

Nuklearmedizin

Basiswissen und klinische Anwendung

Bearbeitet von
Harald Schicha, Otmar Schober

7., überarb. u. aktualisierte Aufl. 2013 2013. Taschenbuch. 456 S. Paperback

ISBN 978 3 7945 2889 9

Format (B x L): 16,5 x 24 cm

[Weitere Fachgebiete > Medizin > Sonstige Medizinische Fachgebiete > Nuklearmedizin, Radiotherapie](#)

Zu [Inhaltsverzeichnis](#)

schnell und portofrei erhältlich bei


DIE FACHBUCHHANDLUNG

Die Online-Fachbuchhandlung beack-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.

6 Qualitätssicherung

In diesem Kapitel geht es um Kosten-Nutzen-Bewertungen und um Begriffe aus der evidenzbasierten Medizin, die bei der Diskussion mit den Kostenträgern um die Ressourcenverteilung im Gesundheitswesen und die Vergütung neuer diagnostischer Verfahren eine zunehmende Rolle spielen. Neben dieser gesellschaftspolitischen Dimension muss jede Praxis und Krankenhausabteilung, die offene radioaktive Stoffe am Menschen einsetzt, in zumindest zweijährigen Intervallen die Einhaltung der Vorgaben nach Strahlenschutzverordnung und Richtlinie Strahlenschutz in der Medizin belegen. Bezüglich der Qualitätskontrolle von Radiopharmaka wird auf Kapitel 3.5 verwiesen.

6.1 Nutzenbewertung diagnostischer Verfahren

In Deutschland wurde das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) geschaffen, um den Nutzen von Behandlungsmaßnahmen zu bewerten. Die Bewertung des IQWiG für eine Behandlungsmaßnahme wird als Vorbericht im Internet veröffentlicht, mit den Fachgesellschaften diskutiert und im nächsten Schritt an den Gemeinsamen Bundesausschuss (G-BA) weitergeleitet. Der G-BA hat eine rechtsverbindliche Entscheidung zu treffen, die dann im Bundesgesetzblatt veröffentlicht wird. Nach dem Sozialgesetzbuch entscheidet das Votum des G-BA über den Leistungsanspruch eines Patienten

gegenüber der Gesetzlichen Krankenversicherung (GKV). Die Bewertungskriterien des IQWiG und des G-BA sind im Internet einsehbar.

Vielfach orientiert sich die Nutzenbewertung diagnostischer Verfahren an der Systematik, die 1991 von Fryback und Thornbury erstellt wurde. Diese Systematik unterscheidet sechs Stufen, die Folgendes beinhalten (s. a. **Tab. 6-1**):

- Stufe 1: technische Parameter eines Untersuchungsverfahrens wie Ortsauflösung, Kantenlänge eines Detektorkristalls in der PET, Pixelgröße in der CT oder Kontrast
- Stufe 2: statistische Kenngrößen eines Testverfahrens wie z. B. Sensitivität, Spezifität oder prädiktiver Wert
- Stufe 3: Änderungen der Diagnosestrategie wie die Verkürzung eines diagnostischen Algorithmus oder der Ersatz anderer, möglicherweise invasiver diagnostischer Verfahren
- Stufe 4: Auswirkung der Diagnostik auf die Therapie-Entscheidung (z. B. Entscheidung für eine systemische Therapie anstelle einer Operation aufgrund des bildgebenden Nachweises einer Fernmetastase oder eine intensivere Chemotherapie aufgrund des Upstagings eines Lymphoms)
- Stufe 5: Effekt auf das rezidivfreie Überleben, das Gesamtüberleben oder die Therapie-assoziierte Morbidität (patientenrelevanter Nutzen)
- Stufe 6: gesellschaftlich-ökonomische beziehungsweise gesundheitspolitische Bewertung der Behandlungsmethode

Tab. 6-1 Evaluation diagnostischer Verfahren nach der Systematik von Fryback und Thornbury (1991). Für die Vergütung von Innovationen werden zunehmend Belege auf der Stufe 5 gefordert.

Stufe	Kriterien	Beispiel
1	technische Effektivität	Ortsauflösung, Kontrast
2	Testgüte	Sensitivität, Spezifität, prädiktiver Wert
3	Diagnosestrategie	⁶⁸ Ga-DOTA-TATE PET/CT statt ¹¹¹ In-Octreoscan®
4	Therapieentscheidung	systemische Therapie anstatt Operation
5	Outcome-Parameter	PFS, OS, weniger Toxizität bei gleichem OS
6	Nutzen für die Gesellschaft	Gesundheitspolitik

OS = overall survival, PFS = progression free survival

Gesundheitsökonomische Methoden

Basierend auf der Systematik nach Fryback und Thornbury gilt ein patientenrelevanter Nutzen als bewiesen, wenn durch eine neue Behandlung die Patientengruppe statistisch ein längeres Gesamtüberleben zeigt (= härtester Outcome-Parameter) oder eine verlängerte rezidivfreie Überlebenszeit aufweist (= relevanter Surrogatparameter) in einem strukturierten Vergleich zu einem Standardarm. Goldstandard für einen systematischen Vergleich ist die Randomisation, also die zufällige Zuordnung zwischen experimenteller Behandlung und Standardtherapie. Damit eine solche Randomisation ethisch vertretbar ist, muss der Standardarm der wissenschaftlich anerkannten Standardtherapie entsprechen. Nur unter diesen Voraussetzungen ist ein Vergleich zu einer neuen Therapie medizinisch überhaupt von Bedeutung.

Sofern es die Datenlage erlaubt, werden die zusätzlich gewonnenen Lebensjahre mit der Lebensqualität multipliziert: Eine hohe Lebensqualität entspricht dem Faktor 1, bei

eingeschränkter Lebensqualität wird ein Faktor zwischen 0 und 1 vergeben. Hierdurch erklärt sich der angelsächsische Begriff des *quality-adjusted life year* (QALY). Das QALY ist eine wichtige Bezugsgröße in der gesundheitsökonomischen Analyse (**Abb. 6-1** und **6-2**). In der Regel bindet ein zusätzliches QALY finanzielle Ressourcen, fällt also in das obere linksseitige Feld der **Abbildung 6-1**, und eine gesundheitsökonomische Evaluation ist vorzunehmen. Dabei werden die zusätzlichen Kosten des neuen Verfahrens im Vergleich zu den Kosten des bisherigen Standardverfahrens (Inkrementalanalyse) und ebenso der zusätzliche Nutzen des neuen Verfahrens gegenüber dem Nutzen des bisherigen Standardverfahrens bewertet. Werden die Zusatzkosten des neuen Behandlungsverfahrens ins Verhältnis zu dem Zusatznutzen gesetzt, erhält man die Kosten pro gewonnenem Lebensjahr. Dieser Quotient entspricht einer *incremental cost-effectiveness ratio* (ICER).

Eine Kosten-Effektivitäts-Analyse (*cost-effectiveness analysis*) kann sich auf zusätzlich gewonnene Lebensjahre beziehen; möglich sind aber auch Kostenquotienten pro

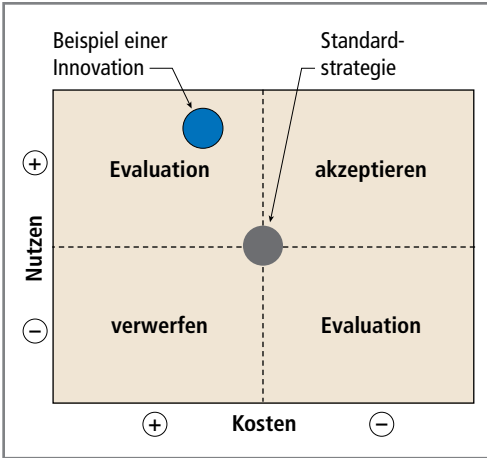


Abb. 6-1 Verglichen mit der Standarddiagnostik kann ein neues diagnostisches Verfahren (Innovation) höhere oder niedrige Kosten beziehungsweise einen besseren oder schlechteren Nutzen haben (vier Felder). Eine gesundheitsökonomische Analyse ist prinzipiell in dem Feld mit den höheren Kosten bei besserem Nutzen erforderlich. Wenn bei nachgewiesenem Nutzen sogar die Gesamtkosten reduziert werden, ist die Innovation in jedem Fall zu akzeptieren.

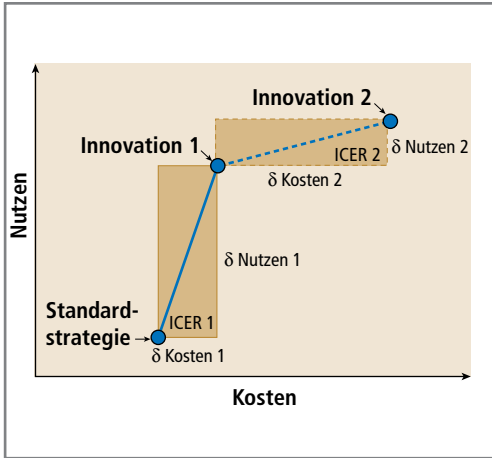


Abb. 6-2 Mitunter sind zwei verschiedene Innovationen zu prüfen, die in der Kosten-Nutzen-Relation dann gegeneinander und gegenüber der Standardstrategie zu bewerten sind. Der Quotient wird ICER genannt. Hierbei weist Innovation 1 einen sehr günstigen Kosten-Nutzen-Quotienten auf (ICER 1). Innovation 2 ist medizinisch überlegen; sofern der ICER 2 in einem gesellschaftspolitisch akzeptierten Bereich liegt, müsste die Innovation 2 implementiert werden. ICER = *incremental cost-effectiveness ratio*.

vermiedenen Ereignis oder Kostenquotienten zu anderen relevanten Parametern (z. B. Kosten pro vermiedener unnötiger Therapiemaßnahme, Kosten für Verbesserung eines Laborwerts, Kosten für die Absenkung eines erhöhten Blutdruckwerts). Werden die Kosten in Relation zu einem QALY gesetzt, so spricht man von einer Kosten-Nutzwert-Analyse (*cost-utility analysis*). Die Kosten-Nutzen-Analyse in der engeren Definition einer *cost-benefit analysis* befasst sich hingegen mit dem Nutzen für die Gesellschaft. Solche quantitativen Betrachtungen gelten als ethisch problematisch (Rationierung der Medizin, medienorientierte Entscheidungen, Partikularinteressen) und spielen in Deutschland keine Rolle.

Datenlage zur PET/CT

Die Nutzenbewertung diagnostischer Verfahren durch das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) und durch den Gemeinsamen Bundesausschuss (G-BA) hat hinsichtlich der Vergütungsregelungen zur PET/CT und des Leistungsanspruchs für die Patienten der Gesetzlichen Krankenversicherung in Deutschland inzwischen unmittelbare Konsequenzen. International wurde aber nur in wenigen Studien getestet, welchen Einfluss ein diagnostisches Verfahren auf den Outcome der Patienten (Stufe 5 in der oben genannten Systematik) hat. Seitens der nuklearmedizinischen Diagnostik sind hierzu

Studien mit PET beziehungsweise PET/CT beim Bronchialkarzinom und beim Hodgkin-Lymphom zu nennen:

In prospektiven randomisierten Multi-centerstudien aus den Niederlanden und aus Dänemark wurde im präoperativen Staging des nichtkleinzelligen Bronchialkarzinoms die konventionelle Diagnostik unter Einschluss der CT (Kontrollgruppe) mit der konventionellen Diagnostik unter zusätzlichem Einsatz der PET beziehungsweise PET/CT (Experimentalgruppe) verglichen. Endpunkt war jeweils der Patientenstatus ein Jahr postoperativ. Der Nachweis von Rezidiv, Metastasen oder ein zwischenzeitliches Versterben des Patienten nach einer kurativ durchgeführten Operation wurde als negatives Ereignis gewertet. Eine solche frustrane Operation fand sich bei etwa 40 % der Patienten in der Kontrollgruppe, aber nur bei etwa 20 % der Patienten in der Experimentalgruppe unter Hinzunahme der PET beziehungsweise PET/CT.

Bei Patienten mit Hodgkin-Lymphom liegt das wissenschaftliche Interesse derzeit auf dem Einsatz der PET beziehungsweise der PET/CT im Therapie-Monitoring, um den Verlauf einer Chemotherapie zu verfolgen, sowie in der Therapiekontrolle nach Abschluss der Chemotherapie, wenn die Entscheidung über eine zusätzliche Strahlentherapie zu treffen ist. Hier konnte in fortgeschrittenen Stadien nach sechs/acht Zyklen einer intensiven Chemotherapie für verbliebenes Restgewebe (partielle Remission) gezeigt werden, dass Patienten mit PET-negativem Restgewebe – das heißt ohne erhöhte Glucoseaufnahme – bei Verzicht auf eine zusätzliche Strahlentherapie die gleiche Überlebensrate haben wie Patienten in kompletter Remission ohne Restgewebe in der CT. Dies ist ein deutlicher Hinweis, dass bei PET-negativem Restgewebe ein Verzicht auf

eine zusätzliche Strahlentherapie und damit eine Toxizitätsreduktion möglich ist. Da das Restgewebe vielfach im Mediastinum gelegen ist, werden kardiopulmonale Späteffekte und die Rate an Zweitmalignomen verringert.

Der Nachweis eines solchen patientenrelevanten Nutzens stößt durch lange Rekrutierungs- und Nachbeobachtungszeiten an Grenzen. Auch gilt der Erkenntnisgewinn zunächst für eine spezielle Tumorentität und ein dezidiertes Therapieprotokoll. Das Fehlen adäquater Studien über den kausalen Nutznachweis der PET beziehungsweise der PET/CT ist kein Beleg dafür, dass die PET beziehungsweise PET/CT nutzlos sei (*the absence of evidence is not the evidence for absence*). Es wird auch in Zukunft neue Substanzgruppen, andere Dosierungsschemata oder eine modifizierte Anzahl an Therapiezyklen geben, sodass für die Bewertung eines diagnostischen Verfahrens (wie z. B. PET, PET/CT) nicht stets eine mehrjährige Rekrutierung und eine mehrjährige Nachbeobachtung abgewartet werden kann, zumal die PET beziehungsweise PET/CT die Güte einer Therapie nur misst, aber nicht bestimmt.

6.2 Evidenzbasierte Medizin (EbM) und Bias-Formen

Bei der evidenzbasierten Medizin richtet sich die Empfehlung zugunsten einer diagnostischen oder therapeutischen Maßnahme nach der methodischen Qualität in den publizierten wissenschaftlichen Studien. Ursprünglich wurde die EbM auf vergleichende Medikamentenstudien bzw. Therapiestu-

dien angewendet. Beispielsweise erhält eine Patientengruppe die Testmedikation (experimenteller Prüfarm), die andere Patientengruppe die bisherige Standardmedikation (Standardarm). Wenn medizinisch sinnvoll, kann die Standardmedikation auch ein Placebo-Präparat sein. Der zum Vergleich der Behandlungsarme dienende Endpunkt wird vor Studienbeginn definiert. Auf dem in der Prüfhypothese abgeschätzten Unterschied zwischen den Behandlungsarmen basiert eine statistisch notwendige Zahl der einzuschließenden Probanden. Die Zuteilung der Patienten zu einem der Behandlungsarme erfolgt nach dem Zufallsprinzip (Randomisation).

In Analogie zu den therapeutischen Studien wird die interne Validität der diagnostischen Studien – z. B. zur PET beziehungsweise PET/CT – zunehmend nach den formalen Kriterien der evidenzbasierten Medizin bewertet. Jede potenzielle Verzerrung (= Bias), die in ihrem Ausmaß nicht quantifiziert werden kann, führt zu einer methodischen Abwertung der Studie. Folgende Bias-Formen werden unterschieden:

- *Selection-Bias*: Erfolgte die Auswahl der Patienten konsekutiv in einem Zeitraum beziehungsweise entsprach die Auswahl einer zufälligen Stichprobe? Wurden die Ein- und Ausschlusskriterien genannt? Waren die demografischen und klinischen Basisdaten der Patienten bekannt?
- *Performance-Bias*: War bei vergleichendem Studiendesign die Begleittherapie/-diagnostik der beiden Gruppen identisch beziehungsweise vorgeschrieben?
- *Attrition-Bias*: Wie wurden die Patienten beschrieben, die in die Studie eingeschlossen wurden, aber nicht vollständig beobachtet werden konnten? Wie hoch war der Prozentsatz der eingeschlossenen Patienten, bei denen das Untersuchungsergebnis der PET nicht bewertet werden konnte? Goldstandard ist das *Intention-to-Treat*-Prinzip.
- *Spectrum-Bias*: Wurde die Studie an Patienten durchgeführt, bei denen die Erkrankung sehr wahrscheinlich war? Entsprach die Studiengruppe einer Population, bei der später in der klinischen Anwendung die PET durchgeführt werden soll? Der Ausschluss eines *Spectrum-Bias* ist für die Beurteilung der Übertragbarkeit der Ergebnisse auf die Gesamtheit der Erkrankten (externe Validität) von großer Bedeutung.
- *Verification-(Reference-)Bias*: Wurde der Goldstandard (Referenztest) zur Diagnosebestätigung bei allen Patienten angewendet oder wurden unterschiedliche Bestätigungsteste eingesetzt? Ein *Verification-Bias* wird angenommen, wenn der Referenztest abhängig vom Ergebnis des Testverfahrens angewendet wurde. Bei diskordanten Befunden (z. B. PET als Testverfahren positiv, Standardverfahren negativ) wurde das Standardverfahren nach vorgegebener Zeit zur Sicherung der Diagnose wiederholt und über die Zahl „zusätzlicher Diagnosen“ berichtet. Die „konkordanten“ Befunde wurden hingegen als „richtig positiv“ oder als „richtig negativ“ akzeptiert. Dies stellt eine Verzerrung zugunsten des Testverfahrens dar, da im Falle eines Minus/Minus-Befundes die Wiederholung des Standardverfahrens unterblieb, wodurch konkordant falsch negative Befunde nicht aufgedeckt werden konnten. Da in einer solchen Konstellation der positive prädiktive Wert der PET nicht zwingend beeinträchtigt wird, führt die Feststellung eines *Verification-Bias* nicht automatisch zu einer Abwertung aller Aussagen einer Studie.