

HTA-Bericht | Kurzfassung

Der Stellenwert von Patient-Reported Outcomes (PRO) im Kontext von Health Technology Assessment (HTA)

Brettschneider C, Lühmann D, Raspe H

Deutsche Agentur für HTA
des DIMDI (DAHTA)
Waisenhausgasse 36-38a
50676 Köln

Tel.: +49 221 4724-525
Fax +49 221 4724-444
dahta@dimdi.de
www.dimdi.de

Gesundheitspolitischer Hintergrund

In Deutschland stellen die gesetzlichen Rahmenbedingungen des Sozialgesetzbuchs (SGB) V die Bewertung des medizinischen Nutzens in das Zentrum der Verfahrensbewertung, wobei eine Operationalisierung des Begriffs „Nutzen“ im Gesetz nicht zu finden ist. Die Formulierungen im § 27 legen jedoch nahe, dass sich die Anforderungen des SGB V über die Parameter Morbidität, Mortalität und Lebensqualität abbilden lassen. Eine weitere Charakterisierung dieser Endpunkte findet sich im Methodenpapier 3.0 vom 27.05.2008 des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Hier wird im Abschnitt 3.1 „patientenrelevant“ beschrieben als „...wie ein Patient fühlt, seine Funktionen und Aktivitäten wahrnehmen kann oder überlebt“. Neben Mortalität, Morbidität (Beschwerden und Komplikationen) und Lebensqualität werden auch der interventions- und erkrankungsbezogene Aufwand und die Zufriedenheit der Patienten mit der Behandlung als potenzielle sekundäre Zielgrößen erwähnt. Für alle Endpunkte wird gefordert, dass sie direkt und zuverlässig Änderungen des Gesundheitszustandes abbilden müssen. Bei der Lebensqualität, dem empfundenen Gesundheitszustand sowie der Zufriedenheit handelt es sich um Endpunkte, über die letztendlich nur der Patient selbst Auskunft geben kann – sie werden als Patient-Reported Outcomes (PRO) bezeichnet.

Alle HTA-Berichte des DIMDI sind in der DAHTA-Datenbank und bei German Medical Science (GMS) kostenlos recherchierbar.

Wissenschaftlicher Hintergrund

PRO wird als Oberbegriff für viele verschiedene Konzepte zur Messung subjektiv empfundener Gesundheitszustände gebraucht. Die gemeinsame Grundlage dieser Konzepte ist, dass der Patient selbst seinen Zustand einschätzt und berichtet. Die Definition der amerikanischen Food and Drug Administration (FDA) drückt dies wie folgt aus: „A PRO is any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else“. PRO werden angewendet, wenn ein Konzept am besten durch den Betroffenen selbst erfasst werden kann. Sie können entweder als feststehender Wert, beispielsweise in Form der Symptomstärke, oder als Veränderung zwischen zwei Zeitpunkten gemessen werden.

Zur Messung von PRO existieren zwei prinzipiell unterschiedliche Ansätze. Mit dem psychometrischen Ansatz werden erlebte Symptome (Vorhandensein, Häufigkeit, Stärke), Fähigkeiten, Verhaltensweisen oder psychische Konstrukte, wie Befindlichkeiten oder Emotionen, erfasst. Einzelne Dimensionen lassen sich zu komplexen Konzepten (z. B. Lebensqualität, siehe unten) zusammenführen. Mithilfe des präferenzbasierten Ansatzes wird der Wert bestimmt, den die Befragten einem definierten Gesundheitszustand

beimessen. Die dazu verwendeten Methoden entstammen dem Bereich der Ökonometrie und basieren auf den Ansätzen der Entscheidungstheorie. Die folgenden Ausführungen beziehen sich ausschließlich auf psychometrische PRO.

Außer in der Gesundheitsökonomie wird im Rahmen der Qualitätssicherung von Gesundheitsleistungen und bei der Bestimmung der Wirksamkeit von Interventionen mit PRO argumentiert. Da das primäre Interesse von Health Technology Assessment (HTA) der Feststellung des Nutzens von gesundheitsrelevanten Technologien gilt, wird sich die Feststellung des Stellenwerts von PRO im vorliegenden Bericht auf das Einsatzgebiet von klinischen Studien konzentrieren.

Als Grundvoraussetzung für die valide Erfassung von PRO in klinischen Studien ist sicher zu stellen, dass die verwendeten Messinstrumente (Fragebögen) den gängigen psychometrischen Gütekriterien entsprechen. Validität, Reliabilität und die Änderungssensitivität gelten dabei als zentrale Merkmale. Weitere Aspekte sind die administrative und ökonomische Umsetzbarkeit sowie die Akzeptanz seitens des Patienten.

Um den Stellenwert von PRO im Kontext von HTA-Verfahren beschreiben zu können, wird zunächst eine Übersicht über Konzepte, Klassifikationen und methodische Messansätze gebraucht, die das sehr breite und heterogene Feld der PRO beschreibt. Grundlage für diese Übersicht bildet eine systematische Literaturrecherche nach Publikationen, die sich aus theoretischer Perspektive mit der Konzeption und der Messung von PRO befassen.

Die Rolle, die PRO derzeit im Kontext von HTA spielen, soll anhand von zwei empirischen Analysen herausgearbeitet werden. Die Häufigkeit der Nutzung von PRO, die Art der genutzten PRO und die anhand von PRO gefunden Studienergebnisse (im Abgleich zu den konventionellen klinischen Ergebnisparametern) aus aktuellen randomisierten kontrollierten Studien (RCT) zu den Beispielthemen rheumatoide Arthritis und Mammakarzinom sollen dokumentiert werden.

Die zweite empirische Analyse befasst sich mit der Verwendung von PRO für HTA im engeren Sinn: Aus der Analyse von Methodenpapieren internationaler HTA-Organisationen wird untersucht, welchen Stellenwert PRO die HTA-Agenturen beimessen. Des Weiteren soll eine Analyse von aktuellen HTA-Berichten zu den beiden Beispielkrankheitsbildern zeigen, inwieweit die Ergebnisse von PRO-Messungen eine Rolle bei der Formulierung von Schlussfolgerungen der HTA-Berichte spielen.

Diese Ziele werden in acht Forschungsfragen präzisiert.

Forschungsfragen

1. Wie sind PRO im Kontext von klinischen Studien definiert?
2. Welche Klassifikationssysteme zur Charakterisierung von Endpunkten für klinische Studien gibt es und welchen Stellenwert nehmen PRO in diesen Systemen ein?
3. Welche methodischen Ansätze zur Messung von PRO gibt es, was sind ihre spezifischen Stärken und Schwächen?
4. Wie häufig wurden PRO als primäre oder nachgeordnete Zielgröße in RCT aus den Indikationsbereichen Mammakarzinom und rheumatoide Arthritis definiert und berichtet?

5. Welche PRO werden in den genannten Studien als Zielgrößen definiert und berichtet?
6. Wie ist die Übereinstimmung von Schlussfolgerungen hinsichtlich der Wirksamkeit der zu bewertenden Therapien basierend auf PRO im Vergleich zu anderen Endpunkten?
7. Wie wird der Stellenwert von PRO für die Nutzenbewertung von medizinischen Technologien in der methodologischen HTA-Literatur beurteilt?
8. Inwieweit werden PRO in HTA-Berichten zu den Beispielthemen berücksichtigt, insbesondere bei der Ableitung von Schlussfolgerungen und Empfehlungen?

Methodik

Die Beantwortung der Forschungsfragen erfordert drei Literaturrecherchen. Die erste dient dem Auffinden von methodisch-theoretischen Publikationen. Die Recherche wird in den medizinnahen Datenbanken des Deutschen Instituts für Medizinische Dokumentation und Information (DIMDI) durchgeführt und umfasst den Zeitraum von 1990 bis 2009. Die Literaturselektion wird in zwei Schritten mittels definierter Ein- und Ausschlusskriterien durchgeführt. Ziel ist die Identifikation von Publikationen, die sich mit der Definition, Klassifikation und Operationalisierung von PRO beschäftigen. Eine standardisierte Datenextraktion mittels Extraktionsbögen ist aufgrund der Beschaffenheit der Literatur nicht möglich. Sie erfolgt durch die Zuordnung der Publikationen zu den Forschungsfragen und im nächsten Schritt der relevanten Inhalte zu Kategorien, die aus der Literatur abgeleitet werden.

Die zweite Recherche dient der Identifikation von klinischen Studien zur Biologikatherapie der rheumatoiden Arthritis und zur Chemotherapie des Brustkrebses. Die Recherche wird ebenfalls in den medizinnahen Datenbanken des DIMDI durchgeführt und umfasst den Zeitraum von 2005 bis 2009. Die Literaturselektion basiert auf definierten Ein- und Ausschlusskriterien. Ziel der Selektion ist die Identifikation von RCT zu den Beispielindikationen. Die relevanten Informationen aus den Studien werden in Extraktionsbögen dokumentiert. Die methodische Qualität der RCT wird mithilfe eines in Anlehnung an das „Risk of Bias Tool“ der Cochrane Collaboration konzipierten Instruments bewertet.

Die dritte Recherche zielt auf das Auffinden von Methodenpapieren und Berichten von HTA-Agenturen und Institutionen der Arzneimittelbewertung. Die Recherche wird als internetgestützte Handrecherche durchgeführt. Auf den Internetseiten der Mitglieder des International Network of Agencies for Health Technology Assessment (INAHTA) und der Institutionen, die im HTA-Bericht „Methoden zur vergleichenden Bewertung pharmazeutischer Produkte“ dargestellt werden, wird nach Methodenpapieren recherchiert. Um HTA-Berichte zu rheumatoider Arthritis und dem Mammakarzinom zu finden, wird eine Recherche in der HTA-Datenbank des Centre for Reviews and Dissemination (CRD) mit kontrollierten Suchbegriffen durchgeführt. Die Literaturselektion erfolgt über definierte Ein- und Ausschlusskriterien. Auch für diesen empirischen Teil werden relevante Informationen aus den Papieren in Extraktionsbögen dokumentiert.

Ergebnisse

Nach dem Prozess der Zweitselektion werden 158 Quellen zur Beantwortung der Forschungsfragen 1 bis 3 zu methodisch-theoretischen Aspekten eingeschlossen. Weiterhin werden 129 Publikationen zur Chemotherapie des Mammakarzinoms und 73 zur Biologikatherapie der rheumatoiden Arthritis identifiziert, die zur Beantwortung der Forschungsfragen 4 bis 6 dienen. Zur Beantwortung der Forschungsfragen 7 und 8 werden 34 HTA-Methodenpapiere, zwölf HTA-Berichte zur Chemotherapie des Mammakarzinoms sowie neun zur Biologikatherapie der rheumatoiden Arthritis gefunden.

Methodisch-theoretische Fragestellungen

Der Begriff PRO wird seit Beginn des neuen Jahrtausends in seiner heutigen Form genutzt. PRO ist eine pragmatische Definition und bezieht sich auf die Herkunft der Daten. Zentrale Kriterien sind der Patientenbericht und die dem Patienten überlassene Deutungshoheit. Der Begriff PRO hat den Begriff „Lebensqualität“ als Oberbegriff für patientenberichtete Endpunkte abgelöst. Eine inhaltliche Ausgestaltung erhält der Begriff PRO durch die Zuordnung von Konstrukten. Hierzu gehören in erster Linie die vom Patienten erlebten Symptome, Funktionsfähigkeiten, Gesundheitswahrnehmungen, Zufriedenheit und (gesundheitsbezogene) Lebensqualität. Die meisten Konstrukte sind noch nicht allgemeingültig definiert.

Mehrere Autorengruppen versuchen, verschiedene Konstrukte in umfassende Endpunktmodelle zu integrieren. Diese Modelle sollen vor allem Beziehungen und Wechselwirkungen zwischen einzelnen Parametern abbilden und externe Einflüsse aufzeigen. Wilson und Cleary entwerfen ein fünfstufiges, lineares Modell. Biologische bzw. physiologische Zielgrößen bilden den objektiven Pol, die Gesamtlebensqualität den subjektiven. Zwischen diesen beiden Extrema sind die Konstrukte Symptom-, Funktionsstatus und allgemeine Gesundheitswahrnehmung angeordnet. Zwischen zwei benachbarten Konstrukten werden gegenseitige Beeinflussungen angenommen.

Ein weiteres Modell stellt die Internationale Klassifikation von Funktionsfähigkeit, Behinderung und Gesundheit (ICF) der Weltgesundheitsorganisation (WHO) dar. Im Zentrum der ICF steht die Beschreibung des „Gesundheitszustandes“ eines Individuums in den Dimensionen „Körperstrukturen und Funktionen“ (Anatomie/Physiologie), „Aktivitäten“ („...die Durchführung einer Aufgabe oder Handlung [Aktion] durch einen Menschen“) und Partizipation („Einbezogenheit in eine Lebenssituation“). Der nach der ICF klassifizierbare Gesundheitszustand steht in unmittelbarer Wechselwirkung mit dem Gesundheitsproblem (klassifiziert nach Internationaler Klassifikation der Krankheiten [ICD]) sowie mit umwelt- und personenbezogenen Kontextfaktoren.

Valderas und Alonso nehmen für ihr Klassifikationssystem eine umfassende Integration der beiden oben beschriebenen Modelle vor. Die Grundlage liefert das Modell von Wilson und Cleary, in das die Kategorien der ICF integriert werden. Im Gegensatz zu den beiden oben beschriebenen Modellen, die lediglich eine Definition und Ordnung von Endpunkten nach inhaltlichen Aspekten liefern, wird das Modell von Alonso und Valderas konzipiert, um Instrumente, die zur Messung von PRO eingesetzt werden,

zu klassifizieren. Die Beschreibung und Gruppierung der Instrumente erfolgt unter drei Gesichtspunkten: den zu messenden Konstrukten, der (Ziel-) Population und den Messeigenschaften.

Der größte Teil der gefundenen Literatur befasst sich mit der Operationalisierung von PRO. Grundsatz einer PRO-Erhebung ist der konzeptionelle Rahmen. Er ermöglicht die Identifikation der für den jeweiligen Kontext angemessenen Konstrukte, die Definition der Zielpopulation und die Darstellung der geplanten Messungen in einem Endpunktmodell. Letzteres umfasst auch die Spezifikation relevanter Variablen und geeigneter Messinstrumente.

Zur Messung von PRO stehen unterschiedliche Instrumententypen zur Verfügung. Psychometrische Instrumente erfragen eine Zustandsbeschreibung aus der Perspektive des Patienten. Präferenzbasierte Instrumente dagegen erfassen die Wertschätzung eines Zustands. Weiterhin sind generische von spezifischen Messinstrumenten zu unterscheiden. Generische Instrumente sind unabhängig vom zu untersuchenden Krankheits- oder Störungsbild einsetzbar, spezifische Instrumente sind für Untersuchungen in definierten Patientenpopulationen konzipiert. Generell gilt, dass spezifische Instrumente präzisere Messergebnisse für die jeweils interessierende Patientenpopulation liefern, jedoch sind die Ergebnisse nicht über Krankheitsgrenzen hinweg vergleichbar. Dies wird erst durch die Verwendung von generischen Instrumenten möglich – eventuell auf Kosten der Präzision. Des Weiteren existieren Profil- und Indexinstrumente. Profilinstrumente geben für jedes gemessene Konstrukt oder einen Teil hiervon einen eigenen Wert aus, während Indexinstrumente die Ergebnisse zu einem einzigen Wert aggregieren.

Im Zuge der Entwicklung von neuen Messinstrumenten kommt der Generierung von Testfragen (Items) eine zentrale Bedeutung zu. Die Fragen bestimmen, ob das neue Instrument in der Lage ist, das interessierende Konstrukt valide und präzise zu messen. Zur Itemgenerierung gilt als die optimale Herangehensweise ein gemischter Ansatz mit Fokusgruppen (mit Patienten), Expertenbefragung und Literaturstudium. Zur Fragebogenentwicklung ist die Unterscheidung zwischen klassischer Testtheorie und Item-Response-Theorie ebenfalls ein relevantes Merkmal. Die meisten Erhebungsinstrumente basieren auf der klassischen Testtheorie. Ein neuerer Ansatz, der kürzere und präzisere Fragebögen liefert, wird mit der Item-Response-Theorie verfolgt.

Entscheidend für die Einsatzfähigkeit eines Instruments sind seine Gütekriterien. Als zentral wichtige Kriterien gelten Validität, Reliabilität und Änderungssensitivität eines Messinstruments. Im Bereich der Validität wird besonders auf die Inhalts-, die Konstrukt- und die Kriteriumsvalidität hingewiesen. Die Reliabilität bezieht sich vor allem auf die Test-Retest-Reliabilität und die interne Konsistenz. Bei der Bestimmung der Änderungssensitivität werden zwei prinzipielle Herangehensweisen unterschieden: Verteilungsbasierte Verfahren ziehen Effektgrößen zur Bewertung der Änderungssensitivität heran, ankerbasierte Verfahren vergleichen die Messwerte und ihre Änderungen mit einem externen (Gold-) Standard.

Zentrale Bedeutung bei der Interpretation von PRO-Ergebnissen hat der kleinste relevante Unterschied (Minimal Important Difference, MID). Hierbei handelt es sich um den kleinsten Unterschied zwischen zwei Messwerten, den ein Patient für bedeutsam hält. Ein Standardverfahren zur Festlegung

der MID existiert nicht. Die gebräuchlichen Verfahren entsprechen denen, die zur Bestimmung der Änderungssensitivität herangezogen werden.

Aus methodischen wie auch aus ökonomischen Gründen ist die Verwendung von bereits bestehenden standardisierten und validierten Instrumenten der Neuentwicklung vorzuziehen. In manchen Situationen kann es aber erforderlich werden, ein bestehendes Instrument an die Erfordernisse eines spezifischen Messkontexts anzupassen. Als besonders kritischer Punkt gilt die Übersetzung eines Instruments in eine andere Sprache. Hierzu liegen Leitlinien vor, die einen zehnschrittigen Prozess mit Vorwärts- und Rückwärtsübersetzungen vorsehen. In jedem Fall muss ein modifiziertes Instrument erneut validiert werden.

PRO in klinischen Studien

Die gefundenen 73 Publikationen zur Biologikatherapie der rheumatoiden Arthritis entstammen 59 Studien. Die Bewertung der Qualität der einzelnen Studien weist keinen Unterschied zwischen Studien aus, die PRO nutzen, und solchen, die dies nicht tun. Die am häufigsten genutzte primäre Zielgröße ist der Effektivitätsparameter des American College of Rheumatology (ACR). Dieses sieben Konstrukte umfassende Endpunktmaß enthält drei patientenberichtete Parameter: Schmerzen, globale Einschätzung des Gesundheitszustands, Funktionsfähigkeit im Alltag. Von diesen muss mindestens ein Parameter positive Veränderungen aufweisen um das Gesamtkriterium zur Erfüllung zu bringen. In 23 RCT wird das ACR20-Kriterium (erfordert 20-prozentige Verbesserung des Kriteriums-werts) als primäre Zielgröße verwendet. Reine PRO werden seltener als primärer Endpunkt genutzt. Am häufigsten werden sie mit dem Health Assessment Questionnaire (HAQ) ermittelt, ein aus 20 Fragen bestehender Fragebogen zum Funktionsstatus. Dieser wird in sechs Studien als primärer Endpunkt verwendet.

Unter den sekundären Endpunkten ist der Disease Activity Score (DAS) der meistgenutzte Endpunkt (in 39 Studien). Der DAS ist eine aus vier Kriterien zusammengesetzte Zielgröße. Eines der Kriterien, die Gesundheitswahrnehmung, ist den PRO zuzurechnen. Reine PRO werden in 34 Studien als sekundäre Endpunkte verwendet. Die dominierenden Konstrukte sind der Funktionsstatus und die Gesundheitswahrnehmung. Die am häufigsten verwendeten PRO-Instrumente sind abermals der HAQ (32 Studien) und der Shortform 36 (SF-36), ein aus 36 Fragen bestehender Fragebogen zur allgemeinen Gesundheitswahrnehmung (16 Studien).

Fünf Studien berichten Widersprüche zwischen patientenberichteten und traditionellen Endpunkten. In einem Fall betrifft dies radiologische Endpunkte (Synovitis, Gelenkerosion, Ödeme) und den HAQ, die für zwei zu vergleichende Interventionen zu gegenläufigen Ergebnissen kommen. In den übrigen vier Studien bilden sich beobachtete unerwünschte Wirkungen nicht in den Ergebnissen von HAQ und SF-36 ab. Hierbei handelt es sich um kurzfristig auftretende Nebenwirkungen wie Fieber, Kopfschmerzen, Asthenie und Infektionen. In den Diskussionsteilen der meisten Studien werden die Ergebnisse von PRO kommentiert. In den 35 Studien, die eine Abwägung zwischen traditionellen und patientenberichteten Endpunkten vornehmen, werden von 19 Studien die PRO als Befund betrachtet, der die Ergebnisse traditioneller Parameter bestätigt. Widersprüchliche Befunde,

wie sie in den Ergebnisteilen der fünf oben angesprochenen Publikationen gezeigt werden, werden nicht diskutiert. Zusammenhänge zwischen PRO und traditionellen Parametern werden in zwölf Studien diskutiert. Häufig beschrieben wird eine positive Korrelation von Krankheitsaktivität und den HAQ-Ergebnissen.

15 Studien nehmen in ihren Schlussfolgerungen Bezug auf PRO – vor allem auf den Funktionsstatus und/oder die Gesundheitswahrnehmung. Hierbei handelt es sich zumeist um Studien, in denen die PRO-Ergebnisse den Resultaten traditioneller Parameter entsprechen und somit gleichlautende Therapieempfehlungen abgeleitet werden können. Zwei Studien verwenden PRO-Ergebnisse, um Therapieempfehlungen zu differenzieren. Die Differenzierung bezieht sich in einem Fall auf die Berücksichtigung von Ergebnissen zur Therapietreue, in einem anderen Fall auf die Begründung einer Therapieempfehlung mit dem Verweis auf positive Effekte im Bereich der Lebensqualität. Sechs Studienautoren formulieren Forschungsbedarf zur Optimierung der Verwendung von PRO. Hier werden in erster Linie die Notwendigkeit von Langzeitdaten und die weitere Ausarbeitung von Zusammenhängen zwischen klinischen und patientenberichteten Endpunkten angeführt.

Den 129 Publikationen zur Chemotherapie des Mammakarzinoms liegen 123 RCT zugrunde. Der am häufigsten verwendete primäre Endpunkt ist die Überlebenszeit (66-mal in 53 Studien). Sechs Studien nutzen PRO als primären Endpunkt, zumeist Lebensqualität. Am häufigsten (in drei Studien) ist das verwendete Instrument der 30 Fragen umfassende Lebensqualitätsfragebogen der European Organisation for Research and Treatment of Cancer (EORTC). Als sekundärer Endpunkt werden PRO in 22 Studien genutzt. Auch hier ist die Lebensqualität das am häufigsten gemessene Konstrukt. In diesem Bereich werden die Lebensqualitätsfragebögen der EORTC am häufigsten eingesetzt (zwölf Studien). 26 Studien nennen PRO-Ergebnisse, von denen 15 signifikante Resultate präsentieren. Eine Studie findet mithilfe von PRO Therapieeffekte, die traditionelle Parameter nicht entdecken. Dies betrifft als traditionelle Parameter die Gesamtüberlebenszeit und das progressionsfreie Überleben sowie als PRO den Lebensqualitätsfragebogen der EORTC. In 24 von 28 Studien, in denen PRO gemessen werden, werden die Ergebnisse auch im Diskussionsteil kommentiert. PRO und traditionelle Parameter werden in 18 Studien verglichen. In drei Diskussionsteilen wird auf Diskrepanzen zwischen PRO und traditionellen Parametern eingegangen. Zusammenhänge werden in sechs Studien diskutiert, von denen drei eine Verbindung zwischen Lebensqualität und beobachteten Nebenwirkungen ausmachen. In neun Studien wird Forschungsbedarf für die weitere Nutzung von PRO formuliert, insbesondere Aspekte der Verwendbarkeit von PRO-Informationen im klinischen Alltag werden angesprochen.

PRO in HTA-Methodenpapieren

Es liegen 34 Dokumente aus 20 Institutionen vor, von denen sich 18 Dokumente eingehend mit der Thematik PRO auseinandersetzen. Schwerpunkte der Darstellungen liegen auf der Begründung für die Verwendung von PRO und auf der Charakterisierung von Messinstrumenten. Einen hohen Stellenwert nimmt weiterhin die Verwendung von Lebensqualitäts-

messungen als Datengrundlage für gesundheitsökonomische Evaluationen ein (Bestimmung von qualitätsadjustierten Lebensjahren (QALY), Kosten-Nutzwert-Analysen).

HTA-Berichte zur rheumatoiden Arthritis: In den Recherchen werden neun Berichte zum Thema Biologikatherapie der rheumatoiden Arthritis gefunden. Durch den Verweis auf die Auswirkungen der Erkrankung auf den Funktionsstatus und die Lebensqualität wird bereits in den Hintergrundkapiteln von sieben Berichten die Bedeutung von PRO als Studienendpunkt herausgehoben.

In den HTA-Berichten werden vor allem die zusammengesetzten Endpunkte ACR20, ACR50 und ACR70 als relevante Zielgrößen für die Nutzenbewertung von Therapien der rheumatoiden Arthritis identifiziert. Weiterhin werden die Schmerz-, Funktionsfähigkeits- und Lebensqualitätsmessung thematisiert. Dabei wird in etwa der Hälfte der Berichte klar, dass die Forderung nach der Berücksichtigung von PRO von den jeweiligen HTA-Autoren formuliert wird.

In allen neun Berichten werden die PRO-Ergebnisse explizit diskutiert, wobei insbesondere auf die Ergebnisse der ACR-Effektivitätsparameter eingegangen wird. Die Ergebnisse von traditionellen Endpunkten und PRO werden nur in zwei Berichten verglichen.

HTA-Berichte zur Chemotherapie des Mammakarzinoms: In drei von zwölf identifizierten Berichten werden bereits im Hintergrundkapitel Lebensbereiche erwähnt, in denen sich durch die Chemotherapie des Mammakarzinoms ausgelöste Veränderungen mittels PRO abbilden lassen. Symptome und Lebensqualität sind die angesprochenen Konstrukte. Die Linderung von Symptomen gilt neben der Lebensverlängerung als Ziel der Chemotherapie. Eine unausweichliche Beeinträchtigung der Lebensqualität bei systemischer Chemotherapie ist der zweite Bereich, für den auf die Bedeutung der Erfassung von PRO hingewiesen wird.

Fünf Berichte verweisen im Methodenteil auf die Bedeutung von Lebensqualitätsdaten, weitere angesprochene Konstrukte sind Schmerz, Angst, Depression und Funktionsstatus. Ein Bericht führt die Nutzung von PRO nur im Rahmen der Kosten-Nutzen-Bewertung an.

Vier Berichte liefern Daten zu PRO, zumeist zur Lebensqualität. Auch zu den anderen, in den Methodenteilen erwähnten Konstrukten finden sich Ergebnisse, die belegen, dass die jeweiligen HTA-Autoren den Studienautoren in der Gewichtung der Endpunkte folgen. Insgesamt werden sieben PRO-Ergebnisse ausgewiesen. Drei Berichte diskutieren die Bedeutung der PRO-Ergebnisse, zwei weitere stellen sie in Relation zu traditionellen Parametern. Die Autoren kommen zum einen zu dem Schluss, dass eine Verbesserung des krankheitsfreien Überlebens sich nicht in der Lebensqualität niederschlägt, zum anderen wird auf die Relation von Toxizität und reduzierter Lebensqualität verwiesen.

Diskussion

Die vorliegende Arbeit unterliegt einigen methodisch-bedingten Limitationen. Für den Bereich der methodisch-theoretischen Darstellungen kann nicht ausgeschlossen werden, dass möglicherweise relevante Publikationen übersehen werden – bedingt durch das Fehlen eines kontrollierten Vokabulars für die Thematik, die Beschränkung der Datenbankauswahl auf medizi-

nische Datenbanken und des Recherchezeitraums auf 1990 bis 2009. Es ist jedoch davon auszugehen, dass fehlende Schlüsselpublikationen über Querverweise in der vorliegenden Literatur gefunden worden wären. Weiterhin werden möglicherweise Publikationen in den Literaturpool aufgenommen, die keine originären Arbeiten repräsentieren. Der Grund ist, dass die methodisch-theoretischen Publikationen nicht nach dem Kriterium eines bestimmten Studientyps selektiert werden können. Dieses Problem scheint jedoch von untergeordneter Bedeutung, da im Bereich der theoretisch-methodischen Literatur keine quantitativen Auswertungen vorgenommen werden.

Zusammenfassend lassen sich die Forschungsfragen wie folgt beantworten: Der Begriff PRO wird aktuell als Oberbegriff für patientenberichtete Zielgrößen verwendet. Der bisherige Oberbegriff der Lebensqualität wird als eines von mehreren Konstrukten (ebenso wie Symptome, Funktions- und Gesundheitsstatus) dem neuen Oberbegriff zugeordnet. Über die Frage, welche Konstrukte neben den oben genannten weiterhin den PRO zuzurechnen sind, besteht noch kein Konsens. Die strikte Trennung von traditionellen Endpunkten und PRO wird nicht mehr für sinnvoll gehalten. Eine integrierte Sichtweise liefert ein ganzheitliches Bild der Erkrankung, da die Abbildung gegenseitiger Beeinflussungen zwischen den einzelnen Endpunkten möglich wird. Auf diesem Weg ist die Klärung von Ursache und Wirkung eines Phänomens einfacher zu erreichen. Mehrere Autorengruppen haben versucht, diese Zusammenhänge in Modellen abzubilden, wobei das Modell von Alonso und Valderas den komplexesten und zugleich pragmatischsten Ansatz liefert.

Für die Messung von PRO stehen verschiedene Modelle zur Verfügung. Psychometrische und nutzentheoretische Ansätze, Profil- und Indexinstrumente, generische und spezifische Instrumente besitzen Stärken und Schwächen und müssen je nach Forschungsvorhaben ausgewählt werden. Validität, Reliabilität und Änderungssensitivität charakterisieren die Güte von Messinstrumenten. In den meisten Fällen ist aus methodischen sowie aus ökonomischen Gründen die Verwendung von standardisierten und validierten Instrumenten einer Neuentwicklung vorzuziehen. Müssen Instrumente auf einen spezifischen Messkontext angepasst werden, z. B. durch Übersetzung, wird gegebenenfalls eine neue Validierung erforderlich. Im Hinblick auf die Interpretation von PRO-Ergebnissen ist der MID eines Instruments entscheidend. Im Bereich der Itementwicklung liegt der Schwerpunkt immer noch auf der klassischen Testtheorie. Jedoch entwickeln sich Ansätze der Item-Response-Theorie weiter, die kürzere und präzisere Instrumente versprechen.

An der Häufigkeit der Nutzung von PRO in Studien zur rheumatoiden Arthritis ist festzustellen, dass PRO ein hoher Stellenwert beigemessen wird. Der Schwerpunkt liegt auf der Verwendung von zusammengesetzten Endpunkten (Effektivitätskriterien des ACR, DAS), die die theoretisch geforderte Integration von PRO mit klinischen Zielgrößen vornehmen. Reine PRO, speziell Funktionsstatus und Gesundheitswahrnehmung (HAQ, SF-36), sind als nachgeordnete Endpunkte relevant. Eine differenzierte Diskussion zum Zusammenhang von PRO und traditionellen Endpunkten in Bezug auf die spezifische Indikation ist den Studien zumeist nicht zu entnehmen.

In klinischen Studien zum Mammakarzinom werden PRO-Ergebnisse nicht

so oft (24 % der gefundenen Studien) und vorwiegend als nachgeordnete Zielgrößen berichtet. Zur Messung kommen ausschließlich Lebensqualitätsfragebögen (vorwiegend der der EORTC) zur Anwendung. Auch bei dieser Indikation findet keine integrierende Diskussion statt.

In etwa der Hälfte der gefundenen Methodenpapiere von HTA-Institutionen wird der Stellenwert von PRO thematisiert und begründet, sowohl im Zusammenhang mit der Nutzenbewertung als auch für gesundheitsökonomische Evaluationen. Konkrete Ausführungen finden sich vor allem zum verfügbaren Messinstrumentarium. HTA-Berichte zu den Beispielindikationen reflektieren den Stellenwert von PRO in den Primärstudien. Hervorzuheben ist aber, dass in den meisten HTA-Berichten Forschungsbedarf im Hinblick auf die Verwendung patientenberichteter Endpunkte formuliert wird, was andeutet, dass die Verwendung von PRO in HTA-Institutionen vorausschauend verfolgt wird.