

Methoden zur vergleichenden Bewertung pharmazeutischer Produkte

Methods for the comparative evaluation of pharmaceuticals

Abstract

Political background

As a German novelty, the Institute for Quality and Efficiency in Health Care (*Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; IGWiG*) was established in 2004 to, among other tasks, evaluate the benefit of pharmaceuticals. In this context it is of importance that patented pharmaceuticals are only excluded from the reference pricing system if they offer a therapeutic improvement.

The institute is commissioned by the Federal Joint Committee (*Gemeinsamer Bundesausschuss, G-BA*) or by the Ministry of Health and Social Security. The German policy objective expressed by the latest health care reform (*Gesetz zur Modernisierung der Gesetzlichen Krankenversicherung, GMG*) is to base decisions on a scientific assessment of pharmaceuticals in comparison to already available treatments. However, procedures and methods are still to be established.

Research questions and methods

This health technology assessment (HTA) report was commissioned by the German Agency for HTA at the Institute for Medical Documentation and Information (DAHTA@DIMDI). It analysed criteria, procedures, and methods of comparative drug assessment in other EU-/OECD-countries. The research question was the following: *How do national public institutions compare medicines in connection with pharmaceutical regulation, i.e. licensing, reimbursement and pricing of drugs?*

Institutions as well as documents concerning comparative drug evaluation (e.g. regulations, guidelines) were identified through internet, systematic literature, and hand searches. Publications were selected according to pre-defined inclusion and exclusion criteria. Documents were analysed in a qualitative matter following an analytic framework that had been developed in advance. Results were summarised narratively and presented in evidence tables.

Results and discussion

Currently licensing agencies do not systematically assess a new drug's added value for patients and society. This is why many countries made post-licensing evaluation of pharmaceuticals a requirement for reimbursement or pricing decisions. Typically an explicitly designated drug review body is involved.

In all eleven countries included (Austria, Australia, Canada, Switzerland, Finland, France, the Netherlands, Norway, New Zealand, Sweden, and the United Kingdom) a drug's *therapeutic* benefit in comparison to treatment alternatives is leading the evaluation. A medicine is classified as a therapeutic improvement if it demonstrates an improved benefit-/risk-profile compared to treatment alternatives. However, evidence of superiority to a relevant degree is requested.

Annette Zentner¹

Marcial

Velasco-Garrido¹

Reinhard Busse¹

¹ Fachgebiet Management im Gesundheitswesen, Institut für Gesundheitswissenschaften, Fakultät VIII, Technische Universität Berlin, Berlin, Deutschland

Health related quality of life is considered as the most appropriate criterion for a drug's added value from *patients' perspective*. Review bodies in Australia, New Zealand, and the United Kingdom have committed themselves to include this outcome measure whenever possible. *Pharmacological or innovative* characteristics (e.g. administration route, dosage regime, new acting principle) and *other advantages* (e.g. taste, appearance) are considered in about half of the countries. However, in most cases these aspects rank as second line criteria for a drug's added value.

All countries except France and Switzerland perform a comparative pharmacoeconomic evaluation to analyse costs caused by a drug intervention in relation to its benefit (preferably by cost utility analysis). However, the question if a medicine is cost effective in relation to treatment alternatives is answered in a political and social context. A range of remarkably varying criteria are considered.

Countries agree that randomised controlled *head-to-head* trials (*head-to-head RCT*) with a high degree of internal and external validity provide the most reliable and least biased evidence of a drug's relative treatment effects (as do systematic reviews and meta-analyses of these RCT). Final outcome parameters reflecting long-term treatment objectives (mortality, morbidity, quality of life) are preferred to surrogate parameters. Following the concept of *community effectiveness*, drug review institutions also explicitly favour RCT in a "natural" design, i.e. in daily routine and country specific care settings.

The countries' requirements for pharmacoeconomic studies are similar despite some methodological inconsistencies, e.g. concerning cost calculation.

Outcomes of clinical and pharmacoeconomic analyses are largely determined by the choice of comparator. Selecting an appropriate comparative treatment is therefore crucial. In theory, the best or most cost effective therapy is regarded as appropriate comparator for clinical and economic studies. Pragmatically however, institutions accept that the drug is compared to the treatment of daily routine or to the least expensive therapy.

If a pharmaceutical offers several approved indications, in some countries all of them are assessed. Others only evaluate a drug's main indication. Canada is the only country which also considers a medicine's off-label use.

It is well known that clinical trials and pharmacoeconomic studies directly comparing a drug with adequate competitors are lacking - in quantitative as well as in qualitative terms. This is specifically the case before or shortly after marketing authorisation. Yet there is the need to support reimbursement or pricing decisions by scientific evidence. In this situation review bodies are often forced to rely on observational studies or on other internally less valid data (including expert and consensus opinions). As a second option they use statistical approaches like indirect adjusted comparisons (in Australia and the United Kingdom) and, commonly, economic modelling. However, there is consensus that results provided by these techniques need to be verified by valid head-to-head comparisons as soon as possible.

Conclusions

In the majority of countries reimbursement and pricing decisions are based on systematic and evidence-based evaluation comparing a drug's clinical and economic characteristics to daily treatment routine. However, further evaluation criteria, requirements and specific methodological issues still lack internationally consented standards.

Zusammenfassung

Gesundheitspolitischer Hintergrund

Erstmals in Deutschland wurde 2004 die Nutzenbewertung von Arzneimitteln gesetzlich verankert, mit der das neu etablierte Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IGWiG) vom Gemeinsamen Bundesausschuss (G-BA) beauftragt werden kann. Eine differenzierte Bewertung von Medikamenten im Vergleich zu bereits vorhandenen Therapiemöglichkeiten ist u.a. bedeutsam, da Arzneimittel mit patentgeschützten Wirkstoffen durch das Gesetz zur Modernisierung der Gesetzlichen Krankenversicherung (GMG) lediglich von der Festbetragsregelung ausgenommen sind, wenn sie eine therapeutische Verbesserung bieten. Jedoch müssen Verfahren und Methoden noch etabliert werden.

Forschungsfragen und Methodik

Vor diesem gesundheitspolitischen Hintergrund untersucht der vorliegende Health Technology Assessment (HTA)-Bericht im Auftrag der Deutschen Agentur für HTA beim Deutschen Institut für Medizinische Dokumentation und Information (DAHTA@DIMDI), wie in anderen Industrieländern Medikamente vergleichend bewertet werden. Die Forschungsfragestellung lautete: *Wie werden Arzneimittel durch öffentliche Institutionen in anderen EU-/OECD-Ländern im Rahmen der Arzneimittelregulierung, d.h. bei der Marktzulassung, bei der Entscheidung zur Aufnahme in den Leistungskatalog und bei der Preisbildung, verglichen (EU = Europäische Union, OECD = Organisation for Economic Co-operation and Development)?*

Durch Internet-, Hand- und systematische Literaturrecherche wurden Dokumente zu Arzneimittelbewertungsverfahren-, -kriterien und -methoden identifiziert, anhand definierter Ein- und Ausschlusskriterien selektiert und qualitativ anhand eines Rasters ausgewertet. Die Ergebnisse wurden in narrativer Form und in Evidenztabellen präsentiert.

Ergebnisse und Diskussion

Bei der Marktzulassung von Medikamenten ist eine systematische Bewertung im Vergleich zu therapeutischen Alternativen nicht vorgesehen. Daher haben viele Länder die Evaluation von Arzneimitteln nach der Marktzulassung zur Voraussetzung für Kostenerstattung oder Preisbildung gemacht. Involviert in diesen Prozess sind ausdrücklich dafür ausgewiesene öffentliche Institutionen oder Fachgremien.

In allen eingeschlossenen elf Ländern (Österreich, Australien, Kanada, Schweiz, Finnland, Frankreich, Niederlande, Neuseeland, Norwegen, Schweden und Großbritannien) ist der zusätzliche *klinisch-therapeutische* Nutzen im Vergleich zu Behandlungsalternativen das leitende Kriterium bei der Bewertung eines Medikaments. Die Klassifikation des Arzneimittels als ein therapeutischer Fortschritt wird mit der Forderung verbunden, dass ein verbessertes Wirksamkeits-/Risikoprofil, d.h. Überlegenheit (Superiorität) in relevantem Ausmaß gegenüber therapeutischen Alternativen, nachgewiesen wird. Als das entscheidende Kriterium eines Mehrnutzens aus *Patientensicht* gilt die verbesserte gesundheitsbezogene Lebensqualität, auf deren Einbeziehung sich die Institutionen in Australien, Neuseeland und Großbritannien festgelegt haben.

Pharmakologische Vorteile (z.B. Applikationsfrequenz, Darreichungsform, Galenik), *innovative* Eigenschaften (z.B. neues Wirkprinzip) oder *andere Merkmale* (z.B. Geschmack, Aussehen) werden in rund der Hälfte der

Länder berücksichtigt, sind jedoch nach Wirksamkeit und Sicherheit nachrangige Kriterien.

In allen Ländern außer in Frankreich und der Schweiz werden als weiteres Bewertungskriterium die durch die Arzneimittelbehandlung ausgelösten *Kosten* in Relation zum Mehrnutzen ermittelt (vorzugsweise mittels Kostennutzwertanalysen). Die Frage eines angemessenen Verhältnisses wird jedoch politisch und im gesellschaftlichen Kontext unter Berücksichtigung einer Reihe - international sehr inhomogener Kriterien - beantwortet.

Als verlässlichste Evidenzgrundlage gelten übereinstimmend randomisierte, kontrollierte *Head-To-Head*-Studien mit hoher interner und externer Validität und/oder systematische Übersichten und Metaanalysen dieses Studientyps. Die Mehrzahl der Länder bevorzugt die Messung finaler Ergebnisparameter i.S. von endgültigen, durch die Therapie ausgelösten Veränderungen des Gesundheitszustands (Mortalität, Morbidität, Lebensqualität) und unter den Bedingungen der landesspezifischen Versorgungspraxis (*Community Effectiveness*). Die Anforderungen an gesundheitsökonomische Studien decken sich in vielen Bereichen, beinhalten aber einige methodische Inkonsistenzen z.B. bei der Kostenkalkulation.

Die Definition und sorgfältige Auswahl von geeigneten Komparatoren ist in der therapeutischen wie gesundheitsökonomischen Analyse von großer Bedeutung für die Ergebnisse einer vergleichenden Evaluation. Als geeigneter Komparator gilt theoretisch die beste bzw. kosteneffektivste Therapie. Als pragmatischen Ansatz akzeptieren die Institutionen jedoch regelhaft die derzeit übliche bzw. günstigste Behandlungspraxis. Überprüft werden teils nur die Hauptindikation, teils alle Anwendungsgebiete und lediglich in Kanada auch der *Off-Label-Use* eines Medikaments.

Auch bei einem quantitativen oder qualitativen Mangel an direkt vergleichenden klinischen Anwendungsstudien (Phase IV) und an gesundheitsökonomischen Analysen - vor allem vor oder kurz nach der Marktzulassung des Produkts - müssen Gutachter Empfehlungen für gesundheitspolitische Entscheidungen aussprechen. In dieser Situation sind sie oftmals gezwungen, auf Beobachtungsstudien oder andere weniger intern valide Daten einschließlich Konsensus- und Expertenmeinungen auszuweichen. Zum zweiten verwenden oder fordern die Institutionen indirekte adjustierte Vergleiche (Australien, Großbritannien) und umfassende gesundheitsökonomische Modellierungstechniken. Übereinstimmend wird formuliert, dass die Ergebnisse solcher Verfahren jedoch frühest möglich durch direkte Vergleichsstudien verifiziert werden sollten.

Schlussfolgerungen

Im Rahmen von gesundheitspolitischen Entscheidungen zum Leistungskatalog oder zur Preisbildung überwiegt international die systematische und evidenzbasierte Bewertung der klinischen und ökonomischen Charakteristika eines Arzneimittels im Vergleich zur üblichen Behandlungspraxis. Jedoch steht ein international konsentierter Standard bei vielen Bereichen des Verfahrens und bei methodischen Einzelaspekten aus.

Executive Summary

1. Political background

According to the Act for Modernisation of the Statutory Health Insurance (*Gesetz zur Modernisierung der Gesetzlichen Krankenversicherung; GMG*) the German Institute for Quality and Efficiency in Health Care (*Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; IQWiG*) was established in 2004 to, among other tasks, evaluate the benefit of pharmaceuticals. The institute is commissioned by the Federal Joint Committee (*Gemeinsamer Bundesausschuss; G-BA*) or by the Ministry of Health and Social Security. In this context it is of importance that patented (and often expensive) pharmaceuticals are only excluded from the reference pricing system if they offer a therapeutic improvement.

2. Research questions

The German health policy objective is to perform a differentiated assessment of pharmaceuticals in comparison to already available treatments but procedures and methods are still to be established. This health technology assessment (HTA) report was commissioned by the German Agency for HTA at the Institute for Medical Documentation and Information (DAHTA@DIMDI). DAHTA@DIMDI asked the question: *Are there methods for a comparative evaluation of pharmaceutical products? Which are these?* The research question was specified as follows: *How do other EU-/OECD-countries compare medicines in connection with pharmaceutical regulation, i.e. licensing, reimbursement and pricing?* This study thus analysed criteria, procedures, and methods of comparative drug assessment used by public institutions in other countries.

3. Methods

As a first step national public institutions in other EU-/OECD-countries as well as documents concerning comparative drug evaluation (e.g. regulations, guidelines) were identified through internet searches.

Additionally a systematic literature search was performed: 27 databases for medical, economic, social, and natural sciences were searched for articles published by February 2005. Synonyms for the term "drug evaluation" in combination with names of the respective country were used as search terms. After duplicates had been removed, 1619 publications were analysed for their relevance in a two step selection process, in accordance with pre-defined inclusion and exclusion criteria. Finally 85 articles were included in the study.

The search was complemented by hand search in bibliographies and, in case of doubt, by email communication with the institutions. All documents were analysed in a qualitative matter and according to an analytic framework that had been developed in advance. Results were summarised narratively and presented in evidence tables.

4. Results and discussion

Currently licensing agencies (i.e national institutions and the European Agency for the Evaluation of Medicinal Products, EMEA) do not systematically assess a new drug's added value for patients and society. This is why many countries made post-licensing evaluation of pharmaceuticals a requirement for reimbursement or pricing decisions. Typically an explicitly designated drug review body is involved.

In this study twelve national public institutions in eleven EU-/OECD-countries performing comparative drug evaluation were included: Commission for Evaluation of Pharmaceuticals (*Heilmittlevaluierungskommission; HEK*) in Austria, Pharmaceutical Benefits Advisory Committee (PBAC) in Australia, Patented Medicine Prices Review Board (PMPRB) and Canadian Expert Drug Advisory Committee (CEDAC) in Canada, Confederate Pharmaceutical Commission (*Eidgenössische Arzneimittelkommission, EAK*) in Switzerland, Pharmaceuticals Pricing Board in Finland, Transparency Commission (*Commission de Transparence*) in France, Committee for Pharmaceutical Aid (*Commissie Farmaceutische Hulp, CFH*) in the Netherlands, Norwegian Medicines Agency (NoMA) in Norway, Pharmaceutical Management Agency (PHARMAC) in New Zealand, Pharmaceutical Benefits Board (PBB) in Sweden, and National Institute for Clinical Excellence (NICE) in the United Kingdom.

In about half of the countries national drug review bodies act as *advisors*. They make reimbursement or pricing *recommendations* to a national or regional government, a ministerial department or to a self-governance body (in Austria, in Australia, in Canada (reimbursement), in France (reimbursement), in the Netherlands, in Switzerland, and in Norway).

The remaining countries have established *regulatory-type* drug review bodies. Though accountable to health ministries, they function as the *decision-maker* responsible for listing or pricing of drugs (in Canada (pricing), in Finland, in France (pricing), in New Zealand, in Sweden, and in the United Kingdom).

Information provided by the documents analysed on drug evaluation procedures and methods varied to a great extent among countries. Australia and Canada belong to the countries which published the most comprehensive and detailed guidelines. Procedures in these countries seem to be characterised by a high degree of transparency - at least in theory.

It can be stated that pharmacoeconomic methodologies are described in more detail than procedures for the clinical review or for assessment of other drug evaluation aspects.

In all countries a drug's *therapeutic* benefit in comparison to treatment alternatives is leading the evaluation. A medicine is classified as a therapeutic improvement if it demonstrates an improved benefit-/risk-profile compared to treatment alternatives. However, evidence of superiority to a relevant degree is requested.

Health related quality of life is considered as the most appropriate criterion for a drug's added value from *patients' perspective*. Review bodies in Australia, New Zealand, and the United Kingdom have committed themselves to include this outcome measure whenever possible.

Pharmacological or innovative characteristics (e.g. administration route, dosage regime, new acting principle) and *other advantages* (e.g. taste, appearance) are considered in about half of countries. However, in most cases they ranked as second line criteria for a drug's added value. The majority of countries perform a comparative pharmacoeconomic evaluation to analyse costs caused by a drug intervention in relation to its benefit (preferably by cost utility analysis). However, the question if a medicine is cost effective in comparison to treatment alternatives is answered in a political and societal context. A range of criteria are considered. They vary remarkably from country to country and include: availability of alternative therapies, social and ethic aspects like equity and solidarity, a drug's potential impact on public health, on organisations, and professionals, the community need and compliance with government-defined priorities, a medicine's estimated budget impact, and manufacturing research and development expenses.

In the case of therapeutic equivalence the drug is either not accepted for public reimbursement or it is subject to a reference pricing system (e.g. in New Zealand and in the Netherlands). In Australia and Canada equivalence is tolerated if a cost minimisation analysis demonstrates fewer costs.

However, there is international consensus that a drug therapy with a benefit-/risk-profile worse than other therapies (inferiority) should not be reimbursed, even in the case of fewer costs.

In Austria, in Australia, in Norway, and in the Netherlands drug evaluation is based on submissions by manufacturers. These comprise a systematic literature search and analysis of clinical and economic studies, in parts including modelling. Institutions have published guidelines to outline the methodological requirements for manufacturers and reviewers which, however, vary in the level of detail.

France, Switzerland, Canada (PMPRB), and Finland require that manufacturers submit a limited number of pivotal clinical trials. Since some details of the evaluation procedure are not made transparent it could not be clarified if these institutions additionally carry out a (systematic) research and analysis on their own.

Finally, in a third group of countries review bodies perform a systematic and comprehensive literature synthesis themselves and independently from manufacturers, including modelling. This is the case in Canada, in New Zealand, in the United Kingdom, and in Sweden.

Assessment of unpublished evidence is explicitly included in Austria, in Australia, in Canada, in the Netherlands, in Sweden, and in the United Kingdom. However, commercial in confidence data is excluded in parts (in Austria and in the United Kingdom).

Corresponding to the procedures' diversity, drug assessments last between a couple of weeks (e.g. in Australia and in France) and a year (e.g. in Sweden and in the United Kingdom). However, for European member states EC transparency directive 89/105/EEC determines that decisions on reimbursement and pricing of new pharmaceuticals are to be made within 180 days after marketing authorisation.

Countries agree that randomised controlled *head-to-head* trials (*head-to-head* RCT) with a high degree of internal and external validity provide the most reliable and least biased evidence for a drug's relative treatment effects (as do systematic reviews and meta-analyses of these RCT). Final outcome parameters reflecting the long-term treatment objective, i.e. final changes in health status caused by the therapy, are preferred (mortality, morbidity, quality of life). For study designs relying on surrogate parameters a strong and scientifically accepted association between intermediate effect and final outcome is required.

Following the concept of *community effectiveness*, nine out of twelve drug review institutions explicitly favour RCT in a "natural" design, i.e. in daily routine and country specific care settings. Concerning literature searches, selection of studies, assessment of internal and external validity of clinical trials and systematic reviews/meta-analyses, the review bodies apply internationally established standards (e.g. guidelines from Cochrane Collaboration, the CONSORT (Consolidation of Standards for Reporting Trials)- and the QUOROM (Quality of Reports of Meta-analyses of Randomised Controlled Trials)-Statement) or use comparable own standards.

The countries' requirements for pharmacoeconomic evaluations are similar. However they include some methodological problems. Cost utility analysis (CUA) measuring health outcomes in terms of Quality Adjusted Life Years (QALY) is most often recommended as appropriate analytic design. The rationale is that CUA can be applied for comparison of *different* health interventions and thus can be used for priority setting. It is considered to be associated with fewer methodological problems as cost-benefit analysis. There is also consensus that pharmacoeconomic studies should be, at least additionally, carried out from a societal perspective. According to the theory of social welfare, costs and benefits *outside* the health sector should be taken into account (direct and indirect non-medical costs).

However, there is no consensus how loss of productivity should be determined (human capital approach or friction cost method). There is also inconsistency if losses in spare time and in housework time are adequately included in changes of quality of life or if these should be additionally measured. Addressing the health economic controversy if costs of other diseases should be included if an intervention results in a prolonged life, guidelines recommend to include only costs that are directly associated with the intervention.

Utility measures should be based on preferences of a country's population. Since data is not always transferable

from one country to another institutions also request that type and quantity of resource consumption and prices should be based on national data. A high degree of transparency in cost calculation is required, i.e. identifying costs accurately, presenting quantity of resources consumed separately from the respective price and accomplished by the sources of data.

Discounting is recommended throughout the guidelines with an annual rate between 2.5% and 10%. From a societal perspective a discounting rate of 5% is typically recommended. Discounting of costs as well as of benefits at identical rates is suggested. As a rule, institutions require sensitivity analyses to determine effects of discounting on the study results by testing different rates and inclusion of benefit discounting.

Outcomes of clinical and pharmacoeconomic analyses are largely determined by the choice of comparator. Selecting an appropriate comparative treatment is therefore crucial. Concerning the institutions analysed in this study two procedures can be distinguished:

A. Some institutions require that a pharmaceutical should be compared with *up to three explicitly defined* comparators (in Australia, in Canada (CEDAC), in Finland, in New Zealand, in Sweden, and in the United Kingdom). In theory, the best or the most cost effective therapy is regarded as appropriate comparator. Pragmatically however, institutions accept if the drug is compared to daily treatment routine or the least expensive therapy. This includes other pharmaceuticals as well as non-pharmaceutical interventions and doing nothing. Concerning drugs the daily treatment routine should be identified by prescription or sales volumes. Dosage and treatment regime of medicines need to be therapeutically equivalent. Few institutions however, provide details how to determine the common practice of a non-pharmaceutical treatment: NICE in England/Wales and CEDAC in Canada specify that the choice of comparator should be based on clinical practice guidelines or on sales volumes. In Australia expert opinions are considered to select the appropriate comparator. Institutions demand that manufacturers and reviewers follow the guidelines as closely as possible when choosing a comparator. It is therefore recommended to contact the institutions early, i.e. in the phase of study design.

B. Other institutions (in Austria, in Canada (PMPRB), and in Switzerland) require that the pharmaceutical should be compared to *all medicines of the same therapeutic group* - based on the ATC (Anatomical-Therapeutic-Chemical)-classification of the World Health Organization. Thus, only currently reimbursed or marketed pharmaceuticals are considered as comparators. For each reference medicine an equivalent dosage form should be defined.

France combines both procedures: From all medicines in the same therapeutic group the most often prescribed, the least expensive and the most recently listed (positive list) are selected for comparison.

If a pharmaceutical provides several approved indications, review bodies in Canada (CEDAC), in Sweden, in the

Netherlands, in Norway, and in the United Kingdom commonly evaluate the drug for all of them. CEDAC in Canada is the only institution which also considers a drug's off-label use. In Austria, in Australia, and Canada (PMPRB) a medicine is only evaluated for its main or primary indication, i.e. either the indication (presumably) prescribed for the largest proportion of the population (in Austria and in Australia) or the indication offering the most therapeutic benefit compared to treatment alternatives.

Transferability of results derived from RCT to a country's daily care setting (external validity) is often limited. This can be due to several reasons: an experimental study design, i.e. artificial treatment setting; a time horizon (i.e. follow up) that is too short; a study population that is not representative for the general population; and inadequate choice of outcome parameter and comparator. Additionally, it is well known that clinical trials and pharmacoeconomic studies that directly compare a drug with adequate competitors are lacking, in quantitative as well as in qualitative terms. This is specifically the case before or shortly after marketing authorisation. Yet there is the need to support reimbursement or pricing decisions by scientific evidence. In this situation review bodies are often forced to rely on observational studies or other internally less valid data (including expert and consensus opinions). As a second option they use statistical approaches like indirect adjusted comparisons (in Australia and the United Kingdom) and, more commonly, economic modelling. Institutions have defined strict requirements for modelling and indirect comparison techniques. However, there is consensus that results provided by these methods need to be verified by valid head-to-head comparisons as soon as possible.

Thus many review bodies perform a re-evaluation after a certain period of time. Re-assessment is either done according to a fixed time schedule (e.g. in Finland and in France) or if new evidence has become available (e.g. in Austria and in the United Kingdom).

5. Conclusions

Most evidence concerning post-licensing evaluation of pharmaceuticals is derived from non-European countries like Australia and Canada. However, an increasing number of EU-countries also look back on several years' experience in comparative drug evaluation in connection with reimbursement or pricing decisions. There is consensus internationally that a new drug which is the first product offering an effective therapy for a particular indication should be classified as a major improvement. If a treatment alternative already exists, the majority of countries perform a systematic and evidence-based evaluation of a drug's clinical and economic characteristics in comparison to daily treatment routine. In the case of lacking or unreliable evidence results of a comparative drug assessment are considered as of preliminary nature. However, further evaluation criteria, requirements and specific

methodological issues still lack internationally consented standards.

Wissenschaftliche Kurzfassung

1. Gesundheitspolitischer Hintergrund

Durch das Gesetz zur Modernisierung der Gesetzlichen Krankenversicherung (GMG) wurde dem Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IGWiG) 2004 die Aufgabe der Nutzenbewertung von Arzneimitteln übertragen. Dazu beauftragt werden kann es u.a. vom Gemeinsamen Bundesausschuss (G-BA) und dem Bundesministerium für Gesundheit und Soziale Sicherung (BMGS; § 35 und § 139a Sozialgesetzbuch (SGB V)). In diesem Zusammenhang ist bedeutsam, dass (meist teure) Arzneimittel mit patentgeschützten Wirkstoffen nur dann von der Festbetragsregelung ausgenommen sind, wenn sie eine therapeutische Verbesserung bieten (§ 35 SGB V).

2. Forschungsfragen

Das Ziel in Deutschland ist, auf Basis der neuen gesetzlichen Grundlagen eine differenzierte Bewertung von Arzneimitteln im Vergleich zu bereits vorhandenen Therapiemöglichkeiten vorzunehmen, jedoch müssen Verfahren und Methoden noch etabliert werden. Im vorliegenden HTA-Bericht im Auftrag der sollte folgende Frage beantwortet werden: *Gibt es Methoden für eine Bewertung pharmazeutischer Produkte, die sich für vergleichende Untersuchungen eignen? Welche Methoden sind dies? Die Forschungsfragestellung des Berichts lautete: Wie werden in anderen Industrieländern Medikamente im Rahmen der Arzneimittelregulierung, d.h. bei der Marktzulassung, bei der Entscheidung zur Aufnahme in den Leistungskatalog und bei der Preisbildung verglichen?* Die vorliegende Studie analysierte komparative Arzneimittelbewertungsverfahren, -kriterien und -methoden bei öffentlichen Institutionen in anderen EU-/OECD-Ländern (EU = Europäische Union; OECD = Organisation for Economic Co-operation and Development).

3. Methodik

Nationale öffentliche Institutionen und dazugehörige Dokumente, die Verfahren einer vergleichenden Evaluation von Arzneimitteln in anderen EU-/OECD-Ländern beschreiben, wurden im ersten Schritt durch Internetrecherche ausfindig gemacht.

Zusätzlich wurde eine systematische Recherche in insgesamt 27 medizinischen, ökonomischen, sozialwissenschaftlichen und naturwissenschaftlichen Datenbanken durchgeführt. Bis Februar 2005 publizierte Literatur wurde mit Synonymen der Suchbegriffe "Arzneimittellevaluation" in Kombination mit Ländernamen sowie mit Namen der identifizierten Institutionen durchsucht. Nach der Entfernung von Duplikaten und einer zweistufigen Selektion anhand zuvor definierter Ein- und Ausschlusskriterien gingen von 1.619 Artikeln 85 Dokumente in die Studie ein.

Die Recherche wurde durch Handsuche in Literaturlisten und vereinzelt durch persönliche Kommunikation mit Institutionen ergänzt. Sämtliche Dokumente wurden qualitativ anhand eines zuvor entwickelten Rasters analysiert. Die Ergebnisse wurden in narrativer Form zusammengefasst und in Evidenztabellen präsentiert.

4. Ergebnisse und Diskussion

Bei der Marktzulassung von Medikamenten ist weder bei nationalen Behörden noch bei der europäischen Arzneimittelbehörde European Agency for the Evaluation of Medicinal Products (EMA) eine systematische Bewertung von Arzneimitteln im Vergleich zu therapeutischen Alternativen vorgesehen.

Im letzten Jahrzehnt haben daher viele Länder die Bewertung von Arzneimitteln nach der Marktzulassung zur Voraussetzung für Entscheidungen bei Kostenerstattung oder Preisbildung gemacht. Involviert in diesen Prozess ist eine ausdrücklich dafür ausgewiesene Institution mit intern oder extern beratenden Fachgremien.

In die Studie eingeschlossen wurden zwölf nationale, öffentliche Institutionen in elf EU-/OECD-Ländern: die Heilmittellevaluierungskommission (HEK) in Österreich, das Pharmaceutical Benefits Advisory Committee (PBAC) in Australien, das Patented Medicine Prices Review Board (PMPRB) und das Canadian Expert Drug Advisory Committee (CEDAC) in Kanada, die Eidgenössische Arzneimittelkommission (EAK) in der Schweiz, das Pharmaceuticals Pricing Board (PPB) in Finnland, die Commission de Transparence (CT) in Frankreich, die Kommission für Pharmazeutische Hilfe in den Niederlanden, die Norwegian Medicines Agency (NoMA) in Norwegen, die Pharmaceutical Management Agency (PHARMAC) in Neuseeland, das Pharmaceutical Benefits Board (PBB) in Schweden sowie das National Institute for Clinical Excellence (NICE) in Großbritannien.

In rund der Hälfte der Länder dienen die *Empfehlungen* der jeweiligen Institutionen zu Arzneimitteln als Grundlage für nachgeschaltete *politische Entscheidungen*, die abhängig von der Struktur des jeweiligen Gesundheitssystems in den nationalen bzw. regionalen Ministerien oder der Selbstverwaltung gefällt werden (Österreich, Australien, Kanada (Leistungskatalog), Frankreich, Niederlande, Norwegen). In den anderen Ländern entfalten die Entscheidungen der Institutionen *unmittelbare Rechtswirksamkeit*. Letztere Einrichtungen sind daher eigenständige Entscheidungsträger, auch wenn sie dem Gesundheitsministerium verantwortlich sind (in Finnland, Kanada (Preisbildung), Neuseeland, Schweden und Großbritannien).

Der Grad, mit dem die Arzneimittelbewertung in den analysierten Dokumenten detailliert beschrieben wird, variiert zwischen den Ländern erheblich. Am ausführlichsten und am ausdrücklichsten sind die Leitlinien aus Australien und Kanada, so dass dort das Vorgehen zumindest theoretisch von hoher Transparenz gekennzeichnet ist. Verallgemeinernd ist festzuhalten, dass Vorgehen und Methodik der gesundheitsökonomischen Evaluation

häufiger und ausführlicher in den analysierten Dokumenten beschrieben werden als der Bereich der klinischen Evaluation oder der Bewertung anderer Aspekte.

In allen Ländern ist der zusätzliche *klinisch-therapeutische Nutzen* im Vergleich zu Behandlungsalternativen das leitende Kriterium zur Bewertung eines Medikaments. Die Klassifikation des Arzneimittels als ein therapeutischer Fortschritt wird mit der Forderung verbunden, ein verbessertes Wirksamkeits-/Risikoprofil, d.h. Überlegenheit (Superiorität) in relevantem Ausmaß gegenüber therapeutischen Alternativen, zu belegen. Als das entscheidende Kriterium eines Mehrnutzens aus *Patientensicht* gilt eine verbesserte gesundheitsbezogene Lebensqualität, auf deren Einbeziehung sich die Institutionen in Australien, Neuseeland und Großbritannien festgelegt haben.

Pharmakologische Vorteile (z.B. Applikationsfrequenz, Darreichungsform, Galenik), *innovative Eigenschaften* (z.B. neues Wirkprinzip) oder *andere Merkmale* (z.B. Geschmack, Aussehen) werden in rund der Hälfte der Länder berücksichtigt, sind jedoch nach Wirksamkeit und Sicherheit nachrangige Kriterien.

Wird ein verbesserter Nutzen nachgewiesen, werden in allen Ländern außer in Frankreich und der Schweiz die durch die Arzneimittelbehandlung ausgelösten *Kosten* in Relation zum Mehrnutzen ermittelt (vorzugsweise mittels Kostennutzwertanalysen und Messung der gesundheitlichen Effekte in qualitätsadjustierten Lebensjahren). Die Frage eines angemessenen Verhältnisses wird jedoch politisch und im gesellschaftlichen Kontext beantwortet und eine Reihe - international sehr inhomogener - Kriterien werden berücksichtigt: Verfügbarkeit von Behandlungsalternativen, soziale und ethische Aspekte wie Verteilungsgerechtigkeit und Solidarität, Auswirkungen auf das Gesundheitsbudget, auf die Bevölkerungsgesundheit sowie auf Organisationen, Strukturen und Fachkräfte des Gesundheitswesens, Bedarf der Gesellschaft, Kosten der Hersteller für Forschung und Entwicklung sowie die Prioritäten der nationalen Regierung.

Therapeutische Gleichwertigkeit oder Non-Superiorität wird für einen Anspruch auf öffentliche Kostenübernahme nicht akzeptiert oder aber das Arzneimittel kann einer Festbetragsregelung unterworfen werden (z.B. in den Niederlanden, Neuseeland). In Australien und Kanada ist eine Äquivalenz nur unter der Prämisse geringerer Kosten, die über eine Kostenminimierungsanalyse ermittelt werden, akzeptabel. Ein schlechteres Wirksamkeits-Risikoprofil (Inferiorität) im Vergleich zu Alternativen wird dagegen international übereinstimmend selbst bei dem gegebenen Vorteil geringerer Kosten nicht toleriert.

In Australien, Österreich, Norwegen und den Niederlanden basiert die Evaluation in erster Linie auf Anträgen der Hersteller, die eine systematische Literatursuche und die Analyse klinischer und ökonomischer Studien z.T. mit Modellierungen umfassen. Die Institutionen haben dafür - allerdings unterschiedlich detailliert - ihre Anforderungen publiziert.

Frankreich, die Schweiz, Kanada (PMPRB) und Finnland fordern von den Unternehmen lediglich die Vorlage einer

meist definierten Anzahl von klinischen Schlüsselstudien. Da einige Details des Bewertungsverfahrens nicht ausreichend transparent gemacht sind, ist nicht ersichtlich, ob die Institutionen ergänzend eine eigene (systematische) Recherche und Analyse durchführen.

In einer dritten Kategorie von Ländern schließlich führen die Einrichtungen eine systematische und umfassende Literaturrecherche und Synthese einschließlich Modellierungen unabhängig von den Herstellern selbst durch, so in Kanada, Neuseeland, Großbritannien und Schweden. Unpublizierte Daten werden ausdrücklich in die Evidenzbewertung in Österreich, Australien, Kanada, den Niederlanden, Schweden und Großbritannien einbezogen, jedoch teilweise unter Ausschluss vertraulicher Informationen der pharmazeutischen Unternehmen (in Österreich, Großbritannien).

Entsprechend der Verschiedenheit der Verfahren ist der Zeitaufwand für die Evaluierung sehr unterschiedlich. Vom Beginn des Bewertungsverfahrens bis zur Empfehlung vergehen einige Wochen (z.B. in Australien, Frankreich) bis zu einem Jahr (z.B. in Schweden, Großbritannien). Für Mitgliedstaaten der EU gilt nach der EG-Transparenzrichtlinie 89/105/EWG jedoch, dass Entscheidungen zur gesetzlichen Erstattungsfähigkeit und Preisbildung von neuen Arzneimitteln binnen 180 Tagen nach der Marktzulassung gefällt werden müssen.

Als verlässlichste Evidenzgrundlage zu relativen Behandlungseffekten mit geringsten Verzerrungen gelten übereinstimmend randomisierte, kontrollierte *Head-To-Head*-Studien (*Head-To-Head*-RCT) mit hoher interner und externer Validität (und/oder systematische Übersichten und Metaanalysen dieser RCT). Bewertet werden sollte das beabsichtigte langfristige Behandlungsziel durch finale Ergebnisparameter, d.h. durch endgültige, durch die Therapie ausgelöste Veränderungen des Gesundheitszustands (Mortalität, Morbidität, Lebensqualität). Für Studienergebnisse, die auf Surrogatparametern beruhen, besteht die Forderung nach einer starken, wissenschaftlich etablierten Assoziation zwischen intermediärem Effekt und finalem Ergebnis.

Von den zwölf Arzneimittel evaluierenden Institutionen fordern neun ausdrücklich RCT in "natürlichem" Design, die nach dem Konzept von *Community Effectiveness* Aussagen zu Auswirkungen der Arzneimittelbehandlung unter den Bedingungen der landesspezifischen Versorgungspraxis machen. Bei der Literaturrecherche und -selektion, der Bewertung der internen und externen Validität von klinischen Studien sowie der Durchführung von systematischen Übersichten bzw. Metaanalysen folgen die Arzneimittel evaluierenden Institutionen - soweit spezifiziert - international etablierten Standards (z.B. den Leitlinien der Cochrane Collaboration, den CONSORT (Consolidation of Standards for Reporting Trials)- und QUOROM (Quality of Reports of Meta-analyses of Randomised Controlled Trials)-Statements) oder legen vergleichbare eigene Maßstäbe zugrunde.

Die Anforderungen an gesundheitsökonomische Studien decken sich in vielen Bereichen, beinhalten aber einige

methodische Inkonsistenzen. So ist die Kostennutzwertanalyse mit Messung der gesundheitlichen Outcomes durch qualitätsadjustierte Lebensjahre die am häufigsten empfohlene pharmakoökonomische Analyseform. Sie lässt den Vergleich zwischen verschiedenen Gesundheitsinterventionen und damit die Priorisierung von konkurrierenden therapeutischen Alternativen zu und ist im Vergleich zur Kostennutzenanalyse mit weniger methodischen Problemen behaftet. Auch besteht Konsens, dass pharmakoökonomische Studien wenn nicht primär, so zumindest zusätzlich aus gesamtgesellschaftlicher Perspektive vorzunehmen sind. Entsprechend der Wohlfahrts-theorie sind Kosten und Nutzen jenseits des Gesundheitswesens (direkte und indirekte nicht-medizinische Kosten) zu berücksichtigen. Es gibt wie in der wissenschaftlichen Literatur keine Übereinstimmung über die zu bevorzugende Methode zur Bestimmung von Produktivitätsverlusten (Humankapital- oder Friktionskostenansatz). Widersprüchlich sehen die Institutionen ebenfalls, ob Freizeit- und Hausarbeitszeitverluste sich erschöpfend in der Veränderung von Lebensqualität spiegeln, oder zusätzlich erfasst werden sollten. Auf die unter Gesundheitsökonomen ungelöste Kontroverse eingehend, ob Kosten anderer Erkrankungen zu berücksichtigen sind, wenn die Behandlung zu einer Lebensverlängerung führt, empfehlen die Leitlinien zum derzeitigen Zeitpunkt, nur Gesundheitskosten einzubeziehen, die in unmittelbarem Zusammenhang mit den zu begutachtenden Interventionen stehen.

Nutzwerte sollten basierend auf Präferenzen der gesamten Bevölkerung des jeweiligen Landes bestimmt werden. Aufgrund der eingeschränkten Übertragbarkeit von Kostendaten aus dem Ausland fordern die Institutionen ebenfalls durchgehend dazu auf, Art und Quantität der Ressourcenverbräuche sowie Preise möglichst auf nationale Gegebenheiten zu basieren. Durchgehend wird eine hohe Transparenz der Kostenkalkulation gefordert, d.h. eine detaillierte Identifikation der Kosten mit getrennter Darstellung von Mengenerfassung und Preisbewertung unter Angabe der jeweiligen Bezugsquellen sowie basierend auf möglichst aktuellen bzw. inflationsbereinigten Preisen. Die Vorgaben zur - grundsätzlich überall empfohlenen - jährlichen Diskontierung variieren mit Raten zwischen 2,5% bis 10%. Für eine Analyse aus gesellschaftlicher Perspektive wird häufig eine Diskontrate von 5% p.a. vorgeschlagen. Einhellig besteht die Forderung, sowohl Kosten als auch Nutzen, und zwar mit identischen Raten zu diskontieren. In der Regel fordern die Institutionen jedoch Sensitivitätsanalysen, um den Einfluss unterschiedlicher Diskontierungssätze und eine Diskontierung des Nutzens auf das Ergebnis zu beurteilen.

Die Definition und sorgfältige Auswahl von geeigneten Komparatoren ist in der therapeutischen wie gesundheitsökonomischen Analyse von großer Bedeutung für die Ergebnisse einer vergleichenden Evaluation. Unterschieden werden können zwei Verfahren:

A. Eine erste Gruppe von Institutionen fordert den Vergleich des zu bewertenden Arzneimittels mit *bis zu drei ausdrücklich definierten Komparatoren* (in Australien, Kanada (CEDAC), Finnland, Neuseeland, Schweden,

Großbritannien). Zum Vergleich geeignet gilt theoretisch zwar die beste bzw. kosteneffektivste Therapie. Als pragmatischen Ansatz akzeptieren die Institutionen jedoch regelhaft die derzeit übliche bzw. günstigste Behandlungspraxis, sei sie eine andere medikamentöse, eine nicht pharmakologische Intervention oder das Nichtstun. Bei Arzneimitteln wird die übliche Behandlungspraxis nach Verordnungshäufigkeit bzw. nach Marktanteilen festgelegt. Dosierungs- und Behandlungsregime von Referenzarzneimitteln und von begutachteten Arzneimitteln sollten therapeutisch äquivalent sein. Kaum eine Institution macht genauere Angaben, wie eine nicht-medikamentöse Routinepraxis zu definieren ist, auch wenn das NICE in Großbritannien und das CEDAC in Kanada spezifizieren, dies vorzugsweise auf Basis von Leitlinien oder anhand des Marktanteils zu bestimmen. In Australien können zur Festlegung des Komparators Expertenmeinungen eingeholt werden. Die Institutionen fordern von den Herstellern oder eigenen Gutachtern, sich bei der Wahl des Komparators so eng wie möglich an die Empfehlungen in den Leitlinien zu halten. Sie empfehlen die Auswahl des Komparators möglichst schon in der Phase des Studiendesigns in Rücksprache mit der Institution zu treffen.

B. Eine zweite Gruppe von Institutionen (in Österreich, Kanada (PMPRB), Schweiz) vergleicht das Arzneimittel mit *allen Medikamenten einer* mehr oder weniger eng gefassten *therapeutischen Gruppe*, orientiert an der zweiten bis vierten Ebene der anatomisch-therapeutisch-chemischen (ATC)-Klassifikation der WHO. Für jedes der Referenzmedikamente wird eine vergleichbare Darreichungsform und Dosierung definiert. Damit sind bei diesem Vorgehen ausschließlich bereits gelistete bzw. vermarktete Arzneimittel mögliche Komparatoren.

Frankreich kombiniert beide Verfahren, indem es aus Medikamenten einer therapeutischen Gruppe das am häufigsten verordnete, das günstigste und das zuletzt auf die Positivliste aufgenommene Produkt als Komparator heranzieht.

Ist ein Arzneimittel bei mehreren Indikationen zugelassen, evaluieren die Institutionen in Kanada (CEDAC), Schweden, den Niederlanden, Norwegen und in der Regel auch das NICE in England alle zugelassenen Indikationen. Lediglich in Kanada (CEDAC) gehen zusätzlich Anwendungsgebiete jenseits der Zulassung (*Off-Label-Use*) in die Bewertung ein. Bei anderen Institutionen (in Österreich, Australien, Kanada (PMPRB)) wird dagegen nur die Haupt- bzw. Primärindikation im Vergleich zu therapeutischen Alternativen geprüft, die entweder als die (voraussichtlich) bei der größten Patientenpopulation verordnete Indikation (Österreich, Australien) oder als diejenige Indikation mit dem größten therapeutischen Vorteil im Vergleich zu adäquaten therapeutischen Alternativen (Kanada (PMPRB)) definiert wird.

Die externe Validität von RCT, d.h. die Übertragbarkeit von Studienergebnissen auf die reale Versorgungspraxis im jeweiligen Land, ist oftmals eingeschränkt. Gründe sind neben einer artifizialen Behandlungssituation, zu

kurze Beobachtungszeiträume, nicht repräsentative Populationen sowie ungeeignete Outcomeparameter oder Komparatoren. Kostenkalkulationen in gesundheitsökonomischen Studien sind aufgrund der landes- und settingspezifischen Situation nur sehr begrenzt auf einen anderen Kontext übertragbar. Trotz dieses quantitativen oder qualitativen Mangels an direkt vergleichenden klinischen Anwendungsstudien (Phase IV) und an gesundheitsökonomischen Analysen - vor allem vor oder kurz nach der Marktzulassung des Produkts - müssen Gutachter Empfehlungen für gesundheitspolitische Entscheidungen aussprechen. Die Institutionen sind daher oftmals gezwungen, auf Beobachtungsstudien oder andere weniger intern valide Daten einschließlich Konsensus- und Expertenmeinungen auszuweichen. Außerdem verwenden oder fordern sie indirekte adjustierte Vergleiche (Australien, Großbritannien) und regelhaft umfassende gesundheitsökonomische Modellierungstechniken. Trotz strikter methodischer Anforderungen an die Qualität von Modellierungen und indirekten Vergleichen, gilt übereinstimmend, dass deren Ergebnisse frühest möglich durch intern und extern valide klinische und ökonomische Vergleichsstudien verifiziert werden sollten.

Daher führen viele Institutionen nach einer bestimmten Zeit eine erneute Bewertung des Arzneimittels durch. Dies erfolgt entweder in festen (z.B. Finnland, Frankreich) oder in variablen zeitlichen Abständen, letztere sind abhängig von der veränderten Evidenzlage bzw. bei erweitertem Anwendungsgebiet des Arzneimittels (z.B. in Österreich, Großbritannien).

5. Schlussfolgerungen

Die umfassendste Evidenz zur vergleichenden Evaluation von Arzneimitteln stammt aus nicht-europäischen Ländern wie Australien und Kanada. Mittlerweile blickt jedoch eine steigende Anzahl von EU-Ländern auf mehrjährige Erfahrung der vergleichenden Arzneimittellevaluation zurück. International besteht im Rahmen von gesundheitspolitischen Entscheidungen zum Leistungskatalog oder zur Preisbildung Konsens, dass ein Medikament, das als erstes Produkt eine effektive Therapie in einer bestimmten Indikation bietet, als wesentliche klinisch-therapeutische Verbesserung klassifiziert werden sollte. Wenn bereits eine Behandlungsalternative existiert, wird in der Mehrzahl der Länder eine systematische und evidenzbasierte Bewertung der klinischen und ökonomischen Charakteristika des Arzneimittels im Vergleich zur üblichen Behandlungspraxis vorgenommen. Bei mangelnder oder ungenügender Evidenz werden die Ergebnisse der Arzneimittellevaluation nur als vorläufig angesehen. Bei vielen anderen Bereichen des Verfahrens durch öffentliche Institutionen und bei methodischen Einzelaspekten steht ein international konsentierter Standard aus.

Korrespondenzadresse:

Annette Zentner
Fachgebiet Management im Gesundheitswesen, Institut
für Gesundheitswissenschaften, Fakultät VIII, Technische
Universität Berlin, Ernst Reuter Platz 7, 10587 Berlin
annette.zentner@tu-berlin.de

Bitte zitieren als

Zentner A, Velasco-Garrido M, Busse R. Methoden zur vergleichenden
Bewertung pharmazeutischer Produkte. *GMS Health Technol Assess.*
2005;1:Doc09.

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/hta/2005-1/hta000009.shtml>

Veröffentlicht: 15.11.2005

**Der vollständige HTA-Bericht in deutscher Sprache steht zum
kostenlosen Download zur Verfügung unter:**

http://portal.dimdi.de/de/hta/hta_berichte/hta122_bericht_de.pdf

Copyright

©2005 Zentner 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.