

Ökonomische Evaluationen der Positronen-Emissions-Tomographie

Ein gesundheitsökonomischer HTA-Bericht

Angelika Müller, Diana Stratmann, Thomas Klose, Reiner Leidl

Vorwort

Dieser Health Technology Assessment (HTA)-Bericht „Ökonomische Analysen der Positronen-Emissions-Tomographie – Gesundheitsökonomischer HTA-Bericht“ ist ein weiterer Beitrag zum Projekt „Aufbau einer Datenbasis zur Evaluation medizinischer Verfahren und Technologien in Deutschland“, das vom Bundesgesundheitsministerium gefördert wird. Auf der Basis einer umfangreichen Literaturrecherche zum Stichtag März 2000 dokumentiert und diskutiert dieser HTA-Bericht den Stand der ökonomischen Evidenz zum Einsatz der Positronen-Emissions-Tomographie in onkologischen, kardiologischen und neuromedizinischen Indikationen. Methodische Grundlagen der ökonomischen Evaluation können der Studie „Ansätze und Methoden der ökonomischen Evaluation - eine internationale Perspektive“ von Leidl R., Schulenburg vd JM. und Wasem J. (Hrsg.) entnommen werden, die als Band 9 in dieser Schriftenreihe erschienen ist. Komplementär zu diesem gesundheitsökonomischen Beitrag wird die medizinische Evidenz zur Positronen-Emissions-Tomographie in einem HTA-Bericht der medizinischen Hochschule Hannover beleuchtet.

Bei der Erarbeitung dieses Beitrags erhielten wir Unterstützung von verschiedenen Seiten. Unser Dank gilt den Reviewern, die im Rahmen der routinemäßigen Qualitätskontrolle, die durch die Medizinische Hochschule Hannover koordiniert wird, sinnvolle und hilfreiche Kommentare einbrachten. Des weiteren danken wir Herrn Prof. Dr. Reske, Leiter der Abteilung Nuklearmedizin des Universitätsklinikums Ulm, für die Bereitstellung von Kostendaten der Positronen-Emissions-Tomographie für diesen HTA-Bericht, die einer gemeinsamen Evaluationsstudie mit der Abteilung Gesundheitsökonomie der Universität Ulm entstammen. Besonderer Dank gebührt Frau Schmid, Sekretärin der Abteilung Gesundheitsökonomie, für ihre unermüdliche Unterstützung diesen HTA-Bericht in eine übersichtliche, den Normen der Buchreihe entsprechende Form zu bringen.

Ulm im Juli 2000

Die Autoren

Inhaltsverzeichnis

Vorwort

A	Abstract	1
B	Executive Summary	5
C	Hauptdokument	10
C.1	Policy Question	10
C.2	Einleitung	11
C.3	Forschungsfragen	12
C.4	Methodik	13
	C.4.1 Literaturrecherche	13
	C.4.1.1 Suchstrategie.....	15
	C.4.1.2 Ein- und Ausschlußkriterien	15
	C.4.2 Auswertung	16
	C.4.2.1 Transparenz- und Qualitätskatalog.....	16
	C.4.2.2 Zusammenfassende Beschreibung und Bewertung der als relevant klassifizierten Studien	18
	C.4.2.3 Typen der ökonomischen Evaluation.....	20
	C.4.2.4 Bewertung der verfügbaren Evidenz	23
C.5	Ergebnisse	25
	C.5.1 Ergebnisse der Literatursuche	25
	C.5.2 Die monatlichen Kosten der Positronen-Emissions- Tomographie	27
	C.5.3 Auswertungsergebnisse mit dem Transparenz- und Qualitätskatalog	32
	C.5.4 Zusammenfassende Beschreibung und Bewertung der als relevant klassifizierten Studien	39
	C.5.4.1 Onkologie	39
	C.5.4.2 Kardiologie	98
	C.5.5 Zusammenfassende Übersicht	137
C.6	Diskussion	146
	C.6.1 Gesundheitsökonomische Evidenz	150
	C.6.2 Beurteilung der Kosteneffektivität des klinischen Einsatzes der PET-Technologie insgesamt	163
C.7	Literaturverzeichnis	168
Anhang	176

Tabellenverzeichnis

- Tabelle 1 Übersicht über die Literaturrecherche (bis Juli 1999)*
- Tabelle 2 Hierarchie der wissenschaftlichen Evidenz bei klinischen Studien nach Antes 1999*
- Tabelle 3 Kosten pro ¹⁸FDG-PET-Untersuchung am Universitätsklinikum Ulm (PET-Zentrum mit klinikinterner Radiotracerproduktion), Stand 1997*
- Tabelle 4 Transparenzkatalog Onkologie: Indikationen NSCLC und SPN*
- Tabelle 5 Qualitätskatalog Onkologie: Indikationen NSCLC und SPN*
- Tabelle 6 Transparenzkatalog Onkologie: Sonstige Indikationsstellungen*
- Tabelle 7 Qualitätskatalog Onkologie: Sonstige Indikationsstellungen*
- Tabelle 8 Transparenzkatalog Kardiologie*
- Tabelle 9 Qualitätskatalog Kardiologie*
- Tabelle 10 Baseline-Annahmen für NSCLC-Studien nach Angaben der Originalstudien*
- Tabelle 11 Beschreibung der alternativen Strategien zur Diagnose und zum Staging von nicht-kleinzelligem Lungenkarzinom (NSCLC) im ECRI-Bericht 1998, Teil 1*
- Tabelle 12 Kosten und Lebenserwartung der alternativen Strategien zur Diagnostik von nicht-kleinzelligem Lungenkarzinom (NSCLC) nach ECRI-Bericht 1998, Teil 2*
- Tabelle 13 Alternative Diagnosestrategien bei Patienten mit Verdacht auf koronare Herzkrankheit nach Angaben von Maddahi & Gambhir 1997*
- Tabelle 14 Strategien der Wahl für die Diagnose von koronarer Herzkrankheit nach Angaben von Maddahi & Gambhir 1997*
- Tabelle 15 Strategien der Wahl für die Prognose von koronarer Herzkrankheit bei Patienten mit normalem bzw. abnormalen Elektrokardiogramm in Ruhe nach Angaben von Maddahi & Gambhir 1997*
- Tabelle 16 Übersicht über die onkologischen Studien*
- Tabelle 17 Übersicht über die kardiologischen Studien*

Abbildungsverzeichnis

Abbildung 1 *Abhängigkeit der Kosteneffektivität der PET bei den Indikationen „Staging von nicht-kleinzelligem Lungenkarzinom“ bzw. „Diagnose von solitären Lungenknoten“ (Indikationsgebiet Onkologie) von der Vortestwahrscheinlichkeit für Metastasenbildung durch den NSCLC-Primärtumor (pNSCLC) bzw. der Vortestwahrscheinlichkeit für maligne SPN (pSPN_{MAL})*

Abbildung 2 *Abhängigkeit der Kosteneffektivität der PET im Indikationsgebiet Kardiologie (Myokardperfusionsmessung) von der Vortestwahrscheinlichkeit für koronare Herzkrankheit (pCAD)*

Abkürzungsverzeichnis

ALND	Axilläre Lymphknotendisektion
CABG	Coronary artery bypass graft, Koronararterien-Bypass
CEA	Colon embryonal antibody
CT	Computer-Tomographie
DGN	Deutsche Gesellschaft für Nuklearmedizin
DRG	Diagnosis related group, US-amerikanische Diagnoseklassifikation
ECRI	Emergency Care Research Institute
EKG	Elektrokardiographie
ExEKG	Elektrokardiographie unter Belastung
¹⁸ FDG	[F-18]2-Deoxy-2-Fluoro-D-Glucose (Radiotracer)
⁶⁷ Ga	Gallium-67-Isotop
HD	Hodgkin-Lymphom, Synonym: Lymphogranulomatose
HTA	Health Technology Assessment
ICP	Institute for Clinical PET
iKER	Inkrementelle Kosten-Effektivitäts-Ratio
iKNR	Inkrementelle Kosten-Nutzwert-Ratio
i.v.	Intravenös
KHK	Koronare Herzkrankheit
Lk	Lymphknoten
MEL	Maligne Melanome
n	Anzahl
NHL	Non-Hodgkin-Lymphom
NPW / PPW	Negativer bzw. positiver prädiktiver Wert
NSCLC	Nicht-kleinzelliges Lungenkarzinom
NT	Nuklearmedizinischer Test
NYHAC	New York Heart Association
p.....	Vortestwahrscheinlichkeit für eine Krankheit
pCAD	Vortestwahrscheinlichkeit für koronare Herzkrankheit
PET	Positronen-Emissions-Tomographie
PTCA	Perkutane transluminale Koronarangioplastie
QALYs	Quality adjusted life years, qualitätsadjustierte Lebensjahre
⁸² Rb	Rubidium-82-Isotop
RCT	Randomisierte kontrollierte Studie
SPECT	Single-Photon-Emissions-Computer-Tomographie
SPN	Solitäre Lungenknoten
²⁰¹ Tl	Thallium-201-Isotop
UKG	Ultraschall-Echokardiographie

A Abstract

Fragestellung: In den 1990er Jahren fand die Positronen-Emissions-Tomographie (PET) in Deutschland zunehmend Eingang in die klinische Diagnostik der Indikationsgebiete Onkologie, Kardiologie und Neuromedizin. Wegen knapper finanzieller Ressourcen im Gesundheitswesen gewinnen Aspekte der Kosteneffektivität des Einsatzes der PET unter klinischen Bedingungen an Bedeutung. Insbesondere ist die Kosteneffektivität der PET in Relation zu eingeführten Alternativmethoden zu klären. Ziel dieses HTA-Berichtes ist die Identifikation und die Bewertung der vorliegenden literaturbasierten Evidenz zur Kosteneffektivität der PET.

Methodik: Hierzu wurde eine breit angelegte, systematische Literaturrecherche in sieben elektronischen Datenbanken durchgeführt. Um die Relevanz der ermittelten Literaturstellen beurteilen zu können, durchliefen diese ein zweistufiges Auswahlverfahren. Anschließend wurden die als relevant klassifizierten Artikel anhand eines Fragenkataloges zu Transparenz- und Qualitätskriterien ausgewertet. Mit Hilfe dieses vorgegebenen Auswertungsschemas wurde die methodische Qualität gesundheitsökonomischer Evaluationen erfaßt. Ergänzend wurde zu jeder ausgewerteten Studie ein Studienbericht, bestehend aus einer Beschreibung des Studiendesigns und der Ergebnisse sowie einer Bewertung der Studienqualität gegeben.

Ergebnisse und Bewertung: Mit Hilfe der Literaturrecherche konnten 14 onkologische Studien (6 begleitende Untersuchungen zu einer klinischen Studie, 8 Modellierungen) und 7 kardiologische Studien (alles Modellierungen) als relevant identifiziert werden. Dabei ist die methodische Qualität der kardiologischen Studien zum großen Teil als hochwertig einzustufen, während sich bei den onkologischen Studien primär die ökonomischen Evaluationen zu den Indikationen „Staging von nicht-kleinzelligem Lungenkarzinom“ und „Diagnose und Management von solitären Lungenknoten“ durch gute bis sehr gute methodische Qualität auszeichnen. Im Bereich Neuromedizin konnten keine relevanten Publikationen ausfindig gemacht werden.

Aufgrund einer Vielzahl von potentiellen Anwendungsgebieten für PET kann keine „globale“ Kosteneffektivität für „die PET“, sondern nur eine anwendungsspezifische Kosteneffektivität zu den einzelnen Indikationen, ermittelt werden. Insgesamt dürfte der Einsatz der PET nach der bisher erschienenen Fachliteratur aus den USA nur bei der Diagnostik von „nicht-kleinzelligem Lungenkarzinom“ als „kosteneffektiv“ beurteilt werden können. Um aus diesen Resultaten Unterstützung für den deutschen Entscheidungskontext ableiten zu können, ist eine adäquate Anpassung der Kosten- und der Effektstruktur notwendig.

Für die übrigen onkologischen Indikationsstellungen ist die derzeit verfügbare Evidenz zur Beurteilung der Kosteneffektivität des klinischen PET-Einsatzes nicht aus-

reichend. Die Resultate im Bereich Kardiologie sind entweder nicht aussagekräftig oder stimmen in ihrer Aussage nicht überein.

Schlußfolgerungen: Nur für einen kleinen Teil der Anwendungsfelder der PET-Diagnostik, speziell für die Diagnostik von Lungenkrebserkrankungen und für die kardiale Perfusionsmessung, liegen qualitativ akzeptable Informationen zur Kosteneffektivität vor.

Um die Lücken der gesundheitsökonomischen Evidenz in den anderen Anwendungsfeldern zu schließen, sind weitere, ökonomische Evaluationen auf der Grundlage von qualitativ hochwertigen klinischen Studien erforderlich. Dringender Bedarf besteht besonders bei dem bisher noch nicht evaluierten Indikationsgebiet Neurologie. Bei der Abschätzung des gesundheitsökonomischen Stellenwertes der PET - für ein Gesundheitssystem als Ganzes, aber auch für einzelne Betreibereinrichtungen - ist zu beachten, wie sich die einzelnen Anwendungsfelder für PET mit unterschiedlicher Kosteneffektivität, die Indikationsgebiete Onkologie, Kardiologie und Neurologie mit ihren spezifischen Indikationsstellungen, prozentual auf den Gesamteinsatz der PET-Technologie in der klinischen Praxis verteilen.

Problem: During the 1990ies positron-emission-tomography (PET) increasingly was introduced into clinical diagnostics in the areas of oncology, cardiology and neuromedicine in Germany. As financial resources in health care are becoming more scarce, issues in cost-effectiveness of applying PET in clinical practice are gaining in relevance. Especially, the cost-effectiveness of PET as compared with common alternative diagnostic methods is of interest. This HTA-report intends to identify and assess the current evidence on cost-effectiveness of PET as it is available in the medical literature.

Methods: A broad based, systematic literature research was conducted in seven electronic data bases. In order to assess the relevance of the identified references, these underwent a two-stage selection procedure. Subsequently, those articles classified as relevant were evaluated according to a list of criteria concerning transparency and quality. Using this standardised evaluation scheme, the methodological quality of the health economic evaluations was assessed. In addition, a short study report was produced for each study evaluated. This report consisted of a description of study design, results, and an evaluation of study quality.

Results and assessment: Based on the literature review, 14 oncological studies (6 studies alongside clinical trials, 8 modelling studies) and 7 cardiological studies (all modelling studies) were identified as relevant. The methodological quality of the cardiological studies to a great extent was judged as high, whereas in the field of oncology, especially the economic evaluation studies on “staging of non-small-cell lung carcinoma” and “diagnosis and management of solitary pulmonary nodules” showed good or very good quality. No relevant publications were identified in the area of neuromedicine.

As PET diagnostics can potentially be applied in a number of areas, it is not possible to derive a “global cost-effectiveness for PET”. Instead, only the cost-effectiveness of a specific use of the diagnostic technique in specified disease areas can be derived.

Summarising the currently published literature, the use of PET is only assessed as “cost-effective” in the diagnostic of “non-small-cell lung carcinoma”. In order to use these results in the support of decision making in the German context it is necessary to adequately adapt the cost as well as the effect side. For diagnostic in other areas of oncology, the current available evidence is not sufficient to judge the cost-effectiveness of the clinical use of PET. In the area of cardiology results are either not assertive or not consistent so far.

Conclusions: Just for a small number of areas, qualitative acceptable information on the cost-effectiveness of PET diagnostics is available. This is true for den diagnostic of lung cancer and for the measurement of perfusing of the heart. In order to overcome the existing gaps in health economic evidence, further economic evaluation studies are needed, at best on the basis of qualitatively high-ranking clinical studies. Significant need exists especially in the area of neuromedicine which has not been

investigated yet. When assessing the economic relevance of PET for a health system in total or for an individual provider, it is important to determine the distribution of PET in clinical practice over its various areas of use, as cost-effectiveness may differ among these.

B Executive Summary

Fragestellung: In den 1990er Jahren fand das nicht-invasive, nuklearmedizinische Diagnoseverfahren der Positronen-Emissions-Tomographie (PET) an den großen, medizinischen Zentren in Deutschland zunehmend Eingang in die klinische Diagnostik der drei Indikationsgebiete Onkologie, Kardiologie und Neuromedizin (Neurologie/Neurochirurgie/Psychiatrie). Wegen knapper finanzieller Ressourcen im Gesundheitswesen ist nicht nur die medizinische Effektivität dieser modernen, technisch sehr aufwendigen Großtechnologie von Interesse, vielmehr sind auch Aspekte der Kosteneffektivität des vermehrten Einsatzes der PET zu berücksichtigen. Insbesondere ist die Kosteneffektivität der PET in Relation zu eingeführten Alternativmethoden zu klären; beispielsweise im Vergleich mit konventionellen nicht-invasiven Diagnoseverfahren, wie der Computer-Tomographie (CT) oder der Single-Photon-Emissions-Computer-Tomographie (SPECT), und mit invasiven Methoden zu diagnostischen Zwecken, wie beispielsweise - im Indikationsgebiet Onkologie - der Entnahme von Gewebeproben zur histologischen Untersuchung mittels Biopsie, die auch als Goldstandard der onkologischen Diagnostik angesehen werden kann.

In diesem Zusammenhang sind folgende Fragen zu beantworten:

Für welche Indikationsgebiete besteht für die PET nach gesundheitsökonomischen Kriterien eine ausreichende literaturbasierte Evidenz, um eine vertretbare Kosteneffektivität im Vergleich zu alternativen, invasiven und nicht-invasiven Diagnosemethoden belegen zu können?

Reicht die gesundheitsökonomische Datenlage auf der Basis der bisher erschienenen Fachliteratur aus, um Entscheidungsträger bei gesundheitspolitischen Überlegungen und Weichenstellungen, beispielsweise bei der Neuordnung der Vergütung von PET-Leistungen oder der Großgeräteplanung, zu unterstützen?

Welche Maßnahmen sind geeignet, etwaige Lücken der gesundheitsökonomischen Evidenz im Bereich der klinischen Diagnostik mittels PET zu schließen?

Methodik: Um die aufgeworfenen Fragen beantworten zu können, wurde die derzeit verfügbare Fachliteratur ausgewertet, welche die ökonomische Bedeutung der PET vergleichend zu alternativen diagnostischen Verfahren bewertet. Dazu wurde eine breit angelegte, systematische Literaturrecherche in sieben elektronischen Datenbanken unter der Suchstrategie „(positron emission tomography or PET) and (cost or costs or cost-effectiveness or economic)“ durchgeführt.

Um die Relevanz der anhand dieser Suchstrategie ermittelten Literaturstellen für die anstehende Thematik beurteilen zu können, durchliefen diese ein zweistufiges Auswahlverfahren. In der ersten Stufe wurden Veröffentlichungen aussortiert, aus deren Titel oder Abstract kein Hinweis auf eine ökonomische Evaluation der PET hervor-

ging. Publikationen, die nicht im Volltext, sondern nur als Poster- oder Konferenzabstract verfügbar sind, wurden ausgeschlossen, da diese keine ausreichenden Informationen zur Beurteilung der Methodik und Relevanz der ökonomischen Studien liefern. Artikel, welche die erste Auswahlrunde passiert hatten, wurden nach folgenden Kriterien weiterselektiert: in die systematische Auswertung eingeschlossen wurden nur ökonomische Studien, in denen eine PET-basierte Diagnosestrategie einer alternativen Diagnosestrategie vergleichend - unter medizinisch-klinischen und ökonomischen Aspekten - gegenübergestellt wird, und die in englischer, französischer oder deutscher Sprache vorliegen. Die Literaturangaben aller in die systematische Analyse einbezogenen Veröffentlichungen wurden auf weitere, relevante Quellen hin durchgesehen.

Artikel, die erfolgreich die zwei Auswahlrunden durchlaufen hatten, wurden zur Einschätzung der gesundheitsökonomischen Evidenz der PET anhand eines Fragenkataloges zu Transparenz- und Qualitätskriterien ausgewertet. Dieser Transparenz- und Qualitätskatalog wurde von den Teilnehmern der ökonomischen Arbeitsgruppe „German Scientific Working Group Technology Assessment for Health Care“ entwickelt. Mit Hilfe dieses vorgegebenen Auswertungsschemas soll die methodische Qualität gesundheitsökonomischer Evaluationen und - soweit dies möglich ist - der zugrundeliegenden klinischen Studien erfaßt werden.

Um dem Leser die Ergebnisse dieser Analyse darzustellen, werden im vorliegenden HTA-Bericht zwei unterschiedliche Zugangswege gewählt: zuerst wird eine zusammenfassende tabellarische Übersicht über alle ausgewerteten Studien gegeben; diese wird sowohl durch Beschreibungen des Studiendesigns und der Ergebnisse als auch durch Bewertungen der Studienqualität zu den einzelnen relevanten Publikationen ergänzt. Dabei sind sowohl die tabellarische Übersicht als auch die Studienbeschreibungen und -bewertungen getrennt nach Indikationsgebieten der klinischen PET-Anwendung aufgeführt.

Ergebnisse und Bewertung: Mit Hilfe der Literaturrecherche konnten 14 onkologische Studien und 7 kardiologische Studien (plus eine kritische Stellungnahme) als relevant identifiziert werden (siehe Literaturverzeichnis Teil I). Im Bereich Neuromedizin konnten keine relevanten Publikationen gefunden werden. Zweiundzwanzig in der ersten Auswahlstufe als relevant erachtete Literaturstellen wurden nach Beurteilung des Volltextes aufgrund der Ausschlusskriterien von der näheren Analyse ausgeschlossen, davon 15 onkologische, 4 kardiologische und 3 neurologische Studien (siehe Literaturverzeichnis Teil II).

Indikationsgebiet Onkologie

Es wurden 14 onkologische Studien, 6 begleitende ökonomische Analysen zu klinischen Studien, darunter eine Übersichtsarbeit zu mehreren onkologischen Indikationsstellungen, und 8 mathematische Modellierungen oder entscheidungstheoretische Modelle, darunter ein HTA-Bericht, als relevant ausgewählt und entsprechend dem vorgegebenen Kriterienkatalog ausgewertet. Aufgrund der vielseitigen

gen Einsetzbarkeit der PET in der Onkologie beschäftigen sich diese Studien mit unterschiedlichen PET-Anwendungsfeldern: Fünf begleitende ökonomische Untersuchungen zu klinischen Studien und 2 Modellierungen verteilen sich auf die Indikationen Staging von malignen Melanomen, Lymphomen bzw. rezidivierendem kolorektalem Karzinom und Screening auf axilläre Lymphknotenmetastasen bei Mammakarzinom, so daß je Indikationsstellung mit 1 bis maximal 3 Untersuchungen nur eine geringe Studienzahl vorliegt. Da diese zudem methodische Mängel im Studiendesign und in der Ausführung der ökonomischen Analysen aufweisen, sind diese Untersuchungen von eingeschränkter Evidenzqualität und daher kaum geeignet gesundheitspolitische Entscheidungen zu unterstützen.

Die meisten Studien, insgesamt 9 Artikel, beschäftigen sich mit der Diagnose und/oder dem Staging und Management von nicht-kleinzelligem Lungenkarzinom (NSCLC) und solitären Lungenknoten (SPN). Bei diesen ökonomischen Studien handelt es sich um zwei begleitende ökonomische Untersuchungen zu klinischen Studien; zwei weitere Artikel beruhen auf einer mathematischen Modellierung des Institute for Clinical PET von eingeschränkter methodischer Qualität; dies ist als Einschränkung ihrer Evidenzqualität zu werten. Die 5 übrigen Analysen sind entscheidungstheoretische Modelle von guter bis sehr guter methodischer Qualität. Aufgrund der Ergebnisse dieser aussagekräftigen Analysen ergeben sich Hinweise auf einen kosteneffektiven Einsatz der PET - als Zusatzuntersuchung zum konventionellen CT-Verfahren - bei der Diagnostik von Lungentumoren (SPN und NSCLC). Speziell für die Indikationsstellung „Staging von NSCLC“ lassen sich erste Resultate zum Einsatz der PET additiv zur Routinemethode CT ableiten, die von gesundheitsökonomischen Entscheidungsträgern als kosteneffektiv gewertet werden dürften. Da diese Studien jedoch sowohl im Studiendesign (inkrementelle oder durchschnittliche Untersuchung) als auch in der zeitlichen Anordnung der diagnostischen Einzelbausteine CT und PET (CT/PET- bzw. PET/CT-Reihenfolge) differieren, sind die Resultate der Studien nicht uneingeschränkt vergleichbar. Um eine gesicherte Aussage zur ökonomischen Wertigkeit der PET in diesen Anwendungsfeldern machen zu können, sollten deshalb die bisherigen Erkenntnisse durch weitere, methodisch hochwertige ökonomische Evaluationen von vergleichbarem Studiendesign (u.a. inkrementeller Ansatz, langfristiger Zeithorizont und Erfassung von patientenbezogenen Effekten und Langzeitkosten, Diskontierung der Kosten und Effekte, Prüfung der Stabilität der Resultate anhand von Sensitivitätsanalysen) bestätigt werden. Um eine Hilfestellung bei gesundheitspolitischen Entscheidungen für Deutschland geben zu können, ist die Durchführung von ökonomischen Evaluationen, die auf die spezifischen deutschen Bedingungen bei der Versorgung von Krebspatienten zugeschnitten sind, oder die Übertragung internationaler Studienergebnisse durch adäquate Anpassung der Kosten- und der Effektstruktur notwendig.

Indikationsgebiet Kardiologie

Bei der Literaturrecherche für diesen HTA-Bericht wurden 8 Publikationen, 7 ökonomische Studien und eine fachliche Stellungnahme zu einer Originalstudie, ermittelt, die sich mit dem Kosteneffektivitätsvergleich zwischen der PET-Technologie und den

in der kardiologischen Diagnostik üblicherweise genutzten konventionellen diagnostischen Verfahren, so dem Goldstandard der Koronarangiographie und nicht-invasiven Alternativmethoden wie der nuklearmedizinischen Konkurrenzmethode SPECT, beschäftigen. Im Gegensatz zu den onkologischen Studien handelt es sich bei allen kardiologischen Evaluationen um Modellansätze, speziell um mathematische Modellierungen und entscheidungstheoretische Modelle (Entscheidungsbäume). Da das Applikationspektrum der PET-Technologie im kardiologischen Indikationsbereich wesentlich homogener ist als in der Onkologie, beschäftigen sich trotz der geringen Studienzahl 5 theoretische Modelle mit der Diagnose der koronaren Herzkrankheit mittels Perfusionsmessung und die restlichen zwei Modelle mit der Bestimmung der Lebensfähigkeit des Myokards bei Herzinfarktpatienten. Da jedoch letztere methodische Mängel aufweisen und somit von begrenzter Aussagekraft sind, können zu diesem Anwendungsbereich der PET noch keine Aussagen gemacht werden. Die meisten entscheidungstheoretischen Modelle und Modellierungen zur Myokardperfusionsmessung dagegen zeichnen sich durch eine gute bis sehr gute methodische Qualität aus. Da aber die Ergebnisse dieser Studien in ihrer Aussage bezüglich des kosteneffektiven Einsatzes der PET als Instrument zur Perfusionsmessung bei Verdacht auf koronare Herzkrankheit nicht übereinstimmen, kann auch zu diesem Punkt keine eindeutige, evidenzbasierte Stellungnahme abgegeben werden. Wie bereits für die Onkologie besteht auch in den kardiologischen Anwendungsbereichen der PET weiterhin ökonomischer Forschungsbedarf.

Schlußfolgerungen: Zusammenfassend beurteilt reicht die derzeitige Datenlage über die ökonomischen Auswirkungen des klinischen Einsatzes der PET auf der Basis der bisher erschienenen Literatur nicht umfassend für das gesamte Indikationsgebiet Onkologie aus, um gesundheitspolitische Entscheidungen zu unterstützen. Im Bereich der Diagnostik von „nicht-kleinzelligem Lungenkarzinom“ liegen jedoch erste Resultate vor, die von gesundheitspolitischen Entscheidungsträgern als kosteneffektiv gewertet werden dürften. Auch für die Diagnostik von „solitären Lungenknoten“ gibt es erste Hinweise auf eine Kosteneffektivität des klinischen Einsatzes der PET, die jedoch durch weitere ökonomische Evaluationen zu erhärten sind. Dringender Forschungsbedarf besteht im Bereich Kardiologie und besonders bei dem bisher noch nicht evaluierten Indikationsgebiet Neuromedizin.

Aufgrund einer Vielzahl von potentiellen Anwendungsgebieten für PET kann keine „globale“ Kosteneffektivität für „die PET“, sondern nur eine indikationsspezifische Kosteneffektivität, ermittelt werden. Zudem ist bei Abschätzung des gesundheitsökonomischen Stellenwertes der PET - für ein Gesundheitssystem als ganzes, aber auch für einzelne Betreibereinrichtungen - zu beachten, wie sich die einzelnen Anwendungsfelder für PET mit unterschiedlicher Kosteneffektivität, die Indikationsgebiete Onkologie, Kardiologie und Neuromedizin mit ihren spezifischen Indikationsstellungen, prozentual auf den Gesamteinsatz der PET-Technologie in der klinischen Praxis verteilen.

Um die bestehenden Lücken der gesundheitsökonomischen Evidenz zu schließen, sind weitere, ökonomische Evaluationen von hoher methodischer Qualität erforderlich. Dabei ist wichtig, daß sich diese auf methodisch gute, prospektive, kontrollierte klinische (Multicenter-)Studien mit hoher Fallzahl und, bei entscheidungstheoretischen Modellen, auf eine repräsentative Datenbasis stützen. Künftige ökonomische Evaluationen sollten möglichst die gesamte Zeitspanne der Folgewirkungen diagnostischer Interventionen sowohl auf den medizinischen Outcome als auch auf die Kostenentwicklung abdecken, langfristige gesundheitsbezogene Effektgrößen und möglichst auch indirekte Kosten einbeziehen sowie die Effekt- und Kostenseite anhand einer inkrementellen Analyse zueinander in Beziehung setzen. Die Stabilität der Studienergebnisse sollte anhand von statistischen Tests oder Sensitivitätsanalysen überprüft werden. Wichtig für die Unterstützung von Entscheidungsträgern innerhalb des deutschen Gesundheitssystems wäre, daß künftig spezifisch auf die Praxis an deutschen Kliniken zugeschnittene ökonomische Evaluationen zur PET-Thematik verfügbar sind.

C Hauptdokument

C.1 Policy Question

Seit den 1970iger Jahren richtet sich die Aufmerksamkeit aller Beteiligten im Gesundheitswesen auf die Verteilung der knapper werdenden Finanzmittel. Auch die zahlreichen Maßnahmen des Gesetzgebers zur Dämpfung der Gesundheitsausgaben konnten bei gleichzeitig fortschreitender Entwicklung der medizinischen Technologie das zunehmend dringliche Problem der Verteilung der begrenzten Ressourcen im Gesundheitssektor nicht dauerhaft entschärfen. Durch die aktuelle öffentliche Diskussion unter den betroffenen Interessengruppen - Ärzteschaft, Krankenhausträger, Krankenkassen und politische Entscheidungsträger - sensibilisiert, rückt die Erkenntnis verstärkt in das gesellschaftliche Bewußtsein, daß nicht alle technisch möglichen und medizinisch wünschenswerten, gesundheitsrelevanten Interventionen künftig auch finanzierbar sein werden. Eine unkontrollierte Diffusion aller Errungenschaften des medizinisch-technischen Fortschritts und deren kritiklose Übernahme in den Regelleistungskatalog der Krankenversicherungen ohne Rücksichtnahme auf die damit verbundenen finanziellen Konsequenzen gefährden das Haushaltsgleichgewicht der Krankenkassen. Die Erhaltung der Beitragssatzstabilität in der gesetzlichen Krankenversicherung jedoch ist, bedingt unter anderem durch deren Einfluß auf die Höhe der Lohnnebenkosten, von großer sozialpolitischer und wirtschaftlicher Relevanz.

In Zeiten enger finanzieller Spielräume im Gesundheitswesen können neue diagnostische oder therapeutische Technologien nicht allein unter dem Gesichtspunkt der medizinischen Wirksamkeit, Sicherheit oder Überlegenheit gegenüber konventionellen Methoden beurteilt werden. Künftig wird bei der Entwicklung und Markteinführung neuer Technologien auch die Wirtschaftlichkeitsprüfung eine immer wichtigere Rolle spielen. Das Wirtschaftlichkeitsgebot ist von Seiten des Gesetzgebers im § 12 des Sozialgesetzbuches V klar vorgegeben: „Leistungen, die nicht notwendig oder unwirtschaftlich sind, können Versicherte nicht beanspruchen, dürfen die Leistungsträger nicht bewirken und die Krankenkassen nicht bewilligen.“ Ferner sieht § 135 des Sozialgesetzbuches vor, daß die Bundesausschüsse der Ärzte und Krankenkassen im vertragsärztlichen Bereich (und wohl künftig auch deren Pendant im Krankenhausbereich) im Zuge der Diskussion über den Umfang des Leistungskataloges der gesetzlichen Krankenkassen nicht nur Empfehlungen hinsichtlich medizinischer Wirksamkeit von Leistungen – also auch diagnostischer Leistungen – geben, sondern auch zu deren Wirtschaftlichkeit. Wissenschaftlich fundierte Grundlagen zur Beurteilung der Wirtschaftlichkeit einer medizinischen Maßnahme werden durch

ökonomische Evaluationsstudien bereitgestellt. Dieser Bericht befaßt sich mit der ökonomischen Evidenz einer diagnostischen Technologie.

Die Beurteilung der Wirtschaftlichkeit einer medizinischen Technologie muß außer den unmittelbar im Rahmen einer Intervention anfallenden Kosten auch jene Kosten - aber auch die Effekte - berücksichtigen, die durch die Maßnahme in der Folgezeit induziert werden. Dies gilt für alle Arten medizinischer Interventionen, seien es Therapien, Präventions- oder Rehabilitationsmaßnahmen. In besonderem Maße trifft dies jedoch für diagnostische Verfahren zu, die bei ihrem Einsatz und zudem durch die Veranlassung weiterer diagnostischer und/oder therapeutischer Interventionen zwar Kosten verursachen, aber per se - ohne nachfolgende, effektive Therapiemaßnahmen - keinen unmittelbaren Einfluß auf den Gesundheitszustand des untersuchten Patienten haben.

Eine Großtechnologie, die in den 1990er Jahren Eingang in die klinische Diagnostik gefunden hat, ist die nuklearmedizinische Bilderfassung mittels Positronen-Emissions-Tomographie (PET). Einerseits liefert PET differenzierte Informationen über physiologische Prozesse im menschlichen Körper, die mit den etablierten anatomischen Bildgebungsverfahren, wie beispielsweise der Computer-Tomographie, nicht gewonnen werden können, andererseits ist ihre Anwendung mit hohen Kosten verbunden. Dieser Sachverhalt wirft die Frage auf, ob der Einsatz der PET-Technologie in der klinischen Routinediagnostik aus wirtschaftlicher Sicht befürwortet werden kann. Die wissenschaftliche Informationsgrundlage für die Beantwortung dieser Frage bereitzustellen, ist Ziel des vorliegenden HTA-Berichtes.

C.2 Einleitung

Die Positronen-Emissions-Tomographie (PET) wurde als wissenschaftliche Methode in den 1970er Jahre eingeführt und hat sich zwischenzeitlich zu einem modernen klinischen Verfahren entwickelt, das heute vor allem zur Diagnostik von onkologischen, neuropsychiatrischen und kardiovaskulären Erkrankungen eingesetzt wird. Die PET stellt das zur Zeit am weitesten entwickelte nuklearmedizinische Bildgebungsverfahren dar. Aufgrund der hohen Investitionskosten und der räumlichen Abhängigkeit von einem Zyklotron für die Produktion der meisten als Radiotracer verwendeten kurzlebigen Radionuklide steht die aufwendige PET-Technik weltweit jedoch nur an einer begrenzten Anzahl medizinischer Zentren zur Verfügung; so waren im Mai 1995 insgesamt nur 134 PET-Geräte registriert. Wegen der limitierten Verfügbarkeit der Geräte hat der klinisch-diagnostische Routineeinsatz der PET bisher nur einen beschränkten Umfang. In den letzten Jahren zeichnet sich jedoch ab, daß PET künftig, zumindest in den großen medizinischen Zentren, einen festen Platz in der klinischen Diagnostik erobern wird; so erhöhte sich die Zahl der PET-Zentren in Deutschland im Zeitraum 1992 - 1997 von 6 auf 26 (König 1993, Daten des Gesund-

heitswesens, Ausgabe 1999). Damit gewinnt auch die Frage der Kosteneffektivität des PET-Einsatzes in der klinischen Diagnostik zunehmend an Bedeutung.

Für rationale Entscheidungen über kollektive Mittel sollten Informationen über die Wirtschaftlichkeit medizinischer Maßnahmen verfügbar sein. Daher sollten, wie bei therapeutischen, so auch bei diagnostischen Verfahren nicht nur deren medizinische Wertigkeit in klinischen Studien, sondern auch ihre Kosteneffektivität mit Hilfe von ökonomischen Evaluationen belegt werden. Gerade für die Großtechnologie PET muß dieser Anspruch erhoben werden, da diese aufgrund ihres hohen apparativen und personellen Aufwandes in der Anschaffung und im Unterhalt teurer ist als die meisten anderen diagnostischen Verfahren. Zur Lösung dieses Problems kann die ökonomische Evaluation von Gesundheitsleistungen wertvolle Informationen zur Verfügung stellen.

Daten über die Kosten und Effekte einer (neuen) Technologie anhand von wissenschaftlich fundierten, aber dennoch praktikablen Methoden und Kriterien bereitzustellen, um deren Kosteneffektivität abschätzen zu können, ist eines der Ziele gesundheitsökonomischer Evaluationen. Kosteneffektivität erfordert, daß eine Versorgungsleistung mit einem Minimum an Kosten erstellt wird, oder daß bei einem vorgegebenen Budget ein Maximum an Versorgung resultiert (Leidl 1998).

Im vorliegenden HTA-Bericht soll versucht werden, durch Sichtung der zur Zeit verfügbaren Originalliteratur an ökonomischen Studien zur PET-Thematik einen Überblick über den derzeitigen Stand der literaturbasierten Evidenz zur Kosteneffektivität der klinischen Anwendung der PET-Technologie zu geben und - soweit dies die aktuelle Datenlage zuläßt - eine erste Einschätzung zur Wertigkeit der innovativen PET-Technologie aus gesundheitsökonomischer Sicht vorzunehmen.

C.3 Forschungsfragen

Bei der Untersuchung der Wirtschaftlichkeit der PET im klinischen Einsatz ergeben sich folgende zentrale Fragen:

Für welche Indikationsgebiete besteht für die Positronen-Emissions-Tomographie (PET) nach gesundheitsökonomischen Kriterien eine ausreichende literaturbasierte Evidenz, um eine vertretbare Kosteneffektivität im Vergleich zu alternativen nicht-invasiven und invasiven Diagnosemethoden belegen zu können?

Reicht die gesundheitsökonomische Datenlage auf der Basis der bisher erschienenen Fachliteratur aus, um Entscheidungsträger bei gesundheitspolitischen Überlegungen und Weichenstellungen, beispielsweise bei der Neuordnung der Vergütung von PET-Leistungen oder im Rahmen der Großgeräteplanung, zu unterstützen?

Welche Maßnahmen sind geeignet, etwaige Lücken der gesundheitsökonomischen Evidenz im Bereich der klinischen Diagnostik mittels PET zu schließen?

C.4 Methodik

Die Technologie der Positronen-Emissions-Tomographie (PET) wird bei einem heterogenen Spektrum von Indikationen aus den Bereichen Onkologie, Kardiologie und Neuromedizin (Neurologie/Neurochirurgie/Psychiatrie) als diagnostisches Instrument verwendet. Der Einsatz der PET ist jedoch mit hohen Kosten verbunden. Um beurteilen zu können, ob die Verwendung der PET aus gesundheitsökonomischer Sicht unterstützt werden kann, wurde eine systematische Literaturrecherche durchgeführt, um ökonomische Evaluationen zur PET-Thematik zu identifizieren. Dabei wurde eine sehr allgemein formulierte Suchstrategie angewandt, um möglichst keine relevanten ökonomischen Publikationen zu übersehen. Die identifizierten Artikel wurden unter Beachtung festgelegter Ein- und Ausschlußkriterien gesichtet.

C.4.1 Literaturrecherche

Um einen möglichst vollständigen Überblick über ökonomische Studien zur klinischen PET-Anwendung zu erhalten, wurde eine breitangelegte Literatursuche in den folgenden Literaturdatenbanken durchgeführt:

- PubMed Suchsystem der National Library of Medicine (NLM), National Institutes of Health (NIH, USA), enthält Medline und Premedline sowie Verweise auf Publikationen in sonstigen Internet-basierten Zeitschriften
- Database of Abstracts of Reviews of Effectiveness (DARE), National Health Services (NHS) Centre for Reviews and Dissemination der University of York
- Health Technology Assessment Database, NHS Centre for Reviews and Dissemination der University of York
- HealthStar im Internet Grateful Med Search Screen, National Library of Medicine (unter Ausschluß von Medline)
- Health Economic Evaluation Database (HEED), Office of Health Economics (OHE) Database Ltd.
- EMBASE Elsevier Science Database Products
- Health Meditec

Zudem wurde im Internet nach relevanten Publikationen zu allgemeinen gesundheitsökonomischen Aspekten der PET gesucht.

Tabelle 1: Übersicht über die Literaturrecherche (bis Juli 1999)

Beschreibung der Datenbank	Eingeschlossener Zeitraum	Suchstrategie	Anzahl der Treffer
PubMed Suchsystem der National Library of Medicine (NLM), National Institutes of Health (NIH, USA), enthält Medline und Premedline sowie Verweise auf Publikationen in sonstigen Internet-basierten Zeitschriften	1966 - Juli 1999	(positron emission tomography OR PET) AND (cost OR costs OR cost-effectiveness OR economic) NOT animal [MH]	527
Database of Abstracts of Reviews of Effectiveness (DARE), National Health Services (NHS) Centre for Reviews and Dissemination der University of York	1994 - Juli 1999	(positron emission tomography OR PET) AND (cost OR costs OR cost-effectiveness OR economic)	6
NHS Economic Evaluation Database (NEED), NHS Centre for Reviews and Dissemination der University of York	1994 - Juli 1999	(positron emission tomography OR PET) AND (cost OR costs OR cost-effectiveness OR economic)	5
Health Technology Assessment Database, NHS Centre for Reviews and Dissemination der University of York	einschließlich Juli 1999	(positron emission tomography OR PET) AND (cost OR costs OR cost-effectiveness OR economic)	5
HealthStar im Internet Grateful Med Search Screen, National Library of Medicine (unter Ausschluß von Medline)	1975 - Juli 1999	[(positron emission tomography OR PET) AND (cost OR costs OR cost-effectiveness OR economic)], Study Group „Human“, Medline overlap „Exclude MEDLINE references“	43
Health Economic Evaluation Database (HEED), Office of Health Economics (OHE) Database Ltd.	1995 - Mai 1999	PET OR positron emission tomography	19
EMBASE Elsevier Science Database Products, Health Meditec	1990 - Juli 1999	(positron emission tomography OR PET) AND (cost OR costs OR cost-effectiveness OR economic), NOT animal, LA=(Engl;Germ), PY>1989, NOT SPECT, (TOMOGRA?; PET)/TI	115

C.4.1.1 Suchstrategie

Für die Literaturrecherche wurde eine Suchstrategie mit den folgenden Begriffen eingesetzt:

positron emission tomography OR PET

AND

cost OR costs OR cost-effectiveness OR economic

Daneben wurden in allen durchsuchten Datenbanken Studien ausgeschlossen, die sich mit Tieren (im Englischen „Pet“) beschäftigen.

Diese allgemeine Suchstrategie wurde bei Sichtung aller elektronischen Literaturdatenbanken mit Ausnahme der Datenbasen EMBASE, HealthStar und HEED eingesetzt. Bei der Recherche in EMBASE und in HealthStar wurden MEDLINE-Zitate ausgeschlossen. Zudem wurde in EMBASE die Suchstrategie noch weiter eingeschränkt, indem beispielsweise nur deutsch- und englischsprachige Artikel zugelassen wurden. Dagegen wurde die Suchstrategie zur Durchsicht der HEED-Datenbank auf den weitgefästeren Suchbefehl „PET OR positron emission tomography“ ausgedehnt. Eine detaillierte Dokumentation der Literatursuche wird in Tabelle 1 gegeben.

C.4.1.2 Ein- und Ausschlußkriterien

Um die Relevanz der im Rahmen der umfangreichen Literaturrecherche identifizierten Publikationen beurteilen zu können, durchliefen diese Artikel ein zweistufiges Auswahlverfahren. In der ersten Stufe wurden Veröffentlichungen aussortiert, aus deren Titel oder Abstract kein Hinweis auf eine ökonomische Evaluation der Positronen-Emissions-Tomographie hervorging. Zudem mußte der Artikel im Volltext zugänglich sein. Publikationen, die nur als Poster- oder Konferenzabstract verfügbar waren, wurden ausgeschlossen, da diese keine ausreichenden Informationen zur Beurteilung der Methodik und Relevanz der ökonomischen Evaluation liefern. In diesen Fällen erfolgte unter dem Namen des Autors eine Suche nach weiteren Veröffentlichungen im Volltext. In der zweiten Stufe des Auswahlprozesses wurden die im Volltext vorliegenden Artikel daraufhin überprüft, ob es sich bei den Studien um vollständige ökonomische Evaluationen handelt. Eine vollständige ökonomische Evaluation zeichnet sich durch zwei Charakteristika aus: Zum einen müssen mindestens zwei alternative medizinische Technologien miteinander verglichen werden, wobei im vorliegenden Fall eine der zu vergleichenden Diagnosetechnologien die Positronen-Emissions-Tomographie sein mußte. Zum andern müssen sowohl Kosten als auch Effekte der verglichenen medizinischen Verfahren bestimmt und in Relation zueinan-

der gesetzt werden. Des weiteren konnten nur deutsch-, englisch- und französischsprachige Artikel berücksichtigt werden.

Nach Beendigung der Literatursuche und des Auswahlprozesses wurden die Referenzlisten aller als relevant erkannten Veröffentlichungen auf weitere interessante Literaturverweise durchsucht. Außerdem wurden die Referenzlisten systematischer Recherchen zu gesundheitsökonomischen Aspekten der PET im Hinblick auf noch nicht bekannte ökonomische Evaluationen überprüft. Das Ergebnis des zweistufigen Auswahlverfahrens ist unter Kapitel C.5.1 zusammengestellt.

Um die während des Reviewverfahrens im Rahmen der Erstellung dieses HTA-Berichtes neu erschienen, relevanten Publikationen zu erfassen, wurde im März 2000 eine erneute Medline-Recherche durchgeführt. Hierbei wurde eine im Februar 2000 veröffentlichte entscheidungstheoretische Studie einer japanischen Arbeitsgruppe zum Indikationsgebiet Onkologie identifiziert und nachträglich in den HTA-Bericht aufgenommen (Kosuda et al. 2000).

Artikel, die erfolgreich die zwei sukzessiven Auswahlverfahren durchlaufen hatten, wurden zur Einschätzung ihrer gesundheitsökonomischen Evidenz bezüglich des klinischen Einsatzes der Positronen-Emissions-Tomographie anhand eines Fragenkataloges zu Transparenz- und Qualitätskriterien analysiert (siehe C.4.2.1).

C.4.2 Auswertung

C.4.2.1 Transparenz- und Qualitätskatalog

Die im Auswahlverfahren thematisch als relevant klassifizierten Artikel wurden nach dem im Anhang aufgelisteten zweiteiligen „Transparenz- und Qualitätskatalog“ ausgewertet. Dieser Kriterienkatalog wurde in enger Kooperation zwischen den Teilnehmern der ökonomischen Arbeitsgruppe der „German Scientific Working Group Technology Assessment for Health Care“ entwickelt (Leidl et al. 1999). Beide Katalogteile setzen sich aus jeweils 56 Fragen zusammen. Ziel des Transparenzkataloges ist eine standardisierte Analyse und Beschreibung der Methodik, die in der zu beurteilenden ökonomischen Evaluation und - soweit dies möglich ist (nur bei begleitenden ökonomischen Analysen zu klinischen Studien) - der zugrundeliegenden klinischen Studie angewandt wurde. Neben allgemeinen Fragen zum Evaluationsrahmen, beispielsweise ob der Typ der ökonomischen Evaluation, die Perspektive und der Zeithorizont angegeben werden, beschäftigen sich die anderen Kriterien mit der Kosten- bzw. der Effektseite, der Präsentation der Ergebnisse, der Behandlung von Unsicherheiten und Aspekten der abschließenden Diskussion. Mit Hilfe des

Qualitätskataloges soll dann in einem zweiten Schritt eine Einschätzung der Qualität der angewandten Methodik ermöglicht werden.

Mit der Festlegung dieser Kriterienkataloge wurde ursprünglich die Intention verfolgt, das Instrumentarium für eine standardisierte Beschreibung und Beurteilung ökonomischer Evaluationen medizinischer Therapiemaßnahmen zur Verfügung zu stellen. Bei der PET handelt es sich jedoch nicht um ein Therapieverfahren, sondern um ein diagnostisches Verfahren. Hieraus ergeben sich Schwierigkeiten bei der Einschätzung einiger Kriterien dieser Fragenkataloge, da diese speziell auf die Beurteilung therapeutischer Behandlungen zugeschnitten sind. Diese Problematik kann besonders gut an der Frage zum Zeithorizont verdeutlicht werden, die im Qualitätskatalog wie folgt lautet:

Wurde ein für die Fragestellung angemessener Zeithorizont für Kosten und Gesundheitseffekte gewählt und angegeben? (Frage 7, Qualitätskatalog)

Aus gesundheitsökonomischer Sicht sind vor allem die Auswirkungen der Diagnose auf den gesundheitlichen Outcome als Folge eines veränderten Therapieregimes von Interesse. Erst durch Änderungen in der Therapiestrategie und somit im Patientenmanagement, eventuell infolge einer verbesserten Diagnosestellung, können relevante Kosteneinsparungen durch vermiedene Interventionen und langfristige Verbesserungen in der Lebenserwartung oder Lebensqualität der versorgten Patienten realisiert werden. In ökonomischen Evaluationen zu diagnostischen Verfahren, die auf klinischen Studien basieren, werden jedoch diese langfristigen Auswirkungen auf den medizinischen Outcome und die Kostenentwicklung zumeist nicht betrachtet. Ein Grund hierfür liegt in der Wahl des Zeithorizontes der zugehörigen klinischen Studie, der meistens lediglich die Zeitspanne der Diagnosefindung und eventuell deren Auswirkungen auf das weitere diagnostische und therapeutische Vorgehen umfaßt (siehe dazu C.4.2.2). Als Effektgrößen werden dabei oftmals nur die diagnostische Genauigkeit des zu evaluierenden Verfahrens (zum Beispiel die Anzahl der korrekt diagnostizierten Patienten) und eventuell deren Folgen auf das Patientenmanagement (zum Beispiel die Zahl der eingesparten diagnostischen und/oder therapeutischen Interventionen im Rahmen der Akutversorgung) betrachtet. Langfristige Wirkungen auf den medizinischen Outcome, wie Veränderungen im Gesundheitszustand der Patienten, bleiben - ebenso wie langfristige Kosten - unberücksichtigt. Da diese ökonomischen Evaluationen aus gesundheitsökonomischer Sicht einen zu kurzen Zeithorizont aufweisen, müßte in diesen Fällen zu der Frage 7 im Qualitätskatalog eine 0-Bewertung vergeben werden. Hatte die Studie jedoch im Extremfall explizit die Durchführung einer reinen Diagnoseevaluation zum Ziel (Definition der Fragestellung/Zielsetzung), in welcher der Zeithorizont per Definition auf die Zeitspanne der diagnostischen Untersuchung begrenzt wird und nur alternative Diagnose- und keine Versorgungsstrategien verglichen werden (Effektmaß: diagnostische Genauigkeit), müßte im Qualitätskatalog für Frage 7 trotz des unzureichenden Zeithorizontes eine 1-Bewertung vergeben werden. Dies würde jedoch zu einer Verzerrung des Bewertungsmaßstabes führen. Um dies zu vermeiden, wurde bei begleitenden ökonomi-

schen Evaluationen zu klinischen Studien, wenn kein Zeithorizont explizit vorgegeben ist (0-Bewertung im Transparenzkatalog), in der Beurteilung zur methodischen Qualität dieser Studien (Qualitätskatalog, Frage 7) nicht die Fragestellung per se zugrunde gelegt, sondern für welchen Zeitraum die Kosten und Effekte erfaßt wurden (zum Beispiel für die Phase der Diagnosestellung oder der diagnostischen und therapeutischen Akutversorgung). Wurden nur Größen ermittelt, die eine relativ kurze Zeitspanne abdecken, wie die diagnostische Genauigkeit oder Änderungen im Patientenmanagement, aber blieb der Einfluß der medizinischen Versorgung auf den gesundheitlichen Outcome und/oder die längerfristige Kostenentwicklung unberücksichtigt, wurde die Frage 7 im Qualitätskatalog mit einer $\frac{1}{2}$ Note bewertet. Bei theoretischen Modellen, deren nicht explizit definierter Zeithorizont sich über die Restlebenszeit erstreckt, wäre dagegen - sofern auch langfristige Kosten erfaßt würden - eine 1-Bewertung zu vergeben.

Die Problematik bei der Beantwortung von Frage 7 erschweren auch die Beantwortung weiterer Fragen des Kriterienkataloges, die sich darauf beziehen. Dies sind die Frage 15 im Qualitätskatalog und die Fragen 26 im Transparenz- und im Qualitätskatalog.

Problematisch bei Diagnosestudien ist auch die Beantwortung der Frage 15 (im Transparenz und Qualitätskatalog):

Wurden die in der Studie untersuchten Gesundheitszustände aufgeführt? (Frage 15, Transparenzkatalog)

In Diagnosestudien werden Effektgrößen, die in therapeutischen klinischen Studien gebräuchlich sind, wie Änderungen einer klinischen Größe (zum Beispiel der Auswurfraction des linken Herzventrikels) oder in der Lebenserwartung zumeist nicht untersucht, sondern oftmals wird als Effektgröße nur die diagnostische Genauigkeit des evaluierten Verfahrens und davon abgeleitet Änderungen im Patientenmanagement betrachtet. In diesen Fällen werden in der Studienbeurteilung Maße der diagnostischen Treffsicherheit, wie die „Anzahl korrekt klassifizierter Patienten“, als gesundheitsrelevante Größen anerkannt, da sie auf der Einstufung des gesundheitlichen Zustandes der Patienten durch die Diagnostik, beispielsweise als richtig-positiv oder als richtig-negativ, und somit als tatsächlich erkrankt oder gesund usw., basieren.

C.4.2.2 Zusammenfassende Beschreibung und Bewertung der als relevant klassifizierten Studien

Bei der Positronen-Emissions-Tomographie handelt es sich um ein modernes nuklearmedizinisches Verfahren, welches vor allem zur Diagnosestellung und zum Therapiemonitoring eingesetzt wird. Der Einfluß der PET auf die Lebensdauer oder Lebensqualität von Patienten erfolgt, wie oben beschrieben, meist nicht über die Diagnose, sondern erst über eine anschließende Therapie (Kuwert et al. 1996, Dietlein

et al. 1999). Der Fall, wo bereits die Diagnose zu einer Verbesserung der Lebensqualität führen kann, liegt dann vor, wenn die Diagnose einen eindeutig negativen Testbefund liefert (Kuwert et al. 1996).

In der Literatur wird ein hierarchisches Modell für die Effektivität eines diagnostischen Verfahrens angegeben, das sich wie folgt gliedert (Kuwert et al. 1996, Fletcher et al. 1998, Dietlein et al. 1999):

1. Technische Genauigkeit (technical efficacy): z.B. Abbildungsgenauigkeit der Methode
2. Diagnostische Genauigkeit (diagnostic accuracy efficacy), z.B. Sensitivität und Spezifität einer Diagnosemethode bei einer definierten Aufgabenstellung oder Anzahl der korrekt diagnostizierten Patienten
3. Einfluß auf die Diagnosestrategie (diagnostic thinking efficacy): z.B. Unterschiede in der subjektiven Einschätzung der Diagnosewahrscheinlichkeit durch den Kliniker in Abhängigkeit vom Informationsstand vor und nach der diagnostischen Untersuchung
4. Einfluß auf die Therapie (therapeutic efficacy): Änderungen im Patientenmanagement durch die diagnostische Information, wie z.B. Unterlassung von Operationen
5. Einfluß auf die gewonnenen Lebensjahre und die Lebensqualität (patient outcome efficacy) aufgrund der diagnostischen Information: z.B. Änderungen des Zugewinns an Lebenszeit oder an gesundheitsbezogener Lebensqualität infolge einer höheren diagnostischen Effektivität und einer daraus resultierenden Verbesserung im Patientenmanagement
6. Nutzen vom Standpunkt der Gesellschaft aus (societal efficacy).

Eine gröbere Einteilung dieser Abstufung der Effektivität eines diagnostischen Verfahrens faßt einzelne dieser sechs Kategorien zusammen. So können auf der Effektseite Größen mit unterschiedlichem Zeitbezug als Maß für die Wertigkeit einer Diagnostetechnologie dienen (Dietlein et al. 1999): Effektgrößen der Kategorien 2 und 3 werden als intermediäre Effekte zusammengefaßt, während sich die Kategorien 4 bis 6 auf finale Effektgrößen beziehen, wobei die Kategorien 5 und 6 primär langfristige finale Effekte widerspiegeln.

Durch die Analyse der Studien anhand des vorgegebenen Transparenz- und Qualitätskataloges wird dem Leser in erster Linie eine standardisierte Übersicht zur Methodik und Qualität der Studien an die Hand gegeben. Um darüber hinaus Detailinformationen über die einzelnen Studien bereitzustellen, zum Beispiel auf welcher Stufe innerhalb des beschriebenen hierarchischen Systems gesundheitsökonomische Evidenz vorliegt, wurden im Rahmen dieses HTA-Berichtes von jeder als relevant klassifizierten Studie zusätzlich ein Studienbericht, das heißt, eine zusammenfassende Beschreibung und Bewertung des Studienkonzeptes und der Studienergebnisse, erarbeitet.

C.4.2.3 Typen der ökonomischen Evaluation

Die Kosten und Effekte einer medizinischen Intervention bilden die Grundkomponenten einer ökonomischen Evaluation. In einer vollständigen ökonomischen Evaluation werden sowohl die Kosten als auch die Effekte, das heißt, die Wirkung oder Folgen einer medizinischen Intervention oder eines Programmes, beispielsweise eine lebensverlängernde Wirkung oder Folgen für die wirtschaftliche Produktivität, von zwei oder mehreren alternativen Strategien ermittelt und gegeneinander abgewogen. Bei der ökonomischen Evaluation einer medizinischen Versorgungsleistung lassen sich mehrere Kategorien von Effekten unterscheiden: klinische Parameter, restliche Lebensdauer, Lebensqualität, in Präferenzeinheiten bewertete Gesundheitszustände und in Geldeinheiten bewertete Effekte (Leidl 1998). Diesen Effektkategorien lassen sich fünf verschiedene Typen der ökonomischen Evaluation zuordnen.

- In *Krankheitskosten-Studien* werden die gesellschaftlichen Kosten einer oder mehrerer Krankheiten ermittelt, das heißt, es wird nur die Kostenseite und nicht auch die Effektseite berücksichtigt. Bei den Kosten lassen sich 3 Kategorien unterscheiden: direkte, indirekte und intangible Kosten. Die direkten Kosten stehen in unmittelbarem Zusammenhang mit einer Krankheit oder einer Gesundheitstechnologie und können unterteilt werden in medizinische Kosten und nicht-medizinische Kosten, welche beispielsweise Unterstützungsleistungen durch die Familie einschließen. Indirekte Kosten stehen in einem mittelbaren Zusammenhang mit einer Erkrankung. In erster Linie handelt es sich bei dieser Kategorie um Kosten durch Produktionsausfall durch Fehlzeiten, verringerte Produktivität oder vorzeitigen Tod. Zu den intangiblen Kosten wird zum Beispiel Leid und Schmerz einer erkrankten Person gezählt.

Eine Gegenüberstellung von Alternativen ist nicht das essentielle Ziel einer Krankheitskostenstudie. Mit diesem Studientyp läßt sich die mit einer bestimmten Krankheit verbundene ökonomische Last beschreiben, die bei gesundheitspolitischen Entscheidungen eine Rolle spielen kann. Da keinerlei Effekte berücksichtigt werden, eignen sich die ökonomischen Analysen dieses Studientyps nicht für die Unterstützung von Allokationsentscheidungen zwischen verschiedenen Interventionen.

- *Kosten-Minimierungs-Studien* untersuchen ebenfalls ausschließlich die Kostenseite einer medizinischen Intervention, allerdings im Vergleich zu einer oder mehreren alternativen Interventionen. Die Effektseite bleibt auch hier von der Betrachtung ausgenommen, denn implizit wird bei diesem Studientyp von einer gleichen Effektivität, sowohl in der Ausprägung als auch im Grad des Effektes (zum Beispiel Erhöhung des Blutdruckes), ausgegangen. Die Richtigkeit dieser Annahme sollte jedoch sorgfältig geprüft werden. Ziel von Kosten-Minimierungs-

Studien ist es, aus zwei oder mehreren Alternativen die kostengünstigste Methode herauszufiltern.

- Auch der Studientyp der *Kosten-Effektivitäts-Analyse* erfordert eine Gegenüberstellung von Alternativen, wobei jedoch sowohl die Kosten als auch die Effekte der miteinander verglichenen Interventionen betrachtet werden (siehe Formel I). Die Effekte werden dabei in „natürlichen“ Einheiten gemessen, wie klinische Parameter, Lebensdauer oder Lebensqualität. Im Mittelpunkt dieser Evaluationsform steht der Vergleich einer Intervention mit einer Basisalternative. Setzt man die Kosten und Effekte zweier Alternativstrategien zueinander direkt ins Verhältnis, ergibt sich eine inkrementelle Kosten-Effektivitäts-Relation, die angibt, in welcher Höhe zusätzliche Kosten entstehen, um eine zusätzliche Einheit des Effekts zu erzeugen; beispielsweise als Mehrkosten pro gewonnenem Lebensjahr. Nur Ergebnisse von Studien mit dem gleichen Effektparameter können untereinander verglichen werden.

Inkrementeller Vergleich zweier Maßnahmen (Kosten und Effekte):

Kosten-Effektivitäts-Analyse

Formel (I) $[(DK_1 + IK_1) - (DK_2 + IK_2)] / (EF_1 - EF_2)$

Kosten-Nutzwert-Analyse

Formel (II) $[(DK_1 + IK_1) - (DK_2 + IK_2)] / (NW_1 - NW_2)$

Kosten-Nutzen-Analyse

Formel (III) $(G_{E1} - G_{E2}) - [(DK_1 + IK_1) - (DK_2 + IK_2)]$

DK_1, DK_2 direkte Kosten der Alternativverfahren 1 und 2

IK_1, IK_2 indirekte Kosten der Alternativverfahren 1 und 2

EF_1, EF_2 durch Alternativverfahren 1 und 2 induzierte Effekte

NW_1, NW_2 durch Alternativverfahren 1 und 2 induzierte Nutzwerte
(z.B. QALYs)

G_{E1}, G_{E2} mit Geld bewertete Effekte (Nutzen) der miteinander
vergleichenen Alternativverfahren 1 und 2

- In *Kosten-Nutzwert-Analysen* (siehe Formel II) wird die Wirkung einer Intervention auf den Gesundheitszustand der behandelten Patienten in die Betrachtung einbezogen. Essentiell bei diesem Evaluationstyp ist, daß der Gesundheitszustand auf der Basis von individuellen Präferenzen in Nutzeinheiten (zwischen 0 für Tod und 1 für vollständige Gesundheit) bewertet wird. Diese Nutzeinheit wird über die Zeit zu Nutzwerten zusammengefaßt (zum Beispiel in Form von qualitätsadjustierten Lebensjahren, QALYs). Ziel einer solchen Analyse ist - analog zur Kosten-Effektivitäts-Analyse - ein inkrementelles Kosten-Nutzwert-Verhältnis. Der Vorteil von Kosten-Nutzwert-Studien liegt darin, daß Interventionen mit verschiedenen medizinischen Effekten miteinander verglichen werden können.

- Bei *Kosten-Nutzen-Analysen* (siehe Formel III auf der vorhergehenden Seite) werden neben den Kosten auch alle gesundheitlichen Effekte, egal ob Auswirkungen auf das Patientenmanagement, wie eine verringerte Anzahl von Operationen, oder Veränderungen in der Lebensdauer bzw. der Lebensqualität, in Geldeinheiten bewertet. Nach einem Literaturreview (Zarnke et al. 1997) werden in mehr als 50% aller sogenannten Kosten-Nutzen-Analysen nur die Veränderungen in den Kosten, die beispielsweise durch vermiedene operative Eingriffe entstehen, als monetär bewerteter Nutzen verbucht. Derartige Änderungen im Patientenmanagement dürfen aber entweder ausschließlich bei der Berechnung der Gesamtkosten pro Versorgungsstrategie auf der Kostenseite veranschlagt und nicht zusätzlich als Nutzen monetär bewertet werden oder auf der Kostenseite dürfen nur die Kosten für die reine Diagnosestellung berücksichtigt und gegen die Einsparungen an Therapiekosten verrechnet werden. Ansonsten würden gemäß Formel (III) bei der Berechnung der Gesamtkosten von Versorgungsstrategien (Kosten für diagnostische + therapeutische Maßnahmen) die Kosten für durch Methode 1 vermeidbare therapeutische Eingriffe (beispielsweise wirkungslose Operationen bei inoperablen Patienten) zweimal, einerseits als Anteil an den direkten medizinischen Kosten DK_2 und andererseits als monetär bewerteter Nutzen G_{E2} , in die Berechnung eingehen. Dies ist jedoch ein methodischer Fehler, den es zu vermeiden gilt.

Eine Kosten-Nutzen-Studie, in der neben diagnosebedingten Änderungen im Versorgungsablauf gesundheitliche Effekte einer Technologie (patientenbezogener Outcome) gar nicht oder nur unvollständig erhoben oder diese nicht monetär bewertet werden, entweder durch die Bestimmung des Humankapitals oder besser und allgemeiner durch die Bestimmung der Zahlungsbereitschaft auf der Basis der individuellen Präferenzen, stellt lediglich eine partielle ökonomische Evaluation dar (Zarnke et al. 1997).

Eine partielle Evaluation liegt allgemein dann vor, wenn entweder eine Maßnahme nicht mit einer alternativen Maßnahmen verglichen wird, oder nur die Kosten oder nur die Effekte einer oder mehrerer Interventionsmaßnahmen (vergleichend) betrachtet werden. Eine ökonomische Evaluation ist dann vollständig, wenn die Kosten und die Effekte von mindestens zwei Maßnahmen einander gegenübergestellt werden. Von den hier vorgestellten Evaluationstypen verkörpern nach Drummond et al. (1997) die Kosten-Minimierungs-Analyse, die Kosten-Effektivitäts-Analyse, die Kosten-Nutzwert-Analyse und die Kosten-Nutzen-Analyse vollständige Evaluationen.

Ökonomische Analysen sollten in der Regel dem inkrementellen Ansatz folgen, das heißt, es werden die zusätzlichen Kosten pro zusätzlicher Effekteinheit der zu evaluierenden Maßnahme im Vergleich zu einer etablierten Alternative angegeben. In der vorliegenden Entscheidungssituation soll die Kosteneffektivität des additiven oder des substitutiven Einsatzes der PET in Relation zu alternativen diagnostischen Maßnahmen, beurteilt werden. Insbesondere wenn mehr als zwei Alternativen vorliegen, kann es dabei zu komplizierten Vergleichsbetrachtungen kommen. Liegt ein derarti-

ger Entscheidungskontext vor, sollte bei einer inkrementellen Analyse, die das Ziel verfolgt, die Kosteneffektivität zu maximieren, folgende Entscheidungsregel beachtet werden (Johannesson 1993):

1. Die Maßnahmen werden nach ihrer medizinischen Wirksamkeit (oder alternativ anhand der aufzubringenden Kosten) aufsteigend geordnet, wobei dominante Maßnahmen aussortiert werden.
2. Dann wird entsprechend der Rangfolge der Effektivität (Kosten) für die aufeinanderfolgenden Maßnahmen die inkrementelle Kosten-Effektivitäts-, Kosten-Nutzwert- bzw. Kosten-Nutzen-Relation berechnet (siehe Formeln (I) – (III)).

Diese Regel wird bei der Durchführung von inkrementellen Analysen häufig nicht angewendet. So werden oftmals mehrere Alternativmaßnahmen zu einer Referenzmaßnahme, beispielsweise der üblichen Routinemethode, und nicht zueinander entsprechend ihrer Effektivitäts- bzw. Kostenrangfolge in Beziehung gesetzt. Auch wenn diese Vorgehensweise im strengen Sinne nicht ganz korrekt ist, wird im vorliegenden HTA-Bericht vereinfachend eine ökonomische Analyse, die diesem Konzept folgt, ebenfalls als inkrementelle Analyse anerkannt und durch die Angabe der Referenzmaßnahme bzw. Referenzstrategie gekennzeichnet.

Studien im Rahmen dieses HTA-Berichtes, die dem Konzept der inkrementellen Analyse (gemäß der Effektivitäts- bzw. Kostenrangfolge oder im Vergleich zu einer Referenzmaßnahme) entsprechen, werden bei der Charakterisierung des Studientyps durch den Vorsatz „inkrementell“ hervorgehoben.

C.4.2.4 Bewertung der verfügbaren Evidenz

Obwohl ein allgemein akzeptierter Maßstab für die qualitative Einschätzung der Resultate der unter C.4.2.3 dargestellten Formen der ökonomischen Evaluation derzeit nicht zur Verfügung steht, wird in einer abschließenden Diskussion versucht, eine Gesamtbewertung und Würdigung der verfügbaren ökonomischen Evidenz zum Einsatz der PET in der klinischen Diagnostik vorzunehmen (siehe dazu Kapitel C.6). Für klinische Studien dagegen wird in der Fachliteratur ein hierarchisches System der Aussagekraft der wissenschaftlichen Evidenz anerkannt, das in Tabelle 2 dokumentiert ist (Antes 1999).

Eine analoge, ebenso einfache Evidenzskala für ökonomische Evaluationen kann es nicht geben, da schon die spezifischen, ökonomisch relevanten Charakteristika des jeweiligen Gesundheitssystems, in dem die Studie durchgeführt wurde, eine Verallgemeinerung einschränken. Ähnlich der Diskussion im medizinischen Bereich werden ferner Designfragen debattiert. Der Goldstandard ist zur Zeit eine ökonomische Evaluation, die auf einer randomisierten kontrollierten Studie (RCT) fußt. In der Literatur werden jedoch Nachteile dieses Goldstandards diskutiert, durch welche die externe Validität der Ergebnisse, das heißt deren Realitätsbezug, eingeschränkt wird

(Baltussen et al. 1999). Als Nachteil wird beispielsweise genannt, daß die durch das künstliche Studiendesign induzierten Kosten von den Kosten, die im klinischen Alltag anfallen, abweichen.

In Fällen, in denen die Durchführung einer RCT-basierten ökonomischen Evaluation nicht praktikabel (sehr langer Zeithorizont) oder unzulässig ist (Patienten kann die neue medizinische Intervention nicht vorenthalten werden), werden häufig theoretische Modellierungen zum Nachweis der Kosteneffektivität verwandt. Die Ergebnisse dieser ökonomischen Evaluationen basieren zwar nicht auf einer speziell auf die untersuchte Fragestellung zugeschnittenen Studie mit experimentellem Design, jedoch wird in eine hochwertige Modellierung die zum Zeitpunkt der Durchführung beste verfügbare systemrelevante klinische und ökonomische Evidenz einbezogen (Baltussen et al. 1999). Aus diesem Grunde weisen Gesundheitsökonomien Resultaten einen hohen Evidenzgrad zu, die einer ökonomischen Evaluation auf der Basis einer qualitativ guten Modellierung entstammen.

Tabelle 2: Hierarchie der wissenschaftlichen Evidenz bei klinischen Studien nach Antes 1999

Stufe	Evidenz-Typ
I	Stärkste Evidenz: wenigstens ein systematischer Review auf der Basis methodisch hochwertiger randomisierter kontrollierter Studien (RCTs)
II	Wenigstens ein ausreichend großer, methodisch hochwertiger RCT
III	Methodisch hochwertige Studien ohne Randomisierung (z.B. Kohorten-Studien) bzw. nicht prospektive Studien (z.B. Fall-Kontroll-Studien)
IV	Mehr als eine methodisch hochwertige, nicht-experimentelle Studie
V	Schwächste Evidenz: Meinungen und Überzeugungen von Autoritäten und Expertenkommissionen (ohne transparente Belege), beschreibende Studien

Neben Problemen bezüglich der richtigen Einordnung der Evidenz ökonomischer Evaluationen müssen weitere spezifische Punkte kritisch beleuchtet werden, die durch den speziellen Untersuchungsgegenstand dieses HTA-Berichtes, der Positronen-Emissions-Tomographie (PET), hervorgerufen werden. Da es sich bei der PET - unter anderem wegen der Vielfalt an verfügbaren radioaktiven Tracersubstanzen - um ein vielseitig einsetzbares diagnostisches Verfahren handelt, wird es in einem heterogenen Spektrum von Indikationsgebieten und zu unterschiedlichen Zwecken, wie Primärdiagnose, Staging/Restaging und Therapiekontrolle, eingesetzt. Demnach muß eine Einschätzung der vorliegenden ökonomischen Evidenz sowohl nach Indikationsgebieten als auch nach Verwendungszweck differenziert erfolgen. Abgesehen davon werden auch in Abhängigkeit von Indikationsgebiet und Verwendungszweck PET-Untersuchungen nach unterschiedlicher Methodik durchgeführt. So ergeben sich Unterschiede bei den eingesetzten Radiopharmaka, bei den untersuchten Kör-

perarealen (Ganzkörper-PET gegenüber PET-Aufnahmen bestimmter Körperregionen) und der technischen Ausstattung der Geräte.

Neben der Berücksichtigung dieser differenzierten Einsatzmöglichkeiten der PET können nur gesundheitsökonomische Resultate in eine abschließende Beurteilung einbezogen werden, die ökonomischen Evaluationen mit einer hinreichend guten methodischen Qualität entstammen.

C.5 Ergebnisse

Zu Beginn dieses Ergebnisteils werden die Resultate der Literaturrecherche kurz vorgestellt (siehe C.5.1). Explizit genannt werden an dieser Stelle nur Veröffentlichungen, die den ersten Auswahlprozeß erfolgreich überstanden haben.

Der Kostenaspekt war und ist ein wichtiger Grund für die jahrzehntelange staatliche Regulierung der Standorte medizinischer Großgeräte, wie der Magnet-Resonanztomographie oder der Positronen-Emissions-Tomographie. Gerade nuklearmedizinische Diagnoseverfahren werden häufig aufgrund ihrer hohen Kosten kritisch betrachtet. Aus welchen Einzelpositionen sich die Kosten für die PET zusammensetzen, wird Gegenstand des Kapitels C.5.2 sein. Falls möglich, werden für die Kostenkomponenten konkrete Beträge für Deutschland genannt. Zusätzlich werden Konzepte vorgestellt, mit denen versucht wird, mögliche Einsparungspotentiale zu realisieren.

Danach wird die gesundheitsökonomische Evidenz zur PET dargelegt, die von den als relevant einzustufenden Studien zur Verfügung gestellt wird. Dies erfolgt auf zwei Arten: zum einen werden die Publikationen anhand der Kriterien des Transparenz- und Qualitätskataloges analysiert und die Resultate dieser Analysen in einer tabellarischen Darstellung übersichtlich präsentiert (siehe C.5.3); zum andern wird jeweils auf der Basis dieser Analysen eine methodische Bewertung der einzelnen Studien gegeben, der - zum besseren Verständnis - jeweils eine zusammenfassende Studienbeschreibung vorangestellt ist (siehe C.5.4).

C.5.1 Ergebnisse der Literatursuche

Zuerst erfolgt eine Bestandsaufnahme der vorliegenden Literatur, aufgegliedert nach Indikationsgebiete des klinischen PET-Einsatzes:

Onkologie

Durch die umfangreiche Literaturrecherche wurden 15 relevante Veröffentlichungen (darunter ein HTA-Bericht) zu verschiedenen onkologischen Indikationen identifiziert.

Davon dokumentieren 6 Publikationen die Resultate ökonomischer Evaluationen, die parallel zu einer klinischen Studie durchgeführt wurden. Die ökonomischen Evaluationen der restlichen 9 Artikel basieren auf mathematischen Modellierungen oder entscheidungstheoretischen Modellen in Form von Entscheidungsbäumen.

Kardiologie

Im Gegensatz zum onkologischen Indikationsgebiet handelt es sich bei den 7 für das Indikationsgebiet Kardiologie als relevant erachteten Publikationen (plus eine kritische Stellungnahme zu einer ökonomischen Studie) ausschließlich um mathematische Modellierungen oder um entscheidungstheoretische Modelle in Form von Entscheidungsbäumen.

Neuromedizin (Neurologie/Neurochirurgie/Psychiatrie)

Für das Indikationsgebiet der Neuromedizin konnten keine aus gesundheitsökonomischer Sicht relevanten Veröffentlichungen ausfindig gemacht werden.

Literatur zu Kostenaspekten der PET

Zusätzlich zu diesen 23 Publikationen wurden noch 3 Artikel und ein Konferenzvortrag als relevant klassifiziert, die sich lediglich mit einer detaillierten Erfassung der Kosten der PET beschäftigen (siehe Literaturverzeichnis Teil I).

Des Weiteren wurde ein Review und ein HTA-Bericht zur ökonomischen Evidenz von PET als relevant eingestuft (siehe Literaturverzeichnis Teil I).

Die folgenden Artikel wurden bei der ersten Durchsicht als mögliche relevante Veröffentlichungen klassifiziert, jedoch bei der näheren zweiten Kontrolle aussortiert:

- zwei Veröffentlichungen in Japanisch und eine in Italienisch aufgrund der Sprache, da keine Übersetzungen vorliegen: Kubota et al. 1997, Kosuda et al. 1998, Giordano et al. 1999
- neun Veröffentlichungen, die nur als Abstract zugänglich sind: Oncology 1996, Duong 1996, Larson 1995, Madar et al. 1995, Mitchell et al. 1998, Morrison & Coates 1995; Steinert et al. 1998a, Steinert et al. 1998b, Yao et al. 1994
- neun Studien, deren Analysen keine ökonomischen Evaluationen enthalten: Di-Bisceglie et al. 1998, Gupta et al. 1995, Holmberg et al. 1997, Kuwert et al. 1998, Pai et al. 1999, Podoloff 1995, Schelbert 1996, Small & Leiter 1998, Stanford 1996
- vier Übersichtsarbeiten zu ökonomischen Aspekten der PET in speziellen Indikationsgebieten und zwei allgemeine Betrachtungen zur Durchführung ökonomischer Evaluationen zum Einsatz von PET: Chiti et al. 1999, Dietlein et al. 1999, Gardner et al. 1992, Kuwert et al. 1996, Lowe & Naunheim 1998, Vock 1997

Die genauen Quellenangaben zu diesen Artikeln sind im Literaturverzeichnis Teil II aufgeführt.

C.5.2 Die monatlichen Kosten der Positronen-Emissions-Tomographie

Die Positronen-Emissions-Tomographie (PET) gehört zur Klasse der „big ticket technologies“. Unter dieser Bezeichnung werden Technologien subsumiert, bei denen die einzelne Anwendung sehr teuer ist. Da diese Begriffsdefinition nicht einheitlich verwendet wird, werden manchmal auch solche Technologien dieser Klasse zugeordnet, die durch wiederholte, regelmäßige und häufige Anwendung hohe Kosten verursachen, wie beispielsweise die medikamentöse Behandlung einer chronischen Erkrankung (König 1993). PET gehört, wie beispielsweise auch die Computertomographie oder die Magnet-Resonanz-Tomographie, zu den Technologien, deren Anwendung an sich bereits kostenintensiv ist. In dem folgenden Kapitel soll aufgezeigt werden, wodurch die hohen Kosten der PET induziert werden. Dazu wird ein Konzept zur Aufschlüsselung der Gesamtkosten in einzelne Kostenpositionen vorgestellt. Soweit monetäre Werte für diese Kostenpositionen für Deutschland der Literatur zu entnehmen sind, werden diese genannt. Zum Abschluß werden noch verschiedene Ansätze dargestellt und diskutiert, mit denen versucht wird, mögliche Einsparpotentiale zu erschließen.

Die Bereitstellung der Positronen-Emissions-Tomographie erfordert eine komplexe und teure Infrastruktur (Conti et al. 1994). Zur Herstellung der meisten Radiopharmaka¹, beispielsweise ¹⁸Fluor-Deoxyglukose (FDG), werden ein Zyklotron und ein radiochemisches Labor benötigt. Mit dem Zyklotron werden radioaktive Nuklide, beim ¹⁸FDG das Fluor-18-Isotop, erzeugt, welche dann im radiochemischen Labor an die physiologisch wirksame Substanz, in diesem Fall an die Deoxyglukose, gebunden werden. Die eigentliche Bilderfassung der spezifischen räumlichen Verteilung des Radiotracers im Körper wird dann mit einem PET-Scanner durchgeführt (Evens et al. 1983, Conti et al. 1994). Diese grobe Beschreibung der Arbeitsabläufe verdeutlicht bereits, daß bei einer Beschreibung der Kostenstruktur der PET neben den Kosten für die Durchführung eines PET-Scans auch die Kosten zur Herstellung des jeweils verabreichten Radiodiagnostikums angemessen berücksichtigt werden müssen.

Die Kostenstrukturen bei der Herstellung der benötigten Radiodiagnostika und bei der Durchführung einer PET-Aufnahme sind vergleichbar. Die Gesamtkosten teilen sich in Fixkosten und in variable Kosten auf. Die Fixkosten der PET fallen auch dann an, wenn keine einzige Aufnahme mit einem PET-Scanner durchgeführt wird; sie sind unabhängig von der Leistungsmenge beziehungsweise vom Auslastungsgrad

¹ Es gibt auch PET-Anwendungen, bei denen Radiotracer appliziert werden, beispielweise im Bereich der Kardiologie Rubidium-82, die nicht mit Hilfe eines Zyklotrons hergestellt werden (Gould et al. 1989). Auf diese zyklotronunabhängige Radionuklidproduktion wird hier nicht näher eingegangen.

des Zyklotrons und des PET-Scanners. Die variablen Kosten sind leistungsmengensabhängig.

Die Anschaffung der benötigten technischen Ausstattung, eine Produktionseinheit zur Herstellung von Radiopharmaka, bestehend aus einem Zyklotron und einem radiochemischen Labor, sowie ein PET-Scanner, ist mit hohen Investitionskosten verbunden. Diese Kosten werden in Kostenanalysen der PET über Abschreibungen einbezogen (Evens et al. 1983, Gould et al. 1989, Conti et al. 1994, Lottes et al. 1998, Lottes & Schober 1999, Klose et al. 1999a). Für Deutschland können konservativ geschätzte monatliche Abschreibungsbeträge (lineare Abschreibung über 10 Jahre, Bezugsjahr 1998) von 27 000 DM für eine Produktionseinheit für Radiopharmaka² und zwischen 23 000 – 33 000 DM für einen PET-Scanner (Klose et al. 1999a; Lottes & Schober 1999) abgeschätzt werden. Falls für diese Investitionen eine Fremdfinanzierung in Form von Krediten notwendig ist, müssen die hierdurch anfallenden Zinsen ebenfalls in den Fixkosten erfaßt werden (Gould et al. 1989; Conti et al. 1994). Vor der Installation einer radiochemischen Produktionseinheit oder eines PET-Scanners sind in vielen Krankenhäusern bauliche Maßnahmen zur Bereitstellung der erforderlichen Räumlichkeiten notwendig. Neben allgemeinen Baumaßnahmen fallen insbesondere auch Maßnahmen an, um strahlenschutzrechtliche Vorschriften zu erfüllen, wie beispielsweise die Errichtung einer Strahlenschutzwand am Aufstellungs-ort des Zyklotrons. Kosten, die durch Neu- oder Umbau anfallen, werden ebenfalls über Abschreibungen den Fixkosten zugerechnet (Conti et al. 1994). Neben diesen anschaffungsbedingten Kosten verursacht die Wartung der Geräte weitere leistungsmengenunabhängige Kosten, die sich bezogen auf Deutschland im Jahr 1998 monatlich in einer Größenordnung von 16 000 DM für die radiochemische Produktionseinheit und von 16 000 – 20 000 DM für den PET-Scanner bewegten (Klose et al. 1999a; Lottes & Schober 1999).

Die Kosten, die von der erbrachten Leistungsmenge abhängen, setzen sich aus den Aufwendungen für Personal und Material zusammen (Evens et al. 1983, Gould et al. 1989, Conti et al. 1994, Lottes et al. 1998, Lottes & Schober 1999, Klose et al. 1999a). In der Radiopharmakaproduktion werden Techniker, Ingenieure und Chemiker beschäftigt (Klose et al. 1999a). Neben dem technischen Personal wird für die Durchführung eines PET-Scans zudem medizinisches Personal benötigt. Insgesamt werden in der Literatur für Deutschland monatlich Personalkosten in Höhe von circa 14 000 DM für die ¹⁸FDG-Produktion (Klose et al. 1999a) und zwischen 19 500 DM und 20 560 DM für die Bedienung des PET-Scanners (Klose et al. 1999a; Lottes & Schober 1999) veranschlagt. Die Materialkosten für die Radiotracerproduktion setzen sich primär aus den Kosten für Chemikalien und Labormaterialien zusammen. Für die PET-Aufnahmen werden Filme und Einmalartikel wie Spritzen und Kanülen benötigt. In der Literatur werden die Materialkosten für die Produktion des Nuklids im Zyklotron und die anschließende Synthese des Radiodiagnostikums doppelt so hoch

² Die Investitionskosten umfassen die Anschaffungskosten für ein Zyklotron und HPLC-Anlagen. Dabei wurde berücksichtigt, daß diese auch für die Produktion von Radiodiagnostika verwendet werden, die anderweitig, z.B. in der Forschung, eingesetzt werden.

angegeben wie für die Durchführung der PET-Scans (Evens et al. 1983, Klose et al. 1999a), wobei dieses Ergebnis durch den Bedarf an teuren Chemikalien zu erwarten war. Absolut verteilten sich die durchschnittlichen monatlichen Materialkosten für ^{18}F FDG-PET-Untersuchungen in Deutschland gemäß den Berechnungen von Klose et al. (1999a) im Jahr 1999 mit 2 150 DM auf die radiochemische Produktionseinheit und mit 1 050 DM auf den PET-Scan, wobei von einem durchschnittlichen Patientendurchsatz von 27,5 Patienten pro Woche ausgegangen wurde.

Als letzter wichtiger Kostenposten müssen die Gemeinkosten betrachtet werden. Darunter werden Kosten verstanden, die nicht direkt der Bereitstellung und dem Betrieb der PET zugerechnet werden können, wie beispielsweise Ausbildungskosten des technischen und medizinischen Personals, Kosten für Verwaltung und Hausreinigung oder Steuern. Für das Universitätsklinikum Ulm wurde beispielsweise zur Berücksichtigung der Gemeinkosten ein pauschaler Aufschlag von 29% auf die variablen Kosten verwendet, der auf Angaben der Klinikverwaltung beruht (Klose et al. 1999b).

Diese Kostenbetrachtungen zeigen auf, daß der größte prozentuale Anteil an den Gesamtkosten fix ist (Evens et al. 1983, Conti et al. 1994, Lottes et al. 1998, Lottes & Schober 1999, Klose et al. 1999a). Klose et al. (1999a) geben bei PET-Ganzkörperuntersuchungen mit ^{18}F FDG einen Fixkostenanteil von 61% für die radiochemische Produktionseinheit (Zyklotron plus radiochemisches Labor) und von 85% für den PET-Scanner an. Demnach ist ein hoher Auslastungsgrad der technischen Geräte für ein Krankenhaus sehr wichtig, um damit einen betriebswirtschaftlich sinnvollen Einsatz dieser Geräte gewährleisten zu können. Dieser Sachverhalt konnte bereits mit Hilfe von Szenarienanalysen demonstriert werden (Klose et al. 1999a). Höhere Auslastungsgrade der radiochemischen Produktionseinheit und des PET-Scanners führten bei diesen modellhaften Berechnungen zu einem wirtschaftlicheren Einsatz der technischen Geräte. Insbesondere die Auslastung der Produktionseinheit beinhaltet ein enormes Einsparpotential, da nicht nur die Zahl der Patienten je PET-Scanner, sondern auch das Verhältnis von Produktionseinheit und Scanner erhöht werden kann. Conti et al. geben ein optimales Verhältnis von 1:8 an, während Klose et al. eine Erhöhung des momentanen Verhältnisses von 1:2 auf ein von Experten empfohlenes Verhältnis von 1:4 als realistisch abschätzen (Conti et al. 1994; Klose et al. 1999a). Diese theoretischen und modellhaften Resultate werden von den empirischen Daten des Universitätsklinikums Münster untermauert. Dort wurden im Zeitraum November 1995 bis Oktober 1996 und Oktober 1997 bis September 1998 zwei Kostenanalysen³ zum Einsatz der PET durchgeführt (Lottes et al. 1998, Lottes & Schober 1999). Anhand der zweiten Kostenanalyse ließ sich aufzeigen, daß durch gestiegene Patientenzahlen, und damit durch einen höheren Auslastungsgrad der technischen Geräte, die durchschnittlichen Kosten gesenkt werden konnten.

³ Die spätere Kostenanalyse wurde als notwendig erachtet, da die erste Analyse in die Einführungsphase der PET in Münster fiel und aus diesem Grund der Patientendurchsatz und das Untersuchungsspektrum untypisch waren.

Tabelle 3: Kosten pro ^{18}F FDG-PET-Untersuchung am Universitätsklinikum Ulm (PET-Zentrum mit klinikinterner Radiotracerproduktion), Stand 1997

Kostenart	^{18}F FDG-Herstellung ⁺	PET-Untersuchung ⁺
Investitionskosten	188 DM	336 DM (5,93 DM/Min.)
Wartungskosten [*]	117 DM	385 DM (6,79 DM/Min.)
Personalkosten [*]	113 DM	Arzt [#] : 1,82 DM; MTA [#] : 0,98 DM
Sachmittelkosten [*]	78 DM	24 DM (0,42 DM/Min.)

*: Angaben zu Wartungs-, Personal- und Sachmittelkosten enthalten einen Gemeinkostenanteil von 29%.

+ : Kosten sind in DM je Untersuchung angegeben. Bei der PET-Untersuchung sind zusätzlich in Klammern die Kosten je Untersuchungsminute angegeben.

#: Personalkosten für PET-Untersuchungen sind je Arbeitsminute angegeben und umfassen neben der reinen Untersuchungszeit auch den Zeitbedarf für Tätigkeiten wie Patientenvorbereitung, Anamnese sowie zur Begutachtung der PET-Aufnahmen.

Tabelle 3 zeigt die Kostensätze je Untersuchung mit ^{18}F FDG-PET am Ulmer Universitätsklinikum, Stand 1997. Neben den entsprechenden Kosten je ^{18}F FDG-PET Untersuchung werden auch die Kosten jeweils pro Minute angegeben. Die hier berichteten Daten beruhen auf einer Kosten-Effektivitäts-Studie am Ulmer Universitätsklinikum (Klose et al. 1999a). Technologische Veränderungen sowie Änderungen in der Organisation und insbesondere in der Kapazitätsauslastung führten zwischenzeitlich gegenüber dem in der Tabelle 3 angegebenen Stand von 1997 zu Verringerungen der Kostensätze.

Um für die einzelne Klinik den hohen Fixkostenanteil zu reduzieren, werden seit langem verschiedene Ansätze zur Realisierung möglicher Einsparpotentiale diskutiert (Evens et al. 1983, Conti et al. 1994). Einer dieser Ansätze besteht darin, daß die Klinik die Produktion der Radiopharmaka an einen externen Hersteller auslagert. Dieses sogenannte Satellitenkonzept wird vom Universitätsklinikum Münster in die Praxis umgesetzt (Lottes et al. 1998, Lottes & Schober 1999). Das Universitätsklinikum Münster bezieht das von ihr benötigte ^{18}F FDG von der Forschungszentrum Jülich GmbH (Lottes et al. 1998). Der externe Bezug von ^{18}F FDG ist möglich, da ^{18}F FDG eine Halbwertszeit von 110 min hat. Bei Radiodiagnostika mit kürzerer Halbwertszeit, die zum Beispiel mit den Radionukliden ^{11}C , ^{13}N oder ^{15}O markiert sind, ist der Ankauf von externen Produzenten wegen der hohen Zerfallsverluste unwirtschaftlich (Lottes et al. 1998). Eine derartige Einschränkung in der Vielfalt der zur Verfügung stehenden Radiodiagnostika muß vor allem aus medizinischer Sicht als Nachteil des Satellitenkonzeptes gewertet werden.

Auf die Kosten der PET wirkt sich das Satellitenkonzept auf unterschiedliche Weise aus. Sämtliche Kostenposten, die der radiochemischen Produktionseinheit zur Herstellung von Radiopharmaka zugerechnet werden müssen, entfallen. Andererseits ergeben sich neue variable Kosten, welche die Aufwendungen für den Ankauf der Radiodiagnostika subsumieren. Da die externen Produzenten auf die Herstellung der Radiopharmaka spezialisiert sind und eine höhere Auslastung der technischen Ge-

räte aufweisen, kann davon ausgegangen werden, daß die externen Produktionskosten für eine Einheit ^{18}F FDG niedriger sind als bei Eigenproduktion in der Klinik. Jedoch fallen neben den Kosten für die nachgefragte Menge an ^{18}F FDG zusätzliche Kosten für Transport und aufwendige Sicherheitsverpackung an. So entstanden dem Universitätsklinikum Münster im Februar 1998 Materialkosten für das Radiopharmakon ^{18}F FDG in Höhe von 45 781 DM und zusätzlich Verpackungs- und Transportkosten von 11 336 DM. Beim Satellitenkonzept muß außerdem beachtet werden, daß bedingt durch den radioaktiven Zerfall des ^{18}F -Isotops größere ^{18}F FDG-Mengen bestellt werden müssen, als tatsächlich bei der Durchführung der PET-Scans appliziert werden. Als Grund für diese Diskrepanz zwischen georderter und applizierter Menge geben Lottes et al. an, daß durch den Transport 59% und durch Zeitverzögerungen in der Klinik weitere 19% der ^{18}F FDG-Aktivität verlorengehen (Lottes et al. 1998). Hierdurch werden Kostenvorteile, die durch die Einsparung des Zyklotrons erzielt werden, teilweise kompensiert. Eine Möglichkeit, einen Teil dieser Kosten aufzufangen, wäre beispielsweise eine verbesserte Ablauforganisation in der Klinik, um so die Zeitverzögerungen innerhalb der Klinik zu minimieren. Abschließend sollte erwähnt werden, daß die Universitätsklinik Münster durch den Bezug des ^{18}F FDG von einem externen Produzenten den prozentualen Anteil der Fixkosten von 61% auf 46% drücken konnte (Lottes & Schober 1999).

Ob durch das Satellitenkonzept im Vergleich zum vollausgestatteten PET-Zentrum (mit Zyklotron, radiochemischem Labor und PET-Scanner) tatsächlich Kosten eingespart werden können, kann im Moment nicht abschließend abgeschätzt werden. Zu beiden Konzepten liegen zwar Kostenbetrachtungen vor, für das Satellitenkonzept eine Kostenanalyse zum nach Indikationsgebieten differenzierten PET-Einsatz am Universitätsklinikum Münster (Lottes et al. 1998, Lottes & Schober 1999) und für ein PET-Zentrum eine Kosten-Effektivitäts-Analyse des Universitätsklinikums Ulm zum Primärstaging bei Lymphomen (Klose et al. 1999a), die jedoch nicht uneingeschränkt vergleichbar sind. Gründe für die fehlende Vergleichbarkeit sind beispielsweise Unterschiede in der Anzahl durchgeführter Ganzkörper- bzw. Teilkörperscans, eine unterschiedliche Verteilung der durchgeführten Untersuchungen auf die Indikationsgebiete Neuromedizin, Kardiologie und Onkologie, Abweichungen in der absoluten Anzahl behandelter Patienten und die verschiedenartige Ausstattung der Einrichtungen (Klose et al. 1999a).

Eine andere Variante des Satellitenkonzeptes stellt die modellhafte Kooperation zwischen dem Universitätsklinikum Freiburg und einer Gruppe niedergelassener Nuklearmediziner dar, die sich zur Euro-PET GmbH zusammengeschlossen haben (Moser 1997). Das Universitätsklinikum Freiburg hatte bis zum Zeitpunkt dieser Zusammenarbeit ^{18}F FDG über externe Produzenten bezogen. Euro-PET richtete auf dem Gelände des Klinikums ein vollständiges PET-Zentrum ein. Im Gegenzug erklärte sich das Klinikum bereit, sämtliche Radiopharmaka über die ortsansässige Firma zu beziehen. Zudem wurde Euro-PET die Nutzung der klinischen Infrastruktur ermöglicht. Durch diese Kooperation kann das Universitätsklinikum Freiburg einigen Nachteilen des „ursprünglichen“ Satellitenkonzeptes begegnen. Denn aufgrund der räumlichen

Nähe zum Zyklotron können Zerfallsverluste minimiert werden. Zudem ermöglicht die kurze Distanz zwischen Produktions- und Verbrauchsort die Arbeit mit den kurzlebigen Radionukliden ^{11}C , ^{13}N oder ^{15}O . Eine explizite Betrachtung der Kosten dieser speziellen Form des Satellitenkonzeptes liegt zum momentanen Zeitpunkt noch nicht vor.

Die Ausführungen verdeutlichen, daß die monatlichen Kosten der PET stark durch die hohen Anschaffungs- und Wartungskosten des Zyklotrons und des PET-Scanners dominiert werden. Ansätze zur Reduktion des hohen prozentualen Anteils an Fixkosten werden auf unterschiedlichen Wegen verfolgt, wie die Beispiele der Universitätskliniken Münster und Freiburg unterstreichen. PET-Zentren mit hausinterner Radiotracerproduktion, aber auch Kliniken, die nach dem Satellitenkonzept arbeiten, können den wirtschaftlichen Einsatz der technischen Geräte nur bei einem hohen Auslastungsgrad gewährleisten. Aus diesem Grunde ist es gerade bei der PET wichtig, daß eine genaue Bedarfsplanung vor der Einrichtung neuer Standorte durchgeführt wird.

C.5.3 Auswertungsergebnisse mit dem Transparenz- und Qualitätskatalog

Die methodische Qualität der für den vorliegenden HTA-Bericht als relevant eingestuften Studien wurde anhand des vorgegebenen Fragenkataloges zu Transparenz- und Qualitätskriterien überprüft. Bereits in Kapitel C.4.2.1, „Transparenz- und Qualitätskatalog“, wurden beide Kriterienkataloge, die jeweils 56 Fragen umfassen, kurz vorgestellt. Anhand des Transparenzkataloges wird aufgezeigt, ob die einzelnen Studien die essentiellen Kriterien erfüllen, die an eine methodisch gute ökonomische Evaluation gestellt werden. In welcher Qualität diese Bestandteile ausgeführt wurden, wird durch die Fragen des Qualitätskataloges erfaßt.

Zur Beantwortung dieser Fragen standen drei Antwortmöglichkeiten zur Verfügung:

Antwortmöglichkeit	Interpretation
1	Das Kriterium wird von der Studie vollständig erfüllt.
$\frac{1}{2}$	Das Kriterium wird teilweise von der Studie erfüllt.
0	Das Kriterium wird in der Studie nicht beachtet oder wird methodisch fehlerhaft behandelt.
n.r.	Das Kriterium ist für diese Studie nicht relevant.

Das Ergebnis dieser Auswertung ist in den Tabellen 4 bis 7 (Onkologie) und in den Tabellen 8 und 9 (Kardiologie) zusammenfassend dargestellt.

C.5.4 Zusammenfassende Beschreibung und Bewertung der als relevant klassifizierten Studien

Nach Einschätzung der methodischen Qualität der relevanten Studien gemäß dem Transparenz- und Qualitätskatalog (siehe Anhang) werden in diesem Berichtsteil die berücksichtigten Studien beschrieben und wichtige Resultate der Analyse anhand des Kriterienkataloges in einer Studienbewertung zusammengefaßt. Die einzelnen Studienberichte werden im folgenden - soweit möglich - nach Indikationen und nach Aktualität der vorzustellenden Studie (Jahr der Veröffentlichung) geordnet besprochen. Wird in einer Publikation der Einsatz der PET bei mehreren Indikationen untersucht, ist die am besten evaluierte Indikationsstellung ausschlaggebend für die Zuordnung.

C.5.4.1 Onkologie

Diagnose und Staging von nicht-kleinzelligem Lungenkarzinom (NSCLC) und solitären Lungenknoten (SPN)

Gambhir SS, Hoh CK, Phelps ME et al. Decision tree sensitivity analysis for cost-effectiveness of FDG-PET in the staging and management of non-small-cell lung carcinoma. J Nucl Med 1996;37(9):1428-1436

Scott WJ, Shepherd J and Gambhir SS. Cost-effectiveness of FDG-PET for staging non-small cell lung cancer: a decision analysis. Ann Thorac Surg 1998; 66(6):1876-1883

Gambhir SS, Shepherd JE, Shah BD et al. Analytical decision model for the cost-effective management of solitary pulmonary nodules. J Clin Oncol 1998; 16(6):2113-2125

Die Mitte und Ende der 1990er Jahre publizierten ökonomischen Evaluationen der kalifornischen Arbeitsgruppe beschäftigen sich alle mit der Diagnostik von Tumorerkrankungen der Lunge. Die frühere Publikation von Gambhir et al. aus dem Jahr 1996 schlägt zwei Alternativstrategien vor, die zum Staging von Lungentumoren, speziell von nicht-kleinzelligem Bronchialkarzinom (NSCLC), eingesetzt werden könnten. Scott et al. (1998) hat diese Untersuchung in modifizierter Form aufgegriffen und durch zwei weitere Versorgungsstrategien ergänzt. Die Veröffentlichung von Gambhir et al. aus dem Jahr 1998 befaßt sich mit der Auswahl einer kosteneffektiven

Strategie aus vier möglichen Strategieansätzen zur Versorgung von Patienten mit solitären Lungenknoten (SPN).

Da alle drei Publikationen eine ähnliche Vorgehensweise und eine vergleichbare formale Struktur aufweisen, werden die drei Evaluationsstudien in einem gemeinsamen Studienbericht abgehandelt.

Beschreibung der Studien

Bei allen drei Studien handelt es sich um entscheidungstheoretische Modelle, in denen mit Hilfe von Entscheidungsbäumen und Sensitivitätsanalysen kosteneffektive Strategien zur diagnostischen Abklärung von NSCLC (Staging durch Untersuchung auf mediastinale Lymphknotenmetastasen) bzw. SPN (Differenzierung zwischen malignen und benignen SPNs) aus mehreren alternativen Strategien ausgewählt werden. Ziel aller Studien ist die Optimierung der Entscheidungsfindung für oder gegen einen kurativen operativen Eingriff zur Resektion von Tumorerkrankungen: so sollen bei der Indikationsstellung NSCLC möglichst viele inoperable Patienten mit mediastinalen Lymphknotenmetastasen und bei der Indikationsstellung SPN Patienten mit gutartigen solitären Lungenknoten mit Hilfe von schonenden bildgebenden Verfahren (CT bzw. PET) bereits im Vorfeld des chirurgischen Eingriffs erkannt und von einer kontraindizierten, operativen Therapie ausgeschlossen werden. In den drei entscheidungstheoretischen Analysen wird geprüft, ob PET, bedingt durch seine bessere diagnostische Treffsicherheit, diesbezüglich eine kosteneffektive Ergänzung zur konventionellen CT (PET + CT-Kombinationsstrategie) darstellt. Dazu wird nicht allein der monetäre Aufwand für die diagnostischen Untersuchungen in die Betrachtung einbezogen, sondern auch deren „downstream“-Wirkung auf die Kosten der nachgeschalteten operativen Versorgung (therapeutic efficacy) und auf die Lebenserwartung der Patienten (patient outcome efficacy, siehe C.4.2.2) betrachtet. In den neueren Veröffentlichungen aus dem Jahr 1998 werden zudem durch die Berechnung von inkrementellen Kosten-Effektivitäts-Quotienten die direkten Kosten medizinischer Interventionen mit deren Auswirkungen auf den patientenbezogenen Outcome unmittelbar in Relation gesetzt.

Wegen der Komplexität der vorzustellenden entscheidungsanalytischen Modelle können die jeweils verglichenen Alternativstrategien hier nur grob skizziert werden.

Indikationsstellung: NSCLC

In der älteren Studie von Gambhir et al. (1996) werden zwei im Aufbau leicht differierende Entscheidungsbäume zum Staging von NSCLC-Patienten vorgestellt. In beiden zweiästigen Entscheidungsbäumen wird eine CT-basierte Strategie, mit einem CT-Scan als einzigem nicht-invasiven Test vor einer Biopsie (nur bei positivem CT-Befund), direkt mit zwei PET/CT-Kombinationsstrategien verglichen, in welchen die

PET-Untersuchung (nur Teilkörperaufnahmen im Brustbereich) als Vortest zur CT vorgegeben ist. Im konservativen PET/CT-Modell werden alle PET-positiven Patienten unabhängig vom Befund der CT-Nachuntersuchung obligatorisch einer Biopsie der metastaseverdächtigen mediastinalen Lymphknoten (u.a. Mediastinoskopie oder Mediastinotomie) unterzogen, während im weniger konservativen PET/CT-Modell nur diejenigen Patienten zur Biopsie überstellt werden, deren PET- bzw. CT-Befunde nicht übereinstimmen.

Aufbauend auf diese entscheidungsanalytischen Modelle von Gambhir et al. (1996) entwerfen Scott et al. in ihrer Veröffentlichung aus dem Jahr 1998 folgende fünf Strategien: eine CT-basierte Baseline-Strategie (identisch mit der CT-basierten Strategie bei Gambhir et al. 1996) und vier Strategien, die sich vor allem in der Kombination von PET mit CT (und der sich eventuell anschließenden Biopsie) unterscheiden. Dabei gibt das Modell die Durchführung eines PET-Scans im Anschluß an die CT-Eingangsuntersuchung (geänderte Reihenfolge der nicht-invasiven Tests gegenüber Gambhir et al. 1996) entweder nur bei einem negativen CT-Eingangsbefund (Modell B) oder obligatorisch unabhängig vom jeweiligen CT-Ergebnis vor (Modell C und Modell D). Die Modelle C und D unterscheiden sich im Einsatz einer mediastinalen Lymphknotenbiopsie zur Verifikation der Befunde aus den vorgeschalteten bildgebenden Verfahren: so erfolgt im Modell C eine Biopsie der CT-positiven Patienten nur bei einem PET-positiven Befund, im Modell D jedoch unabhängig vom Ausgang der PET-Befundung. Im Modell E werden Patienten, die sowohl bei der CT- als auch bei der anschließenden PET-Untersuchung positiv waren, von einer operativen Therapie mittels Thorakotomie ausgeschlossen.

Tabelle 10: Baseline-Annahmen für NSCLC-Studien nach Angaben der Originalstudien

Literaturstelle	Gambhir et al. 1996	Scott et al. 1998
Prävalenz der Metastasenbildung:	31%	31%
Sensitivität:	CT 67% PET 90%	CT 67% PET CT + 89%, PET CT - 76%*
Spezifität:	CT 73% PET 91%	CT 73% PET CT + 81%, PET CT - 97%
Lebenserwartung:	bei erfolgreicher Operation: 7 Jahre bei nicht resektablem Tumor: 1 Jahr	bei erfolgreicher Operation: 7 Jahre bei nicht resektablem Tumor: 1 Jahr
Kostenangaben:	CT: US\$ 700 PET: US\$ 1 200 Biopsie: US\$ 3 000 Chirurgie: US\$ 30 000	Thorax CT: US\$ 378 Thorax PET: US\$ 2 000 Biopsie: US\$ 4 360 Chirurgie: US\$ 18 500

* Die Sensitivität/Spezifität für PET beträgt ...%, unter der Bedingung, daß die vorausgegangene CT-Untersuchung positiv/negativ war.

Die frühere NSCLC-Studie von Gambhir et al. (1996) zeigt, daß in Tabelle 10 aufgelisteten Baseline-Annahmen für die kombinierte PET/CT-Strategie im konservativen

Modell ein durchschnittliches Einsparpotential von US\$ 1 154 pro Patient ohne Verlust an Lebenserwartung (Zunahme von 2,96 Tagen) besteht.

Dabei betragen die durchschnittlichen Kosten pro Untersuchung mit der PET/CT-Kombinationsstrategie US\$ 24 480, die entsprechenden Kosten für die CT-basierte Strategie insgesamt US\$ 25 634. Bei - laut Angaben der nationalen Krebsstatistik der USA - rund 85 000 untersuchten NSCLC-Fällen pro Jahr hätte sich somit auf nationalem Niveau zum Zeitpunkt der Studie ein jährliches Einsparvolumen von rund US\$ 98 Millionen ergeben. Die - theoretische - Kostenminderung pro behandeltem Patient (US\$ 1 154) und der, wenn auch nur marginale, Gewinn an Lebenserwartung (etwa 3 Tage) sind das Resultat eines exakteren präoperativen Staging der NSCLC-Patienten durch den zusätzlichen Einsatz der PET und einer daraus resultierenden Verbesserung des Patientenmanagements, wobei weniger inoperable Patienten einer wirkungslosen operativen Therapie zugeführt werden. Im weniger konservativen Modell (Biopsie erfolgt nur bei nicht übereinstimmendem PET- und CT-Befund) lassen sich die durchschnittlichen Kosten pro Patient um US\$ 2 267 reduzieren, jedoch um den Preis, daß 1,7% der kurablen Patienten von der kombinierten PET/CT-Diagnosestrategie übersehen werden. Aus den Ergebnissen der Sensitivitätsanalysen läßt sich ableiten, daß das beim kombinierten Einsatz von PET/CT resultierende Einsparpotential unter anderem von den Operationskosten und der Prävalenz der Metastasenbildung durch den NSCLC-Primärtumor (im kontralateralen Lungenflügel bzw. in den regionären Lymphknoten) abhängt. Sinken im konservativen Modell die Kosten für die Resektion (Baseline-Annahme: US\$ 30 000) bzw. die Prävalenz (Baseline-Annahme: 31%) unter einen Schwellenwert von US\$ 17 485 bzw. 16,9%, rechnet sich der kombinierte PET/CT-Einsatz im direkten Vergleich zum alleinigen Einsatz von CT nicht mehr. Im Rahmen einer Threshold-Analyse, bei der die Effektgröße Lebenserwartung - bei konstanten Versorgungskosten - maximiert wird, verbessert die PET/CT-Kombination gegenüber der CT-Strategie die Lebenserwartung bei Prävalenzwerten $\geq 5,6\%$.

In der Nachfolgestudie zum Staging von NSCLC-Patienten von Scott et al. (1998) erweist sich die Strategie B (PET nur nach CT-negativem Eingangsbefund) als die Methode der Wahl. So ergibt sich unter den Baseline-Annahmen dieses Modells (siehe Tabelle 10) für eine kombinierte CT/PET-Strategievariante (Modell B) eine inkrementelle Kosten-Effektivitäts-Ratio (iKER) im Vergleich zu Modell A (nur CT-Untersuchung + Biopsie bei CT-positivem Befund) von US\$ 25 286 pro zusätzlich gewonnenem Lebensjahr. Unter der Voraussetzung, daß - laut Aussage der Autoren - in den USA eine iKER von US\$ 50 000 pro gewonnenem Lebensjahr gesellschaftlich akzeptiert wird, wäre somit die Versorgungsstrategie B als kosteneffektiv im Vergleich zum alleinigen Einsatz der CT als Eingangsuntersuchung (Strategie A) zu bewerten. Unter dem Strategie B-Regime würden nach Schätzungen der Autoren jährlich 10% oder circa 10 000 Operationen weniger veranlaßt als durch Strategie A.

Die kritischste Variable für die Stabilität des Scott-Modells ist das Verhältnis der Kosten der Thorakotomie zu den Kosten für die PET-Aufnahme: sinken die PET-Kosten unter US\$ 1 600 oder steigen die Kosten für die Thorakotomie über US\$ 22 000, so wird Strategie B (PET nur nach CT-negativem Eingangsbefund) dominant über Strategie A (nur CT + Biopsie bei CT-positivem Befund); gleichzeitig erreicht auch eine weitere CT/PET-Kombination (Strategie C: PET-Einsatz unabhängig vom CT-Befund) den Bereich der Kosteneffektivität (\leq US\$ 50 000). Verschlechtern sich die diagnostischen Kenndaten für PET (Sensitivität und Spezifität) um 10% und verteuern sich gleichzeitig die PET-Kosten um denselben Betrag, so ist der alleinige Einsatz von CT (Strategie A) aus ökonomischer Sicht die günstigste Vorgehensweise, da bei allen übrigen Strategien die iKER unter diesen Konditionen weit über US\$ 50 000 liegt.

Wie bereits in der Vorstudie von Gambhir et al (1996) variiert auch hier die Kosteneffektivität mit der Prävalenz für N2- oder N3-Metastasenbildung: so ist Strategie B in einem engen Prävalenzbereich von 0,28 bis (ausschließlich) 0,35 kosteneffektiv im Vergleich zu Strategie A und wird für Prävalenzwerte \geq 0,35% gegenüber Strategie A dominant. Die Strategien C und D (PET-Einsatz unabhängig vom CT-Vorbefund, differenzierter Biopsie-Einsatz) werden kosteneffektiv gegenüber Strategie A bei Prävalenzwerten $> 0,40$ bzw. $> 0,43$.

Indikationsstellung: SPN

Zur Bewertung der Kosteneffektivität von unterschiedlichen Strategien zur Diagnose von solitären Lungenknoten und zur Optimierung des Patientenmanagements bei Verdacht auf maligne SPN konstruieren Gambhir et al. (1998) ein entscheidungstheoretisches Modell mit vier alternativen Versorgungsstrategien (Diagnose plus Therapie mittels chirurgischer Resektion):

1. Eine „wait and watch“-Baseline-Strategie, bei der alle Patienten - vor einer Entscheidung für oder gegen eine operative Entfernung von verdächtigen Lungenknoten - über einen zweijährigen Beobachtungszeitraum mit seriellen Röntgen- oder CT-Scans auf Knoten mit maligner Wachstumsgeschwindigkeit untersucht werden,
2. eine Operationsstrategie (konservativstes Strategiemodell), bei der alle Patienten ohne weitere Voruntersuchung mit nicht-invasiven Methoden direkt einer Thorakotomie unterzogen werden, um resektable Lungenknoten chirurgisch zu entfernen,
3. eine CT-basierte Strategie (klinische Routinemethode), bei der die Patienten obligatorisch mit einem hochauflösenden CT-Scanner untersucht werden; die eindeutig CT-negativen Patienten werden mit der „wait and watch“-Strategie weiter beobachtet, während die übrigen Patienten entweder zur Biopsie (+ eventuell Resektion) oder direkt zur Resektion überstellt werden (gemischter Pfad Biopsie/Chirurgie),

4. eine kombinierte CT/PET-Strategie, bei der die Patienten nur bei positivem oder indifferentem CT- Befund einem zusätzlichen PET-Scan unterzogen werden. Eindeutig CT-negative Patienten werden wie in 3. mit der „wait and watch“-Strategie weiterbeobachtet. Bei CT-positiven/PET-positiven Patienten wird wie unter 3. entweder direkt eine Operation veranlaßt oder vorab die Notwendigkeit einer Resektion mittels Biopsie überprüft. CT-positive/PET-negative Patienten werden entweder mit der „wait and watch“-Strategie weiterbeobachtet oder einer klärenden Biopsie überantwortet.

Um die Wirkung der jeweils alternativen Strategien auf die Kosten und den medizinischen Outcome erfassen zu können, werden in allen Studien für die einzelnen Versorgungsstrategien die Interventionskosten (Gesamtkosten einer Strategie für Diagnose + Operation) und die Lebenserwartung der Patienten berechnet (siehe Studienbewertung).

In der SPN-Studie kommen die Autoren Gambhir et al. (1998) zu tendenziell ähnlichen Resultaten wie in den vorab erläuterten NSCLC-Studien. Anhand der Ergebnisse läßt sich eine Abhängigkeit der vier zur Wahl stehenden Strategien von der Vortestwahrscheinlichkeit, daß bei einem Patienten ein maligner SPN vorliegt (Prävalenz der Malignität je nach Alter und Rauchgewohnheiten des Patienten und Größe des verdächtigen Lungenknotens), ableiten:

1. Entsprechend der inkrementellen Kosten-Effektivitäts-Relation (Referenzstrategie: „wait and watch“) wäre für niedrige Vortestwahrscheinlichkeiten bis 0,12 die „wait and watch“-Strategie gegenüber den anderen Strategien zu bevorzugen,
2. während im breiten mittleren Bereich 0,12 - 0,69 die kombinierte CT/PET-Strategie zu favorisieren wäre;
3. im oberen Bereich von 0,69 - 0,90 dagegen wäre die CT-basierte Strategie (ohne additive PET-Untersuchung) vorzuziehen;
4. ab einer Vortestwahrscheinlichkeit von 0,90 dagegen weist die direkte Operation ohne vorherige nicht-invasive Diagnostik die höchste Kosteneffektivität auf.

Dabei zeigt die CT/PET-Kombination im direkten Vergleich zur Referenzstrategie „wait and watch“ im unteren bis mittleren Bereich der Vortestwahrscheinlichkeit für maligne SPNs ($pSPN_{MAL}$), zwischen 0,02 und 0,44, den höchsten Zuwachs an mittlerer Lebenserwartung (inkrementelle Lebenserwartung) und verursacht bis $pSPN_{MAL}$ 0,72 die geringsten zusätzlichen Kosten (inkrementelle Kosten) unter den analysierten Strategien „CT/PET“, „nur CT“ und „direkte Operation“ relativ zur Referenzstrategie „wait and watch“. Im $pSPN_{MAL}$ -Intervall 0,72 - 0,90 ist die CT-Strategie die kostengünstigste Vorgehensweise in Relation zu „wait and watch“. Die Autoren schätzen unter der Annahme, daß die Hälfte der untersuchten SPN-Patienten eine Vortestwahrscheinlichkeit $pSPN_{MAL}$ von 0,1 bis 0,7 haben und sich gleichmäßig über dieses Intervall verteilen, daß in den USA landesweit - bei gleichbleibenden Konditionen - jährlich 15% oder 7 875 Operationen durch den Einbezug der PET in das konventionelle CT-Diagnoseregime (additiver Einsatz der PET) vermeidbar wären.

Bei einem durchschnittlichen Einsparpotential von US\$ 1194 pro Patient (mit einer Variationsspanne von US\$ 91 - 2 200 pro Patient innerhalb des pSPN_{MAL} Intervalls 0,1 - 0,7) ergäbe sich somit auf nationaler Ebene ein Einsparpotential von jährlich US\$ 62,7 Millionen allein an medizinischen Kosten für die Versorgung von SPN-Patienten. Anhand von umfangreichen uni- und multivariaten Sensitivitätsanalysen zeigen die Autoren auf, daß sich - trotz Verschiebung der Schwellenwerte - die Rangfolge der Kosteneffektivität in Abhängigkeit von der Prävalenz nicht ändert: die Kombination CT/PET ist für die Bevölkerungsgruppe mit niedriger und mittlerer Prävalenz und die CT-Strategie für die Bevölkerungsgruppe mit hoher Prävalenz die Methode der Wahl. Im obersten Prävalenzbereich (Übergangsbereich je nach Sensitivitätsanalyse 0,8 - 0,9) ist nach wie vor die invasive Operationsmethode den bildgebenden Verfahren vorzuziehen.

Schlußfolgerungen aus den NSCLC- bzw. SPN-Studien

Die Autoren kommen in ihren Studien zu dem Schluß, daß der kombinierte Einsatz der PET plus CT im Vergleich zum alleinigen Einsatz von CT - zumindest bei einem Teil der Patienten (je nach Vortestwahrscheinlichkeit für die Erkrankung, siehe Studienbeschreibungen) - als kosteneffektiv zu werten ist (\leq US\$ 50 000). Eine breitere Anwendung von PET bei der Diagnostik von NSCLC (Staging durch Untersuchung des Lymphknotenstatus) bzw. SPN (Differenzierung zwischen malignen und benignen SPNs) sei deshalb zu unterstützen.

Gemeinsame Bewertung der Studien (NSCLC und SPN)

Wie aus der tabellarischen Übersicht zur Bewertung der Transparenz- und Qualitätskriterien zu ersehen ist (siehe die Tabellen 4 und 5), handelt es sich bei allen hier untersuchten ökonomischen Studien der US-Arbeitsgruppe am Institute for Biological Imaging, Los Angeles, sowohl aus inhaltlicher als auch aus methodischer Sicht um qualitativ hochwertige entscheidungstheoretische Modelle, welche die wichtigsten im Kriterienkatalog (im Anhang des HTA-Berichtes) gestellten Anforderungen an die methodische Qualität einer ökonomischen Studie erfüllen.

Die Zielsetzungen der Studien werden in allen drei Artikeln deutlich. Zwar sind aus den Fragestellungen selbst die jeweilig zu vergleichenden Alternativtechnologien zur PET nicht direkt zu ersehen, in den Einführungen zur Thematik werden jedoch die heute in der Lungenkrebsdiagnostik üblichen apparativen Verfahren und Vorgehensweisen angesprochen und in den Artikeln von Gambhir et al. (1996 und 1998) zudem ausführlich erläutert.

Die Modellstruktur der Entscheidungsbäume wird anhand von komplexen, aber dennoch verständlichen Baumdiagrammen vorgestellt und die jeweiligen Studienäste, das heißt, die einzelnen Strategien zur diagnostischen und operativen Versorgung

der Patienten, zusätzlich im Text ausführlich kommentiert. Die Auswahl der Einzeltechnologien, Bausteine der alternativen Versorgungsstrategien, wird nicht explizit begründet. Zumindest in den Artikeln von Gambhir et al. (1996 und 1998) ergibt sich jedoch anhand der in der thematischen Einführung dargestellten diagnostischen Vor- und Nachteile der verfügbaren Diagnosetechniken eine schlüssige Begründung für die letztendliche Technologieauswahl (PET, CT, Biopsiemethoden). Eine nähere Spezifikation der technischen Modalitäten der einzelnen diagnostischen Verfahren hätte die Beurteilung, ob die Ergebnisse der Studien untereinander und zu anderen (künftigen) Studien vergleichbar sind, erleichtert.

Der Fachliteratur entnommen wurden die biometrischen Kenndaten, als Maß für die diagnostische Leistungsfähigkeit der alternativen bildgebenden Verfahren PET bzw. CT (und der Biopsietechniken: u.a. transthorakale Nadelaspirationsbiopsie; Gambhir et al. 1998), sowie die Angaben zu folgenden Parametern:

1. Zum Morbiditäts- und Mortalitätsrisiko der diagnostischen und operativen Interventionen,
2. epidemiologische Daten und Daten zur Lebenserwartung der Patienten je nach Resektierbarkeit der malignen Läsionen, und
3. zur monetären Bewertung der diagnostischen und operativen Einzelleistungen (Preisgerüste).

Die epidemiologischen Daten zur Prävalenz (NSCLC-Studien: für N2- oder N3-Metastasen, SPN: Malignität der Tumorknoten), Morbidität, Mortalität und Lebenserwartung beziehen sich auf einen 64jährigen weißen Mann ohne anderweitige Erkrankung. Die bei den Literaturrecherchen angewandten Suchstrategien werden nicht angegeben; die relevanten Ergebnisse der als Datenquellen ausgewählten Publikationen werden jedoch berichtet. Nach Durchsicht der verfügbaren Fachliteratur ermitteln die Autoren Basiswerte (Mittelwerte) und zumeist auch die Varianzspannen der zur Konstruktion der Modelle notwendigen Parameter. Diese sind in allen drei Artikeln akribisch in einer übersichtlichen Tabelle aufgelistet. Die als Datenquellen genutzten Publikationen werden explizit genannt und im Literaturverzeichnis dokumentiert.

Auf der Effektseite werden mit den diagnostischen Kenndaten der Diagnosetechniken (PET, CT, Biopsiemethoden) intermediäre Größen erfaßt, die jedoch nur als Parameter in die Modellberechnungen eingehen. Als Maß für den Einfluß der diagnostischen Treffsicherheit der zur Wahl stehenden Diagnosestrategien - CT versus jeweilige CT + PET-Kombination - auf die Therapie, und damit auf den gesundheitlichen Outcome der behandelten Patienten, wird die aus gesundheitsökonomischer Sicht als aussagekräftig und somit als adäquat einzustufende finale Effektgröße (siehe C.4.2.2) der Lebenserwartung errechnet. Die Extraktion und Synthese dieses Effektparameters (Basislebenserwartung und Streubreite), differenziert nach der Resektabilität des Tumors, aus der Literatur wird, wie bereits dargestellt, ausführlich und nachvollziehbar beschrieben. Die im Modell veranschlagte Lebenserwartung wird mit

Hilfe der DEALE-Methode abgeschätzt (DEALE: declining exponential approximation of life expectancy). In keiner der Studien werden die Gesundheitszustände mit Präferenzen bewertet. Eine qualitative Bewertung der verbliebenen Lebenszeit ist bei einer Erkrankung, welche die Lebensqualität der betroffenen Patienten stark beeinträchtigt, jedoch grundsätzlich wünschenswert.

In den drei Studien ergeben sich das jeweilige Mengengerüst und damit auch die Gesamtkosten jeder Strategie zur diagnostischen und operativen Versorgung von SPN- bzw. NSCLC-Patienten aus den Modellberechnungen und sind von den jeweils eingesetzten Parametern abhängig (Werte der Basisannahmen und der Sensitivitätsanalysen). Das den einzelnen Kostenberechnungen zugrunde liegende Preisgerüst wird hinreichend gründlich beschrieben (u.a. Kosten für technisches und medizinisches Personal). Die Kosten werden ebenfalls in Form eines Basiswertes (Mittelwert) und einer Varianzspanne angegeben. Die Methoden und Quellen zur Ermittlung des Preisgerüsts werden eindeutig genannt, sind in den drei Studien jedoch nicht einheitlich. So werden in der ersten Studie aus dem Jahr 1996 (NSCLC-Studie, Gambhir et al. 1996) die Mittelwerte für die jeweiligen Einzelleistungen auf der Basis von Rechnungsbeträgen („billed costs“), die hausintern an der klinischen Einrichtung der Autoren erfaßt wurden, abgeschätzt. Die Streubreiten basieren auf den Kostangaben von unterschiedlichen - leider nicht näher definierten - medizinischen Einrichtungen in den USA. In den nachfolgenden Publikationen werden die institutionspezifischen Einzelleistungsvergütungen (Rechnungsbeträge) so weit möglich durch nationale Durchschnittswerte der vom staatlichen Vergütungssystem Medicare erstatteten Kosten für die jeweiligen Einzelleistungen ersetzt. Da nunmehr für den operativen Eingriff der Thorakotomie eine administrative monetäre Bewertung anhand der DRG⁴-Vergütung veranschlagt wird, erklärt sich die große Kostenreduktion für die kurative Thorakotomie von ursprünglich US\$ 30 000 (Gambhir et al. 1996) auf nur noch US\$ 18 500 (Scott et al. 1998). Im Jahr 1998 standen Scott et al. (NSCLC-Studie, 1998) zur monetären Bewertung von PET-Scans für die Diagnostik von NSCLC bereits nationale Medicare-Vergütungen zur Verfügung, während Gambhir et al. (SPN-Studie) weiterhin auf aktuelle institutsspezifische PET-Einzelleistungskosten zurückgreifen mußten (83% der klinikspezifischen „billed costs“), da PET-Aufnahmen zur Abklärung von SPNs nach wie vor nicht durch Medicare-Vergütungen abgedeckt wurden. Solange keine verbrauchsgemäßen Kostenzurechnungen vorliegen, ist die Verwendung der von den Autoren gewählten Ersatzgrößen angemessen, obwohl nicht alle untersuchten Patienten zur Medicare-Klientel gehören dürften.

Die Kostenberechnung bezieht sich ausschließlich auf die im Rahmen der akuten Versorgung anfallenden direkten medizinischen Kosten. Sonstige Kostenposten, beispielsweise für Komplikationen oder nachfolgende nicht-operative Therapien wie Chemo- oder Strahlentherapie oder indirekte Kosten durch Produktionsausfall, bleiben unberücksichtigt. Deren Aufnahme hätte zwar die Aussagefähigkeit der Studien

⁴ Diagnosis related groups, US-amerikanische Patientenklassifikation, die u.a. als Grundlage für die Vergütung nach Fallpauschalen im stationären Bereich verwendet wird.

hinsichtlich der Kostenseite noch verbessert, aber die komplexe Modellstruktur zusätzlich kompliziert. Grundsätzlich sind unterschiedliche Zeitfenster auf der Kostenseite (nur Akutkosten) und auf der Effektseite (langfristige Effektgröße der Restlebenszeit) methodisch nicht korrekt. Weitere formale Mängel der Studien sind, daß das Referenzjahr der Kosten nicht angegeben und keine Inflations-/Deflationsbereinigung vorgenommen wird. Der Verzicht auf eine Diskontierung wird nicht begründet. Letztere Unterlassung dürfte jedoch in der Beschränkung der Kosten ausschließlich auf die akuten diagnostischen und operativen Interventionskosten zu suchen sein (enges Zeitfenster für die Kostenberechnung). Da jedoch mit der restlichen Lebenserwartung auf der Effektseite eine langfristige Größe vorliegt, wäre eine Diskontierung wünschenswert gewesen.

Die Kostenseite und die Effektseite werden in den späteren Studien, SPN-Studie bzw. NSCLC-Studie aus dem Jahr 1998, zu einer inkrementellen Kosten-Effektivitäts-Ratio (Kosten pro gewonnenem Lebensjahr) zusammengeführt, wobei in der NSCLC-Studie die alleinige CT-Strategie und in der SPN-Studie die „wait and watch“-Strategie als Bezugsgröße dienen. Mit der Auswahl einer Referenzstrategie, zu der alle übrigen Alternativstrategien in Bezug gesetzt werden, beachten die Autoren nicht die unter Kapitel C.4.2.3 erläuterten Regeln für die korrekte Berechnung von inkrementellen Kosten-Effektivitäts-Ratios. In der früheren Studie von Gambhir et al. 1996 dagegen wird keine Kosteneffektivitätsrelation berechnet, sondern es werden nur die Kosten und die Effekte einander gegenübergestellt (Minimierung der Kosten bei gleichen Effekten bzw. Maximierung der Effekte bei gleichen Kosten anhand einer Threshold-Analyse).

Die Ergebnisse unter den Basiskonditionen werden in allen drei Artikeln anschaulich in Tabellen und Graphiken in disaggregierter und aggregierter Form (das heißt, in Abhängigkeit von der Prävalenz) präsentiert und ausgiebig diskutiert.

Bei der Kosten-Effektivitäts-Analyse von Scott et al. (1998) ist zwischen den einzelnen Strategien praktisch kein Unterschied in der Lebenserwartung festzustellen. Somit hängt die Kosteneffektivität faktisch nur vom Zähler ab, also von der Differenz der Gesamtkosten für die jeweilig zu vergleichenden Strategien. Rechnerisch dagegen wirken sich jedoch gerade kleinste Änderungen im Nenner, in dem die Effektgröße steht, extrem auf die inkrementelle Kosten-Effektivitäts-Ratio (iKER) aus (Problematik des „kleinen“ Nenners). Aus diesem Grund muß bei der Interpretation der Ergebnisse beachtet werden, daß minimale Unterschiede in der Lebenserwartung zwischen den Alternativstrategien (im Nenner des Kosten-Effektivitäts-Quotienten) signifikante Auswirkungen auf die iKER nach sich ziehen.

Der Einfluß unsicherer oder verzerrter Parameterschätzungen auf die Kostenseite (Gesamtkosten pro Versorgungsstrategie), die Effektseite (Lebenserwartung nach Operation) und die iKER (NSCLC-Studie von Scott et al. 1998, SPN-Studie von Gambhir et al. 1998) und damit auf die Stabilität der entscheidungstheoretischen Modelle wird anhand von uni- und multivariaten Sensitivitätsanalysen überprüft.

Variiert werden die für den Entscheidungsausgang relevanten Parameter, wie beispielsweise die Kosten (für PET, CT und Resektion), die diagnostischen Kenndaten von PET und CT, das Mortalitäts- und Morbiditätsrisiko der Interventionen, die Vor-testwahrscheinlichkeit der Erkrankung und die Größe der SPN-Knoten (Gambhir et al. 1998). Die neuen Werte bzw. Wertebereiche der veränderten Variablen werden zusammen mit den Ergebnissen der Sensitivitätsanalysen in Tabellen aufgelistet. Die Resultate werden ausführlich erläutert und die Richtung und Größe des Einflusses unsicherer Parameter auf den Entscheidungsausgang der theoretischen Analysen konsistent diskutiert.

Einschränkungen der vorliegenden Studien werden von den Autoren dargestellt und mögliche Verbesserungen vorgeschlagen. Verteilungsfragen oder ethische Fragen, beispielsweise zur Problematik einer gesellschaftlich noch akzeptierten Obergrenze der inkrementellen Kosten-Effektivitäts-Ratio (z.B. Mehrkosten pro gewonnenem Lebensjahr), werden nicht angesprochen. Die Schlußfolgerungen der Autoren basieren auf den vorgestellten Studienergebnissen und sind in sich schlüssig dargelegt.

Abschließende Bewertung

Die hier besprochenen entscheidungsanalytischen Modelle zur Bewertung des ökonomischen Stellenwertes eines additiven Einsatzes der PET ergänzend zur konventionellen CT (kombinierte CT/PET- bzw. PET/CT-Diagnosestrategien) in den Indikationsgebieten Staging von „nicht-kleinzelligem Lungenkarzinom“ (NSCLC) bzw. Diagnose von „solitären Lungenknoten“(SPN) zeichnen sich durch eine hohe methodische Qualität aus. Diese begründet sich auf der Erfüllung der nachfolgend aufgelisteten Kriterien:

1. Der detaillierten Dokumentation der Modellparameter,
2. der Wahl einer aussagekräftigen Effektgröße (Lebenserwartung),
3. der Berechnung einer inkrementellen Kosten-Effektivitäts-Ratio (Gambhir et al. 1998, Scott et al. 1998, nicht jedoch bei Gambhir et al. 1996),
4. der anschaulichen Darstellung und Diskussion der Ergebnisse und
5. der Durchführung von umfangreichen Sensitivitätsanalysen.

Die Ergebnisse dieser Studien sind somit aussagekräftig im Hinblick auf eine Bewertung der Kosteneffektivität des klinischen Einsatzes der PET in diesen Indikationsstellungen. Vorsicht bei der Interpretation der Ergebnisse ist jedoch bei der Studie von Scott et al. (1998) wegen der Problematik des „kleinen Nenners“ (siehe Studienbewertung) geboten.

ECRI. Positron emission tomography (PET) for the diagnosis and staging of non-small-cell lung cancer, part 1: results of ECRI's meta-analyses. Health Technology Assessment Information Service 1998;72

ECRI. Positron emission tomography (PET) for the diagnosis and staging of non-small-cell lung cancer, part 2: cost-effectiveness analysis. Health Technology Assessment Information Service 1998;73

Studienbeschreibung

Das Emergency Care Research Institute (ECRI, HTA-Einrichtung der USA) legte 1998 eine eigene ökonomische Analyse zur Kosteneffektivität des Einsatzes der PET als Diagnose- und Staginginstrument bei nicht-kleinzelligem Lungenkarzinom (NSCLC) vor. Diese ökonomische Evaluation baut auf einer Meta-Analyse des ECRI auf (1. Teil des ECRI-Berichtes 1998), die vergleichend die diagnostische Effektivität von PET, CT, (verbunden mit Biopsien: transthorakale Nadelbiopsie, Mediastinoskopie oder Mediastinotomie) untersucht. Im 2. Teil des ECRI-Berichtes (1998) werden die durch adäquate diagnostische und therapeutische Versorgung von NSCLC-Patienten erzielbare Verlängerung der Lebenserwartung und die damit verbundenen Kosten (Gesamtversorgungskosten) ermittelt und gegeneinander abgewogen. Aufgrund der Informationsfülle des ECRI-Berichtes kann die zusammenfassende Beschreibung nur auf die zentralen Punkte und Ergebnisse eingehen. An Detailinformationen interessierte Leser werden auf den Originalbericht verwiesen.

Die Autoren entwickeln nach Befragung klinischer Experten und unter Verwendung publizierter Daten ein Entscheidungsbaum-Modell mit neun Alternativstrategien, die in der tabellarischen Auflistung auf der nächsten Seite skizziert werden. Die mehrstufigen Strategieäste setzen sich aus den diagnostischen Bausteinen CT und/oder ¹⁸FDG-PET, plus Biopsiemethoden zur Verifikation des Befundes der nicht-invasiven Untersuchungsmethoden zusammen. Diese Bausteine werden bei den einzelnen Ästen in unterschiedlicher Zusammensetzung und zeitlicher Abfolge kombiniert (siehe dazu Tabelle 11): Strategie A besteht ausschließlich aus einer nicht-invasiven CT-Untersuchung, während die Strategien B, C und D zusätzlich Biopsiemethoden zur Bestätigung des CT-Befundes beinhalten. Dabei entspricht Strategie D dem derzeit gängigen klinischen Routinealgorithmus zur Diagnose und Stadieneinstufung von NSCLC. In den Strategien E - I wird die PET entweder zusätzlich (Strategien E - H) oder substitutiv (Strategie I) zur CT folgendermaßen eingesetzt:

1. Als Eingangsuntersuchung zur Stadieneinstufung der NSCLC-Erkrankung anhand des Lymphknotenstatus (Strategien E, H und I) im Stagingschritt,

Tabelle 11: Beschreibung der alternativen Strategien zur Diagnose und zum Staging von nicht-kleinzelligem Lungenkarzinom (NSCLC) im ECRI-Bericht 1998, Teil 1

Strategiealternativen	Diagnoseschritt zur Differenzierung zwischen malignen und benignen SPNs			Stagingschritt zur Prüfung auf mediastinalen Lymphknotenbefall bei gesichertem NSCLC		
	Eingangstest (ET)	Bestätigungstest, (BT) wenn 1.Test +	Bestätigungstest, (BT) wenn 1.Test –	Ergebnis aus Vortest (ET oder BT) genützt	Bestätigungstest, wenn Vortest +	Bestätigungstest, wenn Vortest –
Strategie A	CT	—	—	CT (ET)	—————	—
Strategie B	CT	—	—	CT (ET)	Mediastinale Biopsie ^c	—
Strategie C	CT	TNB ^b	—	CT (ET)	—————	—
Strategie D^a	CT	TNB	—	CT (ET)	Mediastinale Biopsie	—
Strategie E	CT	PET	—	PET (BT)	—————	—
Strategie F	CT	TNB	—	CT (ET)	Mediastinale Biopsie	PET
Strategie G	CT	TNB	PET	CT (ET)	Mediastinale Biopsie	PET
Strategie H	CT	PET	—	PET (BT)	Mediastinale Biopsie	—
Strategie I	PET	TNB	—	PET (ET)	Mediastinale Biopsie	—

a Strategie D entspricht der üblichen klinischen Praxis

b TNB: transthorakale Nadelbiopsie

c Mediastinale Biopsie: Mediastinoskopie oder ersatzweise Mediastinotomie

2. und/oder zur Verifikation des CT-Eingangsbefundes im Diagnose- und/oder im Stagingschritt (Strategien E - H), und
3. als alleiniges Untersuchungsverfahren sowohl zur Diagnose als auch zur Stadieneinstufung von NSCLC (Strategie I).

Dabei wird, ist die Methodik des Eingangstests im Diagnose- und Stagingschritt identisch, de facto nur eine CT- oder PET-Aufnahme, die beiden Zwecken gleichermaßen dient, angefertigt.

Die Unterschiede zwischen den alternativen Versorgungsstrategien hinsichtlich Gesamtversorgungskosten sind zumeist nicht groß (Ausnahme Strategien A und B) und bezüglich dem gesundheitsbezogenen Outcome eher gering (siehe hierzu Tabelle 12 auf der nächsten Seite). So schwankt die „mittlere“ Lebenserwartung meist nur um wenige Tage (Maximaldifferenz zwischen Strategie F und A: 73 Tage zugunsten von Strategie F). Dennoch sei die Differenz in der Lebenserwartung bezogen auf den einzelnen Patienten - nach Aussage der Autoren - klinisch signifikant, da die, bei den einzelnen Strategien zu erzielende Lebenserwartung über das gesamte Kollektiv gemittelt wird (Durchschnittsbetrachtung über alle untersuchten Patienten mit und ohne NSCLC), aber nur eine Minorität der untersuchten Personen tatsächlich an NSCLC erkrankt ist.

Wie aus Tabelle 12 auf der nächsten Seite zu ersehen ist, verursachen die Strategien A, B, G, H und I bei vergleichbarer oder einer um maximal 69 Tagen verkürzten durchschnittlichen Lebenserwartung höhere Kosten als der aktuelle Routinealgorithmus mit einer mittleren Lebenserwartung von 10,52 Jahren und Gesamtversorgungskosten von rund US\$ 12 700 (Strategie D). Folglich ist die derzeit übliche Vorgehensweise (Strategie D) dominant in Relation zu den genannten Strategien. Die nach dem ECRI-Modell überlegene Strategie F, bei der die PET zur Bestätigung des CT-negativen Vorbefundes im Stagingschritt dient, erzielt gegenüber der Routinestrategie D bei nahezu identischer Lebenserwartung (durchschnittlicher Zugewinn von 1,8 Tagen pro Patient) ein Einsparpotential von US\$ 36 pro untersuchtem Patienten. Hochgerechnet auf das US-Bevölkerungsniveau ergibt sich somit bei circa 130 000 NSCLC-Patienten pro Jahr ein Einsparpotential von jährlich US\$ 4,7 Millionen. Im Vergleich zu Strategie F erzielt Strategie E bei nahezu identischen Kosten auf der Effektseite eine leicht verminderte Lebenserwartung (Differenz: 0,06 Jahre).

Die Auswahl einer geeigneten diagnostischen Strategie beinhaltet immer auch einen Trade-off zwischen der Möglichkeit einerseits im Diagnoseschritt Tumorerkrankungen und im Stagingschritt Patienten mit operablem NSCLC zu übersehen oder andererseits Patienten mit benignen Tumoren bzw. inoperable Patienten mit Metastasen irrtümlich einer unnötigen bzw. sinnlosen Therapie zu unterziehen. Deshalb werden ergänzend zu der Effektgröße Lebenserwartung auch Outcomegrößen wie der prozentuale Anteil an unnötigen (bei benigner Erkrankung) oder sinnlosen (bei inoperablem NSCLC) Resektionen berichtet. So verbessert die von den Autoren

Tabelle 12: Kosten und Lebenserwartung der alternativen Strategien zur Diagnostik von nicht-kleinzelligem Lungenkarzinom nach ECRI-Bericht 1998, Teil 2

Klinische Strategie	Lebenserwartung [Jahre]	Strategiekosten [US\$]	Kosteneffektivität der Strategien
Strategie A	10,38	15.359	A wird dominiert von den Strategien C - I, die effektiver und billiger sind
Strategie B	10,45	16.650	B wird dominiert von den Strategien C - I, die effektiver und billiger sind
Strategie C	10,47	12.406	Strategie mit den geringsten Kosten, aber auch mit geringeren Effekten als die Strategien D – I
Strategie D	10,57	12.689	derzeitige Routinemethode, wird dominiert von F, da diese effektiver und billiger ist
Strategie E	10,52	12.651	E wird dominiert von F, die etwaseffektiver bei nahezu gleichen Kosten ist
Strategie F	10,58	12.653	Dominante Strategie im ECRI-Entscheidungsmodell
Strategie G	10,57	13.375	G wird dominiert von F, die gleich effektiv aber billiger ist
Strategie H	10,55	13.091	H wird dominiert von F, die effektiver und billiger ist
Strategie I	10,57	13.318	Unwahrscheinlich, daß es sich bei I um eine klinisch valide Strategie handelt

Quelle: ECRI-Bericht 1998, Teil 2

favorisierte Strategie F das Patientenmanagement im Hinblick auf unnötige bzw. sinnlose Resektionen um 0,7%- bzw. 2,3%-Punkte gegenüber dem bisherigen Routineverfahren (Strategie D mit 1,2% unnötigen und 3,8% sinnlosen Operationen) bei einem unveränderten Anteil an übersehenen Krebsfällen von 2,5%.

Die Autoren kommen aufgrund ihrer Analysen zu folgenden Schlußfolgerungen:

1. Der Einsatz der PET als Instrument zur *Diagnose von SPNs*, also zur Differenzierung zwischen malignen und benignen Knoten, ist nicht kosteneffektiv. Stattdessen sollte eine CT-Untersuchung, gefolgt von einer Biopsie zur Bestätigung von CT-positiven Befunden, durchgeführt werden.
2. Beim *Staging von gesichertem NSCLC* ist die PET nur kosteneffektiv als Methode zur Überprüfung von *CT-negativer Befundung* auf nicht resektable Lymphknotenmetastasen im Mediastinum.

3. Zur Bestätigung von *CT-positiven Befunden*, daß mediastinale Lymphknotenmetastasen vorliegen, ist eine invasive Untersuchung des mittleren Brustraumes mittels Biopsie der bildgebenden, nicht-invasiven PET-Untersuchung vorzuziehen.

Aus diesen Schlußfolgerungen leiten die Autoren ihre Empfehlung ab, daß die PET nur im Stagingschritt zur Bestätigung von CT-negativen Patienten angewandt wird; also nur bei Patienten, bei denen aufgrund des CT-Eingangsbefundes - bei gesicherter NSCLC-Diagnose - keine inoperablen, metastasierten Lymphknoten zu erwarten sind. Bei Anwendung dieses Strategieregimes in der klinischen Praxis würden in den USA 60% der Patienten mit gesicherter NSCLC-Diagnose, also 30 000 Personen jährlich, einer PET-Untersuchung zu Stagingzwecken unterzogen.

Um diese Empfehlung zu untermauern werden nach Aussage der Autoren 14 Sensitivitäts- und Thresholdanalysen erstellt: u.a. durch Variation der Kosten (Basisannahme: US\$ 1 200 pro PET-Untersuchung) oder durch strukturelle Veränderungen im Szenarienmodell (Variation der Kosten und des Mortalitätsrisikos bei Ersatz der Biopsiemethode Mediastinoskopie durch Mediastinotomie). Die leider nicht näher ausgeführten Ergebnisse der Sensitivitätsanalysen stützen - nach Aussage der Autoren - die oben dargelegten Schlußfolgerungen.

Studienbewertung

Die ECRI-Studie beschreibt ein sehr komplexes entscheidungstheoretisches Modell, das sich dem Leser nicht ohne Mühe erschließt. Insgesamt ist es jedoch als qualitativ gut einzustufen.

Eine Fragestellung wird nicht explizit ausformuliert, jedoch geben die Autoren eine Begründung für den Entschluß, eine eigene ökonomische ECRI-Studie zu erarbeiten. So werden in früheren Fachpublikationen zur Kosteneffektivität des Einsatzes der PET als Staginginstrument bei NSCLC die Verbesserung des Patientenmanagements durch effektivere Erfassung der richtig-positiven NSCLC-Patienten mittels PET im Vergleich zu anderen nicht-invasiven Verfahren wie CT, verbunden mit einer Senkung von Kosten durch Vermeidung von überflüssigen operativen Eingriffen bei inoperablen Patienten, gewürdigt. Aber die Wirkung einer geringeren Quote an fehlklassifizierten, falsch-positiven Patienten im Vergleich zu den eingeführten Methoden bildgebende CT bzw. invasive Biopsie (Goldstandard) auf die Kosteneffektivität der PET war in diesen Studien nicht Gegenstand der Untersuchung. Die Wirkung dieses Effektes der PET-Diagnostik in einer gesundheitsökonomischen Analyse zu bewerten, machte sich die ECRI zur Aufgabe.

Aus dieser Aufgabenstellung ergibt sich bereits die - als relevant einzuordnende - Auswahl der zu vergleichenden Technologien. Dies sind die eingeführten Routineverfahren CT und Biopsiemethoden, wie transthorakale Nadelbiopsie, Mediastinoskopie bzw. Mediastinotomie, und die Ganzkörper-PET mit ¹⁸FDG als Radiotracer.

Diese diagnostischen Hilfsmittel und ihre Verwendbarkeit bei der Diagnose und im nachfolgenden Staging von NSCLC werden im ersten, medizinischen Teil des HTA-Berichtes eingehend erläutert.

Das entscheidungstheoretische Modell, eine Entscheidungsbaum-Analyse, fällt alleine schon wegen der hohen Zahl an alternativen Strategien, neun ein- bis vierstufige Algorithmen, aus dem Rahmen. Die einzelnen Strategien sind in der ökonomischen Analyse (Teil 2 des ECRI-Berichtes) kurz, dafür jedoch sehr ausführlich im ersten Teil des ECRI-Berichtes beschrieben. Die Modellstruktur, das heißt die einzelnen Äste des Entscheidungsbaumes, werden leider nicht graphisch veranschaulicht, wodurch das Verständnis für die komplexen Zusammenhänge erschwert wird. Die Modellparameter, wie diagnostische Kenndaten der Testmethoden, Kostenangaben und Mortalitätsraten der diagnostischen und therapeutischen Interventionen, Lebenserwartung usw., sind vollständig und übersichtlich gemeinsam mit den zugehörigen Datenquellen in einer Tabelle zusammengefaßt. Als adäquate Datenquellen dienen zum einen Fremddaten aus Fachpublikationen, aber auch Daten aus ECRI-internen statistischen Erhebungen, so einem ICP-ECRI-Survey und einer ECRI-Meta-Analyse. Dieser Meta-Analyse, deren Ergebnisse teilweise im ersten Teil des ECRI-Berichtes veröffentlicht sind, entstammen die Kenndaten zur diagnostischen Treffsicherheit von PET und CT. Die elektronischen Datenbasen, die zur Suche von relevanter Originalliteratur genützt wurden, die Ein- und Ausschlusskriterien und das Suchvokabular sind ebenfalls im ersten Teil des ECRI-Berichtes dokumentiert.

Das Mengengerüst und damit auch die Gesamtkosten für die einzelnen Versorgungsstrategien leiten sich aus der Modellrechnung ab. Das Preisgerüst legt sowohl die Kostenkomponenten als auch den jeweiligen Vergütungsbetrag pro Einzelleistung offen. So sind in den Gesamtkosten für die einzelnen Versorgungsstrategien folgende Kostenposten enthalten:

1. Kosten sowohl für diagnostische als auch für operative und nicht-operative therapeutische Interventionen,
2. Kosten für Komplikationen, beispielsweise Pneumothorax, die bei der transthorakalen Nadelbiopsie oder der Mediastinotomie auftreten können,
3. Zusatzkosten für die Wiederholung von erfolglosen Nadelbiopsien (bei Erstentnahme einer für die histologische Weiteruntersuchung ungeeigneten Gewebeprobe),
4. Kosten für die Behandlung von Tumorrezidiven, und
5. Kosten für die Diagnose und Behandlung von NSCLC-Patienten mit unerkannter Tumorerkrankung (falsch-negative Fälle).

Die Berechnung der Kostenposten für die einzelnen diagnostischen und therapeutischen Interventionsschritte wird ausführlich und gut verständlich dargestellt. Für die im Modell veranschlagten Vergütungsbeträge für einzelne Leistungen sind sowohl der Basiswert als auch die Streubreiten (Eckpunkte der Sensitivitätsanalysen) angegeben. Als - in Ermangelung von verursachungsgemäßen Kostenzurechnungen -

akzeptable Quellen für die Kostendaten dienen Patientenabrechnungen, im Rahmen von klinischen Studien erhobene Kostenkalkulationen und Stellungnahmen von klinischen Experten. Die Originalquellen, zumeist Fachpublikationen, sind sorgfältig dokumentiert. Da nur direkte medizinische Kosten in der Kostenbetrachtung berücksichtigt werden, wird die von den Autoren nicht explizit gewählte Sicht des Gesundheitssystems vertreten. Das Bezugsjahr der Kosten wird nicht genannt; eine Inflations-/Deflationsbereinigung erfolgt nicht.

Auf der Effektseite wird die aus gesundheitsökonomischer Sicht als adäquat und aussagekräftig einzuordnende finale Effektgröße der Lebenserwartung errechnet. Dazu geht als gesundheitsbezogener Parameter die vom Erkrankungsstadium, vom Diagnoseerfolg und vom Therapiemodus abhängige Lebenserwartung in das theoretische Modell ein. Die Datenquellen, Fachveröffentlichungen und statistische Erhebungen, sind dokumentiert. Die finale Effektgröße der Restlebenszeit wird nicht mit Präferenzen zur gesundheitsbezogenen Lebensqualität bewertet. Eine gesundheitsökonomische Maßzahl, in der die Versorgungskosten der einzelnen Alternativstrategien mit der Effektgröße Lebenserwartung in einer Kosten-Effektivitäts-Analyse direkt zueinander in Relation gesetzt werden, wird nicht errechnet. Dies ist jedoch bei den minimalen Unterschieden in der Restlebenszeit (zwischen den Strategien D und F) bzw. den Kosten (zwischen den Strategien E und F) auch nicht sinnvoll. Kriterium der Entscheidung für oder gegen die Wahl einer bestimmten Strategie ist vielmehr deren Dominanz über die Alternativen. Das heißt, die zu bevorzugende Vorgehensweise ist besser bei gleichen oder geringeren Kosten als diese. Obwohl durch die Berücksichtigung von längerfristigen Kosten (z.B. durch Fehldiagnose und Rezidivversorgung induzierte Folgekosten) und von langfristigen Effekten (finale Effektgröße der Lebenserwartung) der Zeithorizont der Studie weit über die akute diagnostische und therapeutische Versorgung der Patienten mit NSCLC-Verdacht hinausreicht, werden weder die Kosten noch die Effekte diskontiert; diese Unterlassung wird nicht begründet.

Die Studienergebnisse werden zumeist in disaggregierter Form - also bezogen auf den einzelnen Patienten - präsentiert. Eine Projektion des Zugewinns an Lebenserwartung bzw. der möglichen Einsparungen auf das US-Bevölkerungsniveau erfolgt nur für den - allerdings wichtigsten - Vergleich zwischen Strategie D (derzeit übliches Verfahren) und der vom Modell favorisierten Strategie F (siehe Studienbeschreibung). Die Hauptschwäche der Studie ist die leider nur sehr knappe Darstellung der Behandlung von Unsicherheiten im Modell anhand von Sensitivitätsanalysen. Die variierten Modellparameter und ihre Spannbreiten sind angegeben, die Ergebnisse werden jedoch nur sehr kurz und pauschal abgehandelt und sind leider nicht graphisch veranschaulicht. Es ist dennoch bemerkenswert, daß der Einfluß einer - klinisch relevanten - strukturellen Veränderung im Modellaufbau, so der Ersatz der Biopsiemethode Mediastinoskopie durch Mediastinotomie, auf den Entscheidungsausgang kritisch beleuchtet wird.

Die Datenqualität und die - unserer Meinung nach problematische - Berechnung einer über alle untersuchten Patienten (auch der nicht erkrankten) gemittelten Lebenserwartung wird in der Diskussion nicht kritisch hinterfragt. Dagegen wird das Fehlen einer Bewertung der gesundheitsbezogenen Lebensqualität der restlichen Lebenszeit von den Autoren thematisiert. Ethische und Verteilungsfragen werden nicht diskutiert, sind jedoch auch nicht Teil der Aufgabenstellung. Die Schlußfolgerungen und die darauf basierende Empfehlung eines adäquaten Einsatzes der PET leiten sich konsistent aus den Studienergebnissen ab. Aufgrund der Komplexität des Modells ist jedoch sowohl die Berechnung der Kosten- als auch der Effektseite für den Leser nicht nachvollziehbar und damit auch nicht nachprüfbar.

Abschließende Bewertung

Das komplexe, differenzierte und umfangreiche entscheidungstheoretische Modell des ECRI, in dem neun alternative Strategien zur diagnostischen und therapeutischen Versorgung von Patienten mit Verdacht auf NSCLC vergleichend evaluiert werden, ist insgesamt als gut zu bewerten und somit geeignet, Entscheidungen zum Einsatz der PET als Diagnose- und Staginginstrument bei NSCLC zu unterstützen. Besonders hervorzuheben sind folgende Studienmerkmale:

1. Die Berücksichtigung sowohl falsch-positiver als auch falsch-negativer Fehlklassifikationen von Patienten durch die PET auf den Entscheidungsausgang,
2. auf der Effektseite die Wahl der finalen und weit über den Zeithorizont der akuten diagnostischen und therapeutischen Versorgung hinausreichenden Effektgröße der Lebenserwartung,
3. auf der Kostenseite der Einbezug von Kosten für mögliche Komplikationen und von längerfristigen Kosten (z.B. bei Rezidivbildung) und
4. die Durchführung von (strukturellen) Sensitivitätsanalysen.

Kosuda S, Ichihara K, Watanabe M et al. Decision-tree sensitivity analysis for cost-effectiveness of chest 2-Fluoro-2-D-[¹⁸F]Fluorodeoxyglucose positron emission tomography in patients with pulmonary nodules (non-small cell lung carcinoma) in Japan. Chest 2000;117(2):346-353

Wie in den westlichen Industriestaaten erhöhten sich in den 1990er Jahren auch in Japan die jährlichen Ausgaben für das Gesundheitswesen stetig: so stiegen die Gesamtausgaben für medizinische Versorgung pro Jahr um rund 1 Billion Yen oder US\$ 7,1 Milliarden an und erreichten 1999 schätzungsweise über 30 Billionen Yen (US\$ 214 Milliarden). Dies entspricht einem Anteil von circa 8% des Bruttoinlandsproduktes. Zeitgleich stiegen auch in Japan die Inzidenz- und Mortalitätsrate für Lungenkrebs (davon 75% nicht-kleinzelliges Lungenkarzinom, NSCLC) und damit auch die dafür aufzuwendenden Versorgungskosten steil an. Um die Kosteneffektivität der neuen Diagnosetechnologie Positronen-Emissions-Tomographie (PET) vor ihrer Einführung an japanischen Kliniken - zumindest in Bezug auf die Indikationsstellung „nicht-kleinzelliges Lungenkarzinom“ (NSCLC) - abschätzen zu können, entwarf die interdisziplinäre klinische Arbeitsgruppe Kosuda et al. zu diesem Zweck ein auf die japanischen Verhältnisse zugeschnittenes entscheidungstheoretisches Modell, das im folgenden erläutert wird.

Studienbeschreibung

In dieser gesundheitsökonomischen Studie, eine Entscheidungsbaum-Analyse, werden am Beispiel eines fiktiven Kollektivs von 1 000 Patienten mit Verdacht auf Lungenkrebs für zwei alternative Diagnose- und Stagingstrategien (plus Therapie) die erwarteten Kostenänderungen und der erwartete Zugewinn an Restlebenszeit durch inkrementelle Analyse bestimmt und einander gegenübergestellt. Die alternativen Versorgungsstrategien sind: (a) ein in Japan üblicherweise angewandter, konventioneller CT-basierter Algorithmus, bestehend aus einer nicht-invasiven CT-Eingangsuntersuchung, gefolgt von einer Biopsie (Bronchofiberoendoskopie) zur Bestätigung des CT-Eingangs-befundes und einer chirurgischen Therapie (Thorakotomie), und (b) ein um die ¹⁸FDG-PET-Technologie erweiterter Versorgungsalgorithmus, mit einer kombinierten CT/PET-Diagnostik (mit CT als Eingangstest), gefolgt von einer verifizierenden Biopsie (Bronchofiberoendoskopie oder Mediastinoskopie) und einer Thorakotomie.

Im vorliegenden Fall soll die CT, die als anatomisches Bildgebungsverfahren der Einstufung des aktuellen Tumorstadiums dient (Stadieneinstufung basierend auf der TNM-Klassifikation: Stadium \leq IIIA versus Stadium \geq IIIB) nicht durch PET ersetzt, sondern durch eine nachgeschaltete PET-Untersuchung zur Überprüfung des CT-Eingangsbefundes ergänzt werden (additiver PET-Einsatz). Die Differenzierung des Tumorstadiums erfolgt in der Studie anhand der räumlichen Ausbreitung von

Lymphknotenmetastasen (Bestimmung des Stadiums des Lymphknotenbefalls: N3 positiv \rightarrow Stadium IIIB, N3 negativ \rightarrow \leq Stadium IIIA). Patienten im Stadium \leq IIIA werden einem kurativen, chirurgischen Eingriff, der Thorakotomie, unterzogen, während inoperable Patienten im Stadium \geq IIIB durch Bestrahlung behandelt werden.

Zur nicht-invasiven, diagnostischen Untersuchung von Patienten mit pulmonalen Knoten und Verdacht auf Lungenkrebs werden bei der bewährten CT-Routinestrategie nur Röntgenaufnahmen des Brustkorbes und bei der erweiterten Alternativstrategie ergänzend PET-Teilkörperaufnahmen im Brustbereich, mit ^{18}F FDG als radioaktivem Tracer, im Anschluß an die CT-Teilkörperscans durchgeführt. Die Befunde der nicht-invasiven Untersuchungen werden durch Biopsien histopathologisch abgeklärt.

Im Modell aufbau werden spezifische nationale Besonderheiten in der diagnostischen Vorgehensweise berücksichtigt: so wird in japanischen Kliniken zumeist die Biopsiemethode der Bronchofiberoendoskopie bevorzugt, während in den westlichen Ländern die Methode der Mediastinoskopie favorisiert wird. Folgerichtig ist, wie aus dem obig geschilderten Modell aufbau hervorgeht, die in Japan übliche Bronchofiberoendoskopie fester Bestandteil des konventionellen CT-basierten Algorithmus, während diese im CT/PET-basierten Alternativalgorithmus teilweise durch die Mediastinoskopie verdrängt wird. So ist eine Mediastinoskopie im Ablaufplan vorgesehen, wenn aufgrund des ^{18}F FDG-PET-Befundes ein N3-Lymphknotenstadium vermutet wird. Eine weitere Besonderheit der gesundheitlichen Versorgung in Japan, die einen entscheidenden Einfluß auf die Kosteneffektivität der CT/PET-Strategie im Vergleich zur CT-Strategie haben kann, sind die deutlich geringeren Vergütungen für medizinische Leistungen (z.B. geschätzter Vergütungsbetrag für eine PET-Untersuchung in Japan: umgerechnet US\$ 700 gegenüber einer Spannbreite von US\$ 1 200 - 2 700 in den USA; Mediastinoskopie: US\$ 860 in Japan gegenüber circa US\$ 1 000 - 9 000 in den USA; US-Angaben sind dem ECRI-Report 1998 entnommen).

Die Untersuchung des Brustkorbes bei Patienten mit Verdacht auf Lungenkrebs anhand der kombinierten CT/PET-Strategie ist gemäß den Ergebnissen der vorliegenden Studie im Vergleich zum in Japan bisher praktizierten CT-Verfahren aus medizinischer Sicht von Vorteil. Beim additiven Einsatz der PET im Anschluß an eine CT-Eingangsuntersuchung zur Absicherung des CT-Vorbefundes steigt die Zahl der kurativen Operationen (Thorakotomie), während die Zahl der nicht-kurativen Operationen sinkt: so erhöht sich im Modell die Zahl der Thorakotomien bei kurablen Patienten (Stadium der Metastasierung regionärer Lymphknoten: $N \leq N2$) von 347 (CT-Strategie) auf 462, gleichzeitig sinkt die Zahl der unnötigen Thorakotomien bei Patienten mit benignen Lungenläsionen von ursprünglich 70 (CT-Strategie) auf nunmehr nur noch 19 operative Eingriffe. Zudem werden unter der CT-Strategie 133 Patienten mit malignen, aber resektablen Lungenläsionen (Lymphknotenstadium $N \leq N2$) nicht einer kurativen Thorakotomie, mit einer postoperativen Lebenserwartung von 7 Jahren, unterzogen, sondern irrtümlich mittels Strahlentherapie, mit einer Lebenserwartung von nur 2 Jahren, behandelt; in der CT/PET-Strategie dagegen werden alle Pa-

tienten mit malignem, resektablem Lungentumor auch zur kurativen Operation überstellt. Eine derartige Verbesserung im Patientenmanagement als direkte Folge einer effizienteren Diagnostik verdächtiger Lungenknoten verursacht im Modell eine Verlängerung der Lebenserwartung pro Patient von 7,3 Monaten (durchschnittliche Betrachtung über alle untersuchten Patienten).

Dieser Verbesserung auf der Effektseite in Form einer verlängerten Lebenserwartung steht - bei einer Vergütung von 100 000 Yen pro ^{18}F FDG-PET-Untersuchung (rund US\$ 700) - auf der Kostenseite eine 10,5%ige Zunahme der Gesamtversorgungskosten pro Patient von ursprünglich 1,26 Millionen Yen (rund US\$ 9 000) auf nunmehr 1,39 Millionen Yen (rund US\$ 10 000) gegenüber. Setzt man die Kostensteigerung von 132 000 Yen pro Kopf zum Gesundheitseffekt einer durchschnittlich um 7,3 Monate verlängerten Lebenserwartung pro Patient durch Quotientenbildung direkt ins Verhältnis, ergibt sich eine inkrementelle Kosten-Effektivitäts-Relation (iKER) von 218 000 Yen/Jahr/Patient. Das heißt, für jeden Patienten müssten pro zusätzlich gewonnenem Lebensjahr Mehrkosten von 218 000 Yen (circa US\$ 1 550) aufgebracht werden. Da in Japan zum Zeitpunkt der Veröffentlichung der Studie noch kein Vergütungsbetrag für PET vom Gesundheitsministerium festgelegt war, variieren die Autoren die PET-Kosten über eine Spanne von 30 000 - 200 000 Yen pro Untersuchung. In dieser univariaten Sensitivitätsanalyse ergibt sich selbst bei einer Absenkung der PET-Vergütung auf einen Mindestbetrag von 30 000 Yen (US\$ 214) für die kombinierte CT/PET-Strategie gegenüber der CT-Strategie pro Patient eine iKER von rund 62 000 Yen (circa US\$ 440) pro zusätzlichem Lebensjahr; das heißt, auch bei einem sehr niedrigen PET-Entgelt wirkt sich der additive Einsatz der PET - unter den gegebenen Modellkonditionen - nicht kostensparend in Relation zum üblichen Verfahren aus.

Da die Spezifitätswerte der ^{18}F FDG-PET bei der Diagnose von Lungenkrebs und die Sensitivitätswerte der ^{18}F FDG-PET bei der Einstufung der N-Stadien (Beurteilung der räumlichen Ausbreitung von Metastasen in den regionären Lymphknoten) zwischen einzelnen Untersuchungen und von Klinik zu Klinik nach Aussage der Autoren deutliche Unterschiede aufweisen, werden diese diagnostischen Kenngrößen zur Überprüfung der Robustheit des Modells in univariaten Sensitivitätsanalysen in einem Wertebereich von 50 - 90% variiert. Während bei Absenkung der PET-Spezifität für Lungenkrebs von 90% auf 50% die erwarteten inkrementellen Kosten stetig auf etwa das 1,5fache* bzw. bei Absenkung der PET-Sensitivität für N3-Lymphknotenbefall von 90% auf 50% stetig auf das 2,4fache⁵ steigen, bleibt in beiden Fällen der Gewinn an Lebenserwartung trotz veränderter diagnostischer Treffsicherheit der ^{18}F FDG-PET nahezu stabil. Eine Schlüsselvariable für die Kosteneffektivität von Versorgungsstrategien bei Lungenkrebs ist der Anteil von malignen Lungenknoten relativ zu benignen Knoten (Prävalenz der Malignität). Anhand einer Threshold-Analyse läßt

⁵ Die mit * gekennzeichneten Angaben wurden aus Abbildung 6 im Originalartikel von Kosuda et al. 2000 entnommen.

sich durch stufenweise Erhöhung des prozentualen Anteil maligner Tumore über eine Spanne von 10 - 90% zeigen (Basisannahme: 71,4%):

1. Auf der Effektseite steigt auch der Zugewinn an Lebenserwartung im CT/PET-Modell stetig von 0,2 auf 0,7 Jahre/Patient*⁶ an,
2. gleichzeitig verursacht auf der Kostenseite - bei Vorgabe einer PET-Vergütung von 100 000 Yen - die kombinierte CT/PET-Strategie bei Anteilen an malignen Tumorerkrankungen von $\geq 40\%$ einen stetig wachsenden Mehrkostenbedarf. Nur wenn die Häufigkeit einer malignen Erkrankung unter dem Schwellenwert von 40% läge, ergäben sich bei der CT/PET-Kombination in Relation zur CT-Strategie Nettoeinsparungen in den Gesamtversorgungskosten. Das heißt, die CT/PET-Strategie würde nur für kleine Prävalenzwerte der Malignität bei Verdacht auf Lungenkrebs - also nur für Werte unter der Basisannahme (71,4%) des Modells - dominant und wäre dann die Vorgehensweise der Wahl.

Die Erwartungen der Autoren, daß sich durch eine kombinierte CT/PET-Diagnostik bei Patienten mit Verdacht auf Lungenkrebs unter den spezifisch japanischen Konditionen der Versorgung dieser Patientengruppe ein Einsparpotential erschließen ließe, wurden nicht erfüllt, da der verbesserten Wirksamkeit der Kombinationsdiagnostik gleichzeitig höhere Ausgaben gegenüberstehen (positiver Kosten-Effektivitäts-Quotient).

Studienbewertung

Bei der hier vorgestellten Studie handelt es sich um die einzige japanische Evaluation zur Kosteneffektivität der PET, die bei Abschluß dieses HTA-Berichtes im Volltext in englischer Sprache vorlag. Aufgrund der beschriebenen nationalen Besonderheiten ist die Übernahme von Ergebnissen aus zumeist in den USA erarbeiteten ökonomischen Evaluationen zur Unterstützung von gesundheitspolitischen Entscheidungen nicht ohne weiteres möglich und deshalb eine auf japanische Konditionen adaptierte Evaluation notwendig und zu begrüßen.

Die Evaluation ist insgesamt betrachtet als gut einzustufen, obwohl sie einige methodische Schwächen aufweist, speziell was die Transparenz der Kostenberechnung betrifft (siehe unten). Eine Schwäche der Studie ist, daß sich die Zielsetzung der Analyse dem Leser nicht unmittelbar erschließt. Erst durch eine genaueres Studium der alternativen Entscheidungsbäume wird deutlich, daß die Kosteneffektivität des klinischen Einsatzes der PET bei der Diagnose und beim Staging von nicht-kleinzelligem Lungenkrebs untersucht werden soll.

Durch eine kurze, aber informative Darstellung des medizinischen und ökonomischen Problemkontextes wird der Leser in die Thematik eingeführt und mit den für die Studie relevanten nationalen Besonderheiten vertraut gemacht. Die Zielpopulati-

⁶ Angaben wurden aus Abbildung 6 im Originalartikel von Kosuda et al. 2000 entnommen.

on ist von den Autoren klar definiert: das fiktive Studienkollektiv stellen Patienten mit pulmonalen Knoten, bei denen eine Erkrankung mit nicht-kleinzelligem Lungenkarzinom, im Stadium \leq IIIB / keine Fernmetastasen, vermutet wird. Der Zeithorizont ist von den Autoren nicht vorgegeben, erstreckt sich jedoch durch den Einbezug des Modellparameters „zukünftige Lebenserwartung“ auf der Effektseite über die kurze Zeitspanne der Akutversorgung (diagnostische und operative Interventionen) hinaus.

Die Modellstruktur der Entscheidungsbäume wird anhand von komplexen, aber dennoch anschaulichen Baumdiagrammen vorgestellt. Dennoch hätte eine zusätzliche Textbeschreibung der alternativen Vorgehensweisen das Verständnis der komplexen Abläufe sehr erleichtert. Die im Modell benötigten Modellparameter, Daten zur diagnostischen Effektivität der Diagnosemethoden (CT, PET, Biopsieverfahren), Mortalitätsraten der Interventionen, Angaben zur Lebenserwartung in Abhängigkeit vom Therapiemodus und Kostendaten werden benannt und ihre Quellen dokumentiert. Die meisten Modellparameter werden übersichtlich in Tabellen aufgeführt, ein Teil der Parameter ist jedoch nur aus dem Text zu entnehmen, was dem Leser eine schnelle Orientierung erschwert.

Um den in der Studienbeschreibung angesprochenen nationalen Besonderheiten Rechnung zu tragen, stammen die im theoretischen Modell verwandten Parameter aus sehr unterschiedlichen Quellen. Wie in modellhaften Analysen meist üblich, werden die Angaben zur diagnostischen Treffsicherheit der bildgebenden Verfahren CT bzw. PET (Sensitivität und Spezifität), das Mortalitätsrisiko der diagnostischen Interventionen (keine Angabe zur Bronchofiberoendoskopie) und der invasiven Thorakotomie aus der Fachliteratur entnommen. Die Originalliteraturstellen sind dokumentiert. Wie die Fachliteratur recherchiert (Suchstrategie) und nach welchen Kriterien die im Modell verwandten Parameter identifiziert und ausgewählt wurden, wird nicht berichtet. Ein Teil der Daten, wie die Auftretenshäufigkeit maligner Lungenknoten (Prävalenz der Malignität), die Erfolgsrate der Bronchofiberoendoskopie und die Kostenangaben (außer für PET) entstammen einer hausinternen klinischen Studie. Da diese Daten jedoch nur auf einer kleinen Fallzahl von 56 Patienten mit pulmonalen Knoten beruhen, ist ihre Generalisierbarkeit kritisch zu hinterfragen. Da zum Zeitpunkt der Studie in Japan noch keine Vergütung für eine PET-Untersuchung festgelegt war, legen die Autoren der Kostenrechnung einen Betrag von 100 000 Yen als Basiswert zugrunde. Um einige dieser Unsicherheiten in der Datenbasis zu kompensieren, wird der Einfluß der Modellparameter „Prävalenz der Malignität“ und „Betrag der PET-Vergütung“ auf den Entscheidungsausgang anhand von Sensitivitätsanalysen geprüft.

Auf der Kostenseite werden die direkten, diagnostischen und therapeutischen Gesamtkosten für die Patientenversorgung erfaßt. Das der Kostenberechnung zugrundeliegende Mengengerüst ist zwar nicht im Detail aufgeführt, ergibt sich jedoch aus dem Entscheidungsbaum, da als Endpunkte der einzelnen Modelläste jeweils die Zahl der Patienten pro Versorgungspfad quantifiziert ist. Unterschiede im Mengengerüst zwischen der CT-basierten und der CT/PET-basierten Strategie sind zur besse-

ren Übersichtlichkeit gesondert in einer Tabelle zusammengestellt. Die Definition des Preisgerüsts zur monetären Bewertung des Mengengerüsts stellt die Hauptschwäche der vorliegenden Evaluation dar, da die Kostenangaben für medizinische Leistungen unvollständig und intransparent sind:

1. So ist aus der Kostenberechnung nicht zu ersehen, ob und in welcher Höhe Geldbeträge für die Strahlentherapie bei inoperablen Patienten bzw. für das Monitoring von PET-negativen-Patienten (nur beim CT/PET-Algorithmus) veranschlagt werden, während die Kosten für den Therapiemodus „kurative Thorakotomie“ einbezogen werden.
2. Im Preisgerüst ist ein Leistungspaket für die ambulante Patientenuntersuchung als Kostenkomponente aufgeführt, deren Verwendung im Entscheidungsbaum nicht eindeutig definiert ist.

Um die für Japan charakteristische Vergütungspraxis, gekennzeichnet durch niedrige Vergütungssätze, angemessen berücksichtigen zu können, setzen die Autoren klinikspezifische, durchschnittliche Kosten für erbrachte Leistungen in die Modellrechnung ein. Diese wurden klinikintern über den Zeitraum April 1996 - März 1997 anhand der Entlassungsrechnungen für ambulante und stationäre Patienten ermittelt. Rechnungsbelege sind als adäquate Datenquellen zu werten, solange keine verbrauchsgemäßen Kostenzurechnungen vorliegen. Eine aktuelle Anpassung der Kosten unterbleibt. Da zur Berechnung der Gesamtkosten klinikbezogene Kostendaten verwendet werden und zudem keine indirekten Kostenkomponenten berücksichtigt werden, erfolgt die Analyse aus der - von den Autoren nicht explizit festgelegten - Perspektive des Leistungserbringers. Die Kostenangaben in der Landeswährung Yen werden zumeist, zur besseren Vergleichbarkeit, in US\$ umgerechnet; der Konversionskurs ist angegeben und die Umrechnung korrekt durchgeführt.

Auf der Effektseite geht mit der Größe „Lebenserwartung“ eine aussagekräftige, patientenbezogene Effektgröße (patient outcome efficacy) in das entscheidungstheoretische Modell ein. Die gesundheitsbezogene Qualität der restlichen Lebenszeit wird nicht bewertet (keine QALY-Berechnung) und auch in dieser Lebensspanne möglicherweise auftretende „downstream-events“, wie beispielsweise Rezidivbildung, bleiben sowohl auf der Kosten- als auch auf der Effektseite unberücksichtigt. Die Effektseite wird mit der Kostenseite zu einer sinnvollen Maßzahl, dem inkrementellen Kosten-Effektivitäts-Quotienten „zusätzliche Kosten/gewonnenem Lebensjahr/Patient“ zusammengeführt.

Die Ergebnisse, speziell die Resultate der - relevanten - Sensitivitätsanalysen, werden anschaulich und übersichtlich anhand von Tabellen und Schaubildern (Sensitivitätsanalysen) in disaggregierter Form dargestellt und zudem im Text verständlich erläutert. Sowohl absolute Kosten und klinische Effekte, wie Gesamtversorgungskosten pro Kopf und Verbesserungen im Patientenmanagement (u.a. Anzahl der vermiedenen, unnötigen Operationen), als auch inkrementelle Gesundheitseffekte (Zugewinn an Lebensjahren pro Kopf) und inkrementelle Kosten (Mehrkosten pro ge-

wonnener Lebenszeit pro Kopf) werden berichtet. Die disaggregierten, auf den einzelnen Patienten bezogenen Kosten und Effekte werden nicht auf das japanische Bevölkerungsniveau hochgerechnet.

Die Qualität der klinikintern erhobenen Daten und ihre Generalisierbarkeit wird von den Autoren kritisch diskutiert und damit eine der Hauptschwächen der Studie angesprochen. Auch strukturelle Unterschiede zwischen dem japanischen und den westlichen Gesundheitssystemen werden thematisiert. Ethische und Verteilungsfragen, beispielsweise ob bei einem 1jährigen Zugewinn an Lebenserwartung ein Kosten-Effektivitäts-Verhältnis von 218 000 Yen (US\$ 1 550) pro Patient in Japan gesellschaftlich akzeptabel ist, werden nicht erörtert. Dieser finanzielle Mehraufwand pro zusätzlichem Lebensjahr liegt jedoch deutlich unter dem von Gambhir et al. (1998) berichteten, in den USA gesellschaftlich akzeptierten Wert von US\$ 50 000 pro gewonnenem Lebensjahr.

Abschließende Bewertung

In der japanischen Studie zur Kosteneffektivität des klinischen Einsatzes der PET werden zwei alternative Versorgungsstrategien, ein konventioneller, CT-basierter Algorithmus und ein kombinierter CT/PET-basierter Algorithmus (additiver PET-Einsatz), zur Diagnose mit anschließendem Staging von Patienten mit Verdacht auf nicht-kleinzelliges Lungenkarzinom vorgestellt und deren Kosten und Effekte mit Hilfe einer inkrementellen Kosten-Effektivitäts-Analyse verglichen. Trotz methodischer Mängel, vor allem im Hinblick auf die Transparenz der Gesamtkostenrechnung, ist die Studie, basierend auf einem entscheidungstheoretischen Modell, insgesamt als gut zu bewerten. Sie zeichnet sich durch folgende Punkte aus:

1. In der Studie werden nationale Besonderheiten des japanischen Gesundheitssystems erläutert und im Modellaufbau bzw. bei der Festlegung des Preisgerüsts berücksichtigt.
2. Auf der Effektseite werden durch die Wahl der Effektgröße „Lebenserwartung“ langfristige Effekte in die Betrachtung aufgenommen und damit der Zeithorizont der Studie über die Zeitspanne der reinen Diagnosestellung hinaus ausgedehnt.
3. Zur Überprüfung der Stabilität des Modells werden Sensitivitätsanalysen durch Variation unsicherer Parameter durchgeführt.

Die Ergebnisse der Evaluation von Kosuda et al. sind insgesamt als fundiert anzusehen, so daß die Studie als gute Ausgangsbasis für weitere, auf japanische Verhältnisse zugeschnittene ökonomische Evaluationen dienen kann.

Von Schulthess GK, Steinert HC, Dummer R et al. Cost-effectiveness of whole-body PET imaging in non-small cell lung cancer and malignant melanoma. Acad Radiol 1998;5 (Suppl 2):S300-S302

Beschreibung der klinischen Studien

Die Schweizer Arbeitsgruppe um von Schulthess veröffentlichte 1998 nachträglich eine ökonomische Analyse zu zwei rein klinischen Studien zur Wirksamkeit und Verwendbarkeit der Positronen-Emissions-Tomographie mit ¹⁸FDG (¹⁸FDG-PET) als Staging-Verfahren in den Indikationsgebieten „nicht-kleinzelliges Lungenkarzinom“ (NSCLC) und „maligne Melanome“ (MEL). In beiden klinischen Studien waren PET-Ganzkörperscans mit ¹⁸F-markiertem FDG als Tracer erfolgreich angewandt worden. Zwischenergebnisse dieser klinischen Studien wurden 1995 bzw. 1997 veröffentlicht (NSCLC: Steinert et al. 1997; MEL: Steinert et al. 1995).

In der 1995 publizierten MEL-Studie an 33 Patienten mit bekanntem, metastasierendem oder inzidentem Melanom wurde für PET die diagnostische Treffsicherheit als Staginginstrument, angegeben als Sensitivität, Spezifität und Genauigkeit bestimmt (Referenzverfahren: mehrstufiger Bilderfassungsalgorithmus – CT, MRI, PET – plus Biopsie und histopathologische Untersuchung): Sensitivität für die Detektion von Metastasen 92%, Spezifität ohne bzw. mit zusätzlichen klinischen Informationen 77% bzw. 100%, Genauigkeit 89%. In der NSCLC-Studie von 1997 an 47 Patienten mit vermutetem oder inzidentem Lungenkrebs wurde für PET und für das alternative CT-Verfahren zusätzlich zu diesen diagnostischen Kenngrößen noch der positive und der negative Vorhersagewert (PPW bzw. NPW, in Korrelation zu den Ergebnissen des histopathologischen Stagings) berichtet: Sensitivität der PET (bzw. CT) für die Detektion von mediastinalen Lymphknotenmetastasen (N2- oder N3-Stadium) 89% (57%), Spezifität 99% (94%), PPW / NPW 96% / 97% (76% / 87%) und Genauigkeit 96% (85%). Nach Veröffentlichung dieser Zwischenergebnisse wurden beide Studien fortgeführt und die Fallzahlen durch Aufnahme weiterer Probanden erhöht.

Aus den Ergebnissen dieser klinischen Studien schlußfolgern die Autoren, daß die ¹⁸FDG-PET-Technologie sowohl ein diagnostisch effektives Screeningverfahren bei Verdacht auf Metastasenbildung bei malignen Melanomen darstellt als auch der alternativen CT-Methode beim Staging von Lymphknoten bei NSCLC-Patienten diagnostisch überlegen ist.

Beschreibung der ökonomischen Studie, Indikationsstellung: NSCLC

In die prospektive klinische NSCLC-Studie, auf der die hier vorliegende ökonomische Analyse basiert, waren 62 Patienten mit nicht-kleinzelligem Lungenkrebs einbezogen; 46 Personen des Patientenkollektivs wiesen operable (Stadium der Metastasie-

rung der regionären Lymphknoten: $N < N3$), die restlichen 16 Studienteilnehmer bereits inoperable Läsionen auf.

In der retrospektiven ökonomischen Analyse berechnen die Autoren für zwei alternative, kombinierte Diagnostik/Therapie-Strategien, PET/CT versus CT/Skelettszintigraphie jeweils mit anschließender operativer Resektion der Tumorkläsionen, die jeweilig anfallenden Gesamtversorgungskosten (direkte medizinische Kosten der Primärinterventionen). Neben der - laut Aussage der Autoren - besseren diagnostischen Wirksamkeit der PET (Sensitivität bzw. Spezifität für PET $> 90\%$, diagnostische Kenndaten der übrigen Diagnosemethoden sind nicht angegeben) stellt sich gemäß dieser Kostenrechnung die PET/CT-Kombination auch als die kostengünstigere Verfahrensvariante dar. Aufgrund der höheren diagnostischen Effektivität von PET versus CT, jeweils eingesetzt als Eingangstest, konnten in der klinischen Studie anstatt bei 10 Patienten (CT/Skelettszintigraphie) bei allen 16 Patienten (PET/CT) mit inoperablem Lungenkrebs die Diagnose richtig gestellt werden. Aufgrund dieses treffsicheren PET/CT-Befundes hätten somit 6 unnötige Operationen vermieden werden können. Durch den Einsatz der PET als Eingangsuntersuchung hätten gleichzeitig 16 CT-Scans eingespart werden können, da beim alternativen PET/CT-Verfahren morphologische CT-Scans - zur besseren Vorbereitung einer chirurgischen Resektion - ausschließlich bei Lungenkrebspatienten angefertigt werden, die in der vorausgegangenen PET-Eingangsuntersuchung als operabel eingestuft werden ($n = 46$). Bei der alternativen, konventionellen Vorgehensweise mittels CT/Skelettszintigraphie hingegen werden alle Studienteilnehmer obligatorisch sowohl einer CT-Eingangsuntersuchung ($n = 62$) als auch einer Skelettszintigraphie unterzogen. Nach Berechnung der Autoren hätten - bezogen auf das Patientenkollektiv - somit in der klinischen Studie mit Hilfe der kombinierten PET/CT-Diagnostik aufgrund des verbesserten Patientenmanagements im Vergleich zur konventionellen Kombinationsdiagnostik mittels CT/Skelettszintigraphie letztlich Minderausgaben von 100 200 sfr (Schweizer Franken) verbucht werden können. Diesen Berechnungen liegt ein Gesamtfinanzbedarf von circa 1,3 Millionen sfr für die PET-basierte, diagnostische und therapeutische Versorgung der 62 in die Studie eingeschlossenen NSCLC-Patienten zugrunde, gegenüber 1,4 Millionen sfr für die CT-basierte Versorgung (zur Berechnung siehe Studienbewertung).

Beschreibung der ökonomischen Studie, Indikationsstellung: MEL

Im Verlauf der prospektiven MEL-Studie wurde zwischen 1994 und März 1997 bei 100 Patienten mit malignen Melanomen und erhöhtem Risiko der Metastasierung das aktuelle Erkrankungsstadium bestimmt. Am Beispiel dieses 100köpfigen Studienkollektivs errechnen die Autoren in ihrer retrospektiven ökonomischen Betrachtung die Stagingkosten für drei verschiedenartige Untersuchungsstrategien:

1. Zwei PET-gestützte Strategien: Diagnostik mittels klinischer Daten, Ganzkörper-¹⁸F-DG-PET, mit und ohne zusätzlichen CT-Scan im Kopfbereich;

2. konventionelle Diagnostik mittels klinischer Daten, Röntgenuntersuchung des Brustkorbes und Ultraschalluntersuchung der Unterleibsregion und der regionären Lymphknoten.

Die PET-gestützten Verfahren sind gemäß den Berechnungen der Autoren 1,7% bis 19% teurer als das konventionelle Staging-Verfahren; nur beim Staging von Patienten mit bekannter Metastasierung ergibt sich gegenüber der konventionellen Methode eine Kostenminderung von 11,4% (bei Gesamtkosten des PET-gestützten Algorithmus von rund 201 000 sfr gegenüber rund 227 000 sfr für die konventionelle diagnostische Vorgehensweise; das Bezugsjahr der Kostenberechnung ist unbekannt).

Studienbewertung

Die Arbeitsgruppe um von Schulthess stellt in einer gemeinsamen Veröffentlichung zwei unabhängige Untersuchungen zum Ressourcenverbrauch beim Staging von NSCLC- bzw. MEL-Patienten vor, die in ihrer Qualität unterschiedlich zu beurteilen sind (siehe dazu die Tabellen 6 und 7). Zielsetzung der Autoren ist - nach eigenen Angaben - die Kosteneffektivität von PET als Stagingverfahren in den Indikationsgebieten NSCLC und MEL zu prüfen. Wie die nachfolgenden Ausführungen zeigen werden, handelt es sich bei den zu beurteilenden Untersuchungen nicht um ökonomische Evaluationen im strengen Sinn, sondern um reine Kostenberechnungen für alternative Versorgungsstrategien in den genannten Indikationen (partielle Evaluationen).

Die allgemein sehr knappe Darstellung des Studienrahmens, der klinischen und ökonomischen Ausgangsdaten, des Berechnungsweges zur Bestimmung der Kosten (MEL-Studie) und der Ergebnisse sind grundsätzliche Mängel dieser Veröffentlichung. So werden die in die Studien einbezogenen Technologien, die Studienpopulationen und das jeweilige Studiendesign nur bruchstückhaft beschrieben. Auch eine ausführlichere und detailliertere Darstellung der relevanten klinischen Effekte wäre wünschenswert; speziell die Datenquellen sind unzureichend dokumentiert. Aus den Veröffentlichungen der Zwischenergebnisse der klinischen Studien von 1995 bzw. 1997 können die relevanten Daten ebenfalls nicht entnommen werden, da die dort beschriebenen Studienpopulationen nicht deckungsgleich mit den Patientenkollektiven der hier vorliegenden ökonomischen Studien sind.

Auch die Perspektive der Studien wird von den Autoren nicht angegeben, entspricht jedoch entweder der Sicht des Leistungserbringers, also der Universitätsklinik, oder des Finanzierungsträgers, da keine für die gesellschaftliche Perspektive relevanten indirekten Kosten in die ökonomische Betrachtung eingehen.

In der Studie zum Staging von NSCLC-Patienten werden für die zwei alternativen, kombinierten Diagnostik/Therapie-Strategien sowohl Kosten, in Form von Gesamtversorgungskosten für die akute diagnostische und therapeutische Patientenversor-

gung, als auch Effekte, in Form von vermiedenen diagnostischen und therapeutischen Interventionen erhoben und miteinander verrechnet. Dazu wird auf der Effektseite nicht wie in einer reinen Diagnosestudie die intermediäre Effektgröße der diagnostischen Genauigkeit, sondern - durch Erfassung der Anzahl vermeidbarer Interventionen - eine finale Effektgröße verwendet, die den Einfluß der Diagnosestellung auf die Therapieentscheidung widerspiegelt (therapeutic efficacy, siehe C.4.2.2). Die Kalkulation des Einsparpotentials erfolgt dabei korrekt ohne Doppelberechnung der beim PET-Einsatz eingesparten Aufwendungen für vermiedene Operationen und CT-Scans. Der durch PET-Einsatz erzielbare klinische Nutzen in Form von vermiedenen Operationen ($n = 6$) wird monetär bewertet (Durchschnittskosten pro operativem Eingriff: 25 000 sfr; $\Rightarrow 6 \times 25\,000\text{ sfr} = 150\,000\text{ sfr}$). Von diesen eingesparten Operationskosten in Höhe von 150 000 sfr sind Mehrkosten von 49 800 sfr für die aufwendigere PET-basierte Diagnostik (Differenz der Diagnosekosten der beiden Alternativstrategien) abzuziehen. Zum selben Ergebnis kommt man, wenn man die Gesamtversorgungskosten der alternativen Strategien bilanziert (Gesamtkosten für CT/Skelettszintigraphie von 1 393 000 sfr abzüglich den Gesamtkosten für PET/CT von 1 292 800 sfr \Rightarrow Differenz von 100 200 sfr). Die Kostenberechnung selbst ist anschaulich und gut nachvollziehbar dargestellt, weist jedoch methodische Mängel auf. Zwar können das Mengengerüst, angegeben in verbrauchten Versorgungsmengen (Anzahl der erbrachten Einzelleistungen), und die Vergütungsbeträge pro Einzelleistung der Kostenberechnung problemlos entnommen werden, aber das Preisgerüst ist nur unvollständig beschrieben. So sind weder Art (betriebliche Selbstkosten oder administrative Preise?) und Zusammensetzung, noch die Quellen der monetären Bewertung der diagnostischen und operativen Einzelleistungen bekannt.

Bei der MEL-Evaluationsstudie werden ebenfalls nur Kosten für die medizinische Leistungserstellung und keine daraus resultierenden gesundheitlichen Effekte berücksichtigt (reine Kostenbetrachtung). Dabei werden ausschließlich die beim Staging anfallenden direkten medizinischen Kosten erhoben, wodurch der Zeithorizont der Studie auf die Diagnosestellung eingengt wird. Kosten für die therapeutische Versorgung der MEL-Patienten (wie in der NSCLC-Untersuchung) oder sonstige direkte Kosten, zum Beispiel für die Behandlung von Komplikationen, werden nicht in die Betrachtung aufgenommen. Ein derart zeitlich eng gefaßter Ansatz ist aus gesundheitsökonomischer Sicht nicht zufriedenstellend (siehe C.4.2.1). Die Kostenbetrachtung weist zudem gravierende methodische Mängel auf: so sind Mengen- und Preisgerüst unvollständig und die Datenquellen und der Berechnungsmodus sind aus der Publikation nicht ersichtlich. Somit ist die vergleichende Kostenberechnung für die drei alternativen Vorgehensweisen beim Staging von MEL in der vorliegenden Form für den Leser nicht nachvollziehbar.

Zudem werden für beide Indikationsstellungen weder das Bezugsjahr der Kostenberechnung genannt, noch wird eine Inflations-/Deflationsbereinigung vorgenommen. Auch eine Konversion der Landeswährung sfr (Schweizer Franken) in eine internationale Währung (US\$) zwecks besserer Vergleichbarkeit mit anderen ökonomischen

Studien unterbleibt. Wegen des relativ engen Zeithorizontes erübrigt sich eine Diskontierung.

Die Ergebnisse beider Studien werden kaum diskutiert, methodische Mängel werden nicht kritisch beleuchtet. Unsicherheiten in der Datenlage bleiben unberücksichtigt; so werden keinerlei Sensitivitätsanalysen angestellt. Aufgrund der dargelegten Einschränkungen in der methodischen Qualität sind die Schlußfolgerungen der Autoren nicht ausreichend durch gesicherte Fakten untermauert.

Abschließende Bewertung

Die zwei Kostenvergleiche zum Staging von Patienten mit nicht-kleinzelligem Lungenkarzinom (NSCLC) bzw. malignen Melanomen (MEL) sind nicht als vollständige ökonomische Evaluationen zu betrachten, da die ermittelte Kostenseite nicht zu gesundheitlichen Effekten (Effektseite) in Beziehung gesetzt wird (partielle Evaluation siehe C.4.2.3). Zudem haben die Resultate aufgrund der insgesamt sehr knappen Darstellung der Fakten, der deutlichen methodischen Mängel (vor allem bei der MEL-Studie) und - bei der MEL-Studie - wegen des auf die reine Diagnosestellung beschränkten Zeithorizontes nur eine geringe Aussagekraft und sind somit wenig geeignet, gesundheitspolitische Entscheidungen zu unterstützen.

Sonstige onkologische Indikationsstellungen

Hoh CK, Glaspy J, Rosen P et al. Whole-body FDG-PET imaging for staging of Hodgkin´s Disease and Lymhoma. J Nucl Med 1997;38(3):343-348

Studienbeschreibung

In der zu diskutierenden klinischen Diagnosestudie, die 1997 publiziert wurde, werden zwei Staging-/Restagingstrategien zur Einstufung des Erkrankungsstadiums von Patienten mit malignen Lymphomen hinsichtlich ihrer diagnostischen Wirksamkeit und der anfallenden Kosten verglichen. Als Probanden wurden 18 Patienten im Alter von 17 - 70 Jahren (Durchschnittsalter: 46 ± 14 Jahre) mit der Diagnose „Hodgkin-Lymphom“ (Hodgkin´s Disease, HD, n = 7) bzw. „Non-Hodgkin-Lymphom“ (NHL, n = 11) rekrutiert.

Bei den zwei alternativen Stagingstrategien handelt es sich um einen CT-basierten Diagnosealgorithmus (konventionelles Staging) und einen PET-basierten Diagnosealgorithmus, mit den tomographischen Verfahren CT bzw. PET als jeweilige Kernmethode. Zur Abklärung unsicherer CT- bzw. PET-Befunde wird die CT- bzw. PET-Erstuntersuchung durch konventionelle Zusatzuntersuchungen, wie Kernspintomographie (MRI), radiographische Methoden oder explorative Laparotomie, ergänzt. Im klinischen Versuchsablauf wurde der PET-Scan in einem 8-Ring-Ganzkörper-System angefertigt; als Positronen-emittierender Tracer wurde der Glucoseabkömmling ^{18}F FDG, der den Versuchspersonen 30 min vorab i.v.-injiziert wurde, verwendet. Die PET-Aufnahmen wurden verblindet ausgewertet, das heißt, ohne Zusatzinformationen aus vorausgegangenen CT-Untersuchungen bzw. Biopsien.

Zur Beurteilung der diagnostischen Effektivität der zu evaluierenden Untersuchungsverfahren dient die Anzahl der für jeden Patienten korrekt identifizierten, abnormalen Läsionsstellen, differenziert nach deren regionaler Verteilung im Körper. Dazu wird die Korrektheit der CT- bzw. PET-Untersuchungsergebnisse anhand von nachfolgend durchgeführten Biopsien bzw. Follow-up-Studien verifiziert. Abhängig von Anzahl und Lage der nachgewiesenen Läsionsstellen (primäre Effektgröße) erfolgt eine Einstufung des Erkrankungsgrades der Probanden gemäß der Ann-Arbor-Klassifikation (Staging i.e.S., abgeleitete Effektgröße). Als Referenzwert für die Richtigkeit dieser Stadieneinstufung dient eine abschließende, klinische Bewertung des Gesundheitszustandes der Studienteilnehmer („klinisches“ Staging), in die alle verfügbaren Daten aus dem konventionellen und dem PET-Staging und die Ergebnisse der Biopsie und des klinischen Follow-up eingingen. Als Bewertungsmaßstab zur vergleichenden Beurteilung der alternativen CT- bzw. PET-basierten Algorithmen

wird der Grad der Übereinstimmung an richtig erkannten Läsionsstellen bzw. identischer Stadieneinstufung herangezogen.

Durch „klinisches“ Staging konnten im Patientenkollektiv insgesamt 37 Erkrankungsherde nachgewiesen werden, von denen jeweils 33 (teilweise nicht identische) Läsionsstellen durch die alternativen Diagnosealgorithmen korrekt erfaßt wurden. Beide Untersuchungsstrategien ergaben die gleiche Anzahl ($n = 5$) falsch positiv bzw. falsch negativ diagnostizierter Läsionsstellen. Drei NHL-Patienten wurden mittels PET in ein höheres, ein Patient in ein niedrigeres Tumorstadium eingestuft als durch das konventionelle Staging. Methodische Schwierigkeiten bei der Anwendung der PET ergeben sich (1) bei der eindeutigen Abgrenzung spezifischer ^{18}F FDG-Aufnahme in Tumorknoten, bedingt durch erhöhte Glucose-Stoffwechselaktivität, von unspezifischer ^{18}F FDG-Akkumulation im Darmtrakt, und (2) beim Nachweis von kleinen Läsionen (Knoten $< 1,0$ cm).

Auf der Kostenseite werden ausschließlich die im Rahmen der diagnostischen Leistungserstellung unmittelbar anfallenden direkten Kosten, berechnet als Durchschnittskosten von 5 nicht näher charakterisierten Krankenhäusern, berücksichtigt. Die Gesamtkosten für die Untersuchung der 18 Studienteilnehmer belaufen sich im Rahmen der Studie beim konventionellen CT-Staging auf rund US\$ 68 000; beim Ganzkörper-PET-basierten Alternativverfahren fallen insgesamt Diagnosekosten von circa US\$ 38 000 an.

Die Autoren der Studie ziehen aus ihren Untersuchungsergebnissen den Schluß, daß der PET-basierte Staging-/Restaging-Algorithmus eine genaue und im Vergleich zum konventionellen, CT-basierten Alternativalgorithmus auch eine kosteneffektive Methode darstellt. Beim Einsatz der Ganzkörper-PET könnten zusätzliche Teilkörperaufnahmen mit anatomischen Bildgebungsverfahren, u.a. Röntgenuntersuchungen, MRI, Gallium-Szintigraphie, CT im Kopfbereich, und eventuell auch invasive Diagnosemethoden (Endoskopie, Laparotomie) zur Abklärung eines Tumorverdachts teilweise wegfallen und somit auch die anfallenden Kosten reduziert werden.

Studienbewertung:

Die Fragestellung der Studie und der medizinische Problemkontext sind ausreichend dargestellt, der ökonomische Problemkontext ist aus der Studieneinführung jedoch nicht ersichtlich.

Die Kerntechnologien der miteinander verglichenen Stagingalgorithmen, PET und CT, werden hinreichend gründlich beschrieben, ihre Auswahl schlüssig begründet und der Gesamtablauf der mehrstufigen CT- bzw. PET-basierten Untersuchungsprogramme in einem Fließdiagramm anschaulich dargestellt. Die mit nur 18 Teilnehmern sehr kleine Studienpopulation ist zwar hinsichtlich der HD/NHL-Relation, aber nicht im Hinblick auf die Geschlechterverteilung repräsentativ für die Gesamtbevölkerung,

da männliche Patienten in der Studie unterrepräsentiert sind. Die Perspektive - in der Studie nicht explizit festgelegt - entspricht der Sicht des Leistungserbringers, also hier des Krankenhauses. Der Zeithorizont ist auf die zur Durchführung des Stagingverfahrens benötigte Zeitspanne beschränkt. Diese zeitliche Einschränkung ist jedoch vertretbar, da die vorliegende Studie von den Autoren als reine Diagnosestudie konzipiert wurde und auch als solche zu beurteilen ist. Aus gesundheitsökonomischer Sicht wünschenswert wäre jedoch die Ausweitung des Zeithorizontes über den Zeitraum der diagnostischen Untersuchung hinaus, so daß längerfristige Auswirkungen der Diagnosestellung sowohl auf die Kosten- als auch auf die Effektseite, eventuell durch Änderungen im Patientenmanagement, erfaßt würden.

Die für eine Evaluation benötigten medizinischen und ökonomischen Daten über klinische Effekte und Mengen an verbrauchten Ressourcen (Mengengerüst) wurden den Patientenakten des 18köpfigen Studienkollektivs entnommen (prospektive Datenerhebung). Patientenakten sind als adäquate Datenquellen einzustufen.

Die medizinischen Studienergebnisse sind in disaggregierter Form sehr detailliert beschrieben. Bedingt durch die oben beschriebene Beschränkung des Zeithorizontes allein auf die Diagnosestellung wird als klinischer Effektparameter lediglich der prozedurale Erfolg der alternativen Diagnosealgorithmen, genauer die Anzahl der richtig identifizierten Tumorknoten bzw. der korrekt eingestufteten Tumorstadien, erfaßt. Mögliche mittel- und langfristige Wirkungen, wie Kosteneinsparungen durch vermiedene Operationen als Folge eines verbesserten Stagings oder medizinische Langzeiteffekte auf den gesundheitlichen Outcome, wie Lebenserwartung und Lebensqualität, werden nicht untersucht. Eine zur besseren Vergleichbarkeit der Effektseite, hier der diagnostischen Genauigkeit, hilfreiche Berechnung der Sensitivität bzw. Spezifität der Diagnoseverfahren wird in der Studie leider versäumt.

Auf der Kostenseite werden ausschließlich die direkten medizinischen Diagnosekosten erhoben. Das der Berechnung der Gesamtkosten zugrunde liegende Mengengerüst, differenziert nach Art und Anzahl aller ärztlich angeordneter Einzelleistungen pro Diagnosestrategie, ist detailliert und anschaulich anhand von Balkendiagrammen dargestellt. Das Mengengerüst wird mit Durchschnittspreisen pro Einzelleistung (ermittelt in 5 Referenzkrankenhäusern) bewertet. Die Festlegung des Preisgerüsts ist jedoch für den Leser nicht nachvollziehbar, da die Referenzkrankenhäuser nicht näher charakterisiert (z.B. hinsichtlich ihres Versorgungsniveaus), die Durchschnittskosten nicht nach Kostenarten aufgeschlüsselt und keine Streubreiten angegeben werden.

Aus ökonomischer Sicht handelt es sich bei der vorliegenden Studie um keine vollständige, sondern um eine partielle ökonomische Evaluation: zwar werden Effekte und (direkte medizinische) Kosten erhoben, diese werden jedoch nicht zu einer Kosten-Effektivitäts-Relation zusammengefaßt (siehe dazu C.4.2.3).

Die Währung (US\$) ist genannt. Eine Diskontierung der Effekte bzw. Kosten, ebenso eine inflations-/deflationsbereinigende Anpassung der Kosten ist - sofern aktuelle Kostendaten verwandt wurden - wegen des eingeschränkten Zeitfensters nicht notwendig. Kritisch zu beurteilen ist jedoch, daß aus der Studie auch das Referenzjahr für die Festlegung des Preisgerüsts und damit für die Gesamtkostenkalkulation nicht zu entnehmen ist.

Wesentliche Einschränkungen und Schwächen der Studie im Hinblick auf das klinische Studiendesign und die methodische Durchführung, ebenso die Qualität der klinischen Effektdaten werden eingehend diskutiert. Die Kostenseite wird jedoch nicht kritisch hinterfragt. Um die Aussagekraft der Studie hinsichtlich Kosteneinsparmöglichkeiten besser beurteilen zu können, wäre beispielsweise eine kritische Würdigung des Stellenwertes teurer und somit kostentreibender Zusatzuntersuchungen - wie der Laparotomie (Gesamtkostenanteil beim CT-Algorithmus: 9%) - hilfreich. Sensitivitätsanalysen werden ebenfalls nicht durchgeführt. Sinnvoll wäre eventuell eine Betrachtung der Kostenschwankungen in Abhängigkeit vom Versorgungsniveau der behandelnden Krankenhäuser gewesen.

Ethische Fragen und Fragen zur Verteilung von Gesundheitsleistungen werden in der Diskussion nicht angesprochen, sind aber auch nicht Gegenstand der Aufgabenstellung. Eine Generalisierbarkeit der Ergebnisse ist aufgrund der zu geringen Zahl an Studienteilnehmern (n= 18) nicht gegeben.

Abschließende Bewertung

In der Studie von Hoh et al. (1997) werden zwei alternative Staging-/Restagingstrategien zur Einstufung des Erkrankungsstadiums von Patienten mit malignen Melanomen in erster Linie aus medizinischer und nur begleitend auch aus ökonomischer Sicht evaluiert. Die Studie zeigt ausführlich und verständlich einerseits die methodischen Schwierigkeiten, aber auch die Vorzüge des PET-Verfahrens auf und beschreibt andererseits die im Studienverlauf beobachteten klinischen Effekte detailliert. Aus gesundheitsökonomischem Blickwinkel betrachtet, ist die Studie insgesamt als gute Ausgangsbasis für zukünftige Evaluationen des PET-Verfahrens zur Diagnostik von malignen Lymphomen zu bewerten. Wegen der kleinen Fallzahl (n = 18) und weiterer methodischer Mängel, speziell bei der Erhebung (Preisgerüst) und Kalkulation der Kosten, sind die Ergebnisse der Studie jedoch nicht genügend gesichert, um gesundheitspolitische Entscheidungen zu unterstützen.

Klose T, Leidl R, Buchmann I et al. Primary staging of lymphomas: cost-effectiveness of FDG-PET versus computed tomography. University of Ulm. Department of Health Economics. Working paper series. Working paper number: 2/1999

In einer Bicenterstudie an zwei deutschen Universitätskliniken wurde die klinische Einsatzmöglichkeit der PET-Technologie beim Primärstaging von malignen Lymphomen anhand ihrer diagnostischen Effektivität evaluiert. Die am Studienzentrum Ulm durchgeführte prospektive klinische Wirksamkeitsprüfung wurde zusätzlich durch eine ökonomische Evaluation begleitet. Die Durchführung und die Ergebnisse dieser einzigen in Deutschland erstellten ökonomischen Evaluation zur PET-Thematik, eine inkrementelle Kosten-Effektivitäts-Studie, sind bislang in der Veröffentlichungsreihe „working paper series“ der Abteilung Gesundheitsökonomie der Universität Ulm zusammenfassend dargestellt und werden im Rahmen dieses HTA-Berichtes kurz vorgestellt und diskutiert.

Studienbeschreibung

13 der insgesamt 22 Studienteilnehmer mit malignen Lymphomen waren zu Beginn der klinischen Studie an Lymphogranulomatose (Hodgkin-Lymphom) erkrankt; bei den übrigen 9 Probanden lag eine gesicherte Diagnose auf ein Non-Hodgkin-Lymphom vor. Bei den im Zeitraum April 1997 - Mai 1998 am Universitätsklinikum Ulm zufällig ausgewählten 22 Patienten im Alter von 17 - 66 Jahren (Durchschnittsalter: 39 Jahre) wurde auf der Basis aller zur Verfügung stehenden klinischen Daten, einschließlich der Ergebnisse von PET- bzw. CT-Scans und der nachfolgenden Biopsien, das aktuelle Tumorstadium ermittelt (abschließendes „klinisches Staging“, Goldstandard). Im abschließenden „klinischen Staging“ zeigten die Studienteilnehmer folgende Tumorstadien gemäß der Ann-Arbor-Klassifikation: 3 Patienten waren im Stadium I, 9 im Stadium II, 6 im Stadium III und 4 hatten das Stadium IV erreicht.

Die PET-Scans wurden mit zwei Ganzkörperscannern und ^{18}F FDG als radioaktiver Testsubstanz aufgenommen; zur Röntgenuntersuchung mittels CT wurde pro Patient jeweils eine Aufnahme im Brust-/Nacken- und im Abdomen-/Beckenbereich angefertigt. Insgesamt wurden je Patient 30 Lymphknotenregionen und extralymphatische Organe auf Tumorbefall durchgemustert. Anhand der auf diese Weise gewonnenen Informationen wurde das Tumorstadium - für die CT- bzw. PET-Untersuchung gesondert, also ohne Wissen des jeweils anderen Untersuchungsergebnisses (verblindete Auswertung) - ermittelt.

Auf der Basis der abschließenden Stadieneinstufung durch das „klinische Staging“ wird die diagnostische Genauigkeit der PET bzw. der CT, genauer der jeweilige Anteil an richtig eingestuftem Patienten (Effektgröße der inkrementellen Kosten-Effektivitäts-Analyse), als Maß für die Validität der beiden großtechnischen Bilderfas-

sungsverfahren ermittelt und miteinander verglichen. Während die ^{18}F FDG-PET alle Patienten treffsicher einstuft, erkannte die CT nur bei 13 der 22 Patienten das korrekte Tumorstadium. Dabei ergab sich bei 5 Patienten durch die PET-Untersuchung eine Höhereinstufung als im konventionellen Staging. PET erwies sich folglich mit einer Genauigkeit von 100% korrekt eingestufte Patienten als das diagnostisch überlegene bildgebende Verfahren (CT: 59%).

Ergänzend zur vergleichenden Beurteilung der diagnostischen Effektivität in der klinischen Studie stellt die Publikation von Klose et al. (1999a) sowohl für PET als auch für CT auf der Kostenseite eine detaillierte Berechnung des direkten Finanzbedarfs für Ankauf und Betrieb der Untersuchungsgeräte in Form einer jährlichen Aufstellung der Investitions-, Wartungs- und Betriebskosten vor. In die Kostenkalkulation für PET einbezogen sind die Mittel für den Erwerb und Unterhalt eines Zyklotrons (plus radiochemische Laborausrüstung), das zur Herstellung des ^{18}F -markierten FDG-Radiopharmakons benötigt wird. Als Kostenkomponenten der Betriebsausgaben werden Sach- und Personalkosten für die ^{18}F FDG-Herstellung und die Durchführung der PET- und CT-Scans berücksichtigt (Näheres siehe unter Kapitel C.5.2).

Insgesamt verursacht PET im Vergleich zur CT-Alternativtechnologie mit 961 Euro gegenüber 391 Euro pro Patient das 2,5fache an Ausgaben. Dabei machen die Investitions- und Betriebskosten für die PET - im Gegensatz zum hohen Materialanteil von circa 64% bei CT (29% bei PET inklusive ^{18}F FDG-Produktion) - den Hauptanteil der Geldaufwendungen aus. Wird die Effektgröße „Anzahl korrekt klassifizierter Patienten“ mit den Stagingkosten in Beziehung gesetzt, ergibt sich eine inkrementelle Kosten-Effektivitäts-Relation (iKER) von 662 Euro pro zusätzlich korrekt klassifiziertem Patienten für den Vergleich CT versus „keine diagnostische Untersuchung“ bzw. 1 393 Euro pro zusätzlich korrekt klassifiziertem Patienten für den Vergleich ^{18}F FDG-PET versus CT.

Anhand von Sensitivitätsanalysen, in denen unsichere Parameter variiert werden (Schwankungsbreite im Ressourcenverbrauch, Variation des Auslastungsgrades von Zyklotron und PET-Scanner), konnten Klose et al. zeigen, daß durch optimierten Einsatz der ^{18}F FDG-PET ein hohes Einsparpotential von 8,9 - 32,5% realisierbar ist (in Relation zum Basisszenario).

Die Autoren ziehen aus ihren Ergebnissen das Resümee, daß einerseits die ^{18}F FDG-PET gegenüber der CT das diagnostisch genauere Untersuchungsverfahren darstellt und sich andererseits durch eine Optimierung des Auslastungsgrades der für eine PET-Untersuchung notwendigen Geräte die iKER im Vergleich zur CT senken ließe.

Studienbewertung

In der vorliegenden Evaluation erfüllt die Berechnung und Darstellung der Kostenseite alle wichtigen Kriterien einer gesundheitsökonomischen Analyse. Eine Schwä-

che der Studie sind jedoch methodisch-organisatorische Mängel bei der Erfassung der Effektseite (24% bzw. 1,5 % aller Bewertungen von Lymphknotenregionen und extralymphatischen Organen fehlten für die CT- bzw. PET-Auswertung) sowie die knappe Darstellung des klinischen Studiendesigns und der klinischen Studienergebnisse.

Das Studienziel, die Bestimmung der Kosteneffektivität der PET im Vergleich zum konventionellen CT beim Primärstaging von malignen Lymphomen, ist klar formuliert. Die mit nur 22 Probanden kleine Studienpopulation ist nicht repräsentativ für die Erkrankungsverteilung in der Bevölkerung, da Männer und NHL-Patienten in der Studie unterrepräsentiert sind.

Die Evaluation von Klose et al. (1999a) bezieht sich nicht wie die frühere Studie zum Indikationsgebiet „maligne Lymphome“ von Hoh et al. (1997) auf zwei alternative, aus mehreren Diagnosemethoden zusammengesetzte Diagnosealgorithmen in ihrer Gesamtheit, sondern ausschließlich auf die wahlweise als Algorithmusbausteine einsetzbaren Einzeltechnologien PET bzw. CT. Aus der Beschränkung des Studienrahmens auf die Durchführung der tomographischen Untersuchungen ergibt sich auch eine starke Einengung des Zeithorizontes auf die dafür benötigte Zeitspanne. Aus dem gesundheitsökonomischen Blickwinkel betrachtet, ist eine derartig starke Einengung des Zeitfensters nicht befriedigend. Kurzfristige, additive Effekte und Mehrkosten durch Zusatzuntersuchungen (verifizierende Biopsie) bei unsicherem CT- bzw. PET-Befund in der Studienauswertung bleiben folglich unberücksichtigt. Auch längerfristige Folgeeffekte und Folgekosten der Befundung mittels PET bzw. CT, zum Beispiel der Einfluß der in Abhängigkeit vom Stagingergebnis gewählten Therapiestrategie auf die Behandlungseffekte bzw. Behandlungskosten, werden nicht erfaßt. So wäre zu prüfen, ob sich die inkrementelle Kosten-Effektivitäts-Ratio - 1 393 Euro für PET versus 662 Euro für CT pro korrekt klassifiziertem Patienten - zugunsten der absolut betrachtet teureren PET-Technologie verschiebt, wenn mögliche Kosteneinsparungen durch vermeidbare diagnostische und/oder therapeutische Interventionen aufgrund eines verbesserten Basisstagings mittels PET im Vergleich zur klinisch weniger effektiven CT (diagnostische Genauigkeit in der klinischen Studie: PET 100%, CT 59%) einbezogen würden.

Auf der Effektseite wird mit der Wahl der Diagnosegenauigkeit als Effektmaß eine intermediäre Outcomegröße (diagnostic accuracy efficacy, siehe C.4.2.2) von begrenzter Aussagekraft gewählt. Bei der Beurteilung der klinischen und ökonomischen Wertigkeit einer Diagnosemethode ist - unter Ausweitung des Zeithorizontes - die Erfassung finaler Effektgrößen von weit größerer Aussagekraft; beispielsweise indem die Auswirkungen der Diagnosestellung auf das Patientenmanagement im Rahmen der Therapie (therapeutic efficacy) oder auf die Lebenserwartung oder Lebensqualität der versorgten Patienten (patient outcome efficacy) in die Evaluation eingehen. Ferner ist die Operationalisierung der Bestimmung des klinischen Stadiums als problematisch zu beurteilen, da Lymphknoten bzw. extralymphatische Organe auch dann als befallen eingeordnet werden, wenn nur eines der beiden diagnostischen

Verfahren einen Befall definitiv anzeigt, während das andere diesen nicht definitiv ausschließt. Bei neuen diagnostischen Methoden, die eine hohe Sensitivität, aber eine geringe Spezifität aufweisen, führt diese Operationalisierung der Bestimmung des klinischen Stadiums dazu, daß alle anderen Methoden schlechter abschneiden. Aus diesen Gründen wurde die Operationalisierung der Messung der Effektgröße zwischenzeitlich geändert. Diese Resultate sind noch nicht veröffentlicht. Nach persönlicher Mitteilung der Autoren wird auf der Basis dieser neuen Berechnungsgrundlage die Kosten-Effektivitäts-Relation der PET-Diagnostik größer, die Schlußfolgerungen bleiben aber in der hier beschriebenen Form gültig.

Im ausführlichen ökonomischen Teil der Studie wird sowohl das Mengengerüst (z.B. Anzahl der pro Woche untersuchten Patienten, Zeitbedarf des ärztlichen und des technischen Personals) als auch das Preisgerüst - aufgeschlüsselt nach den Kostenarten Sach-, Personal- und Gemeinkosten - sehr detailliert aufgeführt. Die Datenquellen (z.B. Routineverbrauchsdaten, Expertenschätzungen, Verträge) sind explizit genannt und adäquat gewählt. Anhand einer verursachungsgemäßen Kostenzurechnung, ein Berechnungskonzept das vom tatsächlichen Aufwand der Leistungserstellung ausgeht, wird der benötigte Material- und Personaleinsatz quantifiziert (Mengengerüst) und der Materialbedarf mit Marktpreisen (z.B. Beschaffungskosten laut Buchführung) und der Personalaufwand mit klinikspezifischen Zeitkosten (Kosten pro Minute) bewertet. Die Berechnung des Finanzbedarfs ist ausführlich dargestellt und Schritt für Schritt nachvollziehbar. Der von den Autoren gewählten gesellschaftlichen Perspektive entspricht, daß die von den Bundesländern getragenen Investitionskosten in der Kostenrechnung veranschlagt werden. Die im Referenzjahr 1997 erhobenen Kosten werden in geeigneter Weise auf das Jahr 1999 inflationsbereinigt und zur besseren internationalen Vergleichbarkeit in Euro konvertiert. Wegen des sehr engen, auf die reine Diagnosephase beschränkten Zeithorizontes entfällt - wie auch die Autoren erläutern - die Notwendigkeit zur Diskontierung der Effekte bzw. Kosten.

Der Einfluß von Unsicherheiten in der Datenbasis (z.B. Schwankungen im Ressourcenverbrauch) bzw. von unsicheren Annahmen zur Geräteauslastung (Zyklotron und/oder PET- bzw. CT-Scanner) auf die Kostenentwicklung wird anhand von univariaten, multivariaten und strukturellen Sensitivitätsanalysen überprüft und in einer zusammenfassenden Tabelle dokumentiert.

Die Ergebnisse der vorliegenden Studie werden - auf der Kostenseite ausführlich - diskutiert und mit Ergebnissen aus anderen CT/PET-Evaluationsstudien verglichen. Bei den CT-Untersuchungen offensichtlich aufgetretene methodische und organisatorische Schwierigkeiten (unvollständige Befundung der Lymphknotenregionen und der extralymphatischen Organe, siehe oben) werden nur kurz erläutert. In der Diskussion regen die Autoren an, in künftige ökonomische Analysen auch längerfristige Kosten und Effekte einzubeziehen.

Ethische und Verteilungsfragen werden nicht angesprochen, sind jedoch auch nicht Teil der Aufgabenstellung. Die Ergebnisse können aufgrund der kleinen Studienpo-

pulation (n = 22) nicht verallgemeinert werden, zumal das Studienkollektiv für die Zielpopulation nicht repräsentativ ist (siehe oben).

Abschließende Bewertung

Die inkrementelle Kosten-Effektivitäts-Studie von Klose et al. (1999a) ist derzeit die einzige deutsche Evaluation zur Kosteneffektivität der PET-Technologie. Die Studie ist im Hinblick auf die Berechnung und Darstellung der Kostenseite als methodisch elaboriert zu bezeichnen und kann als gute Datenquelle für künftige ökonomische Evaluationen zum Einsatz der PET in der klinischen Praxis dienen. Ihre Bedeutung als Entscheidungshilfe bezüglich der gesundheitsökonomischen Wertigkeit der PET beim Primärstaging von malignen Lymphomen muß jedoch wegen der beschriebenen Mängel auf der Effektseite und im klinischen Studiendesign kritisch beurteilt werden (z.B. sehr kleine Studienpopulation, Auswirkungen der PET- bzw. CT-Ergebnisse auf das Patientenmanagement bleiben in der Analyse unbeachtet). Die derzeit vorliegenden Ergebnisse werden deshalb als nicht ausreichend erachtet, eine abschließende Entscheidungsunterstützung zum kosteneffektiven Einsatz der PET beim Primärstaging von Lymphomen geben zu können.

Adler LP, Faulhaber PF, Schnur KC et al. Axillary lymph node metastases: screening with [F-18]2-deoxy-2-fluoro-D-glucose (FDG) PET. Radiology 1997; 203:323–327

Studienbeschreibung

Wichtigster prognostischer Faktor, um die lymphogene Ausbreitung eines primären Mammakarzinoms beurteilen und damit die Therapiestrategie festlegen zu können, ist das Auftreten von Metastasen in den regionären Lymphknoten (Lk). Im Rahmen einer größeren, prospektiven klinischen Studie zur Bewertung der diagnostischen Effektivität der Positronen-Emissions-Tomographie (PET) bei Brustkrebsuntersuchungen ging die Arbeitsgruppe am Universitätskrankenhaus von Cleveland, USA, auch der Frage nach, ob PET als Screeninginstrument zum Nachweis von Metastasen in axillären Lymphknoten geeignet ist.

In der vorliegenden klinischen Studie wird die Verwendbarkeit der PET als Screeningverfahren anhand der diagnostischen Genauigkeit ermittelt: hierzu wurden die PET-Schnittbilder direkt am Bildschirm von zwei unabhängigen Gutachtern auf Foci mit abnormal erhöhter ¹⁸FDG-Aufnahme durchgemustert und die Sensitivität, Spezifität und der positive bzw. negative prädiktive Wert berechnet. Die Auswertung erfolgte verblindet, das heißt, die Auswerter erhielten - außer einer Angabe zur untersuchten Körperseite - keine zusätzlichen Informationen. Als Referenzmethode diente das konventionelle, invasive Diagnoseverfahren der Axilladissektion, bei der die Lymphknoten aus der Achselhöhle chirurgisch entfernt und die entnommenen Gewebeproben anschließend histopathologisch untersucht werden. Dieses für die Patientinnen belastende Verfahren ist oftmals von erheblichen gesundheitlichen Nebenwirkungen begleitet.

Eine nicht randomisierte Fallserie von 50⁷ Frauen im Alter von 36 - 79 Jahren mit operablem Primärtumor der Brustdrüse, $\geq 5\text{mm}$, bildete das Patientenkollektiv. Die Patientinnen wurden einem zweistufigen Untersuchungsalgorithmus unterzogen: zuerst wurde ein PET-Scan der Axillaregion durchgeführt (Teilkörperaufnahme 60 min nach i.v.-Gabe des ¹⁸FDG-Radiotracers, 2fache Menge der üblichen Dosis); in einem zweiten Schritt wurde der PET-Befund innerhalb von drei Monaten mittels Exstirpation und histopathologischer Untersuchung der entfernten Lymphknoten überprüft. Als Maßzahlen für die zu prüfende diagnostische Validität der PET als Screeningverfahren für Lk-Tumorbefall (bei bestehender Brustkrebserkrankung) ergaben sich im direkten Vergleich zum Goldstandard der Axilladissektion folgende biometrische

⁷ Zwei weitere Patientinnen, deren PET-Aufnahmen nicht interpretierbar waren, wurden aus der klinischen Studie ausgeschlossen. Die Kosten für die zwei durchgeführten PET-Scans gingen jedoch in die ökonomische Analyse ein.

Kennwerte: Sensitivität, 95%; Spezifität, 66%; negativer bzw. positiver Vorhersagewert, 95% bzw. 63%, Gesamtgenauigkeit 77%.

Anhand dieser diagnostischen Kenndaten schlußfolgern die Autoren:

1. Das PET-Bildgebungsverfahren weist unter Einsatz hoher ^{18}F FDG-Dosen eine ausreichende Sensitivität auf, um als nicht-invasive Screeningmethode bei Verdacht auf axilliäre Lk-Metastasen nutzbar zu sein.
2. Aufgrund des hohen negativen Vorhersagewertes bei negativem PET-Befund kann auf eine Dissektion der axillären Lymphknoten verzichtet werden, da in diesen Fällen nur ein geringes Risiko der Metastasierung von Lymphknoten anzunehmen ist.

Bei positivem PET-Ergebnis bleibt jedoch eine Nachuntersuchung mittels Dissektion zur Verifikation des nicht-invasiv gewonnenen Erstbefundes und zur präoperativen Bestimmung der Anzahl und Lage der befallenen Lymphknoten indiziert (zweistufiges Diagnoseverfahren).

Bei der Studie von Adler et al. handelt es sich in erster Linie um eine klinische Wirksamkeitsprüfung; Kostenbetrachtungen werden nur ergänzend angestellt. Die Autoren berechnen das Einsparpotential beim teilweisen Ersatz der bisher üblichen Axilladissektion, die mit teurer stationärer Unterbringung und postoperativer Nachsorge verbunden ist (US\$ 8 100 pro Eingriff), durch die nicht-invasive PET (US\$ 1 150). Im Rahmen der Studie hätte sich bei 22 PET-negativen Patientinnen des 50köpfigen Kollektivs (darunter eine falsch-negative Patientin, siehe Studienbewertung) eine Dissektion erübrigt, so daß sich das potentielle Einsparvolumen - zum Zeitpunkt der Studie - auf insgesamt US\$ 120 000 beläuft. Zudem bliebe beim Einsatz von PET als Screeninginstrument auf Lk-Metastasen bei primärem Mammakarzinom Patientinnen mit (korrekt) negativem PET-Testergebnis ein belastender und zudem komplikationsanfälliger chirurgischer Eingriff erspart.

Aufgrund dieser Studienergebnisse ziehen die Autoren den Schluß, daß PET eine kostengünstige und diagnostisch sinnvolle Ersatzmethode anstelle einer Axilladissektion darstellt.

Studienbewertung

Die klinische, nicht jedoch die ökonomische Zielsetzung der Studie ist präzise formuliert. Der medizinische Problemkontext ist deutlich dargelegt; der ökonomische Problemkontext wird jedoch nur am Rande angesprochen.

Die beiden alternativen Diagnosemethoden werden ausführlich (PET-Emissions- und Transmissionsmessungen mit ^{18}F FDG, Teilkörperaufnahmen) bzw. hinreichend (Routinemethode der Axilladissektion) vorgestellt. Die Notwendigkeit, eine nebenwirkungsarme, nicht-invasive Ersatztechnik für die chirurgische Dissektion hinsichtlich

ihrer diagnostischen Wirksamkeit zu evaluieren, wird schlüssig begründet. Die Bevorzugung der PET gegenüber anderen, ebenfalls nicht-invasiven Bildgebungsverfahren, wie Mammographie, CT, Sonographie und Immunszintigraphie, wird diskutiert und erweist sich aufgrund der besseren diagnostischen Treffsicherheit der PET-Technologie, speziell bezüglich der für Screeningverfahren wichtigen Sensitivität, als gerechtfertigt. Die Studienpopulation bzw. die Ein- und Ausschlusskriterien sind adäquat gewählt (z.B. liegt der Häufigkeitsgipfel für Mammakarzinom bei 40 - 75 Jahren), wobei das Studienergebnis eine Erweiterung der Ausschlusskriterien (aktuelle systemische Therapie, Vorliegen einer gleichseitigen Axilladissektion, Primärtumor der Mamma < 5 mm) auf Patientinnen mit Adipositas nahelegt. Diese Maßnahme würde eine wichtige Fehlerquelle für falsch-negative Befundung von Lk-Metastasen bei PET-Scans ausschalten. Die Zurechnung einer mittels PET falsch negativ diagnostizierten, adipösen Patientin bei der Berechnung des Einsparpotentials ist problematisch, solange nicht auch die materiellen (und gesundheitlichen) Folgen der Falschbefundung in die Betrachtung eingeschlossen werden.

Da die Autoren explizit die Eignung der PET-Technologie als Screeningmethode, untersuchen, und die Wirkung auf die weitere Vorgehensweise nicht Teil der Zielsetzung ist, ist die Begrenzung des Zeithorizontes allein auf die Phase der Diagnose-durchführung zulässig. Aufgrund dieser zeitlichen Vorgabe werden auf der Effektseite die klinische Effektgröße „korrekt erkannter Lymphknotenstatus“ (diagnostic accuracy efficacy, siehe C.4.2.2) und auf der Kostenseite nur direkte, ausschließlich im Rahmen der Lk-Untersuchung anfallende Diagnosekosten berücksichtigt. Langfristige Effekte und Kosten, beispielsweise Folgekosten und Änderungen im Gesundheitszustand der Patientinnen durch dissektionsbedingte Nebenwirkungen, sind nicht Gegenstand der Studie. Speziell aus gesundheitsökonomischer Sicht hätte jedoch - gerade beim Vergleich einer schonenderen, nicht invasiven Untersuchungsmethode mit einer invasiven, patientenbelastenden Methode - die Erfassung von Folgeschäden durch eine vergleichende Messung der gesundheitsbezogenen Lebensqualität die Studie aufgewertet.

Das zur Berechnung der direkten medizinischen Diagnosekosten notwendige Mengengerüst ergibt sich aus der Anzahl der durchgeführten PET-Scans bzw. Axilladissektionen. Zur monetären Bewertung des Mengengerüsts werden administrative Preise in Form von Gebühren verwendet. Die Gebühren werden nur für die ressourcenaufwendige Dissektionsmethode, nicht jedoch für PET, in einzelne Kostenkomponenten, wie Personalkosten, Kosten für Medikamente und histopathologische Untersuchung usw., aufgeschlüsselt. Die Kostenzurechnung erfolgt aus der Perspektive des Drittfinanzierers, da nur die tatsächlich erstatteten Vergütungen und Einsparungen, die nur für den Kostenträger relevant sind, in die ökonomische Bewertung einbezogen werden.

Die Kosten- und die Effektseite wird nicht durch die Bildung einer Kosten-Effektivitäts-Relation direkt zueinander in Beziehung gesetzt (eventuell als Kosten pro korrekt eingestuftem Patienten). Vielmehr geht die durch den Einsatz der PET als

Screeninginstrument für axilläre Lk-Metastasen induzierte Änderung im Diagnosemanagement (diagnostic thinking efficacy, siehe C.4.2.2), speziell die Effektgröße „eingesparte Dissektionen“, als monetär bewerteter Posten auf der Kostenseite ein (partielle Evaluation, siehe hierzu C.4.2.3).

Die Währung wird genannt, nicht jedoch das Referenzjahr der Kosten/Gebühren. Inflations-/Deflationsanpassungen und Diskontierungen werden nicht vorgenommen; diese sind jedoch, sofern aktuelle Gebühren zur Kostenberechnung herangezogen wurden, wegen des engen Zeithorizontes auch nicht von Relevanz.

Die Qualität der klinischen Daten und wichtige methodische Schwierigkeiten und Einschränkungen des PET-Verfahrens, wie die Gefahr falsch positiver Befunde aufgrund unspezifischer ^{18}F FDG-Aufnahme und die Abhängigkeit der Sensitivität vom Tumorstadium, werden von den Autoren im Diskussionsteil kritisch angesprochen. Die Abhängigkeit der Sensitivität von der Anzahl an tumorbefallenen Lymphknoten und von der Größe der Metastasen, sowie die verminderte Aufnahmequalität bei adipösen Patientinnen werden jedoch nicht näher diskutiert. Auch die Inter-Rater-Reliabilität als Maß für die Übereinstimmung der PET-Befundung durch die zwei unabhängigen Gutachter ist leider nicht angegeben. Unsicherheiten der ökonomischen Daten werden weder diskutiert, noch werden Sensitivitätsanalysen angestellt.

Die abschließende Stellungnahme der Autoren zur Einsetzbarkeit der PET-Technologie als Screeninginstrument zur Detektion von axillären Lymphknotenmetastasen ist in sich schlüssig, die ökonomische Bewertung des PET-Einsatzes jedoch noch ergänzungsbedürftig. Hätten die Verfasser der Studie zusätzlich folgende Punkte erfaßt, hätte sich ein genaueres Bild der Konkurrenzfähigkeit der PET gegenüber der mit erheblichen Nebenwirkungen verbundenen Axilladisektion ergeben:

1. Die materiellen und medizinischen Folgen von unerkanntem Lymphknotenbefall bei falsch-negativem PET-Befund (bei einer adipösen Patientin),
2. mögliche Auswirkungen auf die Therapiestrategie, das heißt, auf die Behandlung des Mammakarzinoms je nach Metastasierungsgrad der regionären Lymphknoten, und
3. Langzeitkosten und dissektionsbedingte Folgeschäden, wie Beeinträchtigungen in der Lebensqualität,

Abschließende Bewertung

Insgesamt handelt es sich bei der vorliegenden Studie vorrangig um eine gute klinische Studie zur Beurteilung der medizinischen Effektivität der ^{18}F FDG-PET als Screeningverfahren zur Erfassung von axillärem Lymphknotenbefall bei Brustkrebspatientinnen; Kostenbetrachtungen werden nur ergänzend angestellt. Da Kosten und Effekte nicht direkt zueinander in Beziehung gesetzt, sondern die durch vermiedene Dissektionen erzielten Einsparungen berechnet werden (Nettovergleich der Unter-

suchungskosten auf aggregierter Ebene des Patientenkollektivs), handelt es sich bei dieser Studie aus gesundheitsökonomischer Sicht um keine vollständige, sondern um eine partielle ökonomische Evaluation (siehe hierzu C.4.2.3). Da die Kostenerhebung einige methodische Mängel aufweist (u.a. ein zu enger Zeithorizont, nur Erfassung von Diagnosekosten, fehlende Sensitivitätsanalysen), ist den Ergebnissen nur eine begrenzte Aussagekraft zuzumessen.

Kritisch zu hinterfragen ist auch die Aktualität dieser Studie, da die kostspielige und von einer relativ hohen Morbidität begleitete Axilladisektion möglicherweise zukünftig durch schonendere diagnostische Verfahren, wie der noch in der Erprobungsphase befindlichen Lymphabflußszintigraphie (Sentinel-Methode), abgelöst werden könnte.

Valk PE, Pounds TR, Tesar RD et al. Cost-effectiveness of PET imaging in clinical oncology. Nucl Med Biol 1996;23(6):737-743

Valk PE, Abella-Columna E, Haseman MK et al. Whole-body PET imaging with [18F]fluorodeoxyglucose in management of recurrent colorectal cancer. Arch Surg 1999;134(5):503-513

Beide - inhaltlich zusammenhängenden - Studien wurden in einer ambulanten, radiologischen Privatpraxis mit Sitz am „Northern California PET Imaging Center“ in Sacramento unter Mitwirkung der ortsansässigen Klinik durchgeführt.

Beschreibung der Übersichtsstudie von Valk et al. 1996

Ziel des erstveröffentlichten Artikels ist die Evaluation der diagnostischen Genauigkeit von PET für fünf verschiedene Indikationsstellungen: Diagnose von metastasierenden malignen Melanomen, Feststellung von rezidivierendem kolorektalem Karzinom und Staging von nicht-kleinzelligem Bronchialkarzinom (NSCLC), fortgeschrittenem Kopf-Hals-Krebs und Lymphogranulomatose (Hodgkin-Lymphom). Mit den prospektiven medizinischen Studien in den genannten Indikationsgebieten verbunden ist jeweils eine retrospektive Beurteilung des Management- und Finanzaufwandes beim Einsatz der PET in Relation zur konventionellen CT-Technik. Der Studienzeitraum ist nicht näher datiert.

Für den klinischen Teil der Studie wurden am Northern California PET Imaging Center in Sacramento, USA, Fall-Serien von 36 (maligne Melanome), 57 (Verdacht auf Kolorektalrezidive), 99 (NSCLC), 29 (Kopf-Hals-Krebs) bzw. 30 Patienten (Hodgkin-Lymphom) mit den bildgebenden Diagnoseverfahren PET und CT vergleichend untersucht. PET-Ganzkörper-Messungen, mit ¹⁸FDG als radioaktivem Tracer, wurden bei den Studienteilnehmern mit NSCLC, Darmkarzinom und Hodgkin-Lymphom ausgeführt. In allen übrigen Fällen sind der Veröffentlichung keine näheren Angaben zur PET-Methodik zu entnehmen. Als verifizierende Referenzdiagnostik diente bei den jeweiligen Erkrankungen: CT und histologische Diagnose bei solitären Lungenknoten und NSCLC, CT und Enddiagnose (mittels Histologie oder wiederholter CT-Untersuchung im zeitlich nicht definierten Follow-up) bei kolorektalem Karzinom und metastasierendem Melanom, CT-basiertes konventionelles Staging beim Hodgkin-Lymphom bzw. Biopsie und CT/Gallium-67-Tumorszintigraphie bei Krebserkrankungen der Kopf-Hals-Region. Um den Behandlungsaufwand - als Grundlage für eine ökonomische Bewertung - ermessen zu können, wurden retrospektiv die Behandlungsunterlagen von Patienten mit solitären Lungenknoten oder NSCLC (n = 72), mit vermutetem oder bekanntem, rezidivierendem Kolorektalkarzinom (n = 68) bzw. metastasierendem Melanom (n = 45) und mit Tumoren im Kopf-Hals-Bereich (n = 22) ausgewertet.

Als Maß für die diagnostische Wertigkeit der Vergleichsalternativen dienen die klinischen Outcomegrößen „diagnostische Genauigkeit“ und „Einfluß auf Behandlungsstrategie“. Gestützt auf die pathologische Überprüfung der PET-Befunde erwies sich in den klinischen Studien die funktionelle Bildgebung mittels PET gegenüber der anatomischen Bildgebung mit CT in den Indikationsgebieten NSCLC, wiederkehrendes kolorektales Karzinom bzw. Kopf-Hals-Krebs und metastasierende Melanome als das diagnostisch überlegene Verfahren mit höherem Trennvermögen für maligne und benigne Läsionen. Aufgrund dieser größeren diagnostischen Effektivität für PET resultiert eine zielgenauere präoperative Stadieneinstufung und somit ein „adäquates“ chirurgisches Management der zur Resektion anstehenden Fälle. Eine „adäquate“ Behandlung liegt dann vor, wenn durch das diagnostisch bessere Untersuchungsverfahren nutzlose und unnötig belastende, chirurgische Eingriffe bei inoperablen Patienten mit nicht-resektierbarem Tumor bzw. bei Patienten mit tumorverdächtigen, aber nachweisbar gutartigen Läsionen verhindert wird. Die Anzahl der zusätzlich adäquat versorgten Fälle, Ausdruck des Einflusses der Diagnosegenauigkeit auf die Versorgungsstrategie, geht als Parameter in die ökonomische Analyse ein.

Diese Effektgröße dient als Grundlage für die Bewertung der ökonomischen Wertigkeit der PET gegenüber der CT, definiert als „eingesparte direkte Behandlungskosten“ infolge vermiedener Operationen. Durch monetäre Bewertung der vermeidbaren operativen Eingriffe anhand der Medicare-Vergütung für stationäre Versorgung und Leistungen der Chirurgie (DRG-Sätze⁸), Anästhesie und Pathologie ergeben sich auf der Ebene des Patientenkollektivs folgende Einsparpotentiale aufgrund unterlassener Interventionen:

1. bei Einsatz der PET zusätzlich zur CT: US\$ 150 000 (NSCLC, n = 72 Patienten), US\$ 178 000 (Kolorektalrezidive, n = 68 Patienten), US\$ 98 000 (metastasierende Melanome, n = 45 Patienten) bzw. US\$ 44 000 (Kopf-Hals-Krebs, n = 22 Patienten),
2. bei Ersatz der CT durch die PET: US\$ 194 000 (NSCLC, n = 72 Patienten), US\$ 232 000 (Kolorektalrezidive, n = 68 Patienten) bzw. US\$ 134 000 (metastasierende Melanome, n = 45 Patienten).

Bei Quotientenbildung - Einsparungen/zusätzliche Kosten für PET - resultieren:

1. Beim Einsatz von PET additiv zu CT „savings/cost ratios“ von 2,2 (NSCLC), 2,5 (Kolorektalrezidive), 2,2 (metastasierende Melanome) und 2,1 (Kopf-Hals-Krebs),
2. bzw. beim Ersatz von CT durch PET „net savings/cost ratios“ von 3,3, 4,4 und 4,0 (bei gleicher Reihenfolge der Indikationen; keine Angabe für Kopf-Hals-Krebs).

Beim Staging von Hodgkin-Lymphomen ergibt sich kein direkter Kostenvorteil; denn Abweichungen in der Stadieneinstufung ziehen in der Studie keine gravierenden Kostenänderungen nach sich, da (1) keine großen und kostspieligen chirurgischen Ein-

⁸ Diagnosis related group, US-amerikanische Diagnoseklassifikation, die als Grundlage für die Vergütung nach Fallpauschalen im stationären Bereich verwendet wird.

griffe ausgeführt wurden und (2) vergleichbare direkte Kosten für die alternativen Bildgebungsverfahren PET versus CT/Gallium-67-Szintigraphie entstanden.

Da gemäß diesen Berechnungen auf der aggregierten Ebene der Studienpopulationen dem Vorteil einer optimierten Patientenversorgung ein negativer Finanzbedarf, das heißt Nettoeinsparungen aufgrund unterlassener Operationen, gegenübersteht, kommen die Autoren in ihrer Publikation zur abschließenden Einschätzung, daß für die untersuchten onkologischen Indikationsstellungen - außer beim Hodgkin-Lymphom - PET das dominante und folglich kosteneffektive Bildgebungsverfahren im Vergleich zur üblichen CT ist. Die Autoren regen zudem die Durchführung prospektiver Kosteneffektivitätsstudien an. Folgerichtig veröffentlichte die Fachgruppe zum Fragenkomplex Patientenmanagement und Kosteneffektivität der PET- bzw. CT-Diagnostik bei Darmkrebspatienten 1999 eine prospektive Vergleichsstudie PET versus CT. Diese Studie wird im folgenden vorgestellt.

Beschreibung der 1999 publizierten Studie von Valk et al. (Staging von rezidivierendem kolorektalem Karzinom)

Die nunmehr prospektive Nachfolgestudie beschränkt sich auf ein Indikationsgebiet, dem Staging von rezidivierendem kolorektalem Karzinom. Mit Hilfe dieser Nachfolgestudie soll die auf frühere Untersuchungsergebnisse beruhende Hypothese der Autoren, daß das auf physiologisch-chemische Stoffwechselfvorgänge basierende metabolische Bildgebungsverfahren PET, mit ^{18}F FDG als Radiotracer, dem üblichen anatomischen Bildgebungsverfahren CT bei der Entdeckung von rezidivierendem kolorektalem Karzinom überlegen sei, überprüft werden. Auf der Effektseite wird zuerst in einer klinischen Studie die diagnostische Genauigkeit der PET in direktem Vergleich mit CT ermittelt, und davon abhängig der Einfluß abweichender PET- bzw. CT-Befunde auf die Therapieentscheidung untersucht; auf der Kostenseite werden in einer begleitenden ökonomischen Analyse mögliche Einsparungen im Ressourcenverbrauch durch verbessertes Operationsmanagement berechnet.

Für eine Gruppe von 115 Patienten, Teil eines größeren Studienkollektivs von 134 Personen (66 Männern und 68 Frauen mit einem Durchschnittsalter von 67 Jahren, Altersspanne: 31 - 93 Jahre), mit Verdacht oder gesicherter Diagnose auf rezidivierendes Darmkarzinom lagen sowohl CT- als auch PET-Scans vor (Zeitraum der Datenerhebung: 2. Hälfte 1992 - 1. Hälfte 1996). Während die CT-Aufnahmen vom Abdomen-/Hüftbereich vorab ambulant an radiologischen Abteilungen mehrerer Kliniken aus der Region Sacramento - mit unterschiedlichem CT-Equipment - erstellt worden waren, wurden anschließend im zeitlichen Abstand von 0 bis 56 Tagen ($\bar{\varnothing}$ 22 Tage) die ^{18}F FDG-PET-Aufnahmen ausschließlich am Northern California PET Imaging Center mit einem 24 Ring-Ganzkörper-System (mit Transmissionsmessung zur Korrektur der Photonenabschwächung) in Kenntnis der CT-Ergebnisse angefertigt. Eine unabhängige CT/PET-Untersuchungsabfolge war somit nicht gegeben. Um zumindest eine einheitliche Interpretation der CT-Scans zu gewährleisten, wurden die CT-

Aufnahmen von zwei erfahrenen Radiologen ohne ergänzende klinische, histologische oder PET-Daten ausgewertet. Da während der Studienphase die PET-Schichtaufnahmen in Kenntnis der verfügbaren CT-Ergebnisse und klinischen Daten interpretiert worden waren, erfolgte eine erneute, diesmal verblindete Auswertung durch 2 - 3 unabhängige Gutachter. Die Ergebnisse wurden anhand von histologischen Befunden, Nachweis der Progression mit seriellen CT-Aufnahmen und klinischem Follow-up verifiziert.

Als Maß für die diagnostische Staginggenauigkeit von PET bzw. CT ergaben sich in der klinischen Studie Sensitivitäts- und Spezifitätswerte von 93% und 98% für PET gegenüber 69% und 96% für CT, wobei für beide Techniken die Sensitivität je nach anatomischer Lage des Rezidivs variierte. Bei der chirurgischen Resektion wurden in 7 von 42 Patienten (17%), in denen PET nur lokale Rezidive erkannte, zusätzliche nicht-resektable Tumorkläsionen gefunden. Durch Veränderung des Therapieregimes aufgrund der PET-Befunde hätten für 25 Patienten mit nicht-resektablen Kläsionen die Kosten für unnötige Operationen (insgesamt US\$ 374 596) vermieden werden können: so erkannte PET aufgrund ihrer höheren Treffsicherheit 12 von 18 Patienten (67%) mit negativem CT-Befund, darunter auch 2 Patienten mit nicht-resektablem Tumor; zudem konnten mit Hilfe der PET in 23 von 78 präoperativen Untersuchungen (29%) im Gegensatz zur konventionellen CT, die nur singuläre Kläsionen ausmachte, zusätzliche Tumorstellen entdeckt werden. Nach Abzug der Kosten für die PET-Diagnostik (US\$ 1 800 pro Scan, Gesamtaufwendungen: US\$ 140 400) beläuft sich das kollektive Einsparpotential auf einen Nettobetrag von US\$ 234 196. Pro Patient ergäbe sich somit ein Betrag von rund US\$ 3 000 (Durchschnittsbetrachtung über alle untersuchten Patienten).

In ihrer Studiendiskussion kommen die Autoren abschließend zu dem Resümee, daß PET wegen der höheren Sensitivität und Spezifität gegenüber der konventionellen Untersuchungsmethode CT eine bessere diagnostische Treffsicherheit für die Detektion von kolorektalen Rezidiven hat. Gleichzeitig würden durch eine verbesserte präoperative Detektion von nicht-resektablen Tumorstellen unnötige chirurgische Eingriffe abgesetzt und somit - bezogen auf das gesamte Patientenkollektiv - die Behandlungskosten gesenkt.

Bewertung der Studien

Während die Publikation von 1996 Untersuchungen zur Kosteneffektivität der PET in wichtigen onkologischen Einsatzgebieten zusammenfaßt, beschränkt sich die Nachfolgepublikation von 1999 auf die Einzelindikation „rezidivierendes kolorektales Karzinom“.

Die Zielsetzung der Studien - direkter Vergleich der bildgebenden Diagnosetechnologien, PET und konventionelles CT, und Erfassung möglicher Einsparpotentiale durch verbessertes Fallmanagement - ist von den Autoren klar vorgegeben. Aus die-

ser Vorgabe ergibt sich unmittelbar die Wahl der beiden Alternativtechnologien, PET versus CT; bei den vorliegenden Studien handelt es sich somit um reine Technologieevaluationen. Die Rahmenbedingungen - Auswahl der Studienteilnehmer und methodische Durchführung der CT- bzw. PET-Diagnostik im Rahmen der klinischen Studie - sind in der Übersichtsveröffentlichung (Valk et al. 1996) nur sehr knapp dargestellt. Außer den kleinen Fallzahlen in den Indikationsgebieten Kopf-Hals-Krebs und Hodgkin-Lymphom ist zu bemängeln, daß nur die Patientengruppen mit diesen zwei Indikationsstellungen im prospektiven klinischen und im retrospektiven ökonomischen Teil der Studie identisch sind; bei den übrigen Indikationsgebieten dagegen dienen unterschiedliche Studienpopulationen als Datenquellen für die medizinische bzw. die ökonomische Evaluation. Bedauerlicherweise versäumen es die Autoren auch, die Vergleichbarkeit der unterschiedlichen Patientensamples und die Repräsentativität der jeweiligen Studienpopulationen zu erläutern. In der Nachfolgeveröffentlichung aus dem Jahr 1999, bei der es sich in erster Linie um eine klinische Studie zur Bewertung der diagnostischen Genauigkeit der PET im Vergleich zur CT bei der Detektion von kolorektalen Rezidiven handelt, wird dagegen der Evaluationsrahmen, wie Studiendesign und methodische Durchführung der Kerntechnologien PET bzw. CT (u.a. Beschreibung der visuellen Bildinterpretation), sehr detailliert beschrieben. Das Studienkollektiv der zugrundeliegenden klinischen Studie, nicht jedoch das verkleinerte Kollektiv der ökonomischen Studie wird vorgestellt.

In beiden Studien werden sowohl medizinische Effekte als auch Kosten betrachtet. Auf der Effektseite werden zwei - miteinander verknüpfte - Effektgrößen berichtet:

1. Biometrische Kenngrößen wie Sensitivität, Spezifität oder Übereinstimmung der PET- bzw. CT-Befunde (z.B. Vergleich der falsch-positiven bzw. falsch-negativen Ergebnisse) zur Beurteilung der diagnostischen Effektivität der alternativen Diagnosemethoden (intermediäre Effektgröße: diagnostic accuracy efficacy), und
2. Effekte der Diagnostik auf die Therapieentscheidung in Form von Änderungen im Operationsmanagement (z.B. Verzicht auf kurative chirurgische Eingriffe bei inoperablen Patienten, finale Effektgröße: therapeutic efficacy).

Dabei besitzt die von der diagnostischen Genauigkeit der Diagnoseverfahren, also von der ersten Effektgröße, abhängige zweite Effektgröße gemäß dem hierarchischen Modell der Effizienz von Effektgrößen (siehe dazu C.4.2.2) einen übergeordneten Stellenwert. Indem in beiden Studien Effekte der Diagnostik auf die Therapie aufgezeigt werden, wird eine finale Effektgröße (siehe C.4.2.2) in die Evaluation mit einbezogen, was aus gesundheitsökonomischer Sicht begrüßenswert ist. Dadurch wird der von den Autoren nicht explizit definierte Zeithorizont über den - vom gesundheitsökonomischen Standpunkt - sehr eng begrenzten Zeitraum der diagnostischen Untersuchung hinaus ausgedehnt. Die finale Effektgröße „vermiedene Operationen“ wird mit Geldeinheiten bewertet und geht in die Berechnung der Nettoeinsparungen (aggregiert auf der Ebene des Studienkollektivs bzw. disaggregiert pro untersuchtem Patienten) bzw. in die Berechnung der Verhältniszahl Nettoeinsparungen/PET-Mehrkosten (Valk et al.1996) ein. Kosten für sonstige Behandlungsmaß-

nahmen bleiben von der Analyse ausgeschlossen, so daß Kostenverschiebungen bei Unterlassung von unwirksamen Operationen bei inoperablen Patienten von den Operationskosten zu anderen Kostenposten wie Kosten für Chemo- oder Strahlentherapie nicht erfaßt werden.

Die monetäre Bewertung der eingesparten Operationen in Form von landesweit gültigen, administrativen Medicare-Vergütungen für operative Leistungskomplexe (Valk et al. 1999) ist in Ermangelung von Marktpreisen angemessen. Da Mitte der 1990er Jahre die PET-Leistung nicht im Medicare-Vergütungskatalog verzeichnet ist, verwenden die Autoren ersatzweise ein praxisintern über die Studienperiode ermitteltes Durchschnittsentgelt von US\$ 1 800. Leider sind nur in der späteren Veröffentlichung die veranschlagten Leistungsentgelte beziffert (quantitative Angabe des Preisgerüsts). In beiden Publikationen wird weder das Bezugsjahr für die Kostenerfassung angegeben, noch wird eine Inflations-/Deflationsbereinigung angestellt. Wegen des auf die Akutversorgung eingeschränkten Zeitfensters ist eine Diskontierung nicht notwendig.

Sensitivitätsanalysen werden in beiden Studien nicht durchgeführt. Da es sich bei der früheren Publikation (1996) um eine Übersichtsarbeit über fünf unterschiedliche Indikationen und bei der späteren Studie (1999) in erster Linie um eine klinische Studie, der eine kurze ökonomische Betrachtung angegliedert ist, handelt, kommen die ökonomischen Aspekte sowohl bei der Darstellung als auch bei der Diskussion der Ergebnisse zu kurz. Die Datenqualität der Studien wird kaum (Valk et al. 1999) oder gar nicht (Valk et al. 1996) diskutiert, die Schwächen der Studien werden nur bei Valk et al. 1996 angesprochen.

Abschließende Bewertung

In den vorliegenden Publikationen von Valk et al. werden retrospektiv (1996) bzw. prospektiv (1999) begleitende Kostenbetrachtungen zu klinischen Studien in den Indikationen nicht-kleinzelliges Lungenkarzinom (1996), rezidivierendes kolorektales Karzinom (1996 und 1999) bzw. Kopf/Hals-Krebs und metastasierendes Melanom (beide 1996) angestellt. Aufgrund der sehr knappen Darstellung der ökonomischen Analysen und wegen methodischer Mängel, die sich teilweise darauf zurückführen lassen, daß es sich bei den Studien in erster Linie um Studien mit klinischer und weniger mit ökonomischer Zielsetzung handelt, sind ihre Ergebnisse für die Beurteilung der ökonomischen Wertigkeit der PET in den angesprochenen Indikationen nur von eingeschränkter Aussagekraft. Dennoch kann die neuere Studie (Valk et al. 1999) als gute Ausgangsbasis für weitere ökonomische Analysen zur Kosteneffektivität der PET bei der Diagnostik von rezidivierendem kolorektalem Karzinom dienen.

Institute for Clinical PET (ICP), Fairfax, USA: Abstracts zum Meeting 1994

Das Institute for Clinical PET (ICP) präsentierte auf ihrem Fachmeeting 1994 vier retrospektive ökonomische Studien, eine Studie im Indikationsgebiet Kardiologie und drei Studien im Indikationsgebiet Onkologie. Die onkologischen Studien werden im folgenden gemeinsam behandelt und bewertet, da sie formal, unter anderem bezüglich der Kostenkalkulation, einheitlich aufgebaut sind. Der Bericht zur kardiologischen Studie ist unter Kapitel C.5.4.2 eingeordnet.

Adler LP, Cascade E, Crowe J et al. Positron emission tomography. Clinical application and economic implications of PET in the assessment of axillary lymph node involvement in breast cancer: A retrospective study. Institute for Clinical PET, 1994, Publication No. CE-BREAST-94.001

Larson SM, Cohen AM, Cascade E et al. Positron emission tomography. Clinical application and economic implications of PET in the assessment of colorectal cancer recurrence: A retrospective study. Institute for Clinical PET, 1994, Publication No. CE-COLOR-94.001

Coleman RE, Cascade E, Gupta NC et al. Positron emission tomography. Clinical application and economic implications of PET in the assessment of solitary pulmonary nodules: A retrospective study. Institute for Clinical PET, 1994, Publication No. CE-SPN-94.001

Deutsche Übersetzung: Coleman RE, Cascade E, Gupta NC et al. Untersuchungen von solitären Lungenknoten mit PET. Electromedica 1995;63(2):48 - 52

Die drei am ICP ausgearbeiteten Studien zum Indikationsgebiet Onkologie beleuchten die Rolle der PET als nicht-invasives Screeninginstrument auf axilläre Lymphknotenmetastasen bei Brustkrebspatientinnen, als Stagingverfahren bei rezidivierendem kolorektalem Karzinom und als Differenzierungsmethode zwischen benignen und malignen solitären Lungenknoten (SPNs) bei inzidenten Lungenkrebspatienten.

Zur Vorbereitung der onkologischen ICP-Studien wurde die diagnostische Effektivität der PET bei Befundung dieser Krebspatientengruppen vorab anhand retrospektiver Auswertungen bislang unveröffentlichter Daten mehrerer US-PET-Zentren empirisch ermittelt und den Ergebnissen aus Literaturrecherchen gegenübergestellt. Diese Informationen wurden von ICP-Fachgruppen zur Entwicklung neuer PET-gestützter Diagnosealgorithmen genutzt, in denen entweder (a) einzelne oder mehrere Bausteine konventioneller Untersuchungsstrategien, gebräuchliche nicht-invasive (CT) und/oder invasive Diagnosetechniken, teilweise oder vollständig durch die PET-Bilderfassung ersetzt sind (Coleman et al. 1994, Larson et al. 1994), oder (b) das

PET-Bildgebungsverfahren zusätzlich in den konventionellen Untersuchungsablauf eingeschoben ist (Brustkrebsstudie von Adler et al. 1994).

Diese neukonzipierten PET-gestützten Algorithmen werden in den ökonomischen Analysen mit den jeweiligen konventionellen Verfahren verglichen. Dazu werden die Auswirkungen der PET-Befundung auf das Patientenmanagement und auf die direkten medizinischen Gesamtkosten untersucht. Eine Verbesserung der Patientenbetreuung durch Optimierung diagnostischer und therapeutischer Maßnahmen könnte sich, eventuell durch Absetzung unnötiger diagnostischer Untersuchungen oder chirurgischer Eingriffe, in erheblichen Kosteneinsparungen niederschlagen. Um derartige Einsparmöglichkeiten aufzuzeigen, stellten die einzelnen Fachgruppen auf die nationale Ebene projizierte ökonomische Modellrechnungen für zwei unterschiedliche, in den USA weitverbreitete Abrechnungssysteme, die Medicare-Vergütung und die private Kostenabrechnung eines nicht näher spezifizierten tertiären Gesundheitszentrums, auf.

Im folgenden werden die drei Studien der fachspezifischen ICP-Arbeitsgruppen im Indikationsgebiet Onkologie, ihre Konzeption und Ergebnisse, kurz dargestellt. Die methodische Qualität der onkologischen Studien wird anschließend gemeinsam bewertet. Die kardiologische ICP-Studie wird unter Kapitel C.5.4.2 behandelt.

Modellierung von Adler et al. 1994 (Screening auf axilläre Lymphknotenmetastasen)

Die Zielpopulation der ersten onkologischen Studie bilden die - bei einer geschätzten Inzidenz von etwa 183 000 Brustkrebsfällen im Bezugsjahr 1993 - in den USA neu entdeckten 100 000 Fälle mit Mammakarzinom der Stadien I und II, bei denen eine brusterhaltende Therapie in Frage kam. Da eine Metastasierung der regionären Lymphknoten, speziell der Achsellymphknoten, durch lymphogene Ausbreitung des Primärtumors der wichtigste Faktor zur Vorhersage der Überlebenswahrscheinlichkeit bei einer Brustkrebserkrankung darstellt, ist eine zuverlässige Methode zur Abklärung des Lymphknotenstatus unerlässlich. Nach Aussage der Autoren war die chirurgische Entfernung der axillären Lymphknoten (Lymphknotendisektion der Axilla, ALND), die eine histopathologische Beurteilung der verdächtigen regionären Lymphknoten erlaubt, zum Zeitpunkt der Studie die Methode der Wahl, da konventionelle, nicht-invasive Screeningtechniken, wie röntgendiagnostische Untersuchung des Thorax (u.a. CT), oder immunologische Tests keine befriedigende diagnostische Genauigkeit aufwiesen. Da mit der ALND neben hohen Kosten auch ein deutliches Morbiditätsrisiko infolge postoperativer Komplikationen in Form seröser Flüssigkeitsansammlungen, Nervenverletzungen und Schulterdysfunktion verbunden ist, wäre bei Brustkrebspatientinnen mit geringem Risiko der Metastasierung axillärer Lymphknoten (3%ige Inzidenz) der Ersatz dieser belastenden, chirurgischen Prozedur durch eine nicht-invasive Screeningtechnik hilfreich. So ist in der ICP-Modellierung bei der alternativen Diagnosestrategie ein PET-Scan der partiellen Mastektomie (zur Entfernung des Primärtumors der Mamma) vorgeschaltet, wobei nur

noch im Anschluß an einen PET-positiven Befund auf axillären Lymphknotenbefall eine ALND vorgesehen ist, während bei einem negativen PET-Befund von einer nebenwirkungsreichen ALND abgesehen wird.

In einer tabellarisch geordneten Literaturübersicht zur diagnostischen Genauigkeit der ^{18}F FDG-PET als Instrument zur Detektion von axillären Lymphknotenmetastasen bei Brustkrebspatientinnen ist eine Sensitivitätsspanne von 67 - 100% und eine Spezifität von 100% angegeben; beim Survey an US-amerikanischen PET-Zentren ergaben sich Sensitivitäts- und Spezifitätswerte von jeweils 96%. Die Wertespanne für die ^{18}F FDG-PET-Sensitivität ist laut Autoren auf die unterschiedlichen Bilderfassungstechniken in den einzelnen Studien zurückzuführen (abschwächungskorrigierte statische oder (un)korrigierte Ganzkörper-Scans).

In die auf das US-Bevölkerungsniveau zugeschnittene ökonomische Modellierung gehen folgende Annahmen ein:

1. Geschätzte Sensitivitäts- bzw. Spezifitätswerte für die PET-Bilderfassung von 90% bzw. 95%,
2. eine Zielpopulation von 100 000 Patientinnen mit Brustkrebs im frühen Stadium und der Option auf eine brusterhaltende Therapie durch partielle Mastektomie,
3. eine Prävalenz der Metastasenbildung in den Achsellymphknoten von 25%, und
4. bei der konventionellen Strategie erhalten 30 000 Frauen, also 30%, eine ALND in Kombination mit einer brusterhaltenden, partiellen Mastektomie, bei den übrigen 70% (also 70 000 Frauen) werden die ALND und die partielle Mastektomie separat durchgeführt (prozentuale Aufteilung erfolgte gemäß Expertenmeinung).

Unter diesen Modellvoraussetzungen fallen - bezogen auf das Jahr 1993 - unter der PET-Strategie bei einer jährlichen Inzidenz von landesweit 100 000 Fällen 73 750 Untersuchungen negativ aus. Da im PET-basierten Testverfahren PET-negative Befunde nicht weiter mittels ALND abgeklärt würden, könnten 51 625 separate ALND und weitere 22 125 ALND, die im konventionellen Behandlungsprogramm in Kombination mit einer partiellen Mastektomie ausgeführt werden, vermieden werden. Den auf nationaler Ebene unterlassenen 73 750 ALND-Eingriffen entspricht - rückblickend auf den Studienzeitraum und je nach Abrechnungssystem - ein einzusparender Gesamtbetrag von US\$ 48,5 Millionen (Medicare⁹-Vergütung) bzw. US\$ 161 Millionen (bei privater Kostenabrechnung); pro Zielperson errechnet sich somit ein durchschnittliches Einsparpotential von US\$ 485 (bei Medicare-Vergütung) bzw. etwa US\$ 1 610 (bei privater Kostenabrechnung). Ein großer Teil dieser Finanzbeträge resultiert aus der Vermeidung von stationärer Versorgung, die zwar bei einer ALND, nicht jedoch bei einer partiellen Mastektomie anfällt. Zudem blieben bei der alternativen PET-Strategie die meisten Frauen mit Brustkrebs im Anfangsstadium - rund 74 000 von 100 000 Patientinnen im Jahr 1993 allein in den USA - durch Einsatz der nicht-

⁹ Staatliches Krankenversicherungsprogramm der USA für über 65jährige, Dialysepflichtige und Behinderte.

invasiven Diagnosemethode von dissektionsbedingter Kurz- und Langzeitmorbidity verschont.

Anhand der Ergebnisse dieser Modellierung sieht die ICP-Fachgruppe es als gegeben an, daß durch den Einsatz der PET als Screeninginstrument auf axilläre Lymphknotenmetastasen eine deutliche Verbesserung im Patientenmanagement durch Vermeidung unnötiger Interventionen bei Patienten mit geringem Risiko für Metastasenbildung in den axillären Lymphknoten zu erzielen ist. Auch das Risiko, eine Patientin mit Lymphknotenbefall irrtümlich zu übersehen (falsch-negativer PET-Befund; falsch-negative Fehlklassifikationsrate: 3%), werde durch die Möglichkeit gerechtfertigt, dafür 33 Patientinnen die Belastungen einer ALND zu ersparen.

Modellierung von S. M. Larson et al. 1994 (Staging von rezidivierendem kolorektalem Karzinom)

In den USA wurden 1992 rund 155 000 Fälle von kolorektalem Karzinom neu erkannt, von denen rund 109 200 Fälle (70%) kurativ mittels Resektion der betroffenen Darmabschnitte behandelt wurden. Die resektierten Krebspatienten wurden im anschließenden Follow-up unter anderem mit CEA-Untersuchungen (colon embryonal antibody), Röntgenuntersuchung im Brustbereich, Kolonoskopie und Leberfunktionstests und - bei Bedarf - mit CT-Scans überwacht. Basierend auf Monitoring-Daten wurden mit Hilfe dieser Überwachungsmethoden bei 44 750 Personen oder circa 41% der 109 200 resektierten Darmkrebspatienten Tumorrezidive erfaßt. Diese Patientengruppe mit rezidivierendem Darmkrebs bildet die Zielgruppe der zweiten Modellierung des ICP.

Um bei einem Wiederauftreten des Tumors entscheiden zu können, für welche Patienten günstige Heilungschancen durch eine erneute Resektion bestehen, müssen anhand einer präoperativen Diagnostik isolierte Krankheitsherde von weit streuenden Metastasen differenziert werden. Dies geschieht konventionell mittels operativer Eröffnung der Bauchhöhle (explorative Laparotomie). Als Ersatzmethode schlägt die ICP-Arbeitsgruppe 1994 auf ihrem Fachmeeting die - in der Publikation methodisch nicht näher spezifizierte - ^{18}F FDG-PET-Bilderfassung vor.

Im Rahmen eines retrospektiven ICP-Surveys an 14 PET-Zentren in den USA wurden die Krankenakten von 267 Patienten mit kolorektalem Karzinom ausgewertet. Anhand von ^{18}F FDG-PET-Scans wurden 54 Läsionen lokalisiert, die operativ bestätigt wurden. Aufgrund dieser Datenbasis ergab sich für die Differenzierung zwischen malignen und benignen kolorektalen Tumorrezidiven laut Autoren eine Sensitivität von 93%, eine Spezifität von 78% und eine Genauigkeit von 88%. Zwei - auf kleine Fallzahlen von 18 bzw. 29 Patienten beruhende - Literaturstellen berichten Sensitivitätswerte von 100% bzw. 95% und Spezifitätswerte von 86% bzw. 100%. In die ökonomische Modellrechnung setzen die Autoren Mittelwerte von 96% (Sensitivität) bzw. 88% (Spezifität) ein.

Im Modellansatz werden von den 44 750 Fällen mit rezidivierendem kolorektalem Karzinom (siehe oben) - unter der Annahme, daß im konventionellen Versorgungsprogramm bei 47% oder 20 900 Fällen eine explorative Laparotomie erfolgt - 20% oder 8 950 Patienten einer erneuten kurativen Resektion unterworfen. Im alternativen Algorithmus - mit PET anstelle der explorativen Laparotomie - dagegen würden bei einer diagnostischen Effektivität der PET von 96% (Sensitivität) bzw. 88% (Spezifität) von 20 900 untersuchten Patienten nur 4 538 Patienten, also 4 412 Personen weniger, erneut mittels Resektion operativ therapiert. Unter den vorgegebenen Konditionen hätten die Kostenträger, durch Reduktion der Zahl unnötiger Resektionen und durch vollständigen Ersatz der invasiven Laparotomie durch nicht-invasive PET-Scans, auf nationaler Ebene - zum Zeitpunkt der Studie - jährliche Kosteneinsparungen von mindestens US\$ 113 Millionen (bei reiner Medicare-Vergütung) bis zu maximal US\$ 671 Millionen (bei rein privater Kostenabrechnung) realisieren können. Somit ergäbe sich pro untersuchtem Darmkrebspatienten (n = 20 901) eine durchschnittliche Kostenreduktion von etwa US\$ 5 420 bis US\$ 32 120. Dabei verursachte - nach Aussage der Autoren - zu Beginn der 1990er Jahre eine überflüssige Resektion von inoperablen Darmkrebsrezidiven mit unerkannten, weitgestreuten Metastasen Kosten von US\$ 20 000 bis 30 000 pro Fall (Erfahrungswert).

Aufgrund der Ergebnisse des Literaturreviews und des retrospektiven ICP-Survey bescheinigt die ICP-Fachgruppe der PET eine höhere diagnostische Effektivität bei der Beurteilung von potentiell resektablen kolorektalen Tumorrezidiven als der chirurgischen Alternativmethode, der explorativen Laparotomie. Diese Aussage, die in einer zukünftigen prospektiven Studie überprüft werden soll, wird von den Autoren leider nicht durch die Angabe der diagnostischen Kenndaten für die explorative Laparotomie belegt. Jedoch erstreckt sich bei einer Ganzkörper-PET-Aufnahme das untersuchte Areal auch auf entfernte Körperregionen (Detektion von Fernmetastasen) und ist nicht nur, wie bei der explorativen Laparotomie, auf das Abdomen beschränkt. Aus den ökonomischen Modellberechnungen folgern die Autoren, daß bei geeigneter Anwendung der PET beim Staging von rezidivierendem kolorektalem Karzinom, das heißt, als Ersatzmethode anstelle der üblichen Laparotomie und nicht als add-on-Technologie, ein erhebliches Kosteneinsparpotential realisierbar sei.

Modellierung von R. E. Coleman et al. 1994 (Untersuchung von solitären Lungenknoten)

Die Studie der US-Expertengruppe beschäftigt sich mit der Untersuchung von solitären Lungenknoten (SPN) mittels PET. Wie bei den vorigen ICP-Studien werden auch hier zwei unterschiedliche Versorgungskonzepte skizziert und der damit verbundene Ressourcenverbrauch aggregiert auf Bevölkerungsniveau bestimmt. Dazu wird das nach Auswertung der Fachliteratur entwickelte „konventionelle“ Behandlungsschema, basierend auf CT, transthorakale Nadelaspirationsbiopsie und Thorakotomie, mit einem alternativen, PET-basierten Algorithmus verglichen. Bei letzterem soll eine

¹⁸FDG-PET-Aufnahme die übliche CT-Untersuchung, plus eine sich eventuell anschließende Nadelbiopsie, ersetzen.

Zu Beginn der 1990er Jahre wurden in den USA bei etwa 172 000 neu aufgetretenen Fällen von Lungenkrebs pro Jahr circa 130 000 SPNs in der Lunge festgestellt. Laut ICP-Modellierung auf US-Bevölkerungsniveau wurden davon 38 600 Fälle (30%) auf der Basis einer Eingangsuntersuchung auf Risikofaktoren und durch konventionelle Röntgenaufnahmen der Lunge als gutartig eingestuft. Die übrigen rund 91 400 Knoten wurden als unbestimmbare SPNs eingeordnet und machten eine weitergehende Untersuchung notwendig. Diese Patientengruppe mit unbestimmten SPNs bildet das Patientenkollektiv für die eigentliche ökonomische Bewertung.

Die retrospektive Datenauswertung im Rahmen einer Multicenter-Studie an 10 US-amerikanischen PET-Zentren weist für ¹⁸FDG-PET eine Sensitivität von 96% und eine Spezifität von 90% aus. In der Literatur finden sich Werte von 95% bzw. 100% für die Sensitivität und 86% bzw. 89% für die Spezifität. Gestützt auf veröffentlichte Daten und empirische Daten der Multicenter-Studie diagnostiziert PET im ICP-Versorgungsmodell 37% der verdächtigen SPNs als benigne und 63% als maligne oder unbestimmbar, mit einer falsch-negativen Rate von 7% und einem positiven Vorhersagewert (PPW) von 94%. Die konventionelle CT/Biopsie-Kombination dagegen klassifiziert 22% als benigne und 78% als unbestimmte SPNs, mit einer falsch-negativen Rate von 8% und einem PPW von 77%. Unter der Vorgabe, daß (a) im konventionellen, CT-basierten Algorithmus 75% der CT-positiven Fälle direkt in die Chirurgie zur Thorakotomie überwiesen werden, während die restlichen 25% mittels Biopsie weiteruntersucht und erst bei einem positiven Biopsieergebnis in die Chirurgie überstellt werden; und (b) daß im alternativen PET-Algorithmus alle PET-positiven Personen (maligne und unbestimmte SPNs) unter Umgehung einer Biopsie direkt der Chirurgie zugeführt werden, erweisen sich im Modell bei CT-basierter Diagnose 20% der entfernten Knoten in der Thorakotomie als benigne, bei PET dagegen nur 6%. Durch den Ersatz von CT durch PET im Diagnosealgorithmus könnte folglich der Anteil der gutartigen an den insgesamt entfernten solitären Lungenknoten gesenkt werden. Dadurch hätten - laut Modellrechnung - zu Beginn der 1990er Jahre circa 10 100 chirurgische Resektionen jährlich wegen benigner SPNs unterbleiben können und zudem wären circa 17 800 Nadelbiopsien pro Jahr entfallen.

Auf nationaler Ebene hätten somit durch verbessertes Patientenmanagement - bezogen auf den Studienzeitraum und je nach Abrechnungssystem - Nettoeinsparungen von US\$ 30 Millionen (Mindestbetrag bei reiner Medicare-Vergütung) bis US\$ 236 Millionen (Maximalbetrag bei ausschließlich privater Kostenabrechnung) erzielt werden können. So wären beim PET-Algorithmus bei privater Kostenabrechnung Mehrausgaben für die Diagnostik von circa US\$ 21 Millionen durch Minderausgaben für unterlassene therapeutische Maßnahmen von US\$ 257 Millionen um mehr als das 12fache ausgeglichen worden. Pro untersuchtem Patienten hätte sich somit eine durchschnittliche Nettoersparnis von mindestens US\$ 325 (nur Medicare-Vergütung) bis höchstens US\$ 2 583 (bei privater Kostenabrechnung) ergeben.

Aufgrund dieser Ergebnisse der retrospektiven Studie zur diagnostischen Treffsicherheit von PET und der darauf aufbauenden Kostenbetrachtung vermuten die Autoren, daß PET gegenüber dem derzeitigen Standardverfahren (CT plus Nadelbiopsie) folgende Vorteile bietet: Durch die höhere Treffsicherheit bei der Differenzierung zwischen malignen und benignen SPNs könnte die Anzahl unnötiger diagnostischer und therapeutischer Maßnahmen und damit zusammenhängende Komplikationen (z.B. perioperative Mortalität oder Pneumothorax) eingeschränkt werden. Gemäß dem Urteil der ICP-Fachgruppe ist PET neben der Thorakotomie zu diagnostischen Zwecken (100%ige Genauigkeit für die Diagnose von SPNs) die beste verfügbare Methode zur Differenzierung zwischen malignen und benignen SPNs. Ihrer Einschätzung nach führt der Einsatz der PET anstelle der Kombination CT/Nadelbiopsie, indem Thorakotomien zu Therapie Zwecken bei malignen SPNs maximiert und bei benignen SPNs minimiert werden, zu einer erheblichen Verbesserung in der Patientenbetreuung, verbunden mit Einsparungen auf der Kostenseite von mindestens US\$ 30 Millionen bis maximal US\$ 236 Millionen.

Gemeinsame Kurzbewertung der drei onkologischen ICP-Einzelstudien

In den ICP-Arbeiten aus dem Jahr 1994 werden Versorgungsszenarien auf nationalem Niveau entworfen. Dabei nutzen die Autoren - neben eigenen Auswertungen von PET-Anwendungsdaten aus Multicenter-Studien - Bezugswerte und Kenngrößen (z.B. Kenndaten diagnostischer Verfahren) aus der Fachliteratur, um die Menge an jährlich erbrachten Versorgungsleistungen für die betroffenen Bevölkerungsgruppen modellieren zu können. Die drei Studien (ebenso die vierte, kardiologische Studie, siehe C.5.4.2) sind formal vergleichbar aufgebaut und die Struktur der ökonomischen Modellrechnungen ist weitgehend identisch. Deshalb ist trotz unterschiedlicher Indikationsstellungen eine gemeinsame Bewertung der Studiengüte nach den Kriterien des Transparenz- bzw. Qualitätskataloges gerechtfertigt.

Die Fragestellung der einzelnen Studien ist entweder direkt formuliert (Larson et al. 1994) oder ergibt sich aus dem inhaltlichen Zusammenhang (Adler et al. 1994 und Coleman et al. 1994). In allen Studien bleibt in der Fragestellung jedoch unklar, welche Alternativmethoden zu PET in die Studien einbezogen sind. Der medizinische Problemkontext wird von allen Autoren verständlich und ausreichend dargelegt; der ökonomische Problemkontext dagegen wird nur bei der ALND-Studie ausreichend besprochen (Adler et al. 1994), ansonsten nur angerissen (Larson et al. 1994) oder ganz übergangen (Coleman et al. 1994).

In allen Veröffentlichungen sind der Evaluationsrahmen und die Vorgehensweise leider nur sehr knapp dargestellt. So sind die alternativen Versorgungsstrategien in Form von übersichtlichen Ablaufschemata veranschaulicht, aber die einzelnen Bausteine der zusammengesetzten Algorithmen, Verfahren zur Bilderfassung wie ¹⁸FDG-PET oder CT und invasive, diagnostische oder therapeutische Eingriffe, sind metho-

disch nicht eingehender beschrieben. Die Perspektive wird nicht genannt, entspricht jedoch der Sicht des Drittfinanzierers.

In allen drei Krebsstudien werden die Ergebnisse der Literaturreviews zur diagnostischen Leistungsfähigkeit konventioneller Diagnoseverfahren und der alternativen PET-Methode und die Ergebnisse der retrospektiven ICP-Auswertungen der Multicenter-Fallserien zur diagnostischen Treffsicherheit der PET gut dokumentiert und übersichtlich in Tabellen zusammengestellt. Diese Auflistungen enthalten leider keinerlei Angaben zu den jeweils rekrutierten Studienpopulationen (Fallzahl, Patientenmerkmale). Deshalb kann nicht beurteilt werden, ob die Studienpopulationen in den angeführten klinischen Studien die jeweilige Zielpopulation in der US-Bevölkerung (siehe Studienbeschreibungen) korrekt repräsentieren und folglich die Ergebnisse der Einzelstudien auf das Bevölkerungsniveau übertragbar sind.

Bei den Modellierungen des ICP handelt es sich nicht um ökonomische Evaluationen im strengen Sinn (siehe dazu C.4.2.3). Die biometrischen Kenndaten der diagnostischen Tests (diagnostic accuracy efficacy, intermediäre Effektgröße) gehen nur als Parameter zur Berechnung des Mengengerüsts in die Modellierung ein. Der Effekt des Einsatzes von PET anstelle von oder zusätzlich zu konventionellen Diagnosemethoden auf die weitere Diagnose- und Therapiestrategie (Effektgröße „Anzahl der durch verbessertes Patientenmanagement vermiedenen diagnostischen und therapeutischen Interventionen“, diagnostic thinking efficacy bzw. therapeutic efficacy, siehe C.4.2.2) als direkte Folge der höheren diagnostischen Treffsicherheit der PET, schlägt sich - monetär bewertet - auf der Kostenseite als einzusparende Minderausgaben nieder. Medizinisch relevante, patientenbezogene Effekte (patient outcome efficacy) bleiben in der Studie unbeachtet. So werden langfristige Effekte, beispielsweise der Einfluß auf finale Effektgrößen wie Lebenserwartung oder qualitätsadjustierte Lebenserwartung, deren Erfassung bei so schwerwiegenden, lebensverkürzenden und belastenden Krankheiten wie einer Tumorerkrankung wichtig wäre, leider nicht in den Modellansatz einbezogen. Damit beschränkt sich der Zeithorizont auf die - aus gesundheitsökonomischer Sicht enge - Phase der Akutversorgung.

In den einzelnen Studien wird das Mengengerüst der Kostenkalkulation, bezogen auf das US-Bevölkerungsniveau und nach diagnostischen und therapeutischen Einzelleistungen aufgeschlüsselt, sowohl in den Ablaufschemata als auch im Kostenvergleich detailliert aufgeführt; der Berechnungsweg wird dargestellt. In alle drei Studien ist die Gesamtkostenberechnung durch methodische Mängel gekennzeichnet, welche die Transparenz beeinträchtigen (siehe die Punkte 1 und 2 in der nachfolgenden Aufzählung).

Folgende Punkte sind als methodische Mängel zu werten:

1. Einige der zur Berechnung des Mengengerüsts verwandten Parameter werden nicht ausreichend hergeleitet und durch Literaturquellen belegt.

2. Das Preisgerüst ist nur unzureichend spezifiziert, da die Vergütungen für die jeweils erbrachten Einzelleistungen nicht explizit aufgeführt werden und das Bezugsjahr für die Festlegung des Preisgerüsts nicht explizit genannt wird.
3. Weitere direkte, medizinische Kostenposten, wie Kosten für nicht-operative Therapien (Behandlung der nicht resektierten Darmkrebspatienten, Larson et al. 1994) oder Folgekosten für Fehldiagnosen bzw. perioperative Komplikationen, sind nicht Bestandteil der Kostenkalkulation. Indirekte Kosten werden nicht erhoben.
4. Es werden keine Sensitivitätsanalysen zur Überprüfung der Stabilität der Modellierungen ausgeführt.

Bedingt durch den engen Zeitrahmen erübrigt sich in allen Studien eine Diskontierung bzw. Inflations-/Deflationsbereinigung. Ein Hauptmangel bei allen drei Studien, besonders jedoch bei der Publikation von Larson et al. (Studie zu rezidivierendem kolorektalem Karzinom), ist die knappe Darstellung und Diskussion der Ergebnisse. Weder die Datenqualität, noch der Einfluß unsicherer Parameter auf das Ergebnis werden diskutiert, noch werden die methodischen Mängel der Studien angesprochen. Aufgrund der meist knappen Zusammenfassung und Diskussion der Ergebnisse in allen drei Studien wird die Argumentationskette, mit der die Autoren ihre Schlußfolgerungen begründen, dem Leser nicht deutlich dargelegt (besonders Larson et al. 1994 und Coleman et al. 1994).

Abschließende Bewertung

Diese drei Studien des Institute for Clinical PET (ICP) wurden bereits 1994 anlässlich einer ICP-Tagung der Fachöffentlichkeit vorgestellt. Sie untersuchen die Auswirkungen auf das Management von Krebspatienten (Patienten mit rezidivierendem Darmkrebs, mit Mammakarzinom bzw. mit solitären Lungenknoten) und auf die dabei anfallenden Gesamtversorgungskosten, die sich durch den teilweisen oder vollständigen Ersatz von konventionellen diagnostischen Verfahren (z.B. CT oder Laparotomie) durch die PET oder durch den zusätzlichen Einschub von PET-Untersuchungen in konventionelle Versorgungsstrategien ergeben. Da die Modellierungen auf US-Bevölkerungsniveau nur die Kostenseite der PET-gestützten bzw. der konventionellen Versorgungsalgorithmen vergleichend bestimmen, diese jedoch nicht zu einer unabhängigen medizinischen Effektgröße in Beziehung setzen, handelt es sich bei den vorliegenden ICP-Studien aus gesundheitsökonomischer Sicht nicht um vollständige, sondern um partielle ökonomische Evaluationen (siehe dazu C.4.2.3). Da alle Studien einen engen Zeithorizont (keine Betrachtung von langfristigen Wirkungen) und weitere methodische Mängel aufweisen (Schwächen in der Dokumentation der verwandten Daten, der Kostenkalkulation und der Darstellung der Ergebnisse, das Fehlen von Sensitivitätsanalysen), ist ihre Aussagefähigkeit begrenzt.

C.5.4.2 Kardiologie

In diesem Kapitel werden 6 ökonomische Modellansätze (6 Originalartikel und zwei ergänzende Publikationen) zur Verwendung der PET-Technologie im Indikationsgebiet Kardiologie vorgestellt. In den zwei ersten Studien dient die PET-Technologie der Messung der Myokardviabilität nach einem Herzinfarkt, in den übrigen vier Studien wird PET zur Perfusionsmessung am Myokard eingesetzt.

Indikationsstellung: Messung der Myokardviabilität

Besozzi MC, Alagona P, Hunter K et al. Positron emission tomography. Clinical application and economic implications of PET in the assessment of myocardial viability: A retrospective study. Institute for Clinical PET, 1994, Publication No. CE-CARDIOLOGY-94.001

Studienbeschreibung

Im Rahmen der einzigen ökonomischen Studie des Institute for Clinical PET zum Indikationsgebiet Kardiologie soll untersucht werden, welche klinische und ökonomische Bedeutung der PET zukommt als Instrument (a) zur Identifikation von lebensfähigen Myokardarealen bei Patienten nach einem akuten Herzinfarkt und (b) zur Prognose, inwieweit die Ventrikelfunktion bei Patienten mit schwerer linksventrikulärer Dysfunktion durch eine Revaskularisation (hier mittels Bypass-Operation) wiederherstellbar ist.

Die aus klinischer Sicht zur Zeit beste Methode (Goldstandard), die Wiederherstellung der linksventrikulären Funktion bei infarktgeschädigtem Myokard vorhersagen zu können, ist die Vitalitätsmessung mit dem funktionellen, bildgebenden Verfahren der ¹⁸FDG-PET. Bei Patienten mit schwerwiegender linksventrikulärer Dysfunktion besteht bei einer operativen Revaskularisation ein Mortalitätsrisiko von 5 – 37% und ein Komplikationsrisiko von 10 – 15% mit geschätzten durchschnittlichen Kosten von etwa US\$ 100 000 pro Komplikationsfall. Aufgabe des PET-Einsatzes ist es, die Gruppe der Hochrisikopatienten, die von einer Revaskularisation nicht profitiert, vorab herauszufiltern.

In der retrospektiven ökonomischen Analyse der kardiologischen ICP-Fachgruppe werden die bei der konventionellen Versorgung von Infarktpatienten mit schwerer Linksherzdysfunktion anfallenden Kosten den Kosten für einen alternativen, PET-gestützten Versorgungsalgorithmus vergleichend gegenübergestellt. Dabei ist im alternativen PET-basierten Untersuchungsablauf die PET-Diagnostik als Vortest in das konventionelle Versorgungsprogramm, zusammengesetzt aus einer Koronararteriographie mittels Herzkatheter (Angiographie) und einer anschließenden Bypass-Operation, eingeschoben (additiver PET-Einsatz).

In den USA erlitten zu Beginn der 1990er Jahre jährlich rund 1,5 Millionen Amerikaner einen akuten Myokardinfarkt. Mehr als die Hälfte der Betroffenen entwickelte eine Linksherzdysfunktion, eventuell verbunden mit einer Stauungsinsuffizienz, an der etwa 30 000 Patienten verstarben. Ein Anteil von etwa 40% der inzidenten Fälle, also 608 400 Infarktpatienten, wurden stationär versorgt. Zur Abklärung der Belastbarkeit dieser Patientengruppe wurden 456 300 Fälle (75% der stationär versorgten Patienten) einem Belastungstest unterzogen. Bei den übrigen 25%, 152 100 Patienten, war ein Belastungstest nicht angezeigt, entweder wegen einer guten klinischen Prognose (bei 38 000 Fällen) oder wegen der gefährlich verringerten linksventrikulären Herzfunktion. Die Patienten mit guter klinischer Prognose wurden ohne weitere diagnostische Untersuchung medikamentös behandelt. Die restlichen 114 075 Fälle, sogenannte Kategorie I-Patienten, mit Herzinsuffizienz als Hauptsymptom und geringer Auswurffraktion des linken Ventrikels (< 35%, NYHAC-Kategorie III oder IV), aber ohne Brustschmerzsymptomatik, bilden die Zielpopulation der theoretischen ICP-Modellierung im Indikationsgebiet Kardiologie. Diese Personengruppe wird im theoretischen Modell den alternativen Diagnosestrategien „kombinierte PET/Angiographie“ versus „nur Angiographie“ zugeführt, um die Lebensfähigkeit des Myokards vor einer Bypass-Operation abzuklären.

Im Rahmen des konventionellen Versorgungsmodells werden alle 114 075 Kategorie I-Patienten präoperativ durch koronare Arteriographie mit dem Herzkatheter (Angiographie) untersucht. 58 178 Infarktpatienten (51%) werden dabei für eine Bypass-Operation als geeignet befunden, während die übrigen Infarktpatienten medikamentös behandelt werden. Postoperativ zeigt sich bei 41 713 der mittels Bypass-Revaskularisation behandelten Patienten eine Verbesserung der Linksherzdysfunktion, bei den restlichen 16 465 operierten Patienten (28%) treten operationsbedingte Komplikationen auf und/oder ist postoperativ keine Verbesserung der Auswurffraktion des linken Ventrikels festzustellen.

Im alternativen, von der ICP-Fachgruppe vorgeschlagenen PET-basierten Algorithmus dagegen werden von den 114 075 Infarktpatienten nach der PET-Untersuchung nur noch 81 717 Personen (72%; anstatt alle 114 075 Personen beim konventionellen Diagnosealgorithmus) zur weiteren Abklärung einer Herzkatheteruntersuchung unterzogen. Davon werden 51 004 (45%; anstatt 58 178 Personen beim konventionellen Diagnosealgorithmus) schließlich zur Bypass-Operation überstellt. Bei nur 9 872 Personen (19%; gegenüber 16 465 beim konventionellen Diagnosealgorithmus), treten in der Modellierung operationsbedingte Komplikationen auf und/oder ist postoperativ keine Verbesserung der Linksherzdysfunktion zu erwarten. Im PET-gestützten Algorithmus würde somit die Zahl der wirkungslosen oder von Komplikationen begleiteten Revaskularisationen reduziert. Medikamentös behandelt werden im Modell insgesamt (einschließlich der 38 000 Fälle mit guter klinischer Prognose) 101 096 Personen gegenüber 93 922 Personen beim konventionellen Untersuchungsverfahren (teilweiser Ersatz der operativen Revaskularisation durch Medikation).

Beim direkten Vergleich der beiden Versorgungsstrategien ergibt sich somit folgendes Ergebnis: Durch den *zusätzlichen* Einsatz von PET (vor der Angiographie) wären - auf dem US-Bevölkerungsniveau - 32 358 Angiographien und 7 174 Bypass-Operationen weniger durchgeführt worden als nach der üblichen Diagnostik ausschließlich per Herzkatheter. Gleichzeitig wären - bei einer Komplikationsrate der Bypass-Operation von 12% (Modellannahme) - beim PET-basierten Algorithmus im Vergleich zum konventionellen Algorithmus für 861 vermiedene Operationen keine zusätzlichen Kosten für operationsbedingte Komplikationen angefallen. Bei einer monetären Bewertung aller vermiedenen diagnostischen und therapeutischen Interventionen und unter Berücksichtigung der finanziellen Aufwendungen für operative Komplikationen (Modellannahme: US\$ 100 000 pro Komplikationsfall) ergibt sich - bezogen auf die frühen 1990er Jahre - für den PET/Herzkatheter-Algorithmus auf der Basis privater Vergütung ein Einsparpotential von maximal US\$ 300 Millionen und auf der Basis der Medicare-Vergütung ein Einsparpotential von minimal US\$ 41 Millionen.

Aus diesen Ergebnissen ziehen die Autoren das Resümee, daß, bedingt durch ein verbessertes Patientenmanagement, der additive Einsatz der PET zur Bestimmung der Myokardviabilität im Vergleich zum konventionellen Vorgehen kosteneffektiv sein kann. Denn Patienten mit verminderter Auswurfraction des linken Herzventrikels werden, wenn keine Aussicht auf Reversibilität dieser Dysfunktion besteht, nicht zur koronaren Arteriographie bzw. zur anschließenden chirurgischen Revaskularisation weitergeleitet und somit fallen durch die Unterlassung von medizinisch ineffektiven Interventionen weniger Gesamtversorgungskosten an.

Studienbewertung

Die erste analysierte Studie zur Thematik des PET-Einsatzes im Indikationsgebiet Kardiologie, eine theoretische Modellierung des ICP, ist im Aufbau, speziell in der Struktur der ökonomischen Modellrechnung, und in ihrer methodischen Qualität den drei im vorangegangenen Berichtsteil vorgestellten onkologischen ICP-Studien vergleichbar (siehe C.5.4.2). Auch in der kardiovaskulären Studie entwerfen die Autoren retrospektiv ein Versorgungsszenarium auf nationalem, US-amerikanischen Niveau.

Wie bereits in den onkologischen Studien ist auch hier der Evaluationsrahmen und die Vorgehensweise nur sehr knapp beschrieben. Die Aufgabenstellung wird dargelegt, definiert jedoch nicht explizit, welche alternative Vorgehensweise zur Identifikation von lebensfähigen Myokardarealen im Vergleich zur PET-gestützten Strategie herangezogen wird. Der medizinische Problemkontext ist ausreichend erläutert, der ökonomische Problemkontext dagegen wird nur angerissen. Die zwei alternativen Versorgungsstrategien werden in Form von Ablaufschemata illustriert; die einzelnen Algorithmusbausteine - wie nicht-invasive (^{18}F FDG-PET) bzw. invasive (Herzkatheter) Diagnoseschritte bzw. Therapieformen (Medikation bzw. Revaskularisation) - werden aber nicht näher spezifiziert. Die Perspektive wird nicht definiert, entspricht jedoch

der Sicht des Drittfinanzierers, da zur Gesamtkostenberechnung wie in den onkologischen Studien Vergütungssätze des Medicare-Systems bzw. privater Kostenabrechnung veranschlagt werden (siehe onkologische ICP-Studien, Kapitel C.5.4.1). Dagegen ist die Zielpopulation klar definiert und hinreichend beschrieben (auch wenn eine Geschlechts- und Alterszuordnung der Kategorie I-Patienten fehlt).

Im Gegensatz zu den onkologischen Studien standen bei der Konzeption des kardiologischen Modells keine Daten aus ICP-Auswertungen von Multicenterstudien zur Verfügung. Deshalb verwendet die Arbeitsgruppe Besozzi et al. zur Berechnung der Menge an jährlich US-weit erbrachten Versorgungsleistungen für inzidente Herzinfarktpatienten neben Bezugswerten und Kenngrößen aus der Fremdliteratur eigene Studienergebnisse einer früheren klinischen Studie. Die Anzahl der jeweils erbrachten diagnostischen und therapeutischen Einzelversorgungsleistungen (Mengengerüst) ist in den alternativen Ablaufschemata vermerkt und in der Kostenkalkulation detailliert aufgeführt. Die zur Berechnung des Mengengerüsts genutzten Modellparameter sind jedoch nicht vollständig dokumentiert und mit Literaturstellen belegt, so daß die einzelnen Schritte zur Berechnung des Mengengerüsts für den Leser nicht ausreichend transparent sind. Das Preisgerüst ist nur unzureichend spezifiziert, da die absoluten Vergütungsbeträge für die erbrachten Einzelleistungen nicht explizit beziffert sind und das Bezugsjahr nicht datiert ist. Vielmehr sind die Kostenposten für die jeweiligen Einzelleistungen (z.B. für PET-Scans) auf das US-amerikanische Bevölkerungsniveau hochgerechnet und zu nationalen Gesamtversorgungskosten aufsummiert (aggregierte Darstellung der Kostenseite).

Der Einschub einer PET-Untersuchung zusätzlich in den üblichen Diagnoseverlauf beeinflußt die weitere Diagnose- bzw. Therapiestrategie (diagnostic thinking efficacy und therapeutic efficacy, siehe C.4.2.2), da bei 32 358 Patienten eine Herzkatheteruntersuchung und bei weiteren 7 174 Patienten eine Bypass-Operation zugunsten einer medikamentösen Behandlung entfällt. Dabei bleibt die teilweise Umwandlung der Aufwendungen für operative in medikamentöse Behandlung unberücksichtigt, indem bei der Kostenkalkulation zwar die Kosten für die chirurgische Therapie, nicht jedoch die Kosten für die medikamentös behandelten Patienten veranschlagt werden. Da - im Gegensatz zu den onkologischen ICP-Studien - auch operationsinduzierte Komplikationen miterfaßt werden, subsumieren sich unter der Effektgröße „Anzahl der unterlassenen diagnostischen und therapeutischen Interventionen“ auch die komplikationsinduzierten Interventionsmaßnahmen, die bei 861 statistisch vermiedenen Komplikationsfällen nicht notwendig würden. Diese im PET-gestützten Algorithmus wegfallenden diagnostischen und therapeutischen Interventionen gehen monetär bewertet in die Gesamtkostenkalkulation des konventionellen Algorithmus - und somit methodisch korrekt auf der Kostenseite - ein und werden mit den Zusatzkosten für PET-Scans im alternativen PET-Algorithmus verrechnet. Da nur die Kostenseite des PET-gestützten bzw. des konventionellen Versorgungsalgorithmus vergleichend bestimmt, diese jedoch nicht zu einer unabhängigen medizinischen Effektgröße in Beziehung setzt, handelt es sich bei der vorliegenden ICP-Studie aus gesundheitsökonomischer Sicht um eine partielle ökonomische Evaluationen (siehe dazu

C.4.2.3). Auf der Effektseite wird, indem die Wirkung der Revaskularisation auf die Wiederherstellung der linksventrikulären Herzfunktion gewürdigt wird, zwar ein längerfristiger Effekt betrachtet (patient outcome efficacy); da dieser jedoch nicht in die Wirtschaftlichkeitsbewertung einfließt, wird das Zeitfenster der gesundheitsökonomischen Analyse auf die relativ kurze Zeitspanne der Akutversorgung verengt. Dies ist aus gesundheitsökonomischer Sicht nicht zufriedenstellend. Die Autoren hätten den Einfluß eines Therapieerfolges, erkennbar an der Verbesserung der linksventrikulären Herzfunktion, bzw. eines Therapieversagens auf die Lebenserwartung oder auf die Lebensqualität untersuchen und durch direkte Bezugsetzung dieser aussagekräftigen finalen Effektgröße zu den errechneten Versorgungskosten eine sinnvolle Kosten-Effektivitäts-Relation ableiten können. Wegen des auf die Phase der Akutversorgung verengten Zeitfensters erübrigt sich eine Diskontierung bzw. eine Inflations-/Deflationsbereinigung (sofern aktuelle Kostendaten verwandt wurden).

Es werden keine Sensitivitätsanalysen zur Prüfung der Stabilität des Modells durch Variation von unsicheren Parameterschätzungen vorgenommen (z.B. Variation der Kosten pro Komplikationsfall oder der Komplikationsrate).

Die Studienergebnisse werden nur in sehr knapper Form dargestellt. Die Datenqualität wird nicht hinterfragt, der Einfluß unsicherer Parameter auf das Ergebnis nicht kritisch beleuchtet, und die methodischen Mängel der Studie nicht angesprochen. Aufgrund der knappen Zusammenfassung und Diskussion der Ergebnisse wird die Argumentationskette, mit der die Autoren ihre Schlußfolgerungen begründen, dem Leser nicht deutlich dargelegt.

Abschließende Bewertung

Die kardiologische Studie des ICP, ebenfalls eine Modellierung auf US-Bevölkerungsniveau, wurde wie die entsprechenden onkologischen Studien bereits 1994 anlässlich einer Tagung des Institute for Clinical PET der Fachöffentlichkeit vorgestellt. Sie untersucht die klinische und ökonomische Bedeutung der PET als Instrument zur Überprüfung der Myokardviabilität bei Herzinfarktpatienten und berechnet das Einsparpotential, das sich durch den Einschub einer PET-Untersuchung vor die konventionelle Koronararteriographie mittels Herzkatheter (Angiographie) ergibt. Die Ergebnisse dieser Studie, die wie die vorab besprochenen onkologischen ICP-Studien einen engen Zeithorizont (kein Einbezug von langfristigen Wirkungen in die ökonomische Analyse) und weitere methodische Mängel aufweist (Schwächen in der Dokumentation der verwandten Daten, der Kostenkalkulation und der Darstellung der Ergebnisse, das Fehlen von Sensitivitätsanalysen), sind in ihrer Aussagefähigkeit begrenzt.

Hunink MGM. Positron emission tomography compared to single photon emission computed tomography in the evaluation of myocardial viability: a preliminary cost-effectiveness analysis. In: Van der Wall et al. (eds.) Cardiac positron emission tomography. 1995 Kluwer, Amsterdam:97-102

Studienbeschreibung

In dem 1995 erschienenen niederländischen Buchbeitrag wird ein erster Entwurf eines entscheidungstheoretischen Modells präsentiert, das die Kosteneffektivität der Positronen-Emissions-Tomographie (PET) im Vergleich zur Routinemethode Single-Photon-Emissions-Computer-Tomographie (SPECT), alternative diagnostische Instrumente zur Wahl einer adäquaten Therapie nach einem Herzinfarkt, vergleichend prüfen soll. Als Zielpopulation der analytischen Studie definiert sind Myokardinfarktpatienten mit möglicher Restischämie, bei denen eine Therapieentscheidung im Hinblick auf eine Revaskularisation, entweder mittels perkutaner transluminaler Koronar-Angioplastie (PTCA) oder mittels Koronararterien-Bypass (CABG), ansteht. Eine Abklärung der Lebensfähigkeit des Herzmuskels nach einem Infarkt, speziell der objektivierte Nachweis von ischämischen Herzmuskelarealen, mit Hilfe einer nicht-invasiven, präoperativen Untersuchung soll die Therapieentscheidung für oder gegen eine Revaskularisation erleichtern.

Unter Verwendung von Ergebnissen bereits publizierter Studien (siehe Studienbewertung) konstruiert die Autorin ein analytisches Modell, das einen Entscheidungsbaum und ein Markov-Modell zur Modellierung einer mehrstufigen Nachbeobachtungszeit (1, 2 und 5jähriges Follow-up) beinhaltet. Der Entscheidungsbaum ist aus zwei Interventionsästen aufgebaut, in denen zwar beide Diagnoseverfahren PET und SPECT in der Reihenfolge PET/SPECT aufeinanderfolgen, jedoch alternativ jeweils nur eine der beiden Technologien über die Durchführung einer Revaskularisation entscheidet:

1. Bei der PET-basierten Interventionsstrategie schließt sich nur bei positivem PET-Ergebnis eine Revaskularisation an die nuklearmedizinische Eingangsuntersuchung an, bei negativem Befund unterbleibt die invasive Therapie;
2. entsprechend hängt bei der SPECT-basierten Strategie die Entscheidung für bzw. gegen eine Revaskularisation von einem positiven bzw. negativen SPECT-Untersuchungsausgang ab.

Auch in die Berechnung der Gesamtkosten der zwei alternativen Interventionsstrategien werden nur die Kosten für die jeweils ausschlaggebende Diagnosemethode aufgenommen. Die aufsummierten Gesamtkosten einer Strategie enthalten somit als Einzelposten: die Kosten für die PET- oder die SPECT-Untersuchung, ferner die Kosten für eine Revaskularisation (sofern durchgeführt) und für im Follow-up auftretende „downstream events“, beispielsweise „kardiale Ereignisse“ wie Herztod, erneuter

Myokardinfarkt, erneute Angina pectoris-Anfälle, Herzversagen, Kammertachykardie und/oder Vorhofflimmern, das eine stationäre Behandlung notwendig macht. Das der Modellierung der Follow-up-Phase dienende Markov-Modell wird in der Publikation nicht detailliert ausgeführt.

Für die in Abschnitte von 1, 2 und 5 Jahren gestaffelte Follow-up-Phase wird die Wahrscheinlichkeit von „downstream events“, das heißt, daß „kardiale Ereignisse“ auftreten und/oder eine erneute Revaskularisation erforderlich wird, für die Alternativstrategien abgeschätzt. So prognostiziert das mathematische Modell, daß bei 88% der mit PET, dagegen nur bei 75% der mit der Alternativtechnologie SPECT untersuchten Herzinfarktpatienten das erste Jahr der Nachbeobachtungszeit „ereignisfrei“ verläuft. Auch 2 bzw. 5 Jahre nach der Bypass-Operation ist laut Modellberechnung die statistische Wahrscheinlichkeit, diese Follow-up-Phasen „ereignisfrei“ zu erleben, für Patienten, bei denen die Therapieentscheidung auf einem PET-Befund basiert, höher als bei mittels SPECT untersuchten Patienten; sie beträgt nach 2 bzw. 5 Jahren für PET 0,77 bzw. 0,51 (77% bzw. 51%) gegenüber 0,61 bzw. 0,38 (61% bzw. 38%) für SPECT. Die bessere klinische Effektivität der PET-Strategie, die sich in einer Reduktion der Anzahl an „downstream events“ in der Nachbeobachtungsphase niederschlägt, ist jedoch in allen drei Zeitabschnitten mit um 1,3 - 1,4mal höheren Gesamtkosten verbunden. So weist die inkrementelle Kosten-Effektivitäts-Ratio für PET gegenüber SPECT im ersten Jahr der Follow-up-Phase Mehrkosten von circa US\$ 22 600 pro gewonnenem „ereignisfreiem“ Lebensjahr aus und sinkt im Verlauf des 5jährigen Follow-up - nach 2 bzw. 5 Jahren - auf rund US\$ 9 000 bzw. US\$ 3 800 pro gewonnenem „ereignisfreiem“ Lebensjahr. Nach Ansicht der Autorin sind zeitlich degressive Mehrausgaben von \leq US\$ 22 600 (\cong 40 000 niederländische Gulden) für ein zusätzliches „ereignisfreies“ Lebensjahr vertretbar und deshalb PET aufgrund des verbesserten gesundheitlichen Outcome gegenüber SPECT zu favorisieren.

Studienbewertung

Die vorliegende Studie zur Anwendung der PET alternativ zur SPECT bei der Festlegung einer Revaskularisationsstrategie für Myokardinfarktpatienten basiert auf einem entscheidungstheoretischen Modell. Nach eigener Aussage der Autorin ist die Studie in der 1995 publizierten Form erst ein vorläufiger Entwurf einer Kosten-Effektivitäts-Analyse; deshalb ist sie auch als solcher zu beurteilen.

Die Aufgabenstellung ist, sieht man von der allgemein gehaltenen Definition der Zielpopulation ab, klar vorgegeben. Eine Einführung in die Thematik durch eine Darstellung des medizinischen und ökonomischen Problemkontextes unterbleibt leider. Auch die Darstellung des Evaluationsrahmens erfüllt die formalen und qualitativen Anforderungen des Transparenz- bzw. Qualitätskataloges nur teilweise. So wäre eine stärkere Spezifizierung der Zielpopulation (z.B. nach Alter, Geschlecht oder Schwere der Erkrankung) wünschenswert, zumal Effektivitätsdaten für PET als

Pfadwahrscheinlichkeiten im Entscheidungsbaum eingesetzt werden, die an Patienten mit verminderter Auswurfraction (im Mittel 35%) und durchschnittlich zwei betroffenen Koronararterien ermittelt wurden. Zudem wird weder die Auswahl der zwei Vergleichstechnologien PET bzw. SPECT begründet, noch werden deren Einsatzbedingungen beschrieben (z.B. keine Angabe der verwendeten Radiodiagnostika). Der vorgegebene Zeithorizont von maximal 5 Jahren ist bei der begrenzten Verfügbarkeit von Langzeitdaten als ausreichend anzusehen, auch wenn eine Ausweitung des Zeithorizontes auf mindestens 10 Jahre erstrebenswert wäre, um die Entwicklung der inkrementellen Kosten-Effektivität-Ratio langfristig verfolgen zu können (degressiver Verlauf in der 5jährigen Follow-up-Phase).

Der erste Teil des mathematischen Modells, der Entscheidungsbaum, ist adäquat gewählt und in einem Baumdiagramm anschaulich dargestellt. Die notwendigen Modellparameter - Wahrscheinlichkeiten für die vier möglichen PET/SPECT Befundkombinationen (SPECT + | PET +, SPECT - | PET +, SPECT + | PET - und SPECT - | PET-), die vom PET-Ergebnis und der Revaskularisationsentscheidung (Revaskularisation durchgeführt oder nicht) abhängige Ereignisrate am Ende des ersten Follow-up-Jahres und die Interventionskosten (Kosten für Diagnostik und für invasive Therapiemaßnahmen, zuzüglich den beim Management von kardialen „downstream“ Ereignissen anfallenden Kosten) - sind im Text aufgelistet, die Pfadwahrscheinlichkeiten zudem im Baumdiagramm abgedruckt. Die Quellen für die Pfadwahrscheinlichkeiten, Fachveröffentlichungen aus den Jahren 1989 bis 1992, sind vermerkt. Kritisch betrachtet werden muß jedoch die geringe Datenbasis, auf die sich die Modellberechnung stützt; so wird für die einzelnen Parameter nur jeweils eine Literaturstelle zitiert. Da weder das Ergebnis der Literatursuche, noch die Suchstrategie selbst erläutert wird, ist nicht nachprüfbar, inwiefern weitere Veröffentlichungen zur Thematik in der ersten Hälfte der 1990er Jahre zugänglich waren. Die schmale Datenbasis ist möglicherweise auch der Grund, daß für die Modellparameter keine empirischen Streubreiten angegeben werden. Der Einfluß unsicherer Parameter auf den Entscheidungsausgang wird im Modellentwurf nicht anhand von Sensitivitätsanalysen überprüft. Als wichtige Schwäche der Studie ist zu vermerken, daß das Markov-Modell, mit dessen Hilfe das Patientenrisiko modelliert wird, während der Follow-up-Phasen ein weiteres kardiales Ereignis zu erleiden, nicht ausreichend erklärt wird. Außer den potentiellen Gesundheitszuständen (siehe Studienbeschreibung) gibt die Autorin keinerlei Modellvorgaben, wie Übergangswahrscheinlichkeiten zwischen den einzelnen Gesundheitszuständen oder Dauer des Verbleibs in einem bestimmten Gesundheitszustand, an. Auch ein Verweis auf eine erläuternde Literaturstelle fehlt. Die einzelnen Gesundheitszustände werden nicht präferenzbewertet.

Bei entscheidungstheoretischen Ansätzen sind Vereinfachungen unumgänglich, um das Modell überhaupt handhaben zu können. Dennoch wäre eine zusätzliche Verzweigung des Entscheidungsbaumes hinsichtlich der gewählten Revaskularisationsmethode, PTCA bzw. CABG, eventuell hilfreich. Eine zusätzliche Astgabelung ließe sich folgendermaßen begründen: PTCA bzw. CABG sind in der Praxis für zwei unterschiedliche Indikationsstellungen die Methoden der Wahl; so wird eine PTCA meist

bei einer 1- oder 2-Gefäßkrankung mit proximalen kurzstreckigen Stenosen, CABG dagegen meist bei 2- oder 3-Gefäßkrankungen bzw. Hauptstammstenosen der linken Koronararterie durchgeführt. Außerdem wäre es wünschenswert, daß auch die Folgekosten und -effekte von Komplikationen, die während des invasiven Eingriffs der Revaskularisation auftreten können, in die Modellkonzeption eingingen.

Wie für eine vollständige Evaluation erforderlich, werden für beide Versorgungsstrategien sowohl die Kosten der medizinischen Interventionen (siehe Studienbeschreibung) als auch die erzielten gesundheitlichen Effekte abgeschätzt und in einer dem Evaluationstyp angemessenen Maßzahl, der inkrementellen Kosten-Effektivitäts-Relation (iKER), zusammengeführt. Beachtenswert dabei ist die Berechnung und tabellarische Darstellung der Kosten und Effekte bzw. des inkrementellen Kosten-Effekte-Quotienten für die unterschiedlichen Follow-up-Phasen; dabei gehen auf der Kostenseite längerfristige Kostenkomponenten in Form von Folgekosten für „downstream events“ und auf der Effektseite die aussagefähige Effektgröße des „ereignisfreien Überlebens“ (siehe unten) in die Berechnungen ein.

Wie bereits erwähnt, wählte die Autorin als gesundheitliche Effektgröße den klinischen Parameter des von kardialen Ereignissen freien Überlebens („event-free survival“), genauer die Wahrscheinlichkeit, mit der die Postinfarktpatienten die Follow-up-Phasen „ereignisfrei“ erleben können. Dieses aggregierte Effektmaß ist - bei vergleichbarer Wirksamkeit gleichwertiger therapeutischer Versorgung - als adäquater Indikator für die klinische Effektivität einer Diagnostiktechnologie geeignet. Denn der Gesundheitszustand im Follow-up steht über die vom PET- bzw. SPECT-Befund abhängige Entscheidung hinsichtlich der Durchführung einer Revaskularisation in direktem Zusammenhang zur angewandten Diagnosemethode. Indem das Modell den Einfluß der diagnostischen Serviceleistungen PET bzw. SPECT auf die Therapieentscheidung (Revaskularisation ja oder nein) und davon abhängig wiederum den Einfluß auf gewonnene „ereignisfreie Lebensjahre“ modelliert, werden finale Outcomegrößen erfaßt (therapeutic efficacy bzw. patient outcome efficacy, siehe C.4.2.2).

Die Quantifizierung und monetäre Bewertung des Ressourcenverbrauchs weist formale Mängel auf: (a) Bei den Kostenangaben handelt es sich um pro Kopf-Kosten; dies geht jedoch nur aus dem inhaltlichen Zusammenhang und nicht direkt aus der Tabellenbeschriftung hervor, (b) das Mengengerüst wird nicht explizit aufgeführt, ergibt sich jedoch aus der Modellrechnung; das Preisgerüst ist angegeben, jedoch werden die Quellen und Methoden zur monetären Bewertung der medizinischen Interventionen (durch Schätzungen bzw. Bildung von Durchschnittspreisen) nur teilweise aufgeführt und erläutert. So läßt sich aus der Veröffentlichung nicht ersehen, ob der Kostenberechnung aus dem tatsächlichen Ressourcenverbrauch abgeleitete Kosten oder administrative Kosten in Form von Gebühren zugrunde liegen. Auch die Festlegung eines einheitlichen Geldbetrages für die Behandlung unterschiedlichster „downstream events“ erscheint fragwürdig (US\$ 8 225).

In die Kostenberechnung gehen ausschließlich die im Rahmen von Primärinterventionen anfallenden direkten medizinischen Kosten ein. Weitere Kostenfaktoren wie Kosten für akute operative Komplikationen (z.B. für einen Herzinfarkt während der invasiven Therapie) oder Ausgaben für nicht invasiv therapierte Patienten, wie beispielsweise Kosten für medikamentöse Behandlung oder für Gesundheitsprogramme zur Reduktion von Risikofaktoren, und Langzeitkosten (> 5 Jahre) bleiben unberücksichtigt. Indirekte Kosten durch krankheitsbedingten Produktionsausfall sind zu vernachlässigen, da als Bezugsperspektive nicht die gesellschaftliche Sicht, sondern explizit die Sicht des Gesundheitssystems gewählt wird. Dabei geht aus der Publikation nicht hervor, ob die Sicht der Leistungserbringer, hier der Kliniken, oder der Drittfinanzierer vertreten wird.

Alle Geldbeträge sind sowohl in holländischen Gulden, als auch in US\$ ausgewiesen. Der im Referenzjahr 1994 gültige Wechselkurs zur Konversion der Landeswährung in die US-Währung ist angegeben. Preisanpassungen wegen Inflation oder Deflation sind nicht notwendig, da mit Kostenangaben aus dem Jahr 1994 aktuelle Daten verwendet werden (Publikationsjahr 1995). Die Kosten und Effekte werden trotz des auf 5 Jahre projizierten - und somit deutlich über die Diagnosestellung hinausreichenden - Zeitfensters nicht diskontiert.

Die Ergebnisse der Studie und deren Generalisierbarkeit, die Datenqualität, die Bedeutung unsicherer Daten auf den Entscheidungsausgang und sonstige Schwächen der Studie werden nicht diskutiert. Vergleiche mit anderen entscheidungstheoretischen Analysen werden nicht gezogen. Ethische Fragen und Verteilungsfragen werden nicht angesprochen, sind jedoch auch nicht Gegenstand der Aufgabenstellung. Dennoch wäre eine Begründung der Schlußfolgerung, daß PET - bei einem Mehraufwand pro gewonnenem „ereignisfreien“ Lebensjahr von circa US\$ 23 000 im ersten Beobachtungsjahr (aber mit degressivem Verlauf in den folgenden vier Jahren) - der günstigeren SPECT-Methode vorzuziehen sei, angemessen gewesen.

Abschließende Bewertung

Bei der vorliegenden Arbeit handelt es sich um einen ersten Entwurf eines entscheidungstheoretischen Modells zur Rolle der alternativen Diagnoseverfahren PET bzw. SPECT bei der Wahl einer geeigneten Therapie für Myokardinfarktpatienten. Auch wenn dieser Entwurf noch einige formale und methodische Mängel aufweist, so kann der Modellansatz dennoch als gute Ausgangsbasis für künftige Analysen dienen. Da in der jetzigen Form jedoch (a) wesentliche Berechnungsgrundlagen, vor allem die Konditionen des verwendeten Markov-Modells und die monetäre Bewertung des Ressourcenverbrauchs, nicht nachvollziehbar sind, (b) die Analyse insgesamt auf eine schmale Datenbasis zurückgreift und (c) die Stabilität des entscheidungstheoretischen Modells nicht anhand von Sensitivitätsanalysen geprüft wird, sind die Ergebnisse dieser ökonomischen Evaluation nicht ausreichend belegt und momentan nur begrenzt geeignet, gesundheitspolitische Entscheidungen zu unterstützen.

Indikationsstellung: kardiale Perfusionsmessung

Gould KL, Goldstein RA & Mullani NA. Economic analysis of clinical positron emission tomography of the heart with rubidium-82. J Nucl Med 1989; 30(5):707–717

Kommentar von Gleason JM & Frick MP mit anschließender Erwiderung von Gould KL. The expected value of cash flows in the economic analysis of clinical positron emission tomography of the heart. J Nucl Med 1990;31(11):1893–1896

Studienbeschreibung

In der wissenschaftlichen Abhandlung aus dem Jahr 1989 über die klinisch-diagnostische Nutzung der PET im Indikationsgebiet „koronare Herzkrankheit“ (KHK) veröffentlichen die drei US-Autoren ökonomische Modellierungen zur PET-basierten KHK-Diagnostik bei drei unterschiedlichen Patientengruppen: Personen mit manifester KHK, speziell Patienten mit Thoraxschmerzen bzw. Postinfarktpatienten, und die Hochrisikogruppe für latente KHK, 45 - 65 Jahre alte Männer mit einem oder mehreren Risikofaktoren wie Rauchen, Bluthochdruck, Fettstoffwechselstörungen und/oder familiärer Prädisposition für KHK. Als Radiotracer dient im Modell das kurzlebige Isotop Rubidium-82, ein Positronen emittierendes Analogon des Erdalkalimetalls Kalium, das mit Hilfe eines Generators klinikintern erzeugt werden kann. Damit verfügt eine Klinik über eine Radionuklidquelle, die sie vom Zugang zu einem Teilchenbeschleuniger (Zyklotron) unabhängig macht, wodurch der routinemäßige Einsatz der PET für Untersuchungen der Herzperfusion und -funktion ermöglicht wird.

Die Autoren verfolgen in ihrer Publikation zwei Zielsetzungen: im ersten Teil der Studie wird ein kostendeckender Geldbetrag pro PET-Untersuchung ermittelt, um dann im zweiten Teil die Kosten für eine PET-basierte, diagnostische und therapeutische Versorgung der drei KHK-Patientengruppen - explizit in Einzelschritten dargestellt am Beispiel einer fiktiven 100köpfigen Kohorte von Patienten mit Thoraxschmerzen - anhand eines mathematischen Modells kalkulieren zu können. Diese - rein hypothetischen - Versorgungsaufwendungen werden zu den Kosten in Relation gesetzt, die bei Versorgung derselben Patientengruppen unter Einsatz der Routinemethode Myokardperfusionsszintigraphie bei pharmakologischer Belastung, mit dem Kalium-Analogon Thallium-201 als Positronenquelle, zum Zeitpunkt der Studie entstanden wären.

Zur Bestimmung einer - für Ende der 1980er Jahre - kostendeckenden Vergütung pro PET-Untersuchung stellen die Autoren eine 5-Jahres-Rentabilitätsberechnung mit folgenden Vorgaben auf:

1. Eine Untersuchung der Blutversorgung des Herzmuskels bestand aus je einem Ruhe- und einem Dipyridamol-induzierten Belastungstest.
2. Der Anschaffungspreis für den PET-Scanner belief sich damals auf US\$ 2,4 Millionen; davon wurden US\$ 1,94 Millionen auf dem Kapitalmarkt zu einem Zinssatz von 12% aufgenommen; der ^{82}Rb -Generator verursachte laufende Kosten von monatlich US\$ 20 000.
3. Die durchschnittliche tägliche Auslastung des PET-Scanners steigerte sich gemäß den Modellvorgaben von 4 Patienten im ersten Jahr, über 6 Patienten im zweiten Jahr auf eine gleichbleibende Auslastung von 8 Patienten ab dem dritten Jahr.
4. Als potentielle Vergütung pro PET-Untersuchung wurden Beträge von US\$ 1 200 bzw. US\$ 1 000 angesetzt:
5. Sonstige Vorgaben: nicht-einbringbare Forderungen von 20%, Steuersatz von 34%, Inflationsrate von 6% und ein Amortisationsbetrag für den PET-Scanner von US\$ 461 000 pro Jahr.

Die für die Leistungserstellung veranschlagten Aufwendungen eines Wirtschaftsjahres enthalten unter anderem Ausgaben für Gehaltszahlungen an das technische und medizinische Personal (ohne Arzthonorare), Sachkosten für den radioaktiven Marker, Abschreibungsbeträge für das PET-Gerät, Unterhaltungskosten und Gemeinkosten.

Unter Vorgabe eines rein „technischen“ Vergütungsbetrages (ohne Arztkostenanteil) von US\$ 1 200 und einer optimalen täglichen Auslastung mit 8 Patienten ergibt sich laut Rentabilitätsberechnung - unter den Ende der 1980er Jahren aktuellen Konditionen (siehe oben) - nach Abzug von Zinsen und Steuern eine jährliche Rendite von 17% (Berechnungszeitraum: 5 Jahre). Sinkt im Modell die tägliche Auslastung dagegen auf 6 Patienten, verringert sich der interne Zinsfuß auf 7%. Diese Zinsrate entsprach nach Aussage der Autoren zum damaligen Zeitpunkt der Rendite einer Bankspareinlage. Eine Absenkung der Entgeltzahlung pro PET-Untersuchung im Kalkulationsansatz von ursprünglich US\$ 1 200 auf nur US\$ 1 000 hat bei einer Grenzauslastung von 8 Patienten pro Tag wiederum eine Verzinsung von 7% zur Folge; bei gleichzeitiger Abnahme der Auslastung auf nur 6 Patienten pro Tag resultiert bei einem internen Zinsfuß von – 3% ein finanzieller Verlust. Unter der Vorgabe einer PET-Grundvergütung von US\$ 1 200 (ohne Arzthonoraranteil) liegt die Gewinnschwelle (break-even point) im ersten Jahr bei einer Geräteauslastung von durchschnittlich 5 Behandlungsfällen pro Tag, mit einer PET-Grundvergütung von US\$ 1 000 dagegen bei circa 6,5 Fällen pro Tag. Aufgrund dieser Rentabilitätskalkulation empfehlen die Autoren in ihrer Veröffentlichung aus dem Jahr 1989, um Kostendeckung zu erreichen, bei einer Geräteauslastung mit 6 bis 8 Untersuchungen am Tag eine „technische“ Vergütung von US\$ 1 200 pro ^{82}Rb -PET. Die Gesamtkosten pro PET-Untersuchung (je eine Aufnahme im Ruhe- bzw. im Belastungszustand) addieren sich somit zusammen mit dem Arzthonorar von US\$ 300 auf insgesamt US\$ 1 500. Dieser Gesamtvergütungsbetrag wird in einer vergleichenden Bewertung der alternativen Diagnoseverfahren ^{201}TI -Belastungsszintigraphie versus ^{82}Rb -PET im zweiten Teil der ökonomischen Analyse für eine PET-Untersuchung veranschlagt.

Im Evaluationsteil der Studie werden am Beispiel einer fiktiven Kohorte von 100 Patienten mit Brustschmerzsymptomatik die Struktur des theoretischen Versorgungsmodells erläutert und die Gesamtkosten für Diagnostik plus Therapie aufsummiert. So setzen sich die alternativen Versorgungsmodelle jeweils aus einem Diagnosteteil und einem Therapieteil, Revaskularisationsmaßnahmen in Form von perkutaner transluminaler Koronar-Angioplastie (PTCA) oder Bypass-Operation, zusammen. Der Diagnosteteil wiederum unterteilt sich in eine Eingangsuntersuchung mittels ^{82}Rb -PET oder alternativ ^{201}Tl -Belastungsszintigraphie, plus - zur Verifikation eines positiven Eingangstestergebnisses - einer Röntgenkontrastdarstellung zum definitiven Nachweis und zur Lokalisation von Stenosen der Koronararterien (Koronararteriographie, Goldstandard).

In die Modellierung gehen folgende, aus der Literatur entnommenen empirischen Daten bzw. konservativen Schätzungen ein:

1. Prävalenz für KHK bei Probanden mit Thoraxschmerzen: 25% (Literaturwert),
2. Kenndaten der diagnostischen Treffsicherheit der nuklearmedizinischen Alternativmethoden ^{201}Tl -Belastungsszintigraphie und ^{82}Rb -PET (Schätzwerte in Anlehnung an publizierte Literaturdaten): Sensitivität (^{201}Tl -Szintigraphie: 85% / ^{82}Rb -PET: 95%; Literaturwerte nach Angabe der Autoren: 70 - 85% für ^{201}Tl -Szintigraphie bzw. 95 - 98% für ^{82}Rb -PET) und Spezifität (^{201}Tl -Szintigraphie: 60% / ^{82}Rb -PET: 95%, Literaturwerte nach Angabe der Autoren: 50 - 70% für ^{201}Tl -Szintigraphie bzw. 95 - 98% für ^{82}Rb -PET),
3. Kosten im Rahmen der Diagnosestellung (einschließlich Personalkosten): US\$ 950 pro ^{201}Tl -Myokardszintigramm bzw. US\$ 1 500 pro ^{82}Rb -PET-Untersuchung (siehe oben), US\$ 6 000 pro Koronararteriographie,
4. 85% der Patienten mit gesicherter KHK (das heißt, bei positiver Koronararteriographie) werden als therapierbar mittels PTCA (US\$ 10 000) oder Bypass-Operation (US\$ 20 000) eingestuft (im Verhältnis 1 : 1).

Anhand der Modellrechnung beläuft sich - bezogen auf die ausgehenden 1980er Jahre - der monetär bewertete Ressourcenverbrauch für medizinische Interventionen im Rahmen der Diagnose und Therapie, also die direkten medizinischen Kosten, (a) für das PET-basierte Management des 100köpfigen Patientenkollektivs auf insgesamt circa US\$ 610 000 (bei 27,6 PET-positiven Patienten*¹⁰, davon 3,8 falsch-positiven Fällen, und jeweils 10 PTCA's bzw. Bypass-Operationen) und (b) für das auf ^{201}Tl -Belastungsszintigraphie basierende Patientenmanagement auf insgesamt circa US\$ 673 000 (bei 51,3 Fällen mit positivem ^{201}Tl -Szintigramm, wovon sich nur 21,3 Fälle als richtig-positiv, aber 30 Fälle als falsch-positiv erweisen, und jeweils 9 PTCA's bzw. Bypass-Operationen).

¹⁰ In der Originalveröffentlichung von Gould et al. werden wegen eines Rechenfehlers irrtümlich 26,6 anstatt 27,6 PET-positive Patienten angegeben. In diesem HTA-Bericht werden diesbezüglich korrigierte Kosten berichtet. Die korrigierten, von der Originalliteraturstelle abweichenden Kostangaben sind mit einem * gekennzeichnet.

Die - trotz des geringeren medizinischen Nutzens von nur 18 gegenüber 20 richtig erkannten und therapierten KHK-Patienten - um circa US\$ 57 000* höheren, direkten medizinischen Gesamtkosten für die ^{201}TI -Szintigraphie basierte Versorgung im Vergleich zur PET-basierten Versorgung sind vor allem auf Zusatzaufwendungen für Koronararteriographien zurückzuführen. Diese ergeben sich aus der Überprüfung der per ^{201}TI -Szintigraphie zusätzlich als falsch positiv eingestuften 26,2 Patienten (30 gegenüber nur 3,8 falsch-positiven Fällen). Diese Mehraufwendungen von rund US\$ 142 000* konnten durch Minderausgaben von insgesamt US\$ 85 000, bedingt durch die per se kostengünstigere ^{201}TI -Diagnostik und die Unterlassung von invasiven Therapiemaßnahmen in zwei Fällen, nicht vollständig kompensiert werden.

Ergänzend zu den direkten medizinischen Kosten für diagnostische und therapeutische Interventionen nehmen die Autoren auch indirekte Kosten in die Kostenberechnung mit auf. Dabei handelt es sich um gesellschaftliche Verluste durch Produktivitätsausfälle bei vorzeitigem Tod von Patienten mit unbehandelter KHK, deren Erkrankung aufgrund falsch-negativer Befundung im nicht-invasiven Eingangstest nicht erkannt wurde. Als Maß für das Mortalitätsrisiko von Patienten mit unbehandelter, schwerer KHK wählen die Verfasser der Studie eine Mortalitätsrate von 7% (Streuung der Literaturwerte laut Autoren: 6 - 25%). Gemäß Berechnung der Autoren belaufen sich - Ende der 1980er Jahre - die indirekten Kosten durch Lohnausfall (einschließlich US\$ 5 000 Sterbekosten pro verstorbenem Patienten), bei einem durchschnittlichen jährlichen Verlust von US\$ 30 000 pro Todesfall, bei 1,31 Verstorbenen innerhalb von 5 Jahren nach Diagnosestellung mittels ^{201}TI -Myokardszintigraphie auf einen Betrag von rund US\$ 202 000; die bei der PET-Diagnostik induzierten Kosten betragen bei nur 0,44 Verstorbenen im gleichen Zeitraum dagegen nur circa US\$ 68 000. Zusammen mit den direkten medizinischen Kosten hätte folglich das fiktive 100köpfige Studienkollektiv Gesamtkosten von fast US\$ 684 000* bei ^{82}Rb -PET-basierter Diagnosestellung (Ruhe- und Belastungsscan) gegenüber annähernd US\$ 874 000 bei Eingangsbefundung mittels ^{201}TI -Belastungsszintigraphie verursacht.

Aufgrund der vorgestellten Modellrechnung ergibt sich somit folgendes Bild: Obwohl die Kosten pro ^{82}Rb -PET-Untersuchung mit US\$ 1 500 gegenüber nur US\$ 950 pro ^{201}TI -Myokardszintigramm zu Buche schlagen, resultieren im Modell - aufgrund der besseren diagnostischen „Ausbeute“ der ^{82}Rb -PET - um US\$ 190 000* niedrigere Gesamtkosten für die Versorgung von 100 Patienten mit Thoraxschmerzen in Relation zur üblichen Methode, der ^{201}TI -Myokardperfusionsszintigraphie unter Belastung (Bezugszeitraum Ende der 1980er Jahre).

Die Autoren verweisen in der vorliegenden Publikation auf vergleichbare Modellierungen zu zwei weiteren Anwendungsgebieten der PET: dem Screening von Patienten mit schwerer koronarer Herzkrankheit (KHK, Hochrisikogruppe der asymptomatischen Männer mittleren Alters mit einem oder mehreren Risikofaktoren für koronare Arteriosklerose) und die Beurteilung der Lebensfähigkeit des Myokards nach einem Myokardinfarkt oder einer Thrombolyse. Nach Aussage der Autoren bestätigen die

nicht im Detail vorgestellten Ergebnisse die bei Brustschmerzpatienten festgestellte diagnostische bzw. ökonomische Überlegenheit der PET über die übliche ^{201}Tl -Myokardperfusionsszintigraphie unter Belastung bzw. über die Alternative „Nichtstun“.

Da in der Veröffentlichung nur die Modellierung der Versorgung von Patienten mit ungeklärter Brustschmerzsymptomatik detailliert vorgestellt wird, während die beiden anderen Einsatzgebiete sehr knapp abgehandelt werden, bezieht sich die weitere Beurteilung der Studie ausschließlich auf die ökonomische Analyse zur Befundung von Thoraxschmerzpatienten.

Bewertung der Modellierung für Patienten mit Thoraxschmerzen

Die vorliegende Modellierung aus dem Jahr 1989 erfüllt aus gesundheitsökonomischer Sicht nicht die Bedingungen einer vollständigen ökonomischen Evaluation, da nur ein Kostenvergleich zwischen zwei Alternativtechnologien durchgeführt wird, während echte Gesundheitseffekte (patient outcome efficacy) unberücksichtigt bleiben. Vielmehr geht eine klinische Effektgröße von untergeordneter Aussagekraft, die von der Treffsicherheit der Diagnoseverfahren abhängige Änderung im Patientenmanagement (Anzahl der diagnostischen und therapeutischen Leistungen, diagnostic thinking efficacy und therapeutic efficacy, siehe C.4.2.2), als monetär bewertetes Effektmaß auf der Kostenseite in das mathematische Modell ein. Um die Problematik der Abgrenzung von Kosten und Effekten zu umgehen, hätten die Autoren die aufgrund der diagnostischen Überlegenheit von ^{82}Rb -PET gegenüber der Routinemethode ^{201}Tl -Szintigraphie zusätzlich richtig diagnostizierten und therapierten (theoretischen) 2,2 Patienten (20,2 gegenüber 18 von 25 an KHK erkrankten Personen) als klinische Effektgröße auswählen und in einer inkrementellen Kosten-Effektivitäts-Analyse mit den - direkten und indirekten - Einsparungen in Höhe von US\$196 500 in Beziehung setzen können (inkrementelle Kosten-Effektivitäts-Ratio: US\$ - 89 318 pro zusätzlich adäquat versorgtem Patienten).

Die in zwei Teile untergliederte Aufgabenstellung der Studie - Kalkulation einer kostendeckenden ^{82}Rb -PET-Vergütung einerseits und vergleichende ökonomische Analyse für ^{82}Rb -PET versus ^{201}Tl -Myokardszintigraphie andererseits - wird in der Publikation präzise dargelegt und konsequent verfolgt. Sowohl der medizinische als auch der ökonomische Problemkontext werden eingehend erläutert. Eine Zielpopulation wird definiert (Patienten mit Schmerzen im Brustbereich), aber nicht näher spezifiziert (hinsichtlich Alters- und Geschlechterverteilung, KHK-relevante Vorerkrankungen, Risikofaktoren usw.). Die Perspektive der Studie ist nicht explizit genannt, durch den Einbezug indirekter Kosten in die Gesamtkostenberechnung wird jedoch die gesellschaftliche Sichtweise vertreten. Die Abfolge der aus jeweils einem Diagnose- und einem Therapieteil zusammengesetzten alternativen Versorgungsalgorithmen (siehe Studienbeschreibung), die sich nur hinsichtlich der Eingangsuntersuchung (mittels ^{201}Tl -Myokardszintigraphie oder ersatzweise mittels ^{82}Rb -PET) unterschei-

den, ist leider nicht in Form eines Ablaufschemas illustriert, wird jedoch aus der Textbeschreibung deutlich. Die Auswahl der beiden Alternativtechnologien wird nicht begründet. Diese ergibt sich jedoch aus der damaligen Bedeutung der untersuchten nuklearmedizinischen Technologien in der klinischen Diagnostik: so eroberte Ende der 1980er Jahre die PET zunehmend - an großen medizinischen Zentren - einen festen Platz in der Diagnostik und Therapiekontrolle kardialer Erkrankungen, gehörte jedoch wegen des beschränkten Angebots an PET-Geräten nicht zu den kardiologischen Routinemethoden, während die Belastungsszintigraphie mit ²⁰¹Thallium als Perfusionsmarker zu diesem Zeitpunkt bereits routinemäßig in der Kardiologie eingesetzt wurde.

Ihr mathematisches Modell zur Abschätzung der beim Management von KHK-Patienten verbrauchten Ressourcen entwickelt die US-Arbeitsgruppe gut nachvollziehbar Schritt für Schritt. Die zur Modellierung eines Versorgungsszenarios notwendigen Vorgaben und Annahmen - diagnostische, epidemiologische und klinische Daten aus der Literatur und konservative Schätzungen - sind ausreichend dokumentiert; die verwerteten Literaturdaten werden jedoch leider nur teilweise durch die Angabe von Quellen belegt. Die Streubreiten der diagnostischen Kennwerte sind für beide Testmethoden angegeben, werden jedoch nicht zur Überprüfung der Stabilität des Modells anhand von Sensitivitätsanalysen genutzt.

Durch die Unterteilung der Kostenberechnung in für den Leser einfach nachvollziehbare Einzelschritte wird das Mengengerüst, das von den verwendeten epidemiologischen und klinischen Kennwerten abhängt, gut deutlich. Auch das der Kostenberechnung zugrundeliegende Preisgerüst (Kosten pro Einzelleistungen) ist angegeben. Wie die Vergütungen für die Einzelleistungen ermittelt wurden, geht jedoch aus der Publikation nicht deutlich hervor; so ist die Art (Gebühren oder Marktpreise ?) und Zusammensetzung (Sach-, Personal- und Gemeinkosten ?) der veranschlagten Kostenposten nur unvollständig aufgeführt. Eine Ausnahme bildet dabei die sehr detaillierte, nach Kostenkomponenten aufgeschlüsselte Berechnung der PET-Vergütung anhand einer detaillierten betrieblichen Kostenkalkulation (Einzelheiten dazu siehe Studienbeschreibung). Ebenfalls hervorzuheben ist, daß der Einfluß der unterschiedlichen diagnostischen Treffsicherheit der Alternativmethoden auf die indirekten Kosten durch Produktionsverluste im Todesfall berücksichtigt wird, wodurch - wenn notwendigerweise auch nur ansatzweise - der Evaluationsrahmen in Richtung einer ökonomischen Technologiefolgenabschätzung auf gesellschaftlichem Niveau ausgeweitet wird. Weitere direkte und indirekte (Langzeit)folgekosten bei gesicherter KHK wären unter anderem: Nachbehandlungskosten, kurz- und langfristige Kosten für Komplikationen, die in Zusammenhang mit einer invasiven Revaskularisation auftreten können (z.B. durch Verletzung der Koronararterien oder bei einem perioperativen Herzinfarkt), und krankheitsbedingte Produktionsausfallkosten, beispielsweise wegen zeitweiser Arbeitsunfähigkeit.

Kritisch anzumerken an der ökonomischen Evaluation sind folgende Punkte:

1. Der von Gould et al. verwandte Modus zur Berechnung der indirekten Kosten impliziert, daß alle unbehandelten KHK-Patienten (falsch-negative Probanden) bereits zu Beginn des 5-Jahres-Follow-up versterben, so daß der Lohnverlust pro Todesfall der Maximalsumme von US\$ 150 000 (5 x US\$ 30 000) entspricht, wodurch die indirekte Kostenkomponente unzulässig in die Höhe getrieben wird. In ihrer Stellungnahme bemängeln Gleason & Frick (1990) diese methodische Schwäche der Studie und legen eine korrigierte Kostenaufstellung vor. So reduzieren sich - bei einer Gleichverteilung der Todesfälle über den 5-Jahres-Zeithorizont - die indirekten Kosten für ^{82}Rb -PET von US\$ 68 200 auf US\$ 41 600 und für ^{201}Tl -Szintigraphie von US\$ 201 500 auf US\$ 125 700, und somit die Differenz von ursprünglich US\$ 133 300 auf nunmehr nur noch US\$ 83 100.
2. Obwohl der auf 5 Jahre projektierte Zeithorizont (Follow-up-Phase) deutlich über die reine Diagnosestellung hinausreicht, werden die Kosten weder diskontiert noch inflations-/deflationsbereinigt (außer bei der Kalkulation einer kostendeckenden PET-Vergütung: Inflationsrate 6%).
3. Richtung und Größe des Einflusses unsicherer Modellannahmen und verzerrter Parameter auf den Entscheidungsausgang, wie beispielsweise die Streubreite für die diagnostische Genauigkeit der bildgebenden Verfahren, werden nicht anhand von Sensitivitätsanalysen überprüft.

In der insgesamt kurzen Zusammenfassung der Studienergebnisse wird weder die Qualität der aus der Literatur übernommenen (z.B. diagnostische und epidemiologische Kenndaten) noch der von den Autoren selbst erhobenen (z.B. das Preisgerüst für medizinische Leistungen) oder errechneten Daten kritisch hinterfragt, ebensowenig werden die Einschränkungen und Schwächen der Studie und die Ergebnisse im Zusammenhang mit anderen Studien diskutiert. Auch ethische Fragen und Verteilungsaspekte werden von den Autoren nicht angesprochen, sind aber auch nicht Gegenstand der Aufgabenstellung.

Die Schlußfolgerungen der Autoren leiten sich in konsistenter Weise aus den berichteten Ergebnissen ab. Denn unter den vorgegebenen Rahmenbedingungen ist die PET die dominante - weil kostensparende und diagnostisch effektivere - Diagnosemethode bei der Versorgung von KHK-Patienten mit Brustschmerzsymptomen und wäre somit aus dem Blickwinkel der Gesundheitsökonomie betrachtet die Methode der Wahl.

Abschließende Bewertung

Die Modellierung zur vergleichenden Abschätzung und monetären Bewertung des Ressourcenverbrauchs bei der Versorgung von Brustschmerzpatienten mittels ^{201}Tl -Belastungsszintigraphie versus ^{82}Rb -PET zeichnet sich durch die Erweiterung des Zeithorizontes über die akute Interventionsphase hinaus (5jährige Follow-up-Phase)

und durch den Einbezug von indirekten Kostenkomponenten in die Gesamtkostenberechnung aus. Die Studie ist aufgrund einer detailliert nach Einzelkomponenten aufgeschlüsselten, betrieblichen Kosten- und Rentabilitätsberechnung für den klinischen Einsatz der PET (mit ^{82}Rb als Positronenstrahler) nach wie vor beispielhaft und kann für künftige Kosteneffektivitätsstudien als gute methodische Vorlage zur Erfassung des tatsächlichen Ressourcenverbrauchs bei PET-Untersuchungen dienen, auch wenn die bereits 1989 publizierte Beschreibung der Kostenseite beim Management von KHK-Patienten in den USA nicht mehr die aktuelle Vergütungs- und Kostensituation widerspiegeln dürfte. Aufgrund fehlender Aktualität (auch bezüglich des technischen Entwicklungsstandes, Vergleichsmethode: ^{201}Tl -Szintigraphie) und methodischer Einschränkungen der Bewertung der gesundheitsökonomischen Wertigkeit der PET ist die Aussagekraft der Studie zur Unterstützung anstehender gesundheitspolitischer Entscheidungen jedoch begrenzt.

Patterson RE, Horowitz SF & Eisner RL. Comparison of modalities to diagnose coronary artery disease. *Semin Nucl Med* 1994;24(4):286-310.

Patterson RE, Eisner RL & Horowitz SF. Comparison of cost-effectiveness and utility of exercise ECG, single photon emission computed tomography, positron emission tomography, and coronary angiography for diagnosis of coronary artery disease. *Circulation* 1995;91(1):54-65

Abweichend von der sonstigen Vorgehensweise werden die zwei vorzustellenden Veröffentlichungen der US-Autorengruppe als Einheit betrachtet und auch als solche nach dem Transparenz- und Qualitätskatalog beurteilt und besprochen, da sie inhaltlich sehr eng zusammengehören. Die Veröffentlichung aus dem Jahr 1994 berichtet in komprimierter Form die Ergebnisse einer Kosten-Nutzwert-Studie, die in der nachfolgenden Publikation von 1995 in ihrer Gesamtheit dargestellt und durch eine Kosten-Effektivitäts-Studie ergänzt wird. Aufgrund der Informationsfülle der Studien kann die zusammenfassende Beschreibung nur auf die zentralen Punkte und Ergebnisse der beiden Artikel eingehen. An Detailinformationen interessierte Leser werden auf die zwei Originalartikel verwiesen.

Studienbeschreibung

Für vier alternative Vorgehensweisen zur Diagnose von obstruktiver koronarer Herzkrankheit (KHK) werden - in Abhängigkeit von verschiedenen Vortestwahrscheinlichkeiten ($p = 0,1; 0,3; 0,5$ und $0,9$), denen nach Alter und Geschlecht differenzierte Subpopulationen in der Bevölkerung zuzuordnen sind - vergleichende Kosten-Effektivitäts- und Kosten-Nutzwert-Analysen angestellt. Dazu wird ein mathematisches Modell, das auf dem Bayes-Theorem aufbaut, verwendet. In die Analyse einbezogen sind die nicht-invasiven Routinetechnologien Belastungs-EKG (ExEKG) und Single-Photon-Emissions-Computer-Tomographie (SPECT), die moderne Emissions-Computer-Tomographie mit Positronen (PET) und die Koronarangiographie (Goldstandard für den Nachweis von Stenosen der Koronararterien und deren Lokalisation). Miteinander verglichen werden drei alternative, zweistufige Diagnosealgorithmen aus je einem nicht-invasiven Eingangstest, wahlweise ExEKG, SPECT oder PET, der mit einer zusätzlichen Angiographie zur Verifikation des Erstbefundes kombiniert wird, wenn der Befund des Eingangstests positiv ausfällt oder nicht auswertbar ist. Die vierte Diagnosestrategie besteht ausschließlich aus einer Koronarangiographie.

Das mathematische Modell ist so konstruiert, daß alle unmittelbar im Verlauf der Diagnosefindung anfallenden Kosten für diagnostische und therapeutische Interventionen einbezogen sind, einschließlich der Aufwendungen für die Angiographie zur diagnostischen Abklärung der im Eingangstest falsch-positiven Probanden (induzierte Kosten); zuzüglich angerechnet werden Aufwendungen für Komplikationen wie myo-

kardialer oder cerebraler Infarkt, die während der diagnostischen Tests und/oder im Verlauf einer 10jährigen Follow-up-Phase bei im Eingangstest falsch negativ diagnostizierten KHK-Patienten als Folge unterlassener Behandlung auftreten können (durch Fehldiagnose induzierte Kosten). Dazu wurden aus der Literatur folgende ökonomischen, epidemiologischen und klinischen Daten gesammelt: u.a. Testgebühren, Mortalitäts- und Komplikationsraten der diagnostischen und therapeutischen Interventionen, Mortalitätsrate in Abhängigkeit vom Therapiemodus (Medikation, PTCA, Bypass-Operation) und die diagnostischen Kenndaten der einzelnen Diagnostiktechnologien (ExEKG, SPECT, PET und Angiographie).

Die klinische Wirksamkeit dieser Diagnosestrategien, definiert als die Anzahl an Patienten mit korrekt diagnostizierter KHK, wird berechnet (diagnostic accuracy efficacy). Diese intermediäre Outcomegröße (siehe C.4.2.2) wird als Effektgröße in die Kosten-Effektivitäts-Analyse eingesetzt.

Um zudem den Einfluß klinischer Variablen auf den medizinischen Outcome der KHK-Versorgungsstrategien anhand der Lebensdauer und Lebensqualität modellieren zu können, schätzen die Autoren basierend auf Literaturdaten die durch eine erfolgreiche Therapie bedingte Verlängerung der Lebenserwartung ab (Bezugszeitraum: 10jähriges Follow-up, in Relation zu „Nichtbehandlung“). Dieser Zugewinn an Lebenszeit wird durch Bewertung der Lebensqualität mit Präferenzen in qualitätsadjustierte Lebensjahre (QALYs) umgewandelt. Um den Einfluß des Komplikations- bzw. Mortalitätsrisikos während der diagnostischen Interventionen und bei unerkannter KHK auf die Lebenserwartung und Lebensqualität im Modell berücksichtigen zu können, wird die finale Outcomegröße „qualitätsadjustierte Lebenserwartung“ hinsichtlich komplikations- und mortalitätsinduzierter Verluste korrigiert. Diese komplikationskorrigierte Effektgröße bildet den Nenner in der Kosten-Nutzwert-Analyse. Im Modell wirksame klinische Einflußfaktoren auf den Outcome sind die unterschiedlichen Therapiemodi, wie medikamentöse Behandlung oder Revaskularisation mittels PTCA bzw. operativem Bypass der Herzkranzgefäße.

Anhand der vorgestellten Modellierung ergeben sich eine Fülle von Ergebnissen, die hier nur auszugsweise dargelegt werden können. Die Wahrscheinlichkeit für einen Patienten mit Verdacht auf KHK, auch tatsächlich an KHK erkrankt zu sein (Vortestwahrscheinlichkeit pCAD), hängt von seiner klinischen Krankengeschichte und seiner persönlichen Belastung mit Risikofaktoren ab. Steigt diese Wahrscheinlichkeit im Modell an, nehmen bei den zweistufigen Diagnosestrategien (nicht-invasive Vortests ExEKG, SPECT bzw. PET kombiniert mit einer Angiographie) sowohl die direkten Kosten pro untersuchtem Patienten als auch die Mortalitätsrate stetig zu, während beide Parameter für den alleinigen Einsatz der Angiographie (ohne weitere nicht-invasive Zusatzuntersuchung) nahezu gleich bleiben. Dagegen nehmen die Kosten pro korrekt diagnostiziertem KHK-Patienten und die Kosten pro gewonnenem, komplikationskorrigiertem und qualitätsadjustiertem Lebensjahr für alle vier Diagnoseansätze ab; das heißt, das Kosten-Effektivitäts- bzw. das Kosten-Nutzwert-Verhältnis verbessert sich mit steigender Vortestwahrscheinlichkeit pCAD.

Anhand der Modellberechnungen ist für die gesamte pCAD-Spanne (0,1 – 1,0) die PET unter Belastung den Alternativtechnologien SPECT und ExEKG - in dieser Reihenfolge - und für pCAD-Werte¹¹ < 0,6* auch der Angiographie als Eingangstest vorzuziehen, da der PET-Diagnosemodus für diese pCAD-Werte sowohl den kleinsten Kosten-Effektivitäts- als auch den kleinsten Kosten-Nutzwert-Quotienten aufweist. Über dem Schwellenwert von 0,6* übersieht PET, wie auch die anderen nicht-invasiven Testmethoden, so viele KHK-Patienten (falsch-negativer Befund), daß mit steigender Komplikationsrate die Kosten pro Effektgröße ebenfalls ansteigen. Bei hohen Werten für die Erkrankungswahrscheinlichkeit (> 0,6*) ist gemäß den Ergebnissen von Patterson et al. (1995) die alleinige Angiographie, trotz der hohen Gebühr von US\$ 4 800, die Methode der Wahl, gefolgt von den zweistufigen Algorithmen PET/Angiographie (Gebühr für PET: US\$ 1 800), SPECT/Angiographie und ExEKG/Angiographie. Im Vergleich zu PET sind die beiden anderen nicht-invasiven Verfahren, SPECT und vor allem ExEKG, aufgrund ihrer geringeren diagnostischen Treffsicherheit und der daraus resultierenden größeren Zahl an nachgeschalteten Angiographien (bei falsch-positiven Probanden) bzw. deutlich erhöhter Komplikations- und Mortalitätsrate (bei falsch-negativen KHK-Fällen) als Eingangstests nicht konkurrenzfähig. Diese pCAD-abhängige Rangfolge ändert sich auch bei Variation einzelner oder mehrerer Modellparameter (uni- und multivariate Sensitivitätsanalysen) nicht: bei Variation der Modellkonditionen durch (a) Absenkung der Angiographiekosten, (b) Verschlechterung der diagnostischen Genauigkeit für alle drei nicht-invasiven Vortestverfahren, (c) Anhebung der PET-Kosten, (d) Reduktion der PET-Genauigkeit oder (e) Halbierung des Modellparameters „qualitätsadjustierte Lebensjahre“ treten zwar Veränderungen auf, ohne jedoch die Rangfolge der Testalgorithmen untereinander substantiell zu verändern.

Auf einzelne Patientengruppen mit unterschiedlicher pCAD übertragen, ergibt sich für die Entscheidung, welche Strategie zur Diagnose von KHK in der klinischen Praxis Anwendung finden sollte, nach Einschätzung der Autoren folgendes Bild: bei Männern mittleren Alters und bei älteren Frauen mit typischer Angina pectoris¹² ist aufgrund der hohen Wahrscheinlichkeit, an koronarer Herzkrankheit zu leiden - sowohl unter medizinischen als auch ökonomischen Gesichtspunkten - die sofortige Durchführung einer Angiographie angeraten, während für Patienten mit geringerem Risiko, beispielsweise bei untypischer Angina pectoris, eher eine PET-basierte Diagnosestrategie in Frage kommt. Weiterhin ergibt sich aus der theoretischen Analyse, daß diagnostisches Testen auf KHK bei symptomfreien Personen im allgemeinen nicht

¹¹ Die mit * gekennzeichneten pCAD-Werte sind aus Abbildung 1 der Originalliteratur Patterson et al. 1995 entnommen.

¹² Typische Angina pectoris ist gekennzeichnet durch die drei charakteristischen Brustschmerzsymptome: substernale Lokalisation, durch Belastung auslösbarer Brustschmerz und Abklingen des Schmerzes bei Ruhe oder Nitroglyceringabe innerhalb von 10 Minuten. Bei atypischer Angina pectoris sind nur zwei der drei charakteristischen Symptome vorhanden (Maddahi & Gambhir 1997).

kosteneffektiv ist, da im Vergleich zu den aufzubringenden Kosten nur ein geringer klinischer Nutzen resultiert. Die Autoren heben in ihrer Veröffentlichung ausdrücklich hervor, daß diese Schlußfolgerungen nicht in klinische Richtlinien umgesetzt werden sollten, ehe sie nicht in klinischen Studien überprüft wurden.

Studienbewertung

Die vorliegende ökonomische Evaluation ist sowohl unter inhaltlichen als auch formalen Gesichtspunkten insgesamt als gut zu beurteilen.

Der Evaluationsrahmen der theoretischen Modellierung ist weitgehend abgesteckt. So wird die Aufgabenstellung von den Autoren explizit auf einen Kosten-Effektivitäts- bzw. einen Kosten-Nutzwert-Vergleich zwischen den vier alternativen Diagnosetechnologien Belastungs-EKG, SPECT, PET und Angiographie beschränkt, so daß es sich bei der vorliegenden Studie um eine reine Diagnoseevaluation handelt. Kosten für therapeutische Interventionen oder Folgekosten, die nicht direkt auf diagnosebedingte Komplikationen oder Fehlklassifikationen zurückzuführen sind, bleiben in der Kostenberechnung außen vor.

Der Zeithorizont ist durch den Einbezug einer 10jährigen Follow-up-Phase weiter gespannt als bei den meisten der im Rahmen dieses HTA-Berichtes vorgestellten Studien. Als Perspektive der ökonomischen Evaluation wird von den Autoren explizit die Sicht der Leistungserbringer gewählt. Da zur Kalkulation der direkten Diagnosekosten jedoch Gebühren, also administrative Preise, verwendet werden, erfolgt die Kostenbetrachtung streng genommen aus der Sicht eines Drittfinanzierers. Die Diagnosealgorithmen werden im Hauptartikel aus dem Jahr 1995 übersichtlich anhand einer tabellarischen Aufstellung erläutert und ihre Bausteine, die einzelnen Diagnosetechnologien, werden zwar nicht im Hauptartikel (1995), aber detailliert mit ihren klinischen Vor- und Nachteilen im ergänzenden Übersichtsartikel (1994) beschrieben. Die Auswahl der nicht-invasiven diagnostischen Vortests, ExEKG, SPECT und PET, und der Angiographie als Verifikationsverfahren spiegelt die Palette der heute zur Verfügung stehenden und - mit Ausnahme von PET - im klinischen Alltag üblicherweise eingesetzten diagnostischen Verfahren in der Kardiologie wider.

Die mathematischen Formeln zur Berechnung der Kosten-Effektivitäts- bzw. der Kosten-Nutzwert-Ratios werden in einem Anhang vorgestellt und gut nachvollziehbar erläutert. Alle für die Modellberechnungen notwendigen Parameter des Basisszenarios, wie diagnostische Kenndaten der einzelnen Testverfahren, Mortalitäts- und Komplikationsraten und Kostenangaben für diagnostische Einzelleistungen, sind in einer klar geordneten Tabelle - jedoch ohne Angabe von Streubreiten - zusammengestellt. Als Datenquellen (Ausnahme: Gebührenangaben) dienten Fachveröffentlichungen, die im Literaturverzeichnis dokumentiert sind. Die bei der Literatursuche angewandte Suchstrategie bzw. das Ergebnis der Literatursuche wird in beiden Artikeln nicht erwähnt. Für Einzelparameter (z.B. Erhöhung oder Senkung der Kosten für

diagnostische Einzelleistungen) oder Parameterkombinationen (z.B. der biometrischen Kenndaten), die in uni- bzw. multivariaten Sensitivitätsanalysen variiert werden, sind in einer gesonderten Tabelle sowohl die ursprünglichen als auch die variierten Werte aufgelistet.

Durch Erfassung der „Anzahl an korrekt diagnostizierten KHK-Patienten“ (diagnostic accuracy efficacy) einerseits und der therapieinduzierten zusätzlichen QALYs (patient outcome efficacy) andererseits wählen die Autoren eine intermediäre und eine finale Effektgröße von unterschiedlicher Aussagekraft entsprechend dem unter C.4.2.2 erläuterten hierarchischen Modell. Die Gesundheitseffektdaten ergeben sich entweder direkt aus der diagnostischen Leistungsfähigkeit der Alternativtechnologien (intermediäre Effektgröße der richtig diagnostizierten KHK-Fälle) oder werden, wie in der Studienbeschreibung ausgeführt, unter Zuhilfenahme von Literaturdaten abgeschätzt (kompliationskorrigierte QALYs). Zur Berechnung der qualitätsadjustierten Lebensjahre der therapierten KHK-Patienten in Relation zu den nicht-therapierten KHK-Patienten werden die Gesundheitszustände mit Präferenzen (Literaturdaten) bewertet. Leider sind weder die Methoden und Meßinstrumente zu dieser Bewertung noch die Literaturquellen angegeben. Trotz dieses formalen Mangels ist jedoch anzuerkennen, daß überhaupt der Einfluß der diagnostischen Serviceleistung (verbunden mit einer Therapie) auf den Patientenoutcome, erfaßt als Lebensqualität und Lebensdauer, in den Modellansatz einbezogen ist.

Anhand des mathematischen Modells und den vorgegebenen Modellparametern ergibt sich das Mengengerüst, das in den Artikeln im Einzelnen nicht aufgeführt ist. Da der interessierte Leser dank der ausführlichen Dokumentation des Modells und der zugehörigen Parameter das Mengengerüst selbst rekonstruieren kann, ist diese Unterlassung nicht zu beanstanden. Das Preisgerüst für die diagnostischen Einzelleistungen ist vollständig angegeben, wobei administrative Kosten in Form von nicht näher spezifizierten Gebühren verwendet werden. Die Literaturquellen für die Kostangaben, Testgebühren und durchschnittliche Aufwendungen für interventions- und durch Fehldiagnose induzierte Komplikationen, sind nicht explizit in der Referenzliste aufgeführt. Auf der Kostenseite ist hervorzuheben, daß die Durchschnittskosten für ein „kardiales (oder cerebrales) Ereignis“ während der Testdurchführung oder im Verlauf des auf 10 Jahre projektierten Follow-up (bei falsch negativ diagnostizierten KHK-Fällen) eine indirekte Kostenkomponente in Form von krankheitsbedingten Produktionsausfallkosten beinhalten, die jedoch nicht gesondert ausgewiesen wird. Da im Verlauf des Follow-up anfallende Kosten einbezogen werden, ist das Zeitfenster für die Kosten- und für die Effektseite zeitlich einheitlich determiniert.

Mit der Berechnung von - durchschnittlichen - absoluten Kosten pro richtig erkanntem KHK-Fall (Kosten-Effektivitäts-Quotient) bzw. pro gewonnenem QALY (Kosten-Nutzwert-Quotient) werden sinnvolle Maßzahlen für die Relation zwischen Kosten und Gesundheitseffekten angegeben. Eine inkrementelle Analyse, welche die zusätzlichen Kosten und zusätzlichen Effekte untersucht, die durch eine Intervention im Vergleich mit einer Alternative verursacht werden, wird in der Studie nicht angestellt.

Trotz des weit über die reine Diagnosestellung hinaus verlängerten Zeithorizontes von 10 Jahren werden die Kosten nicht inflationsbereinigt bzw. die Kosten und Effekte nicht diskontiert. Der Verzicht auf eine Diskontierung wird von den Autoren nicht begründet, läßt sich jedoch - für die Kostenseite - durch den vermutlich geringen Anteil der längerfristigen Folgekosten bei Nichtbehandlung falsch-negativer KHK-Patienten an den Gesamtkosten rechtfertigen. Da jedoch die finalen Effektgrößen Lebenserwartung und qualitätsadjustierte Lebenserwartung auf eine Follow-up-Phase von 10 Jahren projiziert sind, wäre eine Diskontierung dennoch angemessen gewesen. Das Referenzjahr für die Kosten ist nicht datiert.

Um die Stabilität der Modellierung gegenüber unsicheren Parametern zu testen, werden acht verschiedene uni- und multivariate Sensitivitätsanalysen ausgeführt. Durch Variation der Gebühren für diagnostische Einzelleistungen (univariate Sensitivitätsanalysen) und der diagnostischen Kennzahlen (Annahme von geringerer Treffsicherheit der nicht-invasiven Verfahren), durch Reduktion des postdiagnostischen Morbiditäts- bzw. Mortalitätsrisikos und durch Halbierung der qualitätsadjustierten Lebenserwartung werden relevante Parameter in geeigneter Weise verändert.

Die vorliegende Evaluation zur Versorgung von KHK-Patienten (speziell die Veröffentlichung aus dem Jahre 1995) zeichnet sich durch eine umfassende, klar gegliederte und deshalb gut verständliche Darstellung sowohl der Modellstruktur als auch der Ergebnisse aus. Zum besseren Verständnis werden die Ergebnisse, speziell auch die Resultate der uni- und multivariaten Sensitivitätsanalysen, in Schaubildern graphisch aufbereitet. Die Ergebnisse werden ausschließlich in disaggregierter Form, das heißt als absolute Kosten pro Kopf oder pro komplikationskorrigiertem QALY, präsentiert, teilweise differenziert nach Risikoeinstufung der Patienten entsprechend ihrem Alter und Geschlecht, der Symptomatik und der bestehenden Risikofaktoren (KHK-Vortestwahrscheinlichkeit).

Wesentliche Einschränkungen und Schwächen des Modellansatzes, beispielsweise die Verwendung von administrativen anstelle von realen, den tatsächlichen Ressourcenverbrauch widerspiegelnden Preisen, werden von den Autoren kritisch hinterfragt. Vorbehalte, ob der Einsatz einer diagnostisch weniger genauen Methode im Vergleich zum Goldstandard Angiographie ethisch vertretbar ist, und Verteilungsfragen werden nicht erörtert. Die von den Autoren gezogenen Schlußfolgerungen sind in sich konsistent.

Abschließende Bewertung

Die bereits 1995 publizierte Studie, in der die gebräuchlichsten Methoden zur Diagnose von koronarer Herzkrankheit, Belastungs-EKG, SPECT und Angiographie, und die alternative PET auf ihre Kosteneffektivität vergleichend überprüft werden, ist von hoher methodischer Qualität und Informationsdichte. Sie zeichnet sich durch eine langfristige Zeitperspektive (Zeithorizont von 10 Jahren), durch die Erfassung von

aussagekräftigen, finalen Effekten wie Lebenserwartung und qualitätsadjustierter Lebenserwartung, durch die Berücksichtigung von Kosten für Komplikationen (einschließlich eines Kostenanteils an indirekten Kosten infolge Produktionsausfall) und durch eine detaillierte und transparente Darstellung aus. Aufgrund der guten methodischen Qualität dieser Studie kommt den hier gewonnenen Ergebnissen eine große Aussagekraft zu und ist ihnen deshalb ein hoher Stellenwert zuzumessen.

Maddahi J & Gambhir SS. Cost-effective selection of patients for coronary angiography. J Nucl Cardiol 1997;4(2 Pt 2):S141-S151

Studienbeschreibung

Bei Verdacht auf koronare Herzkrankheit (KHK) spielen bei der Auswahl von Patienten, die eine Koronarangiographie benötigen, um entweder eine „definitive“ Diagnose „KHK“ stellen zu können oder als Vorbereitung auf eine Revaskularisation des Myokards, nicht-invasive Untersuchungsverfahren zunehmend eine Rolle. Zu dieser Thematik entwerfen die Autoren Maddahi & Gambhir in ihrer Publikation aus dem Jahr 1997 ein Entscheidungsbaum-Modell mit sechs alternativen Strategien. Diese sind zum besseren Verständnis in der folgenden Übersicht zusammengestellt.

Tabelle 13: Alternative Diagnosestrategien bei Patienten mit Verdacht auf koronare Herzkrankheit, erstellt nach Angaben von Maddahi & Gambhir 1997

Modelläste	Eingangstest	2. Test Bedingung ^d	3. Test Bedingung
Strategie 1	direkte Angiographie		
Strategie 2	PET ^a	Angiographie PET +	
Strategie 3	SPECT ^b	Angiographie SPECT +	
Strategie 4	ExEKG ^c	PET ExEKG +	Angiographie PET +
Strategie 5	ExEKG	SPECT ExEKG +	Angiographie SPECT +
Strategie 6	ExEKG	Angiographie ExEKG +	

a Positronen-Emissions-Tomographie

b Single-Photon-Emissions-Computer-Tomographie

c Belastungs-EKG

d Eine zusätzliche Untersuchung wird unter der Bedingung durchgeführt, daß die vorausgegangene Untersuchung positiv war.

Quelle: zusammenfassende Darstellung nach Angaben von Maddahi & Gambhir 1997

Mit Hilfe dieses mathematischen Modells wird die Wertigkeit der ExEKG und der nuklearmedizinischen Methoden SPECT und PET sowohl aus medizinischer als auch aus ökonomischer Sicht im Hinblick auf eine Entscheidung für oder gegen eine Koronarangiographie bei Patienten mit unterschiedlicher Vortestwahrscheinlichkeit für KHK überprüft. Dazu werden die durch die Alternativstrategien 2 – 6 verursachten direkten medizinischen Ausgaben zu den reinen Angiographiekosten (Strategie 1) in Relation gesetzt („relative Diagnosekosten“). Als Maß für die klinische Effektivität der verwandten Techniken wird der prozentuale Anteil der korrekt diagnostizierten KHK-Patienten ($[\text{richtig-positive} + \text{richtig-negative Patienten}] / \text{Gesamtpatientenzahl}$) be-

stimmt. Die Gesamtkosten bzw. die diagnostische Treffsicherheit der alternativen Diagnosestrategien werden in Abhängigkeit von der KHK-Vortestwahrscheinlichkeit der Patienten untereinander verglichen. Die Vortestwahrscheinlichkeit einer einzelnen Person für KHK (pretest likelihood), die mit ihrem Alter, dem Geschlecht und den vorliegenden Krankheitssymptomen korreliert, ist ein Maß für das Risiko des Einzelnen, tatsächlich an KHK erkrankt zu sein; auf der Bevölkerungsebene spiegelt das epidemiologische Maß der Prävalenz dieses Erkrankungsrisiko wider.

Die wichtigsten Ergebnisse der Studie sind:

1. Im Modell sinkt für die Strategien 2 bis 6 mit steigender KHK-Prävalenz wegen der größer werdenden Zahl falsch-negativer Resultate die Ausbeute an richtig diagnostizierten Patienten. Gleichzeitig steigen auf der Kostenseite die Gesamtkosten für die jeweilige Kombinationsdiagnostik an, da mit steigender KHK-Prävalenz bei einem zunehmenden Anteil der Bevölkerung eine Überstellung zur verifizierenden Angiographie nach einem nicht-invasiven Eingangstest notwendig wird.
2. Bei Überweisung der Patienten zur PET-Untersuchung als nicht-invasivem Vortest zur Angiographie (Strategie 2) übertreffen ab einer KHK-Prävalenz bzw. Vortestwahrscheinlichkeit $> 0,6$ (60%) die dabei anfallenden Gesamtkosten die Kosten für sofortige Angiographie (Strategie 1) bei geringerer Ausbeute an korrekt diagnostizierten Patienten. Das heißt, in diesem Prävalenzbereich dominiert die Angiographie über die PET/Angiographie-Kombination.
3. Entsprechendes gilt bei direkter Überweisung der Patienten zur SPECT-Untersuchung als nicht-invasivem Vortest zur Angiographie (Strategie 3). Auch hier übertreffen bei geringerer Ausbeute an korrekt diagnostizierten Patienten ab einer KHK-Prävalenz/Vortestwahrscheinlichkeit¹³ $> 0,68^*$ (68%) die dabei anfallenden Gesamtkosten die Kosten für die sofortige Angiographie (Strategie 1).
4. Bis zu einer Prävalenz/Vortestwahrscheinlichkeit von $0,42^*$ (42%) dominiert die PET/Angiographie-Kombination über die alternative SPECT/Angiographie-Kombination (bessere diagnostische Ausbeute bei niedrigeren Kosten).
5. Die Gesamtkosten der übrigen Strategien (4 bis 6), mit dem herkömmlichen Belastungs-EKG als Eingangsuntersuchung, bleiben über die gesamte Spanne der Prävalenz/Vortestwahrscheinlichkeit unter dem von den Angiographiekosten vorgegebenen Schwellenwert, jedoch bei deutlich verminderter diagnostischer Ausbeute.

Auf der Basis ihrer eigenen Ergebnisse und Zusatzinformationen (siehe Studienbewertung) aus der Fachliteratur entwickeln die Autoren für den Einsatz der PET bzw. SPECT als Diagnoseinstrumente bei KHK-Verdacht ein in Tabelle 14 (siehe nächste Seite) beschriebenes Strategiemodell.

¹³ Die mit * gekennzeichneten Angaben für die Prävalenz/Vortestwahrscheinlichkeit sind aus Abbildung 2 und Abbildung 3 der Originalliteratur Maddahi & Gambhir 1997 entnommen, da die Autoren in diesen Entscheidungssituationen keine expliziten Angaben machen.

Tabelle 14 Strategien der Wahl für die Diagnose von koronarer Herzkrankheit nach Angaben von Maddahi & Gambhir 1997

Vortestwahrscheinlichkeit ^a	1. Test	2. Test Beding. ^c	Risiko ^e Beding.	Angiographie
Niedrig (20%)	ExEKG ^b ExEKG	NT ^d ExEKG +	niedrig ExEKG – niedrig NT – hoch NT +	nein nein ja
Mittel (50%)	Kombination ExEKG/NT		niedrig NT – hoch NT +	nein ja
Hoch (80%)	direkte Angiographie			ja einziger Test

a Vortestwahrscheinlichkeit eines Patienten, an koronarer Herzkrankheit zu leiden.

b Belastungs-Elektrokardiogramm

c Eine zusätzliche Untersuchung ist angezeigt unter der Bedingung, daß der Vortest + ausgefallen ist.

d Nuklearmedizinischer Test: PET oder SPECT

e Risiko für kardiale Ereignisse, wie beispielsweise Herzinfarkt oder Herzrhythmusstörungen, bei ausschließlich medikamentöser Behandlung.

Quelle: zusammenfassende Darstellung nach Angaben von Maddahi & Gambhir 1997

Tabelle 15: Strategien der Wahl für die Prognose von koronarer Herzkrankheit bei Patienten mit normalem bzw. abnormalem Elektrokardiogramm in Ruhe nach Angaben von Maddahi & Gambhir 1997

Ruhe-EKG	Vortestwahrscheinlichkeit ^a	1. Test	2. Test Beding. ^b	Risiko Beding. ^c	Angiographie
Normal	niedrig			niedrig	nein
Normal	mittel	ExEKG ^d	NT ^e ExEKG +	niedrig ExEKG – niedrig NT – hoch NT +	nein nein ja
Normal	hoch	Kombination ExEKG/NT		niedrig NT – hoch NT +	nein ja
Abnormal	gesamte Bandbreite	Kombination ExEKG/NT		niedrig NT – hoch NT +	nein ja

a Vortestwahrscheinlichkeit eines Patienten, an koronarer Herzkrankheit zu leiden. Zahlenwerte zur Abgrenzung der Abstufungen „niedrig / mittel / hoch“ werden in der Publikation von Maddahi J & Gambhir 1997 nicht angegeben.

b Eine zusätzliche Untersuchung ist angezeigt unter der Bedingung, daß der Vortest + ausgefällt.

c Unter der Bedingung, daß der vorher durchgeführte Test – bzw. + ausgefallen ist, ist das Risiko für kardiale Ereignisse, wie beispielsweise Herzinfarkt oder Herzrhythmusstörungen, niedrig oder hoch. Zahlenangaben zu den von Maddahi & Gambhir 1997 verwendeten Risikoabstufungen „niedrig / hoch“ liegen nicht vor.

d Belastungs-Elektrokardiogramm

e Nuklearmedizinischer Test: PET oder SPECT

Quelle: zusammenfassende Darstellung nach Angaben von Maddahi & Gambhir 1997

Eine Koronarangiographie wird außer zur Diagnose von KHK meist zur Vorbereitung einer Revaskularisation eingesetzt. Eine derartige Intervention ist jedoch nur bei Patienten indiziert, die bei alleiniger Behandlung mit Medikamenten einem hohen Risiko für kardiale Ereignisse ausgesetzt wären. Durch den Einsatz von nuklearmedizinischen Methoden, PET oder SPECT, soll die präoperative Auswahl dieser Patientengruppe erleichtert werden. Die Prognose einer KHK-Erkrankung ist jedoch nicht allein aus der Vortestwahrscheinlichkeit abzuleiten, sondern beruht auch auf dem Status des Ruhe-EKGs. Deshalb entwerfen die Autoren ein zweites Strategiemodell zur Prognose von KHK bei Patienten mit normalem und abnormalen Ruhe-EKG-Befund. Dazu werden Zusatzinformationen aus der Fachliteratur (siehe Studienbewertung) herangezogen. Dieses Strategiemodell ist aus Tabelle 15 auf der vorhergehenden Seite zu ersehen.

Aus den Studienergebnissen ergibt sich für die Autoren, bezogen auf die Vortestwahrscheinlichkeit, somit folgendes Bild: bei Patienten mit hoher Vortestwahrscheinlichkeit für KHK ist die direkte Überweisung zur Angiographie sowohl aus medizinischer als auch aus ökonomischer Sicht die günstigste Vorgehensweise (Kosten für direkte Angiographie < Kosten für Strategien 2 bzw. 3, bei gleichzeitig optimaler diagnostischer Treffsicherheit), während bei Patienten mit mittlerer Vortestwahrscheinlichkeit entweder Strategie 2 oder 3, mit den nuklearmedizinischen Testverfahren PET oder SPECT als Eingangsuntersuchung, vorzuziehen sind. Der Einsatz der nuklearmedizinischen Techniken PET bzw. SPECT bei Patienten mit niedriger Vortestwahrscheinlichkeit ist nur für die Subgruppe der Patienten mit abnormalem Belastungs-EKG angezeigt.

Studienbewertung

Bei der vorliegenden Studie handelt es sich aus gesundheitsökonomischer Sicht nicht um eine ökonomische Evaluation im strengen Sinne, da die hier untersuchten Effekte und Kosten nicht direkt zueinander in Beziehung gesetzt werden. Hauptanliegen der Autoren im Rahmen dieser Publikation ist nicht die Berechnung einer Kosten-Effektivität-Relation, sondern der Entwurf von Strategien zur Diagnose bzw. Prognose von KHK, die sowohl klinische als auch ökonomische Aspekte berücksichtigen (siehe Studienbeschreibung). Aus diesem Grund weicht sowohl der inhaltliche als auch der formale Aufbau der Publikation von der üblichen Form ab.

Der Evaluationsrahmen ist nur unvollständig dargestellt; so werden weder der Zeithorizont noch die Perspektive von den Autoren explizit festgelegt. Der Zeithorizont ergibt sich jedoch aus der Beschränkung auf rein diagnostische Techniken (Diagnosestudie), die Perspektive - in diesem Fall die des Leistungserbringers - aus der Wahl des Preisgerüsts. Zielpopulation der Studie sind Patienten mit KHK, deren Vortestwahrscheinlichkeit für KHK von den Faktoren Alter, Geschlecht und Krankheitssymptome (z.B. für atypische oder typische Angina pectoris) abhängt.

Das hier vorgestellte entscheidungstheoretische Modell von Maddahi & Gambhir aus dem Jahr 1997 ist weniger komplex als die entscheidungsanalytischen Modelle der kalifornischen Arbeitsgruppe bei den Indikationen Staging von nicht-kleinzelligem Bronchiolarkarzinom bzw. solitären Lungenknoten (siehe Studienberichte unter Kapitel C.5.4.1). Der aus sechs Strategieästen aufgebaute Entscheidungsbaum wird zwar im Text beschrieben, jedoch leider nicht anhand eines Baumdiagramms verdeutlicht. Die nicht näher spezifizierten Bausteine der alternativen Diagnosestrategien, Belastungs-EKG (ExEKG), SPECT, PET und Koronarangiographie, entsprechen dem an großen (PET) Zentren heute üblichen Instrumentarium zur präoperativen Diagnostik von KHK-Patienten.

Die Modellparameter, Daten zur diagnostischen Genauigkeit der verglichenen Diagnosemethoden, die Vortestwahrscheinlichkeit für KHK in Abhängigkeit von den Faktoren Alter, Geschlecht und Symptomatik (Literaturdaten) und die „relativen Kosten“ (siehe unten), werden eindeutig genannt und ihre als adäquat zu wertenden Quellen dokumentiert (Ausnahme: Kostenangaben). Für die Sensitivitäts- und Spezifitätswerte (Literaturdaten einer Meta-Analyse) sind nur Basisdaten, aber keine Streubreiten angegeben. Auch die Suchstrategie des Literaturreviews und das Ergebnis der Literatursuche werden nicht berichtet.

Als Maß für die klinische Effektivität der Konkurrenzstrategien dient die biometrische Größe „prozentualer Anteil korrekt diagnostizierter Patienten“. Diese intermediäre Effektgröße (diagnostic accuracy efficacy) hat nach dem unter C.4.2.2 definierten hierarchischen Modell der Effizienz von Effektgrößen nur beschränkte Aussagekraft, solange nicht auch die Auswirkungen der diagnostischen Überlegenheit einer Technologie gegenüber Alternativtechnologien auf finale Größen, beispielsweise als Änderungen im Therapiemanagement oder im medizinischen Outcome, in Betracht gezogen werden. Da in der vorliegenden Studie therapeutische Interventionen und ihre Wirkung auf den Gesundheitszustand der Patienten unberücksichtigt bleiben, sind der Zeithorizont und somit auch die in der Analyse berücksichtigten Kosten und Effekte auf die Phase der Diagnosestellung begrenzt. Für ein entscheidungsanalytisches Modell erscheint dieses Zeitfenster als zu eng gewählt.

Das Mengengerüst zur Berechnung der Gesamtkosten pro Diagnosestrategie ergibt sich aus dem mathematischen Modell. Als Preisgerüst werden institutionsspezifische Kostenangaben für Einzelleistungen, gemittelt über mehrere medizinische Einrichtungen, verwendet. Da diese Institutionen nicht näher spezifiziert werden, ist die Repräsentativität der in der Analyse veranschlagten Kosten für das amerikanische Gesundheitssystem in seiner Gesamtheit fraglich. Für die einzelnen medizinischen Leistungen wird die Vergütung nicht beziffert, sondern es werden nur die „relativen“ Kosten der einzelnen nicht-invasiven Diagnoseverfahren im Vergleich zu den Angiographiekosten angegeben. Daß nicht absolute sondern nur „relative“ Kosten angeführt werden, ist unter dem Gesichtspunkt der Vergleichbarkeit der Studie mit anderen publizierten Studien problematisch. Die für die Effektseite beschriebenen Einschränkungen hinsichtlich der Aussagefähigkeit der Ergebnisse durch die Einengung des

Zeitfensters gelten auch für die Kostenseite. So werden außer den direkten diagnostischen Kosten keine sonstigen mittel- oder längerfristigen Kostenkomponenten, wie Therapiekosten, für Komplikationen aufzubringende Zusatzkosten oder indirekte Kosten durch Produktionsausfall, angerechnet. Wegen des enggesteckten Zeitrahmens erübrigt sich eine Diskontierung und - vorausgesetzt, daß im Rahmen der Studie aktuelle Kostendaten benutzt wurden - ebenso eine Inflations-/Deflationsbereinigung.

Die Kosten und Effekte werden nicht zu einer Maßzahl zusammengeführt. Die relative Ausbeute an korrekt diagnostizierten Patienten und die relativen Gesamtkosten pro Studienast werden getrennt für drei Patientengruppen mit unterschiedlichen Vortestwahrscheinlichkeiten (20%, 50%, 80%) in übersichtlichen Schaubildern präsentiert. Auf dem Bevölkerungsniveau aggregierte Effekte und Kosten werden nicht berichtet, jedoch die Änderung der relativen Effekte und der relativen Kosten über die gesamte Spanne der Vortestwahrscheinlichkeit (0 - 100%) graphisch illustriert. Sensitivitätsanalysen zur Behandlung von Unsicherheiten werden nicht angestellt, aber die Variabilität der relativen Kosten und Effekte in Abhängigkeit von der Vortestwahrscheinlichkeit für koronare Herzkrankheit wird graphisch dargelegt.

Die Studienergebnisse werden in Zusammenhang mit anderen relevanten Studien zur Thematik diskutiert. Basierend auf ihren eigenen Studienergebnissen und zusätzlichen Literaturangaben zur inkrementellen, diagnostischen Power der nicht-invasiven Diagnosemethoden EKG unter Belastung, myokardiale Perfusionsszintigraphie und ^{99m}Technetium-Sestamibi SPECT und ergänzende Informationen zum Einfluß von Perfusionsmessungen mit Radionukliden auf das Patientenmanagement und den gesundheitlichen Outcome entwerfen die Autoren die in der Kurzbeschreibung vorgestellten Strategiemodelle für die Diagnose und Prognose von KHK. Wie und in welchem Maße eigene Studienergebnisse und literaturbasierte Fremddaten in die Strategiemodelle einfließen, ist aus der Publikation nicht zu ersehen.

Abschließende Bewertung

Das entscheidungsanalytische Modell von Maddahi & Gambhir aus dem Jahr 1997, in dem die Rolle der nuklearmedizinischen Verfahren PET und SPECT für die Diagnose und Prognose koronarer Herzkrankheit bei Patienten mit unterschiedlicher Vortestwahrscheinlichkeit sowohl aus dem medizinischen als auch aus dem ökonomischen Blickwinkel beleuchtet wird, ist aufgrund methodischer Defizite - wie fehlender Kosten-Effektivitäts-Relationen und einem zu engen Zeithorizont - nur von eingeschränkter Aussagekraft. Dennoch werden in der abschließenden Diskussion dieses HTA-Berichtes die Ergebnisse bezüglich der Abhängigkeit der Kosteneffektivität der PET von der KHK-Vortestwahrscheinlichkeit in die Gesamtbetrachtung mit einbezogen, da dieses Teilergebnis der Studie eindeutig ist. Hervorzuheben an der Studie sind (a) der Reviewteil, der eine Übersicht über die Literatur zur Studienthematik gibt, und (b) der Ansatz, theoretische ökonomische Erkenntnisse zu nutzen, um für die klinische Praxis relevante diagnostische Ablaufschemata zu entwerfen.

Garber AM & Solomon NA. Cost-effectiveness of alternative test strategies for the diagnosis of coronary artery disease. *Ann Intern Med.* 1999; 130: 719 – 728

Studienbeschreibung

Die 1999 veröffentlichte ökonomische Studie aus den USA hat zum Ziel, für sechs verschiedene Ansätze zur Versorgung von Patienten mit koronarer Herzkrankheit (KHK) den gesundheitlichen Outcome und die dafür aufzubringenden Kosten anhand einer theoretischen Entscheidungsanalyse vergleichend zu evaluieren und eine nach Kosten-Nutzwert-Relation geordnete Rangfolge der evaluierten Strategien aufzustellen.

In die ökonomische Evaluation einbezogen sind fünf zweistufige Diagnosealgorithmen und als sechste Alternative die als Goldstandard anerkannte Methode der Koronarangiographie. Bei den zweistufigen Diagnosestrategien ist jeweils ein nicht-invasiver Eingangstest mit einer nachgeschalteten Koronarangiographie mittels Herzkatheter, die der Verifikation eines positiven Eingangsbefundes dient, kombiniert. Ist der Eingangstest negativ, entfällt die Angiographie. Als initiale, nicht-invasive Testmethoden wählen die Autoren die gebräuchlichen elektrophysiologischen bzw. nuklearmedizinischen Verfahren zur Diagnose einer Koronarinsuffizienz: Nachweis ischämiebedingter Veränderungen im Belastungs-EKG (ExEKG, ergometrische Belastung mittels Tretwerk) und Nachweis von durch körperliche oder pharmakologische Belastung provozierte Ischämie anhand der folgenden vier alternativen bildgebenden Verfahren:

1. Planare Szintigraphie mit ²⁰¹Thallium- oder ^{99m}Technetium-Sestamibi,
2. Single-Photon-Emissions-Computer-Tomographie (SPECT) mit Thallium- oder Technetium-Radiodiagnostika,
3. Ultraschall-Echokardiographie (UKG) unter Belastung durch körperliche Anstrengung oder Gabe der pharmakologischen Stressoren Dobutamin oder Dipyridamol,
4. Untersuchung der myokardialen Perfusion mittels PET.

Patienten mit gesicherter KHK-Diagnose oder mit Verdacht auf KHK, das heißt bei positivem oder unbestimmtem Ergebnis des nicht-invasiven Eingangstests, werden mittels Angiographie weiteruntersucht. Eine KHK liegt definitionsgemäß dann vor, wenn anhand einer Angiographie mindestens eine 50%ige Stenose des Hauptstammes der linken Koronararterie oder eine mindestens 70%ige Stenose einer anderen Koronararterie festzustellen ist. Je nach anatomischer Verteilung der Obstruktionen in den Blutgefäßen werden die KHK-Patienten entsprechend den Modellvorgaben unterschiedlichen Therapieverfahren zugeführt: bei einer 1- bzw. 2-Gefäßerkrankung erfolgt eine medikamentöse Therapie oder eine Revaskularisation in Form einer perkutanen transluminalen Koronar-Angioplastie (PTCA), bei einer 3-Gefäßerkrankung bzw. einer Hauptstammstenose der linken Koronararterie wird entweder medikamentös behandelt oder eine operative Koronarrevaskularisation durch Bypass-

Transplantation (CABG) veranlaßt. Fällt dagegen der Eingangstests oder die (anschließende) Angiographie negativ aus, verbleiben die Patienten bei medikamentöser Behandlung unter Beobachtung.

Die entscheidungstheoretische Analyse ist auf nach Geschlecht stratifizierte Zielpopulationen in der mittleren und fortgeschrittenen Lebensphase (Alterskohorten mit 45, 55 und 65jährigen Patienten) zugeschnitten, die über Brustschmerzen in der Vorgeschichte berichten und aufgrund ihres Geschlechts, ihres Alters, dem Vorliegen von Risikofaktoren und der Ausprägung der Brustschmerzsymptomatik eine mittlere Wahrscheinlichkeit von 25 - 75% aufweisen, an KHK zu leiden (Vortestwahrscheinlichkeit): dies sind die Bevölkerungsgruppen der Frauen mittleren Alters mit typischer Angina pectoris¹⁴ bzw. der Männer im mittleren und im fortgeschrittenen Lebensalter mit atypischer Angina pectoris.

Als Vergleichsmaß für die Effektivität der diagnostischen Verfahren dienen neben der diagnostischen Genauigkeit der Testverfahren (diagnostic accuracy efficacy, siehe C.4.2.2) auch längerfristige gesundheitliche Effekte, wie der Einfluß der Versorgungsstrategien auf Lebenserwartung und Lebensqualität der Patienten (patient outcome efficacy). Verbesserungen der gesundheitsbezogenen Lebensqualität, bewertet als qualitätsgleiche Lebensjahre (quality-adjusted life years, QALYs), resultieren aus der erfolgreichen Behandlung einer Angina pectoris und der Verhinderung von Myokardinfarkten. Um die Wirkung einer korrekten bzw. einer inkorrekten Klassifikation des Erkrankungszustandes auf den gesundheitlichen Outcome ermessen zu können, nutzen die Autoren zur Veranschaulichung des KHK-Progressionsprozesses ein Markov-Modell mit 1jährigen Übergangsperioden zwischen den Gesundheitszuständen „Symptomfreiheit“, „Angina pectoris“, „Myokardinfarkt“ und „Tod“. Auf der Kostenseite werden die für Kostenträger relevanten direkten medizinischen Gesamtkosten der sechs alternativen Versorgungsstrategien erfaßt und in einer inkrementellen Kosten-Nutzwert-Analyse (siehe Studienbewertung) durch Berechnung der zusätzlichen Kosten pro gewonnenem QALY mit den erzielten Gesundheitseffekten direkt ins Verhältnis gesetzt.

Laut Modellberechnungen variiert die Lebenserwartung nur wenig mit der Methodik des diagnostischen Eingangstests. Unter der Basisannahme, daß der untersuchte Patient zur Kohorte der Männer mittleren Alters (55 Jahre), mit einer Vortestwahrscheinlichkeit für KHK (pCAD) von 50%, gehört, ergibt sich beim Vergleich zwischen der besten und der schlechtesten Testmethode ein Gewinn an Lebenszeit von nur 7 Tagen, wobei die Lebenserwartung (qualitätsadjustierte Lebenserwartung) in der Basiskohorte der 55jährigen bei Männern 16,6 (12,2) Jahre und bei Frauen 19,7 (14,1) Jahre beträgt.

Beim Vergleich der alternativen Strategien ergibt sich folgendes Bild:

¹⁴ Definition für typische und atypische Angina pectoris siehe Bericht zur Studie von Patterson et al. 1995.

1. Ultraschall-Echokardiographie (UKG) als Initialtest verbessert deutlich die Gesundheitseffekte relativ zum Initialtest ExEKG bei vergleichbaren oder mäßig erhöhten Kosten (maximale iKER: US\$ 9 000 pro zusätzlichem QALY). Gegenüber der Kombination ²⁰¹Tl-Planarszintigraphie/Angiographie ist die Kombination UKG/Angiographie in den höheren Altersklassen sowohl der männlichen als auch der weiblichen 55- bzw. 65jährigen dominant (nicht jedoch für die Gruppe der 45jährigen).
2. Für die zweistufigen Diagnosealgorithmen - SPECT/Angiographie in Relation zur UKG/Angiographie - liegt die inkrementelle Kosten-Nutzwert-Ratio für die Basis Kohorte der 55jährigen Männer bei US\$ 75 000/QALY (mit einer Variationsspanne von circa US\$ 64 000 pro gewonnenem QALY bei der Altersgruppe der 65jährigen Männer bis nahezu US\$ 150 000 pro gewonnenem QALY bei den 45jährigen Frauen).
3. Der inkrementelle Kosten-Nutzwert-Quotient beträgt unter den Basiskonditionen - bei einem Preis von US\$ 1 500 pro PET-Aufnahme - für die Kombination PET/Angiographie in Relation zur alternativen Kombination SPECT/Angiographie mehr als US\$ 640 000/QALY. Bei Absenkung des PET-Preisniveaus auf US\$ 750 bzw. US\$ 1 000 resultieren immer noch Kosten-Nutzwert-Quotienten von über US\$ 180 000/QALY bzw. US\$ 335 000/QALY. Im Hinblick auf die diagnostische Genauigkeit zeigt PET bei vergleichbarer Sensitivität eine nur leicht verbesserte Spezifität gegenüber SPECT.
4. Die Koronarangiographie als alleinige Testmethode verursacht verglichen mit dem Zweistufenalgorithmus SPECT/Angiographie im analytischen Modell für die Basis Kohorte der 55jährigen Männer Mehrkosten pro gewonnenem QALY von insgesamt US\$ 94 000 (mit einer Variationsspanne von US\$ 80 000/QALY bei den 65jährigen Männern bis fast US\$ 200 000/QALY bei den 45jährigen Frauen).
5. Sofortige Koronarangiographie ohne nicht-invasiven Vortest ist in allen Kohorten (Männern und Frauen aller Altersstufen) gegenüber der PET-basierten Diagnosestrategie eindeutig dominant. Diese Dominanz bleibt auch erhalten, wenn zur Prüfung der Stabilität der Ergebnisse in einer Sensitivitätsanalyse die Mortalitätsrate für Angiographie verdreifacht wird.

Somit ergibt sich - unter der Basisannahme (Kohorte der 55jährigen Männer mit einer pCAD von 50%) - gemäß dem Kosten-Nutzwert-Verhältnis der einzelnen Strategien folgende aufsteigende Rangfolge: UKG/Angiographie, SPECT/Angiographie und sofortige Angiographie. Auch eine Variation der Patientenmerkmale Alter und Geschlecht und der Wahrscheinlichkeit, KHK zu haben (pretest probability), ändert diese Rangfolge nicht qualitativ (Sensitivitätsanalysen).

Anhand dieser Ergebnisse ziehen die Autoren das Resümee, daß die konventionellen Diagnosestrategien UKG/Angiographie, SPECT/Angiographie und sofortige Angiographie kosteneffektive Alternativen zur PET/Angiographie (und zu den diagnostisch weniger wirksamen Methoden ExEKG bzw. Thallium-Szintigraphie) darstellen. Je nach Höhe des - gesellschaftlich - akzeptierten Kosten-Nutzwert-Verhältnisses (Kosten pro zusätzlich gewonnenem QALY) ist der Einsatz der UKG (bei relativ nied-

rigem Kosten-Nutzwert-Verhältnis $< \text{US\$ } 10\,000/\text{QALY}$) bzw. der SPECT (bei mittlerem Kosten-Nutzwert-Verhältnis $\leq \text{US\$ } 80\,000/\text{QALY}$) zusammen mit der Koronarangiographie und der alleinige Einsatz der invasiven Koronarangiographie als Initialdiagnosemethode (bei einem gesellschaftlich akzeptierten Kosten-Nutzwert-Verhältnis von $> \text{US\$ } 80\,000$ pro zusätzlichem QALY, siehe oben) die zu bevorzugende Diagnosestrategie.

Studienbewertung

Die entscheidungstheoretische Modellierung von Garber und Solomon zeigt anhand einer inkrementellen Kosten-Nutzwert-Analyse die Kosten- und Effektunterschiede zwischen mehreren eingeführten Untersuchungsverfahren zur Diagnose von KHK auf. Um einen möglichen Einfluß der Patientenmerkmale „Geschlecht“ und „Alter“ auf die Kosten und Effekte nicht durch Nivellierung von Geschlechts- bzw. Altersunterschieden zu verwischen, erfolgt die Analyse jeweils für Männer und Frauen getrennt und nach Alterskohorten abgestuft (Subgruppenanalyse mit 6 Strata). Die Bevölkerungsgruppe, deren medizinische Versorgung sichergestellt werden soll, also die Zielpopulation, ist klar definiert und die Subgruppen werden hinreichend gründlich beschrieben.

Die Zielsetzung der Studie, Diagnosestrategien anhand ihrer Kosten-Nutzwert-Rangfolge vergleichend zu bewerten, ist eindeutig formuliert und wird in der vorliegenden Studie auch konsequent verfolgt. Der ökonomische Background wird in der Texteingührung nur knapp dargelegt; eine ausführlichere Darstellung des medizinischen Problemkontextes, die der Bedeutung der koronaren Herzkrankheit als häufigste Todesursache in den Industrieländern angemessen ist, wäre wünschenswert.

Das auf einen Zeithorizont von 30 Jahren projizierte Modell ist aufgrund seines anspruchsvollen Studiendesigns sehr komplex: so werden sechs alternative Strategien zur diagnostischen und therapeutischen Versorgung von KHK-Patienten miteinander verglichen und zudem mehrere, über die Zeit variierende Gesundheitszustände berücksichtigt (siehe Markov-Modell in der Studienbeschreibung). Die Modellstrukturen, das heißt die Abfolge der einzelnen Interventionsschritte, aus denen sich die Versorgungsstrategien zusammensetzen (Entscheidungsbaum), und der Progressionsverlauf der KHK im Markov-Modell, sind in graphischen Ablaufschemata anschaulich dargestellt. Dagegen werden die Bausteine der Algorithmen, die einzelnen diagnostischen und therapeutischen Verfahren, methodisch kaum erläutert; so fehlen für die PET jegliche Angaben zur technischen Ausstattung oder zum verwendeten Radiotracer. Die Auswahl der Vergleichstechnologien wird nicht begründet, entspricht jedoch dem heute - zumindest an großen Zentren (im Hinblick auf PET) - üblichen Spektrum an apparativen Standardmethoden zur Diagnose von KHK. Da es sich bei der KHK um eine chronische Erkrankung mit langer Verlaufsdauer handelt, ist der vorgegebene Zeithorizont von 30 Jahren angemessen.

Um das entscheidungsanalytische Modell erstellen zu können, schätzen die Autoren die Parameter der diagnostischen Genauigkeit der beteiligten Testverfahren - nach Schwere und Form der KHK differenziert - aus veröffentlichten Sensitivitäts- und Spezifitätswerten durch Meta-Analyse ab oder nutzen gepoolte Schätzwerte aus bereits publizierten Meta-Analysen (für EKG unter Belastung und planare Szintigraphie mit $^{201}\text{Thallium}$). Die genutzte Datenbank und die Suchstrategie werden berichtet; das Ergebnis der Literatursuche ist dokumentiert, das heißt, die verwerteten Studien sind unter Angabe der Fallzahl, der Testergebnisse und der diagnostischen Kennzahlen akribisch aufgelistet. Das Ergebnis der Meta-Analysen, Basiswerte und Streubreiten der biometrischen Kennzahlen der einzelnen Verfahren, ist in einer Tabelle zusammengestellt. Da zum Zeitpunkt der Studie für die PET keine Sensitivitätsangaben zur Diagnostik von schwerwiegenden KHK-Erkrankungen, wie 3-Gefäßerkrankung bzw. linke Hauptstammstenose, in der Literatur vorlagen, wird dieser Modellparameter von den Autoren unter der Annahme geschätzt, daß die Falschnegativ-Rate bei PET-Untersuchungen im gleichen Maße in Abhängigkeit vom Erkrankungsgrad abnimmt wie bei SPECT. Sonstige Daten, wie Angaben zur Wirksamkeit der Therapieoptionen (Medikation, PTCA und CABG) bzw. zu Erkrankungsprävalenzen, Überlebenswahrscheinlichkeiten und Morbiditätsdaten nach chirurgischen und medikamentösen Interventionen, entstammen ebenfalls der Literatur (z.B. aus Veröffentlichungen zur Coronary Artery Survey Study), die im Literaturverzeichnis ausgewiesen ist.

Auf der Effektseite werden mit den Größen Lebenserwartung und qualitätsadjustierte Lebensjahre aussagekräftige finale Outcomegrößen ausgewählt, die grundsätzlich geeignet sind, den Gesundheitszustand von Patienten im verlängerten Zeithorizont über die reinen diagnostischen und therapeutischen Interventionen hinaus, widerzuspiegeln (patient outcome efficacy). Eine wesentliche Schwäche der Studie ist jedoch, daß weder die Methoden zur Identifikation, Extraktion und Synthese der Effektparameter beschrieben, noch die Übergangswahrscheinlichkeiten des Markov-Modells aufgeführt werden. Die Methoden und Meßinstrumente zur Qualitätsgewichtung der in der Follow-up-Phase verbrachten Lebensjahre werden von den Autoren leider nicht angegeben; aus den in der Referenzliste verzeichneten Literaturquellen ist jedoch zu entnehmen, daß die Bewertung der Lebensqualität anhand der - als geeignet zu bewertenden - Methode der zeitlichen Abwägung (time-trade-off) erfolgte. Aufgrund der beschriebenen formalen Mängel ist die Berechnung der zu erwartenden Lebensdauer bzw. der QALYs für den Leser nicht nachvollziehbar und somit auch nicht nachprüfbar.

Die Gesamtaufwendungen für einen Versorgungsalgorithmus enthalten außer den direkten medizinischen Kosten für die Diagnostik auch die entsprechenden Kosten der therapeutischen Interventionen (einschließlich längerfristiger Kosten für medikamentöse Behandlung), die je nach Diagnosebefund und fortschreitender Erkrankung (gemäß dem Markov-Modell) in der klinischen Praxis veranlaßt würden. Das der Kostenberechnung zugrundeliegende Mengengerüst ist in der Publikation nicht aufgeführt. Dies erübrigt sich jedoch, da für die einzelnen Teststrategien die Kostenbela-

stung pro Kopf, also für den einzelnen Patienten und nicht für ein Patientenkollektiv, berechnet wird. Dagegen ist das Preisgerüst, das heißt die für die einzelnen diagnostischen bzw. therapeutischen Interventionen anzurechnenden Vergütungen, nahezu vollständig aufgelistet (einschließlich der durchschnittlichen Aufwendungen bei einem Myokardinfarkt bzw. für medikamentöse Behandlung; es fehlen jedoch die Kosten für die Überwachung der im nicht-invasiven Eingangstest als negativ diagnostizierten Patienten) und ausreichend beschrieben (u.a. Aufschlüsselung nach stationären Kosten und Kosten für ärztliche Leistungen). Jedoch ist kritisch anzumerken, daß zur Kostenberechnung nationale Medicare-Durchschnittsvergütungen, umgerechnet auf das Bezugsjahr 1996, herangezogen werden. Die Verwendung der Medicare-Vergütungssätze¹⁵ (außer bei PET) ist fragwürdig, denn die Zielpopulation der Studie umfaßt per Definition nur Personen der Altersspanne zwischen 45 und 65 Jahren; also gerade nicht die Medicare-Klientel. Da PET zum Zeitpunkt der Studie nicht - wie die übrigen medizinischen Einzelleistungen - im Medicare-Vergütungskatalog enthalten war, gehen die Kosten für eine PET-Aufnahme als ein auf Privatversicherungsbasis abzurechnender Posten in die Gesamtkostenrechnung ein. Im Interesse einer unverzerrten Bewertung ist jedoch eine einheitliche monetäre Bemessungsgrundlage anzustreben.

Da die Altersklassen der 45 - 65jährigen einen nicht unerheblichen Teil der berufs- und erwerbsfähigen Bevölkerung in den Industrieländern stellen, hätte es die insgesamt als sehr gut einzustufende Evaluation abgerundet, wenn die Autoren in die Gesamtkostenbetrachtung - soweit möglich - auch indirekte Kosten in Form von Produktivitätseinbußen, bei krankheitsbedingtem Arbeitsausfall oder infolge vorzeitiger Berentung oder Tod, einbezogen hätten. Da ohne spezielle Überprüfung nicht davon auszugehen ist, daß die indirekten Kosten für alle KHK-Diagnosestrategien vergleichbar sind, sollten, besonders bei der Evaluation von Maßnahmen zur Versorgung chronischer Erkrankungen mit hoher Prävalenz und großer volkswirtschaftlicher Bedeutung wie der KHK, indirekte Kosten in der Gesamtkostenkalkulation nicht fehlen.

Die Kostenseite wird anhand einer inkrementellen Analyse, die der unter C.4.2.3 dargestellten Entscheidungsregel folgt, mit der Effektseite direkt in Beziehung gesetzt. Dazu werden die alternativen Strategien nach ihrem Kostenaufwand aufsteigend geordnet, wobei teurere, aber weniger effektive Strategien von der weiteren Betrachtung ausgeschlossen werden. Für die verbliebenen Alternativstrategien wird gemäß dieser Kostenrangfolge die Kosten-Nutzwert-Relation berechnet.

Entsprechend der Wahl eines auf 30 Jahren projizierten Zeithorizontes werden zukünftig anfallende Kosten, aber auch zukünftiger Nutzen in Form veränderter Lebenserwartung oder Lebensqualität (\Rightarrow QALYs) auf das Bezugsjahr 1996 abgezinst und inflationsbereinigt. Die inflationsadjustierte Diskontrate, für Kosten und Effekte einheitlich gewählt, beträgt in der Studie 3%. Empfehlungen von Expertengre-

¹⁵ Medicare: staatliches US-Krankenversicherungsprogramm für die Bevölkerungsschicht der über 65jährigen.

mien, mehrere Diskontraten (z.B. 3% und 7%) oder eine Spanne von Diskontraten (z.B. von 0 - 7%) anzusetzen, wird nicht entsprochen.

Die Ergebnisse werden ausschließlich in disaggregierter Form (pro Kopf-Kalkulation), in übersichtlichen Tabellen und Schaubildern zusammengefaßt, präsentiert. In den ausschließlich univariaten Sensitivitätsanalysen werden relevante Parameter, unter anderem Patientenmerkmale wie Alter und Geschlecht, die Prävalenz für KHK oder das Preisniveau für PET, variiert. Zur Prüfung der Stabilität des Modells wären zusätzlich auch Sensitivitätsanalysen mit veränderten diagnostischen Kenndaten sinnvoll gewesen.

Im insgesamt recht kurz gehaltenen Diskussionsteil werden Einschränkungen und Schwächen der Studie und Fragen der Generalisierbarkeit der Ergebnisse nicht diskutiert. Ethische und Verteilungsfragen werden ebenfalls nicht angesprochen, sind jedoch auch nicht Gegenstand der Aufgabenstellung. Die auf den Studienergebnissen basierenden Schlußfolgerungen sind in sich konsistent, werden jedoch in der abschließenden Diskussion leider nicht noch einmal prägnant zusammengefaßt herausgestellt.

Abschließende Bewertung

Das von den Autoren Garber und Solomon entworfene entscheidungstheoretische Modell bestimmt die Kosten und Effekte für sechs unterschiedliche Diagnosestrategien mit wechselnden Eingangstests (fünf nicht-invasive Eingangstests: Belastungs-EKG, planare ²⁰¹Tl-Szintigraphie, SPECT, Belastungs-Echokardiogramm und PET; Goldstandard: Koronarangiographie) und stellt entsprechend der Kosten-Nutzwert-Relation eine - nach Geschlechtern getrennte - Rangfolge ihrer Wertigkeit sowohl unter klinischen als auch unter ökonomischen Aspekten auf. Die insgesamt methodisch als sehr gut einzustufende Kosten-Nutzwert-Analyse zeichnet sich durch folgende Qualitätskriterien aus:

1. Durch die Modellierung eines mit 30 Jahren sehr langfristigen Zeithorizontes,
2. durch die Erfassung von aussagekräftigen, finalen Effektgrößen (siehe C.4.2.2) wie Lebenserwartung und qualitätsadjustierter Lebenserwartung, verbunden mit einer qualitativen Bewertung von Gesundheitszuständen (QALY-Betrachtung),
3. durch die Berechnung von inkrementellen Kosten-Nutzwert-Quotienten,
4. durch die Diskontierung der Kosten und Effekte.

Aufgrund der hohen methodischen Qualität dieser Studie ist den hier dargelegten Ergebnissen eine große Aussagekraft zuzumessen.

C.5.5 Zusammenfassende Übersicht

In den nachfolgenden Tabellen 16 und 17 sind die Zielsetzung und die wesentlichen Ergebnisse der behandelten ökonomischen Studien zusammenfassend dargestellt. Zudem werden die wichtigsten Studienmerkmale und Qualitätsaspekte der einzelnen Studien aufgeführt. Anhand dieser Informationen wird die Relevanz der Ergebnisse der einzelnen Studien für die Beurteilung der Kosteneffektivität des klinischen Einsatzes der PET eingeschätzt, differenziert nach den Indikationsgebieten:

Tabelle 16:

- Diagnose und Management von solitären Lungenknoten
- Diagnose und / oder Staging von nicht-kleinzelligem Lungenkarzinom
- Staging von axillären Lymphknoten bei Mammakarzinom
- Staging von malignen Lymphomen
- Lokalrezidive bei kolorektalem Karzinom
- Staging von metastasierenden, malignen Melanomen

Tabelle 17:

- Perfusionsmessung bei Patienten mit Verdacht auf koronare Herzkrankheit
- Messung der Myokardviabilität bei Herzinfarktpatienten

Tabelle 16: Übersicht über die onkologischen Studien

Autor, Jahr, Land	Studienart	Vergleichs- alternativen ^a	Ergebnisse der Studie ^b	Qualitätsaspekte der Studie	Relevanz der Ergebnisse
Diagnose und Management von solitären Lungenknoten (SPNs)					
Coleman, 1994, USA	Partielle Evaluation (Kostenvergleich), mathematische Modellierung	¹⁸ FDG-PET versus CT + Biopsie	Einsparpotential aufgrund vermiedener Interventionen: US\$ 330 – 2 600 pro Patient, US\$ 30 – 236 Millionen auf US-Bevölkerungsniveau	Zeithorizont: Phase der Akutversorgung, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Anzahl vermiedener Interven- tionen (monetäre Bewertung), Mängel in der Dokumentation der verwand- ten Parameter und in der Darstellung der Ergebnisse, keine Sensitivitätsanalysen	Ergebnisse sind auf- grund methodischer Mängel von begrenz- ter Aussagekraft
Gambhir, 1998, USA	Inkrementelle Kosten-Effektivitäts- Analyse, entscheidungstheo- retisches Modell	Kombination CT/ ¹⁸ FDG-PET (+Biopsie) (CT als Eingangstest), nur CT (+ Biopsie), Thorakotomie versus „wait & watch“ (Referenzstrategie)	Einsparpotential der Kombi- nation CT + PET gegenüber nur CT: US\$ 91 – 2 200 pro Patient (Ø-Wert: US\$ 1 194), US\$ 63 Millionen auf US- Bevölkerungsniveau, Abhängigkeit der Kosten- effektivität von der Vortest- wahrscheinlichkeit siehe in Abbildung 1	Zeithorizont: Restlebenszeit, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Lebenserwartung (keine Prä- ferenzbewertung), umfangreiche uni- und multivariate Sensiti- vitätsanalysen, gute Dokumentation der Modellparameter, anschauliche Darstellung der Ergebnisse	Ergebnisse sind aus- sagekräftig aufgrund der hohen methodi- schen Qualität der Studie
Diagnose und / oder Staging von nicht-kleinzelligem Lungenkarzinom (NSCLC)					
Valk 1996, USA	Partielle Evaluation (Kostenvergleich), begleitende retro- spektive ökonomi- sche Analyse zu einer klinischen Studie (Fallzahl n = 72)	CT + PET, nur PET versus CT + Skelett- szintigraphie	Einsparpotential aufgrund vermiedener Operationen bezogen auf das Studien- kollektiv: US\$ 150 000 (CT + PET), US\$ 170 000 (nur PET)	Zeithorizont: Phase der Akutversorgung, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Anzahl vermiedener Operatio- nen (monetäre Bewertung), sehr knappe Darstellung der ökonomischen Aspekte, Intransparenz der Kostenberech- nung (Preisgerüst), keine Sensitivitätsana- lysen, knappe Darstellung und Diskussion der Ergebnisse	Ergebnisse sind auf- grund methodischer Mängel von begrenz- ter Aussagekraft

a Abkürzungen siehe Abkürzungsverzeichnis

b Angaben gelten bei mathematischen Modellierungen und entscheidungstheoretischen Modellen unter Baseline-Bedingungen

Autor, Jahr, Land	Studienart	Vergleichs- alternativen ^a	Ergebnisse der Studie ^b	Qualitätsaspekte der Studie	Relevanz der Ergebnisse
Fortsetzung: Diagnose und / oder Staging von nicht-kleinzelligem Lungenkarzinom (NSCLC)					
Gambhir, 1996, USA	Gegenüberstellung der Kosten und der Effekte (Minimierung der Kosten bzw. Maximierung der Effekte in einer Threshold-Analyse) entscheidungstheo- retisches Modell	Kombination ¹⁸ FDG- PET/CT (+Biopsie) (mit PET als Eingangstest), versus nur CT (+ Biopsie)	Einsparpotential durch ver- bessertes Patientenmanage- ment: US\$ 1 150 pro Patient bei Zunahme an Lebenserwar- tung von ≈ 3 Tagen, Abhängigkeit der Kosten- effektivität von der Vortest- wahrscheinlichkeit siehe in Abbildung 1	Zeithorizont: Restlebenszeit, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Lebenserwartung (keine Prä- ferenzbewertung), umfangreiche uni- und bivariate Sensitivi- tätsanalysen, gute Dokumentation der Modellparameter und anschauliche Darstellung der Ergeb- nisse	Ergebnisse sind aus- sagekräftig aufgrund der hohen methodi- schen Qualität der Studie
Scott, 1998, USA	Inkrementelle Kosten-Effektivitäts- Analyse, entscheidungs- theoretisches Modell	Vier unterschiedliche Kombinationen aus ¹⁸ FDG-PET, CT, (+ Biopsie) (mit CT als Eingangstest) versus nur CT (+ Biopsie) (Referenzstrategie)	Beste iKER bei der Strategie „PET nur nach CT-negativem Eingangsbefund“: US\$ 25 286 pro zusätzlichem Lebensjahr, Abhängigkeit der Kosten- effektivität von der Vortest- wahrscheinlichkeit siehe in Abbildung 1	Zeithorizont: Restlebenszeit, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Lebenserwartung (keine Prä- ferenzbewertung), umfangreiche uni- und multivariate Sensitivi- tätsanalysen, gute Dokumentation der Modellparameter, anschauliche Darstellung der Ergebnisse	Ergebnisse sind aus- sagekräftig aufgrund der hohen methodi- schen Qualität der Studie, Vorsicht bei der Inter- pretation der iKER (Problematik des „kleinen Nenners“, siehe Studienbe- wertung)
Von Schult- hess, 1998, Schweiz	Partielle Evaluation (Kostenvergleich), begleitende öko- nomische Analyse zu einer prospek- tiven klinischen Studie (Fallzahl n = 62)	¹⁸ FDG-PET/CT (CT nur zur Operations- vorbereitung) versus CT/Skelett- szintigraphie	Einsparpotential aufgrund ver- miedener Operationen und CT-Scans auf der Ebene des Studienkollektivs: 100 200 sfr	Zeithorizont: Phase der Akutversorgung, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Anzahl vermiedener Interven- tionen (monetäre Bewertung), knappe Darstellung der Effekte, Mängel auf der Kostenseite, keine Sensitivitätsanalysen	Ergebnisse sind auf- grund methodischer Mängel von begrenz- ter Aussagekraft

a Abkürzungen siehe Abkürzungsverzeichnis

b Angaben gelten bei mathematischen Modellierungen und entscheidungstheoretischen Modellen unter Baseline-Bedingungen

Autor, Jahr, Land	Studienart	Vergleichs- alternativen ^a	Ergebnisse der Studie ^b	Qualitätsaspekte der Studie	Relevanz der Ergebnisse
Fortsetzung: Diagnose und / oder Staging von nicht-kleinzelligem Lungenkarzinom (NSCLC)					
ECRI, 1998, USA	Gegenüberstellung der Kosten bzw. der Effekte von Alternativstrategien (Entscheidungskriterium: Dominanz), entscheidungstheoretisches Modell	Neun alternative Strategien mit unterschiedlicher Anordnung der Bausteine CT, ¹⁸ FDG-PET, (+ Biopsie) im Diagnose- und im Stagingschritt (siehe Studienbeschreibung)	Strategie F, eine kombinierte Strategie CT/PET + Biopsie, dominiert im Stagingschritt über die Routinestrategie „nur CT + Biopsie“, Einsparpotential durch verbessertes Staging: US\$ 36 pro Patient bei einem Zugewinn an Lebenserwartung von 1,8 Tagen, US\$ 4,7 Millionen auf US-Bevölkerungsniveau	Zeithorizont: Restlebenszeit, direkte medizinische Kosten: diagnostische und therapeutische Interventionen + Kosten für Langzeitversorgung bei Rezidivbildung, Effektgröße: Lebenserwartung (keine Präferenzbewertung) und sonstige patientenbezogene Outcomegrößen (u.a. Zahl der überflüssigen Operationen und der übersehenen NSCLC-Patienten), umfangreiche Sensitivitätsanalysen	Ergebnisse sind aussagekräftig aufgrund der hohen methodischen Qualität der Studie
Kosuda, 2000, Japan	Inkrementelle Kosten-Effektivitäts-Analyse, entscheidungstheoretisches Modell	CT + ¹⁸ FDG-PET versus nur CT	iKER 218 000 Yen (rund US\$ 1 550) pro zusätzlichem Lebensjahr und pro Patient, Zugewinn an Lebenserwartung: 7,3 Monate pro Patient	Zeithorizont: Restlebenszeit, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Lebenserwartung (keine Präferenzbewertung), univariate Sensitivitätsanalysen	Ergebnisse sind aussagekräftig aufgrund der hohen methodischen Qualität der Studie
Staging von axillären Lymphknoten bei Mammakarzinom					
Adler, 1994, USA	Partielle Evaluation (Kostenvergleich), mathematische Modellierung	¹⁸ FDG-PET + ALND versus nur ALND	Einsparpotential aufgrund vermiedener ALND: US\$ 490 – 1 600 pro Patient, US\$ 49 – 160 Millionen auf US-Bevölkerungsniveau	Zeithorizont: Phase der Akutversorgung, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Anzahl vermiedener ALND (monetäre Bewertung), Mängel in der Dokumentation der verwandten Parameter und der Darstellung der Ergebnisse, keine Sensitivitätsanalysen	Ergebnisse sind aufgrund methodischer Mängel von begrenzter Aussagekraft

a Abkürzungen siehe Abkürzungsverzeichnis

b Angaben gelten bei mathematischen Modellierungen und entscheidungstheoretischen Modellen unter Baseline-Bedingungen

Autor, Jahr, Land	Studienart	Vergleichs- alternativen ^a	Ergebnisse der Studie ^b	Qualitätsaspekte der Studie	Relevanz der Ergebnisse
Fortsetzung: Staging von axillären Lymphknoten bei Mammakarzinom					
Adler, 1997, USA	Partielle Evaluation (Kostenvergleich), begleitende ökonomische Analyse zu einer prospektiven klinischen Studie (Fallzahl n = 50)	¹⁸ FDG-PET + ALND versus nur ALND	Einsparpotential aufgrund vermiedener Operationen: US\$ 2 300 pro Person, US\$ 120 000 auf der Ebene des Studienkollektivs	Zeithorizont: Phase der Akutversorgung, direkte medizinische Diagnosekosten, Effektgröße: Anzahl vermiedener ALND (monetäre Bewertung), in erster Linie klinische Studie, sehr knappe Darstellung der Kostenaspekte, keine Sensitivitätsanalysen	Ergebnisse sind aufgrund methodischer Mängel von begrenzter Aussagekraft
Staging von malignen Lymphomen					
Valk, 1996, USA	Partielle Evaluation (Kostenvergleich), retrospektive ökonomische Analyse zu einer klinischen Studie (Fallzahl n = 30)	¹⁸ FDG-PET versus CT + Gallium-67-Tumorszintigraphie	Kein Einsparpotential	Zeithorizont: Phase der Akutversorgung, keine Angaben von Kosten und keine Kostenberechnung, Effektgröße: Änderungen im Patientenmanagement, nur sehr knappe Darstellung der Effekte, keine Diskussion der Ergebnisse, keine Sensitivitätsanalysen	Ergebnisse sind aufgrund vielfältiger methodischer Mängel von sehr begrenzter Aussagekraft
Hoh, 1997, USA	Partielle Evaluation (Kostenvergleich), begleitende ökonomische Analyse zu einer prospektiven klinischen Studie (Fallzahl n = 18)	¹⁸ FDG-PET-basierter Algorithmus versus konventionellen CT-basierten Algorithmus	Kosten bezogen auf das Studienkollektiv: PET-Staging: US\$ 38 000, CT-Staging: US\$ 68 000	Zeithorizont: Phase der Diagnosestellung, direkte medizinische Diagnosekosten, Effektgröße: Anzahl der richtig identifizierten Tumorknoten bzw. eingestuft -stadien, enger Zeithorizont, aussageschwache Effektgröße, kleine Fallzahl (n = 18)	Ergebnisse sind aufgrund methodischer Mängel von begrenzter Aussagekraft; aber gute Ausgangsbasis für künftige Evaluationen
Klose, 1999a, Deutsch- land	Inkrementelle Kosten-Effektivitäts-Analyse; begleitende ökonomische Analyse zu einer prospektiven klinischen Studie (Fallzahl n = 22)	¹⁸ FDG-PET versus nur CT versus „Nichtstun“	Mehrkosten pro zusätzlich richtig klassifiziertem Patienten (iKER): 1 393 Euro für PET versus CT, 662 Euro für CT versus „Nichtstun“:	Zeithorizont: Phase der Diagnosestellung, direkte medizinische Tomographiekosten, verursachungsgemäße Kostenzurechnung, Effektgröße: Anzahl richtig diagnostizierter Patienten, umfangreiche Sensitivitätsanalysen, enger Zeithorizont, aussageschwache Effektgröße, kleine Fallzahl (n = 22)	Ergebnisse sind aufgrund methodischer Schwächen im Studiendesign der klinischen Studie von begrenzter Aussagekraft; aber gute Ausgangsbasis für zukünftige Evaluationen

a Abkürzungen siehe Abkürzungsverzeichnis

b Angaben gelten bei mathematischen Modellierungen und entscheidungstheoretischen Modellen unter Baseline-Bedingungen

Autor, Jahr, Land	Studienart	Vergleichs- alternativen ^a	Ergebnisse der Studie ^b	Qualitätsaspekte der Studie	Relevanz der Ergebnisse
Lokalrezidive bei kolorektalem Karzinom					
Larson, 1994, USA	Partielle Evaluation (Kostenvergleich), mathematische Modellierung	¹⁸ FDG-PET versus Laparotomie	Einsparpotential aufgrund vermiedener Interventionen: US\$ 5 400 – 32 000 pro Patient, US\$ 113 – 671 Millionen auf US-Bevölkerungsniveau	Zeithorizont: Phase der Akutversorgung, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Anzahl vermiedener Interven- tionen (monetäre Bewertung), Mängel in der Dokumentation der verwand- ten Parameter und der Darstellung der Er- gebnisse, keine Sensitivitätsanalysen	Ergebnisse sind auf- grund methodischer Mängel von begrenz- ter Aussagekraft
Valk, 1996, USA	Partielle Evaluation (Kostenvergleich), retrospektive öko- nomische Analyse zu einer klinischen Studie (Fallzahl n = 68)	CT + ¹⁸ FDG-PET, nur ¹⁸ FDG-PET versus nur CT	Einsparpotential aufgrund vermiedener Operationen auf der Ebene des Studien- kollektivs: US\$ 178 000 (CT + PET), US\$ 232 000 (nur PET)	Zeithorizont: Phase der Akutversorgung, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Anzahl vermiedener Operatio- nen (monetäre Bewertung), sehr knappe Darstellung der ökonomischen Aspekte, Intransparenz der Kostenberech- nung (Preisgerüst), keine Sensitivitätsana- lysen, knappe Darstellung und Diskussion der Ergebnisse	Ergebnisse sind auf- grund vielfältiger methodischer Mängel von sehr begrenzter Aussagekraft
Valk, 1999, USA	Partielle Evaluation (Kostenvergleich), begleitende ökono- mische Analyse zu einer prospektiven klinischen Studie (Fallzahl n = 115)	¹⁸ FDG-PET versus nur CT	Einsparpotential aufgrund vermiedener Operationen: US\$ 3 000 pro Patient US\$ 234 000 auf der Ebene des Studienkollektivs	Zeithorizont: Phase der Akutversorgung, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Anzahl vermiedener Operatio- nen (monetäre Bewertung), in erster Linie klinische Studie, knappe Dar- stellung ökonomischer Aspekte, keine Sen- sitivitätsanalysen	Ergebnisse sind auf- grund methodischer Mängel von begrenz- ter Aussagekraft

a Abkürzungen siehe Abkürzungsverzeichnis

b Angaben gelten bei mathematischen Modellierungen und entscheidungstheoretischen Modellen unter Baseline-Bedingungen

Autor, Jahr, Land	Studienart	Vergleichs- alternativen ^a	Ergebnisse der Studie ^b	Qualitätsaspekte der Studie	Relevanz der Ergebnisse
Staging von metastasierenden, malignen Melanomen					
Valk, 1996, USA	Partielle Evaluation (Kostenvergleich), retrospektive ökonomische Analyse zu einer klinischen Studie (Fallzahl n = 45)	CT + PET, nur PET, nur CT	Einsparpotential aufgrund vermiedener Operationen bezogen auf das Studien- kollektiv: US\$ 98 000 (PET + CT) bzw. US\$ 135 000 (nur PET)	Zeithorizont: Phase der Akutversorgung, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Anzahl vermiedener Operatio- nen (monetäre Bewertung), sehr knappe Darstellung der ökonomischen Aspekte, Intransparenz der Kostenberech- nung (Preisgerüst), keine Sensitivitätsana- lysen, knappe Darstellung und Diskussion der Ergebnisse	Ergebnisse sind auf- grund vielfältiger methodischer Mängel von sehr begrenzter Aussagekraft

a Abkürzungen siehe Abkürzungsverzeichnis

b Angaben gelten bei mathematischen Modellierungen und entscheidungstheoretischen Modellen unter Baseline-Bedingungen

Tabelle 17: Übersicht über die kardiologischen Studien

Erstautor, Jahr, Land	Studienart ^a	Vergleichsalternativen	Ergebnisse der Studie ^b	Qualitätsaspekte der Studie	Relevanz der Ergebnisse
Myokardperfusionsmessung bei Patienten mit Verdacht auf koronare Herzkrankheit (KHK)					
Gould, 1989, USA	Partielle Evaluation (Kostenvergleich), mathematische Modellierung	⁸² Rb-PET versus ²⁰¹ Tl-Perfusions-szintigraphie	Einsparpotential: rund US\$ 200 000 für ein fiktives 100köpfiges Patientenkollektiv	Zeithorizont: 5 Jahre, direkte medizinische Kosten: diagnostische und therapeutische Interventionen + indirekte Kosten, Effektgröße: Anzahl vermiedener Versorgungsleistungen (monetäre Bewertung), Mängel in der Kostenberechnung, keine Sensitivitätsanalysen	Ergebnisse sind aufgrund methodischer Einschränkungen und fehlender Aktualität der Daten von begrenzter Aussagekraft
Patterson, 1994, 1995, USA	Kosten-Effektivitäts-Analyse + Kosten-Nutzwert-Analyse (Ø-Betrachtung), mathematische Modellierung	ExEKG/Angiographie, SPECT/ Angiographie, PET/ Angiographie, nur Angiographie	Methode der Wahl: für pCAD < 0,6: PET/Angiographie (unter Belastung) für pCAD > 0,6: nur Angiographie, Abhängigkeit der Kosteneffektivität von der Vortestwahrscheinlichkeit siehe in Abbildung 2	Zeithorizont: 10 Jahre, direkte medizinische Kosten: Diagnose plus Komplikationen (Einbezug indirekter Kosten durch falsch-negative Fehlklassifikation), Effektgrößen: Lebenserwartung und QALYs, uni- und multivariate Sensitivitätsanalysen	Ergebnisse sind aussagekräftig aufgrund der hohen methodischen Qualität der Studie
Garber, 1999, USA	Inkrementelle Kosten-Nutzwert-Analyse, entscheidungstheoretisches Modell (mit Markov-Modell)	Nur Angiographie (AG), PET/AG, SPECT/AG, UKG/AG, ²⁰¹ Tl-Szintigraphie/AG, ExEKG/AG (inkrementelle Analyse entsprechend der Kostenrangfolge)	Kosteneffektive Alternativmethoden zu PET: UKG/AG, SPECT/AG, direkte AG (Reihenfolge gemäß dem Kosten-Nutzwert-Verhältnis), Abhängigkeit der Kosteneffektivität von der Vortestwahrscheinlichkeit siehe in Abbildung 2	Zeithorizont: 30 Jahre, direkte medizinische Kosten: diagnostische und therapeutische Interventionen (einschließlich längerfristiger Kosten für medikamentöse Behandlung), Effektgrößen: Lebenserwartung und QALYs, univariate Sensitivitätsanalysen	Ergebnisse sind aussagekräftig aufgrund der hohen methodischen Qualität der Studie

a Abkürzungen siehe Abkürzungsverzeichnis,

b Angaben gelten bei mathematischen Modellierungen und entscheidungstheoretischen Modellen unter Baseline-Bedingungen

Erstautor, Jahr, Land	Studienart ^a	Vergleichsalternativen	Ergebnisse der Studie ^b	Qualitätsaspekte der Studie	Relevanz der Ergebnisse
Myokardperfusionsmessung bei Patienten mit Verdacht auf koronare Herzkrankheit (KHK)					
Maddahi, 1997, USA	Gegenüberstellung der Kosten bzw. der Effekte der Alternativstrategien (keine KER), entscheidungstheoretisches Modell	Fünf mehrstufige Strategien mit PET, SPECT, ExEKG alternativ als Eingangstest (plus anschließende Angiographie) versus nur Angiographie (mittels Herzkatheter)	Günstigste Vorgehensweise: Patienten mit hoher pCAD (0,8): direkte Angiographie, Patienten mit mittlerer pCAD (0,5): PET oder SPECT als Eingangstest vor der Angiographie, Abhängigkeit der Kosteneffektivität von der Vortestwahrscheinlichkeit siehe in Abbildung 2	Zeithorizont: begrenzt auf Diagnosestellung, direkte medizinische Kosten der Diagnosestrategien, Effektgröße: %-Anteil korrekt diagnostizierter Patienten, keine Kosteneffektivitätsrelation berechnet, sondern Vergleich relativer Kosten und relativer Effekte (Referenz: Angiographie), keine Sensitivitätsanalysen	Ergebnisse sind aufgrund methodischer Defizite von eingeschränkter Aussagekraft
Messung der Myokardviabilität bei Herzinfarkt-Patienten					
Hunink, 1995, Niederlande	Entwurf einer inkrementellen Kosten-Effektivitäts-Analyse, entscheidungstheoretisches Modell (Entscheidungsbaum mit Markov-Modell)	PET versus SPECT	iKER (Kosten pro „ereignisfreiem“ Lebensjahr) im 1, 2, 5jährigen Follow-up: US\$ 22 600 (nach 1 Jahr) US\$ 9 000 (nach 2 Jahren) US\$ 3 400 (nach 5 Jahren) entsprechend einer degressiven Entwicklung der iKER über die 5jährige Follow-up-Phase	Zeithorizont: 1, 2 und 5 Jahre (Follow-up), direkte medizinische Interventionskosten (mit Kosten für „downstream events“), Effektgröße: „event-free survival“, schmale Datenbasis, fehlende Angaben zum Markov-Modell, Mängel der Kostenberechnung und der Ergebnisdarstellung, keine Sensitivitätsanalysen	Ergebnisse sind aufgrund methodischer Einschränkungen von begrenzter Aussagekraft; aber gute Ausgangsbasis für künftige Studien
Besozzi, 1994, USA	Partielle Evaluation (Kostenvergleich), mathematische Modellierung	Kombination PET/Angiographie versus nur Angiographie (mittels Herzkatheter)	Einsparpotential wegen vermiedener Eingriffe: US\$ 355 – 2 627 pro Patient US\$ 40 – 300 Millionen auf US-Bevölkerungsniveau	Zeithorizont: Phase der Akutversorgung, direkte medizinische Kosten: diagnostische und therapeutische Interventionen, Effektgröße: Anzahl vermiedener Versorgungsmaßnahmen, Mängel in der Dokumentation der verwandten Parameter und der Darstellung der Ergebnisse, keine Sensitivitätsanalysen	Ergebnisse sind aufgrund methodischer Mängel von begrenzter Aussagekraft

a Abkürzungen siehe Abkürzungsverzeichnis

b Angaben gelten bei mathematischen Modellierungen und entscheidungstheoretischen Modellen unter Baseline-Bedingungen

C.6 Diskussion

Die Technologie der Positronen-Emissions-Tomographie (PET), die in den frühen 1970er Jahren neu eingeführt wurde, vergrößert die Palette der bereits für die klinische Diagnostik zur Verfügung stehenden nicht-invasiven Bildgebungsverfahren, darunter Röntgenverfahren wie die Computer-Tomographie (CT) oder die Mammographie, die Kernspin-Tomographie, und nuklearmedizinische Methoden wie Planarszintigraphie oder Single-Photon-Emissions-Computer-Tomographie (SPECT). Die nuklearmedizinische PET nutzt kurzlebige Positronen-emittierende Spürsubstanzen, wie beispielsweise das Kalium-Analogon Rubidium-82 oder den Zuckerabkömmling Fluor-18-FDG, die als markierte Substitute von physiologischen Substanzen in Spurdosen in den Organismus eingeschleust werden, um krankhafte Veränderungen biologischer Prozesse qualitativ nachweisen und eventuell quantifizieren zu können, insbesondere zur Messung der Perfusion, des Metabolismus oder der Rezeptorverteilung. Während somit die nuklearmedizinische PET-Technik Informationen zu physiologischen Körperfunktionen und körpereigenen Stoffwechselfvorgängen liefert (funktionelle Bilderfassung), erkennen die konventionellen radiologischen Verfahren wie die CT vor allem morphologische Strukturmerkmale (anatomische Bilderfassung).

Die PET ist aus technischer Sicht das derzeit beste verfügbare nuklearmedizinische Verfahren zur räumlichen Erfassung intrakorporaler Aktivitätsverteilungen bei hoher räumlicher und zeitlicher Auflösung. Sie ermöglicht eine ungefähr 10 – 50mal sensitivere Signalausbeute als die nuklearmedizinische Konkurrenzmethode SPECT. Neben den meßtechnischen Vorteilen für die Bildgebung bietet das Verfahren durch die Möglichkeit der Verwendung einer Vielzahl von anwendungsspezifischen Radiopharmazeutika einen weiteren, oftmals entscheidenden Vorteil gegenüber den herkömmlichen Methoden: Die Vielfalt an verfügbaren Radiotracer ist für die vielseitige Anwendung der PET in Forschung und Klinik mit verantwortlich.

Trotz dieser Vorteile beschränkte sich in den 1970er und 1980er Jahren der Einsatz der mit großem technischen und finanziellen Aufwand verbundenen PET-Technologie auf nur wenige Forschungseinrichtungen. Erst in jüngerer Zeit ist zu beobachten, daß in den Industrienationen an den großen medizinischen Zentren zunehmend PET-Anlagen installiert werden. Im Frühjahr 1995 standen weltweit 134 PET-Geräte in medizinischen Einrichtungen zur Verfügung, davon 60 in den USA, 45 in Europa, 20 in Japan und 9 in anderen Ländern (davon 3 in Australien). In Deutschland verfügten in den Jahren 1990 - 1992 erst 6 Kliniken über PET-Geräte (König 1993); zum Stichtag 1.1.1997 waren es 26 Geräte, davon 23 an Krankenhäusern und 3 in Arztpraxen (Daten des Gesundheitswesens, Ausgabe 1999).

Obwohl bis vor wenigen Jahren die PET hauptsächlich Forschungszwecken diente, schufen die hierbei gewonnenen Erkenntnisse die Voraussetzungen zur Anwendung der neuen Technologie im klinisch-diagnostischen Bereich. Die Weiterentwicklung

der Kamertechnologie, die Zugriffsmöglichkeit auf eine Vielzahl von Positronen-emittierenden Radiotracern und eine stärkere Verbreitung der PET-Geräte führt nunmehr - auch in Deutschland - dazu, daß die PET zunehmend Eingang in den klinischen Alltag findet. Derzeitige Haupteinsatzgebiete für PET sind die auch aus bevölkerungsepidemiologischer Sicht wichtigen Bereiche Onkologie, Kardiologie und die Neuromedizin mit ihren spezifischen Einsatzmöglichkeiten in der Neurologie, Neurochirurgie und Psychiatrie.

Zur Bewertung der Effektivität der innovativen PET-Technik als diagnostisches Instrument wurden unter der Schirmherrschaft der Deutschen Gesellschaft für Nuklearmedizin (DGN) Konsensuskonferenzen zu den drei genannten Haupteinsatzbereichen Onkologie, Kardiologie und Neuromedizin veranstaltet. Auf diesen Konferenzen wurden mögliche klinische Indikationen für den PET-Einsatz von interdisziplinären Expertengremien nach definierten Kriterien bewertet und Konsensuspapiere erarbeitet, in denen die - gemäß dem Wissensstand von 1997 - für eine PET-Anwendung geeigneten Indikationen aufgelistet wurden. Auf diese Leitlinien der DGN soll im folgenden kurz eingegangen werden.

Indikationen im Bereich Onkologie für den Einsatz der PET

Bei vielen Krebsarten sind die Gewebefbiochemie und die physiologischen Prozesse charakteristisch gestört. So zeigen viele Neoplasmen einen überhöhten Glucoseumsatz, der durch intrazelluläres Trapping des nach metabolischer Phosphorylierung vom Zellstoffwechsel nicht weiter verwertbaren Zuckeranalogons Fluor-18-Fluoro-Deoxyglucose (^{18}FDG), dem in Krebsstudien meistgenutzten Radiopharmakon, in PET-Aufnahmen bildhaft dargestellt werden kann. Da die Detektion und Lokalisation von Tumorfoci das vordringlichste Ziel diagnostischer Untersuchungen in der Onkologie ist, wird die PET in der klinischen Praxis bisher hauptsächlich für qualitative oder semiquantitative Untersuchungen genutzt. Quantitative Messverfahren sind derzeit noch in der Erprobung (z.B. zur Quantifizierung des Glucosemetabolismus in Tumorzellen).

Im Laufe der technischen Weiterentwicklung wurde das Verfahren der PET-Bilderfassung auch den Erfordernissen der Krebsdiagnostik angepaßt. Mit der Entwicklung von Ganzkörper-Scannern hat die klinische Nutzung von PET in der Onkologie stark zugenommen, da die Ganzkörper-PET die Möglichkeit bietet, den gesamten Rumpf des Patienten in vertretbarer Zeit (innerhalb von 1 - 1,5 h) auf Primärtumore und streuende (Fern)metastasen durchzuscanen und diese zu lokalisieren. So werden heute international an den meisten PET-Zentren mehr als 70% der Überweisungen für PET-Scans von onkologischen Abteilungen veranlaßt (Robert & Milne 1999, siehe dazu auch Abschnitt C.6.2).

Potentielle Haupteinsatzgebiete für die PET-Technologie in der klinischen Onkologie sind die Lokalisation von Primärtumoren, Differenzierung zwischen benignen und

malignen Tumorerkrankungen, Definition der Progression von Tumorerkrankungen (räumliche Abgrenzung der Tumorausbreitung, Kontrolle auf Metastasierung und Staging), Monitoring der Tumorreaktion auf Chemo- und Strahlentherapie und Kontrolle auf Rezidivbildung (Robert & Milne 1999).

1997 wurde durch den Arbeitsausschuß PET der Deutschen Gesellschaft für Nuklearmedizin im Rahmen einer interdisziplinären Konsensuskonferenz in Ulm die seit 1995 bestehende Indikationsliste zur differenzierten Bewertung des PET-Einsatzes in der Onkologie überarbeitet und aktualisiert. Die Bewertung des Stellenwertes der PET im Rahmen der Versorgung von Krebspatienten erfolgte durch ein Gremium von ausgewiesenen Experten aus unterschiedlichen onkologischen Fachgebieten anhand von Studienergebnissen, die in Form von Originalarbeiten publiziert wurden („literature based evidence“). Das Expertengremium erachtete bei folgenden Indikationen die medizinische Effektivität der PET als erwiesen und somit ihren Einsatz in der klinischen (Routine)diagnostik für sinnvoll, so daß eine Zuordnung in die Klassifikationsstufe 1a (Einsatz angemessen, klinischer Nutzen ist erwiesen) erfolgte (Leitlinien der Deutschen Gesellschaft für Nuklearmedizin, Konsensuspapier Onko-PET 1998; Kongressbericht im Deutschen Ärzteblatt 1998):

- Abklärung eines vermuteten Rezidivs oder von Metastasen bei differenzierten Schilddrüsenkarzinomen (bei negativem Jod-Scan)
- Diagnostik und Erkennung der malignen Entdifferenzierung bei Gliomrezidiven im Gehirn oder Bestimmung des Biopsieortes bei Gliomverdacht
- Restaging von kolorektalen Karzinomen bei begründetem Rezidivverdacht
- Suche nach Primärtumoren im Kopf/Nacken-Bereich
- Lymphknotenstaging und Detektion von Fernmetastasen bei malignen Melanomen im Stadium II und III
- Lymphknotenstaging und Abklärung von Lokalrezidiven bzw. peripheren Lungenrundherden bei nicht-kleinzelligem Bronchialkarzinom
- Differentialdiagnostik von Primärtumoren der Pankreas versus chronischer Pankreatitis.

Indikationen im Bereich Kardiologie für den Einsatz der PET

Die Applikation der PET-Technologie in diesem Indikationsbereich ist wesentlich homogener als in der Onkologie. So gibt es vorrangig zwei unterschiedliche methodische Einsatzmöglichkeiten: einerseits die Perfusionsmessung, zum Beispiel mit Rubidium-82, einem Generator-produzierten Radiotracer, und andererseits die Gewebektivitätsmessung, für die in der Regel das Zuckermanalogon Fluor-18-FDG, das nur unter Einsatz eines Teilchenbeschleunigers hergestellt werden kann, als Spürsubstanz genutzt wird. Auf diesen unterschiedlichen Verfahrensweisen beruhen die Hauptanwendungsgebiete für PET in der kardiologischen Diagnostik: dies ist zum einen die Diagnose der koronaren Herzkrankheit (KHK) mittels Perfusionsmessung

und zum andern die Bestimmung der Lebensfähigkeit des Myokards bei Infarktpatienten.

Indikationen im Bereich Neuromedizin (Neurologie / Neurochirurgie / Psychiatrie) für den Einsatz der PET

Bevor PET mit steigender Tendenz im Bereich Onkologie genutzt wurde, konzentrierten sich die meisten wissenschaftlichen und klinischen Anwendungen auf neurologische und psychiatrische Erkrankungen. Das funktionelle Bildgebungsverfahren mittels PET bietet gegenüber den konventionellen anatomischen Verfahren den Vorteil, daß Einblicke in die biochemischen und physiologischen Eigenschaften einer Gehirnläsion möglich sind und somit ein Rückschluß auf die funktionelle Unversehrtheit der umgebenden Geweberegionen gezogen werden kann. Deshalb ist im Anwendungsbereich Neuromedizin in vielen Fällen eine genauere Diagnose degenerativer und neoplastischer Krankheiten des Gehirns zu erzielen als mit strukturgebenden Verfahren, wie beispielsweise der Kernspin-Tomographie.

Im Rahmen einer 1997 in Köln veranstalteten Konsensuskonferenz der DGN wurde eine Indikationsliste zum klinischen Einsatz der PET bei neurologischen und psychiatrischen Erkrankungen erarbeitet. Als Grundlage für die Bewertung der medizinischen Effektivität der PET dienten Expertisen der Konferenzteilnehmer, basierend auf der damals aktuellen internationalen Fachliteratur und eigenen Erfahrungen, sowie schriftliche Stellungnahmen von Neurologen und Psychiatern. Für folgende Indikationen wurde der klinische Einsatz der Neuro-PET als angemessen oder akzeptabel eingestuft, da ihr medizinischer Nutzen als erwiesen oder aufgrund der bekannten Datenlage als wahrscheinlich beurteilt wurde (Leitlinien der deutschen Gesellschaft für Nuklearmedizin, Konsensuspapier Neuro-PET 1998):

- frühe (Differential)diagnose von Basalganglienerkrankungen (Multisystemdegenerationen, Huntingtonsche Krankheit und Parkinsonismus)
- präoperative Epilepsiediagnostik
- Früh- und Differentialdiagnose von Demenzerkrankungen
- präoperative Diagnostik und Differenzierung von Hirntumoren
- Abgrenzung der kognitiven Beeinträchtigung bei Depression versus Demenz.

Die Wertigkeit der PET als Diagnoseinstrument bei zerebrovaskulären Erkrankungen und Schizophrenien konnte dagegen bisher nicht ausreichend belegt werden. Diese Einstufung gilt nur beim Einsatz von PET-Ganzkörpersystemen und erfolgte ohne Berücksichtigung der Wertigkeit von SPECT.

Eine genauere Diagnosestellung ist für eine angemessene Patientenbetreuung wichtig und kann bei Krankheiten, für die es eine wirksame Therapie gibt, wie beispielsweise für Hirntumore oder akinetisch rigide Syndrome, für die weitere Therapieplanung und somit letztlich für den Therapieoutcome entscheidend sein (Kuwert et al.

1996). Für viele neuropsychiatrischen Störungen stehen bisher jedoch keine effektiven Therapien zur Verfügung (Robert & Milne 1999), so daß eine bessere Diagnostik nicht unmittelbar zu einer Verbesserung im medizinischen Outcome führt.

C.6.1 Gesundheitsökonomische Evidenz der Studien

Die Entscheidung für die Aufnahme in die deutschen Konsensuslisten, und damit für oder gegen die Nutzung der PET in einem bestimmten Indikationsgebiet, beruht bisher allein auf den medizinischen Ergebnissen von klinischen Studien. Da in der Vergangenheit zur PET-Thematik nur eine sehr begrenzte Zahl an gesundheitsökonomischen Studien in der internationalen Literatur vorlag, konnten gesicherte Ergebnisse zur Kosteneffektivität nicht in den Prozeß der Entscheidungsfindung durch die Konsensuskonferenzen eingebunden werden. Im Rahmen dieses HTA-Reports wurden zur PET-Thematik anhand einer umfangreichen Literaturrecherche (bis Juli 1999, Nachrecherche März 2000) 27 verwertbare Veröffentlichungen zur Evaluation der Kosteneffektivität (22 ökonomische Evaluationen, 1 HTA-Bericht; siehe Literaturverzeichnis Teil I) bzw. zum Ressourcenverbrauch (4 indikationsunabhängige Kostenberechnungen) gefunden.

Um die Evidenz zur Kosteneffektivität einer (neuen) Technologie anhand von Publikationen einschätzen zu können, ist die methodische Qualität der veröffentlichten ökonomischen Evaluationen zu beurteilen. Ziel des vorliegenden HTA-Berichtes war es, sowohl eine Übersicht über die bis März 2000 verfügbare ökonomische Originalliteratur zu geben, als auch die methodische Qualität der gesundheitsökonomischen Evaluationsstudien mit Hilfe des Transparenz- und des Qualitätskataloges, der von der ökonomischen Arbeitsgruppe der „German Scientific Working Group Technology Assessment for Health Care“ erarbeitet wurde (Leidl et al. 1999), zu erfassen und zu bewerten. Die Hauptergebnisse dieser nach methodischen und ökonomischen Aspekten ausgerichteten Bewertungen werden in den nachfolgenden drei Abschnitten - nach den Hauptindikationsgebieten Onkologie, Kardiologie und Neuromedizin getrennt - zusammengefaßt und diskutiert.

Zusätzlich ist bei der Beurteilung einer Technologie, die sich hinsichtlich technischer Möglichkeiten rasch weiterentwickelt, auf die Aktualität der untersuchten Studien Wert zu legen, wenn die daraus extrahierten Informationen für den aktuellen Entscheidungskontext Relevanz haben sollen. Die im vorliegenden HTA-Bericht einbezogenen Originalveröffentlichungen von ökonomischen Studien wurden mit Ausnahme der Konferenzberichte des Institute for Clinical PET (ICP) anläßlich des ICP-Meetings 1994 und der Publikation von Gould et al. aus dem Jahr 1989 - in den letzten 5 Jahren publiziert.

Ökonomische Evidenz der Studien im Indikationsgebiet Onkologie

Als nicht-invasives Bildgebungsverfahren mit geringem Komplikationsrisiko für die untersuchten Patienten findet, wie bereits dargestellt, die PET in der Onkologie bei vielen unterschiedlichen Fragestellungen Anwendung. In neuerer Zeit ist zu beobachten, daß sich die PET-Technologie - zumindest an großen medizinischen Zentren - einen festen Platz in der onkologischen Routinediagnostik erobert.

Die zunehmende Verbreitung von PET-Geräten und damit verbunden die wachsende Bedeutung der PET für die klinische Diagnostik förderte die Notwendigkeit, neben der medizinischen auch die ökonomische Effektivität dieser neuen Technologie zu untersuchen. So wurden 11 der 15 im vorliegenden HTA-Bericht eingeschlossenen ökonomischen Veröffentlichungen zu onkologischen Indikationsstellungen - mit Ausnahme der 3 Kongressberichte des amerikanischen Institute for Clinical PET (ICP) aus dem Jahr 1994 (plus deutsche Fassung: Coleman et al. 1995) - erst ab 1996 publiziert.

Die wichtigsten methodischen Merkmale und Ergebnisse der als relevant klassifizierten ökonomischen Studien mit onkologischem Forschungsansatz sind, differenziert nach Indikationsstellungen, in Tabelle 16 zusammengefaßt. Wie daraus zu ersehen ist, sind sechs Veröffentlichungen begleitende ökonomische Analysen zu prospektiven klinischen Studien (Valk et al. 1996 und 1999; Adler et al. 1997, Hoh et al. 1997, von Schulthess et al. 1998, Klose et al. 1999a); darunter auch eine Übersichtsstudie zu 5 unterschiedlichen Indikationen (nicht-kleinzelliges Lungenkarzinom, rezidivierendes kolorektales Karzinom, metastasierende Melanome, Kopf/Nackenkrebs und Hodgkin-Lymphom; Valk et al. 1996). Keine dieser klinischen Studien erfüllt die strengen Bedingungen des Goldstandards einer prospektiven, verblindeten (hinsichtlich der unabhängigen Verwertung von Informationen aus Paralleluntersuchungen mit unterschiedlichen Diagnosemethoden) Multicenter-Studie mit hoher Fallzahl. Eine Randomisierung ist bei bildgebenden Verfahren allerdings nicht erforderlich, da die Patienten gleichermaßen mit allen diagnostischen Alternativmethoden untersucht werden können, und somit ihre eigene Vergleichskontrolle bilden. Bei den restlichen Evaluationen handelt es sich um mathematische Modellierungen (3 Kongressberichte des Institute for Clinical PET, 1994) oder um theoretische, entscheidungsanalytische Modelle (Gambhir et al. 1996 und 1998, Scott et al. 1998, ECRI-Studie 1998, Kosuda et al. 2000).

Entsprechend der Vielzahl von Anwendungsmöglichkeiten der PET gerade im Bereich Onkologie verteilt sich die noch relativ kleine Zahl an ökonomischen Studien auf mehrere zentrale Indikationsgebiete. Deshalb sind, wie aus Tabelle 16 hervorgeht, derzeit pro Indikationsgebiet meist nur wenige ökonomische Studien zugänglich, die sich zudem in der Methodik (begleitende ökonomische Analyse zu einer klinischen Studie oder theoretische Modelle) unterscheiden. Ökonomische Untersuchungen zu anderen Applikationsgebieten der PET in der Onkologie, wie Primärstaging von Sarkomen oder Speiseröhrenkrebs, Abklärung von verdächtigen Knochen-

läsionen, Kontrolle auf Rezidiv- und Metastasenbildung bei Schilddrüsenkarzinom oder Tumorkontrolle bei Chemo- oder Strahlentherapie, liegen zur Zeit nicht vor.

Die ersten ökonomischen Untersuchungen zur PET im Bereich Onkologie, die mathematischen Modellierungen des Institute for Clinical PET (ICP) aus dem Jahr 1994, beschäftigen sich mit der klinischen Anwendung der PET und den daraus resultierenden ökonomischen Auswirkungen auf das Patientenmanagement bei der Beurteilung von solitären Lungenknoten (Coleman et al. 1994, deutschsprachige Veröffentlichung: Coleman et al. 1995), rezidivierendem kolorektalem Karzinom (Larson et al. 1994) und der Untersuchung von Brustkrebspatientinnen auf axilläre Lymphknotenmetastasen (Adler et al. 1994). Gemäß den Ergebnissen dieser Studien könnten durch den Einsatz von PET Verbesserungen im Patientenmanagement erzielt werden, indem operative Eingriffe aufgrund der höheren diagnostischen Treffsicherheit der PET im Vergleich zu den untersuchten Alternativtechnologien vermieden würden. Damit verbunden wäre eine Reduktion der für die Akutversorgung (Diagnose und invasive Therapie) aufzubringenden Gesamtinterventionskosten. Bei der Bewertung der Aussagekraft dieser Ergebnisse im Hinblick auf die gesundheitsökonomische Wertigkeit der PET in der klinischen Diagnostik ist jedoch die eingeschränkte methodische Qualität der Studien zu beachten, die in den jeweiligen Studienbeschreibungen und -bewertungen diskutiert (siehe C.5.4.1) und in Tabelle 16 am Ende des Ergebnisteils (siehe C.5.4) nochmals zusammenfassend dargestellt wurde.

Andere als relevant identifizierte Veröffentlichungen beleuchten ökonomische Aspekte des Einsatzes der PET in den Indikationsgebieten Diagnostik von Kopf/Nacken-Krebs, Staging von malignen Melanomen, Diagnose und Staging von rezidivierendem kolorektalem Karzinom und Screening von axillären Lymphknotenmetastasen bei Brustkrebspatientinnen (Valk et al. 1996, Valk et al. 1999, Adler et al. 1997). Für diese Studien dokumentieren die Einschätzungen nach dem Transparenz- und Qualitätskatalog einige Defizite in der methodischen Qualität (siehe die Tabellen 6 und 7 und die jeweiligen Studienbeschreibungen unter Kapitel C.5.4.1). Zudem muß die Aussagefähigkeit der Studien von Valk et al. (1996) und Adler et al. (1997) kritisch vor dem Hintergrund der niedrigen Fallzahlen und der daraus resultierenden zweifelhaften Repräsentativität der Ergebnisse, die eine Generalisierbarkeit nicht zulassen, gesehen werden. Das Argument der zu kleinen Fallzahl greift auch bei der Beurteilung der Datenlage im Hinblick auf die Verwendung der PET als Stagingmethode bei malignen Lymphomen (Valk et al. 1996, Hoh et al. 1997, Klose et al. 1999a). In diesen drei prospektiven Studien, welche die gesundheitsökonomische Effektivität von FDG-Ganzkörper-PET für diese spezielle Indikationsstellung im direkten Vergleich zum konventionellen Alternativverfahren CT (eventuell ergänzt durch zusätzliche Diagnostik) evaluieren, konnten insgesamt nur 70 Studienteilnehmer rekrutiert werden. Dabei ist gerade die Studie mit der höchsten Fallzahl ($n = 30$, Valk et al. 1996) aufgrund deutlicher methodischer Mängel von sehr eingeschränkter Aussagekraft. Deshalb ist auch für die Indikation „Staging von malignen Lymphomen“ aufgrund der schmalen Datenbasis und aufgrund methodischer Mängel der wenigen,

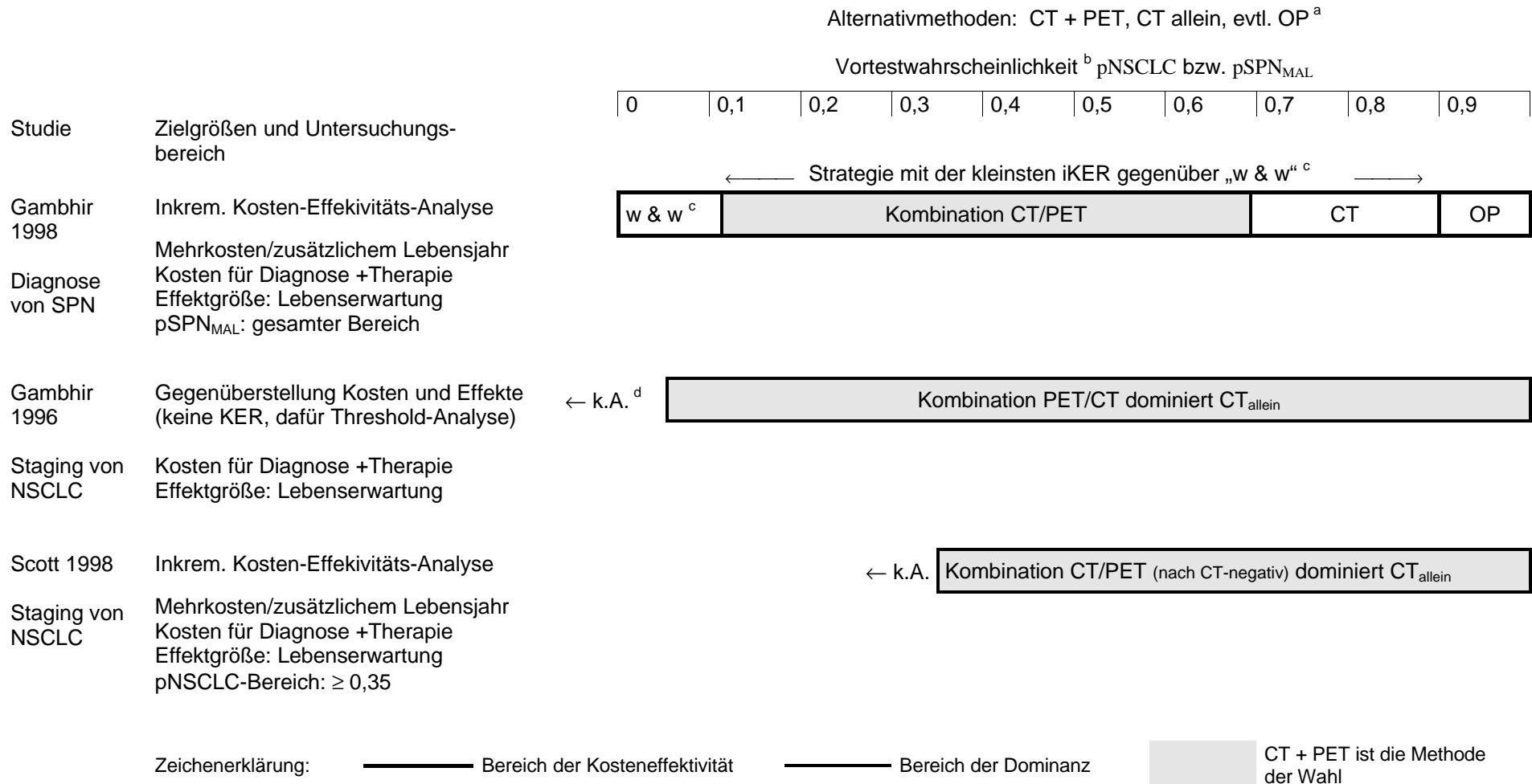
verfügbaren Studien (siehe Tabelle 16) derzeit eine Einschätzung zur Kosteneffektivität der PET problematisch.

Acht - und damit die meisten - Studien untersuchen den Stellenwert der PET aus gesundheitsökonomischer, das heißt, sowohl aus medizinischer als auch aus ökonomischer Sicht, beim Management von Tumorerkrankungen der Lunge (zwei Studien zu solitären Lungenknoten und sechs Studien zu nicht-kleinzelligem Lungenkarzinom). Zwei Studien (Valk et al. 1996, von Schulthess et al. 1998) basieren auf einer klinischen Studie, bei den restlichen Studien handelt es sich um (entscheidungs)theoretische Modelle. Die zwei begleitenden, partiellen ökonomischen Evaluationen zu klinischen Studien, ebenso die bereits angesprochene theoretische Modellierung des Institute for Clinical PET (Coleman et al. 1994), weisen aufgrund methodischer Einschränkungen eine begrenzte Aussagekraft auf (siehe Tabelle 16) und bleiben deshalb von der weiteren Diskussion der Ergebnisse zum Einsatz der PET als Instrument zur Diagnostik von Lungenkrebskrankungen ausgeschlossen.

Die restlichen fünf Studien sind aus gesundheitsökonomischer Sicht als methodisch hochwertig zu betrachten (Gambhir et al. 1996, Gambhir et al. 1998, Scott et al. 1998, ECRI-Bericht 1998, Kosuda et al. 2000) und sind zudem - mit Ausnahme der Evaluation von Gambhir et al. aus dem Jahr 1996 - alle neueren Datums. Diese fünf entscheidungstheoretischen ökonomischen Analysen (Entscheidungsbäume) mit der Thematik Differentialdiagnostik zwischen benignen und malignen Läsionen und/oder Staging (Überprüfung auf Metastasenbildung des Primärtumors) von nicht-kleinzelligem Bronchialkarzinom (NSCLC) und Diagnose und Versorgung von solitären Lungenknoten (SPN) stellen die konventionelle CT als alleinige Testmethode einer kombinierten CT plus ^{18}F FDG-PET-Strategie vergleichend gegenüber. Durch Ausweitung des Zeitfensters über die Phase der Diagnosestellung und operativen Behandlung hinaus wird in allen fünf Studien mit der Berechnung der Lebenserwartung eine aussagekräftige finale Effektgröße (siehe dazu C.4.2.2) in die ökonomischen Evaluationen einbezogen. Anhand der Ergebnisse, die durch Sensitivitätsanalysen auf ihre Robustheit untersucht werden, ergeben sich Hinweise auf eine Kosteneffektivität für den kombinierten Einsatz der bildgebenden Verfahren CT und PET als Hilfsmittel beim Staging und bei der Therapieentscheidung für oder gegen einen chirurgischen Eingriff bei den Indikationsstellungen „Diagnose von solitären Lungenknoten“ bzw. „Staging von nicht-kleinzelligem Lungenkarzinom“.

Die Qualität von theoretischen Modellierungen, die auf in der Fachliteratur veröffentlichte Fremddaten aus klinischen Studien zurückgreifen, ist ihrerseits von der Qualität und Repräsentativität der verwandten Basisdaten abhängig. In den Modellstudien ist jedoch in der Regel das Studiendesign dieser Primärstudien, beispielsweise die Fallzahlen oder Fallmischung, meist nicht dokumentiert. Da eine Sichtung des Ausgangsdatenmaterials im Rahmen dieses HTA-Berichtes nicht zu leisten war, kann über dessen Qualität und Repräsentativität keine Aussage gemacht werden. Dies gilt für alle hier vorgestellten theoretischen Modellstudien.

Abbildung 1: Abhängigkeit der Kosteneffektivität der PET bei den Indikationen „Staging von nicht-kleinzelligem Lungenkarzinom“ bzw. „Diagnose von solitären Lungenknoten“ (Indikationsgebiet: Onkologie) von der Vortestwahrscheinlichkeit für Metastasenbildung durch den NSCLC-Primärtumor (p_{NSCLC}) bzw. der Vortestwahrscheinlichkeit für maligne SPN ($p_{SPN_{MAL}}$) *



a: operativer Eingriff zur Entfernung des Tumorgewebes / b: Vortestwahrscheinlichkeit des einzelnen Patienten, Metastasen (NSCLC), speziell in den regionären Lymphknoten, bzw. maligne SPN zu haben (Bevölkerungsniveau: Prävalenz) / c: wait & watch-Strategie / d: keine Angabe, * Eine detaillierte Aufstellung der in Abbildung 1 verwandten Daten gibt Anhang A

Das von Gambhir et al. (1998) veröffentlichte entscheidungstheoretische Modell zur Versorgung von Patienten mit Verdacht auf maligne solitäre Lungenknoten liefert erste Belege für eine mögliche Kosteneffektivität von ^{18}F FDG-PET in Kombination mit der konventionellen CT-Technologie. In dieser inkrementellen Kosten-Effektivitäts-Analyse ergibt sich über eine breite Prävalenzspanne von 0,12 bis 0,69, die gleichbedeutend damit ist, daß ein einzelner Patient aufgrund seiner persönlichen Merkmale und Risikofaktoren mit einer Vortestwahrscheinlichkeit von 12% bis 69% einen malignen, solitären Knoten in der Lunge hat ($p\text{SPN}_{\text{MAL}}$), für die CT/PET-Kombinationsstrategie im Vergleich zu CT als einzigem nicht-invasiven Test ein Einsparpotential von US\$ 91 bis US\$ 2 200 pro Patient (in Abhängigkeit von der Vortestwahrscheinlichkeit des Patienten je nach Alter, Geschlecht, Risikofaktoren, z.B. Rauchgewohnheiten, Krankengeschichte usw.) bei gleichzeitig erhöhter Lebenserwartung. Hochgerechnet auf das US-amerikanische Bevölkerungsniveau beläuft sich das jährliche Einsparpotential durch den additiven Einsatz der PET bei dieser Patientengruppe auf annähernd US\$ 62,7 Millionen. Dabei ist für $p\text{SPN}_{\text{MAL}} < 0,44$ die Kombination CT/PET gegenüber nur CT dominant, da zur Erzielung eines höheren Zuwachses an Lebenserwartung geringere Mehrkosten aufgebracht werden müssen (Referenzstrategie: wait and watch). Für das $p\text{SPN}_{\text{MAL}}$ -Intervall 0,45 - 0,69 ist CT/PET im Vergleich zu nur CT kosteneffektiv (bei einer gesellschaftlich akzeptierten $i\text{KER} \leq \text{US\$ } 50\,000$). Für Patienten mit höheren Wahrscheinlichkeiten für maligne SPN von 0,7 - 0,9 dagegen ist der alleinige Einsatz der konventionellen CT-Technologie bzw. für Werte $> 0,9$ die direkte Überstellung der Patienten zur Operation die Methode der Wahl, da diese Vorgehensweisen die kleinsten inkrementellen Kosten-Effektivitäts-Relationen aller Alternativstrategien in Relation zur Basisstrategie „wait and watch“ über die $p\text{SPN}_{\text{MAL}}$ -Spanne von 0,7 bis 1,0 ausweisen. Diese Abhängigkeit der Kosteneffektivität der Alternativstrategien von der Vortestwahrscheinlichkeit für maligne SPN ist in Abbildung 1 veranschaulicht.

In der ökonomischen Analyse von Gambhir et al. (1996) zur Indikationsstellung „Staging von NSCLC“ berechnen die Autoren bei einer konservativen PET/CT-Diagnostik (Biopsie bei PET-positiven Patienten unabhängig vom Ausgang der CT-Untersuchung) pro NSCLC-Patient ein Einsparpotential von US\$ 1 154 ohne Verlust an Lebenserwartung gegenüber dem alternativen, alleinigen CT-Einsatz. Bei einer weniger konservativen PET/CT-Diagnostik (Biopsie nur bei von einander abweichenden Ergebnissen der CT- und der PET-Untersuchung) ergäben sich Minderausgaben von US\$ 2 262 pro Patient, wobei jedoch 1,7% der potentiell operablen Patienten unentdeckt blieben und somit keiner adäquaten chirurgischen Behandlung zugeführt würden. Die 1998 publizierte Nachfolgestudie von Scott et al. konnte zeigen, daß eine zweistufige CT/PET-Diagnosestrategie (plus Biopsie), in der PET nur zur Überprüfung negativer CT-Befunde eingesetzt würde, bei einer inkrementellen Kosten-Effektivitäts-Ratio von circa US\$ 25 300 pro gewonnenem Lebensjahr eine kosteneffektive Alternative zur alleinigen Untersuchung mit CT (plus Biopsie) sein kann (bei einer in der US-Gesellschaft diskutierten Obergrenze für Mehrkosten in Höhe US\$ 50 000 pro gewonnenem Lebensjahr). Beim Vergleich dieser zwei Studien ist

jedoch zu beachten: (a) ist die zeitliche Anordnung der Diagnosemethoden CT und PET vertauscht (Gambhir et al. 1996: PET/CT-Kombination mit PET als Eingangsuntersuchung; Scott et al. 1998: CT/PET-Kombination mit CT als Eingangsuntersuchung) und (b) handelt es sich bei der früheren Studie von Gambhir et al. (1996) um eine ökonomische Analyse, bei der für jeweils zwei Alternativstrategien (zwei zweiästige Entscheidungsbäume) nur die Kosten und die Effekte - ohne eine Kosteneffektivitätsrelation zu berechnen - verglichen werden (Threshold-Analyse). Bei der Nachfolgestudie von Scott et al. (1998), eine inkrementelle Kosten-Effektivitäts-Analyse, dagegen werden die zusätzlichen Kosten pro zusätzlichem Effekt in Relation zur Referenzstrategie „nur CT (+ Biopsie)“ (siehe Tabelle 16) berechnet.

Die bei der Gambhir-Studie (1998) zur Differentialdiagnostik von SPN beobachtete Abhängigkeit der Kosteneffektivität des kombinierten CT + PET-Einsatzes von der Vortestwahrscheinlichkeit ist in den NSCLC-Studien nicht gegeben. Denn, wie aus Abbildung 1 abzulesen ist, dominiert - im Gegensatz zum Ergebnis der SPN-Studie - im Wertebereich $> 0,7$ die CT + PET-Kombinationsstrategie (plus Biopsie) den alleinigen Einsatz der CT (plus Biopsie). Somit stimmt nur im mittleren Wertebereich $0,35 - 0,7$ (Gambhir et al. 1996 und 1998, Scott et al. 1998) bzw. im unteren Wertebereich zwischen $0,12$ und $0,35$ (Gambhir et al. 1996 und 1998) die Aussage, daß die CT + PET-Kombinationsstrategie die Methode der Wahl ist (siehe Abbildung 1), für beide Indikationen, SPN und NSCLC, überein. Dabei ist jedoch zu beachten, daß für die Studie von Scott et al. (1998) für pNSCLC-Werte $< 0,35$ keine Angaben vorliegen, und daß sich die ökonomische Analyse von Gambhir et al. (1996) methodisch von den späteren inkrementellen Kosten-Effektivitäts-Analysen unterscheidet, da keinerlei Kosteneffektivitätsrelation berechnet wird (siehe die tabellarische Auflistung der Studienmerkmale in Abbildung 1). Eine Gegenüberstellung der Studienergebnisse ist trotzdem vertretbar, weil bei den Gesamtkostenberechnungen die gleichen Kostenkomponenten berücksichtigt werden und eine einheitliche Beschreibung der Effektseite durch die Outcomegröße Lebenserwartung erfolgt (siehe Abbildung 1).

Zur gleichen Empfehlung wie die Studie von Scott et al. (1998), PET zur Bestätigung eines CT-negativen Vorbefundes im Stagingschritt additiv in den derzeit üblichen, konventionellen CT-basierten Algorithmus zur Diagnose und Staging von Patienten mit NSCLC-Verdacht einzuschieben, kommt die methodisch ebenfalls hochwertige Studie des Emergency Care Research Institute (ECRI) aus dem Jahr 1998. In dieser komplexen Studie ist die ebenfalls im Stagingschritt durch eine PET-Untersuchung ergänzte CT/Biopsie-Strategie (erweitertes Routineverfahren) dominant über das derzeitig verwandte Routineverfahren, bestehend aus einer CT-Untersuchung mit anschließender Biopsie sowohl im Diagnose- als auch im Stagingschritt (siehe Studienbeschreibung). Laut Modell läßt sich pro untersuchtem Patienten ein Einsparpotential von US\$ 36 mit einem Zugewinn an Lebenserwartung von 1,8 Tagen realisieren (Durchschnittsbetrachtung). Bezogen auf das US-amerikanische Bevölkerungsniveau würde sich somit - bei jährlich circa 130 000 Patienten mit Verdacht auf NSCLC - das nationale Einsparpotential auf einen Betrag von US\$ 4,7 Millionen pro Jahr aufsummieren.

Die neueste Studie zur Evaluation der Kosteneffektivität der PET im Rahmen der Versorgung von Patienten mit verdächtigen pulmonalen Läsionen ist die inkrementelle Kosten-Effektivitäts-Analyse von Kosuda et al. (2000). Hier berechnen die Autoren für eine kombinierte CT/PET-Strategie (Teilkörperuntersuchungen im Brustbereich) pro Patient einen Zugewinn von 7,3 Monaten verbunden mit einer inkrementellen Kosten-Effektivitäts-Ratio von 218 000 Yen (etwa US\$ 1 550) pro zusätzlichem Lebensjahr. Die Studie ist auf die spezifischen Konditionen des japanischen Gesundheitssystems zugeschnitten, so werden beispielsweise die in Japan üblichen niedrigen Vergütungssätze für medizinische Leistungen veranschlagt (PET-Vergütungsbetrag: circa US\$ 700). Aus diesem Grunde ist das Ergebnis nicht unmittelbar mit den Ergebnissen der US-Studien vergleichbar.

Das Problem der Übertragbarkeit von zumeist US-amerikanischen Studienergebnissen auf andere Gesundheitssysteme ergibt sich auch für Deutschland. Dabei können sowohl im Mengen- als auch im Preisgerüst gravierende Unterschiede auftreten, die zum Beispiel durch verschiedene Organisationsformen der prä- und postoperativen Versorgung von Patienten oder durch Berücksichtigung unterschiedlicher Kostenkomponenten bei der Festlegung administrativer Vergütungen (zum Beispiel Investitionskosten) hervorgerufen werden können. So ergeben sich beispielsweise durch das zumeist höhere Kostenniveau für medizinische Leistungen in den USA Schwierigkeiten auf der Kostenseite. Aber auch auf der Effektseite können sich durch unterschiedliche Fallmischungen, Prävalenzen und Lebensstilvariablen, wie das Rauchverhalten, länderspezifische Abweichungen ergeben. Eine Hilfestellung zur Beurteilung, ob die Ergebnisse einer ökonomischen Evaluation aus dem Ausland sinnvoll auf Deutschland übertragen werden können, gibt ein Beitrag von Welte und Leidl (1999).

Anhand der von Welte und Leidl (1999) entworfenen Prüfliste zur Übertragbarkeit der Ergebnisse ökonomischer Evaluationen aus dem Ausland auf Deutschland wurden die Resultate der vier im Rahmen dieses HTA-Berichtes verwertbaren US-amerikanischen entscheidungstheoretischen Modelle zur Diagnostik von Lungenerkrankungen beurteilt (Gambhir et al. 1996 und 1998, Scott et al. 1998, ECRI-Bericht 1998). Anhand dieses Prüfkataloges ergab sich eine nur mittlere Übertragbarkeit der vorgestellten Ergebnisse dieser ökonomischen Evaluationen aus dem Studienland USA auf das Entscheidungsland Deutschland, so daß zwar keine quantitative aber dennoch - wenn auch sehr begrenzte - qualitative Übertragbarkeit gegeben ist. Dabei war die Übertragbarkeit der neueren Studien aus dem Jahre 1998 geringfügig besser als die ältere Studie von Gambhir et al. aus dem Jahr 1996. Dies ist auf eine Korrektur des Preisgerüsts (u.a. Senkung der veranschlagten Vergütung für operative Therapiemaßnahmen) zurückzuführen. Um die länderspezifische Übertragbarkeit zu verbessern, sollte zumindest eine Neubewertung des Preisgerüsts und - soweit dies vom Arbeitsaufwand vertretbar ist - eine Anpassung des Mengengerüsts an das deutsche Entscheidungsumfeld vorgenommen werden. Außerdem ist eine zeitliche und währungsspezifische Adjustierung der Studienergebnisse an die aktuellen deutschen Gegebenheiten in Form einer Inflations-/Deflationsbereinigung

bzw. Umrechnung der Kostenangaben in DM mit Hilfe von Kaufkraftparitäten notwendig.

Zusammenfassend ist festzuhalten, daß sich aufgrund der Ergebnisse dieser aus methodischer Sicht als gut bis sehr gut einzustufenden Studien Hinweise auf einen kosteneffektiven - additiven - Einsatz der PET, also zusätzlich zum konventionellen CT-basierten Verfahren, in den Indikationen „Staging von nicht-kleinzelligem Lungenkarzinom“ und „Diagnose von solitären Lungenknoten“ ergeben. Da diese Studien jedoch sowohl im Studiendesign (inkrementelle oder durchschnittliche Untersuchung) als auch in der zeitlichen Anordnung der diagnostischen Einzelbausteine CT und PET (CT/PET- bzw. PET/CT-Reihenfolge) differieren, sind die Resultate der Studien nicht uneingeschränkt vergleichbar. Um eine gesicherte Aussage zur ökonomischen Wertigkeit der PET in diesen Anwendungsfeldern machen zu können, sollten deshalb die bisherigen Erkenntnisse durch weitere, methodisch hochwertige ökonomische Evaluationen von vergleichbarem Studiendesign (u.a. inkrementeller Ansatz, langfristiger Zeithorizont und Erfassung von patientenbezogenen Effekten und Langzeitkosten, Diskontierung der Kosten und Effekte, Prüfung der Stabilität der Resultate anhand von Sensitivitätsanalysen) bestätigt werden.

Zu allen anderen klinischen Anwendungsmöglichkeiten der PET im Bereich der Onkologie liegen entweder keine ökonomischen Evaluationen vor oder die vorhandenen Analysen sind aufgrund methodischer Defizite von geringer Evidenz.

Ökonomische Evidenz der Studien im Indikationsgebiet Kardiologie

Während die Bedeutung der PET für die onkologische Diagnostik in den letzten Jahren zunahm, besteht in Fachkreisen keine einheitliche Einschätzung über ihre zukünftige Rolle im Indikationsgebiet Kardiologie. Gründe hierfür sind die laufende technische Weiterentwicklung der konventionellen nuklearmedizinischen Technologien zur kardialen Perfusionsmessung und der fehlende Nachweis, daß der geringe Zuwachs an diagnostischer Genauigkeit durch den additiven oder substitutiven Einsatz der PET letztlich von Bedeutung für die Diagnostik und Versorgung koronarer Herzkrankheiten ist. So stellte sich bei einer Untersuchung der ECRI (Emergency Care Research Institute, HTA-Einrichtung der USA) heraus, daß - obwohl PET-Aufnahmen von höherer technischer Qualität sind als SPECT-Aufnahmen (besserer Kontrast und höhere räumliche Auslösung) - in der klinischen Alltagssituation PET eine vergleichbare oder nur geringfügig bessere Leistungsfähigkeit als Diagnoseinstrument bei koronarer Herzkrankheit vorweisen konnte als die alternative Routinemethode SPECT (Bailey Merez & Berman 1997). Nach Begutachtung der verfügbaren Daten stufte die American Heart Association den Stellenwert der PET in der kardiologischen Diagnostik nicht höher ein als denjenigen der SPECT (Robert & Milne 1999).

Im Rahmen der Literaturrecherche für den vorliegenden HTA-Bericht wurden 7 Publikationen zu ökonomischen Studien ermittelt, die sich mit dem Kosteneffektivitätsvergleich zwischen der PET-Technologie und konventionellen, bildgebenden Diagnoseverfahren, wie dem in der kardiologischen Diagnostik üblicherweise genutzten Goldstandard der Koronarangiographie (mittels Herzkatheter) oder der nuklearmedizinischen Alternativtechnologie SPECT, beschäftigen. Im Gegensatz zu den onkologischen Studien handelt es sich bei allen kardiologischen Evaluationen um theoretische Modelle, darunter mathematische Modellierungen (Besozzi et al. 1994, Patterson et al. 1994 und 1995, Gould et al. 1989) und entscheidungstheoretische Modelle in Form von Entscheidungsbäumen (Hunink 1995, Maddahi & Gambhir 1997, Garber & Solomon 1999). In Tabelle 17 sind die wichtigsten methodischen Merkmale und Ergebnisse dieser ökonomischen Evaluationen zusammenfassend dargestellt. Begleitende gesundheitsökonomische Studien zu (prospektiven) klinischen Studien zur PET-Thematik fanden sich nicht unter der bis März 2000 veröffentlichten Fachliteratur.

Die frühen Publikationen, die Modellierung des Institute for Clinical PET (Besozzi et al. 1994) und der niederländische Buchbeitrag von Hunink 1995, befassen sich mit der Bestimmung der Lebensfähigkeit des Myokards bei Infarktpatienten mittels ^{18}F FDG-PET. Wegen methodischer Defizite, wie die alleinige Betrachtung der Kostenseite (partielle Evaluation, siehe C.4.2.3), der Einengung des Zeit- und Kostenrahmens auf die Phase der Akutversorgung und fehlender Sensitivitätsanalysen (siehe Einzelbeschreibung), ist diese ICP-Studie jedoch nur von begrenzter Aussagekraft. Die 1995 als Buchbeitrag publizierte Kosten-Effektivitäts-Analyse von Hunink (Entscheidungsbaum mit Markov-Modell) ist aufgrund methodischer Einschränkungen, bedingt vor allem durch die sehr knappe Darstellung der Modellkonditionen und einer schmalen Datenbasis, auch nach Einschätzung der Autorin - als vorläufiger Entwurf eines ausbaufähigen entscheidungstheoretischen Modells zu werten. Da derzeit keine weiteren Publikationen zur Bestimmung der Myokardviabilität mittels ^{18}F FDG-PET verfügbar sind, kann zum jetzigen Zeitpunkt aus gesundheitsökonomischer Sicht keine Einschätzung zur Kosteneffektivität der PET in diesem speziellen Anwendungsbereich gegeben werden.

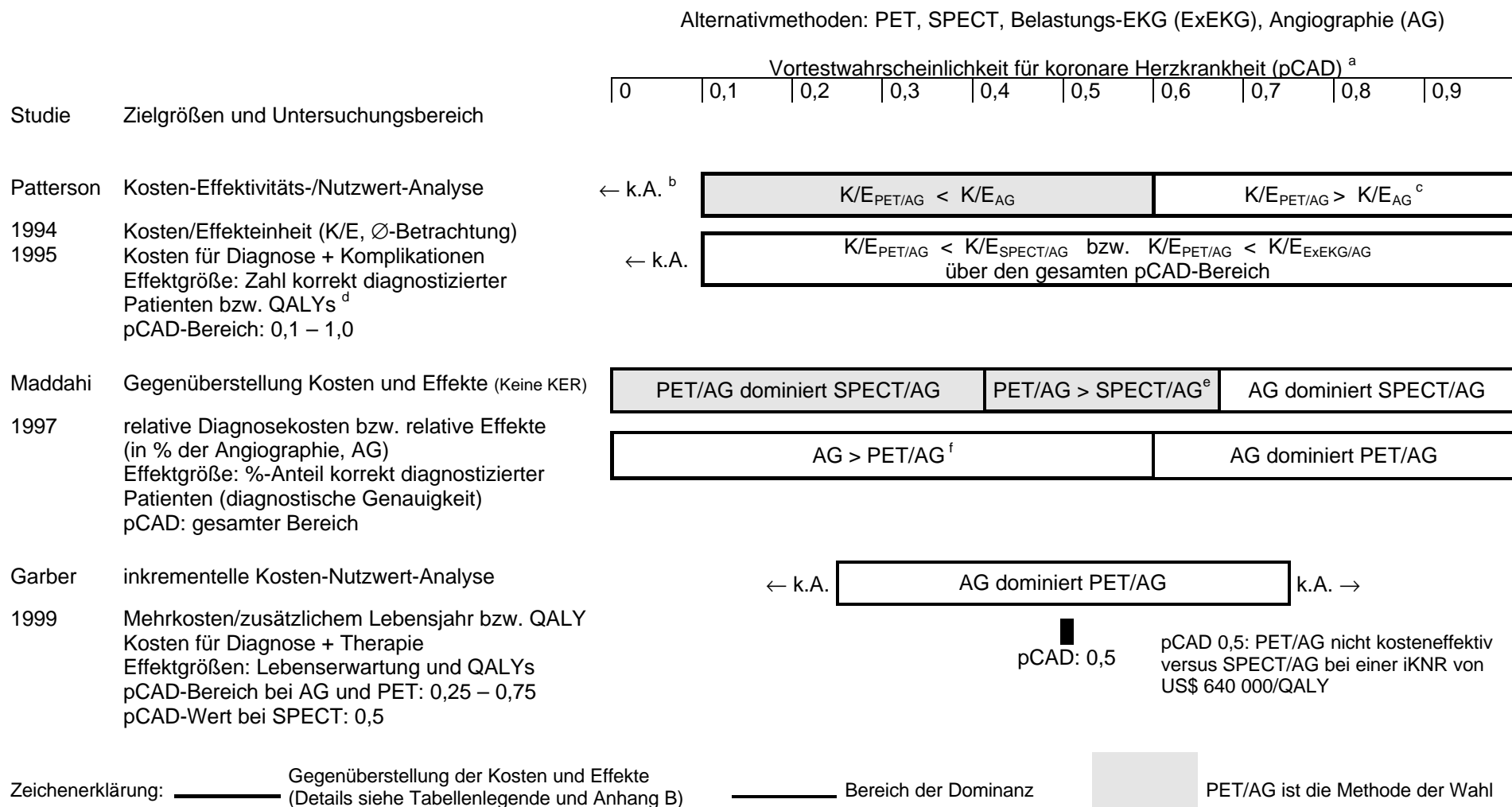
Fünf Artikel analysieren den gesundheitsökonomischen Stellenwert der PET als Instrument zur Diagnose koronarer Herzkrankheit (KHK) mittels Perfusionsmessung. Die bereits 1989 publizierte ökonomische Studie von Gould et al., eine partielle Evaluation, ist trotz fehlendem aktuellen Bezug und methodischer Einschränkungen aufgrund von zwei Besonderheiten dennoch bemerkenswert: Die mit einem Zeitfenster von 5 Jahren über die akute Patientenversorgung hinausreichende Kostenberechnung berücksichtigt außer den direkten Interventionskosten auch indirekte Kosten infolge Produktionsausfall durch vorzeitigen Tod (Humankapitalansatz), und zeichnet sie sich außerdem durch eine detaillierte, betriebliche Kostenkalkulation der beim Erwerb und Betrieb einer PET-Anlage aufzubringenden Finanzmittel aus. Dennoch bleiben die Ergebnisse dieser Studie aufgrund der fehlenden Aktualität und der ein-

geschränkten methodischen Gesamtqualität der Evaluation in der weiteren Diskussion unberücksichtigt.

Die späteren ökonomischen Evaluationen von Patterson et al. (1994 und 1995), Maddahi & Gambhir (1997) und Garber & Solomon (1999) zur Nutzung der PET als Instrument zur Messung der kardialen Perfusion zeichnen sich durch eine gute bis sehr gute methodische Qualität aus (siehe Studienbeschreibungen). Garber & Solomon (1999) vergleichen sechs alternative Diagnosestrategien in einer inkrementellen Kosten-Nutzwert-Analyse miteinander. Dies sind die Kombinationsstrategien (mit nicht-invasivem Eingangstest) Belastungs-EKG/Angiographie, Thallium-201-Szintigraphie/ Angiographie, Ultraschall-Echokardiographie/Angiographie, SPECT/Angiographie, PET/Angiographie und die direkte Angiographie (ohne nicht-invasiven Vortest). Dabei variiert die Lebenserwartung eines 55jährigen Mannes mit einer Vortestwahrscheinlichkeit für KHK von 50% (Basisannahme) nur wenig mit der Methodik des diagnostischen Eingangstests; so beträgt der Gewinn an Lebenserwartung maximal 7 Tage. Die inkrementelle Kosten-Nutzwert-Relation, als Maß für die Kosteneffektivität der einzelnen Vorgehensweisen, beläuft sich beim Vergleich der zweistufigen Algorithmen PET/Angiographie versus SPECT/Angiographie - und einem realistischen Vergütungsbetrag pro PET-Scan von US\$ 1 500 - auf mehr als US\$ 640 000 pro gewonnenem qualitätsadjustiertem Lebensjahr (QALY). Auch bei einer 50%igen Absenkung des PET-Preisniveaus auf US\$ 750 resultiert immer noch ein Kosten-Nutzwert-Quotient von über US\$ 180 000/QALY. Sofortige Koronarangiographie - ohne nicht-invasiven Vortest - ist im Modell in allen nach Alter und Geschlecht stratifizierten Kohorten (45, 55 und 65jährige mit einer KHK-Vortestwahrscheinlichkeit von 25 - 75%) gegenüber der PET-basierten Diagnosestrategie dominant (siehe Abbildung 2). Diese Dominanz bleibt auch bei einer Verdreifachung der Mortalitätsrate für die Angiographie bestehen. Gemäß diesen Ergebnissen ist unabhängig vom Alter und Geschlecht des Patienten und von dessen Vortestwahrscheinlichkeit für KHK (pCAD) im Intervall von 0,25 – 0,75 die Angiographie die Methode der Wahl (für die Wertebereiche pCAD < 0,25 und pCAD > 0,75 liegen keine Angaben vor).

Das entscheidungstheoretische Modell von Maddahi & Gambhir (1997) zeigt jedoch, wie aus Abbildung 2 ersichtlich ist, ein anderes Bild auf. Denn in dieser ökonomischen Evaluation ist die Auswahl einer geeigneten Diagnosestrategie gemäß ihrer Kosteneffektivität abhängig von der KHK-Vortestwahrscheinlichkeit der versorgten Patienten. Für höhere pCAD-Werte ist die direkte Angiographie sowohl über die Kombinationsstrategie SPECT/Angiographie (pCAD \geq 0,68) als auch über die Kombination PET/Angiographie (pCAD-Werte > 0,6) dominant. Im unteren pCAD-Bereich (< 0,42) dominiert dagegen die Kombination PET/Angiographie über die Alternative SPECT/Angiographie. Für kleine (PET/AG versus AG) und mittlere (PET/AG versus AG bzw. PET/AG versus SPECT/AG) Vortestwahrscheinlichkeiten dagegen kann keine definitive Aussage, welche Diagnosestrategie kosteneffektiv und somit die Methode der Wahl ist, getroffen werden, da die Autoren keine Kosten-Effektivitäts-Relationen berechnen (nur Gegenüberstellung von relativen Kosten und relativen Effekten).

Abbildung 2: Abhängigkeit der Kosteneffektivität der PET im Indikationsgebiet Kardiologie (Myokardperfusionsmessung) von der Vortestwahrscheinlichkeit für koronare Herzkrankheit (pCAD) *



a: Vortestwahrscheinlichkeit des einzelnen Patienten, an koronarer Herzkrankheit (KHK) erkrankt zu sein (Bevölkerungsniveau: Prävalenz) / b: keine Angabe / c: Aussage gilt außer bei PET/AG auch für die Kombinationen Belastungs-EKG/AG bzw. SPECT/AG / d: qualitätsadjustierte Lebensjahre / e: PET/AG ist effektiver, aber auch teurer als SPECT/AG / f: Angiographie ist effektiver, aber auch teurer als die Kombination PET/AG ; * Eine detaillierte Aufstellung der in Abbildung 2 verwandten Daten gibt Anhang B, Abkürzungen siehe Abkürzungsverzeichnis

Anhand einer weiteren, qualitativ sehr guten Studie, eine mathematische Modellierung der Arbeitsgruppe Patterson et al. (1994 und 1995), wird PET ebenfalls im Vergleich zu den konventionellen nicht-invasiven Methoden Belastungs-EKG und SPECT und dem Goldstandard, der Koronarangiographie, evaluiert. Die nicht-invasiven Testverfahren werden dabei - wie in der Studie von Maddahi & Gambhir (1997) - der Koronarangiographie als Eingangstest vorgeschaltet. Anhand der Studienergebnisse wäre für die gesamte pCAD-Prävalenzspanne (0,1 – 1,0) der kombinierte PET/Angiographie-Algorithmus unter Belastung den Alternativkombinationen SPECT/Angiographie bzw. Belastungs-EKG/Angiographie - in dieser Reihenfolge – und für pCAD-Werte $< 0,6^{*16}$ auch der direkten Angiographie als Eingangstest vorzuziehen. Liegt jedoch die Vortestwahrscheinlichkeit des Patienten an KHK erkrankt zu sein über dem Schwellenwert von $0,6^*$, wäre bei durchschnittlichen Kosten pro korrekt diagnostiziertem Patienten von $< \text{US\$ } 30\,000$ bzw. durchschnittlichen Kosten pro QALY $< \text{US\$ } 10\,000$ (beachte Durchschnittsbetrachtung, keine inkrementelle Analyse) im Modell die direkte Angiographie die Methode der Wahl (siehe Abbildung 2).

Beim Vergleich der drei Studien im Hinblick auf die Abhängigkeit der Kosteneffektivität der PET von der Vortestwahrscheinlichkeit des einzelnen Patienten, die ihrerseits wiederum mit dessen Alter, Geschlecht, Risikofaktoren und seiner Krankengeschichte variiert, ist jedoch - wie bei den onkologischen Studien - das unterschiedliche Studiendesign bezüglich Studienart und Zielgrößen auf der Kosten- bzw. der Effektseite zu beachten. So stellen Garber & Solomon (1999) eine inkrementelle Kosten-Nutzwert-Analyse auf, während Patterson et al. (1994 und 1995) nur eine Durchschnittsbetrachtung (Kosten pro Effektgröße) anstellen. Die Ergebnisse einer durchschnittlichen Betrachtung sind dabei weniger aussagekräftig als die Ergebnisse einer inkrementellen Analyse (nach der unter C.4.2.3 dargestellten Regel). Maddahi & Gambhir (1997) berechnen keine Kosten-Effekt-Quotienten, sondern vergleichen nur die relativen Kosten bzw. die relativen Effekte (in % der Angiographie) der Alternativstrategien untereinander. Auch die in den Studien verwandten Effektgrößen unterscheiden sich, wie aus der tabellarischen Aufstellung in Abbildung 2 zu entnehmen ist. So erfassen Maddahi & Gambhir (1997) mit dem „%-Anteil korrekt diagnostizierter Patienten“ nur eine intermediäre, wenig aussagekräftige Effektgröße, während Patterson et al. (1994 und 1995) und Garber & Solomon (1999) mit den Outcomegrößen Lebenserwartung bzw. qualitätsadjustierte Lebenserwartung finale, aussagekräftige Effektgrößen in die Betrachtung einbeziehen.

Zusammenfassend ist festzustellen: Die Ergebnisse der ökonomischen Studien von Patterson et al. (1994 und 1995), Maddahi & Gambhir (1997) und Garber & Solomon (1999) stimmen - wie oben dargestellt - für den Wertebereich der Vortestwahrscheinlichkeit $< 0,6$ in ihrer Aussage zum kosteneffektiven Einsatz der PET als Instrument zur Diagnose koronarer Herzkrankheit nicht überein. Für Patienten mit höheren Vortestwahrscheinlichkeiten ist jedoch die Angiographie in allen drei Studien der Kombi-

¹⁶ Die mit * gekennzeichneten Daten sind aus Abbildung 1 der Publikation Patterson et al. 1995 entnommen.

nation PET/Angiographie vorzuziehen. Die differierenden Resultate der vorgestellten Kardiologiestudien machen deutlich, daß aus ökonomischer Sicht eine Empfehlung zur Nutzung der PET zur Diagnostik bei koronarer Herzkrankheit bei der derzeitigen Evidenz nicht gegeben werden kann. Deshalb besteht zur Abklärung der uneinheitlichen Ergebnisse weiterer Forschungsbedarf.

Ökonomische Evidenz anhand von Studien im Indikationsgebiet Neuromedizin (Neurologie / Neurochirurgie / Psychiatrie)

Obwohl es viele klinische Studien zur Anwendung der PET im Bereich Neuromedizin gibt, verlief die Literaturrecherche nach ökonomischen Evaluationsstudien in diesem Bereich erfolglos. Mit Hilfe der unter Kapitel C.4.1.1 angegebenen Suchstrategien konnten in den durchsuchten Datenbanken keine relevanten ökonomischen Evaluationen ermittelt werden. Dies hat zur Folge, daß für die Neuromedizin - dem derzeit immerhin zweitwichtigsten Anwendungsgebiet der PET in der klinischen Diagnostik (siehe dazu C.6.2) - keinerlei Aussagen zur Kosteneffektivität der PET-Technologie gemacht werden können. Obwohl die Onkologie die neuromedizinische Anwendung der PET von der Spitzenposition verdrängt hat, bleibt die Notwendigkeit zur Durchführung von ökonomischen Evaluationen davon unberührt. Zumal die Differentialdiagnostik der Altersdemenz mit einem steigendem Anteil alter Menschen in der Bevölkerung künftig sowohl aus klinischer als auch aus gesamtgesellschaftlicher Sicht an Bedeutung gewinnen dürfte. Eine Ursache für fehlende ökonomische Evaluationen kann möglicherweise darin gesehen werden, daß für viele neuropsychiatrischen Störungen bisher keine effektiven Therapien zur Verfügung stehen (Robert & Milne 1999).

Nach Einschätzung von Kuwert et al. (1996) stellen die nuklearmedizinischen Bildgebungsverfahren SPECT und PET eine im Einzelfall zwar medizinisch sinnvolle, aber derzeit generell keine kosteneffektive Alternative zu den konventionellen Methoden CT und Kernspin-Tomographie dar. Diese Einschätzung wird mit Hilfe von klinischen und gesundheitsökonomischen Studien zu prüfen sein.

C.6.2 Beurteilung der Kosteneffektivität des klinischen Einsatzes der PET-Technologie insgesamt

Bei einer vollständigen Analyse der Kosteneffektivität einer medizinischen Intervention werden sowohl die Kosten als auch die Effekte von mindestens zwei Alternativtechnologien vergleichend untersucht (siehe C.4.2.3). Beide Größen hängen jedoch von einer Vielzahl von Einflußfaktoren ab. So resultieren die Effekte bei einem apparativen, diagnostischen Verfahren unter anderem von dessen technischer Leistungsfähigkeit; bei einem Bildgebungsverfahren wie der PET beispielsweise vom räumlichen und zeitlichen Bildauflösungsvermögen des Detektorsystems und der

angeschlossenen Computereinheit. Aber auch die Kosten für Beschaffung, Unterhalt und Betrieb eines diagnostischen Gerätes variieren entsprechend der technischen Ausstattung, dem zum Gerätebetrieb benötigten Personal- und Materialbedarf, der Art und der Bezugsquelle des verwendeten Positronenstrahlers (Produktion des Nuklids in einem Zyklotron oder in einem Generator; hausinterne Produktion oder Fremdbezug), dem Auslastungsgrad usw., wie in Abschnitt C.5.2 ausführlich erläutert wird. Wie bei den meisten apparativen Techniken wird PET in unterschiedlichen technischen Ausführungen angeboten: zwei- oder dreidimensional arbeitende Ganzring-Scanner, dreidimensional arbeitende Teilring-Rotations-Scanner und Koinzidenz-Gamma-Kameras, Ganzkörper- und Teilkörper-PET. Da aufgrund der aufgezählten Faktoren die tatsächlichen PET-Kosten entsprechend den jeweiligen Modalitäten ihrer spezifischen Verwendung variieren dürften, wäre, um die Qualität der Evidenz und damit die Aussagefähigkeit von künftigen ökonomischen Evaluationen zu erhöhen, die korrekte Erfassung des Ressourcenverbrauches, das heißt der tatsächlich anfallenden Kosten, wünschenswert. Detaillierte, auf betriebliche Kostenkalkulationen basierende Informationen zum Ressourcenverbrauch und zur Kostenstruktur beim klinischen Einsatz der PET geben für die onkologische Diagnostik Klose et al. (1999a) und für die kardiologische Diagnostik Gould et al. (1989). Erste Kostenbetrachtungen zu Deutschland wurden für die PET-Zentren in Münster und Ulm erarbeitet (Lottes et al. 1998; Klose et al. 1999a).

Aufgrund einer Vielzahl von potentiellen Anwendungsgebieten für PET, der unterschiedlichen methodisch-technischen Spezifikationen je nach Applikationsgebiet (z.B. breites Spektrum an Radiopharmaka) und der möglichen Abhängigkeit der gesundheitsökonomischen Wertigkeit der PET von der Prävalenz einer zu diagnostizierenden Erkrankung (siehe C.5.1) kann keine „globale“ Kosteneffektivität für „die PET“, sondern nur eine anwendungsspezifische Kosteneffektivität zu den einzelnen Indikationen, ermittelt werden. Aus diesem Grund ist bei der Abschätzung des gesundheitsökonomischen Stellenwertes der PET - für ein Gesundheitssystem als Ganzes, aber auch für einzelne Betreibereinrichtungen - zu beachten, wie sich die einzelnen Indikationsgebiete mit unterschiedlicher Kosteneffektivität prozentual auf den Gesamteinsatz der PET-Technologie in der klinischen Praxis verteilen. Interne Daten der Nuklearmedizin am Universitätsklinikum Münster, vorgestellt auf der 37. Jahrestagung der Deutschen Gesellschaft für Nuklearmedizin im Frühjahr 1999 in Ulm, zeigen eine sehr ungleiche Verteilung der Anwendung der PET in den drei Hauptindikationsgebieten Onkologie, Kardiologie und Neuromedizin auf. So lag im Zeitraum November 1995 - Oktober 1996 der prozentuale Anteil der onkologischen Untersuchungen an allen durchgeführten PET-Untersuchungen bei 63%, der neuromedizinische Leistungsanteil bei 28% und der kardiologische Leistungsanteil bei nur 9%. Auch wenn diese internen Leistungsdaten keinen Anspruch auf Repräsentativität für die Gesamtheit der deutschen PET-Zentren erheben können, ist die sich hier abzeichnende Ungleichverteilung auf mehrere Anwendungsbereiche bei der Einschätzung der Wertigkeit der PET aus medizinischem und ökonomischem Blickwinkel zu berücksichtigen. Ein ähnlicher Leistungsanteil des Indikationsbereiches Onkologie

wie am PET-Zentrum Münster (63%-Anteil) zeichnet sich auch international ab: so stellten Robert & Milne (1999) fest, daß derzeit international an den meisten PET-Zentren mehr als 70% der Überweisungen für PET-Untersuchungen von onkologischen Abteilungen veranlaßt werden.

Die Bemühungen, trotz methodischer und inhaltlicher Schwierigkeiten letztlich doch zu einer - indikationsgewichteten - Bewertung der gesundheitsökonomischen Effizienz des klinischen PET-Einsatzes zu kommen, wird durch die permanente technologische Weiterentwicklung sowohl der PET-Technologie als auch alternativer, nicht-invasiver (z.B. SPECT) und invasiver Diagnoseverfahren erschwert.

Abschließend läßt sich sagen, daß wegen der begrenzten Anzahl an in der internationalen Literatur verfügbaren ökonomischen Evaluationen mit guter methodischer Qualität, die sich zudem auf die Indikationsbereiche Onkologie und Kardiologie und dort - vor allem in der Onkologie - wiederum in mehrere Unterindikationen aufteilt, zur Zeit die „literature based evidence“ nicht ausreichend ist, um gesicherte Aussagen zur Kosteneffektivität der PET - differenziert nach Indikationen - machen zu können. Für die 19 von der Konsensuskonferenz im Bereich Onkologie als Ia/Ib-klassifizierten Indikationen für den Einsatz der PET in der klinischen Routine (Klutmann et al. 1999) konnten nur für die Indikationen Staging von nicht-kleinzelligem Bronchialkarzinom, Primärstaging von malignen Lymphomen bzw. Melanomen und Restaging von kolorektalen Karzinomen ökonomische Evaluationen durch den Literaturreview identifiziert werden. Erste, durch weitere Studien zu verifizierende positive Aussagen hinsichtlich einer Kosteneffektivität des klinischen PET-Einsatzes können aufgrund der methodischen Qualität der evaluierten Studien nur für die Indikationsstellung „Staging von nicht-kleinzelligem Bronchialkarzinom“ gemacht werden. Eine Kosteneffektivität für die Indikation „Diagnose von solitären Lungenknoten“ deutet sich in der methodisch hochwertigen Studie von Gambhir et al. aus dem Jahr 1998 an. Dieser erste Hinweis ist jedoch durch weitere gute ökonomische Untersuchungen zu überprüfen und gegebenenfalls zu untermauern. Für die übrigen als medizinisch angemessen (Ia) und akzeptabel (Ib) eingestuftem onkologischen Indikationen, z.B. Differentialdiagnostik beim Pankreaskarzinom, liegt derzeit keine ökonomische Evidenz vor.

Im Bereich Kardiologie wurde von der Konsensuskonferenz lediglich die Messung der Myokardlebensfähigkeit nach einem Herzinfarkt als Ia-Indikation eingestuft (Klutmann et al. 1999). Die vorliegenden ökonomischen Evaluationen befassen sich hingegen fast ausschließlich mit der Applikation der PET zur Diagnose von koronarer Herzkrankheit mittels Perfusionsmessung. Die Resultate der derzeit verfügbaren, methodisch hochwertigen Modellstudien unterstützen nicht die uneingeschränkte Nutzung der PET als Instrument zur kardialen Perfusionsmessung.

Trotz des relativ häufigen Einsatzes der PET in der Neuromedizin konnten keinerlei ökonomische Evaluationen hierzu in der umfangreichen Literaturrecherche gefunden

werden. In Anbetracht der relevanten Nutzung der PET in diesem Indikationsgebiet zeigt sich hier eine erhebliche Forschungslücke.

Um die Lücken der gesundheitsökonomischen Evidenz zu schließen, sind weitere, ökonomische Evaluationen von hoher methodischer Qualität erforderlich. Dabei ist wichtig, daß sich diese auf methodisch gute prospektive, kontrollierte klinische (Multi-center-)Studien mit hoher Fallzahl und, bei entscheidungstheoretischen Modellen, auf eine breite Datenbasis stützen. Künftige ökonomische Evaluationen sollten eine möglichst lange Zeitspanne abdecken, sowohl langfristige, gesundheitsbezogene Effektgrößen mit hoher Aussagekraft als auch langfristige Kosten einbeziehen und die Effekt- und Kostenseite anhand einer inkrementellen Analyse (siehe dazu C.4.2.3) direkt zueinander in Beziehung setzen. Auf der Kostenseite sollten möglichst viele Kostenkomponenten, wie zusätzliche Kosten für diagnose- bzw. therapiebedingte Komplikationen und „downstream-events“, durch Fehlklassifikation, also durch falsch positive und falsch negative Befundung, induzierte Kosten (und die zugehörigen Effekte auf der Effektseite) und - soweit möglich - auch indirekte Kosten durch Produktivitätsausfall berücksichtigt werden. Die Stabilität der Studienergebnisse sollte anhand von Sensitivitätsanalysen überprüft werden.

Derzeit liegt lediglich eine ökonomische Evaluation zum klinischen Einsatz der PET aus Deutschland vor (Klose et al. 1999a). Diese Studie zeichnet sich durch eine detaillierte und transparente Darstellung der Kostenseite auf der Basis einer verursachungsgemäßen Kostenzurechnung aus, die sowohl die Produktion des Radiotracers als auch Anschaffung und Betrieb einer PET-Anlage bei verschiedenen Auslastungsgraden berücksichtigt (siehe hierzu Kapitel C.5.2). Diese Daten können als Ausgangsbasis für die Kostenberechnung zukünftiger ökonomischer Analysen zum Einsatz der PET in Deutschland dienen. Alle anderen ökonomischen Evaluationen wurden in der Schweiz, in den Niederlanden, in Japan und primär in den USA durchgeführt. Eine unmittelbare Übertragung der Ergebnisse dieser Studien auf das deutsche Gesundheitssystem ist, wie bereits dargestellt, ohne Anpassung der Resultate des Studienlandes an deutsche Gegebenheiten, speziell an das deutsche Preisgerüst, nur sehr bedingt möglich (siehe dazu die Anmerkungen zur ökonomischen Evidenz der Studien im Indikationsgebiet Onkologie in diesem Kapitel). Aus der Sicht deutscher Entscheidungsträger wäre es sicherlich wünschenswert, auf mehr Studien zugreifen zu können, die unmittelbar auf das deutsche Gesundheitssystem ausgerichtet sind. Dabei könnten begleitende ökonomische Analysen zu geplanten klinischen Studien ebenso wie spezifisch auf Deutschland zugeschnittene Modellanalysen, wie entscheidungstheoretische Modelle, die verfügbare Evidenz deutlich verbessern. Nicht zuletzt ist die Übertragbarkeit der Ergebnisse, die an spezialisierten wissenschaftlichen Einrichtungen gewonnenen wurden, auf die Anwendung der PET-Technologie im klinischen Alltag einzuschätzen. Hinweise auf Differenzen hinsichtlich der Kosteneffektivität der PET-Diagnostik in der klinischen Forschung einerseits und als Routinemethode in der täglichen Praxis andererseits ergab die Untersuchung der American Heart Association zum Stellenwert der PET in der kardiologischen Routinediagnostik (Robert & Milne 1999).

Abschließend ist festzuhalten:

Eine Reihe von Studien, speziell die Analysen zur Diagnostik von Tumorerkrankungen der Lunge bzw. von kardialen Perfusionsstörungen mittels PET, zeigen auf, daß ergänzend zur klinischen Wirksamkeitsprüfung auch eine Beurteilung der Wirtschaftlichkeit des klinischen Einsatzes der PET-Technologie möglich ist.

Für die spezielle Verwendung der PET als Instrument zur diagnostischen Abklärung von solitären Lungenknoten bzw. als Stagingverfahren bei gesichertem NSCLC liegt erste ökonomische Evidenz vor, die erste Hinweise zur wirtschaftlichen Anwendung der PET-Technologie in der klinischen Diagnostik geben können. Diese sollten jedoch durch weitere, speziell auch auf den deutschen Entscheidungskontext bezogene ökonomische Untersuchungen erhärtet werden.

Darüber hinaus besteht, um die Informationsgrundlage für künftige Entscheidungen zu verbessern, Forschungsbedarf in den sonstigen onkologischen PET-Anwendungsfeldern, in der Kardiologie, sowohl in den Einsatzfeldern „Perfusionsmessung“ und insbesondere „Nachweis der Myokardviabilität“, und besonders dringend im bisher nicht evaluierten Bereich Neuromedizin.

C.7 Literaturverzeichnis

C.7.1 Literaturverzeichnis Teil I (berücksichtigte Studien)

Studien, die im Literaturreview als für die HTA-Thematik relevant eingestuft wurden und in die systematische Beurteilung einbezogen wurden.

Onkologie

- 1 Adler LP, Cascade E, Crowe J et al. Clinical application and economic implications of PET in the assessment of axillary lymph node involvement in breast cancer: A retrospective study. Institute for Clinical PET; 1994. Publication No. CE-BREAST-94.001
- 2 Adler LP, Faulhaber PF, Schnur KC et al. Axillary lymph node metastases: screening with [F-18]2-deoxy-2-fluoro-D-glucose (FDG) PET. *Radiology* 1997;203(2):323-7
- 3 Coleman RE, Cascade E, Gupta NC et al. Clinical application and economic implications of PET in the assessment of solitary pulmonary nodules: A retrospective study. Institute for Clinical PET; 1994. Publication No. CE-SPN-94.001
- 4 Coleman RE, Cascade E, Gupta NC et al. Untersuchungen von solitären Lungenknoten mit PET. *Electromedica* 1995; 63(2):48-52
- 5 ECRI. Positron emission tomography (PET) for the diagnosis and staging of non-small-cell lung cancer, part 2: cost-effectiveness analysis. Health technology assessment information service; 1998. No.73
- 6 Gambhir SS, Hoh CK, Phelps ME et al. Decision tree sensitivity analysis for cost-effectiveness of FDG-PET in the staging and management of non-small-cell lung carcinoma. *J Nucl Med* 1996;37(9):1428-36
- 7 Gambhir SS, Shepherd JE, Shah BD et al. Analytical decision model for the cost-effective management of solitary pulmonary nodules. *J Clin Oncol* 1998;16(6):2113-25
- 8 Hoh CK, Glaspy J, Rosen P et al. Whole-body FDG-PET imaging for staging of Hodgkin's disease and lymphoma. *J Nucl Med* 1997;38(3):343-8

- 9 Klose T, Leidl R, Buchmann I et al. Primary staging of lymphomas: cost-effectiveness of FDG-PET versus computed tomography. University of Ulm. Department of Health Economics 1999. Working paper series. Working paper No.: 2/1999. Teilergebnisse der Studie wurden veröffentlicht unter: Leidl R, Klose T, Buchmann I, Brambs HJ, Reske S (1999). Grundlagen und Beispiele der ökonomischen Evaluation. Vortrag auf der 37. Internationalen Jahrestagung der Deutschen Gesellschaft für Nuklearmedizin, Ulm, 14.-17. April 1999a
- 10 Kosuda S, Ichihara K, Watanabe M et al. Decision-tree sensitivity analysis for cost-effectiveness of chest 2-Fluoro-2-D-[¹⁸F]Fluorodeoxyglucose positron emission tomography in patients with pulmonary nodules (non-small cell lung carcinoma) in Japan. *Chest* 2000;117(2):346-53
- 11 Larson SM, Cohen AM, Cascade E et al. Clinical application and economic implications of PET in the assessment of colorectal cancer recurrence: A retrospective study. Institute for Clinical PET; 1994. Publication No. CE-COLOR-94-0001
- 12 Scott WJ, Shepherd J, Gambhir SS. Cost-effectiveness of FDG-PET for staging non-small cell lung cancer: a decision analysis. *Ann Thorac Surg* 1998;66(6):1876-83
- 13 Valk PE, Pounds TR, Tesar RD et al. Cost-effectiveness of PET imaging in clinical oncology. *Nucl Med Biol* 1996;23(6):737-43
- 14 Valk PE, Abella-Columna E, Haseman MK et al. Whole-body PET imaging with [¹⁸F]fluorodeoxyglucose in management of recurrent colorectal cancer. *Archives of Surgery* 1999;134(5):503-13
- 15 von Schulthess GK, Steinert HC, Dummer R et al. Cost-effectiveness of whole-body PET imaging in non-small cell lung cancer and malignant melanoma. *Acad Radiol* 1998;5Suppl.2:S300-2

Kardiologie

- 1 Besozzi MC, Alagona P, Hunter K et al. Clinical application and economic implications of PET in the assessment of myocardial viability: A retrospective study. Institute for Clinical PET; 1994. Publication No.: CE-CARDIOLOGY-94.001
- 2 Garber AM & Solomon NA. Cost-effectiveness of alternative test strategies for the diagnosis of coronary artery disease. *Ann intern med* 1999;130(9):719-28
- 3 Gleason JM & Frick MP (kritische Stellungnahme zur Publikation Gould KL, Goldstein RA, Mullani NA. 1998). The expected value of cash flows in the eco-

- conomic analysis of clinical positron emission tomography of the heart. *J Nucl Med* 1990;31(11):1893-6
- 4 Gould KL, Goldstein RA, Mullani NA. Economic analysis of clinical positron emission tomography of the heart with rubidium-82. *J Nucl Med* 1989;30(5):707-17
 - 5 Hunink MGM. Positron emission tomographie compared to single photon emission tomography in the evaluation of myocardial viability: a preliminary cost-effectiveness analysis. In: Van der Wall et al. (eds.) *Cardiac positron emission tomography*. Amsterdam:Kluwer; 1995. p. 97-102
 - 6 Maddahi J & Gambhir SS. Cost-effective selection of patients for coronary angiography. *J Nucl Cardiol* 1997;4(2Pt2):S141-51
 - 7 Patterson RE, Horowitz SF, Eisner RL. Comparison of modalities to diagnose coronary artery disease. *Sem Nucl Med* 1994;24(4):286-310
 - 8 Patterson RE, Eisner RL, Horowitz SF. Comparison of cost-effectiveness and utility of exercise ECG, single photon emission computed tomography, positron emission tomography, and coronary angiography for diagnosis of coronary artery disease. *Circulation* 1995;91(1):54-65

Kostenanalysen der Positronen-Emissions-Tomographie (allgemeine, indikationsunabhängige Betrachtungen)

- 1 Conti PS, Keppler JS, Halls JM. Positron emission tomography: a financial and operational analysis. *Am J Roentgenol* 1994;162(6):1279-86
- 2 Evens RG, Siegel BA, Welch MJ et al. Cost analyses of positron emission tomography for clinical use. *Am J Roentgenol* 1983;141(5):1073-6
- 3 Lottes G, Gorschluter P, Kuwert T et al. Kosten für F-18-FDG PET: Erfahrungen mit einem Satellitenkonzept. *Nuklearmedizin* 1998;37(5):159-64
- 4 Lottes G & Schober O. Kosten für die FDG-PET bei einem Satellitenkonzept – eine Analyse aktueller Daten. *Nuklearmedizin* 1999;38:A57 und Folien eines Vortrags auf der 37. Internationalen Jahrestagung der Deutschen Gesellschaft für Nuklearmedizin, Ulm

Übersichtsarbeiten

- 1 Dietlein M, Knapp WH, Lauterbach KW et al. Economic evaluation studies in nuclear medicine: the need for standardization. *Eur J Nucl Med* 1999;26(6):663-80
- 2 Robert G, Milne R. Positron emission tomography: establishing priorities for health technology assessment. *Health Technology Assessment* 1999;3(16)

C.7.2 Literaturverzeichnis Teil II (ausgeschlossene Studien)

Literatur, die in der zweiten Stufe des Auswahlprozesses aussortiert wurde.

Onkologie

- 1 Oncology. Pet scanning said to save money and improve staging of lung cancer. *Oncology*, 1996;9(11):1200
- 2 Chiti A, McIntyre WJ, Fouad FM et al. Nuclear medicine procedures in lung cancer. *Eur j nucl med* 1999;25(6):533-55
- 3 DiBisceglie AM, Carithers RL, Gores GJ. Hepatocellular carcinoma. *Hepatology* 1998;28(4 I):1161-5
- 4 Erasmus JJ & Patz EF. Positron emission tomography imaging in the thorax. *Clin Chest Med* 1999;20(4):715-24
- 5 Gupta NC, Maloof J, Gunel E. Probability of malignancy in solitary pulmonary nodules using fluorine-18-FDG and PET. *J Nucl Med* 1995;37:943-48
- 6 Kosuda S, Ichihara K, Watanabe M et al. Decision-tree sensitivity analysis for cost-effectiveness of chest FDG-PET in patients with a pulmonary tumor (non-small cell carcinoma) in Japan. *Kaku Igaku (Jpn J Nucl Med)* 1998;35:395-404
- 7 Kubota, K, Yamada S, Fukuda H et al. Cost effectiveness analysis of FDG-PET in the differential diagnosis and staging of lung cancer *Kaku Igaku (Jpn J Nucl Med)* 1997;35(5):329-35
- 8 Larson, SM. EL4. Cost effectiveness of PET in Oncology. *Kaku Igaku (Jpn J Nucl Med)* 1995;21(5):865

- 9 Lowe VJ & Naunheim KS. Current role of positron emission tomography in thoracic oncology. *Thorax* 1998;53(8):703-12
- 10 Madar I, Hoh C, Figlin RA et al. Cost effective staging of non-small cell lung carcinoma by whole body PET FDG imaging. *J Nucl Med* 1995;36(5):57P
- 11 Mitchell MD, Turkelson C, Doggett D. Cost-effectiveness analysis of PET in lung cancer diagnosis and staging. Annual meeting of international society of technology assessment in health care 1998;14:33
- 12 Podoloff DA. The role of radionuclide scans in sarcoma. *Hematol Oncol Clin North Am* 1995;9(3):605-26
- 13 Steinert, HC, Ullrich SP, Böni R et al. Kosteneffektivität beim Staging des malignen Melanoms: Vergleich der Ganzkörper PET versus konventionelles Staging. *Nuklearmedizin* 1998;37:A37
- 14 Steinert, HC von Schulthess GK, Weder W. Effektivität der Ganzkörper PET mit FDG beim Staging des nicht-kleinzelligen Bronchuskarzinom bei 100 Patienten. *Nuklearmedizin* 1998;37:A37
- 15 Vock P. Bronchuskarzinom: neue radiologische Methoden. *Schweiz Med Wochenschr* 1997;127(47):1945-9
- 16 Yao WJ, Hoh CK, Glaspy JA et al. Whole body FDG-PET imaging for staging of malignant melanoma: is it cost effective? *J Nucl Med* 1994;35:8P

Kardiologie

- 1 Duong T.H., Fonarow G., Laks H. et al. Cost-effectiveness of PET in the management of ischemic cardiomyopathy patients who are referred for cardiac transplantation. *J Am col cardiol* 1996;27:144
- 2 Giordano A, et al. Myocardial SPECT in the study of ischemic heart disease detection of hibernating myocardium and evaluation of cost/benefit ratio. *Rays* 1999;24(1):73-80
- 3 Patterson RE. Cost-effectiveness analysis in diagnosis of cardiac disease: overview of its rationale and method. *J Nucl Cardiol* 1996;3(4):334-41
- 4 Schelbert HR. Cost-effectiveness of cardiac positron emission tomography in the treatment of ischemic cardiomyopathy. *Am J Card Imaging* 1996;10(3):191-7

- 5 Stanford W. Screening of coronary artery disease: is there a cost-effective way to do it? *Am J Card Imaging* 1996;10(3):180-6

Neurologie

- 1 Kuwert T, Bartenstein R, Grunwald F et al. Klinische Wertigkeit der Positronen-Emissions-Tomographie in der Neuromedizin, *Der Nervenarzt* 1998;69(12):1045-60
- 2 Pai M, Park CH, Suh JH et al. Fluorine-18 fluorodeoxyglucose imaging using dual-head coincidence positron emission tomography without attenuation correction in patients with head and neck cancer, *Clin Nucl Med* 1999;24(7):495-500
- 3 Small GW & Leiter F Neuroimaging for diagnosis of dementia. *J clin psychiatry* 1998;58(Suppl.11):4-7

Sonstige Artikel

- 1 Homberg JM, Mohiuddin SM, Hilleman DE et al. Outcomes and costs of positron emission tomography: comparison of intravenous adenosine and intravenous dipyridamole. *Clin Ther* 1997;19(3):570-81

Übersichtsarbeiten

- 1 Gardner SF, Green JA, Bednarczyk EM et al. Principles and clinical applications of positron emission tomography. *Am J Hosp Pharm* 1992;49(6):1499-506
- 2 Morrison A & Coates V Cost-effectiveness of a new diagnostic imaging technology. Annual meeting of international society of technology assessment in health care 1995;11:Abstract No. 198

Allgemeine Betrachtungen zur ökonomischen Evaluation der PET

- 1 Dietlein M, Scheidhauser K, Lauterbach KW et al. Qualitätskriterien für Kosten-Nutzen-Analysen in der onkologischen Nuklearmedizin und Stand ihrer Realisierung. *Z Ärztl Fortbild Qualitätssich* 1999;93(1):49-55

- 2 Kuwert T, Matheja P, Vollet B, et al. Positronen-Emissions-Tomographie. Kosten-Nutzen-Gesichtspunkte. Radiologe 1996;36(4):337-44

C.7.3 Literaturverzeichnis Teil III

Allgemeine Literatur, auf die im Text verwiesen wird.

- 1 Antes G. EBM praktizieren. In: Perleth M, Antes G (Hrsg.) Evidenz-basierte Medizin. München: MMW Medien & Medizin Verlag; 1999. p. 19-27
- 2 Bairy Merez CN, Berman DS. Imaging techniques for coronary artery disease: current status and future directions. Clin Cardiol 1997;20:526-32
- 3 Baltussen R, Leidl R, Ament A. Real world designs in economic evaluation. Pharmacoeconomics 1999;16(5Pt1):449-58
- 4 Bundesministerium für Gesundheit. Daten des Gesundheitswesens. Ausgabe 99. Band 122 Schriftenreihe des Bundesministeriums für Gesundheit. Baden-Baden: Nomos Verlagsgesellschaft; 1999
- 5 Deutsche Gesellschaft für Nuklearmedizin. Konsensus Onko-PET. Stand Januar 1998. <http://www.uni-duesseldorf.de/WWW/AWMF/II/nukl-007.htm>
- 6 Deutsche Gesellschaft für Nuklearmedizin. Konsensus Neuro-PET. Stand Januar 1998. <http://www.uni-duesseldorf.de/WWW/AWMF/II/nukl-008.htm>
- 7 Drummond MF et al. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 1997
- 8 ECRI. Positron emission tomography (PET) for the Diagnosis and staging of non-small-cell lung cancer, part 1: results of ECRI's meta-analyses. Health technology assessment information service; 1998. No.72
- 9 Fletscher JW, Adams E, Flynn K et al. Design of prospective clinical trials for assessment of PET. Clinical Positron Imaging 1998;1(3):201-11
- 10 Johannesson M. On the decision rules of cost-effectiveness analysis. J Health Economics 1993;12:459-467
- 11 Klose T, Herlemann B, Leidl R. Kostenmessung in der ökonomischen Evaluation von Gesundheitsleistungen: die Kosten stationärer Akutversorgung am Universitätskrankenhaus Ulm. Z. f. Gesundheitswiss. 1999b;7(4):351-72
- 12 Klutmann S, Bohuslavizki KH et al. Positronen emission tomography (PET) in clinical routine. Radiol Oncol 1999;33(2):111-7

- 13 König H-H. Steuerung der Anschaffung und Nutzung medizinisch-technischer Großgeräte. [Dissertation]. Tübingen; 1993
- 14 Kuwert T, Matheja P. & Schober O. Effektivität der Diagnostik mit PET und SPECT bei degenerativen Hirnerkrankungen und Hirntumoren, *Der Nuklearmedizin* 1996;1(19):51-9
- 15 Leidl R, Graf von der Schulenburg J-M & Wasem J. Ansätze und Methoden der ökonomischen Evaluation - eine internationale Perspektive. *Health Technology Assessment*. Band 9. Baden-Baden: Nomos Verlagsgesellschaft; 1999
- 16 Leidl R. Der Effizienz auf der Spur: Eine Einführung in die ökonomische Evaluation. In: Schwartz FW (Hrsg.) et al. *Das Public Health Buch*. München: Urban & Schwarzenberg; 1998. p. 346-69
- 17 Moser E. Das PET-Zentrum in Freiburg. *Das Krankenhaus* 1997;4:190-2
- 18 Reske SN. Positronen-Emissions-Tomographie in der Onkologie. *Dt Ärztebl* 1998;30(39):A1847-49
- 19 Steinert, HC et al. Non-Small Cell lung cancer: Nodal Staging with FDG PET versus CT with correlative lymph node mapping and sampling. *Radiology* 1997;202:441-6
- 20 Steinert, HC et al. Malignant Melanoma: Staging with Whole-body positron emission tomography and 2-[F-18]-Fluoro-2-deoxy-D-glucose. *Radiology* 1995;195:705-9
- 21 vom Dahl J. Untersuchungen der Myokardperfusion mit der Positronen-Emissions-Tomographie: eine klinisch nützliche und valide Methode? *Herz* 1997;22(1):1-15
- 22 Welte R, Leidl R. Übertragung der Ergebnisse ökonomischer Evaluationsstudien aus dem Ausland auf Deutschland: Probleme und Lösungsansätze. In: Leidl R (Hrsg.), Schulenburg JM et al. *Ansätze und Methoden der ökonomischen Evaluation – eine internationale Perspektive*. Baden-Baden: Nomos Verlagsgesellschaft; 1999. p. 171-202
- 23 Zarnke KB, Levine MAH & O'Brien BJ. Cost-benefit analyses in the health care literature: Don't judge a study by its label. *J Clin Epidemiol* 1997;50(7):813-22

Anhang

Anhang A: Erläuterungen zu Abbildung 1

Anhang B: Erläuterungen zu Abbildung 2

Anhang C: Transparenz- und Qualitätskatalog

Anhang A: Erläuterungen zu Abbildung 1

Studie, Erstautor	Indikationsstellung	Vergleichsalternativen ^a	Ergebnis der Studie ^a	Anmerkungen
Gambhir 1998	Diagnose von SPN	Kombination CT/ ¹⁸ F-DG-PET (+ Biopsie) nur CT (+ Biopsie), Thorakotomie versus „wait and watch“	pSPN < 0,12: „wait & watch“ ist die bevorzugte Strategie, kleinste iKER bei pSPN [0,12 - 0,69]: CT/PET-Kombination pSPN]0,69 - 0,9]: nur CT pSPN > 0,9: Thorakotomie (Referenzstrategie: „wait and watch“)	Alle iKER liegen unter der gesellschaftlich akzeptierten Obergrenze für die Kosteneffektivität von US\$ 50 000 pro zusätzlichem Lebensjahr (Einschätzung nach Gambhir et al.1998)
Gambhir 1996	Staging von NSCLC	Kombination ¹⁸ F-DG-PET/CT (+ Biopsie) versus nur CT (+ Biopsie)	Kombination PET/CT ist dominant über Strategie „nur CT“ für Prävalenzwerte ≥ 5,6%	
Scott 1998	Staging von NSCLC	Vier unterschiedliche Kombinationen aus ¹⁸ F-DG-PET, CT, Biopsie versus nur CT (+ Biopsie) (Referenzstrategie)	Strategie „PET nach CT-negativem Befund“ ist dominant über Strategie „nur CT“ für pNSCLC-Werte ≥ 0,35	Alle iKER liegen unter der gesellschaftlich akzeptierten Obergrenze für die Kosteneffektivität von US\$ 50 000 pro zusätzlichem Lebensjahr (Einschätzung nach Gambhir et al.1998)

^a Abkürzungen siehe Abkürzungsverzeichnis

Anhang B: Erläuterungen zu Abbildung 2

Studie, Erstautor	Indikationsstellung	Vergleichsalternativen ^a	Ergebnis der Studie ^a	Anmerkungen
Patterson 1995	Perfusionsmessung bei Herzinfarktpatienten	nur Angiographie (AG) Belastungs-EKG/AG SPECT/AG PET/AG	Kosten pro QALY sind für pCAD-Werte $\leq 0,6^*$ bei PET/AG < als bei Angiographie $> 0,6^*$ bei PET/AG > als bei Angiographie über die gesamte pCAD-Spanne (0,1 - 1) gilt: Kosten pro Effekt für PET/AG < SPECT/AG PET/AG < Belastungs-EKG/AG	Im pCAD-Intervall 0,1 - 1 liegen für die Kombination PET/AG die <ul style="list-style-type: none"> • Kosten pro korrekt diagnostiziertem Patienten unter US\$ 30 000 • Kosten pro QALY unter US\$ 10 000
Maddahi 1997	Perfusionsmessung bei Herzinfarktpatienten	Fünf mehrstufige Strategien mit PET, SPECT, ExEKG alternativ als Eingangstest (plus Angiographie) versus nur Angiographie	Vergleich PET/AG gegenüber SPECT/AG: pCAD < 0,42*: PET/AG dominiert SPECT/AG pCAD 0,42 - 0,68*: PET/AG ist besser, aber teurer als SPECT/AG pCAD > 0,68*: AG dominiert SPECT/AG Vergleich PET/AG gegenüber AG: pCAD < 0,6*: AG ist besser, aber teurer als PET pCAD $\geq 0,6^*$: AG dominiert PET	
Garber 1999	Perfusionsmessung bei Herzinfarktpatienten	nur Angiographie (AG) PET/AG SPECT/AG UKG/AG ²⁰¹ Tl-Szintigraphie/AG Belastungs-EKG/AG	Im pCAD-Intervall 0,25 - 0,75 dominiert AG die Kombination PET/AG Bei pCAD 0,5 ist PET/AG versus SPECT/AG nicht kosteneffektiv bei einer inkrementellen Kosten-Nutzwert-Relation von US\$ 640 000/QALY	Für pCAD-Werte < 0,25 bzw. > 0,75 liegen keine Angaben vor

a: Abkürzungen siehe Abkürzungsverzeichnis, * Angaben sind aus Abbildungen der Originalartikel Patterson et al. 1995 bzw. Maddahi & Garber 1997 entnommen

Anhang C: Transparenz- und Qualitätskatalog

Transparenzkatalog		
Autoren, Titel und Publikationsorgan:	1 = Kriterium erfüllt 1/2 = Kriterium teilweise erfüllt 0 = Kriterium nicht erfüllt nr = nicht relevant	1, ½ 0, nr
Fragestellung		
1. Wurde die Fragestellung formuliert? 2. Wurde der medizinische und ökonomische Problemkontext dargestellt?		
Evaluationsrahmen		
3. Wurden alle in die Studie einbezogenen Technologien beschrieben? 4. Werden mindestens zwei alternative Technologien explizit verglichen? 5. Wurde die Auswahl der Vergleichstechnologien begründet? 6. Wurde die Zielpopulation beschrieben? 7. Wurde der Zeithorizont für Kosten und Gesundheitseffekte angegeben? 8. Wurde der Typ der gesundheitsökonomischen Evaluation explizit genannt? 9. Wurden sowohl Kosten als auch Gesundheitseffekte untersucht? 10. Wurde die Perspektive der Untersuchung explizit genannt?		
Analysemethoden und Modellierung		
11. Wurden die statistischen Tests/Modelle zur Analyse der Daten beschrieben? 12. Wurden in entscheidungsanalytischen Modellen die Modellstruktur und alle Parameter dokumentiert (in der Publikation bzw. einem technischen Report)? 13. Wurden Annahmen explizit formuliert? 14. Wurden in entscheidungsanalytischen Modellen die Datenquellen für die Pfadwahrscheinlichkeiten angegeben?		
Gesundheitseffekte		
15. Wurden die in der Studie untersuchten Gesundheitszustände aufgeführt? 16. Wurden die Quellen für die Gesundheitseffekt-daten angegeben? 17. Wurden das epidemiologische Studiendesign und die Auswertungsmethoden beschrieben und wurden die Ergebnisse dargestellt? (falls auf einer einzelnen Studie basierend) 18. Wurden die Methoden zur Identifikation, Extraktion und Synthese der Effektparameter beschrieben? (falls auf einer Informationssynthese basierend) 19. Wurden die verschiedenen Gesundheitszustände mit Präferenzen bewertet und die Methoden und Meßinstrumente dieser Bewertung angegeben? 20. Wurden die Quellen der Bewertungsdaten für die Gesundheitszustände angegeben? 21. Wurden Aussagen zur Evidenz der Gesundheitseffekte gemacht?		

Quelle: Leidl, von der Schulenburg & Wasem 1999

<p>Kosten</p> <p>22. Wurden die den Kosten zugrunde liegenden Mengengerüste angegeben? 23. Wurden die Quellen und Methoden zur Ermittlung der Mengengerüste angegeben? 24. Wurde die den Kosten zugrunde liegenden Preisgerüste angegeben? 25. Wurden die Quellen und Methoden zur Ermittlung der Preise angegeben? 26. Wurden die einbezogenen Kosten anhand der gewählten Perspektive und des gewählten Zeithorizontes begründet? 27. Wurden Daten zu Produktivitätsausfallkosten (falls berücksichtigt) getrennt aufgeführt? 28. Wurde die Währung genannt? 29. Wurden Angaben zu Währungskonversionen gemacht? 30. Wurden Angaben zu Preisanpassungen bei Inflation oder Deflation gemacht?</p>	
<p>Diskontierung</p> <p>31. Wurden zukünftige Gesundheitseffekte <u>und</u> Kosten diskontiert? 32. Wurde das Referenzjahr für die Diskontierung angegeben bzw. bei fehlender Diskontierung das Referenzjahr für die Kosten? 33. Wurden die Diskontraten angegeben? 34. Wurde die Wahl der Diskontraten bzw. der Verzicht auf eine Diskontierung begründet?</p>	
<p>Ergebnispräsentation</p> <p>35. Wurden Maßnahmen zur Modellvalidierung ergriffen und beschrieben? 36. Wurden absolute Gesundheitseffekte und absolute Kosten jeweils pro Kopf bestimmt und dargestellt? 37. Wurden inkrementelle Gesundheitseffekte und inkrementelle Kosten jeweils pro Kopf bestimmt und dargestellt? 38. Wurde eine Maßzahl für die Relation zwischen Kosten und Gesundheitseffekt angegeben? 39. Wurden reine (nicht lebensqualitätsadjustierte) klinische Effekte berichtet? 40. Wurden Ergebnisse in disaggregierter Form dargestellt? 41. Wurden populationsaggregierte Kosten und Gesundheitseffekte dargestellt?</p>	
<p>Behandlung von Unsicherheiten</p> <p>42. Wurden univariate Sensitivitätsanalysen durchgeführt? 43. Wurden multivariate Sensitivitätsanalysen durchgeführt? 44. Wurde strukturelle Sensitivitätsanalysen durchgeführt? 45. Wurden die in den Sensitivitätsanalysen berücksichtigten Werte oder Wertebereiche bzw. Strukturvarianten angegeben? 46. Wurden die Ergebnisse der Sensitivitätsanalysen dargestellt? 47. Wurden statistische Inferenzmethoden (statistische Tests, Konfidenzintervalle) für stochastische Daten eingesetzt und die Ergebnisse berichtet?</p>	

Diskussion 48. Wurde die Datenqualität diskutiert? 49. Wurden Richtung und Größe des Einflusses unsicherer oder verzerrter Parameterschätzung auf das Ergebnis diskutiert? 50. Wurde Richtung und Größe des Einflusses struktureller Modellannahmen auf das Ergebnis diskutiert? 51. Wurden die Einschränkungen und Schwächen der Studie diskutiert? 52. Wurden Angaben zur Generalisierbarkeit der Ergebnisse gemacht? 53. Wurden ethische und Verteilungsfragen diskutiert? 54. Wurde das Ergebnis im Kontext mit unabhängigen Gesundheitsprogrammen diskutiert?	
Schlußfolgerungen 55. Wurden Schlußfolgerungen aus den berichteten Daten/Ergebnissen abgeleitet? 56. Wurde eine Antwort auf die Fragestellung gegeben?	

Qualitätskatalog		
Autoren, Titel und Publikationsorgan	1 = Kriterium erfüllt 1/2 = Kriterium teilweise erfüllt 0 = Kriterium nicht erfüllt nr = nicht relevant	1, ½ 0, nr
Fragestellung		
1. Wurde die Fragestellung präzise formuliert? 2. Wurde der medizinische und ökonomische Problemkontext ausreichend dargestellt?		
Evaluationsrahmen		
3. Wurden alle in die Studie einbezogenen Technologien hinreichend detailliert beschrieben? 4. Wurden alle im Rahmen der Fragestellung relevanten Technologien verglichen? 5. Wurde die Auswahl der Vergleichstechnologien schlüssig begründet? 6. Wurde die Zielpopulation klar beschrieben? 7. Wurde ein für die Fragestellung angemessener Zeithorizont für Kosten und Gesundheitseffekte gewählt und angegeben? 8. Wurde der Typ der gesundheitsökonomischen Evaluation explizit genannt? 9. Wurden sowohl Kosten als auch Gesundheitseffekte untersucht? 10. Wurde die Perspektive der Untersuchung eindeutig gewählt und explizit genannt?		
Analysemethoden und Modellierung		
11. Wurden adäquate statistische Tests/Modelle zur Analyse der Daten gewählt und hinreichend gründlich beschrieben? 12. Wurden in entscheidungsanalytischen Modellen die Modellstruktur und alle Parameter vollständig und nachvollziehbar dokumentiert (in der Publikation bzw. einem technischen Report)? 13. Wurden die relevanten Annahmen explizit formuliert? 14. Wurden in entscheidungsanalytischen Modellen adäquate Datenquellen für die Pfadwahrscheinlichkeiten gewählt und eindeutig genannt?		
Gesundheitseffekte		
15. Wurden alle für die gewählte Perspektive und den gewählten Zeithorizont relevanten Gesundheitszustände berücksichtigt und explizit aufgeführt? 16. Wurden adäquate Quellen für die Gesundheitseffektdaten gewählt und eindeutig genannt? 17. Wurden das epidemiologische Studiendesign und die Auswertungsmethoden adäquat gewählt und beschrieben und wurden die Ergebnisse detailliert dargestellt? (falls auf einer einzelnen Studie basierend) 18. Wurden angemessene Methoden zur Identifikation, Extraktion und Synthese der Effektparameter verwendet und wurden sie detailliert beschrieben? (falls auf einer Informationssynthese basierend) 19. Wurden die verschiedenen Gesundheitszustände mit Präferenzen bewertet und dafür geeignete Methoden und Meßinstrumente gewählt und angegeben? 20. Wurden adäquate Quellen der Bewertungsdaten für die Gesundheitszustände gewählt und eindeutig genannt? 21. Wurde die Evidenz der Gesundheitseffekte ausreichend belegt? (s. ggf. entsprechende Kontextdokumente)		

<p>Kosten</p> <p>22. Wurden die den Kosten zugrunde liegenden Mengengerüste hinreichend gründlich dargestellt?</p> <p>23. Wurden adäquate Quellen und Methoden zur Ermittlung der Mengengerüste gewählt und eindeutig genannt?</p> <p>24. Wurden die den Kosten zugrunde liegenden Preisgerüste hinreichend gründlich beschrieben?</p> <p>25. Wurden adäquate Quellen und Methoden zur Ermittlung der Preise gewählt und eindeutig genannt?</p> <p>26. Wurden die einbezogenen Kosten anhand der gewählten Perspektive und des gewählten Zeithorizontes schlüssig begründet und wurden alle relevanten Kosten berücksichtigt?</p> <p>27. Wurden Daten zu Produktivitätsausfallkosten (falls berücksichtigt) getrennt aufgeführt und methodisch korrekt in die Analyse einbezogen?</p> <p>28. Wurde die Währung genannt?</p> <p>29. Wurden Währungskonversionen adäquat durchgeführt?</p> <p>30. Wurden Preisanpassungen bei Inflation oder Deflation adäquat durchgeführt?</p>	
<p>Diskontierung</p> <p>31. Wurden zukünftige Gesundheitseffekte <u>und</u> Kosten adäquat diskontiert?</p> <p>32. Wurde das Referenzjahr für die Diskontierung angegeben bzw. bei fehlender Diskontierung das Referenzjahr für die Kosten?</p> <p>33. Wurden die Diskontraten angegeben?</p> <p>34. Wurde die Wahl der Diskontraten bzw. der Verzicht auf eine Diskontierung plausibel begründet?</p>	
<p>Ergebnispräsentation</p> <p>35. Wurden Maßnahmen zur Modellvalidierung ergriffen und beschrieben?</p> <p>36. Wurden absolute Gesundheitseffekte und absolute Kosten jeweils pro Kopf bestimmt und dargestellt?</p> <p>37. Wurden inkrementelle Gesundheitseffekte und inkrementelle Kosten jeweils pro Kopf bestimmt und dargestellt?</p> <p>38. Wurde eine für den Typ der gesundheitsökonomischen Evaluation sinnvolle Maßzahl für die Relation zwischen Kosten und Gesundheitseffekt angegeben?</p> <p>39. Wurden reine (nicht lebensqualitätsadjustierte) klinische Effekte berichtet?</p> <p>40. Wurden die relevanten Ergebnisse in disaggregierter Form dargestellt?</p> <p>41. Wurden populationsaggregierte Kosten und Gesundheitseffekte dargestellt?</p>	
<p>Behandlung von Unsicherheiten</p> <p>42. Wurden univariate Sensitivitätsanalysen für die relevanten Parameter durchgeführt?</p> <p>43. Wurden multivariate Sensitivitätsanalysen für die relevanten Parameter durchgeführt?</p> <p>44. Wurde Sensitivitätsanalysen für die relevanten strukturellen Elemente durchgeführt?</p> <p>45. Wurden in den Sensitivitätsanalysen realistische Werte oder Wertebereiche bzw. Strukturvarianten berücksichtigt und angegeben?</p> <p>46. Wurden die Ergebnisse der Sensitivitätsanalysen hinreichend dokumentiert?</p> <p>47. Wurden adäquate statistische Inferenzmethoden (statistische Tests, Konfidenzintervalle) für stochastische Daten eingesetzt und die Ergebnisse berichtet?</p>	

Diskussion 48. Wurde die Datenqualität kritisch beurteilt? 49. Wurden Richtung und Größe des Einflusses unsicherer oder verzerrter Parameterschätzung auf das Ergebnis konsistent diskutiert? 50. Wurde Richtung und Größe des Einflusses struktureller Modellannahmen auf das Ergebnis konsistent diskutiert? 51. Wurden die wesentlichen Einschränkungen und Schwächen der Studie diskutiert? 52. Wurden plausible Angaben zur Generalisierbarkeit der Ergebnisse gemacht? 53. Wurden wichtige ethische und Verteilungsfragen diskutiert? 54. Wurde das Ergebnis sinnvoll im Kontext mit unabhängigen Gesundheitsprogrammen diskutiert?	
Schlußfolgerungen 55. Wurden in konsistenter Weise Schlußfolgerungen aus den berichteten Daten/Ergebnissen abgeleitet? 56. Wurde eine auf Wissensstand und Studienergebnissen basierende Antwort auf die Fragestellung gegeben?	

*Deskriptionskatalog. Kriterienkatalog zur deskriptiven Erfassung und Dokumentation von Merkmalen und Inhalten gesundheitsökonomischer Studien
(entwickelt im Konsensusverfahren von den gesundheitsökonomischen Projektgruppen München, Hannover, Ulm)*

1) Publikation	
Autor(en)	
Titel / Kapitelüberschrift. In: Herausgeber: Buchtitel	
Zeitschrift / Verlag	
Jahrgang	
Band (Nummer): Seiten / Kapelseiten	
Durchführende Institution(en)	
Auftraggeber bzw. Finanzierung	<input type="checkbox"/> Staatliche Stelle <input type="checkbox"/> Versicherung <input type="checkbox"/> Verband, Organisation <input type="checkbox"/> Stiftung <input type="checkbox"/> Universität <input type="checkbox"/> Forschungseinrichtung <input type="checkbox"/> Industrie <input type="checkbox"/> Betrieb <input type="checkbox"/> Anderer: _____ <input type="checkbox"/> Unklar
Hinweis auf Vorhandensein eines technischen Berichts	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Notwendigkeit der Einbeziehung des technischen Berichts zur Studienbeurteilung	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> n/r (nicht relevant)
2) Fragestellung und Evaluationsrahmen	
Studientyp	<input type="checkbox"/> Primärstudie <input type="checkbox"/> Sekundärstudie
Formulierung der Zielsetzung: - Explizite Nennung der interessierenden Technik - Gegenüberstellung von Alternative(n) - Untersuchung von Effekten - Untersuchung von Kosten	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Nein
Explizite Nennung der Perspektive(n)	Ja: <input type="checkbox"/> Gesellschaft <input type="checkbox"/> Kostenträger <input type="checkbox"/> Aufsichtführendes Ministerium <input type="checkbox"/> Patienten <input type="checkbox"/> Angehörige <input type="checkbox"/> Leistungserbringer bzw. Produzenten <input type="checkbox"/> Arbeitgeber <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein

Spezifikation des Personalbedarfs (einschließlich Qualifizierungsgrad)	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
<i>Deskriptive Bemerkungen:</i>		<i>Bewertende Bemerkungen:</i>
3) Gesundheitseffekte (gesundheitsbezogenes Outcome)	<i>Anwendung der Technologie(n):</i>	<i>Anwendung der Vergleichstechnologie(n):</i> <input type="checkbox"/> n/r
Art der interessierenden gesundheits- bezogenen Outcome-Größe(n)	<input type="checkbox"/> Physiologischer oder biochemischer Parameter - Surrogat-Endpunkt <input type="checkbox"/> Morbiditätsbezogener Indikator - Fallrate, Erfolgsrate, Informationswert <input type="checkbox"/> Mortalitätsbezogener Indikator - Lebensdauer, Überlebenszeit <input type="checkbox"/> Ökonomisch orientierter Indikator - Arbeitsunfähigkeits-, KH-Pflegetage <input type="checkbox"/> Gesundheitsbezogene Lebensqualität <input type="checkbox"/> Lebensdauer und Lebensqualität integrierende Ergebnisgröße <input type="checkbox"/> Monetarisierter Ergebnisgröße	<input type="checkbox"/> Physiologischer oder biochemischer Parameter - Surrogat-Endpunkt <input type="checkbox"/> Morbiditätsbezogener Indikator - Fallrate, Erfolgsrate, Informationswert <input type="checkbox"/> Mortalitätsbezogener Indikator - Lebensdauer, Überlebenszeit <input type="checkbox"/> Ökonomisch orientierter Indikator - Arbeitsunfähigkeits-, KH-Pflegetage <input type="checkbox"/> Gesundheitsbezogene Lebensqualität <input type="checkbox"/> Lebensdauer und Lebensqualität integrierende Ergebnisgröße <input type="checkbox"/> Monetarisierter Ergebnisgröße

<p>Beschreibung der Effekte:</p> <ul style="list-style-type: none"> - Spezifikation von unmittelbar durch Einsatz der betreffenden Technik bedingten erwünschten Gesundheitszuständen - Spezifikation von unmittelbar durch Einsatz der betreffenden Technik bedingten unerwünschten Gesundheitszuständen (Nebenwirkungen bzw. Komplikationen) - Spezifikation von im weiteren Verlauf mittelbar durchlaufenen bzw. durchlaufbaren erwünschten Gesundheitszuständen - Spezifikation von im weiteren Verlauf mittelbar durchlaufenen bzw. durchlaufbaren unerwünschten Gesundheitszuständen (Nebenwirkungen bzw. Komplikationen) 	<p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p>
<p>Begründung für Auslassung bzw. Nicht-Berücksichtigung von Gesundheitszuständen</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> n/r</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> n/r</p>
<p><i>Klinische Effekt-Parameter bzw. Effekt-Wahrscheinlichkeiten</i></p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Nein</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Nein</p>
<p>Datenquelle(n) für in Analyse eingehende klinische Effekt-Parameter</p>	<p><input type="checkbox"/> Originalstudie - Primärdaten <input type="checkbox"/> Literatur - Sekundärdaten <input type="checkbox"/> Modellierung <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar</p>	<p><input type="checkbox"/> Originalstudie - Primärdaten <input type="checkbox"/> Literatur - Sekundärdaten <input type="checkbox"/> Modellierung <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar</p>
<p>Untersuchungsansatz bei Erhebung der klinischen Effekt-Parameter</p>	<p><input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar</p>	<p><input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar</p>
<p>Explizite Datierung (Kalenderjahr / Zeitraum) der Erhebung der klinischen Effekt-Parameter</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p>
<p>Explizite geographische Lokalisation der Erhebung von klinischen Effekt-Parametern</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p>

<i>Methodisch-qualitative Beurteilung der Datenquellen:</i>	<u>Checklisten der GERMAN SCIENTIFIC WORKING GROUP TECHNOLOGY ASSESSMENT FOR HEALTH CARE:</u>					
- HTA-Berichte	Checkliste 1a: Kontextdokumente					
- Systematische/narrative Reviews, Metaanalysen, Analysen gepoolter Rohdaten	Checkliste 1b: Systematische Reviews und Meta-Analysen					
- Randomisierte kontrollierte Studien	Checkliste 2a: Primärstudien (RCTs / Fall-Kontrollstudien / Kohortenstudien / Querschnittsstudien / Fallserien)					
- Kohortenstudien	Checkliste 2a: Primärstudien (RCTs / Fall-Kontrollstudien / Kohortenstudien / Querschnittsstudien / Fallserien)					
- Fall-Kontroll-Studien	Checkliste 2a: Primärstudien (RCTs / Fall-Kontrollstudien / Kohortenstudien / Querschnittsstudien / Fallserien)					
- Querschnittsstudien, Fallserien	Checkliste 2a: Primärstudien (RCTs / Fall-Kontrollstudien / Kohortenstudien / Querschnittsstudien / Fallserien)					
- Diagnose-Studien	Checkliste 2b: Diagnosestudien					
- Screening-Studien	(Checkliste in Entwicklung)					
- Modellierungen	(Checkliste in Entwicklung)					
Transparenz der Dokumentation von klinischen Effekt-Parametern bzw. -Wahrscheinlichkeiten hinsichtlich						
- Spezifikation der Input-Parameter	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein
- Beschreibung des(der) zugrundeliegenden Modelle(s)	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein
- Explizierung impliziter Annahmen	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein
- Beschreibung der Patienten-/ Probandenstichproben	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein
Überlebenszeit / (Rest-)Lebensdauer	<input type="checkbox"/> Ja <input type="checkbox"/> Nein		<input type="checkbox"/> Ja <input type="checkbox"/> Nein			
Datenquelle(n) für in Analyse eingehende Überlebenszeit / Lebensdauer	<input type="checkbox"/> Originalstudie - Primärdaten	<input type="checkbox"/> Literatur - Sekundärdaten	<input type="checkbox"/> Modellierung	<input type="checkbox"/> Kombination	<input type="checkbox"/> Originalstudie - Primärdaten	<input type="checkbox"/> Literatur - Sekundärdaten
	<input type="checkbox"/> Unklar				<input type="checkbox"/> Modellierung	<input type="checkbox"/> Kombination
	<input type="checkbox"/> Unklar				<input type="checkbox"/> Unklar	
Untersuchungsansatz bei Erhebung von Daten zu Überlebenszeit / Lebensdauer	<input type="checkbox"/> Retrospektiv	<input type="checkbox"/> Prospektiv	<input type="checkbox"/> Kombination	<input type="checkbox"/> Retrospektiv	<input type="checkbox"/> Prospektiv	<input type="checkbox"/> Kombination
	<input type="checkbox"/> Unklar			<input type="checkbox"/> Unklar		
Explizite Datierung (Kalenderjahr / Zeitraum) der Erhebung von Daten zu Überlebenszeit / Lebensdauer	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein
Explizite geographische Lokalisation der Erhebung von Daten zu Überlebenszeit / Lebensdauer	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein

Ermittlungsansatz für Überlebenszeit / Lebensdauer	<input type="checkbox"/> Sterbetafel <input type="checkbox"/> Gompertz-Funktion <input type="checkbox"/> Cox Proportional Hazards Modell, erweitertes Cox-Modell für zeitveränderliche Kovariablen <input type="checkbox"/> DEALE (Declining Exponential Approximation of Life Expectancy Method) Markov-Verfahren { <input type="checkbox"/> Markov-Kette (Matrixformulierung) <input type="checkbox"/> Markov-Kohorte <input type="checkbox"/> Unklar <input type="checkbox"/> Monte Carlo-Simulation <input type="checkbox"/> Anderer: _____ <input type="checkbox"/> Unklar	<input type="checkbox"/> Sterbetafel <input type="checkbox"/> Gompertz-Funktion <input type="checkbox"/> Cox Proportional Hazards Modell, erweitertes Cox-Modell für zeitveränderliche Kovariablen <input type="checkbox"/> DEALE (Declining Exponential Approximation of Life Expectancy Method) Markov-Verfahren { <input type="checkbox"/> Markov-Kette (Matrixformulierung) <input type="checkbox"/> Markov-Kohorte <input type="checkbox"/> Unklar <input type="checkbox"/> Monte Carlo-Simulation <input type="checkbox"/> Anderer: _____ <input type="checkbox"/> Unklar
Art der in Betracht gezogenen Überlebenszeit / Lebensdauer	a) <input type="checkbox"/> Geschlechtsspezifisch <input type="checkbox"/> Geschlechtsunabhängig <input type="checkbox"/> Unklar b) <input type="checkbox"/> Altersspezifisch <input type="checkbox"/> Altersunabhängig <input type="checkbox"/> Unklar c) <input type="checkbox"/> Krankheitsspezifisch <input type="checkbox"/> Krankheitenüberfassend <input type="checkbox"/> Krankheitsfrei <input type="checkbox"/> Unklar	a) <input type="checkbox"/> Geschlechtsspezifisch <input type="checkbox"/> Geschlechtsunabhängig <input type="checkbox"/> Unklar b) <input type="checkbox"/> Altersspezifisch <input type="checkbox"/> Altersunabhängig <input type="checkbox"/> Unklar c) <input type="checkbox"/> Krankheitsspezifisch <input type="checkbox"/> Krankheitenüberfassend <input type="checkbox"/> Krankheitsfrei <input type="checkbox"/> Unklar
Transparenz der Dokumentation von Überlebenszeit / Lebensdauer hinsichtlich - Spezifikation der Input-Parameter - Beschreibung des(der) zugrundeliegenden Modelle(s) - Explizierung impliziter Annahmen - Beschreibung der Patienten-/ Probandenstichproben	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Gesundheitsbezogene Lebensqualität	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Datenquelle(n) für in Analyse eingehende Lebensqualitätsbewertung	<input type="checkbox"/> Direkte Messung <input type="checkbox"/> Literaturwerte <input type="checkbox"/> Konsenswerte <input type="checkbox"/> Plausibilitätsüberlegungen <input type="checkbox"/> Unklar	<input type="checkbox"/> Direkte Messung <input type="checkbox"/> Literaturwerte <input type="checkbox"/> Konsenswerte <input type="checkbox"/> Plausibilitätsüberlegungen <input type="checkbox"/> Unklar
Untersuchungsansatz bei Erhebung von Lebensqualitätswerten	a) <input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar b) <input type="checkbox"/> Vor Erhebung der Effekt-Parameter <input type="checkbox"/> Nach Erhebung der Effekt-Parameter <input type="checkbox"/> Gleichzeitig zur Erhebung der Effekt-Parameter <input type="checkbox"/> Unklar	a) <input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar b) <input type="checkbox"/> Vor Erhebung der Effekt-Parameter <input type="checkbox"/> Nach Erhebung der Effekt-Parameter <input type="checkbox"/> Gleichzeitig zur Erhebung der Effekt-Parameter <input type="checkbox"/> Unklar

<p>Beschreibung des Bezugsrahmens der Bewertung: - Bewertungsansatz</p> <p>- Spezifikation der bewerteten Gesundheitszustände</p> <p>- Spezifikation der bei Bewertung der Gesundheitszustände berücksichtigten Dimensionen von Gesundheit</p> <p>- Spezifikation des jeweiligen Zeitbezugs der bewerteten Gesundheitszustände</p> <p>- Spezifikation der jeweiligen Zeitdauer der bewerteten Gesundheitszustände</p> <p>- Adjustierung von Lebensqualitätswerten für temporäre Gesundheitszustände bei Vergleich mit Lebensqualitätswerten für chronische Gesundheitszustände bzw. Tod (Standardisierung auf 0/1-Skala)</p> <p>- Berücksichtigung von negativen Lebensqualitätswerten</p> <p>- Spezifikation von Hilfsmitteln bei der Bewertung</p>	<p><input type="checkbox"/> Bewertung lebenslanger Gesundheitszustandspfade <input type="checkbox"/> Bewertung einzelner Gesundheitszustände <input type="checkbox"/> Unklar</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar <input type="checkbox"/> n/r</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p>	<p><input type="checkbox"/> Bewertung lebenslanger Gesundheitszustandspfade <input type="checkbox"/> Bewertung einzelner Gesundheitszustände <input type="checkbox"/> Unklar</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar <input type="checkbox"/> n/r</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p>
<p>Transparenz der Dokumentation von gesundheitsbezogener Lebensqualität hinsichtlich</p> <p>- Spezifikation der Input-Parameter</p> <p>- Beschreibung des(der) zugrundeliegenden Modelle(s)</p> <p>- Explizierung impliziter Annahmen</p> <p>- Beschreibung der Patienten-/ Probandenstichproben</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein</p>

Aggregatmaße (Lebensdauer und Lebensqualität integrierende Ergebnisgrößen)	Ja: <input type="checkbox"/> QALY-Konzept ("quality-adjusted life years") <input type="checkbox"/> DALY-Konzept ("disability-adjusted life years") <input type="checkbox"/> HYE-Konzept ("healthy-years equivalents") <input type="checkbox"/> SAVE-Konzept ("saved young life equivalents") <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein	Ja: <input type="checkbox"/> QALY-Konzept ("quality-adjusted life years") <input type="checkbox"/> DALY-Konzept ("disability-adjusted life years") <input type="checkbox"/> HYE-Konzept ("healthy-years equivalents") <input type="checkbox"/> SAVE-Konzept ("saved young life equivalents") <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein
Monetarisierte Gesundheitszustände	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Datenquelle(n) für in Analyse eingehende monetäre Bewertung von Gesundheitszuständen	<input type="checkbox"/> Direkte Messung <input type="checkbox"/> Literaturwerte <input type="checkbox"/> Konsenswerte <input type="checkbox"/> Plausibilitätsüberlegungen <input type="checkbox"/> Unklar	<input type="checkbox"/> Direkte Messung <input type="checkbox"/> Literaturwerte <input type="checkbox"/> Konsenswerte <input type="checkbox"/> Plausibilitätsüberlegungen <input type="checkbox"/> Unklar
Untersuchungsansatz bei Erhebung von monetären Werten für Gesundheitszustände	a) <input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar b) <input type="checkbox"/> Vor Erhebung der Effekt-Parameter <input type="checkbox"/> Nach Erhebung der Effekt-Parameter <input type="checkbox"/> Gleichzeitig zur Erhebung der Effekt-Parameter <input type="checkbox"/> Unklar	a) <input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar b) <input type="checkbox"/> Vor Erhebung der Effekt-Parameter <input type="checkbox"/> Nach Erhebung der Effekt-Parameter <input type="checkbox"/> Gleichzeitig zur Erhebung der Effekt-Parameter <input type="checkbox"/> Unklar
Explizite Datierung (Kalenderjahr / Zeitraum) der monetären Bewertung von Gesundheitszuständen	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Explizite geographische Lokalisation der monetären Bewertung von Gesundheitszuständen	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Spezifikation der Quellen für monetäre Bewertung von Gesundheitszuständen	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein

<p>Methode der monetären Bewertung von Gesundheitszuständen</p>	<p><input type="checkbox"/> Messung der Zahlungsbereitschaft - "willingness to pay" - <input type="checkbox"/> Direkt (Befragungen) - "contingent valuation"</p> <p><input type="checkbox"/> Messung der Akzeptanzbereitschaft - "willingness to accept/receive" - <input type="checkbox"/> Indirekt (Marktdaten) - "revealed preferences"</p> <p><input type="checkbox"/> Unklar</p> <p>Bewertung von bezahlter Arbeit:</p> <p>a) <input type="checkbox"/> Lohnkosten <input type="checkbox"/> Bruttoeinkommen</p> <p><input type="checkbox"/> Nettoeinkommen</p> <p><input type="checkbox"/> Durchschnittslohnkosten</p> <p><input type="checkbox"/> Durchschnittsbruttoeinkommen</p> <p><input type="checkbox"/> Durchschnittsnettoeinkommen</p> <p><input type="checkbox"/> Unklar</p> <p><input type="checkbox"/> Humankapitalansatz</p> <p>b) <input type="checkbox"/> Altersspezifisch <input type="checkbox"/> Geschlechtsspezifisch</p> <p><input type="checkbox"/> Unspezifisch <input type="checkbox"/> Unklar</p> <p>Bewertung von unbezahlter Arbeit oder Freizeit:</p> <p>Ja: <input type="checkbox"/> Minimaler Opportunitätskostenansatz</p> <p><input type="checkbox"/> Substitutionskostenansatz</p> <p><input type="checkbox"/> Andere: _____</p> <p><input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Unklar</p>	<p><input type="checkbox"/> Messung der Zahlungsbereitschaft - "willingness to pay" - <input type="checkbox"/> Direkt (Befragungen) - "contingent valuation"</p> <p><input type="checkbox"/> Messung der Akzeptanzbereitschaft - "willingness to accept/receive" - <input type="checkbox"/> Indirekt (Marktdaten) - "revealed preferences"</p> <p><input type="checkbox"/> Unklar</p> <p>Bewertung von bezahlter Arbeit:</p> <p>a) <input type="checkbox"/> Lohnkosten <input type="checkbox"/> Bruttoeinkommen</p> <p><input type="checkbox"/> Nettoeinkommen</p> <p><input type="checkbox"/> Durchschnittslohnkosten</p> <p><input type="checkbox"/> Durchschnittsbruttoeinkommen</p> <p><input type="checkbox"/> Durchschnittsnettoeinkommen</p> <p><input type="checkbox"/> Unklar</p> <p><input type="checkbox"/> Humankapitalansatz</p> <p>b) <input type="checkbox"/> Altersspezifisch <input type="checkbox"/> Geschlechtsspezifisch</p> <p><input type="checkbox"/> Unspezifisch <input type="checkbox"/> Unklar</p> <p>Bewertung von unbezahlter Arbeit oder Freizeit:</p> <p>Ja: <input type="checkbox"/> Minimaler Opportunitätskostenansatz</p> <p><input type="checkbox"/> Substitutionskostenansatz</p> <p><input type="checkbox"/> Andere: _____</p> <p><input type="checkbox"/> Nein</p> <p><input type="checkbox"/> Unklar</p>
<p>Nennung der Währung bei monetärer Bewertung von Gesundheitszuständen</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Nein</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Nein</p>
<p>Währungskonversion bei monetärer Bewertung von Gesundheitszuständen</p>	<p>Ja: <input type="checkbox"/> Wechselkurse</p> <p><input type="checkbox"/> Allgemeine Kaufkraftparitäten</p> <p><input type="checkbox"/> Gesundheitssystem-bezogene Kaufkraftparitäten</p> <p><input type="checkbox"/> Nein <input type="checkbox"/> Unklar</p>	<p>Ja: <input type="checkbox"/> Wechselkurse</p> <p><input type="checkbox"/> Allgemeine Kaufkraftparitäten</p> <p><input type="checkbox"/> Gesundheitssystem-bezogene Kaufkraftparitäten</p> <p><input type="checkbox"/> Nein <input type="checkbox"/> Unklar</p>
<p>Nennung des Bezugsjahres für monetäre Bewertung von Gesundheitszuständen</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Nein</p>	<p><input type="checkbox"/> Ja <input type="checkbox"/> Nein</p>
<p>Reflation der monetären Werte für Gesundheitszustände auf Bezugsjahr</p>	<p>Ja: <input type="checkbox"/> BIP-Inflator / BIP-Deflator</p> <p><input type="checkbox"/> General Medical Price-Index</p> <p><input type="checkbox"/> Sectoral Medical Price-Index</p> <p><input type="checkbox"/> Nein <input type="checkbox"/> Unklar <input type="checkbox"/> n/r</p>	<p>Ja: <input type="checkbox"/> BIP-Inflator / BIP-Deflator</p> <p><input type="checkbox"/> General Medical Price-Index</p> <p><input type="checkbox"/> Sectoral Medical Price-Index</p> <p><input type="checkbox"/> Nein <input type="checkbox"/> Unklar <input type="checkbox"/> n/r</p>

Transparenz der Dokumentation von monetisierten Gesundheitszuständen hinsichtlich - Spezifikation der Input-Parameter - Beschreibung des(der) zugrundeliegenden Modelle(s) - Explizierung impliziter Annahmen - Beschreibung der Patienten-/ Probandenstichproben	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
<i>Deskriptive Bemerkungen:</i>		<i>Bewertende Bemerkungen:</i>
4) Kosten (ökonomisches Outcome)	<i>Anwendung der Technologie(n):</i>	<i>Anwendung der Vergleichstechnologie(n):</i> <input type="checkbox"/> n/r
Direkte Kosten	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Spezifikation der berücksichtigten direkten Kosten	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <div style="margin-left: 20px;"> Variable Kosten: <input type="checkbox"/> medizinische Versorgung, Gesundheitsversorgung <input type="checkbox"/> Bereich außerhalb des Gesundheitswesens <input type="checkbox"/> Patient <input type="checkbox"/> Angehörige, Freunde <input type="checkbox"/> Transferkosten <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Fixkosten <input type="checkbox"/> Forschungs- und Entwicklungskosten <input type="checkbox"/> Lehrkosten an akademischen Einrichtungen <input type="checkbox"/> Andere: _____ </div>	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <div style="margin-left: 20px;"> Variable Kosten: <input type="checkbox"/> medizinische Versorgung, Gesundheitsversorgung <input type="checkbox"/> Bereich außerhalb des Gesundheitswesens <input type="checkbox"/> Patient <input type="checkbox"/> Angehörige, Freunde <input type="checkbox"/> Transferkosten <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Fixkosten <input type="checkbox"/> Forschungs- und Entwicklungskosten <input type="checkbox"/> Lehrkosten an akademischen Einrichtungen <input type="checkbox"/> Andere: _____ </div>

Datenquelle(n) für in Analyse eingehenden direkten Ressourcenverbrauch	<input type="checkbox"/> Originalstudie(n) - Primärdaten <input type="checkbox"/> Literatur - Sekundärdaten <input type="checkbox"/> Modellierung <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar	<input type="checkbox"/> Originalstudie(n) - Primärdaten <input type="checkbox"/> Literatur - Sekundärdaten <input type="checkbox"/> Modellierung <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar
Untersuchungsansatz bei Erfassung des direkten Ressourcenverbrauchs	<input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Modellierung <input type="checkbox"/> Unklar	<input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Modellierung <input type="checkbox"/> Unklar
Explizite Datierung (Kalenderjahr / Zeitraum) der Erfassung des direkten Ressourcenverbrauchs	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Explizite geographische Lokalisation der Erfassung des direkten Ressourcenverbrauchs	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Konzept zur Messung des direkten Ressourcenverbrauchs	<input type="checkbox"/> Mengenbestimmung nach medizinischen und nicht-medizinischen Leistungen <input type="checkbox"/> Mengenbestimmung nach Aufwand der Leistungserstellung - Kostenarten <input type="checkbox"/> Mengenbestimmung nach erstellten (Zwischen-)Produkten - Versorgungskontakte, Versorgungszeit, Einzelleistungen, Leistungskomplexe, Falltypen <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar	<input type="checkbox"/> Mengenbestimmung nach medizinischen und nicht-medizinischen Leistungen <input type="checkbox"/> Mengenbestimmung nach Aufwand der Leistungserstellung - Kostenarten <input type="checkbox"/> Mengenbestimmung nach erstellten (Zwischen-)Produkten - Versorgungskontakte, Versorgungszeit, Einzelleistungen, Leistungskomplexe, Falltypen <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar
Spezifikation des Mengengerüsts des direkten Ressourcenverbrauchs (Mengen an verbrauchten Ressourcen)	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Untersuchungsansatz bei monetärer Bewertung des direkten Ressourcenverbrauchs	a) <input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Modellierung <input type="checkbox"/> Unklar b) <input type="checkbox"/> Vor Erhebung des Ressourcenverbrauchs <input type="checkbox"/> Nach Erhebung des Ressourcenverbrauchs <input type="checkbox"/> Gleichzeitig zur Erhebung des Ressourcenverbrauchs <input type="checkbox"/> Unklar	a) <input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Modellierung <input type="checkbox"/> Unklar b) <input type="checkbox"/> Vor Erhebung des Ressourcenverbrauchs <input type="checkbox"/> Nach Erhebung des Ressourcenverbrauchs <input type="checkbox"/> Gleichzeitig zur Erhebung des Ressourcenverbrauchs <input type="checkbox"/> Unklar
Explizite Datierung (Kalenderjahr / Zeitraum) der monetären Bewertung des direkten Ressourcenverbrauchs	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Spezifikation der Quelle(n) für Preisangaben zum direkten Ressourcenverbrauch	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein

Monetäre Bewertung des Mengengerüstes des direkten Ressourcenverbrauchs (Art der Preise pro Einheit verbrauchter Ressource)	<input type="checkbox"/> Opportunitätskosten reflektierende Marktpreise <input type="checkbox"/> Opportunitätskosten reflektierende gewinnbereinigte Marktpreise <input type="checkbox"/> Administrative Preise <input type="checkbox"/> Selbstkosten <input type="checkbox"/> Pauschale Durchschnittspreise <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar	<input type="checkbox"/> Opportunitätskosten reflektierende Marktpreise <input type="checkbox"/> Opportunitätskosten reflektierende gewinnbereinigte Marktpreise <input type="checkbox"/> Administrative Preise <input type="checkbox"/> Selbstkosten <input type="checkbox"/> Pauschale Durchschnittspreise <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar
Spezifikation der Preise pro Einheit direkten Ressourcenverbrauchs	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Nennung der Währung bei Preisen für direkten Ressourcenverbrauch	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Währungskonversion bei monetärer Bewertung des direkten Ressourcenverbrauchs	<input type="checkbox"/> Ja <input type="checkbox"/> Wechselkurse <input type="checkbox"/> Allgemeine Kaufkraftparitäten <input type="checkbox"/> Teilweise <input type="checkbox"/> Gesundheitssystem-bezogene Kaufkraftparitäten <input type="checkbox"/> Unklar <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Wechselkurse <input type="checkbox"/> Allgemeine Kaufkraftparitäten <input type="checkbox"/> Teilweise <input type="checkbox"/> Gesundheitssystem-bezogene Kaufkraftparitäten <input type="checkbox"/> Unklar <input type="checkbox"/> Nein
Nennung des Bezugsjahres für monetäre Bewertung des direkten Ressourcenverbrauchs	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Reflation der monetären Werte für direkten Ressourcenverbrauch auf Bezugsjahr	Ja: <input type="checkbox"/> BIP-Inflator / BIP-Deflator <input type="checkbox"/> General Medical Price-Index <input type="checkbox"/> Sectoral Medical Price-Index <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	Ja: <input type="checkbox"/> BIP-Inflator / BIP-Deflator <input type="checkbox"/> General Medical Price-Index <input type="checkbox"/> Sectoral Medical Price-Index <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein <input type="checkbox"/> Unklar
Produktivitätsausfallkosten (indirekte Kosten)	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Spezifikation der berücksichtigten indirekten Kosten	<input type="checkbox"/> Ja { Morbidität: <input type="checkbox"/> Arbeitsunfähigkeit <input type="checkbox"/> Rehabilitation <input type="checkbox"/> Frühberentung <input type="checkbox"/> Unklar <input type="checkbox"/> Teilweise <input type="checkbox"/> Mortalität <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein	<input type="checkbox"/> Ja { Morbidität: <input type="checkbox"/> Arbeitsunfähigkeit <input type="checkbox"/> Rehabilitation <input type="checkbox"/> Frühberentung <input type="checkbox"/> Unklar <input type="checkbox"/> Teilweise <input type="checkbox"/> Mortalität <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein
Datenquelle(n) für in Analyse eingehende indirekte Ressourcenveränderungen	<input type="checkbox"/> Originalstudie(n) - Primärdaten <input type="checkbox"/> Literatur - Sekundärdaten <input type="checkbox"/> Modellierung <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar	<input type="checkbox"/> Originalstudie(n) - Primärdaten <input type="checkbox"/> Literatur - Sekundärdaten <input type="checkbox"/> Modellierung <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar
Untersuchungsansatz bei Erfassung der indirekten Ressourcenveränderungen	<input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Modellierung <input type="checkbox"/> Unklar	<input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Modellierung <input type="checkbox"/> Unklar

Explizite Datierung (Kalenderjahr / Zeitraum) der Erfassung der indirekten Ressourcenveränderungen	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Explizite geographische Lokalisation der Erfassung der indirekten Ressourcenveränderungen	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Spezifikation des Mengengerüsts der indirekten Ressourcenveränderungen (Mengen an veränderten Ressourcen)	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Untersuchungsansatz bei monetärer Bewertung der indirekten Ressourcenveränderungen	a) <input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar b) <input type="checkbox"/> Vor Erhebung der Ressourcenveränderungen <input type="checkbox"/> Nach Erhebung der Ressourcenveränderungen <input type="checkbox"/> Gleichzeitig zur Erhebung der Ressourcenveränderungen <input type="checkbox"/> Unklar	a) <input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar b) <input type="checkbox"/> Vor Erhebung der Ressourcenveränderungen <input type="checkbox"/> Nach Erhebung der Ressourcenveränderungen <input type="checkbox"/> Gleichzeitig zur Erhebung der Ressourcenveränderungen <input type="checkbox"/> Unklar
Explizite Datierung (Kalenderjahr / Zeitraum) der monetären Bewertung der indirekten Ressourcenveränderungen	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Spezifikation der Quelle(n) für monetäre Bewertung der indirekten Ressourcenveränderungen	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Monetäre Bewertung des Mengengerüsts der indirekten Ressourcenveränderungen (monetärer Wert pro Einheit veränderter Ressource)	Bewertung von bezahlter Arbeit: a) <input type="checkbox"/> Humankapitalansatz <input type="checkbox"/> Friktionskostenansatz <input type="checkbox"/> Unklar b) <input type="checkbox"/> Lohnkosten <input type="checkbox"/> Bruttoeinkommen <input type="checkbox"/> Nettoeinkommen <input type="checkbox"/> Durchschnittslohnkosten <input type="checkbox"/> Durchschnittsbruttoeinkommen <input type="checkbox"/> Durchschnittsnettoeinkommen <input type="checkbox"/> Unklar c) <input type="checkbox"/> Altersspezifisch <input type="checkbox"/> Geschlechtsspezifisch <input type="checkbox"/> Unspezifisch <input type="checkbox"/> Unklar Bewertung von unbezahlter Arbeit oder Freizeit: d) Ja: <input type="checkbox"/> Minimaler Opportunitätskostenansatz <input type="checkbox"/> Substitutionskostenansatz <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	Bewertung von bezahlter Arbeit: a) <input type="checkbox"/> Humankapitalansatz <input type="checkbox"/> Friktionskostenansatz <input type="checkbox"/> Unklar b) <input type="checkbox"/> Lohnkosten <input type="checkbox"/> Bruttoeinkommen <input type="checkbox"/> Nettoeinkommen <input type="checkbox"/> Durchschnittslohnkosten <input type="checkbox"/> Durchschnittsbruttoeinkommen <input type="checkbox"/> Durchschnittsnettoeinkommen <input type="checkbox"/> Unklar c) <input type="checkbox"/> Altersspezifisch <input type="checkbox"/> Geschlechtsspezifisch <input type="checkbox"/> Unspezifisch <input type="checkbox"/> Unklar Bewertung von unbezahlter Arbeit oder Freizeit: d) Ja: <input type="checkbox"/> Minimaler Opportunitätskostenansatz <input type="checkbox"/> Substitutionskostenansatz <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein <input type="checkbox"/> Unklar
Spezifikation der Preise pro Einheit indirekter Ressourcenveränderung	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein

Nennung der Währung bei Preisen für indirekte Ressourcenveränderungen	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Währungskonversion bei monetärer Bewertung der indirekten Ressourcenveränderungen	<input type="checkbox"/> Ja <input type="checkbox"/> Wechselkurse <input type="checkbox"/> Allgemeine Kaufkraftparitäten <input type="checkbox"/> Teilweise <input type="checkbox"/> Gesundheitssystem-bezogene Kaufkraftparitäten <input type="checkbox"/> Unklar <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Wechselkurse <input type="checkbox"/> Allgemeine Kaufkraftparitäten <input type="checkbox"/> Teilweise <input type="checkbox"/> Gesundheitssystem-bezogene Kaufkraftparitäten <input type="checkbox"/> Unklar <input type="checkbox"/> Nein
Nennung eines Bezugsjahres für monetäre Bewertung der indirekten Ressourcenveränderungen	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Reflation der monetären Werte für indirekte Ressourcenveränderungen auf Bezugsjahr	Ja: <input type="checkbox"/> BIP-Inflator / BIP-Deflator <input type="checkbox"/> General Medical Price-Index <input type="checkbox"/> Sectoral Medical Price-Index <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	Ja: <input type="checkbox"/> BIP-Inflator / BIP-Deflator <input type="checkbox"/> General Medical Price-Index <input type="checkbox"/> Sectoral Medical Price-Index <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Nein <input type="checkbox"/> Unklar
Intangible Kosten	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar
Identifikation von intangiblen Kosten	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar
Datenquelle(n) für in Analyse eingehende intangible Ressourcenveränderungen	<input type="checkbox"/> Originalstudie(n) - Primärdaten <input type="checkbox"/> Literatur - Sekundärdaten <input type="checkbox"/> Modellierung <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar	<input type="checkbox"/> Originalstudie(n) - Primärdaten <input type="checkbox"/> Literatur - Sekundärdaten <input type="checkbox"/> Modellierung <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar
Untersuchungsansatz bei Erfassung der intangiblen Ressourcenveränderungen	<input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar	<input type="checkbox"/> Retrospektiv <input type="checkbox"/> Prospektiv <input type="checkbox"/> Kombination <input type="checkbox"/> Unklar
Explizite Datierung (Kalenderjahr / Zeitraum) der Erfassung der intangiblen Ressourcenveränderungen	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Explizite geographische Lokalisation der Erfassung der intangiblen Ressourcenveränderungen	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Berücksichtigung intangibler Ressourcenveränderungen	Ja: <input type="checkbox"/> Separate Aufführung <input type="checkbox"/> Erfassung bei Ermittlung der indirekten Kosten <input type="checkbox"/> Erfassung bei Lebensqualitätsmessung <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	Ja: <input type="checkbox"/> Separate Aufführung <input type="checkbox"/> Erfassung bei Ermittlung der indirekten Kosten <input type="checkbox"/> Erfassung bei Lebensqualitätsmessung <input type="checkbox"/> Nein <input type="checkbox"/> Unklar

Folgekosten gewonnener (bzw. verlorener) Lebensjahre	Ja: <input type="checkbox"/> Gesundheitskosten für Krankheit(en) im Zusammenhang mit Anwendung der Technologie(n) <input type="checkbox"/> Gesundheitskosten für andere Erkrankungen <input type="checkbox"/> Gesundheitsunabhängige Kosten - Nahrung, Unterkunft, Bekleidung <input type="checkbox"/> Produktivität <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Unklar <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	Ja: <input type="checkbox"/> Gesundheitskosten für Krankheit(en) im Zusammenhang mit Anwendung der Technologie(n) <input type="checkbox"/> Gesundheitskosten für andere Erkrankungen <input type="checkbox"/> Gesundheitsunabhängige Kosten - Nahrung, Unterkunft, Bekleidung <input type="checkbox"/> Produktivität <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Unklar <input type="checkbox"/> Nein <input type="checkbox"/> Unklar
Begründung für Nicht-Berücksichtigung von Ressourcenverbrauch bzw. -veränderungen	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> n/r	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> n/r
Transparenz der Dokumentation des Mengengerüsts der berücksichtigten Ressourcenveränderungen hinsichtlich - Spezifikation der Input-Parameter - Beschreibung des(der) zugrundeliegenden Modelle(s) - Explizierung impliziter Annahmen - Beschreibung der Patienten-/Probandenstichprobe	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Transparenz der Dokumentation der monetären Bewertung der berücksichtigten Ressourcenveränderungen hinsichtlich - Spezifikation der Input-Parameter - Beschreibung des(der) zugrundeliegenden Modelle(s) - Explizierung impliziter Annahmen	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
<i>Deskriptive Bemerkungen:</i>	<i>Bewertende Bemerkungen:</i>	

5) Diskontierung	<i>Anwendung der Technologie(n):</i>	<i>Anwendung der Vergleichstechnologie(n):</i> <input type="checkbox"/> n/r
Diskontierung der Effekt-Parameter	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> n/r	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> n/r
Diskontierungssatz für Effekt-Parameter in Basisanalyse	<input type="checkbox"/> 0% <input type="checkbox"/> 3% <input type="checkbox"/> 5% <input type="checkbox"/> __% <input type="checkbox"/> Variabel <input type="checkbox"/> Unklar <input type="checkbox"/> n/r	<input type="checkbox"/> 0% <input type="checkbox"/> 3% <input type="checkbox"/> 5% <input type="checkbox"/> __% <input type="checkbox"/> Variabel <input type="checkbox"/> Unklar <input type="checkbox"/> n/r
Diskontierung der Kosten-Parameter	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> n/r	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein <input type="checkbox"/> n/r
Diskontierungssatz für Kosten-Parameter in Basis-Analyse	<input type="checkbox"/> 0% <input type="checkbox"/> 3% <input type="checkbox"/> 5% <input type="checkbox"/> __% <input type="checkbox"/> Variabel <input type="checkbox"/> Unklar <input type="checkbox"/> n/r	<input type="checkbox"/> 0% <input type="checkbox"/> 3% <input type="checkbox"/> 5% <input type="checkbox"/> __% <input type="checkbox"/> Variabel <input type="checkbox"/> Unklar <input type="checkbox"/> n/r
<i>Deskriptive Bemerkungen:</i>		<i>Bewertende Bemerkungen:</i>
6) Ergebnispräsentation	<i>Anwendung der Technologie(n):</i>	<i>Anwendung der Vergleichstechnologie(n):</i> <input type="checkbox"/> n/r
Zuordnung der indirekten Ressourcenveränderungen (Produktivitätskosten)	<input type="checkbox"/> Effekt-Seite <input type="checkbox"/> Kosten-Seite <input type="checkbox"/> Separate Aufführung <input type="checkbox"/> Unklar <input type="checkbox"/> n/r	<input type="checkbox"/> Effekt-Seite <input type="checkbox"/> Kosten-Seite <input type="checkbox"/> Separate Aufführung <input type="checkbox"/> Unklar <input type="checkbox"/> n/r
Zuordnung der intangiblen Auswirkungen	<input type="checkbox"/> Effekt-Seite <input type="checkbox"/> Kosten-Seite <input type="checkbox"/> Separate Aufführung <input type="checkbox"/> Unklar <input type="checkbox"/> n/r	<input type="checkbox"/> Effekt-Seite <input type="checkbox"/> Kosten-Seite <input type="checkbox"/> Separate Aufführung <input type="checkbox"/> Unklar <input type="checkbox"/> n/r
Dokumentation des Gesamt-(Netto-) Effekts	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Dokumentation des durchschnittlichen Effekts	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Dokumentation der Gesamt-(Netto-) Kosten	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Dokumentation von Durchschnittskosten	<input type="checkbox"/> Ja <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Nein
Dokumentation von summarischen durchschnittlichen Kosten-Effekt-Relationen (Quotienten bzw. Differenzen)	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Dokumentation des inkrementellen / marginalen Effekts (Effektivität bzw. Nutzwert bzw. Nutzen)	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	

Dokumentation der inkrementellen / marginalen Kosten	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein		
Dokumentation von summarischen inkrementellen / marginalen Kosten-Effektivitäts-, -Nutzwert- bzw. -Nutzen-Quotienten bzw. von summarischem Nettonutzen (Differenz)	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein		
Berechnung von aggregierten (populationsbezogenen) Ergebnissen (Kosten und/oder Effekte)	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein		<input type="checkbox"/> Ja	<input type="checkbox"/> Nein
Berechnung von aggregierten (populationsbezogenen) Ergebnissen (inkrementelle / marginale Kosten und/oder inkrementelle / marginale Effekte)	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein		
Dokumentation von differenzierenden disaggregierten Ergebnissen	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein		<input type="checkbox"/> Ja	<input type="checkbox"/> Nein
Nachvollziehbarkeit der Berechnungen	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
Konsistenz der Zahlen (soweit nachprüfbar)	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise	<input type="checkbox"/> Nein	<input type="checkbox"/> Ja	<input type="checkbox"/> Teilweise <input type="checkbox"/> Nein
<i>Subgruppen-Analyse(n)</i>	<i><input type="checkbox"/> Ja</i>	<i><input type="checkbox"/> Nein</i>		<i><input type="checkbox"/> Ja</i>	<i><input type="checkbox"/> Nein</i>
Identifikation von Effekt-Subgruppen	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein	<input type="checkbox"/> Unklar	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein <input type="checkbox"/> Unklar
Identifikation von Präferenz-/Nutzwert- bzw. Nutzen-Subgruppen	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein	<input type="checkbox"/> Unklar	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein <input type="checkbox"/> Unklar
Identifikation von Effektivitäts-Subgruppen	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein	<input type="checkbox"/> Unklar		
Identifikation von Mengengerüst-Subgruppen	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein	<input type="checkbox"/> Unklar	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein <input type="checkbox"/> Unklar
Identifikation von Preis-Subgruppen	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein	<input type="checkbox"/> Unklar	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein <input type="checkbox"/> Unklar
Identifikation von Kosten-Subgruppen	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein	<input type="checkbox"/> Unklar		
Identifikation von Kosten-Effektivitäts- bzw. -Nutzwert- bzw. Nutzen-Subgruppen	<input type="checkbox"/> Ja	<input type="checkbox"/> Nein	<input type="checkbox"/> Unklar		

<i>Deskriptive Bemerkungen:</i>		<i>Bewertende Bemerkungen:</i>	
7) Behandlung von Unsicherheiten	Anwendung der Technologie(n):	Anwendung der Vergleichstechnologie(n):	<input type="checkbox"/> <i>n/r</i>
Sensitivitätsanalyse	<input type="checkbox"/> <i>Ja</i> <input type="checkbox"/> <i>Nein</i>	<input type="checkbox"/> <i>Ja</i> <input type="checkbox"/> <i>Nein</i>	
Sensitivitätsanalyse bezüglich der Effekt-Parameter	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	
Sensitivitätsanalyse bezüglich der Kosten-Parameter	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	
Sensitivitätsanalyse bezüglich des Diskontierungssatzes	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	
Sensitivitätsanalyse bezüglich der Modellstruktur	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	
Art der Sensitivitätsanalyse(n)	<input type="checkbox"/> 1-Weg-Sensitivitätsanalyse <input type="checkbox"/> 2-Weg-Sensitivitätsanalyse <input type="checkbox"/> 3-Weg-Sensitivitätsanalyse <input type="checkbox"/> Unklar	<input type="checkbox"/> 1-Weg-Sensitivitätsanalyse <input type="checkbox"/> 2-Weg-Sensitivitätsanalyse <input type="checkbox"/> 3-Weg-Sensitivitätsanalyse <input type="checkbox"/> Unklar	
Spezifikation der Wertbereiche für die Sensitivitätsanalyse(n)	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	<input type="checkbox"/> Ja <input type="checkbox"/> Teilweise <input type="checkbox"/> Nein	
Extremwert-Analyse	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	
<i>best case / worst case</i> -Analyse	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	
Schwellenwert-Analyse	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	
Identifikation sensibler Parameter	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	<input type="checkbox"/> Ja <input type="checkbox"/> Nein <input type="checkbox"/> Unklar	

<p>Statistische Verfahren</p>	<p>Ja: <input type="checkbox"/> Delta-Methode</p> <p>Monte Carlo-Simulation { <input type="checkbox"/> 1. Ordnung <input type="checkbox"/> 2. Ordnung <input type="checkbox"/> Kombination</p> <p><input type="checkbox"/> Bootstrap-Verfahren <input type="checkbox"/> Bayes-Ansatz <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Unklar</p> <p><input checked="" type="checkbox"/> Nein</p>	<p>Ja: <input type="checkbox"/> Delta-Methode</p> <p>Monte Carlo-Simulation { <input type="checkbox"/> 1. Ordnung <input type="checkbox"/> 2. Ordnung <input type="checkbox"/> Kombination</p> <p><input type="checkbox"/> Bootstrap-Verfahren <input type="checkbox"/> Bayes-Ansatz <input type="checkbox"/> Andere: _____ <input type="checkbox"/> Unklar</p> <p><input checked="" type="checkbox"/> Nein</p>
<p><u>Deskriptive Bemerkungen:</u></p>		<p><u>Bewertende Bemerkungen:</u></p>

n/r = nicht relevant

Abramson N, Cotton S, Eckels R, Baldock J:
Voluntary Screening Program for Prostate Cancer: Detection Rate and Cost.
Southern Medical Journal, 1994; 87(8): 785-788.

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels Bestimmung von Serum-PSA. Kein expliziter Vergleich von Technologiealternativen.

1-2 Fragestellung

Beurteilung der Genauigkeit, Reliabilität und geschätzten Kosten der Früherkennung von Prostatakarzinom auf Gemeindebasis.

Screening-Strategie:

DRU+PSA –[⊕ *und/oder* >10ng/ml (bzw. 2mal >4ng/ml)]→ US *und/oder* Biopsie

Bei Probanden mit PSA-Spiegeln zwischen 4ng/ml und 10ng/ml wurde die PSA-Bestimmung wiederholt; wiederholte PSA-Spiegel von >4ng/ml wurden als positiv betrachtet. Ob alle Probanden mit positivem Vorbefund einer Ultraschalluntersuchung unterzogen wurden bzw. ob ein positiver Ultraschallbefund als Entscheidungskriterium für eine Biopsie galt ist unklar formuliert; die Kostenrechnung nimmt bei jedem positiven Testbefund die Durchführung sowohl einer Sonographie als auch einer Biopsie an.

1-3 Perspektive

Keine Angaben.

1-4 Zeithorizont

Ermittelt wurden Effekte und Kosten für den Zeitraum der Studie (Prävalenz-Screening): 4 Monate.

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Originalstudie.

Datenquelle für Gesundheitseffekt: Originalstudie (Querschnittstudie).

Datenquelle für Ressourcenverbrauch und Mengengerüst: Originalstudie (Querschnittstudie).

Datenquelle für Preise: Werte des lokalen Gesundheitszentrums für die ambulante Versorgung. Keine Angaben zu Quelle(n) für Preisangaben für die stationäre Versorgung bei operativer Behandlung. Auf Medicare-Daten basierende Preisangaben für die Strahlentherapie.

2-2 Datierung der zugrundeliegenden Daten

Keine Angaben bezüglich des Zeitpunktes und des gesamten Zeitraumes der Datenerhebung zum Gesundheitseffekt und Ressourcenverbrauch; PSA-Bestimmung und DRU erstreckten sich über vier Monate.

Die Preisangaben für das Screening beziehen sich auf die Werte des Jahres 1992.

2-3 Studienpopulation / Zielpopulation

Männer im Alter ab 40 Jahren ohne krankheitsspezifische Symptomatik im Einzugsgebiet des lokalen medizinischen Versorgungsnetzes.

2-4 Setting

Tangierte Versorgungsstrukturen: ambulante Versorgung (niedergelassene Mediziner, lokales Gesundheitszentrum), stationäre Versorgung.

Der klinische wie auch der ökonomische Teil der Studie wurden in Florida, USA durchgeführt.

2-5 Spezifikation der Technologie

Bezüglich des eingesetzten PSA-Assays, der DRU und der Ultraschalluntersuchungen wurden keine näheren Angaben gemacht; die Biopsien wurden transrektal ohne Angabe weiterer Spezifikationen durchgeführt. Blutabnahmen und Untersuchungen erfolgten multizentrisch; die Blut- und Gewebeproben wurden im gleichen Labor analysiert.

3 Gesundheitseffekte

Screening-/Diagnosestudie

3-1 Untersuchte Zielgrößen

Primäre Zielgröße war die nach Biopsieabsicherung identifizierte Anzahl Prostatakarzinome im Probandenkollektiv.

3-2 Ein-/Ausschlusskriterien

Als Einschlusskriterien wurden Geschlecht, Alter und das Fehlen krankheitsspezifischer Symptome genannt.

3-3 Rekrutierungsmodus

Kostenloses Screening-Angebot ohne persönliche Aufforderung, Bekanntmachung des Angebots zur Vorsorgeuntersuchung über Leserbrief in der Lokalzeitung.

3-4 Teilnahmerate

Keine Angaben.

3-5 Reproduzierbarkeit der Testergebnisse

Die eingesetzten Diagnosemethoden wurden lediglich umrissen. Die Entscheidungsgrenze bei der PSA-Bestimmung (PSA>10ng/ml bzw. bei wiederholter Bestimmung PSA>4ng/ml) wurde angegeben. Entscheidungskriterien bezüglich der Interpretation von DRU- und Ultraschallbefunden wurden nicht explizit genannt. Vorkehrungen zur Kontrolle der Variabilität von Durchführung und Interpretation der DRU und Ultraschalluntersuchung wurden nicht spezifiziert, die Blut- und Gewebeprobenanalysen erfolgten zentral.

3-6 Auswertung der Testergebnisse

Die Biopsie als Goldstandard wurde lediglich bei Probanden mit positivem DRU-Befund und/oder PSA-Werten von >10ng/ml bzw. bei wiederholter Bestimmung von >4ng/ml durchgeführt. Die Bestimmung der Testergebnisse und des Krankheitsstatus erfolgten nicht in wechselseitiger Verblindung.

3-7 Drop-Outs

Berichtetes Ausscheiden von acht der 572 Freiwilligen vor Beendigung der PSA-Bestimmung und DRU. Berichtetes Ausscheiden von zwei Karzinompatienten bei der Ergreifung von Behandlungsmaßnahmen.

3-8 Ergebnisse der Screening-/Diagnosestudie

Deskriptive Präsentation der Ergebnisse entlang der Interventionspfade ohne Anwendung statistischer Testverfahren; Angabe von Sensitivität und Spezifität als Kenngrößen der Tests.

21% (119/564) der Probanden hatten einen positiven Screening-Befund: 60 (11%) bei der PSA-Bestimmung, 61 (11%) bei der DRU und 18 (3%) sowohl bei der PSA-Bestimmung als auch bei der DRU. Die Biopsie bestätigte den Tumorverdacht in 18 (3.4%) Fällen: bei 4 Probanden mit positivem Befund bei der PSA-Bestimmung, bei 4 Probanden mit positivem DRU-Befund und bei 10 Probanden mit positivem Befund bei der DRU und PSA-Bestimmung. Tumor-Staging: 1 im Stadium A, 14 im Stadium B, 3 im Stadium C (unklar ist, ob dies klinische oder pathologische Stadien sind). Sensitivität und Spezifität der PSA-Bestimmung: 69% bzw. 89%, Sensitivität und Spezifität der DRU: 85% bzw. 89%. Einer Behandlung unterzogen sich 16 Karzinompatienten, 9 davon einer Operation, 7 einer Strahlentherapie.

3-9 Effektmaße für die ökonomische Analyse

Anzahl identifizierter Prostatakarzinome; Anzahl identifizierter und behandelter Prostatakarzinome.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

- Screening/Diagnostik: personeller Zeitaufwand, einmaliges PSA-Assay, DRU, Ultraschalluntersuchung und Biopsie einschließlich Histopathologie
- Behandlung

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach Aufwand der Leistungserstellung (Kostenarten) und nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Keine separate Spezifikation der einzelnen Mengen.

4-3 Monetäre Bewertung des Mengengerüsts

Die Art der Preise ist unklar. Zur Ermittlung der Kosten des Screenings wurden die lokal gültigen Preise für das Jahr 1992 eingesetzt; das Bezugsjahr für die Kosten der Behandlung wurde nicht genannt. Die Preise pro Mengeneinheit wurden aufgeführt.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die beschriebene Screening-Strategie führte zur Entdeckung von 18 Prostatakarzinomen.

6-2 Ermittelte Kosten

Die Gesamtkosten für das Screening allein für 564 Teilnehmer wurden auf US-\$ 130,000 berechnet, diejenigen für das Screening einschließlich der Behandlung von 16 identifizierten Karzinomträgern auf US-\$ 293,800. Die geschätzten Kosten pro Screening-Teilnehmer beliefen sich somit auf US-\$ 231 ohne Berücksichtigung der Behandlungskosten und auf US-\$ 520 bei Berücksichtigung der Behandlungskosten.

6-3 Synthese von Kosten und Effekten

Die Kosten pro entdecktem Prostatakarzinom wurden auf US-\$ 7,240 geschätzt. Die Screening- und Behandlungskosten pro entdecktem und behandeltem Prostatakarzinom wurden auf US-\$ 16,300 geschätzt.

7 Behandlung von Unsicherheiten

Keine Ausführungen.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Es wird ein differentieller Selektionsfehler durch freiwillige Teilnahme diskutiert.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Die Rate der Testpositiven und der entdeckten Karzinomfälle in Jacksonville wurde basierend auf Literatur den Werten aus drei anderen Populationen gegenübergestellt: Seattle/USA, St. Louis/USA, Schweden). Die eingeschränkte Generalisierbarkeit auf Populationen mit anderen Alterszusammensetzungen wird diskutiert.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass weder die Bestimmung von PSA noch die DRU für sich genommen ausreichen: durch eine Kombination beider Maßnahmen können ihrer Ansicht nach die besten diagnostischen Ergebnisse erzielt werden. Bevor sich jedoch Ausgaben für ein Früherkennungsprogramm rechtfertigen lassen, sollte dessen Nutzen hinsichtlich der Überlebenszeit beurteilt und Lebensqualitätsaspekte in Betracht gezogen werden.

9 Kommentar

Studie zu Biopsiekriterien beim Prävalenz-Screening ohne explizite Gegenüberstellung von Technologiealternativen. Als interessierende Zielgröße wurde das intermediäre Outcome entdeckte bzw. behandelte Karzinomfälle gewählt. Die Wahrscheinlichkeit eines selection bias und fehlender Interrater-Reliabilität, die Möglichkeit eines expectation bias bei der Durchführung der untersucherabhängigen Tests sowie die lediglich auf eine selektierte Populationsuntergruppe beschränkte Anwendung des Goldstandards in Verbindung mit einem work-up bias reduzieren die interne Validität, Reproduzierbarkeit und Generalisierbarkeit der klinischen Untersuchungsergebnisse in erheblichem Maße. Der Kostenberechnung mangelt es an Transparenz, die Dokumentation lässt eine Anpassung über die Zeit bzw. zwischen Gesundheitssystemen nicht zu. Bei den Ausführungen bezüglich der Probanden mit positivem Screening-

Befund, bei der Verwendung der Indizes der Testgüte sowie bei den Kostenberechnungen grobe Inkonsistenzen finden und die Darstellung der Studienergebnisse somit nicht nachvollziehbar ist. Die Arbeit weist erhebliche methodische Mängel auf.

Barry MJ, Fleming C, Coley CM, Wasson JH, Fahs MC, Oesterling JE:
Should Medicare Provide Reimbursement for Prostate-Specific Antigen Testing for Early Detection of Prostate Cancer? Part IV: Estimating the Risks and Benefits of an Early Detection Program.
Urology, 1995; 46(4): 445-461.

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels DRU und Bestimmung von Serum-PSA. Explizite Gegenüberstellung der Alternativen "Screening" und "Kein Screening".

1-2 Fragestellung

Schätzung der Risiken, des maximalen Nutzens und der Kosten eines (von der American Cancer Society empfohlenen) Programmes zur Früherkennung von Prostatakarzinom.

Screening-Strategie: PSA+DRU $-\{ >4\text{ng/ml und/oder } \oplus \}$ \rightarrow TRUS+Biopsie

1-3 Perspektive

Perspektive des Kostenträgers.

1-4 Zeithorizont

Ermittelt wurden die lebenslangen Effekte und Kosten im Rahmen des Prävalenz-Screenings.

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Synthese.

Datenquellen für klinische Parameter: klinische Entscheidungsanalyse mit Literaturdaten und Modellierung (Markov-Modell: Markov-Kohorte mit konstanter

Übergangswahrscheinlichkeit (vermutlich für 6-Monats-Zyklen). Das eingesetzte Markov-Modell ist eine Erweiterung des Modells von Fleming et al. (1993)¹.

Datenquelle für Ressourcenverbrauch und Mengengerüst: Modellierung (Entscheidungsanalyse).

Datenquelle für Preise: Medicare fee schedule (Medicare average allowable charge); für die Schätzung der bei terminaler Pflege von (hormonrefraktärem) Prostatakarzinom im Vergleich zu allen anderen Todesursachen zusätzlich anfallenden Kosten wurde auf eine Studie von Riley et al. (1987)² zurückgegriffen.

2-2 Datierung der zugrundeliegenden Daten

Zeitpunkt und Zeitraum der Erhebung von Daten zu klinischen Parametern wurden nicht explizit aufgeführt (die klinischen Parameter leiten sich aus Studien ab, welche bis 1995 veröffentlicht wurden).

Die Preisangaben beziehen sich auf die Werte des Jahres 1992.

2-3 Studienpopulation / Zielpopulation

Kohorten von Männern im Alter von 65, 70 und 75 Jahren.

2-4 Setting

Tangierte Versorgungsstrukturen: ambulante und stationäre Versorgung.

Teilweise Lokalisation der (nordamerikanischen) Daten zu den klinischen Parametern. Der ökonomische Teil der Studie wurde in den USA durchgeführt.

2-5 Spezifikation der Technologie

PSA-Assay: Entscheidungsgrenze bei 4ng/ml.

Bezüglich der DRU und TRUS wurden keine näheren Angaben gemacht; für die transrektale Nadelbiopsie wurde eine Antibiotikaprophylaxe angesetzt.

¹ Fleming C, Wasson JH, Albertsen PC, Barry MJ, Wennberg JE: *A Decision Analysis of Alternative Treatment Strategies for Clinically Localized Prostate Cancer*. JAMA 1993; 269: 2650-2658.

² Riley G, Lubitz J, Prihoda R, Rabey E: *The Use and Costs of Medicare Services by Cause of Death*. Inquiry 1987; 24: 233-244.

3 Gesundheitseffekte

Synthese bereits veröffentlichter Studien

3-1 In der Synthese untersuchte klinische Parameter

Primäre Zielgrößen waren die altersspezifische Prävalenz von Prostatakarzinom, stratifiziert nach Volumen und Ausdehnung (<0.5ml (insignifikant), >0.5ml intrakapsulär (IK: kapsuläre Penetration <1cm), >0.5ml extrakapsulär (EK: kapsuläre Penetration >1cm)), die Verteilung der Differenzierungsgrade in den einzelnen Tumorklassen (<0.5ml, >0.5ml IK und >0.5ml EK), die altersspezifischen Wahrscheinlichkeiten eines positiven Screening-Testbefundes, die altersspezifische Karzinomentdeckungsrate, die Verteilung der entdeckten Karzinome auf die Tumorklassen, die altersspezifische Compliance bei Indikation einer Behandlung, die Wahrscheinlichkeiten von Komplikationen bei Biopsie bzw. Behandlung, die Wahrscheinlichkeit einer späteren Behandlung bei Karzinomprogredienz oder BPH im Falle von Abwarten bzw. bei behandlungsbedingten Komplikationen sowie die alters- und differenzierungsgradspezifische Mortalität bei Behandlung bzw. bei Nichtbehandlung sowie die altersspezifische krankheitsunabhängige Mortalität.

3-2 Annahmen

Die Analyse basierte auf folgenden expliziten, Früherkennung und aggressive Therapie begünstigenden Annahmen: Bei der Alternative "Kein Screening" kommt es zur Karzinomprogredienz, die fatal enden kann. Bei der Alternative "Screening" wird in der Kosten-Effektivitäts-Analyse lediglich die radikale Prostatektomie berücksichtigt; für die Strahlentherapie werden unter Annahme der gleichen Effektivität wie die der radikalen Prostatektomie lediglich die behandlungsbedingte Morbidität sowie die Kosten pro Behandlungsfall ermittelt. Die Biopsie gilt als Goldstandard. Die Compliance bei Indikation einer Biopsie ist implizit in der Karzinomentdeckungsrate enthalten. Das Staging nach positiver Biopsie kann ein Knochenszintigramm, eine pelvine Computertomographie und ein Lymphknoten-Staging (Lymphadenektomie) umfassen. Organbegrenzte Karzinome (Fehlen einer kompletten kapsulären Penetration >1cm²) sind unabhängig von Prognosefaktoren (wie Differenzierungsgrad) durch aggressive Behandlung (radikale Prostatektomie oder Strahlentherapie) zu 100% heilbar. Die Prognose hängt einzig vom Differenzierungsgrad (hochdifferenziert bei Gleason-Score 2 bis 4, mäßigdifferenziert bei Gleason-Score 5 bis 7, geringdifferenziert bei Gleason-Score 8 bis 10) und nicht von der Tumorausdehnung ab; nur hochdifferenzierte Tumoren mit einem Volumen von <0.5ml werden als klinisch nicht signifikant (d.h. ohne

Metastasierungspotential) angesehen. (Mit diesen Annahmen wird die Metastasierungs- und krankheitsspezifische Mortalitätsrate nach radikaler Prostatektomie bei hochdifferenzierten Tumoren um 97%, bei mäßigdifferenzierten Tumoren um 70% bzw. bei geringdifferenzierten Tumoren um 56% gesenkt.) Eine radikale Prostatektomie verhindert vollständig das Risiko einer lokalen Karzinomprogredienz oder einer BPH. Nach einer radikalen Prostatektomie sind außer im Falle von Metastasierung keine zusätzlichen, adjuvanten Behandlungsmaßnahmen erforderlich. An behandlungsbedingten Komplikationen werden bei der Ermittlung der Effektivität perioperativer Tod, Impotenz und Inkontinenz, bei der Ermittlung der Kosten zusätzlich die Kosten für Harnröhrenstrikturen (ohne Diagnosekosten) berücksichtigt. Als impotent angesehen werden lediglich Männer, welche seit der Operation keine partielle oder vollständige Erektion hatten; als inkontinent eingestuft werden nur Männer, bei welchen sich die Schwierigkeiten beim Zurückhalten des Urins auf mehr als ein paar Tropfen pro Tag beziehen. Bei erektiler Dysfunktion werden lediglich die Kosten für penile Implantate, bei Inkontinenz nur die Kosten für künstliche Sphinkter innerhalb von zwei bis vier Jahren nach der Operation berücksichtigt. Diagnose und Behandlung bei Metastasierung umfassen ein Skelettszintigramm sowie eine Orchiektomie (50%) bzw. eine hormonelle Therapie (50%).

Für die einzelnen Gesundheitszustände werden keine Nutzwerte ermittelt: es liegt explizit beim Leser, Risiken und Nutzen der Screening-Entscheidung abzuwägen.

3-3 Berücksichtigung von Primärstudien: Studiendesigns und Ein-/Ausschlusskriterien

Die Krankheitsprävalenz wurde aus Autopsiestudien (unselektierte Fallserien) abgeleitet, in welchen ein klinischer Verdacht auf Prostatakarzinom als Ausschlusskriterium galt und die altersspezifische Prävalenz in 10-Jahres-Altersklassen aufgeführt wurde; in allen Studien wurden histologische Serienschnitte angewendet. Die altersspezifischen Wahrscheinlichkeiten eines positiven Screening-Testbefundes wie auch die altersspezifischen Karzinomentdeckungsraten, der Anteil extrakapsulärer Tumoren (>5ml) an den entdeckten Karzinomen sowie die Compliance bei Indikation einer Behandlung basieren auf den Daten der gemeindebezogenen prospektiven Kohortenstudie von Richie et al. (1993)³. Für die Wahrscheinlichkeiten von Inkontinenz, Impotenz bzw. Harnröhrenstrikturen nach radikaler Prostatektomie wurde auf einen Survey unter

³ Richie JP, Catalona WJ, Ahmann FR, Hudson NA, Scardino PT, Flanigan RC, DeKernion JB, Ratliff, TL, Kavoussi LR, Dalkin BL, et al.: *Effect of Patient Age on Early Detection of Prostate Cancer with Serum-Specific Antigen and Digital Rectal Examination*. Urology 1993; 42: 365-374.

Medicare-Empfängern⁴ zurückgegriffen. Die Wahrscheinlichkeit eines Eingriffs (TUR-P) bei Karzinomprogredienz oder BPH im Falle von Nichtbehandlung wurde aus prospektiven Kohortenstudien abgeleitet. Die altersspezifische krankheitsunabhängige Mortalität basiert auf amerikanischen Statistiken aus dem Jahr 1986, die alters- und differenzierungsgradspezifische Mortalität auf einer Meta-Analyse von Chodak et al. (1994)⁵. Das Design der anderen herangezogenen Studien wurde nicht explizit genannt. Explizit wurde darauf hingewiesen, dass die berücksichtigten Studien vorwiegend Daten zugunsten des Screenings beinhalteten. Ansonsten keine näheren Angaben zu Ein- bzw. Ausschlusskriterien bei der Auswahl der Studien.

3-4 Quellen und Suchstrategie bei der Literaturrecherche

Die genutzten Quellen sowie die Suchstrategie wurden nicht spezifiziert, auch nicht in den vorangegangenen Artikeln der vierteiligen Serie bzw. im HTA-Bericht (1995)⁶, der dieser Artikelreihe (als technischer Bericht) zugrunde liegt.

3-5 Validitäts- bzw. Qualitätskriterien bei der Bewertung der Primärstudien

Keine Angaben.

3-6 Methoden der Bewertung von Relevanz und Validität bzw. Qualität der Primärstudien

Keine Angaben.

3-7 Methoden der Extraktion von Daten aus den Primärstudien

Für die Verteilung der Differenzierungsgrade in den einzelnen Tumorklassen wurde die Publikation von Oesterling et al. (1993)⁷ herangezogen, an deren Rohdaten zusätzliche Analysen durchgeführt wurden. Ansonsten keine näheren Angaben.

⁴ Fowler FJ, Barry MJ, Lu-Yao G, Roman A, Wasson J, Wennberg JE: *Patient-Reported Complications and Follow-Up Treatment Following Radical Prostatectomy: The National Medicare Experience (1988-1990) (Updated June 1993)*. Urology 1993; 42: 622-629.

⁵ Chodak GW, Thisted RA, Gerber GS, Johansson JE, Adolfsson J, Jones GW, Chisholm GD, Moskovitz B, Livne PM, Warner J: *Results of Conservative Management of Clinically Localized Prostate Cancer*. New Engl J Med 1994; 330: 242-248.

⁶ U.S. Congress, Office of Technology Assessment (OTA): *Cost and Effectiveness of Prostate Cancer Screening in Elderly Men*. Washington, DC: U.S. Government Printing Office, 1995 (OTA-BP-H-145).

⁷ Oesterling JE, Suman VJ, Zincke H, Bostwick DG: *PSA-Detected (Clinical Stage T1c or B0) Prostate Cancer: Pathologically Significant Tumors*. Urol Clin North Am 1993; 20: 687-693.

3-8 Anzahl berücksichtigter Primärstudien

Die Ermittlung der altersspezifischen Krankheitsprävalenz basierte auf acht Studien; für die Wahrscheinlichkeiten der nach Volumen und Ausdehnung definierten Tumorklassen wurde auf eine Studie zurückgegriffen. Die perioperative Mortalität, die Wahrscheinlichkeit von Harnwegsinfektionen bei der Biopsie sowie die Wahrscheinlichkeit einer TUR-P bei Karzinomprogredienz oder BPH im Falle von Nichtbehandlung leiteten sich aus jeweils zwei Studien ab. Für die anderen Wahrscheinlichkeiten wurde jeweils eine Veröffentlichung - insgesamt sechs Publikationen - herangezogen.

3-9 Methode der Synthese der gesundheitsbezogenen Parameter

Zur Ermittlung der altersspezifischen Krankheitsprävalenz wurden gewichtete Durchschnittswerte für die einzelnen Alterskategorien gebildet. Die altersspezifischen Prävalenzschätzwerte für die nach Volumen und Ausdehnung definierten Tumorklassen wurden durch Rückgriff auf die Volumendaten der Publikation von McNeal et al. (1986)⁸ und unter der expliziten Annahme einer vom Alter unabhängigen Verteilung der Differenzierungsgrade in den einzelnen Tumorklassen-Strata abgeleitet. Für die Wahrscheinlichkeit einer TUR-P bei Karzinomprogredienz oder BPH im Falle von Nichtbehandlung wurde auf den gewichteten Durchschnitt der Raten der beiden verfügbaren Studien zurückgegriffen. Die altersspezifischen Wahrscheinlichkeiten für Männer im Alter von 70 Jahren werden als Durchschnitt der Wahrscheinlichkeiten für 65- und 75jährige gebildet.

3-10 Untersuchung von Heterogenität der gesundheitsbezogenen Parameter

Nicht untersucht.

3-11 Ergebnisse der Synthese

(Werte für die Basisanalyse; in Klammern Werte im plausiblen Bereich)

Die Krankheitsprävalenz lag für Männer im Alter von 65 Jahren bei 0.22, für Männer im Alter von 75 Jahren bei 0.39. Altersunabhängig wurden 60% der Karzinome als <0.5ml (insignifikant), 73% der Tumoren >0.5ml als intrazellulär (29.2%) und 27% der Tumoren >0.5ml als extrazellulär (10.8%) angesetzt. Die altersunabhängigen Wahrscheinlichkeiten für die Differenzierungsgrade in den einzelnen Tumorklassen waren folgende: Tumoren <0.5ml - hochdifferenziert 0.65, mäßigdifferenziert 0.26, geringdifferenziert 0.09;

⁸ McNeal JE, Bostwick DG, Kindrachuk RA, Redwine EA, Freiha FS, Stamey TA: *Patterns of Progression in Prostate Cancer*. Lancet 1986; i: 60-63.

Tumoren >0.5ml IK - hochdifferenziert 0.33, mäßigdifferenziert 0.56, geringdifferenziert 0.11; Tumoren >0.5ml EK - hochdifferenziert 0.04, mäßigdifferenziert 0.70, geringdifferenziert 0.26. Die Wahrscheinlichkeiten eines positiven Screening-Testbefundes lag für 65jährige bei 0.28, für 75jährige bei 0.40. Die Karzinomentdeckungsrate wurde für 65jährige auf 0.042 angesetzt, bei 75jährigen auf 0.072. Bei Männern im Alter von 65 Jahren waren dabei 11% der entdeckten Karzinome <0.5ml, 65% >0.5ml IK und 24% >0.5ml EK; bei Männern im Alter von 75 Jahren waren dies 11%, 49% bzw. 40%. Die Compliance bei Indikation einer Behandlung lag für 65jährige bei 70%, bei 75jährigen bei 48%. Die Wahrscheinlichkeit von Harnwegsinfektionen bei einer Biopsie wurde altersunabhängig auf 0.056 angesetzt, die einer Urosepsis auf 0.005. Die Wahrscheinlichkeit perioperativen Todes lag im Alter von 65 Jahren bei 0.0062, im Alter von 75 Jahren bei 0.0067. Die altersunabhängige Wahrscheinlichkeit von Inkontinenz bzw. Impotenz war 0.23 bzw. 0.6. 3% der Prostatektomie-Patienten erhielten eine penile Implantation, 6% einen künstlichen Sphinkter, 31% eine Dilatation oder Urethroplastik infolge von Harnröhrenstriktur. Für die Überlebenszeit wurden folgende Werte ermittelt: krankheitsunabhängige Überlebenszeit von 65jährigen 14.45 Jahre, von 75jährigen 8.95 Jahre; Überlebenszeit bei hochdifferenziertem Tumor <0.5ml mit und ohne Behandlung bzw. bei Tumoren >0.5ml (alle Differenzierungsgrade) mit Behandlung 14.45 Jahre im Alter von 65 Jahren, 8.95 Jahre im Alter von 75 Jahren; Überlebenszeit bei hoch- und mäßigdifferenziertem Tumor >0.5ml ohne Behandlung bzw. bei hoch- und mäßigdifferenziertem Tumor >0.5ml EK mit Behandlung 12.64 Jahre im Alter von 65 Jahren und 8.26 Jahre im Alter von 75 Jahren; Überlebenszeit bei geringdifferenziertem Tumor >0.5ml ohne Behandlung bzw. bei geringdifferenziertem Tumor >0.5ml EK mit Behandlung 7.57 Jahre im Alter von 65 Jahren und 6.01 Jahre im Alter von 75 Jahren. Die Wahrscheinlichkeit einer TUR-P bei Karzinomprogredienz oder BPH im Falle von Nichtbehandlung wurde nicht angegeben.

3-12 Effektmaße für die ökonomische Analyse

Gewonnene Lebensjahre. Parallele Darstellung von Inzidenz und den von Inkontinenz und Impotenz betroffenen Lebensjahren.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische ambulante und stationäre Versorgung.

- Screening/Diagnostik: Screening-Untersuchung, TRUS und Biopsie einschließlich Histopathologie, Staging
- Behandlung: Behandlung, Diagnose bei Metastasierung, Behandlung bei Metastasierung (Orchiektomie), TUR-P bei Komplikationen infolge lokaler Karzinomprogredienz oder BPH sowie sekundärer Therapie infolge behandlungsbedingter Komplikationen

Unter Berücksichtigung von Unsicherheiten bezüglich der Art des Ressourcenverzehr und der tatsächlichen Abrechnungspraxis wurden jeweils drei Kostenschätzungen (niedrige, mittlere und hohe Kostenkategorie) aufgestellt.

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach erstellten (Zwischen-)Produkten (CPT-4- / DRG-Positionen). Die sich aus der Entscheidungsanalyse ergebenden Mengen wurden nicht spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Administrative Preise; eingesetzt wurde ein nationaler Durchschnitt der erstatteten (im Gegensatz zu den eingereichten) Gebühren, basierend auf dem Medicare Resource Based Relative Value Scale (RBRVS). Die Preise pro Mengeneinheit wurden aufgeführt.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Zukünftige Kosten und Effekte wurden mit 5% diskontiert.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Im Rahmen des Prävalenz-Screenings wurden in Abhängigkeit vom Alter (65 bis 75 Jahre) folgende Raten ermittelt (jeweils pro 100,000 Screening-Teilnehmer): 19,330 bis 27,200 Biopsien; 18 bis 23 perioperative Todesfälle; 266 bis 317 Fälle mit behandlungsbedingter Inkontinenz, 1,334 bis 1,594 Fälle mit behandlungsbedingter Impotenz bzw. 398 bis 476 Fälle mit behandlungsbedingter Inkontinenz und Impotenz. 653 pro 100,000 Screening-Teilnehmer im Alter von 65 Jahren, 570 pro 100,000

Screening-Teilnehmer im Alter von 70 Jahren und 427 pro 100,000 Screening-Teilnehmer im Alter von 75 Jahren, bei denen ohne Screening und Frühbehandlung eine Metastasierung des Prostatakarzinoms erfolgt wäre, starben infolge anderer Todesursachen. Ohne Diskontierung betrug der Nettonutzen des Screenings pro 100,000 Screening-Teilnehmer 4,421 gewonnene Lebensjahre bei den 65jährigen, 2,844 gewonnene Lebensjahre bei den 70jährigen und 1,482 gewonnene Lebensjahre bei den 75jährigen, entsprechend 16, 10 bzw. 5 gewonnene Lebenstage pro Screening-Teilnehmer oder 1.52, 0.85 bzw. 0.43 gewonnene Lebensjahre pro operativ behandeltem Screening-Teilnehmer mit entdecktem Karzinom. Nach Diskontierung belief sich der Nettonutzen des Screenings pro 100,000 Screening-Teilnehmer auf 2,064 gewonnene Lebensjahre bei den 65jährigen, auf 1,442 gewonnene Lebensjahre bei den 70jährigen und auf 808 gewonnene Lebensjahre bei den 75jährigen, entsprechend 7.5, 5.3 bzw. 2.9 gewonnene Lebenstagen pro Screening-Teilnehmer.

6-2 Ermittelte Kosten

Die im Vergleich zur Alternative "Kein Screening" zusätzlich anfallenden Kosten beim Screening beliefen sich pro Screening-Teilnehmer im Alter von 65 Jahren auf US-\$ 258 in der niedrigen, auf US-\$ 293 in der mittleren und auf US-\$ 312 in der hohen Kostenkategorie (nach Diskontierung; keine Kostenangaben ohne Diskontierung). Pro Screening-Teilnehmer im Alter von 70 Jahren waren dies US-\$ 318 in der niedrigen, US-\$ 365 in der mittleren und US-\$ 390 in der hohen Kostenkategorie; pro Screening-Teilnehmer im Alter von 75 Jahren US-\$ 33.7 in der niedrigen, auf US-\$ 414 in der mittleren und auf US-\$ 45.0 in der hohen Kostenkategorie.

6-3 Synthese von Kosten und Effekten

Im Vergleich zur Alternative "Kein Screening" betrug der inkrementelle Kosten-Effektivitäts-Quotient (Kosten pro gewonnenem Lebensjahr) beim Screening für Männer im Alter von 65 Jahren US-\$ 12,502 in der niedrigen, US-\$ 14,200 in der mittleren und US-\$ 15,123 in der hohen Kostenkategorie. Für Männer im Alter von 70 Jahren waren dies US-\$ 22,028 in der niedrigen, US-\$ 25,289 in der mittleren und US-\$ 27,075 in der hohen Kostenkategorie; für Männer im Alter von 75 Jahren US-\$ 41,672 in der niedrigen, US-\$ 51,267 in der mittleren und US-\$ 55,681 in der hohen Kostenkategorie.

7 Behandlung von Unsicherheiten

Hinsichtlich der differenzierungsgradspezifischen Mortalitätsraten bei Metastasierung (keine explizite Angabe des Wertebereichs: Gegenüberstellung der Werte von Chodak et

al. und Fleming et al.) sowie der Heilungsrate von operativ behandelten organbegrenzten Karzinomen (Tumoren >0.5ml IK; zusätzlich angenommene Heilungsraten 50% und 75%) wurden 1-Weg-Sensitivitätsanalysen durchgeführt. Unter Annahme der niedrigeren krankheitsspezifischen Mortalitätsraten nach Fleming et al. stiegen die Kosten pro gewonnenem Lebensjahr auf US-\$ 42,590 bei 65jährigen bzw. US-\$ 177,094 bei 75jährigen (nach Diskontierung). Unter Annahme einer Heilungsrate von 75% lagen die Kosten pro gewonnenem Lebensjahr zwischen US-\$ 24,556 bei 65jährigen und US-\$ 109,091 bei 75jährigen (nach Diskontierung). Die 2-Weg-Sensitivitätsanalyse mit beiden Parametern ergab Kosten pro gewonnenem Lebensjahr von US-\$ 104,093 bei 65jährigen; bei Männern im Alter von 75 Jahren wurde das Screening durch die Alternative "Kein Screening" dominiert. Eine Erhöhung terminalen Versorgungskosten blieb ohne große Auswirkungen auf die Kosten-Effektivität: jede Erhöhung um US-\$ 10,000 brachte eine Reduzierung des Ergebnisses der Basisanalyse um US-\$ 30 mit sich.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren weisen darauf hin, dass die Kosten-Effektivität sehr von den zugrunde liegenden Annahmen zu unsicheren Parameter - insbesondere bezüglich des natürlichen Verlaufs von entdeckten Karzinomen und der Effektivität aggressiver Therapiemaßnahmen - beeinflusst wird; die für ihre Analyse formulierten Annahmen begünstigen explizit die Alternative "Screening".

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Die Autoren sind der Meinung, dass das Heranziehen der Ergebnisse ihrer Analyse zur Schätzung der bei Einführung eines Screening-Programmes beim Prostatakarzinom zusätzlich auf Medicare zukommenden Kosten problematisch ist, zumal es auch keine Angaben zur Inanspruchnahme eines solchen Programmes gibt. Unter der Annahme, dass 75% der 4.491 Mio. 65- bis 69jährigen Männer in den Vereinigten Staaten, 50% der 3.531 Mio. 70- bis 74jährigen Männer und 25% der 2.482 Mio. 75- bis 79jährigen Männer an einem solchen Früherkennungsprogramm teilnehmen würden und unter Anwendung der Kostenschätzungen für die jeweils jüngsten Männer dieser Altersgruppen würden sich die Kosten des ersten Jahres eines solchen Screeningprogrammes auf ungefähr US-\$ 2.1 Mrd. belaufen, wobei lediglich etwa US-\$ 0.043 Mrd. aufgrund diskontierter zukünftiger Einsparungen aus der Verhinderung von Behandlung bei BPH,

lokaler Karzinomprogredienz und metastasierendem Prostatakarzinom zurückzugewinnen wären.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass die Kosten des Screenings nicht so sehr durch die Screening-Untersuchungen selbst, als durch die abklärende Diagnostik (TRUS, Biopsie), die Staging-Untersuchungen und die therapeutischen Maßnahmen hervorgebracht werden. Die Autoren sind der Meinung, dass es klinischer Studien zum Vergleich der Alternativen "Screening" und "Kein Screening" bedarf, um die unsicheren Risiken und Wahrscheinlichkeiten, von denen die gesundheitspolitische Entscheidung für oder wider ein nationales Screening-Programm abhängt, genau zu bestimmen. Aus der Perspektive von Medicare stellt sich die Frage, ob die große, frühzeitig anfallende Investition für ein auf der PSA-Bestimmung basierendes Früherkennungsprogramm angesichts der Unsicherheit bezüglich des Nutzens und Schadens dieses Programmes gerechtfertigt ist. Auf der Basis ihrer Analyse und der ihr zugrundeliegenden Annahmen glauben die Autoren, dass es verfrüht ist, die PSA-Bestimmung zur Früherkennung von Prostatakarzinom in das Angebot von Medicare aufzunehmen. Hierfür sind Daten aus kontrollierten Studien abzuwarten.

9 Kommentar

Entscheidungsanalyse zum Prävalenz-Screening mit expliziter Gegenüberstellung von Technologiealternativen. Das von den Autoren beschriebene Modell erscheint sorgfältig aufgebaut und transparent. Zur Nachvollziehbarkeit der Effektermittlung und Sensitivitätsanalyse ist der Rückgriff auf die Modellierungsstudie von Fleming et al. sowie auf die diesbezüglichen Verbesserungsvorschläge von Beck et al. (1994)⁹ erforderlich. Nicht nachvollziehbar ist der Berechnungsprozess bei den unmittelbar anfallenden Kosten (TRUS/Biopsie, Staging) sowie bei den mittelbaren Kosten bei Karzinomprogredienz, BPH, Metastasierung und Behandlungskomplikationen. Es muss darauf hingewiesen werden, dass die Parameterschätzer der Basisanalyse nicht aus einer umfassenden Literaturrecherche und -analyse abgeleitet wurden; dieser selection bias bezüglich der in die Analyse eingehenden Resultate des Reviews begünstigt bewusst und explizit die Screening-Alternative. Durch die Sensitivitätsanalyse wurde teilweise auf dieses Problem eingegangen. Präzisierende Angaben zur Art des Screenings (Prävalenz- bzw. Inzidenz-Screening) sowie zur Anwendung des Goldstandards in den

⁹ Beck JR, Kattan MW, Miles BJ: *A Critique of the Decision Analysis for Clinically Localized Prostate Cancer*. J Urol 1994; 152: 1894-1899.

Studienpopulationen, aus deren Daten die Parameterwahrscheinlichkeiten abgeleitet wurden, fehlen. Bei der Kostenermittlung ist unklar, was terminale Versorgungskosten beinhalten und welchen Zeitraum sie abdecken. Die zum Teil nicht nachvollziehbare Ermittlung der oben genannten Kosten deckt sich auch nicht mit den in einer ähnlichen, späteren Publikation der gleichen Autorengruppe (Coley et al., 1997) aufgestellten Kostenberechnungen (siehe Hinweis unten).

10 Ähnliche Publikationen:

Coley CM, Barry MJ, Fleming D, Fahs MC, Mulley AG:

Early Detection of Prostate Cancer. Part II: Estimating the Risks, Benefits, and Costs.

Annals of Internal Medicine, 1997; 126(6): 468-479.

Die Kosten-Effektivitäts-Analyse basierte auf dem gleichen Modell wie die Analyse von Barry et al. Allerdings wurden Männerkohorten im Alter von 50 bis 59 Jahren, 60 bis 69 Jahren und 70 bis 79 Jahren betrachtet. Als Parameterschätzer für Männer im Alter von 60 bis 69 Jahren und von 70 bis 79 Jahren wurden die Wahrscheinlichkeiten für 65- bzw. 75jährige aus der Analyse von Barry et al. herangezogen; die Parameterschätzer für 55jährige als Wahrscheinlichkeiten für Männer im Alter von 50 bis 59 Jahren wurden neu ermittelt. Zur monetären Bewertung des direkten medizinischen Ressourcenverbrauchs wurde auf die Preise der mittleren Kostenkategorie aus der Publikation von Barry et al. zurückgegriffen. Im Abstract werden als Quelle der Literaturrecherche MEDLINE (1966 bis 1995), Reviews, Literaturverzeichnisse von aufgefundenen Publikationen, Dokumentensammlungen der Autoren und Abstracts genannt.

Die angenommenen altersspezifischen Wahrscheinlichkeiten für 55jährige waren folgende: Die Krankheitsprävalenz lag bei 0.15. 60% der Karzinome wurden als <0.5ml (insignifikant), 73% der Tumoren >0.5ml als intrazellulär (29.2%), 27% der Tumoren >0.5ml als extrazellulär (10.8%) angesetzt. Die Wahrscheinlichkeit eines positiven Screening-Testbefundes lag bei 0.15. Die Karzinomentdeckungsrate wurde auf 0.02 angesetzt, wobei 11% der entdeckten Karzinome <0.5ml, 63% >0.5ml IK und 26% >0.5ml EK waren. Die Compliance bei Indikation einer Behandlung wurde auf 71% angesetzt. Die Wahrscheinlichkeit perioperativen Todes lag bei 0.0057. Für die Überlebenszeit wurden folgende Werte ermittelt: krankheitsunabhängige Überlebenszeit 21.40 Jahre; Überlebenszeit bei hochdifferenziertem Tumor <0.5ml mit und ohne Behandlung bzw. bei Tumoren >0.5ml (alle Differenzierungsgrade) mit Behandlung bei 21.40 Jahre; Überlebenszeit bei hoch- und mäßigdifferenziertem Tumor >0.5ml ohne Behandlung bzw. bei hoch- und mäßigdifferenziertem Tumor >0.5ml EK mit Behandlung 17.54 Jahre; Überlebenszeit bei geringdifferenziertem Tumor >0.5ml ohne Behandlung

bzw. bei geringdifferenziertem Tumor >0.5ml EK mit Behandlung 8.60 Jahre. Die Wahrscheinlichkeit einer TUR-P bei Karzinomprogredienz oder BPH im Falle von Nichtbehandlung wurde nicht angegeben.

Daraus ergaben sich für 50-60jährige im Rahmen des Prävalenz-Screenings folgende Effekte für die ökonomische Analyse (jeweils pro 100,000 Screening-Teilnehmer): 10,752 Biopsien; 8 perioperative Todesfälle; 131 Fälle mit behandlungsbedingter Inkontinenz, 657 Fälle mit behandlungsbedingter Impotenz bzw. 196 Fälle mit behandlungsbedingter Inkontinenz und Impotenz. 407 pro 100,000 Screening-Teilnehmer, bei denen ohne Screening und Behandlung eine Metastasierung des Prostatakarzinoms erfolgt wäre, starben infolge anderer Todesursachen.

Ohne Diskontierung errechneten sich als Nettonutzen des Screenings pro 100,000 Screening-Teilnehmer für die Altersgruppen 50-59 bzw. 60-69 bzw. 70-79 Jahre folgende Werte: 4,606 bzw. 4612 bzw. 1618 gewonnene Lebensjahre. Nach Diskontierung belief sich der Nettonutzen des Screenings pro 100,000 Screening-Teilnehmer entsprechend auf 1,727 bzw. 2,064 bzw. 808 gewonnene Lebensjahre.

Die Autoren kamen zu folgenden Ergebnissen bei der Kostenberechnung: Die im Vergleich zur Alternative "Kein Screening" zusätzlich anfallenden Kosten beim Screening beliefen sich pro 100,000 Screening-Teilnehmer im Alter von 50 bis 59 Jahren auf US-\$ 21.6 Mio. (nach Diskontierung; keine Kostenangaben ohne Diskontierung). Pro 100,000 Screening-Teilnehmer im Alter von 60 bis 69 Jahren waren dies US-\$ 38.7 Mio., pro 100,000 Screening-Teilnehmer im Alter von 70 bis 79 Jahren US-\$ 53.2 Mio.

Daraus ergaben sich beim Vergleich der Alternativen "Screening" und "Kein Screening" folgende Kosten-Effektivitäts-Quotienten (Kosten pro gewonnenem Lebensjahr): Männer im Alter von 50 bis 59 Jahren US-\$ 12,491, Männer im Alter von 60 bis 69 Jahren US-\$ 18,769 und Männer im Alter von 70 bis 79 Jahre US-\$ 65,909.

In der 2-Weg-Sensitivitätsanalyse bezüglich der krankheitsspezifischen Mortalität - Werte von Chodak et al. und Fleming et al. - und der Heilungsrate - 50%, 75% und 100% - erwies sich beim Vergleich des Screenings mit der Alternative "Kein Screening" die Alternative "Kein Screening" bei folgenden Annahmekonstellationen als dominant: Altersgruppe der 60- bis 69jährigen - 50%ige Heilungsrate mit den krankheitsspezifischen Mortalitätsraten von Fleming et al.; Altersgruppe der 70- bis 79jährigen - 50%ige Heilungsrate mit den krankheitsspezifischen Mortalitätsraten von Chodak et al. bzw. von Fleming et al. sowie 75%ige Heilungsrate mit den krankheitsspezifischen Mortalitätsraten von Fleming et al. Bei Einsatz der Lebensqualitätsadjustierungen für Impotenz, Inkontinenz sowie metastasierendes bzw.

hormonrefraktäres Karzinom aus der Studie von Fleming et al. (Konsenswerte von Medizinern: 0.95, 0.85, 0.80 bzw. 0.50) erwies sich die Alternative "Kein Screening" zusätzlich bei folgenden Annahmekonstellationen als dominant: Altersgruppe der 50- bis 59jährigen - 50%ige Heilungsrate mit den Mortalitätsraten von Fleming et al.; Altersgruppe der 60- bis 69jährigen - 75%ige Heilungsrate mit den Mortalitätsraten von Fleming et al.; Altersgruppe der 70- bis 79jährigen - 100%ige Heilungsrate mit den Mortalitätsraten von Fleming et al.

Hinweis: in Tabelle 6 erscheinen bei der Altersgruppe 50-60Jährige 1,727 gewonnene Lebensjahre (diskontiert) pro 100,000 Screening-Teilnehmer. Dies entspricht ca. 6 gewonnenen Tagen pro Screening-Teilnehmer. Im Text zu Tabelle 6 ist im Widerspruch hierzu von 11 Tagen die Rede. In unseren Tabellen verwendeten wir einen Wert von 6 Tagen. Ferner ist der Berechnungsweg der Kosten pro Screening-Teilnehmer nicht nachvollziehbar dokumentiert.

Als technischer Report für die Artikelserie diente der HTA-Report des Office of Technology Assessment / US-Congress:

U.S. Congress, Office of Technology Assessment (OTA):

Cost and Effectiveness of Prostate Cancer Screening in Elderly Men.

Washington, DC: U.S. Government Printing Office, 1995 (OTA-BP-H-145).

Benoit RM & Naslund MJ:

An Economic Rationale for Prostate Cancer Screening.

Urology, 1994; 44(6): 795-803.

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels DRU und Bestimmung von Serum-PSA. Explizite Gegenüberstellung von der Früherkennung beim Prostatakarzinom und der Früherkennung beim Mammakarzinom mit Vergleich von jeweils zwei Screening-Strategien.

1-2 Fragestellung

Ermittlung und Vergleich der Kosten pro entdecktem Karzinom (erster Teil der Analyse) sowie der langfristigen Kosten-Effektivität (zweiter Teil der Analyse) der Früherkennung beim Prostatakrebs und beim Brustkrebs unter Rückgriff auf die Daten von jeweils zwei prospektiven Screening-Studien.

Prostatakarzinom: Screening-Strategie 1):

PSA+DRU -[>4ng/ml und/oder \oplus]→ TRUS+Biopsie

Screening-Strategie 2):

PSA -[>4ng/ml]→ DRU+TRUS -[$\oplus^{u/o}$]→ Biopsie

Mammakarzinom: Screening-Strategie 1):

klinische Brustuntersuchung+Mammographie → Biopsie

Screening-Strategie 2):

Mammographie → Biopsie

Keine Angaben der Biopsiekriterien hinsichtlich der Screening-Strategien beim Mammakarzinom.

Anmerkung: im weiteren soll lediglich die Analyse zum Screening des Prostatakarzinoms näher betrachtet werden. Schlussfolgerungen der Autoren bezüglich der Gegenüberstellung von Früherkennung beim Prostatakarzinom und Früherkennung beim Mammakarzinom werden nicht wiedergegeben.

1-3 Perspektive

Keine Angaben. Lediglich in der Diskussion wurden verschiedene Perspektiven erwähnt wie z.B. Kostenträger oder Patient.

1-4 Zeithorizont

Im ersten Teil der Analyse wurden die Effekte und Kosten für das erste Jahr des jeweiligen Früherkennungsprogrammes in der entsprechenden Studienpopulation ermittelt. Im zweiten Teil der Analyse wurden Langzeitüberlebensraten berücksichtigt.

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Synthese.

Datenquellen für klinische Parameter: Literatur.

Datenquelle für Ressourcenverbrauch und Mengengerüst: Literatur.

Datenquelle für Preise: durchschnittliche Gebühren im Rahmen von Medicare nach MAC-Positionen (Medicare allowed charges, Part B Extract and Summary System (BESS)).

2-2 Datierung der zugrundeliegenden Daten

Die Datenerhebung für den ersten Teil der Analyse basiert auf zwei Studien: die Studie von Richie et al.¹⁰ zum Prostatakarzinom-Screening erfolgte 1991 und 1992, die Studie von Catalona et al.¹¹ zum Mammakarzinom-Screening erfolgte von 1989 bis 1992. In die Analyse gingen die Daten des Jahres 1991 bzw. 1989 ein. Der Zeitpunkt und Zeitraum der Datenerhebung bei den im zweiten Teil der Analyse berücksichtigten Studien wurde nicht spezifiziert.

Die Preisangaben beziehen sich auf die Werte des Jahres 1992.

¹⁰ Richie JP, Catalona WJ, Ahmann FR, Hudson NA, Scardino PT, Flanigan RC, DeKernion JB, Ratliff, TL, Kavoussi LR, Dalkin BL, et al.: *Effect of Patient Age on Early Detection of Prostate Cancer with Serum-Specific Antigen and Digital Rectal Examination.* Urology 1993; 42: 365-374.

¹¹ Catalona WJ, Smith DS, Ratliff TL, Basler JW: *Detection of Organ Confined Prostate Cancer is Increased Through Prostate-Specific Antigen-Based Screening.* JAMA 1993; 270: 948-954.

2-3 Studienpopulation / Zielpopulation

Im ersten Teil der Analyse Männer im Alter zwischen 50 und 69 Jahren (Stratifizierung in die Altersgruppen 50 bis 59 Jahre und 60 bis 69 Jahre) (Richie et al.) bzw. im Alter zwischen 50 und 70 Jahren (Catalona et al.). Keine Angaben zur Zielpopulation im zweiten Teil der Studie.

2-4 Setting

In den Analysen wurde davon ausgegangen, dass alle Screening- und diagnostischen Tests bei niedergelassenen Ärzten durchgeführt werden.

Keine explizite Lokalisation der Daten zu den klinischen Parametern (die beiden Studien des ersten Teils der Analyse wurden in Nordamerika durchgeführt: das Design der Studie von Richie et al. wurde als multizentrisch beschrieben, die Studie von Catalona et al. als Washington University-Studie bezeichnet). Der ökonomische Teil der Studie wurde in Maryland, USA durchgeführt.

2-5 Spezifikation der Technologie

In der Studie von Richie et al. wurde die DRU von Urologen durchgeführt; eine palpatorisch unregelmäßige, asymmetrische, indurierte oder knotige Prostata wurde als verdächtig angesehen und einer transrektalen, ultraschallgesteuerten Feinnadelstanzbiopsie (Quadrantenbiopsie) unterzogen.

3 Gesundheitseffekte

Synthese bereits veröffentlichter Studien

3-1 In der Synthese untersuchte klinische Parameter

Primäre Zielgröße im ersten Teil der Analyse war der bei Anwendung des jeweiligen Früherkennungsprogrammes ermittelte Anteil positiver Screening-Testbefunde sowie der Anteil mittels Biopsie bestätigter Prostatakarzinome in der entsprechenden Studienpopulation. Für den zweiten Teil der Analyse wurde neben dem Anteil mittels Screening-Maßnahmen entdeckter organbegrenzter Tumoren der Anteil der beim Nicht-Screening (bzw. Screening mittels DRU allein) diagnostizierten organbegrenzten Karzinome ermittelt.

3-2 Annahmen

Jeweils nur die Ergebnisse des ersten Screening-Jahres der beiden für den ersten Teil der Analyse herangezogenen Studien von Richie et al. (1993, Screening-Strategie 1) und von Catalona et al. (1993, Screening-Strategie 2) wurden berücksichtigt.

3-3 Berücksichtigung von Primärstudien: Studiendesigns und Ein-/Ausschlusskriterien

Die zwei hauptsächlich herangezogenen Studien waren prospektive Kohortenstudien. Bezüglich der Ein- oder Ausschlusskriterien bzw. Beweggründe bei der Berücksichtigung der beiden Studien wurden keine Aussagen gemacht. Das Studiendesign der zur Ermittlung des Anteils organbegrenzter Tumoren beim Nicht-Screening (bzw. Screening mittels DRU allein) berücksichtigten Studien wurde nicht spezifiziert; auch hier wurden keine Angaben zu Ein- bzw. Ausschlusskriterien gemacht.

3-4 Quellen und Suchstrategie bei der Literaturrecherche

Keine Angaben.

3-5 Validitäts- bzw. Qualitätskriterien bei der Bewertung der Primärstudien

Keine Angaben.

3-6 Methoden der Bewertung von Relevanz und Validität bzw. Qualität der Primärstudien

Keine Angaben.

3-7 Methoden der Extraktion von Daten aus den Primärstudien

In der Studie von Richie et al. wurde der Anteil an durchgeführten TRUS und Biopsien sowie der Anteil entdeckter Prostatakarzinome in der Studienpopulation auf eine 100%ige Compliance bei der diagnostischen Abklärung eines positiven Screening-Testbefundes hochgerechnet (ursprünglich 69.2%). Keine weiteren Angaben zur Datenextraktion.

3-8 Anzahl berücksichtigter Primärstudien

Je Screening-Strategie wurde im ersten Teil der Analyse auf eine Studie zurückgegriffen. Für den zweiten Teil der Analyse wurden sechs weitere Studien herangezogen.

3-9 Methode der Synthese der gesundheitsbezogenen Parameter

Die für den ersten Teil der Analyse benötigten Studienergebnisse wurden direkt (Catalona et al.) bzw. modifiziert (Richie et al.: Adjustierung der Compliance) aus der jeweiligen Studie übernommen.

3-10 Untersuchung von Heterogenität der gesundheitsbezogenen Parameter

Nicht untersucht.

3-11 Ergebnisse der Synthese

(Werte für die Basisanalyse; in Klammern Werte im plausiblen Bereich)

Screening-Strategie 1): Der Anteil Männer in der Gesamtstudienpopulation im Alter von 50 bis 69 Jahren mit einem PSA-Wert von $>4\text{ng/ml}$ und/oder einem positiven DRU-Befund lag bei 22.3%. Die auf eine 100%ige Compliance mit der abklärenden TRUS und Biopsie hochgerechnete Karzinomentdeckungsrate lag bei 4.6%. Für die Männer in der Altersgruppe von 50 bis 59 Jahren lagen die entsprechenden Anteile bei 15.3% bzw. 2.9%, für Männer in der Altersgruppe von 60 bis 69 Jahren bei 28.0% bzw. 6.0%.

Screening-Strategie 2): Der Anteil Männer in der Gesamtstudienpopulation im Alter von 50 bis 70 Jahren mit einem PSA-Wert von $>4\text{ng/ml}$ lag bei 7.9%; 7.9% unterzogen sich einer TRUS, 7.5% einer Biopsie. Die Karzinomentdeckungsrate lag bei 2.5%.

Die 10- bis 15-Jahres-Überlebensraten nach Operation bei organbegrenztem Prostatakarzinom wurden auf 0.86 bis 0.92 angesetzt. Der Anteil entdeckter organbegrenzter Tumoren beim Screening lag in den Studien von Catalona et al. und Richie et al. bei 0.63 bzw. 0.74, derjenige beim Nicht-Screening (bzw. Screening mittels DRU allein) wurde auf 0.33 angesetzt; damit wurde für die Früherkennung eine potentielle 30- bis 41%ige Erhöhung der Entdeckungsrate von organbegrenztem Karzinom ermittelt.

3-12 Effektmaße für die ökonomische Analyse

Anzahl identifizierter Prostatakarzinome. Im zweiten Teil der Analyse wurden die Auswirkungen des Screenings auf die Mortalität in Betracht gezogen.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

- Screening/Diagnostik: PSA-Assay, DRU, Überweisung zum Urologen, diagnostische Abklärung positiver Untersuchungsbefunde bei Urologen incl. Sprechstundengebühr, TRUS und Biopsie einschließlich Histopathologie

Für die Kostenberechnung wurde angenommen, dass alle Screening-Teilnehmer mit einem positivem Untersuchungsbefund für die diagnostische Abklärung zu einem Urologen überwiesen würden, wofür eine zusätzliche Sprechstundengebühr angesetzt wurde.

Jeweils nur die Ergebnisse des ersten Screening-Jahres der beiden für den ersten Teil der Analyse herangezogenen Studien wurden berücksichtigt.

Die Kosten der Behandlung von Komplikationen des Screening- und Diagnoseprozesses sowohl beim Prostatakarzinom als auch beim Mammakarzinom wurden explizit nicht berücksichtigt. Sie wurden als ähnlich und hinsichtlich der Gesamtkosten nicht als bedeutsam angesehen.

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Die verrechneten Mengen konnten aus den Ergebnistabellen zu den klinischen Zielgrößen abgeleitet werden.

4-3 Monetäre Bewertung des Mengengerüsts

Administrative Preise: nationaler Durchschnitt der Medicare allowed charges (MAC), basierend auf dem Medicare Resource Based Relative Value Scale (RBRVS). Lediglich für die DRU konnte keine MAC-Gebühr angesetzt werden; hier basiert die monetäre Bewertung auf Annahmen (US-\$ 5 pro Untersuchung). Die Preise pro Mengeneinheit wurden aufgeführt.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Mittels der Screening-Strategie 1) wurden 46 Prostatakarzinome pro 1,000 Screening-Teilnehmern im Alter von 50 bis 69 Jahren entdeckt; die altersspezifische Rate lag in der Gruppe der 50-59jährigen bei 29, in der Gruppe der 60- bis 69jährigen bei 60 Prostatakarzinomen pro 1,000 Screening-Teilnehmern. Mittels der Screening-Strategie 2) wurden 25 Prostatakarzinome pro 1,000 Screening-Teilnehmer im Alter von 50 bis 70 Jahren entdeckt.

Die beim Screening im Vergleich zum Nicht-Screening erwartete Senkung der Mortalität wurde auf 25.8% bis 37.7% berechnet.

6-2 Ermittelte Kosten

Bei der Screening-Strategie 1) beliefen sich die Kosten auf US-\$ 109,096 pro 1,000 Screening-Teilnehmer im Alter von 50 bis 69 Jahren, auf US-\$ 85,636 pro 1,000 Screening-Teilnehmer im Alter von 50 bis 69 Jahren und auf US-\$ 128,199 pro 1,000 Screening-Teilnehmer im Alter von 50 bis 69 Jahren. Bei der Screening-Strategie 2) betragen die Kosten US-\$ 55,116 pro 1,000 Screening-Teilnehmer im Alter von 50 bis 70 Jahren.

6-3 Synthese von Kosten und Effekten

Die Kosten pro entdecktem Prostatakarzinom beliefen sich bei der Screening-Strategie 1) auf US-\$ 2,372 für die Altersgruppe von 50 bis 69 Jahren, auf US-\$ 2,953 für die Altersgruppe von 50 bis 59 Jahren und auf US-\$ 2,137 für die Altersgruppe von 60 bis 69 Jahren. Bei der Screening-Strategie 2) betragen die Kosten pro entdecktem Prostatakarzinom US-\$ 2,205 für die Altersgruppe von 50 bis 70 Jahren.

7 Behandlung von Unsicherheiten

Bei Annahme einer 20% höheren Überweisungsrate von Patienten mit für Allgemeinmediziner verdächtigem, für Urologen normalem Palpationsbefund wären die Kosten pro entdecktem Prostatakarzinom bei den Männern im Alter von 50 bis 69 Jahren der Studie von Richie et al. um US-\$ 31.78 (1.3%) gestiegen. Bei Annahme einer um

20% geringeren Überweisungsrate durch Allgemeinmediziner wäre in der gleichen Studie mit einem Rückgang der Karzinomentdeckungsrate um 2.4% und einer Senkung der Kosten um US-\$ 135.97 (5.7%) zu rechnen gewesen. Eine Verdoppelung angesetzten Kosten für die DRU hätte in der Studie von Richie et al. bei Männern im Alter von 50 bis 59 Jahren die Kosten pro entdecktem Prostatakarzinom um 5.8% auf US-\$ 3,125 erhöht; eine Vervierfachung der Gebühr wäre mit einer Kostensteigerung um 17.5% auf US-\$ 3,470 einhergegangen.

Eine strukturelle Sensitivitätsanalyse, die von einem seriellen Screening in einem 3-Jahres-Zeitraum ausging, ergab eine Kostensteigerung um 26,4% auf US-\$ 2,787 pro entdecktem Krebsfall für Männer zwischen 50 und 70 Jahren.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Auf die untersucherabhängige Rate positiver DRU-Befunde als mögliche Fehlerquelle ihrer Analyse gehen die Autoren in der Sensitivitätsanalyse ein. Die Autoren weisen wiederholt darauf hin, dass die Kosten-Effektivitäts-Berechnung bezüglich der Mortalitätssenkung rein hypothetisch ist, da hierzu bisher keine Daten aus randomisierten klinischen Studien vorliegen.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Die Autoren weisen darauf hin, dass in der klinischen Praxis in bezug auf die PSA-Spiegel auch andere Schwellenwerte bzw. Entscheidungskriterien (PSA-D, PSA-V) zum Einsatz kommen, welche Veränderungen hinsichtlich der Anzahl positiver Screening-Testbefunde und der Kosten pro entdecktem Prostatakarzinom mit sich bringen. Einen weiteren Problempunkt sehen die Autoren im Anteil entdeckter latenter, klinisch nicht signifikanter Prostatakarzinome, der je nach Screening-Strategie differieren kann: ist dieser Anteil hoch, steigen die Kosten pro entdecktem Karzinom.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass das Screening des Prostatakarzinoms mindestens genauso kosten-effektiv wie die Früherkennung des Mammakarzinoms ist. Diese Ansicht beruht auf folgender Ableitung: Die für das Prostatakarzinom ermittelten Kosten pro entdecktem Fall betragen in Abhängigkeit von Alter und Screening-Strategie zwischen 19% und 27% der Kosten pro entdecktem Mammakarzinom. Für das Screening

des Mammakarzinoms bei Frauen im Alter ab 50 Jahren ist eine Verminderung der Mortalität bis zu 30% aufgewiesen worden. Damit sich die Früherkennung des Prostatakarzinoms genauso kosten-effektiv wie die des Mammakarzinoms erweist, müsste die Mortalität beim Prostatakarzinom folglich lediglich um 5.7% bis 8.1% gesenkt werden (0.3×0.19 bzw. 0.27×100). In Anbetracht der screeningbedingten Erhöhung der Entdeckungsrate von organbegrenzten Tumoren halten die Autoren eine Verminderung der Mortalität in dieser Größenordnung für wahrscheinlich. Aus Sicht der Autoren sprechen die Ergebnisse der Studie für eine Berücksichtigung der Kostenübernahme von Screening für Prostatakrebs.

9 Kommentar

Studie zum Prävalenz-Screening; die explizite Gegenüberstellung von Technologiealternativen geschieht auf zwei Ebenen zwischen zwei indikationsgleichen, sich gegenseitig ausschließenden Screening-Strategien sowie zwischen zwei unabhängigen Gesundheitsprogrammen aus unterschiedlichen Bereichen der Medizin. Ein Teil der Analyse bezieht sich auf einen intermediären gesundheitsbezogenen Outcome-Parameter, der andere auf einen finalen; beide gesundheitsbezogene Outcome-Größen wurden dem gleichen intermediären ökonomischen Outcome (Kosten für Screening-Tests und diagnostische Abklärung ohne Berücksichtigung der im Zusammenhang mit Behandlung und Behandlungsergebnissen relevanten Ressourcenveränderungen) gegenübergestellt. Die Effektivitätsermittlung des zweiten Teils der Analyse kann lediglich als Spekulation angesehen werden, zumal aufgrund von stage shift-Daten nicht unbedingt auf einen positiven Effekt hinsichtlich der Morbidität und Mortalität geschlossen werden kann, sondern lead time bias, length bias und die Effektivität der therapeutischen Maßnahmen in Betracht gezogen werden müssen. Die Parameterschätzer wurden nicht aus einer umfassenden Literaturrecherche und -analyse abgeleitet; die Wahrscheinlichkeit eines selection bias ist gegeben. In der Sensitivitätsanalyse wurde nur teilweise auf dieses Problem eingegangen. Die Kostenberechnung der Basisanalyse ist nachvollziehbar; der Vergleich mit der Kostenanalyse beim Screening des Mammakarzinoms ist jedoch äußerst problematisch (z.B. unterschiedlicher Detaillierungsgrad der Kostenaufstellung, unterschiedliche Studiengrößen, Nichtberücksichtigung der Rekrutierungsmodi, Teilnahmeraten und Compliance, fehlende Kostenberechnung für primäre und komplikationsbedingte sekundäre Therapie). Die Berechnungen hinsichtlich einer 20%igen Steigerung bzw. Verminderung der Überweisungsraten sind nicht nachvollziehbar. Die interne Validität,

Reliabilität und Generalisierbarkeit der Ergebnisse ist aufgrund der methodischen Defizite stark eingeschränkt.

**Chadwick DJ, Kemple T, Astley JP, MacIver AG, Gillatt DA, Abrams P, Gingell JC:
Pilot Study of Screening for Prostate Cancer in General Practice.
Lancet, 1991; 338: 613-616.**

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels DRU und Bestimmung von Serum-PSA. Kein expliziter Vergleich von Technologiealternativen.

1-2 Fragestellung

Beurteilung verschiedener Rekrutierungsmodi sowie Ermittlung der relativen Güte der eingesetzten Screening-Tests und der Kosten des Früherkennungsprogrammes anhand einer praxisbezogenen Pilotstudie.

Screening-Strategie: PSA+DRU $-\{ >4\text{ng/ml und/oder } \oplus \}$ \rightarrow TRUS $-\{ \oplus \}$ \rightarrow Biopsie

1-3 Perspektive

Keine Angaben.

1-4 Zeithorizont

Ermittelt wurden Effekte und Kosten für den Zeitraum der Studie (Prävalenz-Screening):
1 Jahr

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Originalstudie.

Datenquelle für Gesundheitseffekt: Originalstudie (Querschnittstudie).

Datenquelle für Ressourcenverbrauch und Mengengerüst: Originalstudie (Querschnittstudie).

Datenquelle(n) für Preise: keine Angaben.

2-2 Datierung der zugrundeliegenden Daten

Die einjährige Datenerhebung zum Gesundheitseffekt und Ressourcenverbrauch begann im August 1989.

Keine Angaben hinsichtlich des Bezugsjahres für die Preisangaben.

2-3 Studienpopulation / Zielpopulation

Männer im Alter zwischen 55 und 69 Jahren im Einzugsgebiet des Bezirksgesundheitszentrums.

2-4 Setting

Tangierte Versorgungsstrukturen: ambulante Versorgung (Bezirksgesundheitszentrum); keine Angaben bezüglich der von der TRUS, Biopsie und Behandlung tangierten Versorgungsstruktur(en). Das Früherkennungsprogramm wurde als Teil einer allgemeineren Gesundheitsuntersuchung durchgeführt.

Der klinische wie auch der ökonomische Teil der Studie wurden in Großbritannien durchgeführt.

2-5 Spezifikation der Technologie

PSA-Assay: Hybritech Tandem-R (monoklonal).

DRU: Palpaton durch verschiedene Allgemeinärzte, Untersuchung in links-lateraler Lage.

TRUS: multipanar; keine Angaben bezüglich Schallkopffrequenz.

Biopsie: transrektale Stanzbiopsie (18G-Nadeln, Biopty-Pistole) unter Ultraschallkontrolle, keine Angaben bezüglich der Anzahl der Stanzen; Antibiotikaphylaxe.

3 Gesundheitseffekte

Screening-/ Diagnosestudie

3-1 Untersuchte Zielgrößen

Zielgrößen waren die Teilnahmerate in Abhängigkeit vom Rekrutierungsmodus sowie die nach Biopsieabsicherung identifizierte Anzahl Prostatakarzinome im Probandenkollektiv.

3-2 Ein-/Ausschlusskriterien

Als Einschlusskriterien wurden Geschlecht, Alter und Wohnort im Einzugsgebiet genannt.

3-3 Rekrutierungsmodus

- Gruppe A: Einladungsschreiben zu einer Gesundheitsuntersuchung beim Allgemeinarzt, fester Termin.
- Gruppe B: Einladungsschreiben zu einer Gesundheitsuntersuchung beim Allgemeinarzt, kein fester Termin.
- Gruppe C: Einladungsschreiben zu einer Gesundheitsuntersuchung in einer well-man-Praxis, fester Termin.
- Gruppe D: opportunistische Gesundheitsuntersuchung, wie Gruppe F: der vorgesehene Rekrutierungsmodus (Einladungsschreiben zu einer Gesundheitsuntersuchung in einer well-man-Praxis ohne festen Termin) wurde verhindert.
- Gruppe E: opportunistische Gesundheitsuntersuchung beim Allgemeinarzt mit Erinnerungsvermerk für den behandelnden Arzt in den betreffenden Patientenakten.
- Gruppe F: opportunistische Gesundheitsuntersuchung beim Allgemeinarzt ohne Erinnerungsvermerk für den behandelnden Arzt in den betreffenden Patientenakten

3-4 Teilnahmerate

82% (671/814) der Männer suchten die Praxis auf, bei 58% (472/814) konnte eine Gesundheitsuntersuchung vorgenommen werden. 86% (407/472) unterzogen sich einer DRU, 93% (437/472) unterzogen sich einer PSA-Bestimmung; keine Angaben darüber, wie viele Probanden sich lediglich einer der beiden bzw. beiden Untersuchungen unterzogen.

- Gruppe A: Screening bei 78% (106/136),
Aufsuchen der Praxis ohne Screening von 8% (11/136) der Männer.
- Gruppe B: Screening bei 64% (86/134),
Aufsuchen der Praxis ohne Screening von 21% (28/134) der Männer.
- Gruppe C: Screening bei 74% (100/136),
Aufsuchen der Praxis ohne Screening von 11% (15/136) der Männer.
- Gruppe D: Screening bei 33% (43/132),
Aufsuchen der Praxis ohne Screening von 49% (65/132) der Männer.
- Gruppe E: Screening bei 67% (95/141),
Aufsuchen der Praxis ohne Screening von 10% (14/141) der Männer.
- Gruppe F: Screening bei 31% (42/135),
Aufsuchen der Praxis ohne Screening von 49% (66/135) der Männer.

3-5 Reproduzierbarkeit der Testergebnisse

Die eingesetzten Diagnosemethoden wurden weitgehend spezifiziert. Die Entscheidungsgrenze bei der PSA-Bestimmung ($>4\text{ng/ml}$) wurde angegeben. Entscheidungskriterien bezüglich der Interpretation der DRU (palpatorisch indurierte oder knotige Prostata) wurden explizit genannt, nicht aber hinsichtlich der TRUS. Die TRUS wurde, im Gegensatz zur DRU, bei allen Probanden vom gleichen Untersucher durchgeführt. Vorkehrungen zur Kontrolle der Variabilität von Blut- und Gewebeprobenanalysen wurden nicht spezifiziert.

3-6 Auswertung der Testergebnisse

Die Biopsie als Goldstandard wurde lediglich bei Probanden mit positivem TRUS-Befund durchgeführt. Die Bestimmung der Testergebnisse und des Krankheitsstatus erfolgten nicht in gegenseitiger Verblindung.

3-7 Drop-Outs

Berichtetes Ausscheiden von sieben Probanden mit positivem Screening-Befund bei der TRUS.

3-8 Ergebnisse der Screening-/Diagnosestudie

Deskriptive Präsentation der Ergebnisse entlang des Interventionspfades ohne Anwendung statistischer Testverfahren; Angabe des positiven Vorhersagewertes (PPV) als Kenngröße der Tests.

15.9% (75/472) der Probanden wiesen einen positiven Screening-Befund auf, 62 bei der PSA-Bestimmung, 12 bei der DRU und 1 sowohl bei der PSA-Bestimmung als auch bei der DRU. Bei 29 der 68 Probanden, die sich einer TRUS unterzogen, fand sich ein verdächtiger Befund. Die Biopsie bestätigte den Tumorverdacht in 7 (1.7%) Fällen: bei 6 Probanden mit positivem Befund bei der PSA-Bestimmung und bei 1 Probanden mit positivem Befund bei der DRU und PSA-Bestimmung; in 14 Fällen wurde eine BPH diagnostiziert. PPV von PSA-Bestimmung und DRU in Kombination 0,10 (7/68), PPV von TRUS: 0,24. 5 Karzinomträger unterzogen sich einer radikalen Prostatektomie, bei 2 Karzinomträgern wurde die Behandlung hinausgeschoben (watchful waiting). Tumor-Staging: 5 im pathologischen Stadium pT2pN0M0, 2 im klinischen Stadium T0N0M0.

3-9 Effektmaße für die ökonomische Analyse

Anzahl identifizierter Prostatakarzinome.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

– Screening/Diagnostik: Sprechstunde, PSA-Assay, TRUS ohne Fixkosten für Ultraschall-Equipment, Biopsie einschließlich Histopathologie.

Behandlungskosten wurden nicht in die Analyse eingeschlossen.

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach Aufwand der Leistungserstellung (Kostenarten) und nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Die sich aus dem Interventionspfad ergebenden Mengen wurden spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Die Art, die Quelle(n) und das Bezugsjahr der Preise wurden nicht spezifiziert. Die Preise pro Mengeneinheit wurden aufgeführt.

4-4 Währung

Englische Pfund (GB-£).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die beschriebene Screening-Strategie führte zur Entdeckung von 7 Prostatakarzinomen.

6-2 Ermittelte Kosten

Die Gesamtkosten für das Screening (ohne Behandlung) beliefen sich auf GB-£ 11,578, entsprechend GB-£ 25 pro tatsächlich untersuchtem Screening-Teilnehmer.

6-3 Synthese von Kosten und Effekten

Die Kosten pro entdecktem Prostatakarzinom beliefen sich auf GB-£ 1,654.

7 Behandlung von Unsicherheiten

Keine Ausführungen.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Mit Bezug auf die im Vergleich zu Literaturangaben niedrigere Karzinomentdeckungsrate bei der DRU wird von den Autoren darauf hingewiesen, dass die Untersuchung durch Allgemeinärzte und nicht durch Urologen durchgeführt wurde. Weiterhin vermerken sie, dass die Sensitivität und Spezifität der Screening-Methoden nicht ermittelt werden können, da die wahre Prävalenz von Prostatakarzinom in der Studienpopulation nicht bekannt ist. Es wird darauf hingewiesen, dass psychologische Effekte nicht berücksichtigt wurden, die aber in die Evaluation eines Screeningprogrammes eingeschlossen werden sollten.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Die Autoren vermerken, dass die in der Studie ermittelten Kosten nur für diese Studie gelten und nicht für ähnliche Studien in anderen Gesundheitszentren.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass der effektivste Rekrutierungsmodus für eine Screening-Untersuchung eine Kombination aus Einladungsschreiben mit festem Termin und Erinnerungsvermerk für den behandelnden Arzt in der Patientenakte ist. Hinsichtlich der Evaluation des Screeningprogrammes vertreten die Autoren die Ansicht, dass nur anhand einer kontrollierten klinischen Studie mit langem follow up ermittelt werden kann, ob eine Verlängerung der Überlebenszeit als Hauptziel der Früherkennung erreichbar ist. Auch ist die Heilung von Patienten mit klinisch begrenztem Karzinom ihrer Meinung nach lediglich eine theoretische Möglichkeit, da der natürliche Verlauf von lokalisiertem Prostatakarzinom nicht bekannt ist.

9 Kommentar

Deskriptive Studie zu Biopsiekriterien beim Prävalenz-Screening ohne explizite Gegenüberstellung von Technologiealternativen. Als interessierende Zielgröße wurde ein intermediärer Outcome-Parameter gewählt. Die Möglichkeit eines selection bias sowie eines expectation bias ist gegeben, fehlende Interrater-Reliabilität ist wahrscheinlich. Zur Abhängigkeit des positiven Vorhersagewertes von der a priori-Wahrscheinlichkeit gibt es keine Äußerungen. Ferner wird die Anwendung des Goldstandards nur auf eine selektierte Populationsuntergruppe beschränkt, was zu einem work-up bias führt. Aus den genannten Gründen sind interne Validität, Reproduzierbarkeit und Generalisierbarkeit der klinischen Untersuchungsergebnisse eingeschränkt. Die Autoren geben sich dementsprechend konservativ in ihren Schlussfolgerungen. Zwei der entdeckten Karzinome waren (nach Angaben der Tabelle II) T0N0M0-Tumoren, nach der TNM-Klassifikation gleichbedeutend mit "Kein Anhalt für Primärtumor". Dies mindert zusätzlich die Qualität der Ergebnisse. Die Kostenberechnung ist nachvollziehbar, doch mangelt es ihrer Dokumentation an Transparenz. Kosten für das Staging wurden nicht berücksichtigt.

Hinweis: in Tabelle III (Kosten) der Publikation ist ein Druckfehler unterlaufen. Bei "Total cost of screening programme" sollte, wie im Text auch angeführt, GB-£ 11,587.72 statt GB-£ 1,157.72 stehen.

**Ciatto S, Bonardi R, Mazzotta A, Lombardi C, Santoni R, Cardini S, Zappa M:
Comparing Two Modalities of Screening for Prostate Cancer: Digital Rectal
Examination + Transrectal Ultrasonography vs. Prostate-Specific Antigen.
Tumori, 1995; 81: 225-229.**

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels (i) Bestimmung von Serum-PSA und (ii) Durchführung von DRU und TRUS. Explizite Gegenüberstellung von zwei Früherkennungsprogrammen.

1-2 Fragestellung

Vergleichende Beurteilung der praktischen Durchführbarkeit und Ausbeute zweier Früherkennungsprogramme bei Prostatakarzinom und Ermittlung deren Kosten

Strategie 1): DRU+TRUS $-\left[\oplus^{u/o} \right] \rightarrow$ PSA+Biopsie

Strategie 2): PSA $-\left[>4\text{ng/ml} \right] \rightarrow$ DRU +TRUS

$-\left[\oplus^{u/o} \text{ oder } \text{PSA}>10\text{ng/ml}+\text{PSA-D}>0.15\text{ng/ml/ccm} \right] \rightarrow$ Biopsie

Probanden mit zweifelhaften Befunden bei der DRU/TRUS bzw. mit einer PSA-Dichte von $>0.15\text{ng/ml/ccm}$ bei absoluten PSA-Werten $\leq 10\text{ng/ml}$ unterzogen sich innerhalb von drei bis sechs Monaten einer wiederholten DRU, TRUS und PSA-Bestimmung zur Abklärung des weiteren Vorgehens.

1-3 Perspektive

Keine Angaben.

1-4 Zeithorizont

Ermittelt wurden Effekte und Kosten für den Zeitraum des Prävalenzscreenings: 3-6 Monate.

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Originalstudie.

Datenquellen für Gesundheitseffekt: Originalstudien (Querschnittstudien).

Datenquellen für Ressourcenverbrauch und Mengengerüst: Originalstudien (Querschnittstudien).

Datenquelle(n) für Preise: keine Angaben.

2-2 Datierung der zugrundeliegenden Daten

Die Datenerhebung zum Gesundheitseffekt und Ressourcenverbrauch für das Screening-Programm 1) erfolgte 1992, diejenige für das Screening-Programm 2) 1993; der jeweilige Studienzeitraum wurde nicht explizit spezifiziert.

Keine Angaben hinsichtlich des Bezugsjahres für die Preisangaben.

2-3 Studienpopulation / Zielpopulation

Gesunde Männer im Alter von 60 bis 75 Jahren, gezogen aus den Patientenlisten der beteiligten Allgemeinärzte zweier Randbezirke einer italienischen Großstadt.

2-4 Setting

Tangierte Versorgungsstrukturen: ambulante Versorgung (niedergelassene Mediziner (Blutabnahme), Tumorzentrum (DRU und TRUS) und Zentrallabor eines Krankenhauses (Blutanalyse)).

Der klinische wie auch der ökonomische Teil der Studie wurden in Italien durchgeführt.

2-5 Spezifikation der Technologie

PSA-Assay: Abbott IMx (monoklonal).

DRU: Palpation durch einen erfahrenen Untersucher.

TRUS: 5MHz-Schallkopf, Durchführung durch einen erfahrenen Untersucher.

Biopsie: transperineale Stanzbiopsie (14-18G Nadeln) unter Ultraschallkontrolle, keine Angaben zur Anzahl der Stanzzyylinder.

3 Gesundheitseffekte

Screening-/Diagnosestudie

3-1 Untersuchte Zielgrößen

Zielgrößen waren die Teilnahmerate, die Rate zurückzurufender Probanden nach einer positiven PSA-Bestimmung (Screening-Programm 2)), die Biopsierate, die Rate nach Biopsieabsicherung identifizierter Karzinome, das klinische Stadium und der Differenzierungsgrad der identifizierten Tumoren sowie das Verhältnis von identifizierten zu erwarteten (altersspezifische Inzidenzen im lokalen Krebsregister) Karzinomen (= Prävalenz:Inzidenz-Verhältnis) im jeweiligen Probandenkollektiv.

3-2 Ein-/Ausschlusskriterien

Als Einschlusskriterien wurden Geschlecht, Alter und Wohnbezirk genannt. Die beiden Wohnbezirke wurden aufgrund der Vergleichbarkeit von Prostatakarzinom-Inzidenz und -Mortalität (lokales Krebsregister) gewählt.

3-3 Rekrutierungsmodus

Kostenloses Screening-Angebot. Vom Allgemeinarzt unterzeichnetes Einladungsschreiben sowie eine Nachfassaktion innerhalb von drei Monaten nach dem ersten Aufgebot; keine näheren Angaben zum Inhalt der Schreiben. In jedem Wohnbezirk wurde ein Screening-Programm durchgeführt (keine individuelle Randomisierung).

3-4 Teilnahmerate

Screening-Strategie 1):

33.7% (1,425/4,228: 1. Einladung 21.5% (908), 2. Einladung 12.2% (517)).

Screening-Strategie 2):

66.9% (1,315/1,965: 1. Einladung 49.8% (978), 2. Einladung 17.1% (337)).

3-5 Reproduzierbarkeit der Testergebnisse

Die eingesetzten Diagnosemethoden wurden spezifiziert. Die Entscheidungsgrenze bei der PSA-Bestimmung ($PSA > 4 \text{ ng/ml}$ bzw. $PSA > (\text{oder } =) 10 \text{ ng/ml}$ und $PSA-D > 0.15 \text{ ng/ml/ccm}$) wurde angegeben. Entscheidungskriterien bezüglich der Interpretation von DRU- und TRUS-Befunden wurden nicht explizit genannt. Die DRU und die TRUS wurden beide vom gleichen Untersucher durchgeführt, die Blutproben zentral analysiert.

3-6 Auswertung der Testergebnisse

Die Biopsie als Goldstandard wurde lediglich bei Probanden mit positivem DRU- und/oder TRUS-Befund bzw. bei Patienten mit PSA-Spiegeln >10ng/ml und einer PSA-Dichte von >0.15 durchgeführt. Die Bestimmung der jeweiligen Testergebnisse und des Krankheitsstatus erfolgten nicht in gegenseitiger Verblindung.

3-7 Drop-Outs

Screening-Strategie 1): Berichtetes Fehlen von je einem Fall beim Staging bzw. beim Grading.

Screening-Strategie 2): Berichtetes Ausscheiden von zwei Probanden bei der abklärenden DRU und TRUS sowie von einem Probanden bei der Biopsie.

3-8 Ergebnisse der Screening-/Diagnosestudie

Deskriptive Präsentation der Ergebnisse entlang der Interventionspfade unter Verwendung von Spezifität und positivem Vorhersagewert (PPV) mit 95%-Konfidenzintervallen als Kenngrößen der Tests.

Screening-Strategie 1): Verdacht auf Prostatakarzinom bei 2.74% (39/1,425: TRUS 38, DRU 25) der Probanden, positiver Biopsiefund bei 26 (1.82%: TRUS 25, DRU 17). Spezifität und PPV von TRUS: 99.1% (98.4-99.5%) bzw. 65.8% (50-81%); Spezifität und PPV von DRU: 99.4% (99.0-99.8%) bzw. 68.0% (49-86%). PSA<4ng/ml in 5, PSA 4-10ng/ml in 10 und PSA>10ng/ml in 11 Karzinomfällen. Tumor-Staging: 14 im Stadium B1, 4 im Stadium B2, 5 im Stadium C, 3 im Stadium D (Staging pathologisch in 7 Fällen nach radikaler Prostatektomie, klinisch in 19 Fällen mit Strahlen- oder Hormontherapie). Tumor-Grading: Gleason-Score 5 in 9, Gleason-Score 6 in 5, Gleason-Score 7 in 10 und Gleason-Score 9 in 1 Fall/Fällen. Verhältnis von identifizierten zu erwarteten Karzinomen 13.8:1 (9.1-20.4:1). Bei der Teilnehmerate und der Karzinomprävalenz war in Abhängigkeit vom Alter ein signifikanter negativer ($p<0.01$, χ^2 -Test auf Trend) bzw. positiver ($p<0.001$) Trend auszumachen. Insgesamt 71 zweifelhafte Befunde mit Untersuchungswiederholung bei der DRU und TRUS.

Screening-Strategie 2): 8.82% (116/1,315: PSA 4-9ng/ml 92, PSA>9ng/ml 22) der Probanden wurden infolge einer positiven PSA-Bestimmung zu einer DRU- und TRUS-Untersuchung bestellt. Verdacht auf Prostatakarzinom bei 2.81% (37/1,315): TRUS 22, DRU 17, PSA allein 13) der Probanden, positiver Biopsiefund bei 22 (1.67%: TRUS 17, DRU 14, PSA allein 3; PSA 4-9ng/ml 9, PSA>9ng/ml 13). Spezifität von PSA: 92.9% (91.4-94.2%); PPV von PSA 4-9ng/ml: 9.8% (3.7-15.8%), PPV von PSA>9ng/ml: 59%

(38.5-79.6%). Tumor-Staging: 3 im Stadium A, 8 im Stadium B1, 9 im Stadium B2, 1 im Stadium C, 1 im Stadium D (Staging pathologisch in 6 Fällen nach radikaler Prostatektomie, klinisch in 16 Fällen mit Strahlen- oder Hormontherapie). Tumor-Grading: Gleason-Score 5 in 5, Gleason-Score 6 in 6, Gleason-Score 7 in 2, Gleason-Score 8 in 6 und Gleason-Score 9 in 3 Fällen. Verhältnis von identifizierten zu erwarteten Karzinomen: 11.3:1 (7.1-17.1:1). Bei der Teilnahmerate und der Karzinomprävalenz war in Abhängigkeit vom Alter ein signifikanter negativer ($p < 0.001$, χ^2 -Test auf Trend) bzw. positiver ($p < 0.05$) Trend auszumachen. Insgesamt 20 zweifelhafte Befunde mit Untersuchungswiederholung bei der DRU und TRUS.

3-9 Effektmaße für die ökonomische Analyse

Anzahl identifizierter Prostatakarzinome je Screening-Strategie.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

- Screening-Untersuchung: Personal- und Materialkosten, Abklärung und Biopsien
- Organisation und Datenauswertung

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach Aufwand der Leistungserstellung (Kostenarten) und nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Keine Spezifikation der einzelnen Mengen.

4-3 Monetäre Bewertung des Mengengerüsts

Keine Angaben bezüglich der Art, der Quelle(n) und des Bezugsjahres der Preise. Keine Spezifikation der Preise pro Mengeneinheit. Die angegebenen Kosten wurden anhand der realen Ausgaben für die beiden Screening-Strategien ermittelt. Für Materialkosten wurde eine jährliche Amortisationsrate von 5% angenommen.

4-4 Währung

Italienische Lire (Lit.).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die Screening-Strategie 1) führte zur Entdeckung von 26 Prostatakarzinomen (1.82%), die Screening-Strategie 2) zur Entdeckung von 22 Prostatakarzinomen (1.67%).

6-2 Ermittelte Kosten

Die Gesamtkosten für die Screening-Strategie 1) beliefen sich auf Lit. 48,210,000 mit einem durchschnittlichen Betrag von Lit. 33,750 pro Screening-Teilnehmer; die Gesamtkosten für die Screening-Strategie 2) beliefen sich auf Lit. 39,986,000 mit einem durchschnittlichen Betrag von Lit. 30,400 pro Screening-Teilnehmer.

6-3 Synthese von Kosten und Effekten

Die Kosten pro entdecktem Prostatakarzinom beliefen sich beim Screening-Programm 1) auf Lit. 1,854,000 und beim Screening-Programm 2) auf Lit. 1,817,000.

7 Behandlung von Unsicherheiten

Strukturelle Sensitivitätsanalyse: ein Cutpoint von > 9 ng/ml würde eine Spezifität von 98.9% und eine Krebsentdeckungsrate von 0.99% bezüglich der Stichprobenpopulation ergeben.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren sind der Meinung, dass die Ergebnisse der vorliegenden Studie keine Schlussfolgerungen bezüglich der Effektivität von Screening erlauben: das hohe Verhältnis von identifizierten zu erwarteten Karzinomen ist für die Autoren eher ein Hinweis auf eine Vorverlegung des Zeitpunktes der Diagnosestellung, bei welcher keine Möglichkeit zur Schätzung des Ausmaßes von overdiagnosis besteht.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Keine Ausführungen.

8-3 Schlussfolgerungen

Nach Meinung der Autoren bestätigt die Studie, dass die Früherkennung von Prostatakarzinom möglich und praktisch durchführbar ist. Die Autoren führen an, dass mit beiden Screening-Strategien vergleichbare Ergebnisse in bezug auf die Karzinomentdeckungsrate erzielt wurden, wobei sich die Sensitivität von TRUS nur wenig von derjenigen der PSA-Bestimmung unterschied (in Screening-Strategie 1) waren in 5 von 26 Karzinomfällen PSA-Werte $<4\text{ng/ml}$ ermittelt worden, in Screening-Strategie 2) war der TRUS-Befund in 5 von 22 Karzinomfällen negativ). Anhand ihrer Studienergebnisse können die Autoren die in der Literatur dokumentierte niedrige Spezifität von der TRUS nicht bestätigen, und auch die Meinung, dass eine auf der TRUS basierende Screening-Strategie zu teuer ist, kann von ihnen nicht geteilt werden. Ein auf der PSA-Bestimmung beruhendes Screening-Programm kommt etwas kostengünstiger und weist eine höhere Teilnahmerate auf. Sie AUtoren schlussfolgern, dass keine Evidenz für die Wirksamkeit von Screening für Prostatakarzinom besteht. Die Empfehlungen für Screening wurden in Frage gestellt. Die Autoren empfehlen die PSA-Bestimmung für randomisierte kontrollierte Studien zur Bestimmung der Kosten-Effektivität und der möglichen negativen Konsequenzen einer Früherkennung von Prostatakarzinom.

9 Kommentar

Studie zu Biopsiekriterien beim Prävalenz-Screening mit Gegenüberstellung von Technologiealternativen. Im Vordergrund des Interesses steht die Ermittlung der Durchführbarkeit einzelner diagnostischer Strategien und weniger der langfristige Nutzen der einzelnen Früherkennungsprogramme bzw. deren Kosten, was sich auch in der Wahl eines intermediären Outcome-Parameters als interessierende klinische Zielgröße ausdrückt. Die Möglichkeit eines selection bias sowie eines expectation bias ist gegeben. Die Bezugsgruppen bei der Berechnung von Sensitivität und Spezifität bleiben unklar und die lediglich auf eine selektierte Populationsuntergruppe beschränkte Anwendung des Goldstandards ist mit einem work-up bias verbunden. Zur Abhängigkeit des positiven Vorhersagewertes von der a priori-Wahrscheinlichkeit gibt es keine Äußerungen. Die interne Validität, Reproduzierbarkeit und Generalisierbarkeit der klinischen Untersuchungsergebnisse sind stark eingeschränkt, die Autoren geben sich dementsprechend vorsichtig und konservativ in ihren Schlussfolgerungen. Der

Kostenberechnung mangelt es an Transparenz und Nachvollziehbarkeit; Protokollkosten (z.B. Datenauswertung) scheinen mit einbezogen worden zu sein.

Hinweis: in der englischsprachigen Publikation der Studie treten einige unstimmmige Prozentzahlen auf. In der italienischsprachigen Originalpublikation sind diese richtig wiedergegeben.

10 Originalpublikation:

Ciatto S, Bonardi R, Mazzotta A, Lombardi C, Santoni R, Cardini S, Zappa M:
Confronto Tra Due Modalità di Screening per il Carcinoma Prostatico. Esplorazione
Rettale ed Ecografia Transrettale vs. Antigene Prostatico Specifico.
Radiologia Medica (Torino), 1994; 88: 453-457.

**de Biasi F, Londero D, Vanuzzo D, Pilotto L, Praturlon S, Guerra UP:
Is Serum PSA Measurement a Good Test for Early Diagnosis of Prostate Cancer?
Results of a Pilot Study and Cost Analysis.
International Journal of Biological Markers, 1996; 11(1): 18-23.**

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels Bestimmung von Serum-PSA. Kein expliziter Vergleich von Technologiealternativen.

1-2 Fragestellung

Ermittlung der regionalen Einsetzbarkeit einer auf der PSA-Bestimmung basierenden Screening-Strategie anhand einer gemeindebezogenen Pilotstudie; Vergleich der festgestellten Prävalenz von durch Screening entdecktes Prostatakarzinom mit Literaturdaten, Abschätzung der Screening-Compliance und Aufstellung einer Kostenanalyse.

Screening-Strategie:

Fragebogen zu urologischen Störungen, auszufüllen mit Hilfe des untersuchenden Arztes,
PSA $-\{ >3\text{ng/ml} \} \rightarrow$ DRU $-\{ \oplus \} \rightarrow$ TRUS $-\{ \oplus \} \rightarrow$ Biopsie

1-3 Perspektive

Keine Angaben.

1-4 Zeithorizont

Ermittelt wurden Effekte und Kosten für den Zeitraum der Studie (Prävalenz-Screening):
12 Monate.

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Originalstudie.

Datenquelle für Gesundheitseffekt: Originalstudie (Querschnittstudie).

Datenquelle für Ressourcenverbrauch und Mengengerüst: Originalstudie (Querschnittstudie).

Datenquelle(n) für Preise: keine Angaben.

2-2 Datierung der zugrundeliegenden Daten

Keine Angaben bezüglich des Zeitpunktes der Datenerhebung zum Gesundheitseffekt und Ressourcenverbrauch; der Studienzeitraum betrug 12 Monate, die Studie wurde im April 1995 eingereicht.

Keine Angaben hinsichtlich des Bezugsjahres für die Preisangaben.

2-3 Studienpopulation / Zielpopulation

Männer im Alter von 55 bis 74 Jahren, alle Einwohner desselben norditalienischen Dorfes.

2-4 Setting

Keine Angaben hinsichtlich der tangierten Versorgungsstruktur(en).

Der klinische wie auch der ökonomische Teil der Studie wurden in Italien durchgeführt.

2-5 Spezifikation der Technologie

PSA-Assay: Hybritech Tandem-R (monoklonal).

DRU: Palpation durch einen erfahrenen Urologen.

TRUS: 6.5MHz-Schallkopf, biplanar.

Biopsie: transrektale Stanzbiopsie (18G-Nadeln, Biopsy-Pistole) unter Ultraschallkontrolle und/oder digitaler Führung, im Durchschnitt vier Stanzzyylinder pro Lappen; Antibiotikaphylaxe.

3 Gesundheitseffekte

Screening-/Diagnosestudie

3-1 Untersuchte Zielgrößen

Primäre Zielgröße war die nach Biopsieabsicherung identifizierte Anzahl Prostatakarzinome (Prävalenz) im Probandenkollektiv.

3-2 Ein-/Ausschlusskriterien

Als Einschlusskriterien wurden Geschlecht, Alter und Wohnort genannt.

3-3 Rekrutierungsmodus

Keine näheren Angaben zum Rekrutierungsmodus; die Probanden waren Teilnehmer einer bereits laufenden prospektiven Kohortenstudie zu kardiovaskulären Parametern.

3-4 Teilnahmerate

91% (472/518).

3-5 Reproduzierbarkeit der Testergebnisse

Die eingesetzten Diagnosemethoden wurden weitgehend spezifiziert. Die Entscheidungsgrenze bei der PSA-Bestimmung ($>3\text{ng/ml}$) wurde angegeben. Entscheidungskriterien bezüglich der Interpretation von DRU- und TRUS-Befunden wurden nicht explizit genannt. Die DRU und die TRUS wurden jeweils vom gleichen Untersucher durchgeführt; Vorkehrungen zur Kontrolle der Variabilität von Blut- und Gewebeprobenanalysen wurden nicht spezifiziert.

3-6 Auswertung der Testergebnisse

Die Biopsie als Goldstandard wurde lediglich bei Probanden mit positivem TRUS-Befund (13/472) durchgeführt. Die Bestimmung der Testergebnisse und des Krankheitsstatus erfolgten nicht in gegenseitiger Verblindung.

3-7 Drop-Outs

Berichtetes Ausscheiden von zwei Probanden mit $\text{PSA} > 3\text{ng/ml}$ bei der DRU.

3-8 Ergebnisse der Screening-/Diagnosestudie

Deskriptive Präsentation der Ergebnisse entlang des Interventionspfades ohne Anwendung statistischer Testverfahren oder Angabe von Kenngrößen des Tests.

18.3% (86/472) der Probanden hatten einen PSA-Wert $>3\text{ng/ml}$; von den 84 Probanden mit DRU wiesen 26 (5.5%) einen für BPH sprechenden Befund und 31 (6.6%) einen für Prostatakarzinom sprechenden Befund auf. Die TRUS war für 13 Probanden (2.7%) positiv, die Biopsie bestätigte diesen Befund in 11 Fällen. Die Prävalenz von durch Screening entdecktes Prostatakarzinom im untersuchten Probandenkollektiv lag somit bei 2.3%. Tumor-Grading: Gleason-Score 2 in 2, Gleason-Score 3 in 6 (5 im Stadium A, 1 im Stadium B), Gleason-Score 4 in 1 und Gleason-Score 5 in 2 (1 im Stadium A, 1 im

Stadium B) Fall/Fällen. Altersstratifizierung der Karzinomträger: 55-59 Jahre 0, 60-64 Jahre 5, 65-69 Jahre 1, 70-74 Jahre 5 Fall/Fälle. Das durchschnittliche Alter der Probanden lag bei 63.7 Jahren (Median 61.2 Jahre).

3-9 Effektmaße für die ökonomische Analyse

Anzahl identifizierter Prostatakarzinome.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

– Screening: PSA-Assay, DRU, TRUS und Biopsie

Die Kosten für die Rekrutierung wurden explizit nicht einbezogen.

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Die sich aus dem Interventionspfad ergebenden Mengen wurden spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Administrative Preise: Standard-Krankenhausgebühren für die ambulante Versorgung. Die Quelle(n) und das Bezugsjahr der Preise wurden nicht spezifiziert. Die Preise pro Mengeneinheit wurden aufgeführt.

4-4 Währung

Italienische Lire (Lit.).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die beschriebene Screening-Strategie führte zur Entdeckung von 11 Prostatakarzinomen bei 472 Teilnehmern (2.3%).

6-2 Ermittelte Kosten

Die Gesamtkosten beliefen sich auf Lit. 10,250,000; ein durchschnittlicher Betrag von Lit. 22,000 pro Screening-Teilnehmer wurde ermittelt.

6-3 Synthese von Kosten und Effekten

Die Kosten pro entdecktem Prostatakarzinom beliefen sich auf Lit. 932,000.

7 Behandlung von Unsicherheiten

Es wurde eine Zusatzanalyse mit einer Entscheidungsgrenze von $>4\text{ng/ml}$ bei der PSA-Bestimmung durchgeführt: 66 DRU (14.0%), 28 TRUS (5.9%) und 12 Biopsien (2.5%) hätten hierbei zu der gleichen Anzahl identifizierter Prostatakarzinome (11: Prävalenz 2,3%) geführt. Die Gesamtkosten hätten Lit. 9,450,000 betragen mit durchschnittlichen Kosten von Lit. 20,000 pro Screening-Teilnehmer bzw. Lit. 860,000 pro entdecktem Prostatakarzinom.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren weisen hinsichtlich der hohen Teilnahmerate auf einen möglichen selection bias bei den bereits an einer anderen epidemiologischen Studie teilnehmenden Probanden hin. In bezug auf die Nichtberücksichtigung von Rekrutierungskosten wird auf die Variabilität der Kosten in Abhängigkeit von Rekrutierungsmodus und Teilnehmerzahl hingewiesen; in bezug auf die Nichtberücksichtigung von Behandlungskosten wird darauf hingewiesen, dass das Aufsuchen eines Urologen ohne Screening-Angebot oftmals erst bei Erreichen eines metastasierten Stadiums geschieht.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Die Autoren sind der Meinung, dass eine Screening-Strategie mit der PSA-Bestimmung als Vorfelddiagnostik in einer größeren Population einsetzbar ist. Angesichts der Altersverteilung der Prostatakarzinomträger halten es die Autoren für angebracht, in eine nächste Screening-Runde eine jüngere Population (45-55 Jahre) mit einzubeziehen.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass die PSA-Bestimmung als Vorfelddiagnostik beim Screening des Prostatakarzinoms eingesetzt werden kann. Die geringe Kostendifferenz zwischen der Screening-Strategie mit einem PSA-Schwellenwert von $>3\text{ng/ml}$ und derjenigen mit einem PSA-Schwellenwert von $>4\text{ng/ml}$ rechtfertigt in ihren Augen die Annahme einer Entscheidungsgrenze von $>3\text{ng/ml}$. Die Vergleichbarkeit der gefunden Krankheitsinzidenz mit den Literaturangaben, die Kosten-Effektivität sowie die mögliche Lebensqualitätsverbesserung der Teilnehmer (bei 26 Probanden konnte eine BPH diagnostiziert und behandelt werden) rechtfertigen die Screening-Maßnahmen.

9 Kommentar

Studie zu Biopsiekriterien beim Prävalenz-Screening ohne explizite Gegenüberstellung von Technologiealternativen. Im Vordergrund steht eher die Durchführbarkeit der diagnostischen Strategie und weniger der langfristige Nutzen der Früherkennung, was sich auch in der Wahl eines intermediären Outcome-Parameters als interessierende klinische Zielgröße ausdrückt. Ein selection bias ist offenbar; die Möglichkeit eines diagnostic suspicion bias sowie eines expectation bias bei der Durchführung der untersucherabhängigen Tests sowie die lediglich auf eine selektierte Populationsuntergruppe beschränkte Anwendung des Goldstandards in Verbindung mit einem work-up bias reduzieren die interne Validität, Reproduzierbarkeit und Generalisierbarkeit der klinischen Untersuchungsergebnisse in erheblichem Maße. Die Kostenberechnung ist nachvollziehbar, doch mangelt es der Dokumentation insgesamt an Transparenz.

Dorr VJ, Williamson SK, Stephens RL:

An Evaluation of Prostate-Specific Antigen as a Screening Test for Prostate Cancer.

Archives of Internal Medicine, 1993; 153: 2529-2537.

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels Bestimmung von Serum-PSA. Kein expliziter Vergleich von Technologiealternativen; implizite Gegenüberstellung der Alternativen "Screening" und "Kein Screening".

1-2 Fragestellung

Beurteilung der Nützlichkeit der PSA-Bestimmung als Früherkennungstest bei Prostatakarzinom und Berechnung der im Rahmen eines Massen-Screening-Programmes zu erwartenden Kosten pro Jahr durch Reanalyse der klinischen Entscheidungsanalyse von Optenberg & Thompson (1990)¹², welche die Ergebnisse neuerer Studien berücksichtigt.

Screening-Strategie: PSA $\geq 4\text{ng/ml}$ → Biopsie

1-3 Perspektive

Keine expliziten Angaben.

1-4 Zeithorizont

Keine expliziten Angaben. Ermittelt wurden die Effekte und Kosten im ersten Jahr des Screening-Programmes (Prävalenz-Screening) sowie die Effekte bei hypothetischer Senkung der Mortalität um 30%.

1-5 Art der ökonomischen Evaluation

Kosten-Konsequenzen-Beschreibung; Kosten-Effektivitäts-Analyse.

¹² Optenberg SA & Thompson IM: *Economics of Screening for Carcinoma of the Prostate*. Urol Clin North Am 1990; 17(4): 719-737.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Narrativer Review mit zusätzlicher Kostenaufstellung.

Datenquellen für klinische Parameter: klinische Entscheidungsanalyse mit Literaturdaten.

Datenquelle für Ressourcenverbrauch und Mengengerüst: Modellierung (Entscheidungsanalyse).

Datenquelle für Preise: keine Angaben (implizit Publikation von Optenberg & Thompson).

2-2 Datierung der zugrundeliegenden Daten

Der Schätzwert für die Größe der männlichen Bevölkerung der Vereinigten Staaten in der untersuchten Altersgruppe basiert auf im Jahre 1990 veröffentlichten Zensusdaten. Zeitpunkt und Zeitraum der Erhebung von Daten zu klinischen Parametern wurden nicht explizit aufgeführt (die klinischen Parameter leiten sich aus Daten ab, welche bis 1992 veröffentlicht wurden).

Keine Angaben hinsichtlich des Bezugsjahres für die Preisangaben.

2-3 Studienpopulation / Zielpopulation

Männer der amerikanischen Gesamtbevölkerung im Alter von 50 bis 75 Jahren ohne Komorbidität.

2-4 Setting

Tangierte Versorgungsstrukturen: keine Angaben.

Keine explizite Lokalisation der Daten zu den klinischen Parametern. Der ökonomische Teil der Studie wurde in Kansas, USA durchgeführt.

2-5 Spezifikation der Technologie

Keine Präzisierungen bezüglich des PSA-Assays oder der Biopsie.

3 Gesundheitseffekte

Synthese bereits veröffentlichter Studien

3-1 In der Synthese untersuchte klinische Parameter

Primäre Zielgrößen waren die Anzahl Männer im Alter von 50 bis 75 Jahren in den Vereinigten Staaten, der Anteil Männer in der männlichen Bevölkerung mit Komorbidität als Ausschlusskriterium, die Testpositiv- und Richtigpositiv-Raten der PSA-Bestimmung, die Stadienverteilung der entdeckten Karzinome, die Wahrscheinlichkeiten von behandlungsbedingten Komplikationen und sekundären Therapien sowie die für 1993 erwartete Anzahl krankheitsspezifischer Todesfälle.

3-2 Annahmen

Die Analyse basierte auf folgenden expliziten Annahmen: Als Staging-Methoden werden eine TRUS und ein Knochenszintigramm eingesetzt. Patienten mit einem klinischen Stadium A1-Karzinom werden nicht (sofort) behandelt (watchful waiting). Patienten mit einem Stadium A2- oder B-Karzinom unterziehen sich einer radikalen Prostatektomie, Patienten mit klinischem Stadium C-Karzinom einer Strahlentherapie (66-70Gy), Patienten mit klinischem Stadium D-Karzinom einer Orchiektomie. Als Komplikationen der radikalen Prostatektomie wurden berücksichtigt (in Klammern therapeutische Maßnahmen): Impotenz (penile Prothese); rektale Verletzung (Kolostomie); Inkontinenz (künstlicher Sphinkter); Harnröhrenstriktur (Korrektur); postoperativer Tod. Als Komplikationen der Strahlentherapie wurden berücksichtigt (in Klammern therapeutische Maßnahmen): Impotenz (penile Prothese); rektale Verletzung (Kolostomie ohne spätere Anastomose); Inkontinenz (künstlicher Sphinkter); Harnröhrenstriktur (Korrektur); behandlungsbedingter Tod. Für die bilaterale Orchiektomie wurden keine Komplikationen angenommen.

Implizite Annahme war, dass bei 50% der Patienten mit einem Stadium A-Karzinom jeweils ein Stadium A1-bzw. ein Stadium A2-Tumor vorliegt. Außerdem wird implizit von einer 100%igen Compliance mit der Früherkennung ausgegangen.

3-3 Berücksichtigung von Primärstudien: Studiendesigns und Ein-/Ausschlusskriterien

Die Anzahl Männer im Alter von 50 bis 75 Jahren in den Vereinigten Staaten basierte auf Zensusdaten, die für 1993 erwartete Anzahl krankheitsspezifischer Todesfälle auf einer Krebsstatistik. Auf welchen Daten der Schätzwert für den Anteil der Männer mit

Komorbidität basiert, ist unklar. Das Design der Studien, die zur Schätzung der interessierenden klinischen Parameter herangezogen wurden, wurde nicht spezifiziert. Bis auf drei Ausnahmen wurden die Wahrscheinlichkeiten für die berücksichtigten behandlungsbedingten Komplikationen und sekundären Therapien implizit und ohne Präzisierung des zugrundeliegenden Studiendesigns der Studie von Optenberg & Thompson entnommen.

3-4 Quellen und Suchstrategie bei der Literaturrecherche

MEDLINE wurde als genutzte Quelle spezifiziert. Keine Angaben zur Suchstrategie.

3-5 Validitäts- bzw. Qualitätskriterien bei der Bewertung der Primärstudien

Keine Angaben.

3-6 Methoden der Bewertung von Relevanz und Validität bzw. Qualität der Primärstudien

Keine Angaben.

3-7 Methoden der Extraktion von Daten aus den Primärstudien

Keine Angaben.

3-8 Anzahl berücksichtigter Primärstudien

Die Testpositiv- und Richtigpositiv-Raten der PSA-Bestimmung wurden aus neun Studien abgeleitet. Für die Stadienverteilung der mittels der PSA-Bestimmung entdeckten Karzinome wurde auf fünf Studien zurückgegriffen. Die Wahrscheinlichkeiten von behandlungsbedingten Komplikationen und sekundären Therapien wurden implizit größtenteils der Studie von Optenberg & Thompson entnommen.

3-9 Methode der Synthese der gesundheitsbezogenen Parameter

Für die Testpositiv- und Richtigpositiv-Raten wurden, soweit ersichtlich, die ungewichteten Durchschnittswerte der neun verfügbaren Studien ermittelt. Zur Ermittlung der Stadienverteilung der diagnostizierten Karzinome wurden implizit die in den fünf berücksichtigten Publikationen mittels der PSA-Bestimmung (mit unterschiedlichen Entscheidungsgrenzen) entdeckten Stadium A- bis D-Karzinome aufaddiert und der jeweilige Anteil an der Gesamtsumme bestimmt.

3-10 Untersuchung von Heterogenität der gesundheitsbezogenen Parameter

Nicht untersucht.

3-11 Ergebnisse der Synthese

(Werte für die Basisanalyse; in Klammern Werte im plausiblen Bereich)

Die Anzahl Männer im Alter von über 50 Jahren in den Vereinigten Staaten betrug auf Basis von im Jahre 1989 veröffentlichten Daten 28,038,000. Nach Ausschluss des auf 10% angesetzten Anteils an Männern mit Komorbidität wurde von einer Screening-Population mit 21,266,100 Männern im Alter von 50 bis 75 Jahren ausgegangen. Die Testpositiv-Rate bei der PSA-Bestimmung wurde auf 0.161 angesetzt, die Richtigpositiv-Rate auf 0.306. 3% der diagnostizierten Karzinome wurden als Stadium A-, 35% als Stadium B-, 32% als Stadium C- und 30% als Stadium D-Tumoren angesetzt. Die für Komplikationen und komplikationsbedingte sekundäre Therapien bei radikaler Prostatektomie angesetzten Raten wurden größtenteils der Publikation von Optenberg & Thompson entnommen und waren folgende: Impotenz 0.30 bzw. 0.075; rektale Verletzung 0.03 bzw. 0.01; Inkontinenz 0.06 bzw. 0.03; Harnröhrenstriktur 0.10 bzw. 0.10; postoperativer Tod 0.01. Die Raten für Komplikationen und komplikationsbedingte sekundäre Therapien bei Strahlentherapie waren: Impotenz 0.25 bzw. 0.0625; rektale Verletzung 0.01 bzw. 0.01; Inkontinenz 0.03 bzw. 0.015; Harnröhrenstriktur 0.08 bzw. 0.08; behandlungsbedingter Tod 0.005. Die für 1993 erwartete Anzahl krankheitsspezifischer Todesfälle wurde auf 34,000 angesetzt.

3-12 Effektmaße für die ökonomische Analyse

Anzahl identifizierter Prostatakarzinome, Anzahl behandlungsbedingter Todesfälle, Anzahl verhinderter Todesfälle infolge einer (hypothetischen) Senkung der krankheitsspezifischen Mortalität um 30%.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

- Screening/Diagnostik: Screening-Untersuchung, Biopsie, Staging
- Behandlung: primäre und sekundäre Behandlung

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Das Mengengerüst wurde spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Keine Angaben. (Bis auf zwei Ausnahmen wurde für die monetäre Bewertung auf die in der Publikation von Optenberg & Thompson aufgeführten administrativen Preise für das Jahr 1988 zurückgegriffen). Die Preise pro Mengeneinheit wurden aufgeführt.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die Anzahl der mittels des Früherkennungsprogrammes identifizierten Karzinome wurde auf 1,047,695 berechnet. Die behandlungsbedingten Todesfälle wurden auf 3,842 bei radikaler Prostatektomie und auf 1,676 bei Strahlentherapie ermittelt. Die Anzahl verhinderter Todesfälle infolge einer 30%igen Senkung der krankheitsspezifischen Mortalität wurde auf 10,200 geschätzt (ohne Berücksichtigung der behandlungsbedingten Todesfälle).

6-2 Ermittelte Kosten

Die Gesamtkosten der untersuchten Screening-Strategie betragen US-\$ 12.7 Mrd., wobei sich die Kosten für die PSA-Bestimmung auf US-\$ 0.9 Mrd., diejenigen für die Biopsie auf US-\$ 1.1 Mrd. und diejenigen für die primäre Behandlung auf US-\$ 9.3 Mrd. belaufen.

6-3 Synthese von Kosten und Effekten

Die Kosten pro diagnostiziertem Prostatakarzinom wurden auf US-\$ 1,924 ermittelt. Pro potentiell verhindertem Todesfall bei 30%iger Senkung der krankheitsspezifischen Mortalität (ohne Berücksichtigung der behandlungsbedingten Todesfälle) beliefen sich

die Kosten bei Berücksichtigung der PSA-Bestimmung und der Biopsie auf US-\$ 197,647.

7 Behandlung von Unsicherheiten

Keine Ausführungen.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren bemerken, dass die verschiedenen Publikationen u.a. wegen unterschiedlicher Rekrutierungsmodi, unterschiedlicher PSA-Assays, unterschiedlicher Entscheidungsgrenzen und unterschiedlicher Indikationen für die Durchführung der Biopsie nur schwer vergleichbar sind. Zusätzlich wird die Ableitung von Indizes bezüglich der Testgüte dadurch eingeschränkt, dass zumeist nicht die Gesamtstudienpopulation einer Biopsie als Referenzstandard unterzogen wurde. Weiterhin erinnern die Autoren an die Möglichkeit von lead time bias und length bias als Fehler, die zu einer Überschätzung des Screeningeffektes führen.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Keine Ausführungen.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass der PSA-Test als reproduzierbares, objektives, untersucherunabhängiges Entscheidungsinstrumentarium bei der Vorfelddiagnostik gesehen werden kann, als alleiniger Screening-Test jedoch nicht als effektiv betrachtet werden kann. Die langfristige Wirksamkeit der Behandlung muss ihrer Ansicht nach noch bewiesen werden. Bevor die Früherkennung von Prostatakarzinom mittels der PSA-Bestimmung empfohlen werden kann, bedarf es nach Meinung der Autoren weiterer Studien, die insbesondere die Auswirkungen des Screenings auf die Mortalität untersuchen. In der Zwischenzeit sollte vor der Durchführung der PSA-Bestimmung eine eingehende Aufklärung der einzelnen Patienten über Nutzen, Risiken und Kosten der Früherkennung erfolgen.

9 Kommentar

Qualitative Informationssynthese zum PSA-Screening mit Kostenrechnung für das Prävalenz-Screening in expliziter Anlehnung an die Entscheidungsanalyse von Optenberg & Thompson; dabei implizite Gegenüberstellung von "Screening" vs. "Kein Screening". Hinsichtlich der Ermittlung der Effektivität weist der Review bei einem Großteil der von einer systematischen Literaturrecherche und -analyse zu erfüllenden Qualitäts- und Validitätskriterien Defizite auf. Der Ermittlung der Kosten ist transparent und nachvollziehbar; die Dokumentation der Quellen für Mengen- und Preisgerüst ist jedoch lückenhaft und ohne Rückgriff auf die Publikation von Optenberg & Thompson wertlos. Als problematisch erscheint, dass die Autoren zwar neuere wissenschaftliche Erkenntnisse bezüglich der PSA-Bestimmung berücksichtigen und auch deren Kosten zeitlich anpassen, eine zeitliche Veränderung der Biopsie- und Behandlungseffekte und -kosten - beispielsweise infolge des technischen Fortschritts und/oder aufgrund von Lernkurveneffekten - jedoch nicht in Erwägung ziehen. Die Berechnung eines Kosten-Effektivitäts-Quotienten für die bei hypothetischer Senkung der krankheitsspezifischen Mortalität um 30% geretteten Leben ohne Berücksichtigung von Kosten für Behandlung, behandlungsbedingte Morbidität und Mortalität ist inkonsequent. Die interne Validität und Generalisierbarkeit der Ergebnisse erscheinen aufgrund der methodischen Defizite eingeschränkt.

10 Technischer Report

Bei der Evaluation dieser Studie wurde zusätzlich auf die Publikation von Optenberg & Thompson (1990) zurückgegriffen:

Optenberg SA & Thompson IA:

Economics of Screening for Carcinoma of the Prostate.

Urologic Clinics of North America, 1990; 17(4): 719-737.

Gottlieb RH, Mooney C, Mushlin AI, Rubens DJ, Fultz PJ:
The Prostate. Decreasing Cost-Effectiveness of Biopsy with Advancing Age.
Investigative Radiology, 1996; 31(2): 84-90.

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mit Bestimmung der PSA-Dichte. Explizite Gegenüberstellung der Alternativen "Biopsie" in Abhängigkeit von dem Cutpoint der PSA-Dichte und "Keine Biopsie".

1-2 Fragestellung

Ermittlung der altersabhängigen Kosten-Effektivität der Prostatabiopsie im Vergleich zur Alternative "Keine Biopsie", wobei der altersspezifische Exzess-PSA-Wert als Entscheidungskriterium für eine Biopsie herangezogen wird.

Interventionsstrategie: PSA ≥ 4 ng/ml \rightarrow TRUS } \rightarrow Exzess-PSA $>$ Entscheidungsgrenze \rightarrow
Biopsie, Wiederholung der PSA-Bestimmung nach einem Jahr

Exzess-PSA-Wert = berechneter volumenadjustierter PSA-Wert (ng/ml) - gemessener PSA-Wert (ng/ml)

Volumenadjustierter PSA-Wert = $0.12 \text{ ng/ml/ccm} \times \text{Prostata volumen (ccm)}$.

Es bleibt unklar, aufgrund welcher Entscheidungskriterien (Veränderung bei PSA-D oder PSA-V) und Entscheidungsgrenzen eine Biopsie im zweiten Screening-Jahr durchgeführt wird.

1-3 Perspektive

Gesellschaftliche Perspektive.

1-4 Zeithorizont

Ermittelt wurden die Effekte hinsichtlich der qualitätsadjustierten Lebenserwartung sowie die Kosten für einen 2-Jahres-Zeitraum.

1-5 Art der ökonomischen Evaluation

Kosten-Nutzwert-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Synthese.

Datenquellen für klinische Parameter: klinische Entscheidungsanalyse mit teils eigenen Screeningdaten, vorwiegend Literaturdaten und Modellierung (DEALE).

Datenquellen für Präferenzen: Lebensqualität bei Impotenz, Inkontinenz und gastrointestinalen Beschwerden infolge von Darmverletzungen: Literatur. Alle Komplikationen: Schätzungen der Autoren mit Sensitivitätsanalysen. Weitere Spezifikationen fehlen.

Datenquelle für Ressourcenverbrauch und Mengengerüst: Modellierung (Entscheidungsanalyse).

Datenquelle für Preise: Institutionseigene Daten.

2-2 Datierung der zugrundeliegenden Daten

Zeitpunkt und Zeitraum der Erhebung von Daten zu klinischen Parametern wurden nicht explizit aufgeführt (die klinischen Parameter leiten sich aus Studien ab, welche bis 1994 veröffentlicht wurden).

Keine Angaben hinsichtlich des Bezugsjahres für die Preisangaben.

2-3 Studienpopulation / Zielpopulation

Keine expliziten Angaben. Ermittelt wurden die Kosten-Effektivität der Biopsiestrategien für Männer ab 50 Jahren.

2-4 Setting

Keine Angaben bezüglich der tangierten Versorgungsstruktur(en).

Größtenteils keine Lokalisation der Daten zu den klinischen Parametern. Der ökonomische Teil der Studie wurde in New York, USA durchgeführt.

2-5 Spezifikation der Technologie

Bei der für die Schätzung der Wahrscheinlichkeit einer positiven Biopsie in Abhängigkeit vom Exzess-PSA-Wert herangezogenen Fallserie wurden ein monoklonaler PSA-Assay mit einer Entscheidungsgrenze bei 4ng/ml eingesetzt und transrektale randomisierte (Sextantenbiopsie) wie auch ultraschallkontrollierte Biopsien durchgeführt.

3 Gesundheitseffekte

Synthese bereits veröffentlichter Studien

3-1 In der Synthese untersuchte klinische Parameter

Zielgrößen waren die Wahrscheinlichkeit einer positiven Biopsie in Abhängigkeit vom Exzess-PSA-Wert, die altersspezifische Wahrscheinlichkeit eines klinisch signifikanten Prostatakarzinoms bei positiver Biopsie, die Wahrscheinlichkeit der Entdeckung eines klinisch signifikanten Prostatakarzinoms nach anormaler Erhöhung des PSA-Spiegels (PSA-V), die Wahrscheinlichkeit von extrakapsulärem Karzinom, die Sensitivität und Spezifität der Staging-Tests (CT, MRI), die Komplikationsraten bei Behandlung (Tod, Impotenz, Inkontinenz und Darmverletzungen), die krankheitsstadiumspezifische Mortalität sowie der Anteil Patienten mit klinisch signifikantem Prostatakarzinom bei der Alternative "Keine Biopsie".

3-2 Annahmen

Im Analysemodell wurde der Einfluss anderer Parameter (z.B. Familienanamnese, DRU, TRUS) auf die Entscheidungsfindung für oder wider eine Biopsie explizit unberücksichtigt gelassen. Gleichzeitig wurde die Möglichkeit der Entdeckung und Behandlung von Prostatakarzinom im klinisch nicht signifikanten Stadium vorausgesetzt. Für Stadium A- oder B-Tumoren wurde eine Operation (mit Staging) oder Strahlentherapie (ohne Staging) angesetzt. Bei Stadium C-Tumoren wurde eine Strahlentherapie (ohne Staging), bei Stadium D-Tumoren eine Strahlentherapie und endokrine Therapie (ohne Staging) angenommen. Es wurde explizit davon ausgegangen, dass ein Staging ausschließlich der Triage von Patienten mit lokalisiertem Ca (Stadium A oder B) dienlich ist. Als Staging-Methoden wurden ein Knochenszintigramm sowie entweder eine Computertomographie (CT) oder eine Kernspinresonanztomographie (MRI) angesetzt. Der Gleason-Score als Prognosefaktor wurde explizit nicht berücksichtigt.

3-3 Berücksichtigung von Primärstudien: Studiendesigns und Ein-/Ausschlusskriterien

Die Wahrscheinlichkeit einer positiven Biopsie in Abhängigkeit vom Exzess-PSA-Wert wurde aus einer Fallserie abgeleitet. Zur Schätzung der krankheitsspezifischen Mortalität bei Nichtbehandlung von klinisch signifikantem Prostatakarzinom wurde auf Daten aus der prospektiven Kohortenstudie zum natürlichen Verlauf von Prostatakarzinom von

Johansson et al. (1992)¹³ zurückgegriffen. Das Design der Studien, die zur Schätzung der anderen interessierenden klinischen Parameter herangezogen wurden, wurde nicht spezifiziert. Keine Angaben zu Ein- bzw. Ausschlusskriterien bei der Berücksichtigung der Studien.

3-4 Quellen und Suchstrategie bei der Literaturrecherche

Keine Angaben.

3-5 Validitäts- bzw. Qualitätskriterien bei der Bewertung der Primärstudien

Keine Angaben zu Suchstrategien. Für die klinischen Parameter wurden vereinzelt Quellenangaben gemacht.

3-6 Methoden der Bewertung von Relevanz und Validität bzw. Qualität der Primärstudien

Keine Angaben.

3-7 Methoden der Extraktion von Daten aus den Primärstudien

Keine Angaben.

3-8 Anzahl berücksichtigter Primärstudien

Jeweils eine Publikation wurde herangezogen, um Schätzungen für die interessierenden klinischen Parameter abzuleiten. Hierbei wurden insgesamt sechs Publikationen berücksichtigt.

3-9 Methode der Synthese der gesundheitsbezogenen Parameter

Diese wurde im allgemeinen nicht spezifiziert. Die Autoren bezogen sich auf einzelne Literaturstellen. Die altersspezifische Wahrscheinlichkeit eines klinisch signifikanten Prostatakarzinoms bei positiver Biopsie basiert für Männer im Alter von <70 Jahren auf Annahmen. Für die krankheitsspezifische Mortalität bei Nichtbehandlung von klinisch signifikantem Prostatakarzinom (alle Stadien) wurde die Rate der Untergruppe mit der schlechtesten Prognose aus der Studie von Johansson et al. gewählt.

3-10 Untersuchung von Heterogenität der gesundheitsbezogenen Parameter

Nicht untersucht.

¹³ Johansson JE, Adami HO, Andersson SO, Bergstrom R, Holmberg L, Krusemo UB: *High 10-Year Survival Rate in Patients with Early, Untreated Prostatic Cancer*. JAMA 1992; 267:2191-2196.

3-11 Ergebnisse der Synthese

(Werte für die Basisanalyse; in Klammern Werte im plausiblen Bereich)

Die Wahrscheinlichkeit einer positiven Biopsie stieg mit zunehmenden Exzess-PSA-Werten von 0.2 bei 0ng/ml auf 0.63 bei 10ng/ml an und blieb bei Exzess-PSA-Werten zwischen 10ng/ml und 20ng/ml annähernd konstant bei 0.63. Für die altersspezifische Wahrscheinlichkeit eines klinisch signifikanten Prostatakarzinoms bei positiver Biopsie wurde bei Männern im Alter von 50 Jahren 1.0 angesetzt; bis zum Alter von 70 Jahren wurde eine lineare Abnahme dieser Wahrscheinlichkeit auf 0.2 angenommen. Für die Wahrscheinlichkeit der Entdeckung eines klinisch signifikanten Prostatakarzinoms nach anormaler Erhöhung des PSA-Spiegels innerhalb eines Jahres (PSA-V) wurde 1.0 (0.5-1.0) angesetzt. Die Wahrscheinlichkeit von extrakapsulärem Karzinom (Stadium C- oder D-Tumoren) wurde auf 0.05 in der ersten Screening-Runde, auf 0.45 in der zweiten Screening-Runde geschätzt. Die Sensitivität und Spezifität der Staging-Tests (CT, MRI) wurde bei 0.72 bzw. 0.90 angesetzt. In der Basisanalyse waren die Komplikationsraten bei Operation folgende: perioperativer Tod 0.01, Impotenz 0.5 und Inkontinenz 0.005. Die Komplikationsraten bei Strahlentherapie waren 0.05 für Darmverletzungen und 0.22 für Impotenz. Für die endokrine Therapie wurden keine Komplikationen angenommen. Die krankheitsspezifische Mortalität bei Nichtbehandlung von klinisch signifikantem Prostatakarzinom (alle Stadien) wurde bei 0.09 angesetzt. Die krankheitsspezifische Mortalität in den Stadien A/B, C und D bei Strahlentherapie war 0.015, 0.09 bzw. 0.29, in den Stadien A und B bei radikaler Prostatektomie 0.05 und im Stadium D bei endokriner Therapie 0.29.

3-12 Effektmaße für die ökonomische Analyse

Qualitätsadjustierte Überlebenszeit (QALE bzw. QALYs). Die krankheitsstadiumspezifische, altersadjustierte Überlebenszeit wurde mittels der DEALE-Methode ermittelt. Gewichtungsfaktoren zur Adjustierung der Überlebenszeit nach der Lebensqualität bei Impotenz, Inkontinenz und gastrointestinalen Beschwerden infolge von Darmverletzungen wurden der Literatur entnommen; weitere Spezifikationen fehlen. Der Gewichtungsfaktor wurde in der Basisanalyse für alle Komplikationsarten auf 0.8 (plausibler Wertebereich 0.5-0.9) angesetzt.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

- Screening/Diagnostik: PSA-Assay, ultraschallgesteuerte Biopsie einschließlich Histopathologie, Staging
- Behandlung: radikale Prostatektomie, Strahlentherapie, endokrine Therapie

Intangible Kosten wurden bei der Ermittlung der Lebensqualität auf der Effektseite erfasst.

Kosten bzw. inkrementelle Kosten-Nutzwert-Quotienten wurden lediglich berechnet, wenn sich die Alternative "Biopsie" im Vergleich zur Alternative "Keine Biopsie" als effektiver erwies.

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Die sich aus der Entscheidungsanalyse ergebenden Mengen wurden nicht spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Die Art der Preise ist unklar, es wurden institutionsspezifische Preise verwendet. Das Bezugsjahr der Preise wurde nicht spezifiziert. Die Preise pro Mengeneinheit wurden aufgeführt.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die qualitätsadjustierte Überlebenszeit nahm in Abhängigkeit vom Alter ab. Unter den Annahmen der Basisanalyse ergab sich für Männer im Alter ab 70 Jahren beim Vergleich der Alternative "Biopsie" mit der Alternative "Keine Biopsie" ein Verlust an QALYs bei allen Exzess-PSA-Werten zwischen 0ng/ml und 20ng/ml. Für Männer im Alter von 50 und 60 Jahren war die Biopsie bei allen Exzess-PSA-Werten ≥ 0 ng/ml mit einem Gewinn

an QALYs verbunden. Aufgeführt wurden lediglich die altersspezifischen Werte für einen Exzess-PSA-Wert von 2ng/ml.

6-2 Ermittelte Kosten

Gesamtkosten, Durchschnittskosten oder inkrementelle Kosten wurden nicht spezifiziert.

6-3 Synthese von Kosten und Effekten

Im Vergleich zur Alternative "Keine Biopsie" betrug der inkrementelle Kosten-Nutzwert-Quotient unter den Annahmen der Basisanalyse für Männer im Alter von 60 Jahren mit Biopsie zwischen US-\$ 13,558 pro QALY und US-\$ 8,974 pro QALY bei einem Exzess-PSA-Wert von 0ng/ml bzw. von 20ng/ml. Für Männer im Alter von 50 Jahren war die Biopsie bei allen Exzess-PSA-Werten von ≥ 0 ng/ml kostengünstiger und effektiver (keine Angaben von monetären Werten; für Männer im Alter von 70 Jahren wurden die Kosten nicht berechnet, da die Biopsie bei allen Exzess-PSA-Werten von ≥ 0 ng/ml mit einem Verlust an QALYs verbunden war).

7 Behandlung von Unsicherheiten

Sensitivitätsanalysen (ohne nähere Kennzeichnung) sollen für alle Hauptvariablen durchgeführt worden sein; Werte bzw. Wertbereiche wurden nur teilweise genannt. Die Erhöhung der Wahrscheinlichkeit eines klinisch signifikanten Prostatakarzinoms bei positiver Biopsie von 0.2 auf 0.4 für Männer im Alter von 70 Jahren ging mit einem kleinen Gewinn an QALYs bei allen Exzess-PSA-Werten einher (welcher jedoch geringer ausfiel als derjenige für Männer im Alter von 50 und 60 Jahren unter den Annahmen der Basisanalyse): Der inkrementelle Kosten-Nutzwert-Quotient der Alternative "Biopsie" im Vergleich zur Alternative "Keine Biopsie" betrug dabei zwischen US-\$ 275,881 pro QALY bei einem Exzess-PSA-Wert von 0ng/ml und US-\$ 68,362 pro QALY bei einem Exzess-PSA-Wert von 20ng/ml.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren weisen darauf hin, dass in der Basisanalyse außer der Annahme, dass die Wahrscheinlichkeit eines klinisch signifikanten Prostatakarzinoms bei positiver Biopsie bei Männern im Alter ab 70 Jahren bei 0.2 liegt, alle anderen Annahmen die Alternative "Biopsie" begünstigen. In der mangelnden Verfügbarkeit von Daten für die Wahrscheinlichkeit eines klinisch signifikanten Prostatakarzinoms bei positiver Biopsie

sehen die Autoren Einschränkungen ihrer Studie. Als weitere Schwächen werden genannt: die Nichtberücksichtigung des Gleason Score, die Schätzung von Wahrscheinlichkeiten bezüglich der Entdeckung eines klinisch signifikanten Prostatakarzinoms nach anormaler Erhöhung des PSA-Spiegels innerhalb eines Jahres (PSA-V), die Nichtberücksichtigung des Einflusses anderer Parameter bei der Entscheidungsfindung für oder wider eine Biopsie sowie die lediglich für einen 2-Jahres-Zeitraum angesetzten Kosten.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Keine Ausführungen.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass die Früherkennung von Prostatakarzinom bei Männern im Alter von über 50 Jahren vernünftig erscheint, jedoch mehr Schaden als Nutzen bringen kann, wenn sich die Entscheidung für eine Biopsie lediglich auf einen erhöhten PSA-Spiegel stützt. Nach Ansicht der Autoren trifft dies insbesondere auf Männer im Alter ab 70 Jahren zu, bei denen sich die Biopsie bei allen Exzess-PSA-Werten von ≤ 20 ng/ml in ihrer Studie als nicht kosten-effektiv und mit möglichen nachteiligen Auswirkungen verbunden erwies.

9 Kommentar

Entscheidungsanalyse zu altersabhängigen Biopsiekriterien beim Prävalenz- und Inzidenz-Screening mit expliziter Gegenüberstellung von Technologiealternativen. Die Transparenz des beschriebenen Modells ist eingeschränkt, die Formulierung der Interventionsstrategien ist nicht präzise. Zur Nachvollziehbarkeit der Berechnungen fehlt die Explizierung einiger Parameter und Annahmen. Die Parameterschätzer der Basisanalyse wurden nicht aus einer umfassenden Literaturrecherche abgeleitet (selection bias), allerdings wurde auf dieses Problem teilweise durch die Sensitivitätsanalyse eingegangen. Die Präsentation der ermittelten Effekte erfolgte vorwiegend graphisch. Detaillierte Kostenrechnungen wurden keine aufgeführt. Die in der Analyse berücksichtigten Kosten entsprechen nur teilweise den für die explizit genannte Perspektive relevanten Ressourcenveränderungen, der für die Kosten gewählte Zeitrahmen unterscheidet sich vom Zeithorizont des in der Analyse berücksichtigten Effekts. Die Schwächen von Studienführung und Dokumentation schränken die interne Validität, Reliabilität und Generalisierbarkeit der Analyse erheblich ein.

Gustafsson O, Carlsson P, Norming U, Nyman CR, Svensson H:
Cost-Effectiveness Analysis in Early Detection of Prostate Cancer: An Evaluation of Six Screening Strategies in a Randomly Selected Population of 2,400 Men.
Prostate, 1995; 26: 299-309.

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels Bestimmung von Serum-PSA, DRU und TRUS.
 Explizite Gegenüberstellung von verschiedenen Früherkennungsprogrammen.

1-2 Fragestellung

Ermittlung der Kosten-Effektivität verschiedener Früherkennungsprogramme auf Basis der Daten einer bereits veröffentlichten Querschnittstudie¹⁴.

Screening-Strategie 1): DRU $-\{ \oplus \} \rightarrow$ Biopsie

Screening-Strategie 2): TRUS $-\{ \oplus \} \rightarrow$ Biopsie

Screening-Strategie 3): DRU $-\{ \oplus \} \rightarrow$ Biopsie *und*

DRU $-\{ \emptyset \} \rightarrow$ TRUS $-\{ \oplus \} \rightarrow$ Biopsie *und*

TRUS $-\{ \emptyset \} \rightarrow$ PSA $-\{ \geq 7\text{ng/ml} \} \rightarrow$ Biopsie

Screening-Strategie 4): PSA $-\{ \geq 4\text{ng/ml} \} \rightarrow$ DRU $-\{ \oplus \} \rightarrow$ Biopsie

Screening-Strategie 5): PSA $-\{ \geq 4\text{ng/ml} \} \rightarrow$ TRUS $-\{ \oplus \} \rightarrow$ Biopsie

Screening-Strategie 6): DRU $-\{ \oplus \} \rightarrow$ Biopsie *und/oder*

PSA $-\{ \geq 4\text{ng/ml} \} \rightarrow$ TRUS $-\{ \oplus \} \rightarrow$ Biopsie

1-3 Perspektive

Keine Angaben.

1-4 Zeithorizont

Ermittelt wurden Effekte und Kosten für den Zeitraum der Studie (Prävalenz-Screening).
 Ein Zeitraum wurde nicht explizit genannt.

¹⁴ Gustafsson O, Norming U, Almgard LE, Fredriksson A, Gustavsson G, Harvig B, Nyman CR: *Diagnostic Methods in the Detection of Prostate Cancer: A Study on a Randomly Selected Population of 2,400 Men.* J Urol 1992; 148: 1827-1831.

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Originalstudie.

Datenquelle für Gesundheitseffekt: Originalstudie (Querschnittstudie) mit retrospektiver Entscheidungsanalyse.

Datenquelle für Ressourcenverbrauch und Mengengerüst: Originalstudie (Querschnittstudie) und Modellierung.

Datenquelle(n) für Preise: Institutionseigene Daten; Statistisches Jahrbuch für Schweden; für Komplikationskosten Daten unbekannter Herkunft zur ambulanten und stationären Versorgung.

2-2 Datierung der zugrundeliegenden Daten

Die Datenerhebung zum Gesundheitseffekt und Ressourcenverbrauch erfolgte zwischen Januar 1988 und Juni 1989.

Die Preisangaben beziehen sich auf die Werte des Jahres 1990.

2-3 Studienpopulation / Zielpopulation

Männer im Alter von 55 bis 70 Jahren im größtenteils städtischen Einzugsgebiet des Krankenhauses.

2-4 Setting

Tangierte Versorgungsstrukturen: ambulante Krankenhausversorgung.

Der klinische wie auch der ökonomische Teil der Studie wurden in Schweden durchgeführt.

2-5 Spezifikation der Technologie

PSA-Assay: Hybritech Tandem-R (monoklonal), Entscheidungsgrenze bei 4ng/ml bzw. 7ng/ml.

DRU: Palpation durch Urologen, Untersuchung in stehender, nach vorwärts gebeugter Position.

TRUS: Real-Time-Verfahren, 7.0MHz-Schallkopf, biplanar (für die 450 ersten Untersuchungen) bzw. multiplanar (für die restlichen Untersuchungen).

Biopsie: transrektale Stanzbiopsie (18G-Nadeln, Biopty-Pistole) mit 2 bis 4 Stanzen und transrektale Saugbiopsie (22G-Nadeln) mit 3 Aspiraten unter Ultraschallkontrolle und/oder digitaler Führung bzw. randomisiert bei PSA-Werten $\geq 10\text{ng/ml}$ ohne verdächtiger DRU/TRUS; Antibiotikaphylaxe bei etwa der Hälfte der Probanden. Bei Nichtverwertbarkeit der Gewebeproben bzw. fortbestehendem Tumorverdacht bei negativer Gewebeprobenanalyse Biopsiewiederholung.

3 Gesundheitseffekte

Screening-/Diagnostestudie

3-1 Untersuchte Zielgrößen

Primäre Zielgrößen waren die nach Biopsieabsicherung identifizierte Anzahl Prostatakarzinome im Probandenkollektiv, die Anzahl identifizierter Stadium \leq T2a-Karzinome und die Anzahl identifizierter und kurativ behandelter Karzinome.

3-2 Ein-/Ausschlusskriterien

Als Einschlusskriterien wurden Geschlecht, Alter und Wohnort im Einzugsgebiet genannt. Ausschlusskriterium war eine frühere Diagnose von Prostatakarzinom (Registrierung im schwedischen Krebsregister).

3-3 Rekrutierungsmodus

Randomisierte Ziehung von 2,400 Männern aus den Volkszählungslisten (9% der entsprechenden Altersgruppe). Kostenloses Screening-Angebot. Informations- und Einladungsschreiben zur Untersuchung (PSA-Assay, DRU, TRUS) mit näherer schriftlicher Information und Terminfestsetzung drei Wochen später.

3-4 Teilnahmerate

74% (1,782/2,400).

3-5 Reproduzierbarkeit der Testergebnisse

Die eingesetzten Diagnosemethoden wurden spezifiziert. Die Entscheidungsgrenze bei der PSA-Bestimmung ($\geq 4\text{ng/ml}$, im ursprünglichen Interventionspfad zusätzlich $\geq 10\text{ng/ml}$ bei negativem DRU/TRUS-Befund) wurde angegeben. Entscheidungskriterien

bezüglich der Interpretation der DRU (palpatorisch indurierte, knotige oder asymmetrische Prostata) bzw. der TRUS (nicht durch Zysten bedingte echoarme oder asymmetrische Areale) wurden explizit genannt. Die DRU wurde von fünf, die TRUS von zwei Urologen durchgeführt. Vorkehrungen zur Kontrolle der Variabilität von Blut- und Gewebeprobenanalysen wurden nicht spezifiziert. Stadium T1-Tumoren wurden entgegen der TNM-Klassifikation (UICC 1987) nicht als inzidentelle Karzinome definiert, sondern als Tumoren der Transitionalzone, T2a- und T2b-Karzinome als organbegrenzte Tumoren mit einer Ausdehnung von $\leq 1.5\text{cm}$ bzw. $> 1.5\text{cm}$.

3-6 Auswertung der Testergebnisse

Alle Probanden unterzogen sich einer PSA-Bestimmung, DRU und TRUS. Die Biopsie als Goldstandard wurde lediglich bei Probanden mit positivem Screening-Testbefund durchgeführt. Die Bestimmung der einzelnen Testergebnisse erfolgte in gegenseitiger Verblindung, nicht aber die Festlegung des Krankheitsstatus.

3-7 Drop-Outs

Keine Angaben.

3-8 Ergebnisse der Screening-/Diagnosestudie

Deskriptive Präsentation der Ergebnisse entlang des ursprünglichen Interventionspfades sowie entlang der modellierten Screening-Strategien. Angabe von bedingt abhängigen positiven Vorhersagewerten (PPV) für verschiedene Testkombinationen. In der vorliegenden Veröffentlichung wurden keine statistischen Testverfahren aufgeführt.

Ursprünglicher Interventionspfad (PSA \rightarrow DRU+TRUS $-\oplus$ ^{w/o} oder PSA $\geq 10\text{ng/ml}$) \rightarrow Biopsie): 19.9% (355/1,782) der Studienpopulation hatten positive Befunde bei der DRU und/oder TRUS. Insgesamt konnten 65 (3.6%) Karzinome entdeckt werden. 42 (2.4%) Karzinome wurden aufgrund einer positiven DRU identifiziert, 20 (1.1%) aufgrund einer positiven TRUS; 3 (0.2%) zusätzliche Karzinome wurden bei Wiederholung der Untersuchung aufgrund von PSA-Werten $\geq 7\text{ng/ml}$ entdeckt. 17% der Gesamtstudienpopulation bzw. 80% der Karzinomträger hatten PSA-Spiegel $\geq 4\text{ng/ml}$. Zur Absicherung der Diagnose wurden bei 371 Probanden 413 Biopsien durchgeführt (Wiederholung der Biopsie in durchschnittlich 11.5% der Fälle).

Die modellierten Screening-Strategien beziehen sich jeweils auf 1,000 Screening-Teilnehmer:

Screening-Strategie 1): Bei 109 Probanden war die DRU verdächtig. Mit 122 durchgeführten Biopsien wurden 24 Karzinome entdeckt. Sechs Karzinomträger hatten Stadium \leq T2a-Karzinome, 15 wurden kurativ behandelt.

Screening-Strategie 2): Bei 137 Probanden war die TRUS verdächtig. Mit 153 durchgeführten Biopsien wurden 33 Karzinome entdeckt. Zehn Karzinomträger hatten Stadium \leq T2a-Karzinome, 20 wurden kurativ behandelt.

Screening-Strategie 3): Bei 109 Probanden war die DRU verdächtig. Von den 891 Patienten mit negativem DRU-Befund hatten 90 einen positiven TRUS-Befund. Von den 801 Patienten mit negativem TRUS-Befund hatten 34 PSA-Werte \geq 7ng/ml. Mit 243 durchgeführten Biopsien wurden 36 Karzinome entdeckt. 12 Karzinomträger hatten Stadium \leq T2a-Karzinome, 23 wurden kurativ behandelt.

Screening-Strategie 4): Bei 170 Probanden war der PSA-Spiegel \geq 4ng/ml, wobei sich die DRU bei 22 als verdächtig erwies. Mit 24 durchgeführten Biopsien wurden 20 Karzinome entdeckt. 4 Karzinomträger hatten Stadium \leq T2a-Karzinome, 12 wurden kurativ behandelt.

Screening-Strategie 5): Bei 170 Probanden war der PSA-Spiegel \geq 4ng/ml, wobei sich die TRUS bei 49 als verdächtig erwies. Mit 55 durchgeführten Biopsien wurden 26 Karzinome entdeckt. 6 Karzinomträger hatten Stadium \leq T2a-Karzinome, 18 wurden kurativ behandelt.

Screening-Strategie 6): Bei 109 Probanden war die DRU verdächtig. Von den 891 Patienten mit negativem DRU-Befund hatten 130 PSA-Werte \geq 4ng/ml, wobei sich die TRUS bei 20 als verdächtig erwies. Mit 144 durchgeführten Biopsien wurden 32 Karzinome entdeckt. 9 Karzinomträger hatten Stadium \leq T2a-Karzinome, 21 wurden kurativ behandelt.

Die Extrapolation der Effekte des Prävalenz-Screenings auf ein nationales Screening-Programm (640,000 Männer im Alter von 55 bis 70 Jahren) ergibt bei der Screening-Strategie 4) 9,500 entdeckte Karzinome bzw. bei der Screening-Strategie 3) 17,000 entdeckte Karzinome.

3-9 Effektmaße für die ökonomische Analyse

Anzahl identifizierter Prostatakarzinome, Anzahl identifizierter Stadium \leq T2a-Prostatakarzinome, Anzahl identifizierter und kurativ behandelter Prostatakarzinome.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten und z.T. Fixkosten für die medizinische Versorgung.

- Screening/Diagnostik: Personalkosten (1 Urologe, 1 Krankenschwester, 1 Sekretärin halbtags), Einrichtung (TRUS-Ausstattung, Räumlichkeiten), PSA-Assay, Biopsie einschließlich Histopathologie/Zytologie
- Behandlung: Behandlung von Komplikationen im Rahmen von Screening/Diagnostik (Infektionen), explizit keine Berücksichtigung des damit verbundenen Produktivitätsverlusts

Die Rekrutierungskosten wurden berücksichtigt.

Ferner wurden für Patienten anfallenden Kosten berücksichtigt, diese wurden als indirekte Kosten definiert. Sie wurden ermittelt anhand von 1,099 Probandenfragebögen zu Transportkosten, Zeitaufwand für die Untersuchung und Berufstätigkeit.

- Patienten-Transport
- untersuchungsbedingter Zeitaufwand für Patienten

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach Aufwand der Leistungserstellung (Kostenarten) und nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Die sich aus den Interventionspfaden ergebenden Mengen wurden spezifiziert. Biopsiewiederholungen in 11.5% der Fälle wurden berücksichtigt. Hinsichtlich der Materialkosten bei der TRUS wurde implizit eine anderweitige Auslastung der Apparatur außerhalb ihrer studienbedingten Einsatzzeit unterstellt. Die Komplikationskosten wurden als Durchschnittswert pro Biopsie angesetzt.

4-3 Monetäre Bewertung des Mengengerüsts

Pauschales Verfahren. Die Preise pro Mengeneinheit wurden aufgeführt. Die Berechnungen der untersuchungsbedingten Zeitkosten für die Probanden basieren auf dem durchschnittlichen altersspezifischen Bruttostundenlohn; der monetäre Wert von Freizeit (bei Berentung) wurde auf 25% desjenigen von bezahlter Arbeit angesetzt.

4-4 Wahrung

Amerikanische Dollar (US-\$), basierend auf der Konversion von Schwedischen Kronen (skr) zum durchschnittlichen Wechselkurs des Jahres 1990: US-\$ 1.00 = skr 5.90.

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die Anzahl identifizierter Prostatakarzinome / \leq T2a-Karzinome / kurativ behandelbare Karzinome betrug fur Strategie 1) 24/6/15, fur Strategie 2) 33/10/20, fur Strategie 3) 36/12/23, fur Strategie 4) 20/4/12, fur Strategie 5) 26/6/18 und fur Strategie 6) 32/9/21.

6-2 Ermittelte Kosten

Die Gesamtkosten pro 1,000 Screening-Teilnehmer beliefen sich bei der Screening-Strategie 1) auf US-\$ 74,500, bei der Screening-Strategie 2) auf US-\$ 97,500, bei der Screening-Strategie 3) auf US-\$ 160,900, bei der Screening-Strategie 4) auf US-\$ 71,200, bei der Screening-Strategie 5) auf US-\$ 82,600 und bei der Screening-Strategie 6) auf US-\$ 116,100. Bei allen Screening-Strategien betrug dabei die Summe der sog. indirekten Kosten (untersuchungsbedingte Transport- und Zeitkosten fur die Probanden) und der Kosten bei Komplikationen etwa die Halfte der Gesamtkosten.

Die Extrapolation der Kosten des Pravalenz-Screenings auf ein nationales Screening-Programm (640,000 Manner im Alter von 55 bis 70 Jahren) ergibt fur die Screening-Strategie 4) Kosten in Hohe von US-\$ 35.2 Mio. bzw. fur die Screening-Strategie 3) Kosten in Hohe von US-\$ 76.1 Mio.

6-3 Synthese von Kosten und Effekten

Die Kosten pro entdecktem Prostatakarzinom, pro entdecktem Stadium \leq T2a-Prostatakarzinom bzw. pro entdecktem und kurativ behandeltem Prostatakarzinom beliefen sich bei der Screening-Strategie 1) auf US-\$ 3,100, US-\$ 12,420 bzw. US-\$ 4,970; bei der Screening-Strategie 2) auf US-\$ 2,950, US-\$ 9,750 bzw. US-\$ 4,880; bei der Screening-Strategie 3) auf US-\$ 4,470, US-\$ 13,410 bzw. US-\$ 7,000; bei der Screening-Strategie 4) auf US-\$ 3,560, US-\$ 17,800 bzw. US-\$ 5,930; bei der Screening-Strategie 5) auf US-\$ 3,180, US-\$ 13,770 bzw. US-\$ 4,590; bei der Screening-Strategie

6) auf US-\$ 3,630, US-\$ 12,900 bzw. US-\$ 5,530. Die Gesamtkosten nahmen in der Reihenfolge Strategie 4)-1)-5)-2)-6)-3) zu. Die inkrementellen Kosten bezogen auf die in der Rangfolge nächstkostengünstigere Strategie pro zusätzlich entdecktem und kurativ behandeltem Prostatakarzinom betragen bei der Screening-Strategie 1) US-\$ 1,100; bei der Screening-Strategie 5) US-\$ 2,700; bei der Screening-Strategie 2) US-\$ 7,450; bei der Screening-Strategie 6) US-\$ 18,600 und bei der Screening-Strategie 3) US-\$ 22,400.

7 Behandlung von Unsicherheiten

Keine Ausführungen.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren weisen darauf hin, dass als Effektmaß in einer Screeningstudie die gewonnenen Lebensjahre bzw. lebensqualitätsadjustierte Lebensjahre heranzuziehen sind. Aufgrund fehlender Daten wurden nur Kosten-Effektivitäts-Quotienten für intermediäre Outcomes berechnet: die Kosten pro entdecktem Prostatakarzinomfall, die Kosten pro entdecktem kleinen Prostatakarzinom und die Kosten pro behandeltem Prostatakarzinom.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Die Autoren sind der Ansicht, dass die Absolutzahlen ihrer Kostenrechnung nicht ohne weiteres auf die Situation in anderen Ländern übertragbar sind. Sie nehmen jedoch an, dass die Verhältnisse zwischen den unterschiedlichen Kosten unabhängig vom allgemeinen Kostenniveau sind und valide für die westliche Welt sind. Einen weiteren Problempunkt bei der Übertragbarkeit ihrer Ergebnisse sehen die Autoren bei der von Ort zu Ort unterschiedlichen Verfügbarkeit der verschiedenen diagnostischen Verfahren.

8-3 Schlussfolgerungen

Mittels der Screening-Strategie 5) (TRUS bei PSA-Werten ≥ 4 ng/ml) konnten 80% der kurativ behandelten Karzinome entdeckt werden; sie erwies sich damit als kosteneffektivste Screening-Strategie, weshalb sie von den Autoren neben der Screening-Strategie 6) bevorzugt wird. Mittels der Screening-Strategie 6) (DRU und/oder PSA-Bestimmung mit TRUS bei PSA-Werten ≥ 4 ng/ml) konnten 90% der kurativ behandelten Karzinome identifiziert und der Anteil der mittels TRUS untersuchten Personen vermindert werden. Allerdings fällt die marginale Kosten-Effektivität bei dieser Strategie

relativ hoch aus. Die Autoren sind dennoch der Ansicht, dass die Höhe des Kosten-Effektivitäts-Quotienten akzeptabel ist. Die anderen Screening-Strategien werden von den Autoren als ungünstig angesehen und folglich abgelehnt.

9 Kommentar

Entscheidungsanalyse zu Biopsiekriterien beim Prävalenz-Screening mit expliziter Gegenüberstellung von Technologiealternativen. Als interessierende Zielgröße wurden intermediäre Outcome-Parameter gewählt. Trotz randomisierter Rekrutierung ist aufgrund der Teilnahmerate von 74% die Möglichkeit eines selection bias gegeben; ein expectation bias bei der Durchführung der Screening-Tests wurde durch gegenseitige Verblindung weitgehend kontrolliert, ist hinsichtlich der Durchführung und Auswertung der Biopsie jedoch möglich. Zur Interrater-Reliabilität wurden keine Angaben gemacht. Die lediglich auf eine selektierte Populationsuntergruppe beschränkte Anwendung des Goldstandards bedingt einen work-up bias. Ein PSA-Wert zwischen 4ng/ml und 10ng/ml allein galt bei der ursprünglichen Interventionsstrategie nicht als Entscheidungskriterium für eine Biopsie; der Einbezug dieses verification bias in die Modellierung wurde durch die Wahl entsprechender Testsequenzen vermieden. Die Kostenberechnung ist nachvollziehbar. Protokollkosten sind teilweise mit berücksichtigt worden. Hervorzuheben ist der Versuch, bei den Patienten anfallende direkte Kosten mit einzubeziehen und Mehrfachdurchführungen der Biopsie, wie sie auch in der Praxis vorkommen, zu berücksichtigen. Die Analyse erscheint sorgfältig durchgeführt, weitgehend transparent und nachvollziehbar, jedoch müssen hinsichtlich der internen Validität, Reliabilität und Generalisierbarkeit ihrer Ergebnisse die genannten Verzerrungen in Betracht gezogen werden.

Kantrowitz W, Doyle J, Semeraro J, Krane RJ:
Prostate Cancer Screening in a Large Corporation Population.
Journal of Occupational and Environmental Medicine, 1995; 37(10): 1193-1198.

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels Bestimmung von Serum-PSA und DRU. Kein expliziter Vergleich von Technologiealternativen.

1-2 Fragestellung

Beurteilung der Durchführbarkeit sowie Ermittlung der Kosten von Massen-Screening im Rahmen der betriebsärztlichen Versorgung.

Screening-Strategie:

WHO-Fragebogen (WHO Urinary Tract Symptom Checklist, 1991),

PSA+DRU^{BA} $-\left[\geq 4\text{ng/ml und/oder } \oplus \right] \rightarrow \text{DRU}^{\text{U}}$ $-\left[\oplus \text{ und/oder PSA} \geq 4\text{ng/ml} \right] \rightarrow \text{TRUS}^{\text{U}}$
 $-\left[\oplus \text{ und/oder PSA} > 10\text{ng/ml} \right] \rightarrow \text{Biopsie}$

BA = Betriebsarzt, U = Urologe

1-3 Perspektive

Keine Angaben.

1-4 Zeithorizont

Ermittelt wurden Effekte und Kosten für den Zeitraum der Studie (Prävalenz-Screening). Das Prävalenz-Screening war nach 6 Wochen abgeschlossen.

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Originalstudie.

Datenquelle für Gesundheitseffekt: Originalstudie (Querschnittstudie).

Datenquelle für Ressourcenverbrauch und Mengengerüst: Originalstudie (Querschnittstudie).

Datenquelle(n) für Preise: keine Angaben.

2-2 Datierung der zugrundeliegenden Daten

Keine Angaben bezüglich des Zeitpunktes der Datenerhebung zum Gesundheitseffekt und Ressourcenverbrauch; die PSA-Bestimmung und DRU durch die Betriebsärzte nahm sechs Wochen in Anspruch.

Keine Angaben hinsichtlich des Bezugsjahres für die Preisangaben.

2-3 Studienpopulation / Zielpopulation

Männer im Alter zwischen 50 und 65 Jahren, alle Mitarbeiter des gleichen Unternehmens.

2-4 Setting

Tangierte Versorgungsstrukturen: ambulante Versorgung (betriebsärztliche Versorgung und niedergelassene Mediziner); keine Angaben bezüglich der von der Behandlung tangierten Versorgungsstruktur(en).

Der klinische wie auch der ökonomische Teil der Studie wurden in Massachusetts, USA durchgeführt.

2-5 Spezifikation der Technologie

PSA-Assay: Abbott IMx (monoklonal).

DRU: Palpation durch zwei Betriebsärzte sowie verschiedene Urologen.

TRUS: keine näheren Angaben.

Biopsie: transrektale Stanzbiopsie unter Ultraschallkontrolle, Quadrantenbiopsie; Antibiotikaprophylaxe.

3 Gesundheitseffekte

Screening-/Diagnosestudie

3-1 Untersuchte Zielgrößen

Primäre Zielgröße war die Anzahl der identifizierten Prostatakarzinome im Probandenkollektiv.

3-2 Ein-/Ausschlusskriterien

Als Einschlusskriterien wurden Geschlecht, Alter und Betriebszugehörigkeit genannt.

3-3 Rekrutierungsmodus

Einladungsschreiben an die Männer der Zielpopulation nach einer sechsmonatigen Informationskampagne bestehend aus persönlichen Anschreiben, Artikeln in der Betriebszeitschrift und einer Telefon-Hotline. Persönliche Kontaktaufnahme als Nachfassaktion.

3-4 Teilnahmerate

54% (1,219/2,241).

3-5 Reproduzierbarkeit der Testergebnisse

Die eingesetzten Diagnosemethoden wurden weitgehend spezifiziert. Die Entscheidungsgrenze bei der PSA-Bestimmung ($>4\text{ng/ml}$) wurde angegeben. Entscheidungskriterien bezüglich der Interpretation der TRUS wurden nur zum Teil explizit genannt (echoarme Areale; keine Angaben bezüglich der Entscheidungsgrenze bei Rückgriff auf die PSA-D), nicht aber hinsichtlich der DRU-Befunde. Vorkehrungen zur Kontrolle der Variabilität der DRU und TRUS sowie von Blut- und Gewebeprobenanalysen wurden nicht spezifiziert.

3-6 Auswertung der Testergebnisse

Die Biopsie als Goldstandard wurde lediglich bei Probanden mit positivem TRUS-Befund durchgeführt. Die Bestimmung der Testergebnisse und des Krankheitsstatus erfolgten nicht in gegenseitiger Verblindung.

3-7 Drop-Outs

Keine Angaben.

3-8 Ergebnisse der Screening-/Diagnosestudie

Deskriptive Präsentation der Ergebnisse entlang des Interventionspfades ohne Anwendung statistischer Testverfahren oder Angabe von Kenngrößen der Tests.

5.0% (61/1,219) der Probanden wiesen einen positiven DRU-Befund bei der betriebsärztlichen Untersuchung auf. Die DRU durch einen Urologen erwies sich bei 27 der 61 überwiesenen Probanden als suspekt. Die Anzahl verdächtiger Befunde bei der TRUS ist unklar; die Biopsie bestätigte den Tumorverdacht in 4 (0.3%) Fällen. 37

Probanden wiesen PSA-Spiegel von ≥ 4 ng/ml auf, 5 davon Werte von > 10 ng/ml; einer Biopsie unterzogen sich 18 dieser Probanden, wobei sich der Tumorverdacht in 10 (0,8%) Fällen bestätigte. Zwei Karzinomträger wiesen sowohl erhöhte PSA-Spiegel als auch einen positiven DRU-Befund auf. Insgesamt wurde in 12 (1%) Fällen ein Tumor diagnostiziert. Tumor-Grading: Gleason-Score 2 in 1, Gleason-Score 3 in 1, Gleason-Score 4 in 1, Gleason-Score 5 in 1, Gleason-Score 6 in 2, Gleason-Score 7 in 4 und Gleason-Score 9 in 1 Fall; in einem Fall konnte der Gleason-Score nicht angegeben werden. 10 Karzinomträger unterzogen sich einer radikalen Prostatektomie, wobei für alle ein N0-Stadium festgestellt wurde; 1 Patient erhielt Strahlentherapie, bei 1 Karzinomträger wurde die Behandlung hinausgeschoben (watchful waiting). Zwischen dem Score des WHO-Fragebogens und der Karzinomentdeckung schien keine Korrelation zu bestehen. Das durchschnittliche Alter der Probanden lag bei 56,2 Jahren. Die Abwesenheit vom Arbeitsplatz betrug pro teilnehmenden Angestellten ungefähr 15 Minuten.

3-9 Effektmaße für die ökonomische Analyse

Anzahl identifizierter Prostatakarzinome.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

- Screening/Diagnostik: PSA-Assay, zusätzliches Personal (ohne Kosten für reguläres Personal, bestehend aus 2 Arbeitsmedizinern, 2 Schwestern, 3 Assistentinnen; aber einschließlich zusätzlich benötigtes Personal wie z.B. Konsultation Urologen), medizinisches Material, Informationsmaterial, TRUS, Biopsie einschließlich Histopathologie, Antibiotika (vor Biopsie).

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach Aufwand der Leistungserstellung (Kostenarten) und nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Einzelne Mengen wurden nur zum Teil spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Die Art, die Quelle(n) und das Bezugsjahr der Preise wurden nicht spezifiziert. Die Preise pro Mengeneinheit wurden zum Teil aufgeführt.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die beschriebene Screening-Strategie führte zur Entdeckung von 12 Prostatakarzinomen bei 1,219 Screening-Teilnehmern (1.0%).

6-2 Ermittelte Kosten

Die Gesamtkosten für das Screening (ohne Behandlung) beliefen sich auf US-\$ 72,138, entsprechend US-\$ 59 pro Screening-Teilnehmer.

6-3 Synthese von Kosten und Effekten

Die Kosten pro entdecktem Prostatakarzinom beliefen sich auf US-\$ 6,012.

7 Behandlung von Unsicherheiten

Keine Ausführungen.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Angesichts der im Vergleich zu Literaturangaben geringeren Karzinomentdeckungsrate wird von den Autoren vermerkt, dass bereits vorher im Rahmen der regulären betrieblichen Gesundheitsuntersuchungen die Möglichkeit einer DRU und/oder einer PSA-Bestimmung genutzt werden konnte.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Es wird auf die altersmäßig eingeschränkte Studienpopulation hingewiesen. Die Tatsache, dass eine sich im Arbeitsleben befindende Teilpopulation untersucht wurde (< 65 J.), wird als eine weitere Erklärung für die niedrige Entdeckungsrate angeführt.

8-3 Schlussfolgerungen

Nach Ansicht der Autoren sind die DRU und die PSA-Bestimmung für sich genommen nicht ausreichend als Screening Tests. Sie kommen zu dem Schluss, dass eine Früherkennungsmaßnahme mit Kombination der beiden Untersuchungen im betrieblichen Rahmen durchführbar, effektiv und effizient ist.

9 Kommentar

Studie zu Biopsiekriterien beim Prävalenz-Screening ohne explizite Gegenüberstellung von Technologiealternativen. Im Vordergrund steht eher die Durchführbarkeit einer Screening-Strategie im betrieblichen Rahmen und weniger der langfristige Nutzen der Früherkennung, was sich auch in der Wahl eines intermediären Outcome-Parameters als interessierende klinische Zielgröße ausdrückt. Die Möglichkeit eines selection bias, eines diagnostic suspicion bias sowie eines expectation bias bei der Durchführung der untersucherabhängigen Tests ist gegeben. Die lediglich auf eine selektierte Populationsuntergruppe beschränkte Anwendung des Goldstandards in Verbindung mit einem work-up bias reduziert weiterhin die interne Validität, Reliabilität und Generalisierbarkeit der klinischen Untersuchungsergebnisse. Der Kostenberechnung mangelt es an Transparenz, Personalkosten sind nach dem Fixkostenprinzip verrechnet, d.h. nicht berücksichtigt worden. Aus welcher impliziten Perspektive der Kostenermittlung zu sehen ist, ist unklar: eine Abgrenzung zwischen den Kosten, die für den Arbeitgeber anfallen und solchen, die für die allgemeine Gesundheitsversorgung entstehen, ist nicht gegeben: es erscheint dabei eher als unwahrscheinlich, dass ein Arbeitgeber neben den Kosten für die Früherkennung im betrieblichen Rahmen auch die Abklärungs- und Behandlungskosten übernehmen würde. Es sei darauf hingewiesen, dass zwischen den Ausführungen im Text und der Entscheidungsmatrix in Abbildung 2 Inkongruenzen auftreten und die Darstellung der Studienergebnisse damit teilweise nicht nachvollziehbar ist. Im Rahmen der Beurteilung des Screening-Verfahrens wird auf dessen Effizienz hingewiesen, ohne dabei jedoch den Effizienzbegriff im gesundheitsökonomischen Sinne zu verwenden bzw. anderweitig zu spezifizieren.

Hinweis: in Tabelle 4 (Kosten) der Publikation ist ein Fehler unterlaufen. Bei "Ultrasounds, 53 x US-\$ 405.00 (hospital and professional component)" sollte US-\$ 21,465.00 statt US-\$ 21,456.00 stehen. Damit belaufen sich die Gesamtkosten (Total) auf US-\$ 72,147.00 (statt auf US-\$ 72,138.00); die Kosten pro identifiziertem Karzinom (US-\$ 6,012) bleiben gleich.

**Krahn MD, Mahoney JE, Eckman MH, Trachtenberg J, Pauker SG, Detsky AS:
 Screening for Prostate Cancer: A Decision Analytic View.
 Journal of the American Medical Association, 1994; 272(10): 773-780.**

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels DRU, Bestimmung von Serum-PSA und TRUS. Explizite Gegenüberstellung der Alternativen "Screening" mittels verschiedener Früherkennungsprogramme und "Kein Screening" (Behandlung bei auftretender Symptomatik).

1-2 Fragestellung

Entscheidungsanalytische Kosten-Nutzwert-Analyse zur Ermittlung der klinischen und ökonomischen Auswirkungen von vier Früherkennungsprogrammen in der Allgemeinbevölkerung sowie in bestimmten Untergruppen.

Screening-Strategie 1): DRU $-\{ \oplus \} \rightarrow$ TRUS $-\{ \oplus \} \rightarrow$ Biopsie

Screening-Strategie 2): PSA $-\{ >4\text{ng/ml} \} \rightarrow$ DRU+TRUS $-\{ \oplus^{u/o} \} \rightarrow$ Biopsie

Screening-Strategie 3): DRU $-\{ \oplus \} \rightarrow$ Biopsie *und/oder* PSA $-\{ >4\text{ng/ml} \} \rightarrow$ TRUS $-\{ \oplus \} \rightarrow$ Biopsie

Screening-Strategie 4): PSA+DRU+TRUS $-\{ >4\text{ng/ml und/oder} \oplus \} \rightarrow$ Biopsie

1-3 Perspektive

Perspektive des Kostenträgers.

1-4 Zeithorizont

Ermittelt wurden die lebenslangen Effekte und Kosten im Rahmen des Prävalenz-Screenings.

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse; Kosten-Nutzwert-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Synthese.

Datenquellen für klinische Parameter: klinische Entscheidungsanalyse mit Literaturdaten und Modellierung (Markov-Modell: Markov-Kohorte).

Datenquellen für Präferenzen: Chronische Gesundheitszustände: Methode der zeitlichen Abwägung (time-trade-off) an einer Gruppe von Urologen, Radioonkologen und Internisten (insgesamt 10 Personen). Lebensqualität nach Harnstauung: durch Schätzung der Autoren.

Datenquelle für Ressourcenverbrauch und Mengengerüst: Modellierung (Entscheidungsanalyse).

Datenquelle für Preise: Clinical Cost Manager des New England Medical Center (Transition Systems Inc., Boston, Massachusetts, USA).

2-2 Datierung der zugrundeliegenden Daten

Zeitpunkt und Zeitraum der Erhebung von Daten zu klinischen Parametern wurden nicht explizit aufgeführt (die klinischen Parameter leiten sich aus Studien ab, welche bis 1994 veröffentlicht wurden).

Die Preisangaben beziehen sich auf die Werte des Jahres 1992.

2-3 Studienpopulation / Zielpopulation

Kohorte von Männern der Allgemeinbevölkerung (keine expliziten Altersangaben; es wurden die Ergebnisse für Männer im Alter von 50 bis 70 Jahren berichtet). Subgruppenanalysen bezogen sich auf Männer mit krankheitsspezifischer Familienanamnese sowie auf Schwarze.

2-4 Setting

Tangierte Versorgungsstrukturen: ambulante und stationäre Versorgung.

Größtenteils keine explizite Lokalisation der (nordamerikanischen) Daten zu den klinischen Parametern. Der ökonomische Teil der Studie wurde in Ontario, Kanada durchgeführt.

2-5 Spezifikation der Technologie

PSA-Assay: monoklonal, Entscheidungsgrenze bei 4ng/ml.

DRU: positiver Befund bei palpatorisch asymmetrischer, indurierter oder knotiger Prostata.

TRUS: positiver Befund bei fokalen, nicht durch normale anatomische Strukturen, Zysten oder Artefakte bedingten echoarmen Arealen.

Biopsie: transrektale Biopsie unter Ultraschallkontrolle oder digitaler Führung.

3 Gesundheitseffekte

Synthese bereits veröffentlichter Studien

3-1 In der Synthese untersuchte klinische Parameter

Primäre Zielgrößen waren die Prävalenz von Prostatakarzinom, die Sensitivität und Spezifität der Tests, die Stadienverteilung der entdeckten Karzinome in Abhängigkeit der Tests und ihrer Einsatzstrategie (einfach oder sequentiell), die Wahrscheinlichkeit von kurz- und langfristigen Behandlungskomplikationen, die Krankheitsprogressionsraten (lokal und metastatisch) sowie die krankheitsstadiumspezifische Mortalität.

3-2 Annahmen

Die Analyse basierte auf folgenden expliziten Annahmen: In Betracht gezogen wird das Prävalenz-Screening. Multiple Tests sind bedingt abhängig. Die Biopsie gilt als Goldstandard. Das Staging nach positiver Biopsie umfasst eine PSA-Bestimmung, ein Knochenszintigramm und eine pelvine Computertomographie. In Verbindung mit einer radikalen Prostatektomie bei klinisch begrenztem Karzinom wird zusätzlich ein Lymphknoten-Staging nach Whitmore-Jewett durchgeführt. Die Behandlung bei Karzinomen der klinischen Stadien A und B besteht in einer radikalen Prostatektomie, bei Karzinomen des Stadiums C in einer Strahlentherapie, bei Karzinomen des Stadiums D in einer Orchiektomie. Die Ermittlung des Nutzens der Früherkennung wird auf die Entdeckung von Patienten mit potentiell heilbarem Karzinom der Stadien A und B beschränkt, da angenommen wird, dass sich Patienten mit anfänglich unentdecktem Karzinom der Stadien C und D im selben Jahr aufgrund von Symptomen in Behandlung begeben hätten. Die Behandlung von schwerer Harnstauung bei lokal oder metastatisch progrediente Erkrankung besteht in einer TUR-P, wobei Harnstauung lediglich bei Nichtbehandlung (Alternative "Kein Screening") auftreten kann. Kurzfristige wie auch

langfristige Komplikationen der Behandlung (rektale Verletzungen, Wundinfektion, Lungenembolie, tiefe Beinvenenthrombose, Anastomosenstriktur, systemische Sepsis, akute Zystitis, akute Proktitis bzw. Impotenz (vollständig und partiell), Inkontinenz (vollständig und partiell), Verletzung des Nervus obturatorius, chronische Zystitis, chronische Proktitis und Tod) bzw. der Nichtbehandlung (Harnstauung) werden in Betracht gezogen. Die Stadienverteilung der entdeckten Karzinome wie auch die perioperative Mortalität sind altersunabhängig.

Aus dem technischen Bericht geht die implizite Annahme hervor, dass für negative Screening-Testbefunde eine gleiche Karzinomwahrscheinlichkeit und -stadienverteilung besteht wie im Falle der Alternative "Kein Screening". Außerdem wurde implizit von einer 100%igen Compliance mit Biopsie und Therapie ausgegangen.

3-3 Berücksichtigung von Primärstudien: Studiendesigns und Ein-/Ausschlusskriterien

Die Krankheitsprävalenz in der Gesamtbevölkerung wurde aus prospektiven Kohortenstudien zum Screening des Prostatakarzinoms abgeleitet, welche alle untersuchten Screening-Methoden (PSA, DRU, TRUS) eingesetzt hatten. Die Krankheitsprävalenz in den untersuchten Subgruppen wurde anhand einer Fall-Kontroll-Studie (Männer mit krankheitsspezifischer Familienanamnese) bzw. anhand einer Studie nicht näher genannten Designs (Schwarze) geschätzt. Veröffentlichungen im Rahmen des American Cancer Society National Prostate Cancer Detection Project (ACS-NPCDP-Studie: prospektive Kohortenstudie)^{15,16,17} wurden für die Sensitivität und Spezifität der Tests sowie für die Stadienverteilung herangezogen. Die Wahrscheinlichkeiten von kurzfristigen und langfristigen Komplikationen der Behandlung basieren auf Daten aus publizierten Fallserien. Zur Schätzung der Krankheitsprogredienzraten und der krankheitsstadiumspezifischen Mortalität wurde auf Studien unterschiedlichsten Designs (RCTs, Kohortenstudien, Fallserien) zurückgegriffen; zur Ermittlung der krankheitsstadiumspezifischen Mortalität wurden nur Studien, die eine prostatakarzinombedingte Mortalität berichteten, berücksichtigt. Keine weiteren Angaben zu Ein- bzw. Ausschlusskriterien bei der Berücksichtigung der Studien.

¹⁵ Mettlin C, Lee F, Drago J, Murphy GP, et al.: *The American Cancer Society National Prostate Cancer Detection Project. Findings on the Detection of Early Prostate Cancer in 2,425 Men.* Cancer 1991; 67:2949-2958.

¹⁶ Babaian RJ, Mettlin CJ, Kane R, Murphy GP, et al.: *The Relationship of Prostate-Specific Antigen to Digital Rectal Examination and Transrectal Ultrasonography in Detecting Adenocarcinoma of the Prostate.* Cancer 1992; 69: 1195-2000.

¹⁷ Kane RA, Littrup PJ, Babaian RJ, Drago J, et al.: *Prostate-Specific Antigen Levels in 1,695 Men Without Evidence of Prostate Cancer.* Cancer 1992; 69:1201-1207.

3-4 Quellen und Suchstrategie bei der Literaturrecherche

Die genutzten Quellen - elektronische Datenbanken und Literaturverzeichnisse von berücksichtigten Publikationen - wurden lediglich bezüglich der Literaturrecherche zu Krankheitsprogression und krankheitsspezifischer Mortalität spezifiziert. Keine Angaben zur Suchstrategie.

3-5 Validitäts- bzw. Qualitätskriterien bei der Bewertung der Primärstudien

Keine Angaben.

3-6 Methoden der Bewertung von Relevanz und Validität bzw. Qualität der Primärstudien

Keine Angaben.

3-7 Methoden der Extraktion von Daten aus den Primärstudien

Bei der Ermittlung der Stadienverteilung von mittels sequentieller Tests entdeckten Karzinomen wurde von einer bedingten Unabhängigkeit der einzelnen Tests ausgegangen; die berechnete Karzinomentdeckungsrate wurde durch einen konstanten Korrekturfaktor an die Rate angeglichen, welche aufgrund der Daten zu Sensitivität und Spezifität (ACS-NPCDP-Studie) zu erwarten waren. Langfristige Komplikationen wurden nach dem Schweregrad stratifiziert und nach der Prävalenz vorbestehender Erkrankung in der Population adjustiert. Die Ermittlung der stadiumspezifischen Krankheitsprogressionsraten bzw. der stadiumspezifischen Mortalität in 5-Jahres-Abschnitten basierte auf der Annahme einer konstanten Ereignisrate; die ermittelten Ereignisraten wurden gegebenenfalls nach zensierten Daten adjustiert.

3-8 Anzahl berücksichtigter Primärstudien

Die Berechnung der Prävalenz in der Gesamtbevölkerung basierte auf drei Studien; zur Schätzung der subgruppenspezifische Prävalenz wurde jeweils eine Studie berücksichtigt. Für die Werte zur Sensitivität und Spezifität der Tests sowie die Ableitung der Stadienverteilung wurden zwei Publikationen (zu derselben Studie) herangezogen. Die Wahrscheinlichkeiten für Komplikationen basieren auf insgesamt 25 Publikationen. Für die Ermittlung von Krankheitsprogredienzraten und krankheitsspezifischer Mortalität wurde auf 23 Veröffentlichungen zurückgegriffen.

3-9 Methode der Synthese der gesundheitsbezogenen Parameter

Die altersspezifische Prävalenz von Prostatakarzinom wurde mittels einer auf den Daten der drei herangezogenen Studien basierenden Regressionsanalyse ermittelt. Die subgruppenspezifischen Prävalenzen sind Vielfache der ermittelten Krankheitsprävalenz in der Gesamtpopulation. Für die Krankheitsprogressionsraten bzw. die krankheitsspezifische Mortalität der Alternative "Kein Screening" wurden die Daten der besten Studie (Johansson et al., 1992¹⁸) anstelle der gewichteten durchschnittlichen Ereignisrate aller Studien gewählt. Die besten Schätzer für die Behandlungswirksamkeit wurden als Parameterschätzer in der Basisanalyse eingesetzt.

3-10 Untersuchung von Heterogenität der gesundheitsbezogenen Parameter

Nicht untersucht.

3-11 Ergebnisse der Synthese

(Werte für die Basisanalyse; in Klammern Werte im plausiblen Bereich)

Die Krankheitsprävalenz in der Gesamtpopulation ergab sich aus der Gleichung $P = e^{(\text{Alter} \times 0.098625 - 9.91734)}$ ($r^2 = 0.97$), für Männer mit krankheitsspezifischer Familienanamnese wurde sie auf das 2.13fache angesetzt, für Schwarze auf das 1.5fache. Ermittelte Werte für die Sensitivität von PSA, DRU und TRUS waren 0.67 (0.67-0.81), 0.50 (0.58-0.89) bzw. 0.81(0.77-0.91), für die Spezifität 0.97 (0.82-0.97), 0.94 (0.84-0.98) bzw. 0.84 (0.82-0.92). Die Wahrscheinlichkeiten für die wesentlichen Komplikationen nach Operation waren folgende: perioperativer Tod 0.011 (0-0.061); Impotenz bei vorheriger Potenz (Anteil vollständiger Impotenz bei 0.75) 0.44 bei <60jährigen, 0.65 bei 60-70jährigen und 1.0 bei >70jährigen (0.29-1.0); vollständige Inkontinenz 0.02 (0-0.10); partielle Inkontinenz 0.06 (0-0.32). Bei Strahlentherapie waren die Wahrscheinlichkeiten für die wesentlichen Komplikationen folgende: Tod 0.002 (0-0.02); Impotenz 0.40 (0.27-0.60); Inkontinenz 0.01 (0-0.01).

3-12 Effektmaße für die ökonomische Analyse

Qualitätsadjustierte Überlebenszeit (QALE bzw. QALYs). Die krankheitsstadiumspezifische, altersadjustierte Überlebenszeit wurde mittels des Markov-Kohorten-Verfahrens berechnet. Dabei wurden Gesundheitszustandsänderungen in Verbindung mit der lokalen und metastatischen Krankheitsprogression in Betracht gezogen; die Länge der - wahrscheinlich einjährigen - Zyklen wurde nicht spezifiziert. Anhand der Methode der

¹⁸ Johansson JE, Adami HO, Andersson SO, Bergstrom R, Holmberg L, Krusemo UB: *High 10-Year Survival Rate in Patients with Early, Untreated Prostatic Cancer*. JAMA 1992; 267:2191-2196.

zeitlichen Abwägung (time-trade-off) wurden Nutzwerte für chronische Gesundheitszustände in einer Gruppe von Urologen, Radioonkologen und Internisten (insgesamt 10 Personen) ermittelt; Lebensqualitätswerte für temporäre Zustände wurden nach der 0/1-Standardskala adjustiert. Werte für die Lebensqualität nach Harnstauung basieren auf Schätzungen der Autoren. Mit den Schätzern der Basisanalyse wurde für Patienten mit Stadium A1- bis B2-Tumoren ohne Behandlung ("Kein Screening") eine krankheitsspezifische 10-Jahres-Überlebensrate von 0.87 bis 0.92 und bei radikaler Prostatektomie von 0.89 bis 0.99 ermittelt. Der Gewinn an Lebenszeit nach radikaler Prostatektomie bewegte sich, in Abhängigkeit von Alter, klinischem Stadium und Behandlungswirksamkeit, im Bereich von -0.1 und +3.5 Jahren, der Gewinn an qualitätsadjustierter Überlebenszeit (QALE) zwischen -2.0 und +1.7 Jahren.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische ambulante und stationäre Versorgung unter Berücksichtigung von Kurzzeit- und Langzeitfolgen.

- Screening/Diagnostik: DRE, PSA, TRUS, TRUS + Biopsie, Staging
- Behandlung: primäre Behandlung, Kurzzeitkomplikationen der primären Behandlung, Langzeitbehandlung von Komplikationen (partielle/komplette Impotenz, partielle/komplette Inkontinenz, Urethraobstruktion)
- Nachsorge: posttherapeutisches Follow-Up (nach radikaler Prostatektomie: jährliche Konsultation beim Urologen, PSA), Behandlung eines Lokalrezidivs (PSA, TRUS, Biopsie, zwei Konsultationen), Nachsorge des Lokalrezidivs (PSA, Konsultation alle drei Monate), Diagnostik bei Verdacht auf Metastasierung (PSA, Knochenszintigramm, zwei Konsultationen), Nachsorge Metastasierung (Konsultationen, Blutabnahme, Medikation)

Intangible Kosten wurden bei der Ermittlung der Lebensqualität auf der Effektseite erfasst.

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte, soweit ersichtlich, nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Die sich aus der Entscheidungsanalyse ergebenden Mengen wurden nicht spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Selbstkostenrechnung. Die Preise pro Mengeneinheit wurden in Auszügen aufgeführt.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Zukünftige Kosten und Effekte wurden mit 5% diskontiert.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die durchschnittliche Lebenserwartung bei der Alternative "Kein Screening" betrug pro Screening-Teilnehmer im Alter von 50, 60 und 70 Jahren 27.8 Jahre, 19.8 Jahre bzw. 12.7 Jahre (ohne Diskontierung). Dem entsprach eine qualitätsadjustierte Lebenserwartung von 26.9 Jahren, 18.9 Jahren bzw. 11.7 Jahren (ohne Diskontierung). Beim Vergleich mit der Alternative "Kein Screening" betrug die durchschnittliche inkrementelle Lebenserwartung pro Screening-Teilnehmer in den gleichen Altersgruppen (in Klammern qualitätsadjustiert) bei der Screening-Strategie 1) 0.0 (-1.8) Tage, 0.1 (-4.1) Tage bzw. 0.2 (-7.1) Tage; bei der Screening-Strategie 2) 0.6 (-1.5) Tage, 1.1 (-3.7) Tage bzw. 1.6 (-7.2) Tage; bei der Screening-Strategie 3) 0.6 (-2.1) Tage, 1.1 (-5.2) bzw. 1.7 (-9.5) Tage; bei der Screening-Strategie 4) 0.8 (-2.9) Tage, 1.6 (-7.1) Tage bzw. 2.2 (-12.8) Tage (jeweils ohne Diskontierung). Der Gewinn erhöhte sich mit dem Alter aufgrund der höheren Karzinomentdeckungsrate bzw. -prävalenz.

6-2 Ermittelte Kosten

Bei allen Screening-Programmen überstiegen die Kosten der Früherkennung und -behandlung die durch Herabsetzung der lokalen und metastatischen Krankheitsprogression erzielten Kostenminderungen. Die Kosten der Alternative "Kein Screening" beliefen sich pro Screening-Teilnehmer im Alter von 50, 60 und 70 Jahren auf US-\$ 82, US-\$ 184 bzw. US-\$ 396 (ohne Diskontierung). Beim Vergleich mit der Alternative "Kein Screening" lagen die inkrementellen Kosten pro Screening-Teilnehmer in den gleichen Altersgruppen für die Screening-Strategie 1) bei US-\$ 64, US-\$ 149 bzw. US-\$ 371; für die Screening-Strategie 2) bei US-\$ 77, US-\$179 bzw. US-\$ 448; für die

Screening-Strategie 3) bei US-\$ 104, US-\$ 233 bzw. US-\$ 572; für die Screening-Strategie 4) bei US-\$ 253, US-\$ 423 bzw. US-\$ 874 (jeweils ohne Diskontierung).

6-3 Synthese von Kosten und Effekten

Von den Autoren wird ein Programm bei einem Kosten-Nutzwert-Quotienten von <US-\$ 100,000 pro QALY als attraktiv erachtet.

Im Vergleich zur Alternative "Kein Screening" betrug der inkrementelle Kosten-Effektivitäts-Quotient (Kosten pro gewonnenem Lebensjahr) der Früherkennung bei Männern im Alter von 50 Jahren US-\$ 113,000 bei der Screening-Strategie 2) und US-\$ 729,000 bei der Screening-Strategie 4); bei Männern im Alter von 60 Jahren US-\$ 127,000 bei der Screening-Strategie 2) und US-\$ 475,000 bei der Screening-Strategie 4); bei Männern im Alter von 70 Jahren US-\$ 189,000 bei der Screening-Strategie 2) und US-\$ 466,000 bei der Screening-Strategie 4) (jeweils mit Diskontierung von Kosten und Effekten). Die Screening-Strategien 1) und 3) erübrigten sich jeweils wegen erweiterter Dominanz. Da jedes Programm mit einem Verlust an QALYs verbunden war, wurden alle Screening-Strategien durch die Alternative "Kein Screening" dominiert. Unter der Annahme, dass der natürliche Verlauf von Prostatakarzinom und die Stadienverteilung unabhängig von Ethnie und Familienanamnese sind, wurde der erwartete Nutzen des Screenings für die untersuchten Subgruppen eine Funktion der Krankheitsprävalenz: der Gewinn an Lebenserwartung stand in Verbindung mit einem Verlust an QALE, womit alle Screening-Strategien auch bei den Subgruppen von der Alternative "Kein Screening" dominiert wurden.

7 Behandlung von Unsicherheiten

1-Weg-Sensitivitätsanalysen wurden für alle Hauptvariablen durchgeführt, wobei auch eine mögliche Adaptation an langfristige Komplikationen in Betracht gezogen wurde; auch eine Sensitivitätsanalyse bezüglich der Modellstruktur (Strahlentherapie anstelle von radikaler Prostatektomie und jährliches Screening) wurde durchgeführt. Als Werte für die Sensitivitätsanalyse wurden Vielfache der Werte der Basisanalyse gewählt; Wertebereiche und Ergebnisse der Sensitivitätsanalyse wurden in Auszügen für Männer im Alter von 60 Jahren spezifiziert. Als sensible Parameter erwiesen sich lediglich die Schätzer für die Behandlungswirksamkeit der radikalen Prostatektomie: bei Annahme des optimistischsten Szenarios (worst case / best case-Sensitivitätsanalyse) waren alle Screening-Strategien mit einem Gewinn an QALE von <3 Tagen pro Screening-Teilnehmer in allen Altersgruppen verbunden; für Männer im Alter von 60 Jahren betrug der inkrementelle Kosten-Effektivitäts-Quotient der Screening-Strategie 2) US-\$ 16,000

pro gewonnenem Lebensjahr (mit Diskontierung). Weiterhin erwies sich die Screening-Strategie 2) für Männer im Alter von 50 Jahren mit einem Kosten-Nutzwert-Quotienten von US-\$ 42,000 pro QALY (mit Diskontierung) als attraktives Programm. Kein weiteres Programm konnte in dieser Altersgruppe als attraktiv betrachtet werden; für Männer im Alter von ≥ 60 Jahren wurde kein Programm als attraktiv angesehen.

Zusätzliche Analysen bei vorbestehender Impotenz, geringer Komorbidität oder rascher Adaptation an behandlungsbedingte Komplikationen erbrachten unter den Annahmen der Basisanalyse QALY-Gewinne zwischen 0.8 und -7.0 Tagen in der Gesamtpopulation und zwischen 2.6 und -9.1 Tagen bei Schwarzen; alle Programme für alle Altersgruppen wurden entweder durch die Alternative "Kein Screening" dominiert oder ihr Kosten-Nutzwert-Quotient lag $> \text{US-}\$ 100,000$. Unter den Annahmen des optimistischen Szenarios ergaben diese Analysen bei der Screening-Strategie 2) für alle Männer einen Nettogewinn an QALYs und attraktive Kosten-Nutzwert-Quotienten.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren weisen darauf hin, dass die Evidenz der den Schätzern zugrundeliegenden Daten gering ist. Die Annahmen, dass die Stadienverteilung unabhängig von Alter und Ethnie sei, überschätzt ihrer Meinung nach den Nutzen des Screenings für Männer in höherem Alter bzw. für Schwarze. Weiterhin merken die Autoren an, dass in ihrer Analyse nicht alle Auswirkungen des Screenings und der Behandlung auf die Lebensqualität und Kosten berücksichtigt wurden, so z.B. psychologische Effekte der Krebsdiagnose. Die Nichtberücksichtigung psychologischer Effekte führt eher zu einer Überschätzung eines Screening-Effektes. Auch sind die Autoren der Ansicht, dass die Präferenzen der Patienten für die verschiedenen Gesundheitszustände von den ermittelten differieren können. Die Autoren sind weiterhin der Meinung, dass Entscheidungsanalysen und andere Modelle auch mit aus RCTs abgeleiteten Schätzern zur Behandlungseffektivität keine endgültige Vorhersage des Nutzens der Früherkennung von Prostatakarzinom erlauben: length bias und selection bias bringen einen Unsicherheitsfaktor in den Zusammenhang zwischen Karzinomentdeckung und Nutzen für den Patienten.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Keine Ausführungen.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass bei Männern im Alter zwischen 50 und 70 Jahren mittels der vier untersuchten Screening-Strategien eine geringe Senkung der krankheitsspezifischen Mortalität erzielt werden kann. Unter Berücksichtigung der Lebensqualität jedoch sind ihrer Meinung nach alle vier Früherkennungsprogramme insgesamt mit mehr Schaden als Nutzen verbunden und bringen höhere Kosten als die herkömmliche Diagnosepraxis. Selektive Früherkennung bei Gruppen mit hoher Krankheitsprävalenz bringt nach Ansicht der Autoren keine Verbesserung der Kosten-Effektivität.

9 Kommentar

Entscheidungsanalytische Studie zum Prävalenz-Screening mit expliziter Gegenüberstellung von Technologiealternativen. Das von den Autoren beschriebene Modell erscheint sorgfältig aufgebaut, transparent und weitgehend nachvollziehbar. Allerdings muss darauf hingewiesen werden, dass die Parameterschätzer der Basisanalyse nicht aus einer umfassenden Literaturrecherche und -analyse abgeleitet wurden; hier ist die Möglichkeit eines selection bias bezüglich der in die Analyse eingehenden Resultate des Reviews gegeben. Auch fehlen präzisierende Angaben zu den berücksichtigten Untersuchungen sowie zu den Kriterien für die Anwendung des Goldstandards in den Studienpopulationen, auf deren Daten die Ermittlung der Sensitivität und Spezifität der Tests basiert (die ACS-NPCDP-Studie ist eine Kohortenstudie mit mehrmaliger Untersuchungswiederholung (Prävalenz- und Inzidenz-Screening), in welcher eine Biopsie lediglich positivem DRU- und/oder TRUS-Befund nicht aber bei alleiniger Erhöhung des PSA-Wertes durchgeführt wurde (work-up bias)). Die implizite Annahme, dass für negative Screening-Testbefunde eine gleiche Karzinomwahrscheinlichkeit und -stadienverteilung besteht wie im Falle der Alternative "Kein Screening", erscheint problematisch. Krahn et al. beziehen sich in ihrer Studie auf das Staging-System nach Whitmore-Jewett; in der ACS-NPCDP-Studie wurde allerdings ein Modifikation dieser Klassifikation angewandt, so dass hier Diskrepanzen entstehen können. Dem natürlichen Verlauf von Prostatakarzinom bei Nichtbehandlung liegen Daten zugrunde, welche einer Studienpopulation mit einem durchschnittlichen Alter von 72 Jahren entstammen; diese Datenquelle erscheint für Männer im Alter von 50 bis 70 Jahren als eher ungeeignet. Einige dieser Problempunkte wurden durch die eingehende Sensitivitätsanalyse aufgegriffen. Der Dokumentation hinsichtlich des Mengengerüsts sowie der Ermittlung des dazugehörigen Preisgerüsts mangelt es an Transparenz, die Kostenberechnungen sind

nicht nachvollziehbar. U.a. ist unklar, was terminale Versorgungskosten beinhalten und welchen Zeitraum sie abdecken.

10 Technischer Report

Ein detaillierter technischer Report wurde angefordert und in der Evaluation berücksichtigt.

**Kramer BS, Brown ML, Prorok PC, Potosky AL, Gohagan JK:
Prostate Cancer Screening: What We Know and What We Need to Know.
Annals of Internal Medicine, 1993; 119(9): 914-923.**

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels DRU und Bestimmung von Serum-PSA. Kein expliziter Vergleich von Technologiealternativen; implizite Gegenüberstellung der Alternativen "Screening" und "Kein Screening".

1-2 Fragestellung

Evaluation der wissenschaftlichen Erkenntnis hinsichtlich der auf der PSA-Bestimmung basierenden Früherkennung von Prostatakarzinom bei asymptomatischen Männern. Beschreibende Aufstellung der im ersten Jahr eines Massen-Screening-Programmes zu erwartenden Kosten.

Screening-Strategie: DRU+PSA $-\{ PSA \geq 4 \text{ng/ml} \} \rightarrow$ TRUS+Biopsie

Inwieweit die DRU die Entscheidung für eine TRUS bzw. Biopsie mitbestimmt ist nicht präzisiert; angenommen werden kann, dass auf einen positiven DRU-Befund hin eine diagnostische Abklärung erfolgt. Auch wurde unklar formuliert, ob ein positiver Ultraschallbefund als Entscheidungskriterium für eine Biopsie gilt oder ob eine Biopsie unabhängig vom TRUS-Befund durchgeführt wird.

1-3 Perspektive

Keine expliziten Angaben.

1-4 Zeithorizont

Ermittelt wurden die Effekte und Kosten des ersten Jahres des Screening-Programmes (Prävalenz-Screening). Teils wurde auf Daten Bezug genommen, denen ein Zeithorizont von 1 Jahr zugrunde liegt.

1-5 Art der ökonomischen Evaluation

Kosten-Konsequenzen-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Narrativer Review mit zusätzlicher Kostenaufstellung.

Datenquellen für klinische Parameter: klinische Entscheidungsanalyse mit Literaturdaten.

Datenquelle für Ressourcenverbrauch und Mengengerüst: Modellierung (Entscheidungsanalyse).

Datenquelle für Preise: Publikation von Optenberg & Thompson (1990)¹⁹ ohne weitere Explizierung.

2-2 Datierung der zugrundeliegenden Daten

Der Schätzwert für die Größe der männlichen Bevölkerung der Vereinigten Staaten in der untersuchten Altersgruppe basierte auf Zensusdaten des Jahres 1990. Zeitpunkt und Zeitraum der Erhebung von Daten zu klinischen Parametern wurden nicht explizit aufgeführt (die klinischen Parameter leiten sich aus Daten ab, welche bis 1993 veröffentlicht wurden).

Keine Angaben hinsichtlich des Bezugsjahres für die Preisangaben.

2-3 Studienpopulation

Männer der amerikanischen Gesamtbevölkerung im Alter von 50 bis 74 Jahren ohne ernste kardiovaskuläre Vorerkrankung. Eine Subgruppenanalyse bezog sich auf Männer im Alter von 50 bis 64 Jahre.

2-4 Setting

Tangierte Versorgungsstrukturen: keine Angaben.

Keine explizite Lokalisation der Daten zu den klinischen Parametern. Der ökonomische Teil der Studie wurde in Maryland, USA durchgeführt.

2-5 Spezifikation der Technologie

Keine Präzisierungen bezüglich des PSA-Assays, der DRU, der TRUS oder der Biopsie.

¹⁹ Optenberg SA & Thompson IM: *Economics of Screening for Carcinoma of the Prostate*. Urol Clin North Am 1990; 17(4): 719-737.

3 Gesundheitseffekte

Synthese bereits veröffentlichter Studien

3-1 In der Synthese untersuchte klinische Parameter

Primäre Zielgrößen waren die Anzahl Männer im Alter von 50 bis 74 Jahren in den Vereinigten Staaten, der Anteil Männer an der männlichen Bevölkerung mit kardiovaskulärer Vorerkrankung als Ausschlusskriterium, die altersspezifische Entdeckungsrate von Prostatakarzinom, die klinische Stadienverteilung der entdeckten Karzinome, der Anteil an Patienten mit Stadium A- oder B-Karzinom und operativer Behandlung bzw. Strahlentherapie sowie die Wahrscheinlichkeiten von behandlungsbedingten Komplikationen und sekundären Therapien.

3-2 Annahmen

Die Analyse basierte auf folgenden expliziten Annahmen: Die Biopsie ist eine Nadelbiopsie. Als Staging-Methoden werden serologische Tests, eine Ausscheidungsurographie, eine Zystoskopie sowie ein Knochenszintigramm eingesetzt. Patienten mit einem klinischen Stadium A- oder B-Karzinom unterziehen sich einer radikalen Prostatektomie oder einer Strahlentherapie, Patienten mit klinischem Stadium C-Karzinom einer Strahlentherapie und Patienten mit klinischem Stadium D-Karzinom einer hormonellen Therapie (Orchiektomie oder Diethylstilböstrol). Als Komplikationen der radikalen Prostatektomie wurden berücksichtigt (in Klammern therapeutische Maßnahmen): Impotenz (penile Prothese); rektale Verletzung; Inkontinenz; Harnröhrenstriktur; postoperativer Tod. Als Komplikationen der Strahlentherapie wurden berücksichtigt (in Klammern therapeutische Maßnahmen): Impotenz (penile Prothese); rektale Verletzung; Inkontinenz; Harnröhrenstriktur; behandlungsbedingter Tod.

Die Analyse orientiert sich in Teilbereichen an der Publikation von Optenberg & Thompson. Daher ist von folgenden impliziten Annahmen auszugehen: Therapeutische Maßnahme bei rektaler Verletzung ist eine Kolostomie, bei Inkontinenz das Legen eines künstlichen Sphinkters, bei Harnröhrenstriktur eine Urethrotomia interna. Bei der hormonellen Therapie wurden keine Komplikationen erwartet.

Implizit wurde die Karzinomentdeckungsrate und behandlungsbedingte Mortalität im Rahmen des Früherkennungsprogrammes der Prostatakarzinom-Inzidenz und behandlungsbedingten Mortalität, welche auf Basis von SEER-Daten (Surveillance, Epidemiology, and End Results Program) zu erwarten waren, gegenübergestellt.

Außerdem wird implizit von einer 100%igen Compliance mit der Früherkennung ausgegangen.

3-3 Berücksichtigung von Primärstudien: Studiendesigns und Ein-/Ausschlusskriterien

Die Anzahl Männer im Alter von 50 bis 74 Jahren in den Vereinigten Staaten basierte auf Zensusdaten. Auf welchen Daten der Schätzwert für den Anteil der Männer mit kardiovaskulärer Vorerkrankung basiert ist unklar. Der Anteil Patienten mit Stadium A- oder B-Karzinom wurde aus prospektiven Kohortenstudien abgeleitet. Zur Schätzung des Anteils an Patienten mit Stadium A- oder B-Karzinom, welche sich einer operativen Behandlung unterziehen, sowie zur Ermittlung der Mortalität bis vier Monate nach einer radikalen Prostatektomie wurden SEER-Daten herangezogen. Die Wahrscheinlichkeiten für die berücksichtigten behandlungsbedingten Komplikationen und sekundären Therapien wurden der Studie von Optenberg & Thompson entnommen. Das Design der Studien, aus denen die Schätzer zu den interessierenden klinischen Parametern abgeleitet wurde, wurde nicht spezifiziert. Keine Angaben zu Ein- bzw. Ausschlusskriterien bei der Berücksichtigung der Studien.

3-4 Quellen und Suchstrategie bei der Literaturrecherche

Die genutzten Quellen - MEDLINE, Dokumentensammlungen der Autoren und Literaturverzeichnisse von aufgefundenen Publikationen - wurden im Abstract des Reviews erwähnt. Keine Angaben zur Suchstrategie.

3-5 Validitäts- bzw. Qualitätskriterien bei der Bewertung der Primärstudien

Keine Angaben.

3-6 Methoden der Bewertung von Relevanz und Validität bzw. Qualität der Primärstudien

Keine Angaben.

3-7 Methoden der Extraktion von Daten aus den Primärstudien

Keine Angaben.

3-8 Anzahl berücksichtigter Primärstudien

Die Annahme zum Anteil an Patienten mit (lokal begrenztem) Stadium A- oder B-Karzinom basierte auf zwei Studien; für die Schätzwerte der anderen interessierenden

klinischen Parameter wurde, soweit ersichtlich, auf jeweils eine Veröffentlichung zurückgegriffen.

3-9 Methode der Synthese der gesundheitsbezogenen Parameter

Keine Angaben.

3-10 Untersuchung von Heterogenität der gesundheitsbezogenen Parameter

Nicht untersucht.

3-11 Ergebnisse der Synthese

(Werte für die Basisanalyse; in Klammern Werte im plausiblen Bereich)

Die Anzahl Männer im Alter von 50 bis 74 Jahren in den Vereinigten Staaten betrug im Jahre 1990 23.8 Mio. Nach Ausschluss des auf 11% angesetzten Anteils an Männern mit kardiovaskulärer Vorerkrankung wurde von einer Screening-Population mit 21,187,000 Männern ausgegangen. Die Schätzer für die linear ansteigenden altersspezifischen Karzinomentdeckungsraten reichten von 1.1% in der Altersgruppe der 50- bis 54jährigen bis 5.5% in der Altersgruppe der 70- bis 74jährigen, mit einer gewichteten durchschnittlichen Entdeckungsrate von 3.1%. 95% der diagnostizierten Karzinome wurden als lokal begrenzte Stadium A- oder B-Karzinome angesetzt, jeweils 2.5% der Tumore als Stadium C- bzw. D-Karzinome. Es wurde davon ausgegangen, dass sich bei den Patienten mit einem Stadium A- oder B-Karzinom 67% einer radikalen Prostatektomie und 33% einer Strahlentherapie unterziehen würden. Die für Komplikationen und komplikationsbedingte sekundäre Therapien angesetzten Raten bei radikaler Prostatektomie waren folgende: Impotenz 0.25 bzw. 0.0625; rektale Verletzung 0.03 bzw. 0.03; Inkontinenz 0.06 bzw. 0.03; Harnröhrenstriktur 0.18 bzw. 0.18; postoperativer Tod 0.012. Die Raten für Komplikationen und komplikationsbedingte sekundäre Therapien bei Strahlentherapie waren: Impotenz 0.25 bzw. 0.0625; rektale Verletzung 0.01 bzw. 0.01; Inkontinenz 0.03 bzw. 0.015; Harnröhrenstriktur 0.08 bzw. 0.08; behandlungsbedingter Tod 0.005.

3-12 Effektmaße für die ökonomische Analyse

Den Kosten wurden die identifizierten Prostatakarzinome sowie die Anzahl behandlungsbedingter Todesfälle gegenübergestellt (ohne sie in Relation zu setzen: Kosten-Konsequenzen-Analyse).

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

- Screening/Diagnostik: Screening-Untersuchung, TRUS, Biopsie, Staging
- Behandlung: primäre und sekundäre Behandlung

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Das Mengengerüst wurde nicht hinreichend spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Administrative Preise. Die Preise pro Mengeneinheit wurden aufgeführt.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die Anzahl der mittels des Früherkennungsprogrammes diagnostizierten Karzinome wurde auf 654,305 berechnet, 591,829 (bzw. 10.5mal) mehr als auf Basis der SEER-Daten (62,476) zu erwarten gewesen wären. Die behandlungsbedingten Todesfälle wurden auf zusätzlich 5,510 Fälle ermittelt. Die Anzahl der Patienten, welche eine komplikationsbedingte sekundäre Therapie benötigten, wurde auf mehr als 250,000 geschätzt. In der Altersgruppe der 50- bis 64jährigen war im Vergleich zur Altersgruppe der 50- bis 74jährigen mit einer um 47% niedriger liegenden Karzinomentdeckungsrate zu rechnen; die Anzahl zusätzlicher behandlungsbedingter Todesfälle wurde bei den 50-64-Jährigen auf 2,293 berechnet.

6-2 Ermittelte Kosten

Die Gesamtkosten für Früherkennung und Behandlung betragen US-\$ 11.9 Mrd.; die Kosten für die Screening-Tests einschließlich der TRUS beliefen sich dabei auf US-\$ 1.6 Mrd., diejenigen für die Biopsie auf US-\$ 2.2 Mrd. und diejenigen für die primäre Behandlung auf US-\$ 6.0 Mrd. beliefen. Ohne Früherkennungsprogramm betragen die berechneten Behandlungskosten US-\$ 1.4 Mrd. Für Männer im Alter von 50 bis 64 Jahren wurden Gesamtkosten für das Früherkennungsprogramm von US-\$ 7.0 Mrd. ermittelt.

6-3 Synthese von Kosten und Effekten

(Nicht zutreffend: Kosten-Konsequenzen-Analyse.)

7 Behandlung von Unsicherheiten

Bei Annahme einer PSA-Entscheidungsgrenze von ≥ 3 ng/ml und einer um 13% über der Basisanalyse liegenden Karzinomentdeckungsrate wurden 6,241 zusätzliche behandlungsbedingte Todesfälle zu Kosten von US-\$ 14.1 Mrd. erwartet. Bei Annahme einer PSA-Entscheidungsgrenze von ≥ 10 ng/ml und einer um 62% unter der Basisanalyse liegenden Karzinomentdeckungsrate waren dies 2,087 zusätzliche behandlungsbedingte Todesfälle zu US-\$ 5.5 Mrd.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren weisen darauf hin, dass die bisher vorliegenden wissenschaftlichen Daten keine genaue Ermittlung von Indizes bezüglich der Testgüte erlauben, da die Biopsie als Goldstandard nie bei allen Studienteilnehmern angewendet wurde. Auch lag für ihren Review keine kontrollierte klinische Studie vor, so dass sie auf potentiellen selection bias, lead time bias und/oder length bias einschließlich overdiagnosis bei den bislang vorliegenden Daten hinweisen.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Die Autoren bemerken, dass es unmöglich ist, die genauen Kosten eines Massen-Screening-Programmes zu ermitteln, solange keine Daten einer kontrollierten klinischen

Studie vorliegen. Lediglich deskriptive Schätzungen zur Größenordnung der Kosten im 1. Jahr wurden angestellt.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass die Früherkennung von Prostatakarzinom potentiell Leben retten kann; wegen der Möglichkeit von overdiagnosis können sich das Screening und die therapeutischen Konsequenzen jedoch auch ungünstig im Hinblick auf die Lebensqualität und die Mortalität auswirken. Ein ausgedehntes Früherkennungsprogramm könnte mit Kosten von mehreren Milliarden verbunden sein, wobei der Nettonutzen unklar bleibt. Angesichts der Bedeutung des Themas und der damit verbundenen kontroversen Diskussion werden prospektive Trials von den Autoren befürwortet. Ähnliche Kriterien wie bei der Zulassung von Arzneimitteln werden gefordert. Die Autoren, wie auch das National Cancer Institute, dem sie angehören, empfehlen deshalb, bis zum Vorliegen besserer wissenschaftlicher Erkenntnisse vom Einsatz der PSA-Bestimmung als routinemäßiges Screening-Verfahren abzusehen.

9 Kommentar

Qualitative Informationssynthese zum PSA-Screening mit Kostenrechnung für das Prävalenz-Screening in impliziter Anlehnung an die Entscheidungsanalyse von Optenberg & Thompson; dabei implizite Gegenüberstellung von Technologiealternativen. Hinsichtlich der Ermittlung der Effektivität weist der Review in vielen Punkten der systematischen Literaturrecherche und -analyse bzw. deren Dokumentation Defizite auf. Es besteht das Potential eines selection bias bezüglich der den Kosten gegenübergestellten Screening-Konsequenzen. Der Ermittlung von Kosten mangelt es an Transparenz und Nachvollziehbarkeit. Ein Teil der Preise aus der Vorlage von Optenberg & Thompson scheint über die Zeit angepasst worden zu sein, ein anderer nicht, so dass hier zeitliche Unstimmigkeiten und Verzerrungen aufgetreten sein können. Die interne Validität, Reliabilität und Generalisierbarkeit der Ergebnisse erscheinen aufgrund der methodischen Defizite stark eingeschränkt.

Launois R:**Cost-Effectiveness Analysis of Strategies for Screening Prostatic Cancer.**

In: Zweifel P, Frech III HE (eds.): Health Economics Worldwide [auch: Developments in Health Economics and Public Policy, Vol. I]. Kluwer Academic Publishers, Dordrecht-Boston-London, 1992: 81-108.

1 Fragestellung und Evaluationsrahmen**1-1 Technologie**

Screening des Prostatakarzinoms mittels DRU, Bestimmung von Serum-PSA und TRUS. Explizite Gegenüberstellung von verschiedenen Früherkennungsprogrammen.

1-2 Fragestellung

Klinische Entscheidungsanalyse zur Ermittlung der Kosten-Effektivität von sechs Früherkennungsprogrammen in der Allgemeinbevölkerung.

Strategie 1): $DRU^{AM} -[\oplus] \rightarrow DRU^U -[\oplus] \rightarrow US -[\oplus] \rightarrow$ Biopsie *und*
 $US -[\emptyset] \rightarrow PSA -[\oplus] \rightarrow$ Biopsie

Strategie 2): $DRU^{AM} -[\oplus] \rightarrow DRU^U -[\oplus] \rightarrow US -[\oplus \text{ oder } \emptyset] \rightarrow$ Biopsie *und*
 $DRU^U -[\emptyset] \rightarrow US -[\oplus] \rightarrow$ Biopsie

Strategie 3): $DRU^{AM} -[\oplus] \rightarrow DRU^U -[\oplus] \rightarrow PSA -[\oplus \text{ oder } \emptyset] \rightarrow$ Biopsie *und*
 $DRU^U -[\emptyset] \rightarrow PSA -[\oplus] \rightarrow$ Biopsie

Strategie 4): $PSA -[\oplus] \rightarrow DRU^U -[\oplus] \rightarrow US -[\oplus] \rightarrow$ Biopsie *und*
 $US -[\emptyset] \rightarrow PSA -[\oplus] \rightarrow$ Biopsie

Strategie 5): $DRU^{AM} + PSA -[\oplus] \rightarrow DRU^U -[\oplus] \rightarrow US -[\oplus] \rightarrow$ Biopsie *und*
 $US -[\emptyset] \rightarrow PSA -[\oplus] \rightarrow$ Biopsie

Strategie 6): $PSA -[\oplus] \rightarrow DRU^U -[\oplus \text{ oder } \emptyset] \rightarrow$ Biopsie

AM = Allgemeinmediziner, U = Urologe

1-3 Perspektive

Perspektive des Kostenträgers (nationale Krankenversicherung).

1-4 Zeithorizont

Ermittelt wurden die Effekte und Kosten in Verbindung mit dem Prävalenz-Screening. U.a. wurde die Restlebenserwartung in die Analyse einbezogen.

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Synthese.

Datenquellen für klinische Parameter: klinische Entscheidungsanalyse mit Literaturdaten und Modellierung (DEALE).

Datenquelle für Ressourcenverbrauch und Mengengerüst: Modellierung (Entscheidungsanalyse).

Datenquelle für Preise: keine Angaben.

2-2 Datierung der zugrundeliegenden Daten

Zeitpunkt und Zeitraum der Erhebung von Daten zu klinischen Parametern wurden nicht explizit aufgeführt. Die klinischen Parameter leiten sich aus Studien ab, welche bis 1990 veröffentlicht wurden.

Keine Angaben hinsichtlich des Bezugsjahres für die Preisangaben.

2-3 Studienpopulation / Zielpopulation

Männer, keine expliziten Altersangaben. Es wurden die Ergebnisse für Männer ab einem Alter von 50 Jahren berichtet.

2-4 Setting

Keine Angaben bezüglich der tangierten Versorgungsstruktur(en).

Die Daten zu den klinischen Parametern stammen explizit aus Frankreich, Schweden und Nordamerika. Der ökonomische Teil der Studie wurde in Frankreich durchgeführt

2-5 Spezifikation der Technologie

Bei der Aufstellung der Screening-Strategien wurde lediglich eine Unterscheidung zwischen der Durchführung der DRU durch Allgemeinmediziner und derjenigen durch Urologen gemacht. Die Entscheidungsgrenze bei der PSA-Bestimmung wurde als >5ng/ml oder >10ng/ml spezifiziert und bleibt damit im konkreten Fall unklar. Die Art der Sonographieuntersuchung wie auch die der Biopsie wurde nicht spezifiziert.

3 Gesundheitseffekte

Synthese bereits veröffentlichter Studien

3-1 In der Synthese untersuchte klinische Parameter

Primäre Zielgrößen waren die Prävalenz von Prostatakarzinom, die Sensitivität und Spezifität der Tests, die Stadienverteilung der entdeckten Karzinome sowie die krankheitsstadiumspezifische Mortalität.

3-2 Annahmen

Die Base-Case-Analyse basierte auf der expliziten Annahme, dass die Biopsie als Goldstandard ein perfekter Test mit einer Sensitivität und Spezifität von jeweils 1.0 ist. Die Effekte und Kosten für die implizite Alternative "Kein Screening" wurden gleich Null gesetzt.

Implizite Annahme war, dass für die verschiedenen Screening-Strategien der Anteil potentiell heilbarer Karzinome an allen entdeckten Karzinomen gleich ist.

3-3 Berücksichtigung von Primärstudien: Studiendesigns und Ein-/Ausschlusskriterien

Die für die Basisanalyse gewählte klinische Krankheitsprävalenz basiert auf den Daten einer französischen Fallserie. Die bedingt abhängige Sensitivität und Spezifität der Tests wurde aus französischen Fallserien abgeleitet; in einer dieser Studien war die Biopsie als Goldstandard bei der gesamten Studienpopulation angewandt worden, die andere war mit einem work-up bias behaftet. Für die Stadienverteilung der entdeckten Karzinome wie auch für die krankheitsstadiumspezifische Mortalität bei Nichtbehandlung wurden Daten der prospektiven Kohortenstudie von Johansson et al. (1989)²⁰ herangezogen; die krankheitsstadiumspezifische Mortalität nach radikaler Prostatektomie basiert auf Daten

²⁰ Johansson JE, Adami HO, Andersson SO, Bergstrom R, Krusemo UB, Kraaz W: *Natural History of Localized Prostatic Cancers. A Population Based Study in 223 Untreated patients.* Lancet 1989; i: 799-803.

einer retrospektiven Kohortenstudie aus USA. Keine Angaben zu Ein- bzw. Ausschlusskriterien bei der Berücksichtigung der Studien.

3-4 Quellen und Suchstrategie bei der Literaturrecherche

Keine Angaben.

3-5 Validitäts- bzw. Qualitätskriterien bei der Bewertung der Primärstudien

Keine Angaben.

3-6 Methoden der Bewertung von Relevanz und Validität bzw. Qualität der Primärstudien

Keine Angaben.

3-7 Methoden der Extraktion von Daten aus den Primärstudien

Die klinische Krankheitsprävalenz basierte auf den zu diesem Zeitpunkt aktuellsten Studienergebnissen für Frankreich. Die Karzinomstadien A2, B1 und hochdifferenziertes B2 aus der Studie von Johansson et al. wurden zur hauptsächlichen Zielgruppe für das Screening des Prostatakarzinoms und dessen Behandlung mittels radikaler Prostatektomie zusammengefasst, für welche auch die altersspezifischen Überlebenszeiten berechnet wurden. Die Sensitivität und Spezifität von multiplen Tests wurde durch Aufstellung von Vierfeldertafeln mit den Daten der entsprechenden Studien abgeleitet; für die US nach einer DRU (positiv oder negativ) wurden bedingt abhängige Maßzahlen für die Testgüte ermittelt.

3-8 Anzahl berücksichtigter Primärstudien

Jeweils eine Publikation wurde herangezogen, um Schätzungen für die interessierenden klinischen Parameter abzuleiten. Hierbei wurden insgesamt fünf Publikationen berücksichtigt.

3-9 Methode der Synthese der gesundheitsbezogenen Parameter

Keine Angaben.

3-10 Untersuchung von Heterogenität der gesundheitsbezogenen Parameter

Nicht untersucht.

3-11 Ergebnisse der Synthese

(Werte für die Basisanalyse; in Klammern Werte im plausiblen Bereich)

Die Krankheitsprävalenz wurde auf 0.03 (50-59jährige 0.0013-0.29, 60-69jährige 0.009-0.3, ≥ 70 jährige 0.036-0.4) angesetzt. Die Sensitivität und Spezifität der Tests war folgende: DRU^{AM} 0.428 (0.100) bzw. 0.756 (0.950); PSA(1. Runde) 0.947 (0.676) bzw. 0.432 (0.889); PSA (2. Runde) 0.676 bzw. 0.889; DRU^{AM}+PSA 0.971 (0.708) bzw. 0.327 (0.847); DRU^U 0.710 bzw. 0.859; US | DRU[⊕] 0.966 bzw. 0.267; US | DRE[⊗] 0.335 bzw. 0.628. Der Anteil Karzinome der Stadien A2, B1 und hochdifferenziertes B2 an den entdeckten Prostatatumoren wurde auf 0.173 geschätzt.

3-12 Effektmaße für die ökonomische Analyse

Anzahl identifizierter Prostatakarzinome pro 100,000 Screening-Teilnehmer; altersspezifische Überlebenszeit. Die altersspezifische Überlebenszeit wurde mittels der DEALE-Methode ermittelt.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

- Screening/Diagnostik: Konsultation, PSA-Assay, TRUS und Biopsie einschließlich Histopathologie sowie Antibiotikaprophylaxe.

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Die sich aus der Entscheidungsanalyse ergebenden Mengen wurden spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Administrative Preise. Die Preise pro Mengeneinheit wurden aufgeführt.

4-4 Währung

Französische Franken (FF).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Pro 100,000 Screening-Teilnehmer wurden bei der Screening-Strategie 1) 3,447 Biopsien durchgeführt: in 902 Fällen wurde ein Karzinom diagnostiziert, 2,545 Fälle erwiesen sich als falschpositiv. Bei der Screening-Strategie 2) waren dies 11,937 Biopsien mit 1,037 richtigpositiven bzw. 10,900 falschpositiven Befunden; bei der Screening-Strategie 3) 16,150 Biopsien mit 1,264 richtigpositiven bzw. 14,886 falschpositiven Befunden; bei der Screening-Strategie 4) 7,919 Biopsien mit 1,995 richtigpositiven bzw. 5,924 falschpositiven Befunden; bei der Screening-Strategie 5) 9,066 Biopsien mit 2,046 richtigpositiven bzw. 7,020 falschpositiven Befunden; bei der Screening-Strategie 6) 57,937 Biopsien mit 2,841 richtigpositiven bzw. 55,096 falschpositiven Befunden. Die im Vergleich zur Nichtbehandlung gewonnene Überlebenszeit pro entdecktem Karzinom betrug im Alter von 50 Jahren 0.18 Jahre, im Alter von 65 Jahren 0.09 Jahre und im Alter von 75 Jahren 0.04 Jahre und drückte das kurative Potential der radikalen Prostatektomie aus.

6-2 Ermittelte Kosten

Die Gesamtkosten pro 100,000 Screening-Teilnehmer beliefen sich bei der Screening-Strategie 1) auf FF 16,560,825, bei der Screening-Strategie 2) auf FF 32,025,429, bei der Screening-Strategie 3) auf FF 30,143,646, bei der Screening-Strategie 4) auf FF 39,407,034, bei der Screening-Strategie 5) auf FF 42,369,276 und bei der Screening-Strategie 6) auf FF 83,487,834.

6-3 Synthese von Kosten und Effekten

Die durchschnittlichen Kosten pro entdecktem Karzinom betragen bei der Screening-Strategie 1) FF 18,360, bei der Screening-Strategie 2) FF 30,883, bei der Screening-Strategie 3) FF 23,848, bei der Screening-Strategie 4) FF 19,753, bei der Screening-Strategie 5) FF 20,708 und bei der Screening-Strategie 6) FF 29,387. Die Screening-Strategie 2) wurde durch die Screening-Strategie 3) dominiert. Die inkrementellen Kosten pro zusätzlich entdecktem Karzinom betragen bei der Screening-Strategie 1) im Vergleich zur Alternative "Kein Screening" FF 18,360, bei der Screening-Strategie 4) im Vergleich zur Screening-Strategie 2) FF 20,902 und bei der Screening-Strategie 6) im Vergleich zur Screening-Strategie 4) FF 52,104; bei den Screening-Strategien 3) und 5) bestand erweiterte Dominanz durch die Screening-Strategien 4) bzw. 6). Beim Vergleich mit der Alternative "Kein Screening" wurden die inkrementellen Kosten pro zusätzlich gewonnenem Lebensjahr für die Screening-Strategie 1) auf FF 102,230, für die

Screening-Strategie 4) auf FF 115,970 und für die Screening-Strategie 6) auf FF 288,110 berechnet (ohne Spezifikation der Altersgruppe: bezieht sich eigenen Berechnungen nach auf die Gruppe der 50jährigen).

7 Behandlung von Unsicherheiten

Die 1-Weg-Sensitivitätsanalyse bezüglich der altersspezifischen Krankheitsprävalenz hatte Auswirkungen auf die Höhe der (inkrementellen) Kosten-Effektivitäts-Quotienten: mit Zunahme der Krankheitsprävalenz im plausiblen Wertebereich verminderte sich der (inkrementelle) Kosten-Effektivitäts-Quotient. Dies änderte jedoch nichts an der relativen Rangfolge der einzelnen Screening-Strategien. Adjustierung der Werte für die Sensitivität und Spezifität von Tests in den Studien, in welchen die Biopsie als Goldstandard nicht bei allen Probanden angewendet worden war (work-up bias), veränderte die relative Rangfolge der einzelnen Screening-Strategien: die Screening-Strategie 6) dominierte alle anderen Strategien in bezug auf die Kosten-Effektivität der Karzinomentdeckung wie auch in bezug auf die Kosten-Effektivität des Lebenszeitgewinns (50-59jährige bei einer Prävalenz von 0.085: FF 122,600 pro gewonnenem Lebensjahr; 60-69jährige bei einer Prävalenz von 0.11: FF 196,500 pro gewonnenem Lebensjahr; ≥ 70 jährige bei einer Prävalenz von 0.24: FF 243,500 pro gewonnenem Lebensjahr).

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Keine Ausführungen.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Keine Ausführungen.

8-3 Schlussfolgerungen

Die Autor kommt zum Schluss, dass bei hoher Krankheitsprävalenz auch invasive Screening-Programme in Erwägung gezogen werden sollten: direkte Biopsien bei verdächtigen PSA-Werten scheinen seiner Ansicht nach effizienter zu sein, da die höheren Kosten durch eine geringere Rate an falschpositiven Screening-Befunden wettgemacht werden. Letztlich hat der Entscheidungsträger unter Berücksichtigung der kollektiven Zahlungsbereitschaft festzulegen, welche Technologien zu implementieren sind.

9 Kommentar

Entscheidungsanalyse zu Biopsiekriterien beim Prävalenz-Screening mit expliziter Gegenüberstellung von Technologiealternativen. Das vom Autor beschriebene theoretische Modell zur Ermittlung der intermediären Kosten-Effektivität (Kosten pro entdecktem Karzinom) der Früherkennung von Prostatakarzinom erscheint transparent und nachvollziehbar, während das theoretische Modell zur Ermittlung der finalen Kosten-Effektivität (Kosten pro gewonnenem Lebensjahr) ein Defizit an Transparenz und Nachvollziehbarkeit aufweist. Die Studienqualität wird dadurch geschmälert, dass Entscheidungskriterien bezüglich der Interpretation des PSA-Assays sowie der DRU- und US-Befunde weder für das Modell noch für die der Ermittlung der Sensitivität und Spezifität der Tests zugrunde liegenden Datenquellen präzisiert wurden. Ein PSA-Wert von 5ng/ml als Entscheidungsgrenze ist zudem eher ungewöhnlich und wäre zu begründen. Als problematisch erscheint die (implizite) Annahme einer gleichen Stadienverteilung der entdeckten Karzinome für die verschiedenen Screening-Strategien. Bezüglich der Studienpopulationen, deren Daten herangezogen wurde, wurden keine näheren Angaben gemacht. Die bedingt abhängige Sensitivität und Spezifität wurde zwar für die US nach einer DRU ermittelt, für die anderen sequentiellen Testkombinationen jedoch nicht in Betracht gezogen. Die Parameterschätzer der Basisanalyse wurden nicht aus einer umfassenden Literaturrecherche abgeleitet; hier ist die Möglichkeit eines selection bias bezüglich der berücksichtigten Datenquellen gegeben. Dem natürlichen Verlauf von Prostatakarzinom bei Nichtbehandlung liegen Daten zugrunde, welche einer Studienpopulation mit einem durchschnittlichen Alter von 72 Jahren entstammen; diese Datenquelle erscheint für Männer im Alter von 50 bis 70 Jahren als ungeeignet. Die Daten, auf denen die Ermittlung der Sensitivität und Spezifität der Tests basiert, wurden aus Fallserien abgeleitet und können damit für die Analyse des (implizit angenommenen) Prävalenz-Screenings als adäquat angesehen werden. Das Problem eines work-up bias bei den Quelldaten wurde in der Sensitivitätsanalyse aufgegriffen: die Korrektur wäre methodisch korrekt, wenn eine randomisierte Aufteilung der biopsierten Populationsuntergruppen auf die einzelnen durch Testbefund und tatsächlichen Krankheitsstatus definierten Kategorien (richtigpositiv, falschpositiv, richtignegativ, falschnegativ) gegeben wäre. Die Kostenberechnung ist nachvollziehbar und transparent. Als problematisch erscheint, intermediäre Kosten in Relation zu finalen gesundheitsbezogenen Outcome-Größen zu setzen, ohne den dafür erforderlichen Ressourcenverbrauch zu berücksichtigen.

**Littrup PJ, Goodman AC, Mettlin CJ, the Investigators of the American Cancer Society National Prostate Cancer Detection Project:
The Benefit and Cost of Prostate Cancer Early Detection.
CA - A Cancer Journal for Clinicians, 1993; 43(3): 134-149.**

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels DRU, Bestimmung von Serum-PSA und TRUS. Explizite Gegenüberstellung von verschiedenen Früherkennungsprogrammen.

1-2 Fragestellung

Evaluation der Effektivität von neun verschiedenen Früherkennungsprogrammen unter Berücksichtigung ökonomischer Gesichtspunkte, Ermittlung der marginalen Kosten-Effektivität sowie Aufstellung einer Kosten-Nutzen-Analyse auf Basis der Daten einer bereits veröffentlichten Longitudinalstudie²¹.

Screening-Strategie 1): PSA → DRU^U+TRUS –[⊕^{wo}]→ Biopsie

Screening-Strategie 2): TRUS –[⊕]→ Biopsie

Screening-Strategie 3): DRU^U –[⊕]→ Biopsie

Screening-Strategie 4): PSA+DRU^U –[>2ng/ml und/oder ⊕]→ Biopsie

Screening-Strategie 5): PSA+DRU^U –[>3ng/ml und/oder ⊕]→ Biopsie

Screening-Strategie 6): PSA+DRU^U –[>4ng/ml und/oder ⊕]→ Biopsie

Screening-Strategie 7): PSA –[>2ng/ml]→ Biopsie

Screening-Strategie 8): PSA –[>3ng/ml]→ Biopsie

Screening-Strategie 9): PSA –[>4ng/ml]→ Biopsie

U = Urologe, erfahrener Untersucher, AM = Allgemeinmediziner, weniger erfahrener Untersucher (s. Sensitivitätsanalysen)

Die Screening-Strategie 1) stellt die in der Studie ursprünglich eingesetzte Strategie dar: der PSA-Spiegel wurde, bis auf einzelne Ausnahmen, nicht als Entscheidungskriterium für eine Biopsie herangezogen.

Unklar ist, ob in Verbindung mit einer Biopsie jedes Mal auch eine TRUS durchgeführt wird.

²¹ Mettlin C, Lee F, Drago J, Murphy GP, the Investigators of the American Cancer Society National Prostate Cancer Detection Project: *The American Cancer Society National Prostate Cancer Detection Project. Findings on the Detection of Early Prostate Cancer in 2,425 Men.* Cancer 1991; 67:2949-2958.

1-3 Perspektive

Keine Angaben.

1-4 Zeithorizont

Ermittelt wurden Effekte und Kosten für die drei ersten Untersuchungsjahre der Studie (Prävalenz- und Inzidenz-Screening).

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse; Kosten-Nutzen-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Originalstudie.

Datenquelle für Gesundheitseffekt: Originalstudie (Longitudinalstudie) mit retrospektiver Entscheidungsanalyse.

Datenquelle für Ressourcenverbrauch und Mengengerüst: Modellierung (Entscheidungsanalyse); Einkommenschätzungen basierend auf Statistiken (U.S. Department of Commerce, Bureau of the Census); im Rahmen der Kosten-Nutzen-Analyse teilweise explizite Annahmen.

Datenquelle(n) für Preise: Health Care Finance Administration und American College of Radiology für die monetäre Bewertung des direkten medizinischen Ressourcenverbrauchs; für die monetäre Bewertung von intangiblen Ressourcenveränderungen wurde auf eine Studie von Goddeeris & Bronken (1985)²² zurückgegriffen.

2-2 Datierung der zugrundeliegenden Daten

Die Datenerhebung zum Gesundheitseffekt erfolgte zwischen 1987 und Januar 1993; keine expliziten Angaben zu Zeitpunkt und/oder Zeitraum der Datenerhebung zu den berücksichtigten indirekten und intangiblen Ressourcenveränderungen.

Keine Angaben hinsichtlich des Bezugsjahres für die Preisangaben.

²² Goddeeris JH & Bronken TP: *Benefit-Cost Analysis of Screening: A Comparison of Tests for Gonorrhea*. Med Care 1985; 23: 1242-1255.

2-3 Studienpopulation / Zielpopulation

Männer, für welche die Ergebnisse von drei aufeinanderfolgenden Untersuchungsjahren vorlagen. Keine expliziten Altersangaben: in die ACS-NPCDP-Studie wurden Männer im Alter von 55 bis 70 Jahren ohne krankheitsspezifische Vorgeschichte oder Verdacht auf Prostatakarzinom aufgenommen.

2-4 Setting

Tangierte Versorgungsstrukturen: verschiedenartige ambulante Versorgungseinrichtungen.

Der klinische Teil der Studie wurde multizentrisch (zwölf Zentren) in den USA und in Kanada durchgeführt; der ökonomische Teil der Studie wurde in den USA aufgestellt.

2-5 Spezifikation der Technologie

PSA-Assay: keine expliziten Angaben (Mettlin et al.: Abbott IMx, Hybritech Tandem-R und Diagnostic Products IRMA-Count (monoklonal) sowie Yang Laboratories Pro-Check (polyklonal)).

DRU: keine expliziten Angaben (Mettlin et al.: Palpation durch Urologen).

TRUS: keine expliziten Angaben (Mettlin et al.: Real-Time-Verfahren, 7.0MHz-Schallkopf, biplanar; Untersuchung durch Radiologen).

Biopsie: keine expliziten Angaben (Mettlin et al.: transrektale Stanzbiopsie (Biopsy-Pistole) unter Ultraschallkontrolle oder digitaler Führung; Antibiotikaprophylaxe).

3 Gesundheitseffekte

Screening-/Diagnostiestudie

3-1 Untersuchte Zielgrößen

Primäre Zielgröße für die Kosten-Effektivitäts-Analyse war die pro Untersuchungsjahr Anzahl diagnostizierter Prostatakarzinome im Probandenkollektiv. Primäre Zielgrößen für die Kosten-Nutzen-Analyse waren die Prävalenz von Prostatakarzinom im Probandenkollektiv, die Sensitivität und Spezifität der Tests, der Anteil an Stadium A- und B- bzw. C- und D-Karzinomen sowie die Wahrscheinlichkeit einer Behandlung nach positivem Biopsiebefund.

3-2 Ein-/Ausschlusskriterien

Als Einschlusskriterium wurde in der Publikation von Littrup et al. lediglich das Geschlecht genannt (Mettlin et al.: weiteres Einschlusskriterium war das Alter; Ausschlusskriterien waren eine krankheitsspezifische Symptomatik bzw. eine krankheitsspezifische Vorgeschichte).

3-3 Rekrutierungsmodus

Keine expliziten Angaben (Mettlin et al.: unabhängige Rekrutierung durch die einzelnen Studienzentren über Werbeanzeigen, Vorträge, Berichterstattung in den lokalen Medien, Teilnahme an nationalen prostate cancer awareness-Programmen; mit Ausnahme eines Studienzentrums kostenloses Screening-Angebot).

3-4 Teilnahmerate

Keine Angaben. Für die Analyse lagen die 3-Jahres-Ergebnisse von 1,449 Männern vor (insgesamt waren zu diesem Zeitpunkt 2,425 Männer rekrutiert worden).

3-5 Reproduzierbarkeit der Testergebnisse

Keine expliziten Angaben.

Mettlin et al.: Adjustierung der PSA-Werte des polyklonalen PSA-Assays zur Gegenüberstellung mit den Werten der monoklonalen PSA-Assays, jedoch keine Aussagen bezüglich der Vergleichbarkeit der eingesetzten PSA-Assays und PSA-Antikörpertypen. Als Entscheidungskriterien bezüglich der Interpretation der DRU galt eine palpatorisch indurierte, knotige oder asymmetrische Prostata. Bezüglich der Interpretation der TRUS galten nicht durch Zysten, Artefakte oder vaskuläre bzw. andere normale anatomische Strukturen bedingte echoarme Areale mit einer durchschnittlichen Größe (33% der Summe aus Länge, Breite und Höhe) von >5-7cm als Entscheidungskriterien. Die Gewebeprobeanalysen wurden von einem Pathologen an jedem Studienzentrum durchgeführt, Blutprobenanalysen an einem Zentrallabor. Vorkehrungen der Qualitätskontrolle bei TRUS-Interpretationen und Gewebeprobeanalysen waren periodische Treffen und eine Überprüfung aller Gewebeproben durch einen Gutachter.

3-6 Auswertung der Testergebnisse

Keine expliziten Angaben (Mettlin et al.: die PSA-Bestimmung diente nicht als Entscheidungskriterium für eine Biopsie und wurde nicht bei allen Probanden durchgeführt; alle Probanden unterzogen sich einer DRU und TRUS, welche von

unterschiedlichen Untersuchern durchgeführt wurde; die Biopsie als Goldstandard wurde lediglich bei Probanden mit positivem DRU- und/oder TRUS-Befund durchgeführt).

3-7 Drop-Outs

Keine Angaben.

3-8 Ergebnisse der Screening-/Diagnosestudie

Deskriptive Präsentation der Ergebnisse für den ursprünglichen Interventionspfad sowie für die modellierten Screening-Strategien. Für die Screening-Strategien 3) bis 9) wurden im Rahmen der Kosten-Nutzen-Analyse Sensitivität und Spezifität der Screening-Tests (einzeln oder in Kombination) angegeben, wobei sich diese Werte auf die 99 mittels DRU/TRUS in drei Jahren entdeckten Karzinome bezogen (inzidentelle Karzinome oder allein aufgrund von erhöhten PSA-Werten entdeckte Tumoren wurden in der Analyse nicht berücksichtigt). In der Veröffentlichung von Littrup et al. zu der Studie wurden keine statistischen Testverfahren aufgeführt.

Ursprünglicher Interventionspfad: Für 1,449 Männer lagen die Ergebnisse von drei aufeinanderfolgenden Untersuchungsjahren vor. Im ersten Jahr wurden insgesamt 78 Karzinome entdeckt: 61 über die DRU/TRUS, 6 über eine TUR-P (inzidentelle Tumoren), 7 allein über die PSA-Bestimmung und 4 über die DRU/TRUS bei fehlendem PSA-Wert. Im zweiten Jahr waren dies 38 (29, 1, 7 bzw. 1) Karzinome, im dritten Jahr 13 (9, 0, 3 bzw. 1) Karzinome. Nach drei Studienjahren lagen Ergebnisse für 2,425 Männer mit unterschiedlicher Anzahl Untersuchungen bzw. Untersuchungsbefunden vor: die Karzinomentdeckungsrate lag gesamt gesehen bei 5.3% (im ersten Jahr bei 3.2%, im zweiten Jahr bei 1.6% und im dritten Jahr bei 0.5%). Der Anteil der diagnostizierten Stadium A- und B- Karzinome lag bei 95% (im Gegensatz zu den auf Daten des American College of Surgeons basierenden 56% im Jahre 1986).

Die modellierten Screening-Strategien beziehen sich jeweils auf 1,000 Screening-Teilnehmer:

Screening-Strategie 1): Im ersten Jahr wurden 132.0 Biopsien durchgeführt und 42.1 Karzinome entdeckt. Im zweiten Jahr waren dies 82.8 Biopsien bzw. 20.0 entdeckte Karzinome, im dritten Jahr 46.9 Biopsien bzw. 6.2 entdeckte Karzinome. Insgesamt wurden 261.6 Biopsien durchgeführt und 68.3 Karzinome diagnostiziert

Screening-Strategie 2): Im ersten Jahr wurden 115.2 Biopsien durchgeführt und 35.9 Karzinome entdeckt. Im zweiten Jahr waren dies 62.8 Biopsien bzw. 14.5 entdeckte Karzinome, im dritten Jahr 39.3 Biopsien bzw. 5.5 entdeckte Karzinome. Insgesamt wurden 217.4 Biopsien durchgeführt und 55.9 Karzinome diagnostiziert.

Screening-Strategie 3): Im ersten Jahr wurden 57.2 Biopsien durchgeführt und 26.9 Karzinome entdeckt. Im zweiten Jahr waren dies 31.8 Biopsien bzw. 12.4 entdeckte Karzinome, im dritten Jahr 13.8 Biopsien bzw. 2.1 entdeckte Karzinome. Insgesamt wurden 102.8 Biopsien durchgeführt und 41.4 Karzinome diagnostiziert.

Screening-Strategie 4): Im ersten Jahr wurden 92.2 Biopsien durchgeführt und 39.4 Karzinome entdeckt. Im zweiten Jahr waren dies 57.3 Biopsien bzw. 20.0 entdeckte Karzinome, im dritten Jahr 32.4 Biopsien bzw. 6.2 entdeckte Karzinome. Insgesamt wurden 181.5 Biopsien durchgeführt und 65.6 Karzinome diagnostiziert

Screening-Strategie 5): Im ersten Jahr wurden 80.9 Biopsien durchgeführt und 37.3 Karzinome entdeckt. Im zweiten Jahr waren dies 49.7 Biopsien bzw. 17.9 entdeckte Karzinome, im dritten Jahr 27.6 Biopsien bzw. 6.2 entdeckte Karzinome. Insgesamt wurden 158.0 Biopsien durchgeführt und 61.4 Karzinome diagnostiziert

Screening-Strategie 6): Im ersten Jahr wurden 76.8 Biopsien durchgeführt und 35.2 Karzinome entdeckt. Im zweiten Jahr waren dies 45.6 Biopsien bzw. 17.9 entdeckte Karzinome, im dritten Jahr 24.2 Biopsien bzw. 6.2 entdeckte Karzinome. Insgesamt wurden 146.3 Biopsien durchgeführt und 59.4 Karzinome diagnostiziert.

Screening-Strategie 7): Im ersten Jahr wurden 75.0 Biopsien durchgeführt und 38.0 Karzinome entdeckt. Im zweiten Jahr waren dies 46.2 Biopsien bzw. 18.6 entdeckte Karzinome, im dritten Jahr 23.5 Biopsien bzw. 5.5 entdeckte Karzinome. Insgesamt wurden 144.2 Biopsien durchgeführt und 62.1 Karzinome diagnostiziert.

Screening-Strategie 8): Im ersten Jahr wurden 54.0 Biopsien durchgeführt und 33.1 Karzinome entdeckt. Im zweiten Jahr waren dies 35.2 Biopsien bzw. 15.2 entdeckte Karzinome, im dritten Jahr 16.6 Biopsien bzw. 4.8 entdeckte Karzinome. Insgesamt wurden 105.6 Biopsien durchgeführt und 53.1 Karzinome diagnostiziert.

Screening-Strategie 9): Im ersten Jahr wurden 43.0 Biopsien durchgeführt und 26.9 Karzinome entdeckt. Im zweiten Jahr waren dies 26.2 Biopsien bzw. 13.8 entdeckte Karzinome, im dritten Jahr 12.4 Biopsien bzw. 4.8 entdeckte Karzinome. Insgesamt wurden 81.4 Biopsien durchgeführt und 45.6 Karzinome diagnostiziert.

Die Werte für die Sensitivität und Spezifität der Screening-Tests waren folgende: Screening-Strategie 3) 0,606 bzw. 0.963; Screening-Strategie 4) 0.960 bzw. 0.688; Screening-Strategie 5) 0.899 bzw. 0.827; Screening-Strategie 6) 0.869 bzw. 0.896; Screening-Strategie 7) 0.909 bzw. 0.705; Screening-Strategie 8) 0.778 bzw. 0.853; Screening-Strategie 9) 0.667 bzw. 0.927.

3-9 Effektmaße für die ökonomische Analyse

Kosten-Effektivitäts-Analyse: Anzahl identifizierter Prostatakarzinome pro Untersuchungsjahr, Anzahl identifizierter Prostatakarzinome in drei Jahren.

Kosten-Nutzen-Analyse (Bezug auf den Gesamteffekt der drei Untersuchungsjahre):
Effektmaß war die nach folgender Gleichung ermittelte Differenz aus Nutzen und Kosten:

$$V = pSewB + (1-p)SpR - S - [pSe + (1-p)(1-Sp)]w(T+A) - (1-p)(1-Sp)F$$

mit: V = Nettonutzen pro Screening-Teilnehmer

B = Monetärer Wert des Nutzens aus der Karzinomfrüherkennung und -behandlung
 = (Medizinische Kostenersparnis aufgrund vermehrter Behandlung von Stadium A- und B-Karzinomen* + Vermindertes physisches und psychisches Leiden aufgrund der Früherkennung + Einkommensverlust) x Anstieg der Entdeckungsrate von Stadium A- und B-Karzinomen in %

* (Behandlungskosten für Stadium C- und D-Karzinome) -
 (Behandlungskosten für Stadium A- und B-Karzinome)

R = Monetärer Wert von Erleichterung und Sicherheit bei negativem Testbefund

S = Kosten für die eingesetzten Screening-Tests

T = Kosten für Staging und Behandlung als Resultat des Screenings

= (Anteil Stadium A- und B-Karzinome x Behandlungskosten von Karzinomen im Frühstadium) + (Anteil Stadium C- und D-Karzinome x Behandlungskosten von Karzinomen in fortgeschrittenem Stadium) + Staging-Kosten

A = Medizinische Kosten bei falschpositivem Screening-Befund sowie für diagnose- und behandlungsbedingte Morbidität

= (Anteil testpositiver Befunde) x (Kosten der TRUS) + [(Anteil Biopsien)(Kosten der Biopsie) + Kosten der DRU* + Kosten von biopsiebedingten Komplikationen + Kosten von unnötiger Therapie bei Überdiagnose + Kosten von behandlungsbedingten Komplikationen]

* nur bei Screening-Strategien mit alleinigem Einsatz der PSA-Bestimmung

F = Potentielle Kosten der psychischen Belastung bei falschpositivem Screening-Testbefund

p = Prävalenz von Prostatakarzinom = Anteil der Patienten mit Prostatakarzinom im Rahmen von Screening und TUR-P

Se = Sensitivität

Sp = Spezifität

w = Wahrscheinlichkeit einer Behandlung nach positivem Biopsiefund

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

In der Kosten-Effektivitäts-Analyse wurden die direkten variablen Kosten für die medizinische Versorgung berücksichtigt.

- Screening/Diagnostik: Screening-Test(s), Biopsie.

In der Kosten-Nutzen-Analyse wurde zusätzlich erfasst:

- Screening/Diagnostik: Staging-Kosten (Knochenszintigramm, CT, MRI)
- Behandlung: primäre Behandlungskosten
- vermiedene Produktivitätskosten (Einkommensverluste)
- die unter Punkt 3.9 aufgeführten direkten Kosten (A)
- die unter Punkt 3.9 aufgeführten intangiblen Kosten (R, F, teilweise B)

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte, soweit ersichtlich, nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Das Mengengerüst wurde nicht spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Soweit ersichtlich pauschales Verfahren hinsichtlich der direkten variablen Kosten. Die monetäre Bewertung des Einkommensverlustes (Produktivitätskosten) ist unklar. Bezüglich der monetären Bewertung der screeningbedingten Verminderung von physischem und psychischem Leiden (Teil von B) wurde auf einen Konsenswert aus der Publikation von Goddeeris & Bronken zurückgegriffen. Der monetäre Wert von R wurde gleich dem Wert von S gesetzt. Für den monetären Wert von F wurde der Wert von A übernommen.

Die Preise pro Mengeneinheit wurden nur teilweise aufgeführt.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die Screening-Strategie 1) (ursprünglicher Interventionspfad) führte im ersten Untersuchungsjahr zur Entdeckung von 61 Prostatakarzinomen, im zweiten Untersuchungsjahr zur Entdeckung von 29 Prostatakarzinomen und im dritten Untersuchungsjahr zur Entdeckung von 9 Prostatakarzinomen. Insgesamt wurden in 3 Jahren 99 Karzinome entdeckt. Bei der Screening-Strategie 2) waren dies 52, 21 bzw. 8 Karzinome, insgesamt 81 Karzinome in drei Jahren. Bei der Screening-Strategie 3) waren dies 39, 18 bzw. 3 Karzinome, insgesamt 60 Karzinome in drei Jahren. Bei der Screening-Strategie 4) waren dies 57, 29 bzw. 9 Karzinome, insgesamt 95 Karzinome in drei Jahren. Bei der Screening-Strategie 5) waren dies 54, 26 bzw. 9 Karzinome, insgesamt 89 Karzinome in drei Jahren. Bei der Screening-Strategie 6) waren dies 51, 26 bzw. 9 Karzinome, insgesamt 86 Karzinome in drei Jahren. Bei der Screening-Strategie 7) waren dies 55, 27 bzw. 8 Karzinome, insgesamt 90 Karzinome in drei Jahren. Bei der Screening-Strategie 8) waren dies 48, 22 bzw. 7 Karzinome, insgesamt 77 Karzinome in drei Jahren. Bei der Screening-Strategie 9) waren dies 39, 20 bzw. 7 Karzinome, insgesamt 66 Karzinome in drei Jahren.

Für die Kosten-Nutzen-Analyse wurden folgende Anpassungen bzw. Annahmen getroffen: Die Gesamtprävalenz wurde aufgrund eines angenommenen sampling bias in der Untergruppe der Männer mit drei Untersuchungsergebnissen von 8.9% (129/1,449) auf 5.3% (129/2,425) adjustiert. Die Wahrscheinlichkeit einer Behandlung nach positivem Biopsiebefund wurde auf 0.85 angesetzt, basierend auf Daten eines Survey des American College of Surgeons aus dem Jahre 1990.

6-2 Ermittelte Kosten

Keine Angaben.

6-3 Synthese von Kosten und Effekten

Die durchschnittlichen Kosten pro entdecktem Karzinom (Quotient aus der Summe der dreijährigen Diagnosekosten und der Anzahl in drei Jahren entdeckter Karzinome) beliefen sich bei der Screening-Strategie 1) auf US-\$ 11,135, bei der Screening-Strategie 2) auf US-\$ 9,994, bei der Screening-Strategie 3) auf US-\$ 3,415, bei der Screening-Strategie 4) auf US-\$ 6,594, bei der Screening-Strategie 5) auf US-\$ 5,895, bei der Screening-Strategie 6) auf US-\$ 5,516, bei der Screening-Strategie 7) auf US-\$ 5,592, bei

der Screening-Strategie 8) auf US-\$ 4,780 und bei der Screening-Strategie 9) auf US-\$ 4,482.

Die marginalen Kosten pro jährlich zusätzlich entdecktem Karzinom (Quotient aus den diagnostischen Kosten im entsprechenden Untersuchungsjahr und der Anzahl der im gleichen Jahr entdeckten Karzinome) betragen im ersten Untersuchungsjahr für Strategie 1) US-\$ 6,556, für Strategie 2) US-\$ 5,784, für Strategie 3) US-\$ 2,178, für Strategie 4) US-\$ 4,174, für Strategie 5) US-\$ 3,731, für Strategie 6) US-\$3,618, für Strategie 7) US-\$ 3,501, für Strategie 8) US-\$ 2,924 und für Strategie 9) US-\$ 2,905.

Die marginalen Kosten im zweiten Untersuchungsjahr betragen zwischen US-\$ 3,693 (Screening-Strategie 3)) und US-\$ 12,562 (Screening-Strategie 1). Die marginalen Kosten im dritten Untersuchungsjahr betragen zwischen US-\$ 12,079 und US-\$ 37,588 (Screening-Strategie 1). Dabei war zwischen dem ersten und dritten Untersuchungsjahr bei allen Screening-Strategien mit Ausnahme der Strategie 3) (marginale Kosten pro zusätzlich im dritten Untersuchungsjahr entdecktem Karzinom von US-\$ 17,823) ein bei 5% liegender Anstieg der Kosten auszumachen.

Bei der Kosten-Nutzen-Analyse wurden die Screening-Strategien 1) und 2) nicht berücksichtigt. Für die Screening-Strategie 3) ergab die Kosten-Nutzen-Analyse einen Nettonutzen V von US-\$ -474, für die Screening-Strategie 4) von US-\$ -5,544, für die Screening-Strategie 5) von US-\$ -2,936, für die Screening-Strategie 6) von US-\$ -1,646, für die Screening-Strategie 7) von US-\$ -5,216, für die Screening-Strategie 8) von US-\$ -2,473, für die Screening-Strategie 9) von US-\$ -1,126.

7 Behandlung von Unsicherheiten

Für die Screening-Strategien mit einer PSA-Bestimmung wurden ROC-Teilkurven für PSA-Spiegel von 2ng/ml, 3ng/ml und 4ng/ml aufgezeichnet: Bei einer Screening-Strategie mit alleiniger PSA-Bestimmung lag das Optimum für die Sensitivität und Spezifität bei einem PSA-Wert nahe 2ng/ml; bei einer Screening-Strategie mit einer Kombination von PSA-Bestimmung und DRU^U lag das Optimum bei einem PSA-Wert von 3ng/ml. Bezüglich der Durchschnittskosten wurden für jede Screening-Strategie 1-Weg-Sensitivitätsanalysen mit variablen Einzeltestkosten im Bereich von 33% bis 333% der Werte der Basisanalyse durchgeführt. Dabei erwiesen sich für eine PSA-Entscheidungsgrenze von 2ng/ml (PSA-Test allein oder in Kombination mit der DRU) die Kosten für die TRUS und bei einer PSA-Entscheidungsgrenze von 4ng/ml (PSA-Test allein oder in Kombination mit der DRU) die Kosten für die PSA-Bestimmung als bestimmend.

In der Kosten-Nutzen-Analyse erwiesen sich die Spezifität, der Nutzen aus der Karzinomfrüherkennung und -behandlung sowie die Prävalenz in absteigender Reihenfolge ($Sp > B > p$) als sensibelste Parameter für den Nettonutzen V . Die Wertebereiche für die Sensitivitätsanalyse wurden nur für einzelne Parameter spezifiziert. Für die DRU^U wurde weiterhin eine um jeweils 20% niedrigere Sensitivität und Spezifität bei Durchführung im breiteren Rahmen angesetzt (DRU^{AM}), wobei die Screening-Strategie 3) von der kostengünstigsten Technologiealternative zu einer der teuersten mutierte ($V_{DRU^{AM}} = \text{US-}\$ -4,112$).

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren weisen darauf hin, dass für die einzelnen Parameter keine Konfidenzintervalle ermittelt wurden und auch keine statistischen Testverfahren bezüglich Unterschieden zwischen den Screening-Strategien angewendet wurden. Nach Meinung der Autoren kann sich die Annahme eines langfristigen Nutzens der Früherkennung und -behandlung von Prostatakarzinom als Mangel erweisen. Die wahre Prävalenz der im Rahmen von Screening entdeckbaren Krebsfällen ist unbekannt.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Die Autoren sind der Ansicht, dass die Testgüte der untersucherabhängigen DRU^U und TRUS, wie sie für die ACS-NPCDP-Studie ermittelt wurde, nicht unbedingt für ein Früherkennungsprogramm im breiteren Rahmen gelten kann, da evtl. nicht genügend qualifizierte Untersucher für ein Massen-Screening zur Verfügung stehen. Die ACS-NPCDP-Studie ist in diesem Zusammenhang als best case-Szenario zu verstehen. Auch die angenommene Krankheitsprävalenz von 5.3% ist in ihren Augen nicht unbedingt generalisierbar.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass die Empfehlung einer jährlichen Durchführung der DRU^U zur Früherkennung von Prostatakarzinom nach dem dritten Untersuchungsjahr von geringem klinischen Nutzen ist, insbesondere wenn zusätzlich auch die Variabilität der Untersuchung berücksichtigt wird. Angesichts einer für den Referenzbereich des PSA-Assays allgemein geltenden oberen Grenze von 4ng/ml sind die Autoren der Meinung, dass eine Preisbegrenzung der PSA-Bestimmung entscheidend

hinsichtlich der Vermeidung eines Kostenanstiegs ist. Die Autoren glauben, dass die Entwicklung eines PSA-Assays mit höherer Spezifität zu einer Senkung der Kosten für die Früherkennung von Prostatakarzinom führen wird. Unter den Annahmen ihrer Analyse erscheint den Autoren die Kombination von PSA-Bestimmung und DRU^U aufgrund verbesserter Sensitivität und ausreichender Spezifität als ethisches und kostengünstiges Mittel der Wahl für die Karzinomfrüherkennung bei individuellen Patienten. Zur Quantifizierung der Signifikanz von Unterschieden zwischen den einzelnen Screening-Strategien bedarf es ihrer Meinung nach weiterer Forschung.

9 Kommentar

Entscheidungsanalyse zu Biopsiekriterien beim Prävalenz- und Inzidenz-Screening mit expliziter Gegenüberstellung von Technologiealternativen. Als interessierende Zielgrößen wurden intermediäre Outcome-Parameter gewählt. Die Formulierung der Screening-Strategien mit PSA-Bestimmung ist irreführend, da bei der ursprünglichen Interventionsstrategie zwischen PSA-Bestimmung und Biopsie selektierende Tests (DRU^U und/oder TRUS) zwischengeschaltet waren. Hinsichtlich der Ermittlung der Effektivität ist die Möglichkeit eines selection bias und eines expectation bias gegeben. Das Problem der Interrater-Reliabilität wurde zum Teil in der Sensitivitätsanalyse und Diskussion aufgegriffen. Die Ermittlung der jährlichen bzw. marginalen Testpositiv- und Biopsierate erfolgte jeweils mit der ursprünglichen Studienpopulation als Referenzgruppe ohne Abzug der Probanden mit bereits entdecktem Karzinom. Die lediglich auf eine selektierte Populationsuntergruppe beschränkte Anwendung des Goldstandards steht in Verbindung mit einem work-up bias. Die auf den Ergebnissen einer dreimaligen Untersuchungswiederholung ermittelten Werte zur Sensitivität und Spezifität der Tests erscheint als äußerst problematisch. Bezüglich der Studienpräsentation ist ein Rückgriff auf die Publikation von Mettlin et al. in einigen Punkten unabdingbar. Angaben zur stadiumspezifischen Entdeckungsrate in Abhängigkeit von der Screening-Strategie fehlen. In der Kosten-Nutzen-Analyse mit der über die verschiedenen Screening-Strategien hinweg unverändert bleibenden Differenz zwischen den Behandlungskosten bei Karzinomen im fortgeschrittenen Stadium und denen bei Karzinomen im Frühstadium implizit von einer bei allen Screening-Strategien gleich ausfallenden Stadienverteilung der entdeckten Tumoren ausgegangen. Auch scheinen mittels der Früherkennung ausschließlich Karzinome im Frühstadium entdeckt zu werden. Der Dokumentation bezüglich der Ermittlung von Effektivität und insbesondere von Kosten mangelt es erheblich an Transparenz. Auch weist die Publikation ein großes Defizit an Nachvollziehbarkeit auf: zwar wird auf eine detailliertere, zu einem späteren Zeitpunkt zu

erscheinende Veröffentlichung verwiesen, bis August 1998 ist eine solche jedoch in keiner elektronischen Datenbank (und auch nicht in anderen HTA-Berichten) vermerkt. Die monetäre Bewertung von A wie auch das dabei zugrundeliegende Mengengerüst ist unklar. Bei der Kosten-Nutzen-Analyse wurde für die intangiblen Kosten auf Konsenswerte bzw. arbiträre Annahmewerte zurückgegriffen. Obwohl 9 Strategien für jeweils 1, 2 und 3 Jahre analysiert wurden, wurde mein Inkrement zwischen den Strategien gebildet. Lediglich bezüglich der Zeit wurde inkrementell vorgegangen, d.h. die im 2. und 3. Jahr zusätzlich anfallenden Effekte, Kosten und Benefits wurden getrennt berichtet. Aufgrund der methodischen Einschränkungen der Publikation sind die interne Validität, Reproduzierbarkeit und Generalisierbarkeit der Analyseergebnisse erheblich reduziert. Erstaunlich ist, dass die Autoren keine expliziten Folgerungen aus ihren durchgehend negativen Analyseergebnissen (d.h. negativer Nutzen mit Nettokosten) ziehen, welche nahelegen, dass die Früherkennung von Prostatakarzinom keinen Nutzen bringt. Außerdem sei darauf hingewiesen, dass der Wertebereich der Sensitivitätsanalyse bezüglich der Einzelpreise der Screening-Tests zwischen 33% und 166% des Wertes der Basisanalyse betrug (und nicht, wie angegeben, zwischen 33% und 333%).

10 Ähnliche Publikationen:

Littrup PJ, Goodman AC, Mettlin CJ, Murphy GP, the Investigators of the American Cancer Society National Prostate Cancer Detection Project:

Cost Analyses of Prostate Cancer Screening: Frameworks for Discussion.

Journal of Urology, 1994; 152: 1873-1877.

Littrup PJ, Goodman AC, the Investigators of the American Cancer Society National Prostate Cancer Detection Project:

Costs and Benefits of Prostate Cancer Screening.

In vivo, 1994; 8: 423-428.

In diesen Publikationen wird für Subgruppenanalysen in die ursprüngliche Gleichung der Kosten-Nutzen-Analyse die Wahrscheinlichkeit der Teilnahme an einem Früherkennungsprogramm j mit einbezogen:

$$V_j = pSeWB_j + (1-p)SpR_j - S_j - [pSe + (1-p)(1-Sp)]w_j(T+A) - (1-p)(1-Sp)F_j$$

Unter Verdoppelung der in der Basisanalyse angenommenen Werte von p , B und A nimmt der Nettonutzen der Screening-Strategie 3) ($j = 1.0$ bzw. 0.41), der Screening-Strategie 6) ($j = 1.0$ bzw. 0.51) sowie der Screening-Strategie 9) ($j = 1.0$ bzw. 0.62) positive Werte an.

Littrup PJ, Kane RA, Mettlin CJ, Murphy GP, Lee F, Toi A, Badalament R, Babaian R, for the Investigators of the American Cancer Society National Prostate Cancer Detection Project:
Cost-Effective Prostate Cancer Detection. Reduction of Low-Yield Biopsies.
Cancer, 1994; 74(12): 3146-3158.

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels DRU, Bestimmung von Serum-PSA und TRUS. Explizite Gegenüberstellung von verschiedenen Früherkennungsprogrammen.

1-2 Fragestellung

Analyse der Reliabilität und Karzinomentdeckungsrate von verschiedenen PSA-Entscheidungsparametern für eine Biopsie (absoluter PSA-Wert, altersabhängiger PSA-Wert, PSA-D und PSA-V) auf Basis der Daten einer bereits veröffentlichten Longitudinalstudie^{23,24,25,26,27}: retrospektive Identifizierung einer Screening-Strategie mit minimaler Biopsie- bzw. Falschpositiv-Rate bei optimaler Karzinomentdeckungsrate. Zusätzlich entscheidungsanalytische Ermittlung der Kosten-Effektivität verschiedener Screening-Strategien.

Screening-Strategie 1): PSA+DRU $-[>4\text{ng/ml und/oder } \oplus] \rightarrow$ systematische Biopsie

Screening-Strategie 2): PSA+DRU $-[>4\text{ng/ml und/oder } \oplus] \rightarrow$ adaptierte Biopsie

Screening-Strategie 3): DRU $-[\oplus] \rightarrow$ adaptierte Biopsie

Screening-Strategie 4): PSA $-[>3\text{ng/ml }] \rightarrow$ adaptierte Biopsie

Screening-Strategie 5): PSA $-[>4\text{ng/ml }] \rightarrow$ adaptierte Biopsie

²³ Mettlin C, Lee F, Drago J, Murphy GP, et al.: *The American Cancer Society National Prostate Cancer Detection Project. Findings on the Detection of Early Prostate Cancer in 2,425 Men.* Cancer 1991; 67:2949-2958.

²⁴ Babaian RJ, Mettlin CJ, Kane R, Murphy GP, et al.: *The Relationship of Prostate-Specific Antigen to Digital Rectal Examination and Transrectal Ultrasonography in Detecting Adenocarcinoma of the Prostate.* Cancer 1992; 69: 1195-2000.

²⁵ Kane RA, Littrup PJ, Babaian RJ, Drago J, et al.: *Prostate-Specific Antigen Levels in 1,695 Men Without Evidence of Prostate Cancer.* Cancer 1992; 69:1201-1207.

²⁶ Mettlin CJ, Murphy GP, Lee F, Littrup PJ, et al.: *Characteristics of Prostate Cancers Detected in Multi-Modality Early Detection Program.* Cancer 1993; 72: 1701-1708.

²⁷ Mettlin CJ, Littrup PJ, Kane RA, Murphy GP, et al.: *Relative Sensitivity and Specificity of Serum PSA Level Compared to Age-Referenced PSA, PSA Density and PSA Change: Data from the American Cancer Society National Prostate Cancer Detection Project.* Cancer 1994; 74: 1615-1620.

Screening-Strategie 6): PSA –[>altersabhängiger PSA-Grenzwert]→ systematische Biopsie

Screening-Strategie 7): PSA –[>altersabhängiger PSA-Grenzwert]→ adaptierte Biopsie

Screening-Strategie 8): PSA+DRU –[>altersabhäng. PSA-Grenzwert *und/oder* ⊕]
→ systematische Biopsie

Screening-Strategie 9): PSA+DRU –[>altersabhäng. PSA-Grenzwert *und/oder* ⊕]
→ adaptierte Biopsie

Adaptierte Biopsie: Biopsie bei Probanden mit einem PSA-D-Wert von >0.12ng/ml/ccm oder gleichzeitig positivem DRU- und TRUS-Befund.

Systematische Biopsie: Biopsie bei Probanden mit PSA-Werten von >PSA-Grenzwert und/oder positivem DRU-Befund.

Altersabhängiger PSA-Grenzwert: PSA=2.5ng/ml bei 40-49jährigen, PSA=3.5ng/ml bei 50-59jährigen, PSA=4.5ng/ml bei 60-69jährigen, PSA=6.5ng/ml bei ≥70jährigen nach Oesterling et al. (1993)²⁸.

1-3 Perspektive

Keine Angaben.

1-4 Zeithorizont

Ermittelt wurden die Effekte und Kosten von Screening und Diagnose im Rahmen des Prävalenz-Screenings. Keine explizite Angabe zum Zeithorizont. Datengrundlage war ein jährliches Follow-Up bis zu 5 Jahren.

1-5 Art der ökonomischen Evaluation

Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Originalstudie.

Datenquelle für Gesundheitseffekt: Originalstudie (Longitudinalstudie) mit retrospektiver Entscheidungsanalyse.

Datenquelle für Ressourcenverbrauch und Mengengerüst: Modellierung (Entscheidungsanalyse).

²⁸ Oesterling JE, Jacobson SJ, Chute CG, Guess HA, Girman CJ, et al.: *Prostate Specific Antigen in a Community Based Population of Healthy Men: An Establishment of Age-Specific Reference Ranges.* JAMA 1993; 270:860-864.

Datenquelle(n) für Preise: keine expliziten Angaben.

2-2 Datierung der zugrundeliegenden Daten

Die Datenerhebung zum Gesundheitseffekt erfolgte zwischen 1987 und einem nicht explizit genannten Zeitpunkt.

Keine Angaben hinsichtlich des Bezugsjahres für die Preisangaben.

2-3 Studienpopulation / Zielpopulation

Männer im Alter zwischen 55 und 70 Jahren ohne krankheitsspezifische Vorgeschichte oder Verdacht auf Prostatakarzinom zum Zeitpunkt ihres Eintritts in die Studie, für welche die Daten von bis zu fünf aufeinanderfolgenden Untersuchungsjahren vorlagen.

2-4 Setting

Tangierte Versorgungsstrukturen: verschiedenartige ambulante Versorgungseinrichtungen.

Der klinische Teil der Studie wurde multizentrisch (zwölf Zentren) in den USA und in Kanada durchgeführt; der ökonomische Teil der Studie wurde den USA aufgestellt.

2-5 Spezifikation der Technologie

PSA-Assay: Abbott IMx, Hybritech Tandem-R und Diagnostic Products IRMA-Count (monoklonal) sowie Yang Laboratories Pros-Check (polyklonal).

DRU: keine expliziten Angaben (Mettlin et al.: Palpation durch Urologen).

TRUS: keine expliziten Angaben bezüglich Durchführung (Mettlin et al.: Real-Time-Verfahren, 7.0MHz-Schallkopf, biplanar; Untersuchung durch Radiologen). Volumenbestimmung größtenteils nach dem Modell des Rotationsellipsoids, bei einer kleinen Gruppe planimetrisch.

Biopsie: keine expliziten Angaben (Mettlin et al.: transrektale Stanzbiopsie (Biopty-Pistole) unter Ultraschallkontrolle oder digitaler Führung; Antibiotikaprophylaxe).

3 Gesundheitseffekte

Screening-/Diagnosestudie

3-1 Untersuchte Zielgrößen

Primäre Zielgrößen waren die Entscheidungsgrenzen mit optimaler Testgüte für den absoluten PSA-Wert (allein oder in Kombination mit der DRU), für die PSA-D, für den altersabhängigen PSA-Wert und für die PSA-V. Bei der retrospektiven Entscheidungsanalyse interessierten der Anteil positiver Screening-Testbefunde, der Anteil der Probanden mit Biopsie, die Karzinomentdeckungsrate sowie der positive Vorhersagewert (PPV) der unmittelbaren Biopsieindikatoren.

3-2 Ein-/Ausschlusskriterien

Als Einschlusskriterien wurden das Geschlecht und das Alter genannt. Ausschlusskriterien waren eine krankheitsspezifische Symptomatik bzw. eine krankheitsspezifische Vorgeschichte.

3-3 Rekrutierungsmodus

Keine expliziten Angaben (Mettlin et al.: unabhängige Rekrutierung durch die einzelnen Studienzentren über Werbeanzeigen, Vorträge, Berichterstattung in der lokalen Medien, Teilnahme an nationalen prostate cancer awareness-Programmen; mit Ausnahme eines Studienzentrums kostenloses Angebot an die Studienteilnehmer).

3-4 Teilnahmerate

Keine Angaben. Für die Analyse lagen die Daten von 2,930 Männern mit unterschiedlicher Anzahl Untersuchungen vor von insgesamt zu diesem Zeitpunkt 2,999 rekrutierten Männern.

3-5 Reproduzierbarkeit der Testergebnisse

Die PSA-Werte des polyklonalen PSA-Assays wurden zur Vergleichbarkeit mit den Werten der monoklonalen PSA-Assays adjustiert. Bezüglich der Vergleichbarkeit der eingesetzten PSA-Assay- und PSA-Antikörpertypen wurden keine Aussagen gemacht. Die Blutprobenanalysen wurden in einem Zentrallabor durchgeführt, die Gewebeproben durch denselben Gutachter überprüft. Zu den Entscheidungskriterien bezüglich der Interpretationen der DRU und TRUS wurden keine expliziten Angaben gemacht (Mettlin et al.: bei der DRU palpatorisch indurierte, knotige oder asymmetrische Prostata bzw. bei

der TRUS nicht durch Zysten, Artefakte oder vaskuläre bzw. andere normale anatomische Strukturen bedingte echoarme Areale mit einer durchschnittlichen Größe (33% der Summe aus Länge, Breite und Höhe) von >5-7cm).

3-6 Auswertung der Testergebnisse

Hinweis darauf, dass die PSA-Bestimmung nicht als Entscheidungskriterium für eine Biopsie galt: die Biopsie als Goldstandard wurde lediglich bei Probanden mit positivem DRU- und/oder TRUS-Befund durchgeführt. Bei negativem Befund für beide Untersuchungen Klassifikation als "nicht erkrankt".

Ermittlung der optimalen Entscheidungsgrenzen sowie signifikanter Unterschiede bezüglich der Testgüte der einzelnen PSA-Entscheidungsparameter mittels ROC-Kurven (AUC-Verfahren nach Hanley & McNeil, 1982). Für die Bestimmung der Entscheidungsgrenzen mit optimaler Testgüte wurden die zuletzt bestimmten PSA-Werte herangezogen; für die retrospektive Entscheidungsanalyse wurde auf den Wert der ersten PSA-Bestimmung von 2,900 Männern zurückgegriffen.

3-7 Drop-Outs

Es wird lediglich erwähnt, dass wegen mangelnder bzw. mit der Zeit abnehmender Compliance nicht für alle Männer Ergebnisse über den ganzen Zeitraum seit ihrer Rekrutierung vorlagen.

3-8 Ergebnisse der Screening-/Diagnosestudie

Hinsichtlich der Reliabilität von Volumenadjustierungen des PSA-Wertes wurden Bestimmungen der intraindividuellen Variabilität des mittels der TRUS bestimmten Prostatavolumens und der PSA-Werte durchgeführt. Sowohl die intraindividuelle Variabilität des Prostatavolumens als auch diejenige des absoluten PSA-Wertes nahmen in Abhängigkeit des Prostatavolumens zu. Im Vergleich zum durchschnittlichen Volumen (34.5ccm) lag die durchschnittliche intraindividuelle Variabilität des Prostatavolumens (+5.5ccm) bei 16% und fiel damit geringer aus als die 28%ige relative intraindividuelle Variabilität des absoluten PSA-Wertes (0.64ng/ml zu 2.25ng/ml). Die PSA-D blieb über alle Volumen-Strata stabil (95%-Perzentil bei 0.16ng/ml/ccm).

Verglichen mit einer systematischen Biopsie bei allen Probanden mit erhöhten absoluten bzw. altersabhängigen PSA-Werten konnte bei Verwendung der PSA-D mit einer Entscheidungsgrenze bei 0.12ng/ml/ccm im Wertebereich von 2.5ng/ml bis 10ng/ml (0.15ng/ml/ccm bei PSA-Werten von >2.5ng/ml, 0.07ng/ml/ccm im gesamten PSA-Wertebereich) kein signifikanter Unterschied festgestellt werden. Die optimale

Entscheidungsgrenze für einen absoluten PSA-Wert allein lag nahe 3ng/ml, in Kombination mit der DRU bei 5ng/ml. Hinsichtlich der PSA-V wies die zeitliche prozentuale Veränderung des PSA-Spiegels eine signifikant ($p < 0.05$) niedrigere Testgüte auf als die zeitliche absolute Veränderung des PSA-Spiegels (optimale Entscheidungsgrenze bei 0.6ng/ml/Jahr im gesamten PSA-Wertebereich bzw. bei 1.0ng/ml/Jahr bei PSA-Werten von >4 ng/ml) oder die zeitliche Veränderung der PSA-D (optimale Entscheidungsgrenze bei 0.02ng/ml/ccm/Jahr).

Insgesamt waren mit der Zeit 172 Prostatakarzinome in der Studienpopulation entdeckt worden (Karzinomentdeckungsrate 0.059). Die Entscheidungsanalyse bezog sich auf 171 entdeckte Karzinome und 2,900 Männer.

Screening-Strategie 1): 24.4% der Screening-Teilnehmer wiesen einen positiven Screening-Testbefund auf, 24.4% wurden biopsiert. Die Karzinomentdeckungsrate lag bei 0.046, der PPV wurde auf 0.195 ermittelt.

Screening-Strategie 2): 24.4% der Screening-Teilnehmer wiesen einen positiven Screening-Testbefund auf, 16.1% wurden biopsiert. Die Karzinomentdeckungsrate lag bei 0.041, der PPV wurde auf 0.281 ermittelt.

Screening-Strategie 3): 12.4% der Screening-Teilnehmer wiesen einen positiven Screening-Testbefund auf, 12.4% wurden biopsiert. Die Karzinomentdeckungsrate lag bei 0.025, der PPV wurde auf 0.212 ermittelt.

Screening-Strategie 4): 22.7% der Screening-Teilnehmer wiesen einen positiven Screening-Testbefund auf, 12.8% wurden biopsiert. Die Karzinomentdeckungsrate lag bei 0.037, der PPV wurde auf 0.297 ermittelt.

Screening-Strategie 5): 15.6% der Screening-Teilnehmer wiesen einen positiven Screening-Testbefund auf, 11.6% wurden biopsiert. Die Karzinomentdeckungsrate lag bei 0.034, der PPV wurde auf 0.308 ermittelt.

Screening-Strategie 6): 12.0% der Screening-Teilnehmer wiesen einen positiven Screening-Testbefund auf, 12.0% wurden biopsiert. Die Karzinomentdeckungsrate lag bei 0.033, der PPV wurde auf 0.288 ermittelt.

Screening-Strategie 7): 12.0% der Screening-Teilnehmer wiesen einen positiven Screening-Testbefund auf, 9.8% wurden biopsiert. Die Karzinomentdeckungsrate lag bei 0.032, der PPV wurde auf 0.338 ermittelt.

Screening-Strategie 8): 21.4% der Screening-Teilnehmer wiesen einen positiven Screening-Testbefund auf, 21.4% wurden biopsiert. Die Karzinomentdeckungsrate lag bei 0.044, der PPV wurde auf 0.214 ermittelt.

Screening-Strategie 2): 21.4% der Screening-Teilnehmer wiesen einen positiven Screening-Testbefund auf, 14.8% wurden biopsiert. Die Karzinomentdeckungsrate lag bei 0.039, der PPV wurde auf 0.274 ermittelt.

3-9 Effektmaße für die ökonomische Analyse

Rate der entdeckten Prostatakarzinome.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden, soweit ersichtlich, die direkten variablen Kosten für die medizinische Versorgung.

- Screening/Diagnostik: PSA-Assay, DRU, TRUS und Biopsie

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte, soweit ersichtlich, nach erstellten (Zwischen-)Produkten (Versorgungsmengen). Das Mengengerüst für die Biopsien wurde nicht spezifiziert, kann jedoch abgeleitet werden.

4-3 Monetäre Bewertung des Mengengerüsts

Soweit ersichtlich pauschales Verfahren. Die Preise pro Mengeneinheit wurden nicht explizit aufgeführt, wurden jedoch der Studie von Littrup et al. (1993)²⁹ entnommen.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Keine Angaben.

²⁹ Littrup PJ, Goodman AC, Mettlin CJ, et al.: *The Benefit and Cost of Prostate Cancer Early Detection*. CA 1993; 43(3): 134-149.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die Rate entdeckter Prostatakarzinome lag für die Screening-Strategie 1) bei 4.6%. Für Strategie 2) war sie 4.1%, für Strategie 3) 2.5%, für Strategie 4) 3.7%, für Strategie 5) 3.4%, für Strategie 6) 3.3%, für Strategie 7) 3.2%, für Strategie 8) 4.4% und für Strategie 9) 3.9%.

6-2 Ermittelte Kosten

Die Gesamtkosten pro 1,000 Screening-Teilnehmer beliefen sich bei der Screening-Strategie 1) auf US-\$ 218,600, wobei die Kosten für die Biopsie US-\$ 122,000 ausmachten. Bei der Screening-Strategie 2) betragen die entsprechenden Kosten US-\$ 177,100 bzw. US-\$ 80,500, bei der Screening-Strategie 3) US-\$ 114,320 bzw. US-\$ 62,000, bei der Screening-Strategie 4) US-\$ 134,860 bzw. US-\$ 64,000, bei der Screening-Strategie 5) US-\$ 116,080 bzw. US-\$ 58,000, bei der Screening-Strategie 6) US-\$ 111,600 bzw. US-\$ 60,000, bei der Screening-Strategie 7) US-\$ 100,600 bzw. US-\$ 49,000, bei der Screening-Strategie 8) US-\$ 199,100 bzw. US-\$ 107,000 und bei der Screening-Strategie 9) US-\$ 166,100 bzw. US-\$ 74,000.

6-3 Synthese von Kosten und Effekten

Die Kosten pro entdecktem Karzinom beliefen sich bei der Screening-Strategie 1) auf US-\$ 4,752, bei der Screening-Strategie 2) auf US-\$ 4,320, bei der Screening-Strategie 3) auf US-\$ 4,511, bei der Screening-Strategie 4) auf US-\$ 3,645, bei der Screening-Strategie 5) auf US-\$ 3,414, bei der Screening-Strategie 6) auf US-\$ 3,382, bei der Screening-Strategie 7) auf US-\$ 3,144, bei der Screening-Strategie 8) auf US-\$ 4,525 und bei der Screening-Strategie 9) auf US-\$ 4,259.

7 Behandlung von Unsicherheiten

Bei Verdoppelung des Preises für die Biopsie oder bei gleichzeitiger Verringerung des Preises für den PSA-Assay um 33% und desjenigen für die DRU um 66% kam es zu keinen Veränderungen der relativen Rangfolge der untersuchten Screening-Strategien.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren weisen darauf hin, dass in ihrer Studie ein möglicher Bias in Richtung Entdeckung von palpablen oder in bildgebenden Verfahren erkennbaren Karzinomen besteht. In bezug auf das Konzept der altersabhängigen PSA-Grenzwerte vermuten sie in ihrer Studienpopulation mit einem relativ höheren Anteil von Probanden in den mittleren Altersklassen als in der Studie von Oesterling et al. einen signifikanten population bias: für die Probanden der Studienpopulation im Alter von 50 bis 60 Jahren, 60 bis 70 Jahren und >70 Jahren wurden altersabhängige Grenzwerte (95%-Perzentile) von 6.3ng/ml, 6.8ng/ml bzw. 6.9ng/ml abgeleitet.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Nach Meinung der Autoren sind Bedenken hinsichtlich der Reliabilität der PSA-D unbegründet; eher sollte die natürliche intraindividuelle Variabilität der absoluten PSA-Werte mehr Berücksichtigung finden. Zur Verbesserung der Nützlichkeit und Einsetzbarkeit der PSA-D in der klinischen Praxis schlagen die Autoren eine modifizierte, vereinfachende Gleichung vor (erwarteter PSA-Wert = (Volumen x 0.1ng/ml/ccm) + 1ng/ml). Bezüglich des altersabhängigen PSA-Wertes (ab einem Alter von 40 Jahren) wurde eine weitere Formel entwickelt (oberer altersabhängiger PSA-Grenzwert = 2.5ng/ml + [(Alter - 40) x 0.1ng/ml/Jahr]).

Die Kosten für DRU, PSA, TRUS und Biopsie variieren zwischen verschiedenen Ländern. Aufgrund der diesbezüglichen Sensitivitätsanalyse wird jedoch eine stabile Rangfolge der Strategien bezüglich der Kosten pro entdecktem Krebsfall erwartet.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass sich die verschiedenen PSA-Entscheidungsparameter hinsichtlich der Testgüte nicht signifikant voneinander unterscheiden. Bei einem positivem Screening-Testbefund bietet sich als kosten-effektive Strategie aus Sicht der Autoren anstelle der systematischen Biopsie eine adaptierte Biopsie mit der PSA-D als unmittelbarem Entscheidungskriterium für die Biopsie an. Eine adaptierte Biopsie bei Patienten mit erhöhtem PSA-Spiegel oder positivem DRU-Befund bringt dabei die größte Reduktion an erforderlichen Biopsien mit einer akzeptablen Karzinomentdeckungsrate.

9 Kommentar

Studie zum Prävalenz- und Inzidenz-Screening mit retrospektiver Entscheidungsanalyse; explizite Gegenüberstellung von Technologiealternativen. Als interessierende Zielgrößen wurden intermediäre Outcome-Parameter gewählt. Hinsichtlich der Ermittlung der Effektivität ist die Möglichkeit eines selection bias und eines expectation bias gegeben. Die lediglich auf eine selektierte Populationsuntergruppe beschränkte Anwendung des Goldstandards steht in Verbindung mit einem work-up bias, welcher sich in den Berechnungen auswirken kann. Problematisch erscheint u.a., dass die Werte einer PSA-Bestimmung auf die mit wiederholten Untersuchungen erzielten Gesamtergebnisse der Studie bezogen wurden. Insgesamt erscheint problematisch, dass keine näheren Angaben dazu gemacht wurden, in welchen Subgruppen welche klinischen Befunde (insb. PSA-Werte) vorliegen und ob hier mit systematischen Verzerrungen zu rechnen ist. Die auch in der Entscheidungsanalyse aufgegriffenen Grenzwerte für altersabhängige PSA-Werte und für die PSA-D wurde lediglich auf der Basis von PSA-Werten zwischen 2.5ng/ml und 10ng/ml ermittelt, was nicht ausreichend hervorgehoben wird. Problematisch bei den statistischen Analysen der ROC-Kurven erscheint zweierlei: (i) die Auswahl des optimalen Cutpoints wird ausschließlich anhand von Sensitivität und Spezifität vorgenommen, ohne die Berücksichtigung klinischer oder ökonomischer Konsequenzen; (ii) bezüglich der großenteils nicht signifikanten Tests fehlen Angaben zur klinischen Relevanz und der Power. Die Entscheidungsanalysen weisen ein großes Defizit an Nachvollziehbarkeit und Transparenz auf. Dabei ist unklar, warum sie sich auf 2,900 Probanden und 171 entdeckte Karzinome und nicht auf die gesamte Studienpopulation beziehen. Problematisch kann sich auch der work-up bias in den die verschiedenen PSA-Parameter evaluierenden Entscheidungsanalysen auswirken. Der Kostenberechnung bzw. ihrer Dokumentation mangelt es erheblich an Transparenz und Nachvollziehbarkeit. Aufgrund der methodischen Einschränkungen der Publikation sind die interne Validität, Reproduzierbarkeit und Generalisierbarkeit der Analyseergebnisse stark reduziert.

10 Ähnliche Publikation:

Littrup PJ, Goodman AC, for the Investigators of the American Cancer Society National Prostate Cancer Detection Project: Economic Considerations of Prostate Cancer. The Role of Detection Specificity and Biopsy Reduction. *Cancer (Supplement)*, 1995; 75(7): 1987-1993.

In dieser Publikation greifen die Autoren die Gleichung für den Nettonutzen der von Littrup et al. (1993) veröffentlichten Kosten-Nutzen-Analyse auf:

$$V = pSeWB + (1-p)SpR - S - [pSe + (1-p)(1-Sp)]w(T+A) - (1-p)(1-Sp)F$$

mit: V = Nettonutzen pro Screening-Teilnehmer

B = Monetärer Wert des Nutzens aus der Karzinomfrüherkennung und -behandlung
 = (Medizinische Kostenersparnis aufgrund vermehrter Behandlung von Stadium A- und B-Karzinomen* + Vermindertes physisches und psychisches Leiden aufgrund der Früherkennung + Einkommensverlust) x Anstieg der Entdeckungsrate von Stadium A- und B-Karzinomen in %

* (Behandlungskosten für Stadium C- und D-Karzinome) -
 (Behandlungskosten für Stadium A- und B-Karzinome)

R = Monetärer Wert von Erleichterung und Sicherheit bei negativem Testbefund

S = Kosten für die eingesetzten Screening-Tests

T = Kosten für Staging und Behandlung als Resultat des Screenings

= (Anteil Stadium A- und B-Karzinome x Behandlungskosten von Karzinomen im Frühstadium) + (Anteil Stadium C- und D-Karzinome x Behandlungskosten von Karzinomen in fortgeschrittenem Stadium) + Staging-Kosten

A = Medizinische Kosten bei falsch positivem Screening-Befund sowie für diagnose- und behandlungsbedingte Morbidität

= (Anteil testpositiver Befunde) x (Kosten der TRUS) + [(Anteil Biopsien)(Kosten der Biopsie) + Kosten der DRU* + Kosten von biopsiebedingten Komplikationen + Kosten von unnötiger Therapie bei Überdiagnose + Kosten von behandlungsbedingten Komplikationen]

* nur bei Screening-Strategien mit alleinigem Einsatz der PSA-Bestimmung

F = Potentielle Kosten der psychischen Belastung bei falsch positivem Screening-Testbefund

p = Prävalenz von Prostatakarzinom = Anteil der Patienten mit Prostatakarzinom im Rahmen von Screening und TUR-P

Se = Sensitivität

Sp = Spezifität

w = Wahrscheinlichkeit einer Behandlung nach positivem Biopsiefund

Auf Basis der Daten der oben abgehandelten Publikation wurde eine Kosten-Effektivitäts-Analyse und eine Kosten-Nutzen-Analyse durchgeführt. Es wurden die Prävalenz sowie die Sensitivität und Spezifität der verschiedenen Screening-Alternativen berechnet. Außerdem wurden Preisanpassungen vorgenommen: Der Preis des PSA-Assays in der Basisanalyse (US-\$ 30) wurde halbiert, die Kosten für die DRU (US-\$ 30) auf ein Drittel reduziert. Weiterhin wurde angenommen, dass die Kosten für die Behandlung von Stadium A- und B-Karzinomen (radikale Prostatektomie) gesunken wären, so dass die Kosten für Staging und Behandlung als Resultat des Screenings (T) um 25% gesenkt wurden, womit sich auch geringe Veränderungen im Nutzen aus Früherkennung und -behandlung (B) ergaben. Die Kosten für die eingesetzten Screening-Tests (S) wurden auf einen Untersuchungseinsatz beschränkt, womit sich auch der monetäre Wert der Erleichterung und Sicherheit bei negativem Testbefund (R) verringerte. Daraus ergab sich folgendes Bild:

Screening-Strategie 1): Mit einer Prävalenz von 0.046, einer Sensitivität von 0.807 und einer Spezifität von 0.791 belief sich der Nettonutzen V pro Screening-Teilnehmer auf US-\$ -2,663, d.h. die Kosten waren höher als der zu erzielende Nutzen.

Screening-Strategie 2): Mit einer Prävalenz von 0.041, einer Sensitivität von 0.725 und einer Spezifität von 0.884 belief sich der Nettonutzen V pro Screening-Teilnehmer auf US-\$ -1,386.

Screening-Strategie 3): nicht aufgeführt.

Screening-Strategie 4): Mit einer Prävalenz von 0.037, einer Sensitivität von 0.643 und einer Spezifität von 0.905 belief sich der Nettonutzen V pro Screening-Teilnehmer auf US-\$ -1,149. Für die gleiche Strategie mit systematischer anstelle von adaptierter Biopsie betragen die Nettokosten US-\$ 2,575.

Screening-Strategie 5): Mit einer Prävalenz von 0.034, einer Sensitivität von 0.602 und einer Spezifität von 0.915 belief sich der Nettonutzen V pro Screening-Teilnehmer auf US-\$ -1,037. Für die gleiche Strategie mit systematischer anstelle von adaptierter Biopsie betragen die Nettokosten US-\$ 1,589.

Screening-Strategie 6): Mit einer Prävalenz von 0.033, einer Sensitivität von 0.585 und einer Spezifität von 0.909 belief sich der Nettonutzen V pro Screening-Teilnehmer auf US-\$ -1,141.

Screening-Strategie 7): Mit einer Prävalenz von 0.032, einer Sensitivität von 0.561 und einer Spezifität von 0.931 belief sich der Nettonutzen V pro Screening-Teilnehmer auf US-\$ -1,093.

Screening-Strategie 8): Mit einer Prävalenz von 0.044, einer Sensitivität von 0.778 und einer Spezifität von 0.821 belief sich der Nettonutzen V pro Screening-Teilnehmer auf US-\$ -2,259.

Screening-Strategie 9): Mit einer Prävalenz von 0.039, einer Sensitivität von 0.690 und einer Spezifität von 0.886 belief sich der Nettonutzen V pro Screening-Teilnehmer auf US-\$ -952.

Die Autoren bemerken, dass der Absolutwert des Nettonutzens V weniger wichtig ist als der beobachtbare Trend der Gleichungsergebnisse: die Nettokosten bewegen sich im Bereich der Kosten, die in der 3-Jahres-Analyse ermittelt wurden, zeigen jedoch eine positive Auswirkung der adaptierten Biopsie auf. Die Autoren kommen zu dem Schluss, dass mit geringen Verbesserungen der Spezifität der Screening-Tests - ca. 5% durch das adaptierte Biopsieverfahren - die Kosten der Frühdiagnose von Prostatakarzinom bis zu 50% verringert werden können.

Auch dieser Publikation mangelt es an Nachvollziehbarkeit und Transparenz. Ohne einen Rückgriff auf die oben abgehandelte Publikation und diejenige von Littrup et al. (1993) wäre sie nicht nachvollziehbar.

Optenberg SA & Thompson IA:
Economics of Screening for Carcinoma of the Prostate.
Urologic Clinics of North America, 1990; 17(4): 719-737.

1 Fragestellung und Evaluationsrahmen

1-1 Technologie

Screening des Prostatakarzinoms mittels DRU, Bestimmung von Serum-PSA mit verschiedenen Cutpoints und TRUS. Explizite Gegenüberstellung der Alternativen "Screening" mittels verschiedener Früherkennungsprogramme und "Kein Screening" (herkömmliche Erkennungs- bzw. Diagnosepraxis).

1-2 Fragestellung

Bestimmung der auf nationaler Ebene zu erwartenden Gesamt-, morbiditäts- und mortalitätsbedingten Kosten und Gesamtkosten der Implementierung von 4 verschiedenen Früherkennungsprogrammen beim Prostatakarzinom.

Screening-Strategie 1): DRU $-\{ \oplus \} \rightarrow$ Biopsie

Screening-Strategie 2): TRUS $-\{ \oplus \} \rightarrow$ Biopsie

Screening-Strategie 3): PSA $-\{ >4\text{ng/ml} \} \rightarrow$ Biopsie

Screening-Strategie 4): PSA $-\{ >10\text{ng/ml} \} \rightarrow$ Biopsie

kein Screening: herkömmliche Erkennungs- bzw. Diagnosepraxis

1-3 Perspektive

Keine expliziten Angaben.

1-4 Zeithorizont

Ermittelt wurden im Rahmen des Prävalenz-Screenings die Effekte und Kosten bis 5 Jahre nach Diagnose und Therapie.

1-5 Art der ökonomischen Evaluation

Kosten-Konsequenzen-Analyse; Kosten-Effektivitäts-Analyse.

2 Studiendesign und Studien- bzw. Zielpopulation

2-1 Studientyp

Synthese.

Datenquellen für klinische Parameter: klinische Entscheidungsanalyse mit Literaturdaten.

Datenquelle für Ressourcenverbrauch und Mengengerüst: Modellierung (Entscheidungsanalyse).

Datenquelle für Preise: Civilian Health and Medical Program for the Uniform Services (Prevailing Rates); Department of Health and Human Services, Health Care Financing Administration (Prospective Payment System Bill Data).

2-2 Datierung der zugrundeliegenden Daten

Der Schätzwert für die Größe der männlichen Bevölkerung der Vereinigten Staaten in der untersuchten Altersgruppe basiert auf im Jahre 1989 veröffentlichten Zensusdaten. Zeitpunkt und Zeitraum der Erhebungen von Daten zu klinischen Parametern wurden nicht explizit aufgeführt (die klinischen Parameter leiten sich aus Studien ab, welche bis 1994 veröffentlicht wurden).

Die Preisangaben beziehen sich auf die Werte des Jahres 1988.

2-3 Studienpopulation / Zielpopulation

Männer der amerikanischen Gesamtbevölkerung im Alter von 50 bis 70 Jahren ohne ernste Komorbidität (approximiert durch ernste kardiale Symptomatik).

2-4 Setting

Tangierte Versorgungsstrukturen: ambulante und stationäre Versorgung.

Keine Lokalisation der Daten zu den klinischen Parametern. Der ökonomische Teil der Studie wurde in Texas, USA durchgeführt.

2-5 Spezifikation der Technologie

Keine Präzisierungen bezüglich des PSA-Assays, der DRU, der TRUS und der Biopsie: implizite Annahme ist ein Nebeneinander der vorhandenen Verfahren und Entscheidungskriterien.

3 Gesundheitseffekte

Synthese bereits veröffentlichter Studien

3-1 In der Synthese untersuchte klinische Parameter

Primäre Zielgrößen waren die Anzahl Männer im Alter von 50 bis 70 Jahren in den Vereinigten Staaten, die altersspezifische Prävalenz von kardialer Symptomatik als Ausschlusskriterium, die altersspezifische Prävalenz von Prostatakarzinom, die Richtigpositiv-, Falschpositiv-, Richtignegativ- und Falschnegativ-Raten der Tests, die klinische Stadienverteilung der entdeckten Karzinome in Abhängigkeit des eingesetzten Testverfahrens, die pathologische Stadienverteilung von mittels radikaler Prostatektomie behandelten (klinisches Stadium B-)Karzinomen, die Wahrscheinlichkeiten von behandlungsbedingten Komplikationen und sekundären Therapien sowie die therapie- und krankheitsstadiumspezifischen 5-Jahres-Überlebensraten.

3-2 Annahmen

Die Analyse basierte auf folgenden expliziten Annahmen: Die Biopsie als Referenzstandard ist ein perfekter Test mit einer Sensitivität und Spezifität von jeweils 1.0. Als Staging-Methoden nach positiver Biopsie werden serologische Tests (PAP, PSA), ein Ausscheidungsurogramm, eine Zystoskopie sowie ein Knochenszintigramm eingesetzt. Bei 60% der Patienten mit einem klinischen Stadium A-Karzinom ist dies ein Stadium A1-Tumor, bei 40% ein Stadium A2-Tumor. Bei Patienten mit einem klinischen Stadium A1-Karzinom werden keine sofortigen Therapiemaßnahmen ergriffen (watchful waiting). Karzinompatienten mit einem klinischen Stadium A2 oder B unterziehen sich einer radikalen Prostatektomie mit anschließendem pathologischen Staging; Patienten mit klinischem Stadium C-Karzinom einer Computertomographie des Beckens und anschließenden externen Strahlentherapie (66-70Gy); Patienten mit klinischem Stadium D-Karzinom einer bilateralen Orchiektomie. Als Komplikationen der radikalen Prostatektomie wurden berücksichtigt (in Klammern therapeutische Maßnahmen): Impotenz (penile Prothese); rektale Verletzung (Kolostomie mit späterer Anastomose); Inkontinenz (künstlicher Sphinkter); Harnröhrenstriktur (Urethrotomia interna); postoperativer Tod. Als Komplikationen der Strahlentherapie wurden berücksichtigt (in Klammern therapeutische Maßnahmen): Impotenz (penilen Prothese); rektale Verletzung (Kolostomie ohne spätere Anastomose); Inkontinenz (künstlicher Sphinkter); Harnröhrenstriktur (Urethrotomia interna); behandlungsbedingter Tod. Für die bilaterale Orchiektomie wurden keine Komplikationen angenommen.

Implizite Annahme der Analyse war, dass Stadium A-Karzinome nur mittels der herkömmlichen Erkennungspraxis, nicht aber durch die verschiedenen Früherkennungsprogramme entdeckt würden. Implizit wurde außerdem von einer 100%igen Compliance bei Screening, Biopsie und Therapie ausgegangen.

3-3 Berücksichtigung von Primärstudien: Studiendesigns und Ein-/Ausschlusskriterien

Die Anzahl Männer im Alter von 50 bis 70 Jahren in den Vereinigten Staaten basierte auf Zensusdaten. Die altersspezifische Prävalenz von kardialer Symptomatik wurde aus der Framingham-Studie (prospektive Kohortenstudie) abgeleitet. Die altersspezifischen Prävalenzwerte von Prostatakarzinom beruhten auf Autopsiestudien (Fallserien). Für die Ermittlung von Werten für die Testpositiv- und Richtigpositiv-Raten der DRU wurden zur Kontrolle eines referral bias lediglich Studien, die auf Massen-Screening beruhten oder mit einer Aufforderung zur Teilnahme verbunden waren, herangezogen. Schätzer bezüglich der bei der herkömmlichen Erkennungspraxis auftretenden klinischen Inzidenz und Stadienverteilung von Prostatakarzinom basierten, soweit ersichtlich, auf den Daten des American College of Surgeons (1982)³⁰. Das Design der Studien, die zur Schätzung der anderen interessierenden klinischen Parameter herangezogen wurden, wurde nicht spezifiziert. Keine Angaben zu Ein- bzw. Ausschlusskriterien bei der Berücksichtigung der Studien.

3-4 Quellen und Suchstrategie bei der Literaturrecherche

Keine Angaben.

3-5 Validitäts- bzw. Qualitätskriterien bei der Bewertung der Primärstudien

Keine Angaben.

3-6 Methoden der Bewertung von Relevanz und Validität bzw. Qualität der Primärstudien

Keine Angaben.

3-7 Methoden der Extraktion von Daten aus den Primärstudien

Keine Angaben.

³⁰ Murphy GP, Natarajan N, Pontes JE, et al.: *The National Survey of Prostatic Cancer in the United States by the American College of Surgeons*. J Urol 1982; 127: 928.

3-8 Anzahl berücksichtigter Primärstudien

Die Werte für die altersspezifische Prävalenz von kardialer Symptomatik beruhen auf einer Studie. Für die Ermittlung der altersspezifischen Krankheitsprävalenz in der Gesamtbevölkerung wurden sechs Publikationen herangezogen. Die Ermittlung von Testpositiv- und Richtigpositiv-Raten der DRU, der TRUS sowie der PSA-Bestimmung basierte auf jeweils acht (insgesamt 22) Veröffentlichungen. Für die Ableitung der klinischen Stadienverteilung der bei der DRU entdeckten Karzinome wurden sechs Publikationen zitiert, für die Ermittlung der pathologischen Stadienverteilung klinischer Stadium B-Karzinome nach radikaler Prostatektomie fünf. Für die Ermittlung der Komplikationsraten wurde auf insgesamt acht Publikationen zurückgegriffen. Die therapie- und krankheitsstadiumspezifischen Überlebensraten basieren auf acht Veröffentlichungen. Für die Schätzer bezüglich der bei der gegenwärtigen Erkennungspraxis auftretenden klinischen Inzidenz und Stadienverteilung von Prostatakarzinom wurde, soweit ersichtlich, eine Publikation herangezogen.

3-9 Methode der Synthese der gesundheitsbezogenen Parameter

Zur Ermittlung der altersspezifischen Prävalenz von Prostatakarzinom wurde eine Gerade an die vorhandenen Datenpunkte angelegt. Die Richtigpositiv-, Falschpositiv-, Richtignegativ- und Falschnegativ-Raten der jeweiligen Tests wurden anhand der durchschnittlichen Testpositiv-Raten und der durchschnittlichen Richtigpositiv-Raten der entsprechenden Studien sowie der ermittelten Krankheitsprävalenz in der Gesamtbevölkerung abgeleitet (Vierfeldertafelmethode). Die Stadienverteilung der durch TRUS bzw. PSA-Bestimmung entdeckten Karzinome begründete auf Annahmen, die sich an der Stadienverteilung der durch DRU identifizierten Tumoren orientierten: die TRUS wurde als doppelt so sensitiv wie die DRU eingeschätzt, wobei die zusätzlich entdeckten Tumoren alle dem Stadium B angehörten; die Stadienverteilung bei PSA-Werten $>4\text{ng/ml}$ wurde gleich der Stadienverteilung bei der TRUS gesetzt; die Stadienverteilung bei PSA-Werten $>10\text{ng/ml}$ wurde gleich der Stadienverteilung bei der DRU gesetzt. Der Anteil Patienten mit behandlungsbedingten Komplikationen, der sich einer sekundären Therapie unterziehen würde, basierte auf expliziten Annahmen, ebenso die 5-Jahres-Überlebensrate der Patienten mit Stadium A1-Karzinom. Bei der Berechnung von bevölkerungsbezogenen Überlebensjahren in einem 5-Jahres-Zeitraum, welche auf den ermittelten therapie- und krankheitsstadiumspezifischen 5-Jahres-Überlebensraten basierte, wurde von einer konstanten Ereignisrate ausgegangen.

3-10 Untersuchung von Heterogenität der gesundheitsbezogenen Parameter

Nicht untersucht.

3-11 Ergebnisse der Synthese

(Werte für die Basisanalyse; in Klammern Werte im plausiblen Bereich)

Die Anzahl Männer im Alter von 50 bis 70 Jahren in den Vereinigten Staaten betrug auf der Basis von im Jahre 1989 veröffentlichten Daten 19,675,380. Nach Ausschluss der Männer mit kardialer Symptomatik wurde von einer Screening-Population mit 17,496,288 Männern ausgegangen und mit 5,701,115 (32.6%) Karzinomträgern gerechnet. Folgende Richtigpositiv-, Falschpositiv-, Richtignegativ- und Falschnegativ-Raten wurden angesetzt: bei der DRU 0.012 (0.015), 0.053 (0.040), 0.621 (0.497) bzw. 0.314 (0.251); bei der TRUS 0.073 (0.091), 0.185 (0.139), 0.489 (0.391) bzw. 0.253 (0.202); bei PSA-Spiegeln >4ng/ml 0.093 (0.116), 0.177 (0.133), 0.497 (0.398) bzw. 0.233 (0.186); bei PSA-Spiegeln >10ng/ml 0.044 (0.055), 0.036 (0.027), 0.638 (0.510) bzw. 0.282 (0.226). Die Komplikationsraten nach radikaler Prostatektomie waren folgende: Impotenz 0.25, davon 25% mit sekundärer Therapie; rektale Verletzung 0.03, davon 33% mit sekundärer Therapie; Inkontinenz 0.06, davon 50% mit sekundärer Therapie; Harnröhrenstriktur 0.18, davon 100% mit sekundärer Therapie; postoperativer Tod 0.02. Folgende Komplikationsraten bei Strahlentherapie wurden angesetzt: Impotenz 0.25, davon 25% mit sekundärer Therapie; rektale Verletzung 0.01, davon 100% mit sekundärer Therapie; Inkontinenz 0.03, davon 50% mit sekundärer Therapie; Harnröhrenstriktur 0.08, davon 100% mit sekundärer Therapie; behandlungsbedingter Tod 0.005. Die Stadienverteilung von mittels der DRU bzw. PSA-Werten >10ng/ml entdeckten Karzinomen war folgende: Stadium B 0.57, Stadium C 0.13, Stadium D 0.30; diejenige von mittels der TRUS bzw. PSA-Werten >4ng/ml identifizierten Tumoren war: Stadium B 0.79, Stadium C 0.06, Stadium D 0.15. Für die herkömmliche Erkennungspraxis wurden 24,393 neu diagnostizierte Karzinome im Jahr festgelegt, mit einem Anteil von 0.293 Stadium A-, 0.309 Stadium B-, 0.133 Stadium C- und 0.265 Stadium D-Tumoren. Folgende 5-Jahres-Überlebensraten wurden ermittelt: Stadium A1 mit watchful waiting 0.95; klinisches Stadium A2 bzw. B bei radikaler Prostatektomie 0.94 (pathologisches Stadium B), 0.81 (pathologisches Stadium C) bzw. 0.75 (pathologisches Stadium D); Stadium C bei Strahlentherapie 0.64; Stadium D bei bilateraler Orchiektomie 0.24.

3-12 Effektmaße für die ökonomische Analyse

Anzahl behandelter Patienten bzw. Anzahl Patienten pro Testergebniskategorie (richtigpositiv, falschpositiv, richtignegativ, falschnegativ); Anzahl der Todesfälle innerhalb von 5 Jahren; Anzahl Überlebensjahre in einem 5-Jahres-Zeitraum nach Diagnose und Therapie.

4 Kosten

4-1 Berücksichtigte Ressourcenveränderungen

Berücksichtigt wurden die direkten variablen Kosten für die medizinische Versorgung.

- Screening/Diagnostik: Screening-Untersuchung, Biopsie, Staging
- Behandlung: primäre und sekundäre Behandlung

4-2 Beschreibung des Mengengerüsts

Die Erstellung des Mengengerüsts erfolgte nach sehr detailliert aufgeschlüsselten (Zwischen-)Produkten (CPT-4-/DRG-Positionen). Die sich aus der Entscheidungsanalyse ergebenden Mengen wurden nicht spezifiziert.

4-3 Monetäre Bewertung des Mengengerüsts

Administrative Preise; eingesetzt wurde ein nationaler Durchschnitt der eingereichten (im Gegensatz zu den erstatteten) Gebühren. Die Preise pro Mengeneinheit wurden aufgeführt.

4-4 Währung

Amerikanische Dollar (US-\$).

5 Diskontierung

Keine Angaben.

6 Ergebnisse

6-1 Ermittelte Gesundheitseffekte

Die Anzahl behandelter (richtigpositiver: RP) Patienten bzw. falschpositiver (FP), richtignegativer (RN) und falschnegativer (FN) Screening-Testbefunde, die Anzahl

Gestorbener innerhalb von 5 Jahren sowie die Anzahl Überlebensjahre in einem 5-Jahres-Zeitraum bei 17,496,288 potentiellen Screening-Teilnehmern waren folgende:

Screening-Strategie 1): RP 209,956; FP 927,303; RN 10,865,195; FN 5,493,834; 76,666 Todesfälle; 98,141,841 Überlebensjahre.

Screening-Strategie 2): RP 1,277,229; FP 3,236,813; RN 8,555,685; FN 4,426,561; 332,647 Todesfälle; 97,337,831 Überlebensjahre.

Screening-Strategie 3): RP 1,627,155; FP 3,096,843; RN 8,695,655; FN 4,076,635; 423,784 Todesfälle; 97,053,155 Überlebensjahre.

Screening-Strategie 4): RP 769,837; FP 629,866; RN 11,162,632; FN 4,933,953; 281,110 Todesfälle; 97,515,017 Überlebensjahre.

Herkömmliche Erkennungspraxis: RP 24,393; FP 0; RN 0; FN 0; 36,556 Todesfälle; 98,115,585 Überlebensjahre.

6-2 Ermittelte Kosten

Die Gesamtkosten der Screening-Strategie 1) beliefen sich auf US-\$ 3.8 Mrd., wobei die Screening- und Behandlungskosten in der Kategorie der richtig positiven Testbefunde US-\$ 2.8 Mrd. ausmachten. Bei der Screening-Strategie 2) betragen diese Kosten US-\$ 23.6 Mrd. bzw. US-\$ 20.2 Mrd. bei der Screening-Strategie 3) US-\$ 28.0 Mrd. bzw. US-\$ 25.5 Mrd.; bei der Screening-Strategie 4) US-\$ 11.3 Mrd. bzw. US-\$ 9.5 Mrd. Die mit der herkömmlichen Erkennungspraxis verbundenen Kosten beliefen sich auf US-\$ 255 Mio.

6-3 Synthese von Kosten und Effekten

Die durchschnittlichen Kosten pro Überlebensjahr im 5-Jahres-Zeitraum beliefen sich bei der Screening-Strategie 1) auf US-\$ 39, bei der Screening-Strategie 2) auf US-\$ 242, bei der Screening-Strategie 3) auf US-\$ 288, bei der Screening-Strategie 4) auf US-\$ 116 und bei der herkömmlichen Erkennungspraxis auf US-\$ 2.6. Inkrementelle Kosten-Effektivitäts-Quotienten wurden nicht berichtet.

7 Behandlung von Unsicherheiten

Bei Annahme einer um 25% höheren Sensitivität bzw. Spezifität sowie bei Annahme von um 6% bis 10% höheren 5-Jahres-Überlebensraten (Stadium A1 +5%, Stadium B +6%, Stadium C bzw. D +10%) für die vier Screening-Strategien ergaben sich sowohl hinsichtlich der Kosten pro Lebensjahr im 5-Jahres-Zeitraum wie auch hinsichtlich der

Anzahl Gestorbener innerhalb von 5 Jahren keine Änderungen in der Rangfolge der Früherkennungsprogramme.

8 Diskussion und Schlussfolgerungen der Autoren

8-1 Bemerkungen hinsichtlich Einschränkungen/Schwächen/Bias der Analyse

Die Autoren weisen auf die teils mangelhafte und heterogene Datenlage zu den verwendeten Wahrscheinlichkeitsparametern aus der verwendeten Literatur hin. Ferner wird festgestellt, dass die Biopsie keine tatsächliche Sensitivität von 100% besitzt und damit keinen perfekten Goldstandard darstellt. Genaue Parameter sind nur aus langjährigen Screeningstudien zu erwarten.

Außerdem weisen sie darauf hin, dass die angeführten Kosten eher Unterschätzungen sind, da lediglich sich gegenseitig ausschließende Komplikationen angenommen wurden. Ihrer Ansicht nach ist bei einem sich anschließenden Inzidenz-Screening-Programm mit einer Verminderung der Mortalität, Morbidität und Kosten zu rechnen.

8-2 Bemerkungen hinsichtlich der Generalisierbarkeit der Ergebnisse (externe Validität)

Die Autoren bemerken, dass ihre Schlussfolgerungen nur bei allgemeiner Akzeptanz der für ihr Entscheidungsmodell herangezogenen Wahrscheinlichkeiten gelten können. Ferner sind die Ergebnisse nicht auf ein Follow-Up-Screening übertragbar.

8-3 Schlussfolgerungen

Die Autoren kommen zu dem Schluss, dass eine auf den alleinigen Kosten der Screening-Untersuchungen basierende Kostenanalyse irreführend ist, da weitaus mehr Kosten bei der Behandlung als bei der Diagnose von Prostatakarzinom anfallen. Die Annahme eines auf der TRUS basierenden Screening-Programmes würde den aktuellen Anteil der Prostatakarzinomdiagnose und -behandlung am Gesamtbudget des amerikanischen Gesundheitswesens von 0.06% auf 5% erhöhen. Gleichzeitig würde die behandlungsbedingte Mortalität um das 100fache ansteigen. Die extrem hohen Gesamtkosten der Screening-Programme beim Prostatakarzinom in Verbindung mit einer bedeutsamen behandlungsbedingten Morbidität und Mortalität und die fehlende Erkennbarkeit einer gewinnbringenden Kontrolle des Prostatakarzinoms vorhersagen zu können, lassen die Früherkennung in den Augen der Autoren als recht problematisch erscheinen.

9 Kommentar

Entscheidungsanalyse zum Prävalenz-Screening mit expliziter Gegenüberstellung von Technologiealternativen. Hinsichtlich der Ermittlung der Effektivität erscheint das von den Autoren beschriebene Modell sorgfältig aufgebaut und weitgehend nachvollziehbar. Die Transparenz in bezug auf die Ableitung der mit den einzelnen Screening-Tests verbundenen Wahrscheinlichkeiten weist Defizite auf. Text und Tabellen sind zum Teil inkongruent, wobei die Angaben der Tabellen größtenteils korrekter erscheinen. Bei der Berechnung der 5-Jahres-Überlebenszeit wurde für die Verstorbenen eine mittlere Überlebenszeit von zwei (anstelle von zweieinhalb) Jahren angesetzt, was einer konstanten Ereignisrate über vier Jahre entspricht. Es muss darauf hingewiesen werden, dass die Parameterschätzer der Basisanalyse nicht aus einer umfassenden Literaturrecherche und -analyse abgeleitet wurden; hier ist die Möglichkeit eines selection bias bezüglich der in die Analyse eingehenden Resultate der Synthese gegeben. Auch fehlen präzisierende Angaben zur Art des Screenings (Prävalenz- bzw. Inzidenz-Screening) in den Studienpopulationen, auf deren Daten die Ermittlung der einzelnen Testwahrscheinlichkeiten basiert. Die Ermittlung der Kosten ist weitgehend transparent. Hier erscheint es fraglich, für eine Zielpopulation von 50- bis 70jährigen durchschnittliche Versorgungspreise für >69jährige heranzuziehen. Palliative Kosten wurden für den 5-Jahreszeitraum nicht ermittelt. Die Berechnungen sind zum Teil nicht nachvollziehbar bzw. es finden sich Inkonsistenzen zwischen den Kostenangaben in Text, Abbildungen und Tabellen. Hinsichtlich der Diagnostik bei der herkömmlichen Diagnosepraxis wurde außer den Kosten für die Biopsie kein weiterer Ressourcenverbrauch angesetzt. Bei der Behandlung von Unsicherheiten mangelt es an Transparenz hinsichtlich der veränderten Testwahrscheinlichkeiten: die angegebenen Werte für die Richtigpositiv- und Falschnegativ-Raten bzw. für die Falschpositiv- und Richtignegativ-Raten führen nicht zu der angeführten 25%igen Erhöhung der Sensitivität bzw. Spezifität der Tests. Das entscheidende Ergebnis der Kosten-Effektivitäts-Analyse, die Angabe des Kosten-Effektivitäts-Quotienten fehlt. Da keine multiplen (sequentiellen und parallelen) Teststrategien in Betracht gezogen wurden, ist die Übertragbarkeit auf die Anwendung realistischer Screening-Strategien stark eingeschränkt.

Tabelle 4: Transparenzkatalog Onkologie: Indikationsstellungen NSCLC und SPN

Transparenzkriterien	Nicht-kleinzelliges Lungenkarzinom Solitäre Lungenknoten						
	USA Coleman 1994	USA Gambhir 1996	USA Scott 1998	USA Gambhir 1998	CH Von Schulthess 1998	USA ECRI 1998	Japan Kosuda 2000
Fragestellung							
1. Fragestellung formuliert	½	1	1	1	1	1	1
2. Darstellung des Problemkontextes	½	1	½	½	½	½	1
Evaluationsrahmen							
3. Technologiespezifikation	1	1	1	1	½	1	1
4. Anzahl Alternativtechnologien	1	1	1	1	1	1	1
5. Begründung Alternativtechnologien	½	1	1	1	½	½	1
6. Beschreibung der Zielpopulation	1	1	½	1	½	½	1
7. Zeithorizont	0	0	0	0	0	0	0
8. Ökonomischer Evaluationstyp	0	½	1	1	½	1	1
9. Kosten + Effekte	½	1	1	1	½	1	1
10. Perspektive	0	0	0	0	0	0	0
Analysemethoden und Modellierung							
11. Modellbeschreibung	1	1	1	1	n.r.	1	1
12. Modelldokumentation	½	1	1	1	n.r.	1	1
13. (Modell)annahmen	1	1	1	1	n.r.	1	1
14. Pfadwahrscheinlichkeiten	½	1	1	1	n.r.	1	1
Gesundheitseffekte							
15. Gesundheitszustände	1	1	1	1	½	1	1
16. Quellen der Effektdaten	½	1	1	1	1	1	1
17. Design + Auswertungsmethoden	n.r.	n.r.	n.r.	n.r.	½	n.r.	n.r.
18. Effektparameter	½	1	1	1	n.r.	0	0
19. Präferenz: Bewertungsmethodik	n.r.	0	0	0	n.r.	0	0
20. Präferenz: Datenquellen	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
21. Evidenz	½	½	½	½	½	½	½
Kosten							
22. Angabe des Mengengerüsts	1	1	1	1	1	1	1
23. Mengengerüst: Quellen/Methoden	½	1	1	1	0	1	1
24. Angabe des Preisgerüsts	½	1	1	1	1	1	1
25. Preisgerüst: Quellen + Methoden	½	1	1	1	0	1	1
26. Perspektive/Zeithorizont	0	0	0	0	0	0	0
27. Indirekte Kosten	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
28. Währung	1	1	1	1	1	1	1
29. Währungskonversion	n.r.	n.r.	n.r.	n.r.	0	n.r.	1
30. Preisanpassung	0	0	0	0	0	0	0
Diskontierung							
31. Effekte und Kosten	n.r.	0	0	0	n.r.	0	0
32. Referenzjahr	0	0	½	0	0	0	1
33. Angabe der Diskontrate	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
34. Begründung	0	0	0	0	0	0	0
Ergebnispräsentation							
35. Modellvalidierung	0	0	0	½	n.r.	½	0
36. Absolute Effekte/Kosten pro Kopf	0	½	0	1	0	1	1
37. Inkrementelle Effekte/Kosten	½	1	0	½	0	1	1
38. Relation Kosten – Effekte	0	0	1	1	0	0	1
39. Rohe klinische Effekte	1	1	1	1	½	1	1
40. Disaggregierte Darstellung	1	1	1	1	0	1	1
41. Populationsaggregierte Darstellung	1	½	0	1	½	1	½
Behandlung von Unsicherheiten							
42. Univariate Sensitivitätsanalyse	0	1	1	1	0	1	1
43. Multivariate Sensitivitätsanalyse	0	1	1	1	0	1	0
44. Strukturelle Sensitivitätsanalyse	n.r.	0	0	1	0	1	0
45. Variation der Werte bzw. Strukturen	n.r.	1	1	1	n.r.	1	1
46. Ergebnisse Sensitivitätsanalysen	n.r.	1	1	1	n.r.	½	1
47. Statistische Inferenzmethoden	0	0	0	0	0	½	0
Diskussion							
48. Datenqualität	0	1	½	½	0	½	½
49. Parameterschätzung	0	1	½	1	0	½	1
50. Strukturelle Modellannahmen	n.r.	½	½	1	n.r.	1	½
51. Einschränkungen/Schwächen	0	1	1	½	0	½	½
52. Generalisierbarkeit	½	0	0	0	0	0	½
53. Ethische Aspekte/Verteilungsfragen	0	0	0	½	0	0	0
54. Kontext mit Gesundheitsprogrammen	0	0	½	½	0	0	½
Schlussfolgerungen							
55. Basierend auf Ergebnissen	1	1	1	1	1	1	1
56. Antwort auf Fragestellung	½	1	1	1	1	1	1

1 – Kriterium erfüllt, ½ - Kriterium teilweise erfüllt, 0 – Kriterium nicht erfüllt oder fehlerhaft, n.r. - nicht relevant

Tabelle 5: Qualitätskatalog Onkologie: Indikationsstellungen NSCLC und SPN

Qualitätskriterien	Nicht klein-zelliges Lungenkarzinom Solitäre Lungenknoten						
	USA Coleman 1994	USA Gambhir 1996	USA Scott 1998	USA Gambhir 1998	CH Von Schulthess 1998	USA ECRI 1998	Japan Kosuda 2000
Fragestellung							
1. Präzision der Fragestellung	½	½	½	½	½	½	½
2. Darstellung des Problemkontextes	½	½	½	½	½	½	1
Evaluationsrahmen							
3. Technologiespezifikation	½	1	½	1	0	½	0
4. Anzahl Alternativtechnologien	1	½	½	1	½	1	1
5. Begründung Alternativtechnologien	0	1	1	1	0	½	1
6. Beschreibung der Zielpopulation	½	1	½	1	½	½	1
7. Zeithorizont	½	½	½	½	½	1	½
8. Ökonomischer Evaluationstyp	0	½	1	1	½	1	1
9. Kosten + Effekte	½	1	1	1	½	1	1
10. Perspektive	½	½	½	½	0	½	½
Analysemethoden und Modellierung							
11. Modellbeschreibung	1	1	1	1	n.r.	1	1
12. Modelldokumentation	½	1	½	1	n.r.	½	1
13. (Modell)annahmen	½	1	1	1	n.r.	1	1
14. Pfadwahrscheinlichkeiten	½	½	1	1	n.r.	1	1
Gesundheitseffekte							
15. Gesundheitszustände	½	½	½	½	0	1	1
16. Quellen der Effektdaten	½	½	1	1	1	1	1
17. Design/ Auswertungsmethoden	n.r.	n.r.	n.r.	n.r.	0	n.r.	n.r.
18. Effektparameter	½	1	1	1	n.r.	0	0
19. Präferenz: Bewertungsmethodik	n.r.	0	0	0	n.r.	0	0
20. Präferenz: Datenquellen	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
21. Evidenz	½	½	½	½	0	½	½
Kosten							
22. Angabe des Mengengerüsts	1	½	½	½	1	½	1
23. Mengengerüst: Quellen/Methoden	½	1	1	1	0	1	½
24. Angabe des Preisgerüsts	½	1	1	1	0	1	½
25. Preisgerüst: Quellen + Methoden	½	1	1	1	0	1	1
26. Perspektive/Zeithorizont	0	0	0	0	0	½	0
27. Indirekte Kosten	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
28. Währung	1	1	1	1	1	1	1
29. Währungskonversion	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	1
30. Preisanpassung	0	0	0	0	0	0	0
Diskontierung							
31. Effekte und Kosten	n.r.	0	0	0	n.r.	0	0
32. Referenzjahr	0	0	½	0	0	0	1
33. Angabe der Diskontrate	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
34. Begründung	0	0	0	0	0	0	0
Ergebnispräsentation							
35. Modellvalidierung	0	0	0	½	n.r.	½	0
36. Absolute Effekte/Kosten pro Kopf	0	½	0	1	0	1	1
37. Inkrementelle Effekte/Kosten	½	1	0	½	0	1	1
38. Relation Kosten – Effekte	0	0	1	½	n.r.	0	1
39. Rohe klinische Effekte	1	1	1	1	½	1	1
40. Disaggregierte Darstellung	0	1	1	1	0	1	1
41. Populationsaggregierte Darstellung	1	½	0	1	½	1	½
Behandlung von Unsicherheiten							
42. Univariate Sensitivitätsanalyse	0	1	1	1	0	1	1
43. Multivariate Sensitivitätsanalyse	0	1	1	1	0	½	0
44. Strukturelle Sensitivitätsanalyse	n.r.	0	0	½	0	1	0
45. Variation der Werte bzw. Strukturen	n.r.	1	1	1	n.r.	1	1
46. Ergebnisse Sensitivitätsanalysen	n.r.	1	1	1	n.r.	½	1
47. Statistische Inferenzmethoden	0	0	0	0	0	½	0
Diskussion							
48. Datenqualität	0	1	0	½	0	0	½
49. Parameterschätzung	0	1	½	1	0	½	½
50. Strukturelle Modellannahmen	n.r.	½	½	½	n.r.	½	½
51. Einschränkungen+Schwächen	0	½	½	½	0	½	½
52. Generalisierbarkeit	0	0	0	0	0	0	0
53. Ethische Aspekte/Verteilungsfragen	0	0	0	½	0	0	0
54. Kontext mit Gesundheitsprogrammen	0	0	½	0	0	0	½
Schlußfolgerungen							
55. Basierend auf Ergebnissen	½	1	1	1	0	1	1
56. Antwort auf Fragestellung	½	1	1	1	½	1	1

1 – Kriterium erfüllt, ½ – Kriterium teilweise erfüllt, 0 – Kriterium nicht erfüllt oder fehlerhaft, n.r. - nicht relevant

Tabelle 6: Transparenzkatalog Onkologie: Sonstige Indikationsstellungen

Transparenzkriterien	Lymphoretikuläres System				Kolorektales Karzinom		Melanom	Über-sicht
	USA Adler 1994	USA Adler 1997	USA Hoh 1997	D Klose (1999a)	USA Larson 1994	USA Valk 1999	CH Von Schult- hess 1998	USA Valk 1996
Fragestellung								
1. Fragestellung formuliert	½	1	1	1	1	1	1	1
2. Darstellung des Problemkontextes	1	1	½	½	½	1	½	1
Evaluationsrahmen								
3. Technologiespezifikation	1	1	1	1	1	1	½	1
4. Anzahl Alternativtechnologien	1	½	1	1	1	1	1	1
5. Begründung Alternativtechnologien	1	1	1	½	½	½	½	½
6. Beschreibung der Zielpopulation	1	1	1	1	1	1	½	½
7. Zeithorizont	0	0	0	½	0	½	0	½
8. Ökonomischer Evaluationstyp	1	0	0	1	0	0	½	1
9. Kosten + Effekte	½	1	1	1	½	1	½	1
10. Perspektive	0	0	0	1	0	0	0	0
Analysemethoden und Modellierung								
11. Modellbeschreibung	1	n.r.	n.r.	n.r.	1	n.r.	n.r.	n.r.
12. Modelldokumentation	1	n.r.	n.r.	n.r.	½	n.r.	n.r.	n.r.
13. (Modell)annahmen	1	n.r.	n.r.	1	½	1	n.r.	n.r.
14. Pfadwahrscheinlichkeiten	1	n.r.	n.r.	n.r.	½	n.r.	n.r.	n.r.
Gesundheitseffekte								
15. Gesundheitszustände	1	1	1	1	1	1	0	1
16. Quellen der Effektdaten	1	1	1	1	½	1	½	1
17. Design + Auswertungsmethoden	n.r.	1	1	1	n.r.	1	0	0
18. Effektparameter	½	n.r.	n.r.	n.r.	½	n.r.	n.r.	n.r.
19. Präferenz: Bewertungsmethodik	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
20. Präferenz: Datenquellen	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
21. Evidenz	1	1	1	1	½	1	½	1
Kosten								
22. Angabe des Mengengerüsts	1	½	1	1	1	1	0	1
23. Mengengerüst: Quellen/Methoden	1	½	1	1	½	1	0	1
24. Angabe des Preisgerüsts	½	1	1	1	½	1	½	½
25. Preisgerüst: Quellen + Methoden	½	½	½	1	½	1	0	1
26. Perspektive/Zeithorizont	0	0	0	0	0	0	0	0
27. Indirekte Kosten	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
28. Währung	1	1	1	1	1	1	1	1
29. Währungskonversion	n.r.	n.r.	n.r.	1	n.r.	n.r.	0	n.r.
30. Preisanpassung	0	0	0	1	0	0	0	0
Diskontierung								
31. Effekte und Kosten	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
32. Referenzjahr	0	0	0	1	0	1	0	0
33. Angabe der Diskontrate	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
34. Begründung	0	0	0	1	0	0	0	0
Ergebnispräsentation								
35. Modellvalidierung	0	n.r.	n.r.	n.r.	0	n.r.	n.r.	n.r.
36. Absolute Effekte/Kosten pro Kopf	0	½	½	½	0	0	0	½
37. Inkrementelle Effekte/Kosten	½	½	0	½	½	½	0	0
38. Relation Kosten – Effekte	0	0	0	1	0	1	0	1
39. Rohe klinische Effekte	1	1	1	1	1	1	0	1
40. Disaggregierte Darstellung	1	1	½	1	1	½	0	½
41. Populationsaggregierte Darstellung	1	1	1	½	1	1	½	1
Behandlung von Unsicherheiten								
42. Univariate Sensitivitätsanalyse	0	0	0	1	0	0	0	0
43. Multivariate Sensitivitätsanalyse	0	0	0	1	0	0	0	0
44. Strukturelle Sensitivitätsanalyse	n.r.	0	0	1	n.r.	0	0	n.r.
45. Variation der Werte bzw. Strukturen	n.r.	n.r.	n.r.	1	n.r.	n.r.	n.r.	n.r.
46. Ergebnisse Sensitivitätsanalysen	n.r.	n.r.	n.r.	1	n.r.	n.r.	n.r.	n.r.
47. Statistische Inferenzmethoden	0	½	0	½	0	1	0	0
Diskussion								
48. Datenqualität	0	½	½	1	0	½	0	0
49. Parameterschätzung	½	n.r.	0	1	0	n.r.	n.r.	0
50. Strukturelle Modellannahmen	n.r.	n.r.	n.r.	1	n.r.	n.r.	n.r.	n.r.
51. Einschränkungen/Schwächen	0	½	½	1	0	½	0	1
52. Generalisierbarkeit	0	0	½	½	0	½	0	0
53. Ethische Aspekte/Verteilungsfragen	0	0	0	0	0	0	0	0
54. Kontext mit Gesundheitsprogrammen	½	½	0	1	0	0	½	0
Schlussfolgerungen								
55. Basierend auf Ergebnissen	1	1	1	1	1	1	1	1
56. Antwort auf Fragestellung	½	1	1	1	½	1	1	1

1 – Kriterium erfüllt, ½ - Kriterium teilweise erfüllt, 0 – Kriterium nicht erfüllt oder fehlerhaft, n.r. - nicht relevant

Tabelle 7: Qualitätskatalog Onkologie: Sonstige Indikationsstellungen

Qualitätskriterien	Lymphoretikuläres System				Kolorektales Karzinom		Melanom	Über- sicht
	USA	USA	USA	D	USA	USA	CH	USA
	Adler 1994	Adler 1997	Hoh 1997	Klose (1999a)	Larson 1994	Valk 1999	Von Schult hess 1998	Valk 1996
Fragestellung								
1. Präzision der Fragestellung	½	½	1	1	½	1	½	1
2. Darstellung des Problemkontextes	1	½	½	½	½	½	½	½
Evaluationsrahmen								
3. Technologiespezifikation	½	1	1	½	½	1	½	½
4. Anzahl Alternativtechnologien	1	1	1	1	1	1	½	½
5. Begründung Alternativtechnologien	1	1	1	½	½	½	0	½
6. Beschreibung der Zielpopulation	½	1	½	1	½	½	½	½
7. Zeithorizont	½	½	½	½	½	½	½	½
8. Ökonomischer Evaluationstyp	1	0	0	1	0	0	½	1
9. Kosten + Effekte	½	1	1	1	½	1	½	1
10. Perspektive	½	½	½	½	½	1	0	½
Analysenmethoden und Modellierung								
11. Modellbeschreibung	½	n.r.	n.r.	n.r.	1	n.r.	n.r.	n.r.
12. Modelldokumentation	½	n.r.	n.r.	n.r.	½	n.r.	n.r.	n.r.
13. (Modell)annahmen	½	n.r.	n.r.	1	½	1	n.r.	n.r.
14. Pfadwahrscheinlichkeiten	½	n.r.	n.r.	n.r.	½	n.r.	n.r.	n.r.
Gesundheitseffekte								
15. Gesundheitszustände	½	1	1	1	½	1	0	1
16. Quellen der Effektdaten	½	1	1	1	0	1	½	1
17. Design/ Auswertungsmethoden	n.r.	½	1	½	n.r.	1	0	0
18. Effektparameter	½	n.r.	n.r.	n.r.	½	n.r.	n.r.	n.r.
19. Präferenz: Bewertungsmethodik	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
20. Präferenz: Datenquellen	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
21. Evidenz	½	1	1	1	0	1	0	1
Kosten								
22. Angabe des Mengengerüsts	1	½	1	1	1	1	0	½
23. Mengengerüst: Quellen/Methoden	½	½	1	1	½	1	0	1
24. Angabe des Preisgerüsts	½	½	0	1	½	½	0	½
25. Preisgerüst: Quellen + Methoden	½	½	0	1	½	1	0	1
26. Perspektive/Zeithorizont	0	0	½	½	0	0	0	0
27. Indirekte Kosten	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
28. Währung	1	1	1	1	1	1	1	1
29. Währungskonversion	n.r.	n.r.	n.r.	½	n.r.	n.r.	n.r.	n.r.
30. Preisanpassung	0	0	0	1	0	0	0	0
Diskontierung								
31. Effekte und Kosten	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
32. Referenzjahr	0	0	0	1	0	1	0	0
33. Angabe der Diskontrate	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
34. Begründung	0	0	0	1	0	0	0	0
Ergebnispräsentation								
35. Modellvalidierung	0	n.r.	n.r.	n.r.	0	n.r.	n.r.	n.r.
36. Absolute Effekte/Kosten pro Kopf	0	½	½	½	0	0	0	½
37. Inkrementelle Effekte/Kosten	½	½	0	½	½	½	0	0
38. Relation Kosten – Effekte	0	n.r.	n.r.	1	0	1	n.r.	½
39. Rohe klinische Effekte	1	1	1	1	1	1	½	1
40. Disaggregierte Darstellung	½	½	½	1	0	½	0	½
41. Populationsaggregierte Darstellung	1	1	1	½	1	1	½	1
Behandlung von Unsicherheiten								
42. Univariate Sensitivitätsanalyse	0	0	0	1	0	0	0	0
43. Multivariate Sensitivitätsanalyse	0	0	0	1	0	0	0	0
44. Strukturelle Sensitivitätsanalyse	n.r.	0	0	1	n.r.	0	0	n.r.
45. Variation der Werte bzw. Strukturen	n.r.	n.r.	n.r.	1	n.r.	n.r.	n.r.	n.r.
46. Ergebnisse Sensitivitätsanalysen	n.r.	n.r.	n.r.	1	n.r.	n.r.	n.r.	n.r.
47. Statistische Inferenzmethoden	0	½	0	½	0	1	0	0
Diskussion								
48. Datenqualität	0	½	½	1	0	½	0	0
49. Parameterschätzung	½	n.r.	0	1	0	n.r.	n.r.	0
50. Strukturelle Modellannahmen	n.r.	n.r.	n.r.	1	n.r.	n.r.	n.r.	n.r.
51. Einschränkungen+Schwächen	0	½	½	1	0	½	0	½
52. Generalisierbarkeit	0	0	½	1	0	½	0	0
53. Ethische Aspekte/Verteilungsfragen	0	0	0	0	0	0	0	0
54. Kontext mit Gesundheitsprogrammen	½	½	0	1	0	0	½	0
Schlußfolgerungen								
55. Basierend auf Ergebnissen	½	1	1	1	½	½	0	½
56. Antwort auf Fragestellung	½	1	1	1	½	1	½	1

1 – Kriterium erfüllt, ½ - Kriterium teilweise erfüllt, 0 – Kriterium nicht erfüllt oder fehlerhaft, n.r. - nicht relevant

Tabelle 8: Transparenzkatalog Kardiologie

Transparenzkriterien	Myokardviabilität		Kardiale Perfusionsmessung			
	USA Besozzi 1994	NL Hunink 1995	USA Gould 1989	USA Patterson 1995	USA Maddahi 1997	USA Garber 1999
Fragestellung						
1. Fragestellung formuliert	1	1	1	1	1	1
2. Darstellung des Problemkontextes	1	0	1	1	1	½
Evaluationsrahmen						
3. Technologiespezifikation	1	0	½	1	1	1
4. Anzahl Alternativtechnologien	1	1	1	1	1	1
5. Begründung Alternativtechnologien	½	0	½	½	½	0
6. Beschreibung der Zielpopulation	1	1	½	½	1	1
7. Zeithorizont	0	1	1	1	0	1
8. Ökonomischer Evaluationstyp	0	1	1	1	0	½
9. Kosten + Effekte	1	1	½	1	1	1
10. Perspektive	0	1	0	1	0	½
Analysemethoden und Modellierung						
11. Modellbeschreibung	1	½	1	1	1	1
12. Modelldokumentation	n.r.	½	1	1	1	½
13. (Modell)annahmen	½	1	1	1	½	1
14. Pfadwahrscheinlichkeiten	n.r.	½	½	1	1	1
Gesundheitseffekte						
15. Gesundheitszustände	1	1	½	1	1	1
16. Quellen der Effektdaten	½	½	1	1	1	1
17. Design + Auswertungsmethoden	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
18. Effektparameter	½	½	½	1	1	½
19. Präferenz: Bewertungsmethodik	0	0	0	½	n.r.	½
20. Präferenz: Datenquellen	n.r.	n.r.	n.r.	1	n.r.	1
21. Evidenz	½	½	n.r.	1	½	1
Kosten						
22. Angabe des Mengengerüsts	1	0	1	1	1	1
23. Mengengerüst: Quellen + Methoden	½	n.r.	1	1	1	1
24. Angabe des Preisgerüsts	½	1	1	1	½	1
25. Preisgerüst: Quellen + Methoden	½	½	½	½	1	1
26. Perspektive/Zeithorizont	0	0	½	1	0	0
27. Indirekte Kosten	n.r.	n.r.	1	0	n.r.	n.r.
28. Währung	1	1	1	1	0	1
29. Währungskonversion	n.r.	1	n.r.	n.r.	n.r.	n.r.
30. Preisanpassung	0	0	½	0	0	1
Diskontierung						
31. Effekte und Kosten	n.r.	0	0	0	n.r.	1
32. Referenzjahr	0	1	0	0	0	1
33. Angabe der Diskontrate	n.r.	n.r.	n.r.	n.r.	n.r.	1
34. Begründung	0	0	0	0	0	0
Ergebnispräsentation						
35. Modellvalidierung	0	0	0	½	0	0
36. Absolute Effekte/Kosten pro Kopf	0	½	0	1	0	1
37. Inkrementelle Effekte/Kosten	½	0	0	0	0	0
38. Relation Kosten – Effekte	0	1	0	1	0	1
39. Rohe klinische Effekte	1	1	½	1	1	1
40. Disaggregierte Darstellung	1	½	0	1	0	1
41. Populationsaggregierte Darstellung	1	½	½	½	1	0
Behandlung von Unsicherheiten						
42. Univariate Sensitivitätsanalyse	0	0	0	1	½	1
43. Multivariate Sensitivitätsanalyse	0	0	0	1	0	0
44. Strukturelle Sensitivitätsanalyse	n.r.	0	0	0	n.r.	0
45. Variation der Werte bzw. Strukturen	n.r.	n.r.	n.r.	1	1	1
46. Ergebnisse Sensitivitätsanalysen	n.r.	n.r.	n.r.	1	1	1
47. Statistische Inferenzmethoden	n.r.	0	0	0	0	0
Diskussion						
48. Datenqualität	0	0	0	½	0	½
49. Parameterschätzung	0	0	0	1	0	0
50. Struktureller Modellannahmen	n.r.	0	0	0	n.r.	n.r.
51. Einschränkungen + Schwächen	0	0	0	½	0	½
52. Generalisierbarkeit	0	0	0	½	1	0
53. Ethische Aspekte/Verteilungsfragen	0	0	0	0	0	0
54. Kontext mit Gesundheitsprogrammen	0	0	0	0	1	0
Schlußfolgerungen						
55. Basierend auf Ergebnissen	1	1	1	1	1	1
56. Antwort auf Fragestellung	½	1	1	1	1	1

1 – Kriterium erfüllt, ½ - Kriterium teilweise erfüllt, 0 – Kriterium nicht erfüllt oder fehlerhaft, n.r. - nicht relevant

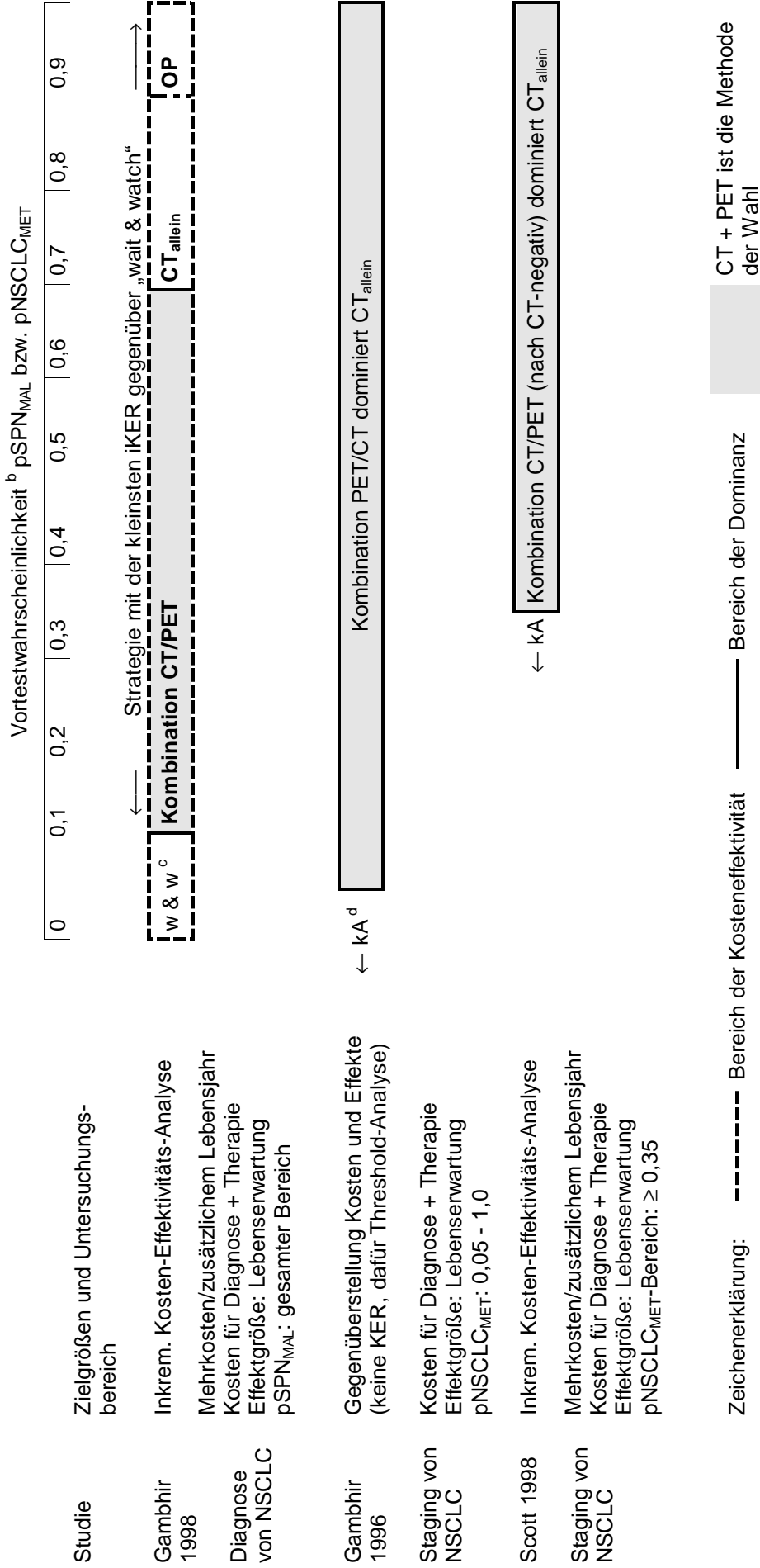
Tabelle 9: Qualitätskatalog Kardiologie

Qualitätskriterien	Myokardviabilität		Kardiale Perfusionsmessung			
	USA Besozzi 1994	NL Hunink 1995	USA Gould 1989	USA Patterson 1995	USA Maddahi 1997	USA Garber 1999
Fragestellung						
1. Präzision der Fragestellung	½	1	1	½	½	1
2. Darstellung des Problemkontextes	½	0	1	1	½	½
Evaluationsrahmen						
3. Technologiespezifikation	½	0	½	½	½	½
4. Anzahl Alternativtechnologien	½	1	1	1	1	1
5. Begründung Alternativtechnologien	½	0	0	½	½	0
6. Beschreibung der Zielpopulation	1	½	½	½	1	1
7. Zeithorizont	0	½	1	1	½	1
8. Ökonomischer Evaluationstyp	0	1	1	1	0	½
9. Kosten + Effekte	1	1	½	1	1	1
10. Perspektive	½	½	½	½	½	½
Analysemethoden und Modellierung						
11. Modellbeschreibung	½	½	½	1	½	1
12. Modelldokumentation	n.r.	½	1	1	½	½
13. (Modell)annahmen	½	½	1	1	½	1
14. Pfadwahrscheinlichkeiten	n.r.	½	½	½	½	1
Gesundheitseffekte						
15. Gesundheitszustände	½	½	½	1	1	½
16. Quellen der Effektdaten	½	½	1	1	1	1
17. Design + Auswertungsmethoden	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
18. Effektparameter	½	½	½	1	½	½
19. Präferenz: Bewertungsmethodik	n.r.	0	0	0	n.r.	½
20. Präferenz: Datenquellen	n.r.	n.r.	n.r.	1	n.r.	1
21. Evidenz	½	½	n.r.	1	½	½
Kosten						
22. Angabe des Mengengerüsts	1	0	1	½	½	½
23. Mengengerüst: Quellen + Methoden	½	n.r.	1	1	1	1
24. Angabe des Preisgerüsts	½	0	½	½	½	½
25. Preisgerüst: Quellen + Methoden	½	½	½	½	½	1
26. Perspektive/Zeithorizont	0	0	½	1	½	½
27. Indirekte Kosten	n.r.	n.r.	½	0	n.r.	n.r.
28. Währung	1	1	1	1	0	1
29. Währungskonversion	n.r.	½	n.r.	n.r.	n.r.	n.r.
30. Preisanpassung	0	0	½	0	0	½
Diskontierung						
31. Effekte und Kosten	n.r.	0	0	0	n.r.	½
32. Referenzjahr	0	1	0	0	0	1
33. Angabe der Diskontrate	n.r.	n.r.	n.r.	n.r.	n.r.	1
34. Begründung	0	0	0	0	0	0
Ergebnispräsentation						
35. Modellvalidierung	0	0	0	½	0	0
36. Absolute Effekte/Kosten pro Kopf	0	½	0	1	0	1
37. Inkrementelle Effekte/Kosten	½	0	0	0	0	0
38. Relation Kosten – Effekte	n.r.	1	0	½	0	1
39. Rohe klinische Effekte	1	1	½	1	1	1
40. Disaggregierte Darstellung	0	½	0	1	0	1
41. Populationsaggregierte Darstellung	1	½	½	½	1	0
Behandlung von Unsicherheiten						
42. Univariate Sensitivitätsanalyse	0	0	0	1	½	1
43. Multivariate Sensitivitätsanalyse	0	0	0	1	0	0
44. Strukturelle Sensitivitätsanalyse	n.r.	0	0	0	n.r.	0
45. Variation der Werte bzw. Strukturen	n.r.	n.r.	n.r.	1	1	1
46. Ergebnisse Sensitivitätsanalysen	n.r.	n.r.	n.r.	1	1	1
47. Statistische Inferenzmethoden	n.r.	0	0	0	0	0
Diskussion						
48. Datenqualität	0	0	0	½	0	½
49. Parameterschätzung	0	0	0	1	0	0
50. Struktureller Modellannahmen	n.r.	0	0	0	n.r.	n.r.
51. Einschränkungen + Schwächen	0	0	0	½	0	½
52. Generalisierbarkeit	0	0	0	½	½	0
53. Ethische Aspekte/Verteilungsfragen	0	0	0	0	0	0
54. Kontext mit Gesundheitsprogrammen	0	0	0	0	1	0
Schlußfolgerungen						
55. Basierend auf Ergebnissen	½	½	1	1	½	1
56. Antwort auf Fragestellung	½	1	1	1	1	1

1 – Kriterium erfüllt, ½ - Kriterium teilweise erfüllt, 0 – Kriterium nicht erfüllt oder fehlerhaft, n.r. - nicht relevant

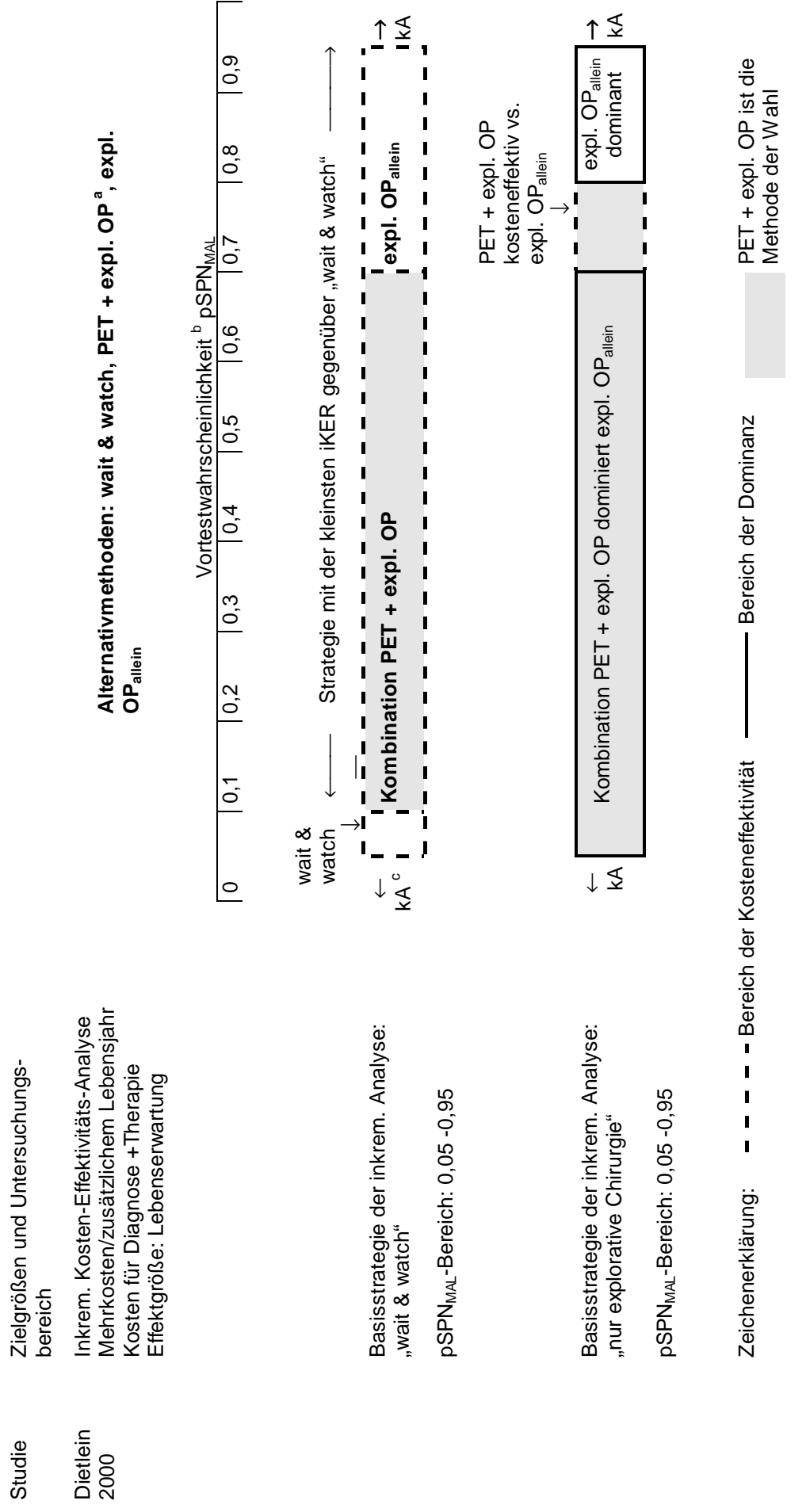
Abbildung 1a: Indikationsgebiet Onkologie: Abhängigkeit der Kosteneffektivität der Kombination aus PET und CT versus „nur CT“ bei den Indikationen „Diagnose und/oder Staging von nicht-kleinzelligem Lungenkarzinom“ (NSCLC) von der Vortestwahrscheinlichkeit für maligne SPN ($pSPN_{MAL}$) bzw. für Metastasenbildung durch den NSCLC-Primärtumor in den mediastinalen Lymphknoten ($pNSCLC_{MET}$)^{*}

Alternativmethoden: CT + PET, CT allein, (wait & watch, direkte OP)^a



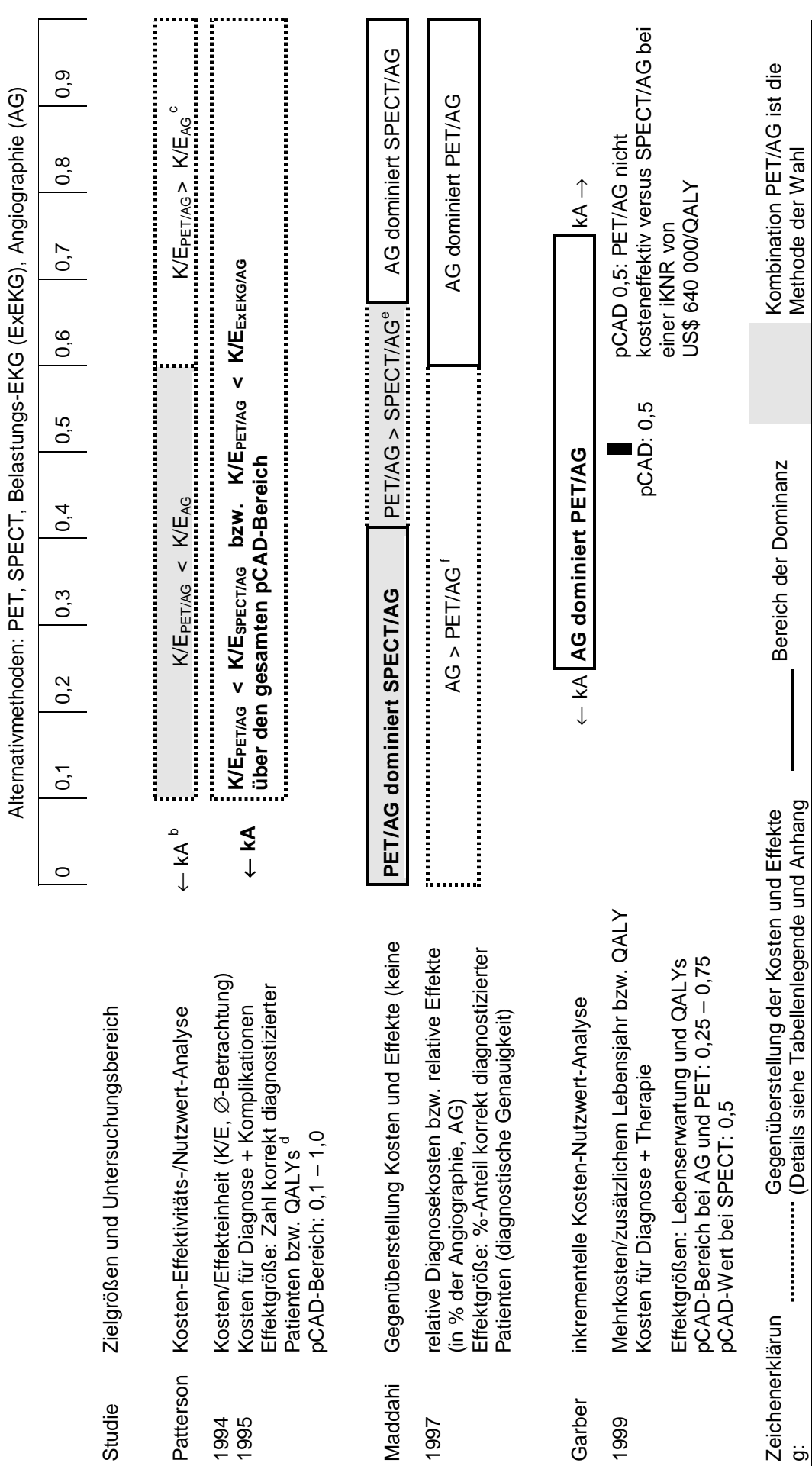
a: Thorakotomie: operativer Eingriff zur Entnahme des Tumorgewebes, „wait and watch“ bzw. direkte Thorakotomie nur bei Gambhir et al. 1998 / b: Vortestwahrscheinlichkeit des einzelnen Patienten, Metastasen, speziell in den regionären Lymphknoten, bzw. einen malignen SPN zu haben (Bevölkerungsniveau: Prävalenz) / c: wait & watch-Strategie / d: keine Angabe
^{*} Eine detaillierte Aufstellung der in Abbildung 1a verwendeten Daten gibt Anhang A1, Abkürzungen siehe Abkürzungsverzeichnis.

Abbildung 1b: Indikationsgebiet Onkologie: Abhängigkeit der Kosteneffektivität der Kombination PET + explorative Chirurgie versus „nur explorative Chirurgie“ bei der Diagnostik von nicht-kleinzelligem Lungenkarzinom (Diagnose und Staging) von der Vortestwahrscheinlichkeit für maligne SPN ($pSPN_{MAL}$) *



a: operativer Eingriff zur Entnahme von Tumorgewebe zur histologischen Weiteruntersuchung / b: Vortestwahrscheinlichkeit des einzelnen Patienten, maligne SPNs und damit NSCLC zu haben / c: keine Angabe * Eine detaillierte Auflistung der in Abbildung 1b verwendeten Daten gibt Anhang A2, Abkürzungen siehe Abkürzungsverzeichnis.

Abbildung 2: Indikationsgebiet Kardiologie (Myokardperfusionssmessung): Abhängigkeit der Kosteneffektivität der Kombination PET/Angiographie' versus konventionelle Diagnoseverfahren/Angiographie bzw. „nur Angiographie“ von der Vorstestwahrscheinlichkeit für koronare Herzkrankheit *



a: Vorstestwahrscheinlichkeit des einzelnen Patienten, an koronarer Herzkrankheit (KHK) erkrankt zu sein (Bevölkerungsniveau: Prävalenz) / b: keine Angabe / c: Aussage gilt außer bei PET/AG uch für die Kombinationen Belastungs-EKG/AG bzw. SPECT/AG / d: qualitätsadjustierte Lebensjahre / e: PET/AG ist effektiver, aber auch teurer als SPECT/AG / f: Angiographie ist effektiver, aber auch teurer als die Kombination PET/AG; * Eine detaillierte Aufstellung der in Abbildung 2 verwendeten Daten gibt Anhang B, Abkürzungen siehe Abkürzungsverzeichnis