

Informative value of Patient Reported Outcomes (PRO) in Health Technology Assessment (HTA)

Abstract

Background

“Patient-Reported Outcome” (PRO) is used as an umbrella term for different concepts for measuring subjectively perceived health status e. g. as treatment effects. Their common characteristic is, that the appraisal of the health status is reported by the patient himself. In order to describe the informative value of PRO in Health Technology Assessment (HTA) first an overview of concepts, classifications and methods of measurement is given. The overview is complemented by an empirical analysis of clinical trials and HTA-reports on rheumatoid arthritis and breast cancer in order to report on type, frequency and consequences of PRO used in these documents.

Methods

For both issues systematic reviews of the literature have been performed. The search for methodological literature covers the publication period from 1990 to 2009, the search for clinical trials of rheumatoid arthritis and breast cancer covers the period 2005 to 2009. Both searches were performed in the medical databases of the German Institute of Medical Documentation and Information (DIMDI). The search for HTA-reports and methodological papers of HTA-agencies was performed in the CRD-Databases (CRD = Centre for Reviews and Dissemination) and by handsearching the websites of INAHTA member agencies (INAHTA = International Network of Agencies for Health Technology Assessment). For all issues specific inclusion and exclusion criteria were defined. The methodological quality of randomized controlled trials (RCT) was assessed by a modified version of the Cochrane Risk of Bias Tool. For the methodological part information extraction from the literature is structured by the report’s chapters, for the empirical part data extraction sheets were constructed. All information is summarized in a qualitative manner.

Results

Concerning the methodological issues the literature search retrieved 158 documents (87 documents related to definition or classification, 125 documents related to operationalisation of PRO). For the empirical analyses 225 RCT (rheumatoid arthritis: 77; breast cancer: 148) and 40 HTA-reports and method papers were found.

The analysis of the methodological literature confirms the role of PRO as an umbrella term for a variety of different concepts. The newest classification system facilitates the description of PRO measures by construct, target population and the method of measurement. Steps of operationalisation involve defining a conceptual framework, instrument development, exploration of measurement properties or, possibly, the modification of existing instruments.

Seven out of 59 RCT analysing the effects of antibody therapy for rheumatoid arthritis define PRO as the primary endpoint, 38 trials utilize composite measures (ACR, DAS) and ten trials report clinical or radiological parameters as the primary endpoint. Six out of 123 chemotherapy

**Christian
Brettschneider¹
Dagmar Lühmann²
Heiner Raspe²**

1 University Medical Center
Hamburg-Eppendorf,
Department of Medical
Sociology and Health
Economics, Hamburg,
Germany

2 University Lübeck, Institute
for Social Medicine, Lübeck,
Germany

trials for breast cancer define PRO as the primary endpoint, while 98 trials report clinical endpoints (survival, tumour response, progression) in their primary analyses. Discrepancies in the number of trials result from inaccurate specifications of endpoints in the publications. This distribution is reflected in the HTA-reports: while almost all reports on rheumatoid arthritis refer to PRO, this is only the case in about half of the reports on breast cancer.

Conclusions

As definition and classification of PRO are concerned, coherent concepts are found in the literature. Their operationalisation and implementation must be guided by scientific principles. The type and frequency of PRO used in clinical trials largely depend on the disease analysed. The HTA-community seems to pursue the utilization of PRO proactively – in case of missing data the need for further research is stated.

Keywords: patient-reported outcome, patient reported outcome, quality of life, rheumatoid arthritis, carcinoma of the breast, breast cancer, clinical trials, clinical studies, Health Technology Assessment, concept, HTA report, HTA-report, endpoint determination, efficiency, efficacy, effectiveness, cost-effectiveness, costs, cost analysis, cost-benefit-analyses, cost control, medical costs, sickness costs, cost-cutting, cost reduction, systematic review, HTA, technology assessment, medical assessment, technology evaluation, medical evaluation, health economics, health economic studies, evidence based medicine, EBM, ethics, juridical, social economic factors, socioeconomics, economic aspect, pharmacoeconomics, diagnosis, prevention, rehabilitation, therapy, treatment, review, academic review, review literature, research article, research-article, report, technical report, methods, care, meta analysis, meta-analysis, randomised controlled trial, randomized controlled trial, randomised controlled study, randomized controlled study, randomised clinical trial, randomized clinical trial, randomised trial, randomized trial, randomised clinical study, randomized clinical study, randomised study, randomized study, RCT, randomisation, randomization, random allocation, random, accident, controlled clinical trial, controlled clinical study, blinded, blinding, blinded study, blinded trial, single-blind, single blind, single-blind method, single blind procedure, single-blind procedure, doubleblind, double blind, double-blind method, double-blind procedure, double blind procedure, triple blind, tripleblind, triple blind method, tripleblind method, triple blind procedure, triple-blind procedure, placebo, placebo effect, validation studies, multicenter studies, multicenter trials, cross-over studies, cross-over trials, crossover procedure, cross-over procedure, sensitivity, specificity, patient-relevant, patient-relevant endpoint, endpoint, patient report, patient statement, arthritis, rheumatoid, breast neoplasms, clinical trials as topic, technology assessment, biomedical, humans, patients, patient satisfaction, evidence-based medicine, biomedical technology assessment, economics, health policy, technology, medical, review literature as topic, costs and cost analysis, cost-benefit analyses, cost effectiveness, rights, decision making, risk assessment, technology, evaluation studies as topic, health, health status, judgement, peer review, meta analysis as topic, randomized controlled trials as topic, single blind method, double blind method, placebos, controlled clinical trials as topic, prospective studies, trial, cross-over, trial, crossover, validation studies as topic, multicenter trial, multicenter studies as topic, models, economic, economics, medical, socioeconomic factors

Summary

Health political background

In Germany the legal framework of the SGB V puts the evaluation of the “medical benefit” in the centre of a technology assessment, although the term “benefit” is not defined by legislation. However, the wording of § 27 suggests, that the requirements of SGB V can be fulfilled by reporting a technology’s effects on morbidity, mortality and quality of life. Further characterisations of these endpoints are found in the “Methodenpapier 3.0 (05/27/2008)” of the Institute for Quality and Efficiency in Health Care (IQWiG). In chapter 3.1. the expression “patient-relevant” is defined as “... how a patient feels, is able to function and participate or survive.” Aside from morbidity, mortality and quality of life, intervention- or disease-specific efforts and patients’ satisfaction with treatment may constitute secondary endpoints. It is mandatory, that all kinds of endpoints directly and reliably indicate changes in health status. Finally, the only way to validly assess patients’ quality of life, some aspects of morbidity and patient satisfaction is by direct questioning of the patients – that is why they are called Patient-Reported Outcomes (PRO).

Scientific background

PRO is used as an umbrella term for different concepts aiming at the measurement of subjectively perceived health status e. g. as treatment effects. Their common characteristic is, that the appraisal of the health status is reported by the patient himself. The American Food and Drug Administration (FDA) defines PRO as follows: “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 are used if a concept may be assessed best by the patient himself. PRO may be elicited either as a single value (e. g. severity of pain) or as a value of change between two measurements (e. g. pain cessation). The measurement of PRO is based on two different approaches. The psychometrical approach refers to the reporting of perceived symptoms (e. g. their presence and severity), capabilities, behaviours, or emotional or mental state. Single dimensions may be summarised forming complex concepts (e. g. quality of life). The preference-based approach refers to the value a patient assigns to a specific health status. Methodological approaches to elicit preferences are based on econometrics and decision theory. The following text exclusively refers to the psychometrical approach.

Aside from health economics, there are two main applications for PRO measurements: as indicators for quality management in health care and as an endpoint in clinical trials. Assessing the benefit of a health care technology on the basis of published clinical trials is the main goal of Health Technology Assessment (HTA). Therefore, de-

termining the role of PRO in HTA will focus on their application in clinical trials.

Psychometric properties such as sufficient validity (content, construct, and criterion), reliability and responsiveness are basic requirements for instruments measuring PRO in clinical trials. Other aspects are administrative and economic feasibility and acceptability.

In order to outline the informative value of PRO in HTA, first of all an overview of concepts, classifications and methods of measurement for PRO is needed. Strengths and weaknesses of the different concepts and measurement approaches need to be pointed out, especially with regard to their psychometric properties and interpretability of results. The overview is based on a systematic analysis of publications that deal with the theoretical framework for the evaluation of PRO.

The current role of PRO measurements in HTA is outlined by two empirical analyses. The first one determines the frequency, the type and consequences of PRO measurement in randomized controlled trials (RCT). Treatment of rheumatoid arthritis with biologicals and chemotherapy of breast cancer were chosen as examples. These indications – treatment pairs were selected in order to represent one chronic disease and one life-threatening acute disease.

The second empirical analysis focuses on the role of PRO in the work of HTA-agencies. HTA-reports on the two example conditions are analysed to determine to which extent conclusions of the reports are based on PRO. Furthermore, method guides of international HTA-agencies are screened in order to determine the role that these agencies attribute to PRO measurements. These overarching objectives are broken down into eight research questions.

Research questions

Theoretical background

1. How are PRO defined in the context of clinical trials?
2. What are the classification systems to characterise endpoints for clinical trials and how are PRO fitted into these systems?
3. What are the methodological approaches for PRO measurements, what are their specific strengths and weaknesses?

Empirical analyses

1. How frequently are PRO chosen and reported as primary and/or secondary endpoints in RCT analysing the effects of antibody treatment of rheumatoid arthritis and chemotherapy of breast cancer?
2. Which PRO are utilized in these trials?
3. Do conclusions on the effectiveness of treatments based on PRO match with the conclusions based on clinical or radiological endpoints?
4. As concerns assessing the health benefit of medical interventions, what is the informative value attributed

to PRO by the methodological literature, especially methods guides of HTA-agencies?

5. Do HTA-reports on rheumatoid arthritis and breast cancer refer to PRO results – especially in their conclusions and recommendations?

Methods

Answering the research questions requires three different literature search strategies.

The first one aims at identifying publications, which report on definition, classification and operationalisation of PRO. It covers the publication period from 1990 to 2009 and is performed in the medical databases of the German Institute of Medical Documentation and Information (DIMDI). The selection of relevant publications is executed by a two-step approach based on defined inclusion and exclusion criteria. A standardised extraction of data employing extraction sheets is impossible, because of the narrative nature of the publications. The extraction of relevant information is performed by matching the publications to a research question in a first step and then by matching relevant information to categories derived from the literature.

The second search strategy aims at identifying RCT that examine the efficacy of biologicals for the treatment of rheumatoid arthritis and chemotherapy for the treatment of breast cancer. Again the search is performed in the medical databases of DIMDI. It covers the publication period from 2005 to 2009. The two-step selection procedure is based on two sets of inclusion and exclusion criteria. Methodological quality of RCT is assessed by a modified version of the Cochrane Risk of Bias Tool. All relevant data are extracted into extraction sheets.

The third search strategy is constructed to retrieve methodological guidance documents from HTA-agencies and institutions of pharmaceutical evaluation as well as HTA-reports on rheumatoid arthritis and breast cancer assessing the treatments specified above. Methodological guidance documents are searched for by hand on the websites of the INAHTA member agencies and institutions (INAHTA = International Network of Agencies for Health Technology Assessment), which are identified by the HTA-report “Methods for the comparative evaluation of pharmaceuticals”. HTA-reports on the two example topics are searched for in the HTA-database of the Centre for Review and Dissemination. The selection of the literature is based on inclusion and exclusion criteria, data extraction is performed into extraction sheets.

Results

After the selection 158 relevant publications remained to answer the methodological research questions (1 to 3). Furthermore, 129 RCT reporting results of chemotherapy in breast cancer patients and 59 RCT reporting results of treatment of rheumatoid arthritis with biologicals are included to answer the empirical research questions related to clinical trials (4 to 6). To answer the empirical

questions concerning HTA-reports (7, 8) 34 methodological guidance documents, twelve reports on chemotherapy of breast cancer and nine reports on biological therapy of rheumatoid arthritis are included.

Theoretical issues

In its current meaning the term PRO is used since about the year 2000. Its definition is pragmatic and refers to the source of information. Essential aspects of a PRO are the patient report and his or her prerogative of interpretation of the health status. PRO substituted the term “quality of life” as the umbrella term for endpoints reported by patients themselves. Constructs subsumed under the umbrella term PRO are perceived symptoms, functioning, health perception, satisfaction and (health related) quality of life. In the methodological literature there is no consensus regarding the definition of these constructs. Several research groups try to integrate the different constructs into comprehensive models, which display the relationships and interactions between parameters and illustrate external influences. Wilson and Cleary created a linear, five-step model. Its poles represent biological and physiological variables on the one side and overall quality of life on the other side. Symptom status, functional status and general health perceptions bridge the gap between the two poles. Between two neighbouring constructs interdependencies are assumed.

Another comprehensive model is the International Classification of Functioning, Disability and Health (ICF) of the World Health Organisation (WHO). The key intention of the ICF is the description of the health state of an individual with regard to the dimensions body structures, body functions and activities and participation. The health state assigned by the ICF is influenced directly by the underlying disease (classified by International Classification of Diseases (ICD)) and environmental as well as personal factors.

Valderas and Alonso integrate both approaches. Their model is designed to classify instruments measuring patient reported endpoints. Definition and classification take three aspects into consideration: construct, target population and measurement approach.

The major part of the methodological literature provides information related to the operationalisation of PRO. Any PRO measurement should be based on a conceptual framework, which specifies the relevant constructs, the definition of the target population and an endpoint model specifying relevant variables and instruments.

There is a variety of instruments available to measure PRO. Most of them are based on classical test theory, some of them on a newer approach called item response theory. Instruments based on item response theory are shorter and produce more focussed results. Psychometric instruments document the description of a health status from the perspective of the patient. Preference based instruments require the valuation of a health status by the patient. Generic instruments may be applied independently of population characteristics while specific in-

struments may only be utilized in populations with certain characteristics (e. g. disease specific instruments). Usually results from measurements with specific instruments are more precise than those from generic instruments, while the latter allow comparisons across patient populations with different diseases. Furthermore, there are profile and index instruments. While profile instruments present separate values for every construct measured the results of index instruments are summarised into one figure.

During the development of new instruments item generation is a pivotal step. According to the literature a mixed approach consisting of literature studies, expert input and focus groups is considered state of the art.

Measurement properties are crucial for the applicability of an instrument. Main properties are validity, reliability and responsiveness. Validity may be determined as content, construct or criterion validity – depending on the type of instrument and the available comparisons. Reliability primarily refers to test-retest-reliability and internal consistency. Responsiveness describes the ability of an instrument to record changes of an endpoint. It is determined by a distribution based or an anchor based approach. The former utilizes effect size while the latter requires comparisons with an external anchor (gold standard).

A pivotal concept for the interpretation of results from PRO measurements is the minimal important difference (MID) – describing the smallest difference between two measurement results that a patient considers relevant. There is no standard approach to determine the MID. The most frequently used approaches correspond with the approaches used to determine responsiveness.

For methodological and economical reasons the use of existing standardised and validated instruments is preferable to the development of new instruments. In some situations it may be necessary to modify an instrument to match a specific context. A particularly critical point is the translation of instruments into a foreign language. Guidelines suggest a ten-step process of forward and backward translations. In any case a modified instrument needs to be revalidated.

PRO in clinical trials

The literature searches retrieved 73 publications reporting on 59 RCT that investigate the effects of treating rheumatoid arthritis with biological drugs. Most frequently the American College of Rheumatology (ACR) core set of disease activity measures is used as the primary endpoint. This composite endpoint consists of seven criteria, among them three PRO (pain, global assessments of health status and functioning). For a positive score, improvements in at least one PRO criterion are required. The 20-percent criterion, which requires an overall 20-percent improvement is used as primary endpoint in 23 studies. Pure PRO are less frequently used as primary endpoints. The most frequently applied PRO is the health assessment questionnaire (HAQ), a questionnaire investigating functional status by 20 items. It is defined as the

primary endpoint in six trials. The disease activity score (DAS) is the most frequently studied secondary endpoint (39 trials). The DAS again is a composite endpoint consisting of four criteria, among them one PRO (health perception). Pure PRO are defined secondary endpoints in 34 trials. The dominant constructs are functional status and health perception. The HAQ (32 trials) and the SF-36 (16 trials) are the most frequently applied questionnaires. The SF-36 measures health perception by asking 36 questions across eight domains. Stratified analysis by methodological quality of trials indicates that trial quality does neither predict the use or non-use of PRO nor the type of constructs investigated.

Five trials report contradictory results of clinical and patient reported outcomes. In one trial worsening of radiological endpoints (synovitis, oedema, and erosions) is not reflected in HAQ results. In the other four trials the occurrence of adverse events (fever, headache, asthenia, and infections) is not reflected by the results of the HAQ or SF-36 respectively. These findings are not commented by the authors. In general, PRO results are commented in the discussion parts of most trials that applied PRO measures. Of 35 trials that discuss the comparison of PRO results and results from traditional endpoints 19 authors come to the conclusion that PRO results are supportive. Interdependencies between PRO and traditional endpoints are discussed in twelve studies. Findings are not always coherent, e. g. between progression of radiologically attested joint erosions and HAQ results. Still, frequently interdependency between disease activity and the HAQ results is reported.

15 studies take reference to PRO in their conclusions – mostly to functional status and health perception. This is particularly true for trials where PRO results and results from traditional endpoints are congruent and therefore suggest identical conclusions. In two trials results from PRO (adherence to therapy, quality of life) are used to refine recommendations. Six trials identify the need for further research to optimise the use of PRO. They recommend the generation of long-term data and the further analysis of interdependencies between clinical and patient reported outcomes.

The literature searches retrieved 129 publications reporting on 123 RCT that investigate the effects of chemotherapy in breast cancer patients. In these trials the most frequently used primary endpoint is survival (66 times in 53 studies). Only six studies utilize PRO as primary endpoint, mostly quality of life. Three of these studies apply the cancer specific EORTC questionnaire. In 22 trials PRO are defined as secondary endpoints. Again, the quality of life questionnaire of the EORTC is the most commonly applied instrument (twelve trials). As for the rheumatology trials, stratified analysis by methodological quality again indicates that trial quality does neither predict the use or non-use of PRO nor the type of constructs investigated. Of 26 studies, which present results of PRO, 15 report statistically significant differences between treatment groups. One trial detects improvements of quality of life that are not reflected in overall survival and progression

free survival. 24 of 28 trials reporting PRO results comment them in the discussion section: 18 trials by comparing PRO and traditional endpoint results, three trials by pointing out inconsistencies. Interdependencies are discussed in six studies of which three find a strong influence of adverse events on quality of life. Nine trials identify the need for further research, especially on how to implement PRO information into routine patient management.

PRO in methodological guidance documents of HTA-agencies

The searches retrieved 34 methodological guidance documents from the websites of 20 HTA-agencies and pharmaceutical regulatory bodies. 18 of these papers address the topic of PRO to a relevant extent. Three aspects are discussed in some detail: The role of PRO in decision making – some agencies equally weigh PRO results and clinical results while others regard results from PRO measurements as supplementary information only. Furthermore, technical aspects of PRO measurement (choice of instrument, psychometric properties) and the role of quality of life measurement in economic evaluation are addressed.

PRO in HTA-reports (rheumatoid arthritis)

Handsearches and database searches identified nine HTA-reports addressing the treatment of rheumatoid arthritis with biological drugs. In their background section seven reports justified the use of PRO for this indication by referring to the disease's impact on functional status and quality of life.

Analyses of the methods section of reports revealed that nearly half of the documents prospectively specified PRO as relevant endpoints for the assessment. Predominantly the composite endpoints ACR20, ACR50 and ACR70 are referred to but single measures of symptoms (pain scale), functioning (HAQ) or quality of life are mentioned as well. In all reports results of PRO measurements are discussed explicitly, with a special focus on the ACR criteria. Two reports compare the results from traditional endpoints with PRO results. In the majority of reports the results from PRO measurements are regarded insufficient yet for decision making. The authors point out the need for more research in order to bridge the gap.

PRO in HTA-reports (breast cancer)

Twelve HTA-reports were found that assess different types of chemotherapy for breast cancer. In their background section only three of these twelve reports refer to aspects of everyday life that may be impacted by chemotherapy. These reports regard symptom relief and quality of life along with prolongation of life as the main goals of chemotherapy.

As specified in the methods sections, quality of life data and other constructs such as pain, fear, depression and functional status are taken into consideration by five re-

ports. One report refers to PRO for input into cost-utility-analyses.

PRO results, mainly quality of life data, are presented in the results section of four reports. In their discussion section, three reports refer to quality of life data. Of these, two reports discuss the fact that improvements in disease-free survival not necessarily translate into improvements of quality of life. On the other hand the inevitable relation of toxicity and reduced quality of life is pointed out.

Discussion

The results of the presented report may be compromised by some methodological limitations. First, since there is no controlled vocabulary covering the issue of PRO and the range of databases and search period was limited, some theoretical and methodological publications may be missed. Still it may be assumed that key publications would have been found in reference lists of the publications available. Second, some of the theoretical publications may not represent original research. The selection was hampered by the fact, that there is no specific study type that can guide a selection of methodological literature. This problem may be subordinate though, since no quantitative analyses of the theoretical literature have been performed.

In summary, the research questions may be answered as follows:

Currently the term PRO is used as an umbrella term for patient reported endpoints. Quality of life, which has formerly been used as the umbrella term, is now one of several constructs subordinated to PRO (e. g. symptoms, functional status, health perception, quality of life). There is actually no consensus, which other constructs are part of PRO. Most authors find a rigorous separation of traditional outcomes and PRO unreasonable. Models offer an integrated perspective which presents a holistic impression of the disease and its consequence. Furthermore, interdependencies between endpoints may be visualised. The most complex and most pragmatic approach is presented by Valderas and Alonso.

There are different approaches for the measurement of PRO. Psychometric and preference-based approaches, profiles and indices, generic and specific instruments all have their own strengths and weaknesses and must be chosen depending on the context of the research project. Validity, reliability and responsiveness represent the psychometric properties of an instrument. For methodical and economic reasons the use of existing standardised and validated instruments is preferable to the development of new instruments. If it is necessary to modify an existing instrument, a revalidation of the modified version is required. Relating to the interpretation of results the MID is the pivotal aspect. Item generation for new instruments is still based on the principles of the classical test theory. But new approaches, e. g. the item response theory, which provides shorter and more precise instruments, are improving.

The frequency of PRO in RCT relating to rheumatoid arthritis indicates that the informative value of PRO is regarded high. Composite endpoints (ACR criteria, DAS) which combine clinical endpoints and PRO are most frequently used. Pure PRO, especially functioning and health perception (HAQ, SF-36) are often found as secondary endpoints. A sophisticated discussion of the interdependencies of PRO and traditional endpoints, referring to the specific condition is seldom found.

The RCT of breast cancer use PRO only to a limited extent (24 % of studies) and mainly as secondary endpoints. Quality of life is the only construct that is referred to in these trials. Again, an integrating discussion of PRO and traditional endpoints is hardly found.

About half of the retrieved methodological guidance papers address the informative value of PRO for the evaluation of health benefits as well as for economic evaluations. Comprehensive information is presented for measurement instruments. Reports relating to the indications reflect the informative value of PRO in primary studies. It needs to be pointed out that in most reports a need for further research concerning PRO is stated – this implies, that HTA-agencies promote the use of PRO in a proactive manner.

Corresponding author:

Dr. med. Dagmar Lühmann
University Lübeck, Institute for Social Medicine,
Ratzeburger Allee 160, 23538 Lübeck, Germany, phone:
0049(0)451/5005881
Dagmar.Luehmann@uk-sh.de

Please cite as

Brettschneider C, Lühmann D, Raspe H. Informative value of Patient Reported Outcomes (PRO) in Health Technology Assessment (HTA). *GMS Health Technol Assess.* 2011;7:Doc14.
DOI: 10.3205/hta000092, URN: urn:nbn:de:0183-hta0000920

This article is freely available from

<http://www.egms.de/en/journals/hta/2011-7/hta000092.shtml>

Published: 2011-02-02

The complete HTA Report in German language can be found online at: http://portal.dimdi.de/de/hta/hta_berichte/hta267_bericht_de.pdf

Copyright

©2011 Brettschneider et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en>). You are free: to Share – to copy, distribute and transmit the work, provided the original author and source are credited.

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

Zusammenfassung

Hintergrund und Zielsetzung

Patient-Reported Outcome (PRO) wird als Oberbegriff für unterschiedliche Konzepte zur Messung subjektiv empfundener Gesundheitszustände verwendet. Sie sind dadurch gekennzeichnet, dass der Patient selbst seine Einschätzung berichtet. Um den Stellenwert von PRO im Kontext von HTA-Verfahren (HTA = Health Technology Assessment) zu beschreiben, wird zunächst eine Übersicht über Konzepte, Klassifikationen und methodische Messansätze erstellt. Diese Übersicht wird ergänzt um eine empirische Analyse von klinischen Studien und HTA-Berichten mit dem Ziel, Art, Häufigkeit und Konsequenzen der verwendeten PRO zu dokumentieren.

Methodik

Beide Fragestellungen werden mithilfe von systematischen Literaturübersichten bearbeitet. Für den methodischen Teil wird in den medizinischen Datenbanken des Deutschen Instituts für Medizinische Dokumentation und Information (DIMDI) mit einer Recherchestrategie aus drei Modulen im Zeitraum von 1990 bis 2009 gesucht. Die Recherche nach randomisierten klinischen Studien (RCT) zu den Krankheitsbildern rheumatoide Arthritis (RA) und Mammakarzinom erfolgt ebenfalls in den Datenbanken des DIMDI, für den Zeitraum von 2005 bis 2009. Die Recherche nach HTA-Berichten und -Methodenpapieren umfasst die Datenbanken des Centre for Reviews and Dissemination (CRD) und Handsuchen. Für alle Fragestellungen werden spezifische Ein- und Ausschlusskriterien zur Literaturselektion definiert. Die methodische Qualität der RCT wird mithilfe eines in Anlehnung an das „Risk of Bias Tool“ der Cochrane Collaboration konzipierten Instruments bewertet. Die Informationsextraktion erfolgt für den methodischen Teil strukturiert durch die Kapitelgliederung, für den empirischen Teil in Extraktionsbögen. Alle Informationen werden qualitativ beschreibend zusammengefasst.

Ergebnisse

Aus den Recherchen können 158 Dokumente zur Bearbeitung der methodischen Fragestellungen (87 Dokumente zu Definition/Klassifikation; 125 Dokumente zur Operationalisierung) und 225 RCT (77 RA, 148 Mammakarzinom) sowie 40 HTA-Berichte zur Bearbeitung der empirischen Fragestellungen gewonnen werden. Die Analysen zu Definitionen bestätigen PRO als Oberbegriff für eine Vielzahl von patientenberichteten Endpunkten. Das neueste Klassifikationssystem ermöglicht die Beschreibung der PRO nach dem Konstrukt, der Zielpopulation und der Messmethode. Ausführungen zur Operationalisierung beziehen sich auf den Konzeptrahmen, die Instrumentenentwicklung, -eigenschaften und -modifikationsmöglichkeiten. Von 59 Studien zur Antikörpertherapie der RA verwenden sieben ausschließlich PRO, 38 gemischte Zielgrößen (American College of Rheumatology [ACR], Disease Activity Score [DAS]) und zehn rein klinische bzw. radiologische Parameter als primäre Outcomes. Von 123 Studien zur Chemotherapie des Mammakarzinoms

**Christian
Brettschneider¹
Dagmar Lühmann²
Heiner Raspe²**

1 Universitätsklinikum
Hamburg-Eppendorf, Institut
für Medizinische Soziologie,
Sozialmedizin und
Gesundheitsökonomie,
Hamburg, Deutschland

2 Universitätsklinikum
Schleswig-Holstein, Campus
Lübeck, Institut für
Sozialmedizin, Lübeck,
Deutschland

stützen sich nur sechs auf PRO als primäre Zielgröße; 98 Studien gebrauchen klinische Parameter (Überlebenszeit, Tumoransprechen, Progression). Abweichungen von der Gesamtzahl resultieren aus ungenauen Angaben der Zielgrößen. Diese Verteilung spiegelt sich auch in den analysierten HTA-Berichten wieder. In den Berichten zur RA werden durchweg PRO-Zielgrößen berichtet, während in den Berichten zum Mammakarzinom dies in knapp der Hälfte der Publikationen der Fall ist. Zusammenhänge zwischen Studienqualität und der Verwendung von PRO sind nicht erkennbar.

Schlussfolgerungen

Für die Definition und die Klassifikation von PRO existieren inzwischen schlüssige Konzepte, deren Umsetzung wissenschaftlichen Kriterien genügen muss. Die Häufigkeit und die Art der in klinischen Studien verwendeten PRO variieren abhängig vom untersuchten Krankheitsbild. Im Kontext von HTA wird die Notwendigkeit zur Erfassung von PRO wahrgenommen, bei fehlenden Daten wird Forschungsbedarf formuliert.

Schlüsselwörter: patientenrelevante Endpunkte, patienten-relevante Endpunkte, patienten-relevant, patientenrelevant, Endpunkt, Patientenbericht, Patientenangabe, Patientenaussage, Selbsteinschätzung, Patientenbeurteilung, Patientensicht, Patientenurteil, patientenberichtete Zielgröße, Lebensqualität, rheumatoide Arthritis, Mammakarzinom, klinische Studien, Health Technology Assessment, Konzept, HTA-Bericht, systematische Übersicht, Übersichtsarbeit, Effektivität, Wirksamkeit, Kosteneffektivität, Kosten Effektivität, Kosten-Effektivität, Kosten, Kostenanalyse, Kostenkontrolle, Krankheitskosten, Kostensenkung, Kostenminimierung, Kostenreduktion, Gesundheitsfinanzierung, HTA, Technologiebeurteilung, medizinische Bewertung, medizinische Technologie, medizinische Beurteilung, Gesundheitsökonomie, gesundheitsökonomische Studien, Übersichts-literatur, Forschungsartikel, HTA Bericht, evidenzbasierte Medizin, EBM, Diagnose, Ethik, Recht, sozio-ökonomische Faktoren, sozioökonomische Faktoren, Sozioökonomie, ökonomischer Aspekt, Pharmaökonomie, Prävention, Vorsorge, Rehabilitation, Reha, Therapie, Behandlung, Methoden, Methodik, Pflege, Meta-Analyse, Metaanalyse, Plazebo, Placebo, Placeboeffekt, RCT, randomisierte Studie, randomisierte kontrollierte Studie, randomisierte klinische Studie, randomisierter Versuch, randomisierte Zuordnung, Randomisierung, Zufall, CCT, CT, kontrollierte klinische Studie, klinische Studie, verblindet, Verblindung, verblindete Studie, Genauigkeitsstudie, Multicenter Studie, Multizenter Studie, Multicenter-Studie, Multizenter-Studie, Multicenterstudie, Multizenterstudie, Sensitivität, Spezifität, Crossover-Studien, Cross-over-Studien, Crossoverstudien, Cross-over-Verfahren, Crossover-Verfahren, Crossoververfahren, Arthritis, rheumatoide, Mammatumoren, Mensch, Technikfolgen-Abschätzung, biomedizinische, Mensch, Patienten, Endpunktbestimmung, Patientenzufriedenheit, gutachtenbasierte Medizin, Technologie, medizinische, Ökonomie, Technologie, Gesundheitspolitik, Gesundheit, Gesundheitszustand, Kosten und Kostenanalyse, Kosten-Nutzen-Analyse, Risikoabschätzung, Beurteilung, Evaluationsstudien, Peer Review, randomisierte kontrollierte Studien, Plazebos, Placeboeffekt, kontrollierte klinische Studien, kontrollierte klinische Versuche, prospektive Studien, Validierungsstudien, multizentrische Studien, sozialökonomisch, Modelle, ökonomische, Ökonomie, ärztliche, Effizienz, Rechte, Entscheidungsfindung

Kurzfassung

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.

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 Recher-

che 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 Gesund-

heitsproblem (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 Kriteriumswerts) 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 Lebensqualitätsmessungen 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 Methodenkapitel 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 medizinische 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 Verwen-

dung von PRO in HTA-Institutionen vorausschauend verfolgt wird.

Korrespondenzadresse:

Dr. med. Dagmar Lühmann
Universitätsklinikum Schleswig-Holstein, Campus Lübeck,
Institut für Sozialmedizin, Ratzeburger Allee 160, 23538
Lübeck, Deutschland, Tel.: 0049(0)451/5005881
Dagmar.Luehmann@uk-sh.de

Bitte zitieren als

Brettschneider C, Lühmann D, Raspe H. Informative value of Patient Reported Outcomes (PRO) in Health Technology Assessment (HTA). *GMS Health Technol Assess.* 2011;7:Doc14.
DOI: 10.3205/hta000092, URN: urn:nbn:de:0183-hta0000920

Artikel online frei zugänglich unter
<http://www.egms.de/en/journals/hta/2011-7/hta000092.shtml>

Veröffentlicht: 02.02.2011

Der vollständige HTA-Bericht in deutscher Sprache steht zum kostenlosen Download zur Verfügung unter:
http://portal.dimdi.de/de/hta/hta_berichte/hta267_bericht_de.pdf

Copyright

©2011 Brettschneider et al. Dieser Artikel ist ein Open Access-Artikel und steht unter den Creative Commons Lizenzbedingungen (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.de>). Er darf vervielfältigt, verbreitet und öffentlich zugänglich gemacht werden, vorausgesetzt dass Autor und Quelle genannt werden.