Euro Observer

The Health Policy Bulletin
of the European Observatory on Health Systems and Policies

Winter 2010

Volume 12, Number 4

Contents

The impact of health technology assessments: an international comparison

Panos Kanavos, Elena Nicod, Stacey van den Aardweg and Stephen Pomedli

With the rising costs of health care due to an ageing population and a growing number of new and expensive technologies, an increasing number of countries have implemented health technology assessments (HTAs) as a means of informing the decision process based on clinical and economic evidence. In an environment where resources are scarce, HTA agencies' objective is to ensure access to safe and effective medicines, while managing health care expenditure in an efficient way by reimbursing clinically cost-effective treatments. In this discourse, pharmaceutical products are the main – but by no means the only – subjects of such appraisals.

Different studies show that the impact of HTAs varies greatly across countries, even though they are assessing the same drug for the same indication.[1,2,3] These differences occur because of a number of considerations, such as the national priorities of the moment, the responsibilities and membership of HTA bodies, the differences in processes and timeframes, the implementation or not of the HTA recommendations, or even the ability to engage in price negotiation.[4,5]

In this issue of *Euro Observer* we undertake an analysis of health technology appraisals conducted across six agencies with a view to better understanding the similarities and differences in the appraisal process and the recommendations that follow. The agencies selected are the Common Drug Review (CDR) in Canada, the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia, the National Institute for Health and Clinical Excellence (NICE) in England, the Scottish Medicines Consortium (SMC) in Scotland, the Dental and Pharmaceutical Benefits Board (TLV) in Sweden and the High Health Authority (HAS) in France. In this comparison, we analyze the type of evidence considered, the timing of each appraisal and implications for pricing of the pharmaceutical treatments studied, as well as identify similarities and differences in terms of all appraisals conducted during 2007–2009 and the recommendations made. Finally, we showcase these in terms of general trends, but also by elaborating on specific disease areas (cancer, orphan and central nervous system drugs).

Methodology

Data on all HTA appraisals were collected between 2007–2009 from the six HTA agencies (Box 1). The data were collected from the public websites of each agency. Appraisals for all drugs and specific indications that were completed during the study period were considered, together with their corresponding recommendations, even if the latter were made outside the study period.

Analysis of these appraisals was conducted in two phases. The first phase examined all HTAs made by different agencies and aimed to identify general similarities and differences

This research was conducted by the Medical Technology Research Group at LSE Health, The London School of Economics under the direction of Panos Kanavos. No conflicts of interest arise.

The Observatory is a partnership between the WHO Regional Office for Europe, the Governments of Belgium, Finland, Ireland, the Netherlands, Norway, Slovenia, Spain, Sweden and the Veneto Region of Italy, the European Commission, the European Investment Bank, the World Bank, UNCAM (French National Union of Health Insurance Funds), the London School of Economics and Political Science and the London School of Hygiene & Tropical Medicine.

Box 1 Health Technology Assessment in six countries

In **Australia**, the Pharmaceutical Benefits Advisory Committee (PBAC) undertakes a centralized review process on innovative drugs and generics on an ad hoc basis. Final recommendations are considered by the Minister for Health and Ageing who is responsible for the final coverage decisions. Only drugs with a positive PBAC appraisal can be listed. The process takes about 17 weeks. 100+ appraisals are conducted annually.

In **Canada**, the Common Drug Review (CDR), undertaken by the Canadian Agency for Drugs and Technologies in Health (CADTH), operates a decentralized evaluation process. Only new chemical entities (NCEs) and new combination products (excluding anti-cancer drugs) are reviewed, on a first-come, first-served basis. Individual drug plans are not compelled to follow CDR recommendations; however, 90% of the time there is concordance. The process takes about 26 weeks. Roughly 25 appraisals are conducted annually. Since 2007, oncology drugs have not been assessed by CDR, but rather are appraised by individual provincial agencies. This process tends to be led by the CED in Ontario.

In **England**, the National Institute for Health and Clinical Excellence (NICE) provides a centralized HTA function for those technologies expected to have "major health implications, budgetary impact, or controversy over effectiveness".[1] Recommendations made by NICE are required to be implemented by the Primary Care Trusts (PCTs) within three months of being published (though in practice this may not be the case, and coverage administered at local level may be differential). Standard Multiple Technology Assessments (MTAs) take 52–62 weeks. Approximately 11 individual or class appraisals are conducted annually. A new Single Technology Assessment (STA) process was implemented in September 2006 to 'fast-track' appraisals for urgently needed drugs (examining only manufacturer submitted evidence) reducing assessment time to 39 weeks. Approximately 25 such appraisals have been completed, primarily on cancer drugs.

In **France**, the Haute Autorité de Santé (HAS) operates at the central level and provides coverage recommendations on every new drug, procedure and medical device. HAS stipulates a drug's SMR (Service medical rendu) level reflecting the medical and public health benefits and therapeutic value it provides. Based on this the Minister of Health decides whether or not to place the drug on the positive list for reimbursement. UNCAM (Union Nationale des Caisses d'Assurance Maladie) also uses this scale to define the drug reimbursement rate. Drugs are then evaluated against their comparators to assess the improvement in medical services rendered (Amélioration du service médical rendu, ASMR), used as the basis for price negotiations by the Comité Economique des Produits de Santé (CEPS). HAS does not conduct economic analysis but is primarily concerned with the drug's efficiency based on clinical (safety and efficacy) endpoints. It ranks treatments from I – V based on their (incremental) therapeutic benefit, with V indicating no improvement in therapeutic benefit and I indicating a breakthrough treatment.

In **Scotland**, the Scottish Medicines Consortium (SMC), reviews only new chemical entities (NCEs) and combination products to be marketed in Scotland. In contrast to NICE, SMC operates at arms-length, supplying advice only; listing decisions are undertaken by individual NHS Boards. Listing is not compulsory following a positive recommendation, except for "unique" drugs which are made available "uniformly" across Scotland. SMC endeavours to publish a recommendation within 12 weeks of products being made available.

In **Sweden**, the Dental and Pharmaceutical Benefits Board (TLV), has reviewed every new drug since 2002. Functioning as an arms-length agency, TLV's recommendations inform both the pricing and reimbursement decisions undertaken as part of the HTA process. Prices are not negotiated. Thereafter, only dugs approved by TLV are eligible for funding at county council level. By law, pricing and reimbursement recommendations are required to be provided within 120 days of receipt of a complete application.

in appraisals across agencies (see results below). The second phase focused on a subgroup of drugs with indications pertaining to diseases of the central nervous system (CNS), cancer and orphan drugs and conducted a more detailed comparative analysis to identify the rationale for decision-making and to understand the reasons for differences in appraisals (See case studies).

All HTA appraisal materials collected were compiled into a database; all drugs were classified according to their generic drug name, indication, and HTA outcome. For each HTA agency, HTA outcomes per drug and indication were divided into: (a) positive recommendations – List (L); (b) positive recommendations with restrictions – List with criteria (LWC); (c) negative recommendations – Do not List (DNL) (AMSR V for France); or (d) no appraisals completed by the agencies.*

All products were then categorized depending on their appraisal outcomes, into products that (a) have been accepted unconditionally as applied for (L), or with restrictions (LWC), (b) have been rejected by all agencies (DNL), and (c) have received different recommendations across agencies, where some have been rejected, while others have been accepted. The aim is to identify homogeneous and different recommendations across agencies.

For each drug, a number of endpoints, summarized in Table 1, were examined to identify similarities and differences across agencies in the evidence they request, and the type of evidence used to reach the final recommendation.

A detailed analysis of the above endpoints was conducted in three case studies, which are discussed in this issue. The objectives were to analyze the main reasons for recommendations and the concerns of the HTA agencies regarding the submissions, while assessing the clinical, safety and economic evidence submitted.

The types of clinical studies submitted (placebo-controlled, head-to-head comparisons, indirect comparisons, or other) were listed to examine the type of evidence requested by the respective agencies. All clinical, safety and economic evidence considered also was listed, so as to identify the type of evidence requested, as well as similarities and differences across agencies.

Results

Here we outline the results pertaining to the general analysis on all drugs appraised between 2007 and 2009. A significant number of appraisals (293) were completed by the six HTA agencies during this period. The evidence suggests that not all agencies appraised the same drugs, nor did they make the same recommendations for a drug if it had been subjected to an appraisal.

Common appraisals across agencies and homogeneity of outcomes

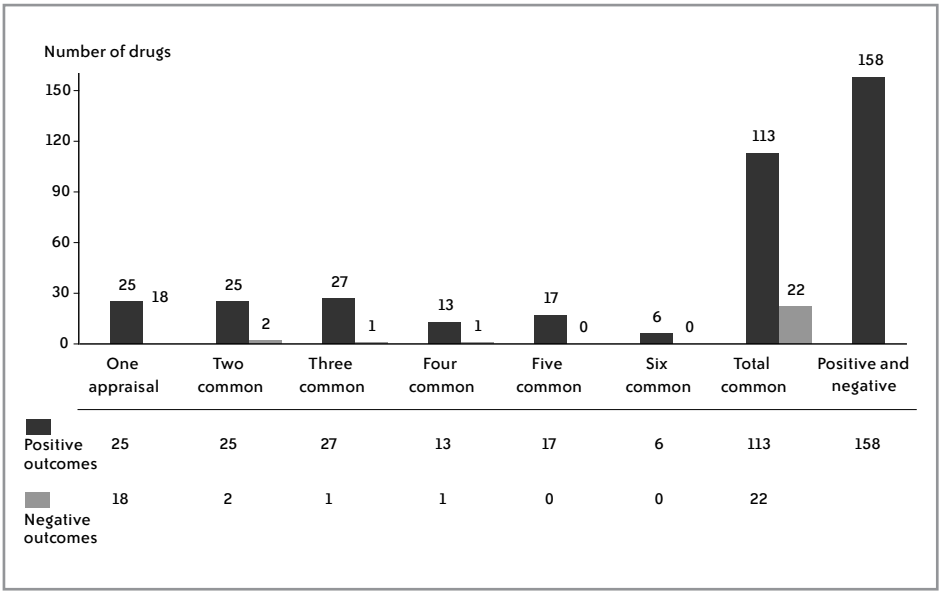
Only 7% of the drugs examined were ap-

* The recommendation system for HAS differs from the other agencies in that its system classifies drugs according to ASMR level, where in the absence of further data, ASMR V is considered as a negative appraisal in this analysis.

Table 1. Endpoints for analysis

Endpoints	Data collected/by agency	Description/reason
Indication/ ICD 10 code	ICD code	To determine whether the same drug has several indications, and whether they are common across countries.
Result or outcome	(a) accept as applied for (b) accept with criteria (c) accept/accept with criteria following rejection/resubmission (d) reject	To highlight the potential differences in outcomes across countries.
Rationale for outcomes	Primary reasons of recommendation	The main reasons used by each agency to make their recommendation.
Main studies cited	Types of studies conducted	Considered as the requested evidence by each agency.
Main comparators	(a) placebo, (b) therapeutic alternatives (current practice, best practice, routine therapy, etc.)	Considered as the requested evidence by each agency.
Clinical evidence	Listing of clinical evidence	To determine the type of clinical evidence requested for evaluation of effectiveness, and to assess differences across agencies.
Safety evidence	Listing of safety considerations	To determine the type of safety considerations requested for safety assessments, and assess differences across agencies.
Economic justification	Listing of economic justifications (i.e. None, ICER, cost/QALY, etc.).	To determine the type of economic justifications requested by each agency, and assess differences across agencies.

Figure 1. Number of appraisals and HTA outcomes



praised by all six agencies, 19% by five agencies, 18% by four agencies, 26% by three agencies, 17% by two agencies, and 13% by only one of the agencies. This shows that, surprizingly, HTA agencies do not necessarily appraise the same drugs, or the same drug for the same indication.

We can now ask ourselves whether these appraisals are homogeneous across agencies. Figure 1 shows the number of drugs with homogeneous outcomes (all positive or negative), or with different appraisals (positive and negative). Unexpectedly, 54% (N=158) of all drugs received a mix of positive and negative recommendations. Since 15% (N=25+18) of the drugs were appraised only once and thus cannot be accounted for as homogeneous appraisals, these 54% actually represent almost 2/3 of cases (against the remaining 31% with common appraisals).

Appraisals by agency and outcome

We looked at the number of drugs appraised by each agency, as well as their outcomes (Figure 2). In almost 95% of cases, TLV's recommendations were positive, followed by PBAC and NICE with 74% and 72% respectively. HAS and SMC, on the other hand, were slightly more strict with 64% and 68% positive appraisals, whereas CDR was by far the agency with the most negative recommendations, occurring in 48% of cases. In France 37% of all cases were considered to offer no improvement (ASMR V) and only 3% were recognized as providing a major innovation (ASMR I). The remaining 60% were classified with ASMR level II-IV.

Appraisals per ICD codes

Drugs were classified by ICD code (WHO, ICD10) and Figure 3 shows the number of drugs per ICD. The one class that sticks out by a substantially higher proportion in comparison to all others is cancer drugs (C00–D48). Additionally, the three other ICD classes with a high proportion of appraisals for each agency are infectious and parasite diseases (classes A00–B99), endocrine, nutritional and metabolic diseases (E00–E99) and muskuloskeletal system and connective tissue (M00–M99).

Figure 2. Appraisals by agency and outcome

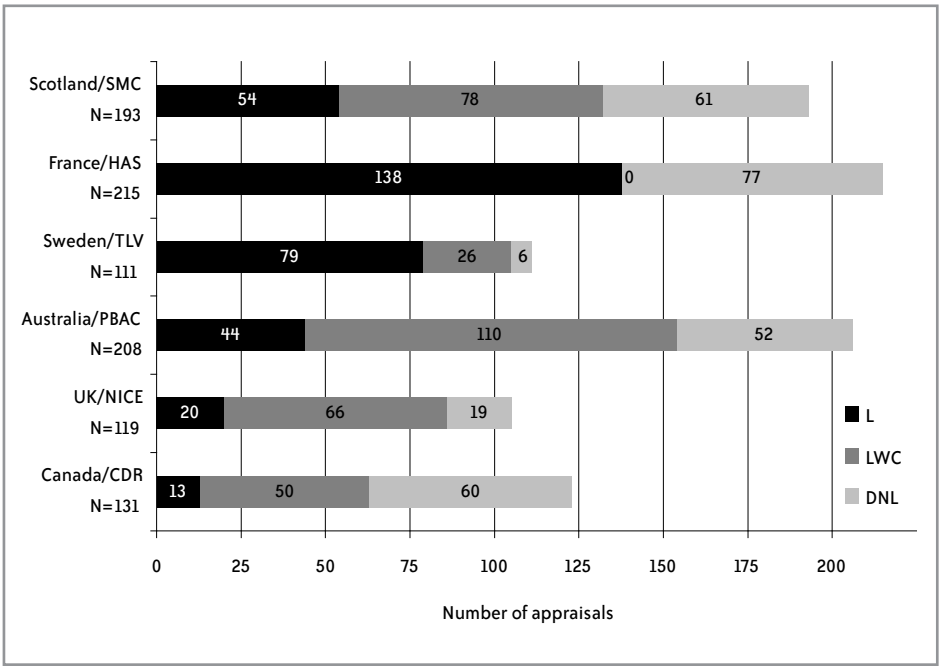
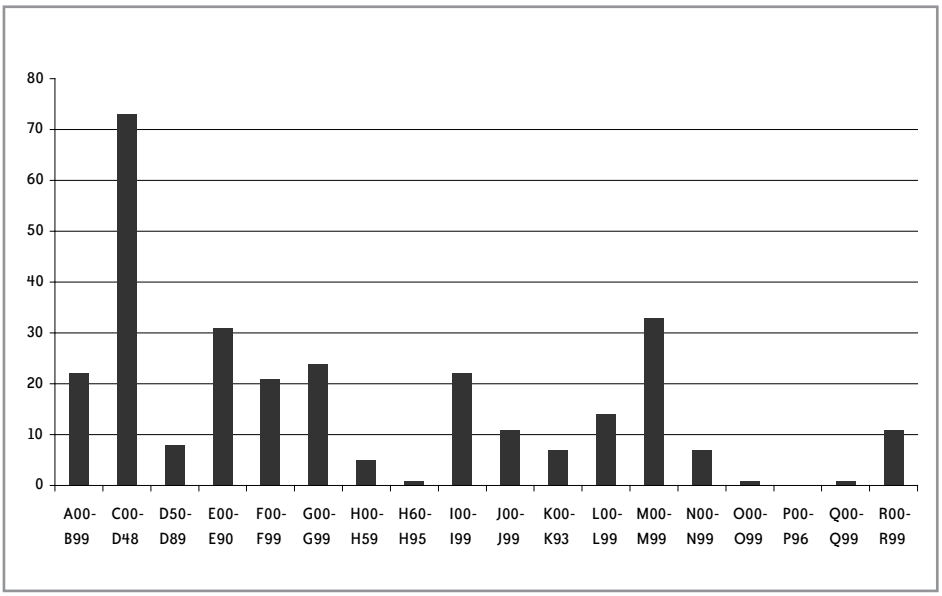


Figure 3. Number of HTA appraisals per ICD code



Selection criteria and national priorities

The number of appraisals completed by each agency varies, and the drugs assessed were not necessarily the same across agencies. In terms of the types of drugs classified per ICD code, we saw that certain classes were appraised more often than others. Moreover, when looking at the three classes with the highest proportion of appraisals per agency, these also differed from one agency to another.

These differences, in part, are a result of the selection criteria for HTA appraisals established by each agency, as well as each country's national priorities. Table 2 summarizes these selection criteria for each agency. For example, when considering that HAS appraises all drugs, the fact that they have appraised the highest number of drugs is more easily understandable. Similarly, NICE, having appraised the lowest number of drugs, focuses only on those which are deemed to fulfill the highest need.

National priorities may be reflected to some extent in the ICD codes of the appraised drugs. For example, in England the majority of cancer treatments (C00–D48) have been appraised as this is a priority because of the need, the severity of illnesses they treat and the cost implications for the National Health Service. In contrast, TLV, which selects all out-patient drugs for HTA appraisals, has a more balanced amount of appraisals across indications, which may suggest that all drugs are considered equally important, regardless of the price, need, or severity of disease.

An important finding is that HTA outcomes between agencies differed in more than half of the cases, whereby some agencies accepted these drugs in most cases, while others rejected them in almost 50% of cases. In the case studies that follow the authors focus on understanding why such differences occured, by elaborating on the type of decisions made in oncology, CNS, and orphan drugs.

Discussion

Although there is some crossover in terms of the clinical, safety and economic information considered by different HTA agencies, there are considerable disparities in the information required, interpretation of evidence, rigour of the appraisal process and stated motivations for listing or not listing drugs.

Clinical and economic evidence

A preference for robust Phase III trial data (particularly head-to-head trials where available) is visible across all agencies, where (possibly due to a scarcity of evidence) the same trials were examined by all. NICE and PBAC examined numerous additional Phase II, extension and open-label trials, while HAS focused, among others, on pharmacovigilance information, case studies and retrospective surveys. Where Phase III trials had been conducted, SMC rarely considered any additional clinical information. Only NICE explicitly considered clinical and patient expert opinions. TLV did not explicitly list any clinical studies examined.

Table 2 Selection criteria for HTA appraisals per agency

	Canada CDR/CED	England NICE	Australia PBAC	Sweden TLV	France HAS	Scotland SMC
All pharmaceuticals					x	
Highest need for guidance		x				
All out-patient pharmaceuticals				x		
All newly licensed drugs and formulations						x
New indications						x
Submission by manufacturer	x		x			
Submission by another entity	x					

Source: The author from the literature

Considering clinical endpoints – CDR, TLV and PBAC focused on general endpoints*, while NICE, HAS and SMC examined the full spectrum of primary, secondary and general endpoints. Primary trial endpoints were usually identified by all the agencies in some manner. All give some consideration to quality of life indicators such as SF-36 and 15D measures.

The economic dimensions of treatment are commonly assessed by looking at cost-effectiveness or budget implications. It appears that NICE accepts only cost-effectiveness models with QALY outcomes. There is a similar emphasis by PBAC, although alternative analyses and outcome measures were accepted in some cases. PBAC explicitly requires that incremental cost-effectiveness ratios (ICERs) are quality-adjusted and manufacturers' data was rejected when this was not the case. SMC received predominantly cost-utility analyses specifying QALY outcomes. There was little commonality in the few models detailed by CDR/CED. HAS focused on the assessment of clinical efficacy and conducted no economic evaluations. Commonly in the case of orphan drugs, high and uncertain ICERs are driven primarily by the high cost of treatment. However, in some cases this is exacerbated by limited evidence of additional efficacy, particularly over best supportive care.

HAS placed the strongest emphasis on drug safety and detailed every adverse event (AE) listed in the trial data, consistently identifying not only the most common, but also the most serious AEs. CDR, PBAC and SMC listed the majority of common AEs, while NICE provided a comprehensive list of AEs arising in common medical practice and those depicted by patient experience. TLV rarely listed specific AEs.

HTA timing and interpretation of data

In some instances, where the same Phase III data was considered key by different agencies, there was a marked difference in the resulting recommendations, suggesting that the interpretation of clinical data is not uniform across agencies. In the case of the orphan drug idursulfase, all appraising agencies focused on the same Phase III placebo-controlled random controlled trial (RCT). HAS concluded that, in the absence of alternative treatment, it demonstrated "significant superiority" compared to placebo on 6MWD** and all other secondary endpoints. In contrast, CDR indicated that the drug's clinical significance had not been established in relation to clinically relevant endpoints, yielding marginal improvements in 6MWD, with no improvements in others (for example, QoL, pain). PBAC concluded that while it presented improved survival outcomes, longer-term effectiveness and toxicity outcomes were not known. SMC concluded that the drug was "significantly more effective" than placebo, but rejected it based on insufficiently robust economic evidence.

There is little uniformity in the time taken by each agency to assess a drug subsequent to receiving marketing authorization (MA). Recommendations were commonly narrowed to a population sub-group within the broader MA indication, suggesting a discrepancy between the two processes, with MA requiring proof of quality, safety and efficacy only, while reimbursement decisions include broader, and often more subjective criteria.

Recommendations and thresholds

CDR/CED focused on evidence of clinical benefit, non-inferiority versus comparators (where this information is available) and value for money (usually proved through cost-effectiveness or comparative cost-analysis); the quality of trial data was also important. CDR recommended up to a threshold of £59 000 when justified or when high clinical need was demonstrated.

The key driver for NICE was cost-effectiveness, with cost implications frequently outweighing evident clinical benefit in instances where the ICER estimate lay outside the 'threshold' of £20 000–30 000 discussed in the literature.[6,7] Yet, closer examination of individual ICER estimates submitted to NICE suggests that this threshold may not be a rigidly adhered to for orphan and cancer treatments. In some cases, drugs with base case ICERs up to £59 000 per QALY were recommended even if they considered the drug to not be cost-effective, although this just suggests that, for some medicines, greater weight is placed on other factors (patient need, ethics and lack of alternative treatments).

PBAC recommendations were predominantly based on non-inferior efficacy and cost versus a comparator. The key driver

* i.e. endpoints that were considered, but not focused on or given additional weight to the extent that primary and secondary endpoints were considered key indicators of clinical efficacy by some agencies.

** Six-minute walking distance.

behind rejections was high and uncertain cost-effectiveness, with heavy criticism of economic models' design. Additionally, treatment safety and population need were regularly cited. Closer examination of ICER estimates indicates that PBAC was willing to support treatments up to an ICER of £45 000 (AUD75 000). TLV was relatively inconsistent in supplying recommendation criteria. A weak pattern indicates that clinical benefit, population need and the efficacy/safety ratio may hold greatest importance. An examination of cases in which ICER information was published suggests that TLV did not reject any of the orphans it considered on the basis of a high cost-effectiveness ratio; rather, the criteria for usage were restricted to reduce the budget impact. TLV seems to be more driven by need than cost and it accepted up to £75 000 when there is a high clinical need in certain sub-populations.

As mentioned before, HAS does not consider economic criteria; the full weight of its decisions are based on the drug's clinical benefit and efficacy/safety ratio, with a higher ASMR classification resulting from evidence of superior efficacy over comparators. SMC emphasizes the need for a demonstrated economic case for a drug. As such, model and clinical trial design are heavily scrutinized. The efficacy/safety ratio is frequently cited as an additional motivation for recommendation. The threshold value of SMC seems to be under £30 000, with rejections of higher values.

Given the scarcity of adequate clinical trial and cost data for some orphan drugs, agencies frequently restricted criteria for reimbursement to isolate patient subgroups in order to increase drug efficacy in these populations, reducing the cost-effectiveness outcomes to within acceptable levels (particularly NICE and PBAC).

Rigour of process

Across indications NICE, PBAC and HAS require the greatest amount of clinical evidence and most rigorously assess it. HAS and TLV strongly emphasize treatment safety and AEs. NICE conducts the most thorough cost-effectiveness examinations, assessing the manufacturer's model submission, frequently re-running the model with modifications and in every case building their own economic model. There does not appear to be a correlation between requirement stringency and the resulting recommendations, although the time-frame is undoubtedly positively related to the rigour of assessment.

Comparators

The choice of the appropriate comparator has shown to have an important effect on the HTA outcome. Most often the choice of comparator reflects agencies' requirements or preferences, which varies depending on the agency and may also help to understand differences in the outcomes achieved. The requirements for the number and type of comparators vary. Most include the current best alternative or relevant comparator, whereas CDR/CED and HAS also make requests for the cheapest available comparator, and TLV for a placebo comparison. All agencies request evidence on clinical efficacy and cost-effectiveness, except for HAS, that requires evidence on clinical efficacy and safety.

Efficacy and safety

With regard to the type of evidence considered, it appears that the agencies put different emphasis on different endpoints. In terms of efficacy, PBAC and TLV most often consider the primary endpoints, whereas SMC and HAS look at all primary and secondary endpoints. CDR and NICE have been shown to appraise the main endpoints while including an assessment of QoL. Thus, we may conclude that the value judgment of a drug's efficacy may vary according to whether it is based on one endpoint or several endpoints (i.e. if the effect on QoL is included in one assessment and not in the other, the judgment may be different).

Generally, safety is considered by all agencies but at different levels. HAS and SMC seem to put more weight on the drug's toxicity profile, and request the full list of most common adverse effects, whereas CDR, NICE and PBAC make a general safety assessment while highlighting the most relevant cases. TLV, on the other hand, most often makes a general assessment, and in some cases no mention of the drug's safety profile has been found. Although the drug's safety profile was only included by both HAS and SMC in the recommendation justifications, it does not seem to have an important impact on the end result.

Decision pricing and spillover effects

Recommendations by CDR/CED were usually associated with (upward) price volatility, both prior to publication and for several quarters thereafter. The trend effect of positive recommendations by both NICE and SMC was an immediate increase in price which moderated after 6–9 months. Prices in Sweden exhibited some volatility in either direction following a TLV recommendation, which is surprizing given that prices for reimbursement are fixed during the recommend- ation process. Positive and negative recommendations by HAS had little or no visible effect on French prices. There was too little evidence in the sample examined to indicate effects related to PBAC decisions.

Equity implications

Overall, this research suggests that in a majority of cases, recommendations are not homogeneous across agencies, and that the proportion of positive variations also varies greatly across agencies (i.e. TLV with the highest proportion, and CDR with the lowest). The proportion of resubmissions is fairly high, especially for SMC and TLV, with a significant proportion of appraisals being resubmissions. As a result, the cost of having to resubmit a reimbursement application, most often after a negative recommendation, can be very high and time consuming.

Given the evident disparity in time lapse between MA and HTA recommendation, the diverse criteria (and narrow subgroups) dictating reimbursement eligibility and inconsistencies in appraisal outcomes across countries, there is a strong indication that an international "post-code" lottery exists in terms of access to medicines. Not only does this have broad repercussions for cost, media attention and public opinion, it also highlights an area of ongoing debate regarding whether

citizens with conditions for which treatment is not reimbursed (or not yet assessed) in their home country should be refunded (by their national health system) for seeking care in other EU Member States, or in fact, seek treatment elsewhere, where it may be available.

References

1. Morgan SG, McMahon M, Mitton C, Roughead E, Kirk R, Kanavos P, Menon D. Centralized Drug Review Processes in Australia, Canada, New Zealand, and the United Kingdom. *Health Affairs* 2006;25(2):337 -47.

2. Drummond M. More haste, less speed? The emerging practice of HTA in the United Kingdom. *Euro Observer* 2009;11(1):9-10.

3. Cairns J. Providing guidance to the NHS: The Scottish Medicines Consortium and the National Institute of Clinical Excellence. *Health Policy* 2006;76, 134–43.

4. Sorenson C, Drummond M, Kanavos P. *Ensuring value for money in health care. The role of health technology assessment in the European Union*. World Health Organization, on behalf of the European Observatory on Health Systems and Policy, 2008.

5. Clement F, Harris A, Li JJ, Yong K, Lee K, Manns B. Using effectiveness and cost-effectiveness to make drug coverage decisions, a comparison of Britain, Australia and Canada. *JAMA* 2010;302(13): 1437–43.

6. Kanavos P, Manning J, Taylor D, Schurer W, Checchi K. *Implementing Value-based Pricing for Pharmaceuticals in the UK*, 2020 Health, Final Report, March 2010.

7. Rawlins M, Culyer A. National Institute for Clinical Excellence and its value judgments, *BMJ* 2004;329:224–27.

Panos Kanavos is Senior Lecturer in International Health Policy at the London School of Economics & Political Science and Programme Director of the Medical Technology Research Group at LSE Health.

HTA and access to cancer medicines

Stephen R. Pomedli

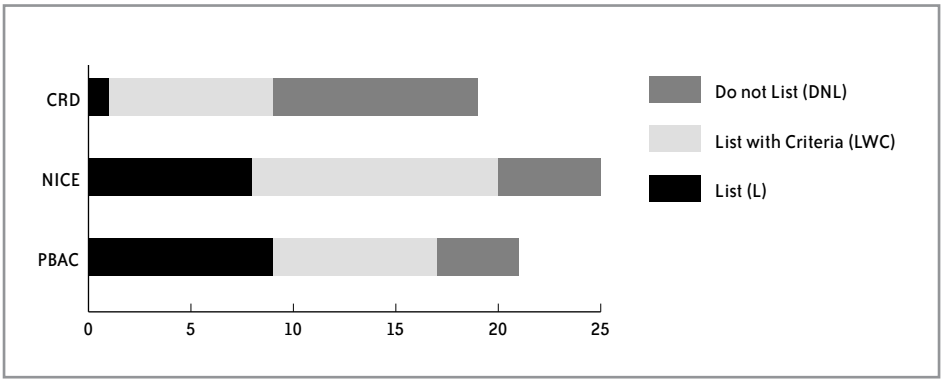
Anti-cancer pharmaceuticals historically have received much attention from HTA agencies,[1] primarily because of the high number of anti-cancer drugs obtaining marketing authorization, especially in recent years, the application of these drugs to novel indications, increasing clinical need, and the high costs associated with most of the newer anti-cancer compounds.[2,3,4] For this case study, HTAs related to cancer therapeutics (ICD-10 designator 'C') were compiled from three HTA agencies, namely, NICE (England), PBAC (Australia) and CDR/CED (Canada). Orphan drugs, even if given a 'C' ICD-10 designator, were excluded from this study due to the special considerations usually afforded them because of their unique status.*

Outcomes of HTA reviews

The methods described above yielded a total of 37 compound-indication pairings for anti-cancer, non-orphan drugs. This list of 37 pairings included 21 unique chemical entities. Of the 21, 11 were submitted for appraisal for more than one indication. The most frequent indications were for gastrointestinal cancers (nine) followed by breast cancers (seven), hematologic cancers (six) and cancers of the lung (five), skin (four), ovary and cervix, prostate, and kidney (two each).

Of the 37 compound-indication pairings, 51% were reviewed by CDR/CED, 68% by NICE, and 57% by PBAC. The majority of compound-indications were approved across agencies (45 of 65 final assessments or 71%). Of the reviews by the respective agencies, 47%, 80% and 81% of the final outcomes were considered positive**. Comparatively, CDR/CED were more likely to issue a negative HTA, and if recommending reimbursement, were more likely to require criteria for listing (LWC). Figure 1 illustrates the proportion of HTA outcomes by agency. Of the 21 pairings evaluated by at least two of the agencies, 14 had convergent outcomes, while 7 were divergent.

Figure 1. Number of HTA outcomes by type for Canada, England and Australia



*As defined by the European Medicines Agency or Australia's Therapeutic Goods Administration. Canada has no formal listing of orphan drugs.

** In the case of three compounds evaluated by CED, despite an initial negative appraisal and rejection on unacceptable ICERs, these products were eventually reimbursed after subsequent pricing agreements with the Executive Officer (List: bevacizumab, pemetrexed; LWC: sunitinib). If these were to be considered 'positive outcomes', the rate of positive outcomes from CDR/CED would rise from 47% to 63%, with the overall rate of positive assessments rising from 71% to 74%.

Evidence considered

Qualitative examination of the assessments suggested that the agencies tend to, on the whole, rely on the same studies and published evidence to provide information on efficacy and rates and types of adverse events. This is likely to be simply a by-product of the limited amount of high-quality evidence present in the published literature on which the manufacturer submissions rely due to the short time between marketing authorization and HTA in most cases. Further, the similarities in the evidence cited by the agencies are likely to be due to the close temporal proximity of the evaluations, precluding the development of substantial additional evidence between the assessment cycles. However, some of the differences in reimbursement decisions were the result of variations in interpretation of the same key trials, rather than reliance on different evidence per se. Among convergent and divergent outcomes, there were several main factors that influenced the reimbursement decisions of the various agencies.

Use of evidence

As expected, due to the lengthy time period required to assess long-term outcomes (especially relevant to evaluating survival benefits with cancer therapies) as well as to monitor adverse events, and collect complete economic data, in most cases adequate studies were not available in order to calculate robust estimates of cost-effectiveness, and resulted in high degrees of uncertainty in the incremental cost-effectiveness ratio (ICER). As a result, agencies were faced with a number of options: using sub-group analysis to limit reimbursement to narrower groups or indications for which the estimates were more certain; undertaking indirect comparisons across trials to supplement available evidence; employing expert opinion to reach a decision when the evidence was equivocal; deferral of appraisal decisions until further evidence is developed; or, rejecting reimbursement due to the high level of uncertainty. The agencies seemed to have differing degrees of willingness to employ these different strategies, which probably resulted in the differences in reimbursement outcomes. For example, in the case of erlotinib for non-small cell lung carcinoma, CDR/CED limited its use to patients with EGFR*-positive tumors or tumors of unknown status (LWC), based on putative improved response. In contrast, NICE considered the subgroup evidence too weak for this type of restriction (L), while PBAC conducted further analysis and concluded that EGFR status was in fact not predictive of response, but still restricted listing based on patient performance status (LWC).

Special considerations

Overall, the acceptability of the calculated ICER was a strong predictor of whether the compound was approved for reimbursement, and there were few cases in which a drug was approved for reimbursement despite a high, and otherwise unacceptable, ICER. These few decisions were made by NICE concerning the use of topotecan in small cell lung cancer, and in ovarian cancer. Nonetheless, many of the compound-indications were given special consideration in the HTAs due to the severity of disease, the relatively few people affected by the specific cancer, lack of other treatment options, or the relative therapeutic benefit that the drug offered. These special considerations appeared to contribute to reimbursement approvals for several compounds, notably sunitinib (to treat renal cell carcinoma: NICE, PBAC) and erlotinib (non-small cell lung cancer: NICE) due to lack of alternative effective therapies currently available; and for docetaxel (prostate cancer: PBAC, NICE) and trastuzumab (metatstatic breast cancer: NICE), for providing significant clinical gains, thereby implicitly rewarding innovation.

However, in other instances, even when such extenuating circumstances were considered, this did not always overcome significantly high ICERs, as pemetrexed (malignant pleural mesothelioma, PBAC) and cetuximab (metastatic colorectal cancer, PBAC) were initially rejected. It is unclear what exactly separates the compounds in that some, despite their excessive ICERs, receive positive HTA outcomes, while others do not, when additional criteria are taken into consideration. However, certain considerations may have more weight in this regard – for example, whether an indication has few as opposed to no effective treatments available.

Pricing

Through the HTA process, the agencies were noted to exert pressure on manufacturers to decrease the pricing of the compound in question in order to improve ICERs and increase the likelihood of reimbursement approval. Similarly, this pressure was also reflected in the development of various forms of risk-sharing agreements, in response to the lack of adequate evidence, requiring the manufacturer to bear a portion of the inherent risk when future costs were uncertain. Though the sample size here is small, for pricing negotiations the trends suggested that the Canadian process may lean towards putting pressure on the manufacturer to adopt price decreases in order to be approved for reimbursement, the English process favours development of risk-sharing agreements, and the Australian model variably applies both strategies.

Other factors

The results of an HTA assessment occasionally seemed contingent upon other factors external to the HTA process itself, such as current prices and market authorization of other comparators, release of new data, and changes in clinical practice. Similarly, while patients often have a consultative role during an HTA, patient preferences for certain aspects of the therapy (for example, method of administration, frequency of doses, or relative detriments of certain side-effect profiles) played a pivotal role in some instances. For example, in the case of docetaxel for prostate cancer (NICE), the assessment, while noting the occurrence of potentially debilitating side effects, acknowledged that patients generally considered these to be outweighed by the possible benefits of therapy.

Overall, the agencies tended to approach the use of less-than-ideal evidence with differing strategies: CDR/CED were likely to reject an application if

* Epidermal growth factor receptor, a biologic marker

inadequate evidence was submitted, but also engaged in pricing negotiations to reach positive outcomes; NICE tended to navigate uncertainty and poor evidence by using indirect comparisons and expert opinion as necessary, along with the development of risk-sharing agreements; while PBAC tended to encourage price negotiations and the development of risk-sharing agreements to overcome informational uncertainty.

Policy implications

This study suggests that poor evidence remains a critical obstacle in the HTA process. Across agencies, the evaluations analyzed here tend to rely on similar studies, but are usually limited by evidence that does not sufficiently address questions of clinical effectiveness, impact on quality-of-life, rates of adverse events or costs, relative to pertinent comparators.

Because of the need for better clinical and economic evidence, the HTA process would benefit from more transparent guidelines for manufacturers as to the types of data needed by HTA agencies to make rapid decisions, or by stipulating that certain data requirements must be available at the time of marketing authorization that fulfill these HTA needs. As this need seems to transcend the specifics of context or country, the formal development of standardized methodologies for HTA, international harmonization of data requirements for new therapeutics, and sharing of HTA expertise and results across countries would further develop the field, reduce duplicative effort in collecting and analyzing HTA-relevant data, and help address the data gaps that currently persist. Altogether, this would contribute to the rapid, safe and fair dissemination of cost-effective anti-cancer therapeutics.

References

1. Drummond MF, Mason AR. European perspectives on the costs and cost-effectiveness of cancer therapies. *Journal of Clinical Oncology* 2007;25:191–95.

2. Uyl-de Groot CA. Economic evaluation of cancer therapies: more and better studies will lead to better choices in cancer care. *European Journal of Cancer* 2006;42:2862-66.

3. Meropol NJ, Schulman KA. Cost of cancer care: Issues and implications. *Journal of Clinical Oncology* 2007;25:180–86.

4. Hillner BE, Smith TJ. Efficacy does not necessarily translate to cost effectiveness: a case study in the challenges associated with 21st century cancer drug pricing. *Journal of Clinical Oncology* 2009;27:2111–13.

Stephen R. Pomedli is a Research Assistant at the London School of Economics & Political Science and Medical Student at Harvard Medical School.

Harmonizing the outcomes of HTA: An application to Central Nervous System treatments

Elena Nicod

Health Technology Assessments (HTAs) have a substantial impact on which treatments are made available to patients.[1] Different studies suggest that the impact of HTAs varies greatly across countries, partly because of differences in the HTA appraisals.[2,3,4] This case study focuses on Central Nervous System (CNS) treatments and compares the HTA outcomes of these treatments across the six countries' HTA agencies (see Overview).

The Central Nervous System is one of two parts of the nervous system, together with the peripheral nervous system, and consists of the brain and the spinal cord.[5] The aim here is to identify similarities and differences of appraisals for CNS treatments completed between 2007–2009 across the study countries and to understand the rationale used by each agency to issue their recommendations by analyzing in depth the evidence that was considered.

Methodology

Data was collected from the 2007-2009 HTA reports available on the agencies' respective public websites. These reports include information on the evidence and main reasons for the recommendations issued. All drugs with indications F, G or R* pertaining to CNS treatments were selected. The analysis focused on treatments for Schizophrenia, Major Depressive Disorders (MDD), Parkinson's disease, Alzheimer 's disease, epilepsy, and chemo-induced nausea and vomiting.

HTA outcomes per agency, drug and indication were divided into (a) positive recommendation (List or 'L'), (b) positive recommendation with criteria (List with criteria or 'LWC'), and (c) negative recommendation (Do not list or 'DNL'"). In France, drugs are classified into five levels according to the drug's relative medical benefit (ASMR**) , where ASMR level V is considered as a negative recommendation in this analysis. To enable comparability, all drugs appraised by at least two HTA agencies were included, resulting in 25 unique molecule-indication pairs, of which four drugs had two indications each (see Table 1).

Results

Variations exist in the number of appraisals completed by each agency, in the drugs and indications that were appraised, and in the HTA outcomes. HAS and PBAC appraised almost all

* The indications by ICD10 codes included are classified as F, G or R.

** Amélioration du service médical rendu. ASMR levels = I major innovation; II important improvement; III significant improvement; IV minor improvement; V no improvement.

Table 1. Drugs by generic name, indication, ICD code, and appraisal outcome

				Appraisal Outcome					
	GENERIC NAME	INDICATION	ICD CODE	Canada CDR	England NICE	Australia PBAC	Sweden TLV	France HAS	Scotland SMC
1	Olanzapine paomate monohydrate	Schizophrenia	F20			LWC	L	V	
2	Aripiprazole	Schizophrenia	F20	DNL		DNL		III-IV	LWC
3	Paliperidone	Schizophrenia	F20	DNL		L	L	V	DNL
4	Ziprasidone hydrochloride	Schizophrenia	F20	LWC		LWC			
5	Escitalopram oxalate	Major depressive disorder	F33	DNL		LWC		IV	L
6	Duloxetine hydrochloride	Major depressive disorder	F33	DNL		LWC	LWC	V	LWC
7	Desvenlafaxine succinate	Major depressive disorder	F33	DNL		L			
8	Pramipexole hydrochloride	Parkinson's disease	G20			LWC		V	
9	Rotigotine, mono	Parkinson's disease	G20			DNL	LWC	V	L
10	Rotigotine, add-on	Parkinson's disease	G20			DNL	LWC	V	LWC
11	Levodopa / carbidopa	Parkinson's disease	G20	DNL		DNL	LWC	IV	DNL
12	Carbidpoa, levodopa, entacapone	Parkinson's disease	G20	L		LWC	L	V	L
13	Rasagiline Mesylate	Parkinson's disease	G20	DNL			LWC	V	DNL
14	Galantamine	Alzeihmer's disease	G30		LWC	DNL	L	V	
15	Memantine hydrochloride	Alzeihmer's disease	G30	DNL	LWC	LWC		IV	DNL
16	Rivastigmine, patch	Alzeihmer's disease	G30	DNL		LWC	L	IV	LWC
17	Rivastigmine, capsules	Alzeihmer's disease	G30		LWC			IV	
18	Interferon beta 1b	Multiple Sclerosis	G35		DNL	DNL		I-II	DNL
19	Natalizumab	Multiple Sclerosis	G35	LWC	L	LWC	L	III	LWC
20	Levetiracetam	Epilepsy, JME, PGTCS	G40			LWC		IV	LWC
21	Levetiracetam	Epilepsy, JME, PGTCS	G40			LWC		III	L
22	Zonisamide	Epilepsy	G40			LWC	LWC	V	LWC
23	Lacosamide	Epilepsy	G40			LWC	LWC	V	LWC
24	Aprepitant	Chemo-induced N/V, HEC	R11	L		L	L	III	LWC
25	Aprepitant	Chemo-induced N/V, MEC	R11	DNL		L	L	V	DNL

Table 2. HTA outcomes for Paliperidone and Rasagiline Mesylate across agencies

HTA outcomes		Canada CDR	United Kingdom NICE	Australia PBAC	Sweden TLV	France HAS	Scotland SMC	Restrictions
Paliperidone	L			X	X			
	LWC							
	DNL	X				X	X	
Rasagiline Mesylate	L							
	LWC				X			To patient sub-group for the treatment of on-off syndrome
	DNL	X				X	X	

drugs, while SMC, TLV and CDR appraised slightly more than half, and NICE only 5 of the 25 drugs. However, even though it did not appraise drugs for Schizophrenia and Parkinson's disease, NICE issued guidance for the treatment and management of these indications. NICE does not appraise all drugs, only those with the highest need for guidance,[6] which usually are innovative and costly treatments with significant potential impact on National Health Service resources. Because of this limited number of drugs appraised by NICE, and considering the ASMR ratings from HAS, where 78% received a very low ASMR rating (level IV or V) indicating marginal or no innovation, this category of drugs may be perceived as lacking innovative products, and, consequently, comprising largely me-too drugs*.

Of the 25 drugs and indications, only 24% had a homogeneous recommendation across all agencies (all positive or all negative), while 76% had a mix of positive and negative recommendations (Table 1). Thus, the number of drugs with divergent recommendations was substantial. Two case studies are used to understand why such differences occurred.

Focus on two drugs

An in-depth analysis of the type of evidence considered was conducted based on two case studies: an atypical antipsychotic for the treatment of Schizophrenia (Paliperidone), and a MAO-inhibitor for the treatment of Parkinson's disease (Rasagiline mesylate).

HTA recommendations

In both cases, HTA recommendations are not uniform (Table 2). Paliperidone was rejected by CDR, HAS and SMC because of insufficient evidence of a clinical advantage compared to other antipsychotics. In contrast, PBAC and TLV recommended paliperidone on the basis that it was deemed similar in terms of efficacy and price to olanzapine and risperidone respectively.

TLV accepted to list rasagiline because it was considered cost-effective in comparison with entecapone and tolcapone, but was restricted to the treatment of 'on-off' syndrome since its therapeutic alternative (selegiline) was deemed inappropriate for this sub-population. CDR and SMC rejected the application because no comparison with a less expensive therapeutic alternative was presented. Similarly, HAS's rating (ASMR V) was mainly due to the conclusion that the treatment did not have additional benefit compared to existing alternatives.

NICE did not appraise either drug, but instead issued guidance for the treatment and management of Schizophrenia (March 2009) and Parkinson's disease (June 2006). In these guidelines, all antipsychotics are considered to have equal effects, and similarly MAO-inhibitors are one of the accepted treatment options for early Parkinson's disease. In both cases, the choice of treatment depends on patient tolerability and preferences.

Decisional differences

In order to address the reasons for the variation in appraisal recommendations, the type of evidence considered in terms of comparators, clinical studies, and the drug's efficacy, safety and cost-effectiveness were examined.

Comparators and clinical trials
First, the trials considered and their designs vary, especially in the comparators used (Table 3). This is particularly true for paliperidone, where HAS considered placebo-controlled trials only, SMC and CDR considered placebo-controlled trials and an indirect comparison to quetiapine, while PBAC and TLV placed a greater emphasis on head-to-head comparisons. HAS, CDR and SMC rejected the treatment because of the lack of evidence demonstrating its clinical benefit in comparison to other therapeutic alternatives, mainly due to having used the placebo as a comparator.

Second, the number and type of studies also differ across the agencies. SMC, HAS and CDR seemed to place greater emphasis on longer study durations. This has been identified in the case of SMC,

* A me-too drug is a new molecule having a similar mechanism of action compared to existing drugs.

Table 3. Comparators and Clinical Trials

Comparators and Clinical Trials		Canada CDR	England NICE	Australia PBAC	Sweden TLV	France HAS	Scotland SMC	Comparator
Paliperidone	Placebo	x				x	x	
	Therapeutic comparator	x		x	x		x	
	Placebo-controlled	5				5	3	3 with olanzapine as active control arm
	Head-to-head comparisons			4	1			Olanzapine
	Indirect comparisons	1			1*		1	Quetiapine (* risperidone)
	Other			10				9 supportive trials and 1 pooled analysis of 3 RCTs
	Extensions					2	2	Placebo
Rasagiline Mesylate	Placebo	x			x	x	x	
	Therapeutic comparator	x			x	x	x	
	Placebo-controlled	3			3	3	4	1 RCT with a study arm with entecapone
	Head-to-head comparisons							
	Indirect comparisons							
	Other					1		Observational study
	Extensions						2	6-month extension and an open-label extension to one of the studies for patients treated with Rasagiline for up to 6.5 years, with additional therapy for Parkinson's Disease.

where the manufacturer submitted trial extensions for both paliperidone and rasagiline; on the other hand, HAS included the extension for paliperidone only. Both HAS and CDR expressed some concerns regarding the absence of long term data for rasagiline. PBAC placed greater demand on the volume of available evidence to arrive at a decision (four pivotal head-to-head trials, nine supportive trials and one pooled analysis of three random controlled trials (RCTs)), whereas this appeared to have been less relevant for TLV which considered only two trials for paliperidone.

Efficacy

With regard to the type of evidence considered, it appears that the agencies put different emphasis on different endpoints (Table 4). Efficacy for paliperidone was assessed by PBAC based on the primary endpoint, the PANSS score*, which determines the level of symptoms of schizophrenia. CDR considered three endpoints: the PANSS score, the personal and social performance scale (PSP), and schizophrenia quality of life scale (SQLS). HAS and SMC considered all primary and secondary endpoints, respectively 5 and 8 endpoints**. In contrast, TLV assessed the drug's general efficacy. The same tendency was observed for rasagiline as in paliperidone.

Safety

Safety is considered by all agencies, but only PBAC and HAS included the safety assessment as one of the reasons for the recommendation. Moreover, differences in the reporting of adverse effects were identified. The two case studies show that HAS and SMC report the most frequent adverse events, and CDR and PBAC the most relevant. In contrast, TLV made a general assessment of paliperidone's safety, stating that it was demonstrated. For rasagiline it is not clear in the appraisal report whether safety was considered, although TLV requested that the manufacturer demonstrates the safety of rasagiline in comparison to entecapone and selegiline.

Cost-effectiveness

Surprizing differences have also been identified in the type of economic model

* The Positive and Negative Syndrome Score.

** The additional endpoints considered were: treatment response rates, CGI-S variations, day/time drowsiness, quality of sleep, and the maintenance of efficacy.

Table 4. Type of evidence considered

Type of evidence considered			Canada CDR	England NICE	Australia PBAC	Sweden TLV	France HAS	Scotland SMC
Paliperidone	Efficacy	General assessment				x		
		Main endpoints	3		1			
		Primary and secondary endpoints					5	8
	Safety	General assessment	x		x	x	x	
		Detailed assessment			2		9	8
	Economical	Cost/patient/year			< $15 000			
		Cost/day	Paliperidone 3–12mg v. risperidone 4–10mg			Paliperidone all doses v. risperidone 5mg	N/A	
		Cost-utility analysis					N/A	v. three comparators, olanzapine as main comparator
Rasagiline Mesylate	Efficacy	General assessment	x			x		
		Main endpoints	4					
		Primary and secondary endpoints					7	6
	Safety	General assessment				?		
		Detailed assessment	5				18	3
	Economical	Cost/day	v. entecapone 10mg			v. entecapone and tolcapone	N/A	v. entecapone
		Cost-utility analysis					N/A	v. ripinerole

used, the choice of comparator and dosages. CDR and TLV both considered the cost per day of the treatment compared to other alternatives. For PBAC, the yearly cost per patient was estimated at less than AUS$15 000. No further details are given on how this estimation was arrived at. SMC, on the other hand, considered a cost-utility analysis using a different comparator (ripinerole) than the other agencies.

The number and comparators used varied between agencies. CDR used entecapone as a comparator for rasagiline, whereas TLV used entecapone and tolcapone, while SMC considered another comparator, ripinerole. This variance is due to differences in agency requirements; some request the cheaper alternative (i.e. CDR, HAS), while others the current best practice (i.e. CDR, NICE, PBAC, SMC), or the routinely used alternative (i.e. TLV).

Policy implications

In this case study we identify differences in HTA recommendations for twenty-five CNS treatments across six agencies. These differences may be caused not only by differing judgments from the expert committees appraising the drugs, but also due to the different permutations of evidence considered by the agencies in terms of study design, type of economic model and choice of comparator.

The low number of drugs in this category appraised by NICE, the low ASMR rating received by most drugs through HAS and the dominance of 'Do Not List' or 'List with Criteria' decisions are indicative of drugs being perceived by the payers as offering modest value added. Nevertheless, significant variation exists among agencies and the main causes identified include the choice and number of comparators, the (economic) models used, the different weights given to the clinical endpoints, and the variety of endpoints considered by the agencies. Consequently, there are many reasons for the variations in the recommendations made, even though the majority of evidence considered is based on robustly produced scientific data (mainly on RCTs).

It is likely that more homogeneous and standardized procedures across agencies (for example, data requirements pertain-

ing to comparators and equi-effective doses), may go some way to help manufacturers design clinical research that better fits the agencies' requirements, as well as minimize the cost of appraising treatments by each and every agency. A variation in the interpretation of evidence is unavoidable, even when the evidence is robust. Stakeholder involvement (i.e. patients, physicians) may play an important role to minimize these variations by providing complementary evidence on patient preferences or physician experience, although the perception is that such involvement is currently variable.

References

1. Sorenson C, Drummond M, Kanavos, P. *Ensuring Value for Money in Health Care. The Role of Health Technology Assessment in the European Union*. World Health Organization, on behalf of the European Observatory on Health Systems and Policy, 2008.

2. Morgan SG, McMahon M, Mitton C, Roughead E, Kirk R, Kanavos P, Menon D. Centralized drug review processes in Australia, Canada, New Zealand and the United Kingdom. *Health Affairs* 2006;25:2,337–47.

3. Drummond M. More haste, less speed? The emerging practice of HTA in the United Kingdom. *Euro Observer* 2009;11:1,9–10.

4. Cairns J. Providing guidance to the NHS: The Scottish Medicines Consortium and the National Institute of Clinical Excellence. *Health Policy* 2006;76:134–43.

5. MedTerms.com. Definition of Central Nervous System diseases. http://www.medterms.com/script/main/art.asp?articlekey=2667

6. Drummond M, Sorenson C. Nasty or Nice? A perspective on the use of health technology assessment in the United Kingdom. *International Society for Pharmacoeconomics and Outcomes Research (ISPOR)* 2009;12 supplement 2: S08-S13.

Elena Nicod is a Research Officer at LSE Health, the London School of Economics and Political Science.

HTA of orphan drugs across six countries: fair, flawed or failing?

Stacey van den Aardweg

The unique constraints in designing and undertaking robust clinical trials for orphan drugs present difficulties for HTA agencies in terms of the criteria and processes utilized in their appraisals.[1] Moreover, where orphan drugs are assessed using standard HTA processes that include economic evaluation they are unlikely to prove cost-effective and hence receive limited coverage.[2] This results in substantial uncertainty regarding the long-term value and level of innovation of novel technologies, duplication of assessment processes, inefficient use of national resources, and restricted and inequitable access to much needed new technologies.

Compounds assessed

In this case study, drugs with European Union (EU) orphan indications, based on the EMEA Register of designated Orphan Medicinal Products,[3] were identified and selected for study if they were appraised between January 2007 and December 2009. To enable comparison, only compounds that were assessed by at least two of the six agencies were included in the final detailed analysis, resulting in a set of 23 recommendations for different compound-indication pairs of orphan drugs, including 17 drugs with cancer-related indications (Table 1).

Results

Although there is some crossover in terms of the clinical, safety and economic information considered in the assessment of orphan drugs by different HTA agencies (Table 2), there are considerable disparities in the information required, interpretation of evidence, rigour of the appraisal process and stated motivations for listing or not listing drugs, in general and in orphan drugs in particular.

Evidence considered

The number and final outcomes of recommendations pertaining to treatment for rare diseases varied across countries. Cost-effectiveness is clearly a focus, with the exception of HAS, which does not examine economic evidence; however, other factors such as population need and adverse treatment effects are also significant. A preference for robust Phase III trial data (particularly head-to-head trials where available) is visible across all agencies; however, some agencies (particularly NICE and PBAC) examine additional clinical trial data.

There were instances in which the interpretation of a single study was inconsistent, resulting in disparate recommendations across agencies. For example for idursulfase, all appraising agencies focused on the same Phase III placebo-controlled Random Controlled Trial (RCT); while HAS concluded that, in the absence of alternative treatment, it demonstrated 'significant superiority' compared to placebo for primary and secondary endpoints, CDR indicated that the drug's clinical significance had not been established in relation to clinically relevant endpoints. Only NICE explicitly considers clinical and patient expert opinions.

NICE accepts only cost-effectiveness models with QALY outcomes. There is a similar emphasis by PBAC, although alternative analyses and outcome measures were accepted in some cases. HAS focused on the assessment of clinical efficacy, placing the strongest emphasis on drug safety, but conducted no economic evaluations. Commonly in the case of orphan drugs, high and uncertain incremental cost-effectiveness ratios (ICERs), are driven primarily by the high cost of treatment. However, in some cases this is exacerbated by limited evidence of

Table 1. Orphan drug appraisals – recommendations

Drug / Indication	ICD10	Appraisal outcome					
		Canada CDR/CED	England NICE	Australia PBAC	Sweden TLV	France HAS	Scotland SMC
Ambrisentan Pulmonary arterial hypertension	I27	LWC	-	LWC	LWC	DNL	LWC
Arsenic trioxide Acute promyelocytic leukemia	C92.4	-	-	LWC	-	LWC	-
Azacitidine Acute myeloid leukemia	C92	-	-	LWC	-	LWC	DNL
Azacitidine Myelodysplastic syndrome	D46	-	-	LWC	-	LWC	DNL
Dasatinib Acute lymphoblastic leukemia	C91.0	-	-	LWC	L	LWC	DNL
Dasatinib Chronic myeloid leukemia	C92.1	LWC	-	-	L	LWC	LWC
Eculizumab Paroxysmal nocturnal haemoglobinuria	D59.5	DNL	-	DNL	-	LWC	DNL
Idursulfase Hunter Syndrome (MPS II)	E76.1	DNL	-	DNL	-	LWC	DNL
Imatinib mesylate Acute lymphoblastic leukemia	C91.0	-	-	LWC	-	LWC	DNL
Imatinib mesylate Chronic eosinophilic leukaemia	D47.5	-	-	LWC	-	LWC	DNL
Imatinib mesylate Chronic myeloid leukaemia	C92.1	LWC	LWC	-	-	LWC	LWC
Imatinib mesylate Dermatofibros sarcoma protuberans (DFSP)	C49	-	-	LWC	-	LWC	DNL
Imatinib mesylate Gastrointestinal stromal tumour (GIST)	C26.9	case by case	LWC	-	L	LWC	DNL
Imatinib mesylate Myelodysplastic syndrome/ myeloproliferative disorders	D46	-	-	LWC	-	LWC	DNL
Lenalidomide Multiple myeloma	C90	-	LWC	LWC	LWC	LWC	DNL
Levodopa/carbidopa Parkinson's	G20	DNL	-	DNL	LWC	LWC	DNL
Nilotinib Chronic myeloid leukemia	C92.1	-	ongoing	LWC	L	LWC	LWC
Paclitaxel Ovarian cancer	C56	LWC	LWC	-	-	LWC	-
Sildenafil citrate Pulmonary arterial hypertension	I27	LWC	-	LWC	L	LWC	LWC
Sitaxentan sodium Pulmonary arterial hypertension	I27	DNL	-	LWC	L	DNL	LWC
Sorafenib tosylate Hepatocellular carcinoma	C22	LWC	DNL	LWC	case by case	LWC	DNL
Sorafenib tosylate Renal cell carcinoma (RCC)	C64	DNL	DNL	DNL	L	LWC	DNL
Temsirolimus Renal cell carcinoma (RCC)	C64	DNL	DNL	DNL	-	LWC	-

Key: L = List
LWC = List with criteria
DNL = Do not list

Table 2. Trends in evidence submitted and accepted by HTA agencies

HTA	Clinical evidence	Economic evaluation			Safety information
	Preferred trial data	Preferred economic model	Preferred ICER units	Budget impact considered	Emphasis on adverse effects
Canada CDR/CED	Phase III RCT	CEA (CMA, CUA)	LYG, LYS, QALY	No	Some
England NICE	All available evidence including: Phase III RCT (head to head where available) Phase II Clinical and patient expert opinion	CUA	QALY	Yes	Some
Australia PBAC	Phase III RCT (head to head where available) Open-label trials Comparator RCTs to enable indirect comparison Confidential efficacy data	CEA, CUA, CMA (CA, CC, DES)	QALY, LYG, ER, MG, CR, DA	Yes	Some
France HAS	Phase III RCT Pharmacovigilence information Observational studies	n/a	n/a	No	Strong
Sweden TLV	Trial data used rarely specified in pubic documentation	CMA (CEA, CUA, CA)	QALY	No	Weak
Scotland SMC	Phase III RCT	CUA (CEA, CMA, DES)	QALY, LYG	Yes	Some

additional efficacy, particularly over best supportive care.

Timing

There is little uniformity in the time taken by each agency to assess a drug subsequent to receiving marketing authorization. It is no surprise that HAS is observed as a 'leader' in terms of appraisals, since this process is pivotal to pricing and reimbursement decisions.[4] The sizeable lag between marketing authorization and HTA appraisal observed for CDR and PBAC is interesting given that, in principle, both agencies assess the majority of drugs that come onto the market.[5]

Basis of recommendations

The main reasons for individual agency recommendations of orphan drugs are summarized in Table 3. In the recommendations for orphan drugs provided, CDR/CED focused on evidence of clinical benefit, non-inferiority versus comparators and value for money. The key driver for NICE was cost-effectiveness. Contrary to indications that treatments with an ICER greater than £30 000/QALY carry a very high risk of being rejected by NICE,[6] there were instances in which base case ICERs up to £59 000/QALY were recommended even if they considered the drug to not be cost-effective, suggesting that, for orphans, greater weight is placed on other factors (patient need, efficacy, ethics and lack of alternative treatments).

PBAC recommendations were predominantly based on non-inferior efficacy and cost versus a comparator. The key driver behind rejections was high and uncertain cost-effectiveness, with heavy criticism of economic models' design. Closer examination of ICER estimates indicates that PBAC is willing to support treatments up to an ICER of c. £45 500. TLV is inconsistent in supplying recommendation criteria. A weak pattern indicates that clinical benefit, population need and the efficacy/safety ratio may hold greatest importance.

The full weight of HAS's decisions are based on the drug's clinical benefit and efficacy/safety ratio, with a higher Amélioration du service médical rendu (ASMR) classification resulting from evidence of superior efficacy over comparators. SMC emphasises the need for a demonstrated economic case for a drug; model and clinical trial design are heavily scrutinized. NICE, PBAC and HAS perform the most rigorous HTAs.

Due to the scarcity of adequate clinical trial and cost data for some orphan drugs, agencies frequently restricted criteria for reimbursement to isolate patient subgroups in order to increase drug efficacy in these populations, reducing the cost-effectiveness outcomes to within acceptable levels, whilst at the same time retaining the list price (particularly NICE and PBAC). For example, in the case of imatinib mesylate for Chronic myeloid leukaemia (CML), CRD/CED and NICE both narrowed their recommendations to a first line treatment for patients with Philadelphia-chromosome-positive CML in the chronic or accelerated phases or those with blast crisis, while HAS specifies support for newly diagnosed patients only.

Policy implications

Given the evident disparity in time lapse between market authorization and HTA recommendation, the diverse criteria (and

Table 3 Main criteria on which recommendations are based

		Canada CDR/CED	England NICE	Australia PBAC	France HAS	Sweden TLV	Scotland SMC
Clinical efficacy	Non-inferiority/superiority	Non-inferiority v. placebo		Non-inferiority v. comparator	Superiority v. therapeutic comparators	Non-inferiority to comparator	Superiority/ non-inferiority v. comparators
	Clinical benefit	X	X	X	X	X	X
	Strength of trial design	X	X	X			
Economic evaluation	Cost-effectiveness (low, certain)	X	X	X		X	X
	Cost vs comparators	X		X		X	
	Economic model validity (inputs, methods)	X	X	X		X	X
	Value for money	X					
	Budget impact		X	X			
Need	Rx alternatives available/not	X		X	X	X	
	Population medical need		X	X		X	
Safety	Toxicity/safety profile			X			X
	Efficacy/safety ratio (high)				X	X	X

narrow sub-groups) dictating reimbursement eligibility and inconsistencies in appraisal outcomes across countries, there is a strong indication that a form of international 'postcode lottery' exists in terms of access to medicines. Conducting HTAs earlier and more rapidly, and merging the requirements for marketing authorization and HTA into trial design could help to optimize resource use, maximise health benefits and enhance access to and availability of orphan drugs for needy populations.

By removing budget considerations from decisions through the provision of earmarked or central funding for specific rare diseases and creating a separate 'orphan drug' protocol policy-makers can ensure objective, accurate and timely HTA of orphan drugs. Manufacturers can maximize their effectiveness and increase the probability of an orphan drug receiving a positive recommendation by designing trials to provide more comparative data and structuring economic models from both a health and societal perspective, applying the agency preferred methods for discounting and quality-adjusting utility values. From their perspective, HTA agencies need to recognize that robust modelling and adequate power – however desirable – may not always be possible in the case of orphan drugs.

Streamlining and standardizing HTA processes, increasing international HTA collaboration and improving communication between HTA agencies and manufacturers will undoubtedly increase the efficiency and effectiveness of orphan drug appraisal procedures. Ensuring that recommendations can be implemented optimally and uniformly, to enhance uptake and ensure geographical equity in access requires additional work and attention by decision makers.

References

1. McCabe C, Claxton K, Tsuchiya A. Orphan drugs and the NHS: should we value rarity? *British Medical Journal* 2005;331:1016–19.

2. Drummond M, Wilson D, Kanavos P, Ubel and Rovira J. Assessing the economic challenges posed by orphan drugs. *International Journal of Technology Assessment in Health Care* 2007;23(1):36–42.

3. European Medicines Agency (EMEA). Register of designated Orphan Medicinal Products [Online] (Updated: 2010) Available at: http://ec.europa.eu/health/documents/community-register/html/orphreg.htm (Accessed 5 August 2010).

4. Chevreul K, Durand-Zaleski I. HTA in coverage and reimbursement decisions in France: toward a new paradigm? *Euro Observer* 2009;11(1):5–6.

5. Menon D, Stafinski. Health Technology Assessment in Canada: 20 years strong? *Value in Health* 2009;12 (Supplement 2):14–19.

6. Kanavos P, Manning J, Taylor D, Schurer W, Checchi K. *Implementing Value-based Pricing for Pharmaceuticals in the UK*, 2020 Health, Final Report, March 2010.

Stacey van den Aardweg is a Research Assistant at LSE Health, the London School of Economics & Political Science.

Observatory–Health Evidence Network Joint Policy Briefs

Published for the Belgian European Union Presidency Ministerial Conference on the European Health Workforce

Two joint HEN-Observatory policy briefs and two policy summaries were prepared at the invitation of the Belgian Federal Public Service–Health, Food Chain Safety and the Environment, for the Belgian EU Presidency Ministerial Conference on 'Investing in Europe's health workforce of tomorrow: scope for innovation and collaboration' (La Hulpe, 9–10 September 2010). They reflect key priority areas for European policy/decision-makers in respect of future health workforce needs, and where learning from comparative experience is crucial to informing future policy choices.

How to create conditions for adapting physicians' skills to new needs and lifelong learning

HEN-OBS joint policy brief No.14 (2010)

Policy issue and context

The knowledge and skills acquired at the end of formal undergraduate and post-graduate professional medical education are insufficient to sustain competence and performance over a career, thus physicians are expected to effectively engage in lifelong learning strategies.

There is increasing scrutiny of professional and public concerns related to the variability in the quality of care provided, the safety of the health system, and the frequency of adverse events.

Within Europe there is currently no commonly accepted approach to lifelong learning. However, there is broad agreement that patients are best served when those who care for them maintain competence by engaging in continuous learning and assessment strategies. There are currently no standards governing the following lifelong learning strategies: the organization and management of activities; incentive structures for participation; classification systems for activities or credits; accreditation standards; physician discretion regarding choice of learning activities; accreditation ex ante for providers; and industry sponsorship.

Policy options

Discordance between the expectations of patients and the abilities of physicians are prompting the profession to strengthen assertions of "professionalism". To increase accountability, compulsory engagement in continuing professional development (CPD) systems or programmes can be considered.

It will be important to both enhance and ensure the quality and rigour of the providers or programmes that physicians depend on to develop and implement a practice-specific, needs-based learning plan. The development of a common CPD accreditation system for providers and programmes is deemed essential.

To address barriers within the health care system and to optimize the benefits of lifelong learning for patient care and outcomes, physicians, providers of CPD, and the health care system itself need to take a "shared responsibility" approach to lifelong learning and

Implementation considerations

If the goal of CPD systems is to improve the delivery of good-quality patient care and thus improve patient outcomes, the environment in which physicians practice should be both supportive and constructed in a way that promotes and enhances learning.

In the EU, the diversity of CPD systems is increasingly becoming a barrier to those in pursuit of harmonization of CPD across Member States. In order to build equivalent and successful national CPD systems, infrastructure considerations must include the following: mutual agreement and recognition of CPD; uniformity of accreditation standards; efficient and accessible delivery mechanisms for CPD; equivalent standards for industry sponsorship allowances; and performance-assessment metrics.

How to create an attractive and supportive working environment for health professionals

HEN-OBS joint policy brief No.15 (2010)

Policy issue and context

Poor work environments compromise health workforce supply and quality of care. Health policy-makers face the challenge of matching increasing demand for health care with a supply of health professionals in times of existing and projected health-workforce shortages.

The work environment constitutes an important factor in the recruitment and retention of health professionals, and the characteristics of the work environment affect the quality of care both directly and indirectly. Addressing the work environment, therefore, plays a critical role in ensuring both the supply of a health workforce and the enhancement, effectiveness and motivation of that workforce.

The purpose of providing attractive and supportive work environments is to create incentives for entering – and remaining in – the health professions, and to provide conditions that enable health workers to perform effectively (to achieve high-quality health services).

Policy options

Given the complexity of the work-environment issues to be addressed, policy responses need to be multidimensional, cross-cutting and inclusive. For coherent policies, action has to be considered at four levels: international/regional level; national level; sectoral level; and local/organizational level. Effective solutions are context-related and therefore priority has to be given to the local and organizational level. The other levels provide the legislative and regulatory framework and

All are available at: www.euro.who.int/en/what-we-do/data-and-evidence/health-evidence-network-hen/publications/joint-policy-briefs-and-policy-summaries/published-for-the-belgian-european-union-presidency-ministerial-conference-on-the-european-health-workforce

provide guidance and support for the development of workplace policies.

Two examples of what can be done to improve the quality of the work environment in the health professions include policy approaches to promote a healthy balance between family life and work, and the enhancement of the protection of workers' health.

In order to encourage health-sector employers to make a commitment to positive work environments, the development of workplace assessment/recognition programmes could be considered.

Implementation considerations

As many factors influencing the work environment operate outwith the health sector, intersectoral collaboration is required. In particular, the interface between labour and health-policy mandates needs to be strengthened. Here, the use of social dialogue can help to ensure sustainable and cross-sectional implementation with multiple stakeholders.

Assessing future health workforce needs

Policy summary 2 (2010)

The choice of a strategy to assess the future health workforce (HW) is value-based and depends on what health outcomes and service objectives policymakers have set. Various models, approaches and toolkits have been proposed and tried over the years by international agencies, as well as by individual countries. Workforce situation analysis for determining future staff requirements typically builds upon variables such as expected population growth, technological and social change, skills mix, individual performance and health policy.

There is little benefit in educating adequate numbers of doctors or nurses, and then seeing them migrate to other countries because the labour market cannot integrate them, or because working conditions are not attractive enough.

Assessing future HW needs is not only about projecting the numbers. Policymakers need also to address the issues of recruiting, educating, distributing, retaining, motivating and managing the HW, which implies improving the knowledge about the expectations and behaviours of health workers. Addressing needs implies more than producing more workers; scaling up can be achieved by improving competences, changing skills mix, and by augmenting productivity.

It is important to see HW planning as a process that engages the main stakeholders in assessing needs for change and in devising strategies to achieve those changes. The better the information base and the technical capacity to use it, the better the diagnosis and the selection of interventions will be. Monitoring is essential to adjust interventions to a changing environment.

Sufficient and predictable funding must be available to invest in workforce development. The benefits will soon be apparent in terms of better access to services, more efficient utilization of resources and higher satisfaction of citizens.

Using audit and feedback to health professionals to improve the quality and safety of health care

Policy summary 3 (2010)

There is now extensive evidence demonstrating that there is a gap between the health care that patients receive and the practice that is recommended. In both primary and secondary care there are unwarranted variations in practice and in resulting outcomes that cannot be explained by the characteristics of patients.

While it is difficult to find examples of measures for addressing this issue from all 53 countries of the World Health Organization's European Region, there are interventions that can be identified in the 27 Member States of the European Union. However, the nature of these measures and the extent to which they are implemented vary considerably.

Audit and feedback defined as "any summary of clinical performance of health care over a specified period of time aimed at providing information to health professionals to allow them to assess and adjust their performance" is an overarching term used to describe some of the measures that are used to improve professional practice. It can be used in all health care settings, involving all health professionals, either as individual professions or in multiprofessional teams. In practical terms, health professionals can receive feedback on their performance based on data derived from their routine practice. Health professionals involved in audit and feedback may work either in a team or individually and in primary, secondary or tertiary care.

While it seems intuitive that health care professionals would be prompted to modify their clinical practice if receiving feedback that it was inconsistent with that of their peers or accepted guidelines, this is in fact not always the case.

The available evidence suggests that audit and feedback may be effective in improving professional practice but that the effects are generally small to moderate. Nonetheless, depending on the context, such small effects, particularly if shown to be cost-effective, may still be regarded as worthwhile. The benefits are most likely to occur where existing practice is furthest away from what is desired, and when feedback is more intensive.

Even on the basis of the best evidence available, no strong recommendations can be given regarding the best way to introduce audit and feedback into routine practice. However, decisions about if, and how, this approach can be used to improve professional practice must be guided by pragmatism and the consideration of local circumstances. The following scenarios, for example, might indicate suitability for such an approach: the known (or anticipated) level of initial adherence to guidelines or desired practice is low; it is feasible to conduct an audit and the associated costs of collecting the data are low; routinely collected data are reliable and appropriate for use in an audit; and small to moderate improvements in quality would be worthwhile.

The cost of audit and feedback is highly variable and is determined by local conditions, including the availability of reliable routinely collected data and personnel costs. Its impact, with or without additional interventions, should be monitored routinely by auditing practice after the intervention.

Volume 12, Number 4

HiTs to be included in MEDLINE

The Observatory is delighted to announce that its Health System in Transition (HiT) country profiles will now be included in MEDLINE.

MEDLINE is the US National Library of Medicine's premier bibliographic database covering the fields of medicine, nursing, dentistry, veterinary medicine, the health care system, and the pre-clinical sciences. The database contains more than 20 million records from approximately 5 000 selected publications and more than 80 countries.

The decision to include HiT reports follows a rigorous selection process by a panel of scientific experts which assessed not only the quality of the HiTs' content and methodology but also their importance to research and policy communities.

Using MEDLINE, researchers and policy-makers from all over the world will now be able to find HiT reports much more easily. This will increase dissemination and ensure health system information is available to all those who need and want it most. It will also reinforce the Observatory's commitment to supporting and promoting evidence-based policy-making in health.

Download Hit country profiles at: www.euro.who.int/en/home/projects/observatory/publications/health-system-profiles-hits

Observatory/BMJ collaboration

Recent Observatory work is being published in the *British Medical Journal* under a new occasional series called Health Systems Perspectives. The series focuses specifically on developments in the European Union and the EU health policy agenda with two articles already published on the antibiotic pipeline and user charges.

Forthcoming HiTs

New HiTs for Spain, Greece and France are due to be published shortly.

To join the mailing list, please contact
Anna Maresso
Observatory – London Hub
Email: a.maresso@lse.ac.uk

Euro Observer is published quarterly by the European Observatory on Health Systems and Policies, with major funding provided by a grant from Merck & Co., Inc., Whitehouse Station, New Jersey, USA.

For information and ordering details on any of the Observatory publications mentioned in this issue, please contact:

The European Observatory on Health Systems and Policies
WHO ECHP
Rue de l'Autonomie, 4
B - 1070 Brussels, Belgium
Tel: +32 2 525 09 33
Fax: +32 2 525 0936
Email: info@obs.euro.who.int
www.euro.who.int/observatory

Design and production by Westminster European

ISSN: 1020-7481

THE EUROPEAN OBSERVATORY ON HEALTH SYSTEMS AND POLICIES PARTNERS

- World Health Organization Regional Office for Europe
- Government of Belgium
- Government of Finland
- Government of Ireland
- Government of the Netherlands
- Government of Norway
- Government of Slovenia
- Government of Spain
- Government of Sweden
- Veneto Region of Italy
- European Commission
- European Investment Bank
- World Bank
- UNCAM
- London School of Economics and Political Science
- London School of Hygiene & Tropical Medicine