹⁸F-FDG-PET) in the identification of responders to chemoradiation therapy for the treatment of locally advanced esophageal cancer. *Ann Surg* 2006; 243(4): 472-478.

70. Ott K, Weber WA, Lordick F, Becker K, Busch R, Herrmann K et al. Metabolic imaging predicts response, survival, and recurrence in adenocarcinomas of the esophagogastric junction. *J Clin Oncol* 2006; 24(29): 4692-4698.
71. Song SY, Kim JH, Ryu JS, Lee GH, Kim SB, Park SI et al. FDG-PET in the prediction of pathologic response after neoadjuvant chemoradiotherapy in locally advanced, resectable esophageal cancer. *Int J Radiat Oncol Biol Phys* 2005; 63(4): 1053-1059.
72. Weber WA, Ott K, Becker K, Dittler HJ, Helmberger H, Avril NE et al. Prediction of response to preoperative chemotherapy in adenocarcinomas of the esophagogastric junction by metabolic imaging. *J Clin Oncol* 2001; 19(12): 3058-3065.
73. Westerterp M, Omloo JM, Sloof GW, Hulshof MC, Hoekstra OS, Crezee H et al. Monitoring of response to pre-operative chemoradiation in combination with hyperthermia in oesophageal cancer by FDG-PET. *Int J Hyperthermia* 2006; 22(2): 149-160.
74. Wieder HA, Brücher BL, Zimmermann F, Becker K, Lordick F, Beer A et al. Time course of tumor metabolic activity during chemoradiotherapy of esophageal squamous cell carcinoma and response to treatment. *J Clin Oncol* 2004; 22(5): 900-908.
75. Cheze-Le Rest C, Metges J-P, Teyton P, Jestin-Le Tallec V, Lozac'h P, Volant A et al. Prognostic value of initial fluorodeoxyglucose-PET in esophageal cancer: a prospective study. *Nucl Med Commun* 2008; 29(7): 628-635.
76. Chung HW, Lee KH, Lee EJ, Lee SJ, Cho YS, Choi JY et al. Comparison of uptake characteristics and prognostic value of ^{201}Tl and ^{18}F -FDG in esophageal cancer. *World J Surg* 2008; 32(1): 69-75.
77. Kim MK, Ryu JS, Kim S-B, Ahn JH, Kim SY, Park SL et al. Value of complete metabolic response by ^{18}F -fluorodeoxyglucose-positron emission tomography in oesophageal cancer for prediction of pathologic response and survival after preoperative chemoradiotherapy. *Eur J Cancer* 2007; 43(9): 1385-1391.
78. Guo H, Zhu H, Xi Y, Zhang B, Li L, Huang Y et al. Diagnostic and prognostic value of ^{18}F -FDG PET/CT for patients with suspected recurrence from squamous cell carcinoma of the esophagus. *J Nucl Med* 2007; 48(8): 1251-1258.
79. Buchmann I, Hansen T, Brochhausen C, Kneist W, Oberholzer K, Junginger T et al. FDG-PET in the initial staging of squamous cell oesophageal carcinoma. *Nucl Med (Stuttg)* 2006; 45(6): 235-241.

80. Choi J, Kim SG, Kim JS, Jung HC, Song IS. Comparison of endoscopic ultrasonography (EUS), positron emission tomography (PET), and computed tomography (CT) in the preoperative locoregional staging of resectable esophageal cancer. *Surg Endosc* 2010; 24(6): 1380-1386.
81. Flamen P, Lerut A, Van Cutsem E, De Wever W, Peeters M, Stroobants S et al. Utility of positron emission tomography for the staging of patients with potentially operable esophageal carcinoma. *J Clin Oncol* 2000; 18(18): 3202-3210.
82. Lerut T, Flamen P, Ectors N, Van Cutsem E, Peeters M, Hiele M et al. Histopathologic validation of lymph node staging with FDG-PET scan in cancer of the esophagus and gastroesophageal junction: a prospective study based on primary surgery with extensive lymphadenectomy. *Ann Surg* 2000; 232(6): 743-752.
83. Heeren PAM, Jager PL, Bongaerts F, Van Dullemen H, Sluiter W, Plukker JTM. Detection of distant metastases in esophageal cancer with ¹⁸F-FDG PET. *J Nucl Med* 2004; 45(6): 980-987.
84. Itoh K, Hosokawa M, Fujita M. ¹⁸FDG-PET/CT in esophageal cancer: comparison between PET and histopathological evaluation on metastasis to locoregional lymph node [Japanisch]. *Japanese Journal of Clinical Radiology* 2008; 53(4): 501-507.
85. Jager PL, Que TH, Vaalburg W, Pruijm J, Elsinga P, Plukker JT. Carbon-11 choline or FDG-PET for staging of oesophageal cancer? *Eur J Nucl Med* 2001; 28(12): 1845-1849.
86. Kato H, Miyazaki T, Nakajima M, Takita J, Kimura H, Faried A et al. Comparison between whole-body positron emission tomography and bone scintigraphy in evaluating bony metastases of esophageal carcinomas. *Anticancer Res* 2005; 25(6C): 4439-4444.
87. Kato H, Kimura H, Nakajima M, Sakai M, Sano A, Tanaka N et al. The additional value of integrated PET/CT over PET in initial lymph node staging of esophageal cancer. *Oncol Rep* 2008; 20(4): 857-862.
88. Kneist W, Schreckenberger M, Bartenstein P, Menzel C, Oberholzer K, Junginger T. Prospective evaluation of positron emission tomography in the preoperative staging of esophageal carcinoma. *Arch Surg* 2004; 139(10): 1043-1049.
89. Kneist W, Schreckenberger M, Bartenstein P, Grünwald F, Oberholzer K, Junginger T. Positron emission tomography for staging esophageal cancer: does it lead to a different therapeutic approach? *World J Surg* 2003; 27(10): 1105-1112.
90. Kneist W, Schreckenberger M, Bartenstein P, Grünwald F, Oberholzer K, Junginger T. Positronenemissionstomographie oder Computertomographie zur präoperativen Diagnostik von Fernmetastasen bei Ösophaguskarzinom. *Viszeralchirurgie* 2003; 38(1): 32-38.

91. Junginger T, Kneist W, Schreckenberger M, Menzel C, Oberholzer K, Bartenstein P. Positronen-Emissions-Tomographie zum präoperativen Staging des Ösophaguskarzinoms. *Dtsch Med Wochenschr* 2002; 127(38): 1935-1941.
92. Kneist W, Schreckenberger M, Bartenstein P, Grünwald F, Oberholzer K, Junginger T. Die Positronenemissionstomographie zur präoperativen Lymphknotendiagnostik bei Ösophaguskarzinom. *Chirurg* 2003; 74(10): 922-930.
93. Kole AC, Plukker JT, Nieweg OE, Vaalburg W. Positron emission tomography for staging of oesophageal and gastroesophageal malignancy. *Br J Cancer* 1998; 78(4): 521-527.
94. Liberale G, Van Laethem JL, Gay F, Goldman S, Nagy N, Coppens E et al. The role of PET scan in the preoperative management of oesophageal cancer. *Eur J Surg Oncol* 2004; 30(9): 942-947.
95. Lowe VJ, Booya F, Fletcher JG, Nathan M, Jensen E, Mullan B et al. Comparison of positron emission tomography, computed tomography, and endoscopic ultrasound in the initial staging of patients with esophageal cancer. *Molecular Imaging & Biology* 2005; 7(6): 422-430.
96. Noble F, Bailey D, Tung K, Byrne JP. Impact of integrated PET/CT in the staging of oesophageal cancer: a UK population-based cohort study. *Clin Radiol* 2009; 64(7): 699-705.
97. Rankin SC, Taylor H, Cook GJ, Mason R. Computed tomography and positron emission tomography in the pre-operative staging of oesophageal carcinoma. *Clin Radiol* 1998; 53(9): 659-665.
98. Schreurs LM, Verhoef CC, Van der Jagt EJ, Van Dam GM, Groen H, Plukker JT. Current relevance of cervical ultrasonography in staging cancer of the esophagus and gastroesophageal junction. *Eur J Radiol* 2008; 67(1): 105-111.
99. Sihvo EIT, Rasanen JV, Knuuti MJ, Minn HRI, Luostarinen MES, Viljanen T et al. Adenocarcinoma of the esophagus and the esophagogastric junction: positron emission tomography improves staging and prediction of survival in distant but not in locoregional disease. *J Gastrointest Surg* 2004; 8(8): 988-996.
100. Räsänen JV, Sihvo EIT, Knuuti MJ, Minn HRI, Luostarinen MES, Laippala P et al. Prospective analysis of accuracy of positron emission tomography, computed tomography, and endoscopic ultrasonography in staging of adenocarcinoma of the esophagus and the esophagogastric junction. *Ann Surg Oncol* 2003; 10(8): 954-960.
101. Yoon YC, Lee KS, Shim YM, Kim BT, Kim K, Kim TS. Metastasis to regional lymph nodes in patients with esophageal squamous cell carcinoma: CT versus FDG PET for presurgical detection prospective study. *Radiology* 2003; 227(3): 764-770.

102. Teyton P, Metges JP, Jestin-Le Tallec V, Lozac'h P, Volant A, Visviki D et al. Prognostic value of initial FDG PET in esophageal cancer [Französisch]. *Medecine Nucleaire* 2008; 32(6): 323-331.
103. Higuchi I, Yasuda T, Yano M, Doki Y, Miyata H, Tatsumi M et al. Lack of fludeoxyglucose F 18 uptake in posttreatment positron emission tomography as a significant predictor of survival after subsequent surgery in multimodality treatment for patients with locally advanced esophageal squamous cell carcinoma. *J Thorac Cardiovasc Surg* 2008; 136(1): 205-212.
104. Roedl JB, Halpern EF, Colen RR, Sahani DV, Fischman AJ, Blake MA. Metabolic tumor width parameters as determined on PET/CT predict disease-free survival and treatment response in squamous cell carcinoma of the esophagus. *Mol Imaging Biol* 2009; 11(1): 54-60.
105. Roedl JB, Harisinghani MG, Colen RR, Fischman AJ, Blake MA, Mathisen DJ et al. Assessment of treatment response and recurrence in esophageal carcinoma based on tumor length and standardized uptake value on positron emission tomography-computed tomography. *Ann Thorac Surg* 2008; 86(4): 1131-1138.
106. Teyton P, Metges JP, Atmani A, Jestin-Le Tallec V, Volant A, Visvikis D et al. Use of positron emission tomography in surgery follow-up of esophageal cancer. *J Gastrointest Surg* 2009; 13(3): 451-458.
107. Zhang JD, Yu JM, Guo HB, Sun XR, Xing LG, Li G. Clinical value of positron emission tomography-CT for the diagnosis of postoperative recurrence and metastasis in the patients with esophageal cancer [Chinesisch]. *Zhonghua Wei Chang Wai Ke Za Zhi* 2006; 9(1): 56-58.
108. Hara T, Yuasa M. Automated synthesis of [¹¹C]choline, a positron-emitting tracer for tumor imaging. *Appl Radiat Isot* 1999; 50(3): 531-533.
109. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 1934; 26(4): 404-413.
110. Lordick F, Ott K, Krause BJ, Weber WA, Becker K, Stein HJ et al. PET to assess early metabolic response and to guide treatment of adenocarcinoma of the oesophagogastric junction: the MUNICON phase II trial. *Lancet Oncol* 2007; 8(9): 797-805.
111. Zum Buschenfelde CM, Herrmann K, Schuster T, Geinitz H, Langer R, Becker K et al. ¹⁸F-FDG PET-guided salvage neoadjuvant radiochemotherapy of adenocarcinoma of the esophagogastric junction: the MUNICON II trial. *J Nucl Med* 2011; 52(8): 1189-1196.
112. Markman M. Use of positron emission tomography scans in ovarian cancer: a diagnostic technique in search of an indication. *J Clin Oncol* 2005; 23(30): 7385-7387.

113. Van Vliet EP, Heijenbrok-Kal MH, Hunink MG, Kuipers EJ, Siersema PD. Staging investigations for oesophageal cancer: a meta-analysis. *Br J Cancer* 2008; 98(3): 547-557.
114. Weber WA, Ziegler SI, Thodtmann R, Hanauske AR, Schwaiger M. Reproducibility of metabolic measurements in malignant tumors using FDG PET. *J Nucl Med* 1999; 40(11): 1771-1777.
115. Lordick F, Ruers T, Aust DE, Collette L, Downey RJ, El Hajjam M et al. European Organisation of Research and Treatment of Cancer (EORTC) Gastrointestinal Group: workshop on the role of metabolic imaging in the neoadjuvant treatment of gastrointestinal cancer. *Eur J Cancer* 2008; 44(13): 1807-1819.
116. Chen YM, Pan XF, Tong LJ, Shi YP, Chen T. Can ^{18}F -fluorodeoxyglucose positron emission tomography predict responses to neoadjuvant therapy in oesophageal cancer patients? A meta-analysis. *Nucl Med Commun* 2011; 32(11): 1005-1010.
117. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; 50(6): 683-691.
118. Gebiski V, Burmeister B, Smithers BM, Foo K, Zalcborg J, Simes J. Survival benefits from neoadjuvant chemoradiotherapy or chemotherapy in oesophageal carcinoma: a meta-analysis. *Lancet Oncol* 2007; 8(3): 226-234.
119. Sjoquist KM, Burmeister BH, Smithers BM, Zalcborg JR, Simes RJ, Barbour A et al. Survival after neoadjuvant chemotherapy or chemoradiotherapy for resectable oesophageal carcinoma: an updated meta-analysis. *Lancet Oncol* 2011; 12(7): 681-692.
120. Sargent DJ, Conley BA, Allegra C, Collette L. Clinical trial designs for predictive marker validation in cancer treatment trials. *J Clin Oncol* 2005; 23(9): 2020-2027.
121. Law SY, Fok M, Wong J. Pattern of recurrence after oesophageal resection for cancer: clinical implications. *Br J Surg* 1996; 83(1): 107-111.
122. Carlisle JG, Quint LE, Francis IR, Orringer MB, Smick JF, Gross BH. Recurrent esophageal carcinoma: CT evaluation after esophagectomy. *Radiology* 1993; 189(1): 271-275.
123. Esteves FP, Schuster DM, Halkar RK. Gastrointestinal tract malignancies and positron emission tomography: an overview. *Semin Nucl Med* 2006; 36(2): 169-181.
124. Kato H, Miyazaki T, Nakajima M, Fukuchi M, Manda R, Kuwano H. Value of positron emission tomography in the diagnosis of recurrent oesophageal carcinoma. *Br J Surg* 2004; 91(8): 1004-1009.

125. Moehler M, Al-Batran SE, Andus T, Anthuber M, Arends J, Arnold D et al. S3-Leitlinie "Magenkarzinom": Diagnostik und Therapie der Adenokarzinome des Magens und ösophago-gastralen Übergangs (AWMF-Regist.-Nr. 032-009-OL). *Z Gastroenterol* 2011; 49(4): 461-531.
126. Peeters M, Lerut T, Vlayen J, Mambourg F, Ectors N, Deprez P ea. Wetenschappelijke ondersteuning van het College voor Oncologie: een nationale praktijkrichtlijn voor de aanpak van slokdarm- en maagkanker [online]. 2008 [Zugriff: 09.01.2012]. (KCE reports; Band 75A). URL: http://kce.fgov.be/sites/default/files/page_documents/d20081027316.pdf.
127. Flamen P, Lerut T, Haustermans K, Van Cutsem E, Mortelmans L. Position of positron emission tomography and other imaging diagnostic modalities in esophageal cancer. *Q J Nucl Med Mol Imaging* 2004; 48(2): 96-108.
128. Swisher SG, Maish M, Erasmus JJ, Correa AM, Ajani JA, Bresalier R et al. Utility of PET, CT, and EUS to identify pathologic responders in esophageal cancer. *Ann Thorac Surg* 2004; 78(4): 1152-1160.
129. Facey K, Bradbury I, Laking G, Payne E. Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers. *Health Technol Assess* 2007; 11(44): iii-iv, xi-267.
130. Block MI, Patterson GA, Sundaresan RS, Bailey MS, Flanagan FL, Dehdashti F et al. Improvement in staging of esophageal cancer with the addition of positron emission tomography. *Ann Thorac Surg* 1997; 64(3): 770-776.
131. Flanagan FL, Dehdashti F, Siegel BA, Trask DD, Sundaresan SR, Patterson GA et al. Staging of esophageal cancer with 18F-fluorodeoxyglucose positron emission tomography. *Am J Roentgenol* 1997; 168(2): 417-424.
132. Jingu K, Kaneta T, Nemoto K, Takeda K, Ogawa Y, Ariga H et al. ¹⁸F-fluorodeoxyglucose positron emission tomography immediately after chemoradiotherapy predicts prognosis in patients with locoregional postoperative recurrent esophageal cancer. *Int J Clin Oncol* 2010; 15(2): 184-190.
133. Wong RJ, Walker-Dilks C, Raab FJ. PET imaging in esophageal cancer [online]. 30.11.2010 [Zugriff: 30.03.2012]. (PET Recommendation Reports; Band 4 Version 2). URL: <https://www.cancercare.on.ca/common/pages/UserFile.aspx?fileId=43135>.
134. Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med* 2004; 140(3): 189-202.

135. Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. *J Med Libr Assoc* 2006; 94(4): 451-455.

Anhang A – Suchstrategien**Primärstudien****1. EMBASE****Suchoberfläche: Ovid**

- EMBASE 1980 to 2010 February 16

#	Searches
1	exp Melanoma/
2	(melanom* or melanotic).ab,ti.
3	(naevocarcinoma* or nevocarcinoma* or melanocarcinoma* or melanoameloblastoma*).ti,ab.
4	((melanocyt* or pigmentar*) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
5	malignant lentigo.ab,ti.
6	or/1-5
7	exp Esophagus cancer/
8	((esophag* or oesophag*) adj6 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
9	or/7-8
10	exp Esophagus/
11	(esophag* or oesophag*).ti,ab.
12	Squamous cell carcinoma/
13	Adenocarcinoma/
14	(adenocarcinom* or scc).ab,ti.
15	(squam* adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
16	(10 or 11) and (12 or 13 or 14 or 15)
17	or/9,16
18	exp ovary cancer/
19	Endometrioid carcinoma/
20	Dysgerminoma/
21	((ovarian* or ovarial* or ovary or ovaries or ovarium or adnexa* or endometr*) adj3 (mass* or teratom* or tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
22	(granulosa cell tumor* or luteoma* or meigs syndrom* or sertoli-leydig* or thecoma*

#	Searches
	or dysgerminoma*).ab,ti.
23	or/18-22
24	exp Soft tissue disease/
25	exp Sarcoma/
26	exp Bone tumor/
27	exp Bone cancer/
28	exp Muscle tumor/
29	((Desmoplastic or small round cell) adj3 (tumor* or tumour*)).ti,ab.
30	Neuroblastoma*.ab,ti.
31	(sarcoma* or osteosarcoma*).ab,ti.
32	(connective tissue* adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
33	(Carcinosarcoma* or Chondrosarcoma* or Fibrosarcoma* or Dermatofibrosarcoma* or Neurofibrosarcoma* or Neurofibroma* or Hemangiosarcoma* or Haemangiosarcoma* or Leiomyosarcoma* or Liposarcoma* or Lymphangiosarcoma* or Myosarcoma* or Rhabdomyosarcoma* or Myxosarcoma* or Phyllodes* or Myxofibrosarcom* or Enchondroma*).ab,ti.
34	(Lipoblastoma* or Myolipoma* or Hibernoma* or Elastofibroma* or Fibroblastoma* or Myofibroblastoma* or Angiofibroma* or Lipofibromatos* or Angioleiomyoma* or Myopericytoma* or Lymphangioma* or Haemangioendothelioma* or Hemangioendothelioma* or Angioendothelioma* or Angiosarcoma* or Thymoma* or Myoepithelioma* or Mesenchymoma* or Hamartoma* or Adamantinoma*).ab,ti.
35	((Proliferative or ossifican* or nodular or ischaem*) adj5 (fasciitis or myositis)).ab,ti.
36	(Angiolipoma* or Angiom* or Lipoma* or Myelolipoma* or Chondroblastoma* or Chondroma* or Mastocytosis* or Myofibroma* or Myxoma* or Mastocytoma* or Neurothekeoma* or Osteoblastoma* or Osteochondroma* or Osteoma* or Fibroma* or Adenofibroma* or Fibroadenoma* or Fibroid* or Leiomyoma* or Rhabdomyoma* or Lymphangioleiomyomatosis* or Osteofibrous* or Angiomyoma* or Angiomyolipoma*).ab,ti.
37	(Schwannoma or Histiocytosis or Histiocytoma*).ab,ti.
38	((Giant Cell* or Plasma Cell* or Fibroepithelial* or Brenner or Granular Cell* or Myofibroblasti* or Osseous* or Ossifying or Fibrohistiocy* or Glomus* or Angiectatic* or Fibromyxoid or Rhabdoid or Neuroectodermal) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom* or granulom* or histiocytoma* or myeloma*)).ab,ti.
39	(soft adj3 tissue adj3 (mass* or tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
40	((muscle* or muscul* or vascular* or perivascular*) adj3 (tumor* or tumour* or

#	Searches
	neoplas* or malignan* or cancer* or carcinom*).ab,ti.
41	((bone* or skeletal*) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom* or metas*).ab,ti.
42	(bone* adj6 (cyst* or lymphoma*).ab,ti.
43	(fibrous adj3 (tumor* or tumour*).ab,ti.
44	((femoral* or jaw* or nose* or orbital* or skull* or spinal* or mandib* or maxilla* or palatal* or palate* or paravertebral*) adj3 (mass* or tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom* or metas*).ab,ti.
45	((Erdheim-Chester or Adenomatous or Beckwith-Wiedemann or McCune-Albright or Retinoblastoma or Rothmund-Thomson or Werner) adj3 (polyposis or syndrom*).ab,ti.
46	or/24-45
47	positron emission tomography/
48	(pet or petscan* or positron).ab,ti.
49	((pet* or tracer*) and (deoxy* or de oxy* or fluor* or fdg* or amino*).ab,ti.
50	or/47-49

2. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations
- Ovid MEDLINE(R) Daily
- Ovid MEDLINE(R) 1950 to Present

#	Searches
1	exp Melanoma/
2	(melanom* or melanotic).ab,ti.
3	(naevocarcinoma* or nevocarcinoma* or melanocarcinoma* or melanoameloblastoma*).ti,ab.
4	((melanocyt* or pigmentar*) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*).ab,ti.
5	malignant lentigo.ab,ti.
6	or/1-5
7	exp Esophageal Neoplasms/
8	((esophag* or oesophag*) adj6 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*).ab,ti.

#	Searches
9	or/7-8
10	exp Esophagus/
11	(esophag\$ or oesophag\$).ti,ab.
12	Carcinoma, Squamous Cell/
13	Adenocarcinoma/
14	(adenocarcinom* or scc).ab,ti.
15	(squam* adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
16	(10 or 11) and (12 or 13 or 14 or 15)
17	or/9,16
18	exp Ovarian Neoplasms/
19	Carcinoma, Endometrioid/
20	Dysgerminoma/
21	((ovarian* or ovarial* or ovary or ovaries or ovarium or adnexa* or endometr*) adj3 (mass* or teratom* or tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
22	(granulosa cell tumor* or luteoma* or meigs syndrom* or sertoli-leydig* or thecoma* or dysgerminoma*).ab,ti.
23	or/18-22
24	exp "Neoplasms, Connective and Soft Tissue"/
25	exp Soft Tissue Neoplasms/
26	exp Sarcoma/
27	exp Bone Neoplasms/
28	((Desmoplastic or small round cell) adj3 (tumor* or tumour*)).ti,ab.
29	Neuroblastoma*.ab,ti.
30	(sarcoma* or osteosarcoma*).ab,ti.
31	(connective tissue* adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
32	(Carcinosarcoma* or Chondrosarcoma* or Fibrosarcoma* or Dermatofibrosarcoma* or Neurofibrosarcoma* or Neurofibroma* or Hemangiosarcoma* or Haemangiosarcoma* or Leiomyosarcoma* or Liposarcoma* or Lymphangiosarcoma* or Myosarcoma* or Rhabdomyosarcoma* or Myxosarcoma* or Phyllodes* or Myxofibrosarcoma* or Enchondroma*).ab,ti.
33	(Lipoblastoma* or Myolipoma* or Hibernoma* or Elastofibroma* or Fibroblastoma* or Myofibroblastoma* or Angiofibroma* or Lipofibromatos* or Angioleiomyoma* or

#	Searches
	Myopericytoma* or Lymphangioma* or Haemangioendothelioma* or Hemangioendothelioma* or Angioendothelioma* or Angiosarcoma* or Thymoma* or Myoepithelioma* or Mesenchymoma* or Hamartoma* or Adamantinoma*).ab,ti.
34	((Proliferative or ossifican* or nodular or ischaem*) adj5 (fasciitis or myositis)).ab,ti.
35	(Angiolipoma* or Angiom* or Lipoma* or Myelolipoma* or Chondroblastoma* or Chondroma* or Mastocytosis* or Myofibroma* or Myxoma* or Mastocytoma* or Neurothekeoma* or Osteoblastoma* or Osteochondroma* or Osteoma* or Fibroma* or Adenofibroma* or Fibroadenoma* or Fibroid* or Leiomyoma* or Rhabdomyoma* or Lymphangiroleiomyomatosis* or Osteofibrous* or Angiomyoma* or Angiomyolipoma*).ab,ti.
36	(Schwannoma or Histiocytosis or Histiocytoma*).ab,ti.
37	((Giant Cell* or Plasma Cell* or Fibroepithelial* or Brenner or Granular Cell* or Myofibroblasti* or Osseous* or Ossifying or Fibrohistiocy* or Glomus* or Angiectatic* or Fibromyxoid or Rhabdoid or Neuroectodermal) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom* or granulom* or histiocytoma* or myeloma*)).ab,ti.
38	(soft adj3 tissue adj3 (mass* or tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
39	((muscle* or muscul* or vascular* or perivascular*) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
40	((bone* or skeletal*) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom* or metas*)).ab,ti.
41	(bone* adj6 (cyst* or lymphoma*)).ab,ti.
42	(fibrous adj3 (tumor* or tumour*)).ab,ti.
43	((femoral* or jaw* or nose* or orbital* or skull* or spinal* or mandib* or maxilla* or palatal* or palate* or paravertebral*) adj3 (mass* or tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom* or metas*)).ab,ti.
44	((Erdheim-Chester or Adenomatous or Beckwith-Wiedemann or McCune-Albright or Retinoblastoma or Rothmund-Thomson or Werner) adj3 (polyposis or syndrom*)).ab,ti.
45	or/24-44
46	Positron-Emission Tomography/
47	Fluorodeoxyglucose F18/du
48	(pet or petscan* or positron).ab,ti.
49	((pet* or tracer*) and (deoxy* or de oxy* or fluor* or fdg* or amino*)).ab,ti.
50	or/46-49
51	and/6,50 [malignes Melanoma]
52	and/17,50 [Oesophaguskarzinom]

#	Searches
53	and/23,50 [Ovarialkarzinom]
54	and/45,50 [Knochen und Weichteiltumor]
55	or/51-54
56	animals/ not (animals/ and humans/)
57	55 not 56

3. Pubmed

Suchoberfläche: NLM

- PubMed – as supplied by publisher
- PubMed – in process
- PubMed – OLDMEDLINE
- PubMed – pubmednotmedline

Search	Most Recent Queries
#1	Search melanom*[TIAB] OR melanotic[TIAB]
#2	Search naevocarcinoma*[TIAB] OR nevocarcinoma*[TIAB] OR melanocarcinoma*[TIAB] OR melanoameloblastoma*[TIAB]
#3	Search (melanocyt*[TIAB] OR pigmentar*) AND (tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB])
#4	Search malignant lentigo
#5	Search #1 OR #2 OR #3 OR #4
#6	Search (esophag*[TIAB] OR oesophag* [TIAB]) AND (tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB])
#7	Search esophag*[TIAB] OR oesophag*[TIAB]
#8	Search adenocarcinom*[TIAB] OR scc[TIAB]
#9	Search squam*[TIAB] AND (tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB])
#10	Search #7 AND (#8 OR #9)
#11	Search #6 OR #10
#12	Search (ovarian*[TIAB] OR ovarial*[TIAB] OR ovary[TIAB] OR ovaries[TIAB] OR ovarium[TIAB] OR adnexa*[TIAB] OR endometr*) AND

Search	Most Recent Queries
	(mass*[TIAB] OR teratom*[TIAB] OR tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB])
#13	Search granulosa cell tumor*[TIAB] OR granulosa cell tumour*[TIAB] OR luteoma*[TIAB] OR meigs syndrom*[TIAB] OR sertoli-leydig*[TIAB] OR thecoma*[TIAB] OR dysgerminoma*[TIAB]
#14	Search #12 OR #13
#15	Search (Desmoplastic[TIAB] OR small round cell[TIAB]) AND (tumor*[TIAB] OR tumour*[TIAB])
#16	Search Neuroblastoma*[TIAB]
#17	Search sarcoma*[TIAB] OR osteosarcoma*[TIAB]
#18	Search connective tissue*[TIAB] AND (tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB])
#19	Search Carcinosarcoma*[TIAB] OR Chondrosarcoma*[TIAB] OR Fibrosarcoma*[TIAB] OR Dermatofibrosarcoma*[TIAB] OR Neurofibrosarcoma*[TIAB] OR Neurofibroma*[TIAB] OR Hemangiosarcoma*[TIAB] OR Haemangiosarcoma*[TIAB] OR Leiomyosarcoma*[TIAB] OR Liposarcoma*[TIAB] OR Lymphangiosarcoma*[TIAB] OR Myosarcoma*[TIAB] OR Rhabdomyosarcoma*[TIAB] OR Myxosarcoma*[TIAB] OR Phyllodes*[TIAB] OR Myxofibrosarcom*[TIAB] OR Enchondroma*[TIAB]
#20	Search Lipoblastoma*[TIAB] OR Myolipoma*[TIAB] OR Hibernoma*[TIAB] OR Elastofibroma*[TIAB] OR Fibroblastoma*[TIAB] OR Myofibroblastoma*[TIAB] OR Angiofibroma*[TIAB] OR Lipofibromatos*[TIAB] OR Angioleiomyoma*[TIAB] OR Myopericytoma*[TIAB] OR Lymphangioma*[TIAB] OR Haemangioendothelioma*[TIAB] OR Hemangioendothelioma*[TIAB] OR Angioendothelioma*[TIAB] OR Angiosarcoma*[TIAB] OR Thymoma*[TIAB] OR Myoepithelioma*[TIAB] OR Mesenchymoma*[TIAB] OR Hamartoma*[TIAB] OR Adamantinoma*[TIAB]
#21	Search (Proliferative[TIAB] OR ossifican*[TIAB] OR nodular[TIAB] OR ischaem*[TIAB]) AND (fasciitis[TIAB] OR myositis[TIAB])
#22	Search Angiolipoma*[TIAB] OR Angiom*[TIAB] OR Lipoma*[TIAB] OR Myelolipoma*[TIAB] OR Chondroblastoma*[TIAB] OR Chondroma*[TIAB] OR Mastocytosis*[TIAB] OR Myofibroma*[TIAB] OR Myxoma*[TIAB] OR Mastocytoma*[TIAB] OR Neurothekeoma*[TIAB] OR Osteoblastoma*[TIAB] OR Osteochondroma*[TIAB] OR Osteoma*[TIAB] OR Fibroma*[TIAB] OR Adenofibroma*[TIAB] OR Fibroadenoma*[TIAB] OR Fibroid*[TIAB] OR Leiomyoma*[TIAB] OR Rhabdomyoma*[TIAB] OR Lymphangioleiomyomatosis*[TIAB] OR Osteofibrous*[TIAB] OR

Search	Most Recent Queries
	Angiomyoma*[TIAB] OR Angiomyolipoma*[TIAB]
#23	Search Schwannoma[TIAB] OR Histiocytosis[TIAB] OR Histiocytoma*[TIAB]
#24	Search soft[TIAB] AND tissue[TIAB] AND (mass*[TIAB] OR tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB])
#25	Search (muscle*[TIAB] OR muscul*[TIAB] OR vascular*[TIAB] OR perivascular*[TIAB]) AND (tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB])
#26	Search (bone*[TIAB] OR skeletal*[TIAB]) AND (tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB] OR metas*[TIAB])
#27	Search fibrous[TIAB] AND (tumor*[TIAB] OR tumour*[TIAB])
#28	Search (femoral*[TIAB] OR jaw*[TIAB] OR nose*[TIAB] OR orbital*[TIAB] OR skull*[TIAB] OR spinal*[TIAB] OR mandib*[TIAB] OR maxilla*[TIAB] OR palatal*[TIAB] OR palate*[TIAB] OR paravertebral*[TIAB]) AND (mass*[TIAB] OR tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB] OR metas*[TIAB])
#29	Search (Erdheim-Chester[TIAB] OR Adenomatous[TIAB] OR Beckwith-Wiedemann[TIAB] OR McCune-Albright[TIAB] OR Retinoblastoma[TIAB] OR Rothmund-Thomson[TIAB] OR Werner[TIAB]) AND (polyposis[TIAB] OR syndrom*[TIAB])
#30	Search (Giant Cell*[TIAB] OR Plasma Cell*[TIAB] OR Fibroepithelial*[TIAB] OR Brenner[TIAB] OR Granular Cell*[TIAB] OR Myofibroblasti*[TIAB] OR Osseous*[TIAB] OR Ossifying[TIAB] OR Fibrohistiocy*[TIAB] OR Glomus*[TIAB] OR Angiectatic*[TIAB] OR Fibromyxoid[TIAB] OR Rhabdoid[TIAB] OR Neuroectodermal[TIAB]) AND (tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB] OR granulom*[TIAB] OR histiocytoma*[TIAB] OR myeloma*[TIAB])
#31	Search bone* [TIAB] AND (cyst*[TIAB] OR lymphoma*[TIAB])
#32	Search #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31
#33	Search pet[TIAB] OR petscan*[TIAB] OR positron[TIAB]
#34	Search (pet*[TIAB] OR tracer*[TIAB]) and (deoxy*[TIAB] OR de oxy*[TIAB] OR fluor*[TIAB] OR fdg*[TIAB] OR amino*[TIAB])
#35	Search #33 OR #34
#36	Search in process[sb] OR publisher[sb]

4. The Cochrane Library

Suchoberfläche: Ovid

- EBM Reviews - Cochrane Central Register of Controlled Trials 1st Quarter 2010

#	Searches
1	exp Melanoma/
2	(melanom* or melanotic).ab,ti.
3	(naevocarcinoma* or nevocarcinoma* or melanocarcinoma* or melanoameloblastoma*).ti,ab.
4	((melanocyt* or pigmentar*) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
5	malignant lentigo.ab,ti.
6	or/1-5
7	exp Esophageal Neoplasms/
8	((esophag* or oesophag*) adj6 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
9	or/7-8
10	exp Esophagus/
11	(esophag\$ or oesophag\$).ti,ab.
12	Carcinoma, Squamous Cell/
13	Adenocarcinoma/
14	(adenocarcinom* or scc).ab,ti.
15	(squam* adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
16	(10 or 11) and (12 or 13 or 14 or 15)
17	or/9,16
18	exp Ovarian Neoplasms/
19	Carcinoma, Endometrioid/
20	Dysgerminoma/
21	((ovarian* or ovarial* or ovary or ovaries or ovarium or adnexa* or endometr*) adj3 (mass* or teratom* or tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
22	(granulosa cell tumor* or luteoma* or meigs syndrom* or sertoli-leydig* or thecoma* or dysgerminoma*).ab,ti.
23	or/18-22

#	Searches
24	exp "Neoplasms, Connective and Soft Tissue"/
25	exp Soft Tissue Neoplasms/
26	exp Sarcoma/
27	exp Bone Neoplasms/
28	((Desmoplastic or small round cell) adj3 (tumor* or tumour*)).ti,ab.
29	Neuroblastoma*.ab,ti.
30	(sarcoma* or osteosarcoma*).ab,ti.
31	(connective tissue* adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
32	(Carcinosarcoma* or Chondrosarcoma* or Fibrosarcoma* or Dermatofibrosarcoma* or Neurofibrosarcoma* or Neurofibroma* or Hemangiosarcoma* or Haemangiosarcoma* or Leiomyosarcoma* or Liposarcoma* or Lymphangiosarcoma* or Myosarcoma* or Rhabdomyosarcoma* or Myxosarcoma* or Phyllodes* or Myxofibrosarcom* or Enchondroma*).ab,ti.
33	(Lipoblastoma* or Myolipoma* or Hibernoma* or Elastofibroma* or Fibroblastoma* or Myofibroblastoma* or Angiofibroma* or Lipofibromatos* or Angioleiomyoma* or Myopericytoma* or Lymphangioma* or Haemangioendothelioma* or Hemangioendothelioma* or Angioendothelioma* or Angiosarcoma* or Thymoma* or Myoepithelioma* or Mesenchymoma* or Hamartoma* or Adamantinoma*).ab,ti.
34	((Proliferative or ossifican* or nodular or ischaem*) adj5 (fasciitis or myositis)).ab,ti.
35	(Angiolipoma* or Angiom* or Lipoma* or Myelolipoma* or Chondroblastoma* or Chondroma* or Mastocytosis* or Myofibroma* or Myxoma* or Mastocytoma* or Neurothekeoma* or Osteoblastoma* or Osteochondroma* or Osteoma* or Fibroma* or Adenofibroma* or Fibroadenoma* or Fibroid* or Leiomyoma* or Rhabdomyoma* or Lymphangioliomyomatosis* or Osteofibrous* or Angiomyoma* or Angiomyolipoma*).ab,ti.
36	(Schwannoma or Histiocytosis or Histiocytoma*).ab,ti.
37	((Giant Cell* or Plasma Cell* or Fibroepithelial* or Brenner or Granular Cell* or Myofibroblasti* or Osseous* or Ossifying or Fibrohistiocy* or Glomus* or Angiectatic* or Fibromyxoid or Rhabdoid or Neuroectodermal) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom* or granulom* or histiocytoma* or myeloma*)).ab,ti.
38	(soft adj3 tissue adj3 (mass* or tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
39	((muscle* or muscul* or vascular* or perivascular*) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
40	((bone* or skeletal*) adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom* or metas*)).ab,ti.

#	Searches
41	(bone* adj6 (cyst* or lymphoma*)).ab,ti.
42	(fibrous adj3 (tumor* or tumour*)).ab,ti.
43	((femoral* or jaw* or nose* or orbital* or skull* or spinal* or mandib* or maxilla* or palatal* or palate* or paravertebral*) adj3 (mass* or tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom* or metas*)).ab,ti.
44	((Erdheim-Chester or Adenomatous or Beckwith-Wiedemann or McCune-Albright or Retinoblastoma or Rothmund-Thomson or Werner) adj3 (polyposis or syndrom*)).ab,ti.
45	or/24-44
46	Positron-Emission Tomography/
47	Fluorodeoxyglucose F18/du
48	(pet or petscan* or positron).ab,ti.
49	((pet* or tracer*) and (deoxy* or de oxy* or fluor* or fdg* or amino*)).ab,ti.
50	or/46-49

Evidenzsynthesen

1. Embase

Suchoberfläche: Ovid

EMBASE 1980 to 2010 February 12

Es wurde folgender Filter übernommen:

Systematic Review: Wong [135] - High specificity strategy

#	Searches
1	positron emission tomography/
2	(pet or petscan* or positron).ab,ti.
3	((pet* or tracer*) and (deoxy* or de oxy* or fluor* or fdg* or amino*)).ab,ti.
4	or/1-3
5	(meta analysis or systematic review or MEDLINE).tw.
6	4 and 5

2. Medline

Suchoberfläche: Ovid

Es wurde folgender Filter übernommen:

Systematic Review: Wong [135] - High specificity strategy

#	Searches
1	Positron-Emission Tomography/
2	Fluorodeoxyglucose F18/du
3	(pet or petscan* or positron).ab,ti.
4	((pet* or tracer*) and (deoxy* or de oxy* or fluor* or fdg* or amino*)).ab,ti.
5	or/1-4
6	cochrane database of systematic reviews.jn.
7	(search or MEDLINE or systematic review).tw.
8	meta analysis.pt.
9	or/6-8
10	and/5,9
11	animals/ not (animals/ and humans/)
12	10 not 11
13	<i>[Abgleich mit Referenzen aus Pubmed]</i>
14	12 not 13

3. Pubmed

Suchoberfläche: NLM

Search	Most Recent Queries
#1	Search Pet[TIAB] or petscan*[TIAB] or positron[TIAB]
#2	Search ((pet*[TIAB] or tracer*[TIAB]) and (deoxy*[TIAB] or de oxy*[TIAB] or fluor*[TIAB] or fdg*[TIAB] or amino*[TIAB]))
#3	Search Search[TIAB] or MEDLINE[TIAB] or systematic review[TIAB]
#4	Search (#1 OR #2) AND #3
#5	Search in process[sb] OR publisher[sb]
#6	Search #4 AND #5

4. The Cochrane Library

Suchoberfläche: Centre for Reviews and Dissemination

- Database of Abstracts of Reviews of Effects
- Health Technology Assessment Database

	Search
# 1	MeSH Positron-Emission Tomography EXPLODE 1 2 3 4 5
# 2	MeSH Fluorodeoxyglucose F18 QUALIFIERS DU EXPLODE 1
# 3	(pet OR petscan* OR positron)
# 4	((pet* OR tracer*) AND (deoxy* OR oxy* OR fluor* OR fdg* OR amino*))
# 5	#1 OR #2 OR #3 OR #4

Suchoberfläche: Wiley

- Cochrane Database of Systematic Reviews

ID	Search
#1	MeSH descriptor Positron-Emission Tomography, this term only
#2	MeSH descriptor Fluorodeoxyglucose F18 explode all trees with qualifier: du
#3	(pet or petscan* or positron):ab,ti
#4	((pet* or tracer*) and (deoxy* or de oxy* or fluor* or fdg* or amino*)):ab,ti
#5	(#1 OR #2 OR #3 OR #4)
#6	(#5)

- Health Technology Assessment Database

ID	Search
#13	MeSH descriptor Fluorodeoxyglucose F18 explode all trees with qualifier: DU
#14	MeSH descriptor Tomography, Emission-Computed explode all trees
#15	MeSH descriptor Positron-Emission Tomography explode all trees
#16	(pet or petscan or (positron and emission) or (Positron and tomography)):ti,ab
#17	pet* and (deoxy* or fluor* or 18fluor* or fdg* or 18fdg* or fludeoxy*):ti,ab
#18	(#13 OR #14 OR #15 OR #16 OR #17)
#19	(#18)

Anhang B – Liste der im Volltext überprüften, aber ausgeschlossenen Evidenzsynthesen mit Ausschlussgründen

1. Agencia de Evaluación de Tecnologías Sanitarias. PET in non-neurological clinical oncology: IPE-97/11 [Spanisch]. Madrid: AETS; 1997. (Informe de Evaluación de Tecnologías Sanitarias). URL: http://www.isciii.es/htdocs/investigacion/publicaciones_agencia/11PEToncologia.pdf. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
2. Basu S, Nair N. 18FDG PET in primary oat cell carcinoma of the esophagus. Indian J Cancer 2005; 42(1): 60-62. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
3. Blue Cross Blue Shield Association. FDG positron emission tomography for evaluating esophageal cancer. Chicago: BCBS; 2002. – **Ausschlussgrund: E6** (Vollpublikation nicht verfügbar)
4. Bourguet P. Standards, options and recommendations 2002 for the use of positron emission tomography with [18F]-FDG (PET-FDG in cancerology (integral connection) [Französisch]. Bull Cancer (Paris) 2003; 90(Spec No): S5-S17. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
5. Bradbury I, Bonell E, Boynton J, Cummins E, Facey K, Iqbal K et al. Positron emission tomography (PET) imaging in cancer management. Glasgow: Health Technology Board for Scotland; 2002. URL: http://www.nhshealthquality.org/nhsqis/files/pet_full_report.doc. – **Ausschlussgrund: E1** (Population entspricht nicht den Einschlusskriterien)
6. Comite d'Evaluation et de Diffusion des Innovations Technologiques. Positron emission tomography [online]. 03.03.1999 [Zugriff: 05.11.2010]. URL: <http://cedit.aphp.fr/servlet/siteCeditGB?Destination=reco&numArticle=97.02.2>. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
7. Comite d'Evaluation et de Diffusion des Innovations Technologiques. Positron emission tomography coupled with computer tomography (PET-CT) [online]. 12.12.2002 [Zugriff: 18.10.2010]. URL: <http://cedit.aphp.fr/servlet/siteCeditGB?Destination=reco&numArticle=02.06.1>. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
8. Cuerva Carvajal A, Villegas Portero R. Positron Emission Tomography combined with Magnetic Resonance (PET/RM). Sevilla: Andalusian Agency for Health Technology Assessment; 2006. URL: http://aunets.isciii.es/ficherosproductos/164/2006_AETSA_F6_PETRM.pdf. – **Ausschlussgrund: E2** (Keine Prüfung der PET bzw. PET/CT oder kein geeigneter Referenztest)

9. Department of Science and Technology: Brazilian Health Technology Assessment General Coordination. 18-FDG positron emission tomography in oncology: HTA reports by international HTA agencies [Portugiesisch] [online]. 08.2004 [Zugriff: 04.10.2010]. URL: <http://portal.saude.gov.br/portal/arquivos/pdf/Relat%201%20PET%20-%20Sintese%20das%20Agencias%20Internacionais%20de%20ATS.pdf>. –

Ausschlussgrund: E4 (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)

10. Facey K, Bradbury I, Laking G, Payne E. Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers. Health Technol Assess 2007; 11(44): iii-iv, xi-267. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)

11. Gillham CM, Reynolds J, Hollywood D. Predicting the response of localised oesophageal cancer to neo-adjuvant chemoradiation. World J Surg Oncol 2007; 5: 97. –

Ausschlussgrund: E4 (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)

12. Institute for Clinical Evaluative Sciences. Health technology assessment of PET (positron emission tomography): a systematic review; ICES investigative report [online]. 05.2001.

URL: http://www.ices.on.ca/WebBuild/site/ices-internet-upload/file_collection/Health_Technology_Assessment-PET_May-2001.pdf. –

Ausschlussgrund: E3 (Endpunkte entsprechen nicht den Einschlusskriterien)

13. Institute for Clinical Evaluative Sciences. Health technology assessment of positron emission tomography in oncology: a systematic review. Toronto: ICES; 2002. –

Ausschlussgrund: E6 (Vollpublikation nicht verfügbar)

14. Institute for Clinical Evaluative Sciences. Health technology assessment of positron emission tomography in oncology: a systematic review; ICES investigative report; quarterly update september 2003 [online]. 09.2003 [Zugriff: 01.10.2010]. URL:

http://www.ices.on.ca/WebBuild/site/ices-internet-upload/file_collection/ACF17B.pdf. –

Ausschlussgrund: E4 (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)

15. Medical Services Advisory Committee. Positron emission tomography [part 2(i)]: assessment report; MSAC reference 10. Canberra: MSAC; 2001. URL:

[http://www.health.gov.au/internet/msac/publishing.nsf/Content/8FD1D98FE64C8A2FCA2575AD0082FD8F/\\$File/Ref%2010%20part%202i%20-%20PET%20Report.pdf](http://www.health.gov.au/internet/msac/publishing.nsf/Content/8FD1D98FE64C8A2FCA2575AD0082FD8F/$File/Ref%2010%20part%202i%20-%20PET%20Report.pdf). –

Ausschlussgrund: E3 (Endpunkte entsprechen nicht den Einschlusskriterien)

16. Medicare Services Advisory Committee. Positron emission tomography: MSAC assessment report. Canberra: MSAC; 2000. URL:

[http://www.msac.gov.au/internet/msac/publishing.nsf/Content/8FD1D98FE64C8A2FCA2575AD0082FD8F/\\$File/Ref%2002%20-%20Positron%20emission%20tomography%20Report.pdf](http://www.msac.gov.au/internet/msac/publishing.nsf/Content/8FD1D98FE64C8A2FCA2575AD0082FD8F/$File/Ref%2002%20-%20Positron%20emission%20tomography%20Report.pdf). –

Ausschlussgrund: E3 (Endpunkte entsprechen nicht den Einschlusskriterien)

17. Morland B. Positron emission tomography (PET): diagnostic and clinical use [Norwegisch]. Oslo: Norwegian Knowledge Centre for the Health Services; 2003. (SMM-rapport; Band 6/2003). URL: [ftp://ftp.nokc.no/SMM/SMM-Rapport_03-06_Positronemisjonstomografi%20\(PET\) %96 diagnostisk%20og%20klinisk%20nytteverdi.pdf](ftp://ftp.nokc.no/SMM/SMM-Rapport_03-06_Positronemisjonstomografi%20(PET)%20%96_diagnostisk%20og%20klinisk%20nytteverdi.pdf). – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
18. Netherlands Organisation for Health Research and Development. Accuracy of FDG-PET and spiral CT for the early prediction of non-response to preoperative chemoradiotherapy in patients with esophageal cancer (project record) [online]. In: CRD HTA Database. 20.08.2005 [Zugriff: 21.09.2010]. URL: <http://www.crd.york.ac.uk/CRDWeb/ShowRecord.asp?ID=32005001063>. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
19. Pan L, Gu P, Huang G, Xue H, Wu S. Prognostic significance of SUV on PET/CT in patients with esophageal cancer: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol 2009; 21(9): 1008-1015. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
20. Perez Lozano J, Briones Perez de la Blanca E. Positron emission tomography: research synthesis on effectiveness in specific indications; systematic review [Spanisch] [online]. 02.2000 [Zugriff: 01.10.2010]. URL: http://www.juntadeandalucia.es/salud/servicios/contenidos/nuevaaetsa/up/AETSA_P_2000_P ET.pdf. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
21. Perleth M, Gerhardus A, Velasco-Garrido M. Positronen-Emissions-Tomographie: systematische Übersichten zur Wirksamkeit bei ausgewählten Indikationen. St. Augustin: Asgard-Verlag; 2003. (DIMDI health technology assessment; Band 24). – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
22. Pichon Riviere A, Augustovski F, Cernadas C, Ferrante D, Regueiro A, Garcia Marti S. Positron emission tomography (PET): diagnostic usefulness and indications. Buenos Aires: Institute for Clinical Effectiveness and Health Policy; 2003. – **Ausschlussgrund: E6** (Vollpublikation nicht verfügbar)
23. Rebollo Aguirre AC, Villegas Portero R, Ramos Font C. Utility of Positron Emission Tomography (PET) to assess the reaction of neoadjuvant therapy in breast, oesophagus, and lungs cancer [Spanisch]. Sevilla: Agencia de Evaluación de Tecnologías Sanitarias de Andalucía; 2009. URL: http://aunets.isciii.es/ficherosproductos/186/AETSA_2006-15.pdf. – **Ausschlussgrund: E5** (Mehrfachpublikation ohne relevante Zusatzinformation)

24. Reske SN, Bares R, Bull U, Guhlmann A, Moser E, Wannemacher MF. Klinische Wertigkeit der Positronen-Emissions-Tomographie (PET) bei onkologischen Fragestellungen: Ergebnisse einer interdisziplinären Konsensuskonferenz. Nucl Med (Stuttg) 1996; 35(2): 42-52. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
25. Rodriguez Garrido M, Asensio del Barrio C. PET-CT: indications, systematic review and meta-analysis [Spanisch]. Madrid: Agencia de Evaluacion de Tecnologias Sanitarias; 2004. (Informe de Evaluación de Tecnologías Sanitarias; Band 41). URL: http://www.isciii.es/htdocs/investigacion/publicaciones_agencia/41PET-TAC.pdf. – **Ausschlussgrund: E1** (Population entspricht nicht den Einschlusskriterien)
26. Rodriguez Garrido M, Asensio del Barrio C, Alcazar Alcazar R. Use of FDG-PET in clinical situations not addressed in the monitored use of this technology: IPE-06/49 [Spanisch]. Madrid: Agencia de Evaluacion de Tecnologias Sanitarias; 2006. URL: http://www.isciii.es/htdocs/investigacion/publicaciones_agencia/AETS_49.pdf. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
27. Rodriguez Garrido M, Asensio del Barrio C, Gomez Martinez MV, Carreras Delgado JL, Martin Moreno JM. Positron emission tomography (PET) with 18FDG on clinical oncology: IPE-01/30 [Spanisch]. Madrid: Agencia de Evaluacion de Tecnologias Sanitarias; 2001. URL: http://www.isciii.es/htdocs/investigacion/publicaciones_agencia/30PETOnco.pdf. – **Ausschlussgrund: E1** (Population entspricht nicht den Einschlusskriterien)
28. Suttie SA, Welch AE, Park KGM. Positron emission tomography for monitoring response to neoadjuvant therapy in patients with oesophageal and gastro-oesophageal junction carcinoma. Eur J Surg Oncol 2009; 35(10): 1019-1029. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
29. Unidad de Evaluacion de Tecnologias Sanitarias Agencia Lain Entralgo. Efficacy, safety and indications of hybrid PET-CT system [Spanisch] [online]. 10.2004 [Zugriff: 04.10.2010]. URL: <http://www.madrid.org/cs/Satellite?blobtable=MungoBlobs&blobcol=urldata&blobkey=id&blobheadervalue1=filename%3DPET-TAC.pdf&blobwhere=1158607783819&blobheadername1=Content-Disposition&ssbinary=true&blobheader=application%2Fpdf>. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
30. Van Vliet EPM, Heijenbrok-Kal MH, Hunink MGM, Kuipers EJ, Siersema PD. Staging investigations for oesophageal cancer: a meta-analysis. Br J Cancer 2008; 98(3): 547-557. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)

31. Van Westreenen HL, Westerterp M, Bossuyt PMM, Pruijm J, Sloof GW, Van Lanschot JJB et al. Systematic review of the staging performance of 18F-fluorodeoxyglucose positron emission tomography in esophageal cancer. *J Clin Oncol* 2004; 22(18): 3805-3812. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
32. Westerterp M, Van Westreenen HL, Sloof GW, Plukker JTM, Van Lanschot JJB. Role of positron emission tomography in the (re-)staging of oesophageal cancer. *Scand J Gastroenterol Suppl* 2006; (243): 116-122. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)
33. Wong WL, Chambers RJ. Role of PET/PET CT in the staging and restaging of thoracic oesophageal cancer and gastro-oesophageal cancer: a literature review. *Abdom Imaging* 2008; 33(2): 183-190. – **Ausschlussgrund: E4** (Keine Evidenzsynthese von ausreichender Qualität / Dokumentation)

Anhang C – Liste der ausgeschlossenen Primärstudien mit Ausschlussgründen

1. FDG positron emission tomography for evaluating esophageal cancer. Technologica MAP Suppl 2001: 13-16. – **Ausschlussgrund: E6** (Vollpublikation nicht verfügbar)
2. Surgical Oncology Group announces PET study. J Nucl Med 2000; 41(2): 26N. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
3. Alexander M, Brasic JR. The diagnosis of esophageal cancer by 2-deoxy-2-F-18 fluoro-D-glucose positron emission tomography (F-18 FDG PET). Clin Nucl Med 2006; 31(9): 566-567. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
4. Balink H, Gemmel FAY, Van Schelven WD. The value of positron emission tomography in the diagnosis and treatment of oesophageal cancer [Niederländisch]. Ned Tijdschr Geneesk 2008; 152(20): 1177-1178. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
5. Berger AC, Scott WJ. Noninvasive staging of esophageal carcinoma. J Surg Res 2004; 117(1): 127-133. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
6. Berrisford RG, Wong W-L, Day D, Toy E, Napier M, Mitchell K et al. The decision to operate: role of integrated computed tomography positron emission tomography in staging oesophageal and oesophagogastric junction cancer by the multidisciplinary team. Eur J Cardiothorac Surg 2008; 33(6): 1112-1116. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
7. Block MI, Patterson GA, Sundaresan RS, Bailey MS, Flanagan FL, Dehdashti F et al. Improvement in staging of esophageal cancer with the addition of positron emission tomography. Ann Thorac Surg 1997; 64(3): 770-776. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
8. Bollschweiler E, Holscher AH. Prediction of tumour response by FDG-PET in patients with adenocarcinomas of the oesophagogastric junction. Eur J Nucl Med Mol Imaging 2008; 35(9): 1742-1743. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
9. Bombardieri E. The added value of metabolic imaging with FDG-PET in oesophageal cancer: prognostic role and prediction of response to treatment. Eur J Nucl Med Mol Imaging 2006; 33(7): 753-758. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien) (Studiendesign entspricht nicht den Einschlusskriterien)

10. Bourguet P, Blanc-Vincent MP, Boneu A, Bosquet L, Chauffert B, Corone C et al. Use of positron emission tomography with [¹⁸F]-FDG in oncology: results in digestive cancers [Französisch]. Bull Cancer (Paris) 2003; 90(2 Spec Iss): S56-S66. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)

11. Brücher BL, Weber W, Bauer M, Fink U, Avril N, Stein HJ et al. Neoadjuvant therapy of esophageal squamous cell carcinoma: response evaluation by positron emission tomography. Ann Surg 2001; 233(3): 300-309. – **Ausschlussgrund: E7** (Durch Literatursuche der eingeschlossenen Evidenzsynthesen abgedeckt)

12. Cerfolio RJ, Bryant AS, Ohja B, Bartolucci AA, Eloubeidi MA. The accuracy of endoscopic ultrasonography with fine-needle aspiration, integrated positron emission tomography with computed tomography, and computed tomography in restaging patients with esophageal cancer after neoadjuvant chemoradiotherapy. J Thorac Cardiovasc Surg 2005; 129(6): 1232-1241. – **Ausschlussgrund: E7** (Durch Literatursuche der eingeschlossenen Evidenzsynthesen abgedeckt)

13. Chatterton BE, Ho Shon I, Baldey A, Lenzo N, Patrikeos A, Kelley B et al. Positron emission tomography changes management and prognostic stratification in patients with oesophageal cancer: results of a multicentre prospective study. Eur J Nucl Med Mol Imaging 2009; 36(3): 354-361. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

14. Downey RJ, Akhurst T, Ilson D, Ginsberg R, Bains MS, Gonen M et al. Whole body ¹⁸F-FDG-PET and the response of esophageal cancer to induction therapy: results of a prospective trial. J Clin Oncol 2003; 21(3): 428-432. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

15. Duong CP, Demetriou H, Weih L, Thompson A, Williams D, Thomas RJS et al. Significant clinical impact and prognostic stratification provided by FDG-PET in the staging of oesophageal cancer. Eur J Nucl Med Mol Imaging 2006; 33(7): 759-769. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

16. Duong CP, Hicks RJ, Weih L, Drummond E, Leong T, Michael M et al. FDG-PET status following chemoradiotherapy provides high management impact and powerful prognostic stratification in oesophageal cancer. Eur J Nucl Med Mol Imaging 2006; 33(7): 770-778. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

17. Fiore D, Baggio V, Ruol A, Bocus P, Casara D, Corti L et al. Multimodal imaging of esophagus and cardia cancer before and after treatment. Radiol Med (Torino) 2006; 111(6): 804-817. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

18. Flamen P, Lerut A, Van Cutsem E, Cambier JP, Maes A, De Wever W et al. The utility of positron emission tomography for the diagnosis and staging of recurrent esophageal cancer. *J Thorac Cardiovasc Surg* 2000; 120(6): 1085-1092. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)

19. Flamen P, Van Cutsem E, Lerut A, Cambier JP, Haustermans K, Bormans G et al. Positron emission tomography for assessment of the response to induction radiochemotherapy in locally advanced oesophageal cancer. *Ann Oncol* 2002; 13(3): 361-368. –
Ausschlussgrund: E7 (Durch Literatursuche der eingeschlossenen Evidenzsynthesen abgedeckt)

20. Flanagan FL, Dehdashti F, Siegel BA, Trask DD, Sundaresan SR, Patterson GA et al. Staging of esophageal cancer with 18F-fluorodeoxyglucose positron emission tomography. *Am J Roentgenol* 1997; 168(2): 417-424. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)

21. Gananadha S, Hazebroek EJ, Leibman S, Berry H, Osgood L, Shon IH et al. The utility of FDG-PET in the preoperative staging of esophageal cancer. *Dis Esophagus* 2008; 21(5): 389-394. – **Ausschlussgrund: E2** (Keine Prüfung der PET bzw. PET/CT oder kein geeigneter Referenztest)

22. Hong D, Lunagomez S, Kim EE, Lee JH, Bresalier RS, Swisher SG et al. Value of baseline positron emission tomography for predicting overall survival in patient with nonmetastatic esophageal or gastroesophageal junction carcinoma. *Cancer* 2005; 104(8): 1620-1626. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)

23. Hsu WH, Hsu PK, Wang SJ, Lin KH, Huang CS, Hsieh CC et al. Positron emission tomography-computed tomography in predicting locoregional invasion in esophageal squamous cell carcinoma. *Ann Thorac Surg* 2009; 87(5): 1564-1568. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)

24. Hu Q, Wang W, Zhong X, Yuan S, Fu Z, Guo H et al. Dual-time-point FDG PET for the evaluation of locoregional lymph nodes in thoracic esophageal squamous cell cancer. *Eur J Radiol* 2009; 70(2): 320-324. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

25. Huguier M, Barrier A, Zacharias T, Valinas R. Positron emission tomography of gastrointestinal carcinomas [Französisch]. *Bull Acad Natl Med* 2006; 190(1): 75-84. –
Ausschlussgrund: E2 (Keine Prüfung der PET bzw. PET/CT oder kein geeigneter Referenztest)

26. Imdahl A, Hentschel M, Kleimaier M, Hopt UT, Brink I. Impact of FDG-PET for staging of oesophageal cancer. *Langenbecks Arch Surg* 2004; 389(4): 283-288. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)

27. Inoue T, Oriuchi N, Koyama K, Ichikawa A, Tomiyoshi K, Sato N et al. Usefulness of dual-head coincidence gamma camera with thick NaI crystals for nuclear oncology: comparison with dedicated PET camera and conventional gamma camera with thin NaI crystals. *Ann Nucl Med* 2001; 15(2): 141-148. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
28. Jadvar H, Henderson RW, Conti PS. 2-deoxy-2-[F-18]fluoro-D-glucose-positron emission tomography/computed tomography imaging evaluation of esophageal cancer. *Mol Imaging Biol* 2006; 8(3): 193-200. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
29. Javeri H, Xiao L, Rohren E, Komaki R, Hofstetter W, Lee JH et al. Influence of the baseline 18F-fluoro-2-deoxy-D-glucose positron emission tomography results on survival and pathologic response in patients with gastroesophageal cancer undergoing chemoradiation. *Cancer* 2009; 115(3): 624-630. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
30. Javeri H, Xiao L, Rohren E, Lee JH, Liao Z, Hofstetter W et al. The higher the decrease in the standardized uptake value of positron emission tomography after chemoradiation, the better the survival of patients with gastroesophageal adenocarcinoma. *Cancer* 2009; 115(22): 5184-5192. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
31. Kato H, Fukuchi M, Miyazaki T, Nakajima M, Tanaka N, Inose T et al. Prediction of response to definitive chemoradiotherapy in esophageal cancer using positron emission tomography. *Anticancer Res* 2007; 27(4C): 2627-2633. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
32. Kato H, Miyazaki T, Nakajima M, Takita J, Kimura H, Faried A et al. The incremental effect of positron emission tomography on diagnostic accuracy in the initial staging of esophageal carcinoma. *Cancer* 2005; 103(1): 148-156. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
33. Katsoulis IE, Wong WL, Mattheou AK, Damani N, Chambers J, Livingstone JJ. Fluorine-18 fluorodeoxyglucose positron emission tomography in the preoperative staging of thoracic oesophageal and gastro-oesophageal junction cancer: a prospective study. *Int J Surg* 2007; 5(6): 399-403. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
34. Kim K, Park SJ, Kim BT, Lee KS, Shim YM. Evaluation of lymph node metastases in squamous cell carcinoma of the esophagus with positron emission tomography. *Ann Thorac Surg* 2001; 71(1): 290-294. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

35. Klaeser B, Nitzsche E, Schuller JC, Koberle D, Widmer L, Balmer-Majno S et al. Limited predictive value of FDG-PET for response assessment in the preoperative treatment of esophageal cancer: results of a prospective multi-center trial (SAKK 75/02). *Onkologie* 2009; 32(12): 724-730. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
36. Kabori O, Kirihara Y, Kosaka N, Hara T. Positron emission tomography of esophageal carcinoma using ^{11}C -choline and ^{18}F -fluorodeoxyglucose: a novel method of preoperative lymph node staging. *Cancer* 1999; 86(9): 1638-1648. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
37. Koga H, Sasaki M, Kuwabara Y, Nakagawa M, Hayashi K, Kaneko K et al. Lesion detectability of a gamma camera based coincidence system with FDG in patients with malignant tumors: a comparison with dedicated positron emission tomography. *Ann Nucl Med* 2004; 18(2): 131-136. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
38. Korst RJ, Kansler AL, Port JL, Lee PC, Kerem Y, Altorki NK. Downstaging of T or N predicts long-term survival after preoperative chemotherapy and radical resection for esophageal carcinoma. *Ann Thorac Surg* 2006; 82(2): 480-484. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
39. Kroep JR, Van Groeningen CJ, Cuesta MA, Craanen ME, Hoekstra OS, Comans EFI et al. Positron emission tomography using 2-deoxy-2- ^{18}F -fluoro-D-glucose for response monitoring in locally advanced gastroesophageal cancer; a comparison of different analytical methods. *Mol Imaging Biol* 2003; 5(5): 337-346. – **Ausschlussgrund: E7** (Durch Literatursuche der eingeschlossenen Evidenzsynthesen abgedeckt)
40. Lan XL, Zhang YX, Wu ZJ, Jia Q, Wei H, Gao ZR. The value of dual time point ^{18}F -FDG PET imaging for the differentiation between malignant and benign lesions. *Clin Radiol* 2008; 63(7): 756-764. – **Ausschlussgrund: E1** (Population entspricht nicht den Einschlusskriterien)
41. Leong T, Everitt C, Yuen K, Condron S, Hui A, Ngan SYK et al. A prospective study to evaluate the impact of FDG-PET on CT-based radiotherapy treatment planning for oesophageal cancer. *Radiother Oncol* 2006; 78(3): 254-261. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
42. Levine EA, Farmer MR, Clark P, Mishra G, Ho C, Geisinger KR et al. Predictive value of 18-fluoro-deoxy-glucose-positron emission tomography (^{18}F -FDG-PET) in the identification of responders to chemoradiation therapy for the treatment of locally advanced esophageal cancer. *Ann Surg* 2006; 243(4): 472-478. – **Ausschlussgrund: E7** (Durch Literatursuche der eingeschlossenen Evidenzsynthesen abgedeckt)

43. Li W, Fan W. Value of esophageal abnormal metabolism lesions of 18F-FDG PET/CT for the clinical diagnosis of esophageal malignancy [Chinesisch]. Chinese Journal of Medical Imaging Technology 2009; 25(5): 901-904. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
44. Little SG, Rice TW, Bybel B, Mason DP, Murthy SC, Falk GW et al. Is FDG-PET indicated for superficial esophageal cancer? Eur J Cardiothorac Surg 2007; 31(5): 791-796. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
45. Lordick F, Ott K, Krause BJ, Weber WA, Becker K, Stein HJ et al. PET to assess early metabolic response and to guide treatment of adenocarcinoma of the oesophagogastric junction: the MUNICON phase II trial. Lancet Oncol 2007; 8(9): 797-805. – **Ausschlussgrund: E7** (Durch Literatursuche der eingeschlossenen Evidenzsynthesen abgedeckt)
46. Lorenzen S, Herrmann K, Weber WA, Wieder H, Hennig M, Ott K et al. Visualisation of metastatic oesophageal and gastric cancer and prediction of clinical response to palliative chemotherapy using ¹⁸F-FDG PET. Nucl Med (Stuttg) 2007; 46(6): 263-270. – **Ausschlussgrund: E1** (Population entspricht nicht den Einschlusskriterien)
47. Luketich JD, Friedman DM, Weigel TL, Meehan MA, Keenan RJ, Townsend DW et al. Evaluation of distant metastases in esophageal cancer: 100 consecutive positron emission tomography scans. Ann Thorac Surg 1999; 68(4): 1133-1136. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
48. Luketich JD, Schauer PR, Meltzer CC, Landreneau RJ, Urso GK, Townsend DW et al. Role of positron emission tomography in staging esophageal cancer. Ann Thorac Surg 1997; 64(3): 765-769. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
49. Malik V, Keogan M, Gilham C, Duffy G, Ravi N, Reynolds JV. FDG-PET scanning in the management of cancer of the oesophagus and oesophagogastric junction: early experience with 100 consecutive cases. Ir J Med Sci 2006; 175(4): 48-54. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
50. McAteer D, Wallis F, Couper G, Norton M, Welch A, Bruce D et al. Evaluation of 18F-FDG positron emission tomography in gastric and oesophageal carcinoma. Br J Radiol 1999; 72(858): 525-529. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
51. McDonough PB, Jones DR, Shen KR, Northup PG, Galysh RL, Hernandez A et al. Does FDG-PET add information to EUS and CT in the initial management of esophageal cancer? A prospective single center study. Am J Gastroenterol 2008; 103(3): 570-574. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

52. McLoughlin JM, Melis M, Siegel EM, Dean EM, Weber JM, Chern J et al. Are patients with esophageal cancer who become PET negative after neoadjuvant chemoradiation free of cancer? *J Am Coll Surg* 2008; 206(5): 879-886. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
53. Meyers BF, Downey RJ, Decker PA, Keenan RJ, Siegel BA, Cerfolio RJ et al. The utility of positron emission tomography in staging of potentially operable carcinoma of the thoracic esophagus: results of the American College of Surgeons Oncology Group Z0060 trial. *J Thorac Cardiovasc Surg* 2007; 133(3): 738-745. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
54. Miyata H, Doki Y, Yasuda T, Yamasaki M, Higuchi I, Makari Y et al. Evaluation of clinical significance of ¹⁸F-fluorodeoxyglucose positron emission tomography in superficial squamous cell carcinomas of the thoracic esophagus. *Dis Esophagus* 2008; 21(2): 144-150. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
55. Montravers F, Grahek D, Kerrou K, de Beco V, Younsi N, Barrier A et al. FDG CDET (2D dual-head coincidence gamma camera) in the primary staging of oesophageal cancer: histopathological correlation. *Clin Positron Imaging* 2000; 3(4): 168. – **Ausschlussgrund: E2** (Keine Prüfung der PET bzw. PET/CT oder kein geeigneter Referenztest)
56. Montravers F, Grahek D, Kerrou K, Gutman F, Talbot JN. Positron emission tomography (PET) in gastrointestinal cancer [Französisch]. *Gastroenterol Clin Biol* 2006; 30(5): 737-741. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
57. Moureau-Zabotto L, Touboul E, Lerouge D, Deniaud-Alexandre E, Grahek D, Foulquier JN et al. Impact of computed tomography (CT) and ¹⁸F-deoxyglucose positron emission tomography (FDG-PET) image fusion for conformal radiotherapy in esophageal carcinoma [Französisch]. *Cancer Radiother* 2005; 9(3): 152-160. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
58. Nanni C, Marzola MC, Rubello D, Fanti S. Positron emission tomography for the evaluation of soft-tissue sarcomas and bone sarcomas. *Eur J Nucl Med Mol Imaging* 2009; 36(12): 1940-1943. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
59. Omloo JMT, Van Heijl M, Smits NJ, Phoa SSKS, Van Berge Henegouwen MI, Sloof GW et al. Additional value of external ultrasonography of the neck after CT and PET scanning in the preoperative assessment of patients with esophageal cancer. *Dig Surg* 2009; 26(1): 43-49. – **Ausschlussgrund: E2** (Keine Prüfung der PET bzw. PET/CT oder kein geeigneter Referenztest)

60. Paul AK, Tatsumi M, Higuchi I, Fukunaga H, Yasuda T, Nishimura T. Gamma camera coincidence imaging with [¹⁸F]fluorodeoxyglucose in the pretreatment evaluation of patients with oesophageal cancer. *Nucl Med Commun* 2003; 24(9): 963-970. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
61. Roedl JB, Blake MA, Holalkere NS, Mueller PR, Colen RR, Harisinghani MG. Lymph node staging in esophageal adenocarcinoma with PET-CT based on a visual analysis and based on metabolic parameters. *Abdom Imaging* 2009; 34(5): 610-617. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
62. Roedl JB, Colen RR, Holalkere NS, Fischman AJ, Choi NC, Blake MA. Adenocarcinomas of the esophagus: response to chemoradiotherapy is associated with decrease of metabolic tumor volume as measured on PET-CT: comparison to histopathologic and clinical response evaluation. *Radiother Oncol* 2008; 89(3): 278-286. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
63. Roedl JB, Colen RR, King K, Fischman AJ, Mueller PR, Blake MA. Visual PET/CT scoring for nonspecific 18F-FDG uptake in the differentiation of early malignant and benign esophageal lesions. *Am J Roentgenol* 2008; 191(2): 515-521. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
64. Roedl JB, Prabhakar HB, Mueller PR, Colen RR, Blake MA. Prediction of metastatic disease and survival in patients with gastric and gastroesophageal junction tumors: the incremental value of PET-CT over PET and the clinical role of primary tumor volume measurements. *Acad Radiol* 2009; 16(2): 218-226. – **Ausschlussgrund: E2** (Keine Prüfung der PET bzw. PET/CT oder kein geeigneter Referenztest)
65. Roedl JB, Sahani DV, Colen RR, Fischman AJ, Mueller PR, Blake MA. Tumour length measured on PET-CT predicts the most appropriate stage-dependent therapeutic approach in oesophageal cancer. *Eur Radiol* 2008; 18(12): 2833-2840. – **Ausschlussgrund: E2** (Keine Prüfung der PET bzw. PET/CT oder kein geeigneter Referenztest)
66. Schreurs LM, Pultrum BB, Koopmans KP, Verhoef CC, Jager PL, Van Dam GM et al. Better assessment of nodal metastases by PET/CT fusion compared to side-by-side PET/CT in oesophageal cancer. *Anticancer Res* 2008; 28(3B): 1867-1873. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
67. Sepesi B, Raymond DP, Polomsky M, Watson TJ, Litle VR, Jones CE et al. Does the value of PET-CT extend beyond pretreatment staging? An analysis of survival in surgical patients with esophageal cancer. *J Gastrointest Surg* 2009; 13(12): 2121-2127. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

68. Shimizu S, Hosokawa M, Itoh K, Fujita M, Takahashi H, Shirato H. Can hybrid FDG-PET/CT detect subclinical lymph node metastasis of esophageal cancer appropriately and contribute to radiation treatment planning? A comparison of image-based and pathological findings. *Int J Clin Oncol* 2009; 14(5): 421-425. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
69. Shreve PD, Steventon RS, Deters EC, Kison PV, Gross MD, Wahl RL. Oncologic diagnosis with 2-[fluorine-18]fluoro-2-deoxy-D-glucose imaging: dual-head coincidence gamma camera versus positron emission tomographic scanner. *Radiology* 1998; 207(2): 431-437. – **Ausschlussgrund: E1** (Population entspricht nicht den Einschlusskriterien)
70. Shuto K, Saito H, Ohira G, Natsume T, Kono T, Tohma T et al. Diffusion-weighted MR imaging for postoperative nodal recurrence of esophageal squamous cell cancer in comparison with FDG-PET [Japanisch]. *Gan To Kagaku Ryoho* 2009; 36(12): 2468-2470. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)
71. Song SY, Kim JH, Ryu JS, Lee GH, Kim SB, Park SI et al. FDG-PET in the prediction of pathologic response after neoadjuvant chemoradiotherapy in locally advanced, resectable esophageal cancer. *Int J Radiat Oncol Biol Phys* 2005; 63(4): 1053-1059. – **Ausschlussgrund: E7** (Durch Literatursuche der eingeschlossenen Evidenzsynthesen abgedeckt)
72. Sun L, Su XH, Guan YS, Pan WM, Luo ZM, Wei JH et al. Clinical usefulness of 18F-FDG PET/CT in the restaging of esophageal cancer after surgical resection and radiotherapy. *World J Gastroenterol* 2009; 15(15): 1836-1842. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
73. Swisher SG, Erasmus J, Maish M, Correa AM, Macapinlac H, Ajani JA et al. 2-Fluoro-2-deoxy-D-glucose positron emission tomography imaging is predictive of pathologic response and survival after preoperative chemoradiation in patients with esophageal carcinoma. *Cancer* 2004; 101(8): 1776-1785. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
74. Van Heijl M, Omloo JMT, Van Berge Henegouwen MI, Busch ORC, Tilanus HW, Bossuyt PMM et al. NEOadjuvant therapy monitoring with PET and CT in Esophageal Cancer (NEOPEC-trial). *BMC Med Phys* 2008; 8: 3. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)
75. Van Vliet EPM, Steyerberg EW, Eijkemans MJC, Kuipers EJ, Siersema PD. Detection of distant metastases in patients with oesophageal or gastric cardia cancer: a diagnostic decision analysis. *Br J Cancer* 2007; 97(7): 868-876. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)

76. Van Westreenen HL, Heeren PAM, Jager PL, Van Dullemen HM, Groen H, Plukker JTM. Pitfalls of positive findings in staging esophageal cancer with F-18-fluorodeoxyglucose positron emission tomography. *Ann Surg Oncol* 2003; 10(9): 1100-1105. –

Ausschlussgrund: E4 (Studiendesign entspricht nicht den Einschlusskriterien)

77. Van Westreenen HL, Westerterp M, Sloof GW, Groen H, Bossuyt PMM, Jager PL et al. Limited additional value of positron emission tomography in staging oesophageal cancer. *Br J Surg* 2007; 94(12): 1515-1520. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

78. Vrieze O, Haustermans K, De Wever W, Lerut T, Van Cutsem E, Ectors N et al. Is there a role for FGD-PET in radiotherapy planning in esophageal carcinoma? *Radiother Oncol* 2004; 73(3): 269-275. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

79. Weber WA, Ott K, Becker K, Dittler HJ, Helmberger H, Avril NE et al. Prediction of response to preoperative chemotherapy in adenocarcinomas of the esophagogastric junction by metabolic imaging. *J Clin Oncol* 2001; 19(12): 3058-3065. – **Ausschlussgrund: E7** (Durch Literatursuche der eingeschlossenen Evidenzsynthesen abgedeckt)

80. Westerterp M, Omloo JMT, Sloof GW, Hulshof MCCM, Hoekstra OS, Crezee H et al. Monitoring of response to pre-operative chemoradiation in combination with hyperthermia in oesophageal cancer by FDG-PET. *Int J Hyperthermia* 2006; 22(2): 149-160. – **Ausschlussgrund: E7** (Durch Literatursuche der eingeschlossenen Evidenzsynthesen abgedeckt)

81. Williams RN, Ubhi SS, Sutton CD, Thomas AL, Entwisle JJ, Bowrey DJ. The early use of PET-CT alters the management of patients with esophageal cancer. *J Gastrointest Surg* 2009; 13(5): 868-873. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)

82. Wren SM, Stijns P, Srinivas S. Positron emission tomography in the initial staging of esophageal cancer. *Arch Surg* 2002; 137(9): 1001-1006. – **Ausschlussgrund: E4** (Studiendesign entspricht nicht den Einschlusskriterien)

83. Yuan S, Yu Y, Chao KSC, Fu Z, Yin Y, Liu T et al. Additional value of PET/CT over PET in assessment of locoregional lymph nodes in thoracic esophageal squamous cell cancer. *J Nucl Med* 2006; 47(8): 1255-1259. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

84. Yuan SH, Yu JM, Yu YH, Fu Z, Guo HB, Liu TH et al. FDG PET/CT versus PET alone for pre-surgical detection of lymph node metastasis in esophageal carcinoma [Chinesisch]. *Zhonghua Zhong Liu Za Zhi* 2007; 29(3): 221-224. – **Ausschlussgrund: E3** (Endpunkte entsprechen nicht den Einschlusskriterien)

85. Zacharias T, Barrier A, Montravers F, Houry S, Lacaine F, Huguier M. Cardio-esophageal cancer: is ¹⁸Fluorodeoxyglucose positron emission tomography worthwhile?
Hepatogastroenterology 2004; 51(57): 741-743. – **Ausschlussgrund: E2** (Keine Prüfung der PET bzw. PET/CT oder kein geeigneter Referenztest)

Anhang D– Beurteilung der Qualität von Evidenzsynthesen**Anhang D.1 – Qualitätsbeurteilung von Evidenzsynthesen nach Oxman und Guyatt-Kriterien [38,39]**

Item	Frage	ja	teilweise	nein	unklar
1	Wurde die Suchmethodik berichtet?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	--
2	Wurde eine umfassende Literatursuche durchgeführt?	<input type="checkbox"/>	--	<input type="checkbox"/>	<input type="checkbox"/>
3	Wurden die Einschlusskriterien für die Primärstudien benannt?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	--
4	Wurde Selektionsbias vermieden?	<input type="checkbox"/>	--	<input type="checkbox"/>	<input type="checkbox"/>
5	Wurden die Kriterien, die zur Erfassung der Validität eingeschlossener Studien herangezogen wurden, berichtet?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	--
6	Wurde die Validität der eingeschlossenen Studien angemessen bewertet?	<input type="checkbox"/>	--	<input type="checkbox"/>	<input type="checkbox"/>
7	Wurden die Methoden, die zum Pooling verwendet wurden, berichtet?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	--
8	Wurden die Ergebnisse der Studien, auf denen die Schlussfolgerung beruht, angemessen gepoolt (in Bezug auf die primäre Fragestellung des Reviews)?	<input type="checkbox"/>	--	<input type="checkbox"/>	<input type="checkbox"/>
9	Wurden die Schlussfolgerungen der Autoren für die Primärfrage durch die Daten und / oder die Analyse unterstützt?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	--
Gesamtbewertung (1–7)		1–4 (keine adäquate Evidenzsynthese) 5–7 (adäquate Evidenzsynthese)			

Anhang D.2 – Ergebnisse der Qualitätsbewertung der Evidenzsynthesen, die nach Oxman und Guyatt bewertet wurden, aber nicht in den Bericht eingeschlossen wurden

Evidenz-synthesen	OG 1	OG 2	OG 3	OG 4	OG 5	OG 6	OG 7	OG 8	OG 9	Gesamt-punkte
van Westreenen 2004	+	+	+	+	+	+	+	+	+/-	5
van Vliet 2008	+	+/-	+/-	+/-	-	-	+	+	+	3
OG1 bis OG9: Qualitätskriterien nach Oxman und Guyatt; +: Kriterium erfüllt; -: Kriterium nicht erfüllt; +/-: Kriterium teilweise erfüllt;										

Die Evidenzsynthese von van Westreenen 2004 [64] wurde nicht aufgrund der Bewertung mithilfe des Index von Oxman und Guyatt ausgeschlossen, sondern, weil keine der von van Westreenen 2004 eingeschlossenen Primärstudien den Einschlusskriterien des vorliegenden Berichts entsprach. Die Autoren in van Westreenen 2004 gaben an, dass keine Interessenskonflikte bestanden. In der Evidenzsynthese von van Vliet 2008 [113] konnten keine Angaben zu Interessenskonflikten gefunden werden. Es wurde angegeben, dass der Erstautor durch den „Doelmatigheidsonderzoek“ der Erasmus MC – University Medical Center Rotterdam, The Netherlands finanziert wurde.

Anhang D.3 – Bewertung des Verzerrungspotenzials von Primärstudien auf Basis von QUADAS-Kriterien

Item	Frage	QUADAS-Item	ja	nein	unklar
1	Referenztest akzeptabel?	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	Adäquate Zeitspanne zwischen Referenz- und Indextest?	4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Interdependenz der Tests adäquat berücksichtigt?	ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	Partielle Verifikation vermieden?	5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Differenzielle Verifikation vermieden?	6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	Incorporation Bias?	7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	Verblindung bei Interpretation der Testergebnisse?	10 und 11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	Intention-To-Diagnose(ITD)-Analyse adäquat?	13 und 14	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	Ergebnisunabhängige Berichterstattung?	ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	Sonstige Aspekte, die das Verzerrungspotenzial beeinflussen?	ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	--
Gesamtbewertung des Verzerrungspotenzials			hoch	niedrig	
			<input type="checkbox"/>	<input type="checkbox"/>	

Anhang D.4 – Einschätzung der Übertragbarkeit der Studienergebnisse von diagnostischen Primärstudien

Item	Frage	QUADAS-Item	ja	nein	unklar
1	Repräsentativität des Patientenspektrums?	1	<input type="checkbox"/>	<input type="checkbox"/>	--
2	Übertragbarkeit Index- und Referenztest?	8 und 9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Übertragbarer Entscheidungsprozess bei Testinterpretation?	ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	--
4	Übertragbare Hintergrundinformationen?	12	<input type="checkbox"/>	<input type="checkbox"/>	--
5	Stand der Technik übertragbar?	ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	--
6	Qualifikation der Untersucher übertragbar?	ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	--
7	Qualität der Geräte übertragbar?	ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	--
Gesamtbewertung der Übertragbarkeit		Die Übertragbarkeit wird als gegeben angesehen	<input type="checkbox"/>		
		Es gibt Anhaltspunkte für eine eingeschränkte Übertragbarkeit	<input type="checkbox"/>		

Anhang D.5 – Bewertung des Verzerrungspotenzials von prognostischen Primärstudien

Item	Frage	ja	nein	unklar
Studienpopulation				
1	Waren die Selektionskriterien für die Patienten klar beschrieben?	<input type="checkbox"/>	<input type="checkbox"/>	--
2	Befanden sich die Patienten in einer vergleichbaren Phase der Erkrankung?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Ist die Studienpopulation repräsentativ für die Zielpopulation?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Nachbeobachtung / Patientenfluss				
4	War die Nachbeobachtungszeit adäquat?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Wurde der Patientenfluss lückenlos und nachvollziehbar beschrieben?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	War die Rate der Teilnehmer an den Nachbeobachtungen adäquat? ^a	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Messung prognostischer Faktoren				
7	Waren die prognostischen Faktoren klar definiert?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	Wurden die prognostischen Faktoren adäquat gemessen?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	Standen die prognostischen Faktoren für einen ausreichend großen Teil der Patienten zur Verfügung?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	Waren die Teilnehmer bezüglich der prognostischen Information verblindet?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Messung der Endpunkte				
11	Waren die Endpunkte klar definiert?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	Wurden die Endpunkte adäquat gemessen?	<input type="checkbox"/>	<input type="checkbox"/>	--
13	Wurden die Endpunkte verblindet bezüglich prognostischer Faktoren festgelegt?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Statistische Analyse / Umgang mit Störgrößen (Confounder)				
14	Wurden alle wichtigen Störgrößen adäquat berücksichtigt?	<input type="checkbox"/>	<input type="checkbox"/>	--
15	Wurden alle Behandlungen, die während der Nachbeobachtungszeit durchgeführt wurden, standardisiert oder randomisiert?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sonstige Aspekte				
16	Ergebnisunabhängige Berichterstattung?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17	Keine sonstigen Aspekte, die das Verzerrungspotenzial beeinflussen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gesamtbewertung des Verzerrungspotenzials		hoch	niedrig	
		<input type="checkbox"/>	<input type="checkbox"/>	
a: Eine Follow-up-Rate $\geq 80\%$ wird als adäquat betrachtet.				

Anhang D.6 – Einschätzung der Übertragbarkeit Studienergebnisse der prognostischen Primärstudien

Item	Frage	QUADAS-Item	ja	nein	unklar
1	Repräsentativität des Patientenspektrums?	1	<input type="checkbox"/>	<input type="checkbox"/>	--
2	Übertragbarkeit Index- und Referenztest?	8 und 9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Übertragbarer Entscheidungsprozess bei Testinterpretation?	Ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	--
4	Übertragbare Hintergrundinformationen?	12	<input type="checkbox"/>	<input type="checkbox"/>	--
5	Stand der Technik übertragbar?	Ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	--
6	Qualifikation der Untersucher übertragbar?	Ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	--
7	Qualität der Geräte übertragbar?	Ergänzt	<input type="checkbox"/>	<input type="checkbox"/>	--
Gesamtbewertung der Übertragbarkeit		Die Übertragbarkeit wird als gegeben angesehen	<input type="checkbox"/>		
		Es gibt Anhaltspunkte für eine eingeschränkte Übertragbarkeit	<input type="checkbox"/>		

Anhang E – Ergebnisse aus einfaktorieller und multifaktorieller Analyse der eingeschlossenen Prognosestudien

Studie	Zielgröße	Ergebnisse aus einfaktorieller und multifaktorieller Analyse
Primärstaging		
Chez-Le Rest 2008	Gesamtüberleben	1) Behandlung nur OP: einfaktorielle Analyse [p = 0,02]; multifaktoriell [p = 0,07]; 2) Grad: einfaktorielle Analyse [p = 0,03]; multifaktoriell [k. A.]. Ebenfalls einfaktoriell modelliert, aber mit p-Wert > 0,05: Geschlecht, Alter, Lokation, Typ (SCC oder ADK), klinisches Stadium, FDG positiv bzgl. Fernmetastasen. Welche Variablen für die multifaktorielle Analyse herangezogen wurden, ist nicht genannt.
Choi 2004	Krankheitsfreies Überleben	1) Behandlung: einfaktorielle Analyse [p = 0,011]; multifaktoriell [k. A.]; 2) pathologisches Stadium: einfaktorielle Analyse [p = 0,004]; multifaktoriell [k. A.]
	Gesamtüberleben	1) Klinisches Stadium: einfaktorielle Analyse [p = 0,034]; multifaktoriell [p = 0,039]; 2) pathologisches Stadium: einfaktorielle Analyse [p < 0,001]; multifaktoriell [p = 0,002]; 3) Behandlung: einfaktorielle Analyse [p = 0,003]; multifaktoriell [k. A.].
Chung 2008	Krankheitsfreies Überleben	1) Stadium: einfaktorielle Analyse [p < 0,0001]; multifaktoriell [p = 0,03]; 2) Tumorgröße: einfaktorielle Analyse [p = 0,02]; multifaktoriell [p = 0,16]; 3) Retentionsindex: einfaktorielle Analyse [p = 0,01]; multifaktoriell [p = 0,23]. Ebenfalls einfaktoriell modelliert, aber mit p-Wert > 0,05: Alter, Verhältnis der TI-Aufnahme von Tumor zu Mediastinum in der SPECT zu früherem Zeitpunkt, Lokation. Diese Variablen gehen nicht in die multifaktorielle Analyse ein.
Omloo 200 ^l	Gesamtüberleben	1) Histologie: einfaktorielle Analyse [p = 0,5]; multifaktoriell [p = k. A.]; 2) Tumorumfang: einfaktorielle Analyse [p = 0,1]; multifaktoriell [p = k. A.]; 3) Tumorort: einfaktorielle Analyse [p = 0,005]; multifaktoriell [p = 0,07]; 4) EUS T-Stadium: einfaktorielle Analyse [p < 0,001]; multifaktoriell [p = 0,007]; 5) EUS N-Stadium: einfaktorielle Analyse [p = 0,006]; multifaktoriell [p = 0,5]; 6) klinisches Stadium: einfaktorielle Analyse [p < 0,006]; multifaktoriell [k. A.].
Shenfine 2009	Gesamtüberleben	1) Pathologisches Stadium: einfaktorielle Analyse [p = 0,001]; multifaktoriell [p = 0,001]; 2) histologischer Grad: einfaktorielle Analyse [k. A.]; multifaktoriell [p = 0,001]; 3) endoskopische Länge (< 6 cm): einfaktorielle Analyse [p = 0,005]; multifaktoriell [k. A.]; 4) endoskopische Länge (> 6 cm): einfaktorielle Analyse [p = 0,04]; multifaktoriell [k. A.]; 5) T-stage: einfaktorielle Analyse [p = 0,002]; multifaktoriell [k. A.]; 6) N-stage: einfaktorielle Analyse [p < 0,001]; multifaktoriell [k. A.]; 7) AJCC: einfaktorielle Analyse [p = 0,001]; multifaktoriell [k. A.]; 8) Tumor differentiation: einfaktorielle Analyse [p = 0,001]; multifaktoriell [k. A.]. Ebenfalls einfaktoriell modelliert, aber mit p-Wert > 0,05: Geschlecht, Alter, Lokation (primary site, Klinisches Stadium, CT Diagnose. Diese Variablen gehen nicht in die multifaktorielle Analyse ein.

(Fortsetzung)

Anhang E – Ergebnisse aus einfaktorieller und multifaktorieller Analyse der eingeschlossenen Prognosestudien (Fortsetzung)

Studie	Zielgröße	Ergebnisse aus einfaktorieller und multifaktorieller Analyse
Shenfine 2009 (Fortsetzung)	Krankheitsfreies Überleben	<p>1) Pathologisches Stadium: einfaktorielle Analyse [p = 0,001]; multifaktoriell [p = 0,001]; 2) histologischer Grad: einfaktorielle Analyse [k. A.]; multifaktoriell [p = 0,08]; 3) Endoskopische Länge (< 6 cm): einfaktorielle Analyse [p = 0,03]; multifaktoriell [k. A.]; 4) endoskopische Länge (> 6 cm): einfaktorielle Analyse [p = 0,007]; multifaktoriell [k. A.]; T-stage: einfaktorielle Analyse [p = 0,001]; multifaktoriell [k. A.]; 6) N-stage: einfaktorielle Analyse [p < 0,001]; multifaktoriell [k. A.]; 7) AJCC: einfaktorielle Analyse [p = 0,001]; multifaktoriell [k. A.]; 8) Tumor differentiation: einfaktorielle Analyse [p = 0,001]; multifaktoriell [k. A.].</p> <p>Ebenfalls einfaktoriell modelliert, aber mit p-Wert > 0,05: Geschlecht, Alter, Lokation (primary site, Klinisches Stadium, CT Diagnose, PET Diagnose).</p> <p>Diese Variablen gehen nicht in die multifaktorielle Analyse ein. Es ist unklar, ob sich die Werte der multifaktoriellen Analyse auf das Gesamtüberleben oder das krankheitsfreie Überleben beziehen</p>
Restaging		
Kim 2007	Krankheitsfreies Überleben	Ebenfalls einfaktoriell modelliert, aber mit p-Wert > 0,05: Geschlecht, Alter, Lokation (primary site, klinisches Stadium, histologische Differenzierung, SUV _{max} des Primärtumors).
Ott 2006	Gesamtüberleben	<p>Es wurden außerdem modelliert: Geschlecht, grading, lauren classification, localization, aber p-Wert > 0,1,(vermutlich einfaktoriell).</p> <p>Multifaktorielle und einfaktorielle Analyse, weitere Faktoren: Zielgröße Gesamtüberleben: 1) ypN0: einfaktorielle Analyse [p = 0,005]; multifaktoriell [p = 0,005]; 2) ypT-Kategorie kleiner als ypT3: einfaktorielle Analyse [p = 0,019]; multifaktoriell [p > 0,05]; 3) histopathologisches Ansprechen: einfaktorielle Analyse [p = 0,03]; multifaktoriell [p > 0,05];</p>
	Rezidivfreies Überleben	ypN0: einfaktorielle Analyse [p = 0,09]; multifaktoriell [p > 0,05].
Rezidivdiagnostik		
Guo 2007	Gesamtüberleben	<p>1) Lokalisation des Primärtumors im oberen Drittel: einfaktorielle Analyse [p= 0,019]; multifaktoriell [p = 0,074]; 2) Behandlung: einfaktorielle Analyse [p = 0,004]; multifaktoriell [p = 0,044]; 3) Krankheitsstatus: einfaktorielle Analyse [p = 0,013]; multifaktoriell [p = 0,045]; 4) Lokalisation des Primärtumors im mittleren Drittel: einfaktorielle Analyse [p = 0,067]; multifaktoriell [p = 0,463];</p> <p>Ebenfalls einfaktoriell modelliert, aber mit p-Wert > 0,05: Geschlecht, Alter, Tumordifferenzierung, klinisches Stadium. Diese Variablen gehen nicht in die multifaktorielle Analyse ein.</p>
AJCC = American Joint Committee on Cancer; CT = Computertomographie; EUS = Endosonografie; FDG = 2-[¹⁸ F]Fluor-Deoxy-D-Glukose ; SUV = standardized uptake value; OP = Operation		