

IQWiG-Berichte – Nr. 172

# **Positronenemissionstomographie (PET) und PET/CT bei Ösophaguskarzinom**

## **Abschlussbericht**

Auftrag: D06-01H  
Version: 1.0  
Stand: 24.06.2013

# 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 (KölnTurm)  
50670 Köln

Tel.: +49 (0)221 – 35685-0

Fax: +49 (0)221 – 35685-1

E-Mail: [berichte@iqwig.de](mailto:berichte@iqwig.de)

Internet: [www.iqwig.de](http://www.iqwig.de)

**ISSN: 1864-2500**

Dieser Bericht wurde unter Beteiligung externer Sachverständiger erstellt. Der Vorbericht wurde einem externen Review unterzogen.

Für die Inhalte des Berichts ist allein das IQWiG verantwortlich.

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. Die Selbstangaben der externen Sachverständigen und der externen Reviewer zu potenziellen Interessenkonflikten sind in Anhang G dargestellt. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden.

#### **Externe Sachverständige:**

- Brigitta G. Baumert, Maastricht University Medical Centre, Maastricht, Niederlande
- Jos Kleijnen, Kleijnen Systematic Reviews, York, Großbritannien, School for Public Health and Primary Care (CAPHRI), Maastricht University, Maastricht, Niederlande
- Marie Westwood, York, Kleijnen Systematic Reviews, York, Großbritannien
- Robert Wolff, York, Kleijnen Systematic Reviews, York, Großbritannien
- Weiterer Sachverständiger, Nuklearmediziner (keine Zustimmung zur Veröffentlichung des Namens), Deutschland

#### **Externes Review des Vorberichts:**

- Florian Lordick, Klinikum Braunschweig, Medizinische Klinik III, Braunschweig (bis 31.08.2012); Universitätsklinikum Leipzig, Universitäres Krebszentrum, Leipzig (aktuell)

#### **Mitarbeiter des IQWiG<sup>1</sup>:**

- Milly Schröer-Günther
- Elke Hausner
- Tatjana Hermanns
- Stefan Sauerland
- Fülöp Scheibler
- Anke Schulz
- Wiebke Sieben

---

<sup>1</sup> Aufgrund gesetzlicher Datenschutzbestimmungen haben Mitarbeiter das Recht, ihrer Namensnennung nicht zuzustimmen.

# Inhaltsverzeichnis

	Seite
<b>Tabellenverzeichnis</b> .....	<b>x</b>
<b>Abbildungsverzeichnis</b> .....	<b>xiv</b>
<b>Abkürzungsverzeichnis</b> .....	<b>xv</b>
<b>Kurzfassung</b> .....	<b>xvii</b>
<b>1 Hintergrund</b> .....	<b>1</b>
<b>1.1 Definition des Krankheitsbildes</b> .....	<b>1</b>
<b>1.2 Epidemiologie und individuelle Krankheitslast</b> .....	<b>1</b>
<b>1.3 Ursache der Erkrankung</b> .....	<b>1</b>
<b>1.4 Verlauf der Erkrankung</b> .....	<b>2</b>
<b>1.5 Klassifikation</b> .....	<b>3</b>
<b>1.6 Diagnostische Verfahren</b> .....	<b>5</b>
<b>1.7 Derzeit übliche therapeutische Verfahren</b> .....	<b>5</b>
<b>1.8 PET und PET/CT</b> .....	<b>6</b>
<b>1.9 Empfehlungen wissenschaftlicher Fachgesellschaften zum Einsatz der PET bei Ösophaguskarzinomen</b> .....	<b>7</b>
<b>2 Ziele der Untersuchung</b> .....	<b>10</b>
<b>3 Projektbearbeitung</b> .....	<b>12</b>
<b>3.1 Zeitlicher Verlauf des Projekts</b> .....	<b>12</b>
<b>3.2 Dokumentation der Änderungen im Projektverlauf</b> .....	<b>13</b>
3.2.1 Änderungen im Vorbericht im Vergleich zum Berichtsplan.....	13
3.2.2 Änderungen im Abschlussbericht im Vergleich zum Vorbericht.....	13
<b>4 Methoden</b> .....	<b>15</b>
<b>4.1 Ermittlung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT</b> .....	<b>16</b>
4.1.1 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT anhand von Evidenzsynthesen.....	17
4.1.2 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT anhand von Primärstudien.....	17
4.1.2.1 Population.....	17
4.1.2.2 Prüf- und Vergleichsinterventionen.....	17
4.1.2.3 Zielgrößen.....	18
4.1.2.4 Studientypen.....	18
4.1.2.5 Sonstige Studiencharakteristika.....	19

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

4.4.2	Bewertung von Evidenzsynthesen zur diagnostischen und prognostischen Güte .....	28
4.4.3	Bewertung von Primärstudien zur diagnostischen und prognostischen Güte .....	29
<b>4.5</b>	<b>Informationssynthese und -analyse .....</b>	<b>29</b>
4.5.1	Gegenüberstellung der Ergebnisse der Einzelstudien .....	30
4.5.1.1	Gegenüberstellung der Ergebnisse von (Primär-)Studien zur Bewertung des patientenrelevanten Nutzens und Schadens.....	30
4.5.1.2	Gegenüberstellung der Ergebnisse von Evidenzsynthesen zur diagnostischen und prognostischen Güte.....	31
4.5.1.3	Gegenüberstellung der Ergebnisse von Primärstudien zur diagnostischen und prognostischen Güte („Ergänzungsrecherche“).....	31
4.5.1.4	Gegenüberstellung der Ergebnisse aus Evidenzsynthesen und der Ergänzungsrecherche („Robustheitsprüfung“) .....	32
4.5.2	Meta-Analysen .....	32
4.5.2.1	Meta-Analysen für vergleichende Studien .....	32
4.5.2.2	Meta-Analysen für Prognosestudien.....	33
4.5.2.3	Meta-Analysen für Diagnosestudien .....	33
4.5.3	Sensitivitätsanalysen .....	34
4.5.4	Subgruppenmerkmale und andere Effektmodifikatoren .....	34
<b>4.6</b>	<b>Änderungen der Methodik.....</b>	<b>35</b>
4.6.1	Vorbericht im Vergleich zum Berichtsplan.....	35
4.6.2	Abschlussbericht im Vergleich zum Vorbericht .....	36
<b>5</b>	<b>Ergebnisse.....</b>	<b>37</b>
<b>5.1</b>	<b>Ergebnisse der Informationsbeschaffung .....</b>	<b>37</b>
5.1.1	Ergebnis der bibliografischen Literaturrecherche .....	37
5.1.2	Weitere publizierte und nicht publizierte Studien.....	39
5.1.2.1	Evidenzsynthesen .....	39
5.1.2.2	Studienregister .....	40
5.1.2.3	Suche in den Datenbanken der Leitlinienanbieter .....	40
5.1.2.4	Unterlagen des G-BA .....	40
5.1.2.5	Kongressbände.....	40
5.1.2.6	Informationen aus der Anhörung.....	41
5.1.2.7	Zusätzliche Informationen zu relevanten Studien .....	41
5.1.3	Resultierender Studienpool .....	41
5.1.3.1	Studien zur Bewertung des patientenrelevanten Nutzens.....	41
5.1.3.2	Evidenzsynthesen zur diagnostischen und prognostischen Güte.....	42
5.1.3.3	Primärstudien zur diagnostischen und prognostischen Güte (Ergänzungsrecherche) .....	43
<b>5.2</b>	<b>Teilziel 1: patientenrelevanter Nutzen .....</b>	<b>46</b>

<b>5.3</b>	<b>Teilziel 2a: diagnostische und prognostische Güte beim Primärstaging .....</b>	<b>47</b>
5.3.1	Diagnostische und prognostische Güte auf Grundlage in die Bewertung eingeflossener Evidenzsynthesen beim Primärstaging.....	47
5.3.2	Diagnostische Güte der eingeschlossenen Primärstudien beim Primärstaging....	47
5.3.2.1	Studiendesigns und Studienpopulationen der eingeschlossenen Primärstudien zum Primärstaging.....	50
5.3.2.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien beim Primärstaging .....	63
5.3.2.3	Einschätzung der Übertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien auf den deutschen Versorgungskontext beim Primärstaging ..	64
5.3.2.4	Ergebnisse zur diagnostischen Güte der eingeschlossenen Studien beim Primärstaging .....	66
5.3.3	Prognostische Güte auf Grundlage von Primärstudien beim Primärstaging (Ergänzungsrecherche) .....	88
5.3.3.1	Studiendesigns und Studienpopulationen der eingeschlossenen Prognosestudien beim Primärstaging .....	90
5.3.3.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Prognosestudien beim Primärstaging .....	95
5.3.3.3	Einschätzung der Übertragbarkeit der Ergebnisse der eingeschlossenen Prognosestudien beim Primärstaging auf den deutschen Versorgungskontext.....	98
5.3.3.4	Ergebnisse zur prognostischen Güte der eingeschlossenen Primärstudien beim Primärstaging .....	99
<b>5.4</b>	<b>Teilziel 2b: diagnostische und prognostische Güte beim Restaging.....</b>	<b>105</b>
5.4.1	Diagnostische und prognostische Güte beim Restaging auf Grundlage in die Bewertung eingeflossener Evidenzsynthesen.....	105
5.4.1.1	Studiendesigns und Studienpopulationen der eingeschlossenen Evidenzsynthesen beim Restaging .....	105
5.4.1.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Evidenzsynthesen .....	108
5.4.1.3	Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien aus den Evidenzsynthesen.....	108
5.4.1.4	Ergebnisse zur diagnostischen Güte auf Grundlage der eingeschlossenen Evidenzsynthesen .....	109
5.4.2	Diagnostische Güte beim Restaging auf Grundlage von Primärstudien (Ergänzungsrecherche) .....	115
5.4.2.1	Studiendesigns und Studienpopulationen der eingeschlossenen Primärstudien beim Restaging.....	115
5.4.2.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien beim Restaging (Ergänzungsrecherche).....	122

5.4.2.3	Einschätzung der Übertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien auf den deutschen Versorgungskontext beim Restaging (Ergänzungsrecherche) .....	122
5.4.2.4	Ergebnisse zur diagnostischen Güte der eingeschlossenen Primärstudien beim Restaging (Ergänzungsrecherche) .....	124
5.4.3	Prognostische Güte der eingeschlossenen Primärstudien beim Restaging (Evidenzsynthesen und Ergänzungsrecherche) .....	131
5.4.3.1	Studiendesigns und Studienpopulationen der eingeschlossenen Prognosestudien beim Restaging (Evidenzsynthesen und Ergänzungsrecherche).....	131
5.4.3.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Prognosestudien beim Restaging (Evidenzsynthesen und Ergänzungsrecherche).....	136
5.4.3.3	Einschätzung der Übertragbarkeit der Ergebnisse der Prognosestudien auf den deutschen Versorgungskontext beim Restaging (Evidenzsynthesen und Ergänzungsrecherche).....	138
5.4.3.4	Ergebnisse der Primärstudien zur prognostischen Güte beim Restaging (Evidenzsynthesen und Ergänzungsrecherche) .....	138
5.4.3.5	Studiendesign und Studienpopulation der Studie zur „prognostic accuracy“ beim Restaging (Ergänzungsrecherche) .....	142
5.4.3.6	Einschätzung des Verzerrungspotenzials der Studie zur „prognostic accuracy“ beim Restaging .....	144
5.4.3.7	Einschätzung der Übertragbarkeit der Studie zur „prognostic accuracy“ beim Restaging .....	144
5.4.3.8	Ergebnisse der Studie zur „prognostic accuracy“ beim Restaging.....	145
<b>5.5</b>	<b>Teilziel 2c: diagnostische und prognostische Güte bei der Rezidivdiagnostik..</b>	<b>146</b>
5.5.1	Diagnostische und prognostische Güte auf Grundlage in die Bewertungen eingeflossener Evidenzsynthesen bei der Rezidivdiagnostik .....	146
5.5.2	Diagnostische Güte auf Grundlage der eingeschlossenen Primärstudien bei der Rezidivdiagnostik .....	146
5.5.3	Studiendesigns und Studienpopulationen der eingeschlossenen Primärstudien bei der Rezidivdiagnostik .....	147
5.5.3.1	Einschätzung des Verzerrungspotenzials der eingeschlossenen Primärstudien bei der Rezidivdiagnostik.....	150
5.5.3.2	Einschätzung der Übertragbarkeit der Ergebnisse der Primärstudien auf den deutschen Versorgungskontext .....	151
5.5.3.3	Ergebnisse der Primärstudien zur diagnostischen Güte bei der Rezidivdiagnostik .....	152
5.5.4	Prognostische Güte auf Grundlage der Primärstudien bei der Rezidivdiagnostik .....	155
5.5.4.1	Studiendesign und Studienpopulation der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	155



5.5.4.2	Einschätzung des Verzerrungspotenzials der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	158
5.5.4.3	Einschätzung der Übertragbarkeit der Ergebnisse der Prognosestudie auf den deutschen Versorgungskontext bei der Rezidivdiagnostik.....	160
5.5.4.4	Ergebnisse der Primärstudien zur prognostischen Güte bei der Rezidivdiagnostik .....	160
<b>5.6</b>	<b>Zusammenfassung der Ergebnisse .....</b>	<b>162</b>
<b>6</b>	<b>Diskussion.....</b>	<b>164</b>
<b>6.1</b>	<b>Gesamtbewertung der Evidenzlage.....</b>	<b>164</b>
6.1.1	Studien zur Untersuchung des patientenrelevanten Nutzens.....	164
6.1.2	Diagnostische und prognostische Güte der PET und PET/CT beim Primärstaging.....	164
6.1.3	Diagnostische und prognostische Güte der PET und PET/CT beim Restaging.	166
6.1.4	Diagnostische und prognostische Güte der PET und PET/CT bei der Rezidivdiagnostik .....	174
<b>6.2</b>	<b>Vergleich mit den Empfehlungen wissenschaftlicher Fachgesellschaften zum Einsatz der PET und PET/CT beim Ösophaguskarzinom .....</b>	<b>174</b>
<b>6.3</b>	<b>Strahlenschutzüberlegungen.....</b>	<b>177</b>
<b>6.4</b>	<b>Allgemeine methodische Probleme.....</b>	<b>178</b>
<b>6.5</b>	<b>Würdigung der Anhörung zum Vorbericht .....</b>	<b>178</b>
6.5.1	Anmerkungen zum Hintergrundkapitel.....	178
6.5.2	Anmerkungen zu den Methoden .....	179
6.5.3	Patientenrelevanter Nutzen.....	182
6.5.4	Indikation Primärstaging .....	184
6.5.5	Indikation Restaging.....	186
6.5.6	Primärstudien bzw. Evidenzsynthesen mit Einschlussforderung.....	188
6.5.7	Ausschluss von Ergebnissen aufgrund fehlender Werte größer als 30 % .....	189
6.5.8	Anmerkungen zum Ausschluss von Studien .....	189
6.5.9	Einschätzung der Übertragbarkeit der Ergebnisse .....	190
<b>7</b>	<b>Fazit.....</b>	<b>191</b>
<b>8</b>	<b>Liste der eingeschlossenen Studien .....</b>	<b>192</b>
<b>9</b>	<b>Literatur .....</b>	<b>200</b>
<b>Anhang A</b>	<b>– Suchstrategien .....</b>	<b>216</b>
<b>Anhang B</b>	<b>– Liste der im Volltext überprüften, aber ausgeschlossenen Evidenzsynthesen mit Ausschlussgründen .....</b>	<b>223</b>
<b>Anhang C</b>	<b>– Liste der ausgeschlossenen Primärstudien mit Ausschlussgründen .....</b>	<b>227</b>
<b>Anhang D</b>	<b>– Beurteilung der Qualität von Evidenzsynthesen.....</b>	<b>245</b>
<b>Anhang D.1</b>	<b>– Qualitätsbeurteilung von Evidenzsynthesen nach den Kriterien von Oxman und Guyatt .....</b>	<b>245</b>

<b>Anhang D.2 – Ergebnisse der Qualitätsbewertung der Evidenzsynthesen, die nach Oxman und Guyatt bewertet wurden, aber nicht in den Bericht eingeschlossen wurden.....</b>	<b>246</b>
<b>Anhang D.3 – Bewertung des Verzerrungspotenzials von Primärstudien auf Basis von QUADAS-Kriterien .....</b>	<b>247</b>
<b>Anhang D.4 – Einschätzung der Übertragbarkeit der Studienergebnisse von diagnostischen Primärstudien.....</b>	<b>248</b>
<b>Anhang D.5 – Bewertung des Verzerrungspotenzials von prognostischen Primärstudien.....</b>	<b>249</b>
<b>Anhang D.6 – Einschätzung der Übertragbarkeit der Studienergebnisse von prognostischen Primärstudien .....</b>	<b>250</b>
<b>Anhang E – Ergebnisse aus einfaktorieller und multifaktorieller Analyse der eingeschlossenen Prognosestudien .....</b>	<b>251</b>
<b>Anhang F – TNM-7-Klassifikation .....</b>	<b>253</b>
<b>Anhang G – Darlegung potenzieller Interessenkonflikte der externen Sachverständigen und der externen Reviewer .....</b>	<b>255</b>

## Tabellenverzeichnis

	<b>Seite</b>
Tabelle 1: Inzidenz und Mortalität nach Altersgruppen in Deutschland 2008 (Fälle pro 100 000).....	2
Tabelle 2: Stadieneinteilung nach der TNM-Klassifikation in der Version 6.....	4
Tabelle 3: Stadieneinteilung nach AJCC / UICC.....	4
Tabelle 4: Anwendung der FDG-PET bei Ösophaguskarzinomen. Bewertung durch die 3. Interdisziplinäre Konsensuskonferenz Onko-PET III.....	9
Tabelle 5: Indikationen für den Einsatz der PET im Rahmen der Diagnostik von Ösophaguskarzinomen .....	11
Tabelle 6: Klassifikationsschemata für Evaluierungsstudien zu diagnostischen Testverfahren.....	15
Tabelle 7: Evidenzklassifizierung des G-BA für Unterlagen zu diagnostischen Methoden....	16
Tabelle 8: Tabellarische Darstellung der Einschlusskriterien für Primärstudien zur Nutzenbewertung der PET bzw. PET/CT .....	19
Tabelle 9: Tabellarische Darstellung der Einschlusskriterien für Evidenzsynthesen zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT .....	23
Tabelle 10: Tabellarische Darstellung der Einschlusskriterien für Primärstudien zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT (Ergänzungsrecherche).....	24
Tabelle 11: Suche in Kongressbänden .....	41
Tabelle 12: Studienpool zur Beurteilung der diagnostischen und prognostischen Güte der PET bzw. PET/CT zum Restaging .....	42
Tabelle 13: Überblick über die in den Evidenzsynthesen eingeschlossenen Primärstudien (n = 11) zur diagnostischen Güte beim Restaging (alphabetisch geordnet).....	43
Tabelle 14: Studienpool zur diagnostischen und prognostischen Güte der eingeschlossenen Primärstudien (Ergänzungsrecherche).....	44
Tabelle 15: Designs, primäre Studienziele und untersuchte Zielgrößen der eingeschlossenen Primärstudien beim Primärstaging .....	50
Tabelle 16: Index- und Referenztest der eingeschlossenen Primärstudien beim Primärstaging .....	54
Tabelle 17: Eingeschlossene Patienten der Primärstudien beim Primärstaging (Ergänzungsrecherche).....	60
Tabelle 18: Verzerrungspotenzial der eingeschlossenen Primärstudien beim Primärstaging..	64

Tabelle 19: Indizien für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien beim Primärstaging .....	66
Tabelle 20: Ergebnisse der eingeschlossenen Primärstudien zur diagnostischen Güte der PET bzw. PET/CT und anderer Indextests beim Primärstaging .....	68
Tabelle 21: Schätzung der Sensitivität und Spezifität aus der bivariaten Meta-Analyse beim N-Staging .....	84
Tabelle 22: Ergänzend betrachtete Ergebnisse zu Änderungen im klinischen Management im Vergleich zur konventionellen Diagnostik (CT und EUS) beim Primärstaging.....	87
Tabelle 23: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudien beim Primärstaging .....	90
Tabelle 24: Index- und Referenztest der eingeschlossenen Prognosestudien beim Primärstaging .....	92
Tabelle 25: Eingeschlossene Patienten der eingeschlossenen Prognosestudien beim Primärstaging .....	93
Tabelle 26: Verzerrungspotenzial der eingeschlossenen Prognosestudien beim Primärstaging .....	97
Tabelle 27: Indizien für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien zum Primärstaging .....	98
Tabelle 28: Ergebnisse der Primärstudien zur prognostischen Güte beim Primärstaging .....	100
Tabelle 29: Fragestellung und Methodik der Evidenzsynthesen zur diagnostischen und prognostischen Güte der PET bzw. PET/CT beim Restaging.....	106
Tabelle 30: Ergebnisse der Qualitätsbewertung der eingeschlossenen Evidenzsynthesen nach Oxman und Guyatt.....	108
Tabelle 31: Ergebnisse der Studien aus den Evidenzsynthesen zum Restaging .....	111
Tabelle 32: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche).....	116
Tabelle 33: Index- und Referenztest der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche).....	118
Tabelle 34: Eingeschlossene Patienten der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche).....	120
Tabelle 35: Verzerrungspotenzial der eingeschlossenen Studien beim Restaging (Ergänzungsrecherche).....	122
Tabelle 36: Indizien für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Studien beim Restaging (Ergänzungsrecherche).....	124

Tabelle 37: Ergebnisse der Studien zur diagnostischen Güte der PET und anderer Indextests beim Restaging von Ösophaguskarzinomen (Ergänzungsrecherche).....	125
Tabelle 38: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudien beim Restaging .....	133
Tabelle 39: Index- und Referenztest der eingeschlossenen Prognosestudien beim Restaging.....	134
Tabelle 40: Eingeschlossene Patienten der eingeschlossenen Prognosestudien beim Restaging.....	135
Tabelle 41: Verzerrungspotenzial der eingeschlossenen Prognosestudien beim Restaging ..	137
Tabelle 42: Indizien für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Prognosestudien bei der Restaging .....	138
Tabelle 43: Ergebnisse der Studien zur prognostischen Güte der PET beim Restaging.....	140
Tabelle 44: Design, primäres Studienziel und untersuchte Zielgröße der Studie zur „prognostic accuracy“ beim Restaging .....	142
Tabelle 45: Index- und Referenztest der Studie zur „prognostic accuracy“ beim Restaging	142
Tabelle 46: Eingeschlossene Patienten der Studie zur „prognostic accuracy“ beim Restaging.....	143
Tabelle 47: Verzerrungspotenzial der Studie zur „prognostic accuracy“ bei der Rezidivdiagnostik.....	144
Tabelle 48: Indizien für eine Nichtübertragbarkeit der Ergebnisse der Studie zur „prognostic accuracy“ beim Restaging .....	144
Tabelle 49: Ergebnisse der Studie zur „prognostic accuracy“ bei der Restaging .....	145
Tabelle 50: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Studien bei der Rezidivdiagnostik.....	147
Tabelle 51: Index- und Referenztest der eingeschlossenen Studien bei der Rezidivdiagnostik.....	148
Tabelle 52: Eingeschlossene Patienten der eingeschlossenen Studien bei der Rezidivdiagnostik.....	149
Tabelle 53: Verzerrungspotenzial der eingeschlossenen Studien bei der Rezidivdiagnostik	150
Tabelle 54: Indizien für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Studien bei der Rezidivdiagnostik.....	151
Tabelle 55: Ergebnisse der Studien zur diagnostischen Güte der PET bzw PET/CT bei der Rezidivdiagnostik.....	153
Tabelle 56: Ergänzend betrachtete Ergebnisse zu Änderungen im klinischen Management im Vergleich zur CT bei der Rezidivdiagnostik.....	155

Tabelle 57: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik .....	156
Tabelle 58: Index- und Referenztest der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	156
Tabelle 59: Eingeschlossene Patienten der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	157
Tabelle 60: Verzerrungspotenzial der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik.....	159
Tabelle 61: Indizien für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik .....	160
Tabelle 62: Ergebnisse der Studien zur prognostischen Güte der PET bei der Rezidivdiagnostik.....	161
Tabelle 63: Ergebnisse der Evidenzsynthesen Van Westreenen 2004, Van Vliet 2008 und des IQWiG-Abschlussberichts beim N-Staging bei Ösophaguskarzinomen .....	166
Tabelle 64: Schlussfolgerungen der eingeschlossenen Evidenzsynthesen beim Restaging von Ösophaguskarzinomen .....	168
Tabelle 65: Stadieneinteilung nach dem TNM-System 7 .....	253
Tabelle 66: Stadieneinteilung von Plattenepithelkarzinomen der Speiseröhre nach dem AJCC .....	254
Tabelle 67: Stadieneinteilung von Adenokarzinomen der Speiseröhre nach dem AJCC .....	254

## Abbildungsverzeichnis

	<b>Seite</b>
Abbildung 1: Ergebnisse der bibliografischen Literaturrecherche und des Literaturscreenings .....	37
Abbildung 2: Forestplots der univariaten Ergebnisse für Sensitivität und Spezifität beim N-Staging bei Ösophaguskarzinomen .....	82
Abbildung 3: Darstellung der direkten Vergleiche der PET und CT beim N-Staging bei Ösophaguskarzinomen .....	83
Abbildung 4: 95 %-Konfidenzregionen und Punktschätzungen für Sensitivität und Spezifität aus der bivariaten Meta-Analyse beim N-Staging bei Ösophaguskarzinomen .....	84
Abbildung 5: Forestplots der univariaten Ergebnisse für Sensitivität und Spezifität beim M-Staging bei Ösophaguskarzinomen .....	85
Abbildung 6: Forestplots der univariaten Ergebnisse für Sensitivität und Spezifität beim Interimsstaging von Ösophaguskarzinomen .....	129
Abbildung 7: Forestplots der univariaten Ergebnisse für Sensitivität und Spezifität beim Restaging von Ösophaguskarzinomen nach Abschluss der Therapie .....	129
Abbildung 8: Forestplot der Ergebnisse für die diagnostische Güte der PET bei der Rezidiverkennung von Ösophaguskarzinomen .....	154

**Abkürzungsverzeichnis**

<b>Abkürzung</b>	<b>Bedeutung</b>
AEG	Adenokarzinome des gastroösophagealen Übergangs
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-[ <sup>18</sup> 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



<b>Abkürzung</b>	<b>Bedeutung</b>
PICO	Patient, intervention, comparison (Vergleichsintervention), outcome (patientenrelevanter Endpunkt)
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

## **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 Abschlussbericht 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 letzte Suche fand am 18.05.2012 statt.

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 und der Rezidivdiagnostik ermöglicht hätte. Auch die Suche in Kongressbänden lieferte keine Hinweise auf noch laufende vergleichende Studien. In den Studienregistern konnten 4 laufende Studien identifiziert werden.

### ***Diagnostische und prognostische Güte***

3 Evidenzsynthesen und 37 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 48 Primärstudien: Primärstaging (n = 25), Restaging (n = 19) und Rezidivdiagnostik (n = 4).

#### ***Primärstaging***

Für das N-Staging wurden direkte Vergleiche zwischen PET und CT in 12 Studien untersucht. In den 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 7 Studien untersucht. Die Ergebnisse dieser Meta-Analyse wurden nicht dargestellt, da die Daten dieser 7 Studien

nicht geeignet waren, um präzise Schätzungen zu berechnen. Auch in Sensitivitätsanalysen ergaben sich keine präzisen Schätzungen.

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*

Für das Restaging nach Abschluss einer neoadjuvanten Therapie wurden in 3 Studien direkte Vergleiche durchgeführt:

Der Vergleich PET/CT versus Endosonographie (EUS) wurde in der Studie von Cerfolio 2005 untersucht. Es zeigte sich für die PET/CT eine Sensitivität von 86,7 %, (95 %-KI [59,7; 98,3]) und eine Spezifität von 87,9 %, (95 %-KI [71,8; 96,6]). Im Gegensatz dazu hatte die EUS eine Sensitivität von 20 %, (95 %-KI [4; 48]) und eine Spezifität von 94 %, (95 %-KI [79,8; 99,3]). In der Studie von Roedl 2009 wurde PET versus PET/CT untersucht. Hier zeigte sich für die PET eine Sensitivität von 59 %, (95 %-KI [36,4; 79,3]) und eine Spezifität von 100 %, (95 %-KI [87,2; 100]). Die PET/CT erreichte eine Sensitivität von 68 %, (95 %-KI [45,1; 86,1]) und eine Spezifität von 100 %, (95 %-KI [87,2; 100]). In der Studie von Kroep 2003 wurde die PET mit der EUS und der CT basierend auf 10 bzw. 11 eingeschlossenen Patienten verglichen. Für die PET zeigte sich eine Sensitivität von 100 %, (95 %-KI [39,8; 100]) und eine Spezifität von 100 % (95 %-KI [54,1; 100]). Die Vergleichstechnologie EUS hatte ebenfalls eine Sensitivität von 100 % (95 %-KI [47,8; 100]) und eine Spezifität von 100 % (95 %-KI [54,1; 100]). Im Gegensatz dazu hatte die CT eine Sensitivität von 50 % (95 %-KI [6,8; 93,2]) und eine Spezifität von 71 % (95 %-KI [29,0; 96,3]).

### *Interimsstaging*

Für das Interimsstaging lag 1 direkter Vergleich diagnostischer Technologien vor:

In der Studie von Kroep 2003 wurde die PET mit der EUS und der CT basierend auf 11 eingeschlossenen Patienten verglichen. Die PET hatte eine Sensitivität von 100 % (95 %-KI [39,8; 100]) und eine Spezifität von 85,7 % (95 %-KI [42,1; 99,6]). Die Vergleichstechnologie EUS hatte eine Sensitivität von 100 % (95 %-KI [47,8; 100]) und eine Spezifität von 100 % (95 %-KI [54,1; 100]). Im Gegensatz dazu hatte die CT eine Sensitivität von 50 % (95 %-KI [6,8; 93,2]) und eine Spezifität von 71 % (95 %-KI [29,0; 96,3]).

Da es in der Indikation Interimsstaging während einer neoadjuvanten Chemotherapie bzw. Chemoradiotherapie nur eine Studie gibt, die einen direkten Vergleich auf Basis von nur 11 Patienten untersucht, werden hier auch ergänzend die Ergebnisse der nicht-vergleichenden Studien dargestellt. In den 7 Studien lagen die Punktschätzungen der Sensitivität für die PET bzw. PET/CT im Bereich von 44 % (zugehörige Spezifität 52 %) bis 88,9 % (zugehörige Spezifität 61,4 %). Die Punktschätzungen der Spezifität lagen im Bereich von 52 %

(zugehörige Sensitivität 44 %) bis 78 % (zugehörige Sensitivität 80 %). Für alle Punktschätzungen waren weite Konfidenzintervalle zu beobachten.

### ***Rezidivdiagnostik***

Im Rahmen der Ergänzungsrecherche wurden 4 Primärstudien eingeschlossen. Alle 4 Studien untersuchten die diagnostische Güte der PET bzw. PET/CT. Eine Studie untersuchte zusätzlich 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 Interventionsstudien zum patientenrelevanten Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen identifiziert werden. Ob der Einsatz der PET bzw. PET/CT auch zu einer Verbesserung patientenrelevanter Endpunkte führt, ist bisher in keiner der 3 Indikationen untersucht worden.

Zu der 2. Fragestellung des Berichts, der diagnostischen und prognostischen Güte, konnten aus 3 eingeschlossenen Evidenzsynthesen und der Ergänzungsrecherche insgesamt 48 Primärstudien eingeschlossen werden. Über alle Indikationen hinweg, wurde in 19 Studien ein direkter Vergleich der PET bzw. PET/CT mit einem anderen bildgebenden Verfahren durchgeführt. In einer Studie wurde der Tracer FDG mit dem Tracer <sup>11</sup>C-Cholin direkt verglichen.

Die meisten Studien wurden in der Indikation Primärstaging durchgeführt. Zwölf Studien verglichen die PET direkt mit der CT beim N-Staging und 7 Studien untersuchten den direkten Vergleich von PET mit CT beim M-Staging. Beim N-Staging konnte in der bivariaten Meta-Analyse kein statistisch signifikanter Unterschied zwischen der diagnostischen Güte von PET und CT festgestellt werden. Beim M-Staging ergab die bivariate Meta-Analyse keine präzisen Schätzungen, sodass auch hier keine Aussage bezüglich der diagnostischen Güte von PET und CT getroffen werden kann. 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. Bisher konnten keine abgeschlossenen parallel vergleichende Studien zum patientenrelevanten Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen identifiziert werden. Methodisch hochwertige Studien sind (insbesondere bezüglich der Fragestellung des Behandlungsansprechens neoadjuvanter Therapien) dringend erforderlich, um den patienten-relevanten Nutzen oder Schaden der PET bzw. PET/CT von Ösophaguskarzinomen verlässlich bewerten zu können.

**Schlagwörter:** Positronenemissionstomographie, Computertomographie, Ösophagustumoren, Staging, Rezidiv, Systematische Übersicht

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

## **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 50 % bis 60 %) und das Adenokarzinom (25 % bis 30 %) [1].

Während Plattenepithelkarzinome überwiegend im mittleren und oberen Drittel der Speiseröhre entstehen, finden sich Adenokarzinome am häufigsten im unteren Drittel und im gastroösophagealen Übergang [1,2]. 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 erkrankten 2008 nach den Daten der deutschen bevölkerungsbezogenen Krebsregister etwa 4800 Männer und etwa 1380 Frauen an Speiseröhrenkrebs [1,8]. Das mittlere Erkrankungsalter liegt bei Männern bei 67 Jahren, das der Frauen bei 69 Jahren. Laut Angaben des RKI erkranken in Deutschland Männer etwa 4- bis 5-mal häufiger als Frauen. Der Speiseröhrenkrebs verursacht ca. 3 % aller Krebstodesfälle in Deutschland bei Männern und 1 % bei Frauen [1].

### **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 [1,6,9]. Eine familiäre Häufung von Erkrankungsfällen ist bekannt, könnte jedoch auf gemeinsame lebensstilbezogene Risiken zurückzuführen sein [1].

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 [10]. Ebenso ist eine Ernährung, die wenig frisches Obst und Gemüse beinhaltet, mit einem erhöhten Risiko der

Entstehung von Plattenepithel- und Adenokarzinomen assoziiert [6,8]. 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 [1,6,11,12].

Eine Barrett-Metaplasie, eine Schleimhautveränderung infolge von häufigem Rückfluss von Mageninhalt in die Speiseröhre (Refluxösophagitis), wird als Präkanzerose für Adenokarzinome angesehen [1,6].

#### 1.4 Verlauf der Erkrankung

Der Verlauf der Erkrankung unterscheidet sich zwischen Adeno- und Plattenepithelkarzinomen. Nach einer Resektion sind das 5- und das 10-Jahres-Überleben für Adenokarzinome tendenziell besser als für Plattenepithelkarzinome [13]. Die absoluten bzw. relativen 5-Jahres-Überlebensraten für ein Karzinom der Speiseröhre werden für Deutschland mit 14 % bzw. 16 % für Männer und 17 % bzw. 20 % für Frauen angegeben [1]. Eine Unterscheidung nach Adeno- und Plattenepithelkarzinomen erfolgt dabei nicht. Tabelle 1 stellt die Inzidenz und Mortalität der Erkrankung nach Altersgruppen dar.

Tabelle 1: Inzidenz und Mortalität nach Altersgruppen in Deutschland 2008 (Fälle pro 100 000) [8]

Alter in Jahren	Männer		Frauen	
	Inzidenz	Mortalität	Inzidenz	Mortalität
0-4	0,0	0,0	0,0	0,0
5-9	0,0	0,0	0,0	0,0
10-14	0,0	0,0	0,0	0,0
15-19	0,1	0,0	0,0	0,0
20-24	0,1	0,0	0,0	0,0
25-29	0,1	0,1	0,0	0,0
30-34	0,1	0,1	0,1	0,0
35-39	0,8	0,4	0,1	0,3
40-44	2,3	1,4	0,3	0,3
45-49	6,5	4,2	1,9	0,9
50-54	14,2	10,3	1,8	1,6
55-59	22,2	15,8	7,2	3,5
60-64	34,2	23,1	9,5	6,2
65-69	37,0	30,0	8,0	5,6
70-74	42,1	31,2	7,5	7,1
75-79	38,5	33,2	8,1	8,5
80-84	42,5	37,5	10,9	10,4
> 85	33,7	32,4	11,6	14,4



## 1.5 Klassifikation

Klinisch-pathologisch lassen sich Tumoren der Speiseröhre nach dem TNM-6-Schema einteilen [14]. 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. Die Stadieneinteilungen für das Ösophaguskarzinom werden in Tabelle 2 und Tabelle 3 dargestellt. Im Jahr 2010 wurde die neue TNM-Klassifikation in der Version 7 publiziert [15]. In der neuen Klassifikation wurde insbesondere das N-Stadium neu sortiert. Das N-Stadium wird gemäß der TNM-Klassifikation in der Version 7 nach der Anzahl der befallenen Lymphknoten unterteilt. Darüber hinaus entfiel die Kategorie M1a. Ein lymphonodaler Befall der Knoten am Truncus coeliacus, welcher zuvor als M1a bewertet wurde, wird nach der TNM-Klassifikation in der Version 7 nicht mehr als fernmetastasiertes Stadium gewertet [16]. Darüber hinaus wird ein Großteil der Adenokarzinome des gastroösophagealen Übergangs als Ösophaguskarzinome kategorisiert, welche zuvor als Magenkarzinome klassifiziert wurden [15]. Da keine der eingeschlossenen Studien des vorliegenden Berichts die TNM-Klassifikation in der Version 7 anwandte, wird in diesem Abschnitt die alte TNM-Klassifikation in der Version 6 dargestellt. Die neue TNM-Klassifikation in der Version 7 wird in Anhang F präsentiert.

Tabelle 2: Stadieneinteilung nach der TNM-Klassifikation in der Version 6 [14]

<b>T Primärtumor</b>	
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
<b>N Lymphknotenbefall</b>	
NX	regionäre Lymphknoten nicht beurteilbar
N0	Lymphknoten tumorfrei
N1	Befall regionaler Lymphknoten
<b>M Fernmetastasen</b>	
Mx	Fernmetastasen nicht beurteilbar
M0	keine Fernmetastasen
<b>Tumoren des oberen Drittels</b>	
M1a	Metastasen in Lymphknoten im Bereich des Halses
M1b	andere Fernmetastasen
<b>Tumoren des mittleren Drittels</b>	
M1a	nicht anwendbar
M1b	Befall nicht regionaler Lymphknoten oder / und andere Fernmetastasen
<b>Tumoren des unteren Drittels</b>	
M1a	Metastasen nahe dem Truncus coeliacus
M1b	andere Fernmetastasen

Entsprechend den Ergebnissen der TNM-Klassifikation in der Version 6 werden Tumoren der Speiseröhre in verschiedene Grade eingeteilt.

Tabelle 3: Stadieneinteilung nach AJCC / UICC [14]

<b>Stadium</b>	<b>T</b>	<b>N</b>	<b>M</b>
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

Adenokarzinome des gastroösophagealen Übergangs (AEG) werden nach Siewert eingeteilt [17]. Die Klassifikation basiert im Wesentlichen auf der anatomischen Lokalisation, wobei es Überschneidungen mit den zuvor dargestellten Systemen gibt [2].

## **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 [11,12].

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.

Die European Society of Medical Oncology (ESMO) empfiehlt, dass die endgültige Diagnose anhand einer endoskopisch gewonnenen Biopsie gestellt werden sollte, die nach den Kriterien der Weltgesundheitsorganisation befundet wurde [18].

Aufgrund der therapieassoziierten Mortalität wird eine genaue Erfassung des Gesamtzustandes des Patienten empfohlen [11,12]. Dabei können verschiedene Scoringsysteme wie POSSUM (physiological and operative severity score for the enumeration of mortality and morbidity) und ASA (American Society of Anesthesiologists) eingesetzt werden [11,12].

## **1.7 Derzeit übliche therapeutische Verfahren**

Die Therapie von Tumoren der Speiseröhre richtet sich nach der Tumorausbreitung und dem allgemeinen Gesundheitszustand des Patienten [19,20]. In über der Hälfte der Patienten wird eine Operation aufgrund des unzureichenden Gesundheitszustandes abgelehnt [11,12]. 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]) [19-21]. Dieser Eingriff sollte in Einrichtungen, die bereits ausreichende Erfahrungen mit ähnlichen Operationen haben, durch erfahrene Chirurgen durchgeführt werden [11,12]. Eine ablative endoskopische Therapie für Tumoren mit geringer Ausbreitung (T1 / N0 / M0) wird von dem SIGN und der ESMO empfohlen [11,12,18]. Für lokalisierte Tumoren kann unter Umständen Radiotherapie als Primärtherapie zum Einsatz kommen [11,12]. Ein multimodaler Therapieansatz unter Einbeziehung von Operation, Chemotherapie und Strahlentherapie ist einfachen Ansätzen überlegen [11,19-22].

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 [11,12]. Die Radiotherapie kann zum Einsatz kommen, um bei Patienten in schlechtem Allgemeinzustand Schluckbeschwerden zu vermindern [11,12]. In der weiteren Behandlung wird ein symptomorientiertes Vorgehen empfohlen [11,12].

## 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 Nichtvollringsysteme 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 Fluordeoxyglukose ( $[^{18}\text{F}]$ -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 Untersuchungsgang durch die beiden Detektorringsysteme 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 [23].

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 [24].

### **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 veraltet [19]. Parallel zur Erstellung des Abschlussberichts wurde allerdings eine S3-Leitlinie zur „Diagnostik und Therapie der Plattenepithelkarzinome und Adenokarzinome des Ösophagus“ auf der Homepage der AWMF angemeldet. Die Fertigstellung dieser Leitlinie ist für Juni 2015 geplant [25].

Die Empfehlungen in der bereits publizierten S3-Leitlinie zu Adenokarzinomen des Magens und des ösophagogastralen Übergangs sind teilweise für dieses Projekt relevant [26]. In der Leitlinie 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 offen empfohlen. Darüber hinaus wies die Leitlinie darauf hin, dass die PET als

Technologie für die Prognose bezüglich des Gesamtüberlebens bei Patienten ohne Metastasen Verwendung finden sollte. Ebenfalls wurde der Einsatz der PET auch für das Behandlungsansprechen einer Chemotherapie vor einer Operation empfohlen.

Auch international liegen relativ wenige Leitlinien zur Diagnostik von Ösophaguskarzinomen vor, was unter anderem aus einer orientierenden Suche nach dem Begriff „oesophageal cancer“ auf der Homepage des National Guideline Clearinghouse hervorgeht [27]. Oftmals sind diese älter und die PET bzw. PET/CT findet darin keine Erwähnung [9].

Die European Society for Medical Oncology (ESMO) sieht die PET als „hilfreich“ bei der Identifikation von bisher nicht beschriebenen Metastasen und bei der Diagnose bei Verdacht auf ein Rezidiv an [18]. Zudem könne die PET in der Kontrolle des Therapieansprechens eine Rolle spielen [18]. Dagegen wird von dem Scottish Intercollegiate Guidelines Network (SIGN) der routinemäßige Einsatz der PET im Staging von Speiseröhrentumoren nicht empfohlen [11,12].

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 [10]. 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 National Comprehensive Cancer Network (NCCN) empfohlen [28]. Darüber hinaus wurde der Einsatz der PET und der PET/CT für das M-Staging, basierend auf einer Meta-Analyse von van Westreenen 2004 [29], empfohlen.

Basierend auf Expertenmeinungen empfiehlt die kanadische Saskatchewan Cancer Agency den Einsatz der PET im Falle einer multimodalen Therapie für das T- und M-Staging zusammen mit 4 weiteren diagnostischen Verfahren [30].

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 [31]. 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 [31]

<b>Indikation</b>	<b>Anwendung der FDG-PET</b>
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 sollten 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 [32].



### **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 <sup>a</sup>	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 Abschlussbericht 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) [33]. Der vorliegende Abschlussbericht 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 Abschlussbericht 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.

Die vorläufige Bewertung, der Vorbericht in der Version 1.0 vom 15.05.2012, wurde am 23.05.2012 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 21.06.2012 konnten schriftliche Stellungnahmen eingereicht werden. Unklare Aspekte aus den schriftlichen Stellungnahmen zum Vorbericht wurden am 29.08.2012 in einer wissenschaftlichen Erörterung mit den Stellungnehmenden diskutiert. Die Dokumentation der Anhörung zum Vorbericht ist auf der Website des IQWiG veröffentlicht. Die in den Stellungnahmen vorgebrachten Argumente werden im Kapitel „Diskussion“ des vorliegenden Abschlussberichts gewürdigt.

Der vorliegende Abschlussbericht beinhaltet die Änderungen, die sich aus der Anhörung ergeben haben.

Der Vorbericht wurde zusätzlich einem externen Review unterzogen.

Im Anschluss an die Anhörung erstellte das IQWiG den vorliegenden Abschlussbericht, der 8 Wochen nach Übermittlung an den G-BA auf der Website des IQWiG veröffentlicht wird. Die zum Vorbericht eingegangenen Stellungnahmen und das Protokoll der wissenschaftlichen Erörterung werden in einem gesonderten Dokument „Dokumentation und Würdigung der Anhörung zum Vorbericht“ zeitgleich mit dem Abschlussbericht im Internet bereitgestellt.

## **3.2 Dokumentation der Änderungen im Projektverlauf**

### **3.2.1 Änderungen im Vorbericht im Vergleich zum Berichtsplan**

- 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.

### **3.2.2 Änderungen im Abschlussbericht im Vergleich zum Vorbericht**

- In Kapitel 1 wurden mehrere Referenzen aktualisiert und Leitlinien ergänzt.
- In Abschnitt 4.4.3 wurde zur besseren Nachvollziehbarkeit ein projektspezifischer Algorithmus zur Einschätzung des Verzerrungspotenzials dargestellt.
- In Abschnitt 4.5.2 zu den Methoden der Meta-Analysen wurde ein Absatz zu Prognosestudien ergänzt. Ferner wurde die metaanalytische Methodik für Diagnosestudien aktualisiert bezüglich der bivariaten Modellierung und des statistischen Testens auf Unterschiede in der diagnostischen Güte. Darüber hinaus wurden ebenfalls in den Abschnitten 4.5.3 und 4.5.4 Aspekte zu diagnostischen und prognostischen Studien ergänzt.
- In Abschnitt 5.3.2.4 wurden die Ergebnisse der Vierfeldertafel der Studie Kato 2008 für die PET/CT dargestellt.
- In Abschnitt 5.3.2.4 wurden die Ergebnisse der Vierfeldertafeln der Studie Liberale 2004 [34] nicht mehr dargestellt aufgrund der 30-%-Regel.

- In Abschnitt 5.3.2.4 wurden die Ergebnisse zum M1a-Staging aus der Studie Lerut 2000, welche eine Zweitpublikation von Flamen 2000 ist, nachextrahiert und dargestellt.
- In Abschnitt 5.3.2.4 wurden die Ergebnisse der Vierfeldertafeln der Studie Kole 1998 für das M-Staging nachextrahiert und dargestellt.
- In Abschnitt 5.3.3.2 wurde eine Aussage bezüglich des Verzerrungspotenzials zur Studie Chung 2008 korrigiert.
- In Abschnitt 5.4.2.4 wurde im Rahmen der Beschreibung der Abbildung explizit darauf hingewiesen, dass es für das histologische Ansprechen unterschiedliche Klassifikationen gab.
- In der Stellungnahme der DGN wurde der Hinweis gegeben, dass die Übertragbarkeit von Studien, welche asiatische Patienten untersuchten, in bestimmten Aspekten eingeschränkt sei. Aufgrund dieses Hinweises wurde die Einschätzung der Übertragbarkeit folgender Studien in „Indiz für eine Nichtübertragbarkeit“ geändert: Choi 2000, Choi 2004, Choi 2010, Chung 2008, Guo 2007, Higuchi 2008, Itho 2008, Kato 2005, Kato 2008, Kim 2007, Ma 2012, Guo 2007, Yoon 2003, Zhang 2006.
- Im gesamten Abschlussbericht wurde eine redaktionelle Änderung von „Anhaltspunkt für eine Nichtübertragbarkeit“ in „Indiz für Nichtübertragbarkeit“ vorgenommen.

Darüber hinaus ergaben sich im Abschlussbericht im Vergleich zum Vorbericht lediglich redaktionelle Änderungen.

#### 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 [35]. 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 [36,37]. Fryback und Thornbury haben dazu in ihrer Publikation „The efficacy of diagnostic imaging“ von 1991 [36] 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 [37]	Fryback & Thornbury [36]	Evidenzklassifizierung G-BA [38]
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 Denken („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		
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 (2. Kapitel § 11 Abs. 2) dieser Sichtweise angeschlossen [38]. 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 1990 [37] mit aufgenommen.

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

<b>Evidenzklasse nach G-BA [38]</b>	<b>Kriterien</b>
<b>I a</b>	Systematische Übersichtsarbeiten von Studien der Evidenzstufe I b
<b>I b</b>	Randomisierte kontrollierte Studien
<b>I c</b>	Andere Interventionsstudien
<b>II a</b>	Systematische Übersichtsarbeiten von Studien zur diagnostischen Testgenauigkeit der Evidenzstufe II b
<b>II b</b>	Querschnitt- und Kohortenstudien, aus denen sich alle diagnostischen Kenngrößen zur Testgenauigkeit (Sensitivität und Spezifität, Wahrscheinlichkeitsverhältnisse, positiver und negativer prädiktiver Wert) berechnen lassen
<b>III</b>	Andere Studien, aus denen sich die diagnostischen Kenngrößen zur Testgenauigkeit (Sensitivität und Spezifität, Wahrscheinlichkeitsverhältnisse) berechnen lassen
<b>IV</b>	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 [39]. 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 [39].

#### **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 Beobachtungs-



gruppen 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 könnten möglicherweise nicht randomisierten kontrollierten Interventionsstudien Informationen entnommen werden, die bei der Ausgestaltung zukünftiger aussagekräftigerer Studien nützlich sein könnten.

#### 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 <sup>a</sup>
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- [40] oder TREND-Statements [41] genügte und eine Bewertung der Studie ermöglichte. CONSORT: Consolidated Standards of Reporting Trial; TREND: Transparent Reporting of Evaluation with Nonrandomized Designs	

## 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 Variabilitätsschätzungen kommen konnte (sogenannter Verifikationsbias oder Inkorporationsbias) [42-46].

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. [37])
- 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 [47].

#### 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 [48-50].
- 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

<b>Einschlusskriterien zur Evaluierung der diagnostischen und prognostischen Güte anhand von Evidenzsynthesen</b>	
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 <sup>a</sup>
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 [51] oder des MOOSE-Statemats [52] 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 <sup>a</sup>
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 [53] oder STROBE-Statements [54] 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 Recherche wurde während der Projektbearbeitung aktualisiert. Die letzte Suche erfolgte am 18.05.2012.

#### **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**

Da im Rahmen von Berichten auf Basis von Evidenzsynthesen die Nachrecherche dazu dient, die Ergänzungsrecherche nach Primärliteratur zu aktualisieren, wurden Evidenzsynthesen, die im Rahmen der Nachrecherche identifiziert wurden, nicht zur Ergebnisextraktion genutzt. Sie 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 12.06.2012 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 und des Vorberichts erfolgte eine Anhörung, die sich unter anderem auch auf in die Nutzenbewertung einzubeziehende Informationen beziehen konnte. Relevante Informationen aus diesen Anhörungen 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 [49,50]. Diese Kriterien wurden um die Angabe von Interessenkonflikten ergänzt (vgl. 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 [45]. 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).

Obwohl das QUADAS-I-Instrument [45] die Verwendung von Algorithmen nicht empfiehlt, wurde für den vorliegenden Bericht zur besseren Nachvollziehbarkeit ein projektspezifischer Algorithmus zur Einschätzung des Verzerrungspotenzials entwickelt: Wenn eines der Items (bis auf Item 5 „Differenzielle Verifikation“) mit „nein“ angekreuzt wurde, hatte die Studie immer ein hohes Verzerrungspotenzial.

Wenn mindestens 2 der Items mit „unklar“ bewertet wurden, war in den meisten Fällen die Unsicherheit zu groß, als dass der Studie noch ein niedriges Verzerrungspotenzial hätte attestiert werden können. Ausnahmen waren die Items 7, 8 und 10. Wenn diese mit „unklar“ bewertet wurden, fiel dies nicht ins Gewicht.

Für die Bewertung des Verzerrungspotenzials von prognostischen Studien war zusätzlich zu QUADAS die adäquate Berücksichtigung von möglichen Confoundern als ein weiteres Qualitätskriterium vorgesehen [55]. 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. [55] und Altman et al. [56] 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 [57]. 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 [58].

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 [59].

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 [60]. In begründeten Ausnahmefällen sollten Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzungen 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 Effektschätzungen und Konfidenzintervalle aus den Studien sollten mittels Forestplots 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 [61]. 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 Prognosestudien

Die Meta-Analysen von Prognosestudien sollten entsprechend dem Vorgehen bei der Bewertung des patientenrelevanten Nutzens erstellt werden.

#### 4.5.2.3 Meta-Analysen für Diagnosestudien

Die Punktschätzungen und dazugehörigen univariaten 95%-Konfidenzintervalle [62] aus den Studien wurden mittels Forest Plots zusammenfassend dargestellt. Außerdem wurde, sofern die dafür nötigen Anforderungen erfüllt waren, für die Testgütekriterien eine Meta-Analyse anhand der Sensitivität und Spezifität in einer bivariaten Meta-Analyse durchgeführt [63]. Die Schätzung der Modellparameter erfolgte über ein generalisiertes lineares gemischtes Modell [64,65]. Der Algorithmus zum Schätzen der Parameter im bivariaten Modell kann zu unpräzisen Schätzungen führen, das heißt zu Schätzungen mit zu großen Standardfehlern und entsprechenden Konfidenzregionen. Auch kann der Algorithmus gegebenenfalls keine Schätzungen liefern, wenn das Maximum-Likelihood-Verfahren nicht konvergiert. In beiden Fällen fehlen brauchbare Schätzungen. Die Gründe hierfür können beispielsweise sein, dass zu wenige Studien vorliegen oder dass einzelne Studien extreme Werte aufweisen. Sind die resultierenden Schätzungen unpräzise, werden die Ergebnisse der bivariaten Meta-Analysen in der Regel nicht dargestellt. Wenn zu wenige vergleichende Studien für eine bestimmte Technologie vorlagen, konnten Technologien in den bivariaten Meta-Analysen zusammengelegt werden, falls die Analyse eine sinnvolle Interpretation zuließ.

Falls die bivariate Meta-Analyse präzise Schätzungen lieferte, so wurden bei diagnostischen Studien die beobachteten Paare aus Sensitivität und Spezifität zweidimensional grafisch dargestellt. Ergebnisse verschiedener Technologien, die aus derselben Studie stammen, wurden durch eine Verbindungslinie gekennzeichnet. Des Weiteren wurden die aus der bivariaten Meta-Analyse gewonnenen Schätzungen für die Erwartungswerte als gepoolte Paare der Sensitivität und der Spezifität mit den dazugehörigen 95%-Konfidenzregionen dargestellt [66].

Es wurde auf einen Unterschied in der diagnostischen Güte zwischen 2 Technologien mit dem Likelihood-Ratio-Test (LRT) getestet (Test auf Gleichheit der Erwartungswerte von Sensitivität und Spezifität der beiden Technologien). Die Differenzierung zwischen einer „höheren diagnostischen Güte“ und einer „vermutlich höheren diagnostischen Güte“ wurde anhand der grafischen Darstellungen vorgenommen.

Lieferte die bivariate Meta-Analyse unpräzise Schätzungen, die zu sehr großen Konfidenz- und Prädiktionsregionen führten, so wurden keine Aussagen über einen Unterschied in der diagnostischen Güte zwischen 2 Technologien abgeleitet. In Ausnahmefällen, wie beispielsweise beim Vorliegen von mehreren großen Studien mit niedrigem Verzerrungspotenzial, sollten die Ergebnisse geeigneter univariater statistischer Tests, das heißt für die Sensitivität und Spezifität getrennt, dargestellt werden. Zusammen mit den

grafischen Darstellungen sollten gegebenenfalls Aussagen zu einem Unterschied in der diagnostischen Güte zwischen 2 Technologien abgeleitet werden.

Das Vorliegen von Heterogenität wurde anhand von Sensitivitätsanalysen untersucht.

#### **4.5.3 Sensitivitätsanalysen**

Zur Einschätzung der Robustheit der 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. Beim Einschluss von mehreren Evidenzsynthesen sollten, wenn möglich, Sensitivitätsanalysen anhand der Beurteilung mit dem Index von Oxman und Guyatt [49,50] 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, zum Beispiel die Festlegung von Cut-off-Werten und Erhebungszeitpunkten oder die Wahl des Effektmaßes.

Bei Nutzenstudien kann das Ergebnis der Sensitivitätsanalysen die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestuftes Effekt kann zum Beispiel dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen Nutzen oder Schaden attestiert wird. Bei Studien zur diagnostischen Güte wurden 95%-Prädiktionsregionen für den Erwartungswert einer neuen Studie bestimmt, um Ausreißer zu identifizieren. Sensitivitätsanalysen nach dem Leave-one-out-Prinzip wurden durchgeführt, um den Einfluss jeder einzelnen Studie auf das Ergebnis zu bewerten.

Bei Studien zur diagnostischen oder prognostischen Güte können die Ergebnisse von Sensitivitätsanalysen den Aussagen der Primäranalyse ([vermutlich] vorhandene oder nicht vorhandene unterschiedliche oder höhere Güte) widersprechen, ihnen nicht widersprechen oder sie unterstützen. Es muss individuell entschieden werden, ob die Ergebnisaussage verändert wird.

#### **4.5.4 Subgruppenmerkmale und andere Effektmodifikatoren**

Für Studien zum patientenrelevanten Nutzen oder Schaden und für Prognosestudien sollten die Ergebnisse hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies konnten direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen sein. Im Gegensatz zu den in Abschnitt 1.1.3 beschriebenen methodischen Faktoren für Sensitivitätsanalysen war hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken.

Für den Nachweis unterschiedlicher Effekte sollte die auf einem Homogenitäts- bzw. Interaktionstest basierende statistische Signifikanz Voraussetzung sein. In die Untersuchung



von Effektmodifikatoren sollten die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalteten, und aus Subgruppenanalysen einbezogen werden. 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:

- Alter
- Geschlecht
- Tumorentitäten
- 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.

Für Studien zur diagnostischen Güte werden keine Subgruppenanalysen durchgeführt. Da Diagnosestudien keine Effektschätzungen für den Unterschied zwischen 2 diagnostischen Technologien liefern, kann der in der univariaten Meta-Analyse verfolgte Ansatz zur Untersuchung von Effektmodifikationen mittels Meta-Regression und Homogenitäts- bzw. Interaktionstests nicht auf den bivariaten Fall übertragen werden. In der Literatur findet sich kein etablierter Ansatz zur Untersuchung der Modifikation von diagnostischer Güte.

## **4.6 Änderungen der Methodik**

### **4.6.1 Vorbericht 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.

#### **4.6.2 Abschlussbericht im Vergleich zum Vorbericht**

- In Abschnitt 4.5.2 zu den Methoden der Meta-Analysen wurde ein Absatz zu Prognosestudien ergänzt.
- Ferner wurde die metaanalytische Methodik für Diagnosestudien aktualisiert bezüglich der bivariaten Modellierung und des statistischen Testens auf Unterschiede in der diagnostischen Güte.

## 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.

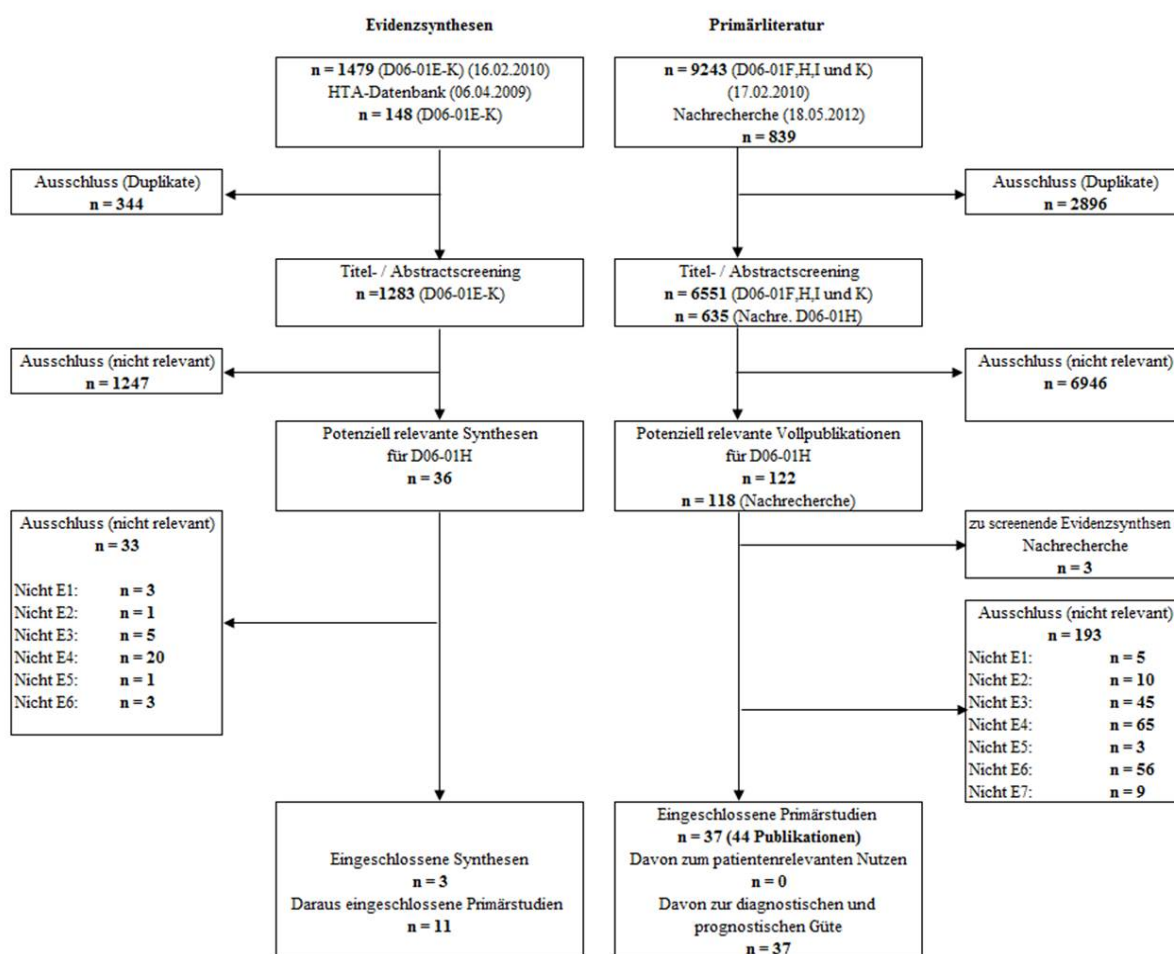


Abbildung 1: Ergebnisse der bibliografischen Literaturrecherche und des Literaturscreenings

#### 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 Primärrecherche wurden für die Krankheitsbilder malignes Melanom (D06-01F), Ösophaguskarzinom (D06-01H), Ovarialkarzinom (D06-01I) und Knochen- und Weichteiltumoren (D06-01K) insgesamt 9243 Referenzen identifiziert. Durch die Nachrecherche kamen 839 Treffer für das Krankheitsbild Ösophaguskarzinom hinzu.

Nach Ausschluss von 2896 Duplikaten ergab sich eine Gesamtzahl von 7186 zu screenenden Treffern.

6946 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 240 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.

In der Nachrecherche fanden die Suche und das Screening nach Evidenzsynthesen gemeinsam mit der nach Primärliteratur statt. Da im Rahmen von Berichten auf Basis von Evidenzsynthesen die Nachrecherche dazu dient, die Ergänzungsrecherche nach Primärliteratur zu aktualisieren, wurden Evidenzsynthesen, die im Rahmen der Nachrecherche identifiziert wurden, nicht zur Ergebnisextraktion genutzt. Sie dienten vielmehr der Identifikation weiterer Primärstudien und sollten darüber hinaus in der Diskussion gewürdigt werden. Durch die Nachrecherche wurden 3 Evidenzsynthesen identifiziert (siehe Abschnitt 5.1.2.1).

## **Ergänzungsrecherche zur diagnostischen und prognostischen Güte**

Für die Indikation Restaging verfügte die Arbeit von Ngamruengphong 2010 [67] 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 240 potenziell relevante Treffer, die im Volltext gesichtet wurden. Hiervon wurden 193 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 befassten, 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 37 Primärstudien (44 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 Nachrecherche 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 Studienregistern ergab 4 laufende Studien:

- NCT 01333033: PET Scan Imaging in Assessing Response in Patients With Esophageal Cancer Receiving Combination Chemotherapy
- NCT 01608464: A Trial Comparing Pre-operative Chemo-radiotherapy With Cisplatin and Fluorouracil Versus Chemotherapy With Docetaxel and Irinotecan in PET Non Responders Resectable Cancer Esophagus: a Multicenter Study
- NCT 01551589: Elective or Prophylactic Nodal Irradiation for Esophageal Cancer (ENI or PNI)
- ACTRN 12609000665235: A Randomised Phase II Trial of Pre-operative cisplatin, 5 fluorouracil and docetaxel +/- Radiotherapy based on poor early response to standard chemotherapy for resectable adenocarcinoma of the oesophagus and/or Oesophago-gastric Junction.

Die Studienleiter wurden angeschrieben und um Zusendung des Studienprotokolls gebeten. Die eingegangenen Studienprotokolle wurden dem G-BA bereits übersandt. Mit der Studiengruppe NCT 01551589 ist der Autorenkontakt nicht gelungen. Die Studiengruppe ACTRN 12609000665235 bat um strenge vertrauliche Behandlung des Protokolls und untersagte die Weiterleitung an Dritte.

### 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 Berichtsfragstellungen geprüft. Von diesen befasste sich eine Studie, Brücher 2001 [68], 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 <sup>a</sup>	1573	0	0
	2009 <sup>a,b</sup>	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 den Anhörungen zum Berichtsplan und zum Vorbericht 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[69,70], Duong 2006a [71], Duong 2006b [72], Omluo 2008 [73], Ott 2006 [74] und Shenfine 2009 [75]) zu erhalten. Zwei Autorenanfragen wurden beantwortet und enthielten ausreichende Informationen. Ebenfalls wurden die Studienleiter der 4 laufenden Studien (vgl. 5.1.2.2) angeschrieben und um die Zusendung des Studienprotokolls gebeten. 3 Studienprotokolle sind eingegangen.

## 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) [67,76,77]. 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 [29] 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	[67]
Rebollo Aquirre 2009	Rebollo Aquirre AC et al. Ann Surg 2009; 250 (2): 247-254.	[76]
Westerterp 2005	Westerterp M et al. Radiology 2005; 236(3): 841-851.	[77]

Ein Studienpiegel 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	[68]	X	-	X
Cerfolio	2005	[78]	X	X	-
Flamen	2002	[79]	X	-	X
Gillham	2006	[80]	X	-	-
Kroep	2003	[81]	X	-	X
Levine	2006	[82]	X	X	-
Ott	2006	[74]	X	-	-
Song	2005	[83]	X	X	-
Weber	2001	[84]	-	-	X
Westerterp	2006	[85]	X	X	-
Wieder	2004	[86]	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 37 Primärstudien (44 Publikationen).

Für das Primärstaging konnten 25 Studien (32 Publikationen) eingeschlossen werden. Von diesen 25 Studien untersuchten 19 ausschließlich die diagnostische Güte der PET bzw. PET/CT. Eine Studie, Shenfine 2009, untersuchte sowohl die diagnostische als auch prognostische Güte, und 5 Studien, Cheze-Le Rest 2008, Choi 2004, Chung 2008, Gillies 2012a und Omloo 2008, 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 8 weitere Studien eingeschlossen werden. 2 dieser Studien (Kim 2007 und Ma 2012) untersuchten dabei ausschließlich die prognostische Güte der PET bzw. PET/CT.

Vier Studien wurden für die Rezidivdiagnostik eingeschlossen. Eine dieser Studie (Guo 2007) 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)

<b>Studie</b>	<b>Referenz</b>
<b>Primärstaging (diagnostische Kohortenstudien)</b>	
Buchmann 2006	[87]
Choi 2000 <sup>a</sup>	[70]
Choi 2010	[88]
Flamen 2000	[89]
	[90]
Heeren 2004	[91]
Itoh 2008	[92]
Jager 2001	[93]
Kato 2005	[94]
Kato 2008	[95]
Kneist 2004	[96]
	[97]
	[98]
	[99]
	[100]
Kole 1998	[101]
Liberale 2004	[34]
Lowe 2005	[102]
Noble 2009	[103]
Rankin 1998	[104]
Schreurs 2008	[105]
Shenfine 2009	[75]
Sihvo 2004	[106]
	[107]
Vyas 2011	[108]
Yoon 2003	[109]
<b>Primärstaging (prognostische Kohortenstudien)</b>	
Cheze-Le Rest 2008	[110]
	[111]
Choi 2004 <sup>a</sup>	[69]
Chung 2008	[112]
Gillies 2012a	[113]
Omloo 2008	[73]
Shenfine 2009	[75]

(Fortsetzung)

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

<b>Restaging (diagnostische Kohortenstudien)</b>	
Gillies 2012b	[114]
Higuchi 2008	[115]
Ilson 2011	[116]
Malik 2010	[117]
Roedl 2009	[118]
<b>Restaging (prognostische Kohortenstudien)</b>	
van Heijl 20 11	[119]
Kim 2007	[120]
Higuchi 2008	[115]
Ma 2012	[121]
<b>Rezidivdiagnostik (diagnostische Kohortenstudien)</b>	
Guo 2007	[122]
Roedl 2008	[123]
Teyton 2009	[124]
Zhang 2006	[125]
<b>Rezidivdiagnostik (prognostische Kohortenstudien)</b>	
Guo 2007	[122]

## **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 daher auf Basis der Primärstudien durchgeführt.

#### 5.3.2 Diagnostische Güte der eingeschlossenen Primärstudien beim Primärstaging

In der Recherche wurden 20 Primärstudien zur diagnostischen Güte der PET und PET/CT beim Primärstaging identifiziert (vgl. Tabelle 14). In 16 Studien wurden Angaben zum Geschlecht der eingeschlossenen Patienten gemacht, die in der Summe 1048 Männer und 180 Frauen ergeben. In 13 von 20 Studien wurden Patienten mit einem Plattenepithelkarzinom oder einem Adenokarzinom der Speiseröhre untersucht. In 4 Studien wurden ausschließlich Patienten mit einem Plattenepithelkarzinom [70,87,95,109] und in 3 Studien nur Patienten mit einem Adenokarzinom der Speiseröhre eingeschlossen [75,106,108]. In den Studien Rankin 1998 und Kato 2005 fehlten die Angaben bzgl. der Primärdiagnose. 60 % der eingeschlossenen Studien wurden in Europa (2 Studien in Deutschland), 30 % im asiatischen Raum und 10 % der Studien in den USA oder Australien durchgeführt. Der Einsatz der nicht integrierten FDG-PET wurde in 16 von 20 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 <sup>11</sup>C-Cholin eingesetzt. Beim Vergleich der Ziele der Studien fiel auf, dass sie unterschiedliche Fragestellungen untersuchten. 8 der eingeschlossenen Studien analysierten ausschließlich das N-Staging, 3 Studien (Kato 2005, Noble 2009, Vyas 2012) das M-Staging und 9 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** 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 Chemoradiotherapie 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** 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 [89,90], wobei sich die Publikation von **Lerut 2000** 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 der Definition des Abschlussberichts 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 zum N-Staging konnten mithilfe der Tabelle 1 der Publikation von Lerut 2000 rekonstruiert und im vorliegenden Bericht dargestellt werden. Darüber hinaus konnten die Daten der Vierfeldertafel für das M1a-Staging ebenfalls aus Tabelle 1 der Publikation Lerut 2000 rekonstruiert werden. Laut Lerut 2000 entsprach auch hier der Referenzstandard den Einschlusskriterien des vorliegenden Berichts.

Zu **Itoh 2008** 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** untersuchte als einzige neben dem Tracer FDG auch den Einsatz des Tracers <sup>11</sup>C-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**, die zwischen 1999 und 2005 rekrutiert wurden (n = 117), lagen getrennte Ergebnisse für PET und CT vor. Die Patienten, die 2006 oder 2007 nachrekrutiert wurden (n = 50), wurden mittels PET/CT untersucht. Die Studie berichtete die Ergebnisse für PET, CT und PET/CT. Allerdings waren die Daten der Vierfeldertafel der PET/CT nicht mit den Daten der PET und CT vergleichbar, da es sich aufgrund der Nachrekrutierung um unterschiedliche Patienten handelte .

Die Ergebnisse der Arbeit von **Kneist 2004** wurden in 5 verschiedenen Publikationen berichtet [96-100]. 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** 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. Aufgrund der 30-%-Regel wurden allerdings die Daten der Vierfeldertafel für beide Indextests nicht dargestellt. 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 **Shenfine 2009** 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 **Sihvo 2004** gibt es eine weitere Arbeit, Rasanen 2003 [107], in der Zwischenergebnisse präsentiert wurden. Sihvo 2004 wurde als Hauptpublikation der Studie angesehen. Die im vorliegenden Bericht dargestellten Daten stammen aus Sihvo 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** 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 <sup>a</sup>	Zielkriterium	Evidenzstufe <sup>b</sup>	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 <sup>a</sup>	Zielkriterium	Evidenzstufe <sup>b</sup>	Verzerrungspotenzial
Jager 2001	Diagnostische Kohortenstudie Keine Angaben zum Rekrutierungs- und Untersuchungszeitraum. Niederlande	Bewertung von <sup>11</sup> 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 und Fernmetastasen 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 <sup>a</sup>	Zielkriterium	Evidenz- stufe <sup>b</sup>	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
Sihvo 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 <sup>a</sup>	Zielkriterium	Evidenz- stufe <sup>b</sup>	Verzerrungs- potenzial
Vyas 2011	Diagnostische Kohortenstudie Keine Angaben zum Rekrutierungs- und Untersuchungszeitraum. Vereinigtes Königreich	Bewertung der diagnostischen Güte von FDG-PET/CT zum Staging von Adenokarzinomen der Speiseröhre und des gastroösophagealen Übergangs.	Diagnostische Güte	I Ib	hoch
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	I Ib	hoch
<p>a: Studienziel gemäß Formulierung der Autoren  b: entspricht der Evidenzstufe des G-BA  CT = Computertomographie; EUS = Endoskopischer Ultraschall; FDG = <sup>18</sup>Fluorodeoxyglukose; M = Metastase(n); md-CT = Multidetector computed tomography; N = Nodus ((Lymph-) Knoten)); ÖGD = Ösophagogastroduodenoskopie; 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	<b>PET</b> (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 Chemoradiotherapie erhielten) <sup>a</sup>
Choi 2000	<b>PET</b> (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. <b>CT</b> (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. <b>EUS</b> (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 <sup>b</sup> , PET und Histologie verblindet war.	Histologie nach OP
Choi 2010	<b>PET</b> (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. <b>CT</b> (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. <b>EUS</b> (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 <sup>c</sup>	<p><b>PET</b> (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><b>CT</b> (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 &gt; 10 mm (für runde LK) bzw. &gt; 15 mm (ovale LK) wurden als bösartig bewertet. Auswertung durch Radiologen, der gegenüber den Ergebnissen der anderen bildgebenden Verfahren verblindet war.</p> <p><b>EUS</b> (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><b>PET</b> (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><b>CT</b> (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><b>EUS</b> (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	<b>PET/CT</b> (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 <sub>max</sub> ) 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><b>PET</b> (ECAT Exact 951 / 31 oder HCAT HR+ positron camera, Siemens, Deutschland) 5 Minuten nach i.v.-Injektion von 370 MBq (Spannweite 232-403) <sup>11</sup>C-Cholin Herstellung nach Hara und Yuasa<sup>d</sup>) 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	<b>PET</b> (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.  Knochenszintigraphie (E.CAM Gammakamera, Toshiba Corporation, Japan) 3–4 Stunden nach i.v.-Injektion von 740 MBq <sup>99m</sup> Tc-Technetium-Methyldiphosphonat. Keine Angaben zur Befundung.	Histologie nach OP oder klinisch-radiologische Nachbeobachtung, inkl. CT, MRT, Knochenszintigraphie, FDG-PET und spezifische konventionelle Röntgenuntersuchungen über mindestens 6 Monate.  Keine Angaben zur Befundung
Kato 2008	<b>PET</b> (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.  <b>PET/CT</b> (GE Discovery ST8, GE, USA). Zunächst CT-Untersuchung, dann PET-Untersuchung desselben Volumens.  <b>CT</b> -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.	Histologie nach OP.  Keine Angaben zur Befundung
Kneist 2004	<b>PET</b> (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.  <b>CT</b> (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.  Keine Angaben zur Expertise der Untersucher oder zur Konsensusfindung.	Histologie nach OP, klinische Nachbeobachtung  Keine Angaben zur Befundung

(Fortsetzung)

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

Studie	Indextests	Referenztest
Kole 1998	<p><b>PET</b> (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><b>CT</b> (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><b>PET</b> (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><b>CT</b> (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><b>EUS</b> (keine Angaben zum Gerät) durch 1 Endoskopiker basierend auf cTNM-Klassifikation. Lymphknoten &gt; 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><b>PET</b> (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><b>CT</b> (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><b>EUS</b> 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	<b>PET/CT</b> (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	<b>PET</b> (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. <b>CT</b> (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	<b>FDG-PET</b> 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. <b>CT</b> 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. <b>EUS</b> 7,5-13,5 MHz Linear-Array-Schallkopf (Siemens Antaris, Siemens Medical Systems, Erlangen, Deutschland)	Histologie nach OP und / oder Nachbeobachtung über 6 Monate. Keine Angabe zur Befundung
Shenfine 2009	<b>PET</b> (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
Sihvo 2004	<b>PET</b> (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
Vyas 2011	PET/CT (Discovery DST 16, GE Medical Systems, USA) 60 bis 90 Minuten nach i.v.-Injektion von 555–740 MBq FDG. Erhöhte FDG-Aufnahme im Vergleich zum Hintergrund, die nicht durch natürliche Verteilung erklärt werden konnte, wurde als bösartig gewertet. Befundung durch 2 erfahrene Nuklearmediziner, die Zugriff auf Ergebnisse der vorangegangenen Bildgebung hatten.	Histologie nach OP oder FNAB oder radiologische Nachbeobachtung (keine weiteren Angaben) für 6 Monate. Keine Angaben zur Befundung
Yoon 2003	<b>PET</b> (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. <b>CT</b> (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.	Histologie nach OP, klinische Nachbeobachtung (Spanne 3 bis 19 Monate. Keine Angaben zur Befundung
<p>a: Es blieb unklar, ob PET-negative Patienten nachbeobachtet wurden.</p> <p>b: Angabe im Text („unaware of the PET, ES and histologic findings“) wurde als Fehler gedeutet und entsprechend korrigiert.</p> <p>c: Die Angaben stammen aus der Publikation von Lerut 2000 [90].</p> <p>d: siehe [126]</p> <p>AJCC = American Joint Committee on Cancer; CT = Computertomographie; EUS = Endoskopischer Ultraschall; FDG = <sup>18</sup>Fluorodeoxyglukose; FNAB = Feinnadelaspirationsbiopsie; GE = General Electric; i.v. = intravenös; kg = Kilogramm; KG = Körpergewicht; KM = Kontrastmittel; kVP = Peak kilovoltage; M = Metastase(n); mAs = Milliampere Sekunde(n); MBq = Megabecquerel; mCi = Millicurie; MHz = Megahertz; MRT = Magnetresonanztomographie; N = nodus ((Lymph-) Knoten); ÖGD = Ösophagogastroduodenoskopie; 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	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	42	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	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	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	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
Sihvo 2004	55	55 (42 / 13)	60,5 <sup>a</sup> ; k. A.; k. A.	Histologie: Adenokarzinom (20) oder gastroösophagealen Übergang (35). Keine weiteren Angaben.	k. A.
Vyas 2011	104 <sup>b</sup>	104 (91 / 13)	63,7; k. A.; 26,7–85,2	Histologie: Adenokarzinom (104) Lokalisation: oberes Drittel (6), mittleres Drittel der Speiseröhre (18), unteres Drittel der Speiseröhre oder gastroösophagealer Übergang (80).	k. A.
Yoon 2003	81 <sup>c</sup>	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: eigene Berechnung  b: Die Studie rekrutierte ursprünglich 142 Patienten. 38 Patienten wurden jedoch nicht in die Analyse eingeschlossen.  c: Die Studie rekrutierte ursprünglich 136 Patienten. 55 Patienten wurden jedoch nicht in die Analyse eingeschlossen.  k. A. = keine Angabe(n); m = männlich; 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**

15 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 16 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 war es unklar, ob die Befunder bei der Interpretation der Ergebnisbilder der Indextests und / oder Referenztests verblindet waren. Bei den Studien Choi 2000, Heeren 2004, Jager 2001 und Lowe 2005 wurde verblindet ausgewertet. Die Ergebnisse bei Vyas 2011 und 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	Verzerrungspotenzial 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
Sihvo 2004	+	?	+	+	+	+	?	-	+	?	hoch
Vyas 2011	?	?	+	+	+	+	-	+	+	-	hoch
Yoon 2003	+	+	+	-	+	+	-	-	+	?	hoch

+: Kriterium erfüllt; -: Kriterium nicht erfüllt; ?: unklar  
 F1: Beschreibung des Referenztests; F2: Zeitspanne zwischen dem Index- und Referenztest; F3: Interdependenz der Tests; F4: partielle Verifikation; F5: differenzielle Verifikation; F6: Incorporation Bias; F7: Verblindung bei Interpretation der Testergebnisse; F8: Intention-to-diagnose-/Intention-to-screen-Analyse; F9: ergebnisunabhängige Berichterstattung; F10: sonstige Aspekte

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

In 3 Studien konnten keine Indizien für eine Nichtübertragbarkeit gefunden werden. Eine generelle Übertragbarkeit dieser Studien auf den deutschen Versorgungskontext wird daher angenommen. Bei 17 Studien (**Buchmann 2006, Choi 2000, Choi 2010, Flamen 2000, Heeren 2004, Itoh 2008, Jager 2001, Kato 2008, Kato 2005, Kneist 2004, Liberale 2004, Lowe 2005, Noble 2009, Rankin 1998 und Schreurs 2008, Vyas 2011, Yoon 2003**) konnten Indizien gefunden werden, die eine Übertragbarkeit der Ergebnisse für das Primärstaging auf den deutschen Versorgungskontext einschränken könnten:

- Die Bewertung der Indextests in den Studien **Buchmann 2006, Heeren 2004, Jager 2001, Kato 2008, Noble 2009, Rankin 1998 und Schreurs 2008, Vyas 2011** wurde gemeinsam bzw. auch unabhängig voneinander durch mehrere Bewerter gemacht.

- In den Studien **Choi 2000**, **Flamen 2000**, **Jager 2001**, **Kneist 2004** und **Lowe 2005** wurden die Ergebnisse der Indextests ohne Kenntnis vorhergehender klinischer Untersuchungen beurteilt.
- Die Studie **Liberale 2004** schloss ursprünglich 58 Patienten ein. Es wurden aber schließlich nur 24 Patienten ausgewertet. Die Daten der Vierfeldertafel wurden aufgrund der 30-%-Regel nicht dargestellt.
- In der Studie **Yoon 2003** 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.
- In den Studien **Choi 2000**, **Choi 2010**, **Itoh 2008**, **Kato 2005**, **Kato 2008** und **Yoon 2008** wurden ausschließlich asiatische Patienten untersucht. Durch die Stellungnahme der DGN wurde der Hinweis gegeben, dass aufgrund der bekannten tumorbiologischen Unterschiede zwischen asiatischen und europäischen Patienten die Übertragbarkeit der Studienergebnisse einschränkt sein könnte.

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 8 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: Indizien für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Primärstudien beim Primärstaging

Studie	F1	F2	F3	F4	F5	F6	F7	Generelle Einschätzung der Übertragbarkeit
Buchmann 2006	-	-	+	-	-	-	-	Indiz für Nichtübertragbarkeit
Choi 2000	+	-	-	+	-	-	-	Indiz für Nichtübertragbarkeit
Choi 2010	+	-	-	-	-	-	-	Indiz für Nichtübertragbarkeit
Flamen 2000	-	-	-	+	-	-	-	Indiz für Nichtübertragbarkeit
Heeren 2004	-	-	+	-	-	-	-	Indiz für Nichtübertragbarkeit
Itoh 2008	+	-	-	-	-	-	-	Indiz für Nichtübertragbarkeit
Jager 2001	-	-	+	+	-	-	-	Indiz für Nichtübertragbarkeit
Kato 2005	+	-	-	-	-	-	-	Indiz für Nichtübertragbarkeit
Kato 2008	+	-	+	-	-	-	-	Indiz für Nichtübertragbarkeit
Kneist 2004	-	-	-	+	-	-	-	Indiz für Nichtübertragbarkeit
Kole 1998	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Liberale 2004	+	-	-	-	-	-	-	Indiz für Nichtübertragbarkeit
Lowe 2005	-	-	+	+	-	-	-	Indiz für Nichtübertragbarkeit
Noble 2009	-	-	+	-	-	-	-	Indiz für Nichtübertragbarkeit
Rankin 1998	-	-	+	-	-	-	-	Indiz für Nichtübertragbarkeit
Schreurs 2008	-	-	+	-	-	-	-	Indiz für Nichtübertragbarkeit
Shenfine 2009	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Sihvo 2004	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Vyas 2011	-	-	+	-	-	-	-	Indiz für Nichtübertragbarkeit
Yoon 2003	+	-	-	-	-	-	-	Indiz für Nichtübertragbarkeit
+ = konkretes Indiz für eine Nichtübertragbarkeit lag vor - = kein Indiz für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden) F1: Repräsentativität Patientenspektrum; F2: Übertragbarkeit Index- und Referenztest; F3: übertragbarer Entscheidungsprozess bei Testinterpretation; F4: übertragbare Hintergrundinformation; F5: technologische Entwicklung; F6: Untersucherqualifikation; F7: Qualität der Geräte								

#### 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 20 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 Schätzungen wurden auch die Konfidenzintervalle berichtet. Sofern diese nicht in der Studie präsentiert wurden, wurden sie eigenständig nach der Methode von Clopper und Pearson berechnet [127]. 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 <sup>11</sup>C-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, <sup>11</sup>C-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.

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

Studie	n (N)	Indikation	Index- test	Trennwerte (Definition Indextest positiv)	Referenz- test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>
Buch- mann 2006	16 <sup>b</sup> (20)	N-Staging	FDG- PET	FDG-Aufnahme im Vergleich zum umgebenden Gewebe erhöht	Histologie	k. A.	2	0	8	6	20 <sup>a</sup> [2,5; 55,6]	100 <sup>a</sup> [54,1; 100]
	17 <sup>c</sup> (20)	M-Staging	FDG- PET				3	1	1	12	75 <sup>a</sup> [19,4; 99,4]	92,3 <sup>a</sup> [64,0;99,8]
	17 <sup>c</sup> (20)						2	1	2	12	50 <sup>a</sup> [6,8; 93,2]	92,3 <sup>a</sup> [64,0;99,8]
Choi 2000	48 (48) <sup>d</sup>	N-Staging	FDG- PET	FDG-Aufnahme in LK in 2 Schnitten > FDG- Aufnahme im Lungen- parenchym	Histologie	k. A.	26	2	6	14	81 [63,6; 92,8]	87,5 [61,7;98,4]
			CT	Kurze Achse des LK > 10 mm			13	0	19	16	41 [23,7; 59,4]	100 [79,4; 100]
	45 (48)		EUS	LK scharf begrenzt, rund, hypoechogen und > 10 mm	15	4	15	11	50 [31,3; 68,7]	73,3 [44,9;92,2]		
Choi 2010	109 (109)	N-Staging	FDG- PET	Aufnahme im Vergleich zum umgebenden Gewebe erhöht.	Histologie	N1 (+); N0 (-)	27	7	28	47	49 [35,4; 62,9]	87 [75,1;94,6]

(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	Index- test	Trennwerte (Definition Indextest positiv)	Referenz- test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>
Choi 2010		M-Staging	FDG- PET	FDG-Aufnahme im Vergleich zum umgebenden Gewebe erhöht	Histologie	M1 (+); M0 (-)	2	1	3	103	40 [5,3; 85,3]	99 [94,8; 100]
		N-Staging	CT	LK > 8 mm		N1 (+); N0 (-)	19	4	36	50	35 [22,2;48,6]	93 [82,1;97,9]
		M-Staging	CT	k. A.		M1 (+); M0 (-)	1	1	4	103	20 [0,5;71,6]	99 [94,8;100]
		N-Staging	EUS	LK scharf begrenzt, rund, hypoechogen und > 10 mm.		N1 (+); N0 (-)	23	5	32	49	42 [28,7; 55,9]	91 [79,7;96,9]
		N-Staging	Kombi- nation von FDG- PET, CT und EUS <sup>e</sup>	k. A.		N1 (+); N0 (-)	36	11	19	43	65 [51,4; 77,8]	80 [66,5;89,4]
		M-Staging	Kombi- nation von FDG- PET und CT <sup>e</sup>	k. A.		M1 (+); M0 (-)	2	2	3	102	40 [5,3; 85,3]	98 [93,2;99,8]

(Fortsetzung)

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

Studie	n (N)	Indikatio n	Index- test	Trennwerte (Definition Indextest positiv)	Referenz- test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>
Flamen 2000	39 (42) <sup>f, g</sup>	N-Staging	FDG- PET	N1-2 (+); N0 (-)	Histologie nach OP	N1-2 (+); N0 (-)	12	1	15	11	44,4 [25,5; 64,7]	91,7 [61,5;99,8]
			CT			13	1	14	11	48,1 [28,7; 68,1]	91,7 [61,5;99,8]	
			EUS			19	7	3	5	86,4 [65,1; 97,1]	41,7 [15,2;72,3]	
	42 (42) <sup>g</sup>	M1a- Staging	FDG- PET	M+Ly (+); M0 (-)	M+Ly (+); M0 (-)	10	3	3	26	76,9 [46,2; 95,0]	89,7 [72,6;97,8]	
			CT			2	5	11	24	15,4 [1,9; 45,4]	82,8 [64,2;94,2]	
			EUS			5	5	5	22	50,0 [18,7; 81,3]	81,5 [61,9;93,7]	
37 (42) <sup>g</sup>		EUS										

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

(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	Index-test	Trennwerte (Definition Indextest positiv)	Referenz-test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>	
Jager 2001	17 (18)	M-Staging	FDG-PET	M1b (+)	Laparotomie, OP oder klinische Nachbeobachtung	Als positiv definiert: stage M1b	4	0	0	13	100 <sup>a</sup> [39,8; 100]	100 [75,3; 100]	
	16 (18)		<sup>11</sup> C-Cholin-PET				3	0	1	12	75,0 <sup>a</sup> [19,4; 99,4]	100 <sup>a</sup> [73,5; 100]	
	17 (18)	M-Staging	FDG-PET	M1a (+)		Als positiv definiert: stage M1a	4	0	2	11	66,7 [22,3; 95,7]	100 <sup>a</sup> [71,5; 100]	
	16 (18)		<sup>11</sup> C-Cholin-PET				3	0	2	11	60,0 [14,7; 94,7]	100 <sup>a</sup> [71,5; 100]	
	17 (18)	N- und M-Staging	FDG-PET	Jede Metastase (N oder M +) <sup>k</sup> ; N0M0 (-)		Laparotomie, OP oder klinische Nachbeobachtung	N1 oder M1 (+); N0M0 (-)	4	0	8	5	33,3 [9,9; 65,1 <sup>a</sup> ]	100 <sup>a</sup> [47,8; 100 <sup>a</sup> ]
	16 (18)		<sup>11</sup> C-Cholin-PET					3	0	8	5	27,3 <sup>a</sup> [6,0; 61,0]	100 <sup>a</sup> [47,8; 100]
Kato 2005	44 (44)	M-Staging	FDG-PET	Nachweis von Knochenmetastasen (+)	Histologie oder klinisch-radiologische Nachbeobachtung	Nachweis von Knochenmetastasen (+)	12	2	1	29	92 [64,0; 99,8]	94 [78,6; 99,2]	
			Knochen-szintigraphie				10	5	3	26	77 [46,2; 95,0]	84 [66,3; 94,5]	

(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	Index- test	Trennwerte (Definition Indextest positiv)	Referenz- test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>	
Kato 2008	117 (117)	N-Staging	FDG- PET	N1 (+), N0 (-)	Histologie nach OP	k. A.	33	8	27	49	55,0 [41,6; 67,9]	86,0 [74,2; 93,7]	
			CT				29	15	31	42	48,3 [35,2; 61,6]	73,7 [60,3; 84,5]	
	50 (50)		PET/CT				22	4	7	17	75,9 [56,5; 89,7]	81,0 [58,1; 94,6]	
Kneist 2004	81 (81)	M-Staging	FDG- PET	FDG-Aufnahme > umgebendes Gewebe	Histologie nach OP oder Biopsie	k. A.	16	8	4	53	80,0 <sup>a</sup> [56,3; 94,3]	86,9 <sup>a</sup> [75,8; 94,2]	
		M-Staging	CT	Für Fern-LK > 6 mm <sup>n</sup> bzw. Verklumpung			16	8	4	53	80,0 <sup>a</sup> [56,3; 94,3]	86,9 <sup>a</sup> [75,8; 94,2]	
	41 (81) <sup>m</sup>	M-Staging	FDG- PET (AJCC- Stadium IV)	FDG-Aufnahme > umgebendes Gewebe				7	0	10	24	41 [24; 68]	100 [87; 100]
	42 <sup>l</sup> (81) <sup>m</sup>	M-Staging	CT (AJCC- Stadium IV)	Für Fern-LK > 6 mm bzw. Verklumpung				12	6	4	20	75 [50; 92]	77 [58; 89]
	42 <sup>l</sup> (81) <sup>m</sup>	N-Staging	FDG- PET (Thora- kale LK)	FDG-Aufnahme > umgebendes Gewebe				11	1 <sup>a</sup>	14	16 <sup>a</sup>	42 [23; 63]	94 [72; 100]

(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	Index-test	Trennwerte (Definition Indextest positiv)	Referenz-test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>
Kneist 2004	41 <sup>l</sup> (81) <sup>m</sup>	N-Staging	CT (Thorakale LK)	LK > 6 mm <sup>n</sup> bzw. Verklumpung (mehr als 3 kleinere Lymphknoten in einer Gruppe)	Histologie nach OP oder Biopsie		18	5 <sup>a</sup>	6	12 <sup>a</sup>	75 [55; 89]	71 [46; 88]
	39 <sup>l</sup> (81) <sup>m</sup>	N-Staging	FDG-PET (LK kleine Magenkurvatur)	FDG-Aufnahme > umgebendes Gewebe			_h	_h	_h	_h	_h	_h
	39 <sup>h,l</sup> (81) <sup>m</sup>	N-Staging	CT (LK kleine Magenkurvatur)	LK > 6 mm <sup>n</sup> bzw. Verklumpung			_h	_h	_h	_h	_h	_h
	39 <sup>l</sup> (81) <sup>m</sup>	N-Staging	FDG-PET (LK Truncus coeliacus)	FDG-Aufnahme > umgebendes Gewebe			_h	_h	_h	_h	_h	_h
	39 <sup>h,l</sup> (81) <sup>m</sup>	N-Staging	CT (LK Truncus coeliacus)	LK > 6 mm <sup>n</sup> bzw. Verklumpung			_h	_h	_h	_h	_h	_h
	10 <sup>h,l</sup> (81) <sup>m</sup>	N-Staging	FDG-PET (Zervikale LK)	FDG-Aufnahme > umgebendes Gewebe			_h	_h	_h	_h	_h	_h

(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	Index- test	Trennwerte (Definition Indextest positiv)	Referenz- test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>
Kneist 2004	4 <sup>h,1</sup> (81) <sup>m</sup>	N-Staging	CT (Zervikale LK)	LK > 6 mm bzw. Verklumpung	Histologie nach OP oder Biopsie		<sub>h</sub>	<sub>h</sub>	<sub>h</sub>	<sub>h</sub>	<sub>h</sub>	<sub>h</sub>
Kole 1998	21 (26) <sup>o</sup>	N-Staging	FDG- PET	FDG-Aufnahme > Hintergrund	Histologie nach OP	k. A.	12	1	1	7	92 [64,0;99,8]	88 [47,3;99,7]
	22 (26) <sup>o</sup>		CT	Kurze Achse des LK > 10 mm (+)			5	0	8	9	38 [13,9;68,4]	100 [66,4;100]
	21 (26) <sup>o</sup>	M-Staging	FDG- PET	FDG-Aufnahme > Hintergrund			1	3	0	17	100,0 <sup>a</sup> [2,5; 100,0 <sup>a</sup> ]	85,0 <sup>a</sup> [62,1; 96,8 <sup>a</sup> ]
	21 (26) <sup>o</sup>		CT	„Läsionen“ (keine weiteren Angaben)			0	2	1	18	0,0 <sup>a</sup> [0,0; 97,5 <sup>a</sup> ]	90,0 <sup>a</sup> [68,3; 98,8 <sup>a</sup> ]
Liberale 2004	24 (58) <sup>h</sup>	N-Staging	FDG- PET	N1-2 (+); N0 (-)	Histologie nach Biopsie oder OP	N1-2 (+); N0 (-)	<sub>h</sub>	<sub>h</sub>	<sub>h</sub>	<sub>h</sub>	<sub>h</sub>	<sub>h</sub>
			CT+EUS	Bösartiger Befund in min. 1 Untersuchung.			<sub>h</sub>	<sub>h</sub>	<sub>h</sub>	<sub>h</sub>	<sub>h</sub>	<sub>h</sub>

(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	Index- test	Trennwerte (Definition Indextest positiv)	Referenz- test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>
Lowe 2005	59 (69) <sup>p</sup>	N-Staging	FDG- PET	FDG-Aufnahme > Hintergrund	Histologie nach Biopsie oder OP	N1 (+); N0 (-)	36	6	8	9	82 [67;92]	60 [32;84]
	48 (69) <sup>h</sup> <sub>p</sub>	M-Staging	FDG- PET	FDG-Aufnahme > Hintergrund		M1 (+); M0 (-)	_h	_h	_h	_h	_h	_h
	59 (69) <sup>p</sup>	N-Staging	CT	k. A.	N1 (+); N0 (-)	37	5	7	10	84 [69;93]	67 [38;88]	
	48 (69) <sup>h</sup> <sub>p</sub>	M-Staging	CT		M1 (+); M0 (-)	_h	_h	_h	_h	_h	_h	
	59 (69) <sup>p</sup>	N-Staging	EUS	k. A.	N1 (+); N0 (-)	38	5	6	10	86 [73;95]	67 [38;88]	
	48 (69) <sup>h</sup> <sub>p</sub>	M-Staging	EUS		M1 (+); M0 (-)	_h	_h	_h	_h	_h	_h	
Noble 2009	191 (191)	M-Staging	FDG- PET/CT	Unphysiologisch erhöhte FDG- Aufnahme	Histologie nach Biopsie oder OP. Klinische Nachbeo- bachtung (inkl. MRT, Kolonos- kopie oder Laparos- kopie)	M1 (+); M0 (-)	21	10	2	158	91,3 <sup>a</sup> [72,0;98,9]	94,0 <sup>a</sup> [89,3;97,1]

(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	Index- test	Trennwerte (Definition Indextest positiv)	Referenz- test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>
Rankin 1998	18 (25)	N-Staging	FDG- PET	Kurze Achse von peri- ösophagealen LK > 10 mm (+)	Histologie nach OP  (periösop- hageal)	N1 (+); N0 (-)	3	1	5	9	37,5 <sup>a</sup> [8,5;75,5]	90,0 <sup>a</sup> [55,5;99,7]
	19 (25)			Kurze Achse von LK der kleinen Magenkurvatur > 8 mm (+)		N1 (+); N0 (-)  (kleine Magen- kurvatur)	1	1	8	9	11,1 <sup>a</sup> [0,3;48,2]	90,0 <sup>a</sup> [55,5;99,7]
	18 (25)		CT	k. A.	N1 (+); N0 (-)  (periösop- hageal)	4	2	4	8	50,0 <sup>a</sup> [15,7;84,3]	80,0 <sup>a</sup> [44,4;97,59]	
	19 (25)				N1 (+); N0 (-)  (kleine Magen- kurvatur)	5	0	4	10	55,6 <sup>a</sup> [21,2;86,3]	100 <sup>a</sup> [69,2;100]	
Schreurs 2008	55 (55)	N-Staging	CT	Axialer LK Durchmesser > 11 mm für jugulodigastric	Histologie und / oder klinische Anzeichen einer Krankheit während der ersten 6 Monate Follow-up		4	3	1	47	80 [28,4;99,5]	94 [83,5;98,7]
	44 (55)		FDG- PET	Knoten und > 10 mm für alle anderen			4	2	0	38	100 [39,8;100]	95 [83,1;99,4]
	44 (55)		CT + FDG- PET	Halslymphknoten			4	1	0	39	100 [39,8;100]	98 [86,8;99,9]

(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	Index- test	Trennwerte (Definition Indextest positiv)	Referenz- test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>
Shenfine 2009	45 (45)	N-Staging	FDG- PET	Fortgeschrittene Erkrankung ( $T \geq 3$ oder N1), SUV $\geq 5$ (+), frühes Krank- heitsstadium ( $T \leq 2$ und N0), SUV $< 5$ (-)	Histologie nach OP	Fortge- schrittene Erkrankung ( $T \geq 3$ oder N1), SUV $\geq 5$ (+), frühes Krankheits- stadium ( $T \leq 2$ und N0), SUV $< 5$ (-)	22	4	9	10	71 [52,0;85,8]	71,4 [41,9;91,6]
Sihvo 2004	43 (55)	N-Staging	FDG- PET	Unphysiologisch erhöhte FDG- Aufnahme	Histologie nach OP	N1 (+); N0 (-)	9	0	17	17	35 [17,2;55,7]	100 [80,5;100]
	55 (55)	M-Staging	FDG- PET			M1 (+); M0 (-)	10	4	9	32	53 [28,9;75,6]	89 [73,9;96,9]
	42 <sup>q</sup> (55)						7	3	8	24	47 [21,3;73,4]	89 [70,8;97,6]
	43 (55)	N-Staging	CT	k. A.		N1 (+); N0 (-)	11	3	15	14	42 [23,4;63,1]	82 [56,6;96,2]
	55 (55)	M-Staging	CT			M1 (+); M0 (-)	6	1	13	35	32 [12,6;56,6]	97 [85,5;99,9]
	42 <sup>q</sup> (55)						5	1	10	26	33 [11,8;61,6]	96 [81,0;99,9]
	43 (55)	N-Staging	EUS			N1 (+); N0 (-)	22	8	4	9	85 [65,1;95,6]	53 [27,8;77,0]

(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	Index-test	Trennwerte (Definition Indextest positiv)	Referenz-test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>
Sihvo 2004	43 (55)	N-Staging	FDG-PET und CT	k. A., wie Ergebnisse kombiniert wurden	Histologie nach OP	N1 (+); N0 (-)	13	0	13	17	50 [29,9;70,1]	100 [80,5;100]
	55 (55)	M-Staging	FDG-PET und CT			M1 (+); M0 (-)	12	0	7	36	64 <sup>r</sup> [38,4;83,7]	100 [90,3;100]
	43 (55)	N-Staging	EUS and CT	k. A., wie Ergebnisse kombiniert wurden		N1 (+); N0 (-)	22	3	4	14	85 [65,1;95,6]	82 [56,6;96,2]
	55 (55)	M-Staging	EUS und CT			M1 (+); M0 (-)	8	0	11	36	42 [20,3;66,5]	100 [90,3;100]
	43 (55)	N-Staging	FDG-PET, CT und EUS	k. A., wie Ergebnisse kombiniert wurden		N1 (+); N0 (-)	22	0	4	17	85 [65,1;95,6]	100 [80,5;100]
	55 (55)	M-Staging	FDG-PET, CT und EUS			M1 (+); M0 (-)	14	0	5	36	74 [48,8;90,9]	100 [90,3;100]
Vyas 2011	104 (142) <sup>s</sup>	M-Staging	FDG-PET/CT	FDG-Aufnahme > Hintergrund (+)	Histologie nach OP, FNAP oder radiologische Nachbarbeobachtung	M1 (+); M0 (-)	4	15	3	82	57,1 [18,4;90,1]	84,5 [75,8;91,1]

(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	Index- test	Trennwerte (Definition Indextest positiv)	Referenz- test	Trennwerte	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %-KI] <sup>a</sup>
Yoon 2003	81 (136) <sup>f</sup>	N-Staging	FDG- PET	FDG-Aufnahme > Hintergrund (+)	Histologie nach OP	N1 (+); N0Mx (-)	25	13	14	29	64 [47,2;78,8]	69 [52,9;82,4]
		M-Staging	FDG- PET	FDG-Aufnahme > Hintergrund (+)		M1 (+); M0 (-)	3	1	4	73	43 [9,9;81,6]	99 [92,7;100]
		N-Staging	CT	LK > 10 mm (+)		N1 (+); N0Mx (-)	12	6	27	36	31 [17,0;47,6]	86 [71,5;94,6]
		M-Staging	CT	k. A.		M1 (+); M0 (-)	1	4	6	70	14 [0,4;57,9]	95 [86,7;98,5]
<p>a Werte wurden selbst berechnet</p> <p>b: Die Daten stammen aus Tabelle 2 der Publikation. 3 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</p> <p>c: Nach Pathologie, falsch- positiver Befund (Metastase im Lungenapex) und richtig- positiver Befund (bronchohilar) bei einem Patienten. Hier sind die 2 Möglichkeiten dargestellt, die sich ergeben, wenn dieser Patient nur in jeweils einer der Kategorien gezählt wird: In der oberen Reihe als RP, unten als FN gewertet. Es ergeben sich zum Paper abweichende Sensitivitäten. In Tabelle 3 der Publikation wird der Schätzwert der Sensitivität mit 60 % angegeben. Möglicherweise wurde der Grenzfall jeweils zur Hälfte als RP und FN gewertet.</p> <p>d: In die Studie wurden konsekutiv 61 Patienten eingeschlossen, von denen 5 die OP ablehnten und 8 nicht-operationsfähig waren.</p> <p>e: Keine Angabe, wie diese Tests kombiniert wurden.</p> <p>f : 74 Patienten eingeschlossen. Für das N-Staging lagen Daten für 39 Patienten vor. Die Daten für das M-Staging wurden nicht verwendet, weil kein adäquater Referenztest verwendet wurde</p> <p>g : Die Daten wurden aus der Publikation Lerut 2000 (zweite Publikation zur Studie) extrahiert.</p> <p>h: Die Ergebnisse wurden nicht dargestellt, weil der Anteil fehlender Werte größer als 30 % war.</p> <p>i: Von den 72 in der Studie eingeschlossenen Patienten erhielten 52 EUS.</p> <p>j: Im Text der Publikation wird der Teiler (RN+FP) zur Berechnung der Spezifität mit 47 angegeben, was zu einer falschen Anzahl ausgewerteter Patienten führen würde. Tabelle 2 gibt die Zahl mit 37 an, was zur richtigen Anzahl ausgewerteter Patienten führt. Hier wurden diese Daten verwendet. Somit hatten 4 Patienten einen falsch positiven Befund.</p>												

(Fortsetzung)

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

k: Außerdem Angaben zur Sensitivität für N1, M1a und M1b. Jedoch wurden die entsprechenden Schätzwerte der Spezifität nicht berichtet.

l: Diese Zwischenauswertung von Kneist 2004 untersuchte anhand einer Teilmenge von 58 Probanden das M-Staging.

m: 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.

n: Für Lymphknoten nahe der Vena azygos und unterhalb der Luftröhren-gabelung > 10 mm.

o: Die Studie schloss 26 Patienten ein. 4 Patienten wurden nicht operiert und daher von der Analyse ausgeschlossen. Ein weiterer Patient wurde von der Bewertung der PET ausgeschlossen.

p: Es wurden zunächst 75 Patienten eingeschlossen. 6 Patienten wurden wieder ausgeschlossen wegen anderer Tumoren.

q: Räsänen et al. 2003 war eine Zwischenauswertung von Sihvo 2004 [107]. Räsänen enthält eine Teilmenge der Patienten von Sihvo und untersucht das N-Staging.

r: Rechnerisch ergibt sich hier eine leicht abweichende Sensitivität von 63,2.

s: 142 bestätigte Fälle mit Ösophaguskarzinom, davon 121 ohne entfernte Metastasen, davon 104 Adenokarzinome.

t: 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)

AJCC = American Joint Committee on Cancer; CT = Computertomographie; EUS = endoskopischer Ultraschall; FDG = <sup>18</sup>Fluorodeoxyglukose; 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); ÖGD = Ösophagogastroduodenoskopie; PET = Positronenemissionstomographie; RN = richtig negativ; RP = richtig positiv; SUV = standardised 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 Forestplot 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.

Ösophaguskarzinom

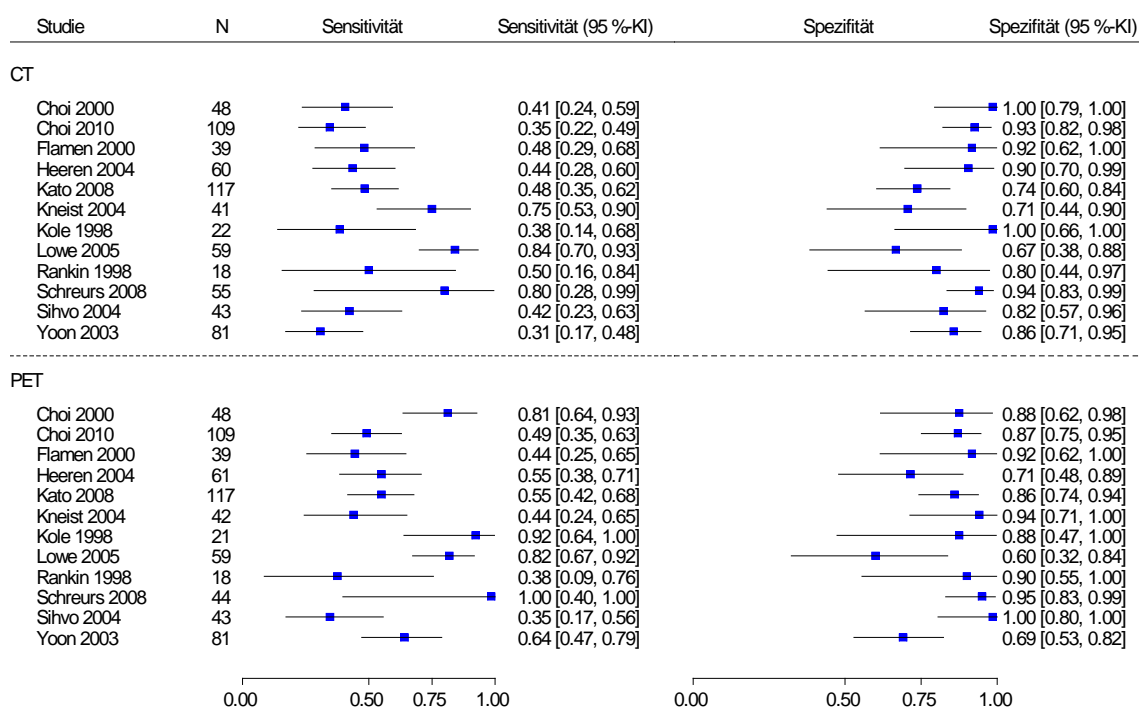


Abbildung 2: Forestplots der univariaten Ergebnisse für Sensitivität und Spezifität beim N-Staging bei Ösophaguskarzinomen

95 %-KIs wurden berechnet nach Leemis und Travis 1996 [62].

In Abbildung 3 und Abbildung 4 werden zum einen die direkten Vergleiche der PET und CT und zum anderen die 95 %-Konfidenzregionen und Punktschätzungen für die Sensitivität und Spezifität aus der bivariaten Meta-Analyse beim N-Staging dargestellt. Für den Vergleich der PET mit der CT wurde eine bivariate Meta-Analyse mit den oben beschriebenen 12 Studien durchgeführt. In Tabelle 21 werden die Schätzungen für die Sensitivität und Spezifität präsentiert.

Der Unterschied in der diagnostischen Güte zwischen den beiden Technologien wurde mithilfe eines Likelihood-Ratio-Tests untersucht. Dieser ergab einen p-Wert von 0,370 und konnte damit keinen Unterschied zwischen den Technologien feststellen.

Die Studien Kole 1998, Schreurs 2008 und Sihvo 2004 wurden anhand der 95%-Prädiktionsregion für die PET, Kole 1998 und Choi 2000 anhand der 95%-Prädiktionsregion



für die CT als Ausreißer identifiziert. Die Sensitivitätsanalysen lieferten keine Ergebnisse, die der Hauptanalyse entgegenstanden.

### Sensitivitätsanalysen

Es wurden Sensitivitätsanalysen bezüglich des Verzerrungspotenzials durchgeführt. Hierzu wurde eine bivariate Meta-Analyse mit den vier niedrig verzerrten Studien durchgeführt. Es ergaben sich keine präzisen Schätzungen, so dass die Ergebnisse der Hauptanalyse weder bestätigt noch in Frage gestellt werden.

Es ist zu bemerken, dass von den vier Studien mit niedrigem Verzerrungspotenzial zwei als Ausreißer identifiziert wurden.

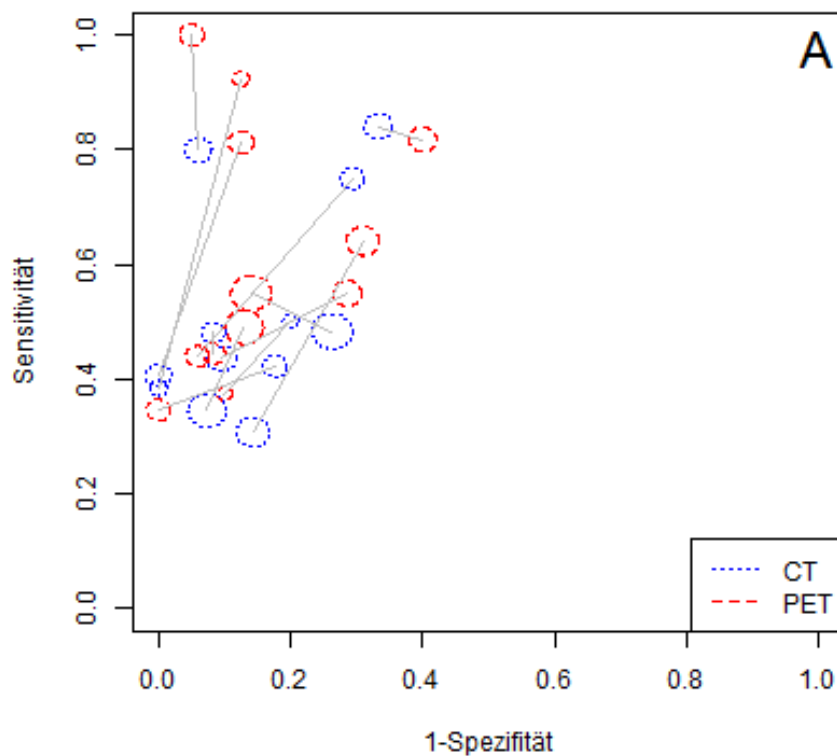


Abbildung 3: Darstellung der direkten Vergleiche der PET und CT beim N-Staging bei Ösophaguskarzinomen

Durchmesser der Kreise sind proportional zur jeweiligen Studiengröße. Die Sensitivität und Spezifität sind ohne Korrekturfaktor berechnet worden.

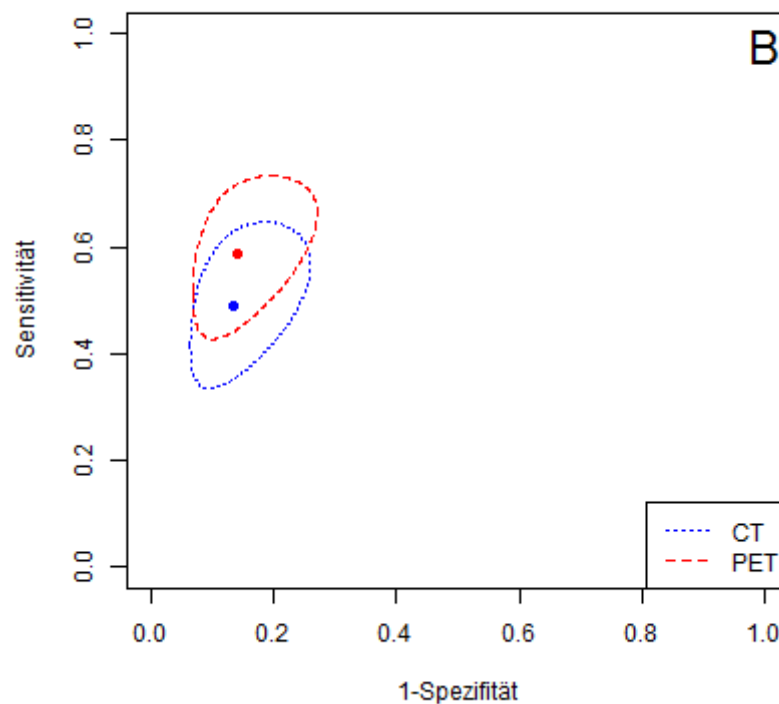


Abbildung 4: 95 %-Konfidenzregionen und Punktschätzungen für Sensitivität und Spezifität aus der bivariaten Meta-Analyse beim N-Staging bei Ösophaguskarzinomen

Tabelle 21: Schätzung der Sensitivität und Spezifität aus der bivariaten Meta-Analyse beim N-Staging

Test	Parameter	Schätzwert (%)	95 %-KI in %
CT	Sensitivität	49	[38;60]
	Spezifität	87	[79;92]
PET	Sensitivität	59	[48;69]
	Spezifität	86	[78;91]

### Direkte Vergleiche zwischen PET und CT für das M-Staging

In Abbildung 5 wird der Forestplot der Sensitivität und der Spezifität für alle Studien, die einen direkten Vergleich zwischen PET und CT für das M-Staging durchführten dargestellt.

Ösophaguskarzinom

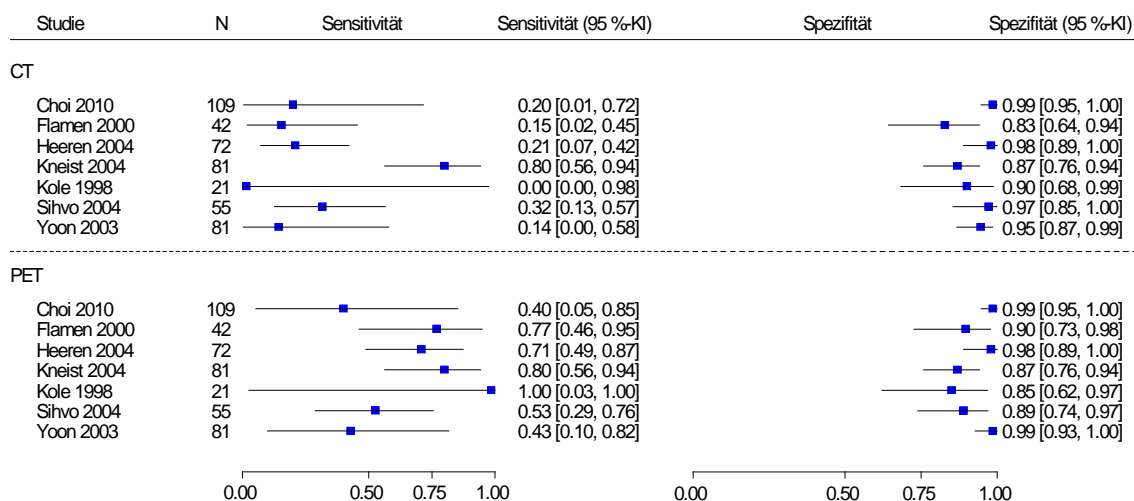


Abbildung 5: Forestplots der univariaten Ergebnisse für Sensitivität und Spezifität beim M-Staging bei Ösophaguskarzinomen

95 %-KIs wurden berechnet nach Leemis und Travis 1996 [62]

Für den Vergleich der PET mit CT wurde eine bivariate Meta-Analyse mit den oben beschriebenen 7 Studien durchgeführt. Die Ergebnisse dieser Meta-Analyse werden nicht dargestellt, Beim M-Staging ergab die bivariate Meta-Analyse keine präzisen Schätzungen. Auch in Sensitivitätsanalysen ergaben sich keine präzisen Schätzungen.

**Subgruppenmerkmale und andere Effektmodifikatoren**

In der Studie von Kneist 2004 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. Es lag nur eine niedrig verzerrte Studie vor.

**Ergänzend betrachtete Ergebnisse zu Änderungen im klinischen Management**

In der Studie **Buchmann 2006** wurde eine Änderung im klinischen Management basiert auf dem Einsatz der PET (siehe Tabelle 22). 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 **Kneist 2004** 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 22).

Auch in der Studie **Noble 2009** 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 22).

In der Studie **Vyas 2011** kam es bei 13 von 104 eingeschlossenen Patienten zu klinischen Managementänderungen. Die genaueren Beschreibungen der Änderungen sind Tabelle 22 zu entnehmen.

Tabelle 22: 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 Management-änderung: 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.
Vyas 2011	FDG-PET	104	13 (12,5)	Upstaging zu M1 (nicht durch CT diagnostiziert): 1 Patient mit Lebermetastasen (durch MRT, US und Biopsie bestätigt) erhielt palliative Chemotherapie. 1 Patient mit Lungenmetastasen (durch Thorakotomie und Biopsie bestätigt) erhielt palliative Chemotherapie. 2 weitere Patienten wurden ebenfalls auf eine palliative Behandlung umgestellt (keine weiteren Angaben). 1 Patient mit mediastinalem Befall erhielt eine Ivor-Lewis-Ösophagektomie (zusätzlich zur neoadjuvanten Chemotherapie). 3 Patienten erhielten durch Einsatz der PET eine verzögerte Behandlung. 5 Patienten hatten kolorektale Karzinome (2 davon Polypen) und wurden zusätzlich zur kurativen Therapie operiert. 1 Patient mit synchronem Pankreaskarzinom wurde operativ behandelt. 1 Patient mit synchronem Lungenkarzinom wurde operativ behandelt. 1 Patient mit lokal fortgeschrittenem synchronem Prostatakarzinom wurde auf palliative Therapie umgestellt.
CT = Computertomographie; FDG = <sup>18</sup> 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 6 Arbeiten (**Cheze-Le Rest 2008**, **Choi 2004**, **Chung 2008**, **Omloo 2008** und **Shenfine 2009**, **Gillies 2012a**). In 4 der 6 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**, 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 **Choi 2004** 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** 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 Tl-SPECT 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 **Gillies 2012a** untersuchte 121 Patienten mit potenziell heilbaren Adenokarzinomen des Ösophagus vor einer neoadjuvanten Chemotherapie und OP. Ziel der Arbeit war die Bewertung der prognostischen Güte der FDG-PET/CT zur Vorhersage des Überlebens von Patienten nach neoadjuvanter Chemotherapie und Operation. In einem multifaktoriellen Modell, welches pathologische und radiologische Faktoren einschloss, sollten Parameter bestimmt werden, die das Gesamtüberleben (Zeitpunkt der PET/CT-Aufnahme bis zum tumorbedingten Versterben) vorhersagen können.

Die Studie **Omloo 2008** 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** 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 23: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudien beim Primärstaging

Studie	Studiendesign	Studienziel <sup>a</sup>	Zielkriterium	Evidenzstufe <sup>b</sup>	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 <sup>c</sup>	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 <sup>c</sup>	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 <sup>c</sup>	hoch
Gillies 2012a	Prognostische Kohortenstudie Rekrutierung der Patienten zwischen April 2006 und April 2009. Vereinigtes Königreich	Bewertung der prognostischen Güte von FDG-PET/CT zur Vorhersage des Überlebens von Patienten mit Ösophaguskarzinom nach neoadjuvanter Chemotherapie und Operation.	Prognostische Güte von FDG-PET zur Vorhersage des krankheitsfreien Überlebens und des Gesamtüberlebens.	II <sup>c</sup>	hoch

(Fortsetzung)



Tabelle 23: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudien beim Primärstaging (Fortsetzung)

Studie	Studiendesign	Studienziel <sup>a</sup>	Zielkriterium	Evidenzstufe <sup>b</sup>	Verzerrungs- potenzial
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 <sup>c</sup>	hoch
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 Adenokarzinomen der Speiseröhre ist.	Prognostische Güte von FDG-PET zur Vorhersage des krankheitsfreien Überlebens und des Gesamtüberlebens.	II <sup>c</sup>	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 24: Index- und Referenztest der eingeschlossenen Prognosestudien beim Primärstaging

Studie	Indextest	Referenztest
Cheze-Le Rest 2008	<b>PET</b> (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	<b>PET</b> (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	<b>PET</b> (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)
Gillies 2012a	<b>PET/CT</b> (GE Discovery STE 16, General Electric Co.) 60 Minuten nach i.v.-Injektion von 5 MBq/kg KG (min. 300 MBq) FDG. Bestimmung von $SUV_{max}$ , FDG-aufnehmende Lymphknoten sowie Länge des FDG-aviden Teils des Primärtumors. Befundung durch 2 Radiologen.	Überleben vom Zeitpunkt der PET/CT-Aufnahme bis zum krebisbedingten Versterben (durch Krankenakten des Krankenhauses oder des behandelnden Arztes). Rezidive wurden histologisch oder nur durch Bildgebung (keine weiteren Angaben) befundet, wenn nach Einschätzung einer multidisziplinären Tumorkonferenz ein Rezidiv vorlag.
Omloo 2008	<b>PET</b> (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 $130 \pm 810$ MBq, je nach Körpergewicht) durchgeführt. <b>EUS</b> Radialscanner (GF-UM130 oder GF-UM160, $5 \pm 20$ MHz, Olympus Medical Systems, Tokio, Japan). EUS wurde bei Patienten in einer linken Dekubitusposition und unter Sedierung mit $2,5 \pm 10$ mg Midazolam intravenös durchgeführt.	Gesamtüberleben, Follow-up mindestens 25 Monate

(Fortsetzung)

Tabelle 24: Index- und Referenztest der eingeschlossenen Prognosestudien beim Primärstaging (Fortsetzung)

Studie	Indextest	Referenztest
Shenfine 2009	<b>PET</b> (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 = <sup>18</sup>Fluordeoxyglukose; GE = General Electric; i.v. = intravenös; kg = Kilogramm; KG = Körpergewicht; MAV = metabolisches Aktivitätsvolumen; MBq = Mega Becquerel; SUV = standardized uptake value; TI =Thallium

Tabelle 25: 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 <sup>a</sup>	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 <sup>b</sup>	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)

(Fortsetzung)

Tabelle 25: Eingeschlossene Patienten der eingeschlossenen Prognosestudien beim Primärstaging (Fortsetzung)

Studie	Patienten [n]	Geschlecht [n (m/w)]	Alter (MW, Median, Spannweite)	Diagnose	Behandlung
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 Fern- metastasierung (23).
Gillies 2012a	121	121 (97 / 24)	63; k. A.; k. A.	Patienten mit potenziell heilbaren Adeno- karzinomen. Lokalisation: mittlere (4), untere Speiseröhre (69), gastroösophagealer Übergang (48)	Keine Angaben zur Vorbehandlung. Nach Studienbeginn: NAC
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.
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. b: 89 Patienten eingeschlossen. Ein Patient mit unvollständiger Resektion 9 Patienten, die eine OP ablehnten, und 10 nicht OP-fähige Patienten wurden ausgeschlossen. 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 **Cheze-Le Rest 2008** 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** 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** wurde die Behandlung, die möglicherweise einen Einfluss haben könnte, nicht berücksichtigt, sodass ein hohes Verzerrungspotenzial besteht. Die Anzahl der in der multifaktoriellen Analyse eingeschlossenen Faktoren (7) scheint im Verhältnis zu den Ereignissen für die multifaktorielle Analyse zu hoch zu sein, sodass hier das Problem der Überparametrisierung vorlag. Die eingeschlossenen Patienten waren darüber hinaus zu Studienbeginn in unterschiedlichen Krankheitsstadien. Es gab keine separaten Auswertungen nach Krankheitsstadien.

In der Studie **Gillies 2012a** könnten die unterschiedlichen Krankheitsstadien der Patienten zu Studienbeginn und potenziell relevante Faktoren (Art der Chemotherapie, Tumordifferenzierung und -ausbreitung), die in der Analyse nicht berücksichtigt wurden, die Aussagekraft der multivariaten Analyse mindern. Darüber hinaus wird in der Studie der Patientenfluss unvollständig beschrieben.

In der Studie **Omloo 2008** 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 **Shenfine 2009** 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 26: 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
Gillies 2012a	+	-	+	+	-	+	+	-	?	?	+	+	?	-	-	+	+	hoch
Omloo 2008	+	-	+	+	+	+	+	-	+	?	+	+	?	-	?	+	+	hoch
Shenfine 2009	+	-	+	+	+	+	?	?	+	?	+	+	?	-	?	+	-	hoch

+: Kriterium erfüllt; -: Kriterium nicht erfüllt; ?: unklar  
F1: Beschreibung Selektionskriterien; F2: Vergleichbarkeit der Krankheitsphase der Patientinnen; F3: Repräsentativität Studienpopulation; F4: Nachbeobachtungszeit adäquat; F5: Patientenfluss nachvollziehbar; F6: Rate der Teilnehmer an Nachbeobachtung adäquat; F7: Definition prognostischer Faktoren; F8: Messung prognostischer Faktoren adäquat; F9: prognostische Faktoren für ausreichenden Teil der Patientinnen vorhanden; F10: Teilnehmer bzgl. prognostischer Information verblindet; F11: Definition Endpunkte; F12: Messung Endpunkte adäquat; F13: Festlegung Endpunkte verblindet bezüglich prognostischer Faktoren; F14: Berücksichtigung von Störgrößen; F15: Behandlungen während Nachbeobachtungszeit standardisiert oder randomisiert; F16: ergebnisunabhängige Berichterstattung; F17: keine sonstigen Aspekte

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

In 3 Studien konnten keine Indizien für eine Nichtübertragbarkeit gefunden werden. Eine generelle Übertragbarkeit dieser Studien auf den deutschen Versorgungskontext wird angenommen. Bei 3 von 6 eingeschlossenen Prognosestudien konnten Indizien gefunden werden, die eine Übertragbarkeit der Ergebnisse der Prognosestudien auf den deutschen Versorgungskontext einschränken:

- In den Studien **Choi 2004 und Chung 2008** wurden ausschließlich asiatische Patienten untersucht. Durch die Stellungnahme der DGN wurde der Hinweis gegeben, dass aufgrund der bekannten tumorbiologischen Unterschiede zwischen asiatischen und europäischen Patienten die Übertragbarkeit der Studienergebnisse einschränkt sein könnte.
- Die Bewertung der Ergebnisse der Technologien erfolgte in den Studien **Cheze-Le Rest 2008, Choi 2004 und Chung 2008** gemeinsam bzw. auch unabhängig voneinander durch mehrere erfahrene Bewerter.
- In den Studien **Cheze-Le Rest 2008** 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 Indizien dafür, dass die Ergebnisse nicht auf die Anwendungssituation in Deutschland gelten.

Tabelle 27: Indizien 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	-	-	+	+	-	-	-	Indiz für eine Nichtübertragbarkeit
Choi 2004	+	-	+	+	-	-	-	Indiz für eine Nichtübertragbarkeit
Chung 2008	+	-	+	-	-	-	-	Indiz für eine Nichtübertragbarkeit
Gillies 2012a	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Omloo 2008	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Shenfine 2009	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
+: konkretes Indiz für eine Nichtübertragbarkeit lag vor -: kein Indiz für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden) F1: Repräsentativität Patientenspektrum; F2: Übertragbarkeit Index- und Referenztest; F3: übertragbarer Entscheidungsprozess bei Testinterpretation; F4: übertragbare Hintergrundinformation; F5: technologische Entwicklung; F6: Untersucherqualifikation; F7: Qualität der Geräte								



#### **5.3.3.4 Ergebnisse zur prognostischen Güte der eingeschlossenen Primärstudien beim Primärstaging**

In Tabelle 28 werden die Ergebnisse der einfaktoriellen und multifaktoriellen Analyse der eingeschlossenen Prognosestudien dargestellt.

Von 6 eingeschlossenen Prognosestudien hatten 5 Studien das Ziel, das Gesamtüberleben zu untersuchen. 3 dieser 5 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 6 eingeschlossenen Studien waren 3 PET-relevante Merkmale im multifaktoriellen Modell statistisch signifikant.

Tabelle 28: Ergebnisse der Primärstudien zur prognostischen Güte beim Primärstaging

Studie	Zielgröße	n (N)	Prognostisches Merkmal	Cut-off zum prognostischen Merkmal	Patienten mit prog. Merkmal / alle Patienten	Ergebnisse aus einfaktorierter Analyse		Ergebnisse aus multifaktorieller Analyse	
						p-Wert	HR	p-Wert	HR
Chez-Le Rest 2008	Gesamtüberleben	47 <sup>a</sup> (52)	SUVmax im Primärtumor	> 9 <sup>b</sup>	k. A.	0,02	k. A.	0,01	3,16
			Lymphknoten mit erhöhter FDG-Aufnahme	≥ 1	k. A.	0,03	k. A.	0,05	2,43
				≥ 2	k. A.	0,004	k. A.	0,01	3,16
				Anzahl abnormaler FDG foci	> 1	25/47	0,006	k. A.	k. A.
Choi 2004	Krankheitsfreies Überleben	69 <sup>c</sup> (89)	Anzahl Lymphknoten mit erhöhter FDG-Aufnahme	0	26/69	< 0,001 <sup>d</sup>	–	< 0,001	1,87
				1	16/69				
				2	16/69				
				≥ 3	11/69				
			Tumurlänge in der PET	< 3 cm	14/69	0,018 <sup>d</sup>	–	k. A.	k. A.
				3 bis 5 cm	26/69				
				≥ 5 cm	29/69				
			PET Stadium	N0 M0	26/69,	< 0,001 <sup>d</sup>	–	k. A.	k. A.
				N1 M0	28/69				
				M1	15/69				

(Fortsetzung)

Tabelle 28: Ergebnisse der Primärstudien zur prognostischen Güte beim Primärstaging (Fortsetzung)

Studie	Zielgröße	n (N)	Prognostisches Merkmal	Cut-off zum prognostischen Merkmal	Patienten mit prog. Merkmal / alle Patienten	Ergebnisse aus einfaktorieller Analyse		Ergebnisse aus multifaktorieller Analyse	
						p-Wert	HR	p-Wert	HR
Choi 2004	Gesamtüberleben	69 <sup>c</sup> (89)	Anzahl Lymphknoten mit erhöhter FDG-Aufnahme	0	26/69	< 0,001 <sup>d</sup>	–	0,024	1,71
				1	16/69				
				2	16/69				
				≥ 3	11/69				
			Tumurlänge in der PET	< 3 cm	14/69	0,001 <sup>d</sup>	–	0,010	2,74
				3 bis 5 cm	26/69				
				≥ 5 cm	29/69				
			Maximum SUV of primary tumor	< 6,3	14/69	0,035 <sup>d</sup>	–	k. A.	k. A.
				6,3 bis 13,7	27/69				
				≥ 13,7	28/69				
			PET-Stadium	N0 M0	26/69	0,001 <sup>d</sup>	–	k. A.	k. A.
				N1 M0	28/69				
M1	15/69								
Chung 2008	Krankheitsfreies Überleben	100 (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 <sup>e</sup>
			Visualisierung durch TI-SPECT	Visualisierung	85/100	0,01	k. A.	0,72	4,88 <sup>e</sup>
			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 <sup>e</sup>

(Fortsetzung)

Tabelle 28: Ergebnisse der Primärstudien zur prognostischen Güte beim Primärstaging (Fortsetzung)

Studie	Zielgröße	n (N)	Prognostisches Merkmal	Cut-off zum prognostischen Merkmal	Patienten mit prog. Merkmal / alle Patienten	Ergebnisse aus einfaktorieller Analyse		Ergebnisse aus multifaktorieller Analyse	
						p-Wert	HR	p-Wert	HR
Gillies 2012a	Krankheitsfreies Überleben	95 <sup>f</sup> (121)	FDG-Aufnahme in Lymphknoten	ja / nein	28	k. A.	k. A.	0,004	2,97
			Länge des FDG-aviden Teils des Primärtumors	≥ 5 cm; < 5 cm	k. A.	k. A.	k. A.	0,032	0,46
			SUV <sub>max</sub>	≥ 8,5; < 8,5	k. A.	k. A.	k. A.	0,362	0,73
	Gesamtüberleben	95 <sup>f</sup> (121)	FDG-Aufnahme in Lymphknoten	ja / nein	28	k. A.	k. A.	< 0,001	4,75
			Länge der FDG-Aufnahme im Primärtumor	≥ 5 cm; < 5 cm	k. A.	k. A.	k. A.	0,117	0,53
			SUV <sub>max</sub>	≥ 8,5; < 8,5	k. A.	k. A.	k. A.	0,230	0,63
Omloo 2008	Gesamtüberleben	125 (125)	SUV niedrig versus hoch, cut-off: Median	k.A.	k.A.	0,04	Unklar	0,6	Unklar <sup>g</sup>
Shenfine 2009	Gesamtüberleben	45 (45)	SUV <sub>max</sub>	≤ 5	19/45	0,006	k. A.	Unklar <sup>g</sup>	Unklar <sup>g</sup>
				> 5	26/45	0,007	k. A.	Unklar <sup>g</sup>	Unklar <sup>g</sup>
			FDG PET qualitatives Ergebnis	Negative Lymphknoten	37/45	0,02	k. A.	k. A.	k. A.
				Positive Lymphknoten	8/45				
	Krankheitsfreies Überleben	45 (45)	SUV <sub>max</sub>	≤ 5	19/45	0,013	k. A.	Unklar <sup>g</sup>	Unklar <sup>g</sup>
> 5				26/45	0,015	k. A.	Unklar <sup>g</sup>	Unklar <sup>g</sup>	

(Fortsetzung)

Tabelle 28: Ergebnisse der Primärstudien zur prognostischen Güte beim Primärstaging (Fortsetzung)

- a: 52 Patienten eingeschlossen. 5 Patienten, die innerhalb von 30 Tagen nach OP nicht krebsbedingt verstarben, wurden nicht in die Analyse eingeschlossen.
- b: In Tabelle 3 “< 9”, in Tabelle 5 “> 9”.
- c: 89 Patienten eingeschlossen. 1 Patient mit unvollständiger Resektion 9 Patienten, die eine OP ablehnten, und 10 nicht OP-fähige Patienten wurden ausgeschlossen.
- d: Log-Rank-Test, daher keine Angaben zu den Hazard Ratios
- e: Korrespondenz mit den Autoren der Studie
- f: Aus der Publikation wird nicht klar, ob alle 201 oder nur 103 Patienten in die finalen Analysen eingeflossen sind. In Tabelle 3 der Publikation werden nur 95 Patienten angegeben.
- g: 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; MAV = metabolisches Aktivitätsvolumen; SUV = standardized uptake value

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

Ein direkter Vergleich zwischen FDG-PET und  $^{201}\text{Tl}$  SPECT wurde in der Studie **Chung 2008** durchgeführt. Alle PET- sowie auch die SPECT-relevanten Merkmale waren nach multiplen Testen nicht mehr statistisch signifikant. Die Autoren der Studien wiesen selbst darauf hin, dass die Zunahme der  $^{201}\text{Tl}$  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** [67], **Rebollo Aquirre 2009** [76] und **Westerterp 2005** [77]).

#### 5.4.1.1 Studiendesigns und Studienpopulationen der eingeschlossenen Evidenzsynthesen beim Restaging

Die Evidenzsynthese **Ngamruengphong 2010** 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 **Rebollo Aquirre 2009** 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 Abschlussbericht nicht als Referenztest akzeptiert.

Die Evidenzsynthese **Westerterp 2005** 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 29 zusammengefasst.

Tabelle 29: 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
<b>Ngamruengphong 2010</b> PET und PET/CT bei Patienten mit Ösophaguskarzinom	<u>Quellen:</u> MEDLINE, EMBASE, Cochrane Library, <u>Suchstrategie:</u> selbst entwickelte Suchstrategie, <u>Zeitraum:</u> Jan 1980 bis Feb 2008	<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.	Nach den Kriterien des QUADAS-Instruments	-
<b>Rebollo Aquirre 2009</b> PET bzw. PET/CT bei Patienten mit Ösophaguskarzinomen	<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	<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.	Nach den Kriterien des QUADAS-Instruments	-

(Fortsetzung)



Tabelle 29: Fragestellung und Methodik der Evidenzsynthesen zur diagnostischen und prognostischen Güte der PET bzw. PET/CT beim Restaging (Fortsetzung)

<b>Evidenzsynthese / Fragestellung</b>	<b>Literaturrecherche</b>	<b>Ein- / Ausschlusskriterien, Studienselektion</b>	<b>Studienbewertung</b>	<b>Informations- synthese</b>
<b>Westerterp 2005</b> 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 Indizien für einen Selektionsbias. Da die Evidenzsynthese von Rebollo Aquirre 2009 ansonsten alle Kriterien des Qualitätsindex erfüllte, wurde sie in die Bewertung des vorliegenden Berichts eingeschlossen.

Tabelle 30: Ergebnisse der Qualitätsbewertung der eingeschlossenen Evidenzsynthesen nach Oxman und Guyatt [49,50]

Evidenzsynthesen	OG1	OG2	OG3	OG4	OG5	OG6	OG7	OG8	OG9	Total
Ngamruengphong 2010	+	+	+	+	+	+	+	+	+	7
Westerterp 2005	+	+	+	+	+	+	+	+	+	7
Rebollo Aquirre 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 [68,74,78-81,83-86,128].

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 Studien, die im vorliegenden Bericht eingeschlossen wurden, erzielten bei Ngamruengphong 2010 10 bis 13 Punkte [68,74,78-81,83,85,86,128]. 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 [73,78,82,83,85,86]. 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 [68,79,81] eingeschlossen. Die Studie von Weber 2001 war nur in Westerterp 2005 enthalten. Drei der untersuchten Studien [68,79,84] erhielten 6 von 7 möglichen Punkten, die Studie von Kroep 2003 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 31 werden die Ergebnisse zur diagnostischen Güte auf Basis der extrahierten Evidenzsynthesen dargestellt. Bezüglich der Studie von Song 2005 [83] 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 Schätzwerten für Sensitivität und Spezifität werden in Tabelle 31 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 [127] 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 untersuchten den Einsatz der PET während einer neoadjuvanten Chemotherapie. Hier zeigte sich für die PET eine Spannweite der Punktschätzungen der Sensitivitäten von 80 % (bei einer zugehörigen Spezifität von 78 % bzw. 100 %) bis 88 % (zugehörige Spezifität 75 %). Für die Punktschätzungen der Spezifitäten ergab sich eine Spannweite von 75 % (zugehörige Sensitivität 88 %) bis 100 % (zugehörige Sensitivität 80 %). Für alle Punktschätzungen waren weite Konfidenzintervalle zu beobachten. Die Studien von Gillham 2006, Westerterp 2006 und Wieder 2004 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 Punktschätzungen der Sensitivitäten von 27 % (zugehörige Spezifität 95 %) bis 100 % (zugehörige Spezifität 55 % bzw. 100 %). Für die Punktschätzungen 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 31: Ergebnisse der Studien aus den Evidenzsynthesen zum Restaging

Studie	Anzahl Patienten	Index-test	Referenztest	Population	Cut-Off	Indikation	RP	FN	FP	RN	Sensitivität % (95 %-KI)	Spezifität % (95 %-KI)
Brücher 2001	24 <sup>a</sup>	PET	Histologie, Tumoran-sprechen nach Mandard mit Modifikation	AJCC II und III	Reduction des FDG-PET uptake 52 %	Restaging nach Abschluss der NAC / NACR	13	0	5	6	100 (75,3 bis 100) <sup>b</sup>	55 (23,4 bis 83,3) <sup>b</sup>
Cerfolio 2005	48 <sup>c</sup>	PET/CT	Histologie, Tumoran-sprechen wenn TONOMO	AJCC II und IV	SUV $\geq$ 2,5	Restaging nach Abschluss der NACR	13	2	4	29	86,7 (59,5 bis 98,3) <sup>b</sup>	87,9 (71,8 bis 96,6) <sup>b</sup>
	48 <sup>c</sup>	EUS		AJCC II und IV	k. A.	Restaging nach Abschluss der NACR	3	12	2	31	20 (4,3 bis 48,1) <sup>b</sup>	94 (79,8 bis 99,3) <sup>b</sup>
Flamen 2002	36	PET	Histologie, Tumoran-sprechen wenn TONOMO	AJCC III	Reduktion der Tumor-to-liver uptake ratio > 80 %	Restaging nach Abschluss der NACR	10	4	4	18	71 (41,9 bis 91,6) <sup>b</sup>	82 (59,7 bis 94,8) <sup>b</sup>
Gillham 2006	32	PET und integrierte PET/CT <sup>e</sup>	Histologie, Tumoran-sprechen nach Mandard TRG 1+2	AJCC II und III	Reduktion der SUV > 20 %	Eine Woche nach Beginn der NAC / NACR (Interimsstaging)	4	5	11	12	44 (13,7 bis 78,8) <sup>b</sup>	52 (30,6 bis 73,2) <sup>b</sup>

(Fortsetzung)

Tabelle 31: Ergebnisse der Studien aus den Evidenzsynthesen zum Restaging (Fortsetzung)

Studie	Anzahl Patienten	Index-test	Referenztest	Population	Cut-Off	Indikation	RP	FN	FP	RN	Sensitivität % (95 %-KI)	Spezifität % (95 %-KI)
Kroep 2003	11 <sup>f</sup>	PET	Histologie, Tumoran-sprechen nach Mandard mit Modifikation	AJCC III	SUV > 40 %	Während der NAC (Interimsstaging)	4	0 <sup>b, g</sup>	1 <sup>b, g</sup>	6	100 <sup>b, g</sup> (39,8 bis 100) <sup>b</sup>	85,7 <sup>g</sup> (42,1 bis 99,6) <sup>b</sup>
	10 <sup>f</sup>				SUV > 60 %	Restaging nach Abschluss der NAC	4	0	0	6	100 (39,8 bis 100) <sup>b</sup>	100 (54,1 bis 100) <sup>b</sup>
	11 <sup>f</sup>	EUS			k. A.	Interimsstaging	5	0	0	6 <sup>b, g</sup>	100 (47,8 bis 100) <sup>b</sup>	100 (54,1 bis 100) <sup>b</sup>
						Restaging	5	0	0	6 <sup>b, g</sup>	100 (47,8 bis 100) <sup>b</sup>	100 (59,0 bis 100) <sup>b</sup>
	11 <sup>f</sup>	CT			k. A.	Interimsstaging	2	2	2	5	50 (6,8 bis 93,2) <sup>b</sup>	71 (29,0 bis 96,3) <sup>b</sup>
						Restaging	2	2	2	5	50 (6,8 bis 93,2) <sup>b</sup>	71 (29,0 bis 96,3) <sup>b</sup>
Levine 2006 <sup>h</sup>	41	PET	Histologie Tumoran-sprechen wenn kein Restgewebe vorhanden	AJCC I und IV	SUV: 4, Post-ChT + RT	Restaging nach Abschluss der NAC / NACR	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>
Ott 2006	56 <sup>i</sup>	PET	Histologie, Tumoran-sprechen nach Mandard mit Modifikation	AJCC II und III	≥ 35 % nach nach 2-wöchiger Chemotherapie	2 Wochen nach Beginn der NAC (Interimsstaging)	8	2	10	36	80 (44,4 bis 97,5) <sup>b</sup>	78 (63,6 bis 89,1) <sup>b</sup>

(Fortsetzung)

Tabelle 31: Ergebnisse der Studien aus den Evidenzsynthesen zum Restaging (Fortsetzung)

Studie	Anzahl Patienten	Index-test	Referenztest	Population	Cut-Off	Indikation	RP	FN	FP	RN	Sensitivität % (95 %-KI)	Spezifität % (95 %-KI)
Song 2005	32	PET	Histologie, Tumoran-sprechen wenn, Reduktion des Tumorumfanges (< 50%)	AJCC II und III	Maximale SUV < 2,5	Restaging nach Abschluss der NACR	3 <sup>a</sup>	8 <sup>a</sup>	1 <sup>a</sup>	20 <sup>a</sup>	27 <sup>d</sup> (6,0 bis 61,0) <sup>b</sup>	95 <sup>d</sup> (76,2 bis 99,9) <sup>b</sup>
Weber 2001	37 <sup>j</sup>	PET	Histologie, Tumoran-sprechen nach Mandat TRG 1+2	AJCC III	k. A.	2 Wochen nach Beginn der NAC (Interimsstaging)	8	1	7	21	88 <sup>b</sup> (51,8 bis 99,7) <sup>a, b</sup>	75 <sup>b</sup> (55,1 bis 89,3) <sup>b</sup>
Westertorp 2006 <sup>h</sup>	26	PET	Histologie, Tumoran-sprechen nach Mandat mit Modifikation	AJCC II und III	SUV: ≥ 31 %, 2 Wochen Behandlung	Während der NACR (Interimsstaging)	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>
Wieder 2004	23 <sup>h</sup>	PET	Histologie, Tumoran-sprechen nach Mandat mit Modifikation	AJCC II und III	SUV: ≥ 30 %, 2 Wochen Behandlung	Während der NACR (Interimsstaging)	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>	- <sup>h</sup>
	33	PET		AJCC II und III	SUV: 52 %, präoperativer PET scan	Restaging nach Abschluss der NACR	17	2	6	8	89 (66,9 bis 98,7) <sup>b</sup>	57 (28,9 bis 82,3) <sup>b</sup>

(Fortsetzung)

Tabelle 31: Ergebnisse der Studien aus den Evidenzsynthesen zum Restaging (Fortsetzung)

- a: Es wurden ursprünglich 27 Patienten in die Studie eingeschlossen. Daten liegen nur von 24 Patienten vor.
- b: Eigene Berechnung.
- 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.
- d: Daten kommen aus der Primärstudie.
- 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.
- f: Es wurden ursprünglich 13 Patienten in die Studien eingeschlossen.
- g: Die Angaben für Sensitivität und Spezifität wurden von den Autoren Ngamruengphong 2009 nicht korrekt dargestellt
- h: Die Ergebnisse wurden nicht dargestellt, weil der Anteil fehlender Werte größer als 30 % war.
- i: Ott 2006 schlossen insgesamt 65 Patienten in ihre Studien ein. Es liegen allerdings nur Daten von 56 Patienten vor.
- j: Weber 2001 schlossen insgesamt 40 Patienten in ihre Studien ein. Es liegen allerdings nur Daten von 37 Patienten vor.



## **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 6 weitere Diagnosestudien für die Indikation des Restagings identifiziert (Gillies 2012b, Higuchi 2008, Ilson 2011, Malik 2010, Roedl 2009 und van Heijl 2011).

Alle Studien machten Angaben zum Geschlecht der eingeschlossenen Patienten, die in der Summe 303 Männer und 81 Frauen ergaben. In 3 von 6 Studien wurden Patienten mit einem Plattenepithelkarzinom oder einem Adenokarzinom der Speiseröhre untersucht. In 2 Studien wurden ausschließlich Patienten mit einem Plattenepithelkarzinom [115,118] und in 1 Studie nur Patienten mit einem Adenokarzinom der Speiseröhre eingeschlossen [117]. 3 von 6 eingeschlossenen Studien wurden in Europa (keine Studie in Deutschland), 1 Studie im asiatischen Raum und 2 von 6 Studien in den USA durchgeführt. Der Einsatz der nicht integrierten FDG-PET wurde in 3 von 6 Studien, die integrierte FDG-PET/CT in 3 Studien untersucht. Als Tracer wurde in allen Studien Fluorodeoxyglukose (FDG) benutzt. 2 Studien untersuchten das Restaging nach Abschluss einer neoadjuvanten Chemoradiotherapie, 1 Studie das Restaging nach Abschluss einer neoadjuvanten Chemotherapie und 3 Studien untersuchten das frühe Behandlungsansprechen (Interimsstaging) während einer neoadjuvanten Chemoradiotherapie. Als Referenztest wurde die Histologie eingesetzt. Für das Therapieansprechen wurden unterschiedliche Klassifikationen des histologischen Ansprechens verwendet.

Tabelle 32: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche)

Studie	Studiendesign	Studienziel <sup>a</sup>	Zielkriterium	Indikation	Evidenzstufe <sup>b</sup>	Verzerrungspotenzial
Gillies 2012b	Diagnostische Kohortenstudie Rekrutierung der Patienten zwischen April 2008 und Februar 2010. Vereinigtes Königreich	Bewertung, ob metabolisches Ansprechen das pathologische Ansprechen auf eine standardisierte neoadjuvante Chemotherapie vorhersagen kann.	Diagnostische Güte von FDG-PET zur Bestimmung des histologischen Ansprechens auf Chemotherapie.	Restaging nach Abschluss der NAC	I Ib	hoch
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.	Restaging nach Abschluss der NAC / NACR	I Ib	hoch
Ibson 2011	Diagnostische Kohortenstudie Rekrutierung der Patienten zwischen Dezember 2002 und Oktober 2005. USA	Bestimmung des Anteils des pathologisch vollständigen Ansprechens von Patienten mit Karzinomen der Speiseröhre oder des gastroösophagealen Übergangs auf Chemoradiotherapie und Induktionschemo mit eingebetteter Evaluation der diagnostischen Güte der PET.	Diagnostische Güte von FDG-PET nach Induktionschemo zur Vorhersage des histologischen Ansprechens auf Chemoradiotherapie.	Frühes Behandlungsansprechen (Interimsstaging) während NAC(R)	I Ib	hoch
Malik 2010	Diagnostische Kohortenstudie Rekrutierung der Patienten zwischen September 2003 und Dezember 2007. Irland	Bestimmung, ob FDG-PET in der 2. Woche einer 3-wöchigen Chemoradiotherapie das histologische Behandlungsansprechen auf und Überleben nach einer 6-wöchigen NACR vorhersagen kann.	Diagnostische Genauigkeit von FDG-PET bei der Vorhersage des Ansprechens auf Chemoradiotherapie.	Frühes Behandlungsansprechen (Interimsstaging) während NACR	I Ib	hoch

(Fortsetzung)

Tabelle 32: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche)  
(Fortsetzung)

Studie	Studiendesign	Studienziel <sup>a</sup>	Zielkriterium	Indikation	Evidenzstufe <sup>b</sup>	Verzerrungspotenzial
Roedl 2009	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 krankheitsfreien Überlebens von Patienten mit Ösophaguskarzinomen.	Diagnostische Genauigkeit von FDG-PET bzw. PET/CT zur Bestimmung des Ansprechens auf Chemoradiotherapie. Angabe der Ergebnisse verschiedener Parameter.	Restaging nach Abschluss der NACR	IIb	hoch
van Heijl 2011	Diagnostische Kohortenstudie (in RCT eingebettet) Rekrutierung der Patienten zwischen Januar 2005 und November 2008. Niederlande	Untersuchung, inwieweit eine Unterscheidung von Patienten mit Ösophaguskarzinomen hinsichtlich des Ansprechens bzw. Nichtansprechens auf eine neoadjuvante Chemoradiotherapie möglich ist.	Diagnostische Genauigkeit von FDG-PET bei der Bestimmung des Ansprechens auf Chemoradiotherapie.	Frühes Behandlungsansprechen (Interims-staging) während NACR	IIb	hoch
<p>a: Studienziel gemäß Formulierung der Autoren  b: entspricht der Evidenzstufe des G-BA  FDG = <sup>18</sup>Fluorodeoxyglukose; NAC = neoadjuvante Chemotherapie; NACR = neoadjuvante Chemoradiotherapie; NAC/R = Induktionschemotherapie gefolgt von NACR; PET = Positronenemissionstomographie; USA = Vereinigte Staaten von Amerika</p>						

Tabelle 33: Index- und Referenztest der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche)

Studie	Indextest	Referenztest
Gillies 2012b	<b>PET/CT</b> (GE Discovery STE 16, General Electric Co.) 60 Minuten nach i.v.-Injektion von 5 MBq/kg KG (mind. 300 MBq) FDG. Aufnahme zu Studienbeginn und 24–28 Tage nach dem Beginn des 2. Zyklus der Chemotherapie. Bestimmung von SUVmax sowie des Vorliegens von Lymphknoten- und Fernmetastasen, Befundung durch 2 Radiologen.	Histologie nach OP. Pathologisches Ansprechen nach Mandard (1–3 Ansprechen, 4+5 kein Ansprechen). Bewertung durch 2 Untersucher im Konsens. Bei aufgrund von Progression nicht resektierten Patienten wurde ein Nichtansprechen angenommen.
Higuchi 2008	<b>PET</b> (HEADTOME/SET 2400 W, Shimadzu Co, Japan) 60 Minuten nach i.v.-Injektion von ca. 370 MBq FDG. Läsionen mit SUVmax $\geq 2,5$ 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.
Ilsou 2012	<b>PET</b> (keine Angaben zum Gerät) 60 Minuten nach i.v.-Injektion von mind. 7 mCi FDG in einem von mehreren Studienzentren. Aufnahme zu Studienbeginn und nach Ende der 5-wöchigen Induktionstherapie. Bestimmung des SUV sowie Berechnung der prozentualen Änderung zwischen 1. und 2. Aufnahme. Keine Angaben zum Bewerter oder zur Entscheidungsfindung.	Histologie nach OP oder ÖGD bzw. CT und Bewertung von pathologischem Ansprechen. Keine Angaben zur Befundung.
Malik 2010	<b>PET</b> (PET Advance Scanner [n=12] oder Discovery ST PET/CT Scanner [n=25, GE Healthcare, USA]) ca. 60 Minuten nach i.v.-Injektion von 350–450 MBq FDG vor Beginn und in der 2. Woche der Chemoradiotherapie. Patienten wurden bei beiden Terminen mit demselben Gerät untersucht. Aufnahme von Schädel bis Oberschenkel. Bestimmung von SUVmax in einer „region of interest“ und Berechnung der Änderung zwischen den beiden Aufnahmen. Keine Angaben zum Bewerter oder zur Entscheidungsfindung.	Histologie Tumoransprechen < 10 % wachstumsfähige Tumorzellen (entspricht Mandard 1+2). Auswertung durch einen erfahrenen Histopathologen verblindet.
Roedl 2009	<b>PET/CT</b> (Biograph Sensation 16 scanner, Siemens, Erlangen, Deutschland) ca. 60 Minuten nach i.v.-Injektion von 15 mCi (555 MBq) FDG vor Beginn und nach Ende der Chemoradiotherapie. Tumoransprechen durch Abnahme von Tumorgröße, Tumordurchmesser, SUV des Tumors, Index (Tumordurchmesser x SUV des Tumors) und SUVmax. 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.

(Fortsetzung)

Tabelle 33: Index- und Referenztest der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche) (Fortsetzung)

Studie	Indextest	Referenztest
van Heijl 2011	<p><b>PET/CT</b> (Ecat Exact, Siemens, USA; Biograph PET/CT, Siemens, USA; Gemini PET/CT, Philips Medical Systems, USA; Gemini TF PET/CT, Philips, USA) ca. 90 Minuten nach i.v.-Injektion von 5 MBq/kg KG FDG in einem von 5 beteiligten PET-Zentren. Beide Aufnahmen (tumorfokussiert, statisch; zu Studienbeginn und 14 Tage nach Beginn der Chemoradiotherapie) mit demselben Gerät durchgeführt.</p> <p>Zentrale Befundung (VU Medical Center, Amsterdam) nach interinstitutioneller Kalibrierung. Keine Angaben zu den Bewertern, zur Verblindung und zur Entscheidungsfindung.</p> <p>Bestimmung von SUVmax korrigiert für Körpergewicht und Serumblutzucker und Berechnung der Änderung zwischen 1. und 2. Aufnahme.</p>	<p>Histologie nach OP</p> <p>Tumoransprechen &lt; 10 % wachstumsfähige Tumorzellen.</p> <p>Auswertung durch 1 von 2 erfahrenen Histopathologen.</p>
<p>a: Für mindestens 3 Patienten mit Tumoransprechen, die nicht operiert wurden, war offensichtlich die Nachbeobachtung der Referenzstandard. Dies wird in der Publikation nicht näher beschrieben.</p> <p>FDG = 18Fluorodeoxyglukose; MBq = Megabecquerel; mCi = Millicurie; ÖGD = Ösophagogastrroduodenoskopie; PET = Positronenemissionstomographie; SUV = standardized uptake value; USA = Vereinigte Staaten von Amerika</p>		

Tabelle 34: Eingeschlossene Patienten der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche)

Studie	Patienten [N]	Geschlecht [N (m/w)]	Alter (MW, Median, Spannweite)	Diagnose	Behandlung
Gillies 2012b	48	48 (37/11)	65; k. A.; k. A.	Histologie: Adenokarzinom (8), Plattenepithelkarzinom (38), adenosquamöser (1) und schlecht differenzierter Tumor (1) Lokalisation: mittlere (6), untere Speiseröhre (27) und gastroösophagealer Übergang (Sievert 1 + 2) (15). Klinisches Stadium (vor Behandlung): II (22), III (26)	Nach Studienbeginn: NAC (Oxaliplatin and 5-Fluorouracil, 2 Zyklen im Abstand von 3 Wochen)
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 NAC und 15 NACR. 11 zervikale, 39 subtotale Ösophagektomie.
Ilson 2011	60	55 (47/8)	k. A.; 59; 21–74	Histologie: Adenokarzinom (41), Plattenepithelkarzinom (12), schlecht differenzierter Tumor (2) Lokalisation: obere (1), mittlere (7), untere Speiseröhre (29) und gastroösophagealer Übergang (18).	Nach Studienbeginn: NAC (Cisplatin, Irinotecan Woche 1–5) . NACR (50,4 Gy, Woche 7–11).
Malik 2010	37	37 (31/6)	58; k. A.; 37–73	37 Patienten mit neu diagnostiziertem, unbehandeltem, lokal fortgeschrittenem (cT3/T4, N0/N1 basierend auf CT) Adenokarzinom der Speiseröhre oder des gastroösophagealen Übergangs.	Nach Studienbeginn: NACR (40 Gy, 5-Fluorouracil, Cisplatin) über 6 Wochen.

(Fortsetzung)

Tabelle 34: Eingeschlossene Patienten der eingeschlossenen Studien zum Restaging (Ergänzungsrecherche) (Fortsetzung)

Studie	Patienten [N]	Geschlecht [N (m/w)]	Alter (MW, Median, Spannweite)	Diagnose	Behandlung
Roedl 2009	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: NACR (50,4 Gy, Cisplatin und 5-Fluorouracil nach 1 und 5 Woche(n) nach Beginn).
van Heijl 2011	108 <sup>a</sup>	145 (108/37)	60; k. A.; 37–79	Histologie: Adenokarzinom (119), Plattenepithelkarzinom (26). T 1 (1), T2 (13), T 3 (131); N0 (50), N1 (95)	Nach Studienbeginn: NACR (41,4 Gy, Paclitaxel, Carboplatin) über 5 Wochen.

a: 100 Patienten wurden in die Analyse eingeschlossen. Die Anzahl der eingeschlossenen Patienten (N) war 108. Charakteristika wurden für 145 Patienten berichtet. CT = Computertomographie; Gy = Gray; OP = Operation; LK = Lymphknoten; m = männlich; N = Lymphknoten; NAC = neoadjuvante Chemotherapie; NACR = neoadjuvante Chemoradiotherapie; T = Tumor; UICC = Union internationale contre le cancer (dt.: Internationale Vereinigung gegen Krebs); 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. Alle eingeschlossenen Studien hatten ein hohes Verzerrungspotenzial.

Generell wurde festgestellt, dass bei 2 Studien (Gillies 2012b und Roedl 2009) keine genauen Angaben zur Dauer zwischen dem Index- und Referenztest gemacht wurden. Bei Ilson 2011 wurde die Zeitspanne mit 4 bis 8 Wochen als zu lang eingestuft. Bei 5 Studien war es unklar, ob die Befunder bei der Interpretation der Ergebnisbilder der Indextests und / oder Referenztests verblindet waren. Die Ergebnisse bei Malik 2010 wurden verblindet ausgewertet.

Tabelle 35: Verzerrungspotenzial der eingeschlossenen Studien beim Restaging (Ergänzungsrecherche)

Studie	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Verzerrungspotenzial der Ergebnisse
Gillies 2012b	+	?	+	+	+	-	?	+	-	+	hoch
Higuchi 2008	+	+	+	+	?	+	?	-	+	?	hoch
Ilson 2011	+	-	+	+	-	?	?	+	-	+	hoch
Malik 2010	+	+	+	+	+	+	+	-	-	+	hoch
Roedl 2009	+	?	+	+	?	+	?	+	-	?	hoch
van Heijl 2011	+	+	+	+	+	+	?	-	?	+	hoch
+ = Kriterium erfüllt; - = Kriterium nicht erfüllt; ? = unklar											
F1: Beschreibung des Referenztests; F2: Zeitspanne zwischen dem Index- und Referenztest; F3: Interdependenz der Tests; F4: partielle Verifikation; F5: differenzielle Verifikation; F6: Incorporation Bias; F7: Verblindung bei Interpretation der Testergebnisse; F8: Intention-to-diagnose- / Intention-to-screen-Analyse; F9: ergebnisunabhängige Berichterstattung; F10: sonstige Aspekte											

### 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 untersucht. In 2 Studien (**Gillies 2012b** und **van Heijl 2011**) wurden keine Indizien für eine Nichtübertragbarkeit gefunden. Eine generelle Übertragbarkeit dieser Studien auf den deutschen Versorgungskontext wird daher angenommen.



Bei 4 Studien (**Higuchi 2008**, **Ibson 2011**, **Malik 2010** und **Roedl 2009**) lagen Indizien vor, die eine Übertragbarkeit der Ergebnisse für das Primärstaging auf den deutschen Versorgungskontext einschränken könnten.

In der Studie **Higuchi 2008** wurden ausschließlich asiatische Patienten untersucht. Wie bereits erwähnt, wurde durch die Stellungnahme der DGN der Hinweis gegeben, dass aufgrund der bekannten tumorbiologischen Unterschiede zwischen asiatischen und europäischen Patienten die Übertragbarkeit der Studienergebnisse einschränken sein könnte. Darüber hinaus wurde in der Stellungnahme der DGN beschrieben, dass der in der Studie von Higuchi 2008 verwendete Referenztest eine in Europa nicht übliche Klassifikation des histologischen Ansprechens verwendet. Darüber hinaus wurden die Untersuchungen mit einem PET-Gerät durchgeführt, welches ebenfalls in Europa nicht verfügbar ist, wie aus der Stellungnahme der DGN entnommen werden konnte.

Aus der Studienpublikation **Ibson 2011** gingen zu wenige Informationen zur jeweiligen Testdurchführung hervor, um eine Replikation zu ermöglichen.

In der Studie **Malik 2010** wurde angegeben, dass die Interpretation der PET-Bilder ohne Kenntnis der klinischen Patientendaten erfolgte. Den Beurteilern der PET bzw. aller weiteren Tests sind im klinischen Alltag sehr wahrscheinlich die Ergebnisse einer vorangegangenen Diagnostik bekannt, sodass plausible Indizien dafür vorliegen, dass die Ergebnisse möglicherweise nicht auf die Anwendungssituation in Deutschland gelten.

In der Studie **Roedl 2009** wurden 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 36).

Genannte Aspekte entsprechen möglicherweise nicht dem klinischen Alltag, sodass es plausible Gründe dafür gibt, nicht auf die Anwendungssituation in Deutschland gelten.

Tabelle 36: Indizien für eine Nichtübertragbarkeit der Ergebnisse der eingeschlossenen Studien beim Restaging (Ergänzungsrecherche)

Studie	F1	F2	F3	F4	F5	F6	F7	Generelle Einschätzung der Übertragbarkeit
Gillies 2012b	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
Higuchi 2008	+	+	-	-	-	-	-	Indiz für Nichtübertragbarkeit
Ilson 2011	-	+	-	-	-	-	-	Indiz für Nichtübertragbarkeit
Malik 2010	+	-	-	+	-	-	-	Indiz für Nichtübertragbarkeit
Roedl 2009	-	-	+	+	-	+	-	Indiz für Nichtübertragbarkeit
van Heijl 2011	-	-	-	-	-	-	-	Übertragbarkeit wird angenommen
+ = konkretes Indiz für eine Nichtübertragbarkeit lag vor - = kein Indiz für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden) F1: Repräsentativität Patientenspektrum, F2: Übertragbarkeit Index- und Referenztest, F3: übertragbarer Entscheidungsprozess bei Testinterpretation, F4: übertragbare Hintergrundinformation, F5: technologische Entwicklung, F6: Untersucherqualifikation, F7: Qualität der Geräte								

#### 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 37) aus der Ergänzungsrecherche dargestellt. Insgesamt wurden noch 6 weitere Primärstudien, die nicht durch die Evidenzsynthesen abgedeckt wurden, eingeschlossen.

Tabelle 37: 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	FP	FN	RN	Sensitivität in % [95 %- KI] <sup>a</sup>	Spezifität in % [95 %- KI] <sup>a</sup>
Gillies 2012b	48 (48)	Restaging nach Abschluss der NAC <sup>b</sup>	FDG-PET/CT	keine FDG-Aufnahme nach Behandlung	Histologie nach OP <sup>c</sup>	Mandard-Klassifikation 1-3 (Ansprechen), 4+5 (kein Ansprechen) <sup>d</sup>	13 <sup>e</sup>	3	12	20	52,0 [31,3;72,2]	87,0 [66,4;97,2]
				$\Delta\text{SUV} > 42\%$ = Median <sup>f</sup>		17	7	8	16	68,0 [46,5;85,1]	69,6 [47,1;86,8]	
				$\Delta\text{SUV} > 30\%$ = PERCIST <sup>g</sup>		18	8	7	15	72,0 [50,6;87,9]	65,2 [42,7;83,6]	
				keine FDG-Aufnahme nach Behandlung		11	3	9	15	55,0 [31,5;76,9]	83,3 [58,6;96,4]	
				$\Delta\text{SUV} > 50\%$ = Median		14	5	6	13	70,0 [45,7;88,1]	72,2 [46,5;90,3]	
Higuchi 2008	50 (50)	Restaging nach Abschluss der NAC / NACR	FDG-PET (nach Behandlung)	$\text{SUV}_{\max} < 2,5$	Histologie	> 2/3 Verschwinden von Tumorzellen	27 <sup>a</sup>	2 <sup>a</sup>	3 <sup>a</sup>	18 <sup>a</sup>	90,0 [73,5;97,9] <sup>i</sup>	90,0 [68,3;98,8] <sup>j</sup>
Ilson 2011	53 (60) <sup>j, k</sup>	Frühes Behandlungsansprechen (Interimsstaging) während NAC/R	FDG-PET	Abnahme SUV > 35 %, entspricht medianem SUV	Histologie	komplettes Ansprechen (+)	8	17	1	27	88,9 [51,8;99,7]	61,4 [45,5;75,6]
Malik 2010	37 (37) <sup>l</sup>	Frühes Behandlungsansprechen (Interimsstaging) während NACR	FDG-PET	Abnahme $\text{SUV}_{\max} \geq 26,4\%$ (entspricht dem Median)	Histologie	< 10 % lebende Tumorzellen (entspricht Mandard 1+2)	10	6	6	15	62,5 [35,5;84,7]	71,4 [47,8;88,6]

(Fortsetzung)

Tabelle 37: Ergebnisse der Studien zur diagnostischen Güte der PET und anderer Indextests beim Restaging von Ösophaguskarzinomen (Ergänzungsrecherche) (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Cut-off	Referenz- test	Cut-off	RP	FP	FN	RN	Sensitivi- tät in % [95 %- KI] <sup>a</sup>	Spezifität in % [95 %- KI] <sup>a</sup>
Roedl 2009	49 (49)	Restaging nach Abschluss der NACR	FDG-PET (Abnahme der Werte zwischen Prä- und Post- Behandlung)	59 % (Reduktion des Bereiches)	Histologie	< 10 % lebende Tumorzellen	17 <sup>a</sup>	4 <sup>a</sup>	5 <sup>a</sup>	23 <sup>a</sup>	77 [54,6;92,2]	85 [66,3;95,8]
				32 % (Reduktion des Durchmessers)			18 <sup>a</sup>	3 <sup>a</sup>	4 <sup>a</sup>	24 <sup>a</sup>	82 [59,7;94,8]	89 [70,8;97,6]
				43 % (Reduktion des SUV des Bereichs)			13 <sup>a</sup>	0 <sup>a</sup>	9 <sup>a</sup>	27 <sup>a</sup>	59 [36,4;79,3]	100 [87,2; 100]
				56 % (Index = Durchmesser – SUV)			20 <sup>a</sup>	2 <sup>a</sup>	2 <sup>a</sup>	25 <sup>a</sup>	91 [70,8;98,9]	93 [75,7;99,1]
			FDG- PET/CT (Abnahme der Werte Prä- und Post- Behandlung)	54 % (Bereich)			15 <sup>a</sup>	2 <sup>a</sup>	7 <sup>a</sup>	25 <sup>a</sup>	68 [45,1;86,1]	93 [75,7;99,1]
			23 % (Durch- messer)	20 <sup>a</sup>			3 <sup>a</sup>	2 <sup>a</sup>	24 <sup>a</sup>	91 [70,8;98,9]	89 [70,8;97,6]	
			42 % (SUV des Bereichs)	15 <sup>a</sup>			0 <sup>a</sup>	7 <sup>a</sup>	27 <sup>a</sup>	68 [45,1;86,1]	100 [87,2; 100]	
			55 % (Index = Durchmesser – SUV)	20 <sup>a</sup>			2 <sup>a</sup>	2 <sup>a</sup>	25 <sup>a</sup>	91 [70,8;98,9]	93 [75,7;99,1]	
			PET und / oder PET/CT <sup>m</sup>	42 % (SUV <sub>max</sub> ) <sup>m</sup>			18 <sup>a</sup>	8 <sup>a</sup>	4 <sup>a</sup>	19 <sup>a</sup>	82 [59,7;94,8]	70 [49,8;86,2]

(Fortsetzung)

Tabelle 37: Ergebnisse der Studien zur diagnostischen Güte der PET und anderer Indextests beim Restaging von Ösophaguskarzinomen (Ergänzungsrecherche) (Fortsetzung)

Studie	n (N)	Indikation	Indextest	Cut-off	Referenz- test	Cut-off	RP	FP	FN	RN	Sensitivität in % [95 %- KI] <sup>a</sup>	Spezifität in % [95 %- KI] <sup>a</sup>
van Heijl 2011	100 (108) <sup>n</sup>	Frühes Behandlungs- ansprechen (Interimsstaging) während NACR	PET (SUV <sub>max</sub> - Abnahme zwischen Studien- beginn und während Therapie)	> 0 %; ≤ 0 %	Histologie	< 10 % lebende Tumorzellen	58	18	6	18	90,6 [80,7;96,5]	50,0 [32,9;67,1]
				> 10 %; ≤ 10 %			52	16	12	20	81,3 [69,5;89,9]	55,6 [38,1;72,1]
				> 20 %; ≤ 20 %			45	13	19	23	70,3 [57,6;81,1]	63,9 [46,2;79,2]
				> 30 %; ≤ 30 %			35	12	29	24	54,7 [41,7;67,2]	66,7 [49,0;81,4]
<p>a: eigene Berechnung</p> <p>b: Abnahme von SUV<sub>max</sub> zwischen der 1. Aufnahme (Studienbeginn) und 2. Aufnahme (24-28 Tage nach 2. Zyklus der Chemotherapie)</p> <p>c: Bei 2 Patienten mit starker Progression (in der PET) wurde keine OP durchgeführt und ein Nichtansprechen angenommen.</p> <p>d: Bei 2 Patienten mit starker Progression (in der PET) wurde keine OP durchgeführt und ein Nichtansprechen angenommen.</p> <p>e: Angaben aus Tabelle 4 der Publikation. Eigene Berechnung von Punktschätzungen und KIs.</p> <p>f: Abnahme von SUV<sub>max</sub> zwischen der 1. Aufnahme (Studienbeginn) und 2. Aufnahme (24-28 Tage nach 2. Zyklus der Chemotherapie) &gt; 42 %</p> <p>g: Abnahme von SUV<sub>max</sub> zwischen der 1. Aufnahme (Studienbeginn) und 2. Aufnahme (24-28 Tage nach 2. Zyklus der Chemotherapie) &gt; 30 %</p> <p>h: Diese Auswertungen beziehen sich auf die Subgruppe der Adenokarzinome.</p> <p>i: 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.</p> <p>j: 60 Patienten eingeschlossen. 5 Patienten (3 mit anderen Primärtumoren, 1 fehlende Compliance, 1 Fernmetastasen) von der Bewertung ausgeschlossen. 55 Patienten ausgewertet</p> <p>k: 2 Patienten nicht ausgewertet (1 kein SUV zu Studienbeginn berichtet, 1 kein SUV-Aufnahme des Primärtumors)</p> <p>l: Ursprünglich 70 Patienten. Unklar, ob 33 Patienten ausgeschlossen wurden, weil sie den Einschlusskriterien nicht entsprachen oder ob sie als Studienabbrecher zu werten sind.</p> <p>m: Es wurde nicht beschrieben, auf welchen Indextest oder eventuell auf eine Kombination beider Tests sich dieser Cut-off bezieht.</p> <p>n: 100 Patienten wurden in die Analyse eingeschlossen. Die Anzahl der eingeschlossenen Patienten (N) war 108. Charakteristika sind für alle 145 eingeschlossenen Patienten berichtet.</p> <p>FDG = <sup>18</sup>Fluorodeoxyglukose; FN = falsch negativ; FP = falsch positiv; KI = Konfidenzintervall; NAC = neoadjuvante Chemotherapie; NACR = neoadjuvante Chemoradiotherapie; NAC/R = Induktionschemotherapie gefolgt von NACR; PET = Positronenemissionstomographie; RN = richtig negativ; RP = richtig positiv; SUV = Standardized Uptake Value</p>												

### **Univariate Ergebnisse für die diagnostische Güte der PET bzw. 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 wurden nur die Studien dargestellt, aus denen eindeutig hervorging, dass die PET bzw. PET/CT nach Beendigung der Therapie oder während der Therapie erfolgte. Die eingeschlossenen Studien zum Restaging verwendeten unterschiedliche Klassifikationen zum histologischen Ansprechen. Diese werden in Tabelle 31 dargestellt. Die Verwendung der verschiedenen Klassifikationen als Referenzstandards erschwert eine Vergleichbarkeit der Studien. Dies sollte bei der Interpretation von Abbildung 6 mit in Betracht gezogen werden.

8 Studien untersuchten die PET bzw. PET/CT in der Indikation des Interimsstaging (siehe Abbildung 6). 3 dieser Studien (Weber 2001, Ott 2006, Kroep 2003) untersuchten den Einsatz der PET während einer neoadjuvanten Chemotherapie. Hier zeigte sich für die PET eine Spannweite der Punktschätzungen der Sensitivitäten von 80 % (bei einer zugehörigen Spezifität von 78 % bzw. 100 %) bis 88 % (zugehörige Spezifität 75 %). Für alle Punktschätzungen waren weite Konfidenzintervalle zu beobachten.

In den weiteren 4 Studien (Westerterp 2006, Malik 2010, Wieder 2004 und van Heijl 2011) wurde der Einsatz der PET während einer neoadjuvanten Chemoradiotherapie untersucht. Allerdings konnten für 2 Studien (Westerterp 2006 und Wieder 2004) die Ergebnisse nicht dargestellt werden, weil der Anteil fehlender Werte größer als 30 % war. Die Punktschätzungen der Sensitivität der 2 noch bleibenden Studien lagen bei 54,7 % (zugehörige Spezifität 66,7 % bei einem Cut-off von > 30 %) und 62,5 % (zugehörige Spezifität 71,4 % bei einem Cut-off von 26,4 %). In 2 Studien (Gillham 2006 und Ilson 2011) wurde der Einsatz der PET während beider Therapieregimes untersucht. Die Punktschätzungen der Sensitivität der 2 Studien lagen bei 44 % (zugehörige Spezifität 52 %) und 88,9 % (zugehörige Spezifität 61,4 %). Für alle Punktschätzungen waren weite Konfidenzintervalle zu beobachten. Für den Einsatz der PET nach Beendigung der Therapie lagen direkte Vergleiche vor, welche im folgenden Abschnitt beschrieben werden.

Ösophaguskarzinom

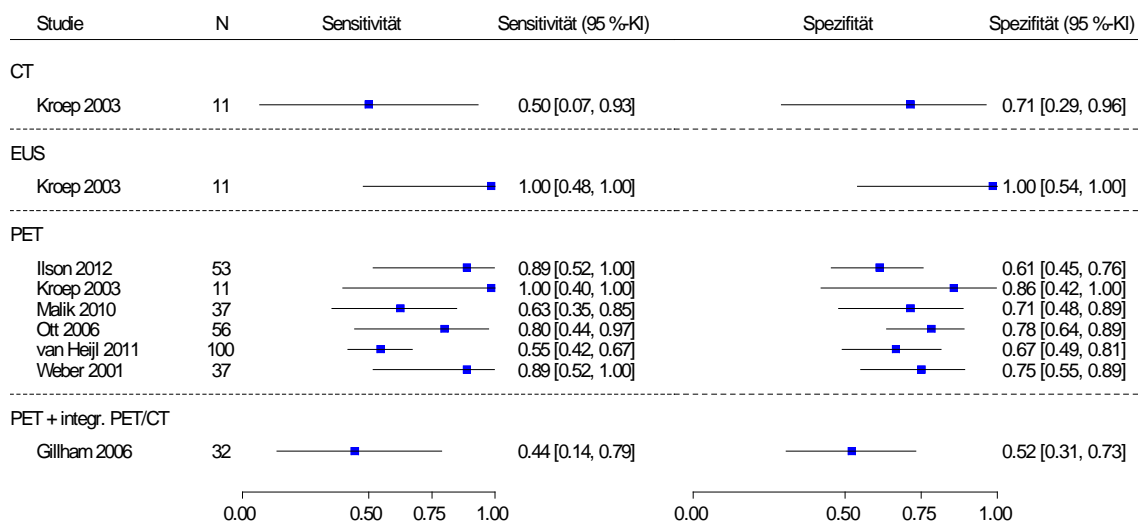


Abbildung 6: Forestplots der univariaten Ergebnisse für Sensitivität und Spezifität beim Interimsstaging von Ösophaguskarzinomen

Ösophaguskarzinom

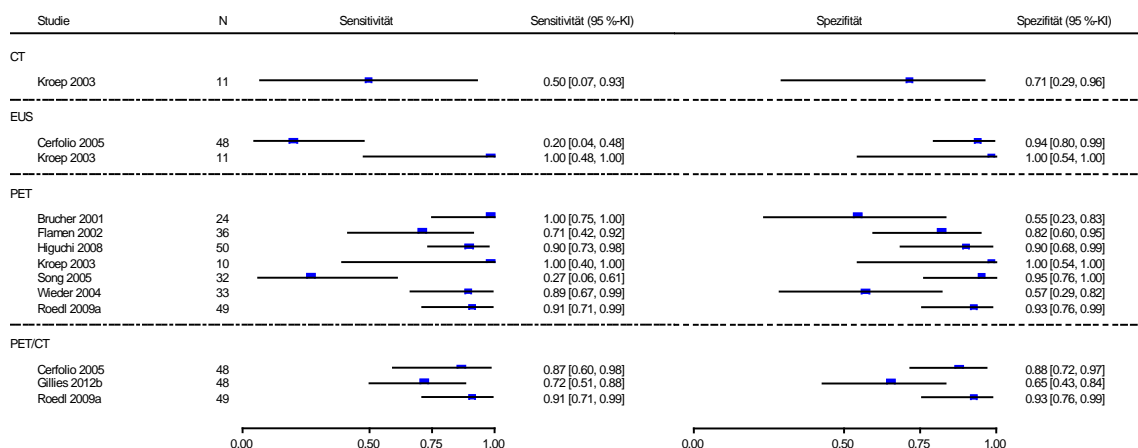


Abbildung 7: Forestplots der univariaten Ergebnisse für Sensitivität und Spezifität beim Restaging von Ösophaguskarzinomen nach Abschluss der Therapie

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

In den 17 für das Restaging eingeschlossenen Diagnosestudien wurden 3 direkte Vergleiche zwischen PET und anderen bildgebenden Verfahren für das Restaging und 1 direkter Vergleich für das Interimsstaging 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 einer neoadjuvanten Therapie wurde in 3 Studien (Kroep 2003, Roedl 2009 und Cerfolio 2005) ein direkter Vergleich diagnostischer Technologien untersucht:

Der Vergleich PET/CT versus Endosonografie (EUS) wurde in der Studie **Cerfolio 2005** untersucht. Es zeigte sich für die PET/CT eine Sensitivität von 86,7 % (95 %-KI [59,7; 98,3]) und eine Spezifität von 87,9 % (95 %-KI [71,8; 96,6]). Im Gegensatz dazu hatte die EUS eine Sensitivität von 20 % (95 %-KI [4; 48]) und eine Spezifität von 94 % (95 %-KI [79,8;99,3]). Die beiden Verfahren wurden bezüglich der Sensitivität und Spezifität je in einem nicht genau beschriebenen statistischen Test einander gegenübergestellt.

In der Studie **Roedl 2009** wurde der Vergleich PET versus PET/CT untersucht. Hier zeigte sich für die PET eine Sensitivität von 59 % (95 %-KI [36,4; 79,3]) bei einem Cut-off von 43 %, eine SUV-Reduktion und eine Spezifität von 100 % (95 %-KI [87,2;100]). Die PET/CT erreichte eine Sensitivität von 68 % (95 %-KI [45,1; 86,1]) bei einem Cut-off von 42 %, eine SUV-Reduktion und eine Spezifität von 100 % (95 %-KI [87,2; 100]). In der Studie wurde nicht berichtet, ob ein für den vorliegenden Bericht relevanter statistischer Test gerechnet wurde.

In der Studie **Kroep 2003** wurde die PET mit der EUS und der CT basierend auf 10 bzw. 11 eingeschlossenen Patienten verglichen. Für die PET zeigte sich eine Sensitivität von 100 %, (95 %-KI [39,8;100]) und eine Spezifität von 100 % (95 %-KI [54,1;100]). Die Vergleichstechnologie EUS hatte ebenfalls eine Sensitivität von 100 % (95 %-KI [47,8;100]) und Spezifität von 100 % (95 %-KI [54,1;100]). Im Gegensatz dazu hatte die CT eine Sensitivität von 50 % (95 %-KI [6,8;93,2]) und eine Spezifität von 71 % (95 %-KI [29,0;96,3]).

Für das Interimsstaging lag ein direkter Vergleich diagnostischer Technologien vor:

In der Studie von **Kroep 2003** wurde die PET mit der EUS und der CT ebenfalls für das Interimsstaging verglichen. Die PET hatte eine Sensitivität von 100 % (95 %-KI [39,8;100]) und eine Spezifität von 85,7 % (95 %-KI [42,1;99,6]). Die Vergleichstechnologie EUS hatte eine Sensitivität von 100 % (95 %-KI [47,8;100]) und Spezifität von 100 % (95 %-KI [54,1;100]). Im Gegensatz dazu hatte die CT eine Sensitivität von 50 % (95 %-KI [6,8;93,2]) und eine Spezifität von 71 % (95 %-KI [29,0;96,3]).

### **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 diese Primärstudien im Volltext gesichtet [68,74,79,86,129].

In 2 dieser Studien wurde kein multifaktorielles Modell berechnet, sodass die Einschlusskriterien nicht erfüllt waren [68,86].

In der Studie von Flamen 2002 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 **Ott 2006** 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 3 weitere Prognosestudien, **Higuchi 2008**, **Kim 2007** und **Ma 2012**, für die Indikation Restaging, die den Einschlusskriterien des vorliegenden Berichts entsprachen, eingeschlossen werden. Bei der Studie 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 Prognosestudien (**Ott 2006**, **Kim 2007** und **Ma 2012**) beim Restaging präsentiert. Im Anschluss erfolgt die Darstellung der „prognostic accuracy study“ **Higuchi 2008**.

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

Die Studie von **Kim 2007** 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 einfaktoriel und multifaktoriel 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. 4 Patienten verstarben an postoperativen Komplikationen und wurden nicht in die Analyse eingeschlossen.

In die Studie **Ma 2012** wurden 50 Patienten mit nachgewiesenen Plattenepithelkarzinomen eingeschlossen. Ziel der Arbeit war die Bewertung, ob eine Abnahme der FDG-Aufnahme bei Patienten mit Ösophaguskarzinomen mit dem Gesamtüberleben und lokalen Kontrollraten korreliert ist. Die Bildgebung mittels FDG-PET wurde zu Studienbeginn sowie während und nach Abschluss einer Chemoradiotherapie durchgeführt. Multivariate Analysen ermittelten den Einfluss verschiedener PET-Parameter (Änderungen zwischen unterschiedlichen Messzeitpunkten) bezüglich lokaler Kontrollraten und des Gesamtüberlebens.

In der Studie **Ott 2006** 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 38: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudien beim Restaging

Studie	Studiendesign	Studienziel <sup>a</sup>	Zielkriterium	Evidenzstufe <sup>b</sup>	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.	IIC	hoch
Ma 2012	Prognostische Kohortenstudie Rekrutierung der Patienten zwischen März 2006 und Oktober 2008. China	Bewertung, ob eine Abnahme der FDG-Aufnahme bei Patienten mit Ösophaguskarzinomen mit dem Gesamtüberleben und lokalen Kontrollraten korreliert ist.	Prognostische Güte von FDG-PET zur Vorhersage des Gesamtüberlebens und lokaler Kontrollraten.	IIC	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.	IIC	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 mit Stufe II gekennzeichnet.  FDG = 18Fluorodeoxyglukose; PET = Positronenemissionstomographie</p>					

Tabelle 39: Index- und Referenztest der eingeschlossenen Prognosestudien beim Restaging

Studie	Indextest	Referenztest
Kim 2007	<b>PET</b> (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 (SUVmax) durch erfahrenen Nuklearmediziner. Bewertung der metabolischen Antwort nach EORTC-Kriterien.	Überleben (Datum der Operation bis Tod)
Ma 2012	<b>PET/CT</b> (Discovery LS PET-CT scanner, General Electric Co., USA) 45 Minuten nach i.v.-Injektion von durchschnittlich 350 (Spannweite 259–444) MBq FDG. Aufnahme von Hals und Brustkorb.  3 Aufnahmen (Studienbeginn, nach 1. Bestrahlung und nach Ende der Strahlentherapie). Bestimmung des SUVmax und Berechnung von Änderungen zwischen den Aufnahmen. Keine Angaben zum Bewerter und zur evtl. Entscheidungsfindung.	Gesamtüberleben (Zeit vom Beginn der Strahlentherapie bis Tod oder letzte Nachbeobachtung, Nachbeobachtung 3 Jahre)  Lokale Kontrolle (Barium-Breischluck-Untersuchung oder CT alle 3 Monate für 2 Jahre, dann alle 6 Monate zur Diagnose von lokalen und regionalen Rezidiven sowie Fernmetastasen).
Ott 2006	<b>PET</b> (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
CT = Computertomographie; FDG = 18Fluorodeoxyglukose; i.v. = intravenös; kg = Kilogramm; KG = Körpergewicht; MBq = Mega Becquerel; PET = Positronenemissionstomographie; SUV = standardised uptake value; TI = 201Thaliumchlorid; USA = Vereinigte Staaten von Amerika		

Tabelle 40: Eingeschlossene Patienten der eingeschlossenen Prognosestudien beim Restaging

Studie	Patienten [N]	Geschlecht [N (m/w)]	Alter (MW, Median, Spannweite)	Diagnose	Vorbehandlung
Kim 2007	62	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 NAC (9: 5-FU, Cisplatin und Bestrahlung; 53: Capecitabin, Cisplatin und Bestrahlung)
Ma 2012	50	50 (36/14)	k. A.; k. A.; 41–74	Lokalisation: zervikal (6), obere (12) und mittlere Speiseröhre (32) Tumorgrad: T1 (2), T2 (7), T3 (29), T4 (12) Lymphknotenstatus: N0 (45), N1 (5)	Keine Angaben zur Vorbehandlung. Nach Studienbeginn: Strahlentherapie, welche nach dem Ergebnis der 1. PET-Untersuchung individuell festgelegt wurde; NAC (5- Fluorouracil 3 Tage/Woche).
Ott 2006	56	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 <sup>a</sup> : 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.
a: Siewert 1998 [17] AEG = adenocarcinomas of the esophagogastric junction; m = männlich; MW = Mittelwert; n = Anzahl; NAC = neoadjuvante Chemotherapie; OP = Operation; PET = Positronenemissionstomographie; w = weiblich					

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

In der Studie **Kim 2007** 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 der Studie **Ma 2012** fehlten Angaben zur Auswahl der Patienten und zu den Auswertern des Indextests. Patienten wurden in unterschiedlichen Krankheitsstadien eingeschlossen, was eine Interpretation der Ergebnisse erschwerte. Zudem wurde die Anzahl der Ereignisse nicht berichtet. Angesichts der Anzahl der eingeschlossenen Patienten (50) und betrachteten Faktoren (6) wurde davon ausgegangen, dass das Modell nur eine begrenzte Aussagekraft hatte (Überparametrisierung).

In **Ott 2006** wurden einige potenziell relevante Faktoren, wie Alter und Art der Behandlung, in der multifaktoriellen Analyse nicht berücksichtigt.

Tabelle 41: 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
Ma 2012	-	-	-	+	?	?	?	-	?	?	+	+	?	+	?	+	-	hoch
Ott 2006	+	+	+	+	+	+	+	+	+	?	+	+	+	-	?	+	-	hoch
+ = Kriterium erfüllt; - = Kriterium nicht erfüllt; ? = unklar F1: Beschreibung Selektionskriterien; F2: Vergleichbarkeit der Krankheitsphase der Patientinnen; F3: Repräsentativität Studienpopulation; F4: Nachbeobachtungszeit adäquat; F5: Patientenfluss nachvollziehbar; F6: Rate der Teilnehmer an Nachbeobachtung adäquat; F7: Definition prognostischer Faktoren; F8: Messung prognostischer Faktoren adäquat; F9: prognostische Faktoren für ausreichenden Teil der Patientinnen vorhanden; F10: Teilnehmer bzgl. prognostischer Information verblindet; F11: Definition Endpunkte; F12: Messung Endpunkte adäquat; F13: Festlegung Endpunkte verblindet bezüglich prognostischer Faktoren; F14: Berücksichtigung von Störgrößen; F15: Behandlungen während Nachbeobachtungszeit standardisiert oder randomisiert; F16: ergebnisunabhängige Berichterstattung; F17: keine sonstigen Aspekte																		

### 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 die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext evaluiert. In allen 3 Studien fanden sich Indizien für eine Nichtübertragbarkeit auf den deutschen Versorgungskontext.

Bei Ott 2006 wurden die Ergebnisse der Tests ohne Kenntnis vorhergehender klinischer Untersuchungen und ohne Kenntnis der Ergebnisse weiterer Tests beurteilt (siehe Tabelle 61).

In den Studien Ma 2012 und Kim 2007 wurden ausschließlich asiatische Patienten untersucht, was aufgrund der bekannten tumorbiologischen Unterschiede zwischen asiatischen und europäischen Patienten die Übertragbarkeit der Studienergebnisse einschränken könnte, laut Stellungnahme der DGN. In der Studie Ma 2012 wurde ebenfalls kein Cut-off genannt, sodass die Reproduzierbarkeit des Index- und Referenztests damit eingeschränkt ist.

Genannte Aspekte aus den Studien Kim 2007, Ott 2006 und Ma 2012 entsprachen möglicherweise nicht dem klinischen Alltag. Damit gibt es Indizien für eine Nichtübertragbarkeit.

Tabelle 42: Indizien 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	+	-	-	-	-	-	-	Indiz für eine Nichtübertragbarkeit
Ma 2012	+	+	-	-	-	-	-	Indiz für eine Nichtübertragbarkeit
Ott 2006	-	-	-	+	-	-	-	Indiz für eine Nichtübertragbarkeit
+ = konkretes Indiz für eine Nichtübertragbarkeit lag vor - = kein Indiz für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden) F1: Repräsentativität Patientenspektrum, F2: Übertragbarkeit Index- und Referenztest, F3: übertragbarer Entscheidungsprozess bei Testinterpretation, F4: übertragbare Hintergrundinformation, F5: technologische Entwicklung, F6: Untersucherqualifikation, F7: Qualität der Geräte								

### 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 Tabelle 43 nur die PET-relevanten Merkmale dargestellt werden. Alle weiteren Merkmale werden im Anhang E präsentiert.

In der Studie **Kim 2007** 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.

In der Studie **Ma 2012** wurden die Zielgrößen Gesamtüberleben und lokale Kontrollrate untersucht. Für das Gesamtüberleben und für die Zielgröße lokale Kontrollrate wurden jeweils 2 PET-relevante Merkmale in das multifaktorielle Modell integriert. Für beide Zielgrößen fehlten Angaben zu den Ergebnissen der einfaktoriellen Analyse. In der multifaktoriellen Analyse war für die Zielgröße Gesamtüberleben ein PET-relevantes Merkmal statistisch signifikant. Für die Zielgröße lokale Kontrollrate waren beide Merkmale statistisch signifikant. Außerdem war für den Endpunkt Gesamtüberleben anhand der Angaben im Text und der Ergebnistabelle nicht klar zu erkennen, welches statistische Modell für die multifaktorielle Analyse verwendet wurde. Für das Gesamtüberleben bei **Ott 2006** 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 unterscheiden sich die Ergebnisse für die uni- und multifaktorielle Analyse nicht.

Tabelle 43: Ergebnisse der Studien zur prognostischen Güte der PET beim Restaging

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	
Kim 2007	Krankheitsfreies Überleben	62 <sup>a</sup>	SUV-Aufnahme	Referenz: < 80 %	k. A. <sup>b</sup>	0,016	k. A.	0,025	0,25	
			Komplettes metabolisches Ansprechen	Referenz: Ja		k. A.	k. A.	0,006	3,58	
	Gesamtüberleben	62 <sup>a</sup>	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		
Ma 2012	Gesamtüberleben	50	Metabolisches Ansprechen (Reduktion von SUV <sub>max</sub> im Zentrum des Primärtumors zwischen Erst- und Interimsauswertung)	k. A.	k. A.	k. A.	k. A.	0,001 <sup>c</sup>	k. A.	
			Metabolisches Ansprechen (Reduktion von SUV <sub>max</sub> zwischen Erst- und Interimsauswertung)			k. A.	k. A.	0,503	k. A.	
			Lokale Kontrollrate	Metabolisches Ansprechen (Reduktion von SUV <sub>max</sub> im Zentrum des Primärtumors zwischen Erst- und Interimsauswertung)			k. A.	k. A.	< 0,001	k. A.
			Metabolisches Ansprechen (Reduktion von SUV <sub>max</sub> zwischen Erst- und Interimsauswertung) <sup>d</sup>			k. A.	k. A.	0,03	k. A.	

(Fortsetzung)

Tabelle 43: Ergebnisse der Studien zur prognostischen Güte der PET beim Restaging (Fortsetzung)

Studie	Zielgröße	n (N)	Prognostisches Merkmal	Cut-off	Patienten mit prog. Merkmal	Ergebnisse aus einfaktorierter Analyse		Ergebnisse aus multifaktorieller Analyse	
						p-Wert	HR	p-Wert	HR
Ott 2006	Gesamtüberleben	50 <sup>e</sup>	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 <sup>g</sup>	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: verwendetes statistisches Modell für die Schätzung unklar

d: Vorhersage der lokalen Kontrollrate unter „Total radiation therapy“; unklare Operationalisierung.

e: Außerdem einfaktorierter modelliert, aber p-Wert > 0,1: baseline FDG uptake, FDG uptake at day 14, R0-category.

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

FDG = <sup>18</sup>Fluorodeoxyglukose; HR = Hazard Ratio; KI = Konfidenzintervall; LK = Lymphknoten; PET= Positronenemissionstomographie; SUV = standardised uptake value

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

Die diagnostische als auch prognostische Kohortenstudie **Higuchi 2008** 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 44: Design, primäres Studienziel und untersuchte Zielgröße der Studie zur „prognostic accuracy“ beim Restaging

Studie	Studiendesign	Studienziel <sup>a</sup>	Zielkriterium	Evidenzstufe <sup>b</sup>	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 <sup>c</sup>	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 = <sup>18</sup>Fluordeoxyglukose; PET = Positronenemissionstomographie</p>					

Tabelle 45: 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 = <sup>18</sup>Fluordeoxyglukose; GE = General Electric; i.v. = intravenös; MBq = Mega Becquerel; PET = Positronenemissionstomographie; SUV = standardized uptake value</p>		

Tabelle 46: 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 **Higuchi 2008** 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 47: 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 F1: Beschreibung des Referenztests; F2: Zeitspanne zwischen dem Index- und Referenztest; F3: Interdependenz der Tests; F4: partielle Verifikation; F5: differenzielle Verifikation; F6: Incorporation Bias; F7: Verblindung bei Interpretation der Testergebnisse; F8: Intention-to-diagnose- / Intention-to-screen-Analyse; F9: ergebnisunabhängige Berichterstattung; F10: sonstige Aspekte											

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

In der Studie Higuchi 2008 konnten Indizien für eine Nichtübertragbarkeit identifiziert werden. Wie unter 5.4.2.3 ausführlich erläutert, wurden die Ergebnisse als nicht übertragbar auf den deutschen Versorgungskontext angesehen (vgl. Tabelle 48).

Tabelle 48: Indizien 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	+	+	-	-	-	-	-	Indiz für Nichtübertragbarkeit
+ = konkretes Indiz für eine Nichtübertragbarkeit lag vor -: kein Indiz für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden) F1: Repräsentativität Patientenspektrum, F2: Übertragbarkeit Index- und Referenztest, F3: übertragbarer Entscheidungsprozess bei Testinterpretation, F4: übertragbare Hintergrundinformation, F5: technologische Entwicklung, F6: Untersucherqualifikation, F7: Qualität der Geräte								

### 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 Schätzwerte der Spezifität höher (70,8 % bis 87,5 %) lagen. Die Daten wurden kontrolliert und noch einmal selbst berechnet.

Tabelle 49: Ergebnisse der Studie zur „prognostic accuracy“ bei der Restaging

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

a: alle Angaben selbst berechnet  
b: Punktschätzungen und Konfidenzintervalle wurden selbst berechnet  
c: 50 Patienten wurden eingeschlossen. 2 Patienten wurden von der Analyse ausgeschlossen, da sie eine R2-Resektion bei fortgeschrittener metastasierter Erkrankung hatten.  
FDG = 18Fluorodeoxyglukose; PET = Positronenemissionstomographie; SUV = standardised 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 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 **Guo 2007** 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 **Teyton 2009** 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 50: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Studien bei der Rezidivdiagnostik

Studie	Studiendesign	Studienziel <sup>a</sup>	Zielkriterium	Evidenzstufe <sup>b</sup>	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 2008	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 Rezidiv-erkennung.	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 = <sup>18</sup>Fluorodeoxyglukose; M = Metastase(n); N = Nodus (Knoten); PET = Positronenemissionstomographie; T = Tumor</p>					

Tabelle 51: Index- und Referenztest der eingeschlossenen Studien bei der Rezidivdiagnostik

Studie	Indextest	Referenztest
Guo 2007	<b>PET/CT</b> (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 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 ( $SUV \geq 2,5$ ).	Histologie oder klinische Nachbeobachtung zur Erkennung von Rezidiven (lokal, regional und Fernmetastasen)
Roedl 2008	<b>PET/CT</b> (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. 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.	Auffällige Befunde wurden mit Biopsie abgeklärt. Negative Befunde wurden nachbeobachtet und währenddessen mit EUS untersucht. Keine weiteren Angaben.
Teyton 2009	<b>PET</b> (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. PET-Untersuchungen alle 6 Monate während der ersten 2 Jahre der Nachbeobachtung und ab dann jährlich.	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	<b>PET/CT</b> (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.	Histologie (keine weiteren Angaben) oder klinische Nachbeobachtung über mindestens 6 Monate (keine weiteren Angaben).
CT = Computertomographie; FDG = <sup>18</sup> Fluorodeoxyglukose; 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		

Tabelle 52: Eingeschlossene Patienten der eingeschlossenen Studien bei der Rezidivdiagnostik

Studie	Patienten [N]	Geschlecht [N (m/w)]	Alter	Diagnose	Vorbehandlung
Guo 2007	56 <sup>a</sup>	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 2008	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 **Teyton 2009** wurde mit einem niedrigen Verzerrungspotenzial bewertet. Die methodischen Charakteristika der Diagnosestudien werden in Tabelle 53 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 **Guo 2007** 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 **Roedl 2008** 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** 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 **Zhang 2006** 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 53: 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 2008	+	?	+	+	-	+	?	+	+	?	hoch
Teyton 2009	+	?	+	+	-	?	+	+	+	?	niedrig
Zhang 2006	?	?	+	+	?	?	?	?	+	?	hoch

+: Kriterium erfüllt; -: Kriterium nicht erfüllt; ?: unklar  
 F1: Beschreibung des Referenztests, F2: Zeitspanne zwischen dem Index- und Referenztest, F3: Interdependenz der Tests, F4: partielle Verifikation, F5: differenzielle Verifikation, F6: Incorporation Bias, F7: Verblindung bei Interpretation der Testergebnisse, F8: Intention-to-diagnose- / Intention-to-screen-Analyse, F9: ergebnisunabhängige Berichterstattung, F10: sonstige Aspekte

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

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

- In den Studien **Guo 2007** und **Zhang 2006** wurden ausschließlich asiatische Patienten untersucht, was aufgrund der bekannten tumorbiologischen Unterschiede zwischen asiatischen und europäischen Patienten die Übertragbarkeit der Studienergebnisse einschränken könnte, laut Stellungnahme der DGN.
- In **Zhang 2006** 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 **Guo 2007**, **Roedl 2008** und **Teyton 2009** 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 54 stellt die Ergebnisse der Einschätzung der Übertragbarkeit der Ergebnisse der Primärstudien auf den deutschen Versorgungskontext bei der Rezidivdiagnostik dar.

Tabelle 54: Indizien 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	+	-	+	-	-	-	-	Indiz für eine Nichtübertragbarkeit
Roedl 2008	-	-	+	-	-	-	-	Indiz für eine Nichtübertragbarkeit
Teyton 2009	-	-	+	-	-	-	-	Indiz für eine Nichtübertragbarkeit
Zhang 2006	+	+	-	-	-	-	-	Indiz für eine Nichtübertragbarkeit
+: konkretes Indiz für eine Nichtübertragbarkeit lag vor -: kein Indiz für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden) F1: Repräsentativität Patientenspektrum, F2: Übertragbarkeit Index- und Referenztest, F3: übertragbarer Entscheidungsprozess bei Testinterpretation, F4: übertragbare Hintergrundinformation, F5: technologische Entwicklung, F6: Untersucherqualifikation, F7: Qualität der Geräte								

### **5.5.3.3 Ergebnisse der Primärstudien zur diagnostischen Güte bei der Rezidivdiagnostik**

In Tabelle 55 werden die Ergebnisse zur diagnostischen Güte der Primärstudien bei der Rezidivdiagnostik dargestellt. Neben den jeweiligen Schätzungen wurden auch die Konfidenzintervalle berichtet. Sofern diese nicht in der Studie präsentiert wurden, wurden sie eigenständig nach der Methode von Clopper und Pearson 1934 [127] 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 2008 und Teyton 2009) untersucht, wohingegen die PET/CT in 3 Studien untersucht wurde (Guo 2007, Roedl 2008 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 Schätzungen der Sensitivität waren gering und die Konfidenzintervalle entsprechend weit. Außerdem waren die Konfidenzintervalle der Schätzungen der Spezifität sehr gering und die Konfidenzintervalle entsprechend sehr weit.

Tabelle 55: Ergebnisse der Studien zur diagnostischen Güte der PET bzw PET/CT bei der Rezidivdiagnostik

Studie	n (N)	Indextest	Cut-off	Referenztest	Cut-off	RP	FP	FN	RN	Sensitivität in % [95 %-KI] <sup>a</sup>	Spezifität in % [95 %- KI] <sup>a</sup>
Guo 2007	56 (56) <sup>b</sup>	FDG-PET/CT	Rezidiv pro Patient	Histologie oder klinische Nachbeobachtung (min. 6 Monate)	Jedes Rezidiv	43	5	2	6	95,6 [84,9; 99,5]	54,5 [23,4; 83,3]
Roedl 2008	47 (47)	FDG-PET	Rezidiv pro Patient	Biopsie, klinische Nachbeobachtung (Median 25 Monate), nach 18 Monaten: EUS	Jedes Rezidiv	24	9	3	11	89 [70,8; 97,6]	55 [31,5; 76,9]
		FDG-PET/CT				24	5	3	15	89 [70,8; 97,6]	75 [50,9; 91,3]
Teyton 2009	41 (41)	FDG-PET	Rezidiv pro Patient	Biopsie oder klinische Nachbeobachtung	Jedes Rezidiv	23	2 <sup>d</sup>	0	16 <sup>d</sup>	100 [85,2; 100]	88,9 [65,3; 98,6]
Zhang 2006	31 (31) <sup>c</sup>	FDG-PET/CT	Läsionen pro Patient mit SUV > 2,5	Histologie oder klinische Nachbeobachtung > 6 Monate	Jedes Rezidiv	23 <sup>d</sup>	1 <sup>d</sup>	0 <sup>d</sup>	6 <sup>d</sup>	100 [85,2; 100]	85,7 [42,1; 99,6]
	31 (31)	CT	k. A.	Histologie oder klinische Nachbeobachtung > 6 Monate	Jedes Rezidiv	- <sup>e</sup>	- <sup>e</sup>	- <sup>e</sup>	- <sup>e</sup>	- <sup>e</sup>	- <sup>e</sup>
<p>a: Punktschätzungen und Konfidenzintervalle selbst berechnet</p> <p>b: 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>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 krebsbedingt verstarb und daher nicht in der Vierfeldertafel eingeschlossen wurde.</p> <p>d: Werte selbst berechnet</p> <p>e: 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</p> <p>CT = Computertomographie; FDG = <sup>18</sup>Fluorodeoxyglukose; 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)</p>											

### 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. Im Forestplot 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.

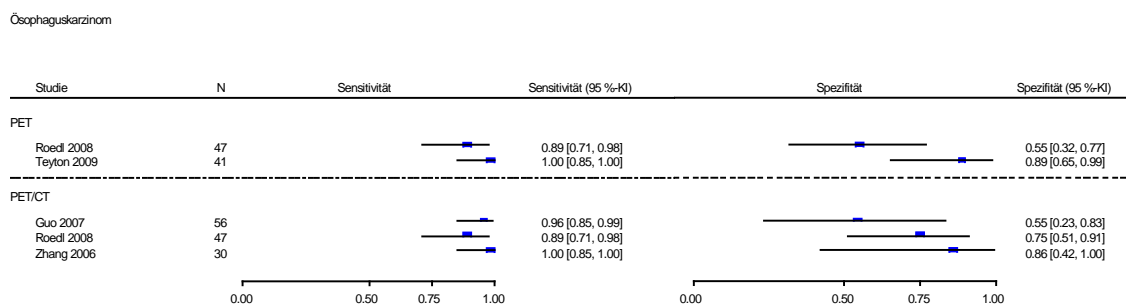


Abbildung 8: Forestplot der Ergebnisse für die diagnostische Güte der PET bei der Rezidiverkennung von Ösophaguskarzinomen

### 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 2008 [123] durchgeführt. Dieser zeigte, dass die PET und die PET/CT die gleichen Werte für Sensitivität erzielten (89 %), die Konfidenzintervalle für die Schätzwerte der Spezifität schließen jeweils den Schätzwert der anderen Technologie ein. Zhang 2006 untersuchte den direkten Vergleich zwischen der PET/CT und CT. Aus der Studie konnten allerdings nur die Punktschätzungen 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 **Teyton 2009** wurde eine Änderung im klinischen Management basierend auf dem Einsatz der PET (siehe Tabelle 56) 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 56: 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 = <sup>18</sup> Fluordeoxyglukose; 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 [122]) für die Indikation Rezidivdiagnostik identifiziert.

##### 5.5.4.1 Studiendesign und Studienpopulation der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik

Wie erwähnt, betrachtete die Arbeit **Guo 2007** 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 57: Design, primäres Studienziel und untersuchte Zielgröße der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik

Studie	Studiendesign	Studienziel <sup>a</sup>	Zielkriterium	Evidenzstufe <sup>b</sup>	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 <sup>c</sup>	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.                      CT = Computertomographie; FDG = <sup>18</sup>Fluorodeoxyglukose; PET = Positronenemissionstomographie;</p>					

Tabelle 58: Index- und Referenztest der eingeschlossenen Prognosestudie bei der Rezidivdiagnostik

Studie	Indextest	Referenztest
Guo 2007	<b>PET/CT</b> (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 ≥ 2,5).	Gesamtüberleben (Zeit von Behandlungsende bis Tod jeglicher Ursache). Mediane Nachbeobachtung 24,5 Monate (Spannweite 4 bis 34,5 Monate).
<p>CT = Computertomographie; FDG = <sup>18</sup>Fluorodeoxyglukose; GE = General Electric; i.v. = intravenös; mA = Milliampere; mAs = Milliamperesekunde; MBq = Mega Becquerel; PET = Positronenemissionstomographie; SUV = standardized uptake value; USA = Vereinigte Staaten von Amerika</p>		

Tabelle 59: 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 **Guo 2007** 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 60: 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 F1: Beschreibung Selektionskriterien; F2: Vergleichbarkeit der Krankheitsphase der Patientinnen; F3: Repräsentativität Studienpopulation; F4: Nachbeobachtungszeit adäquat; F5: Patientenfluss nachvollziehbar; F6: Rate der Teilnehmer an Nachbeobachtung adäquat; F7: Definition prognostischer Faktoren; F8: Messung prognostischer Faktoren adäquat; F9: prognostische Faktoren für ausreichenden Teil der Patientinnen vorhanden; F10: Teilnehmer bzgl. prognostischer Information verblindet; F11: Definition Endpunkte; F12: Messung Endpunkte adäquat; F13: Festlegung Endpunkte verblindet bezüglich prognostischer Faktoren; F14: Berücksichtigung von Störgrößen; F15: Behandlungen während Nachbeobachtungszeit standardisiert oder randomisiert; F16: ergebnisunabhängige Berichterstattung; F17: keine sonstigen Aspekte																		

### 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 61).

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

Tabelle 61: Indizien 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	-	-	+	-	-	-	-	Indiz für eine Nichtübertragbarkeit
+: konkretes Indiz für eine Nichtübertragbarkeit lag vor -: kein Indiz für eine fehlende Übertragbarkeit (sie kann daher als gegeben angesehen werden) F1: Repräsentativität Patientenspektrum, F2: Übertragbarkeit Index- und Referenztest, F3: übertragbarer Entscheidungsprozess bei Testinterpretation, F4: übertragbare Hintergrundinformation, F5: technologische Entwicklung, F6: Untersucherqualifikation, F7: Qualität der Geräte								

### 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 62: 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 Primärstagings, des Restagings und der Rezidivdiagnostik ermöglicht hätte. Auch die Suche in Kongressbänden lieferte keine Hinweise auf noch laufende vergleichende Studien. In den Studienregistern konnten 4 laufende Studien identifiziert werden.

### *Diagnostische und prognostische Güte*

3 Evidenzsynthesen und 37 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 48 Primärstudien: Primärstaging (n = 25), Restaging (n = 19) und Rezidivdiagnostik (n = 4).

#### Primärstaging

Für das N-Staging wurden direkte Vergleiche zwischen PET und CT in 12 Studien untersucht. In den 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 7 Studien untersucht. Die Ergebnisse dieser Meta-Analyse wurden nicht dargestellt, da die Daten dieser 7 Studien nicht geeignet waren, um präzise Schätzungen zu berechnen. Auch in Sensitivitätsanalysen ergaben sich keine präzisen Schätzungen.

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

Für das Restaging nach Abschluss einer neoadjuvanten Therapie wurden in 3 Studien direkte Vergleiche durchgeführt:

Der Vergleich PET/CT versus Endosonographie (EUS) wurde in der Studie von Cerfolio 2005 untersucht. Es zeigte sich für die PET/CT eine Sensitivität von 86,7 %, (95 %-KI [59,7;98,3]) und eine Spezifität von 87,9 %, (95 %-KI [71,8; 96,6]). Im Gegensatz dazu hatte die EUS eine Sensitivität von 20 %, (95 %-KI [4;48]) und eine Spezifität von 94 %, (95 %-KI [79,8;99,3]). In der Studie von Roedl 2009 wurde PET versus PET/CT untersucht. Hier zeigte sich für die PET eine Sensitivität von 59 %, (95 %-KI [36,4; 79,3]) und eine Spezifität von 100 %, (95 %-KI [87,2;100]). Die PET/CT erreichte eine Sensitivität von 68 %, (95 %-KI [45,1;86,1]) und eine Spezifität von 100 %, (95 %-KI [87,2;100]). In der Studie von Kroep



2003 wurde die PET mit der EUS und der CT basierend auf 10 bzw. 11 eingeschlossenen Patienten verglichen. Für die PET zeigte sich eine Sensitivität von 100 %, (95 %-KI [39,8;100]) und eine Spezifität von 100 % (95 %-KI [54,1;100]). Die Vergleichstechnologie EUS hatte ebenfalls eine Sensitivität von 100 % (95 %-KI [47,8;100]) und Spezifität von 100 % (95 %-KI [54,1;100]). Im Gegensatz dazu hatte die CT eine Sensitivität von 50 % (95 %-KI [6,8;93,2]) und eine Spezifität von 71 % (95 %-KI [29,0;96,3]).

### Interimsstaging

Für das Interimsstaging lag 1 direkter Vergleich diagnostischer Technologien vor:

In der Studie von Kroep 2003 wurde die PET mit der EUS und der CT basierend auf 11 eingeschlossenen Patienten verglichen. Die PET hatte eine Sensitivität von 100 % (95 %-KI [39,8;100]) und eine Spezifität von 85,7 % (95 %-KI [42,1;99,6]). Die Vergleichstechnologie EUS hatte eine Sensitivität von 100 % (95 %-KI [47,8;100]) und Spezifität von 100 % (95 %-KI [54,1;100]). Im Gegensatz dazu hatte die CT eine Sensitivität von 50 % (95 %-KI [6,8;93,2]) und eine Spezifität von 71 % (95 %-KI [29,0;96,3]).

Da es in der Indikation Interimsstaging während einer neoadjuvanten Chemotherapie bzw. Chemoradiotherapie nur eine Studie gibt, die einen direkten Vergleich auf Basis von nur 11 Patienten untersucht, werden hier auch ergänzend die Ergebnisse der nicht-vergleichenden Studien dargestellt. In den 7 Studien lagen die Punktschätzungen der Sensitivität für die PET bzw. PET/CT im Bereich von 44 % (zugehörige Spezifität 52 %) bis 88,9 % (zugehörige Spezifität 61,4 %). Die Punktschätzungen der Spezifität lagen im Bereich von 52 % (zugehörige Sensitivität 44 %) bis 78 % (zugehörige Sensitivität 80 %). Für alle Punktschätzungen waren weite Konfidenzintervalle zu beobachten.

### Rezidivdiagnostik

Im Rahmen der Ergänzungsrecherche wurden 4 Primärstudien eingeschlossen. Alle 4 Studien untersuchten die diagnostische Güte der PET bzw. PET/CT. Eine Studie untersuchte zusätzlich 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 [87,96,103,124]. Diese Problematik bestand beispielsweise auch in den MUNICON-I- und II-Studien [129,130], 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 [131].

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

Für die Indikation Primärstaging wurden 25 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 [29] und van Vliet 2008 [132] für die Indikation identifiziert werden, sie wurden aber nicht eingeschlossen, da beide Mängel auf dem Qualitätsindex von Oxman und Guyatt 1991 [49,50] 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 [29], der wie die vorliegende Arbeit nur nach Studien recherchierte, die den Einsatz der PET bzw. PET/CT untersuchten, suchte van Vliet 2008 [132] 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 [132] 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 [89,91,106,109] in die bivariate

Meta-Analyse des vorliegenden Abschlussberichts eingeschlossen, weil in diesen Studien ein direkter Vergleich zwischen der PET und einer anderen Technologie durchgeführt wurde. Wie bereits in Abschnitt 5.3.2.4 erläutert, wurden die Ergebnisse der Meta-Analyse für das M-Staging im Abschlussbericht nicht dargestellt, da die Daten nicht geeignet waren, um präzise Schätzwerte zu berechnen. Auch in Sensitivitätsanalysen ergaben sich keine präzisen Schätzungen. Deshalb wurden in Tabelle 63 nur die Ergebnisse für das N-Staging präsentiert. Für das N-Staging poolten van Vliet et al. 2008 10 Studien, von denen ebenfalls 6 den Einschlusskriterien des vorliegenden Berichts entsprachen. Fünf [70,91,102,106,109] 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 [29] poolten 12 Studien, die N-Staging und ebenfalls M-Staging untersuchten. Von diesen wurden 7 für das N-Staging in den vorliegenden Bericht eingeschlossen.

Vergleicht man die gepoolten Ergebnisse der PET des vorliegenden Abschlussberichts beim N-Staging mit den gepoolten Ergebnissen der 2 Evidenzsynthesen (vgl. Tabelle 63), so kann festgestellt werden, dass die gepoolten Schätzungen 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 Schätzwerte für die CT beim N-Staging aus dem vorliegenden Bericht mit denen aus van Vliet 2008 [132] 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 gepoolten Ergebnisse für das N-Staging des vorliegenden Berichts werden in Tabelle 63 dargestellt. Der Unterschied der Konfidenzregionen wurde mithilfe des Likelihood-Ratio-Tests untersucht und ergab einen p-Wert von  $p = 0,370$ . Es kann daher keine Aussage darüber getroffen werden, ob sich die Verfahren in der erwarteten Sensitivität und Spezifität unterscheiden.

Tabelle 63: Ergebnisse der Evidenzsynthesen Van Westreenen 2004, Van Vliet 2008 und des IQWiG-Abschlussberichts beim N-Staging bei Ösophaguskarzinomen

		Sensitivität in % [95 %-KI]	Spezifität in % [95 %-KI]
<i>N-Staging</i>			
<b>IQWiG-Abschlussbericht<sup>a</sup></b>	<b>PET</b>	<b>59 [48; 69]</b>	<b>86 [78; 91]</b>
	<b>CT</b>	<b>49 [38; 60]</b>	<b>87 [79; 92]</b>
Van Westreenen 2004	PET	51 [34;69]	84 [76;91]
Van Vliet 2008	PET	57 [43;70]	85 [76;95]
	CT	50[41;60]	83 [77;89]
	EUS	80 [75;84]	70 [65;75]
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. Erwähnenswert ist in diesem Zusammenhang, dass die Mehrzahl der Managementänderungen (siehe Tabelle 22) sich nicht aus der eigentlich beabsichtigten Erkennung von Fernmetastasen, sondern der Entdeckung von Zweitmalignomen ergab, sodass diese Ergebnisse im Grunde die Fragestellung des vorliegenden Berichts teilweise verlassen.

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

Für die Indikation Restaging konnten insgesamt 19 Studien eingeschlossen werden. Von diesen stammen 11 Primärstudien aus den eingeschlossenen Evidenzsynthesen [67,76,77]. 8 Studien wurden im Rahmen der Ergänzungsrecherche gesichtet und eingeschlossen. 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.

Eine meta-analytische Zusammenfassung war nicht möglich, das zu wenige direkte Vergleiche in den Studien durchgeführt wurden. Darüber hinaus verwendeten die eingeschlossenen Studien eine Vielzahl an unterschiedlichen Klassifikationen zum histologischen Ansprechen. Die Verwendung der verschiedenen Klassifikationen als Referenzstandards erschwerte eine Vergleichbarkeit der Studien.

Alle 3 eingeschlossenen Evidenzsynthesen hatten eine hohe Qualität nach Oxman und Guyatt 1991 [49,50]. Zwei der eingeschlossenen Evidenzsynthesen [67,77] verwendeten das QUADAS-Instrument, eine Evidenzsynthese [76] 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 Rebollo Aquirre 2009 [76] untersuchte nur die diagnostische Güte der FDG-PET beim Restaging. Im Gegensatz dazu verfolgte die Evidenzsynthese Westerterp 2005 [77] das Ziel, den Einsatz der CT, EUS und FDG-PET beim Restaging von Ösophaguskarzinomen zu analysieren. Ngamruengphong 2010 [67] 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. Im Rahmen der Nachrecherche zum Abschlussbericht konnten für die Indikation Restaging die 3 Evidenzsynthesen Chen 2011, Kwee 2010 und Zhu 2012 [133-135] identifiziert werden. Aus diesen ergab sich keine zusätzliche relevante Primärstudie, die nicht im Rahmen der Ergänzungsrecherche ohnehin identifiziert worden war. Die

Evidenzsynthese Zhu 2012 [135] würde aufgrund des Ausschlusskriteriums E4 den Einschlusskriterien des vorliegenden Berichts nicht entsprechen. Die Evidenzsynthesen Kwee 2010 und Chen 2011 würden den Einschlusskriterien entsprechen [133,134], weisen jedoch gemäß dem Qualitätsindex von Oxman und Guyatt 1991 [49,50] Mängel auf. In der folgenden Tabelle 64 werden die Schlussfolgerungen der Autoren der 3 eingeschlossenen Evidenzsynthesen und der 2 im Rahmen der Nachrecherche identifizierten Evidenzsynthesen zusammengefasst.

Tabelle 64: Schlussfolgerungen der eingeschlossenen Evidenzsynthesen beim Restaging von Ösophaguskarzinomen

<b>Evidenzsynthese</b>	<b>Schlussfolgerungen zur Indikation des Restagings bei Ösophaguskarzinomen<sup>a</sup></b>
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.
Kwee 2010	Die eingeschlossenen Studien hatten eine moderate methodische Qualität und ihre Ergebnisse waren sehr heterogen. Studien, welche außerhalb der USA durchgeführt wurden, und Studien, welche eine höhere methodische Qualität hatten, zeigten eine höhere Genauigkeit. Auf Basis der vorliegenden Evidenz sollte die FDG-PET noch nicht im routinemäßigen klinischen Umfeld für die Entscheidung einer neoadjuvanten Therapie bei Patienten mit Ösophaguskarzinomen eingesetzt werden. Auf der Basis einer zu standardisierenden PET-Diagnostik sind gut geplante multizentrische Studien notwendig.
Chen 2011	Die FDG-PET hat einen Stellenwert in der Beurteilung des Ansprechens einer neoadjuvanten Therapie bei Patienten mit einem Ösophaguskarzinom. Eine 50%ige Reduktion des SUV zwischen der Prä- und Posttherapie, wenn die PET 2 Wochen nach Beginn der Therapie durchgeführt wird, ist ein optimales Kriterium für die Vorhersage des Behandlungsansprechens einer neoadjuvanten Therapie.
a: Schlussfolgerungen gemäß Formulierung der Autoren	

Insgesamt äußerten sich die Autoren der eingeschlossenen Evidenzsynthesen eher zurückhaltend. Nur die Evidenzsynthese Chen 2011 [133] empfahl den Einsatz der PET. Alle weiteren Autorengruppen gaben keine ausdrückliche Empfehlung zum Einsatz der PET bei Ösophaguskarzinomen ab. In 2 Evidenzsynthesen wurden weitere qualitativ hochwertige

Studien mit höheren Patientenzahlen gefordert. In der Evidenzsynthese Kwee 2010 [134] wurde die Durchführung einer multizentrischen Studie gefordert.

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

In der Indikation des Interimsstaging bei neoadjuvanter Chemotherapie (NAC) und neoadjuvanter Chemoradiotherapie (NACR)<sup>2</sup> 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 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 Weber 2001 und Ott 2006 untersuchten mit einem sehr homogenen Patienten- und Therapiespektrum den Einsatz der PET während einer NAC an Patienten mit Adenokarzinomen des gastroösophagealen 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 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

---

<sup>2</sup> NAC oder NACR wird im Folgenden mit NAC(R) abgekürzt

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 Ott 2006 wurden 65 Patienten eingeschlossen, von denen 56 in die Analyse gingen. 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 % (95 %-KI [21,5;67,4]). Die Nonresponder erzielten dagegen eine histologische Responserate nach Abschluss der NAC von 5 % (95 %-KI [0;12,4]).

In der Studie 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 95 % (95 %-KI [82;99] für Ott 2006) und bei 96 % (95 %-KI [77;100] für Weber 2001).

Es wurden weitere Studien in diesen Bericht eingeschlossen, die die PET im Interimsstaging untersuchten. Nicht alle diese Studien beschränkten sich auf AEG I+II. Neben Weber 2001 und Ott 2006 untersuchte auch die Studie Kroep 2003 den Einsatz der PET während einer NAC. Die Studien Gillham 2006, Westerterp 2006, Wieder 2004, Malik 2010 und van Heijl 2011 untersuchten den Einsatz der PET während einer NACR. In allen Studien bis auf van Heijl 2011 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.

Die Studie Ilson 2011 untersuchte den Einsatz der PET während einer NAC und einer anschließenden NACR im Interimsstaging bei einem Cut-off von > 35 %. Keine der Studien war randomisiert. Alle Patienten wurden unabhängig vom PET-Ergebnis weiter mit einer NAC(R) behandelt.

Die Studie Kroep 2003 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  $100\%$  (95 %-KI [40;100]).

Die Studie Gillham 2006 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  $76\%$  (95 %-KI [53;92]).

In der Studie Westerterp 2006 wurde angegeben, dass 26 Patienten mit einem Adenokarzinom oder Plattenepithelkarzinom des Ösophagus mit  $\leq 2$  cm Involvierung des Magens konsekutiv eingeschlossen wurden. Nach Abschluss einer 5-wöchigen NACR hatten 9 von 17 Patienten kein relevantes histopathologisches Ansprechen. Der NPV (bei  $> 35\%$ ) lag bei  $64\%$  95 %-KI [31;89].

In der Studie Wieder 2004 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  $78\%$  (95 %-KI [40;97]).

In der Studie Malik 2010 wurden 37 Patienten mit einem Adenokarzinom der Speiseröhre oder des gastroösophagealen Übergangs eingeschlossen. Nach Abschluss der NACR hatten 21 von 37 Patienten kein Ansprechen in der Histologie. Der NPV (bei  $> 35\%$ ) lag in dieser Studie bei  $64\%$  (95 %-KI [43;82]).

Die Studie Ilson 2011 schloss 60 Patienten mit einem Adenokarzinom oder Plattenepithelkarzinom oder einem Tumor des gastroösophagealen Übergangs ein. Ausgewertet wurden schließlich nur 53 Patienten. Die Gründe für den Ausschluss aus der Analyse wurden in Tabelle 37 erläutert. Ziel der Studie war die Bestimmung des pathologisch vollständigen Ansprechens auf eine Chemoradiotherapie und Induktionschemotherapie (NAC/R) mit eingebetteter Evaluation der diagnostischen Güte der PET. Nach Abschluss der NAC/R hatten 18 von 53 ausgewerteten Patienten kein relevantes Ansprechen in der Histologie. Der NPV (bei  $> 35\%$ ) lag bei  $96\%$  (95 %-KI [82;100]).

Zusammenfassend kann also festgestellt werden, dass in den 3 Studien zur diagnostischen Güte der PET beim Interimsstaging einer NAC die Punktschätzungen der NPV bei  $95\%$  (38 Patienten mit negativem PET-Ergebnis) und  $100\%$  (4 Patienten mit negativem PET-Ergebnis) und jene der 4 Studien zur NACR zwischen  $64\%$ <sup>3</sup> [85] und  $64\%$  [117] (11 [bei

---

<sup>3</sup> In den Studien Westerterp 2006 und Malik 2010 lag der NPV bei jeweils  $64\%$ . Das 95 %-KI bei Westerterp 2006 war [31;89], jenes bei Malik 2010 war [43;82]).

Westerterp 2006] bzw. 25 Patienten [bei Malik 2010] mit negativem PET-Ergebnis) und 96 % (28 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

- Die 2 aufeinander aufbauenden Studien Weber 1999 und Weber 2001 mit einem sehr homogenen Patienten- und Therapiespektrum untersucht wurden, aus denen ein Grenzwert von  $> 35\%$  hervorging, der in Ott 2006 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 für die NAC bzw. Gillham 2006, Westerterp 2006, Wieder 2004 und Malik 2010 für die NACR),
- 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 [136].

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 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 [18]. 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 [137]. 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 [137]. 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 Weber 1999 und 2001 [84,138], Ott 2006 [74] und MUNICON 2007 [129] möglich, einen indirekten Vergleich nach Bucher 1997 [139] 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 Weber 2001 und Ott 2006, 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 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 [140,141]. In den gesichteten Studien scheint die PET in der Lage zu sein, Ansprecher von Nichtansprechern zu einem frühen Zeitpunkt zu unterscheiden (Weber 2001, Ott 2006, MUNICON I [74,84,129]). 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 [39,142].

### **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 [143]. 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 [144,145]. Durch den Einsatz der PET erhofft man sich, diese Hindernisse zu überwinden [146].

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 [27]. Leitlinien, die gefunden

wurden, waren oftmals älter und die PET bzw. PET/CT fand darin keine Erwähnung [9]. In der internationalen Leitlinie des NCCN wurde der Einsatz der PET und der PET/CT für das M-Staging empfohlen [28]. 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 2004 [29], welche nicht in dem vorliegenden Bericht eingeschlossen wurde, weil die Primärstudien nicht den Einschlusskriterien des Abschlussberichts entsprachen. Besonders für das N-Staging wurde der Einsatz der PET basierend auf der Studie von Flamen 2000 [89] empfohlen. Die genannte Studie wurde im Rahmen des Berichts für die Indikation Primärstaging eingeschlossen. Wie bereits erwähnt, gab es zur Studie Flamen 2000, in welcher das N- und M-Staging untersucht wurde, eine zweite Publikation von Lerut 2000 [90]. Das N-Staging wurde bei 39 Patienten durchgeführt. Die Daten der Vierfeldertafel für das N- und das M1a-Staging 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 offen empfohlen [26]. Die Empfehlung basierte auf einer nichtsystematischen Leitlinie [147].

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 [148]. 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 [149], Levine 2006 [82], Weber 2001 [84] und Flamen 2001 [79]. Von diesen 4 Studien wurden 3 Studien in den vorliegenden Bericht eingeschlossen [79,82,84]. Die Studie Swisher 2004 [149] entsprach nicht den Einschlusskriterien des Abschlussberichts, 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 [78] hin, in der die PET/CT eine höhere diagnostische Güte als der Ultraschall und die CT hatte. Genannte Studie wurde ebenfalls in den Abschlussbericht 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 [11]. Die Empfehlung basierte auf

der Evidenzsynthese von van Westreenen 2004, die wie erwähnt nicht in den vorliegenden Bericht eingeschlossen werden konnte [29].

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 [10]. 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 [132] und Facey 2007 [150]. Die Evidenzsynthese von van Vliet 2008 wurde nicht im Abschlussbericht eingeschlossen, weil diese nur 3 von 7 Punkten auf dem Qualitätsindex von Oxman und Guyatt 1991 [49,50] erreichte. Die Ergebnisse der Untersuchung wurden allerdings im Abschnitt 6.1.2 diskutiert. Die Arbeit von Facey 2007 [150] wurde im Rahmen des Abschlussberichts 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 [18]. Darüber hinaus könnte die PET in der Kontrolle des Therapieansprechens eine Rolle spielen [18]. Worauf diese Empfehlung basierte, wurde in der Leitlinie nicht deutlich.

In der 2011 publizierten Leitlinie der kanadischen Saskatchewan Cancer Agency [30] wurde der Einsatz der PET im Falle einer multimodalen Therapie für das T- und M-Staging zusammen mit 4 weiteren diagnostischen Verfahren empfohlen. Diese Empfehlung basiert auf Expertenmeinungen.

Von der Deutschen Gesellschaft für Nuklearmedizin e. V., der Deutschen Röntgen-gesellschaft 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 [31]. 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 [68,89,90,101,151,152]. Vier dieser Arbeiten [68,89,90,101] wurden auch in den vorliegenden Abschlussbericht eingeschlossen. Allerdings handelte es sich bei den Arbeiten von Flamen 2000 und Lerut 2000 [89,90] um dieselbe Studie. Im Abschlussbericht wurden nur die Ergebnisse für das N-Staging (aus Flamen 2000 und Lerut 2000 [89,90]) sowie die Ergebnisse des M1a-Stagings aus der Publikation von Lerut 2000 dargestellt. Die Vierfeldertafeln für das M-Staging in der Publikation von Flamen 2000 konnten nicht verwendet werden, weil die Autoren für das M-Staging keinen, nach Definition des Abschlussberichts, validen Referenztest verwendeten. Die Studie von Brücher 2001 [68] wurde in den Bericht zwar eingeschlossen, jedoch untersuchte sie nicht das Primärstaging,

sondern das Therapieansprechen (Restaging). Die Studien Block 1997 und Flanagan 1997 [151,152] 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) [18]. 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 [31]. Die Cancer-Care-Ontario-Gesellschaft unterstützte diese Ansicht, indem sie erläuterten, dass sie bislang nur 2 Studien [122,153] zum Einsatz der PET bei der Rezidivdiagnostik identifizieren konnten und hieraus keine Empfehlung ableiten konnten [154]. Eine dieser Studien (Guo 2007 [122]) wurde auch im Rahmen des vorliegenden Berichts eingeschlossen. Die zweite Publikation von Jingu 2010 konnte nicht im Abschlussbericht eingeschlossen werden, weil in der Studie kein multifaktorielles Modell berechnet wurde (E3).

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) [46]. Insgesamt wäre daher eine präzisere Beschreibung des Studiendesigns für das Verständnis, aber auch die weitere quantitative Zusammenfassung zukünftiger Studien sehr hilfreich.

## 6.5 Würdigung der Anhörung zum Vorbericht

Insgesamt wurden 2 Stellungnahmen zum Vorbericht frist- und formgerecht eingereicht.

Die im Rahmen der Anhörung vorgebrachten Aspekte wurden hinsichtlich valider wissenschaftlicher Argumente für eine Änderung des Vorberichts überprüft. Die wesentlichen Argumente werden im Folgenden diskutiert. Neben projektspezifischen wissenschaftlichen Aspekten wurden auch übergeordnete Punkte, z. B. zu rechtlichen Vorgaben für das Institut, angesprochen. Auf solche Punkte wird im Rahmen dieser projektspezifischen Würdigung der Anhörung nicht weiter eingegangen. Verschiedene Stellungnahmen bezogen sich auf Punkte der projektspezifischen Methodik, die bereits im Rahmen der Anhörung zum Berichtsplan ausführlich diskutiert wurden (siehe Dokumentation und Würdigung der Anhörung zum Berichtsplan D06-01E-K) [155]. Auf diese wird nicht erneut eingegangen.

Die 2 Stellungnahmen werden in den nachfolgenden Abschnitten gewürdigt.

Die Zusammenfassung aller Änderungen des Abschlussberichts gegenüber dem Vorbericht, die sich unter anderem durch die Anhörung zum Vorbericht ergeben haben, ist in Abschnitt 3.2 dargestellt.

### 6.5.1 Anmerkungen zum Hintergrundkapitel

*In einer Stellungnahme wurde darauf hingewiesen, dass die Leitlinie der Deutschen Krebsgesellschaft und der Deutschen Gesellschaft für Chirurgie „Ösophaguskarzinom“ (AWMF-Registernr. 032 / 008, Stand: 10/2003) entgegen der Behauptung im Vorbericht sehr wohl verfügbar sei.*

Bedauerlicherweise lag der Stellungnahme keine Kopie der erwähnten Leitlinie bei. Die erwähnte Leitlinie konnte weder im Rahmen des Vorberichts noch im Rahmen des Abschlussberichts gefunden werden. Damit war die Feststellung im Vorbericht, dass die Leitlinie nicht verfügbar war, korrekt.



Es ergab sich daher keine Änderung im Abschlussbericht.

*In einer Stellungnahme wurde angemerkt, dass die Darstellung der verfügbaren Leitlinien und Empfehlungen der wissenschaftlichen Fachgesellschaften große Mängel aufweise oder nicht aktuell sei. Darüber hinaus wurde darauf hingewiesen, dass die aktuelle AWMF-S3-Leitlinie zu Adenokarzinomen des Magens und des ösophagogastralen Übergangs im Vorbericht nicht erwähnt worden sei.*

Grundsätzlich ist festzuhalten, dass der vorliegende Bericht nicht den Anspruch erhebt, vorhandene Leitlinien unabhängig von ihrer Methodik systematisch zu erfassen. Sofern eine Leitlinie die vorhandene wissenschaftliche Literatur systematisch recherchiert, bewertet, zusammengefasst und transparent dargestellt hat und somit den festgelegten Einschlusskriterien für systematische Übersichten genügt, wird sie regelhaft eingeschlossen und gewürdigt. Leitlinien, die diesen Kriterien (vgl. 4.2.1) nicht genügen, können nur punktuell im Rahmen des Hintergrundkapitels oder der Diskussion gewürdigt werden. Die in der Stellungnahme erwähnte S3-Leitlinie wurde ausführlich im Vorbericht in Abschnitt 6.2 dargestellt und diskutiert. Im Rahmen des Abschlussberichts wurden alle Leitlinien, die im Diskussionsabschnitt erläutert wurden, auch in Abschnitt 1.9 aufgenommen. Darüber hinaus wurden weitere Leitlinien im Rahmen des Diskussionsteils des vorliegenden Abschlussberichts gewürdigt.

## **6.5.2 Anmerkungen zu den Methoden**

### **Ergänzungen im Methodenabschnitt**

*In einer Stellungnahme wurde angemerkt, dass in Abschnitt 4.5.2.2 ergänzt werden sollte, dass im Falle einer Überlappung der Konfidenzregionen andere statistische Testverfahren herangezogen werden sollten (z. B. McNemar-Test), die in der Publikation der Studie beschrieben wurden.*

Konfidenzregionen sind nur eine grafische Darstellung der Ergebnisse einer bivariaten Meta-Analyse von mehreren direkten Vergleichen. Im vorliegenden Abschlussbericht wurde keine Signifikanzaussage aufgrund der Lage der Konfidenzregionen gemacht, sondern anhand eines Likelihood-Ratio-Tests eine Signifikanzaussage zur Nullhypothese der Gleichheit der Erwartungswerte der zu vergleichenden diagnostischen Tests abgeleitet. Univariate Tests, wie zum Beispiel der McNemar-Test, würden nur durchgeführt werden, falls eine bivariate Meta-Analyse nicht zu präzisen Schätzungen führen würde oder nicht möglich wäre, unter der Voraussetzung, dass Studien mit hohen Patientenzahlen und mit niedrigem Verzerrungspotenzial vorlägen. Dieses war im Abschlussbericht nicht der Fall.

*In einer Stellungnahme wurde erwähnt, dass im Methodenabschnitt 4.1.2.2 ergänzt werden sollte, dass eine Therapie, die zwischen der Prüftechnologie und Vergleichstechnologie einerseits und dem Referenztest andererseits bei allen Patienten durchgeführt wurde, nicht als Verzerrungspotenzial angesehen werden sollte, da eine differenzielle Wirkung dieser Therapie*

*auf die der Prüf- bzw. Vergleichstechnologie unterzogenen Patientengruppen nicht angenommen werden könne.*

Eine aktuelle Meta-Analyse bestätigt, dass die Chemo- bzw. Chemoradiotherapie bei Ösophaguskarzinomen durchaus wirkt [141]. Ob eine differenzielle Wirkung auf die beiden Indextests vorliegt, kann nur schwer beurteilt werden. In der Literatur wird dieses diskutiert, da die Radiochemotherapie möglicherweise zu Entzündungsreaktionen führt, die die Interpretation der PET-Bilder erschweren könnte. Die morphologische Bildgebung ist hiervon deutlich weniger betroffen [141].

Es ergab sich daher keine Änderung im Abschlussbericht.

### **Analyse von Primärstudien aus Evidenzsynthesen**

*In einer Stellungnahme wurde die methodische Vorgehensweise bei der Analyse von Meta-Analysen abgelehnt, da diese in besonderem Maße intransparent sei. Es wurde argumentiert, dass es durch die Beschränkung der Literaturanalyse auf einerseits Meta-Analysen und andererseits Primärliteratur, welche nicht in diesen Meta-Analysen enthalten sei, zu einer unzulässigen Begrenzung der eingeschlossenen Primärliteratur komme.*

Die Methode des Review of Reviews mit einer ergänzenden Update-Recherche U ist mittlerweile nicht nur unter den Erstellern von HTA-Berichten ein gängiges Verfahren (vgl. z. B. Whitlock 2008 [156]). Im Bereich der PET basiert etwa ein Drittel der zusammenfassenden Literatur auf derartigen Designs (vgl. z. B. Facey 2007 [150]). Das IQWiG hält diese Methode für valide und hat sie daher bereits in den IQWiG-Methoden 4.0 implementiert [32]. Die Review-of-Reviews-Methode wurde für die Indikation Restaging angewandt, weil 1. Evidenzsynthesen von ausreichend hoher Qualität vorhanden waren, weil sie 2. die gleiche Fragestellung untersuchten wie der vorliegende Bericht und weil sie 3. qualitativ hochwertige Primärstudien enthielten und diese transparent und nachvollziehbar darstellten. Es konnte davon ausgegangen werden, dass, wenn diese 3 Kriterien erfüllt waren, alle für die Fragestellung relevanten Studien auch in der Evidenzsynthese untersucht wurden. Außerdem basieren die im Bericht dargestellten Meta-Analysen, sofern sie durchführbar waren, auf den Ergebnissen aller Einzelstudien (sowohl von jenen aus den Evidenzsynthesen als auch von jenen aus der Ergänzungsrecherche). Alle vorhandenen Primärstudien werden daher transparent dargestellt und adäquat berücksichtigt. Eine unzulässige Begrenzung der Primärliteratur liegt daher nicht vor.

Es ergab sich daher keine Änderung im Abschlussbericht.

### **Anmerkungen zur Literaturrecherche**

*In einer Stellungnahme wurde darauf hingewiesen, dass sich die Suchstrategien in Anhang A des Vorberichts nicht nachvollziehen ließen, da die Suche nicht nur Ösophaguskarzinome eingeschlossen habe.*

Die Literaturrecherche des Vorberichts wurde für 4 Projekte (D06-01F, H, I und K) gleichzeitig durchgeführt, da diese Projekte von einer Sachverständigengruppe gemeinsam bearbeitet wurden. Der Aufbau der Suchstrategie gliederte sich dabei in verschiedene Suchblöcke. Die Suche zur Erkrankung in MEDLINE war zum Beispiel in den Zeilen 7 bis 17 des Vorberichts abgebildet. Das Suchergebnis zum Ösophaguskarzinom befand sich in Suchzeile 52 in Anhang A des Vorberichts. Im Abschlussbericht befindet sich ausschließlich die Suchstrategie für das Krankheitsbild Ösophaguskarzinom.

Es ergab sich daher keine Änderung im Abschlussbericht.

*In einer weiteren Stellungnahme wurde angemerkt, dass der Begriff „Karzinom des gastroösophagealen Übergangs“ keinen Eingang in die Literatursuchstrategie gefunden habe. Damit könnte nicht sicher ausgeschlossen werden, dass relevante Literatur nicht in Betracht gezogen wurde.*

Die Suchen des IQWiG bilden auch Karzinome des gastroösophagealen Übergangs ab. Beispielsweise zeigte der Abgleich mit den eingeschlossenen Evidenzsynthesen, dass die Erweiterung der Suchstrategie nicht notwendig war. Im Rahmen der Anhörung zum Vorbericht war es den Stellungnehmenden möglich, diesbezüglich zusätzliche relevante Studien vorzulegen, die nicht durch die Informationsbeschaffung des IQWiG gefunden wurden. Es wurden vonseiten der Stellungnehmenden keine zusätzlichen relevanten Studien eingereicht.

Es ergab sich daher keine Änderung im Abschlussbericht.

*In derselben Stellungnahme wurde weiter darauf verwiesen, dass es sinnvoll sei, die Literatursuchstrategie bereits im Berichtsplan darzulegen.*

Eine Veröffentlichung der Suchstrategie bereits im vorläufigen Berichtsplan wird auch bei der Cochrane Collaboration [25] nur optional umgesetzt, zum Beispiel, falls die Suche zum Zeitpunkt der Veröffentlichung des Protokolls bereits durchgeführt wurde. Da das IQWiG erst mit der Veröffentlichung des vorläufigen Berichtsplans die Suchstrategien qualitätssichert, finalisiert und in den bibliografischen Datenbanken durchführt, wird die Abbildung der Suchstrategien weiterhin im Vorbericht vorgenommen. Eine Kommentierung ist wie bisher zum Vorbericht möglich. Dies sollte anhand von zusätzlich identifizierten relevanten Studien geschehen, die nicht durch die abgebildete Suchstrategie gefunden werden konnten.

Es ergab sich daher keine Änderung im Abschlussbericht.

### **Einbeziehung von Patientenvertretern**

*In einer Stellungnahme wurde angemerkt, dass sich kein Hinweis dazu finde, ob im Zuge der Erstellung des Vorberichts erneut versucht worden sei, themenbezogene Patientenvertreter zu gewinnen. Falls das nicht erfolgt sei, werde um eine Erläuterung der Gründe und*

*gegebenenfalls die Information, ob Patientenvertreter eine Beteiligung abgelehnt hätten, gebeten.*

Die Methodik des IQWiG sieht es aktuell nicht vor, im Zuge der Erstellung von Vorberichten Patientenvertreter zu beteiligen. Die Beteiligung von Patientenvertretern wird im Rahmen der Berichtsplanerstellung angestrebt, weil zu diesem Zeitpunkt die Festlegung patientenrelevanter Endpunkte und eventueller weiterer methodischer Aspekte erfolgt. Wie in Abschnitt 3.1 beschrieben wird, fand sich im Rahmen der Berichtsplanerstellung kein Patientenvertreter für das Krankheitsbild Ösophaguskarzinom. Patientenvertreter haben jedoch stets die Möglichkeit, im Rahmen des Anhörungsverfahrens der Vorberichte Stellungnahmen abzugeben. Im aktuellen Fall wurde keine Stellungnahme von Patientenvertretern eingereicht.

Es ergab sich daher keine Änderung im Abschlussbericht.

### **6.5.3 Patientenrelevanter Nutzen**

*In einer Stellungnahme wurde darauf hingewiesen, dass die vom IQWiG geforderten Studien zum Nachweis des Nutzens in der Indikation Tumorstaging sowohl ethisch problematisch als auch kostenaufwendig seien. Es wurde argumentiert, dass RCTs für diese Indikation kaum durchführbar seien, da verschiedene internationale Leitlinien, vor allem die NCCN-Leitlinie, den Einsatz der PET/CT zum Staging bereits empfehlen.*

In Abschnitt 6.2 wird die NCCN-Leitlinie [28] ausführlich diskutiert. Die Leitlinie empfiehlt den Einsatz der PET für das M-Staging und bezieht sich dabei auf die Meta-Analyse von Westreenen 2004 [29], in welcher ausschließlich die diagnostische Güte der PET bzw. PET/CT untersucht wurde. Anhand dieser Datenbasis ließ sich kein patientenrelevanter Nutzen der PET bzw. PET/CT für die Indikation Staging ableiten bzw. voraussetzen. Die Diskussion um die Notwendigkeit der Durchführung von vergleichenden Interventionsstudien zur Ermittlung des patientenrelevanten Nutzens wurde bereits im Rahmen anderer Projekte (z. B. D06-01A) geführt.

Dass RCTs praktisch nicht durchführbar seien, weil sich die PET oder PET/CT bereits als klinischer Standard etabliert habe, erscheint kaum plausibel. Sowohl die Daten des US-amerikanischen PET-Registers [157] als auch Umfragen unter Nuklearmedizinern [158,159] zeigen, dass eine metabolische Bildgebung des Ösophaguskarzinoms nur bei einem kleinen Teil der Patienten erfolgt, was jedoch auch teilweise in der Stadienverteilung begründet sein kann.

*In einer anderen Stellungnahme wurde ebenfalls darauf hingewiesen, dass die Forderung nach randomisierten, simultanen Vergleichen zwischen einem Vorgehen mit und ohne PET in vielen Konstellationen nicht sachgerecht sei und diese von verantwortlichen Fachärzten nicht durchgeführt werden könnten.*

An dieser Stelle sollte erwähnt werden, dass im Rahmen des Abschlussberichts 4 laufende RCTs zum Restaging von Ösophaguskarzinomen identifiziert werden konnten. Weiter wurde in einer aktuellen systematischen Übersicht [160] gezeigt, dass bereits 12 RCTs zur PET in verschiedenen Krankheitsbildern publiziert wurden. 42 RCTs wurden registriert und werden in wenigen Jahren robuste Ergebnisse liefern [160]. Bedacht werden muss hierbei auch, dass nicht jeder RCT, der den Nutzen der PET bzw. PET/CT prüft, notwendigerweise ein Vorgehen mit und ohne PET vergleichen muss, da auch Interaktionsdesigns mit einer Randomisierung nicht der diagnostischen, sondern der therapeutischen Intervention denkbar sind [39]. Die von den Stellungnehmenden wiederholt geäußerte Meinung, dass die Durchführung von RCTs zur PET aus ethischen Gründen problematisch sei, ist damit nicht haltbar. Vielmehr ist zu hinterfragen, ob die Ablehnung der Überprüfung des Nutzens der PET bzw. PET/CT in RCTs noch ethisch vertretbar ist, wenn selbst die Autoren der führenden Studien zum Restaging von Ösophaguskarzinomen die Notwendigkeit der Durchführung von RCTs explizit fordern [129].

Es ergab sich daher keine Änderung im Abschlussbericht.

*In einer Stellungnahme wurde argumentiert, dass der klinische Stellenwert einer allein mit der PET bzw. PET/CT nachgewiesenen Metastasierung ähnlich sei wie der einer Metastasierung, die bereits mit konventioneller Diagnostik erkennbar sei. Ferner seien die Diagnostik und Therapie von Metastasen insgesamt – über alle konventionellen diagnostischen Modalitäten hinweg – von großer und allgemein akzeptierter Bedeutung für die Lokalthherapie des Ösophaguskarzinoms.*

Die Stellungnahme suggeriert, dass aufgrund von Analogieschlüssen eine Nutzenbewertung der PET bzw. PET/CT unnötig ist, weil die etablierten onkologischen Standards der Metastasierungsdiagnostik (beispielsweise CT, MRT oder Röntgen-Thorax) auch auf die PET bzw. PET/CT übertragen werden könnten. Es ist jedoch fraglich, ob eine immer weiter gehende Steigerung des technischen Detektionsvermögens von Metastasen tatsächlich eine immer in gleichem Maße effektive Behandlung der gefundenen Metastasen erlaubt. So zeigten sich beispielsweise bislang für die Mikrometastasierung des Ösophaguskarzinoms uneinheitliche Ergebnisse zur klinischen Relevanz [161,162]. Auch das Argument der Stellungnehmenden, dass beispielsweise eine konventionell-radiologisch und eine computertomografisch detektierte Metastase die gleiche prognostische und damit therapeutische Relevanz hätten, ist lediglich nicht widerlegt, damit aber noch nicht im Umkehrschluss belegt worden.

Es ergab sich daher keine Änderung im Abschlussbericht.

#### 6.5.4 Indikation Primärstaging

##### **Umgang mit N- und M-Stadien vor dem Hintergrund der Änderung der TNM-Klassifizierung**

*In einer Stellungnahme wurde darauf hingewiesen, dass die überwiegende Zahl der in Deutschland diagnostizierten Ösophaguskarzinome im Stadium N1 diagnostiziert werde. Deshalb sei die Feststellung eines N1-Stadiums keine Fragestellung von hoher klinischer Relevanz. Der Nachweis von distanten Lymphknotenmetastasen sei wichtig, weil dieser das weitere therapeutische Vorgehen beeinflusse. Deshalb sei in der Indikation Primärstaging die Differenzierung der Stadien N1 und M1a entscheidend. Dabei sei die Aufteilung in zervikale, thorakale und abdominelle Lymphknotenstationen wichtig.*

Dieser Punkt wurde ausführlich in der Erörterung am 29.08.2012 diskutiert. Die Aussage der Stellungnehmenden, dass die Feststellung eines N1-Stadiums (und damit die Anzahl und Lokalisation befallener Lymphknoten) keine Fragestellung von klinischer Relevanz sei, wurde von den anwesenden Experten nicht bestätigt. In der Erörterung bestand Einigkeit darüber, dass besonders die Erkennung einer Fernmetastasierung (M1b-Stadium nach der TNM-Klassifikation in der Version 6) durch den Einsatz der PET von hoher klinischer Relevanz ist. Die Forderung der Stellungnehmenden nach einer expliziten Differenzierung der Stadien N1 und M1a, so wie sie die TNM-Klassifikation in der Version 6 vorsah, wurde bereits, falls die Studiendaten es erlaubten, im Vorbericht umgesetzt. Keine der eingeschlossenen Studien verwendete die TNM-Klassifikation in der Version 7. Allerdings sei darauf hingewiesen, dass genau die von den Stellungnehmenden geforderte Differenzierung zwischen N1 und M1a in der TNM-Klassifikation in der Version 7 aufgehoben worden ist (vgl. Abschnitt 1.5).

Aufgrund der Diskussionen in der Erörterung wurden die eingeschlossenen Studien zum Primärstaging nochmals daraufhin geprüft, ob eine Differenzierung zwischen M1a und M1b möglich gewesen wäre. Dazu konnten nur Studien herangezogen werden, welche die gesamten Patientendaten darstellten. Von diesen war es lediglich bei Choi 2000 möglich, eine Differenzierung der Stadien zu machen; allerdings ließen sich die Vierfeldertafeln zur diagnostischen Genauigkeit nicht einfach ableiten: Bei der Erstellung einer Vierfeldertafel für M1b stellte sich neben einem möglichen Problem hinsichtlich einer Überdiagnose (Indextest: M1b, Referenztest: M1a) auch die Frage nach den fehlenden nicht nodalen Fernmetastasen, da diese in der Arbeit nicht berichtet wurden.

Es ergab sich daher keine Änderung im Abschlussbericht.

*In diesem Zusammenhang wurde in einer Stellungnahme angemerkt, dass in der Studie Flamen 2000 [89] distante Lymphknotenmetastasen separat von lokalen Lymphknotenmetastasen und separat vom Stadium IV ausgewertet würden. Diese Ergebnisse seien im Vorbericht jedoch nicht dargestellt worden, weil die Referenztests radiologische Verfahren beinhalteten. Diese Bewertung sei aber für das M1a-Stadium nicht korrekt.*

*Darüber hinaus wurde darauf hingewiesen, dass in der Studie Lerut 2000 [90] die Datenbasis des M1a-Stagings im Vorbericht nicht dargestellt worden sei.*

Die Studie Flamen 2000 [89] befasste sich mit dem M-Staging sowie auch mit dem N-Staging. Die Ergebnisse der Studie wurden in 2 Arbeiten (Flamen 2000 und Lerut 2000 [89,90]) berichtet, wobei sich die Publikation Lerut 2000 ausschließlich mit dem N-Staging beschäftigte. Die in der Arbeit Flamen 2000 dargestellten Ergebnisse zum M-Staging konnten auch nach einer weiteren Durchsicht 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 der Definition des Vorberichts entsprechen radiologische Verfahren nicht den Anforderungen an 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. Bezüglich der Studie Flamen 2000 ergab sich keine Änderung im Abschlussbericht. Bezüglich der Studie Lerut 2000 wurde festgestellt, dass die Ergebnisse zum M1a-Staging extrahierbar waren und die Angaben zum Referenzstandard in der Publikation Lerut 2000 den Einschlusskriterien des Berichts entsprachen. Die Daten wurden dementsprechend im Abschlussbericht dargestellt.

### **Beurteilung des Verzerrungspotenzials generell und nach den QUADAS-Kriterien für Studien zum M-Staging**

*In einer Stellungnahme wurde betont, dass die Beurteilung des Verzerrungspotenzials nach den QUADAS-Kriterien insbesondere für Studien zum M-Staging problematisch sei, da eine große Zahl der in der QUADAS postulierten Qualitätskriterien notwendigerweise verletzt werde. In diesem Zusammenhang wurde auf die Publikation Weber 2011 [163] verwiesen.*

Die Kritikpunkte von Weber wurden in einem Leserbrief der QUADAS-Autoren zu seinem Artikel [164] und im Rahmen der Erörterung zum Vorbericht ausführlich diskutiert. Man kam zu der Übereinkunft, dass es durchaus diagnostische Fragestellungen (z. B. M-Staging) gibt, in welchen die Kriterien des QUADAS-Instruments zwangsläufig verletzt werden. Dies ist jedoch kein Problem des QUADAS-Instruments. In solchen Situationen ist es eben nicht möglich, die diagnostische Güte unverzerrt zu erfassen. Die Konsequenz aus dieser Feststellung kann nicht sein, dass hier das QUADAS-Instrument „angepasst“ werden muss. Vielmehr resultiert aus der Feststellung, dass diagnostische Studien ohne Verzerrungspotenzial nicht durchführbar sind, die Notwendigkeit, die entsprechenden Fragestellungen im Rahmen von anderen Studiendesigns (z. B. RCTs) zu evaluieren.

Es ergab sich daher keine Änderung der Methodik des Abschlussberichts.

*In einer Stellungnahme wurde weiter angemerkt, dass die Einschätzungen des Verzerrungspotenzials bei den Studien Buchmann 2006 und Schreus 2008 [87,105] nicht nachvollziehbar seien.*

Für die Bewertung des Verzerrungspotenzials wurde für den vorliegenden Bericht ein Algorithmus festgelegt, nach dem die Studien bewertet wurden. Dieser Algorithmus muss laut den Vorgaben der QUADAS-Autoren fragestellungsspezifisch angepasst werden. Der Algorithmus wurde im betreffenden Methodenabschnitt hinzugefügt.

*In einer Stellungnahme wurde kritisch angemerkt, dass eine einheitliche Linie hinsichtlich der Bewertung des Verzerrungspotenzials der Prognosestudien Ott 2006 und Chung 2008 nicht zu erkennen sei. Bei der Bewertung von Chung 2008 [112] wurde die Aussage „Die Anzahl der in der multifaktoriellen Analyse eingeschlossenen Faktoren (7) scheint hinsichtlich der Gesamtzahl der eingeschlossenen Patienten (100) zu hoch zu sein“ (vgl. Abschnitt 5.3.3.2) kritisiert.*

In der Studie Chung 2008 wurde eine Cox-Regression gerechnet. Im Abschlussbericht wurde die oben angeführte Aussage korrigiert in: „Die Anzahl der Faktoren war im Verhältnis zu den Ereignissen für die multifaktorielle Analyse zu hoch, sodass hier aber weiterhin das Problem der Überparametrisierung vorlag.“

In der Studie Ott 2008 wurden wichtige Faktoren in der multifaktoriellen Analyse nicht berücksichtigt.

Die Problemlagen beider Studien, so wie sie in der Stellungnahme angemerkt wurden, sind damit nicht vergleichbar.

Bezüglich der Bewertung des Verzerrungspotenzials ergab sich daher keine Änderung im Abschlussbericht.

### **6.5.5 Indikation Restaging**

#### **Definition: histopathologisches Ansprechen**

*In einer Stellungnahme wurde angemerkt, dass die Indikation Restaging auf den Nachweis des lokalen Tumoransprechens reduziert werde. Der Nachweis oder Ausschluss von unter der Therapie aufgetretenen Fernmetastasen sei ebenso wichtig.*

Im vorliegenden Bericht wurde das Restaging als Behandlungsansprechen nach abgeschlossener Therapie definiert. Das Interimsstaging wurde als eine Diagnostik bei noch laufender Therapie zur Ermittlung des Behandlungsansprechens definiert. Dabei wurde es nicht nur auf den Nachweis des lokalen Tumoransprechens reduziert, sondern es wären auch Studien berücksichtigt worden, die den Nachweis oder Ausschluss von unter der Therapie aufgetretenen Fernmetastasen untersucht hätten. Diese Fragestellung wurde allerdings in den eingeschlossenen Studien nicht explizit untersucht. Es ergab sich daher keine Änderung im Abschlussbericht.

*In einer Stellungnahme wurde erläutert, dass im Bericht global vom „histologischen Ansprechen“ als Referenzstandard gesprochen werde. Es wurde darauf hingewiesen, dass*



*das „histologische Ansprechen“ ein Sammelbegriff sei, der in den Studien unterschiedlich definiert werde. In diesem Zusammenhang wurde weiter angemerkt, dass aufgrund dieser unterschiedlichen Definitionen des Referenzstandards die Zusammenstellung der Sensitivität und Spezifität der Studien in Form eines Forest Plots nicht sinnvoll sei.*

Das histologische Ansprechen wurde in Tabelle 31 in der Spalte „Cut-off“ dargestellt. Da die Ergebnisse der Studien ohnehin nicht gepoolt, sondern deskriptiv im Forest Plot dargestellt wurden, konnte nicht nachvollzogen werden, warum die Abbildung nicht weiterhin im Bericht dargestellt werden sollte. Allerdings wird im Abschlussbericht in der Beschreibung der Abbildung explizit darauf hingewiesen, dass es für das histologische Ansprechen unterschiedliche Klassifikationen gab.

*In der Stellungnahme wurde weiter argumentiert, dass der „Goldstandard Histopathologie“ (siehe Abschnitt 6.1.3) und seine klinische Relevanz bei der Beurteilung des Ansprechens auf eine Induktionsbehandlung derzeit sehr umstritten seien und hinterfragt werden sollten. In diesem Zusammenhang wurde die Arbeit Javeri 2009 [165] zitiert.*

In der Stellungnahme wird die Anwendung dieses Endpunktes im Rahmen der Nutzenbewertung kritisiert. Hier liegen offensichtlich 2 Missverständnisse der Stellungnehmenden vor: Zum einen ist die Arbeit Javeri 2009 [165] eine multifaktorielle Prognosestudie. Hier wurde das pathologische Ansprechen als prognostischer Faktor für das Gesamtüberleben und das krankheitsfreie Überleben eingesetzt. Die Autoren konstatieren in ihrer Diskussion, dass die Unterschiede im pathologischen Ansprechen im multifaktoriellen Modell möglicherweise aufgrund der geringen Patientenzahlen nicht signifikant wurden. Diese Arbeit eignet sich daher weder für die Validierung noch für das Widerlegen der Validität dieses Endpunktes als Surrogat für patientenrelevante Endpunkte im Rahmen einer Nutzenbewertung. Das zweite Missverständnis der Stellungnehmenden bezieht sich auf die Interpretation dieses Endpunktes im Bericht. Im vorliegenden Bericht wurde das pathologische Ansprechen als Referenzstandard für das Ansprechen einer neoadjuvanten Chemotherapie im Rahmen von Diagnose- und Prognosestudien verwendet. Ziel war es hier nicht, aus diesem Ansprechen unmittelbar auf die Verbesserung patientenrelevanter Zielgrößen und damit auf eine klinische Relevanz zu schließen. Entsprechende Studien sind daher auch nicht im Abschnitt zur Nutzenbewertung gelistet, sondern im Abschnitt zur diagnostischen und prognostischen Güte.

Es ergab sich daher keine Änderung im Abschlussbericht.

#### **Bewertung der Studie Ott 2006 [74]**

*In einer Stellungnahme wurde bemängelt, dass die Studie Ott 2006 [74] aufgrund von folgenden nicht nachvollziehbaren Aspekten abgewertet worden sei: a) Angaben zum PET-Gerät, b) Entscheidungsfindung anhand eines prospektiven, quantitativen Schwellenwertes beschrieben, c) Angaben zur Vorbehandlung (neoadjuvante Therapie von neu*

*diagnostizierten Tumoren), d) unterschiedliches Krankheitsstadium der Patienten, e) Art der Chemotherapie nicht in der multifaktoriellen Analyse berücksichtigt.*

Eine Studie wurde nur aufgrund der Bewertung ihres Verzerrungspotenzials abgewertet. Die genannten Aspekte a) bis c) beziehen sich lediglich auf die Beschreibung der Studie und führten in keinem Fall zu einer Abwertung der Studie. Die genannten Kritikpunkte wurden gemeinsam mit der Autorin der Studie besprochen und entsprechende Überarbeitungen im Abschlussbericht vorgenommen.

### **6.5.6 Primärstudien bzw. Evidenzsynthesen mit Einschlussforderung**

*In einer Stellungnahme wurde darauf hingewiesen, dass unklar bleibe, weshalb die MUNICON-Studien und die HICON-Studie nicht angemessen im Vorbericht berücksichtigt wurden (Lordick 2007, zum Büschenfelde 2011, Lorenzen BMC Cancer 2011).*

Dieser Punkt wurde ausführlich in der Erörterung, welche am 29.08.2012 im IQWiG durchgeführt wurde, erläutert und diskutiert. Die MUNICON-I-Studie wurde im Rahmen der systematischen Literaturrecherche zum Vorbericht, die MUNICON-II- und die HICON-Studie im Rahmen der systematischen Literaturrecherche zum Abschlussbericht gesichtet. Keine der genannten Studien entsprach den Einschlusskriterien des vorliegenden Berichts. Die ausführlichen Gründe für den Ausschluss der Studien sind dem Protokoll der Erörterung zu entnehmen. Darüber hinaus wurden die Ergebnisse der MUNICON-I- und -II-Studie in Abschnitt 6.1.3 des Berichts ausführlich diskutiert.

Es ergab sich daher keine Änderung im Abschlussbericht.

*In einer Stellungnahme wurde gefragt, warum der HTA-Bericht der australischen Gesundheitsbehörde Medical Services Advisory Committee (MSAC) „Positron emission tomography for oesophageal and gastric cancer“ [166] keine Erwähnung im Vorbericht gefunden habe.*

Der HTA-Bericht des MSAC wurde vom IQWiG durch die systematische Literaturrecherche im Rahmen der Erstellung des Vorberichts gefunden und gesichtet. Der Bericht wurde aber bereits im Titel-Abstract-Screening ausgeschlossen. Es handelte sich um einen Review of Reviews plus Update-Recherche. Damit entsprach der Bericht nicht den Einschlusskriterien des Vorberichts.

Es ergab sich daher keine Änderung im Abschlussbericht.

*Im gleichen Abschnitt wurde in der Stellungnahme angemerkt, dass auf Grundlage der gleichen Datenbasis wie die des IQWiG der Bericht des MSAC zu einer konträren Schlussfolgerung komme.*

Der vorliegende Abschlussbericht verfolgte das Ziel, den patientenrelevanten Nutzen für die PET bei Patienten mit Ösophaguskarzinom zu untersuchen. Da bezüglich dieser Fragestellung

keine Studien gefunden wurden, konnte kein patientenrelevanter Nutzen abgeleitet werden. Der MSAC-Bericht aus 2008 [166] formulierte diesbezüglich ein ähnliches Fazit wie der IQWiG-Bericht: „There was no direct evidence available at the time to demonstrate that improvements in diagnostic accuracy provided by PET, and any subsequent management changes, led to improvements in health outcomes for patients“ [166, S.14].

Es ergab sich daher keine Änderung im Abschlussbericht.

### **6.5.7 Ausschluss von Ergebnissen aufgrund fehlender Werte größer als 30 %**

*In einer Stellungnahme wurde darauf hingewiesen, dass es intransparent bleibe, wann Studien aufgrund fehlender Werte größer als 30 % ausgeschlossen würden. In diesem Zusammenhang wurden die Studien Kato 2008, Liberale 2004, Ott 2006 und Wieder 2004 [34,74,86,95] genannt.*

Es wurde keine Studie aufgrund der 30%-Regel ausgeschlossen, sondern es wurden lediglich die Ergebnisse der Vierfeldertafeln der betroffenen Studien nicht im Bericht dargestellt.

Im Rahmen der Überarbeitung des Abschlussberichts wurden die Angaben noch einmal geprüft und es wurde festgestellt, dass die Daten der Vierfeldertafel der Studie Liberale 2004 [34] im Vorbericht nicht hätten dargestellt werden dürfen aufgrund der 30%-Regel. Die Daten wurden im Abschlussbericht nicht mehr dargestellt.

Entgegen der Studie Liberale 2004 hätten die Daten der Vierfeldertafel für die PET/CT in der Studie Kato 2008 im Vorbericht dargestellt werden müssen. Diese wurden im Abschlussbericht nachextrahiert und dargestellt. An dieser Stelle sollte erwähnt werden, dass in der Studie Kato 2008 50 Patienten nachrekrutiert wurden, die ausschließlich eine PET/CT bekamen. Die Gesamtzahl der eingeschlossenen Patienten war damit nicht 167, sondern 117, sodass die ~30%-Regel hier nicht angewendet werden durfte. Allerdings sind die Daten der Vierfeldertafel der PET/CT nicht mit den Daten der PET und CT vergleichbar, da es sich aufgrund der Nachrekrutierung um unterschiedliche Patienten handelte.

### **6.5.8 Anmerkungen zum Ausschluss von Studien**

*In einer Stellungnahme wurde bemängelt, dass 37 Primärstudien mit dem Ausschlussgrund „Studiendesign entspricht nicht den Einschlusskriterien“ ausgeschlossen worden seien. Es wurde darum gebeten, die Definition der Einschlusskriterien zu beleuchten.*

Der Ausschlussgrund „Studiendesign entspricht nicht den Einschlusskriterien“ wurde nur dann vergeben, wenn die Studie weder für den patientenrelevanten Nutzen noch für die diagnostische und prognostische Güte verwertbar war. Viele der Studien mit diesem Ausschlussgrund E4 waren beispielsweise retrospektive Analysen zur diagnostischen Güte oder es handelte sich um läsionsbasierte Analysen. Aus Gründen der Übersichtlichkeit wurde auf eine detailliertere Erläuterung des Ausschlussgrundes E4 verzichtet.

Es ergab sich daher keine Änderung im Abschlussbericht.

### **6.5.9 Einschätzung der Übertragbarkeit der Ergebnisse**

*In einer Stellungnahme wurde die Übertragbarkeit der Untersuchungsergebnisse der Studie Higuchi 2008 [115] auf die Patientenversorgung in Deutschland bemängelt. Begründet wurde die Einschätzung zum einen mit der Untersuchung asiatischer Patienten (Unterschiede in der Tumorbiologie von Ösophaguskarzinomen bei asiatischen und europäischen Patienten seien ausführlich dokumentiert, z. B. Rice 2009) und zum anderen mit dem Referenztest, welcher eine in Europa nicht eingesetzte Klassifikation des histologischen Ansprechens verwende. Darüber hinaus sei der in der Studie verwendete PET-Gerätetyp in Europa nicht verfügbar.*

Das IQWiG konnte zwar auf Basis der zitierten Arbeit von Rice 2008 die genannten Aspekte nicht rekonstruieren, wird der Argumentation der DGN jedoch folgen und bewertet die genannten Aspekte als Indizien für eine Nichtübertragbarkeit der Studie.

Die Bewertung dieser Studie und aller weiteren Studien, welche ausschließlich asiatische Patienten untersuchten, wurde im Abschlussbericht geändert.

*In einer Stellungnahme wurde angemerkt, dass, während zur Verhinderung eines Bias verlangt werde, verblindete Studien durchzuführen, eine Verblindung zu einer negativen Beurteilung der Studie im Hinblick auf die Übertragbarkeit auf den Versorgungsalltag führe. Im gleichen Zusammenhang wurde in einer anderen Stellungnahme argumentiert, dass man bei konsequenter Anwendung dieses Kriteriums zu dem Schluss kommen müsse, dass kein einziger RCT auf den deutschen Versorgungskontext übertragbar sei.*

Für das Verzerrungspotenzial von diagnostischen Studien ist es entscheidend, ob die Bewerter des Indextests die Ergebnisse des Referenztests kennen (und umgekehrt). Letzterer wird im klinischen Alltag in der Regel nicht oder erst nach dem Indextest erhoben. Für die Übertragbarkeit war es entscheidend, ob den Bewertern des Indextests dieselben klinischen Vorinformationen zur Verfügung standen, wie sie im deutschen Versorgungsalltag zur Verfügung stehen. Wenn die Bewerter gegenüber Vorabinformationen vollständig verblindet waren, entspricht dies höchstwahrscheinlich nicht der Versorgungsrealität und es gab ein Indiz für eine Nichtübertragbarkeit der Studie. Die Verblindung bezieht sich also auf die Kenntnis einerseits des Referenztests und andererseits der klinischen Vorinformationen. Da beide Verblindungen unabhängig voneinander sind, entsteht hierdurch kein Dilemma zwischen interner und externer Validität.

Es ergab sich daher keine Änderung im Abschlussbericht.

## 7 Fazit

Der patientenrelevante Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen ist nicht belegt. Es konnten weder laufende noch abgeschlossene vergleichende Interventionsstudien zum patientenrelevanten Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen identifiziert werden. Ob der Einsatz der PET bzw. PET/CT auch zu einer Verbesserung patientenrelevanter Endpunkte führt, ist bisher in keiner der 3 Indikationen untersucht worden.

Zu der 2. Fragestellung des Berichts, der diagnostischen und prognostischen Güte, konnten aus 3 eingeschlossenen Evidenzsynthesen und der Ergänzungsrecherche insgesamt 48 Primärstudien eingeschlossen werden. Über alle Indikationen hinweg, wurde in 19 Studien ein direkter Vergleich der PET bzw. PET/CT mit einem anderen bildgebenden Verfahren durchgeführt. In einer Studie wurde der Tracer FDG mit dem Tracer <sup>11</sup>C-Cholin direkt verglichen.

Die meisten Studien wurden in der Indikation Primärstaging durchgeführt. Zwölf Studien verglichen die PET direkt mit der CT beim N-Staging und 7 Studien untersuchten den direkten Vergleich von PET mit CT beim M-Staging. Beim N-Staging konnte in der bivariaten Meta-Analyse kein statistisch signifikanter Unterschied zwischen der diagnostischen Güte von PET und CT festgestellt werden. Beim M-Staging ergab die bivariate Meta-Analyse keine präzisen Schätzungen, sodass auch hier keine Aussage bezüglich der diagnostischen Güte von PET und CT getroffen werden kann. 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. Bisher konnten keine abgeschlossenen parallel vergleichende Studien zum patientenrelevanten Nutzen der PET bzw. PET/CT bei Ösophaguskarzinomen identifiziert werden. Methodisch hochwertige Studien sind (insbesondere bezüglich der Fragestellung des Behandlungsansprechens neoadjuvanter Therapien) dringend erforderlich, um den patienten-relevanten Nutzen oder Schaden der PET bzw. PET/CT von Ösophaguskarzinomen verlässlich bewerten zu können.

## 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.  $^{18}\text{F}$ -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. <sup>18</sup>F-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-[<sup>18</sup>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.

Levine EA, Farmer MR, Clark P, Mishra G, Ho C, Geisinger KR et al. Predictive value of 18-fluoro-deoxy-glucose-positron emission tomography (<sup>18</sup>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.

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, Brucher 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. *Nuklearmedizin* 2006; 45(6): 235-241.

#### **Cheze-Le Rest 2008**

Cheze-Le Rest C, Metges JP, 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. <sup>18</sup>F-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 <sup>201</sup>Tl and <sup>18</sup>F-FDG in esophageal cancer. *World J Surg* 2008; 32(1): 69-75.



### **Flamen 2000**

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.

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.

### **Gillies 2012a**

Gillies RS, Middleton MR, Han C, Marshall REK, Maynard ND, Bradley KM et al. Role of positron emission tomography-computed tomography in predicting survival after neoadjuvant chemotherapy and surgery for oesophageal adenocarcinoma. *Br J Surg* 2012; 99(2): 239-245.

### **Gillies 2012b**

Gillies RS, Middleton MR, Blesing C, Patel K, Warner N, Marshall RE et al. Metabolic response at repeat PET/CT predicts pathological response to neoadjuvant chemotherapy in oesophageal cancer. *Eur Radiol* 2012; 22(9): 2035-2043.

### **Guo 2007**

Guo H, Zhu H, Xi Y, Zhang B, Li L, Huang Y et al. Diagnostic and prognostic value of <sup>18</sup>F-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 <sup>18</sup>F-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.

### **Iison 2011**

Iison DH, Minsky BD, Ku GY, Rusch V, Rizk N, Shah M et al. Phase 2 trial of induction and concurrent chemoradiotherapy with weekly irinotecan and cisplatin followed by surgery for esophageal cancer. *Cancer* 2011; 118(11): 2820-2827.

**Itoh 2008**

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

**Jager 2001**

Jager PL, Que TH, Vaalburg W, Pruim 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 SB, Ahn JH, Kim SY, Park SL et al. Value of complete metabolic response by <sup>18</sup>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.

**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, 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.

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.

#### **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. *Mol Imaging Biol* 2005; 7(6): 422-430.

#### **Ma 2012**

Ma JB, Song YP, Yu JM, Zhou W, Cheng EC, Zhang XQ. Linear correlation between patient survival and decreased percentage of tumor [<sup>18</sup>F]fluorodeoxyglucose uptake for late-course accelerated hyperfractionated radiotherapy for esophageal cancer. *Int J Radiat Oncol Biol Phys* 2012; 82(4): 1535-1540.

#### **Malik 2010**

Malik V, Lucey JA, Duffy GJ, Wilson L, McNamara L, Keogan M et al. Early repeated <sup>18</sup>F-FDG PET scans during neoadjuvant chemoradiation fail to predict histopathologic response or survival benefit in adenocarcinoma of the esophagus. *J Nucl Med* 2010; 51(12): 1863-1869.

#### **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.

**Sihvo 2004**

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.

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.

**Van Heijl 2011**

Van Heijl M, Omloo JM, Van Berge Henegouwen MI, Hoekstra OS, Boellaard R, Bossuyt PM et al. Fluorodeoxyglucose positron emission tomography for evaluating early response during neoadjuvant chemoradiotherapy in patients with potentially curable esophageal cancer. *Ann Surg* 2011; 253(1): 56-63.

**Vyas 2011**

Vyas S, Markar SR, Iordanidou L, Read S, Stoker D, Hashemi M et al. The role of integrated F-18-FDG-PET scanning in the detection of M1 disease in oesophageal adenocarcinoma and impact on clinical management. *J Gastrointest Surg* 2011; 15(12): 2127-2135.

**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. Kaatsch P, Spix C, Katalinic A, Hentschel S, Baras N, Barnes B et al. Krebs in Deutschland 2007/2008. Berlin: Robert Koch-Institut; 2012. URL: [http://www.rki.de/Krebs/DE/Content/Publikationen/Krebs\\_in\\_Deutschland/kid\\_2012/krebs\\_in\\_deutschland\\_2012.pdf?\\_\\_blob=publicationFile](http://www.rki.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/kid_2012/krebs_in_deutschland_2012.pdf?__blob=publicationFile).
2. Reynolds JV, Ravi N, Muldoon C, Larkin JO, Rowley S, O'Byrne K et al. Differential pathologic variables and outcomes across the spectrum of adenocarcinoma of the esophagogastric junction. *World J Surg* 2010; 34(12): 2821-2829.
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. Zentrum für Krebsregisterdaten. ICD-10 C15 (Speiseröhre): Neuerkrankungen; Inzidenzraten; Sterbefälle; Sterberaten [online]. 06.03.2012 [Zugriff: 31.01.2013]. URL: [http://www.krebsdaten.de/Krebs/DE/Home/Datenbankabfrage/c15\\_speiseroehre.xls?\\_\\_blob=publicationFile](http://www.krebsdaten.de/Krebs/DE/Home/Datenbankabfrage/c15_speiseroehre.xls?__blob=publicationFile).
9. 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.
10. 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](http://www.mdl.nl/uploads/240/809/Richtlijn_Oesofaguscarcinoom_definitief_december_2010.pdf).
11. Scottish Intercollegiate Guidelines Network. Management of oesophageal and gastric cancer: a national clinical guideline. Edinburgh: SIGN; 2006. (SIGN Publications; Band 87). URL: <http://www.sign.ac.uk/pdf/sign87.pdf>.

12. SIGN. Proposed review of sign guideline: consultation summary; SIGN 87; management of oesophageal and gastric cancer [online]. 12.01.2012 [Zugriff: 31.01.2013]. URL: [http://www.sign.ac.uk/pdf/SIGN87\\_gastric\\_cancer\\_review.pdf](http://www.sign.ac.uk/pdf/SIGN87_gastric_cancer_review.pdf).
13. Gertler R, Stein HJ, Langer R, Nettelmann M, Schuster T, Hoefler H et al. Long-term outcome of 2920 patients with cancers of the esophagus and esophagogastric junction: evaluation of the new Union Internationale Contre le Cancer/American Joint Cancer Committee staging system. *Ann Surg* 2011; 253(4): 689-698.
14. Sobin LH, Wittekind C (Ed). *TNM classification of malignant tumours*. New York: Wiley; 2002.
15. Wittekind C, Meyer HJ (Ed). *TNM Klassifikation maligner Tumoren*. Weinheim: Wiley-Blackwell; 2010.
16. Kahlert C, Büchler MW, Weitz J. Extendierte Lymphknotendisektion und Gefäßresektion beim Pankreaskarzinom. *Chirurg* 2008; 79(12): 1115-1122.
17. Siewert JR, Stein HJ. Classification of adenocarcinoma of the oesophagogastric junction. *Br J Surg* 1998; 85(11): 1457-1459.
18. Stahl M, Budach W, Meyer HJ, Cervantes A. Esophageal cancer: clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2010; 21(Suppl 5): v46-v49.
19. 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](http://www.tumorzentrum-freiburg.de/tumorzentrum/live/Medizin-Info/Leitlinien/oesophagus_karzinom.pdf).
20. Stahl M, Oliveira J. Esophageal cancer: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann Oncol* 2009; 20(Suppl 4): iv32-iv33.
21. Malthaner R, Collin S, Fenlon D, Rhodes S. Preoperative chemotherapy for resectable thoracic esophageal cancer. *Cochrane Database Syst Rev* 2006; (3): CD001556.
22. 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.
23. 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](http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/wgi/download/pet_ct_stellungnahme_050209.pdf).

24. 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.
25. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Diagnostik und Therapie der Plattenepithelkarzinome und Adenokarzinome des Ösophagus: Registernummer 021 - 023OL [online]. In: AWMF online: angemeldete Leitlinien. 31.10.2012 [Zugriff: 15.04.2013]. URL: <http://www.awmf.org/leitlinien/detail/anmeldung/1/ll/021-023OL.html>.
26. 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 ösophagogastralen Übergangs (AWMF-Regist.-Nr. 032-009-OL). *Z Gastroenterol* 2011; 49(4): 461-531.
27. Agency for Healthcare Research and Quality. National Guideline Clearinghouse [online]. [Zugriff: 29.05.2013]. URL: <http://www.guideline.gov>.
28. National Comprehensive Cancer Network. NCCN clinical practice guidelines in oncology: esophageal and esophagogastric junction cancers (excluding the proximal 5 cm of the stomach); version 2.2012 [online]. 31.05.2012 [Zugriff: 06.03.2013]. URL: <http://www.nccn.org>.
29. Van Westreenen HL, Westerterp M, Bossuyt PM, Pruijm J, Sloof GW, Van Lanschot JJ et al. Systematic review of the staging performance of <sup>18</sup>F-fluorodeoxyglucose positron emission tomography in esophageal cancer. *J Clin Oncol* 2004; 22(18): 3805-3812.
30. Saskatchewan Cancer Agency. Provincial esophageal cancer and gastro-esophageal junction cancer treatment guidelines [online]. 05.2011 [Zugriff: 12.02.2013]. URL: <http://www.saskcancer.ca/Esophageal%20Guidelines>.
31. 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.
32. 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](https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf).
33. 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](https://www.iqwig.de/download/D06-01_E-K_Berichtsplan_PET_bei_onkologischen_Indikationen_.pdf).



34. 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.
35. How to read clinical journals; II: to learn about a diagnostic test. *Can Med Assoc J* 1981; 124(6): 703-710.
36. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991; 11(2): 88-94.
37. Köbberling J, Trampisch HJ, Windeler J. Memorandum for the evaluation of diagnostic measures. *J Clin Chem Clin Biochem* 1990; 28(12): 873-879.
38. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. 11.06.2013 [Zugriff: 13.06.2013]. URL: [http://www.g-ba.de/downloads/62-492-711/VerfO\\_2013-06-11.pdf](http://www.g-ba.de/downloads/62-492-711/VerfO_2013-06-11.pdf).
39. Lijmer JG, Bossuyt PM. Various randomized designs can be used to evaluate medical tests. *J Clin Epidemiol* 2009; 62(4): 364-373.
40. 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.
41. 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.
42. 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.
43. Kirby AM, George MN. The role of FDG PET in the management of lymphoma: practical guidelines. *Nucl Med Commun* 2007; 28(5): 355-357.
44. 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.
45. 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.

46. 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.
47. Pepe MS, Alonzo TA. Comparing disease screening tests when true disease status is ascertained only for screen positives. *Biostatistics* 2001; 2(3): 249-260.
48. Jadad AR, Murray WE. *Randomized controlled trials: questions, answers and musings*. Malden: Blackwell; 2007.
49. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991; 44(11): 1271-1278.
50. 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.
51. 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.
52. 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.
53. 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.
54. 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.
55. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med* 2006; 144(6): 427-437.
56. Altman DG. Systematic reviews of evaluations of prognostic variables. *BMJ* 2001; 323(7306): 224-228.
57. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; 359(9308): 781-785.
58. 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.

59. Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *CMAJ* 1997; 156(10): 1411-1416.
60. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7(3): 177-188.
61. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414): 557-560.
62. Leemis L, Trivedi K. A comparison of approximate interval estimators for the Bernoulli parameter. *Am Stat* 1996; 50(1): 63-68.
63. 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.
64. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol* 2006; 59(12): 1331-1332.
65. Menke J. Bivariate random-effects meta-analysis of sensitivity and specificity with SAS PROC GLIMMIX. *Methods Inf Med* 2010; 49(1): 54-64.
66. Hotelling H. The generalization of student's ratio. *Annals of Mathematical Statistics* 1931; 2(3): 360-378.
67. 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.
68. 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.
69. Choi JY, Jang HJ, Shim YM, Kim K, Lee KS, Lee KH et al. <sup>18</sup>F-FDG PET in patients with esophageal squamous cell carcinoma undergoing curative surgery: prognostic implications. *J Nucl Med* 2004; 45(11): 1843-1850.
70. 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.
71. 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.

72. 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.
73. 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.
74. 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.
75. 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.
76. Rebollo Aguirre AC, Ramos-Font C, Villegas Portero R, Cook GJR, Llamas Elvira JM, Tabares AR. <sup>18</sup>F-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.
77. 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.
78. 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.
79. 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.
80. Gillham CM, Lucey JA, Keogan M, Duffy GJ, Malik V, Raouf AA et al. <sup>18</sup>FDG uptake during induction chemoradiation for oesophageal cancer fails to predict histomorphological tumour response. *Br J Cancer* 2006; 95(9): 1174-1179.

81. Kroep JR, Van Groeningen CJ, Cuesta MA, Craanen ME, Hoekstra OS, Comans EF et al. Positron emission tomography using 2-deoxy-2-[<sup>18</sup>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.
82. Levine EA, Farmer MR, Clark P, Mishra G, Ho C, Geisinger KR et al. Predictive value of 18-fluoro-deoxy-glucose-positron emission tomography (<sup>18</sup>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.
83. 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.
84. 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.
85. 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.
86. Wieder HA, Brucher 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.
87. 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. *Nuklearmedizin* 2006; 45(6): 235-241.
88. 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.
89. 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.
90. 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.

91. Heeren PAM, Jager PL, Bongaerts F, Van Dullemen H, Sluiter W, Plukker JTM. Detection of distant metastases in esophageal cancer with <sup>18</sup>F-FDG PET. *J Nucl Med* 2004; 45(6): 980-987.
92. Itoh K, Hosokawa M, Fujita M. <sup>18</sup>F-FDG-PET/CT in esophageal cancer: comparison between PET and histopathological evaluation on metastasis to locoregional lymph node [Japanisch]. *Rinsho Hoshasen* 2008; 53(4): 501-507.
93. 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.
94. 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.
95. 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.
96. 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.
97. 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.
98. 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.
99. 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.
100. 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.
101. 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.
102. 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. *Mol Imaging Biol* 2005; 7(6): 422-430.

103. 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.
104. 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.
105. 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.
106. 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.
107. 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.
108. Vyas S, Markar SR, Iordanidou L, Read S, Stoker D, Hashemi M et al. The role of integrated F-18-FDG-PET scanning in the detection of M1 disease in oesophageal adenocarcinoma and impact on clinical management. *J Gastrointest Surg* 2011; 15(12): 2127-2135.
109. 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.
110. Cheze-Le Rest C, Metges JP, 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.
111. 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.
112. Chung HW, Lee KH, Lee EJ, Lee SJ, Cho YS, Choi JY et al. Comparison of uptake characteristics and prognostic value of <sup>201</sup>Tl and <sup>18</sup>F-FDG in esophageal cancer. *World J Surg* 2008; 32(1): 69-75.

113. Gillies RS, Middleton MR, Han C, Marshall REK, Maynard ND, Bradley KM et al. Role of positron emission tomography-computed tomography in predicting survival after neoadjuvant chemotherapy and surgery for oesophageal adenocarcinoma. *Br J Surg* 2012; 99(2): 239-245.
114. Gillies RS, Middleton MR, Blesing C, Patel K, Warner N, Marshall RE et al. Metabolic response at repeat PET/CT predicts pathological response to neoadjuvant chemotherapy in oesophageal cancer. *Eur Radiol* 2012; 22(9): 2035-2043.
115. 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.
116. Ilson DH, Minsky BD, Ku GY, Rusch V, Rizk N, Shah M et al. Phase 2 trial of induction and concurrent chemoradiotherapy with weekly irinotecan and cisplatin followed by surgery for esophageal cancer. *Cancer* 2011; 118(11): 2820-2827.
117. Malik V, Lucey JA, Duffy GJ, Wilson L, McNamara L, Keogan M et al. Early repeated <sup>18</sup>F-FDG PET scans during neoadjuvant chemoradiation fail to predict histopathologic response or survival benefit in adenocarcinoma of the esophagus. *J Nucl Med* 2010; 51(12): 1863-1869.
118. 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.
119. Van Heijl M, Omloo JM, Van Berge Henegouwen MI, Hoekstra OS, Boellaard R, Bossuyt PM et al. Fluorodeoxyglucose positron emission tomography for evaluating early response during neoadjuvant chemoradiotherapy in patients with potentially curable esophageal cancer. *Ann Surg* 2011; 253(1): 56-63.
120. Kim MK, Ryu JS, Kim SB, Ahn JH, Kim SY, Park SL et al. Value of complete metabolic response by <sup>18</sup>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.
121. Ma JB, Song YP, Yu JM, Zhou W, Cheng EC, Zhang XQ. Linear correlation between patient survival and decreased percentage of tumor [<sup>18</sup>F]fluorodeoxyglucose uptake for late-course accelerated hyperfractionated radiotherapy for esophageal cancer. *Int J Radiat Oncol Biol Phys* 2012; 82(4): 1535-1540.



122. Guo H, Zhu H, Xi Y, Zhang B, Li L, Huang Y et al. Diagnostic and prognostic value of  $^{18}\text{F}$ -FDG PET/CT for patients with suspected recurrence from squamous cell carcinoma of the esophagus. *J Nucl Med* 2007; 48(8): 1251-1258.
123. 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.
124. 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.
125. 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.
126. Hara T, Yuasa M. Automated synthesis of [ $^{11}\text{C}$ ]choline, a positron-emitting tracer for tumor imaging. *Appl Radiat Isot* 1999; 50(3): 531-533.
127. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binominal. *Biometrika* 1934; 26(4): 404-413.
128. Levine EA, Farmer MR, Clark P, Mishra G, Ho C, Geisinger KR et al. Predictive value of 18-fluoro-deoxy-glucose-positron emission tomography ( $^{18}\text{F}$ -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.
129. 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.
130. Zum Buschenfelde CM, Herrmann K, Schuster T, Geinitz H, Langer R, Becker K et al.  $^{18}\text{F}$ -FDG PET-guided salvage neoadjuvant radiochemotherapy of adenocarcinoma of the esophagogastric junction: the MUNICON II trial. *J Nucl Med* 2011; 52(8): 1189-1196.
131. 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.
132. 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.

133. Chen YM, Pan XF, Tong LJ, Shi YP, Chen T. Can  $^{18}\text{F}$ -fluorodeoxyglucose positron emission tomography predict responses to neoadjuvant therapy in oesophageal cancer patients? A meta-analysis. *Nucl Med Commun* 2011; 32(11): 1005-1010.
134. Kwee RM. Prediction of tumor response to neoadjuvant therapy in patients with esophageal cancer with use of  $^{18}\text{F}$  FDG PET: a systematic review. *Radiology* 2010; 254(3): 707-717.
135. Zhu W, Xing L, Yue J, Sun X, Sun X, Zhao H et al. Prognostic significance of SUV on PET/CT in patients with localised oesophagogastric junction cancer receiving neoadjuvant chemotherapy/chemoradiation: a systematic review and meta-analysis. *Br J Radiol* 2012; 85(1017): e694-e701.
136. 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.
137. Chen YM, Pan XF, Tong LJ, Shi YP, Chen T. Can  $^{18}\text{F}$ -fluorodeoxyglucose positron emission tomography predict responses to neoadjuvant therapy in oesophageal cancer patients? A meta-analysis. *Nucl Med Commun* 2011; 32(11): 1005-1010.
138. 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.
139. 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.
140. GebSKI V, Burmeister B, Smithers BM, Foo K, Zalcbberg J, Simes J. Survival benefits from neoadjuvant chemoradiotherapy or chemotherapy in oesophageal carcinoma: a meta-analysis. *Lancet Oncol* 2007; 8(3): 226-234.
141. Sjoquist KM, Burmeister BH, Smithers BM, Zalcbberg 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.
142. 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.
143. Law SY, Fok M, Wong J. Pattern of recurrence after oesophageal resection for cancer: clinical implications. *Br J Surg* 1996; 83(1): 107-111.

144. 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.
145. Esteves FP, Schuster DM, Halkar RK. Gastrointestinal tract malignancies and positron emission tomography: an overview. *Semin Nucl Med* 2006; 36(2): 169-181.
146. 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.
147. Peeters M, Lerut T, Vlayen J, Mambourg F, Ectors N, Deprez P et al. 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](http://kce.fgov.be/sites/default/files/page_documents/d20081027316.pdf).
148. 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.
149. 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.
150. 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.
151. 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.
152. 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.
153. Jingu K, Kaneta T, Nemoto K, Takeda K, Ogawa Y, Ariga H et al. <sup>18</sup>F-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.
154. Wong RJ, Walker-Dilks C, Raifu AO. 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>.

155. 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; Dokumentation und Würdigung der Anhörung zum Berichtsplan; Aufträge: D06-01E-K [online] [online]. [Zugriff: 18.07.2012]. URL: [https://www.iqwig.de/download/D06-01E-K\\_DWA-Berichtsplan\\_PET\\_und\\_PET-CT\\_bei\\_onkologischen\\_Indikationen.pdf](https://www.iqwig.de/download/D06-01E-K_DWA-Berichtsplan_PET_und_PET-CT_bei_onkologischen_Indikationen.pdf).
156. Whitlock EP, Lin JS, Chou R, Shekelle P, Robinson KA. Using existing systematic reviews in complex systematic reviews. *Ann Intern Med* 2008; 148(10): 776-782.
157. Hillner BE, Tosteson AN, Song Y, Tosteson TD, Onega T, Goodman DC et al. Growth in the use of PET for six cancer types after coverage by medicare: additive or replacement? *J Am Coll Radiol* 2012; 9(1): 33-41.
158. Beyer T, Czernin J, Freudenberg LS. Variations in clinical PET/CT operations: results of an international survey of active PET/CT users. *J Nucl Med* 2011; 52(2): 303-310.
159. Kotzerke J, Oehme L, Lindner O, Hellwig D. Positronenemissionstomographie 2008 in Deutschland: Ergebnisse der Erhebung und Standortbestimmung. *Nuklearmedizin* 2010; 49(2): 58-64.
160. Scheibler F, Zumbé P, Janssen I, Viebahn M, Schröer-Günther M, Grosselfinger R et al. Randomized controlled trials on PET: a systematic review of topics, design, and quality. *J Nucl Med* 2012; 53(7): 1016-1025.
161. Gray RT, O'Donnell ME, Verghis RM, McCluggage WG, Maxwell P, McGuigan JA et al. Bone marrow micrometastases in esophageal carcinoma: a 10-year follow-up study. *Dis Esophagus* 2012; 25(8): 709-715.
162. Prenzel KL, Hölscher AH, Drebber U, Agavonova M, Gutschow CA, Bollschweiler E. Prognostic impact of nodal micrometastasis in early esophageal cancer. *Eur J Surg Oncol* 2012; 38(4): 314-318.
163. Weber WA. Is there evidence for evidence-based medical imaging? *J Nucl Med* 2011; 52(Suppl 2): 74S-76S.
164. Kleijnen J, Westwood M, Wolff R, Whiting P, Raatz H, Kunz R. Evidence-based assessment of PET in Germany. *J Nucl Med* 2012; 53(7): 1166-1167.
165. 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.

166. Medical Services Advisory Committee. Positron emission tomography for oesophageal and gastric cancer: assessment report. Canberra: MSAC; 2008. (MSAC References; Band 35b(i)). URL:

[http://www.msac.gov.au/internet/msac/publishing.nsf/Content/8FD1D98FE64C8A2FCA2575AD0082FD8F/\\$File/MSAC\\_35b\\_i\\_PET\\_gastroesophageal\\_toprint\\_FINAL%20to%20print\\_301008.pdf](http://www.msac.gov.au/internet/msac/publishing.nsf/Content/8FD1D98FE64C8A2FCA2575AD0082FD8F/$File/MSAC_35b_i_PET_gastroesophageal_toprint_FINAL%20to%20print_301008.pdf).

167. 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

### A.1 Primärstudien

#### 1. EMBASE

##### Suchoberfläche: Ovid

- Embase 1980 to 2012 May 17

#	Searches
1	exp Esophagus cancer/
2	((esophag* or oesophag*) adj6 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
3	or/1-2
4	exp Esophagus/
5	(esophag* or oesophag*).ti,ab.
6	Squamous cell carcinoma/
7	Adenocarcinoma/
8	(adenocarcinom* or scc).ab,ti.
9	(squam* adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
10	(4 or 5) and (6 or 7 or 8 or 9)
11	or/3,10
12	positron emission tomography/
13	(pet or petscan* or positron).ab,ti.
14	((pet* or tracer*) and (deoxy* or de oxy* or fluor* or fdg* or amino*)).ab,ti.
15	or/12-14
16	and/11,15

## 2. MEDLINE

### Suchoberfläche: Ovid

- Ovid MEDLINE(R) 1946 to May Week 2 2012
- Ovid MEDLINE(R) Daily Update May 17, 2012
- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations May 17, 2012

#	Searches
1	exp Esophageal Neoplasms/
2	((esophag* or oesophag*) adj6 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
3	or/1-2
4	exp Esophagus/
5	(esophag\$ or oesophag\$).ti,ab.
6	Carcinoma, Squamous Cell/
7	Adenocarcinoma/
8	(adenocarcinom* or scc).ab,ti.
9	(squam* adj3 (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)).ab,ti.
10	(4 or 5) and (6 or 7 or 8 or 9)
11	or/3,10
12	Positron-Emission Tomography/
13	Fluorodeoxyglucose F18/du
14	(pet or petscan* or positron).ab,ti.
15	((pet* or tracer*) and (deoxy* or de oxy* or fluor* or fdg* or amino*)).ab,ti.
16	or/12-15
17	and/11,16
18	animals/ not (animals/ and humans/)
19	17 not 18

## 3. PubMed

### Suchoberfläche: NLM

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

Search	Query
#1	Search (esophag*[TIAB] OR oesophag* [TIAB]) AND (tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB])
#2	Search esophag*[TIAB] OR oesophag*[TIAB]
#3	Search adenocarcinom*[TIAB] OR scc[TIAB]
#4	Search squam*[TIAB] AND (tumor*[TIAB] OR tumour*[TIAB] OR neoplas*[TIAB] OR malignan*[TIAB] OR cancer*[TIAB] OR carcinom*[TIAB])
#5	Search #2 AND (#3 OR #4)
#6	Search #1 OR #5
#7	Search pet[TIAB] OR petscan*[TIAB] OR positron[TIAB]
#8	Search (pet*[TIAB] OR tracer*[TIAB]) and (deoxy*[TIAB] OR de oxy*[TIAB] OR fluor*[TIAB] OR fdg*[TIAB] OR amino*[TIAB])
#9	Search #7 OR #8
#10	Search #6 AND #9
#11	Search #10 NOT medline[sb]

#### 4. The Cochrane Library

##### Suchoberfläche: Centre for Reviews and Dissemination

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

ID	Search
#1	MeSH DESCRIPTOR Positron-Emission Tomography EXPLODE ALL TREES WITH QUALIFIER undefined IN DARE,HTA
#2	MeSH DESCRIPTOR Fluorodeoxyglucose F18 EXPLODE ALL TREES WITH QUALIFIER undefined IN DARE,HTA
#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
#6	MeSH DESCRIPTOR Esophageal Neoplasms EXPLODE ALL TREES
#7	((esophag* or oesophag*) AND (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*))
#8	#6 OR #7
#9	MeSH DESCRIPTOR Esophagus EXPLODE ALL TREES
#10	(esophag* or oesophag*)
#11	MeSH DESCRIPTOR Carcinoma, Squamous Cell EXPLODE ALL TREES
#12	MeSH DESCRIPTOR Adenocarcinoma EXPLODE ALL TREES
#13	(adenocarcinom* or scc)
#14	squam* AND (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)
#15	#9 OR #10



ID	Search
#16	#11 OR #12 OR #13 OR #14
#17	#15 AND #16
#18	#8 OR #17
#19	#5 AND #18

### Suchoberfläche: Wiley

- Cochrane Database of Systematic Reviews (Cochrane Reviews), Issue 5, 2012
- Cochrane Central Register of Controlled Trials (Clinical Trials), Issue 5, 2012

ID	Search
#1	MeSH descriptor Positron-Emission Tomography explode all trees
#2	MeSH descriptor Fluorodeoxyglucose F18 explode all trees
#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	MeSH descriptor Esophageal Neoplasms explode all trees
#7	((esophag* or oesophag*) and (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)):ab,ti
#8	(#6 OR #7)
#9	MeSH descriptor Esophagus explode all trees
#10	(esophag* or oesophag*):ti,ab
#11	MeSH descriptor Carcinoma, Squamous Cell explode all trees
#12	MeSH descriptor Adenocarcinoma explode all trees
#13	(adenocarcinom* or scc):ab,ti
#14	(squam* and (tumor* or tumour* or neoplas* or malignan* or cancer* or carcinom*)):ab,ti
#15	(( #9 OR #10 ) AND ( #11 OR #12 OR #13 OR #14 ))
#16	(#8 OR #15)
#17	(#5 AND #16)

## A.2 Evidenzsynthesen

### 1. EMBASE

#### Suchoberfläche: Ovid

- Embase 1980 to 2010 February 12

Es wurde folgender Filter übernommen:

Systematic Review: Wong [167] - 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

- Ovid MEDLINE(R) 1946 to February Week 2 2010
- Ovid MEDLINE(R) Daily Update February 12, 2010
- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations February 12, 2010

Es wurde folgender Filter übernommen:

Systematic Review: Wong [167] - 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

## 3. PubMed

### Suchoberfläche: NLM

- PubMed - as supplied by publisher

- PubMed - in process
- PubMed – OLDMEDLINE
- PubMed – pubmednotmedline

Search	Query
#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

ID	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 (Cochrane Reviews), Issue 2, 2010

ID	Search
#1	MeSH descriptor Positron-Emission Tomography explode all trees
#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 (Technology Assessments), Issue 1, 2010

<b>ID</b>	<b>Search</b>
#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**

### **E1**

1. 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.healthcareimprovementscotland.org/his/idoc.ashx?docid=6c5b129c-a6e9-43a8-b731-c59229d89cad&version=-1>.
2. 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://gesdoc.isciii.es/gesdoccontroller?action=download&id=13/12/2012-6a26a98083>.
3. Rodriguez Garrido M, Asensio del Barrio C, Gomez Martinez MV, Carreras Delgado JL, Martin Moreno JM. Positron emission tomography (PET) with <sup>18</sup>FDG on clinical oncology: IPE-01/30 [Spanisch]. Madrid: Agencia de Evaluacion de Tecnologias Sanitarias; 2001. URL: <http://gesdoc.isciii.es/gesdoccontroller?action=download&id=19/10/2012-581d45027a>.

### **E2**

1. 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](http://aunets.isciii.es/ficherosproductos/164/2006_AETSA_F6_PETRM.pdf).

### **E3**

1. Institute for Clinical Evaluative Sciences. Health technology assessment of positron emission tomography (PET): a systematic review; ICES investigative report [online]. 05.2001 [Zugriff: 29.05.2013]. URL: [http://www.ices.on.ca/WebBuild/site/ices-internet-upload/file\\_collection/Health\\_Technology\\_Assessment-PET\\_May-2001.pdf](http://www.ices.on.ca/WebBuild/site/ices-internet-upload/file_collection/Health_Technology_Assessment-PET_May-2001.pdf).
2. 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%20i%20-%20PET%20Report.pdf](http://www.health.gov.au/internet/msac/publishing.nsf/Content/8FD1D98FE64C8A2FCA2575AD0082FD8F/$File/Ref%2010%20part%20i%20-%20PET%20Report.pdf).
3. Medicare Services Advisory Committee. Positron emission tomography: MSAC assessment report. Canberra: MSAC; 2000. URL: [http://www.msac.gov.au/internet/msac/publishing.nsf/Content/8FD1D98FE64C8A2FCA2575AD0082FD8F/\\$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).

4. 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: 29.05.2013]. URL: [http://www.juntadeandalucia.es/salud/servicios/contenidos/nuevaaetsa/up/AETSA\\_P\\_2000\\_PET.pdf](http://www.juntadeandalucia.es/salud/servicios/contenidos/nuevaaetsa/up/AETSA_P_2000_PET.pdf).
5. Unidad de Evaluacion de Tecnologias Sanitarias Agencia Lain Entralgo. Efficacy, safety and indications of hybrid PET-CT system [Spanisch] [online]. 10.2004 [Zugriff: 29.05.2013]. 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>.

#### E4

1. Basu S, Nair N. <sup>18</sup>F-DG PET in primary oat cell carcinoma of the esophagus. Indian J Cancer 2005; 42(1): 60-62.
2. Bourguet P. Standards, options and recommendations 2002 for the use of positron emission tomography with [<sup>18</sup>F]-FDG (PET-FDG in cancerology (integral connection) [Französisch]. Bull Cancer (Paris) 2003; 90(Spec No): S5-S17.
3. Comite d'Evaluation et de Diffusion des Innovations Technologiques. Positron emission tomography [online]. In: Hta. 03.03.1999 [Zugriff: 28.05.2013]. URL: <http://ancien-credit.aphp.fr/servlet/siteCeditGB?Destination=reco&numArticle=97.02.2>.
4. Comite d'Evaluation et de Diffusion des Innovations Technologiques. Positron emission tomography coupled with computer tomography (PET-CT) [online]. 12.12.2002 [Zugriff: 29.05.2013]. URL: <http://ancien-credit.aphp.fr/servlet/siteCeditGB?Destination=reco&numArticle=02.06.1>.
5. 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: 29.05.2013]. URL: <http://portal.saude.gov.br/portal/arquivos/pdf/Relat%201%20PET%20-%20Sintese%20das%20Agencias%20Internacionais%20de%20ATS.pdf>.
6. 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.
7. Gillham CM, Reynolds J, Hollywood D. Predicting the response of localised oesophageal cancer to neo-adjuvant chemoradiation. World J Surg Oncol 2007; 5: 97.

8. Institute for Clinical Evaluative Sciences. Health technology assessment of positron emission tomography (PET) in oncology: a systematic review; ICES investigative report; quarterly update September 2003 [online]. 09.2003 [Zugriff: 29.05.2013]. URL: [http://www.ices.on.ca/WebBuild/site/ices-internet-upload/file\\_collection/ACF17B.pdf](http://www.ices.on.ca/WebBuild/site/ices-internet-upload/file_collection/ACF17B.pdf).
9. Morland B. Positron emission tomography (PET): diagnostic and clinical use [Norwegisch]. Oslo: Norwegian Knowledge Centre for the Health Services; 2003. (SMM-Rapports; 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)_%96_diagnostisk%20og%20klinisk%20nytteverdi.pdf).
10. 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>.
11. 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.
12. 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 Assessments; Band 24).
13. Reske SN, Bares R, Bull U, Guhlmann A, Moser E, Wannenmacher MF. Klinische Wertigkeit der Positronen-Emissions-Tomographie (PET) bei onkologischen Fragestellungen: Ergebnisse einer interdisziplinären Konsensuskonferenz. Nuklearmedizin 1996; 35(2): 42-52.
14. Rodriguez Garrido M, Asensio del Barrio C, Alcazar Alcazar R. Use of FDG-PET in clinical situations not addressed in the monitorized use of this technology: IPE-06/49 [Spanisch]. Madrid: Agencia de Evaluacion de Tecnologias Sanitarias; 2006. URL: <http://gesdoc.isciii.es/gesdoccontroller?action=download&id=13/12/2012-c2fa9a5546>.
15. Rodriguez Garrido M, Conde Olasagasti J, Garcia Comas L, Alcaide JF. PET in non-neurological clinical oncology: IPE-97/11 [Spanisch]. Madrid: Agencia de Evaluacion de Tecnologias Sanitarias; 1997. URL: <http://gesdoc.isciii.es/gesdoccontroller?action=download&id=19/10/2012-c074231767>.
16. 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.
17. 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.

18. Van Westreenen HL, Westerterp M, Bossuyt PMM, Pruijm J, Sloof GW, Van Lanschot JJB et al. Systematic review of the staging performance of  $^{18}\text{F}$ -fluorodeoxyglucose positron emission tomography in esophageal cancer. *J Clin Oncol* 2004; 22(18): 3805-3812.
19. 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.
20. 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.

## E5

1. 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](http://aunets.isciii.es/ficherosproductos/186/AETSA_2006-15.pdf).

## E6

1. Blue Cross Blue Shield Association. FDG positron emission tomography for evaluating esophageal cancer. Chicago: BCBS; 2002.
2. Institute for Clinical Evaluative Sciences. Health technology assessment of positron emission tomography in oncology: a systematic review. Toronto: ICES; 2002.
3. 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.



**Anhang C – Liste der ausgeschlossenen Primärstudien mit Ausschlussgründen****E1**

1. Fischer MA, Nanz D, Hany T, Reiner CS, Stolzmann P, Donati OF et al. Diagnostic accuracy of whole-body MRI/DWI image fusion for detection of malignant tumours: a comparison with PET/CT. *Eur Radiol* 2011; 21(2): 246-255.
2. Lan XL, Zhang YX, Wu ZJ, Jia Q, Wei H, Gao ZR. The value of dual time point <sup>18</sup>F-FDG PET imaging for the differentiation between malignant and benign lesions. *Clin Radiol* 2008; 63(7): 756-764.
3. 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 <sup>18</sup>F-FDG PET. *Nucl Med (Stuttg)* 2007; 46(6): 263-270.
4. Nakaminato S, Torihara A, Makino T, Kawano T, Kishimoto S, Shibuya H. Prevalence of esophageal cancer during the pretreatment of hypopharyngeal cancer patients: routinely performed esophagogastroduodenoscopy and FDG-PET/CT findings. *Acta Oncol* 2012; 51(5): 645-652.
5. 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.

**E2**

1. Barber TW, Duong CP, Leong T, Bressel M, Drummond EG, Hicks RJ. <sup>18</sup>F-FDG PET/CT has a high impact on patient management and provides powerful prognostic stratification in the primary staging of esophageal cancer: a prospective study with mature survival data. *J Nucl Med* 2012; 53(6): 864-871.
2. 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.
3. Huguier M, Barrier A, Zacharias T, Valinas R. Positron emission tomography of gastrointestinal carcinomas [Französisch]. *Bull Acad Natl Med* 2006; 190(1): 75-84.
4. 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.
5. 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.

6. 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.
7. 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.
8. Smit JK, Pultrum BB, Van Dullemen HM, Van Dam GM, Groen H, Plukker JTM. Prognostic factors and patterns of recurrence in esophageal cancer assert arguments for extended two-field transthoracic esophagectomy. *Am J Surg* 2010; 200(4): 446-453.
9. Tixier F, Cheze Le Rest C, Hatt M, Albarghach N, Pradier O, Metges JP et al. Intratumor heterogeneity characterized by textural features on baseline <sup>18</sup>F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med* 2011; 52(3): 369-378.
10. Zacharias T, Barrier A, Montravers F, Houry S, Lacaine F, Huguier M. Cardio-esophageal cancer: is <sup>18</sup>Fluorodeoxyglucose positron emission tomography worthwhile? *Hepatogastroenterology* 2004; 51(57): 741-743.

### **E3**

1. Abdelsalam M, Bazarbashi S, Abouzied M, Amin T, Soudy H, Rahal M et al. Whole body <sup>18</sup>F-FDG pet predicts progression free and overall survival in squamous cell carcinoma of the esophagus: results of a prospective trial. *Hematol Oncol Stem Cell Ther* 2010; 3(4): 179-184.
2. Aoyagi T, Shuto K, Okazumi S, Shimada H, Nabeya Y, Kazama T et al. Evaluation of the clinical staging of esophageal cancer by using diffusion-weighted imaging. *Exp Ther Med* 2010; 1(5): 847-851.
3. Blom RL, Vliegen RF, Schreurs WM, Belgers HJ, Stohr I, Oostenbrug LE et al. External ultrasonography of the neck does not add diagnostic value to integrated positron emission tomography-computed tomography (PET-CT) scanning in the diagnosis of cervical lymph node metastases in patients with esophageal carcinoma. *Dis Esophagus* 2012; 25(6): 555-559.
4. Blom RLGM, Schreurs WMJ, Belgers HJ, Oostenbrug LE, Vliegen RFA, Sosef MN. The value of post-neoadjuvant therapy PET-CT in the detection of interval metastases in esophageal carcinoma. *Eur J Surg Oncol* 2011; 37(9): 774-778.
5. 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.
6. Downey RJ, Akhurst T, Ilson D, Ginsberg R, Bains MS, Gonen M et al. Whole body <sup>18</sup>FDG-PET and the response of esophageal cancer to induction therapy: results of a prospective trial. *J Clin Oncol* 2003; 21(3): 428-432.

7. 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.
8. 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.
9. 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.
10. Han D, Yu J, Zhong X, Fu Z, Mu D, Zhang B et al. Comparison of the diagnostic value of 3-deoxy-3-<sup>18</sup>F-fluorothymidine and <sup>18</sup>F-fluorodeoxyglucose positron emission tomography/computed tomography in the assessment of regional lymph node in thoracic esophageal squamous cell carcinoma: a pilot study. *Dis Esophagus* 2011; 25(5): 416-426.
11. 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.
12. 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.
13. Javeri H, Xiao L, Rohren E, Komaki R, Hofstetter W, Lee JH et al. Influence of the baseline <sup>18</sup>F-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.
14. Jingu K, Kaneta T, Nemoto K, Takeda K, Ogawa Y, Ariga H et al. <sup>18</sup>F-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.
15. 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.
16. 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.
17. Kobori O, Kirihara Y, Kosaka N, Hara T. Positron emission tomography of esophageal carcinoma using <sup>11</sup>C-choline and <sup>18</sup>F-fluorodeoxyglucose: a novel method of preoperative lymph node staging. *Cancer* 1999; 86(9): 1638-1648.

18. 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.
19. Li W, Fan W. Value of esophageal abnormal metabolism lesions of <sup>18</sup>F-FDG PET/CT for the clinical diagnosis of esophageal malignancy [Chinesisch]. *Chinese Journal of Medical Imaging Technology* 2009; 25(5): 901-904.
20. 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.
21. Lorenzen S, Panzram B, Keller G, Lordick F, Herrmann K, Becker K et al. Association of the VEGF 936C>T polymorphism with FDG uptake, clinical, histopathological, and metabolic response in patients with adenocarcinomas of the esophagogastric junction. *Mol Imaging Biol* 2011; 13(1): 178-186.
22. Lorenzen S, Von Gall C, Stange A, Haag GM, Weitz J, Haberkorn U et al. Sequential FDG-PET and induction chemotherapy in locally advanced adenocarcinoma of the oesophago-gastric junction (AEG): the Heidelberg imaging program in cancer of the oesophago-gastric junction during neoadjuvant treatment; HICON trial. *BMC Cancer* 2011; 11: 266.
23. 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.
24. 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.
25. 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.
26. 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.
27. Miyata H, Doki Y, Yasuda T, Yamasaki M, Higuchi I, Makari Y et al. Evaluation of clinical significance of <sup>18</sup>F-fluorodeoxyglucose positron emission tomography in superficial squamous cell carcinomas of the thoracic esophagus. *Dis Esophagus* 2008; 21(2): 144-150.
28. Moureau-Zabotto L, Touboul E, Lerouge D, Deniaud-Alexandre E, Grahek D, Foulquier JN et al. Impact of computed tomography (CT) and <sup>18</sup>F-deoxyglucose positron emission tomography (FDG-PET) image fusion for conformal radiotherapy in esophageal carcinoma [Französisch]. *Cancer Radiother* 2005; 9(3): 152-160.

29. Paul AK, Tatsumi M, Higuchi I, Fukunaga H, Yasuda T, Nishimura T. Gamma camera coincidence imaging with [<sup>18</sup>F]fluorodeoxyglucose in the pretreatment evaluation of patients with oesophageal cancer. *Nucl Med Commun* 2003; 24(9): 963-970.
30. Schreurs LM, Janssens AC, Groen H, Fockens P, Van Dullemen HM, Van Berge Henegouwen MI et al. Value of EUS in determining curative resectability in reference to CT and FDG-PET: the optimal sequence in preoperative staging of esophageal cancer? *Ann Surg Oncol* 06.05.2011 [Epub ahead of print].
31. 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.
32. 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.
33. 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.
34. Solomon N, Mezentsev D, Reis I, Lima M, Rios J, Avisar E et al. A phase II study of neoadjuvant and adjuvant chemotherapy with 5-fluorodeoxyuridine, leucovorin, oxaliplatin and docetaxel in the treatment of previously untreated advanced esophageal adenocarcinoma. *Jpn J Clin Oncol* 2011; 41(4): 469-476.
35. Van Heijl M, Omloo JM, Van Berge Henegouwen MI, Van Lanschot JJ, Sloof GW, Boellaard R. Influence of ROI definition, partial volume correction and SUV normalization on SUV-survival correlation in oesophageal cancer. *Nucl Med Commun* 2010; 31(7): 652-658.
36. 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.
37. 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.
38. Walker AJ, Spier BJ, Perlman SB, Stangl JR, Frick TJ, Gopal DV et al. Integrated PET/CT fusion imaging and endoscopic ultrasound in the pre-operative staging and evaluation of esophageal cancer. *Mol Imaging Biol* 2011; 13(1): 166-171.
39. Yan YY, Huang Y, Li WW. Research on correlation among maximum FDG PET standardized uptake value, expression of Ki-67 and pathology grading of esophageal carcinoma [Chinesisch]. *Shi Yong Zhong Liu Za Zhi* 2011; 26(5): 477-480.

40. 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.
41. 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.
42. Yue J, Chen L, Cabrera AR, Sun X, Zhao S, Zheng F et al. Measuring tumor cell proliferation with <sup>18</sup>F-FLT PET during radiotherapy of esophageal squamous cell carcinoma: a pilot clinical study. *J Nucl Med* 2010; 51(4): 528-534.
43. Zhang DQ, Fang N, Cui XJ. Applications of <sup>18</sup>F-FDG PET/CT in staging of esophageal carcinoma [Chinesisch]. *Zhonghua Zhong Liu Fang Zhi Za Zhi* 2009; 16(22): 1808-1810.
44. Zhang JD, Yu JM, Guo HB, Xing LG, Cao CJ. Positron emission tomography for diagnosis and staging of advanced esophagus carcinoma [Chinesisch]. *Chinese Journal of Cancer Prevention and Treatment* 2009; 16(23): 1875-1877.
45. Zum Buschenfelde CM, Herrmann K, Schuster T, Geinitz H, Langer R, Becker K et al. <sup>18</sup>F-FDG PET-guided salvage neoadjuvant radiochemotherapy of adenocarcinoma of the esophagogastric junction: the MUNICON II trial. *J Nucl Med* 2011; 52(8): 1189-1196.

#### E4

1. Surgical Oncology Group announces PET study. *J Nucl Med* 2000; 41(2): 26N.
2. Abate E, DeMeester SR, Zehetner J, Oezcelik A, Ayazi S, Costales J et al. Recurrence after esophagectomy for adenocarcinoma: defining optimal follow-up intervals and testing. *J Am Coll Surg* 2010; 210(4): 428-435.
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.
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.
5. Berger AC, Scott WJ. Noninvasive staging of esophageal carcinoma. *J Surg Res* 2004; 117(1): 127-133.
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.
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.

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.
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.
10. Bourguet P, Blanc-Vincent MP, Boneu A, Bosquet L, Chauffert B, Corone C et al. Use of positron emission tomography with [<sup>18</sup>F]-FDG in oncology: results in digestive cancers [Französisch]. *Bull Cancer (Paris)* 2003; 90(2 Spec Iss): S56-S66.
11. Brown C, Howes B, Jamieson GG, Bartholomeusz D, Zingg U, Sullivan TR et al. Accuracy of PET-CT in predicting survival in patients with esophageal cancer. *World J Surg* 2012; 36(5): 1089-1095.
12. Centro de Estudos e Pesquisa em Saude Coletiva. Rapid HTA on the use of positron emission tomography (PET) in the diagnosis, staging and re-staging of esophageal cancer [Portugiesisch] [online]. 2009 [Zugriff: 28.05.2013]. URL: <http://portal2.saude.gov.br/rebrats/visao/estudo/detEstudo.cfm?codigo=190&evento=6&v=true>.
13. 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.
14. Flanagan FL, Dehdashti F, Siegel BA, Trask DD, Sundaresan SR, Patterson GA et al. Staging of esophageal cancer with <sup>18</sup>F-fluorodeoxyglucose positron emission tomography. *Am J Roentgenol* 1997; 168(2): 417-424.
15. Ganeshan B, Skogen K, Pressney I, Coutroubis D, Miles K. Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. *Clin Radiol* 2012; 67(2): 157-164.
16. 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.
17. Hsu PK, Lin KH, Wang SJ, Huang CS, Wu YC, Hsu WH. Preoperative positron emission tomography/computed tomography predicts advanced lymph node metastasis in esophageal squamous cell carcinoma patients. *World J Surg* 2011; 35(6): 1321-1326.
18. 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.
19. 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.

20. 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.
21. Ishihara R, Yamamoto S, Iishi H, Nagai K, Matui F, Kawada N et al. Predicting the effects of chemoradiotherapy for squamous cell carcinoma of the esophagus by induction chemotherapy response assessed by positron emission tomography: toward PET-response-guided selection of chemoradiotherapy or esophagectomy. *Int J Clin Oncol* 2011; 17(3): 225-232.
22. 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.
23. Jayachandran P, Pai RK, Quon A, Graves E, Krakow TE, La T et al. Postchemoradiotherapy positron emission tomography predicts pathologic response and survival in patients with esophageal cancer. *Int J Radiat Oncol Biol Phys* 2012; 84(2): 471-477.
24. 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.
25. 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.
26. 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.
27. Kauppi JT, Oksala N, Salo JA, Helin H, Karhumaki L, Kemppainen J et al. Locally advanced esophageal adenocarcinoma: response to neoadjuvant chemotherapy and survival predicted by  $^{18}\text{F}$ FDG-PET/CT. *Acta Oncol* 2012; 51(5): 636-644.
28. Kobayashi M, Kaida H, Kawahara A, Hattori S, Kurata S, Hayakawa M et al. The relationship between GLUT-1 and vascular endothelial growth factor expression and  $^{18}\text{F}$ -FDG uptake in esophageal squamous cell cancer patients. *Clin Nucl Med* 2012; 37(5): 447-452.
29. 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.



30. 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.
31. Li J, Zhu S, Liu Z, Su J, Shen W. Comparison of  $^{18}\text{F}$ -FDG-PET and  $^{18}\text{F}$ -FDG-CT scans for the clinical staging of esophageal cancer [Chinesisch]. *Chinese Journal of Clinical Oncology* 2011; 38(23): 1449-1452.
32. 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.
33. Makino T, Miyata H, Yamasaki M, Fujiwara Y, Takiguchi S, Nakajima K et al. Utility of response evaluation to neo-adjuvant chemotherapy by  $^{18}\text{F}$ -fluorodeoxyglucose-positron emission tomography in locally advanced esophageal squamous cell carcinoma. *Surgery* 2010; 148(5): 908-918.
34. Marzola MC, De Manzoni G, Grassetto G, Cordiano C, Al-Nahhas A, Alavi A et al. Extended staging of oesophageal cancer using FDG-PET: a critical appraisal. *Eur J Radiol* 2012; 81(1): 21-30.
35. McAteer D, Wallis F, Couper G, Norton M, Welch A, Bruce D et al. Evaluation of  $^{18}\text{F}$ -FDG positron emission tomography in gastric and oesophageal carcinoma. *Br J Radiol* 1999; 72(858): 525-529.
36. 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.
37. Monjazez AM, Riedlinger G, Aklilu M, Geisinger KR, Mishra G, Isom S et al. Outcomes of patients with esophageal cancer staged with [ $^{18}\text{F}$ ]fluorodeoxyglucose positron emission tomography (FDG-PET): can postchemoradiotherapy FDG-PET predict the utility of resection? *J Clin Oncol* 2010; 28(31): 4714-4721.
38. 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.
39. Muijs CT, Beukema JC, Pruim J, Mul VE, Groen H, Plukker JT et al. A systematic review on the role of FDG-PET/CT in tumour delineation and radiotherapy planning in patients with esophageal cancer. *Radiother Oncol* 2010; 97(2): 165-171.
40. Murthy SB, Patnana SV, Xiao L, Rohren E, Hofstetter WL, Swisher SG et al. The standardized uptake value of  $^{18}\text{F}$ -fluorodeoxyglucose positron emission tomography after chemoradiation and clinical outcome in patients with localized gastroesophageal carcinoma. *Oncology* 2010; 78(5-6): 316-322.
41. 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.

42. Omloo JMT, Van Heijl M, Hoekstra OS, Van Berge Henegouwen MI, Van Lanschot JJB, Sloof GW. FDG-PET parameters as prognostic factor in esophageal cancer patients: a review. *Ann Surg Oncol* 2011; 18(12): 3338-3352.
43. Rankin S. The value of [<sup>18</sup>F]fluorodeoxyglucose-PET/CT in oesophageal cancer. *Cancer Imaging* 2011; 11(Special Issue A): S156-S160.
44. 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.
45. 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.
46. 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.
47. 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.
48. Sharma NK, Silverman JS, Li T, Cheng J, Yu JQ, Haluszka O et al. Decreased posttreatment SUV on PET scan is associated with improved local control in medically inoperable esophageal cancer. *Gastrointest Cancer Res* 2011; 4(3): 84-89.
49. Sohda M, Kato H, Suzuki S, Tanaka N, Sano A, Sakai M et al. <sup>18</sup>F-FAMT-PET is useful for the diagnosis of lymph node metastasis in operable esophageal squamous cell carcinoma. *Ann Surg Oncol* 2010; 17(12): 3181-3186.
50. Sun L, Su XH, Guan YS, Pan WM, Luo ZM, Wei JH et al. Clinical usefulness of <sup>18</sup>F-FDG PET/CT in the restaging of esophageal cancer after surgical resection and radiotherapy. *World J Gastroenterol* 2009; 15(15): 1836-1842.
51. Sun L, Wan Y, Lin Q, Sun YH, Zhao L, Luo ZM et al. Multiple primary malignant tumors of upper gastrointestinal tract: a novel role of <sup>18</sup>F-FDG PET/CT. *World J Gastroenterol* 2010; 16(31): 3964-3969.
52. Suzuki A, Xiao L, Hayashi Y, Blum MA, Welsh JW, Lin SH et al. Nomograms for prognostication of outcome in patients with esophageal and gastroesophageal carcinoma undergoing definitive chemoradiotherapy. *Oncology* 2012; 82(2): 108-113.
53. 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.

54. Tanabe S, Naomoto Y, Shirakawa Y, Fujiwara Y, Sakurama K, Noma K et al. F-18 FDG PET/CT contributes to more accurate detection of lymph nodal metastasis from actively proliferating esophageal squamous cell carcinoma. *Clin Nucl Med* 2011; 36(10): 854-859.
55. 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.
56. 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.
57. 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.
58. Wang LD, Du XY, Feng R, Zhang Q, Wang YS, Sun XD et al. Correlation between FDG PET/CT metabolic activity volume and prognosis in esophageal carcinoma [Chinesisch]. *Chinese Journal of Cancer Prevention and Treatment* 2011; 18(15): 1186-1188.
59. Welsh J, Settle SH, Amini A, Xiao L, Suzuki A, Hayashi Y et al. Failure patterns in patients with esophageal cancer treated with definitive chemoradiation. *Cancer* 2012; 118(10): 2632-2640.
60. 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.
61. Wren SM, Stijns P, Srinivas S. Positron emission tomography in the initial staging of esophageal cancer. *Arch Surg* 2002; 137(9): 1001-1006.
62. Yanagawa M, Tatsumi M, Miyata H, Morii E, Tomiyama N, Watabe T et al. Evaluation of response to neoadjuvant chemotherapy for esophageal cancer: PET response criteria in solid tumors versus response evaluation criteria in solid tumors. *J Nucl Med* 2012; 53(6): 872-880.
63. Yang Y, Yang Z, Han A, Feng R, Ma Y, Kong L et al. Prediction of the outcome of definitive chemoradiation by decrease in F-18 FDG uptake in nonsurgical esophageal squamous cell cancer. *Clin Nucl Med* 2011; 36(10): 860-866.
64. Yasuda T, Higuchi I, Yano M, Miyata H, Yamasaki M, Takiguchi S et al. The impact of <sup>18</sup>F-fluorodeoxyglucose positron emission tomography positive lymph nodes on postoperative recurrence and survival in resectable thoracic esophageal squamous cell carcinoma. *Ann Surg Oncol* 2012; 19(2): 652-660.
65. Yu W, Fu XL, Zhang YJ, Xiang JQ, Shen L, Chang JY. A prospective evaluation of staging and target volume definition of lymph nodes by <sup>18</sup>FDG PET/CT in patients with squamous cell carcinoma of thoracic esophagus. *Int J Radiat Oncol Biol Phys* 2011; 81(5): e759-e765.

## E5

1. Abdelsalam M, Bazarbashi S, Amin T, Soudy H, Memon M, Rahal M et al. Whole body <sup>18</sup>F-FDG-PET predicts progression free and overall survival in patients with squamous cell carcinoma of the esophagus: a prospective trial. *Ann Oncol* 2008; 19(Suppl 8): viii171.
2. Lordick F, Meyer zum Bueschenfelde C, Herrmann K, Geinitz H, Schuster T, Friess H et al. PET-guided treatment in locally advanced adenocarcinoma of the esophagogastric junction (AEG): the MUNICON-II study. *J Clin Oncol* 2011; 29(4 Suppl): 3.
3. Van Heijl M, Omloo JM, Van Berge Henegouwen MI, Hoekstra OS, Boellaard R, Bossuyt PM et al. Early response assessment of neoadjuvant therapy with positron emission tomography in patients with oesophageal cancer. *Gastroenterology* 2010; 138(5 Suppl 1): S512.

## E6

1. FDG positron emission tomography for evaluating esophageal cancer. *Tecnologica MAP Suppl* 2001: 13-16.
2. Aigner C, Moons J, Naftoux P, Coosemans W, Decaluwe H, Decker G et al. The value of PET scan in the clinical N staging of locally advanced oesophageal cancer. *Dis Esophagus* 2010; 23(Suppl s1): 45A.
3. Aigner C, Moons J, Naftoux P, Coosemans W, Decaluwe H, Decker G et al. The value of PET scan in the clinical N staging of locally advanced oesophageal cancer. *Eur Surg* 2011; 43(Suppl 241): 31-32.
4. Alam N, Bayam F, Kneale E, Wiesmann H. The association between PET CT SUVmax values and patient mortality in primary oesophageal cancer. *Cancer Imaging* 2011; 11(Special Issue A): S176.
5. Alan MG, Sallum RAA, Meneghetti JS, Szachnowicz S, Takeda FR, Cecconello I. Impact of FDG-PETscan in esophageal cancer staging: analysis of consecutive 96 cases; a prospective study. *Dis Esophagus* 2010; 23(Suppl s1): 88A-89A.
6. Atturni K, Shioyama Y, Ohga S, Yoshitake T, Nonoshita T, Ohnishi K et al. Prediction of outcome with <sup>18</sup>F-FDG-PET in definitive chemoradiotherapy for locally esophageal cancer. *Radiother Oncol* 2010; 96(Suppl 1): S287.
7. Attia A, Vern-Gross T, Garg PK, Garg S, Thomas A, Smith H et al. The utility of <sup>11</sup>C-choline positron emission tomography computerized tomography (PET-CT) for staging locally advanced esophageal cancer. *Int J Radiat Oncol Biol Phys* 2011; 81(2 Suppl): S314.
8. Bain G, Murray G, Denison A, Brooks M, Collie-Duguid E, McKiddie F et al. Aberdeen microarray and PET in optimising oesophagogastric cancer response-1 (AMPETOOR-1): preliminary results. *Ann Oncol* 2010; 21(Suppl 6): vi43.

9. Blom R, Schreurs W, Lammering G, Oostenbrug L, Nap M, Sosef M. Neoadjuvant therapy in esophageal carcinoma followed by a long interval until resection: a high incidence of complete pathological response. *Dis Esophagus* 2010; 23(Suppl s1): 20A-21A.
10. Celiz P, Galan M, Robles JJ, Sanchez JJ, Gamez C, Liberal JM et al. Role of FDG-PET/CT in predicting pathological response and survival in locally advanced esophageal cancer patients (p) treated with neoadjuvant chemoradiotherapy. *European Journal of Cancer Supplements* 2009; 7(2): 381.
11. Chang JT, Lin CY, Fan KH, Wang HM, Liao CT, Kang K et al. The treatment result of advanced stage oropharyngeal cancer by radiotherapy with or without chemotherapy: the impact of intensity modulation radiotherapy and FDG-PET. *Eur J Cancer* 2011; 47(Suppl 1): S555.
12. Chionh F, Handolias D, Poon A, Gebiski V, Dayan S, Aly A et al. The decrease in standardised uptake value (SUVmax) on <sup>18</sup>F-FDG-PET after neoadjuvant chemotherapy is a prognostic factor for survival and recurrence in patients with gastro-oesophageal cancer. *Asia Pac J Clin Oncol* 2009; 5(Suppl 1): A61-A62.
13. Cools-Lartigue J, Hickeson M, Edwards M, Alcindor T, Ades S, Marcus V et al. PET response after neo-adjuvant chemotherapy predicts recurrence post resection of esophago-gastric adenocarcinoma: the utility of volume based parameters of metabolic activity. *Gastroenterology* 2011; 140(5 Suppl 1): S671.
14. Cruz A, Chuong M, Weber J, Eikman E, Lomas H, Pimiento JM et al. Post-radiation metabolic tumor volume (MTV) and survival in esophageal cancer. *Ann Surg Oncol* 2012; 19(1 Suppl): S162.
15. Cuenca X, Hindie E, Quero L, Hennequin V, Moretti JL, Maylin C et al. Metabolic response with [<sup>18</sup>F] fluorodeoxyglucose (FDG) positron emission tomography (PET) scanning during chemoradiotherapy (RT-CT) of oesophageal cancers: feasibility and prognostic value. *EJC Supplements* 2009; 7(2): 374.
16. Emi M. The value of FDG-PET in assessment of induction therapy for esophageal squamous cell carcinoma. *Dis Esophagus* 2010; 23(Suppl s1): 65A.
17. Emi M, Hihara J, Hamai Y, Aoki Y, Sakogawa K, Okada M. Post-chemoradiotherapy standardized uptake value of FDG-PET as a significant predictor of survival after subsequent surgery in multimodality treatment for patients with locally advanced esophageal squamous cell carcinoma. *Eur J Cancer* 2011; 47(Suppl 1): S450.
18. Fiorillo G, Evangelista L, Alfieri R, Castoro C, Gregianin M, Saladini G et al. Changes in total lesion glycolysis (TLG) and in standardized uptake value (SUV) on <sup>18</sup>F-FDG PET/CT and the prognosis in esophageal cancer. *Eur J Nucl Med Mol Imaging* 2011; 38(Suppl 2): S195.
19. Gockel I, Sgourakis G, Domeyer M, Hansen T, Junginger T, Lang H. Prediction of lymph node metastases in early esophageal cancer. *Gastroenterology* 2009; 136(5 Suppl 1): A912.

20. Guo H, Yu J, Zhang B, Li H, Wakelee HA. The influence of FDG-PET/CT on CT-based surgical management for patients with squamous-cell cancer at thoracic esophagus. *J Clin Oncol* 2010; 28(15 Suppl): e13550.
21. Han D, Yu J, Zhong X, Mu D, Fu Z, Zhang B et al. Compare the diagnostic value of FLT and FDG PET/CT in assessment of regional lymph nodes in thoracic esophageal squamous cell carcinoma. *J Clin Oncol* 2009; 27(15 Suppl): e15620.
22. Han D, Yu JM, Fu Z, Wang X, Zhao SQ. Correlation of 3'-deoxy-3'-<sup>18</sup>F-fluorothymidine PET/CT uptake to DNA ploid, S phase fraction, and proliferation index in esophageal squamous cell carcinoma. *J Clin Oncol* 2011; 29(15 Suppl): e21035.
23. Hirose Y, Kaida H, Kurata S, Kobayashi M, Kawahara A, Hattori S et al. The relationship between <sup>18</sup>F-FDG uptake and GLUT-1, GLUT-3, CD34, CD68 expressions in esophageal squamous cell carcinoma patients. *Eur J Nucl Med Mol Imaging* 2011; 38(Suppl 2): S361-S362.
24. Hui J, Yu J, Song P, Zhang B, Zhang W, Mu D et al. Comparison of <sup>18</sup>F-fluorothymidine and <sup>18</sup>F-fluorodeoxyglucose PET/CT in assessment of chemoradioresponse in esophageal squamous cell cancer: a pilot study. *Int J Radiat Oncol Biol Phys* 2011; 81(2 Suppl): S310.
25. Jayachandran P, Krakow T, La T, Loo BW, Koong AC, Chang DT. Postchemoradiotherapy positron emission tomography for predicting survival in patients with esophageal cancer. *Int J Radiat Oncol Biol Phys* 2011; 81(2 Suppl): S310.
26. Jung K, Romero Y, Prasad GA, Wang KK, Dunagan KT. The effectiveness of FDG-PET in the detection of metastasis: a specific emphasis on intramucosalesophageal cancer. *Dis Esophagus* 2010; 23(Suppl s1): 71A.
27. Kaida H, Kobayashi M, Ishibashi M, Kurata S, Hattori S, Kawahara A et al. The relationship between biological factors and <sup>18</sup>F-FDG uptake in esophageal squamous cell carcinoma patients. *Intern Med J* 2010; 40(Suppl 2): 30.
28. Karashima R, Watanabe M, Nagai Y, Kurashige J, Kinoshita K, Saito S et al. Advantages of positron emission tomography/computed tomography in preoperative lymph node assessment of esophageal cancer. *Dis Esophagus* 2010; 23(Suppl s1): 27A.
29. Koyanagi K, Tabuchi S, Nagata K, Koyama I. Clinical utility of FDG-PET/CT for the response evaluation of neoadjuvant chemotherapy in patients with esophageal squamous cell carcinoma. *Dis Esophagus* 2010; 23(Suppl s1): 34A.
30. Liu L, Huang Y, Guo J, Wang ZH, Song DG. Correlation between FDG PET/CT and the expression of Ki-67, MMP-2, micro-vessel density (MVD), and pathological grading in squamous cell carcinoma of the esophagus. *J Clin Oncol* 2009; 27(15 Suppl): e15556.
31. Lomas H, Weber J, Chuong M, Hoffe SE, Almhanna K, Karl RC et al. Post chemoradiation SUV is highly predictive of disease free survival and overall survival in esophageal cancer. *Int J Radiat Oncol Biol Phys* 2011; 81(2 Suppl): S308.

32. Moon J. Utility of response evaluation to neo-adjuvant chemotherapy by <sup>18</sup>F-fluorodeoxyglucose-positron emission tomography in locally advanced esophageal squamous cell carcinoma. *Dis Esophagus* 2010; 23(Suppl s1): 100A.
33. Nafteux P, Moons J, Coosemans W, Decaluwe H, Decker G, De Leyn P et al. Comparison of clinical lymph node staging (cN) with pathological lymph node staging (pN) in locally advanced (cT2-3) distal third adenocarcinoma of the esophagus. *Dis Esophagus* 2010; 23(Suppl s1): 69A.
34. Narushima K. Histological evaluation of carbon ion radiotherapy and chemoradiotherapy for esophageal squamous cell carcinoma using PET. *Dis Esophagus* 2010; 23(Suppl s1): 92A.
35. Okazumi S, Kato R, Kinoshita T, Shimada H, Shuto K, Kono T et al. Clinical application of qualitative diagnostic method using MD-CT and FDG-PET to the decision making of treatment strategy for the advanced esophageal cancer. *Dis Esophagus* 2010; 23(Suppl s1): 40A.
36. Park S, Rhu J, Park S, Kim Y, Jung H, Lee J et al. <sup>18</sup>F-FLT PET-CT for early treatment response after 2 cycles of induction chemotherapy with S-1/oxaliplatin in patients with resectable esophageal cancer treated with preoperative chemoradiotherapy. *Dis Esophagus* 2010; 23(Suppl s1): 54A.
37. Sepesi B, Polomsky M, Watson TJ, Jones CE, Litle VR, Raymond D et al. Does the value of pet-ct extend beyond pretreatment staging? An analysis of survival in patients with esophageal cancer. *Gastroenterology* 2009; 136(5 Suppl 1): A898.
38. Sergeant G, Deroose C, De Hertogh G, Moons J, Coosemans W, Nafteux P et al. Early metabolic response evaluation with PET-CT after a single cycle of chemotherapy in patients with advanced oesophageal cancer subsequently treated by neoadjuvant chemoradiotherapy. *Dis Esophagus* 2010; 23(Suppl s1): 69A.
39. Sergeant G, Deroose C, De Hertogh G, Moons J, Coosemans W, Nafteux P et al. Early metabolic response evaluation on PET-CT after a single cycle of chemotherapy in patients with cT3-4N0/+ oesophageal or GE-junction cancer subsequently treated by neoadjuvant chemoradiotherapy. *J Clin Oncol* 2010; 28(15 Suppl): e14505.
40. Shuto K, Saito H, Kono T, Ohira G, Natsume T, Toma T et al. Preoperative N-staging of esophageal squamous cell cancer by DWIBS in comparison with PET. *Dis Esophagus* 2010; 23(Suppl s1): 85A.
41. Singhal R, Garman K, Vamvakopoulos J, Guest P, Tucker O, Alderson D. The role of CT-PET in staging junctional oesophagogastric adenocarcinomas. *Dis Esophagus* 2010; 23(Suppl s1): 54A.
42. Small C, Joseph K, McEwan A, Wachowicz K, Riauka T, Bedard E et al. Diffusion weighted magnetic resonance, FDG-PET and <sup>18</sup>F FAZA PET imaging in predicting response to neoadjuvant chemo-radiotherapy in oesophageal carcinoma. *Radiother Oncol* 2010; 96(Suppl 2): S44.

43. Tabata K, Miyata K, Yuasa N, Takeuchi E, Goto Y, Miyake H et al. Correlation between proliferating cell index and clinicopathological features and prognosis in squamous cell carcinoma of the esophagus. *Dis Esophagus* 2010; 23(Suppl s1): 27A.
44. Vallbohmer D, A HH, Brabender J, Metzger R, Monig S, Schmidt M et al. The value of early and late response evaluation by [18F]-fluorodeoxyglucosepositron emission tomography in the neoadjuvant therapy of patients with esophageal cancer. *Dis Esophagus* 2010; 23(Suppl s1): 47A.
45. Vallbohmer D, Holscher AH, Brabender J, Metzger R, Monig S, Schmidt M et al. Evaluation of response in the multimodality treatment of esophageal cancer by [18F]-fluorodeoxyglucose-positron emission tomography: a systematic review and meta-analysis. *J Clin Oncol* 2009; 27(15 Suppl): e15520.
46. Van Heijl M, Omloo JMT, Van Berge Henegouwen MI, Hoekstra OS, Boellaard R, Bossuyt PMM et al. FDG-PET for evaluating response early during neoadjuvant chemoradiotherapy in patients with potentially curable esophageal cancer. *Dis Esophagus* 2010; 23(Suppl s1): 11A-12A.
47. Vecchione L, De Vita F, Farella A, Martinelli E, Orditura M, Innocente R et al. The role of PET-TC in predicting the pCR in locally-advanced esophageal cancer (LAEC) after a preoperative CT-RT treatment: data from B152 trial. *EJC Supplements* 2009; 7(2): 380.
48. Watanabe M, Nagai Y, Iwagami S, Miyamoto Y, Hayashi N, Baba H. Phase II study of induction chemotherapy with docetaxel, cisplatin, and 5-fluorouracil for patients with node-positive esophageal cancer. *J Clin Oncol* 2010; 28(15 Suppl): e14535.
49. Wilson VL, Ballantyne S, Fullarton G. The role of PET/CT in identifying advanced disease in patients with potentially curable oesophageal carcinoma. *Gastroenterology* 2010; 138(5 Suppl 1): S899.
50. Won Lee H, Wook Hwang J, Hyun Kim J, Jong Baek H, Ho Park J. Clinical review of necessity of cervical lymph node dissection for thoracic esophageal squamous cell carcinoma. *Dis Esophagus* 2010; 23(Suppl s1): 55A-56A.
51. Yu J, Han D, Zhong X, Mu D, Fu Z, Zhang B et al. The optimal threshold of <sup>18</sup>F-FLT PET and <sup>18</sup>F-FDG PET to estimate the length of gross tumor volume in patients with squamous cell carcinoma of the thoracic esophagus verified by pathological examination. *J Clin Oncol* 2009; 27(15 Suppl): e15665.
52. Yu J, Han D, Zhong X, Mu D, Fu Z, Zhang B et al. Comparing the diagnostic value of FLT and FDG PET/CT in assessment of regional lymph nodes in thoracic esophageal squamous cell carcinoma. *Biomed Pharmacother* 2009; 63(5): 322-323.
53. Yue J, Yu J, Cabrera A, Chen L, Sun X, Zhao S et al. Measuring tumor cell proliferation and predicting clinical response with <sup>18</sup>F-FLT PET during radiotherapy of esophageal squamous cell carcinoma: a pilot clinical study. *Int J Radiat Oncol Biol Phys* 2010; 78(3 Suppl): S162.



54. Zemanova M, Harustiak T, Fencel P, Pechacova Z, Pazdrova G, Petruzela L. FDG-PET/CT evaluation of early metabolic response after neoadjuvant chemotherapy in patients with esophageal and GEJ cancer: a pilot study. *Ann Oncol* 2010; 21(Suppl 8): viii253.

55. Zhong X, Han D, Yu J, Fu Z, Mu D, Yang W. The assessment value of FLT and FDG PET/CT for lymph node staging in thoracic esophageal squamous cell carcinoma. *J Clin Oncol* 2010; 28(15 Suppl): e14599.

56. Zhu WQ, Li MH, Sun XR, Sun XD, Xing LG, Kong L et al.  $^{18}\text{F}$ -FDG uptake by primary tumor as a predictor of lymph node involvement in clinical N<sub>0</sub> patients with esophageal squamous cell carcinoma. *Int J Radiat Oncol Biol Phys* 2010; 78(Suppl 3): S332.

## E7

1. Brucher 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.

2. 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.

3. 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.

4. Kroep JR, Van Groeningen CJ, Cuesta MA, Craanen ME, Hoekstra OS, Comans EFI et al. Positron emission tomography using 2-deoxy-2- $^{18}\text{F}$ -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.

5. Levine EA, Farmer MR, Clark P, Mishra G, Ho C, Geisinger KR et al. Predictive value of 18-fluoro-deoxy-glucose-positron emission tomography ( $^{18}\text{F}$ -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.

6. 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.

7. 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.

8. 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.

9. Westerterp M, Omlöo 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.

**Anhang D – Beurteilung der Qualität von Evidenzsynthesen**

**Anhang D.1 – Qualitätsbeurteilung von Evidenzsynthesen nach den Kriterien von Oxman und Guyatt [49,50]**

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**

<b>Evidenz-synthesen</b>	<b>OG 1</b>	<b>OG 2</b>	<b>OG 3</b>	<b>OG 4</b>	<b>OG 5</b>	<b>OG 6</b>	<b>OG 7</b>	<b>OG 8</b>	<b>OG 9</b>	<b>Gesamt-punkte</b>
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 [29] 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 [132] 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"/>	--
<b>Gesamtbewertung des Verzerrungspotenzials</b>			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"/>	--
<b>Gesamtbewertung der Übertragbarkeit</b>		Die Übertragbarkeit wird als gegeben angesehen	<input type="checkbox"/>		
		Es gibt Indizien für eine eingeschränkte Übertragbarkeit	<input type="checkbox"/>		

**Anhang D.5 – Bewertung des Verzerrungspotenzials von prognostischen Primärstudien**

Item	Frage	ja	nein	unklar
<b>Studienpopulation</b>				
1	Waren die Selektionskriterien für die Patienten klar beschrieben?	<input type="checkbox"/>	<input type="checkbox"/>	--
2	Befanden sich die Patienten in einer vergleichbaren Stadium 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"/>
<b>Nachbeobachtung / Patientenfluss</b>				
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? <sup>a</sup>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>Messung prognostischer Faktoren</b>				
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"/>
<b>Messung der Endpunkte</b>				
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"/>
<b>Statistische Analyse / Umgang mit Störgrößen (Confounder)</b>				
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"/>
<b>Sonstige Aspekte</b>				
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"/>
<b>Gesamtbewertung des Verzerrungspotenzials</b>		<input type="checkbox"/>	<input type="checkbox"/>	
		hoch	niedrig	
a: Eine Follow-up-Rate $\geq 80$ % wird als adäquat betrachtet.				

**Anhang D.6 – Einschätzung der Übertragbarkeit der Studienergebnisse von 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"/>	--
<b>Gesamtbewertung der Übertragbarkeit</b>		Die Übertragbarkeit wird als gegeben angesehen			<input type="checkbox"/>
		Es gibt Indizien 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
<b>Primärstaging</b>		
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.
Gillies 2012a	Krankheitsfreies Überleben	1) Alter: einfaktorielle Analyse [k. A.]; multifaktoriell [p = 0,018]; 2) pathologisches Stadium: einfaktorielle Analyse [k. A.]; multifaktoriell [p < 0,001]; 3) lymphathische Streuung: einfaktorielle Analyse [k. A.]; multifaktoriell [p = 0,012].
	Gesamtüberleben	1) Alter: einfaktorielle Analyse [k. A.]; multifaktoriell [p = 0,041]; 2) pathologisches Stadium: einfaktorielle Analyse [k. A.]; multifaktoriell [p < 0,001]; 3) lymphathische Streuung: einfaktorielle Analyse [k. A.]; multifaktoriell [p = 0,029].
Omloo 200 <sup>j</sup>	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	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.].  Ebenfalls einfaktoriell modelliert, aber mit p-Wert > 0,05: Geschlecht, Alter, Lokation (primary site, Klinisches Stadium, CT Diagnose, PET Diagnose).  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
<b>Restaging</b>		
Kim 2007	Krankheitsfreies Überleben	Ebenfalls einfaktoriell modelliert, aber mit p-Wert > 0,05: Geschlecht, Alter, Lokation (primary site, klinisches Stadium, histologische Differenzierung, SUV <sub>max</sub> des Primärtumors).
Ott 2006	Gesamtüberleben	Es wurden außerdem modelliert: Geschlecht, grading, lauren classification, localization, aber p-Wert > 0,1,(vermutlich einfaktoriell).  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];
	Rezidivfreies Überleben	ypN0: einfaktorielle Analyse [p = 0,09]; multifaktoriell [p > 0,05].
<b>Rezidivdiagnostik</b>		
Guo 2007	Gesamtüberleben	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];  Ebenfalls einfaktoriell modelliert, aber mit p-Wert > 0,05: Geschlecht, Alter, Tumordifferenzierung, klinisches Stadium. Diese Variablen gehen nicht in die multifaktorielle Analyse ein.
AJCC = American Joint Committee on Cancer; CT = Computertomographie; EUS = Endosonografie; FDG = 2-[ <sup>18</sup> F]Fluor-Deoxy-D-Glukose ; SUV = standardized uptake value; OP = Operation		

**Anhang F – TNM-7-Klassifikation**

**Tabelle 65: Stadieneinteilung nach dem TNM-System 7 [15]**

<b>T Primärtumor</b>	
TX	Primärtumor nicht beurteilbar
T0	kein Nachweis eines Primärtumors
Tis	hochgradige Dysplasie
T1	Tumor infiltriert Lamina propria, Lamina muscularis mucosae und Lamina submucosa
T1a	Tumor infiltriert Lamina propria oder Lamina muscularis mucosae
T1b	Tumor infiltriert Lamina submucosa
T2	Tumor infiltriert die Muscularis propria
T3	Tumor infiltriert die Adventitia
T4	Tumor infiltriert angrenzende extraösophageale Strukturen
T4a	resektabler Tumor mit Infiltration von Pleura, Perikard oder Zwerchfell
T4b	nicht resektabler Tumor mit Infiltration von anderen angrenzenden Strukturen wie Aorta, Wirbelkörpern, Luftröhre etc.
<b>N Lymphknotenbefall</b>	
NX	regionale Lymphknoten nicht beurteilbar
N0	regionale Lymphknoten tumorfrei
N1	Befall 1–2 regionaler Lymphknoten
N2	Befall 3–6 regionaler Lymphknoten
N3	Befall $\geq 7$ regionaler Lymphknoten
<b>M Fernmetastasen</b>	
M0	keine Fernmetastasen
M1	Fernmetastasen

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

Tabelle 66: Stadieneinteilung von Plattenepithelkarzinomen der Speiseröhre nach dem AJCC [15]

Stadium	T	N	M	G	Tumorlokalisation
0	Hochgr. Dysplasie	N0	M0	1, X	Jede
IA	T1	N0	M0	1, X	Jede
IB	T1	N0	M0	2–3	Jede
	T2–3	N0	M0	1, X	Unterer Anteil, X
IIA	T2–3	N0	M0	1, X	Oberer, mittlerer Anteil
	T2–3	N0	M0	2–3	Unterer Anteil, X
IIB	T2–3	N0	M0	2–3	Oberer, mittlerer Anteil
	T1–2	N1	M0	Jedes	Jede
IIIA	T1–2	N2	M0	Jedes	Jede
	T3	N1	M0	Jedes	Jede
	T4a	N0	M0	Jedes	Jede
IIIB	T3	N2	M0	Jedes	Jede
IIIC	T4a	N1–2	M0	Jedes	Jede
	T4b	Jedes	M0	Jedes	Jede
	Jedes	N3	M0	Jedes	Jede
IV	Jedes	Jedes	M1	Jedes	Jede

Tabelle 67: Stadieneinteilung von Adenokarzinomen der Speiseröhre nach dem AJCC [15]

Stadium	T	N	M	G
0	Hochgr. Dysplasie	N0	M0	1, X
IA	T1	N0	M0	1–2, X
IB	T1	N0	M0	3
	T2	N0	M0	1–2, X
IIA	T2	N0	M0	3
IIB	T3	N0	M0	Jedes
	T1–2	N1	M0	Jedes
IIIA	T1–2	N2	M0	Jedes
	T3	N1	M0	Jedes
	T4a	N0	M0	Jedes
IIIB	T3	N2	M0	Jedes
IIIC	T4a	N1–2	M0	Jedes
	T4b	Jedes	M0	Jedes
	Jedes	N3	M0	Jedes
IV	Jedes	Jedes	M1	Jedes

## Anhang G – Darlegung potenzieller Interessenkonflikte der externen Sachverständigen und der externen Reviewer

Im Folgenden sind die potenziellen Interessenkonflikte der externen Sachverständigen und des externen Reviewers dargestellt. Alle Informationen beruhen auf Selbstangaben der einzelnen Personen anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte“ mit Stand 03/2009, 06/2011 oder 12/2011. Das aktuelle Formblatt ist unter [www.iqwig.de](http://www.iqwig.de) abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung

### Externe Sachverständige

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
Baumert, Brigitta <sup>4</sup>	nein	nein	ja	nein	nein	nein
Kleijnen, Jos <sup>4</sup>	nein	nein	ja	ja	nein	nein
Westwood, Marie <sup>4</sup>	nein	nein	nein	ja	nein	nein
Wolff, Robert <sup>4</sup>	nein	nein	nein	ja	nein	nein
Weiterer Sachverständiger, Nuklearmediziner (keine Zustimmung zur Veröffentlichung des Namens) <sup>4</sup>	nein	nein	nein	nein	nein	nein

### Externes Review

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
Lordick, Florian <sup>5</sup>	nein	ja	ja	ja	nein	nein

<sup>4</sup> Formblatt zur Offenlegung potenzieller Interessenkonflikte; Stand 03/2009

<sup>5</sup> Formblatt zur Offenlegung potenzieller Interessenkonflikte; Stand 06/2011

Im „Formblatt zur Offenlegung potenzieller Interessenkonflikte“ wurden folgende 6 Fragen gestellt (Version 06/2011):

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor angestellt bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband direkt oder indirekt beraten?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischem Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband Honorare erhalten für Vorträge, Stellungnahmen oder Artikel?

Frage 4: Haben Sie und / oder hat die Einrichtung<sup>6</sup>, die Sie vertreten, abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischem Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Frage 5: Haben Sie und/oder hat die Einrichtung<sup>6</sup>, bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren ohne wissenschaftliche Gegenleistung) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist?

---

<sup>6</sup> Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.

Im „Formblatt zur Offenlegung potenzieller Interessenkonflikte“ wurden folgende 6 Fragen gestellt (Version: 03/2009):

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor bei einem Interessenverband im Gesundheitswesen oder einem vergleichbaren Interessenvertreter<sup>7</sup> abhängig (angestellt) beschäftigt? Falls ja, wo und in welcher Position?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor einen Interessenverband im Gesundheitswesen oder einen vergleichbaren Interessenvertreter direkt oder indirekt beraten? Falls ja, wen und wie hoch ist / war die Zuwendung / das Honorar?

Frage 3: Haben Sie abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres oder der 3 Kalenderjahre davor im Auftrag eines Interessenverbands im Gesundheitswesen oder eines vergleichbaren Interessenvertreters Honorare für Vorträge, Stellungnahmen, Ausrichtung und / oder Teilnahme an Kongressen und Seminaren – auch im Rahmen von Fortbildungsveranstaltungen, oder für (populär-)wissenschaftliche oder sonstige Aussagen oder Artikel erhalten? Falls ja, von wem, für welche Tätigkeiten und wie hoch war die Zuwendung / das Honorar?

Frage 4: Haben Sie abseits einer Anstellung oder Beratungstätigkeit und / oder hat die Institution<sup>8</sup>, bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Interessenverband im Gesundheitswesen oder einem vergleichbaren Interessenvertreter finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten? Falls ja, von wem, für welche Tätigkeit und in welcher Höhe?

Frage 5: Haben Sie und / oder hat die Institution, bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres oder der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Reisekostenunterstützung ohne wissenschaftliche Gegenleistungen) von einem Interessenverband im Gesundheitswesen oder einem vergleichbaren Interessenvertreter erhalten? Falls ja, von wem, aus welchem Anlass und in welcher Höhe?

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile (auch in Fonds) von einer Firma oder Institution, die zu einem Interessenverband im Gesundheitswesen oder einem vergleichbaren Interessenvertreter gehört? Falls ja, von wem und welchen Wert haben diese aktuell?

---

<sup>7</sup> Dieses Formblatt erfasst finanzielle Beziehungen zu Interessenverbänden im Gesundheitswesen oder vergleichbaren Interessenvertretern, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie.

<sup>8</sup> Sofern Sie in einer ausgedehnten Institution tätig sind, ist es ausreichend, die geforderten Angaben auf Ihre Arbeitseinheit (z. B.: Klinikabteilung, Forschungsgruppe etc.) zu beziehen.