

**Positronenemissionstomographie  
(*PET und PET/CT*) bei malignen  
Lymphomen**

**Berichtsplan (vorläufige Version)**

Auftrag D06-01A  
Version 0.1  
Stand: 12.09.2007

# Impressum

**Herausgeber:**

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

**Thema:**

Positronenemissionstomographie (PET und PET/CT) bei malignen Lymphomen

**Auftraggeber:**

Gemeinsamer Bundesausschuss

**Datum des Auftrags:**

21.12.2006

**Interne Auftragsnummer:**

D06-01A

**Anschrift des Herausgebers:**

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Dillenburger Str. 27

51105 Köln

Tel: 0221/35685-0

Fax: 0221/35685-1

Berichte@iqwig.de

www.iqwig.de

Im folgenden Text wurde bei der Angabe von Personenbezeichnungen jeweils die männliche Form angewandt. Dies erfolgte ausschließlich zur Verbesserung der Lesbarkeit.

Zu allen Dokumenten, auf die via Internet zugegriffen wurde und die entsprechend zitiert sind, ist das jeweilige Zugriffsdatum angegeben. Sofern diese Dokumente zukünftig nicht mehr über die genannte Zugriffsadresse verfügbar sein sollten, können sie im Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen eingesehen werden.

Zu diesem Berichtsplan (vorläufige Version) können Stellungnahmen abgegeben werden, die gegebenenfalls zu einer Überarbeitung des Berichtsplans führen können. Die Frist für den Eingang der Stellungnahmen finden Sie auf der Internetseite des Instituts ([www.iqwig.de](http://www.iqwig.de)), ebenso wie die dafür notwendigen Formblätter und einen Leitfaden.

**Schlagwörter:** Positronenemissionstomographie, Computertomographie, maligne Lymphome, Staging, Rezidiv, Behandlungsresiduen, Systematische Übersicht

Der vorliegende Berichtsplan (vorläufige Version) soll wie folgt zitiert werden:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Positronenemissionstomographie (PET und PET/CT) bei malignen Lymphomen. Berichtsplan D06-01A (vorläufige Version 0.1). Köln: IQWiG; 2007.

# Inhaltsverzeichnis

	Seite
<b>Inhaltsverzeichnis</b> .....	<b>iii</b>
<b>Tabellenverzeichnis</b> .....	<b>v</b>
<b>Abkürzungsverzeichnis</b> .....	<b>vi</b>
<b>1 Hintergrund</b> .....	<b>1</b>
<b>1.1 Definition des Krankheitsbildes</b> .....	<b>1</b>
<b>1.2 Epidemiologie und Krankheitslast</b> .....	<b>1</b>
<b>1.3 Ursache der Erkrankung</b> .....	<b>2</b>
<b>1.4 Verlauf der Erkrankung (natürlicher Verlauf)</b> .....	<b>2</b>
<b>1.5 Klassifikationsschemata</b> .....	<b>3</b>
<b>1.6 Diagnostische Verfahren</b> .....	<b>5</b>
<b>1.7 Derzeit übliche Therapiepraxis</b> .....	<b>5</b>
<b>1.8 Positronenemissionstomographie</b> .....	<b>6</b>
<b>2 Ziele der Untersuchung</b> .....	<b>8</b>
<b>2.1 Ermittlung des patientenrelevanten Nutzens von PET bzw. PET/CT</b> .....	<b>8</b>
<b>2.2 Bewertung der diagnostischen und prognostischen Güte von PET bzw. PET/CT</b> .....	<b>8</b>
<b>3 Projektablauf</b> .....	<b>10</b>
<b>4 Methoden</b> .....	<b>11</b>
<b>4.1 Ermittlung des patientenrelevanten Nutzens von PET bzw. PET/CT</b> .....	<b>12</b>
4.1.1 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens von PET bzw. PET/CT anhand von systematischen Übersichten bzw. HTA-Berichten	13
4.1.2 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens von PET bzw. PET/CT anhand von Primärstudien.....	13
<b>4.2 Bewertung der diagnostischen und prognostischen Güte von PET bzw. PET/CT</b> .....	<b>16</b>
4.2.1 Einschlusskriterien zur Ermittlung der diagnostischen und prognostischen Güte anhand von systematischen Übersichten bzw. HTA-Berichten.....	16
4.2.2 Einschlusskriterien zur Ermittlung der diagnostischen und prognostischen Güte anhand von Primärstudien.....	18
<b>4.3 Informationsbeschaffung</b> .....	<b>20</b>
4.3.1 Literaturrecherche .....	20

4.3.2	Identifizierung relevanter Publikationen bzw. Studien .....	22
4.3.3	Suche nach zusätzlichen Informationen zu relevanten Publikationen .....	22
4.3.4	Informationen zu relevanten unpublizierten und publizierten Studien aus den Anhörungsverfahren.....	22
<b>4.4</b>	<b>Informationsbewertung .....</b>	<b>23</b>
4.4.1	Bewertung des Verzerrungspotenzials von Studien zur Bewertung des patientenrelevanten Nutzens .....	23
4.4.2	Bewertung der Qualität von Primärstudien zur Bewertung der diagnostischen und prognostischen Güte.....	24
4.4.3	Bewertung der Qualität von systematischen Übersichten und HTA-Berichten zur Bewertung der diagnostischen und prognostischen Güte .....	24
<b>4.5</b>	<b>Informationssynthese und -analyse .....</b>	<b>25</b>
4.5.1	Charakterisierung und Darstellung der Ergebnisse von einzelnen (Primär-)Studien zur Bewertung des patientenrelevanten Nutzens .....	25
4.5.2	Charakterisierung und Darstellung der Ergebnisse von systematischen Übersichten oder HTA-Berichten .....	25
4.5.3	Sensitivitätsanalyse .....	26
4.5.4	Subgruppenanalysen.....	26
<b>5</b>	<b>Literaturverzeichnis.....</b>	<b>27</b>

**Tabellenverzeichnis**

Tabelle 1: Inzidenz und Mortalität pro 100 000 Patienten-Jahre von Hodgkin- und Non-Hodgkin-Lymphomen in Deutschland 2002 nach Geschlecht [2] .....	1
Tabelle 2: Lymphoide Tumoren nach WHO-Subtypen [3] .....	4
Tabelle 3: Zubrod-ECOG-WHO-Skala zur Bewertung des Allgemeinzustandes .....	5
Tabelle 4: Klassifikationsschemata für Evaluierungsstudien zu diagnostischen Testverfahren .....	12

**Abkürzungsverzeichnis**

<b>Abkürzung<sup>a</sup></b>	<b>Bedeutung</b>
ABVD	Chemotherapieschema mit Adriamycin (syn. Doxorubicin (-hydrochlorid)), Bleomycin, Vinblastin und Dacarbazin
AIDS	Acquired immunodeficiency syndrome
BEACOPP	Chemotherapieschema mit Bleomycin, Etoposid, Adriamycin (syn. Doxorubicin(-hydrochlorid)), Cyclophosphamid, Vincristin (entspr. Oncovin), Procarbazin, Prednisolon (gesteigertes BEACOPP enthält noch zusätzlich G-CSF)
CLL	Chronisch lymphatische Leukämie
CONSORT	Consolidated Standards of Reporting Trials
CT	Computertomographie
DGHO	Deutsche Gesellschaft für Hämatologie und Onkologie
EBV	Epstein-Barr-Virus
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
FDA	Food and Drug Association
FLIPI	Follicular Lymphoma International Prognostic Index
FDG	Fluor-Deoxy-Glukose
G-BA	Gemeinsamer Bundesausschuss
G-CSF	Granulocyte colony stimulating factor
HAART-Therapie	Highly active antiretroviral therapy (= hoch aktive antiretrovirale Kombinationstherapie gegen HIV)
HHV-8	Human Herpesvirus 8
HIV	Humanes Immundefizienz-Virus
HL	Hodgkin-Lymphom
HTA	Health technology assessment
HTLV-1	T-Zell-Leukämievirus
ICD	International Classification of Diseases
IPI	International Prognostic Index

IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-treat
LDH	Lactatdehydrogenase
MALTom	Lymphom aus MALT (mucosa associated lymphoid tissue)
Max	maximal
NHL	Non-Hodgkin-Lymphom
NK-Zellen	Natürliche Killerzellen
NNT	Number needed to treat
NOS	Not otherwise specified
PET	Positronenemissionstomographie
PET/CT	Integration von PET (Positronenemissionstomographie) und CT (Computertomographie) in einem Untersuchungsgerät
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
QUOROM	Quality Of Reporting Of Meta-analyses
RCT	Randomised controlled trial (Randomisierte kontrollierte Studie)
REAL-Klassifikation	Revised European American Lymphoma Classification
STARD	Standards for Reporting of Diagnostic Accuracy
SLL	Kleinzelliges lymphozytisches Lymphom
SUV	Standardized Uptake Value
TREND	Transparent Reporting of Evaluations with Nonrandomized Designs
WHO	World Health Organization
Fußnoten:	
a: Diese Tabelle enthält keine Abkürzungen für elektronische Literaturdatenbanken	



## 1 Hintergrund

### 1.1 Definition des Krankheitsbildes

Unter der Bezeichnung „Lymphome“ wird ein breites Spektrum an Erkrankungen zusammengefasst, die aus entarteten Zellen des lymphatischen Systems entstehen. Bis vor Kurzem unterschied man zwischen Hodgkin-Lymphomen (HL) und Non-Hodgkin-Lymphomen (NHL) [1]. Heute werden die unterschiedlichen Entitäten unter ‚Malignen Lymphomen‘ zusammengefasst. Da die Unterteilung in Hodgkin- und Non-Hodgkin-Lymphome in der Literatur noch vielfach benutzt wird, kommt diese Nomenklatur auch in dem vorliegenden Berichtsplan teilweise zur Anwendung. Des weiteren unterteilt man Lymphome nach ihrem Wachstumsverhalten: indolente Lymphome wachsen langsam, während die aggressiven Lymphome rasch progredient sind [2]. Gemäß Auftragskonkretisierung (s. Kapitel 3) werden das multiple Myelom und alle leukämisch verlaufenden lymphatischen Neoplasien bei der Berichterstellung nicht berücksichtigt. Unter ‚leukämisch verlaufende lymphatische Neoplasien‘ fallen u. a. die chronisch lymphozytische Leukämie sowie das kleinzellige lymphozytische Lymphom und die Neoplasien mit Abstammung von den Vorläuferzellen (d. h. Vorläuferzell-B-lymphoblastische Leukämie bzw. –Lymphom und Vorläuferzell-T-lymphoblastische Leukämie bzw. –Lymphom).

### 1.2 Epidemiologie und Krankheitslast

Die Inzidenz und Mortalität von Hodgkin- und Non-Hodgkin-Lymphomen in Deutschland 2002 ist in Tabelle 1 aufgeführt. Während Hodgkin-Lymphome mit etwa 0,4 % aller Krebsneuerkrankungen eher seltenere Tumoren hierzulande darstellen, treten Non-Hodgkin-Lymphome ca. 7-mal häufiger auf. In den USA stehen NHL an der 5. (Frauen) bzw. 6. Stelle (Männer) aller Krebsdiagnosen [1].

Tabelle 1: Inzidenz und Mortalität pro 100 000 Patienten-Jahre von Hodgkin- und Non-Hodgkin-Lymphomen in Deutschland 2002 nach Geschlecht (nach Robert-Koch-Institut [3])

Erkrankung	Männer		Frauen	
	Inzidenz	Mortalität	Inzidenz	Mortalität
Hodgkin-Lymphom (ICD-10 C81)	2.2	0.4	2.0	0.4
Non-Hodgkin-Lymphom (ICD-10 C82-85)	14.6	6.9	14.8	6.5

Im Vergleich zu den anderen europäischen Ländern liegen in Deutschland die Inzidenzraten für HL bei Männern im unteren, bei Frauen im mittleren Bereich, während sie sich für NHL bei Männern im mittleren und bei Frauen im oberen Bereich befinden [3].

Von 1970 bis 1990 hatte man eine 50%ige Zunahme der NHL in den USA beobachtet, was zum Teil auf die veränderte Diagnostik und Klassifikation sowie die AIDS-Epidemie zurückgeführt wird [4,5]. Obwohl ein Teil der Zunahme auch durch Veränderungen in der Verteilung in der Alterspyramide bedingt ist, erklärt diese veränderte Altersstruktur die Zunahme nicht gänzlich [5]. Die Inzidenz des HL hat sich dagegen in den vergangenen 10 Jahren nicht wesentlich geändert [6].

### **1.3 Ursache der Erkrankung**

Bei den meisten Lymphomen ist wenig über die Ätiologie bekannt. Die Risikofaktoren für die Entwicklung von Lymphomen sind nur teilweise identifiziert und treffen für die verschiedenen Entitäten nicht gleichermaßen zu [3]. Lymphome treten gehäuft im Zusammenhang mit einer Suppression des Immunsystems auf, wie zum Beispiel nach Transplantation oder im Kontext von HIV-Infektionen. Bei transplantierten Patienten scheint die Infektion mit dem Epstein-Barr-Virus (EBV) eine zentrale Rolle zu spielen, wobei die genauen Zusammenhänge nicht geklärt sind. Die erhöhte Inzidenz von Lymphomen, die man in den 90er Jahren bei HIV-Patienten beobachtete, ist seit Einführung der HAART-Therapie wieder stark rückläufig. [7]. Auch Infekte mit anderen Viren und Erregern, z. B. mit Human Herpesvirus 8 (HHV 8), T-Zell-Leukämievirus (HTLV-1) und Helicobacter pylori, sind mit der Entwicklung von Lymphomen assoziiert [7]. Eine Lymphomerkkrankung in der Familienanamnese scheint das Risiko, an einem malignen Lymphom zu erkranken, zu erhöhen [1,3,6]. Ferner werden die Expositionen mit Pestiziden, Schwermetallen, UV-Licht oder Autoimmunkrankheiten als potenzielle Risikofaktoren diskutiert [1,5-7].

### **1.4 Verlauf der Erkrankung (natürlicher Verlauf)**

Die Beschwerden und Symptome bei malignen Lymphomen sind oft unspezifisch [7]. Bei Erstdiagnose präsentieren sich die Patienten mit HL zumeist mit vergrößerten, kaum schmerzhaften, oberflächlichen Lymphknoten in der Zervikal-, Supraklavikular- oder Axillarregion [5,7]. Maligne Lymphome können sich mit retroperitonealen, mesenterialen und inguinalen Lymphknotenvergrößerungen präsentieren. Follikuläre Lymphome sind zum Zeitpunkt der Diagnose oft schon disseminiert und haben bereits häufig das Knochenmark befallen.

Husten, Dyspnoe oder retrosternale Schmerzen können bei intrathorakalen Lymphomen auftreten, aber selbst große mediastinale Lymphommassen können asymptomatisch sein [5]. Das typische undulierende Pel-Epstein-Fieber wird gelegentlich in den Endstadien gesehen, und klassischer Nachtschweiß weist auf ein fortgeschrittenes Stadium hin. Pruritus wird häufiger beim HL beobachtet. Alkoholschmerz mit ziehenden Schmerzen in befallenen Regionen nach Alkoholkonsum tritt in weniger als 5 % der Fälle auf und ist nicht auf das HL beschränkt [7]. Häufig befallene extranodale Regionen sind Gastrointestinaltrakt (einschließlich Magen), ZNS, Hoden, Skelett und viszerale Organe wie Lunge und Leber [7]. Patienten

mit Burkitt-Lymphom haben oft Bauchschmerzen oder ein abdominales Völlegefühl, da der Tumor häufig den Bauchraum befällt.

Bei den Lymphomen finden sich erhebliche prognostische Variationen in Abhängigkeit von der Histologie. Während aggressive Tumorarten unbehandelt innerhalb kurzer Zeit (Wochen bis Monate) zum Tod führen, aber oft gut auf die Therapie ansprechen und gute Heilungschancen haben, sind weniger aggressive Tumoren mit einer deutlich besseren Überlebenszeit (im Rahmen von Jahren) assoziiert, sprechen aber vergleichsweise schlecht auf Behandlungen an und sind im Allgemeinen nicht heilbar

## 1.5 Klassifikationsschemata

Im Jahr 2001 wurde basierend auf der REAL (Revised European American Lymphoma)-Klassifikation die neue WHO-Klassifikation der Lymphome entwickelt (Tabelle 2) [3]. Diese gilt derzeit als Goldstandard [1,8]. Die malignen Lymphome werden gemäß WHO-Klassifikation in 5 Gruppen unterteilt: B-Zell, T-Zell und NK-Zell Tumoren (NK=natürliche Killer Zellen), histiozytische und dendritische Tumoren sowie Mastocytose. Morphologie, Immunphänotypisierung, zytogenetische Veränderungen sowie die Klinik werden benutzt, um die verschiedenen Lymphom-Entitäten voneinander zu unterscheiden.

Die Ann-Arbor-Klassifikation dient der Stadieneinteilung maligner Lymphome und unterteilt sie aufgrund der Lymphknotenregionen bzw. der extralymphatischen Organe, die befallen sind, und ihrer Lage in Bezug auf das Zwerchfell. Das Vorliegen von B-Symptomen (Nachtschweiß, Fieber  $> 38\text{ °C}$ , Gewichtsverlust  $> 10\%$  in 6 Monaten) wird für jedes Stadium zusätzlich angegeben. [5].

Die Ann-Arbor-Klassifikation ist Bestandteil des International Prognostic Index (IPI) zur Prognosebeurteilung von malignen Lymphomen. Der IPI war primär für das diffuse großzellige B-Zell-Lymphom entwickelt worden und kennt 5 klinische Risikofaktoren: Alter  $\geq 60$  Jahre, LDH (Lactatdehydrogenase) über Normwert, ECOG performance status (Eastern Cooperative Oncology Group; Skala zur Bewertung des Allgemeinzustandes, Tabelle 3)  $\geq 2$  oder Karnofsky Index  $\leq 70$ , Ann-Arbor Stadium III oder IV, Zahl der betroffenen extranodalen Krankheitsherde  $> 1$  [6]. Auch im Rahmen eines Rezidivs erlaubt das Staging mit dem IPI eine Einschätzung der Prognose. Inzwischen wurde der IPI auch für andere maligne Lymphome, z. B. als FLIPI für folliculäre Lymphome, adaptiert.

Tabelle 2: Lymphoide Tumoren nach WHO-Subtypen [nach 4]

<b>B-Zell-Lymphome</b>		<b>T-Zell-Lymphome</b>	
<b>WHO-Klassifikation</b>	<b>Rate / 100.000 Personen- jahre</b>	<b>WHO-Klassifikation</b>	<b>Rate / 100.000 Personen- jahre</b>
Alle lymphoiden B-Zell Neoplasmen*	26.13	Alle lymphoiden T-/NK-Zellen Neoplasmen*	1.79
<i>Vorläuferzell-Lymphome</i>		<i>Vorläuferzell-Lymphome</i>	
Vorläuferzell-B- lymphoblastische Leukämie/Lymphom	0.22	Vorläuferzell-T- lymphoblastische Leukämie/Lymphom	0.63
<i>Reifzellige Neoplasien</i>		<i>Reifzellige Neoplasien</i>	
Diffuses großzelliges B-Zell- Lymphom*	7.14	Mycosis fungoides/Sézary Syndrom*	0.52
Marginalzonen-B-Zell- Lymphom*	0.97	Peripheres T-Zell-Lymphom*	0.52
Follikuläres Lymphom*	3.18	Angioimmunoblastisches Lymphom	0.05
CLL/SLL (chronisch lymphozytische Leukämie, kleinzelliges lymphozytisches Lymphom)*	5.17	Anaplastisches großzelliges Lymphom	0.25
Mantelzell-Lymphom	0.51	Peripheres T-Zell-Lymphom, NOS**	0.30
Burkitt-Lymphom/Leukämie*	0.30	T-/NK-Zell-Lymphome, NOS*,**	0.44
Plasmazellneoplasmen*	5.66		
Haarzell-Leukämie	0.33		
Lymphoplasmozytisches Lymphom	0.27		
Waldenström Macroglobulinämie	0.35		
B-Zell lymphoide Neoplasmen, NOS*,**	1.45		
Hodgkin Lymphome*			2.67
Lymphoide Neoplasmen unbekanntem Typs*			2.4
Alle lymphoiden Neoplasmen***			33.65

\* Diese Gruppe umfasst mehrere ICD-0-3 Codes,

\*\* NOS: not otherwise specified

\*\*\* Diese Gruppe umfasst alle in der SEER-Studie erfassten lymphoiden Neoplasmen

Tabelle 3: Zubrod-ECOG-WHO-Skala zur Bewertung des Allgemeinzustandes

Zubrod-ECOG-WHO-Skala	
0	Volle Aktivität, Ausübung der normalen Tätigkeit möglich
1	Eingeschränkte Aktivität, Ausübung von leichter Tätigkeit möglich
2	Arbeitsunfähigkeit, Selbstversorgung, keine dauerhafte Bettlägerigkeit
3	Starke Einschränkung der Selbstversorgung, zusätzliche Pflegebedürftigkeit, überwiegende Bettlägerigkeit
4	Bettlägerigkeit, vollständige Pflegebedürftigkeit

## 1.6 Diagnostische Verfahren

In den Leitlinien der Deutschen Gesellschaft für Hämatologie und Onkologie (DGHO) vom Februar 2007 werden folgende Untersuchungen bei der Erstuntersuchung vorgeschlagen [2]: Anamnese insbesondere von B-Symptomen, körperliche Untersuchung und Blutuntersuchungen, eine Computertomographie von Hals, Thorax und Abdomen, welche zur Verlaufskontrolle mit einer Sonographie ergänzt werden sollte. Bei hochzervikalem Lymphknotenbefall werden eine Liquorpunktion und eine HNO-ärztliche Untersuchung empfohlen, welche bei Mycosis fungoides und CLL entbehrlich sind. Je nach klinischer Symptomatik sollten Endoskopien, Röntgenuntersuchungen und/oder Skelettszintigraphien durchgeführt werden. Eine Abklärung mit einer Positronenemissionstomographie (PET) empfiehlt die DGHO nur bei fraglichen Befunden in der bildgebenden Diagnostik und bei klinischen Konsequenzen, wobei gegebenenfalls eine histologische Sicherung obligat sei [2]. Andere beurteilen die Rolle des PET je nach Tumorphistologie differenziert. So sei die PET z. B. bei großzelligen B-Lymphomen etabliert [9]. Lymphome mit niedrigem Malignitätsgrad würden jedoch aufgrund des geringeren Stoffwechsels mit Fluor-Deoxy-Glukose (FDG) nicht immer mit der PET erkannt [9,10]. Eine PET könne hingegen eine Unterscheidung zwischen Narbengewebe und Resttumor nach Therapie erlauben [10]. Unter Therapie und direkt nach Abschluss der Therapie werden von der DGHO zur Verlaufs- und Nebenwirkungskontrolle Laboruntersuchungen und ggf. weitere Spezialuntersuchungen empfohlen [2].

Zur Nachsorge wird von der DGHO eine Wiederholung der Bildgebung nach Abschluss der Therapie in 3- bis 6-monatigen Abständen und ab dem 4. Jahr in zwölfmonatigen Abständen empfohlen. Auch hier soll eine Abklärung mit PET oder PET/CT nur dann erfolgen, wenn dies klinische Konsequenzen hat oder im Rahmen einer klinischen Studie gemacht wird [2].

## 1.7 Derzeit übliche Therapiepraxis

Für die Behandlung der Lymphome steht ein weites Spektrum an Therapien zur Verfügung, das von der einfachen Eradikationstherapie für *Helicobacter pylori* bei sehr frühen Stadien des

MALToms bis zu komplexen Therapieschemata wie dem gesteigerten BEACOPP-Schema (Cyclophosphamid, Doxorubicin-Hydrochlorid, Etoposid, Procarbazin, Bleomycin, Vincristin, Prednisolon, G-CSF), der Hochdosistherapie mit autologer Stammzellreinfusion und Immuntherapien mit monoklonalen Antikörpern (B-Zell Antikörpern) reicht. Maligne Lymphome sind sehr strahlensensibel, sodass die Bestrahlung ein Bestandteil vieler Therapieschemata ist [2,11].

Je nach Tumorart und Stadium bzw. Klinik wird die Therapie angepasst, bzw. es werden eventuell verschiedene Therapien miteinander kombiniert, wobei sich die Therapieziele im Verlauf ändern können [5]. So präsentiert sich das Mantelzell-Lymphom von der Histologie her als ein relativ benigner Tumor, es ist jedoch nur in den seltenen gefundenen, frühen Stadien heilbar, und eine befriedigende Standardtherapie existiert nicht [5,9]. Beim Hodgkin-Lymphom können dagegen mit einer stadienadaptierten Therapie gute Erfolge erzielt werden. Je nach Stadium wird mit ABVD (Doxorubicin-Hydrochlorid, Bleomycin, Vinblastin, Dacarbazin i.v.) bzw. BEACOPP oder eskaliertem BEACOPP, jeweils bei Bedarf in Kombination mit Radiotherapie, eine Heilung auch in fortgeschrittenen Fällen bei über 80 % aller Patienten erreicht [9,12].

## 1.8 Positronenemissionstomographie

Die PET ist ein nichtinvasives diagnostisches Verfahren, das geeignet ist, mittels kurzlebiger Radioisotope metabolische Prozesse abzubilden. Im Gegensatz zu den klassischen bildgebenden Verfahren, bei denen Informationen über die anatomische Struktur und die Lokalisation von Läsionen erhoben werden, kann die PET komplementär dazu Funktion, Metabolismus und biochemische Prozesse der Organe charakterisieren. Wegen ihres verstärkten Glukosemetabolismus wird bei Lymphomen in der Regel, aber nicht ausschließlich, das Glukoseanalogon FDG als Tracer eingesetzt. Die von dem Zerfall des radioaktiven Tracers freigesetzten Positronen werden von einem PET-Scanner entdeckt und in ein Bild umgewandelt. Die Befundung ergibt sich aus dem Vergleich zwischen der Aktivität eines bestimmten Areals verglichen mit der Aktivität aus dem umliegenden „normalen“ Gewebe. Aufgrund der Unterschiede an physiologischer Glukoseaufnahme in den verschiedenen Körperregionen kann die Zuordnung einer bestimmten Region als „erhöhter Uptake“ schwierig sein und erfordert erhebliche Expertise. Um bei der Interpretation eine gewisse Standardisierung zu erreichen, wird der Glukosemetabolismus semiquantitativ als *Standardized Uptake Value* (SUV) wiedergegeben. Damit kann die morphologisch orientierte CT- und MRT-Bildgebung bei spezifischen Fragestellungen um weitere Informationen ergänzt werden. Allerdings ist ein positiver Befund nicht malignomspezifisch, vielmehr können sich auch andere Prozesse mit erhöhtem Glukoseumsatz, wie z. B. Infektionen, dahinter verbergen. PET-Vollringscanner werden im Allgemeinen als Technologie der Wahl angesehen, sodass sich offenbar in der Vergangenheit insbesondere deutschsprachige HTAs bei ihren Bewertungen auf diese Technologie beschränkt haben [13,14].

Eine technische Weiterentwicklung stellt die PET/CT dar, eine Integration von PET und CT in einem Untersuchungsgerät. Hier wird der Patient in einem Untersuchungsgang durch die beiden Detektorringe von CT und PET gefahren. Die entstehenden Bilder werden später im Computer fusioniert, wobei die CT-Information meist in Graustufen abgebildet und von der PET-Information dann farbig überlagert wird. Mit dieser Methode wird die hohe Auflösung der CT (ca. < 1 mm) mit den Stoffwechsel-Informationen der PET verknüpft.

## 2 Ziele der Untersuchung

Die vorliegende Untersuchung verfolgt zwei Ziele:

### 2.1 Ermittlung des patientenrelevanten Nutzens von PET bzw. PET/CT

Zum einen geht es um die Beschreibung des patientenrelevanten Nutzens, den Ärzte und Patienten von der Bildgebung mit PET bzw. PET/CT bei malignen Lymphomen erwarten können. Als Indikationen für den Einsatz werden a) die Bestimmung des Tumorstadiums („staging“), b) das Ansprechen des Lymphoms auf die Behandlung („residual disease evaluation“/„re-staging“) und c) der Nachweis eines Rezidivs bei begründetem Verdacht betrachtet. Unter Nutzen seien hierbei Veränderungen verstanden, die für den Patienten fassbare Konsequenzen haben, wie z. B. die Auswirkung auf die Mortalität und Morbidität, die optimierte Auswahl von vorhandenen Behandlungsoptionen mit mehr oder weniger toxischen Nebenwirkungen, das allgemeine klinische Management der Patienten und Veränderungen in der Lebensqualität.

Folgende Vergleiche sind dabei denkbar und sollen – soweit möglich – betrachtet werden:

- PET bzw. PET/CT versus konventionelle Diagnostik (wie z. B. die Kombination aus mehreren Verfahren wie Computertomographie, Ultraschall, MRI, Knochenszintigraphie, Labordiagnostik, Biopsie; vgl. Abschnitt 4.1.2.2)
- PET bzw. PET/CT versus einzelne Komponenten der konventionellen Diagnostik
- PET bzw. PET/CT + konventionelle Diagnostik versus konventionelle Diagnostik
- PET vs. PET/CT

### 2.2 Bewertung der diagnostischen und prognostischen Güte von PET bzw. PET/CT

Läge ausreichend Evidenz zu Ziel 1 vor, wäre eine Nutzenbewertung im Prinzip alleine anhand der Beantwortung dieser Fragestellung möglich. Vorrecherchen haben jedoch gezeigt, dass zu diesem Ziel kaum Primärstudien vorliegen. Sollte sich dieser Mangel an aussagekräftigen Primärstudien nach der systematischen Literaturrecherche bestätigen, werden als zweites Ziel auch die Ergebnisse von Diagnosestudien zur PET bzw. PET/CT zusammengefasst.

Bei der Bewertung der diagnostischen und prognostischen Güte von PET bzw. PET/CT geht es primär um die Frage, inwieweit PET bzw. PET/CT den diagnostischen Standardverfahren ohne PET überlegen ist. Das bedeutet: Verbessert sich durch den Einsatz von PET bzw. PET/CT die korrekte Zuordnung in die jeweilige Staginggruppe mit den unterschiedlichen prognostischen Konsequenzen, das korrekte Erkennen von Patienten mit bzw. ohne Residuen nach Behandlungsabschluss oder die korrekte Diagnose bzw. der korrekte Ausschluss von Rezidiven? Ähnlich gelagert ist die Frage, ob sich mittels PET und PET/CT zuverlässigere



prognostische Aussagen über die Wahrscheinlichkeit für ein Rezidiv machen lassen, als es mit den bisherigen diagnostischen Standardverfahren möglich war.

Dieses Ziel setzt den Vergleich mit einem sogenannten Goldstandard voraus, der für eine adäquate Interpretation der Ergebnisse unabhängig von einem zu prüfenden Diagnoseverfahren erhoben werden muss, da es anderenfalls zu Verzerrungen der Güte- und/oder Variabilitätsschätzer kommen kann (sogenannter Verifikations-Bias oder Inkorporations-Bias) [15-19].

Vorrecherchen lassen erkennen, dass in der Vergangenheit in Studien, die augenscheinlich das Ziel hatten, die diagnostische Güte von PET bzw. PET/CT zu ermitteln, allein die Übereinstimmung mit der konventionellen Diagnostik untersucht wurde. Da in solchen Studien die konventionelle Diagnostik als „Goldstandard“ interpretiert werden muss, lassen die resultierenden Ergebnisse nur eine sehr eingeschränkte Interpretation zu, da dann PET bzw. PET/CT definitionsgemäß bestenfalls als „gleichwertig“ zur konventionellen Diagnostik betrachtet werden kann.

Für die im Prinzip relevantere Frage, ob die diagnostische bzw. prognostische Güte von PET bzw. PET/CT der konventionellen Diagnostik überlegen ist (d. h., die konventionelle Diagnostik fungiert als sogenannter Komparator, analog den unter 2.1 aufgeführten Fragestellungen), müssen Studien so konzipiert sein, dass neben den beiden zu vergleichenden Diagnosemodalitäten zusätzlich noch ein davon unabhängiger Goldstandard erhoben wird. Es muss betont werden, dass dafür aus den vorgenannten Gründen die Biopsie nicht Bestandteil der konventionellen Diagnostik (als „Komparator“) sein kann, da sie gemeinhin als der „Goldstandard“ (Referenzstandard) für die endgültige Diagnosesicherung angesehen wird [18].

### 3 Projektablauf

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 bzw. insgesamt 21 Indikationen beauftragt. Der vorliegende Berichtsplan (vorläufige Version) bezieht sich auf den Teilauftrag zu den 3 Indikationen:

- Bestimmung des Tumorstadiums von malignen Lymphomen
- Beurteilung des Behandlungsansprechens bei malignen Lymphomen
- Nachweis von Rezidiven (bei begründetem Verdacht) bei malignen Lymphomen

Das abschließende Gespräch zur Auftragskonkretisierung hierzu erfolgte mit der zuständigen Themengruppe im G-BA am 28.03.2007. Hierbei wurde u. a. festgelegt, dass für den zu erstellenden Bericht alle leukämisch verlaufenden lymphatischen Neoplasien sowie das multiple Myelom ausgeschlossen werden sollen.

Zu dem vorliegenden Berichtsplan (vorläufige Version) können Stellungnahmen eingereicht werden. Das Ende der Stellungnahmefrist wird auf den Internetseiten des Instituts unter [www.iqwig.de](http://www.iqwig.de) bekannt gegeben. Stellungnahmen können von allen interessierten Personen, Institutionen und Gesellschaften, einschließlich Privatpersonen, Fachgesellschaften und Industrieunternehmen, abgegeben werden. Die Stellungnahmen müssen bestimmten formalen Anforderungen genügen, die ebenfalls auf den Internetseiten des Instituts in einem entsprechenden Leitfaden dargelegt sind. Die Stellungnahmen können gegebenenfalls zu einer Überarbeitung des Berichtsplans in seiner vorläufigen Version führen.

Die vorläufige Bewertung des IQWiG wird in einem Vorbericht veröffentlicht. Der Vorbericht wird zusätzlich einem externen Review unterzogen. Im Anschluss an die Veröffentlichung des Vorberichts erfolgt eine Anhörung zum Vorbericht mittels schriftlicher Stellungnahmen, die sich u. a. auch auf die Vollständigkeit der Informationsbeschaffung beziehen können. Relevante Information aus der Anhörung kann in die Nutzenbewertung einfließen. Gegebenenfalls wird eine wissenschaftliche Erörterung zur Klärung unklarer Aspekte aus den schriftlichen Stellungnahmen durchgeführt.

Nach der wissenschaftlichen Erörterung wird das IQWiG einen Abschlussbericht erstellen. Dieser Bericht wird an den G-BA übermittelt und 8 Wochen später im Internet veröffentlicht.

## 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. seine Fähigkeit, kranke von gesunden Personen zu unterscheiden, [20] in der impliziten Erwartung, dass sich eine bessere diagnostische Güte automatisch in eine bessere Patientenversorgung umsetzt. 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 [21,22]. Fryback und Thornbury haben dazu in ihrer Meilensteinpublikation „The efficacy of diagnostic imaging“ von 1991 [22] sechs Stufen identifiziert, die von den technischen Auswirkungen („technical efficacy“) als unterste Stufe über die diagnostische Genauigkeit („diagnostic accuracy efficacy“) als vorletzte Stufe hin zu den therapeutischen Auswirkungen („therapeutic efficacy“) als vierte und patientenrelevanten Auswirkungen („Patient outcome efficacy“) als fünfte Stufe reichen. Als oberste Stufe betrachteten sie schließlich die gesellschaftlichen Konsequenzen („societal efficacy“) der diagnostischen Technologie (s. auch Tabelle 4). 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 ziehen Überlegungen über die Konsequenzen und den Nutzen diagnostischer Tests für Arzt, Patient und Gesellschaft in Betracht, wenn sie über den Wert eines diagnostischen Verfahrens nachdenken. Auch der Gemeinsame Bundesausschuss hat sich in seiner Verfahrensordnung von 2006 bei der Bewertung diagnostischer Tests (§ 18 Abs. 2) dieser Sichtweise angeschlossen [23]. In Tabelle 4 sind verschiedene Klassifikationsschemata für Evaluierungsstudien zu diagnostischen Testverfahren einander gegenübergestellt. Neben der Einteilung von Fryback und Thornbury und der Evidenzklassifizierung des G-BA wurde auch das Phasenschema nach Köbberling et al. [21] mit aufgenommen.

Tabelle 4: Klassifikationsschemata für Evaluierungsstudien zu diagnostischen Testverfahren

	Köbberling et al. [21]	Fryback & Thornbury [22]	Evidenzklassifizierung G-BA [23]
Patientenrelevante Zielgrößen	Phase 4: Wirksamkeit	6. Stufe: Auswirkung auf Systemebene	Evidenzstufe I
		5. Stufe: Auswirkung im Hinblick auf das Outcome der Patienten	
		4. Stufe: Auswirkung im Hinblick auf das therapeutische Denken des Behandelnden	
Likelihood-Ratio (LR); Nachtestwahrscheinlichkeit		3. Stufe: Auswirkung im Hinblick auf das (differenzial-)diagnostische Denken ( <i>„Nachtestwahrscheinlichkeit“</i> )	
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 II
Sensitivität (SN), Spezifität (SP), Likelihood-Ratio (LR)	Phase 2: Diagnostische Genauigkeit bei bekanntem Krankheitsstatus		Evidenzstufe III
Analytische Sensitivität, Spezifität, Reproduzierbarkeit	Phase 1: Technische Voruntersuchungen	1. Stufe: Technische Auswirkungen	

#### 4.1 Ermittlung des patientenrelevanten Nutzens von PET bzw. PET/CT

Um den patientenrelevanten Nutzen der bildgebenden diagnostischen Technologien PET bzw. PET/CT für die Behandlung von Patienten mit Lymphomen zu bewerten, genügt es nicht, festzustellen, inwieweit die PET bzw. die PET/CT den gängigen diagnostischen Methoden überlegen sind, um ein Lymphom (bzw. seine Ausbreitung im Körper) korrekt zu identifizieren. Vielmehr geht es bei einer erweiterten Bewertung darum, festzustellen, inwieweit sich eine präzisere Diagnosestellung auch 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, Rezidivraten, Krankheitsprogression oder Lebensqualität, beeinflusst.

Diese Fragen lassen sich zuverlässig beantworten, wenn die verschiedenen diagnostischen Vorgehensweisen im Kontext von Managementstrategien betrachtet werden, bei denen der Patient neben der konventionellen Diagnostik zusätzlich die PET bzw. die PET/CT (oder ggf. auch alternativ zur konventionellen Diagnostik) erhält oder nicht und überprüft wird, welche

Änderungen sich dadurch für den Patienten ergeben: Ändern sich die Behandlungsoptionen (aggressiveres oder weniger aggressives Vorgehen, Chemo- oder Radiotherapie oder beides), und welche Auswirkungen haben die unterschiedlichen Strategien auf patientenrelevante Endpunkte wie Lebensqualität, Morbidität und Mortalität? Solche Fragen lassen sich am ehesten durch vergleichende Längsschnittstudien beantworten (s. a. Abschnitt 4.1.2.4).

#### **4.1.1 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens von PET bzw. PET/CT anhand von systematischen Übersichten bzw. HTA-Berichten**

Vorrecherchen haben ergeben, dass zur Bewertung des patientenrelevanten Nutzens von PET bzw. PET/CT nur eine geringe Anzahl von Primärstudien und kaum qualitativ ausreichende systematische Übersichtsarbeiten bzw. HTA-Berichte vorliegen. Daher soll der patientenrelevante Nutzen in dem zu erstellenden Bericht im Gegensatz zur Bewertung der diagnostischen und prognostischen Güte (s. Abschnitt 4.2) lediglich auf Basis der Primärliteratur ermittelt werden.

#### **4.1.2 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens von PET bzw. PET/CT anhand von Primärstudien**

##### **4.1.2.1 Population**

Folgende Patientenpopulationen werden eingeschlossen:

- Patienten mit nachgewiesenem malignem Lymphom, die ein Staging für die Primärtherapie erhalten
- Patienten nach Behandlung zur Erfassung des Behandlungsansprechens
- Patienten mit begründetem Verdacht auf ein Rezidiv

Unter Lymphomen im Sinne der Fragestellung werden alle nicht leukämisch verlaufenden lymphatischen Neoplasien verstanden. Ebenfalls ausgeschlossen werden Patienten mit multiplem Myelom.

Eingeschlossen werden Studien, bei denen der Anteil der Patienten mit malignen Lymphomen (wie zuvor definiert) mindestens 80 % beträgt oder bei denen Subgruppenanalysen für Patienten mit malignen Lymphomen durchgeführt wurden.

##### **4.1.2.2 Prüf- und Vergleichsintervention**

Prüfintervention ist eine diagnostische Strategie mit Anwendung einer (Vollring-)PET bzw. PET/CT. Als Vergleichsintervention wird eine diagnostische Strategie ohne Anwendung von (Vollring-)PET bzw. PET/CT betrachtet. Grundsätzlich vorstellbar sind die in Abschnitt 2.1 beschriebenen Vergleiche, was als Prüfintervention eine diagnostische Strategie mit

Anwendung von PET/CT und als dazugehörige Vergleichsintervention eine diagnostische Strategie mit Anwendung von (Vollring-)PET ermöglicht.

Ggf. bestehende Unterschiede in den verwendeten diagnostischen Maßnahmen – außer der Anwendung von PET bzw. PET/CT – zwischen den zu vergleichenden Strategien bestimmen die mögliche Interpretation der Ergebnisse. Eine Restriktion auf bestimmte (erlaubte) Begleitdiagnostik bzw. erlaubte Verteilung von Begleitdiagnostik zwischen Prüf- und Vergleichsintervention erscheint zum Zeitpunkt der Berichtsplanung nicht sinnvoll.

#### **4.1.2.3 Zielgrößen**

Als Zielgrößen werden die nachfolgend aufgeführten Kriterien verwendet, die eine Beurteilung patientenrelevanter Interventionsziele ermöglichen:

- Mortalität, krankheitsfreies Überleben
- Morbidität, z. B. Auftreten von Rezidiven, Nebenwirkungen durch aggressivere/weniger aggressive Behandlung
- Gesundheitsbezogene Lebensqualität
- Änderungen im klinischen Management der Patienten (wie z. B. Wahl der Therapie [intensiver oder weniger intensiv] aufgrund geänderter Prognoseschätzung, Größe des Bestrahlungsfeldes, Durchführung oder Unterlassung von weiteren, insbesondere invasiven diagnostischen Verfahren (z. B. Laparoskopie, Knochenmarksbiopsie))

#### **4.1.2.4 Studientypen**

Randomisierte klinische Studien (RCTs) liefern für die Bewertung des Nutzens einer medizinischen Intervention die zuverlässigsten Ergebnisse, weil sie, sofern methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt, mit der geringsten Ergebnisunsicherheit behaftet sind. Es werden deshalb primär RCTs betrachtet. Finden sich keine bzw. nur eine geringe Zahl von relevanten RCTs, aus denen der patientenrelevante Nutzen von PET bzw. PET/CT nicht mit ausreichender Ergebnissicherheit geschätzt werden kann (z. B. kleine Fallzahl, wenige Ereignisse [geringe Präzision], hohes Verzerrungspotenzial), werden gegebenenfalls auch verfügbare eindeutig prospektiv geplante nichtrandomisierte kontrollierte (Interventions-)Studien herangezogen, sofern das Problem einer möglichen Strukturungleichheit (unfairer Vergleich) adäquat in der Planung und Auswertung der entsprechenden Studien berücksichtigt wurde.

**4.1.2.5 Sonstige Charakteristika**

Es erfolgt keine weitere Einschränkung.

**4.1.2.6 Ein-/Ausschlusskriterien**

Es werden alle Studien herangezogen, die

- alle nachfolgenden Einschlusskriterien und
- keines der nachfolgenden Ausschlusskriterien erfüllen.

<b>Einschlusskriterien</b> zur Bewertung des patientenrelevanten Nutzens	
E1a	Patienten mit malignem Lymphom (Hodgkin/Non-Hodgkin), wie in Abschnitt 4.1.2.1 beschrieben
E2a	<ul style="list-style-type: none"> <li>▪ Prüfintervention: diagnostische Strategie mit Anwendung von (Vollring-) PET bzw. PET/CT; Vergleichsintervention: diagnostische Strategie ohne Anwendung von (Vollring-)PET bzw. PET/CT</li> <li>oder</li> <li>▪ Prüfintervention: PET/CT; Vergleichsintervention: (Vollring-)PET</li> </ul>
E3a	Patientenrelevante Endpunkte wie unter 4.1.2.3 beschrieben
E4a	Randomisierte kontrollierte Studien (und ggf. prospektive nichtrandomisierte kontrollierte Studien, wie in Abschnitt 4.1.2.4 beschrieben)
E5a	Sprachen: Englisch, Deutsch, Französisch, Spanisch

<b>Ausschlusskriterien</b> zur Bewertung des patientenrelevanten Nutzens	
A1a	Mehrfachpublikationen ohne relevante Zusatzinformation
A2a	Keine Vollpublikation verfügbar <sup>a</sup>
<p>a: Als Vollpublikation gilt 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- [24] oder TREND-Statements [25] genügt und eine Bewertung der Studie ermöglicht.</p>	

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

In den vergangenen Jahren haben sich viele westliche Gesundheitssysteme mit der Bedeutung der PET bzw. PET/CT für die Patientenversorgung auseinandergesetzt, und es sind zahlreiche systematische Bewertungen dazu publiziert worden. Orientierende Vorrecherchen haben z. B. mehr als 14 systematische Übersichten und HTA-Berichte mit einer Gesamtzahl von etwa 100 Primärstudien gefunden, die die Frage der diagnostischen Genauigkeit und prognostischen Güte von PET bzw. PET/CT näher untersucht haben. Um Redundanzen und Doppelarbeit zu vermeiden, soll die Frage der diagnostischen Genauigkeit von PET und PET/CT primär auf der Grundlage dieser Publikationen erfolgen, sofern sichergestellt ist, dass die Berichte bestimmte Mindestanforderungen erfüllen, um als tatsächlich systematische Übersichten gewertet werden zu können. Es soll sich also um Arbeiten bzw. deren Publikationen handeln, die auf systematische, reproduzierbare und transparente Weise Ergebnisse von Primärstudien zusammenfassend dargestellt und bewertet haben. Dazu müssen entsprechende Publikationen auf dem Qualitätsindex von Oxman/Guyatt mindestens 5 von 7 möglichen Punkten erreichen [26,27]. Weitere Voraussetzung ist, dass mindestens 2 durch voneinander unabhängige Arbeitsgruppen erstellte systematische Übersichten bzw. HTA-Berichte mit einem „Oxman/Guyatt-Score“ von mindestens 5 Punkten existieren, die die jeweilige Fragestellung bearbeitet haben.

Sollten nach Abschluss der Recherchen der eingeschlossenen systematischen Übersichten bzw. HTA-Berichte noch weitere relevante Primärstudien zu den Fragestellungen publiziert worden sein, so werden diese separat aufgeführt. Liegt der Verdacht nahe, dass die neuen Primärstudien das Ergebnis der Evidenzsynthesen möglicherweise maßgeblich verändern könnten, werden sie ausführlich analysiert und systematisch in die Gesamtbewertung einbezogen.

Fragestellungen, für die keine 2 systematischen Übersichten bzw. HTA-Berichte entsprechend obiger Definition identifiziert werden können, sollen allein auf der Basis von Primärstudien bearbeitet werden.

Nachfolgend werden zunächst die Kriterien für einen Einschluss von systematischen Übersichten bzw. HTA-Berichten in die geplante Bewertung aufgeführt. Anschließend erfolgt die Darstellung für ggf. (s. o.) einzuschließende Primärstudien.

### **4.2.1 Einschlusskriterien zur Ermittlung der diagnostischen und prognostischen Güte anhand von systematischen Übersichten bzw. HTA-Berichten**

#### **4.2.1.1 Population**

Ergebnisse für eine Population, wie unter 4.1.2.1 definiert, müssen extrahierbar sein.



#### **4.2.1.2 Prüftechnologie, ggf. Vergleichstechnologie („Komparator“) und „Goldstandard“**

Prüftechnologie ist die (Vollring-)PET bzw. PET/CT, ggf. dient als Vergleichstechnologie („Komparator“) die konventionelle Diagnostik. Prüf- und ggf. Vergleichstechnologie müssen zu einem sogenannten Goldstandard in Beziehung gebracht worden sein (s. Abschnitt 2.2). In den einzuschließenden systematischen Übersichten bzw. HTA-Berichten muss ein solches Ordnungsprinzip erkennbar und die Ergebnisse aus den Einzelstudien entsprechend dargestellt sein. Es erfolgt keine Eingrenzung auf einen bestimmten Goldstandard.

#### **4.2.1.3 Zielgrößen**

- Beschreibung der diagnostischen Genauigkeit mittels Sensitivität/Spezifität (ggf. ROC-Kurven) oder positivem/negativem Likelihood-Quotienten, daraus abgeleitet ggf. prädiktive Aussagen mittels z. B. positivem/negativem prädiktiven Wert. Dabei muss die vollständige Darstellung der Information aus Vierfeldertafeln der Primärstudien gewährleistet sein. Es ist z. B. nicht ausreichend, allein Werte für die Sensitivität anzugeben.
- Prognostische Aussagen bzgl. des Überlebens (z. B. gesamt, progressionsfrei) anhand geeigneter statistischer Maßzahlen (z. B. Kaplan-Meier-Schätzungen)

#### **4.2.1.4 Den systematischen Übersichten bzw. HTA-Berichten zugrunde liegende Studientypen**

- Querschnittsstudien, ggf. mit klinischer Nachbeobachtung, durchgeführt in der Anwendungssituation (gemäß Phase 3 nach Köbberling et al. [21])
- Prospektive Kohortenstudien

#### **4.2.1.5 Sonstige Charakteristika**

- Erreichen eines Mindestpunktwertes von 5 auf dem Qualitätsindex von Oxman/Guyatt
- Adressierung eines möglicherweise in den Primärstudien zu konstatierenden Verifikations- bzw. Inkorporations-Bias

#### **4.2.1.6 Ein-/Ausschlusskriterien**

Es werden alle systematischen Übersichten bzw. HTA-Berichte herangezogen, die

- alle nachfolgenden Einschlusskriterien und
- keines der nachfolgenden Ausschlusskriterien erfüllen.

<b>Einschlusskriterien</b> zur Evaluierung von diagnostischer und prognostischer Güte anhand von systematischen Übersichten bzw. HTA-Berichten	
E1b	Wie E1a (s. Abschnitt 4.1.2.1)
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	Systematische Übersichten und HTA-Berichte aus Querschnitts- oder prospektiven Kohortenstudien wie in den Abschnitten 4.2.1.4 und 4.2.1.5 beschrieben
E5b	Sprachen: Englisch, Deutsch, Französisch, Spanisch

<b>Ausschlusskriterien</b> zur Evaluierung von diagnostischer und prognostischer Güte anhand von systematischen Übersichten bzw. HTA-Berichten	
A1b	Mehrfachpublikationen ohne relevante Zusatzinformation
A2b	Keine Volltextpublikation verfügbar <sup>a</sup>
a: Als Vollpublikation gilt 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 QUOROM-Statements [28] genügt und eine Bewertung der Studie ermöglicht.	

## 4.2.2 Einschlusskriterien zur Ermittlung der diagnostischen und prognostischen Güte anhand von Primärstudien

### 4.2.2.1 Population

Folgende Patientenpopulationen werden eingeschlossen:

- Patienten mit nachgewiesenem malignem Lymphom, die ein Staging für die Primärtherapie erhalten
- Patienten nach Behandlung zur Erfassung des Behandlungsansprechens („Residual mass“)
- Patienten mit begründetem Verdacht auf ein Rezidiv

Unter Lymphomen im Sinne der Fragestellung werden alle nicht leukämisch verlaufenden lymphatischen Neoplasien verstanden. Ebenfalls ausgeschlossen werden Patienten mit multiplem Myelom.

Eingeschlossen werden Studien, bei denen der Anteil der Patienten mit malignen Lymphomen (wie zuvor definiert) mindestens 80 % beträgt oder bei denen Subgruppenanalysen für Patienten mit malignen Lymphomen durchgeführt wurden.

#### **4.2.2.2 Prüftechnologie, ggf. Vergleichstechnologie („Komparator“) und „Goldstandard“**

Prüftechnologie ist die (Vollring-)PET bzw. PET/CT, als Vergleichstechnologie („Komparator“) dient ggf. die konventionelle Diagnostik. Prüf- und ggf. Vergleichstechnologie müssen zu einem sogenannten Goldstandard in Beziehung gebracht worden sein. In den einzuschließenden Studien muss ein solches Ordnungsprinzip erkennbar und die Ergebnisse entsprechend dargestellt sein. PET bzw. PET/CT und ggf. die konventionelle Diagnostik (s. Abschnitt 2.2) dürfen nicht Bestandteil des „Goldstandards“ sein, bzw. das Problem eines möglicherweise resultierenden Verifikations- oder Inkorporations-Bias muss adäquat adressiert sein (z. B. durch den Versuch einer Adjustierung).

#### **4.2.2.3 Zielgrößen**

- Beschreibung der diagnostischen Genauigkeit mittels Sensitivität/Spezifität (ggf. ROC-Kurven) oder positivem/negativem Likelihood-Quotienten, daraus abgeleitet ggf. prädiktive Aussagen mittels z. B. positivem/negativem prädiktiven Wert. Dabei muss die vollständige Darstellung der Information aus Vierfeldertafeln gewährleistet sein. Es ist z. B. nicht ausreichend, allein Werte für die Sensitivität anzugeben.
- Prognostische Aussagen bzgl. Überleben (z. B. gesamt, progressionsfrei) anhand geeigneter statistischer Maßzahlen (z. B. Kaplan-Meier-Schätzungen)

#### **4.2.2.4 Studientypen**

- Querschnittsstudien, ggf. mit klinischer Nachbeobachtung, durchgeführt in der Anwendungssituation (gemäß Phase 3 nach Köbberling et al. [21])
- Prospektive Kohortenstudien

#### **4.2.2.5 Sonstige Charakteristika**

Es erfolgt keine weitere Einschränkung.

#### 4.2.2.6 Ein-/Ausschlusskriterien

Es werden alle Studien einbezogen, die

- alle nachfolgenden Einschlusskriterien und
- keines der nachfolgenden Ausschlusskriterien erfüllen.

<b>Einschlusskriterien</b> für Primärstudien zur Evaluierung von diagnostischer und prognostischer Güte	
E1c	Patienten mit malignem Lymphom (Hodgkin/Non-Hodgkin) wie in Abschnitt 4.2.2.1 beschrieben
E2c	Prüftechnologie: (Vollring-)PET bzw. PET/CT, Evaluierung wie in Abschnitt 4.2.2.2 beschrieben
E3c	Zielgrößen wie in Abschnitt 4.2.2.3 beschrieben
E4c	Querschnitts- oder prospektive Kohortenstudien wie in Abschnitt 4.2.2.4 beschrieben
E5c	Sprachen: Englisch, Deutsch, Französisch, Spanisch

<b>Ausschlusskriterien</b> für Primärstudien zur Evaluierung von diagnostischer und prognostischer Güte	
A1b	Mehrfachpublikationen ohne relevante Zusatzinformation
A2b	Keine Volltextpublikation verfügbar <sup>a</sup>
a: Als Vollpublikation gilt 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 [29] oder TREND-Statements [25] genügt und eine Bewertung der Studie ermöglicht.	

### 4.3 Informationsbeschaffung

#### 4.3.1 Literaturrecherche

Die Suche nach relevanten Studien bzw. systematischen Übersichten bzw. HTA-Berichten wird in folgenden Quellen durchgeführt:

Datenbanken bzw. Organisationen	
Bibliographische Datenbanken	<ul style="list-style-type: none"> <li>• EMBASE</li> <li>• MEDLINE</li> <li>• Cochrane Central Register of Controlled Trials (Clinical Trials)</li> <li>• Cancerlit</li> <li>• SciSearch</li> </ul>
HTA-Organisationen; Datenbanken für systematische Reviews	<ul style="list-style-type: none"> <li>• The Cochrane Database of Systematic Reviews (CDSR)</li> <li>• Database of Reviews of Effects (DARE)</li> <li>• Health Technology Assessment Database (HTA)</li> <li>• NHS Economic Evaluation Database (NHS EED)</li> <li>• International Health Technology Assessment Database (IHTA)</li> <li>• Health Services Technology Assessment Texts (HSTAT) der National Library of Medicine</li> <li>• TRIP Database</li> <li>• Datenbank der Deutschen Agentur für Health Technology Assessment (DAHTA)</li> <li>• WHO Health Evidence Network</li> <li>• Current Contents Medizin (CC Med) deutscher und deutschsprachiger Zeitschriften</li> </ul> <p>Zur Identifizierung (noch) nicht in den o. a. Datenbanken befindlicher Publikationen, laufender Projekte und in Bearbeitung befindlicher systematischer Übersichten bzw. HTA-Berichte werden durchsucht:</p> <ul style="list-style-type: none"> <li>• die Webseiten ausgewählter internationaler HTA-Institutionen</li> <li>• das Regional and National Research Program des National Research Register</li> <li>• das Register of Reviews des National Research Register</li> </ul>
Hersteller	<ul style="list-style-type: none"> <li>• Siemens, Philips, General Electric</li> <li>• Suche in Studienregistern der Hersteller, sofern verfügbar</li> </ul>
Unterlagen des G-BA	Es wurden (keine) Unterlagen durch den G-BA übermittelt.
Sonstiges	<ul style="list-style-type: none"> <li>• Suche in öffentlich zugänglichen Dokumenten der Zulassungsbehörden EMEA und FDA</li> <li>• Ggf. Kontaktaufnahme mit Sachverständigen/Experten/Fachgesellschaften</li> <li>• Ggf. Kontaktaufnahme mit Autoren einzelner Publikationen, z. B. zur Frage nach nicht publizierten Teilaspekten</li> </ul>

#### **4.3.2 Identifizierung relevanter Publikationen bzw. Studien**

Die Selektion relevanter Publikationen bzw. Studien aus den in der bibliographischen Literaturrecherche identifizierten Literaturzitate wird auf Basis des Titels und, sofern vorhanden, Abstracts von 2 Gutachtern unabhängig voneinander vorgenommen. Es werden in einem 1. Selektionsschritt alle Zitate aussortiert, die sicher als nicht für die Bewertung relevante Dokumente zum Projekt klassifiziert werden können. Alle verbleibenden Dokumente werden in einem 2. Selektionsschritt im Volltext gesichtet und abschließend als relevante bzw. nicht relevante Dokumente eingestuft. Diskrepanzen im Selektionsprozess zwischen den Gutachtern werden durch Diskussion aufgelöst.

Zusätzlich werden für die zwei folgenden Fragestellungen bzw. Situationen vorhandene systematische Übersichten oder HTA-Berichte nach weiteren relevanten Primärpublikationen durchsucht: wenn keine Bearbeitung auf der Basis von systematischen Übersichten oder HTA-Berichten vorgesehen (Bewertung des patientenrelevanten Nutzens) oder keine Bearbeitung möglich ist (beim Vorliegen von weniger als 2 systematischen Übersichten oder HTA-Berichten zu einer Fragestellung). Die Volltexte dieser aus den systematischen Übersichten oder HTA-Berichten identifizierten Publikationen werden von 2 Gutachtern unabhängig voneinander bezüglich ihrer Relevanz bewertet. Diskrepanzen zwischen den Gutachtern werden durch Diskussion aufgelöst.

Die Suche nach zusätzlichen relevanten Studien auf Webseiten ausgewählter internationaler HTA-Institutionen, in (Online-)Registern und Dokumenten von Zulassungsbehörden erfolgt durch einen Gutachter. Deren Relevanz wird wiederum von 2 Gutachtern unabhängig voneinander beurteilt.

#### **4.3.3 Suche nach zusätzlichen Informationen zu relevanten Publikationen**

Sofern in den vorhandenen Informationen (Publikationen) zu einer Studie bzw. zu einer systematischen Übersicht oder einem HTA-Bericht wesentliche, für die Nutzenbewertung relevante Aspekte offenbleiben, werden die zusätzlichen Informationen bei den Autoren angefragt, allerdings in der Regel nur dann, wenn das Publikationsalter nicht höher als 10 Jahre ist.

#### **4.3.4 Informationen zu relevanten unpublizierten und publizierten Studien aus den Anhörungsverfahren**

Die an die Veröffentlichung des (vorläufigen) Berichtsplans (bzw. der möglichen jeweiligen Amendments) sowie des Vorberichts sich jeweils anschließenden Anhörungsverfahren dienen ebenfalls der Informationsbeschaffung. Dabei beziehen sich das Anhörungsverfahren zum Berichtsplan insbesondere auf die Frage nach bislang unveröffentlichten Daten zu Fragestellungen der Nutzenbewertung und das Anhörungsverfahren zum Vorbericht vor allem

auf die Frage nach relevanten, in der Nutzenbewertung nicht berücksichtigten Studien bzw. systematischen Übersichten oder HTA-Berichten, unabhängig vom Publikationsstatus.

#### **4.4 Informationsbewertung**

Die Dokumentation der Studiencharakteristika und -ergebnisse bzw. der Ergebnisse von systematischen Übersichten oder HTA-Berichten sowie deren Bewertung erfolgt anhand von standardisierten Datenextraktionsbögen und Evidenztabelle. Ein Gutachter führt die Datenextraktion unter Verwendung des Extraktionsbogens durch. Ein zweiter Gutachter überprüft die Extraktion. Etwaige Diskrepanzen in der Bewertung werden durch Diskussion zwischen den Gutachtern, ggf. unter Hinzuziehung einer dritten Person, aufgelöst.

##### **4.4.1 Bewertung des Verzerrungspotenzials von Studien zur Bewertung des patienten-relevanten Nutzens**

Um darzustellen, mit welcher Sicherheit sich ein Effekt oder das Fehlen eines Effekts aus einer Studie ableiten lässt, erfolgt für die in die Nutzenbewertung eingeschlossenen Studien eine Bewertung des Verzerrungspotenzials der Ergebnisse.

Diese erfolgt sowohl für die Studie insgesamt als auch für die einzelnen Endpunkte. Zur Bewertung des Verzerrungspotenzials der Ergebnisse auf Studienebene im Falle von randomisierten kontrollierten Studien werden die folgenden Angaben zum Studiendesign und der Studiendurchführung systematisch extrahiert:

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuweisung
- Patientenfluss
- Verblindungsmethodik des Patienten sowie des Behandlers
- Umsetzung der prospektiven Planung
- Spezielle Aspekte, die Bias verursachen können (bspw. Dateninkonsistenzen)

Für nichtrandomisierte kontrollierte Studien werden die folgenden Angaben zum Studiendesign und der Studiendurchführung systematisch extrahiert:

- Gruppenzuteilung
- Störgrößenkontrolle
- Patientenfluss

- Verblindungsmethodik des Patienten sowie des Behandlers
- Umsetzung der prospektiven Planung
- Spezielle Aspekte, die Bias verursachen können (bspw. Dateninkonsistenzen)

Zur Bewertung des Verzerrungspotenzials der Ergebnisse der Endpunkte werden die folgenden Angaben systematisch extrahiert:

- Angaben zum Endpunkt für die daraus folgende Bewertung der Anfälligkeit für systematische Verzerrung (Messsicherheit)
- Verblindungsmethodik des Endpunkterhebers
- Umsetzung des ITT-Prinzips
- Relevante Dateninkonsistenz
- Statistische Auswertung
- Ergebnisdarstellung

Anhand der Bewertungen der oben genannten Aspekte wird das Verzerrungspotenzial der Ergebnisse der Studien in die Ausprägungen „hoch“ und „niedrig“ eingeteilt.

Die Ausprägungen werden vorab wie folgt definiert: Ein „niedriges Verzerrungspotenzial“ liegt vor, wenn davon ausgegangen wird, dass das Ergebnis (des Endpunktes bzw. der Studie insgesamt) durch die möglicherweise aufgezeigten Defizite im Studiendesign bzw. der Studiendurchführung nicht wesentlich verzerrt/beeinflusst wird. Ein „hohes Verzerrungspotenzial“ liegt vor, wenn davon ausgegangen wird, dass das Ergebnis (des Endpunktes bzw. der Studie insgesamt) infrage gestellt ist.

#### **4.4.2 Bewertung der Qualität von Primärstudien zur Bewertung der diagnostischen und prognostischen Güte**

Die Qualitätsbewertung wird mithilfe des QUADAS-Instruments erfolgen [16]. Dieses kann bei Bedarf an die Fragestellungen angepasst werden.

#### **4.4.3 Bewertung der Qualität von systematischen Übersichten und HTA-Berichten zur Bewertung der diagnostischen und prognostischen Güte**

Die Bewertung der allgemeinen Qualität einer systematischen Übersicht bzw. eines HTA-Berichts erfolgt anhand des validierten Qualitätsindex für systematische Übersichten von Oxman/Guyatt (s. Abschnitt 4.2). Dieser kann bei Bedarf an die Fragestellungen angepasst werden, z. B. über den Aspekt, wie in den systematischen Übersichten bzw. HTA-Berichten



mit dem Problem eines möglichen Verifikations- bzw. Inkorporationsbias umgegangen wurde (bzw. wie die Qualität der zugrunde liegenden Primärstudien generell berücksichtigt wurde), oder auch zur Frage nach der Qualität der Suchstrategie und/oder Aktualität.

#### **4.5 Informationssynthese und -analyse**

Die Informationen werden einer Informationssynthese und -analyse, wenn möglich bei Bewertungen anhand von Primärstudien unter Zuhilfenahme einer Meta-Analyse, unterzogen. Eine abschließende zusammenfassende Bewertung der Informationen erfolgt darüber hinaus in jedem Fall. Eine quantitative Synthese von Meta-Analysen ist allerdings nicht vorgesehen.

##### **4.5.1 Charakterisierung und Darstellung der Ergebnisse von einzelnen (Primär-) Studien zur Bewertung des patientenrelevanten Nutzens**

Im Bericht werden die Studien anhand von Designcharakteristika (Studiendesign, Studiendauer, Zahl der randomisierten/zugeteilten Patienten, Zuteilungsmechanismus, Ort und Zeitraum der Studiendurchführung, relevante Zielkriterien) beschrieben. Darüber hinaus werden die in der jeweiligen Studie verglichenen diagnostischen Strategien (z. B. eingesetzte Technologien, Vorselektion) dargestellt. Die Studienpopulationen werden weiterhin durch demographische Daten (Alter, Geschlecht), durch Charakteristika der betrachteten Krankheitsentitäten, durch prognostische Angaben und durch die Zahl der Studienabbrecher beschrieben. Die Studienpopulationen in den einzelnen Studien sowie die Interventionsgruppen innerhalb der einzelnen Studien werden deskriptiv verglichen. Unterschiede bezüglich der Studienpopulationen oder Interventionsgruppen werden im Bericht dargestellt.

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

##### **4.5.2 Charakterisierung und Darstellung der Ergebnisse von systematischen Übersichten oder HTA-Berichten**

Bei der Bewertung der diagnostischen und prognostischen Güte anhand von systematischen Übersichten oder HTA-Berichten wird ein Studienspiegel, d. h. eine Übersicht der jeweils eingeschlossenen Studien mit den wesentlichen Studiencharakteristika, erstellt. Die in den systematischen Übersichten oder HTA-Berichten dargestellten Ergebnisse aus den zugrunde liegenden Primärstudien werden dabei – soweit notwendig und möglich – zunächst nach den Fragestellungen bzw. Ordnungskriterien „Prüftechnologie“, ggf. „Komparator“ und „Goldstandard“ klassifiziert und qualitativ zusammengefasst. Weiterhin erfolgt eine

Abschätzung der Heterogenität der Ergebnisse innerhalb dieser Klassen, entweder direkt anhand von Angaben aus den systematischen Übersichten oder HTA-Berichten oder von eigenen Analysen. Beim Vorliegen inkonsistenter Ergebnisse aus mehreren Arbeiten zu einer Fragestellung werden mögliche Erklärungen für die beobachtete Heterogenität (z. B. Unterschiede bei der Berücksichtigung eines möglichen Verifikations- oder Inkorporations-bias) beschrieben [30].

Sollten nach Abschluss der Recherchen der eingeschlossenen systematischen Übersichten bzw. HTA-Berichte noch weitere relevante Primärstudien zu den Fragestellungen publiziert worden sein, so werden diese separat und entsprechend den (Primär-)Studien zur Bewertung des patientenrelevanten Nutzens aufgeführt (s. Abschnitt 4.5.1), falls davon ausgegangen werden kann, dass sie das Gesamtergebnis (aus den systematischen Übersichten bzw. HTA-Berichten) wesentlich beeinflussen, oder falls für die bearbeitete Fragestellung keine systematischen Übersichten bzw. HTA-Berichte in ausreichender Zahl und Güte vorliegen. Anderenfalls werden solche (Primär-)Studien nur kurz aufgelistet und begründet, warum nicht davon ausgegangen wird, dass durch sie das Gesamtergebnis wesentlich verändert werden könnte.

#### **4.5.3 Sensitivitätsanalyse**

Sensitivitätsanalysen sind insbesondere für die Bewertung des Verzerrungspotenzials der Ergebnisse der Endpunkte bei (Primär-)Studien zum patientenrelevanten Nutzen bzw. für die Qualitätsbewertung bei Studien zur diagnostischen und prognostischen Güte geplant.

#### **4.5.4 Subgruppenanalysen**

Subgruppenanalysen sind – soweit durchführbar und neben der ohnehin geplanten Aufteilung nach (Einzel-)Fragestellungen, insbesondere bei der Bewertung der diagnostischen und prognostischen Güte (bezüglich „Prüftechnologie“, ggf. „Komparator“ und „Goldstandard“) – für die folgenden Merkmale vorgesehen:

- Geschlecht
- Alter
- Krankheitsentitäten

## 5 Literaturverzeichnis

1. Alexander DD, Mink PJ, Adami HO, Chang ET, Cole P, Mandel JS et al. The non-Hodgkin lymphomas: a review of the epidemiologic literature. *Int J Cancer* 2007; 120 Suppl 12: 1-39.
2. Dreyling M., Hiddemann W., Trümper L., Pfreundschuh M. Leitlinie der DGHO: Maligne Lymphome [online]. Letztes Update Feb. 2007 [Zugriff am: 18. Juli 2007]; Gelesen unter: [http://www.dgho.de/cmsdata/file/file\\_181.pdf](http://www.dgho.de/cmsdata/file/file_181.pdf).
3. Bertz J., Giersiepen K., Haberland J., et.al. Krebs in Deutschland. Häufigkeiten und Trends [online]. Letztes Update 16. Feb. 2006 [Zugriff am: 18. Juli 2007]; Gelesen unter: [http://www.rki.de/cln\\_049/nn\\_204124/DE/Content/GBE/DachdokKrebs/Broschuere/broschuere\\_node.html?\\_nn=true](http://www.rki.de/cln_049/nn_204124/DE/Content/GBE/DachdokKrebs/Broschuere/broschuere_node.html?_nn=true).
4. Morton LM, Wang SS, Devesa SS, Hartge P, Weisenburger DD, Linet MS. Lymphoma incidence patterns by WHO subtype in the United States, 1992-2001. *Blood* 2006; 107(1): 265-276.
5. Hiddemann W., Dreyling M., Stein H. Lymphome. Neue Erkenntnisse und Therapiestrategien. Georg Thieme Verlag; Stuttgart; 2005.
6. Kasper L.D., Braunwald E., Fauci A.S., Hauser S.L., et.al. Harrison's Principles of Internal Medicine McGraw Hill Companies; New York; 2007.
7. Lohri A., Dellas S., Dirnhofer S., Herrmann R., et.al. Die malignen Lymphome (Teil1). *Schweiz Med Forum* 2002;(34): 773-780.
8. Stein H., Hummel M., Jöhrens K., Anagnostopoulos I. Klassifikation und Pathogenese der malignen Lymphome. *Internist* 2007;(48): 351-361.
9. Lohri A., Dirnhofer S., Gregor M., Jost L., et.al. Die malignen Lymphome (Teil2). *Schweiz Med Forum* 2002;(35): 803-809.
10. Sökler M., et.al. Lymphome. Empfehlungen zur Diagnostik Therapie und Nachsorge. [online]. Letztes Update Sept. 2006 [Zugriff am: 18. Juli 2007]; Gelesen unter: <http://www.medizin.uni-tuebingen.de/ccp/pdfinhal/lymphome.pdf>.
11. Lohri A., Jost L., Landmann Ch., Passweg J., Herrmann R. Die malignen Lymphome (Teil3). *Schweiz Med Forum* 2002;(36): 827-834.
12. Krause S.W., Schmitz S., Engert A. Leitlinie der DGHO: Morbus Hodgkin [online]. Letztes Update 02. Okt. 2006 [Zugriff am: 18. Juli 2007]; Gelesen unter: [http://www.dgho.de/cmsdata/file/file\\_177.pdf](http://www.dgho.de/cmsdata/file/file_177.pdf).
13. G-BA. Abschlussbericht des Gemeinsamen Bundesausschusses nach § 91 Abs. 7 SGB V „Krankenhausbehandlung“: Positronenemissionstomographie bei nichtkleinzelligen Lungenkarzinomen [online]. Letztes Update 02. Mär. 2006 [Zugriff am: 22. Aug. 2007]; Gelesen unter: <http://www.g-ba.de/downloads/40-268-86/2005-12-20-khb-PET-NSCLC-Abschluss.pdf>.

14. G-BA. Positronen-Emissions-Tomographie (PET): Zusammenfassender Bericht des Arbeitsausschusses "Ärztliche Behandlung" des Bundesausschusses der Ärzte und Krankenkassen über die Beratungen gemäß §135 Abs.1 SGB V [online]. Letztes Update 23. Mai 2002 [Zugriff am: 22. Aug. 2007];  
Gelesen unter: <http://www.g-ba.de/downloads/40-268-238/2003-05-23-Abschluss-PET.pdf>.
15. Whiting P, Rutjes AW, Dinnes J, Reitsma J, Bossuyt PM, Kleijnen J. Development and validation of methods for assessing the quality of diagnostic accuracy studies [online]. Letztes Update 05. Jan. 2004 [Zugriff am: 22. Aug. 2007];  
Gelesen unter: <http://www.hta.nhsweb.nhs.uk/fullmono/mon825.pdf>.
16. 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
17. 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.
18. Kirby AM, George MN. The role of FDG PET in the management of lymphoma: practical guidelines. Nucl Med Commun 2007; 28(5): 355-357.
19. Cheran SK, Herndon JE, Patz EF, Jr. 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.
20. How to read clinical journals: II. To learn about a diagnostic test. Can Med Assoc J 1981; 124(6): 703-710.
21. Memorandum for the evaluation of diagnostic measures. J Clin Chem Clin Biochem 1990; 28(12): 873-879.
22. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991; 11(2): 88-94.
23. Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. Letztes Update 18. Apr. 2006 [Zugriff am: 22. Aug. 2007];  
Gelesen unter: <http://www.g-ba.de/downloads/36-232-42/2006-07-07-VerfO.pdf>.
24. The CONSORT statement [online]. Letztes Update 17. Juli 2007 [Zugriff am: 22. Aug. 2007];  
Gelesen unter: <http://www.consort-statement.org/>.
25. The TREND statement [online]. Letztes Update 17. Juli 2007 [Zugriff am: 22. Aug. 2007];  
Gelesen unter: <http://www.trend-statement.org>.
26. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol 1991; 44(11): 1271-1278.
27. Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA et al. Agreement among reviewers of review articles  
2. J Clin Epidemiol 1991; 44(1): 91-98.

28. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; 354(9193): 1896-1900.
29. 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): W1-12.
30. Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *CMAJ* 1997; 156(10): 1411-1416.