

**Positronenemissionstomographie  
(PET) und PET/CT zur  
Erkennung von vitalem  
Myokardgewebe bei Patienten mit  
koronarer Herzerkrankung und  
eingeschränkter regionaler und  
globaler linksventrikulärer  
Funktion**

**Vorläufiger Berichtsplan**

Auftrag: D06-01L  
Version: 1.0  
Stand: 05.12.2011

# Impressum

**Herausgeber:**

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

**Thema:**

Positronenemissionstomographie (PET) und PET/CT zur Erkennung von vitalem Myokardgewebe bei Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler und globaler linksventrikulärer Funktion

**Auftraggeber:**

Gemeinsamer Bundesausschuss

**Datum des Auftrags:**

19.12.2006

**Interne Auftragsnummer:**

D06-01L

**Anschrift des Herausgebers:**

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen  
Dillenburger Str. 27  
51105 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

Berichte@iqwig.de

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

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

**Schlagwörter:** Koronararterienerkrankung, Myokardrevaskularisation, Myokard-Stunning, Positronenemissionstomographie, Computertomographie, systematische Übersicht

**Keywords:** Coronary Artery Disease, Myocardial Revascularization, Myocardial Stunning, Positron-Emission Tomography, Tomography, X-Ray Computed, Systematic Review

# Inhaltsverzeichnis

	Seite
<b>Tabellenverzeichnis</b> .....	<b>vi</b>
<b>Abkürzungsverzeichnis</b> .....	<b>vii</b>
<b>1 Hintergrund</b> .....	<b>1</b>
<b>1.1 Definition der koronaren Herzkrankheit</b> .....	<b>1</b>
<b>1.2 Epidemiologie und individuelle Krankheitslast der koronaren Herzerkrankung</b> .....	<b>1</b>
<b>1.3 Ursache der koronaren Herzerkrankung</b> .....	<b>2</b>
<b>1.4 Klassifikation der koronaren Herzerkrankung</b> .....	<b>2</b>
<b>1.5 Verlauf der koronaren Herzerkrankung</b> .....	<b>3</b>
<b>1.6 Therapie der koronaren Herzkrankheit</b> .....	<b>4</b>
<b>1.7 Verfahren der Vitalitätsdiagnostik</b> .....	<b>5</b>
<b>1.8 Empfehlungen wissenschaftlicher Fachgesellschaften zur Vitalitätsdiagnostik mittels PET</b> .....	<b>8</b>
<b>2 Ziele der Untersuchung</b> .....	<b>9</b>
<b>3 Projektbearbeitung</b> .....	<b>11</b>
<b>3.1 Zeitlicher Verlauf des Projekts</b> .....	<b>11</b>
<b>4 Methoden</b> .....	<b>12</b>
<b>4.1 Ermittlung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT</b> .....	<b>13</b>
4.1.1 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT anhand von systematischen Übersichten.....	14
4.1.2 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT anhand von Primärstudien.....	14
4.1.2.1 Populationen.....	14
4.1.2.2 Prüf- und Vergleichsinterventionen .....	14
4.1.2.3 Patientenrelevante Endpunkte.....	15
4.1.2.4 Studientypen.....	15
4.1.2.5 Studiendauer.....	16
4.1.2.6 Sonstige Studiencharakteristika .....	16
4.1.3 Tabellarische Darstellung der Einschlusskriterien („Nutzenstudien“).....	16
<b>4.2 Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT</b> .....	<b>17</b>
4.2.1 Einschlusskriterien zur Ermittlung der diagnostischen und prognostischen Güte anhand von systematischen Übersichten .....	18

4.2.1.1	Populationen.....	18
4.2.1.2	Prüftechnologien, Vergleichstechnologien („Komparatoren“) und Referenztest.....	18
4.2.1.3	Zielgrößen .....	19
4.2.1.4	Den systematischen Übersichten zugrunde liegende Studientypen.....	19
4.2.1.5	Sonstige Charakteristika der systematischen Übersichten.....	20
4.2.1.6	Tabellarische Darstellung der Einschlusskriterien (systematische Übersichten zur diagnostischen und prognostischen Güte).....	21
4.2.2	Einschlusskriterien zur Ermittlung der diagnostischen und prognostischen Güte anhand von Primärstudien („Ergänzungsrecherche“).....	21
4.2.2.1	Populationen.....	21
4.2.2.2	Prüftechnologien (Indextest I), Vergleichstechnologien (Indextest II) und Referenztests .....	21
4.2.2.3	Zielgrößen .....	22
4.2.2.4	Studientypen.....	22
4.2.2.5	Sonstige Studiencharakteristika .....	22
4.2.2.6	Tabellarische Darstellung der Einschlusskriterien („Ergänzungsrecherche“ nach Primärstudien zur diagnostischen und prognostischen Güte) .....	22
4.2.3	Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen.....	22
<b>4.3</b>	<b>Informationsbeschaffung.....</b>	<b>23</b>
4.3.1	Bibliografische Literaturrecherche .....	23
4.3.2	Suche nach weiteren publizierten und nicht publizierten Studien und systematischen Übersichten.....	23
4.3.3	Selektion relevanter Studien und systematischer Übersichten .....	23
<b>4.4</b>	<b>Informationsbewertung .....</b>	<b>24</b>
4.4.1	Bewertung von Studien zum patientenrelevanten Nutzen und Schaden .....	24
4.4.2	Bewertung von systematischen Übersichten zur diagnostischen und prognostischen Güte .....	25
4.4.3	Bewertung von Primärstudien zur diagnostischen und prognostischen Güte ...	25
<b>4.5</b>	<b>Informationssynthese und -analyse .....</b>	<b>26</b>
4.5.1	Gegenüberstellung der Ergebnisse der Einzelstudien.....	27
4.5.1.1	Gegenüberstellung der Ergebnisse von (Primär-)Studien zur Bewertung des patientenrelevanten Nutzens und Schadens .....	27
4.5.1.2	Gegenüberstellung der Ergebnisse von systematischen Übersichten zur diagnostischen und prognostischen Güte.....	28
4.5.1.3	Gegenüberstellung der Ergebnisse von Primärstudien zur diagnostischen und prognostischen Güte („Ergänzungsrecherche“) .....	28

4.5.1.4	Gegenüberstellung der Ergebnisse aus systematischen Übersichten und der Ergänzungsrecherche („Robustheitsprüfung“)	29
4.5.2	Meta-Analysen	29
4.5.3	Sensitivitätsanalyse	30
4.5.4	Subgruppenmerkmale und andere Effektmodifikatoren	31
<b>5</b>	<b>Literaturverzeichnis</b>	<b>32</b>

**Tabellenverzeichnis**

	<b>Seite</b>
Tabelle 1. Klassifikation der Angina pectoris (CCS).....	3
Tabelle 2. Klassifikation der Herzinsuffizienz (NYHA) .....	3
Tabelle 3: Klassifikation für Evaluierungsstudien zu diagnostischen Testverfahren .....	12
Tabelle 4: Evidenzklassifizierung des G-BA für Unterlagen zu diagnostischen Methoden....	13
Tabelle 5: Übersicht der Einschlusskriterien für Primärstudien zur Nutzenbewertung der PET bzw. PET/CT.....	17
Tabelle 6: Übersicht der Einschlusskriterien für systematische Übersichten zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT.....	21
Tabelle 7: Übersicht der Einschlusskriterien für die Primärliteratur zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT (Ergänzungsrecherche).....	22

**Abkürzungsverzeichnis**

<b>Abkürzung</b>	<b>Bedeutung</b>
ASNC	American Society of Nuclear Cardiology
CABG	Coronary artery bypass graft (Koronararterien-Bypass-Operation)
CCS	Canadian Cardiovascular Society
CCT	Controlled clinical trial (kontrollierte klinische Studie)
CT	Computertomographie
DGK	Deutsche Gesellschaft für Kardiologie
DGN	Deutsche Gesellschaft für Nuklearmedizin e. V.
EANM	European Association of Nuclear Medicine
Echo	Echokardiographie
ECNC	European Council of Nuclear Cardiology
EF	Ejektionsfraktion
EKG	Elektrokardiogramm
ESCR	European Society of Cardiac Radiology
FDG	Fluordeoxyglukose
G-BA	Gemeinsamer Bundesausschuss
HFI	Health-Force-Initiative
HSROC	hierarchical summary receiver operating curve (Methode zur Bewertung und Optimierung von Analyse-Strategien)
HTA	Health Technology Assessment
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-Treat
KHK	Koronare Herzkrankheit
LDDE	Low-dose dobutamine stress echocardiography, (Stress-Echokardiographie)
LR	Likelihood Ratio
LVEF	Linksventrikuläre Ejektionsfraktion
MID	Minimal Important Difference
MRT	Magnetresonanztomographie
NVL KHK	Nationale Versorgungsleitlinie Chronische Koronare Herzerkrankung
NYHA	New York Heart Association



<b>Abkürzung</b>	<b>Bedeutung</b>
PCI	Percutaneous coronary intervention
PET	Positronenemissionstomographie
PET/CT	Integration von PET und CT in einem Untersuchungsgerät
PICO	Population, Interventions, Comparators (Vergleichsintervention), Outcomes (Ergebnis – Endpunkt)
PTCA	Percutaneous transluminal coronary angioplasty (Perkutan transluminale koronare Angioplastie)
RCT	Randomized controlled trial (Randomisierte kontrollierte Studie)
RNV	Radionuklidventrikulographie
ROC	Receiver operating characteristic (Methode zur Bewertung und Optimierung von Analyse-Strategien)
SF-36	Short Form-36 Health Survey
SMD	Standardized mean differences (Standardisierte Mittelwertdifferenzen)
SN	Sensitivität
SP	Spezifität
SPECT	Single-Photon Emission Computed Tomography (Einzelphotonenemissions-Computertomographie, Schnittbildverfahren der Nuklearmedizin)
VOPT	Verification of Only Positive Testers

## **1 Hintergrund**

### **1.1 Definition der koronaren Herzkrankheit**

Die koronare Herzkrankheit (KHK) ist nicht einheitlich definiert. Oft wird sie als „die Manifestation der Atherosklerose an den Herzkranzarterien“ angesehen [1,2]. Sie führt häufig zu einem Missverhältnis zwischen Sauerstoffbedarf und Sauerstoffangebot im Herzmuskel. Die KHK ist mit einem erhöhten Morbiditäts- und Mortalitätsrisiko verbunden.

Leitsymptome sind bei Belastung auftretende Enge- oder Druckempfindungen im Brustkorb (Angina pectoris), die in fortgeschrittenem Stadium auch im Ruhezustand auftreten können. Diese Symptome können durch Schmerzen in der Brust, Luftnot, Angst, Schweißausbruch, Blutdruckabfall und gesteigerte Herzfrequenz begleitet werden. Ein Angina-pectoris-Anfall kann beispielsweise durch körperliche oder psychische Belastung ausgelöst werden und dauert in der Regel wenige Minuten. Die Symptome können durch körperliche Ruhe oder Nitroglyzeringabe verbessert werden [1].

Die chronische Minderdurchblutung des Myokards kann dessen Pumpleistung verschlechtern (kontraktile Dysfunktion) und so zu einer Herzinsuffizienz (z. B. Herzleistungsschwäche) führen (ischämische Kardiomyopathie). Für die Quantifizierung der Herzleistung wird oft die Auswurfraction (ejection fraction, EF) verwendet. Sie errechnet sich aus dem prozentualen Anteil des ausgeworfenen Blutvolumens im Vergleich des systolischen zum diastolischen Füllungsvolumen. Der Mittelwert für die Auswurfraction des linken Ventrikels (left ventricular ejection fraction, LVEF) liegt im Normalfall in Ruhe bei 62 %, im unteren Grenzbereich bei 50 % [3]. Eine schwerwiegende Beeinträchtigung der linksventrikulären Funktion wird in Studien und Leitlinien von Fachgesellschaften im Wesentlichen mit einer Auswurfraction von unter 35 % definiert [4-8].

### **1.2 Epidemiologie und individuelle Krankheitslast der koronaren Herzerkrankung**

Die geschätzte Prävalenz für KHK lag im Jahr 2009 bei 7,8 % [9]. Männer erkrankten mit einer Häufigkeit von 9,2 %, Frauen von 6,5 %. Für Personen ab dem 65. Lebensjahr betrug der Anteil 28,2 % für Männer und 19,1 % für Frauen [9].

Die KHK ist Ursache für circa jeden fünften Sterbefall in Deutschland. Laut Todesursachenstatistik starben im Jahr 2010 in Deutschland 7,7 % aller verstorbenen Frauen und 7,9 % aller verstorbenen Männer an einer KHK [10].

Die Mortalität ist allerdings über die vergangenen drei Jahrzehnte in den westlichen Industrienationen deutlich rückläufig [11,12]. Die Lebensqualität der Patienten mit KHK ist in verschiedener Hinsicht eingeschränkt: einerseits durch Symptome (Angina pectoris, Symptome der Herzinsuffizienz), andererseits durch Komorbiditäten (z. B. Diabetes mellitus), und psychische Auswirkungen (Angst, Depression). Letztere sind wiederum eigenständige Risikofaktoren für die Entwicklung einer KHK [13].

### **1.3 Ursache der koronaren Herzerkrankung**

Die KHK ist bei der überwiegenden Mehrheit der Patienten durch eine mit dem Alter und anderen exogenen und endogenen Risikofaktoren bedingte zunehmende Arterienverkalkung (Arteriosklerose) verursacht [1,2]. Makroskopisch imponieren in betroffenen veränderten Gefäßen atherosklerotische „Plaques“. Durch komplexe pathophysiologische Prozesse der beteiligten Zellen (z. B. Endothelzellen, glatte Muskelzellen, Makrophagen) und Faktoren der Gefäßwand (z. B. Wachstumsfaktoren, LDL-Cholesterin) resultieren letztlich über die zunehmende Plaquebildung und Gefäßwandverdickung eine fortschreitende Verkleinerung des Gefäßlumens und somit eine verminderte Kapazität des Herzens, die Sauerstoffversorgung des Myokards an den Bedarf anzupassen. Körperliche Aktivität erfordert eine erhöhte Herzarbeit und somit gesteigerte Koronarperfusion. Liegt jedoch der Bedarf höher als die maximal mögliche Koronarperfusion entsteht eine Minderversorgung, die die Symptome der KHK herbeiführt.

Die atherosklerotisch veränderten Gefäße können lokal zu einer Aktivierung der Blutgerinnung mit Thrombusbildung führen. Wird eine Koronararterie so plötzlich verschlossen, führt dies oft zu dem klinischen Bild eines akuten Myokardinfarktes. Je nach Lokalisation des Ereignisses kann dadurch auch Kammerflimmern und somit der „plötzliche Herztod“ ausgelöst werden. Alternativ können derartige Koronarthrombosen klinisch unbemerkt („stummer Herzinfarkt“) ablaufen und so zu einer chronischen Schädigung des Herzens führen.

Die Wahrscheinlichkeit der Manifestation einer koronaren Herzerkrankung ist deutlich abhängig von Alter und Geschlecht. Wie bereits o. g. ist die Mortalität im Geschlechtervergleich ähnlich, wobei Männer früher an KHK erkranken. Neben Alter und Geschlecht zählen zu den nicht zu beeinflussenden Risikofaktoren noch erbliche Faktoren (Auftreten von KHK bei Verwandten ersten Grades im jüngeren Lebensalter). Zu den klassischen „beeinflussbaren“ Risikofaktoren zählen Nikotinabusus, Hypertonie, Adipositas, Hyperlipoproteinämie, Diabetes mellitus und Bewegungsmangel [14,15].

### **1.4 Klassifikation der koronaren Herzerkrankung**

Angina pectoris, das primäre Symptom der koronaren Herzkrankheit, wird mittels der Kriterien der Canadian Cardiovascular Society klassifiziert (CCS) [16]. Eine deutsche Übersetzung durch die Deutsche Gesellschaft für Kardiologie (DKG) findet sich in Tabelle 1 [2].

Tabelle 1. Klassifikation der Angina pectoris (CCS)

CCS <sup>(a)</sup> -Grad	Definition	Beispiel
I	Keine Angina bei normaler Belastung, Angina bei sehr hoher oder andauernder Anstrengung	Angina z. B. beim Schneeräumen, beim Dauerlauf
II	Geringe Einschränkung bei normalen Tätigkeiten	Angina z. B. beim schnellen Treppensteigen, beim Bergaufgehen, bei Belastung kurz nach dem Aufwachen
III	Deutliche Einschränkung der Leistungsfähigkeit	Angina z. B. beim An- und Ausziehen, längerem langsamen Gehen, leichter Hausarbeit
IV	Angina bei jeder Belastung oder in Ruhe	Angina unterhalb der bei Grad III genannten Belastungen

Herzinsuffizienz wird mittels der Kriterien der New York Heart Association Society (NYHA) klassifiziert [17] (Tabelle 2).

Tabelle 2. Klassifikation der Herzinsuffizienz (NYHA)

NYHA-Grad	Definition
I	Herzerkrankung ohne körperliche Limitation. Alltägliche körperliche Belastung verursacht keine inadäquate Erschöpfung, Rhythmusstörungen, Luftnot oder Angina pectoris.
II	Herzerkrankung mit leichter Einschränkung der körperlichen Leistungsfähigkeit. Keine Beschwerden in Ruhe und bei geringer Anstrengung. Stärkere körperliche Belastung verursacht Erschöpfung, Rhythmusstörungen, Luftnot oder Angina pectoris, z. B. Bergaufgehen oder Treppensteigen.
III	Herzerkrankung mit höhergradiger Einschränkung der körperlichen Leistungsfähigkeit bei gewohnter Tätigkeit. Keine Beschwerden in Ruhe. Geringe körperliche Belastung verursacht Erschöpfung, Rhythmusstörungen, Luftnot oder Angina pectoris, z. B. Gehen in der Ebene.
IV	Herzerkrankung mit Beschwerden bei allen körperlichen Aktivitäten und in Ruhe. Bettlägerigkeit.

## 1.5 Verlauf der koronaren Herzerkrankung

Die koronare Herzerkrankung verläuft mit zunehmendem Alter progredient. Durch die Prädisposition zur erhöhten Blutgerinnung kann plötzlich eine Thrombose einer Koronararterie mit der Folge eines akuten Myokardinfarktes entstehen. Zusätzlich wird die „Plaque-Ruptur“, das plötzliche Aufreißen der atherosklerotischen Gefäßveränderungen und dadurch Initiierung einer Gefäßthrombose, als Auslöser akuter Ereignisse wie Myokardinfarkt diskutiert. Typisch klinische Zeichen derartiger „akuter Koronarsyndrome“ sind starke Schmerzen im Brustkorb beispielsweise mit Ausstrahlung in Schulter, Rücken und linken Arm (instabile Angina pectoris). In der Folge kann es zu einem irreversiblen Gewebeeruntergang kommen (transmuraler oder nicht transmuraler Myokardinfarkt), der von tödlichen

Herzrhythmusstörungen (Kammerflimmern, plötzlicher Herztod) oder einem tödlichen akuten Herzversagen begleitet sein kann [18].

Die chronische KHK kann infolge der Pumpleistungsschwäche des Herzens außerdem zur Herzinsuffizienz mit verminderter körperlicher Leistungsfähigkeit, Luftnot und Beinödemen führen [1]. Es resultiert ein Rückstau des Blutes in die Lungen (Linksherzinsuffizienz), in den restlichen Körper (Rechtsherzinsuffizienz) oder in den gesamten Körper (globale Herzinsuffizienz). Typischerweise steigt der enddiastolische Druck in den betroffenen Herzhöhlen, die Pumpleistung ist parallel dazu herabgesetzt. Der Organismus versucht die chronische Herzinsuffizienz beispielsweise tachykard, d. h. durch eine gesteigerte Herzfrequenz, auszugleichen. Ist eine Kompensation nicht mehr möglich, sammelt sich Wasser im Gewebe an. Die Folgen sind Lungenödem und Unterschenkelödeme [18].

## 1.6 Therapie der koronaren Herzkrankheit

Die medikamentöse Behandlung zielt auf eine Reduktion des Sauerstoffbedarfs zur Kompensation der Minderperfusion ab. Dies kann durch Verkleinerung des Gefäßwiderstandes beispielsweise kurzfristig durch die Verabreichung von Nitraten oder mit Beta-blockern zur Verringerung der Herzfrequenz („negative Chronotropie“) erreicht werden. Betablocker und Kalziumantagonisten können dieses Ziel zusätzlich durch Senkung der Kontraktionskraft des Myokards („negative Inotropie“) erreichen. Dauerhafte Thrombozytenaggregationshemmung beeinflusst die Prognose von KHK-Patienten günstig [1].

Weiterhin wird mit teilweise medikamentösen Behandlungen das kardiovaskuläre Risikoprofil beeinflusst. Insbesondere die Regulierung des Blutdrucks, des Serumcholesterins und des Blutzuckers in die von der Leitlinie vorgegebenen Zielbereiche, aber auch Verhaltensänderungen (Rauchen, Bewegung, Ernährung) kommen als Maßnahmen in Frage.

Neben den medikamentösen Therapien besteht die Möglichkeit, die Perfusion des Myokards entweder minimalinvasiv durch Ballon-Angioplastie (percutaneous transluminal coronary angioplasty, PTCA; percutaneous coronary intervention, PCI [19]) oder chirurgisch durch die Anlage von aorto-koronaren Bypässen (coronary artery bypass graft, CABG) zu verbessern. Bei diesen Verfahren wird jeweils von „Revaskularisation“ gesprochen [14]. Bei hochgradig eingeschränkter LV-Funktion ist die Bypassoperation eine effektive Intervention für Patienten mit koronarer Herzkrankheit und Einschränkung der linksventrikulären Auswurfraction unter 35 % [7,20]. Hintergrund beider Revaskularisations-Verfahren ist, dass gezeigt wurde, dass eine bessere Herzleistung mit einer verbesserten Prognose dieser Patienten einhergehen kann [21].

Dysfunktionales Myokard kann nicht erfolgreich revaskularisiert werden, wenn das Gewebe narbig verändert, d. h. terminal geschädigt ist [22]. Daher wird bei Patienten mit einer höhergradigen Herzleistungsschwäche (linksventrikulären Dysfunktion) eine Vitalitätsdiagnostik zur Differenzierung von terminal geschädigtem (Narbengewebe nach Myokardinfarkt) und sog. „viablen“ Myokard (welches nach Wiederherstellung einer verbesserten Durchblutung potenziell für die Herzarbeit wiedergewonnen werden kann)

empfohlen [1,7]. Viabilität bezeichnet einen reversiblen Zustand einer Dysfunktion des Myokards, bei dem das betreffende Myokard-Segment minderperfundiert aber nicht abgestorben ist. Da solche Segmente durch die Verbesserung der Perfusion die kontraktile Funktion wieder aufnehmen können, spricht man auch von hibernierendem Myokard (Herzmuskelanteile „im Winterschlaf“) [22,23]. Etwa die Hälfte der Patienten mit koronarer Herzkrankheit und chronischer Dysfunktion des Herzmuskels haben hibernierende Myokard-Segmente [24]. Umgekehrt gilt ein dysfunktionelles Myokard-Segment nur dann als viabel, wenn die Funktion nach der Revaskularisation verbessert werden kann [25]. Mit Blick auf den spezifischen Fokus dieses Berichts auf die Unterscheidung zweier Typen dysfunktionalen Gewebes wird im Folgenden je nach Kontext die Bezeichnung „vitales Myokard“ gleichbedeutend mit „viablen bzw. hibernierendem Myokard“ verwendet. Entsprechend ist der Begriff Vitalitätsdiagnostik zu verstehen.

Angesichts einer periinterventionellen bzw. perioperativen Mortalität von ca. 5 % [26] stellt sich die Frage, ob die Vitalitätsdiagnostik bei der Erfolgs- / Risikoabschätzung für die Entscheidung über Revaskularisierungs-Maßnahmen eine wichtige Information darstellen kann. Neben der Indikationsstellung könnte die Vitalitätsdiagnostik das Vorgehen während des Eingriffs steuern (z. B. Anzahl und geeignete Lokalisation von Bypassen).

## 1.7 Verfahren der Vitalitätsdiagnostik

Zur Unterscheidung zwischen vitalem (hibernierendem) und avitalem (fibrosiertem) Myokard stehen verschiedene nicht invasive diagnostische Verfahren zur Verfügung [22,23,25,27,28].

- Stress-Echokardiographie (Low-dose dobutamine stress echocardiography, LDDE)
- SPECT (single-photon emission computed tomography)
- Magnetresonanztomographie (MRT) mit medikamentöser Belastung
- Kardio-Computertomographie (CT) mit medikamentöser Belastung
- PET (Positronenemissionstomographie) oder PET/CT

Verfahren zur Darstellung der Gefäße und Ventrikel (Kardioangiographie oder Radionuklidventrikulographie [RNV]) erlauben keine Beurteilung der Vitalität von Myokard-Segmenten, weil nicht das fragliche Gewebe direkt dargestellt wird [29,30].

### Dobutamin-Stress-Echokardiographie

Die sonographische Darstellung des Herzens (Echokardiographie, Echo) erlaubt die Beurteilung segmentaler Myokardbewegungen. Unter adrenerger Stimulation (z. B. mit Dobutamin) können Wandregionen mit eingeschränkter („hypokinetisch“) oder nicht vorhandener („akinetisch“) Kontraktilität identifiziert werden. Nur bei vitalem Myokard führt eine adrenerge Stimulation zur Verbesserung der systolischen Kontraktion in dysfunktionalen Segmenten. Bei stärkerer Stimulation folgt bei vitalem (aber hibernierendem) Myokard, bedingt durch eine stress-induzierte Ischämie, auf die zu beobachtende Verbesserung eine Verschlechterung der systolischen Kontraktion („biphasisches Verhalten“) [31].

## **SPECT**

Das Myocard-SPECT (Myokardszintigraphie) ist ein nuklearmedizinisches Verfahren zur zweidimensionalen metabolischen Bildgebung des Myokardgewebes [27]. Dabei wird ein Radionuklid (Thallium-201 oder Technetium-99m) intravenös injiziert. In der Bildgebung im zeitlichen Verlauf reflektiert die Traceraufnahme die Durchblutung der Myokardsegmente und zeigt damit deren Vitalität an [31].

## **Magnetresonanztomographie**

Die mit dem Herzzyklus synchronisierte Magnetresonanztomographie (ECG-gated MRI) kann die kontraktile Wandbewegungen aller myokardialen Segmente mit hoher Präzision und räumlicher Auflösung darstellen. Ähnlich der Stressechokardiographie kann die Vitalitätsdiagnostik auch mit der MRT durch medikamentös induzierte Belastung trennschärfer zwischen vitalem und fibrotischem Gewebe differenzieren [27,28,32].

## **Kardio-Computertomographie**

Die ultraschnelle Mehrzeilen-Kardio-Computertomographie (Ultrafast-Cardio CT; ebenfalls Elektrokardiogramm [EKG] synchronisiert) stellt inzwischen die Standard-Technologie innerhalb der CT-diagnostischen Verfahren dar und bietet im Vergleich zu anderen bildgebenden Verfahren in der Kardiologie bei relativ niedriger Strahlenbelastung wegen ihrer hohen räumlichen und zeitlichen Auflösung gute Voraussetzungen, um mithilfe von Kontrastmittel Narbengewebe von vitalem Myokard zuverlässig zu differenzieren. Die Erkennung vitaler Myokardsegmente erfolgt – vergleichbar zur Kardio-MRT – unter medikamentös induzierter Belastung im Vergleich des Kontraktions-Verhaltens zwischen verschiedenen Belastungsstufen.

## **PET und PET/CT**

PET und PET/CT sind nicht invasive diagnostische Verfahren, die freigesetzte Photonen registrieren. Dazu werden radioaktive Substanzen (sogenannte Tracer) über eine Vene appliziert. Durch den Zerfall des Tracers werden Positronen freigesetzt. Diese kollidieren mit Elektronen, die im Körper vorhanden sind. Mittels ringförmig um den Patienten angeordneter Photonendetektoren werden die dabei freigesetzten Photonen erfasst. Aus der räumlichen Verteilung der registrierten Photonen wird auf die räumliche Verteilung der radioaktiven Substanz (des sogenannten Tracers) im Körperinneren geschlossen.

In der Anfangszeit der PET wurden sogenannte Nichtvollringsysteme hergestellt. Die Empfindlichkeit der später entwickelten Vollring-PET-Systeme war den ursprünglichen so stark überlegen, dass nur 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 komplementär 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.

In der kardiologischen Vitalitätsdiagnostik kommt am häufigsten 18F-Fluordesoxyglucose (FDG) zum Einsatz.

Hibernierendes Myokard zeigt sich in der PET dadurch, dass das betreffende Gewebe Merkmale eines Energiestoffwechsels zeigt, dabei aber vermindert durchblutet ist. Fibrotische Segmente zeigen im Vergleich dazu keinen herzmuskeltypischen Energiestoffwechsel und sind minderdurchblutet. Nicht geschädigtes Myokard weist eine gute Perfusion und normalen Energiestoffwechsel auf. Bei hibernierendem Myokard wird wegen der Nichtübereinstimmung der beiden Vitalitätsmerkmale (Perfusion; Metabolismus) in diesem Zusammenhang von „mismatch“ gesprochen [33,34]. Über das Vorliegen von Perfusions-Metabolismus-mismatch wird jedes einzelne von 17 Segmenten (auf Basis unterschiedlicher Analysemethoden) semiquantitativ befundet [34,35]. Ein „mismatch Score“ drückt als Prozentwert den Anteil viablen Myokards am Gesamtmyokard aus. Sowohl die Festlegung, ob ein einzelnes myokardiales Segment hibernierend ist, als auch die Festlegung, ob ein Gesamtmyokard-Score eine Indikation für eine revaskularisierende Maßnahme darstellt, werden subjektiv definiert.

Im Unterschied zu früheren Untersuchungsprotokollen, bei denen für diese doppelte Information nacheinander zwei verschiedene radioaktive Tracer verwendet werden mussten, ist es heute nach den Empfehlungen der American Society of Nuclear Cardiology (ASNC) [34] mit dem Glukose analogen Tracer 18F-Fluordesoxyglucose (FDG) möglich, sowohl Durchblutung (Perfusion) als auch Energiestoffwechsel in einem Untersuchungsablauf abzubilden [36].

Die Aktivitätsanreicherungen bei der PET lassen sich anatomisch nicht immer präzise lokalisieren, da in PET-Bildern in erster Linie Stoffwechselprozesse abgebildet werden; hinzu kommt die begrenzte Ortsauflösung von etwa 4 bis 6 mm. Technische Weiterentwicklungen stellen Kombinationen aus PET und CT oder MRT in einem Untersuchungsgerät (PET/CT bzw. MRT) dar. Die Informationen aus der CT erlauben eine genauere anatomische Bestimmung. Bei einer Kombination der beiden Verfahren (PET/CT) wird der Patient in einem Untersuchungsgang durch die beiden Detektorringssysteme der 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 von CT oder MRT (Ortsauflösung < 1 mm) mit den hochsensitiven Stoffwechselformationen 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 [37]. PET/MRT-Geräte sind erst an einigen wenigen Zentren in Deutschland verfügbar.

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.

### **1.8 Empfehlungen wissenschaftlicher Fachgesellschaften zur Vitalitätsdiagnostik mittels PET**

In der nationalen Versorgungsleitlinie Chronische Koronare Herzerkrankung (NVL KHK) finden die PET oder PET/CT keine Erwähnung [1]. In der Leitlinie der European Society of Cardiology zum Thema „myocardial revascularization“ wird erklärt, dass die verschiedenen Verfahren zur Diagnostik der Viabilität des Herzgewebes sich möglicherweise geringfügig hinsichtlich der Genauigkeitsparameter unterscheiden, dass aber randomisierte Studien fehlten, um einen Zusatznutzen einzelner Verfahren gegenüber anderen erkennen zu lassen [7]. In der entsprechenden Leitlinie derselben Gesellschaft zum Thema „Heart Failure“ werden als viabilitätsdiagnostische Verfahren LDDE, SPECT, PET oder MRI sowie CT angeführt. Es heißt, die Verfahren seien vergleichbar in ihrer Genauigkeit. Während zwar eine allgemeine Empfehlung gegeben wird, der Frage der Vitalität des Myokards im Rahmen der Planung revaskularisierender Maßnahmen Beachtung zu schenken, gibt es keine Empfehlung für ein bestimmtes diagnostisches Verfahren [5].

Die Deutsche Gesellschaft für Nuklearmedizin (DGN) empfiehlt mit dem Verweis auf deren prognostischen Wert zur Vitalitäts-Diagnostik des Myokards SPECT oder PET [38]. Gemäß eines gemeinsamen Positionspapiers der European Association of Nuclear Medicine (EANM), der European Society of Cardiac Radiology (ESCR) und des European Council of Nuclear Cardiology (ECNC) gibt es Hinweise, dass die PET/CT bei Patienten mit intermediärem Risiko für die koronare Herzkrankheit (hybrid cardiac imaging) genauere Informationen zur Diagnose und Prognose ermöglicht [39]. Die Bestätigung einer vermuteten klinisch-praktischen Bedeutung der PET/CT bei diesen Patienten stehe jedoch noch aus und solle in größeren Kohortenstudien untersucht werden [39].

## 2 Ziele der Untersuchung

Die vorliegende Untersuchung soll 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 von vitalem (hibernierendem; synonym: viablem) versus avitalem (fibrotischem) Myokard verbunden ist. Die Fragestellung der Erkennung von vitalem Myokardgewebe bezieht sich auf Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion.

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

Alle Fragestellungen werden auch im Rahmen der deutschen Strahlenschutzgesetzgebung erörtert.

Folgende Technologievergleiche werden betrachtet:

- 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 wird im vorliegenden Bericht die Anwendung der potenziellen Alternativen verstanden, welche die EG, die SPECT und die MRT einschließen. Es sollen aber darüber hinaus alle weiteren in den gefundenen Studien angestellten diagnostischen Vergleiche dargestellt werden.

Es erfolgt keine Eingrenzung der Fragestellungen auf bestimmte Tracer.

### **Allgemeine Voraussetzungen für Studien zur Nutzenbewertung („Nutzenstudien“)**

Unter „patientenrelevantem Nutzen und Schaden“ (im Weiteren kurz als „Nutzen und Schaden“ bezeichnet) werden Veränderungen verstanden, die für den Patienten fassbare Konsequenzen haben, wie z. B. die Auswirkung auf die Mortalität, Morbidität und gesundheitsbezogene Lebensqualität. Mit dem Begriff „Nutzen“ werden kausal begründete positive Effekte, mit dem Begriff „Schaden“ kausal begründete negative Effekte einer medizinischen Intervention auf patientenrelevante Endpunkte (siehe 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 [40].

### **Allgemeine Voraussetzungen für Studien zur Bewertung der diagnostischen und prognostischen Güte**

Falls ausreichend Evidenz zum patientenrelevanten Nutzen und Schaden vorliegt, ist eine Nutzenbewertung alleine anhand der Beantwortung dieser Fragestellung möglich. In diesem Fall ist die Nutzenbewertung abgeschlossen. Falls jedoch keine ausreichende Evidenz zur Bewertung des patientenrelevanten Nutzens und Schadens vorliegt, erfolgt eine Bewertung der diagnostischen und der prognostischen Güte der PET bzw. der PET/CT.

Unter „diagnostischer Güte“ wird eine korrekte Zuordnung der Patienten zum jeweiligen Gesundheitszustand verstanden, unter „prognostischer Güte“ eine korrekte Vorhersage der patientenrelevanten Zielgrößen.

### **Indikationen für den Einsatz der PET bzw. PET/CT laut Konkretisierung des G-BA**

Im Folgenden wird die laut Konkretisierung des G-BA festgelegte Indikation 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 diese Indikationen untersucht werden:

- Vitalitätsdiagnostik bei bestehender KHK und eingeschränkter regionaler oder globaler linksventrikulärer Funktion

### **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 PET und PET/CT bei 14 verschiedenen Krankheitsbildern beauftragt. Der vorliegende Berichtsplan bezieht sich auf den Teilauftrag zur Diagnostik von vitalem Myokardgewebe bei Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion.

Die endgültigen Konkretisierungsunterlagen wurden dem Institut mit dem Schreiben des GBA vom 19.06.2008 übersandt. In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Während der Erstellung des Berichtsplans wurden am 18.05.2011 Patientenvertreter der Selbsthilfe-Initiative HFI [Health Force-Initiative] e. V. zur Festlegung patientenrelevanter Zielgrößen konsultiert.

Der vorliegende vorläufige Berichtsplan (Version 1.0) wird zur Anhörung gestellt. Hierzu können schriftlich Stellungnahmen eingereicht werden. Das Ende der Stellungnahmefrist wird auf der Website des IQWiG ([www.iqwig.de](http://www.iqwig.de)) bekannt gegeben. Stellungnahmen können von allen interessierten Personen, Institutionen und Gesellschaften abgegeben werden. Die Stellungnahmen müssen bestimmten formalen Anforderungen genügen, die ebenfalls auf der Website des IQWiG in einem entsprechenden Leitfaden dargelegt sind. Gegebenenfalls wird eine wissenschaftliche Erörterung zur Klärung unklarer Aspekte aus den schriftlichen Stellungnahmen durchgeführt. Die Anhörung kann zu Änderungen und / oder Ergänzungen des Berichtsplans führen. Im Anschluss an diese Anhörung wird der dann gültige Berichtsplan publiziert.

#### 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 [41]. 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 [42,43]. Fryback und Thornbury haben dazu in ihrer Publikation „The efficacy of diagnostic imaging“ von 1991 [42]. 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 Auswirkungen („therapeutic efficacy“) als 4. Stufe und den patientenrelevanten Auswirkungen („patient outcome efficacy“) als 5. Stufe reichen. Als oberste Stufe betrachten sie schließlich die gesellschaftlichen Konsequenzen („societal efficacy“) der diagnostischen Technologie (siehe auch Tabelle 3).

Tabelle 3: Klassifikation für Evaluierungsstudien zu diagnostischen Testverfahren

	Köbberling et al. [43]	Fryback & Thornbury [42]	Evidenzklassifizierung G-BA[44]
Patientenrelevante Zielgrößen	Phase 4: Wirksamkeit	6. Stufe: Auswirkung auf Systemebene	Evidenzstufe I
		5. Stufe: Auswirkung im Hinblick auf patientenrelevante Endpunkte	
Surrogatzielgröße (Managementänderung)		4. Stufe: Auswirkung im Hinblick auf das therapeutische Denken des Behandelnden	Evidenzstufe II
		3. Stufe: Auswirkung im Hinblick auf das (differenzial-) diagnostische Denken („Nachtstwahrscheinlichkeit“)	
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 (§ 11 Abs. 2) dieser Sichtweise angeschlossen [44]. In Tabelle 3 sind verschiedene Klassifikationsschemata für Evaluierungsstudien zu diagnostischen Testverfahren einander gegenübergestellt. Neben der Einteilung von Fryback und Thornbury und der Evidenzklassifizierung des G-BA (Tabelle 6) wurde auch das Phasenschema nach Köbberling et al. [43] mit aufgenommen.

Tabelle 4: Evidenzklassifizierung des G-BA für Unterlagen zu diagnostischen Methoden

<b>Evidenzklasse nach G-BA [44]</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 Erkennung von vitalem Myokardgewebe bei Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion zu bewerten, genügt es nicht festzustellen, inwieweit die PET bzw. PET/CT den gängigen Methoden hinsichtlich der Differenzierung von lebendem (hibernierendem) versus nicht lebendem (fibrosiertem) Myokard 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 z. B. je nach PET-Befund randomisiert entweder der bisherigen, von der PET unabhängigen Standardtherapie oder einer aufgrund des PET-Befundes adaptierten Therapie zugeteilt werden (sogenannte Studien mit Anreicherungsdesign). Aber auch andere Studiendesigns kommen infrage [45]. Wie sich die Behandlungsoptionen verändern und welche Auswirkungen die unterschiedlichen Strategien schließlich auf patientenrelevante Endpunkte haben, lässt sich am besten durch Studien mit ausreichender Nachbeobachtungszeit beantworten (siehe auch Abschnitt 4.1.2.3).

#### **4.1.1 Einschlusskriterien zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT anhand von systematischen Übersichten**

Im Folgenden werden die Begriffe Sekundärliteratur, systematische Reviews oder HTA-Berichte unter dem Begriff „systematische Übersichten“ zusammengefasst.

Vorabrecherchen haben ergeben, dass zur Bewertung des patientenrelevanten Nutzens und Schadens der PET bzw. PET/CT für die Erkennung von vitalem Myokardgewebe bei Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion vermutlich keine qualitativ ausreichende systematische Übersicht vorliegt. Daher soll der patientenrelevante Nutzen in dem zu erstellenden Bericht nicht auf Basis systematischer Übersichten, 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 Populationen**

Studien mit mindestens 10 Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion werden eingeschlossen.

##### **4.1.2.2 Prüf- und Vergleichsinterventionen**

Als Prüfintervention gilt eine diagnostische Strategie mit Anwendung einer Vollring-PET bzw. einer Vollring-PET/CT (Indextest I). Als Vergleichsintervention (Indextest II) wird eine diagnostische Strategie ohne Anwendung der PET bzw. der PET/CT betrachtet. Ein weiterer Vergleich beinhaltet 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 werden betrachtet:

- 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) sind ebenfalls möglich [45].

#### 4.1.2.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevanten Endpunkte verwendet:

- Mortalität (Gesamtmortalität und kardiovaskuläre Mortalität)
- Morbidität, wie z. B.
  - Überleben ohne kardiovaskuläre Ereignisse (Herztod, Myokard-Infarkt, nicht beabsichtigte Revaskularisation)
  - Unerwünschte Ereignisse
  - Symptomschwere (Belastungsluftnot [NYHA Klassifikation])
  - Hospitalisierungsrate (gesamt und kardiovaskulär bedingt)
- gesundheitsbezogene Lebensqualität

Eine Änderung im Patientenmanagement alleine (ohne Verknüpfung zu den oben genannten patientenrelevanten Endpunkten) stellt keine Zielgröße im Sinne dieses Berichts dar und wird daher nur ergänzend betrachtet.

Subjektive Endpunkte (zum Beispiel gesundheitsbezogene Lebensqualität) werden nur dann berücksichtigt, wenn sie mit validierten Messinstrumenten (zum Beispiel SF-36) erfasst worden sind.

#### 4.1.2.4 Studientypen

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

Ist die auf randomisierten kontrollierten Studien basierende Datenlage nicht hinreichend, 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), werden gegebenenfalls auch verfügbare eindeutig prospektiv geplante, jedoch nicht randomisierte Interventionsstudien mit zeitlich parallelen Kontrollgruppen herangezogen (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.



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

Dieses Vorgehen scheint für die zugrunde liegende Fragestellung gerechtfertigt, da sich erst sehr langsam der Paradigmenwechsel in der Bewertung diagnostischer Verfahren hin zu vergleichenden Studien mit patientenrelevanten Endpunkten vollzieht. Zudem erlauben möglicherweise nicht randomisierte kontrollierte Interventionsstudien – wie oben skizziert – Hinweise auf die Ausgestaltung zukünftiger aussagekräftigerer Studien.

#### **4.1.2.5 Studiendauer**

Hinsichtlich der Studiendauer besteht keine Einschränkung.

#### **4.1.2.6 Sonstige Studiencharakteristika**

Es erfolgt keine weitere Einschränkung.

#### **4.1.3 Tabellarische Darstellung der Einschlusskriterien („Nutzenstudien“)**

In die Nutzenbewertung werden alle Studien einbezogen, die alle nachfolgenden Einschlusskriterien erfüllen (siehe Tabelle 5).

Tabelle 5: Übersicht der Einschlusskriterien für Primärstudien zur Nutzenbewertung der PET bzw. PET/CT

<b>Einschlusskriterien</b>	
E1a	Mindestens 10 Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion (siehe auch Abschnitt 4.1.2.1)
E2a	Vergleich von 2 diagnostischen Strategien wie z. B. <ul style="list-style-type: none"> <li>- PET versus Strategie ohne PET</li> <li>- PET/CT versus Strategie ohne PET/CT</li> <li>- PET/CT versus PET</li> </ul> (siehe auch Abschnitt 4.1.2.2)
E3a	Patientenrelevante Endpunkte wie z. B. Mortalität, Morbidität und gesundheitsbezogene Lebensqualität (siehe auch Abschnitt 4.1.2.3)
E4a	Randomisierte kontrollierte Studien (Bei nicht hinreichend auf randomisierten kontrollierten Studien basierender Datenlage werden ggf. prospektive nicht randomisierte kontrollierte Studien [mit und ohne experimentelle Gruppenzuteilung], wie in Abschnitt 4.1.2.4 beschrieben, eingeschlossen.)
E5a	Keine Mehrfachpublikationen ohne relevante Zusatzinformationen
E6a	Vollpublikation 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 CONSORT-Statements [46] oder des TREND-Statements [47] genügt und eine Bewertung der Studie ermöglicht.	

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

In den orientierenden Vorrecherchen fanden sich systematische Übersichten, die die diagnostische Güte der PET oder PET/CT beim Krankheitsbild der koronaren Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion untersuchten. Mit dem Ziel einer möglichst raschen Ergebniserhebung und Nutzung bereits vorliegender wissenschaftlicher Evidenz auf höchster Evidenzstufe soll die Bewertung der diagnostischen Güte der PET und PET/CT primär auf Grundlage von systematischen Übersichten erfolgen, sofern sichergestellt ist, dass sie bestimmte Mindestanforderungen erfüllen. Findet sich keine geeignete systematische Übersicht, werden die diagnostische und die prognostische Güte anhand von Primärstudien beurteilt. Sind die identifizierten systematischen Übersichten nicht mehr aktuell, wird für den Zeitraum, den sie nicht abdecken, relevante Primärliteratur recherchiert und im Bericht berücksichtigt (Ergänzungsrecherche).

Nachfolgend werden zunächst die Kriterien für einen Einschluss von systematischen Übersichten 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 systematischen Übersichten**

### **4.2.1.1 Populationen**

Studien zu mindestens 10 Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion werden eingeschlossen (siehe Abschnitt 4.1.2.1).

### **4.2.1.2 Prüftechnologien, Vergleichstechnologien („Komparatoren“) und Referenztest**

Die den systematischen Übersichten zugrunde liegenden Primärstudien müssen die Angaben zur diagnostischen Güte anhand folgender Vergleiche ermittelt haben:

Prüftechnologie („Indextest I“) ist die Vollring-PET bzw. PET/CT. Als Vergleichstechnologie („Indextest II“) dient jegliche diagnostische Strategie ohne PET bzw. PET/CT („konventionelle Diagnostik“), die nicht Teil des Referenzstandards ist. Außerdem wird der Vergleich PET vs. PET/CT zugelassen (s. Abschnitt 4.1.2.2).

Indextest I und Indextest II müssen 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 ist, da es anderenfalls zu Verzerrungen der Güte- und / oder Variabilitätsschätzer kommen kann (sogenannter Verifikationsbias oder Inkorporationsbias) [48-50].

Mangels Goldstandard für Vitalitätsdiagnostik wird für die Abschätzung der Genauigkeit von PET-Tests im Vergleich zu Tests, die mit Komparatoren durchgeführt wurden, ein Referenzstandard herangezogen. Basierend auf der Annahme, dass viables Myokard durch Revaskularisation wieder zum Funktionieren gebracht wird, erfolgt die Beurteilung der diagnostischen Güte patientenbezogen anhand der segmentalen myokardialen Funktion nach Revaskularisation. Operationalisiert wird die Funktion gemäß der Definition von hibernierendem Myokard durch die Wandbewegung im Vergleich vor und nach dem Eingriff. Viabilitätspositive Befunde werden im Referenzstandard vor dem Eingriff durch einen Zustand ohne Wandbewegung und nach dem Eingriff mit Wandbewegung bestätigt. Viabilitätsnegative Befunde werden entsprechend patientenbezogen durch das Ausbleiben von segmentaler Bewegung auch nach dem Eingriff bestätigt [35,51]. Die Beurteilung der Funktion betreffender Herzsegmente bedarf der Darstellung des Myokardgewebes (nicht der Ventrikel oder Gefäße) und neben der räumlichen Auflösung auch einer präzisen zeitlichen Auflösung („ECG-gated“). Zudem ist eine Differenzierung nach Belastungsstufen erforderlich. Im vorliegenden Bericht werden Studien zur Beurteilung der diagnostischen Güte herangezogen, die die Messgenauigkeit der PET/PET-CT (im Vergleich mit einem zweiten Indextest) mit den derzeit im Sinne der oben geschilderten Gesichtspunkte besten Verfahren überprüfen. Als geeignete Verfahren werden die folgenden drei Verfahren LDDE, MRT und CT angesehen. RNV und SPECT sind gegenüber MRT und LDDE weniger exakt [52-54]. Echographie ohne Belastungstest ist wegen der fehlenden Induktion von Kontraktilität in viablen Segmenten der Stressechokardiographie unterlegen. PET/PET-CT

Verfahren selbst kommen wegen des zu erwartenden Inkorporationsbias‘ nicht für den Referenzstandard in Frage. Gesucht wird demzufolge nach diagnostischen Studien zur PET, in denen als Referenzstandard die zweimalige Durchführung eines der drei Verfahren LDDE, MRT oder CT sowohl vor als auch nach der Revaskularisation jeweils zur Feststellung der segmentalen Wandbewegung eingesetzt wird, in denen aber der zweite Indextest nicht Teil des Referenzstandards ist.

Als mögliche mit der Wahl des Referenzstandards verbundene Limitierungen sind Verzerrungen durch die Intervention (Behandlungs-Paradox), selektive Dropouts, Remodelling und erneute Gefäßverschlüsse zu berücksichtigen.

#### **4.2.1.3 Zielgrößen**

Es werden systematische Übersichten eingeschlossen, deren Ergebnisse auf Primärstudien beruhen, die relevante Angaben zu folgenden Zielgrößen enthalten:

- 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 müssen die patientenbezogenen Vierfeldertafel-Daten zur Berechnung der diagnostischen Güte der Primärstudien ableitbar sein.
- Prognostische Aussagen (Wahrscheinlichkeit) hinsichtlich zukünftiger klinischer Ereignisse (z. B. Herzversagen, Herzrhythmusstörungen) 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 systematischen Übersichten zugrunde liegende Studientypen**

Berücksichtigt werden Ergebnisse systematischer Übersichten, die auf Daten beruhen, die anhand folgender Primärstudientypen erhoben wurden:

- Diagnostische Güte  
Querschnittstudien, ggf. mit klinischer Nachbeobachtung, durchgeführt in der Anwendungssituation (gemäß Phase 3 nach Köbberling et al. [43]).
- Prognostische Güte  
Prospektiv geplante Kohortenstudien: Um die prognostische Güte der PET möglichst unverzerrt bestimmen zu können, soll eine Gruppe von Patienten, die zu einem bestimmten Zeitpunkt rekrutiert und mit der PET untersucht wurde, im zeitlichen Verlauf untersucht werden (Longitudinalstudie). Dabei sind zum einen ein konsekutiver Einschluss der Patienten und die Dokumentation aller fehlenden Werte notwendig. Zum Zweiten sollen die PET-Positiven und die PET-Negativen zeitlich parallel, prospektiv und im direkten Vergleich beobachtet werden (historische Vergleiche oder der Vergleich positiver und negativer Ergebnisse aus unterschiedlichen Patientenkollektiven liefern

keine verlässlichen Daten). Im Idealfall soll die weitere Behandlung dieser Patienten ohne das Wissen um die Ergebnisse der PET erfolgen (Verblindung).

Für die Interpretation der Wandbewegung ist ein Follow-up von mindestens einer Woche erforderlich. Frühere Nachmessungen treffen die betreffenden Herzareale möglicherweise noch im Zustand des *stunning* an, einer Art reversible Starre als Reaktion auf den chirurgischen Eingriff. Die exakte Differenzierung von hibernierenden und „stunned“ Myokardsegmenten ist nicht möglich [24,34,55,56]. Systematische Übersichten, die sowohl prospektive als auch retrospektive Studien enthalten, werden nur dann berücksichtigt, wenn eine separate Auswertung der in ihnen enthaltenen prospektiv geplanten Studien möglich ist.

#### ■ Diskordanzstudien

Ist die Datenlage aus Studien, welche die diagnostischen Vorgehensweisen mit dem Referenztest direkt vergleichen, unzureichend, können systematische Übersichten der o. g. Studientypen herangezogen werden, die lediglich die diskordanten Befunde zwischen den alternativen diagnostischen Vorgehensweisen untersuchen. Diskordanzstudien müssen 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 [57].

#### **4.2.1.5 Sonstige Charakteristika der systematischen Übersichten**

Neben diesen genannten Charakteristika müssen die systematischen Übersichten folgende Kriterien erfüllen:

- Es muss sich um Übersichten handeln, die auf systematische, reproduzierbare und transparente Weise Ergebnisse von Primärstudien zusammenfassend dargestellt und bewertet haben. Daher müssen die Übersichten auf dem Qualitätsindex von Oxman und Guyatt mindestens 5 von 7 möglichen Punkten erreichen [58-60].
- Die jeweilige Forschungsfrage muss im Format einer sog. PICO-Frage formuliert sein, die Population, Indextest I, ggf. Indextest II, Referenztest und Endpunkte explizit benennt, oder die PICO-Fragestellung lässt sich der Publikation eindeutig entnehmen.
- Ein möglicherweise in den Primärstudien zu konstatierender Verifikations- bzw. Inkorporationsbias (vgl. Abschnitt 4.2.1.2) soll adressiert sein.
- Die Ergebnisse der Qualitätsbewertung, die Kerninformationen über die Population und die Ergebnisse einer jeden Studie sollen transparent und vergleichbar abgebildet sein und die Daten sollen eindeutig mit den zugehörigen Studien verknüpft werden können.

#### 4.2.1.6 Tabellarische Darstellung der Einschlusskriterien (systematische Übersichten zur diagnostischen und prognostischen Güte)

Es werden alle systematischen Übersichten einbezogen, die alle nachfolgenden Einschlusskriterien erfüllen (siehe Tabelle 6).

Tabelle 6: Übersicht der Einschlusskriterien für systematische Übersichten zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT

<b>Einschlusskriterien</b> zur Evaluierung der diagnostischen und prognostischen Güte anhand von systematischen Übersichten	
E1b	Mindestens 10 Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion (siehe auch Abschnitt 4.1.2.1)
E2b	Prüftechnologie (Vollring-PET bzw. PET/CT), Vergleichstechnologie und Referenztests, Evaluierung wie in Abschnitt 4.2.1.2 beschrieben
E3b	Zielgrößen wie z. B. Sensitivität / Spezifität oder positiver / negativer Likelihood-Quotient (siehe auch Abschnitt 4.2.1.3)
E4b	Systematische Übersichten 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 Zusatzinformationen
E6b	Vollpublikation 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 PRISMA-Statements [61] oder des MOOSE-Statements [62] genügt und eine Bewertung der Studie ermöglicht.	

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

Im Rahmen einer Ergänzungsrecherche wird in Ergänzung der systematischen Übersichten Primärliteratur identifiziert und anhand folgender Kriterien eingeschlossen:

##### 4.2.2.1 Populationen

Studien zu mindestens 10 Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion werden eingeschlossen (siehe Abschnitt 4.1.2.1).

##### 4.2.2.2 Prüftechnologien (Indextest I), Vergleichstechnologien (Indextest II) und Referenztests

Prüftechnologie ist die Vollring-PET bzw. PET/CT. Vergleichstechnologie, Referenztest und Evaluierung entsprechen den in Abschnitt 4.2.1.2 beschriebenen.

### 4.2.2.3 Zielgrößen

Für die Zielgrößen der Primärstudien aus der Ergänzungsrecherche gelten dieselben Einschlusskriterien wie für die Zielgrößen der Primärstudien in den systematischen Übersichten (siehe Abschnitt 4.2.1.3).

### 4.2.2.4 Studientypen

Für die Primärstudien der Ergänzungsrecherche gelten dieselben Einschlusskriterien hinsichtlich der Studientypen wie für die in den systematischen Übersichten enthaltenen Primärstudien (siehe Abschnitt 4.2.1.4).

### 4.2.2.5 Sonstige Studiencharakteristika

Es erfolgt keine weitere Einschränkung.

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

Tabelle 7: Übersicht der Einschlusskriterien für die Primärliteratur zur Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT (Ergänzungsrecherche)

<b>Einschlusskriterien</b> für Primärstudien zur Evaluierung der diagnostischen und prognostischen Güte	
E1c	Mindestens 10 Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion (siehe auch Abschnitt 4.1.2.1)
E2c	Prüftechnologie (Vollring-PET bzw. PET/CT), Vergleichstechnologie und Referenztests, Evaluierung wie in Abschnitt 4.2.1.2 beschrieben
E3c	Zielgrößen wie z. B. Sensitivität / Spezifität oder positiver / negativer Likelihood-Quotient (siehe auch Abschnitt 4.2.1.3)
E4c	Querschnitt- oder prospektiv geplante Kohortenstudien wie in Abschnitt 4.2.1.4 beschrieben
E5c	Keine Mehrfachpublikationen ohne relevante Zusatzinformationen
E6c	Vollpublikation verfügbar <sup>a</sup>
E7c	Publikationszeitraum durch Literatursuche der eingeschlossenen systematischen Übersichten nicht abgedeckt
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 [63] oder des STROBE-Statements [64] genügt und eine Bewertung der Studie ermöglicht.	

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

Für das Einschlusskriterium E1 (Population) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patienten dieses Kriterium erfüllt ist. Liegen für solche Studien entsprechende Subgruppenanalysen vor, wird auf diese Analysen zurückgegriffen. Studien, bei denen das Einschlusskriterium E1 bei weniger als 80 % erfüllt ist, werden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorliegen.

Ebenfalls eingeschlossen werden Studien, die zu mindestens 80 % das Einschlusskriterium E2 erfüllen (Prüf- und Vergleichsintervention bzw. Indextests). Beim Studieneinschluss werden unvollständige Erhebungen des Referenzstandards toleriert, wenn Patienten keine revaskularisierende Maßnahme erhalten (z. B. wegen akuter kardiovaskulärer Ereignisse) oder diese Maßnahme eine Revaskularisierung nicht erreicht (z. B. bei postinterventionellem Gefäßverschluss). Ein Ausschluss von Patienten, bei denen aufgrund der diagnostischen Befunde (z. B. unklarer PET-Befund) auf eine revaskularisierende Maßnahme verzichtet wurde, ist hierunter jedoch nicht zu verstehen.

### **4.3 Informationsbeschaffung**

#### **4.3.1 Bibliografische Literaturrecherche**

Die systematische Literaturrecherche nach relevanten Publikationen soll in folgenden Quellen durchgeführt werden:

- Suche nach Primärstudien in den bibliografischen Datenbanken MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials (Clinical Trials)
- Suche nach relevanten systematischen Übersichten: 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)

#### **4.3.2 Suche nach weiteren publizierten und nicht publizierten Studien und systematischen Übersichten**

Zusätzlich zur Suche in bibliografischen Datenbanken sollen folgende Quellen zur Identifizierung publizierter und nicht publizierter Studien herangezogen werden:

- systematische Übersichten
- durch den G-BA übermittelte Unterlagen
- öffentlich zugängliche Studienregister
- im Rahmen der Anhörung zum vorläufigen Berichtsplan und zum Vorbericht eingereichte Informationen

#### **4.3.3 Selektion relevanter Studien und systematischer Übersichten**

Die Selektion relevanter Primärstudien und systematischer Übersichten erfolgt durch 2 Reviewer unabhängig voneinander. Dazu wird das Ergebnis der Recherche in den oben genannten Quellen herangezogen.

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



Die Fragestellung 2 zur diagnostischen und prognostischen Güte wird primär auf Basis von systematischen Übersichten beantwortet. Zunächst wird daher gezielt nach systematischen Übersichten gesucht und die gefundenen Arbeiten mit dem Qualitätsindex von Oxman und Guyatt bewertet (vgl. Abschnitt 4.2.1.5). Anschließend wird für den Zeitraum, der durch den Recherchezeitraum der systematischen Übersichten nicht abgedeckt wird, eine zusätzliche Recherche nach Primärliteratur durchgeführt (Ergänzungsrecherche; vgl. Abschnitt 4.2.2). Sollten nicht ausreichend belastbare Daten aus systematischen Übersichten vorliegen, wird sie ausschließlich auf Basis der Primärliteratur beantwortet.

#### **4.4 Informationsbewertung**

Die Dokumentation und Bewertung der Studiencharakteristika und -ergebnisse der relevanten Publikationen 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 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 hängt 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 werden hinsichtlich ihrer Ergebnissicherheit, bestehend aus dem Verzerrungspotenzial und der Präzision der Ergebnisse, überprüft.

##### **Bewertung des Verzerrungspotenzials der Ergebnisse**

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

##### **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 nicht randomisierten kontrollierten Studien)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (bei nicht randomisierten kontrollierten Studien)
- Verblindung des Patienten sowie des Behandlers (bei randomisierten Studien)
- ergebnisgesteuerte Berichterstattung

**B: Aspekte des Verzerrungspotenzials der Ergebnisse auf Endpunktebene**

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

Das Verzerrungspotenzial wird als „niedrig“ oder „hoch“ eingestuft. 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 wird zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft. Falls diese Einstufung als „hoch“ erfolgt, wird das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet. Ansonsten finden die unter B genannten endpunktspezifischen Aspekte Berücksichtigung.

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

**4.4.2 Bewertung von systematischen Übersichten zur diagnostischen und prognostischen Güte**

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

Die Bewertung des Verzerrungspotenzials einer systematischen Übersicht erfolgt anhand der Kriterien von Oxman und Guyatt [58,59]. Diese Kriterien werden um die Angabe von Interessenkonflikten ergänzt. Die Bewertung erfolgt immer für die gesamte systematische Übersicht und ist daher für mehrere Fragestellungen bzw. Indikationen einer Übersicht immer gleich.

**4.4.3 Bewertung von Primärstudien zur diagnostischen und prognostischen Güte**

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

Die Bewertung des Verzerrungspotenzials der Primärstudien zur diagnostischen Güte wird mithilfe des QUADAS2-Instruments durchgeführt [65].

Für die Bewertung des Verzerrungspotenzials von prognostischen Studien wird ein Instrument eingesetzt, das auf Grundlage der Arbeiten von Hayden et al. [66] und Altman et al. [67] entwickelt wurde [68].

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

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

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

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

Bei der Betrachtung patientenrelevanter Endpunkte, die mithilfe von (komplexen) Skalen operationalisiert werden, ist es in besonderer Weise notwendig, neben der statistischen Signifikanz der Effekte die Relevanz der beobachteten Wirkungen der untersuchten Interventionen zu bewerten, da die Komplexität der Skalen häufig eine sinnvolle Interpretation geringfügiger Unterschiede erschwert. Die Relevanzbewertung kann auf Basis von Mittelwertdifferenzen und Responderanalysen vorgenommen werden. Welches dieser Verfahren gewählt werden kann, hängt insbesondere von der Verfügbarkeit der Daten aus den Primärstudien ab.

Um skalen- und indikationsspezifischen Besonderheiten gerecht zu werden, soll die Relevanzbewertung primär auf Basis validierter (bzw. etablierter) Relevanzkriterien für die jeweilige Skala durchgeführt werden. Für die Betrachtung von Mittelwertdifferenzen ist dies z. B. eine validierte bzw. etablierte Irrelevanzschwelle für den Gruppenunterschied: Liegt das zum beobachteten Effekt korrespondierende Konfidenzintervall vollständig oberhalb dieser Irrelevanzschwelle, wird davon ausgegangen, dass die Effektstärke nicht in einem sicher irrelevanten Bereich liegt. Bei der Betrachtung von Responderanalysen ist es erforderlich, dass bei diesen Analysen ein validiertes bzw. etabliertes Responsekriterium angewendet wurde (im Sinne einer individuellen Minimal Important Difference [MID]). Liegt bei einer solchen Auswertung ein statistisch signifikanter Unterschied der Anteile an Respondern zwischen den Gruppen vor, wird dies als relevanter Effekt angesehen, da die Responserdefinition bereits eine Schwelle der Irrelevanz (nämlich die MID) beinhaltet.

Falls skalenspezifisch validierte bzw. etablierte Relevanzkriterien nicht vorliegen, muss auf ein allgemeines statistisches Maß zur Relevanzbewertung zurückgegriffen werden. In diesem Fall werden standardisierte Mittelwertdifferenzen (SMD in Form von Hedges'  $g$ ) betrachtet. Als Irrelevanzschwelle wird dann 0,2 verwendet: Liegt das zum beobachteten Effekt korrespondierende Konfidenzintervall vollständig oberhalb dieser Irrelevanzschwelle, wird davon ausgegangen, dass die Effektstärke nicht in einem sicher irrelevanten Bereich liegt. Dies soll gewährleisten, dass der beobachtete Effekt hinreichend sicher mindestens als „klein“ angesehen werden kann [69,70].

#### **4.5.1 Gegenüberstellung der Ergebnisse der Einzelstudien**

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

In bestimmten Fällen werden einzelne Ergebnisse aus den Studien zu einem Endpunkt nicht dargestellt bzw. nicht in die Nutzenbewertung oder Bewertung der diagnostischen und prognostischen Güte einbezogen. Dies trifft insbesondere zu, wenn viele Patienten nicht in der Auswertung enthalten sind. Ergebnisse fließen in der Regel nicht in die Bewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patienten basieren, das heißt, wenn der Anteil der Patienten ohne jegliche Berücksichtigung in der Auswertung (Nichtberücksichtigungsanteil) größer als 30 % ist. In der Literatur werden zum Teil bereits Nichtberücksichtigungsanteile größer als 20 Prozentpunkte als nicht mehr aussagekräftig betrachtet [71].

Ausnahmen von dieser Regel können zum Beispiel dann gemacht werden, wenn aus logistischen Gründen für ganze Zentren (ganze Randomisierungsblöcke) keine Daten erhoben wurden und dies bereits bei der Studienplanung vorgesehen war [72].

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

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

Im Bericht werden die Studien anhand des Designs (Studiendesign, Studiendauer, Zahl der randomisierten / zuteilten 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, eventuelle Effekte der Vorselektion der Patienten etc.) dargestellt. Die Studienpopulationen werden weiterhin durch demografische Daten (Alter, Geschlecht), durch Charakteristika der betrachteten Krankheitsentität und durch prognostische Angaben beschrieben. Zusätzlich wird die Zahl der Studienabbrecher 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 Endpunkten (siehe 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.

#### **4.5.1.2 Gegenüberstellung der Ergebnisse von systematischen Übersichten zur diagnostischen und prognostischen Güte**

Bei der Beurteilung der diagnostischen und prognostischen Güte werden zunächst die zu der Fragestellung vorhandenen Ergebnisse der systematischen Übersichten abgebildet. Die Ergebnisse aus den zugrunde liegenden Primärstudien werden 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 erfolgt eine Abschätzung der Heterogenität der Ergebnisse innerhalb dieser Klassen. Beim Vorliegen inkonsistenter Ergebnisse zu einer Fragestellung werden mögliche Erklärungen für die beobachtete Heterogenität (z. B. Unterschiede bei der Berücksichtigung eines möglichen Verifikations- oder Inkorporationsbias) beschrieben [73].

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

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

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

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

#### **4.5.1.4 Gegenüberstellung der Ergebnisse aus systematischen Übersichten und der Ergänzungsrecherche („Robustheitsprüfung“)**

Finden sich sowohl relevante systematische Übersichten als auch relevante Primärstudien in der Ergänzungsrecherche, wird geprüft, ob die (quantitativ zusammengefassten) Ergebnisse der aktuelleren Primärstudien die Ergebnisse der systematischen Übersichten infrage stellen könnten („Robustheitsprüfung“). Nur wenn dies zu erwarten ist, werden 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**

Sofern die Studien hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbar sind, werden die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst. Für die statistische Auswertung vergleichender Studien werden primär die Ergebnisse aus Intention-to-Treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet. Die Meta-Analysen erfolgen in der Regel auf Basis von Modellen mit zufälligen Effekten [56]. In begründeten Ausnahmefällen werden Modelle mit festen Effekten eingesetzt. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, werden diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt.

#### **Studien zum patientenrelevanten Nutzen**

Für kontinuierliche Variablen wird die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges'  $g$ , als Effektmaß eingesetzt. Bei binären Variablen werden Meta-Analysen primär anhand des Odds Ratio und beim Vergleich von Überlebenszeiten anhand des Hazard Ratio durchgeführt. Bei kategorialen Variablen wird ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet [74].

Die Effektschätzer und Konfidenzintervalle aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Anschließend erfolgt die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes  $I^2$  und des statistischen Tests auf Vorliegen von Heterogenität [75]. Ist die Heterogenität der Studienergebnisse nicht bedeutsam ( $p > 0,2$  für Heterogenitätstest), wird der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt. Bei bedeutsamer Heterogenität werden die Ergebnisse nur in begründeten Ausnahmefällen gepoolt. Außerdem wird untersucht, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.5.3) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.5.4).

#### **Prognosestudien**

Die Meta-Analyse prognostischer Studien entspricht dem Vorgehen bei der Bewertung des patientenrelevanten Nutzen.

## Diagnosestudien

Die Effektschätzer und Konfidenzintervalle aus den Studien werden ebenfalls mittels Forest Plots zusammenfassend dargestellt. Außerdem wird, sofern die dafür nötigen Anforderungen als erfüllt gelten können, für die Testgütekriterien eine Meta-Analyse anhand der Sensitivität und Spezifität in einer bivariaten Analyse durchgeführt [76]. Die Schätzung der Modellparameter (Sensitivität, Spezifität und Parameter für die „hierarchical summary receiver operating curve“ [HSROC]) erfolgt über einen Maximum-Likelihood-Ansatz in einem nicht linearen gemischten Modell bzw. in einem generalisierten linearen gemischten Modell [77]. In begründeten Ausnahmefällen kommen auch andere Effektmaße zum Einsatz.

Zusätzlich werden bei diagnostischen Studien die beobachteten Paare aus Sensitivität und Spezifität zweidimensional grafisch dargestellt. Studien, in denen Technologien wie PET und MRT direkt verglichen wurden, werden gekennzeichnet. Des Weiteren werden die aus der bivariaten Meta-Analyse gewonnenen Schätzer für die Erwartungswerte als gepoolte Paare der Sensitivität und der Spezifität mit den dazugehörigen 95 %-Konfidenzregionen grafisch dargestellt. Schließlich werden die geschätzten HSROC-Kurven abgebildet. Die grafischen Darstellungen von Sensitivität und Spezifität und der HSROC-Kurve können Tendenzen zu einer besseren Testgüte in den vorliegenden Studien aufzeigen. Schneiden sich die Konfidenzregionen zweier Technologien nicht, so wird davon ausgegangen, dass in den vorliegenden Studien ein Unterschied in der diagnostischen Güte besteht. Die Richtung des Unterschieds wird ggf. grafisch bestimmt. Überlappen sich zwei Konfidenzregionen oder schneiden sich zwei HSROC-Kurven im Bereich der Grafik, in dem auch Paare aus Sensitivität und Spezifität beobachtet wurden, kann in den vorliegenden Studien kein signifikanter Unterschied in der diagnostischen Güte nachgewiesen werden. Das Vorliegen von Heterogenität wird untersucht anhand von Sensitivitätsanalysen und Subgruppenanalysen.

### 4.5.3 Sensitivitätsanalyse

Zur Einschätzung der Robustheit der Ergebnisse sind Sensitivitätsanalysen hinsichtlich methodischer Faktoren geplant. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ wird für Sensitivitätsanalysen verwendet. Dies betrifft sowohl Primärstudien zum patientenrelevanten Nutzen und Schaden als auch jene zur diagnostischen und prognostischen Güte. Bei Einschluss von mehreren systematischen Übersichten kann eine Sensitivitätsanalyse anhand der Beurteilung mit dem Index von Oxman und Guyatt [58,59] durchgeführt werden (Stufe 5 versus 6 versus 7; alle < 5 werden ausgeschlossen).

Der Algorithmus zum Schätzen der Parameter im bivariaten Modell zur diagnostischen Güte kann ggf. keinen Schätzer liefern, da das Maximum-Likelihood-Verfahren nicht konvergiert, oder zu unrealistischen Schätzern führen, d. h. zu Schätzern mit zu großen Standardfehlern und entsprechenden Konfidenzintervallen. Die Gründe hierfür können beispielsweise darin bestehen, dass zu wenige Studien vorliegen, als dass die Parameter geschätzt werden könnten.

Aber auch einzelne Studien mit extremen Werten können die Ursache sein. Um diese Studien zu identifizieren, werden Sensitivitätsanalysen durchgeführt.

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

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

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, d. h. klinischer Faktoren, die die Effekte beeinflussen, untersucht. Dies können direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen sein. Im Gegensatz zu den in Abschnitt 4.5.3 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- bzw. Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung von Effektmodifikatoren werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Es ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Geschlecht
- Alter
- Ausmaß der linksventrikulären Ejektionsfraktion (EF)

Sollten sich aus den verfügbaren Informationen Anhaltspunkte für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden.

Bei Identifizierung möglicher Effektmodifikatoren erfolgt gegebenenfalls eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen. Beispielsweise kann der Beleg eines Zusatznutzens auf eine spezielle Subgruppe von Patienten eingeschränkt werden.



## 5 Literaturverzeichnis

1. Donner-Banzhoff N, Held K, Laufs U, Trappe HJ, Werdan K, Zerkowski HR. Nationale VersorgungsLeitlinie Chronische KHK: Langfassung; Version 1.11 [online]. 03.2011 [Zugriff: 23.11.2011]. URL: [http://www.versorgungsleitlinien.de/themen/khk/pdf/nvl\\_khk\\_lang.pdf](http://www.versorgungsleitlinien.de/themen/khk/pdf/nvl_khk_lang.pdf).
2. Dietz R, Rauch B. Leitlinie zur Diagnose und Behandlung der chronischen koronaren Herzerkrankung der Deutschen Gesellschaft für Kardiologie - Herz- und Kreislaufforschung (DGK). Z Kardiol 2003; 92(6): 501-521.
3. Pfisterer ME, Battler A, Zaret BL. Range of normal values for left and right ventricular ejection fraction at rest and during exercise assessed by radionuclide angiocardiology. Eur Heart J 1985; 6(8): 647-655.
4. Beanlands RS, Ruddy TD, DeKemp RA, Iwanochko RM, Coates G, Freeman M et al. Positron emission tomography and recovery following revascularization (PARR-1): the importance of scar and the development of a prediction rule for the degree of recovery of left ventricular function. J Am Coll Cardiol 2002; 40(10): 1735-1743.
5. Dickstein K, Cohen-Solal A, Filippatos G, McMurray JJ, Ponikowski P, Poole-Wilson PA et al. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2008. Eur Heart J 2008; 29(19): 2388-2442.
6. Dussault FP, Nguyen VH, Rachet F. Positron emission tomography in Quebec. Montreal: Agence d'Evaluation des Technologies et des Modes d'Intervention en Sante; 2001. URL: <http://www.aetmis.gouv.qc.ca/site/download.php?f=65e3c486983f01767e90e114f7ff1e23>.
7. Wijns W, Kolh P, Danchin N, Di Mario C, Falk V, Folliguet T et al. Guidelines on myocardial revascularization. Eur Heart J 2010; 31(20): 2501-2555.
8. Velazquez EJ, Lee KL, O'Connor CM, Oh JK, Bonow RO, Pohost GM et al. The rationale and design of the Surgical Treatment for Ischemic Heart Failure (STICH) trial. J Thorac Cardiovasc Surg 2007; 134(6): 1540-1547.
9. Robert Koch-Institut. Daten und Fakten: Ergebnisse der Studie »Gesundheit in Deutschland aktuell 2009«. Berlin: RKI; 2011. URL: [http://www.gbe-bund.de/gbe10/abrechnung.prc\\_abr\\_test\\_logon?p\\_uid=&p\\_aid=&p\\_knoten=FID&p\\_sprache=D&p\\_suchstring=13126](http://www.gbe-bund.de/gbe10/abrechnung.prc_abr_test_logon?p_uid=&p_aid=&p_knoten=FID&p_sprache=D&p_suchstring=13126).
10. Statistisches Bundesamt. Gesundheit: Todesursachen in Deutschland 2010 [online]. 23.09.2011 [Zugriff: 23.11.2011]. URL: <http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Publikationen/Fachveroeffentlichungen/Gesundheit/Todesursachen/Todesursachen2120400107004,property=file.pdf>.

11. Ford ES, Ajani UA, Croft JB, Critchley JA, Labarthe DR, Kottke TE et al. Explaining the decrease in U.S. deaths from coronary disease, 1980-2000. *N Engl J Med* 2007; 356(23): 2388-2398.
12. Löwel L, Meisinger C, Heier M, Hörmann A, Von Scheidt W. Herzinfarkt und koronare Sterblichkeit in Süddeutschland: Ergebnisse des bevölkerungsbasierten MONICA/KORA-Herzinfarktregisters 1991 bis 1993 und 2001 bis 2003. *Dtsch Arztebl* 2006; 103(10): A616-A622.
13. Heran BS, Chen JM, Ebrahim S, Moxham T, Oldridge N, Rees K et al. Exercise-based cardiac rehabilitation for coronary heart disease. *Cochrane Database Syst Rev* 2011; (7): CD001800.
14. Meyer J, Breithardt G, Erbel R, Erdmann E, Gohlke H, Hanrath P et al. Leitlinie: Koronare Herzkrankheit / Angina pectoris [online]. [Zugriff: 23.11.2011]. URL: [http://www.medizin.uni-essen.de/cardio/leitlinien/leitlinie\\_khk.pdf](http://www.medizin.uni-essen.de/cardio/leitlinien/leitlinie_khk.pdf).
15. Huxley RR, Woodward M. Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. *Lancet* 2011; 378(9799): 1297-1305.
16. Campeau L. Grading of angina pectoris. *Circulation* 1976; 54(3): 522-523.
17. Hoppe UC, Böhm M, Dietz R, Hanrath P, Kroemer HK, Osterspey A et al. Leitlinien zur Therapie der chronischen Herzinsuffizienz [online]. 2004 [Zugriff: 23.11.2011]. URL: [http://leitlinien.dgk.org/images/pdf/leitlinien\\_volltext/2001-04\\_herzinsuffizienz.pdf](http://leitlinien.dgk.org/images/pdf/leitlinien_volltext/2001-04_herzinsuffizienz.pdf).
18. Libby P, Theroux P. Pathophysiology of coronary artery disease. *Circulation* 2005; 111(25): 3481-3488.
19. Bonzel T, Erbel R, Hamm CW, Levenson B, Neumann FJ, Rupprecht HJ et al. Perkutane Koronarinterventionen (PCI). *Clin Res Cardiol* 2008; 97(8): 513-547.
20. Muhlbaier LH, Pryor DB, Rankin JS, Smith LR, Mark DB, Jones RH et al. Observational comparison of event-free survival with medical and surgical therapy in patients with coronary artery disease: 20 years of follow-up. *Circulation* 1992; 86(5 Suppl): II198-II204.
21. Allman KC, Shaw LJ, Hachamovitch R, Udelson JE. Myocardial viability testing and impact of revascularization on prognosis in patients with coronary artery disease and left ventricular dysfunction: a meta-analysis. *J Am Coll Cardiol* 2002; 39(7): 1151-1158.
22. Camici PG, Prasad SK, Rimoldi OE. Stunning, hibernation, and assessment of myocardial viability. *Circulation* 2008; 117(1): 103-114.

23. Ammirati E, Rimoldi OE, Camici PG. Is there evidence supporting coronary revascularization in patients with left ventricular systolic dysfunction? *Circ J* 2010; 75(1): 3-10.
24. Schinkel AF, Bax JJ, Poldermans D, Elhendy A, Ferrari R, Rahimtoola SH. Hibernating myocardium: diagnosis and patient outcomes. *Curr Probl Cardiol* 2007; 32(7): 375-410.
25. Beller GA. Noninvasive assessment of myocardial viability. *N Engl J Med* 2000; 343(20): 1488-1490.
26. Eagle KA, Guyton RA, Davidoff R, Edwards FH, Ewy GA, Gardner TJ et al. ACC/AHA 2004 guideline update for coronary artery bypass graft surgery: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1999 Guidelines for Coronary Artery Bypass Graft Surgery). *Circulation* 2004; 110(14): e340-e437.
27. Bax JJ, Van Eck-Smit BL, Van der Wall EE. Assessment of tissue viability: clinical demand and problems. *Eur Heart J* 1998; 19(6): 847-858.
28. Kim RJ, Wu E, Rafael A, Chen EL, Parker MA, Simonetti O et al. The use of contrast-enhanced magnetic resonance imaging to identify reversible myocardial dysfunction. *N Engl J Med* 2000; 343(20): 1445-1453.
29. Schinkel AF, Poldermans D, Elhendy A, Bax JJ. Assessment of myocardial viability in patients with heart failure. *J Nucl Med* 2007; 48(7): 1135-1146.
30. Schinkel AF, Bax JJ, Delgado V, Poldermans D, Rahimtoola SH. Clinical relevance of hibernating myocardium in ischemic left ventricular dysfunction. *Am J Med* 2010; 123(11): 978-986.
31. Medical Advisory Secretariat. Magnetic resonance imaging (MRI) for the assessment of myocardial viability: an evidence-based analysis. *Ont Health Technol Assess Ser* 2010; 10(15): 1-45.
32. Glaveckaite S, Valeviciene N, Laucevicius A, Celutkiene J, Rudys A, Tamosiunas A. Cardiovascular magnetic resonance imaging for detection of myocardial viability in chronic ischemic left ventricular dysfunction. *Medicina (Mex)* 2009; 45(8): 585-599.
33. Beanlands R, Thorn S, DaSilva J, Ruddy T, Maddahi J. Myocardial viability. In: Wahl R (Ed). *Principles and practices of positron emission tomography*. Philadelphia: Lippincott Williams and Wilkins; 2008. S. 334-350.
34. Dilsizian V, Bacharach SL, Beanlands RS, Bergmann SR, Delbeke D, Gropler RJ et al. ASNC imaging guidelines for nuclear cardiology procedures: PET myocardial perfusion and metabolism clinical imaging [online]. 2009 [Zugriff: 23.11.2011]. URL: <http://www.asnc.org/imageuploads/ImagingGuidelinesPETJuly2009.pdf>.

35. Bax JJ, Poldermans D, Elhendy A, Boersma E, Rahimtoola SH. Sensitivity, specificity and predictive accuracies of various noninvasive techniques for detecting hibernating myocardium. *Curr Probl Cardiol* 2001; 26(2): 147-181.
36. Ghosh N, Rimoldi OE, Beanlands RS, Camici PG. Assessment of myocardial ischaemia and viability: role of positron emission tomography. *Eur Heart J* 2010; 31(24): 2984-2995.
37. Kotzerke J, Oehme L, Lindner O, Hellwig D. Positronenemissionstomographie 2008 in Deutschland: Ergebnisse der Erhebung und Standortbestimmung. *Nuklearmedizin* 2010; 49(2): 58-64.
38. Knapp WH. Leitlinie zur Myokard-Perfusions-Szintigraphie [online]. [Zugriff: 23.11.2011]. URL: [http://www.nuklearmedizin.de/leistungen/leitlinien/html/myo\\_perf\\_szin.php?navId=53](http://www.nuklearmedizin.de/leistungen/leitlinien/html/myo_perf_szin.php?navId=53).
39. Flotats A, Knuuti J, Gutberlet M, Marcassa C, Bengel FM, Kaufmann PA et al. Hybrid cardiac imaging: SPECT/CT and PET/CT; a joint position statement by the European Association of Nuclear Medicine (EANM), the European Society of Cardiac Radiology (ESCR) and the European Council of Nuclear Cardiology (ECNC). *Eur J Nucl Med Mol Imaging* 2011; 38(1): 201-212.
40. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 4.0 [online]. 23.09.2011 [Zugriff: 23.11.2011]. URL: [https://www.iqwig.de/download/IQWiG\\_Methoden\\_Version\\_4\\_0.pdf](https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf).
41. How to read clinical journals: II. To learn about a diagnostic test. *Can Med Assoc J* 1981; 124(6): 703-710.
42. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991; 11(2): 88-94.
43. Köbberling J, Trampisch HJ, Windeler J. Memorandum for the evaluation of diagnostic measures. *J Clin Chem Clin Biochem* 1990; 28(12): 873-879.
44. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. 17.12.2009 [Zugriff: 18.03.2011]. URL: [http://www.g-ba.de/downloads/62-492-422/VerfO\\_2009-12-17.pdf](http://www.g-ba.de/downloads/62-492-422/VerfO_2009-12-17.pdf).
45. Lijmer JG, Bossuyt PMM. Various randomized designs can be used to evaluate medical tests. *J Clin Epidemiol* 2009; 62(4): 364-373.
46. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c869.

47. 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.
48. 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.
49. 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.
50. 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.
51. Rahimtoola SH. The hibernating myocardium. *Am Heart J* 1989; 117(1): 211-221.
52. Klocke FJ, Baird MG, Lorell BH, Bateman TM, Messer JV, Berman DS et al. ACC/AHA/ASNC guidelines for the clinical use of cardiac radionuclide imaging: executive summary; a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (ACC/AHA/ASNC Committee to Revise the 1995 Guidelines for the Clinical Use of Cardiac Radionuclide Imaging). *Circulation* 2003; 108(11): 1404-1418.
53. Hesse B, Lindhardt TB, Acampa W, Anagnostopoulos C, Ballinger J, Bax JJ et al. EANM/ESC guidelines for radionuclide imaging of cardiac function. *Eur J Nucl Med Mol Imaging* 2008; 35(4): 851-885.
54. Wu E, Judd RM, Vargas JD, Klocke FJ, Bonow RO, Kim RJ. Visualisation of presence, location, and transmural extent of healed Q-wave and non-Q-wave myocardial infarction. *Lancet* 2001; 357(9249): 21-28.
55. Braunwald E, Kloner RA. The stunned myocardium: prolonged, postischemic ventricular dysfunction. *Circulation* 1982; 66(6): 1146-1149.
56. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7(3): 177-188.
57. Pepe MS, Alonzo TA. Comparing disease screening tests when true disease status is ascertained only for screen positives. *Biostatistics* 2001; 2(3): 249-260.
58. 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.

59. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991; 44(11): 1271-1278.
60. Jadad AR, Murray WE. *Randomized controlled trials: questions, answers and musings*. Malden: Blackwell; 2007.
61. 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.
62. 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.
63. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. *Ann Intern Med* 2003; 138(1): 40-44.
64. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.
65. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155(8): 529-536.
66. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med* 2006; 144(6): 427-437.
67. Altman DG. Systematic reviews of evaluations of prognostic variables. *BMJ* 2001; 323(7306): 224-228.
68. Wolff R, Westwood M, Scheibler F, Schröer-Günther M, Janßen I, Kleijnen J. Assessment of risk of bias in prognostic studies. *Cochrane Database Syst Rev* 2010; (Suppl): 23.
69. Cohen J (Ed). *Statistical power analysis for the behavioral sciences*. New York: Academic Press; 1977.
70. Fayers PM, Hays R (Ed). *Assessing quality of life in clinical trials: methods and practice*. Oxford: Oxford University Press; 2005.
71. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; 359(9308): 781-785.
72. 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.

73. Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *CMAJ* 1997; 156(10): 1411-1416.

74. Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

75. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414): 557-560.

76. 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.

77. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Analysing and presenting results [online]. In: Deeks JJ, Bossuyt PM, Gatsonis C (Ed). *Cochrane handbook for systematic reviews of diagnostic test accuracy: version 1.0*. 23.12.2010 [Zugriff: 10.11.2011]. URL: <http://srdta.cochrane.org/sites/srdta.cochrane.org/files/uploads/Chapter%2010%20-%20Version%201.0.pdf>.