

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

**Dokumentation und Würdigung der
Anhörung zum vorläufigen Berichtsplan**

Auftrag: D06-01L
Version: 1.0
Stand: 07.09.2012

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:

21.12.2006

Interne Auftragsnummer:

D06-01L

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)
Im Mediapark 8 (KölnTurm)
50670 Köln

Tel.: +49 (0)221 – 35685-0

Fax: +49 (0)221 – 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

Inhaltsverzeichnis

	Seite
Abkürzungsverzeichnis.....	iv
1 Dokumentation der Anhörung	1
2 Würdigung der Anhörung	2
2.1 Anmerkungen zum Impressum	2
2.2 Anmerkungen zum Kapitel „Hintergrund“ des Berichtsplans	2
2.3 Fragestellung des Berichts.....	3
2.3.1 Vergleich PET und PET/CT.....	3
2.3.2 Nutzen versus diagnostische Güte.....	3
2.4 Methoden	4
2.4.1 Patientenpopulation	4
2.4.2 Geeigneter Referenzstandard.....	4
2.4.3 Informationssynthese.....	6
2.5 Literatur.....	7
3 Offenlegung potenzieller Interessenkonflikte	8
3.1 Potenzielle Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen und Firmen	8
3.2 Potenzielle Interessenkonflikte von weiteren Teilnehmern der wissenschaftlichen Erörterung (externe Sachverständige)	8
4 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll.....	10
4.1 Teilnehmerliste der wissenschaftlichen Erörterung.....	10
4.2 Liste der Stellungnahmen, zu denen kein Vertreter an der wissenschaftlichen Erörterung teilgenommen hat.....	10
4.3 Tagesordnung der wissenschaftlichen Erörterung	11
4.4 Protokoll der wissenschaftlichen Erörterung.....	11
4.4.1 Begrüßung und Einleitung.....	11
4.4.2 Tagesordnungspunkt 1: Adäquater Referenzstandard (nur MRT ausreichend?).....	13
4.4.3 Tagesordnungspunkt 2: Sind diagnostische Studien überhaupt sinnvoll / interpretierbar (Treatment paradox)?	26

4.4.4	Tagesordnungspunkt 3: Ausschluss von (vergleichenden) Studien mit Patienten mit nur geringfügig eingeschränkter regionaler und globaler linksventrikulärer Funktion?	38
4.4.5	Tagesordnungspunkt 4: Verschiedenes	40
Anhang – Dokumentation der Stellungnahmen		44

Abkürzungsverzeichnis

Abkürzung	Bedeutung
CT	Computertomographie
DGN	Deutsche Gesellschaft für Nuklearmedizin
DKG	Deutsche Krankenhausgesellschaft
Echo	Echokardiographie
EF	Ejektionsfraktion
G-BA	Gemeinsamer Bundesausschuss
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
LDDE	Low-dose dobutamine stress echocardiography (Stress-Echokardiographie)
LV	linksventrikulär
MRT	Magnetresonanztomographie
PET	Positronenemissionstomographie
NVL KHK	Nationale Versorgungsleitlinie Chronische Koronare Herzerkrankung
PET/CT	Integration von PET (Positronenemissionstomographie) und CT (Computertomographie) in einem Untersuchungsgerät
RNV	Radionuklidventrikulographie
SPECT	Einzelphotonenemissions-Computertomographie

1 Dokumentation der Anhörung

Am 13.12.2011 wurde der vorläufige Berichtsplan in der Version Nr. 1.0 vom 05.12.2011 veröffentlicht und zur Anhörung gestellt. Bis zum 11.01.2012 konnten schriftliche Stellungnahmen eingereicht werden. Insgesamt wurden 2 Stellungnahmen form- und fristgerecht abgegeben. Diese Stellungnahmen sind im Anhang abgebildet.

Unklare Aspekte in den schriftlichen Stellungnahmen wurden in einer wissenschaftlichen Erörterung am 21.02.2012 im IQWiG diskutiert. Das Wortprotokoll der Erörterung befindet sich in Kapitel 4.

Eine Würdigung der in der Anhörung vorgebrachten Aspekte befindet sich in Kapitel 2. Im überarbeiteten Berichtsplan sind darüber hinaus Änderungen, die sich durch die Anhörung ergeben haben, zusammenfassend dargestellt. Der überarbeitete Berichtsplan ist auf der Website des IQWiG unter www.iqwig.de veröffentlicht.

2 Würdigung der Anhörung

Die im Rahmen der Anhörung vorgebrachten Aspekte wurden hinsichtlich valider wissenschaftlicher Argumente für eine Änderung des Berichtsplans überprüft. Die wesentlichen Argumente werden im Folgenden diskutiert. Neben projektspezifischen wissenschaftlichen Aspekten wurden auch übergeordnete Punkte, z. B. zu rechtlichen Vorgaben für das Institut, angesprochen. Auf solche Punkte wird im Rahmen dieser projektspezifischen Würdigung der Anhörung nicht weiter eingegangen.

2.1 Anmerkungen zum Impressum

Die DGN weist darauf hin, dass das Schlagwort „Myocardial Stunning“ klinisch nicht der Fragestellung des vorliegenden Berichts entspricht. Vielmehr empfiehlt sie, den Begriff „Hibernating Myocardium“ als Schlagwort anzugeben.

Die Schlagwörter auf Seite ii werden abgebildet, um diese später für die Berichtsdistribution in bibliografischen Datenbanken zu verwenden. Hierfür verwenden wir die MeSH Terms der Datenbank Pubmed sowie die deutsche Übersetzung von DIMDI (http://www.dimdi.de/static/de/klassi/mesh_umls/mesh/). Die Begriffe „Hibernating myocardium“ und „Myocardial viability“ werden nicht als MeSH Terms in Pubmed geführt. „Heart failure“ ist zwar ein MeSH Term, erscheint jedoch zu unspezifisch für den vorliegenden Bericht. Deshalb wurde der MeSH Term „Myocardial Stunning“ gewählt, der die Fragestellung noch am besten abbildet. Unter diesem MeSH Term werden im Übrigen auch die beiden Entry Terms „Hibernation, Myocardial“ und „Myocardial Hibernation“ geführt.

Eine Änderung der Schlagwörter wird daher nicht vorgenommen.

2.2 Anmerkungen zum Kapitel „Hintergrund“ des Berichtsplans

Beide stellungnehmenden Organisationen machen in ihren Stellungnahmen Vorschläge zur Präzisierung des Abschnitts 1.8 (Empfehlungen wissenschaftlicher Fachgesellschaften zur Vitalitätsdiagnostik mittels PET). Insbesondere wird darauf hingewiesen, dass in der nationalen Versorgungsleitlinie Chronische Koronare Herzerkrankung (NVL KHK) von 2011 und einer aktuell in Überarbeitung befindlichen S1-Leitlinie der DGN die PET zur Evaluierung von Vitalität in dysfunktionalem Myokard empfohlen wird.

Die Vorschläge wurden detailliert geprüft und der Abschnitt grundlegend überarbeitet. Einzig dem Wunsch, die aktuelle Leitlinie (S1) der Deutschen Gesellschaft für Nuklearmedizin zu berücksichtigen, konnte nicht Rechnung getragen werden, da diese zum Zeitpunkt der Fertigstellung des Berichtsplans noch nicht öffentlich verfügbar war. Sie wird jedoch in den weiteren Berichtsdocumenten Berücksichtigung finden, sobald sie publiziert ist.

2.3 Fragestellung des Berichts

2.3.1 Vergleich PET und PET/CT

Die DGN weist in ihrer Stellungnahme darauf hin, dass der Vergleich der PET mit der PET/CT inhaltlich nicht sinnvoll und praktisch nicht nützlich sei. Sie schlägt daher vor, die Studien unterschiedlicher Geräteklassen zu poolen. Auf der anderen Seite weist die DGN darauf hin, dass die neueren Geräte „dramatisch leistungsfähiger hinsichtlich Empfindlichkeit [...], und insbesondere Volumenauflösung [...]“ geworden seien. „Die technische Entwicklung und zeitgemäße Durchführung der Untersuchung sollten daher im Nutzenbericht hinreichend Berücksichtigung finden.“

Der Vergleich von PET und PET/CT wird in der Konkretisierung zum vorliegenden Projekt ausdrücklich gefordert. Wir teilen die Auffassung, dass beim Vergleich der PET bzw. PET/CT mit den konventionellen Verfahren Studien zu diesen beiden Geräteklassen zunächst gepoolt werden können. Sofern dies aber möglich ist, werden zusätzlich auch die Ergebnisse für PET und PET/CT gesondert dargestellt.

Er ergibt sich daher keine Änderung im Berichtsplan.

2.3.2 Nutzen versus diagnostische Güte

Die DGN stellt fest, dass der vorläufige Berichtsplan lediglich auf Studien zum patientenrelevanten (Zusatz-)Nutzen fokussiere. Sie regt daher an, auch „die diagnostische und prognostische Güte konkurrierender Untersuchungsverfahren umfassend zu ermitteln.“ Außerdem schlägt sie vor, die beiden Absätze „allgemeine Voraussetzungen für Studien zur Nutzenbewertung“ und „Allgemeine Voraussetzungen für Studien zur Bewertung der diagnostischen und prognostischen Güte“ an die Bewertung diagnostischer Studien anzupassen. Sie verweist dabei auf einen aktuellen methodischen Diskurs.

Der Berichtsplan sieht bereits vor, falls aufgrund der Datenlage der patientenrelevante Nutzen oder Schaden mit der bildgebenden Diagnostik mittels PET bzw. PET/CT nicht zu ermitteln ist, auch zu untersuchen, ob sich durch den Einsatz der PET bzw. PET/CT die prognostische bzw. diagnostische Güte der Diagnosestrategie verbessert (vgl. Abschnitt 2). Dem Wunsch der DGN ist damit Rechnung getragen. Bezüglich des patientenrelevanten Nutzens diagnostischer Verfahren haben die DGN und das IQWiG ihre Standpunkte bereits ausführlich ausgetauscht [1,2]. Selbst der Nuklearmedizin nahestehende Methodiker sehen in den meisten Fällen die Notwendigkeit, prospektive, randomisiert- kontrollierte Studien durchzuführen, um den Nutzen oder Schaden der PET eindeutig zu quantifizieren [3].

Es ergibt sich daher keine Änderung im Berichtsplan.

2.4 Methoden

2.4.1 Patientenpopulation

Die DGN regt in ihrer Stellungnahme an, bei den Einschlusskriterien der Patienten (4.1.2.1) klar festzulegen, „wie die ‚eingeschränkte regionale oder globale linksventrikuläre Funktion‘ definiert werde. Eine exakte Definition [sei] für diese Analyse hoch-relevant, da definitionsgemäß der Grad der Verbesserung der linksventrikulären Funktion nach Revaskularisierung vom Grad der Einschränkung vor der Revaskularisierung abhängig“ sei. Die Einschränkung auf einen bestimmten Schweregrad der Erkrankung war auch Gegenstand der Erörterung.

Die bereits in der Konkretisierung festgehaltene Einschränkung der Analyse auf Patienten mit „koronarer Herzerkrankung und eingeschränkter regionaler oder globaler linksventrikulärer Funktion“ wurde daher präziser operationalisiert.

Im Berichtsplan wird daher das Einschlusskriterium E1 (Patientenpopulation) wie folgt geändert: Patienten mit eingeschränkter regionaler oder globaler linksventrikulärer Funktion und einer Ejektionsfraktion von 35 % oder weniger.

2.4.2 Geeigneter Referenzstandard

Die DGN weist in ihrer Stellungnahme außerdem darauf hin, dass „üblicherweise bei Vergleichen in der Bildgebung tatsächlich die FDG-PET als der Goldstandard für hibernierendes Myokard bzw. die myokardiale Vitalität benutzt [wird], da die FDG-PET die krankheitsrelevanten Stoffwechselfvorgänge des Gewebes direkt darstellt“. In beiden Stellungnahmen wird auf aktuelle Leitlinien hingewiesen, die den Einsatz der PET bzw. PET/CT als Referenzstandard empfehlen. Auch die SPECT wird von den Stellungnehmenden als Referenztest empfohlen. Dieser Punkt war auch Gegenstand der Erörterung zum vorläufigen Berichtsplan.

Hier muss zunächst eine begriffliche Differenzierung vorgenommen werden. Die PET kann eingesetzt werden, um vor der Revaskularisation die Vitalität zu bestimmen (= Indextest). In dieser Funktion wird sie in den Leitlinien der Fachgesellschaften auch als das gegenwärtig beste Verfahren empfohlen. In diesem Zusammenhang wird sie bisweilen als diagnostischer Gold- oder Referenzstandard bezeichnet. Ob sie in dieser Funktion tatsächlich das beste Verfahren darstellt, ist u. a. Gegenstand des vorliegenden Berichts.

In der Erörterung war man sich darüber einig, dass der beste zur Verfügung stehende Test zur Erfassung von Veränderungen in der regionalen Wandbewegung des Herzens (= Referenzstandard) aus einer zweizeitigen Messung (Baseline und nach Revaskularisation) und einem standardisierten Analysealgorithmus (beispielsweise entsprechend dem 17-Segmente-Modell der American Heart Association) bestehe. Hierfür seien mehrere Verfahren etwa gleich gut geeignet. Genannt wurden die Echokardiographie (Echo), die Radionuklid-ventrikulographie (RNV), die Computertomographie (CT), die Einzelphotonenemissions-

Computertomographie (SPECT; ggf. als Gated-SPECT), die Magnetresonanztomographie (MRT), die Angiographie (Lävokardiographie) und die PET bzw. PET/CT.

Begrifflich wird daher im vorliegenden Bericht zwischen den Indextests (zur Messung der Vitalität) und den Referenzstandards (zur Messung der Veränderungen in der regionalen Wandbewegung des Herzens) unterschieden. Als Indextests werden alle Verfahren zur Messung der Vitalität berücksichtigt (keine Einschränkungen). Als Referenzstandard werden alle oben genannten Verfahren berücksichtigt, sofern sie entsprechend den oben dargestellten Kriterien zur Messung der Wandbewegung eingesetzt werden.

Auch die Frage, ob eine diagnostische Technologie, wenn sie in derselben Studie sowohl zur Vitalitätsdiagnostik (= Indextest) als auch zur Messung der Kontraktilität vor und nach Revaskularisierung (= Referenzstandard) verwendet wird, zu einem Inkorporationsbias führt und daher ausgeschlossen werden sollte, wurde in der Erörterung diskutiert.

Insbesondere bei Verfahren wie der Ultraschalldiagnostik ist ein solcher Bias unmittelbar ersichtlich, da bei der Messung des Index- und des Referenzstandards dieselben Parameter zur Anwendung kommen. Auch bei Technologien, die unterschiedliche Parameter verwenden, ist ein Inkorporationsbias nicht auszuschließen, sodass der Einschluss derartiger Studien nicht sinnvoll erscheint.

Entsprechend dem Auftrag ist ein Einsatz der PET bzw. PET/CT als Indextest Voraussetzung für den Einschluss einer Studie in den Bericht. Daher werden alle Studien, die die PET oder PET/CT auch als Referenzstandard benutzen, prinzipiell ausgeschlossen. Dies gilt auch, wenn für Indextest und Referenzstandard unterschiedliche PET-Parameter erhoben wurden.

Änderungen ergeben sich im Berichtsplan bezüglich der Technologien, die als Referenzstandard berücksichtigt werden. Entscheidend sind hierfür außerdem eine zweizeitige Messung und ein standardisierter Auswertungsalgorithmus.

Die DGN weist in ihrer Stellungnahme darauf hin, dass der in Abschnitt 4.2.1.4 angegebene Mindestzeitraum für eine Nachbeobachtung nach Revaskularisation von einer Woche „in der Situation der chronischen linksventrikulären Dysfunktion bei Herzinsuffizienz bei weitem nicht ausreichend [sei], um einen Therapieeffekt abschließend beurteilen zu können“.

Die DGN nennt aber in diesem Zusammenhang keine Literaturquelle und schlägt auch keine aus ihrer Sicht adäquate Mindestzeitspanne vor. Im vorläufigen Berichtsplan wurde bei Festlegung dieses Grenzwerts auf die systematische Übersicht von Bax et al. [4] rekurriert. Dort wird (auf S. 159) für eine frühe Beurteilung des Behandlungsansprechens ein Zeitpunkt von mindestens 7 Tagen nach Revaskularisierung angegeben. Um potenziell relevante Arbeiten nicht auszuschließen, wird die Mindestnachbeobachtungsdauer von 7 Tagen daher beibehalten.

Es ergibt sich keine Änderung für den Berichtsplan.

2.4.3 Informationssynthese

Die DKG teilt mit, dass die im Berichtsplan (Abschnitt 4.5) festgelegte Irrelevanzschwelle für Skalen von 0,2 standardisierte Mittelwertdifferenzen (Cohens d) nicht akzeptabel sei. Sie nennt hierzu keine Referenzen und nennt keine Alternative, sondern schlägt vor, die Effekte lediglich zu berichten und die Beurteilung der klinischen Relevanz dem G-BA zu überlassen.

Dieser Punkt betrifft eine Festlegung, die bereits im Rahmen der allgemeinen Methoden des IQWiG getroffen wurde [5]. Einzelne Berichte orientieren sich regelhaft an den Methoden des Instituts. Eine methodische Befassung mit diesem Punkt sollte daher ggf. im Rahmen der nächsten Überarbeitung der Allgemeinen Methoden erfolgen. Diese Schwelle wird im Übrigen nur für Skalen verwendet, die über keine validierten Schwellenwerte verfügen. Die prospektive Festlegung einer Relevanzschwelle dient allein dem Zweck, den Bewertungsprozess objektiv und transparent zu gestalten. Es steht dem G-BA selbstverständlich frei, statistisch signifikante Effekte vergleichender Studien anders zu interpretieren, als dies von IQWiG vorgenommen wird.

Es ergibt sich daher keine Änderung für den Berichtsplan.

2.5 Literatur

1. Scheibler F, Raatz H, Suter K, Janssen I, Grosselfinger R, Schröer-Günther M et al. Nutzenbewertung der PET bei malignen Lymphomen: Standpunkt des IQWiG. Nuklearmedizin 2010; 49(1): 1-5.
2. Kotzerke J, Dietlein M, Grünwald F, Bockisch A. Vorbehalte der DGN zur Nutzenbewertung der PET durch das IQWiG. Nuklearmedizin 2010; 49(1): 6-12.
3. Vach W, Hoiland-Carlsen PF, Gerke O, Weber WA. Generating evidence for clinical benefit of PET/CT in diagnosing cancer patients. J Nucl Med 2011; 52 Suppl 2: 77S-85S.
4. Bax JJ, Poldermans D EA, Boersma E, Rahimtoola SH. Sensitivity, specificity and predictive accuracies of various noninvasive techniques for detecting hibernating myocardium. Current Problems in Cardiology [online] 2001; 26(2).
5. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 4.0 [online]. 23.09.2011 [Zugriff: 19.12.2011]. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf.

3 Offenlegung potenzieller Interessenkonflikte

Im Folgenden sind die potenziellen Interessenkonflikte der Stellungnehmenden sowie weiterer Teilnehmer an der wissenschaftlichen Erörterung zusammenfassend dargestellt. Alle Informationen beruhen auf Selbstangabe der einzelnen Personen anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte“. Das Formblatt ist unter www.iqwig.de abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

3.1 Potenzielle Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen und Firmen

Organisation / Institution / Firma	Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
DGN	Bengel, Frank	nein	ja	ja	ja	nein	nein
	Burchert, Wolfgang	ja	ja	ja	ja	nein	nein
	Kotzerke, Jörg	nein	nein	nein	nein	nein	nein
	Schäfers, Michael	nein	nein	nein	ja	nein	nein
DKG	Acikgöz, Özlem	ja	nein	nein	nein	nein	nein
	Dippmann, Anja	ja	nein	nein	nein	nein	nein
	Hermeyer, Gottfried	ja	nein	nein	nein	nein	nein
	Rüther, Klaus	ja	nein	nein	nein	nein	nein
	Schlottmann, Nicole	ja	nein	nein	nein	nein	nein

3.2 Potenzielle Interessenkonflikte von weiteren Teilnehmern der wissenschaftlichen Erörterung (externe Sachverständige)

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
Bäumer, Anselm	nein	ja	ja	nein	nein	nein
Hoffmann, Frauke	nein	nein	nein	nein	nein	nein
Kasper, Jürgen	nein	nein	nein	nein	nein	nein

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

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

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

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

Frage 4: Haben Sie und / oder hat die Einrichtung¹, die Sie vertreten, abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischem Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Frage 5: Haben Sie und/oder hat die Einrichtung¹, bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren ohne wissenschaftliche Gegenleistung) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

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

¹ Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.

4 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll

4.1 Teilnehmerliste der wissenschaftlichen Erörterung

Name	Organisation / Institution / Firma /privat
Bäumer, Anselm	Praxis für Innere Medizin
Bengel, Frank	Deutsche Gesellschaft für Nuklearmedizin e. V.
Burchert, Wolfgang	Deutsche Gesellschaft für Nuklearmedizin e. V.
Hausner, Elke	IQWiG
Hoffmann, Frauke	Institut für Kommunikation in der Medizin
Kasper, Jürgen	Gesundheitswissenschaften Universität Hamburg
Kotzerke, Jörg	Deutsche Gesellschaft für Nuklearmedizin e. V.
Mennekes, Beate	Sitzungsdokumentarischer Dienst, Landtag NRW
Sauerland, Stefan	IQWiG
Sieben, Wiebke	IQWiG
Scheibler, Fülöp	IQWiG
Windeler, Jürgen	IQWiG (Moderation)

4.2 Liste der Stellungnahmen, zu denen kein Vertreter an der wissenschaftlichen Erörterung teilgenommen hat

In der folgenden Tabelle werden Stellungnahmen genannt, zu denen trotz Einladung kein Stellungnehmender oder Vertreter zur wissenschaftlichen Erörterung erschienen ist.

Organisation / Institution / Firma / Privatperson
Deutsche Krankenhausgesellschaft

4.3 Tagesordnung der wissenschaftlichen Erörterung

	Begrüßung
	Einleitung
TOP 1	Sind diagnostische Studien überhaupt sinnvoll / interpretierbar (Treatment paradox)?
TOP 2	Adäquater Referenzstandard (nur MRT ausreichend?)
TOP 3	Ausschluss von (vergleichenden) Studien mit Patienten mit nur geringfügig eingeschränkter regionaler und globaler linksventrikulärer Funktion?
TOP 4	Verschiedenes

4.4 Protokoll der wissenschaftlichen Erörterung

Datum: 21.02.2012, 11:00 bis 13:00 Uhr

Ort: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG),
Dillenburger Straße 27, 51105 Köln

Moderation: Jürgen Windeler

Beginn: 11:00 Uhr

4.4.1 Begrüßung und Einleitung

Moderator Jürgen Windeler: Ich begrüße Sie zu der Erörterung zu dem Thema „PET, PET/CT bei koronarer Herzerkrankung“, um das erst einmal so allgemein zu formulieren. Ich freue mich, dass Sie gekommen sind.

Am Anfang gibt es immer einige technische Details, die ich noch einmal zum Besten geben möchte, auch wenn sie der eine oder andere schon gehört hat, der bereits an früheren Erörterungen teilgenommen hat. Der erste wichtige Punkt ist, dass die Erörterung aufgezeichnet wird und das Protokoll anschließend auf der Internetseite des IQWiG mit Namensnennung der Beitragenden veröffentlicht wird. Wir haben Sie in der Einladung darauf hingewiesen, dass die Aufzeichnung und insbesondere die Veröffentlichung stattfinden werden. Wir gehen davon aus, dass Sie mit Ihrem Kommen Ihr Einverständnis erklären. Falls nicht, muss ich Sie darauf hinweisen, dass Sie an der Erörterung nicht teilnehmen können. – Ich habe nicht den Eindruck, dass das ein Problem gibt.

Im Zusammenhang mit der Veröffentlichung und der Aufzeichnung, auch für die Protokollantin, möchte ich Sie bitten, vor jeder Wortmeldung Ihren Namen zu sagen, sodass wir die Äußerungen zuordnen können.

Noch einmal zum Ziel der Erörterungen: Wir haben schriftliche Stellungnahmen bekommen, speziell von der Deutschen Gesellschaft für Nuklearmedizin. Die Stellungnahmen haben wir

sorgfältig studiert und glauben, dass wir das meiste verstanden haben. Die Erörterung jetzt dient der Klärung oder Vertiefung von Punkten, die wir vielleicht schon verstanden haben, zu denen wir aber noch Diskussionsbedarf mit Ihnen haben.

Es gibt wie üblich den Tagesordnungspunkt „Verschiedenes“. Wenn Sie über die anderen Tagesordnungspunkte hinausgehend noch Kommentare oder weitere Punkte anzubringen haben, dann können Sie das gerne machen.

Wir hatten Ihnen die Tagesordnung zugeschickt und haben sie Ihnen jetzt in etwas geänderter Form noch einmal vorgelegt. Die geänderte Form bezieht sich nur darauf, dass wir TOP 1 und TOP 2 aus rein organisatorischen Gründen, weil es uns sinnvoller erschien, getauscht haben. Ich frage Sie trotzdem: Haben Sie noch Ergänzungsbedarf, Besprechungsbedarf zur Tagesordnung, Punkte, die Sie vermissen? – Das ist nicht der Fall.

Wenn es von Ihrer Seite keine Unklarheiten mehr gibt, machen wir jetzt ohne Pause durch. Am Ende gibt es einen Imbiss. Ich gehe davon aus, dass wir zwischendurch keine Pause brauchen.

Jörg Kotzerke: Können wir vielleicht eine kleine Vorstellungsrunde machen, weil Herr Prof. Bengel zum ersten Mal von unserer Seite aus dabei ist?

Moderator Jürgen Windeler: Gut, das können wir gerne machen. – Wollen Sie anfangen? Die Protokollantin, das ist klar.

Beate Mennekes: Mein Name ist Beate Mennekes. Ich habe nicht gestern zu viel Rosenmontag gefeiert, sondern bin noch etwas erkältet und führe das Protokoll.

(Fülöp Scheibler: Das kommt ins Protokoll!)

Frank Bengel: Mein Name ist Frank Bengel. Ich bin seit einem Jahr Direktor der Klinik für Nuklearmedizin an der MH in Hannover, war vorher fünf Jahre in den USA an der Johns Hopkins Universität in der kardiovaskulären Nuklearmedizin tätig.

Jörg Kotzerke: Jörg Kotzerke, Ordinarius für Nuklearmedizin in Dresden und Präsident der Deutschen Gesellschaft für Nuklearmedizin.

Wolfgang Burchert: Mein Name ist Wolfgang Burchert. Ich arbeite am Herz- und Diabeteszentrum Nordrhein-Westfalen und beschäftige mich schon lange mit dem Thema Herzbildgebung.

Elke Hausner: Elke Hausner, tätig im Stabsbereich Informationsmanagement im IQWiG.

Wiebke Sieben: Wiebke Sieben. Ich arbeite im IQWiG in der Biometrie.

Stefan Sauerland: Stefan Sauerland. Ich leite das Ressort Nichtmedikamentöse Verfahren am IQWiG.

Moderator Jürgen Windeler: Jürgen Windeler, Leiter des IQWiG und Moderator der Erörterung.

Fülöp Scheibler: Fülöp Scheibler. Ich arbeite im Ressort Nichtmedikamentöse Verfahren und bin zuständig für dieses PET-Projekt.

Andrea Steinzen: Andrea Steinzen. Ich bin auch Mitarbeiterin im Ressort Nichtmedikamentöse Verfahren.

Anselm Bäumer: Anselm Bäumer. Ich bin niedergelassener Kardiologe und Pneumologe und arbeite mit an dem Projekt.

Jürgen Kasper: Jürgen Kasper, Gesundheitswissenschaften Universität Hamburg, externer Sachverständiger.

Frauke Hoffmann: Frauke Hoffmann, ebenfalls externe Sachverständige, tätig für das Institut für Kommunikation in der Medizin.

Moderator Jürgen Windeler: Vielen Dank. – Dann kommen wir jetzt zu Tagesordnungspunkt 1

4.4.2 Tagesordnungspunkt 1: Adäquater Referenzstandard (nur MRT ausreichend?)

Fülöp Scheibler: Im vorläufigen Berichtsplan sind zur Erhebung des Referenzstandards die Verfahren LDDE, MRT und CT vorgesehen. In Ihrer Stellungnahme schlagen Sie vor, auch die SPECT als gated oder nicht gated SPECT mit zu berücksichtigen. Wir möchten nachfragen, ob Sie die SPECT als Referenztest oder als weiteren Indextest in die Bewertung aufnehmen wollen.

Zur Erklärung, weil die Begriffe in der Literatur teilweise unterschiedlich verwendet werden: Für uns sind die Indextests die Tests, die bewertet werden sollen, deren Sensitivität und Spezifität ermittelt werden soll. Referenztest oder Referenzstandard ist für uns der Vergleichstest, der notwendig ist, weil ein Goldstandard nicht zur Verfügung steht, anhand dessen die Daten zur diagnostischen Güte ermittelt werden sollen.

So wie wir Ihre Stellungnahme verstanden haben, wollen Sie gerne, dass wir die SPECT noch zusätzlich als Referenztest, also als Quasigoldstandard mit berücksichtigen. Wir möchten nachfragen, ob das tatsächlich so gemeint war.

Wolfgang Burchert: Ich denke schon, dass das sinnvoll ist. Wir haben im deutschen Gesundheitswesen zwei wesentliche Verfahren: Das sind die Echokardiographie und die SPECT, die es im Moment für diese Indikationen in der gesetzlichen Krankenversorgung gibt.

Die anderen – MRT, vielleicht auch CT, das ist momentan sicherlich noch nicht richtig kliniktauglich – sind in der Krankenversorgung sozusagen nicht verfügbar. Da es um die Fragestellung geht: „Gesetzliche Krankenversorgung ja oder nein?“, war der Gedanke, dass man es schon an den Standards messen müsste, die tatsächlich zur Verfügung stehen. Die MRT – korrigieren Sie mich, wenn das falsch ist – ist meines Wissens im Moment noch nicht zugelassen. Aus wissenschaftlicher Sicht kann man das natürlich auch als Referenzstandard nehmen.

Fülöp Scheibler: Das Problem, das wir dabei sehen, ist: Hier geht es um die Messung der Kontraktilität nach Revaskularisation. Der Referenzstandard besteht ja in diesem Fall – so wie wir es verstanden haben – aus Messungen der Kontraktilität, die nicht vorhanden ist, dann der Revaskularisation und dann einer erneuten Messung der Kontraktilität. Das Delta zwischen den beiden Messungen zeigt: Hat sich da etwas verändert oder nicht? Wenn sich etwas verändert hat, dann war die Revaskularisation erfolgreich, und das Herz pumpt an dieser Stelle wieder. Die Frage ist: Ist die SPECT wirklich dafür geeignet?

Frank Bengel: Man muss das noch einmal klar definieren. Wir müssen auf das zurückkommen, was Sie gerade gesagt haben, und aufpassen, dass wir uns mit dem Wort „Referenzstandard“ nicht verwirren. Es gibt nämlich zwei Ansätze, die man hier verfolgen könnte.

Ich beschreibe einmal den ersten Ansatz, der meines Erachtens – insbesondere der Kommentar zu 5.6 – nicht gemeint ist: Man hat einen diagnostischen Test, so wie die Positronenemissionstomographie, der vor der Revaskularisation angewendet und mit einem zweiten Test verglichen wird, der auch nur vor der Revaskularisation angewendet wird, weil man denkt, dass er der eigentliche Goldstandard ist. Ich verwende jetzt den Begriff „Goldstandard“ für die Vitalitätsdiagnostik. In diesem Falle würden die Tests dann nicht mehr nach der Revaskularisation wiederholt werden, sondern man würde irgendein anderes Verfahren nehmen, zum Beispiel Outcome oder was auch immer, um zu schauen, ob das eine Verfahren, also PET, mit dem Goldstandard oder Referenzstandard, den es in dieser Situation eigentlich nicht gibt, vergleichbar ist. Das wäre ein Ansatz, der hier meines Erachtens nicht gegeben ist, weil eher das Umgekehrte eintritt. Wenn ein neues Verfahren kommt – welches Verfahren auch immer –, dann wird häufig in einem ersten Szenario geschaut, ob dieses neue Verfahren mit der Positronenemissionstomographie als Verfahren, das vor der Revaskularisation angewendet wird, auch wirklich vergleichbar ist, sprich: In diesem Szenario ist PET eigentlich schon als Referenzstandard anerkannt.

Dann gibt es den zweiten Schritt – das ist, glaube ich, das, was Sie gerade gemeint haben: Man hat das eine Verfahren, also den Vitalitätstest, und dann, was wir eigentlich Referenzstandard nennen sollten, das zweite Verfahren, das sowohl vor als auch nach Revaskularisation angewendet wird, um über sequenzielle Veränderungen der linksventrikulären Funktion – darüber reden wir hier – festzustellen, ob durch den anderen Test, den Vitalitätstest, vor der Revaskularisation vorhergesagt werden kann, ob nach der

Revaskularisation eine sequenzielle Veränderung oder Verbesserung der linksventrikulären Pumpfunktion eintritt. Das ist das klassische Szenario, das auch in der Literatur als qualitativ hochwertiges Szenario angesetzt wird, um festzustellen, ob ein Vitalitätsdiagnostiktest wirklich sinnvoll ist. Denn man testet: Kann man eine Vorhersage der Funktionsverbesserung nach der Revaskularisation machen?

Jetzt kommen wir zu der Frage: Welches Verfahren ist dafür geeignet, sequenziell die linksventrikuläre Pumpfunktion zu erfassen? Dafür gibt es ein breites Spektrum an Verfahren, das einsetzbar ist. Letzten Endes sind es alles Verfahren, die in der klinischen kardiovaskulären Praxis genutzt werden, um die linksventrikuläre Pumpfunktion zu messen.

In dem Berichtsplan verwenden Sie unter anderem Terminologien, die meines Erachtens auf solche Verwechslungen schließen lassen. Eine Low-dose dobutamine echocardiography, also LDDE, ist ein Test zur Vitalitätsdiagnostik, der vor der Revaskularisierung angewendet wird. Die Echokardiographie als solche ist aber ein Test, um die linksventrikuläre Pumpfunktion in Ruhe zu beurteilen, die durchaus für serielle Messungen der linksventrikulären Funktion geeignet ist und in der klinischen Praxis auch so eingesetzt wird. Eine Low-dose dobutamine echocardiography ist nicht geeignet als Referenzstandard nach den Kriterien der zweiten Kategorie, wie ich gerade gesagt habe, weil es ein Vitalitätstest ist. Unserer Meinung nach ist eine Echokardiographie aber durchaus geeignet, um serielle Veränderungen der linksventrikulären Pumpfunktion zu erfassen.

Jetzt wieder der Bogen zurück zur SPECT oder zu nuklearmedizinischen Techniken: Es gibt auch nuklearmedizinische Techniken, die sehr gut geeignet sind, die linksventrikuläre Pumpfunktion sequenziell zu beurteilen. Die SPECT zählt dazu, dazu gehören meines Erachtens aber auch alte Verfahren wie die Radionuklidventrikulographie. Auch die ist in der Literatur häufig für diese Belange angewendet worden. Man muss hier definieren. Es gibt ein paar Grundvoraussetzungen, die schon erfüllt sein müssen, wenn man solche Untersuchungen heranziehen will, um zu beurteilen, ob mit PET eine sequenzielle Verbesserung der linksventrikulären Pumpfunktion vorhergesagt werden kann. Man muss als Maßstab angeben, dass vorher und nachher wirklich der gleiche Test verwendet worden ist, der auch unabhängig von dem Test ist, der als Vitalitätsdiagnostik angewendet worden ist, sprich: Es sollte nicht PET sein, die hinterher noch einmal verwendet worden ist.

Jörg Kotzerke: Es scheint mir ein Trugschluss zu sein, dass die Myokard-Perfusions-Szintigraphie (SPECT), mit der primär die Durchblutung des Myokards gesehen wird, nicht als ein Test anerkannt wird – wenn sie zusammen mit einer EKG-Triggerung gemacht wird, also als gated SPECT –, wobei dann auch die Wandbewegung des Myokards beurteilt werden kann. Die Myokard-Perfusions-Szintigraphie ist durchaus in zwei Techniken durchführbar: Eine beurteilt nur die Perfusion. Wenn das Ganze herzphasengetriggert, also die zeitliche Komponente mit drin ist, dann kann man es nicht nur herzachsengerecht rekonstruieren, sondern zusätzlich auch die globale und regionale Wandbewegung beurteilen. Dann ist

natürlich auch gated SPECT eine Referenz, um die Pumpfunktion vor und nach Revaskularisierung einschätzen zu können.

Anselm Bäumer: Man könnte dazu anmerken, dass es aus kardiologischer Sicht keine Leitlinie oder generell gültige Empfehlung gibt, wie eine regionale Wandbewegung des Herzens gemessen werden muss oder soll. Es gibt keine Leitlinie, die sagt: Der Goldstandard ist das und das. Es stehen immer die verschiedenen schon genannten technischen Möglichkeiten zur Diskussion. Von der Genauigkeit, die man für den Referenzstandard fordert – Reproduzierbarkeit, fehlende Abhängigkeit vom Untersucher usw. –, ist die Kernspintomographie, wie es immer heißt, als führende Technik empfohlen und aus meiner Sicht auch so zu sehen. Was die Genauigkeit und Wiederholbarkeit der Ergebnisse angeht, die Möglichkeiten der Beurteilung der Wanddickenzunahme zum Beispiel, ist die Kernspintomographie schon die beste Technik.

Frank Bengel: Die Aussage ist nicht evidenzbasiert; im ersten Teil haben Sie es selber gesagt. In der klinischen Praxis sind sehr viele Verfahren für diese Fragestellung akzeptiert und werden dann auch in der klinischen Praxis durchgeführt. Ich will nur sagen, dass für die routinemäßige klinische Beurteilung der linksventrikulären Pumpfunktion andere Verfahren wahrscheinlich wesentlich häufiger angesetzt werden als eine Magnetresonanztomographie. Ich verstehe zwar Ihren Punkt – das ist der generelle Konsensus, oder viele Leute meinen, es ist das akkurateste Verfahren –, aber es gibt viele Möglichkeiten, wie man das messen kann, und auch viele klinisch akzeptierte Möglichkeiten, wie die linksventrikuläre Funktion seriell gemessen werden kann.

Fülöp Scheibler: Sie haben natürlich vollkommen Recht. Für „evidenzbasiert“ müsste es einen Goldstandard geben, und dagegen müssten dann alle Verfahren getestet worden sein. Wenn wir das hätten, wäre das wunderbar. Das gibt es aber nicht.

Das Problem, das sich uns aus methodischer Sicht stellt, ist: Wenn wir wissen, es gibt unterschiedliche Referenzstandards, und wir gehen davon aus, dass sie von unterschiedlicher Qualität sind, dann kann ein und derselbe Indextest – beurteilt an unterschiedlichen Referenzstandards – zu unterschiedlichen Ergebnissen führen. Das heißt, wenn wir mehrere Referenzstandards berücksichtigen würden, dann hätten wir zumindest eine schwierige Vergleichbarkeit. Nehmen wir einen Indextest, zum Beispiel PET, verglichen mit Referenzstandard 1, und einen anderen Indextest, zum Beispiel SPECT, verglichen mit Referenzstandard 2. Egal was dabei herauskommt, wir können die beiden Ergebnisse nicht vergleichen. Nehmen wir an, PET wurde verglichen mit Referenzstandard 1, und SPECT wurde verglichen mit Referenzstandard 2. Es kommt heraus: Die SPECT schneidet schlechter ab als die PET. Dann wissen wir nicht: Liegt es daran, dass die SPECT wirklich schlechter ist als die PET, oder liegt es daran, dass sich die 2 Referenzstandards unterscheiden?

Insofern: Auch wenn es nicht evidenzbasiert ist, macht es aus klinischer Sicht vielleicht schon Sinn, einen Referenzstandard zu definieren. Das ist natürlich ein Werturteil, aber man legt

vorher fest: Es geht nicht darum, was erstattet wird, sondern wir glauben, technisch ist der sinnvollste Referenzstandard das und jenes. Es gibt zum Beispiel ein Positionspapier der DGN, in dem auch die MRT als bester Test beschrieben wird, um Wandbewegung zu erfassen. Man könnte sich darauf einigen und sagen: Das ist der Referenzstandard, den wir als den besten, technisch am sinnvollsten erachten, und damit vergleichen wir alle Indextests, die es gibt. – Dann hat man zumindest eindeutige Ergebnisse.

Wolfgang Burchert: Dazu muss ich einen Kommentar abgeben. Das bedeutet gleichzeitig, dass man für die FDG-PET nur Patientengruppen testen kann, die auch MR-gängig sind. Gerade Patienten mit eher niedriger linksventrikulärer Pumpfunktion haben häufig automatische Defibrillatoren, Schrittmacher, Implantate usw., alles Dinge, die momentan noch nicht in der Breite MR-gängig sind. Das heißt, durch die Verwendung von MRT würde man diese Patiententeilung von vornherein ausschließen.

Aus diesem Grunde wird MRT ganz häufig nicht als Referenzstandard genommen. Die Bilder sehen zwar wunderbar aus – ich selber mache das auch –, bei der Bestimmung der Volumina gibt es aber durchaus Probleme bei der MRT. Innen hat man zum Beispiel eine trabekuläre Struktur. Nimmt man sie mit oder nicht? Das sind, weil sie außen liegen, erhebliche Volumenunterschiede. Auch die Trabekulierung zwischen Patienten ist unter Umständen sehr unterschiedlich. Ein Volumeneintrittsverfahren wie die Radionuklidventrikulographie zum Beispiel hat das nicht, weil es direkt das Volumen markiert. Das sieht zwar optisch wenig beeindruckend aus, es ist aber, weil es sozusagen den ganzen Raum erfasst, relativ genau. Daher würde ich mich von den vordergründigen Effekten nicht so beeindrucken lassen.

Was die Reproduzierbarkeit und solche Dinge angeht, ist die MRT sehr gut und für viele Fragestellungen auch hervorragend geeignet. Wenn man sie aber hier einsetzt, dann würde man die Anzahl der Studien sofort sehr stark beschneiden; denn in vielen Studien ist MRT einfach nicht eingesetzt worden, war nicht verfügbar bei Multicenter-Studien gleicher Qualität überall usw. Dann kommt man in Probleme, weil dieses Verfahren sozusagen auch multizentrisch erst einmal stabilisiert werden muss. Da gibt es praktisch nur sehr wenige Studien. Das wäre theoretisch ein sehr attraktiver Ansatz, aber ich würde sagen: Man muss damit leben, dass die Referenzverfahren in den Studien, was die Verfügbarkeit und auch die Patientengruppen angeht, sehr unterschiedlich gestaltet sind. Wie gesagt, wenn sie einen Defibrillator oder einen Schrittmacher haben, werden sie nicht im MRT untersucht. Für diese Gruppen fällt das völlig aus, und das ist nicht so selten.

Frank Bengel: Dazu fallen mir ein paar Punkte ein. Vielleicht müssten Sie als Erstes noch einmal definieren, was wirklich Ihr Ziel ist. Wenn ich Sie richtig verstanden habe, scheint Ihr Ziel zu sein, dass Sie einen Vitalitätstest bezüglich des Vorhersagepotenzials zur Verbesserung der linksventrikulären Pumpfunktion mit anderen vergleichen können.

Fülöp Scheibler: Genau.

Frank Bengel: Sie wollen also wirklich in der Analyse auf alle verfügbaren Vitalitätsdiagnostiktests eingehen? Womit wollen Sie PET vergleichen? Mit einem anderen Vitalitätstest?

Fülöp Scheibler: Genau.

Frank Bengel: Mit welchem anderen?

Fülöp Scheibler: Mit jedem anderen. Beim Indextest haben wir uns nicht beschränkt.

Frank Bengel: Dann müsste man erst einmal ganz klar darlegen, welche Verfahren es dafür überhaupt gibt. Ich glaube, das ist in dem Bericht noch nicht ganz klar dargelegt. Man muss dann auch aufpassen, dass man keine anderen Verfahren auslässt. – Das ist der erste Punkt.

Mein zweiter Punkt ist: Ich verstehe, dass man MRT als Referenzstandard nehmen will, weil es unter Umständen zum aktuellen Zeitpunkt der akkurateste Test für eine linksventrikuläre Pumpfunktionserfassung ist. Meines Erachtens schließt das aber aus, dass Sie MRT-Techniken als Vitalitätstests mit berücksichtigen, weil dann die gleiche Untersuchungsmethode a priori zur Vitalitätsdiagnostik verwendet worden ist und post Revaskularisationem noch einmal als Referenzstandard. Das ist eine Vermischung von beiden und würde nicht funktionieren.

Deswegen reicht es meines Erachtens nicht, nur einen Test zu nehmen, der vorher und nachher genommen wird, sondern man muss letzten Endes durch die Literatur schauen und alle verfügbaren Tests nehmen, die dafür verwendet worden sind. Meine Vorhersage wäre auch, dass Sie wahrscheinlich, wenn Sie durch die Literatur schauen, ein relativ breites Spektrum an Techniken finden werden, die zur seriellen Messung der Pumpfunktion verwendet worden sind, sodass Effekte, die Sie angedeutet haben – mit Verfahren A als Vitalitätstest, bei dem dann Pumpfunktionsverbesserungen mit Verfahren B gemessen worden sind, mit Verfahren C als Vitalitätstest, bei dem Pumpfunktionsverbesserungen mit Verfahren D gemessen worden sind –, wahrscheinlich nicht eintreten werden. Sie werden wahrscheinlich A und C vergleichen, und Sie werden Studien finden, in denen sowohl B als auch D zur seriellen Funktionsmessung genommen worden sind. Deswegen würde sich das wahrscheinlich herausschälen. Das wäre meine Vorhersage. Deswegen muss man erst einmal systematisch durch die ganze Literatur schauen, was sie überhaupt hergibt.

Mein letzter Punkt, was das betrifft: Man muss dabei auch berücksichtigen – und wieder aufpassen, dass man kein Bias hineinbringt –, dass Verfahren, die älter sind, also schon vor längerer Zeit etabliert worden sind, wahrscheinlich nicht die aktuellen Goldstandards angewendet haben. Konkret: Die Positronenemissionstomographie zur Vitalitätsdiagnostik ist erstmals 1985 vorgestellt worden. Damals hat es keine kardiovaskuläre serielle Magnetresonanztomographie gegeben. Das heißt, damals ist alles mit Echokardiographie oder Angiographie gemacht worden. In den darauffolgenden 5 oder 10 Jahren sind sehr viele ähnliche Publikationen herausgekommen, die auch Echokardiographie, Angiographie,

Radionuklidventrikulographie verwendet haben. Danach ist das Thema irgendwann einmal ausgearbeitet gewesen und nicht nur wiederholt mit Studien abgedeckt worden. Deswegen werden Sie, wenn Sie systematisch durch die Literatur gehen, auch durch die Selektion des aktuellen Goldstandards, ein gewisses Bias hineinbringen, weil Verfahren, die es schon länger gibt, plötzlich weniger Evidenz haben, obwohl sie eigentlich mehr Evidenz haben. Verstehen Sie, was ich meine?

Fülöp Scheibler: Ja.

Frank Bengel: Auch da gibt es also eine Problematik.

Jörg Kotzerke: Den letzten Punkt von Herrn Bengel möchte ich unterstreichen, weil die myokardiale Vitalitätsmessung mit FDG-PET auf Anfang der 90er – Sie sagten: Erstmessung 1985 – zurückgeht, also einen Zeitpunkt, als die MR-Diagnostik nicht die Qualität hatte, wie es heute der Fall ist, weder von den Sequenzen her noch von der zeitlichen oder der örtlichen Auflösung her.

Beim CT ist auch ganz einfach zu ersehen, dass die Mehrzeilen-CTs, die dann kardiovaskuläre Diagnostik erlauben, Entwicklungen der letzten Jahre sind. Das ist genau das, was Herr Bengel gesagt hat. Wenn ein Thema mit vorhandenen Möglichkeiten abgearbeitet erschien, dann wird das nicht unbedingt wieder aufgegriffen, sondern man akzeptiert es einfach: „So ist es“ und wendet sich dann neuen Fragestellungen zu.

Moderator Jürgen Windeler: Ich habe jetzt den Eindruck – ich fasse das kurz zusammen –, dass wir über ein größeres Spektrum von Indextests sprechen, dass wir über ein größeres Spektrum von Referenztests sprechen, also die Revaskularisationserfolge, dass wir auch über Studien sprechen, in denen die beiden Tests nicht gleichzeitig verwendet werden dürfen – das haben Sie gesagt –, dass wir über ein Altersproblem sprechen in dem Sinne, dass bestimmte Referenztests bei bestimmten Verfahren wahrscheinlich nicht oder nur rudimentär angewendet worden sind. Möglicherweise haben wir auch nicht die Situation, dass A immer mit C und B immer mit D korreliert, aber dass möglicherweise doch in den Mischungen – einfach aufgrund der zeitlichen Reihenfolge – bei A bevorzugt D und bei B bevorzugt C angewendet wird.

Ich habe ein bisschen den Eindruck, dass wir uns bei dieser Vielfalt – Sie sagen dann noch, MRT ist anders als Volumenmessung, sodass wir nicht davon ausgehen können, dass sich die Referenztests untereinander entsprechen – in einem großen Knäuel verheddern. Oder habe ich das jetzt missverstanden?

Frank Bengel: Das ist genau die Problematik, wenn man über die Zulassung von Tests 20, 25 oder 30 Jahre nachdem sie eingeführt sind, entscheiden will. Es ist praktische Realität. Die Tests sind schon längst in die Praxis eingegangen. Sie haben eine wahnsinnige Vielfalt, weil die Dinge, die vielleicht jetzt in die Wege geleitet werden sollen, eigentlich vor 30 Jahren hätten in die Wege geleitet werden sollen. Das ist die Problematik dabei. Aber ob man das

dann gegen den Test auslegt oder es nicht vielleicht eher umgedreht sehen sollte, wenn sich ein Test so vielfältig entwickelt hat – ein Vitalitätstest, welcher auch immer –, dass er dann einen gewissen Sinn macht, ist meines Erachtens eine philosophische Frage. Wir kommen hier wieder auf die grundlegende Problematik bei diagnostischen Tests. Wie kann man Tests, die schon einen gewissen Einzug in die praktische Realität gefunden haben, in dem System so abbilden, dass man entscheiden kann: Stimmt es oder stimmt es nicht, spielen sie wirklich in der praktischen Realität eine Rolle?

Ich muss aber noch einmal auf den ersten Punkt zurückkommen, das ist auch eine philosophische Frage: Wenn Sie wirklich anfangen wollen, den prädiktiven Wert der Positronenemissionstomographie noch mit anderen diagnostischen Verfahren zu vergleichen, ohne dass man ganz genau weiß, was eigentlich der akzeptierte Goldstandard ist – nach landläufiger Meinung ist PET ja eigentlich der Goldstandard –, dann wird es in der Tat sehr divers. Die Frage ist: Ist es nicht ausreichend, wenn man sagt: „Durch einen Test kann vorhergesagt werden, ob sich eine Funktion nach Revaskularisierung verbessert“? Ist das nicht als solches per se ein ausreichendes Kriterium, oder müssen wir noch einen Überkreuzvergleich mit einem anderen Verfahren hineinbringen, von dem man ohnehin nicht genau weiß, ob es eine adäquate Referenz ist?

Anselm Bäumer: Das Problem ist auch, wenn man vor der Maßnahme, also der Revaskularisation, nur testen will und dann ein Tool hat, um etwas zu beeinflussen. Es geht um die Therapie und am Ende um den Patienten. Die Situation ist dazu momentan aus Klinikersicht noch kritisch zu bewerten, weil die Studien widersprüchliche Ergebnisse zeigen und es nicht in die Richtung geht, dass man sagen kann: Habe ich bei der Entscheidungsfindung zu einer Operation oder Maßnahme beim Patienten eine Vitalitätsdiagnostik an der Hand, kann ich das auch günstig beeinflussen. In diagnostischen Studien geht es ja darum, das hibernierende Myokard zu identifizieren und dann am Ende zu sehen – zum Beispiel nach 3 Monaten –: Pumpt es wirklich besser mit Bypass oder Stent? Über diese Frage und die Schwierigkeiten dabei ist schon diskutiert worden.

Noch komplexer wird das Ganze in der klinischen Situation, weil man schon über Jahrzehnte denkt, durch eine gute oder hoch qualitative Vitalitätsdiagnostik, zum Beispiel PET, könnte man das Ergebnis beim Patienten günstig beeinflussen. Das ist momentan unklar, wenn man die STICH-Studie und die wenigen anderen Studien mit größeren Patientengruppen dazunimmt.

Frank Bengel: Jetzt reden wir schon direkt über einzelne Studien. Ich weiß nicht, ob es auf diesem Niveau Sinn macht. Die STICH-Studie hat nichts mit Positronenemissionstomographie zu tun. – So viel nur für das Protokoll. Das kommt da nicht vor.

Fülöp Scheibler: Herr Bengel, das Problem, das Sie ansprechen, ist genau unser technisches Problem. Was machen wir dann? Wir haben eine ganze Menge an Indextests, an Referenzstandards. Die werden dann unterschiedlich miteinander verglichen, und wir

werden – das kann ich jetzt schon prophezeien – eine ganze Fülle von Ergebnissen zur diagnostischen Güte in einer sehr großen Bandbreite bekommen. Die Frage ist: Wem ist damit geholfen? Nur deswegen unsere Überlegung: Kann man sich vielleicht auf einen oder auf wenige Referenzstandards beschränken, um so wenigstens die Bandbreite an Ergebnissen im Vorfeld irgendwie zu reduzieren? Denn, wie gesagt, am Ende sagt man: In den Studien reichen die Sensitivitäten von 20 bis 95 % und die Spezifitäten von 10 bis 95 %. Damit kann niemand etwas anfangen. – Das ist der eine Punkt.

Der andere Punkt, den Sie angesprochen haben, ist der direkte Vergleich. Das ist uns durch die Konkretisierung des G-BA vorgegeben. Da wird explizit danach gefragt: Welche Untersuchung kann die PET oder die PET/CT ersetzen? Das können wir, wenn wir uns auf die diagnostische Güte beschränken, nur in direkten Vergleichen ermitteln. In einer Studie muss die PET als Indextest mit irgendeinem anderen Verfahren als zweitem Indextest verglichen worden sein, und zwar in Bezug auf einen Referenzstandard. Das kann irgendein Referenzstandard sein, er sollte aber möglichst valide sein, nicht dass man dann irgendein Verfahren ungerechterweise als minderwertig bezeichnet, nur weil der Referenzstandard schlecht war. Aber den direkten Vergleich brauchen wir.

Moderator Jürgen Windeler: Ich würde gerne versuchen, weil wir uns im Moment nach meinem Eindruck in einem Meinungsbildungsprozess, in einem Sammlungsprozess bewegen, noch einmal den Punkt zu beschreiben. Wir müssen einen Vergleich von mehreren Indextests machen, jedenfalls von mehr als 2, von mehr als PET selber. Zu PET selber ist ja die Aussage: Sie hat eine hohe Aussagekraft. Das beantwortet nicht die Frage, jedenfalls für uns nicht: Haben andere eine genauso hohe Aussagekraft? Für Sie mag das klar sein, aber für uns nicht. Die müssen wir beantworten.

Der zweite Punkt ist: Eigentlich müsste man, wenn man sich für verschiedene Indextests und deren Vergleiche interessiert, den gleichen Referenztest für alle haben oder aber mehrere Referenztests, von denen man einigermaßen sicher weiß, dass sie das Gleiche abbilden, gleich zuverlässig sind. Ich habe immer noch ein grundsätzliches Fragezeichen auf der Stirn – auch wenn man das nicht sieht –, wie man mit mehreren Referenztests, die unter sich durchaus nicht vergleichbar sein müssen, anscheinend gar nicht sind, umgehen soll – was ich an sich schon schwierig finde – und wie man dann noch mit einer über die Studien – das scheint auch unser Eindruck zu sein – unkalkulierbaren und möglicherweise von sich aus gebiasten Mischung jeweils zwischen Indextest und Referenztest umgehen soll. Die eine Lösung wäre, zu sagen: Es kann nur einen Referenztest geben. Meinetwegen kann es 2 geben, von denen man weiß, dass sie wirklich gut vergleichbar sind, aber es muss jedenfalls eine Standardisierung erfolgen. Sonst müssten wir ein bisschen schulterzuckend sagen: Dabei wird nichts herauskommen.

Wolfgang Burchert: Es gibt noch eine Komponente, über die wir nicht gesprochen haben, die etwas mit dem Alter des Verfahrens zu tun hat. Die therapeutischen Möglichkeiten haben sich natürlich stark entwickelt. Die Verschiebung der Kathetertechniken, die stattgefunden hat

– früher waren bestimmte Dinge sakrosankt, die sofort in einen Bypass gemündet sind, heute aber selbstverständlich sind –, müsste man unter Umständen berücksichtigen, weil das ein nicht geringes Bias in solch eine Untersuchung einführt. Wenn Sie besser revascularisieren können, würde das bedeuten, dass die neueren Tests dadurch alle sozusagen gebiast sind.

Das ist noch ein ganz unabhängiges Problem, das aber vielleicht dazu führt, wie man Diagnostik und Therapie trennen kann. Das ist ja sehr schwierig, weil Sie eine eigene Entwicklung der Therapie haben. Wir kommen jetzt in die Situation, dass wir sozusagen, da die FDG-PET so lange eingeführt ist, eine deutliche zeitliche Entwicklung haben. Das habe ich hier nicht so wiedergefunden. Ich weiß nicht genau, wie man mit Diagnostik versus Therapie umgehen möchte. Ich würde sagen: Wenn wir einen deutlichen therapeutischen Fortschritt haben, dann wird das deutlicher zu sehen sein, als wenn wir einen Fortschritt in der Diagnostik haben. Das ist immer so, weil es sehr stark von dem therapeutischen Potenzial getriggert ist.

Ein zweiter Aspekt zu den Größen, die man nachher anschaut: Wandbewegung oder LVEF global ist ein Parameter, der mit dem Outcome der Patienten korreliert, aber es gibt auch andere Dinge. Nach Infarkten beispielsweise entwickeln die Patienten nicht selten Rhythmusstörungen, weshalb sie mittlerweile in großer Zahl automatisch Defibrillatoren bekommen. Beim FDG-PET sieht man zum Beispiel das Imprint der Ischämie in dem Gewebe immer mit. Das ist der Mismatch, der so schnell apostrophiert wird, der nicht nur einen Teil der Vitalitätsinformationen, sondern auch die Frage abdeckt: Wie anfällig ist der Patient für Rhythmusstörungen? Das hat man zum Beispiel beim MRT nicht, weil da nur die Narbe klar zu erkennen ist, während man das umliegende Myokard nicht ganz so gut definieren kann. Es gibt im Moment Versuche, das zu machen, aber das ist sicherlich noch in der Entwicklung.

Stefan Sauerland: Kurz noch zu dem Punkt der Therapie: Auch uns ist natürlich klar, dass sich die Therapie in diesem Feld enorm entwickelt hat. Nicht umsonst gilt die Kardiologie als ein hoch innovatives Feld der Medizin. Diesem Fakt tragen wir dadurch Rechnung, dass wir vergleichende Diagnostikstudien haben wollen. Dann haben wir nicht eine Studie nur zu FDG-PET von 1985 oder 1995 und eine andere Studie, die auch nicht vergleichbar ist, in der ein anderes diagnostisches Verfahren untersucht wird, sondern wir wollen beide Verfahren innerhalb einer Studie gegen einen Referenztest haben. Innerhalb der Studie kommen dann auch exakt die gleichen Therapiemethoden zum Einsatz, sodass sich das dadurch, dass wir vergleichende Diagnostikstudien betrachten, in dem Fall wieder herausnivelliert. Da sehe ich nicht so ein Problem.

Fülöp Scheibler: Das Problem der sich entwickelnden Interventionen sehen wir natürlich auch. Wir stellen auch die Frage: Inwiefern ist tatsächlich nur der Endpunkt „Wandbewegung“ sinnvoll? Eigentlich ist die Wandbewegung nach der Revascularisation – aus streng methodischer Sicht – ein Surrogatendpunkt. Wir wissen eigentlich gar nichts. Es kann theoretisch sein, dass die Patienten, deren Herzwand sich wunderbar bewegt, trotzdem

früher an einem plötzlichen Herztod sterben. Das kommt noch erschwerend hinzu. Wir sind gar nicht so sicher – Sie sagen es ja selber –, ob wir uns mit diesem einzigen Endpunkt als Referenz begnügen oder ob wir uns noch andere Endpunkte ansehen sollten. Die werden aber selten gleichzeitig in den Studien erfasst.

Dann gäbe es noch ein methodisches Problem. Dann hätte ich einmal eine diagnostische Güte für die Herzwandbewegung, und in derselben Studie mit demselben Index hätte ich einen anderen Referenztest, eine andere prognostische Güte für den plötzlichen Herztod. Sie sehen, die Fülle an möglichen Ergebnissen wird immer größer. Das ist die Schwierigkeit, mit der wir gerade kämpfen. Wir brauchen ja möglichst eindeutige Ergebnisse. Wenn wir uns nicht so sicher sind, was eigentlich der Endpunkt ist, den wir messen, und wenn es auch noch unterschiedliche Endpunkte und unterschiedliche Referenztests gibt, dann fürchten wir, dass das in eine Fülle von Ergebnissen ausartet, die wir nicht mehr interpretieren können.

Jürgen Kasper: Ich habe dem eigentlich kaum noch etwas hinzuzufügen. Ich wollte nur sagen: Aus anderen Gutachten, wo es um den Nutzen – um einen ganz anderen Bereich zu nennen – von diagnostischen Verfahren geht, kenne ich, dass man in Ermangelung entsprechender randomisierter Nutzenstudien zur Beurteilung dann hilfsweise die Messgenauigkeit heranzieht. Das ist nicht die Information, die man eigentlich haben möchte, aber wenigstens etwas, womit man die Evidenz vielleicht zusammensetzen und sagen kann: Das Verfahren misst genau. Wenn man dann noch zeigen kann, dass eine Therapie in diesen Fällen hilft, dann ist das ja etwas.

Aber in diesem Fall können wir bei der Beurteilung der Messgenauigkeit noch nicht einmal die Therapie herauslassen. Wir alle haben eben schon festgestellt, dass eine Beurteilung der Messgenauigkeit der PET ohne eine Revaskularisation dazwischen kaum möglich ist. Deswegen befürchten wir sehr stark, dass wir nur aus dieser Teilfragestellung der Aufarbeitung der Accuracy Studies keine nützliche zusätzliche Information bekommen.

Frank Bengel: Ich muss noch einmal bei dem Methodikvergleichspunkt ansetzen. Sie haben recht klar gesagt, dass die Frage, die hier an PET als Vitalitätstest gestellt ist, lautet: Welchen anderen Vitalitätstest kann sie denn ersetzen? Wenn man zu der Schlussfolgerung kommt, dass PET keinen anderen Test ersetzen kann, würde ich daraus ableiten, dass es dann kein nützlicher Test ist.

Fülöp Scheibler: Ja.

Frank Bengel: Meines Erachtens beißt sich da die Katze in den Schwanz. Die Frage ist: Mit welchem Verfahren vergleicht man es dann, und was sagt man zu den anderen Verfahren? Wenn man es auf die Spitze treibt – gehen wir einmal davon aus, wir haben 4 Vitalitätsdiagnostikverfahren, die derzeit in der Literatur oder auch in der Praxis eingesetzt werden: PET, Echokardiographie, MR und SPECT –, dann kommt man zu der Schlussfolgerung, dass PET keine der anderen 3 ersetzen kann, wir kommen aber auch zu der

Schlussfolgerung, wenn wir es weitertreiben, dass Echo keine der anderen 3 ersetzen kann, MR keine der anderen 3 ersetzen kann und SPECT auch nicht. Also brauchen wir gar keine Vitalitätsdiagnostik. Das wäre letzten Endes die Schlussfolgerung daraus.

Das ist das, was ich nicht verstehe, dabei habe ich ein Problem. Ich weiß aus meiner Praxis in den USA, dass ein Test dann zugelassen wird, weil er mit anderen Tests vergleichbar ist, und nicht, weil er nicht mit anderen Tests vergleichbar ist oder die ersetzen kann. Das wäre meine grundlegende Überlegung. Vielleicht sollten wir darüber erst einmal diskutieren.

Fülöp Scheibler: Man muss dazu sagen: Es ist nicht unsere Aufgabe, ein Urteil über die Erstattung zu fällen. Wir würden sagen: Die Verfahren zur Vitalitätsdiagnostik, die wir untersucht haben, kommen zu gleichen Ergebnissen. Dann müsste der G-BA beurteilen: Unter dieser Voraussetzung wird die PET von der gesetzlichen Krankenversicherung erstattet oder nicht. Aber das ist ein Job, den der G-BA macht. Es wäre ja schon ein deutliches Ergebnis, wenn wir sagen könnten: „Die PET ist genauso gut, sie hat die gleiche Sensitivität und die gleiche Spezifität wie die SPECT“ – das wäre aus unserer Sicht schon ein sehr zufriedenstellendes Ergebnis – oder: Sie ist bezüglich der Spezifität sogar besser oder schlechter oder wie auch immer.

Im Moment befürchten wir, dass das Ergebnis so schwammig wird, dass man dann gar nichts sagen kann. Weder kann man definitiv „gleich gut“ sagen noch kann man „deutlich besser“ oder „schlechter“ sagen. Denn wenn Sie sagen: „Wenn wir die PET als Indextest mit der SPECT als Referenztest vergleichen, dann ist sie besser, wenn wir sie aber mit der Echokardiographie vergleichen, dann schneidet sie wieder schlechter ab als die SPECT“, haben Sie letztendlich ein Ergebnis, aus dem Sie keine klaren Schlüsse ziehen können.

Die Arbeitsteilung zwischen IQWiG und G-BA ist – das muss man klar sagen –, dass wir eigentlich nur den Stand der Forschung darstellen und aufarbeiten. Die Erstattungsentscheidung trifft dann der G-BA.

Frank Bengel: Das mit dem G-BA habe ich verstanden, das ist klar. Mein zweiter Punkt: Ich wollte nur noch einmal die Aussage ein bisschen anzweifeln, die Sie vorhin zu dem Nutzen einer linksventrikulären Pumpfunktion als Surrogatmarker getätigt haben. Der Punkt mit der Arrhythmie und der Auswurfraction ist zwar klar, aber die Entscheidung über die Implantation von ICDs, also von Kardioverter/Defibrillator, wird heutzutage einzig und allein aufgrund der linksventrikulären Pumpfunktion getroffen. Arrhythmierisiko und Pumpfunktion zu entkoppeln, entspricht nicht der gängigen Praxis. Um das umzudrehen: Wenn es überhaupt einen generell akzeptierten Surrogatmarker in der Kardiologie gibt, dann ist es die linksventrikuläre Pumpfunktion. Das wäre meine starke Aussage dazu.

Fülöp Scheibler: Ich zitiere Sie selber: Aber das ist nicht evidenzbasiert. Ich würde Ihnen recht geben: Das ist generell als Surrogatendpunkt akzeptiert, aber dieser Surrogatendpunkt ist nicht validiert.

Frank Bengel: Aber natürlich. Es gibt doch sehr viele Studien, die zeigen, dass die Prognose das Outcome der Patienten bestimmt. Das ist doch ein Surrogatmarker. Ein Surrogatmarker ist ein Vorhersagemarker für das Outcome des Patienten.

Fülöp Scheibler: Nicht nur. Er muss nicht nur mit dem Outcome positiv korrelieren, sondern die Effekte einer Intervention auf Surrogat und Outcome sollten hoch korreliert sein.

Jörg Kotzerke: Ich möchte noch zu Herrn Sauerland Stellung nehmen. Ich kann verstehen, dass Sie gerne die vergleichenden Studien multipler Tests hätten, sodass man hinterher wirklich sauber sagen kann: Sowohl in der Beurteilung der Pumpfunktion als auch in der Beurteilung der Myokardvitalität und damit in der Vorhersage, welcher Patient von einer Revaskularisierungsmaßnahme am deutlichsten profitieren wird, ist dieser und jener Test gleichwertig. Ich habe aber in Erinnerung – Kardiologie ist nicht meine Spezialstrecke, dafür sind andere Leute hier –, dass es diese Studien nicht geben wird. Genauso wie in der Onkologie werden wünschenswerterweise randomisierte kontrollierte Studien gefordert oder erwünscht, aber sie sind nicht vorhanden. Wenn Sie das auf ein solch hohes Niveau einengen, dann werden Sie zum Schluss auch nichts herausbekommen, weil es solche Studien in der Qualität nicht gibt.

Zu Herrn Kasper würde ich gerne äußern: Ich finde es sehr erfreulich, dass Sie frühzeitig darauf zurückkommen, dass man nicht nur in der Wandbewegungszunahme oder in der Lebensverlängerung ein diagnostisches Verfahren messen muss, sondern dass für die Diagnostik eher Level 3 nach Thornbury in Betracht kommt, die Messgenauigkeit der einzelnen Tests in Sensitivität, Spezifität oder in welchen Kategorien auch immer.

Fülöp Scheibler: Ich denke, wir haben den Punkt jetzt relativ ausführlich diskutiert. Ich würde gerne festhalten wollen: Als Referenzstandard würden Sie in dem endgültigen Berichtsplan die SPECT, die MRT und die Echokardiographie beschreiben. Was habe ich noch vergessen?

(Jürgen Kasper: RNV!)

– RNV, genau. Das habe ich auch aufgeschrieben.

(Frank Bengel: Lävokardiographie!)

– Das ist die Angiographie. Ist das der Konsens? Sie wären damit zufrieden, wenn wir das so aufnehmen würden? – Gut. Es ist mir wichtig, dass wir das für uns so festhalten können.

Stefan Sauerland: Ich möchte noch kurz Herrn Kotzerke entgegnen, dass es die vergleichenden Studien natürlich gibt, dass solche Studien durchaus denkbar sind. Wir reden nicht von randomisierten Studien, wir reden von Testgütestudien. Das, worum es uns im Vergleich geht, ist allein der Indextest. Wir wollen nicht, dass in der Referenzdiagnostik, also in der Beurteilung der myokardialen Wandbewegungen, noch irgendwelche Tests verglichen

werden, sondern wir wollen, dass die PET mit einem anderen Test verglichen wird. Es gibt zum Beispiel aus Australien einen Bericht aus 2010. Dort haben sich mehrere Studien mit dieser Frage beschäftigt, die in den Bericht der australischen Kollegen eingeschlossen wurden. Die PET wurde zum Beispiel mit SPECT, aber auch mit anderen Verfahren verglichen. Das Argument, dass diese Studien nicht vorhanden seien oder nicht durchführbar seien, kann ich nicht nachvollziehen.

Frank Bengel: Da stimme ich Ihnen zu. Es gibt solche Vergleichsstudien. Ich wollte nur noch einmal den Punkt streifen, den ich vorhin genannt habe: Man muss bei den Vergleichsstudien immer aufpassen, dass nicht das Verfahren, das nach der Revaskularisation zur Messung der Pumpfunktion verwendet worden ist, identisch ist mit dem Verfahren, das vorher zur Vitalitätsdiagnostik verwendet worden ist, weil das ein Bias für das Verfahren gibt.

Fülöp Scheibler: Das steht im Berichtsplan.

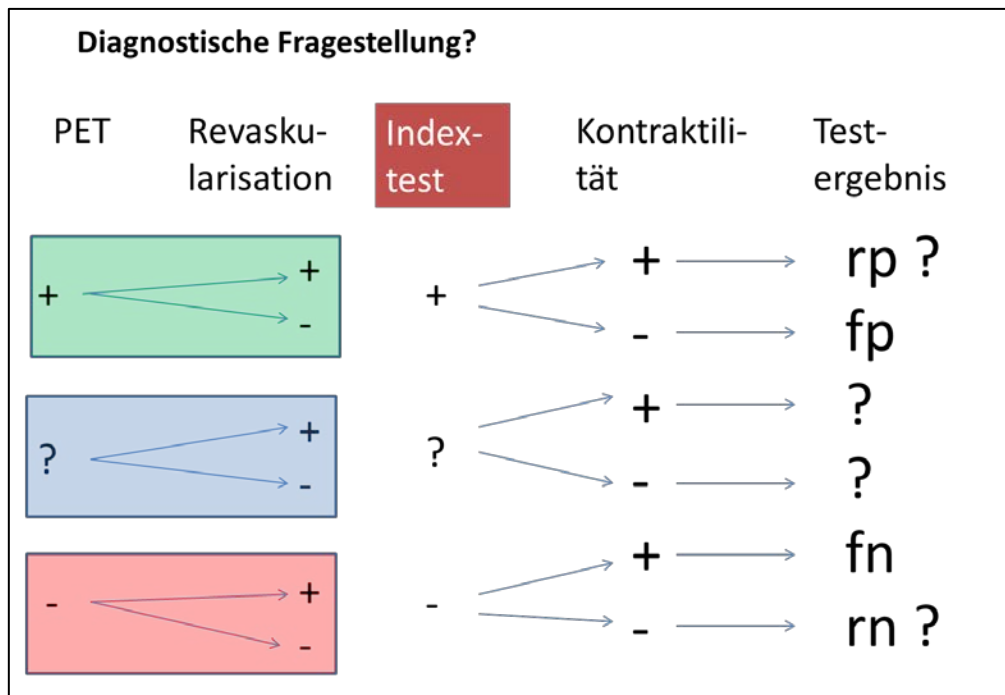
Moderator Jürgen Windeler: Dann scheint über die Frage des Studiendesigns, der Notwendigkeit, an dieser Stelle vergleichende Studien zu haben – auch wegen der nicht analysierten Folgebedingungen einarmiger Studien –, über die Frage, welches die Indextests, welches denkbare Referenztests sind, im Wesentlichen Einigkeit zu bestehen. Das ist mein Eindruck. – Dann können wir TOP 1 verlassen.

Wir kommen zu Tagesordnungspunkt 2.

4.4.3 Tagesordnungspunkt 2: Sind diagnostische Studien überhaupt sinnvoll / interpretierbar (Treatment paradox)?

Fülöp Scheibler: Wir wollen die Sache nicht unnötig verkomplizieren, aber auch das ist eine Schwierigkeit, mit der wir uns herumschlagen. Dazu brauchen wir auch Ihre Einschätzung, wie Sie das sehen.

Bei dem vorliegenden Thema besteht insbesondere das Problem, dass zwischen dem Indextest, also der PET zum Beispiel, und dem Referenzstandard, beispielsweise der MRT vor und nach Revaskularisation eine Intervention durchgeführt wird. Das führt nach unserer Auffassung zu Ergebnissen, die teilweise nicht interpretierbar bzw. schwer zwischen den Studien vergleichbar sind. Wir haben versucht, das auf einer Folie darzustellen.



Das ist die Ausgangssituation. PET sagt: Entweder es ist eine Revaskularisation möglich, oder sie ist nicht möglich. Aber es gibt auch die Fälle – das haben wir in der Literatur gesehen –, wo das PET-Ergebnis nicht ganz eindeutig ist; die lassen wir erst einmal außen vor. Dann findet die Revaskularisation statt. Es gibt aber auch Fälle, dass trotz der positiven Vorhersage der PET keine Revaskularisation durchgeführt wird, und es gibt Fälle, wo sie zwar durchgeführt wird, aber nicht erfolgreich ist. Das haben wir erst einmal unter dem negativen Ausgang zusammengefasst. Der Idealfall wäre: Die PET sagt es voraus, die Revaskularisation findet statt und ist erfolgreich, und am Ende hat sich auch die Kontraktilität verbessert. Dann könnte man von einem richtig positiven Ergebnis sprechen.

Analog dazu ist es im Negativen: Wenn nicht revaskularisiert wird und sich keine verbesserte Kontraktilität zeigt, dann würde man von richtig negativen Testergebnissen sprechen.

Wolfgang Burchert: Nein.

Fülöp Scheibler: Würden Sie nicht?

Wolfgang Burchert: Wenn die Vitalität nicht vorhergesagt wird, dann müssen Sie revaskularisieren. Wenn es sich dann nicht bewegt, ist es richtig negativ.

Fülöp Scheibler: Dann sind wir hier. Danach bewegt sich nichts. Dann sind wir hier.

Wolfgang Burchert: Ja, genau.

Fülöp Scheibler: Und wie würden Sie diese Fälle interpretieren? Die PET sagt: Nein, es wird auch nicht revaskularisiert.

Wolfgang Burchert: Dann werden Sie es nie herauskriegen.

Jörg Kotzerke: Die fallen heraus.

Fülöp Scheibler: Das sind dann unklare Befunde. – Auch einverstanden.

Wolfgang Burchert: Dass wir revaskularisieren, ist ja Voraussetzung. Das Plus/Minus ist da nicht sehr sinnvoll, sondern man muss eigentlich eine Koppelung haben. Die Beobachtungsstudien zeigen häufig, dass man retrospektive Datenanalysen macht, anhand derer man dann sieht: Soundso häufig ist revaskularisiert worden, das ist vorhergesagt worden.

Wenn man das aber prospektiv sieht, dann muss immer revaskularisiert werden. Diejenigen, die technisch nicht revaskularisiert werden, sind eigentlich Dropouts. Die sind weg, die kann man nicht beurteilen. Da ist es an anderen Dingen gescheitert, nicht an der Diagnostik.

Fülöp Scheibler: Das ist jetzt ein bisschen die Frage. Diejenigen, die ausfallen und nicht revaskularisiert werden, sind eben nicht nur die Fälle, die technisch nicht möglich sind, sondern es gibt auch andere Gründe. Wir sind da sehr pingelig. Wir sagen: Jeder Patient muss beurteilt werden. Wir würden auch in der Onkologie nicht akzeptieren, dass wir bestimmte Fälle, die zum Beispiel keine Nachbeobachtungsdaten haben, herausrechnen, sondern die müssen irgendwie berücksichtigt werden.

Unser Vorschlag war eine Art Intention-to-Treat-Analyse, dass man sagt: Egal wie die Revaskularisation durchgeführt, nicht durchgeführt, ausgegangen, nicht ausgegangen ist, wir zählen alle PET-Positiven zu den Indextest-Positiven, und wir zählen alle PET-Negativen zu den Indextest-Negativen. Dann müssen wir nur noch dieses Ergebnis mit dem Ausgang des Referenzstandards vergleichen und hätten die 4 Felder.

Wolfgang Burchert: Für meinen Geschmack können Sie nur diejenigen vergleichen, bei denen vorher „positiv“ oder „negativ“ gesagt worden ist. Dann haben Sie einen Zweig, wo die Revaskularisation gelungen ist, und man kann gucken, ob die Vorhersage positiv oder negativ war. Dann gibt es noch den Zweig, wo nicht revaskularisiert worden ist. Da können Sie gucken: War die Vorhersage „positiv“ oder „negativ“ nachher auch in der Gruppe zu sehen, wenn nicht revaskularisiert worden ist? Denn theoretisch müssten die, wenn Vitalität da ist, eine schlechtere Prognose haben.

Fülöp Scheibler: Dann hätte man aus methodisch-diagnostischer Perspektive das große Problem, dass man eine sehr hohe Anzahl an nicht interpretierbaren Ergebnissen oder Dropouts hätte. Sie hätten dann quasi eine präzise diagnostische Güte, aber nur für den Bruchteil der ursprünglich eingeschlossenen Patienten. – Das wäre die eine Variante.

Aber selbst da sehen wir ein Problem. Bei den sogenannten richtig Positiven wüsste man ja eigentlich nur, ob sie wirklich richtig positiv sind, wenn man randomisiert hätte. Die PET sagt

Ja, und jetzt müsste man eigentlich zwischen Revaskularisation und Nichtrevaskularisation, zum Beispiel medikamentöser Therapie, randomisieren. Wenn dann das Ergebnis positiv ist bei denen, die revaskularisiert wurden, und negativ bei denen, die nicht revaskularisiert wurden, dann könnte man eindeutig sagen: Ja, richtig positiv. Aber so haben Sie bei denen immer das Problem: Sie hätten vielleicht auch bei rein medikamentöser Behandlung eine Differenz der Kontraktilität gehabt.

Wolfgang Burchert: Aber das ist schwierig. Bei der Patientengruppe, die wir anschauen, die eine niedrige EF hat, wird man ethisch gar nicht vertreten können, von vornherein zu sagen: Wir randomisieren auf keine Revaskularisation. Das – ich bin zufällig Mitglied einer Ethikkommission – würden wir nie im Leben durchlassen. In dem Bereich des etwas niedrigeren Risikos, im COURAGE Trial, hat man tatsächlich Revaskularisierung randomisiert. Das sind aber alles Patienten, wenn Sie sich die Kurven anschauen, die wahrscheinlich eine sehr gute EF haben, und sie haben, wenn man sich die Absterberaten ansieht, eine sehr leichte Form der KHK. Die Subgruppenanalyse der COURAGE Trial zeigt: Bei einer deutlichen Ischämie profitieren sie schon von der Revaskularisierung. Das heißt, da ist man auf einer ganz anderen Ebene des Risikos für Patienten.

Bei Patienten mit einer eingeschränkten linksventrikulären Funktion, die wir betrachten, würde ich mich schwertun, das einfach zu randomisieren. Dazu wird es in der Form keine Studien geben.

Fülöp Scheibler: Ich wollte auch nicht sagen, dass man das so machen sollte.

Wolfgang Burchert: Das ist eigentlich keine valide Option.

Fülöp Scheibler: Das eine ist ja, darüber zu sprechen: Wie sollten Studien zukünftig ausgestaltet sein? Sind sie ethisch vertretbar oder nicht? Die andere Frage, die mich umtreibt, ist, dass bei den Ergebnissen, die wir aus einarmigen Beobachtungsstudien erzielen, immer das Fragezeichen offenbleibt: Hätte der hier möglicherweise auch ohne Revaskularisation eine Kontraktilität erzielt?

Noch eine kurze Rückfrage zu STICH: Ist das nicht genau die Frage der Revaskularisation bei hohem Risiko versus medikamentöse Behandlung gewesen?

Wolfgang Burchert: Ja.

Fülöp Scheibler: Die ist doch ethisch unbedenklich, oder?

Anselm Bäumer: Nein, EF durchschnittlich 26 %, also die Pumpleistung des Herzens.

Fülöp Scheibler: In diesem Zusammenhang halte ich die STICH-Studie schon für relevant für unsere Fragestellung.

Frank Bengel: Wir müssen die STICH-Studie eben doch diskutieren. Eines der Probleme der STICH-Studie, die über Jahre bis fast Jahrzehnte gelaufen ist, ist unter anderem genau der Punkt gewesen, dass es sehr schwer war, überhaupt Patienten einzuschließen. Wenn man auf die Anzahl der beteiligten Zentren bei der STICH-Studie schaut, dann sieht man, dass zwar weltweit Zentren eingeschlossen waren, aber teilweise sehr geringe Patientenzahlen, eben weil es so schwierig war, dem Studienprotokoll zu folgen. Genau deswegen wird die STICH-Studie auch so kontrovers diskutiert.

Viele Leute, die die STICH-Studie kritisieren – nicht alle kritisieren sie –, sagen, dass die mangelnden Effekte, die dort gemessen wurden, schlicht und einfach daher kommen, dass dort eine Vielzahl von Zentren eingeschlossen war, mit weltweit sehr unterschiedlichen Standards und auch weltweit sehr unterschiedlichen Praktiken, was letzten Endes dazu führt, dass jeder potenzielle Effekt verwaschen wird. Deswegen sind keine signifikanten Unterschiede zwischen den einzelnen Armen herausgekommen. Das ist ein Kritikpunkt, der in der STICH-Studie angeführt worden ist, der meines Erachtens unter anderem dazu geführt hat, dass die Vitalitätsdiagnostik dort, die im Übrigen mit Echo und SPECT gemischt durchgeführt worden ist, keinerlei signifikante Effekte gezeigt hat.

Aber ich glaube, um es noch einmal auf den Punkt zu bringen, das Problem ist: Wie beurteilt man einen diagnostischen Test? Denn der Test selber beeinflusst das Outcome eigentlich nicht, sondern die Therapie hintendran.

Fülöp Scheibler: Genau.

Moderator Jürgen Windeler: Wobei wir dann im Moment, wenn wir die Revaskularisation als Therapie verstehen, bei dem Design sind, das auch Sie favorisiert haben, bei dem man sagt: Ich muss alle revaskularisieren, damit ich sehe, was da passiert. Damit sind wir noch nicht bei der Therapie oder bei der Therapieentscheidung. Wir sind nur dabei, zu schauen, ob das eine prognostische Bedeutung hat.

Frank Bengel: Ja.

Moderator Jürgen Windeler: Das Design beantwortet die Frage, wenn auch vermutlich mit einem nicht unerheblichen Bias, weil ich bestimmte Patienten weglasse. Das Bias muss man sich dann eventuell ansehen. Das große Problem ist aber, dass man aus dieser prognostischen Studie nicht wirklich weiß, ob man von PET die Revaskularisationsentscheidung abhängig machen sollte. Das weiß man dann noch nicht. Das ist im Grunde eine Frage, die man – das war das, was Herr Scheibler beschrieben hat – beantworten müsste, wenn man weiß: PET ist ein – vermutlich nicht perfekter – gewisser prognostischer Marker mit einem bestimmten prognostischen Sinn. Der Einsatz von PET macht überhaupt keinen Sinn, wenn man sowieso alle Patienten revaskularisiert. Dann ist es zwar ein interessanter prognostischer Marker, aber nicht spannend.

Spannend wird es nur dann, wenn man die Entscheidung zur Revaskularisation vom PET-Ergebnis abhängig macht. Ob das für die Patienten sinnvoll ist, ist durch den Prognosecharakter noch nicht gezeigt. Das ist das Gesamtproblem, vor dem wir stehen. Insofern sind wir noch nicht in der Situation, uns den Entscheidungscharakter, den Entscheidungsbeitrag von PET anzusehen, sondern wir sind erst einmal nur auf der Ebene der Prognose. Auch die Diskussion vorhin zu TOP 1 hat sich eigentlich nur um den Punkt gedreht: Können wir die Prognose – eine Diagnose ist es hier ja nicht – von PET bejahen, auch im Vergleich zu anderen? Die Frage, ob das entscheidungsrelevant eingesetzt werden sollte, um davon die Therapien abhängig zu machen, beantworten uns die Studien noch nicht. – Sie nicken?

Frank Bengel: Im Endeffekt – das ist das, was ich gemeint habe – sind wir mit der Beurteilbarkeit, was die Wertigkeit eines diagnostischen Tests betrifft, sehr stark abhängig von der Therapie, die gemacht wird. Das ist genauso, wie Sie es gerade gesagt haben. Wenn man heutzutage alle revaskularisieren müsste, dann bräuchten wir den diagnostischen Test nicht – das ist ganz klar – oder wenn man alle nicht revaskularisieren müsste ...

Moderator Jürgen Windeler: Genauso.

Frank Bengel: ... so wie es die STICH-Studie sagt. Dann brauchen wir den diagnostischen Test nicht.

(Heiterkeit)

Stefan Sauerland: Ich glaube, dieser Punkt ist essenziell. Man unterscheidet: Ist die PET als prognostischer Marker eingesetzt worden oder als prädiktives, also therapieleitendes Werkzeug verwendet worden? In den Studien wird es sich überwiegend nur um den ersten Schritt handeln. Wir schauen erst einmal, ob die PET prognostisch irgendwelche Informationen liefern kann.

Dann kommt die Frage: Erhalten alle Patienten eine revaskularisierende Maßnahme oder keiner? Das halte ich für nicht so wahrscheinlich. Aber es wird auch Studien geben, in denen ein Teil der Patienten, vielleicht sogar aufgrund des PET-Ergebnisses, revaskularisiert wird, während die anderen nicht revaskularisiert werden. Dann habe ich das Problem einer sich selbst erfüllenden Prophezeiung und sehe: Bei diesen Patienten hat die PET etwas gezeigt. Im Follow-up werden die Lebensqualität, die linksventrikuläre Funktion, die Mortalität besser. Das macht es schwierig. Man muss die beiden Elemente, Diagnose und Therapie, möglichst gleich halten und gleichartig einsetzen, sodass man den Anteil der einzelnen Komponenten am Gesamttherapieerfolg beurteilen kann.

Fülöp Scheibler: Ein weiteres methodisches Problem, das ich sehe, ist: Ich habe einen spezifischen Auswertungsalgorithmus vorgeschlagen. Man könnte das aber auch anders machen. Man könnte, wie Sie vorgeschlagen haben, nur die Fälle, bei denen die Revaskularisation tatsächlich erfolgt ist, auswerten und mit dem Referenzstandard

vergleichen. Man könnte noch einen Schritt weitergehen und sagen: Nur bei erfolgter und erfolgreicher Revaskularisation – auch solche Studien habe ich gesehen – wird überhaupt mit dem Referenztest verglichen.

Wieder aus methodischer Sicht stellt sich mir aber das Problem: Ich habe 2 Studien, die denselben Indextest, denselben Referenzstandard benutzt haben. Die einen verwenden den Auswertungsalgorithmus, die anderen einen anderen und kommen zu unterschiedlichen Ergebnissen. Wie soll ich solche Studien poolen? Die Autoren müssten eigentlich beschreiben, welchen Auswertungsalgorithmus sie verwendet haben. Das wäre eine Möglichkeit. Dann könnte ich sagen: Ich sortiere die Studie nach Auswertungsalgorithmen, nach Indextests, nach Referenztests und versuche dann zu poolen. Oder aber – das gibt es auch manchmal, wenn es nicht sehr viele Patienten sind – man gibt wirklich pro Patient in der Studie an: Was war das Ergebnis des Indextests? Fand eine Revaskularisation statt? Welche? War sie erfolgreich oder nicht? Was war das Ergebnis des Referenzstandards? Dann könnte ich meinen Algorithmus auf diese Daten anwenden, so wie er mir passt.

Erfahrungsgemäß – zugegeben, aus den onkologischen Projekten – werden wir so detaillierte Daten in den wenigsten Fällen tatsächlich vorfinden. Also schlagen wir uns wieder mit einem rein methodischen Problem herum. Wenn wir wissen, dass aufgrund der Problematik der Kombination aus Indextest und Intervention unterschiedliche Algorithmen möglich sind, mit denen man auswertet, dann weiß man: Es wird sehr schwierig sein, 2 Studien miteinander zu vergleichen.

Moderator Jürgen Windeler: Wobei diese Schwierigkeit nicht Ihre und auch nicht unsere Sorge sein muss. Der entscheidende Punkt dabei ist, dass diese Schwierigkeit und die Heterogenität, die wir in TOP 1 schon diskutiert haben, dazu führen können, dass wir nach der Sichtung, Interpretation und Dokumentation der vielen Studien am Ende mit den Schultern zucken und sagen werden: Das ist alles ganz nett, schöne viele Daten, große Sammlung, Bias hier, Bias da, damit kann man nichts anfangen.

Die Frage ist daher – deswegen thematisieren wir das so ausführlich –, welchen Stellenwert Sie den Aussagen in der Bewertung eines solchen Verfahrens wie PET beimessen würden. Wir werden absehbar – sicher kann man nicht sein – in einem Dschungel von vielfältigen Konstellationen landen – wir haben mindestens eine fünf- oder sechsdimensionale Matrix mit Konstellationen aufgebaut – und haben die große Sorge, dass wir im Grunde keine belastbaren und guten Aussagen machen können. Diese werden dann – meine Wortmeldung von vorhin – auch nur etwas über die Prognose der PET-Informationen sagen, aber noch nichts über die Entscheidungsrelevanz. Deswegen noch einmal die Frage an Sie: Welche Bedeutung hätte solch ein Ergebnis?

Wolfgang Burchert: Können Sie das etwas spezifizieren? In welcher Hinsicht meinen Sie das? Welche Bedeutung hat das? Sie erstellen das Gutachten ja nicht für uns.

Moderator Jürgen Windeler: Nein, aber auch für Sie insofern, als dass Sie es lesen und sich eine Meinung bilden, indem Sie es dann auch kommentieren werden. Wir erstellen es vielleicht nicht für Sie, aber Sie haben sich selber aufgrund von Daten, von Studien, von Evidenz eine Meinung gebildet. Die Frage ist: Welche Rolle spielt bei dieser Meinung dieses nach unserer Auffassung im Moment sehr vielfältige Problem?

Um noch einmal auf den Punkt vom Anfang zurückzukommen: Sie haben mehrere Indextests, Sie haben mehrere Referenztests mit unterschiedlichen Kombinationen und einem unterschiedlichen Case Mix, sodass Sie an der Stelle wahrscheinlich schon keine vernünftigen Aussagen herausbekommen werden.

Sie haben hier zusätzlich ein hohes Biaspotenzial, indem Sie möglichst viele revascularisieren sollten, aber einige wahrscheinlich nicht werden, und wir wissen nicht genau, warum, möglicherweise ergebnisabhängig von PET. Dann bricht die ganze Prognose schon in sich zusammen.

Wir hatten noch mindestens 2 Dimensionen mehr, über die wir jetzt diskutiert haben, wo wir am Ende unterschiedliche Algorithmen haben, was die Richtig- und die Falsch-Positiven angeht. Somit haben wir die Befürchtung, dass wir zu der Aussage: „Ist PET im Vergleich zu anderen Verfahren ein geeignetes prognostisches Verfahren?“ wenig sagen können, vermutlich noch weniger als wenig und dann noch gar nicht die Frage beantwortet haben: Ist PET an dieser Stelle als Entscheidungskriterium einzusetzen? Da es, wie ich gerade noch einmal gehört habe, möglicherweise Studien gibt, die die Frage der Entscheidungsunterstützung beantworten oder versuchen zu adressieren, ist die Frage, ob man die nicht sehr in den Vordergrund setzen müsste.

Jörg Kotzerke: Sie führen das für die PET aus als den Test, der Entscheidungshilfe oder prognostische Informationen geben soll. Das gilt aber genauso ...

Moderator Jürgen Windeler: Für die anderen Indextests genauso.

Jörg Kotzerke: ... für die anderen Tests, sodass sich vermutlich auch für keinen der anderen Tests ein Benefit, ein Patientennutzen wird nachweisen lassen. Die methodische Schwierigkeit liegt für alle anderen Tests ganz genauso vor.

Fülöp Scheibler: Ja.

Jürgen Kasper: Ich möchte das noch präzisieren. So wie ich Herrn Windeler verstanden habe, geht es in seinen Ausführungen um die methodische Schwierigkeit der diagnostischen und der Prognosestudien und damit um die Frage, ob nicht die methodisch weniger schwierigen, also die einfacheren Studien zum Nutzen der Verfahren in den Vordergrund gerückt werden sollten. Das ist es, oder?

Moderator Jürgen Windeler: Absolut. Noch einmal: Wenn ich den Unterschied zwischen Prognose und Diagnose – hier geht es um Prognose – von der Frage trenne: „Ist dieses Verfahren entscheidungsrelevant, entscheidungsleitend mit den daraus resultierenden positiven Nutzenkonsequenzen für die Patienten?“, dann reden wir im Moment nur über die erste Schiene: Diagnostik/Prognose. Da taucht das Problem auf. Bei der Frage des Nutzens – ich kann das nicht beurteilen – könnte das Problem geringer werden.

Noch einmal: Wenn wir – auch das haben wir unter TOP 1 schon besprochen – für den Vergleich von Indextests und deren Relation, was die prognostische Information angeht, auf einen einheitlichen Referenztest mit angemessenen Vergleichen gehen würden, dann ist das methodische Problem, abgesehen von solchen Dingen wie Erfolgskriterien, vielleicht noch lösbar. Aber wenn wir Indextests mit der Vielfalt von Referenztests und mit weiteren vielfältigen Dimensionen vergleichen, ist meine Prognose, dass wir über die prognostische Qualität von PET und, in der Tat, den Konkurrenzverfahren genauso wenig werden sagen können.

Deswegen meine Frage: Wenn wir meinen, wenig sagen zu können, dann könnten Sie und andere Fachleute am Ende sagen: „Das ist aber ganz furchtbar wichtig“, oder Sie als Fachleute könnten sagen: Das ist doch gar nicht so interessant. – Dann würden wir die Entscheidung vielleicht davon abhängig machen, ob wir uns sehr intensiv mit dem Thema befassen oder nicht. Wenn Sie sagen: „Das ist die Information, die wir brauchen“, dann werden wir uns vielleicht eher in die Tiefen der Vielfalt stürzen, als wenn Sie sagen: Nein, das brauchen wir überhaupt nicht, sondern uns interessiert eher die Frage, ob es entscheidungsleitend ist und was die Patienten davon haben, als die Frage der Prognose.

Wolfgang Burchert: Ich denke schon, da hier im Grunde alles von vorne aufgerollt wird. Wie gesagt, das Verfahren ist ja im Einsatz. Wir haben auch MR, wenn wir uns aber die Vitalität ansehen, nehmen wir PET. Die Schwierigkeit im Alltag ist: Wir haben hier so schön „Plus/Minus“ stehen. Das sind zwei diskrete Zustände, dazwischen gibt es nichts. Die typische Anmeldung ist: Der Patient hat eine massive Dreifäßerkrankung, die linksventrikuläre Funktion ist eingeschränkt, aber er hat zum Beispiel schon einen Bypass gehabt. Es geht darum, noch einen zweiten Bypass draufzusetzen. Dann fragt man: Macht der ganze Anlauf, den wir hier nehmen, überhaupt Sinn, weil beispielsweise die Risiken einer Revaskularisation sehr hoch sind? Dann wird geschaut, und man sagt: Es sind noch große Bereiche vital, da muss man eben auf Teufel komm raus versuchen, irgendwie zu revaskularisieren. Dann wird er eben operiert. Das sind typische klinische Szenarien.

Es wird ja nicht bei jedem Patienten, der hereinkommt, sofort gesagt: $EF < 35 \%$, wir machen einen Vitalitätstest. – So ist das nicht, sondern das hängt von der Prozedur und von dem Zustand ab. Beim Zuhören hatte ich jetzt ein paar Probleme damit. Ich weiß, dass es für die Analyse schwierig ist, aber wir setzen es ja regelmäßig ein. Ich kann nur sagen: Bei uns operieren die Chirurgen danach. Daher – das ist jetzt subjektiv, niedriger Evidenzlevel, alles,

was dazugehört – habe ich die Schwierigkeit: Wenn man das sehr stark formalisiert, dann sind all diese Dinge, die tatsächlich die klinische Praxis ausmachen, nicht drin.

Es gibt eine Studie von Herrn Beanlands aus Kanada, der zunächst einfach eine methodische Studie aufgesetzt hat, dann eine multizentrische Studie aufgesetzt hat, und dann ist alles auseinandergebrochen, weil die Therapeuten nicht mehr pariert und das Studienprotokoll nicht mehr eingehalten haben. Da, wo er selber die Kontrolle darüber hatte, hat es noch funktioniert. Das ist etwas ganz Typisches. Wenn Sie fragen: „Wie stark beeinflusst es die therapeutische Entscheidung?“, dann geht das sehr ins Philosophische. Wie sehr kann man Chirurgen überhaupt beeinflussen? Wie sehr kann man Kardiologen beeinflussen? Zum Beispiel: Wann übergibt der Kardiologe den Patienten mit den kaputten Koronarien an den Chirurgen und sagt: „Mach du weiter“? Das sind Dinge, die die klinische Versorgungsrealität darstellen, mit der wir arbeiten müssen.

Sie müssen das nicht beurteilen, das ist klar, aber deswegen habe ich ein bisschen Bauschmerzen damit, größere Teile wegzulassen. Irgendwann werden die Fragen wieder zurückkommen, und es heißt: Geht das denn überhaupt? Das ist die Schwierigkeit, weil es schon sehr lange in Betrieb ist. Zu den grundsätzlichen Fragen gibt es aus letzter Zeit auch keine Studien mehr, weil eigentlich klar ist, dass es geht. Das müssen Sie jetzt aber wieder neu beweisen. Deswegen müsste man dabei die Fragen der Prognose mit aufarbeiten, weil jetzt im Grunde alles wieder infrage gestellt wird. Ich denke nicht, dass man es rein auf den Nutzen reduzieren kann – das wäre schön, aber ein bisschen zu einfach –, da wir jetzt, im Jahre 2012, im Grunde wieder ganz neu anfangen, über PET als Vitalitätsdiagnostik zu reden, was seit über 20 Jahren eingeführt ist.

Anselm Bäumer: Sie sehen das natürlich aus der Sicht aus Bad Oeynhausen. Man muss dazusagen, dass deutschlandweit viele operative Zentren kardiochirurgisch seit langer Zeit ohne PET arbeiten und das auch in Zukunft tun werden.

Wolfgang Burchert: Ja, genau.

Anselm Bäumer: Und der diagnostische Standard in der Entscheidungsfindung ist üblicherweise SPECT.

Wolfgang Burchert: Unter anderem. Auch MR, auch Echo.

Anselm Bäumer: Auch Stressecho, genau.

Wolfgang Burchert: Das ist sehr vielfältig und hängt auch mit der Verfügbarkeit zusammen.

Frank Bengel: Wir kommen dann letzten Endes immer zu der gleichen Schlussfolgerung, genauso wie es Herr Kotzerke gesagt hat. Für keinen diagnostischen Test ist irgendetwas aus dem ganzen Wust herauszuholen.

Ich will trotzdem noch einmal auf das, was Sie vorhin unter TOP 1 vorgeschlagen haben, eingehen. Sie haben gesagt: Wir müssen uns vielleicht auf einen Referenzstandard beschränken, um hier ein mögliches Bias herauszunehmen. Ich werde konkreter: Ich bin der Meinung, dass Sie kein Bias introduzieren, wenn Sie die ganzen Tests, die wir hier vorgeschlagen haben, als Referenzstandards nehmen. Ich bin der Meinung, dass eine MRT geeignet ist, um eine linksventrikuläre Auswurfraction zu messen, dass eine Echokardiographie, eine PET, eine SPECT und eine Lävokardiographie dazu geeignet sind. Unter den Bedingungen, zu denen die Baseline nach Revaskularisation verwendet worden ist, ist meines Erachtens das eine durch das andere prinzipiell ersetzbar und deswegen auch zu poolen, ohne dass man sagt: Es ist ein Bias. – Wenn Sie das machen, werden Sie letzten Endes ausreichend Power erreichen, um festzulegen, dass dieser Test – und wahrscheinlich auch andere diagnostische Tests – tatsächlich eine Wertigkeit hat, um eine Verbesserung der linksventrikulären Auswurfractionen vorherzusagen.

Wenn Sie natürlich sagen: „Sie können den einen Test nicht zusammen mit dem anderen poolen, weil da ein Bias hineinkommt“, dann muss man alles auseinanderdividieren. Dann kommt ein zusätzlicher Komplexitätslevel hinein, und man kommt am Ende zu der Schlussfolgerung, dass man gar nichts sagen kann. Das wäre noch einmal mein konkreter Punkt. Ich bin der Meinung, dass man alle zusammen poolen kann. Ich bin nicht der Meinung, dass es ein Bias gibt.

Anselm Bäumer: Zu der Beurteilung der LV-Funktion gibt es auch Leitlinien und das 17-Segment-Modell der American Heart Association, das den methodischen Ansätzen schon als Leitlinie zugrunde gelegt werden muss, zum Beispiel die Lävokardiographie als ganz alte Maßnahme mit 2 Ebenen. Da kann man keine Segmentbeurteilung durchführen. Insofern ist das qualitativ sicherlich kein vergleichbares Verfahren zu den genannten.

Die Radionuklidventrikulographie zeigt natürlich die Wand selber und die Kontraktilitätszunahme nicht, also die Kontraktilitätsreserve. Da sieht man einen Ausguss im Volumenpräparat des Herzens. Das wäre in die Richtung gehend, dass die Segmentbeurteilung möglich ist.

Frank Bengel: Aber Sie haben ja einen ganz guten Vorschlag gemacht, nämlich dass man sich nach akzeptierten Modellen umsieht für die Systematik, dass man sich weniger auf eine Technik versus die andere konzentriert, sondern mehr darauf, dass die linksventrikuläre Pumpfunktion nach einem gewissen standardisierten, den Richtlinien entsprechenden Algorithmus berichtet oder analysiert wird. Das würde ich für einen sinnvolleren Ansatz halten, als einzelne Methoden als Grund für ein potenzielles Bias herauszuholen.

Moderator Jürgen Windeler: Ich kann das mit dem Referenzstandard gut nachvollziehen, aber immer unter der Grundannahme – es scheint mir so zu sein, dass die Grundannahme auch niemand wirklich verifizieren kann –, dass die Verfahren gleich sind. Denn sonst kommen wir im Ergebnis in die Situation, Sie haben es vorhin plakativ gesagt: A mit B und C

mit D. – So einfach wird es nicht sein, aber es kann schon sein, dass mit A schwerpunktmäßig B und mit C schwerpunktmäßig D korreliert. Wenn die dann nicht gleich sind, werden wir unterschiedliche Ergebnisse bekommen. Das macht nichts, aber ich sehe uns dann schon in der Erörterung des Vorberichts hier sitzen. Bestimmte Diskussionen möchte ich möglichst vermeiden.

Fülöp Scheibler: Es mag ja sein, dass das aus klinischer Sicht gleichwertig ist, aber aus methodischer Sicht sehe ich da nach wie vor ein Problem. Sie haben selber geschrieben: Um das Inkorporationsbias zu vermeiden, dass sozusagen die PET als Indextest mit der PET als Referenzstandard verglichen wird und damit womöglich noch ein weiterer Indextest – das wäre ein unfairer Vergleich, auch die PET mit sich selber zu vergleichen, kann nicht zu validen Ergebnissen führen –, bedarf es der Möglichkeit, Referenztests untereinander abzugrenzen. Wenn ich zum Beispiel die MRT für die Vitalitätsdiagnostik als Indextest benutze, dann darf die MRT kein Referenztest mehr sein, aber möglicherweise darf die PET in der Studie Referenztest sein. Dann wird sie zirkulär. Wenn wir das alles akzeptieren, kann ja nur noch das Gleiche herauskommen.

Wir müssen uns entscheiden, was als Referenztest und was als Indextest möglich ist, oder wir sagen: Anything goes. – Dabei wird dann aber auch nichts herauskommen. In einem Projekt zu Sprachentwicklungsstörungen hat sich einmal herausgestellt, dass der Test – das war ein Elternfragebogen – anhand eines klinischen Tests, der durch Kliniker durchgeführt wurde, validiert war. Dann haben wir geguckt, wie er denn validiert war, und festgestellt, dass er an einer Vorläuferversion des Elternfragebogens validiert wurde. Irgendwann beißt sich dann die Katze in den Schwanz.

Ich meine, das Gleiche gilt hier auch. Wenn ich versuche, die diagnostische Güte der MRT als Indextest anhand der PET herauszufinden, und in einer anderen Studie versuche ich, die diagnostische Güte der PET anhand der MRT als Referenzstandard herauszufinden, dann kann dabei nichts Vernünftiges herauskommen. Wir müssen uns schon darauf einigen: Was sind die Indextests, und was sind die Referenzstandards?

Wolfgang Burchert: Ich glaube, das Problem, das Sie gerade genannt haben, existiert nicht, denn Sie messen ja unterschiedliche Parameter. Sie messen Vitalität versus Wandbewegung, Auswurfraction oder was auch immer. Das sind schon unterschiedliche Parameter. Ich würde nicht sagen, dass die Variabilität in der Bestimmung der Auswurfraction bei PET genauso ist wie bei MR und es deshalb nivelliert. Das ist nicht so. Ich denke, das Problem gibt es nicht.

Wenn Sie fragen, wie die untereinander validieren, ob sie voneinander abweichen und vergleichen jeden mit jedem, dann wissen Sie auch nicht genau: Wo liegt die Wahrheit? Man müsste theoretisch ein Phantom oder sonst etwas messen. Das ist schwierig. Das Problem ist letztlich weder mit MR noch mit CT gelöst, obwohl beide Geräte eine extrem hohe morphologische Auflösung haben, CT eigentlich noch besser, außerdem isotrope Voxel

generiert, was die Volumenbetrachtung angeht. Aber auch die Verfahren haben eine gar nicht so geringe Fehlerspannbreite.

Moderator Jürgen Windeler: Ich habe jetzt den Eindruck, dass wir uns auf dem asymptotischen Ende einer Klärung befinden, noch nicht ganz am Ende sind, aber den Rest werden wir dann intern diskutieren. Wenn Sie einverstanden sind und es keine Wortmeldungen mehr gibt, dann würde ich TOP 2 beenden.

Wir kommen zu TOP 3. – Herr Scheibler.

4.4.4 Tagesordnungspunkt 3: Ausschluss von (vergleichenden) Studien mit Patienten mit nur geringfügig eingeschränkter regionaler und globaler linksventrikulärer Funktion?

Fülöp Scheibler: In Ihrer Stellungnahme schlagen Sie vor, nur Studien einzuschließen, die Patienten mit einer linksventrikulären Auswurffraktion von 35 % oder weniger untersuchen. Im Berichtsplan wird das Kriterium der Schwere der Erkrankung im Rahmen der prospektiven Festlegung von Subgruppenanalysen festgelegt. Spricht aus Ihrer Sicht etwas dagegen, zunächst alle vorhandenen Studien in die Analysen einzubeziehen und anschließend Subgruppen zu bilden?

Frank Bengel: Wie wollen Sie das machen? Wenn Sie eine Studie haben, bei der die Auswurffraktion gar kein Einschlusskriterium war, wie kriegen Sie dann heraus, welche Subgruppe innerhalb der Studie in das klassische Kriterium passt und welche nicht? Die bekommt man bei den meisten solcher Untersuchungen nicht heraus. Wie wollen Sie dann die Subgruppen bilden? Das ist mir nicht klar.

Fülöp Scheibler: Die kann man natürlich nur bilden, wenn man diese Angaben in den Studien hat. Ich stelle es mir so vor, dass ich einen Riesenhaufen Studien habe, der sowohl die schwereren als auch die leichteren Fälle enthält. Dann prüfe ich: Kann ich diesen Haufen in kleinere Haufen zerlegen? Am Schluss werden wahrscheinlich 3 kleinere Haufen übrig bleiben: ein Haufen mit den schwereren Fällen, ein Haufen mit den leichteren Fällen und ein Haufen mit unklaren Fällen. Dann habe ich das, was Sie in Ihrer Stellungnahme fordern, nämlich die schwereren Fälle mit einem separaten Ergebnis dargestellt, habe aber zusätzlich auch noch die anderen Ergebnisse.

Die Frage ist – das ist aus meiner Sicht vielleicht naiv –: Gibt es klinisch einen Grund, warum man die PET bei Revaskularisation in leichteren Fällen überhaupt nicht anwenden sollte? Wäre es klinisch falsch, sich diese Untersuchung anzusehen?

Wolfgang Burchert: Ich glaube, dass diese Fragestellung umso häufiger auftritt, je schlechter die linksventrikuläre Pumpfunktion ist. Patienten, die eine sehr gut erhaltene Pumpfunktion haben, haben in der Regel auch eine sehr gute Prognose. Das heißt, sie haben ein geringes Operationsrisiko, ein geringes Interventionsrisiko. Da wird in der Regel, wenn

Befunde da sind, interveniert. Der Wert von 35 % betrifft die Patienten deshalb, weil sie per se eine eingeschränkte Lebenserwartung und ein hohes Risiko haben, dass sie einen ernsten kardialen Zwischenfall erleiden. Aus dem Grund ist es die Gruppe, die für Vitalitätstests interessiert. Einen Vitalitätstest würde man bei einem Patienten, der eine EF von 55 % hat, so nicht durchführen, sondern es wird geguckt, ob es technisch machbar ist, wenn er zum Beispiel einen Infarkt hatte. Es gibt noch andere Verfahren, aber er wird revaskularisiert. Das gibt es heute nicht mehr anders.

Die 35 % sind Patienten, die aufgrund einer Rhythmusstörung, aber auch der eingeschränkten Auswurfraction eine schlechtere Lebenserwartung haben. Diese Gruppe ist die interessante, bei der die Frage entsteht: Was macht man jetzt mit denen? Sie haben natürlich ein höheres Operationsrisiko usw. In der Regel haben sie auch komplexere Gefäßverhältnisse. Da geht es häufig um die technische Machbarkeit und solche Dinge. Die 35 % sind schon das relevante Patientenkollektiv, auf das sich Vitalitätstests generell beschränken. Es ist auch egal, ob das SPECT, Echo oder sonst etwas ist. Niemand wird mit Hingebung ein Dobutamin-Echo machen, wenn man eine sehr gute Auswurfraction hat. Oder?

(Anselm Bäumer: Richtig!)

Das ist einfach unrealistisch.

Frank Bengel: Dazu noch der andere Punkt: Nuklearmedizinische Verfahren haben Strahlenexpositionen und unterliegen, wie wir es dargestellt haben, der rechtfertigenden Indikation. Es gibt andere Vitalitätsdiagnostikverfahren, die vielleicht auch bei Patienten eingesetzt werden, die eine weniger stark eingeschränkte linksventrikuläre Pumpfunktion haben, die nicht mit Strahlenexpositionen verbunden sind, die auch in den Händen des behandelnden Arztes liegen und deswegen häufiger in solchen Situationen eingesetzt werden. Deswegen kommt auch ein gewisses Bias hinein.

Fülöp Scheibler: Absolut einverstanden. Nur noch einmal die Frage: Wenn wir jetzt die Studien auf diese Kollektive einschränken, dann könnte es rein theoretisch sein, dass es eine tolle Studie gibt, die den Nutzen der PET bei Patienten mit < 45 darstellt und keine Subgruppe für $< 35 / > 35$ anbietet. Diese Studie müssten wir ausschließen und würden sie gar nicht interpretieren. Das Gleiche gilt dann auch für die diagnostische Güte. Deswegen frage ich: Sollen wir das wirklich tun?

Frank Bengel: Das kann man trotzdem machen. Ich finde, dass nicht arg viel gegen ein solches Vorgehen spricht. Es ist nur die Frage, welche Schlussfolgerungen man daraus zieht. Mit anderen Worten – wenn man das übersetzt, was Herr Burchert gesagt hat –: Es kann sein, dass die ganze Analyse, die Sie machen, für Patienten mit wenig eingeschränkter Pumpfunktion am Schluss für die Katz ist. Aber Schaden tut es sicherlich nicht.

Stefan Sauerland: Der spannendere Punkt ist doch, inwieweit solche Ergebnisse übertragbar sind. Ich habe völliges Verständnis dafür, dass so etwas klinisch nicht gemacht wird und auch

aus Strahlenschutzgründen sicherlich bedenklich ist. Wenn ich aber eine solche Studie finde, wo tatsächlich aufgrund der PET eine Revaskularisierung gemacht wird, wo ich dann bestimmte Parameter in der Nachuntersuchung beobachte, dann ist das doch ein Ergebnis, das zumindest für diese Subgruppe der Patienten Informationen bieten kann. Diese Informationen haben vielleicht auch eine Validität für die andere Patientengruppe, nämlich die mit einer deutlich eingeschränkten linksventrikulären Ejektionsfraktion.

Frank Bengel: Ich bin jetzt am Schluss ein bisschen zynisch, aber in der praktischen Realität ist es genau umgekehrt. Immer wenn jemand mit einem neuen Test zeigen will, dass er zumindest genauso gut ist wie PET, dann nimmt er Patienten mit möglichst wenig Einschränkung der linksventrikulären Pumpfunktion. Dann klappt es nämlich, und man kann sagen: Bei wenig Einschränkung gibt es ohnehin keinen Unterschied zwischen PET und dem neuen Test, deswegen ist der Test genauso gut wie PET. – So ist es in der praktischen Realität gelaufen. Das ist natürlich meine persönliche Meinung, aber das könnte eine solche Analyse auch unterstützen.

Moderator Jürgen Windeler: Dann haben wir den Punkt auch geklärt.

Wir kommen zu Tagesordnungspunkt 4.

4.4.5 Tagesordnungspunkt 4: Verschiedenes

Gibt es Wortmeldungen von Ihrer Seite? Ist alles gesagt? – Gibt es Wortmeldungen von unserer Seite?

Stefan Sauerland: Ganz hilfreich fand ich die Diskussion, als gesagt wurde, dass ein Herzzentrum aufgrund der PET, das andere Herzzentrum aufgrund der SPECT vorgeht und seine revaskularisierenden Maßnahmen durchführt. Das ist doch eigentlich die klassische Situation, in der es organisatorisch und ethisch machbar wäre, eine randomisierte Studie durchzuführen. – Das nur als kleiner Appell von meiner Seite.

Wolfgang Burchert: Die Fragen, die jetzt aufgeworfen wurden, sind so noch nie gestellt worden. Deswegen sind sie so auch noch nie beforscht worden. Deswegen gibt es überhaupt so viele Probleme. Sonst könnten wir einfach sagen: Wir haben eine wunderbar designte Multicenter-Studie, in der man im Vergleich verschiedene Verfahren verwendet hat. Es gibt ganz kleine Ansätze dazu, aber im Bereich FDG-PET gerade nicht. Im englischen Gesundheitssystem gibt es beispielsweise für Myokard SPECT versus Echo versus MRT, um für die Frage bei KHK und Revaskularisation die Patienten zu identifizieren, die am meisten von der Revaskularisation profitieren. Dazu gibt es Studien, die ein paar Jahre alt sind. Bei solch einer Fragestellung hat man tatsächlich einmal verschiedene Gruppen genommen, aber das ist vom Gesundheitssystem ausgegangen. Die haben gefragt: Wie müssen wir unsere Infrastruktur aufbauen? Dann hat man geschaut und fand, dass Echo sehr gut und sehr günstig war, aber leider relativ viel Follow-up, also neue Untersuchungen indiziert hat. Die

Myokardszintigraphie kam zum Beispiel am besten heraus, weil die richtigen Patienten richtig revascularisiert worden sind.

Es ist wünschenswert, dass solche Studien mit etwas mehr Power gemacht werden, aber die gibt es so wahrscheinlich in der Literatur nicht. Ich denke, es ist eine Kultur, die erst entsteht. Es ist auch die Frage: Wer finanziert diese Studien? Das ist ein ungelöstes Problem.

Moderator Jürgen Windeler: Wobei ich immer darauf hinweise, dass jemand, wenn er die Frage nach der Finanzierung stellt, eigentlich schon gesagt hat: „Solche Studien müsste es geben“, sonst erübrigt sich ja die Frage nach der Finanzierung.

Wolfgang Burchert: Nein. Heute ist es so, dass Sie gar nicht forschen können, wenn Sie kein Geld haben.

Moderator Jürgen Windeler: Ja, klar. Aber die Voraussetzung dafür, dass Sie sich Gedanken machen, ob und wie Sie forschen können, ist erst einmal, dass Sie das Gebiet beforschen möchten. Sie haben am Anfang gesagt: Die Frage hat in der Form noch niemand gestellt. – Erst einmal legt man die Fragen offen, die wichtig sind, die man beforschen möchte. Wenn man das alles festgestellt hat, dann muss man sich in der Tat sehr engagiert und häufiger auch frustrierende Gedanken machen: Wer finanziert das? Die entscheidende Frage ist aber, zunächst einmal zu sagen: Das wollen wir wissen.

Wolfgang Burchert: In Erwiderung dazu: Wir haben lange bei uns in der Ethikkommission diskutiert, ob man Studien an Patienten aus wirtschaftlichen Gründen durchführen darf. Letztlich ist alles eine Frage der Finanzierung, des Reimbursements, wie auch immer, das hinter solchen Untersuchungen steckt. Die verschiedenen Verfahren sind ja alle eingeführt und in der Praxis. MRT und PET sind nicht in der gesetzlichen Krankenversicherung, Echokardiographie und SPECT sind in der gesetzlichen Krankenversicherung. Es geht letztlich um die Frage: Welche Kriterien gibt es dafür, dass es wirtschaftlich eingeführt werden kann? Das ist die Frage, um die es geht und die diese Studien triggert; denn verwendet wird eine breite Vielfalt. Das bedeutet, es gibt keine klare Antwort. Das ist so ähnlich wie in der Juristerei. Man wird das jetzt untersuchen, und es wird wahrscheinlich – das ist meine persönliche Prognose – in einer gewissen Pattsituation enden.

Fülöp Scheibler: Der Wirtschaftlichkeitsaspekt kommt zwar in unserem Namen vor, aber nicht in den PET-Projekten. Das muss ich ganz deutlich klarstellen. Uns geht es um den patientenrelevanten Nutzen solcher Verfahren. Uns würde interessieren: Kann man durch Anwendung der PET einen größeren Nutzen erzielen als durch Anwendung der MRT? Die Frage: „Kann man das auch noch bezahlen?“ stellt sich dann der G-BA. Wir gucken zum Beispiel nicht nach gesundheitsökonomischen Studien, sondern nur nach dem Nutzen: Kann man durch eine präzisere Prognose den Patienten tatsächlich mehr helfen als mit bisherigen Verfahren? Das ist unser Ziel.

Anselm Bäumer: Ich habe noch eine ganz andere Frage für den Kardiologen zum Verständnis. Sie machen ja in Bad Oeynhausen relativ viele PET-Untersuchungen für Deutschland. Für die methodische Durchführung an Patienten gibt es in Deutschland und in Europa keine echte Leitlinie, nur eine Handlungsleitlinie aus den USA aus dem Jahr 2009. In der europäischen Empfehlung der ESC, der Europäischen Gesellschaft für Kardiologie, gibt es für die Vitalitätsdiagnostik 2 Methoden: 2-Tracer-Perfusion und FDG oder das hyperinsulinämische Clamp-Modell. Die Frage ist nur, weil das aus den Studien manchmal nicht so klar hervorgeht, was Sie anwenden und warum diese Methoden so nebeneinanderstehen, zumindest in der europäischen Zusammenfassung.

Wolfgang Burchert: Das hat zwar konkret hiermit nicht unbedingt etwas zu tun, aber weil die Frage gestellt ist: Das hyperinsulinämische Clamp-Modell ist praktisch eine Glukoseinfusion mit etwas Insulin dazu. Die kann man im Prinzip ohne Kontrolle laufen lassen. Den einzelnen Patienten, die einen sehr hohen Blutzuckerspiegel haben, geben wir noch zusätzlich Insulin. Das sorgt dann dafür, dass das Herz auf den Verbrauch von Kohlenhydraten umgestellt wird. Ein Grundgedanke dabei ist, dass man eine optimale Bildgebungsqualität bekommt. Wenn man Patienten lange nüchtern lässt, zum Beispiel 24 Stunden, und würde dann FDG injizieren, würde man eine sehr heterogene Anreicherung bekommen.

Anselm Bäumer: Und wann brauche ich dann 2 Tracer?

Wolfgang Burchert: 2 Tracer braucht man bei der Frage von Vitalität und Outcome eigentlich immer. Wichtig ist ein Perfusionsmetabolismus-Mismatch. Das heißt, dort, wo noch Stoffwechsel nachweisbar ist, aber die Perfusion deutlich erniedrigt ist, sind die Regionen, die tatsächlich wesentlich profitieren, die Hibernating Myocardium repräsentieren. Wenn Sie beispielsweise Stunned Myocardium haben, dann ist die Perfusion eben hoch. Da ist dann die Bewegungsstörung sozusagen nur Folge eines vorangegangenen ischämischen Ereignisses, aber im Normalzustand nicht vorhanden. Das heißt, da ist die Perfusion normal und der Stoffwechsel auch normal.

Anselm Bäumer: Das heißt, Sie nehmen den Ammonium-Tracer?

Wolfgang Burchert: Genau. Wir machen es mit Ammoniak und FDG.

Jörg Kotzerke: Andere Leute machen das mit MIBI-SPECT und FDG-PET und matchen das dann auf einer gemeinsamen Polar Map.

Anselm Bäumer: Gated SPECT plus.

Jörg Kotzerke: Oder gated SPECT plus.

Frank Bengel: Ich muss noch eine versteckte Aussage von Ihnen korrigieren. Es gibt natürlich europäische Richtlinien zur Durchführung der Positronenemissionstomographie. Die

sind in den EANM/ESC-Guidelines for myocardial perfusion imaging, allerdings versteckt. Sie sind 2007 publiziert worden.

Moderator Jürgen Windeler: Gut. Mit diesem ergänzenden Hinweis sehe ich keine weiteren Wortmeldungen mehr. Ich darf mich herzlich bei Ihnen für Ihre Beiträge zur Klärung bedanken.

Damit schließe ich die Erörterung. – Danke schön.

Anhang – Dokumentation der Stellungnahmen

Inhaltsverzeichnis

	Seite
1 Stellungnahmen von Organisationen, Institutionen und Firmen.....	A 2
1.1 Deutsche Gesellschaft für Nuklearmedizin.....	A 2
1.2 Deutsche Krankenhausgesellschaft	A 13

1 Stellungnahmen von Organisationen, Institutionen und Firmen

1.1 Deutsche Gesellschaft für Nuklearmedizin

Autoren:

Bengel, Frank

Burchert, Wolfgang

Kotzerke, Jörg

Schäfers, Michael



DGN e. V. • Nikolaistraße 29 • D-37073 Göttingen

Geschäftsstelle
Nikolaistraße 29
D-37073 Göttingen

Tel. (Mitglieder): +49 (0)551 48857-
Tel. (Presse): +49 (0)551 48857-402
Fax: +49 (0)551 48857-

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

– IQWiG Stellungnahme zum Berichtsplan D06-01L –
Prof. Dr. med. Jürgen Windeler
Dillenburger Straße 27

51105 Köln

– vorab per E-Mail –

Ressort	Info	Rückspr.	Bearb.	Verbleib	Ressort	Info	Rückspr.	Bearb.	Verbleib
iL					GI				
AM					Komm				
NMIV					VW				
VQ					GÖ				
Biometrie									

office@nuklearmedizin.de
www.nuklearmedizin.de

EINGEGANGEN

16. Jan. 2012

Ihre Zeichen

Ihre Nachricht vom

Unser Zeichen

Göttingen

gj/uf

11. Januar 2012

Stellungnahme der Deutschen Gesellschaft für Nuklearmedizin (DGN) zum vorläufigen Berichtsplan D06-01L „Positronenemissionstomographie (PET) und PET/CT zur Erkennung von vitalem Myokardgewebe bei Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler und globaler linksventrikulärer Funktion – Version 1.0“

Sehr geehrter Herr Professor Windeler,
sehr geehrte Damen und Herren,

für die Möglichkeit, zu oben genanntem Berichtsplan vom 5. Dezember 2011 Stellung nehmen zu können, bedanken wir uns recht herzlich.

Wir bitten um Berücksichtigung folgender Anmerkungen:

1. Allgemeine Anmerkungen

- 1.1. Insgesamt finden sich im vorläufigen Berichtsplan sehr viele inhaltliche z. T. substantielle Fehler und Begriffsunschärfen, die im Weiteren beispielhaft benannt werden. Der Autor ist offensichtlich mit der klinischen Situation und den dazugehörigen Begriffen nicht vertraut. Die aufgeworfenen Fragestellungen sind teilweise unscharf formuliert. Es entstand bei allen an der Stellungnahme der Deutschen Gesellschaft für Nuklearmedizin Mitwirkenden der Eindruck einer oberflächlichen Bearbeitung mit substantiellen Fehlern. Eine Publikation von einem solch niedrigen Qualitätsniveau würde bei einer wissenschaftlichen Zeitschrift konsequent abgelehnt werden.
- 1.2. Der Titel und auch die verschiedenen Teile des Berichtsplanes weisen die PET und die PET/CT als verschiedene Verfahren aus, was historisch richtig ist. Dem Autor des Berichtsplans scheinen die methodischen Unterschiede aber nicht geläufig zu sein. Der tatsächliche Unterschied hinsichtlich der hier untersuchten klinischen Fragestellung ist allerdings marginal und bezieht sich auf die Art der angewandten Schwächungskorrektur (Gammaemitter vs. Röntgenstrahlung). Insofern ist es bis zu einem gewissen Grade möglich, die Ergebnisse beider Verfahren zu poolen. Durch die schnelle methodische Weiterentwicklung der PET sind allerdings die modernen PET/CTs (reine PET-Scanner ohne CT sind seit mehrere Jahren nicht mehr kommerziell erhältlich, hier verweisen wir auf unsere früheren Stellungnahmen) häufig dramatisch leistungsfähiger hinsichtlich Empfindlichkeit (Verbesserung der Bildqualität, Erniedrigung der Strahlenexposition, Verkürzung der Untersuchungszeit) und insbesondere der Volumenauflösung (aktuell isotrop ca. 2-3 mm und damit fast in der Größenordnung der kardialen MRT). Die technische Entwicklung und zeitgemäße Durchführung der Untersuchung sollte daher im Nutzenbericht hinreichend Berücksichtigung finden.

IQWiG_D06-01L_vorl_berichtsplan_stellungnahme_dgn_120111.doc



- 1.3. Bei den Schlagwörtern ist „Myocardial Stunning“ sicherlich ein zutreffender Begriff; jedoch wird klinisch unter „Myocardial Stunning“ eine kurzdauernde Episode einer Minderperfusion mit einer konsekutiven Bewegungsstörung des Myokards verstanden, die spontan reversibel ist. Von „Hibernating Myocardium“ (deutsch: winterschlafendes oder hibernierendes Myokard) wird bei chronischer Minderperfusion mit begleitender Bewegungsstörung des Myokards gesprochen, welche nach einer Revaskularisation meistens – zumindest teilweise – reversibel ist. In beiden Fällen liegt bei einer Bewegungsstörung des Myokards vitales Myokard vor („myocardial viability“), welches die pathophysiologische Basis der Reversibilität der Bewegungsstörung darstellt. Weiterhin bezieht sich die klinische Fragestellung des geplanten Nutzenberichtes insbesondere auf Patienten mit eingeschränkter linksventrikulärer Pumpfunktion. Dies ist das klinische Korrelat der kardiovaskulär bedingten Herzinsuffizienz („heart failure“).

Wir würden daher empfehlen, als Schlagwörter auf Grund der treffenderen Beschreibung „hibernating myocardium“, „myocardial viability“ und „heart failure“ mit aufzuführen und bei den Recherchen zu benutzen. Die deutschen Schlagwörter wären dementsprechend „Hibernierendes Myokard“, „Myokardvitalität“ und „Herzinsuffizienz“. (Seite ii)

2. Kommentare zu Kapitel 1: Hintergrund

- 2.1. Das Kapitel 1 beschäftigt sich mit der koronaren Herzkrankheit im Allgemeinen und ist nicht richtungsweisend für die spezielle Situation von Patienten mit eingeschränkter linksventrikulärer Funktion, bei denen die Erkennung von vitalem Myokardgewebe relevant ist. Es sollte betont werden, dass es sich bei Kandidaten für eine myokardiale Vitalitätsdiagnostik um Patienten mit fortgeschrittener koronarer Herzerkrankung und Herzinsuffizienz handelt. Insbesondere die Abschnitte 1.2, 1.3, 1.4 sowie 1.5 sollten auf diese spezielle Situation ausgelegt werden.
- 2.2. Die Gesamtsituation wird in Kapitel 1.2 verharmlost, in dem auf einen Rückgang der Mortalität der KHK hingewiesen wird. Es sollte jedoch vielmehr berücksichtigt werden, dass (im Gegensatz zur Mortalität der KHK allgemein) die Inzidenz der Herzinsuffizienz im Speziellen in den letzten Jahren zunimmt und weiter zunehmen wird. Es geht bei der Nutzenbewertung thematisch um Hoch-Risikopatienten mit bekannter KHK und Herzinsuffizienz, nicht um den Ausschluss einer unkomplizierten KHK. Somit ist auch eine Diagnostik bei solchen Patienten von sehr hoher Relevanz, um künftig zielgenaue Therapieangebote vor dem Hintergrund der knappen Ressourcen im Gesundheitssystem machen zu können (Revaskularisation, Herzersatzsysteme (LVAD), Herztransplantation, automatische Defibrillatoren). Diese Aspekte sollten im Kapitel 1.2 dargestellt sein.
- 2.3. In Kapitel 1.3 sind in den ersten beiden Absätzen die pathophysiologischen Grundlagen der Myokardischämie und als klinisches Symptom die Angina pectoris beschrieben. Die fortschreitende Stenosierung der Gefäße ist tatsächlich ein determinierendes Element der Myokardischämie. Daneben entwickeln sich aber Veränderungen des Gefäßsystems, wie Anastomosen oder Veränderungen im Bereich der kleinen Gefäße (Stichworte: Endotheldysfunktion, hypertensive Herzkrankheit, Remodelling des Myokards). Weiterhin ist zu berücksichtigen, dass bei Reinfarkten häufig auch z. B. arterielle/venöse Bypässe oder auch Aneurysmata der Koronarien vorliegen können. Insbesondere bei Patienten mit eingeschränkter linksventrikulärer Pumpfunktion sind Mehrgefäßerkrankungen die Regel. Daneben ist zu berücksichtigen, dass nach einem Myokardinfarkt in der überwiegenden Anzahl der Fälle auch eine medikamentöse Therapie erfolgt, die auf der myokardialen Seite direkt oder indirekt den Sauerstoffbedarf und andere Parameter beeinflusst. Diese Vorgänge sind mit einer „verminderten Kapazität des Herzens“ leider nicht zutreffend und auch nicht ausreichend beschrieben.

Zur Auslösung eines Myokardinfarktes kommt es in der Regel auf der Grundlage einer arteriosklerotischen Plaqueruptur, die eine Thrombusbildung induziert. Solche Plaquerupturen geschehen in Regionen mit einer starken Verengung des Gefäßes genauso häufig wie auch in Regionen mit nur geringen Einengungen des Gefäßlumens, so dass der Stenosegrad des Gefäßes nicht unbedingt prädiktiv für die Häufigkeit der Infarkte ist, eher ist es die Gesamtmenge der Softplaques.

Die Induktion von Rhythmusstörungen im Infarktgeschehen ist tatsächlich abhängig von der Lokalisation, aber insbesondere auch von der Lage, Größe sowie der Menge und Lokalisation des ischämischen Myokards und der myokardialen sympathischen Innervation.

Wenn eine so umfassende Beschreibung der Pathophysiologie der KHK erfolgen soll, dann sollte diese auch zumindest für die untersuchte Fragestellung hinreichend spezifiziert werden.

- 2.4. In Kapitel 1.5 wird als regelhafte Ursache eines Myokardinfarktes eine Prädisposition zur erhöhten Blutgerinnung beschrieben, während der rupturierte Plaque nur auf der Grundlage einer erhöhten Blutgerinnung einen Infarkt auslösen soll. Es ist zwar bekannt, dass bestimmte Risikogruppen, wie Diabetiker, zusätzlich zu der Atherosklerose auch Veränderungen in der Blutgerinnung aufweisen. Das zentrale Element bei der Pathogenese des Myokardinfarktes ist aber der Progress der Atherosklerose auf der Grundlage einer Fettstoffwechselstörung in Kombination mit einer Entzündungsreaktion in der Gefäßwand.
- 2.5. Bei der Begriffsfindung im Kapitel 1.6 ist es wichtig, nicht nur zwischen vitalem und nicht vitalem Myokard zu unterscheiden, sondern in der Untergruppe des vitalen Myokards auch noch darauf hinzuweisen, dass es gut perfundiertes und normal kontraktiles Myokard gibt, welches vital ist (und nach einer Revaskularisierung wegen fehlender Dysfunktion auch keine Funktionsverbesserung erreichen kann) und dass es auf der anderen Seite dysfunktionierendes, ischämisch kompromittiertes, aber vitales Myokard (hibernating myocardium) gibt, von dem eine Verbesserung der linksventrikulären Dysfunktion nach Revaskularisierung zu erwarten ist.
- 2.6. Bezüglich der ersten beiden Absätze des Kapitels 1.6: Die Darstellung der medikamentösen Therapie, deren Mechanismen und Ziele, ist nur schwer verständlich und ohne klare Struktur. Die symptomatische Therapie der KHK wird beispielsweise nicht von der Sekundärprophylaxe nach Myokardinfarkt getrennt.
- 2.7. Im dritten Absatz von Kapitel 1.6 sollte im letzten Satz „einhergeht“ anstelle des dort verwendeten „einhergehen kann“ eingesetzt werden, da der dort beschriebene Effekt ja in dem zitierten Artikel in einer Metaanalyse gezeigt wurde.
- 2.8. Im vierten Absatz ff. in Kapitel 1.6 wird von „Viabilität“ gesprochen. Diese Eindeutschung von „viability“ ist im klinischen Sprachgebrauch in keiner Weise gebräuchlich. In den Wirtschaftswissenschaften hingegen wird dieser Begriff tatsächlich als Fachbegriff verwendet, aber mit einer völlig anderen Bedeutung. Im klinischen Sprachgebrauch wird im Kontext mit der koronaren Herzkrankheit von Myokardvitalität oder Vitalität des Myokards als Begriff für lebendes Myokard im Gegensatz zur narbigen Defektheilung gesprochen.
Weiterhin wird in der Mitte des vierten Absatzes der Begriff Viabilität völlig falsch verwendet: „...Viabilität bezeichnet einen reversiblen Zustand einer Dysfunktion des Myokards ...“. Diese Definition ist falsch. Es handelt sich hierbei um hibernierendes Myokard, was in der zweiten Satzhälfte auch so benannt wird. Dieser Satz sollte durch eine korrekte Definition ersetzt werden, da dies ja eigentlich das zentrale Thema des Berichts zur Nutzenbewertung darstellt. Bei dieser Gelegenheit könnte er auch mit dem vorangehenden Satz zusammengefasst werden, der die gleiche Definition enthält.
- 2.9. Im drittletzten Satz des vierten Absatzes erfolgt eine Definition der Gewebevitalität, die sich nur auf die Erholung der Wandbewegung bezieht: „...Umgekehrt gilt ein dysfunktionelles Segment nur dann als viabel...“ Das Zitat (25) verweist als Quelle auf ein Editorial, in dem diese Aussage so nicht zu finden ist.
- 2.10. Im Unterkapitel 1.7 ist die Technik der Magnetresonanztomographie ungenau dargestellt. Mit der MRT kann Myokardvitalität durch eine Dobutamin-Stressuntersuchung und Beurteilung der kontraktiven Funktionszunahme erfolgen – oder/und durch eine Messung der Retention von Kontrastmittel in fibrotischem Gewebe in Ruhe (sog. „delayed enhancement“). Die Technik der Kardio-Computertomographie ist ebenfalls ungenau dargestellt.
- 2.11. Die Kardio-CT wird zur Vitalitätsdiagnostik nicht mit medikamentöser Belastung, sondern in sog. „delayed enhancement“ Technik in Ruhe durchgeführt.
- 2.12. Im Kapitel 1.7 im Abschnitt unter der Überschrift SPECT wird das Verfahren als „zweidimensionale metabolische Bildgebung des Myokardgewebes“ bezeichnet. Die Single Photon Emission Computed Tomography (SPECT) ist aber, wie der Name schon sagt ein 3D- bzw. 4D-Verfahren (als gated SPECT zeitaufgelöst über den Ventrikelzyklus).
- 2.13. Weiterhin sollte korrigiert werden, dass die Perfusionstracer (Tc-99m-MIBI, Tc-99m-Tetrofosmin, TI-201) ein Mischbild aus Perfusion und Vitalität ergeben, da die Radiopharmaka

primär perfusionsabhängig, aber auch vitalitätsabhängig mitochondrial getrappt werden (Tc-99m-markierte Radiopharmaka).

Die Aussage, dass die Durchblutung gleich Vitalität ist, stimmt sicher so allgemein nicht und sollte dementsprechend korrigiert werden. Eine Perfusionsmessung im zeitlichen Verlauf wird in der Regel mit Diffusionstracern bei der Positronen-Emissions-Tomographie durchgeführt. Die klinisch am häufigsten angewandten Verfahren mit Tc-99m-markierten Radiopharmaka basieren hingegen auf dem Prinzip des Gewebetappings (wie die Mikrosphärenmethodik) und werden zwischen 15 und 60 Minuten nach der Injektion als gated SPECT tomographisch aufgenommen, woraus später die funktionellen Schnittbilder in herzachsengerechten Ebenen berechnet werden können.

- 2.14. Im Kapitel 1.7 unter der Überschrift Magnetresonanztomographie ist die Aussage des zweiten Satzes unklar. Die medikamentösen Belastungsverfahren mit Dobutamin erlauben eine qualitative Beurteilung der Wandbewegung und deren Änderung in repräsentativen Arealen aus den Myokardsegmenten nach der American Heart Association (AHA). Eine Trennung zwischen Narbe und durchblutetem Myokard erfolgt mit höherer Auflösung in Ruhe nach Gabe eines Gd-basierten Kontrastmittels als sogenanntes Late Enhancement.
- 2.15. Der Abschnitt Computertomographie des Herzens (Was ist Kardio-Computertomographie?) weist mehrere inhaltliche Mängel auf. Die offizielle Bezeichnung lautet Mehrschicht-Computertomographie (MS-CT). Praktisch alle Hersteller bieten Geräte mit 64 oder mehr Schichten mit hohen Rotationsgeschwindigkeiten an, die für die Untersuchung des Herzens geeignet sind. Die Strahlenexposition der CTs bei Herzuntersuchungen hat sich in den letzten Jahren durch die technologische Weiterentwicklung deutlich vermindert und liegt jetzt etwa in einem vergleichbaren Bereich wie andere bildgebende Verfahren in der Herzdiagnostik auch (Koronarangiographie, Myokard-Perfusions-SPECT, PET und PET/CT). Die Vitalitätsdiagnostik erfolgt im CT nicht unter Belastung, sondern in Ruhe nach einem analogen Verfahren wie das Late Enhancement im MRT, hier nur nach Applikation von jodhaltigem Kontrastmittel. Neueste Entwicklungen zeigen hier in der Kombination von Spektral-CT und iterativen Rekonstruktionsverfahren Fortschritte, die möglicherweise die klinische Anwendung einer Late Enhancement Technik in der Vitalitätsdiagnostik in Zukunft klinisch erlauben.
- 2.16. In Kapitel 1.7 ist die Beschreibung des PET und des PET/CT sehr fehlerhaft und unscharf und muss komplett überarbeitet werden. Wichtige Konzepte werden gar nicht dargestellt und andere ausführlich beschrieben, die aber nur historische Bedeutung haben (Teilringe, Koinzidenzkameras). Existierende wichtige Fachbegriffe fehlen vollständig. Der Aufbau und das Prinzip eines Radiopharmakons sind nicht ausreichend dargestellt. Dies ist umso problematischer, da über die Art und die Menge des applizierten Radiopharmakons die eigentliche Spezifität der Bildgebung und die Strahlenexposition des Patienten definiert werden. Es besteht eine sehr große Unsicherheit bei der Verwendung der Begriffe (z. B. wird von einer „kardiologischen Vitalitätsdiagnostik“ gesprochen, gemeint ist aber eine myokardiale Vitalitätsdiagnostik).
- 2.17. Im Kapitel 1.7 unter der Überschrift PET und PET/CT ist insbesondere beim sechsten Absatz erneut zu betonen, dass durch den Mismatch-Score nicht der Anteil des gesamten vitalen Myokards, sondern der Anteil an ischämisch kompromittiertem, aber vitalem Myokard erfasst wird. Die Aussage, dass die Festlegung von Myokardsegmenten als hibernierend rein subjektiv erfolgt, ist falsch. Es sind hierfür verschiedene semiquantitative und auch quantitative Ansätze entwickelt worden. Die besondere Stärke der PET und PET/CT ist gerade die Quantifizierbarkeit der Daten. Besonders zu betonen ist, dass die PET erlaubt, dies auf Pixelbasis zu tun und nicht nur auf dem groben Segmentraster.
- 2.18. Die Aussage im siebten Absatz des Abschnittes PET und PET/CT, Kapitel 1.7 ist falsch. Die American Society of Nuclear Cardiology (ASNC) beschreibt nicht, dass mit FDG allein sowohl Durchblutung als auch Energiestoffwechsel untersucht werden können. Die Guidelines der ASNC, die hier als Referenz 34 zitiert sind, empfehlen explizit, dass sowohl Perfusion als auch Stoffwechsel mit separaten Radiopharmaka untersucht werden sollen. Zu Grunde liegt dieser Aussage offensichtlich eine Verwechslung mit O-15-Wassertechniken, bei denen sowohl Perfusion als auch der sog. perfusable tissue index als Marker der Vitalität in einer einzigen Untersuchung erfasst werden können.

- 2.19. Die letzten beiden Absätze von Kapitel 1.7, auf Seite 7 unten sowie Seite 8 oben sind für die Herzbildgebung nicht relevant und offensichtlich aus der onkologischen Diagnostik entnommen und sollten entsprechend gelöscht werden.
- 2.20. Unterpunkt 1.8 weist mehrere schwerwiegende Fehler auf und sollte komplett überarbeitet und der publizierten Realität angepasst werden:
- 2.20.1. Der erste Satz ist falsch. In der nationalen Versorgungsleitlinie KHK wird PET durchaus erwähnt, und zwar genau im Zusammenhang mit der Vitalitätsdiagnostik, unter Punkt 7.22.
 - 2.20.2. Der zweite Satz ist auch falsch. Die ESC-Guideline zum Thema „Myocardial Revascularisation“ sagt nicht, dass randomisierte Studien fehlen, sondern erwähnt explizit auf Seite 2510, dass die derzeitige Evidenz zum größten Teil auf Observationsstudien beruht, *„jedoch mit der Ausnahme von 2 randomisierten klinischen Studien, die sich beide auf PET beziehen“*.
 - 2.20.3. Der vierte Satz ist ebenfalls inkorrekt wiedergegeben: Die DGN verweist auf den „prädiagnostischen Wert“ (zur Vorhersage einer Funktionsverbesserung), nicht auf den „prognostischen Wert“ von SPECT und PET.
 - 2.20.4. Die beiden letzten Sätze von 1.8, sowie die Referenz 39 tragen zum Thema der Vitalitätsdiagnostik nichts bei, da sie auf das Thema der allgemeinen Diagnose der koronaren Herzerkrankung ausgelegt sind, und nicht auf die spezielle Situation der myokardialen Vitalitätsdiagnostik bei Patienten mit fortgeschrittener KHK.
 - 2.20.5. Die aktuelle Leitlinie (S1) der Deutschen Gesellschaft für Nuklearmedizin [1] steht unmittelbar vor der Aufnahme in das offizielle Leitlinienregister durch die AWMF (im Akkreditierungsprozess bei der AWMF; wird im ersten Quartal 2012 publiziert). Unter dem Abschnitt 2.1.7 Vitalitätsdiagnostik findet sich dort folgender Hinweis zur FDG-PET: „Die F-18-FDG-PET zeigt bei der Vitalitätsdiagnostik eine höhere Genauigkeit als die Myokard-Perfusions-SPECT und ist besonders bei Patienten mit schlechter LVEF (< 30 bis 40%) – sofern verfügbar – als nuklearkardiologisches Verfahren heranzuziehen [2]. In Kombination mit einer Ruheperfusion-SPECT oder -PET gelingt so auch in Segmenten mit einer Ruheperfusionstörung durch eine hochgradige Stenose, aber nachgeschaltet vitalem Myokard (hibernierendes Myokard) ein Vitalitätsnachweis. Es zeigt sich bei dieser Konstellation ein sog. Perfusions-Vitalitäts-Mismatch.“

3. Kommentare zu Kapitel 2: Ziele der Untersuchung

- 3.1. Absatz 3 in Kapitel 2 ist nicht klar. Dieser sollte deshalb gestrichen werden.

Die Anwendung der PET/CT am Menschen unterliegt der Strahlenschutzgesetzgebung (StrlSchV, RöV sowie weitere Regelwerke) und ist am Menschen immer nur dann erlaubt, wenn der Nutzen der Untersuchung eine mögliche (strahlenbiologische) Nebenwirkung deutlich überwiegt. Diese Entscheidung muss daher im Einzelfall abhängig von der jeweiligen Erkrankungssituation des Patienten getroffen werden, und wird vom fachkundigen Arzt in der rechtfertigenden Indikation dokumentiert. Die klinische Anwendung der rechtfertigenden Indikation wird im Rahmen der ärztlichen Qualitätssicherung regelmäßig von den Ärztlichen Stellen überprüft.

Es ist allenfalls sinnvoll, auf Populationsebene das Strahlenrisiko in einer Nutzen/Risikoanalyse mit einzubeziehen. Bei dieser sollte aber beachtet werden, dass mit zunehmendem Alter und schlechterer Prognose der Patienten sich die Risiken für eine strahleninduzierte Nebenwirkung stark reduzieren.

- 3.2. Im vierten Absatz bei der Aufzählung der Technologievergleiche ist nicht klar, welcher wissenschaftliche Wert von einem Vergleich der PET mit der PET/CT erwartet wird (Gleiches gilt für die Tabelle 5 in Kapitel 4.1.3.). Bezüglich der klinischen Effektivität werden sich nur geringe Unterschiede abbilden, die aufgrund der unterschiedlichen Schwächungskorrekturverfahren (Gammastrahlung vs. Röntgenstrahlung) oder des Alters der Geräte herleiten. Insofern kann man die Studien beider Geräteklassen poolen. Die praktische Nutzenanwendung ist aber begrenzt, da nur noch PET/CT-Geräte kommerziell vertrieben werden.
- 3.3. Im fünften Absatz ist ein konventionelles Verfahren mit EG abgekürzt, wofür steht das?

- 3.4. Die Absätze unter der Überschrift „Allgemeine Voraussetzungen für Studien zur Nutzenbewertung“ und „Allgemeine Voraussetzungen für Studien zur Bewertung der diagnostischen und prognostischen Güte“ sind immer noch stark an therapeutischen Maßnahmen orientiert. Hier ist es einfach, die Effektivität und den Nutzen eines Verfahrens darzustellen und zu bewerten. Bei diagnostischen Verfahren ist dies erheblich komplexer und es gibt derzeit eine aktuelle Diskussion, ob solche Bewertungskriterien überhaupt angewendet werden können oder wie eine sinnvolle Evidenzbasierung auch bei diagnostischen Verfahren realisiert werden kann [3]. Der Nutzen erstreckt sich dabei nicht nur auf die unmittelbaren Auswirkungen auf den Patienten, sondern auch auf positive Effekte des klinischen Prozederes und auf ökonomische Effekte.

4. Kommentare zu Kapitel 3: Projektbearbeitung

keine

5. Kommentare zur Kapitel 4 Methoden

- 5.1. Bei den Einschlusskriterien der Patienten (4.1.2.1) muss klar festgelegt werden, wie die „eingeschränkte regionale oder globale linksventrikuläre Funktion“ definiert wird. Eine exakte Definition ist für diese Analyse hoch-relevant, da definitionsgemäß der Grad der Verbesserung der linksventrikulären Funktion nach Revaskularisierung vom Grad der Einschränkung vor der Revaskularisierung abhängig ist. Mit anderen Worten, bei Einschluss von Patienten mit a priori nur gering eingeschränkter linksventrikulärer Pumpfunktion kann die Verbesserung nach Revaskularisierung auch allenfalls gering ausfallen. Die diagnostische Wertigkeit der Vitalitätsbeurteilung wird dann unterschätzt, sofern nicht unverhältnismäßig hohe Fallzahlen in Vergleichsstudien rekrutiert werden. Generell wird eine linksventrikuläre Auswurffraktion von 35 % oder weniger als ausreichendes Kriterium für die Einschränkung der linksventrikulären Funktion, und damit für die Notwendigkeit einer Vitalitätsdiagnostik gesehen.
- 5.2. Zu Kapitel 4.1.2.2 des Berichtsplans siehe Anmerkungen zum Methodenvergleich PET vs. PET/CT unter 3.2. dieser Stellungnahme.
- Teilringsysteme und Koinzidenzgammakameras sind historische Entwicklungslinien, die keine besondere Berücksichtigung mehr finden müssen. Eine einmalige Benennung bei den Einschlusskriterien ist ausreichend.
- 5.3. In den letzten beiden Absätzen des Kapitels 4.1.2.3 (Patientenrelevante Endpunkte) werden diejenigen Endpunkte korrekt aufgelistet, die bei typischen Therapiestudien leicht erreicht werden können. Eine angemessene, auf diagnostische Verfahren angepasste Vorgehensweise wird nicht vorgelegt. Es wird die höchste Evidenzstufe für Therapiestudien nach G-BA als Minimum verlangt. Insofern ist es unbedingt erforderlich, im Berichtsplan zu berücksichtigen, dass es sich bei der Fragestellung um ein diagnostisches Verfahren handelt. Ansonsten ist mit falschen Ergebnissen und Interpretationen zu rechnen, die an der klinischen Realität vorbeigehen. In Tabelle 3 des Abschnittes 4 Methoden des Berichtsplanes sind neben der Klassifikation des G-BA auch Klassifikationen benannt, die auf diagnostische Verfahren ausgerichtet sind (Fryback und Thornbury, Referenz 42 des Berichtsplanes). Obwohl diese Klassifikation schon 20 Jahre alt ist, ist die wissenschaftliche Diskussion immer noch nicht abgeschlossen [3], da die allgemein verwendeten Bewertungsalgorithmen für therapeutische Verfahren offensichtlich keine adäquate Bewertung diagnostischer Verfahren erlauben. Aspekte der Verbesserung klinischer Prozeduren oder wirtschaftlicher Effekte werden nicht berücksichtigt. Dies ist um so mehr ein Problem, als dass die alleinige Verwendung von patientenbezogenen reinen Outcome-Kriterien die Geschwindigkeit der Entwicklung von diagnostischen, hier bildgebenden Verfahren, sich daran bemessen würde, wie groß die therapeutischen Fortschritte sind. In der Realität ist die Situation aber umgekehrt: die bildgebenden Verfahren helfen mit, neue Therapien zu entwickeln und anzuwenden. Dies ist ein typischer Mechanismus eines Innovationshemmnisses. Der Gesetzgeber hat dieses sehr wohl erkannt und in der aktuellen Gesetzgebung berücksichtigt (§137e GKV-VStG). Gerade im Hinblick auf die neue Gesetzgebung soll das Ziel Ihres Nutzenberichtes sein, die derzeit verfügbare Evidenzlage zusammenzutragen und dem G-BA anzudienen. Aus wissenschaftlicher Sicht zählt dazu auch, die diagnostische und prognostische Güte konkurrierender Untersuchungsverfahren umfassend zu ermitteln. Dieser Aufgabe werden Sie bei Auslassen relevanter wissenschaftlicher Literatur durch Fokussierung auf Studien zum Zusatznutzen nicht gerecht.

- 5.4. In Hinblick auf die Einschlusskriterien der Patienten in Abschnitt 4.2.1.1. des Berichtsplans verweisen wir auf Punkt 5.1. dieser Stellungnahme.
- 5.5. Zu Abschnitt 4.2.1.2 des Berichtsplanes, vierter Absatz: Üblicherweise wird bei Vergleichen in der Bildgebung tatsächlich die FDG-PET als der Goldstandard für hibernierendes Myokard bzw. die myokardiale Vitalität benutzt, da die FDG-PET die krankheitsrelevanten Stoffwechselforgänge des Gewebes direkt darstellt.
- 5.6. Zu Abschnitt 4.2.1.2 des Berichtsplanes: Hier liegt im vierten Absatz (Seite 18, untere Hälfte) ein grundlegendes Missverständnis vor. Die Beurteilung der linksventrikulären Pumpfunktion vor und nach Revaskularisation erfolgt typischerweise unter Ruhe-Bedingungen. Hierfür muss auch nicht das Myokardgewebe direkt dargestellt werden, sondern es reicht eine Darstellung des Herzbinnenraums (also des Ventrikels) aus. Eine Differenzierung nach Belastungsstufen ist nicht notwendig, da Belastungsuntersuchungen zur seriellen Verlaufsbeurteilung der linksventrikulären Funktion nach Revaskularisation nicht indiziert sind. Als geeignete Verfahren zur Verlaufsbeurteilung der linksventrikulären Funktion in Ruhe, die für die Beurteilung der diagnostischen und prognostischen Güte der Vitalitätsdiagnostik relevant ist, sind somit MRT, CT, Echokardiographie (ohne Belastung), EKG-getriggerte Myokardperfusions-SPECT, Radionuklid-Ventrikulographie (RNV) sowie Lävokardiographie anzusehen.

Die Aussagen „Als geeignete Verfahren werden die folgenden drei Verfahren LDDE, MRT und CT angesehen. RNV und SPECT sind gegenüber MRT und LDDE weniger exakt [...]“ sowie im nächsten Satz „Echographie ohne Belastungstest ist wegen der fehlenden Induktion von Kontraktilität in viablen Segmenten der Stressechokardiographie unterlegen.“ sind nicht durch ausreichende Evidenz belegt und somit nicht haltbar.

Zusammenfassend ist somit nach diagnostischen Studien zur PET-Untersuchung zu suchen, in denen als Referenzstandard die zweimalige Durchführung eines Verfahrens zur Beurteilung der linksventrikulären Pumpfunktion in Ruhe (also Echokardiographie, MRT, CT, RNV, EKG-getriggerte SPECT oder Lävokardiographie), sowohl vor als auch nach der Revaskularisierung zur Feststellung der segmentalen Wandbewegung eingesetzt wird. Es ist hier außerdem auf den zeitlichen Abstand zwischen der ersten und zweiten Beurteilung der linksventrikulären Pumpfunktion einzugehen, da eine zu frühe Wiederholung der Pumpfunktionsuntersuchung nach Revaskularisierung aufgrund von Stunning-Effekten zu einer Unterschätzung der Funktionsverbesserung führen wird. In der allgemeinen Literatur wird ein Zeitabstand von mindestens drei, besser sechs Monaten nach Revaskularisation gefordert.

- 5.7. Diesbezüglich wird auch auf Abschnitt 4.2.1.4 des Berichtsplans, vierter Absatz, verwiesen: Der hier aufgeführte Zeitraum von einer Woche ist in der Situation der chronischen linksventrikulären Dysfunktion bei Herzinsuffizienz bei weitem nicht ausreichend, um einen Therapieeffekt abschließend beurteilen zu können (s. a. 5.6. dieser Stellungnahme).
- 5.8. Kommentar zum dritten Absatz des Abschnitts 4.2.1.4. des Berichtsplans: Randomisierte Studien zur Fragestellung der Vitalitätsdiagnostik vor geplanter Revaskularisation wären in Deutschland derzeit aus ethischen Erwägungen nicht durchführbar.
- 5.9. Kommentar zu Abschnitt 4.5.2. Diagnosestudien, zweiter Absatz: Es ist bei der Interpretation der Daten später aber zu beachten, dass zum Beispiel die MRT, aber auch die CT des Herzens bisher nicht zur gesetzlichen Krankenversorgung zugelassen sind. Auf wissenschaftlicher Ebene sind diese Vergleiche sicher angezeigt, aber in Hinblick auf die klinische Versorgungsrealität bestehen hier doch erhebliche Unterschiede.

6. Kommentar zu Kapitel 5: Literaturverzeichnis

- 6.1. Referenz Nr. 38 ist veraltet. Die aktuelle Leitlinie der Deutschen Gesellschaft für Nuklearmedizin, die für die Aufnahme in das offizielle Leitlinienregister durch die AWMF vorgesehen ist (s. o. zu Punkt 2.20.5.), finden Sie beigefügt [1].
- 6.2. Die Referenz Nr. 33 ist fehlerhaft.
- 6.3. Die Referenz Nr. 68 ist fehlerhaft.

Wir würden uns freuen, wenn Sie diese Aspekte im abschließenden Berichtsplan berücksichtigten.
Mit freundlichen Grüßen



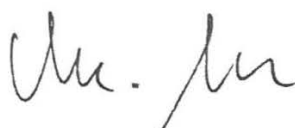
Prof. Dr. med. Jörg Kotzerke
Präsident der Deutschen Gesellschaft für Nuklearmedizin



Prof. Dr. med. Wolfgang Burchert
Mitglied des Vorstands der Deutschen Gesellschaft für Nuklearmedizin sowie
Mitglied der Arbeitsgemeinschaft Kardiovaskuläre Nuklearmedizin
der Deutschen Gesellschaft für Nuklearmedizin



Prof. Dr. med. Frank Bengel
Medizinische Hochschule Hannover – Klinik für Nuklearmedizin



Prof. Dr. med. Michael Schäfers
Vorsitzender der Arbeitsgemeinschaft Kardiovaskuläre Nuklearmedizin
der Deutschen Gesellschaft für Nuklearmedizin

Anlagen:

- Literaturverzeichnis
- Zitierte Literatur (eingereicht in elektronischer Form mit E-Mail vom 11. Januar 2012)
- Formblatt zur schriftlichen Stellungnahme zum Vorbericht
- Formblatt zur Offenlegung potenzieller Interessenkonflikte: Prof. Dr. med. J. Kotzerke
- Formblatt zur Offenlegung potenzieller Interessenkonflikte: Prof. Dr. med. W. Burchert
- Formblatt zur Offenlegung potenzieller Interessenkonflikte: Prof. Dr. med. F. Bengel
- Formblatt zur Offenlegung potenzieller Interessenkonflikte: Prof. Dr. med. W. Schäfers

Literaturverzeichnis

1. Lindner O, Bengel F, Burchert W et al. DGN-Handlungsempfehlung (S1-Leitlinie) Myokard-Perfusions-Szintigraphie, Version 1.3 (Stand 1.12.2011), AWMF-Registernummer 031-006. Im Akkreditierungsprozess bei der AWMF. Text und Abb. sind beigefügt.
2. Bax JJ, Wijns W, Cornel JH, et al. Accuracy of Currently Available Techniques for Prediction of Functional Recovery After Revascularization in Patients With Left Ventricular Dysfunction Due to Chronic Coronary Artery Disease: Comparison of Pooled Data. *J Am Coll Cardiol* 1997;30:1451-60.
3. Gazelle GS, Kessler L, Lee DW, et al. A Framework for Assessing the Value of Diagnostic Imaging in the Era of Comparative Effectiveness Research. *Radiology* 2011;261(3):692-698. Online verfügbar unter: <http://www.ncbi.nlm.nih.gov/pubmed/22095993> [Zugriff: Januar 5, 2012].

1.2 Deutsche Krankenhausgesellschaft

Autoren:

Acikgöz, Özlem

Dippmann, Anja

Hermeyer, Gottfried

Rüther, Klaus

Schlottmann, Nicole



DEUTSCHE
KRANKENHAUS
GESELLSCHAFT

Institut für Qualität und
Wirtschaftlichkeit im Gesundheitswesen
- Stellungnahme zum Berichtsplan <D06-01L>
Prof. Dr. med. Jürgen Windeler
Dillenburger Str. 27
51105 Köln

- vorab per E-mail -

Bundesverband
der Krankenhausträger
in der Bundesrepublik
Deutschland

EINGEGANGEN

13. Jan. 2012

Unser Zeichen

Dez. V/Ac

Telefon

Durchwahl +49 (0) 30 3 98 01 -

Telefax

+49 (0) 30 3 98 01 -

Datum

11.01.2012

Stellungnahme der Deutschen Krankenhausgesellschaft zum vorläufigen Berichtsplan D06-01L „Positronenemissionstomographie (PET und PET/CT) zur Erkennung von vitalem Myokardgewebe bei Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler und globaler linksventrikulärer Funktion“

Sehr geehrter Herr Professor Windeler,
sehr geehrte Damen und Herren,

für die Möglichkeit, zu oben genanntem Berichtsplan Stellung nehmen zu können, bedanken wir uns recht herzlich.

Der Gemeinsamen Bundesausschuss (G-BA) hat das Institut gem. § 139a Abs. 3, 1. SGB V am 21.12.2006 mit der Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissenstandes der PET-Diagnostik zur Erkennung von vitalem Myokardgewebe bei Patienten mit koronarer Herzerkrankung und eingeschränkter regionaler und globaler linksventrikulärer Funktion beauftragt. Der spätere Abschlussbericht soll eine Grundlage für die Beratungen des G-BA zur Bewertung der PET bei o. g. Indikation nach § 137c SGB V (Methodenbewertung stationär) und § 135 SGB V (Methodenbewertung ambulant) darstellen.

Gemäß § 137c SGB V hat der G-BA zu prüfen, ob Untersuchungs- oder Behandlungsmethoden, die zu Lasten der gesetzlichen Krankenkassen im Rahmen einer Krankenhausbehandlung angewendet werden oder angewendet werden sollen, für eine ausreichende, zweckmäßige und wirtschaftliche Versorgung der Versicherten unter Berücksichtigung des allgemein anerkannten Standes der medizinischen Erkenntnisse erforderlich sind.

Die Deutsche Krankenhausgesellschaft (DKG) als eine Trägerorganisation des G-BA nimmt zum vorliegenden Berichtsplan des Instituts wie folgt Stellung:

Postfach 120555 · D-10695 Berlin
Wegelystraße 3 · D-10623 Berlin

Telefon +49 (0) 30 3 98 01
Telefax +49 (0) 30 3 98 01

dkg@mail@dkgev.de · www.dkgev.de

Rue d'Arion 50
B-1000 Brüssel

Telefon +32 (0) 2 282
Telefax +32 (0) 2 282

europa@dkgev.de

Commerzbank AG

1. Die Abkürzung für die Deutsche Gesellschaft für Kardiologie lautet nicht DKG, sondern DGK.
2. Gemäß den *Methoden 4.0* Ihres Instituts hoffen wir, dass sich im endgültigen Berichtsplan eine geschlechtergerechte Sprache wiederfinden wird.
3. Wir begrüßen es, dass Sie in Ihrem Bericht verschiedene Leitlinien berücksichtigen. Allerdings stellen Sie im vorläufigen Berichtsplan fest, dass in der nationalen Versorgungsleitlinie Chronische Koronare Herzerkrankung (NVL KHK) von 2011 die PET oder PET/CT keine Erwähnung finden würde. Tatsächlich ist es so, dass laut Abschnitt „Empfehlungen zur nichtinvasiven Diagnostik bei Verdacht auf KHK“ die PET zur Evaluierung von Vitalität in dysfunktionalem Myokard durchgeführt werden kann.
4. Wie vom IQWiG dargestellt, ist das Ziel Ihrer Untersuchung die Prüfung des patientenrelevanten Nutzens oder Schadens. Wir möchten Sie noch einmal darauf hinweisen, dass der Auftrag an das IQWiG darin besteht, „die Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissensstandes zu PET, PET/CT“ durchzuführen (G-BA-Beschluss vom 19.12.2006). Die Beurteilung eines etwaigen Schadens ist hingegen nicht Gegenstand der Beauftragung. Wir bitten, dies im Berichtsplan zu berücksichtigen.
5. Wir begrüßen, dass bei der Beurteilung der Studien die gesundheitsbezogene Lebensqualität als patientenrelevanten Endpunkt in Ihre Betrachtungen mit einbezogen werden soll. Allerdings beschränken Sie sich hinsichtlich des patientenrelevanten Nutzens erneut auf die Stufe 5 nach Fryback & Thornbury. Die in der Auftragskonkretisierung des G-BA benannten zentralen Fragen beziehen sich aber auch auf die Bewertung, inwieweit sich eine präzisere Diagnosestellung auf die Behandlung und das weitere Management des Patienten und der Patientin auswirkt (4. Stufe). PET ermöglicht das Erkennen von Gewebe, das nicht mehr durchblutet wird, aber noch Glukosestoffwechsel aufweist. In Anbetracht der Tatsache, dass die PET-Untersuchung zur Beurteilung der Vitalität von Myokardgewebe gerade in der präoperativen Diagnostik für Therapiemanagement und OP-Risikoabwägung von entscheidender Bedeutung ist, würde laut Ihrem Berichtsplan eine wesentliche Indikationsstellung nicht angemessen berücksichtigt werden.
Wir möchten in Bezug auf den Begriff der therapeutisch-diagnostischen Einheit gemäß Auftragskonkretisierung nochmals darauf hinweisen, dass eine *direkte* Beeinflussung der von Ihnen angeführten patientenrelevanten Zielgrößen durch ein diagnostisches Verfahren nicht möglich ist. Eine diagnostische Methode kann daher anhand ihrer Befunde nur *indirekt* in Form einer therapeutischen Konsequenz patientenrelevante Endpunkte beeinflussen.
6. Bei Ihrer Bewertung diagnostischer Tests berufen Sie sich auf die „Mehrzahl der westlichen Gesundheitssysteme“, die „Überlegungen zu den Konsequenzen und dem Nutzen und Schaden diagnostischer Tests für Arzt, Patient und Gesellschaft in Betracht zieht“. Dieser Sichtweise habe sich die Verfahrensordnung des G-BA angeschlossen. Es bleibt unklar, was das IQWiG mit dem Ausdruck „westliches Gesundheitssystem“ meint. Die Verfahrensordnung des G-BA gibt unter § 11 , Abs 2 eindeutig Auskunft über die Klassifizierung und Bewertung von Unterlagen, die eine evidenzbasierte Beurteilung des medizinischen Wissensstandes ermöglichen sollen, die Beurteilung von Nutzen oder Schaden ist allerdings nicht Gegenstand dieses von Ihnen zitierten Paragraphen.
7. Darüber hinaus zitieren und beziehen Sie sich auf eine Verfahrensordnung, die nicht mehr gültig ist. Eine aktuelle Version¹, in Kraft seit August 2011, ist online verfügbar.

8. Wir begrüßen es, dass Sie gemäß Auftragskonkretisierung die Genauigkeit des Diagnostikverfahrens, d.h. auch die diagnostische und prognostische Güte bewerten wollen, soweit Primärstudien und Evidenzsynthesen zum patientenrelevanten Nutzen nicht vorhanden sind. Mangels Goldstandard für Vitalitätsdiagnostik beabsichtigen Sie, für die Abschätzung der Testgenauigkeit einen Referenzstandard heranzuziehen. Dazu soll patientenbezogen die segmentale myokardiale Funktion nach Revaskularisation beurteilt werden. Operationalisiert werden soll dies gemäß der Funktion von hibernierendem Myokard durch die Wandbewegung im Vergleich vor und nach einem Eingriff. Als Vergleichsdiagnostikum sollen lediglich LDDE, MRT und CT Akzeptanz finden. Die SPECT wird als Vergleich nicht akzeptiert, obwohl u. a. die nationale Versorgungsleitlinie KHK, die European Society of Cardiology und die Deutsche Gesellschaft für Nuklearmedizin zur Vitalitätsdiagnostik auch die SPECT empfehlen. Laut Ihrem Berichtsplan ist die SPECT „weniger exakt“; dabei berufen Sie sich auf Publikationen von Klocke et al, Hesse et al und Wu et al. Ihre Schlussfolgerung ist allerdings nicht nachvollziehbar, da in den genannten Publikationen diese Aussage nicht auffindbar getroffen wird:

Die EANM/ESC guidelines aus dem Jahre 2008 (Hesse et al) messen der gated SPECT als Diagnostikum bei arterieller koronarer Herzerkrankung eine zunehmende Bedeutung für die linksventrikuläre Funktion bei.

Wu et al, eine Publikation aus dem Jahre 2001, haben diese Thematik gar nicht selbst untersucht, sondern beziehen sich in ihrer Diskussion auf Publikationen von 1988 (Mahmarian et al. ²) und von 1995 (Miller et al ³), die beide SPECT nach nachgewiesenem Herzinfarkt mit einem nicht-vergleichenden Ansatz untersuchten. Mahmarian et al. untersuchten 30 Patienten in Bezug auf eine mögliche Korrelation der SPECT-Ergebnisse mit Enzymaktivitäten. Miller et al. untersuchten eine bestimmte Form der SPECT (Technetium -^{99m}Tc- sestamibi) und kamen zu der Schlussfolgerung, dass die Größe des mittels Myokard-Perfusions-SPECT ermittelten Infarktareals mit erhöhtem Mortalitätsrisiko korreliert. Ihre Nicht-Berücksichtigung der SPECT erschließt sich uns daher nicht.

Auch die in diesem Zusammenhang zitierte „Guidelines for the clinical use of cardiac radionuclide imaging“ (Klocke et al. JACC 2003) sind inzwischen überholt von der interdisziplinären Leitlinie „Appropriate Use Criteria for Cardiac Radionuclide Imaging.“ (Hendel et al 2009 ⁴), die allerdings auch keine Aussage in Ihrem Tenor enthält.

Ein Ausschluss von Studien, die SPECT als Vergleichsdiagnostikum genutzt haben, ist somit nicht gerechtfertigt.

9. Bei der Bewertung der PET als diagnostischer Test beziehen sich u. a. auf das Klassifikationsschema von Fryback und Thornbury, welches die Evidenzstufen von Testgüte und Nutzen hierarchisch integriert (Abschnitt 4, „Methoden“, S. 12). Allerdings beziehen Sie sich hinsichtlich diagnostischer und prognostischer Güte lediglich auf die Stufe 2 und zu den patientenrelevanten Endpunkten auf die Stufe 5 des Klassifikationsschemas nach Fryback und Thornbury. Angesichts der differenzierten, inhaltlich unmittelbar aufeinander aufbauenden Evidenzsteigerung wäre es sinnvoll und stringent, nicht nur Teilaspekte des Schemas, sondern die gesamte Evidenz von Stufe 2 bis 5 (einschließlich Stufe 3 und 4) in die Betrachtungen einzubeziehen und darzustellen.
10. Wir begrüßen die realistische Darstellung des Paradigmenwechsels in der Bewertung diagnostischer Verfahren hin zu vergleichenden Studien mit patientenrelevanten Endpunkten, der sich erst sehr langsam vollzieht und dass nicht-randomisierte, kontrollierte Interventionsstudien möglicherweise Hinweise auf die Ausgestaltung zukünftiger aussa-

gekräftigerer Studien erlauben. Allerdings halten wir die von Ihnen vorab anhand der Ein- und Ausschlusskriterien erfolgte Einschränkung der zu berücksichtigenden Primärliteratur auf Studien der Evidenzstufen I und II nicht für sachgerecht und nicht für konform mit den Vorgaben der Verfahrensordnung des G-BA. Daher bitten wir Sie, alle vergleichenden Studientypen zur Bewertung einzubeziehen, d.h. auch Studien auf Stufe III gemäß Verfahrensordnung des G-BA in Betracht zu ziehen. Zu dieser Problematik haben wir bereits mehrfach im Rahmen von Stellungnahmen u. a. zur allgemeinen Methodik des Instituts hingewiesen und diesen Aspekt in der Auftragskonkretisierung (vom 19.06.2008) bereits thematisiert. Die Einschränkungen führen gerade auch in diesem Fall dazu, dass ein Großteil wissenschaftlicher Erkenntnisse von vornherein ausgeschlossen bleibt. Damit kann insbesondere die „Darstellung und Bewertung des aktuellen medizinischen Wissensstandes“, die ausdrücklich Gegenstand des Auftrages vom 21.12.2006 ist, nicht sachgerecht erfolgen. Die Darstellungen des vorgelegten Berichtsplanes als Ausgangspunkt nehmend, wird der G-BA hier voraussichtlich wieder selbständig weitere Erkenntnisse ermitteln müssen.

11. Darüber hinaus möchten wir darauf hinweisen, dass die prognostische Güte nicht nur anhand des Auftretens zukünftiger klinischer Ereignisse wie Herzversagen und Herzrhythmusstörungen, sondern auch an dem patientenrelevanten Faktor Lebensqualität gemessen werden sollte, und möchten deshalb anregen, dies in die Betrachtungen aufzunehmen.
12. Die von Ihnen geplante Verwendung einer Relevanzschwelle von 0,2 ist international keineswegs unumstritten. Nicht kontrovers ist hingegen, dass die Relevanzschwelle vom Kontext der in diesem Fall klinischen Situation abhängig ist. Die Festlegung einer Relevanzschwelle im Falle fehlender skalenspezifisch validierter bzw. etablierter Relevanzkriterien auf eine standardisierte Mittelwertdifferenz (SMD) von 0.2 ist willkürlich und durch die zitierte Literatur nicht hinreichend belegt. Insbesondere vor dem Hintergrund, dass die im Bericht behandelte Erkrankung – die koronare Herzerkrankung mit eingeschränkter regionaler und globaler linksventrikulären Funktion - potenziell lebensbedrohlich ist, ist der o.g. Grenzwert für die SMD nicht akzeptabel. Da zudem die Entscheidung über Relevanz und Irrelevanz ohnehin vom G-BA getroffen werden muss und nicht Gegenstand der Literaturdarstellung sein sollte, fordern wir Sie auf, die Mittelwertdifferenzen zu benennen und auf die Anwendung jedweder Schwellenkriterien zu verzichten.
13. Die Metaanalysen, die keine bedeutsame Heterogenität in dem von Ihnen definierten Sinne aufweisen, sollten statt mit den von Ihnen in erster Linie geplanten Modellen mit zufälligen Effekten mit solchen mit festen Effekten durchgeführt werden.
14. Wir begrüßen es, dass Sie Patientenvertreter bei der Erstellung des Berichtsplans eingebunden haben. Sie stellen in Aussicht, dass zudem externe Sachverständige in die Bearbeitung des Projekts eingebunden werden. Wir regen in Bezug auf das vorliegende Bewertungsverfahren an, bereits ab Erstellung des Berichtsplanes kardiologische, herzchirurgische und nuklearmedizinische Expertise einzubeziehen.

Wir hoffen, dass wir Ihnen mit unseren Anmerkungen einige Anregungen für die Überarbeitung des Berichtsplans und die weitere Auftragsbearbeitung geben konnten. Im Übrigen behalten wir uns vor, ggf. ergänzende Aspekte in das weitere Verfahren einzubringen.

Bitte haben Sie Verständnis dafür, dass wir mit Blick auf die urheberrechtlichen Aspekte keine Kopien der Volltexte der von uns zitierten Literatur beigefügt haben.

Mit freundlichen Grüßen
Der Hauptgeschäftsführer
Im Auftrag



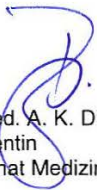
Dr. med. N. Schlottmann
Geschäftsführerin
Dezernat Medizin



Ö. Acikgoz
Referentin
Dezernat Medizin



Prof. Dr. med. K. Rüter
Referent
Dezernat Medizin



Dr. med. A. K. Dippmann
Referentin
Dezernat Medizin



Dr. med. Gottfried Hermeyer
Referent
Dezernat Medizin

Quellenangaben:

¹ http://www.g-ba.de/downloads/62-492-548/VerfO_2011-01-20_2011-08-04.pdf

² „Quantification of infarct size by 201Tl single-photon emission computed tomography during acute myocardial infarction in humans. Comparison with enzymatic estimates.“ Mahmarian JJ, Pratt CM, Borges-Neto S, Cashion WR, Roberts R, Verani MS. *Circulation*. 1988 Oct;78(4):831-9.

³ „Infarct size after acute myocardial infarction measured by quantitative tomographic 99mTc sestamibi imaging predicts subsequent mortality“. Miller TD, Christian TF, Hopfenspirger MR, Hodge DO, Gersh BJ, Gibbons RJ. *Circulation*. 1995 Aug 1;92(3):334-41.

⁴ ACCF/ASNC/ACR/AHA/ASE/SCCT/SCMR/SNM „Appropriate Use Criteria for Cardiac Radionuclide Imaging“. Hendel RC, Berman DS, Di Carli MF, Heidenreich PA, Henkin RE, Pellikka PA, Pohost GM, Williams KA; *J Am Coll Cardiol*. 2009 Jun 9;53 (23):2201-29