Karin H. Cerri, Martin Knapp, Jose-Luis Fernandez

Untangling the complexity of funding recommendations: a comparative analysis of health technology assessment outcomes in four European countries

Article (Accepted version) (Refereed)

Original citation:
Cerri, Karin H., Knapp, Martin and Fernández, José-Luis (2015) *Untangling the complexity of funding recommendations: a comparative analysis of health technology assessment outcomes in four European countries.* Pharmaceutical Medicine. ISSN 1178-2595

DOI: 10.1007/s40290-015-0112-8

Untangling the Complexity of Funding Recommendations: A Comparative Analysis of Health Technology Assessment Outcomes in Four European Countries

Karin H. Cerri,[1,2] Martin Knapp,[1] and Jose-Luis Fernandez[1]

1. London School of Economics and Political Science, London, England
2. Janssen Pharmaceutica NV, Beerse, Belgium

Corresponding author: Dr. Karin H. Cerri
London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK. Tel: +44 (0)20 7405 7686; k.h.cerri@lse.ac.uk

Running head: Comparative analysis of HTA outcomes in four European countries

Abstract

Objectives: Health Technology Assessment (HTA) agencies produce recommendations that guide public funding of pharmaceuticals, based on various criteria. We explored factors that may contribute to explaining differences in coverage decisions by the National Institute for Health and Care Excellence (NICE) in England and Wales, the Scottish Medicines Consortium (SMC), the Dutch College voor Zorgverzekeringen (CVZ), and the French Haute Autorité de Santé (HAS).

Methods: A dataset of 977 HTA decisions made in 2004-2009 was created. A three-category outcome variable was used (decision to 'recommend', 'restrict' or 'not recommend' a technology). Multivariate analyses explored impacts of clinical, economic, process and socio-economic variables in their decision-making.

Results: Relative to the CVZ and adjusting for a range of confounders, technologies were more likely to be recommended by NICE and HAS, and restricted or not-recommended by the SMC. Recommendation was significantly associated ($p \leq 0.10$) with several variables: strength of clinical evidence (number of trials, use of active comparator-arm, demonstration of clinical superiority) orphan status and indication for cancer. Simultaneous assessment of multiple rather than single pharmaceuticals was associated with increased probability of restriction.

Conclusions: In this European multi-HTA study, appraisal outcomes differed significantly across HTA bodies. A range of evidence and non-evidence factors were associated with HTA decisions, confirming the value of comprehensive, multivariate analyses. Nevertheless, a large proportion of variance in HTA decisions remained unexplained, suggesting that greater transparency of decision-making is needed, along with associated further research.

Word count: 234

Key points:

1. European HTA agencies included in this analysis significantly differed in their decision-making, after adjusting for a range of factors including the scoping and decision-making process, evidence considered, and socio-economic differences
2. The strength of the evidence demonstrating the clinical value of the pharmaceutical, together with the therapeutic area which it targets, have a significant role in HTA decisions
3. 'Untangling' the complexity of payer decision-making was limited given that multivariate analyses could only explain between 13% and 30% of the observed variability in decision-making, despite taking into account a wide range of variables

1. Introduction

The health technology assessment (HTA) decision-making process is prominent in several European countries, and advises healthcare systems on the appropriate use of pharmaceuticals (among other technologies) and whether they should be recommended for public funding. HTA decisions have an impact on clinicians and patients by defining to whom a medicine can be made available, for how long, at what price and under what circumstances or conditions. Thus HTA decisions represent a key point within the complex decision-making process that governs funding and access for pharmaceuticals in some countries.

Because HTA decisions have a significant impact on patient care, transparency around HTA decision-making is highly important. For example, the European Commission (1) has launched an up-date of the EU Transparency Directive (89/105/EEC) to further emphasize this notion. The proposed amendment from the European Commission now includes direct reference to the concept of transparency and proposes that determination of price and access requires "transparent, objective and verifiable criteria" [p.11].

It is within this context that we examine the impact of a range of evidence, process and socio-economic factors on HTA decisions and public funding of pharmaceuticals in a selection of European countries, namely UK, France, and the Netherlands:

- National Institute for Health and Care Excellence (NICE), England and Wales);
- Scottish Medicines Consortium (SMC), Scotland
- College voor Zorgverzekeringen (CVZ), the Netherlands
- Haute Autorité de Santé (HAS – French Health Authority), France

Descriptions of these four HTA bodies are provided as supplementary material (part 1).

Specifically, the analysis aimed to (i) assess whether these HTA bodies differ amongst themselves in terms of the coverage decisions they make, while adjusting for a range of confounding factors; and (ii) explore the factors that contribute to explaining the variability in coverage decisions made by the four HTA bodies in 2004-2009.

2. Methods

2.1 Explanatory Variables

It was hypothesised, based on the nature of the HTA bodies selected and the literature available, that decisions are driven by the HTA decision-making process itself, the evidence considered within that process, and by the socio-economic and political context in which decisions are made (further information provided as supplementary material, part 2). Research has shown that the evidence related to the medicine under review (whether clinical, economic or otherwise) can have an impact on HTA decisions (2-8). The literature examining the HTA appraisal process provides insights into a number of process-related factors that can potentially influence decisions (9-12). Indeed, reference in the literature is made to the impact of broader healthcare and welfare characteristics on HTA decision-making, such as healthcare spending per capita, societal willingness to pay, the structure of the healthcare system, as well as ethical and social considerations (13-17). In line with the hypothesised drivers of HTA decision-making, 29 variables were identified, shown in Table I.

2.2 Sample

The choice of HTA body included in the analysis aimed to maximize the chance of obtaining useful data to address the research question, provide a comprehensive platform for analysis of the research question rather

than the examination of a particular factor in isolation, and to allow for the exploration of variation in the implementation and drivers of HTA decision-making. To gather information on variables related to the coverage decision for specific technologies, data were required from HTA bodies that published their appraisal decisions and rationale for those decisions in a comprehensive format that was accessible to the public. To this end, pharmaceutical technology appraisals performed by NICE, SMC, CVZ and HAS were selected to provide the sample for this analysis. To capture a sufficient number of appraisals for both individual and aggregate analyses, a five-year time horizon was implemented. Data extraction by one of the authors (KC) took place in July 2008-December 2009. The sample included all drug technology appraisals (as opposed to medical devices or other interventions) made during the period January 2004-June 2009, indicated for an adult population (aged over 17 years). Technology appraisals were excluded from the analysis for any of the following reasons: if they focused on a non-adult population; if they appraised non-drug interventions; if marketing authorisation was withdrawn; if the Amélioriation du Service Médicale Rendu **(**ASMR) rating was not reported (HAS only); if an abbreviated or independent review panel (IRP) guidance was issued (SMC only); or if the full guidance was not available.

The French HAS issued 2600 recommendations in 2004-2009. Given resource constraints, it was not possible to review all of these recommendations, thus the HAS sample was restricted to technologies appraised by SMC/NICE. While it is understood that this approach may lead to selection bias, the benefit derived was that it increased the opportunity for comparability across agencies, by collecting information on a common list of compounds and streamlining data extraction to those appraisals that were relevant for the research question.

A data extraction form was developed to extract information from public data sources to ensure transparency, reproducibility and consistency. The resulting extracted data was coded and prepared for analysis. The variable definitions and data sources used are shown in Table I.

2.3 Outcome variable

To determine the appropriate outcome variable, the approach to decision-making was assessed for each HTA body. In the period 2004-2009, HAS used a five-point scale known as the ASMR rating, that classified a technology according to the level of incremental medical service rendered, with the highest level (I) indicating a high incremental medical service and the lowest (V) denoting no additional medical benefit. Subsequent pricing and volume negotiations were then based on these ratings. Similarly, the CVZ has several possible HTA options in its armamentarium: the GVS 1A or 1B reimbursement list, or specific reimbursement policies (e.g. the expensive drug list, or list '2'), according to whether the technology is for inpatient or outpatient use, or have special conditions associated with their use. The SMC utilises a three-category approach: accepted for use, accepted for restricted use and not accepted for use in NHS Scotland, therefore clearly not operating in a binary-type HTA decision system. Similar to the SMC, NICE either recommends, recommends with conditions or does not recommend technologies for NHS funding in England and Wales. In all four HTA bodies, decisions are based on more than a binary decision-process. This was a key justification for developing a non-binary modelling approach. In addition, previously published analyses have also utilised non-binary approaches (6,18). The analysis used a standardised three category outcome variable where the new technology could be:

- recommended for routine use
- recommended for restricted use
- not recommended for use

Our analysis focuses on the HAS and its decision-making. Technologies with an ASMR V are considered 'not recommended' because the HAS did not find any evidence to suggest that the technology would offer

incremental value relative to alternative treatment options. In other words, with an ASMR V, the HAS signals it finds no basis upon which to recommend the technology for use. Economic considerations are not considered by the HAS, differing from the CVZ, NICE and SMC. Price, in particular, is known by NICE and SMC at time of appraisal, and officially, this is not the case for the HAS. A variable was included in the analysis to capture where price was known up-front during the appraisal or not (Table I).

While a non-recommendation by CVZ, NICE or SMC generally implies no market entry, a non-recommendation from the HAS (ASMR V) does not preclude market entry, as this is driven by the negotiation between the manufacturer and the CEPS. The CEPS is a separate committee that is responsible for finalising the price and volume agreements for technologies. We were not able to include CEPS decision-making in our analysis as these negotiations are confidential, to our knowledge, and so would not be able to analyse the criteria driving CEPS decision-making. Other research conducted on the ASMR has made a similar assumption whereby technologies with ASMR V were considered to represent non-recommendation (18). Further details on how the outcome variable was defined are provided in supplementary material (part 3).

2.4 Statistical Analyses

Descriptive statistics were calculated for each variable, stratified by outcome group (recommended, restricted, or not recommended) and we adopted a 0.10 level of significance. For categorical variables, we used the Chi-squared test to test for differences in proportions across the three outcomes. For continuous indicators, we used the ANOVA test to test for differences between means for normally distributed indicators and the Kruskal-Wallis test to identify differences between ranks of means for not normally distributed indicators. It was recognised that collinearity could exist as these variables were included in the multivariate analysis, and thus a step-by-step process was followed to look for evidence of collinearity.

Multinomial logistic regression was used in the analysis to model the probabilities associated with the three types of technology appraisal outcome. Two base cases were examined in multivariate analyses: base case model 1 included all four HTA bodies in the pooled analysis accompanied by fixed effects, and excluded economic variables not common across the four (in particular ICER-related variables). Base case model 2 included HTA bodies (NICE and SMC) that consider the cost-effectiveness of technologies to avoid imputation of information that was not formally considered. In the 2004-09 period, HAS did not consider economic variables and the CVZ was excluded due to low reporting of ICERs within the CVZ appraisals (11% of appraisals reported ICERs); this is perhaps in part driven by the fact that cost-effectiveness considerations were formally introduced in the CVZ process only in 2006 and cost-effectiveness results are only reported for those technologies that are associated with an incremental therapeutic benefit to patients.

The 'recommended' outcome was selected as the referent category in the analysis. The objective of the analysis was to identify, ceteris paribus, the effect of a range of factors potentially associated with HTA appraisal decisions, and to assess which combination of factors best explains the pattern of HTA decisions. Given the wide range of factors considered in the analysis (see Table I) a process was developed to determine which explanatory variables would appear in the final specification of the model:

- First, bivariate regression models were run to ascertain the degree of correlation between individual explanatory variables and appraisal decisions.
- On the basis of these models, a subset of indicators was selected which included those variables that showed at least moderate significance levels (indicators with $p<0.25$). A preliminary model was estimated including these indicators.
- The model was reduced by removing those variables with significance levels above the 0.10 threshold. To guarantee its stability, this 'base' model was re-estimated by sequentially removing one variable at a time and verifying the stability of the effects on the coefficient and significance level of the remaining estimates.

- The model was subsequently tested through alternative model specifications to examine its robustness and to assess the sensitivity of the results to different assumptions
- As a final step, the base-case model results were presented to HTA representatives in 1-hour telephone interviews, to seek feedback on the variables identified within the base-case model, the coefficient and level of significance to assess the validity of the model.

The application of the model-specification process outlined above facilitated the interpretation of the results of the models whilst allowing the analysis to explore the impact of the wide range of indicators collected in the study. A step-by-step process was followed to look for evidence of collinearity within the set of regressors.

While significant effort was made to identify information relevant to the variables of interest, a small proportion of the data could not be found. To maximize the sample size, imputation techniques were used to estimate entries for the missing observations. Missing values were replaced with regression imputation estimates using the 'impute' command in STATA software. The imputed values obtained were then checked manually to ensure their face validity. In addition, dummy variables were created to identify observations with missing data to test in the regression models whether the lack of data was significantly associated with differences in the outcome variable.

Sensitivity analyses were performed on the base-case regression model to test for changes in the effects to alternative specifications of the indicators and help evaluate the robustness of the results. The sensitivity analyses included: i) examining the impact of a binary (coded as: covered/not covered) rather than a three-category outcome variable; ii) examining the impact of including only the sub-set of technologies which were appraised by all four HTA bodies; (iii) examining the impact of restricting the analysis to technologies indicated for the treatment of cancer.

Statistical analyses were conducted using Intercooled (IC) STATA (Version 10.1 2009).

3. Results

Univariate Analyses

In total, 1258 appraisals were reviewed and 977 HTA decisions made in 2004-2009 met the inclusion criteria: 118 NICE decisions, 288 SMC decisions, 256 CVZ decisions and 315 HAS decisions (Figure I). 281 appraisals were excluded from the analysis, most commonly because they were abbreviated submissions, or because they focused on non-adult populations (Figure I). Within this data set, 27% of the decisions recommended funding of the technology, 39% restricted funding and 35% did not recommend funding the technology. HTA bodies differed in the pattern of decision-making ($p<0.001$, Figure II a). The most common decision by NICE was to restrict funding (58%), whereas for the CVZ the most common decision was to recommend (51%)[1]. For both the SMC and HAS, the most common decision was to not recommend (46% and 44%, respectively). When a sub-sample of those technologies common across all four HTA bodies was examined (n=192, listed in supplementary material, part 4), differences in the pattern of decision-making between them was maintained ($p<0.05$, Figure IIb). The trends in HTA decisions over time (January 2004 - June 2009) for each HTA body are presented in Figure III. These data suggest that within HAS and NICE there has been a decrease in the proportion of positive recommendations made over time and a corresponding increase in the proportion of

[1] While prices of 1A listed technologies are pre-defined based on already existing reference technologies, it is the manufacturer that submits a proposal with a price in the anticipated range according to the therapeutic differentiation of the product, and it is not the CVZ who asks for a price reduction to meet the reference price rule as a result of the appraisal. In addition, both 1A and 1B result in similar coverage from the patient perspective. This is why we decided to consider both 1A and 1B listings recommendations by the CVZ.

non-recommendations. The CVZ and SMC, on the other hand, appear to have maintained a relatively stable pattern of decision-making over time.

Univariate analyses showed that within the multi-HTA sample, a range of factors may play an important role in determining HTA decision-making. For these variables (Table II), statistically significant differences were observed between interventions that were recommended, restricted and not recommended ($p \leq 0.10$). Univariate analyses stratified by HTA agency can be found in supplementary material (part 5).

Multivariate Analyses

Base Case Model 1

Base case model 1 included 20 variables yielding a pseudo R-squared of 0.13, suggesting that the model explains approximately 13% of the variability in HTA decisions across the multi-HTA sample (Table III).

One of the objectives of this analysis was to assess whether differences between HTA bodies in the pattern of decision-making after adjusting for a broad range of covariates. When the impact of NICE, SMC and HAS on HTA decisions was examined relative to the CVZ, the results suggest that NICE and HAS assessment bodies are strongly associated with a decreased odds of a restriction or non-recommendation. This can be contrasted with the effect of the SMC, which was found to statistically significantly increase the log-odds of both restriction and non-recommendation in all base case models. The impact of the HTA body was highly statistically significant across all assessments, while adjusting for a range of confounders.

With regard to clinical variables, a higher number of RCTs, and the inclusion of an active comparator in the trial design had a significant impact on the log-odds of recommendation versus restriction or non-recommendation. Specifically, if the technology was compared to an active control rather than placebo, the log-odds for restriction relative to recommendation decreased (p=0.011), as did the log-odds of non-recommendation relative to recommendation (p=0.001), while holding all other variables constant. If the technology demonstrated clinical superiority, the log-odds for restriction relative to recommendation decreased (p=0.023), as did the log-odds of non-recommendation relative to recommendation (p=0.001), while holding all other variables constant. Orphan designated pharmaceuticals were less likely to be rejected (p=0.023).

Process factors had a significant impact on HTA decisions in this multi-HTA analysis. An increase in the number of technologies appraised simultaneously exerted a bigger impact on the log-odds of a restriction (p=0.002). The inclusion of patient submissions and patient evidence as part of the process was linked with an increase in the log-odds of a restriction (p=0.008).

Socio-economic factors contributed to explaining the variability in HTA decisions across the HTA bodies. With regard to the size of the population within the HTA body remit, a unit increase in the population size increased the odds of both restriction and non-recommendation, and both effects were statistically significant.

Base Case Model 2

Base case model 2 (NICE and SMC) included 19 variables yielding a pseudo R-squared of 0.16, suggesting that the model explains approximately 16% of the variability in HTA decisions across the multi-HTA sample (Table IV).

With regard to clinical variables, as in base case model 1, a higher number of RCTs, and the inclusion of an active comparator in the trial design had a significant impact on the log-odds of recommendation versus restriction or non-recommendation. Specifically, if the technology was compared to an active control rather than placebo, the log-odds for restriction relative to recommendation decreased (p=0.05), as did the log-odds of non-recommendation relative to recommendation (p=0.022), while holding all other variables constant. No impact of orphan status was identified in base-case model 2.

With regard to economic variables, a unit increase in the ICER was shown to increase the odds of restriction (p=0.011) and the log odds of non-recommendation (p=0.001), relative to recommendation.

Process factors had a significant impact on HTA decisions in this multi-HTA analysis. Relative to recommendation, an increase in the number of technologies significantly increased the odds of both restriction (p=0.001) and non-recommendation (p=0.074). No effect on decision was found for either patient submissions or patient evidence.

Socio-economic factors contributed to explaining the variability in decisions across the HTA bodies. A unit increase in the population size within the HTA body remit increased the odds of both restriction and non-recommendation, and both effects were statistically significant.

Sensitivity Analyses

The results of a series of sensitivity analyses suggest that the base case models are robust. A sensitivity analysis was conducted using a binary outcome variable, providing similar results to the base-case multi-HTA analysis. Sensitivity analyses were also conducted on the sample of technologies indicated for cancer treatment (247 appraisals of 977) as well as on a sub-set of the sample including only those technologies appraised by all four HTA bodies at least once (192 appraisals of 977). The results of these analyses suggest that, compared to the base case model 1, the explanatory power of the combination of clinical, process and socio-economic factors is twice as high when including only those technologies from the same disease area or technologies that all four HTA bodies hold in common (as implied by the higher pseudo R-squared: 0.26 and 0.27 in these sub-analyses respectively, compared with 0.13 in the base-case multi-HTA analysis).

For each HTA agency, a summary of the statistically significant explanatory variables identified in descriptive and multivariate analyses is given in Table V.

4. Discussion

This study aimed to untangle the factors driving funding decisions by conducting multivariate analyses of 977 decisions observed within a multi-HTA data set of appraisals performed by four European HTA agencies: NICE, SMC, CVZ and HAS. Previous comparative analyses across HTA bodies identified in the literature were primarily qualitative or adopted descriptive quantitative methodologies (5,18-20). Such descriptive analytical techniques make it difficult to interpret the relative contribution of each factor, given the absence of adjustment for other factors in the analysis. In contrast, we created a bespoke dataset of HTA coverage decisions from four European HTA bodies over a five-year period and utilised statistical analysis to assess the relative contribution of a comprehensive range of factors on coverage decisions.

Significant variability in decision-making across HTA bodies

Addressing the first objective, the analysis confirmed that observed differences in decision-making by HTA agencies remained after adjusting for a range of factors including scoping, evidence considered, process and socio-economic differences. On average, 27% of the decisions (ranging from 18% in HAS to 51% in CVZ) recommended funding of the technology, 39% restricted funding (range 33% CVZ to 58% NICE) and 35% did not recommend funding the technology (range 14% NICE to 46% SMC) (Figure IIa). Indeed, HTA bodies differ significantly even when the analysis is restricted to coverage decisions made for the same set of technologies (Figure IIb). From one perspective, such differences in coverage patterns and the factors that drive decisions can be explained by the fact that each HTA body is designed to match as closely as possible the specific healthcare system it serves. Therefore, variation observed across HTA bodies could be considered to reflect the reality of healthcare market variations between countries. However, from another perspective, differences in the proportion of recommendations, restrictions and non-recommendations might be seen to be contrary to the principle of equitable access to treatment in Europe.

In particular, non-recommendation was observed in 14% and 16% of NICE and CVZ coverage decisions, and 44% and 46% of HAS and SMC coverage decisions, respectively. These technologies were not recommended for reimbursement despite obtaining a license for use within Europe from regulatory agencies, primarily the European Medicines Agency (EMA). Differences in the criteria used to assess the value of new technologies between the EMA and HTA bodies could be seen as a key driver of this situation. This suggests there could be potential for efficiencies in the regulatory and HTA processes so that effort is not spent on generating marketing authorizations for a significant proportion of technologies that end up not being recommended for reimbursed use by HTA bodies.

Explaining the variability in decision-making – so much for 'untangling'?

The second objective of this analysis was to explore the factors that contribute to explaining the variability in coverage decisions made by the four HTA bodies in 2004-2009. The degree of 'untangling' achieved was limited given that multivariate analyses could only explain between 13% and 30% of the observed variability in decision-making, despite taking into account a wide range of variables. A higher proportion of the variability could be explained when limiting the analysis to a common set of technologies evaluated by all HTA bodies. This suggests that the variability in HTA outcomes may in part be driven by the heterogeneity observed in the technologies selected for appraisal. However, even when this heterogeneity is managed by focusing the multivariate analysis on a subset of technologies that are either commonly appraised across HTAs or focus on a single disease area, more than 70% of the variability in decision-making could not be explained.

Indeed, whether the technology was appraised by one HTA or another was found to have a statistically significant effect on the outcome, even when adjusting for a range of confounding factors. Relative to the CVZ, the results suggest that NICE and HAS are associated with a higher probability of recommendation, while the SMC is associated with a higher probability of both restriction and non-recommendation. This may be due to the fact that the HTA body is surrogate for other variables and thus that additional variables needed to be included, alternative methods of analysis tested, and/or that decision-making is partly random. This result echoes in part results from Clement et al. (5) reporting statistically significant differences in the nature of the HTA decisions made by NICE, PBAC and CDR. Our analysis would support the view that while evidence-based assessment of technologies is a key part of the HTA process, decision-making is ultimately driven by a range of factors that appear to be highly specific to each HTA body.

The strength of the evidence demonstrating the clinical value of the pharmaceutical, together with the therapeutic area which it targets, have a significant role in HTA decisions. Interestingly, there were some specific variables which appeared to have an impact (such as demonstration of clinical superiority), while other clinical/disease evidence variables did not show any effect in the multivariate analyses (e.g. number of observational studies). This raises the question as to whether it would be of benefit for HTA bodies and manufacturer submissions (and clinical trials) to streamline the submission and review of clinical evidence that most drives decision-making, encouraging more efficient HTA assessments. There were no other published cross-HTA multivariate studies available for comparison, although single-HTA analyses have identified a range of clinical evidence characteristics associated with HTA outcomes (21).

The multi-HTA analysis of NICE and SMC appraisals (more than 400 of which were analysed here), representing one of the largest analyses of the effect of cost-effectiveness on HTA decision-making, confirmed the hypothesis that the effect of the ICER is similar to that observed in individual HTA analyses (6,7,22): a higher ICER decreases the probability of recommendation, an effect that is highly statistically significant in our analyses. While CVZ and HAS were excluded from this analysis due to limited consideration (CVZ)[2] or no

[2] It should be noted that since 2005, CVZ has increasingly used cost-effectiveness evidence in its decision-making, in particularly for technologies evaluated for 1B listing and specific instances. However, because of the fact that in the sample of CVZ decisions between 2004-2009 only 11% of technologies (<30) were supported by reported ICER data, it was not included in base case model 2

consideration (HAS) of cost-effectiveness evidence, since the time of this analysis both HTA bodies have increased the role of economic criteria in their decision-making.

This multi-HTA analysis finds support for the role of process factors in explaining decision-making, in particular the role of patient evidence, and the impact of single versus multi-technology appraisals. The multivariate analyses confirmed that an increase in the number of technologies appraised simultaneously increased the probability of restriction and non-recommendation. This would suggest that in circumstances where multiple technologies are being appraised simultaneously, a 'winner' is more likely to be selected, with the remainder being restricted or not recommended for funding. This has implications when considering the efficiencies that may be gained in the appraisal process by coupling the review of multiple technologies together, against the time-lag that this may create in access to new medicines, depending on how frequently multi-technology reviews are conducted.

National population size had a significant impact on HTA decision-making in this multi-HTA analysis, an effect not previously observed in individual HTA analyses (5-6,22). Greater population size decreased the probability of recommendation. It is not clear to what extent this reflects the effect of the absolute size of the patient population eligible for treatment: this variable also increases the odds of restriction and non-recommendation, but its effect is only significant in base case model 1 (on odds of restriction relative to recommendation). It is plausible that population size may serve as a proxy for the budget impact of adopting a technology, but when the budget impact variable was introduced in additional sensitivity analyses, it was not significant.

Limitations

We focused specifically on pharmaceutical technologies as we wanted to generate a sample of technologies representative of the majority of decisions made by the HTA bodies included in the study. Thus results of this analysis are not readily generalizable to non-pharmaceutical technologies. The selection of HTA bodies for inclusion in this analysis was driven by the objectives of the study, and in addition HTA bodies were selected that reported their appraisals in English, French or Dutch (to make it feasible for us to collect our primary data).

The potential for information bias, particularly in how the outcome and explanatory variables were defined and extracted for each HTA body, was recognised, and a data extraction protocol and process was defined to limit inconsistencies in the extraction process. However, there are still limitations. In particular, decisions are communicated differently across HTA bodies: HAS uses ASMR ratings to define incremental therapeutic value, CVZ attributes technologies to different reimbursement lists, SMC uses a three-category system and NICE a binary category system. In addition, the use of the 'restricted' category in this analysis is challenged by the fact that inter- and intra-HTA differences exist in how it is applied.

The generalizability of the results is further constrained by the fact that HTA appraisal processes are subject to change (for example, the potential introduction of value-based pricing in UK in 2014, changes in France with increasing consideration for medico-economic evidence in 2013), and thus the results obtained are specific to 2004-2009. HTA bodies, both within and outside Europe, vary in their objectives, and in the approach and methods used to implement HTA within their jurisdictions.

Socio-economic indicators such as GDP are known to be influenced by many different factors. Indeed, such indicators act as a surrogate for many characteristics of the countries to which they apply. In addition, such indicators that vary at the HTA body level, rather than the technology appraisal level, are unlikely to have a very strong effect, due to the limited number of HTA bodies in this analysis. Therefore, the interpretation of the impact of such broad indicators, such as the percentage of GDP spent on healthcare, will need to take into

consideration the risk that variations observed in such indicators across HTA bodies may be correlated with other factors.

Given the heterogeneity in decision-making observed in this analysis, extrapolating to other HTA bodies not included in this research is not advisable without having a more concrete understanding of their decision-making processes and the outcome definition used.

Implications

This analysis involved an intensive data collection exercise lasting 18 months: information about HTA decision-making and the characteristics of a technology, the appraisal process and the context in which decisions are made, is not readily available. It raises interesting questions about the role of transparency in decision-making, in particular to what extent any transparency which exists at the level of individual technology appraisals should be extended to transparency in the outcome and drivers of decision-making for the decisions made by HTA bodies.

The implications of this analysis in the light of the increasing collaboration between European HTA bodies through EUnetHTA should be addressed. A key project ongoing within EUnetHTA is the *HTA Core Model for Rapid Relative Effectiveness Assessment of Pharmaceuticals*: the "HTA Core Model defines the content elements to be considered in an HTA and facilitates standardised reporting. The aim is to share information, to avoid duplication of work, and to facilitate the adaptation of information in national HTA reports and the co-production of HTA reports (by multiple HTA agencies)" (EUnetHTA accessed 2013). This core-model concept assumes that there is overlap in the timing of when pharmaceutical appraisals take place, and overlap in the scoping practices of HTA bodies. In the sample covered in this study, only 10% of technologies were appraised by all four HTA bodies, and significant differences in the timing of appraisals were observed. This has implications in terms of the co-ordination and timing of centralised HTA Core Model initiatives to ensure that they have maximum relevance to the HTA stakeholders they wish to serve. This analysis may also help to inform which types of evidence might best be included in the Core Model, based on those variables which appear to have the most effect in explaining payer decisions.

The framework adopted in our analysis, notwithstanding its limitations, provides an example of a comprehensive analytical approach for understanding coverage decision-making. While recognising the diversity in scope, objective and context in which different HTA bodies operate, convergence towards an analytical framework for the analysis of coverage decisions in future research may be of value.

5. Conclusions

In this European multi-HTA study, a comprehensive and multivariate analytical approach found that clinical, process, socio-economic factors and an "HTA-agency effect" explained part of the observed variability in HTA decisions, helping to untangle some aspects of decision-making. However, this extensive analysis of HTA decision-making could explain no more than 30% of the observed variability, suggesting that while some aspects of decision-making have been 'untangled', greater transparency of decision-making is needed, along with associated further research.

Acknowledgements

The authors would like to thank Michael Drummond and Alistair McGuire for their valuable comments on the methods used and interpretation of analyses that were performed.

Compliance with Ethical Standards

Funding No funding was received for the conduct or publication of this study.

Conflicts of interest KC was an employee of Bristol-Myers Squibb during the time this research was conducted. MK has received funding from NICE for PSSRU's part in the NICE Collaborating Centre for Social Care. JLF has no conflicts of interest to declare.

Reference List

(1) European Commission. Amended proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the transparency of measures regulating the prices of medicinal products for human use and their inclusion in the scope of public health insurance systems. COM(2013) 168 final/2. 20-3-2013.

(2) Sorenson C, Drummond M, Kanavos P. Ensuring value for Money in Health Care: the role of HTA in the European Union. 2008. Cornwall, UK., World Health Organization 2008, on behalf of the European Observatory on Health Systems and Policies.

(3) Bellanger MM, Cherilova V, Paris V. The "Health Benefit Basket" in France. Eur J Health Econ 2005 December;Suppl:24-9.

(4) Sandier S, Paris V, Polton D. Health Care Systems in Transition – France 2004. 2004. Copenhagen, Denmark, WHO Regional Office for Europe on behalf of the European Observatory on Health Systems and Policies.

(5) Clement FM, Harris A, Li JJ, Yong K, Lee KM, Manns BJ. Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia, and Canada. JAMA 2009 October 7;302(13):1437-43.

(6) Dakin HA, Devlin NJ, Odeyemi IA. "Yes", "No" or "Yes, but"? Multinomial modelling of NICE decision-making. Health Policy 2006 August;77(3):352-67.

(7) Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Econ 2004 May;13(5):437-52.

(8) Mason AR, Drummond MF. Public funding of new cancer drugs: Is NICE getting nastier? Eur J Cancer 2009 May;45(7):1188-92.

(9) Al MJ, Feenstra T, Brouwer WB. Decision makers' views on health care objectives and budget constraints: results from a pilot study. Health Policy 2004 October;70(1):33-48.

(10) Menon D, Stafinski T, Stuart G. Access to drugs for cancer: Does where you live matter? Can J Public Health 2005 November;96(6):454-8.

(11) OECD. Health Technologies and Decision Making. 2005. OECD.

(12) Vuorenkoski L, Toiviainen H, Hemminki E. Decision-making in priority setting for medicines--a review of empirical studies. Health Policy 2008 April;86(1):1-9.

(13) Bryan S, Williams I, McIver S. Seeing the NICE side of cost-effectiveness analysis: a qualitative investigation of the use of CEA in NICE technology appraisals. Health Econ 2007 February;16(2):179-93.

(14) Packer C, Simpson S, Stevens A. International diffusion of new health technologies: a ten-country analysis of six health technologies. Int J Technol Assess Health Care 2006;22(4):419-28.

(15) Owens DK. Interpretation of cost-effectiveness analyses. J Gen Intern Med 1998 October;13(10):716-7.

(16) Buxton MJ. How much are health-care systems prepared to pay to produce a QALY? Eur J Health Econ 2005 December;6(4):285-7.

(17) Ross J. The use of economic evaluation in health care: Australian decision makers' perceptions. Health Policy 1995 February;31(2):103-10.

(18) Kanavos P, Nicod E, van den Aardweg S, Pomedli S. The impact of health technology assessments: an international comparison. Euro Observer 4. 2010.

(19) Barbieri M, Hawkins N, Sculpher M. Who does the numbers? The role of third-party technology assessment to inform health systems' decision-making about the funding of health technologies. Value Health 2009 March;12(2):193-201.

(20) Lexchin J, Mintzes B. Medicine reimbursement recommendations in Canada, Australia, and Scotland. Am J Manag Care 2008 September;14(9):581-8.

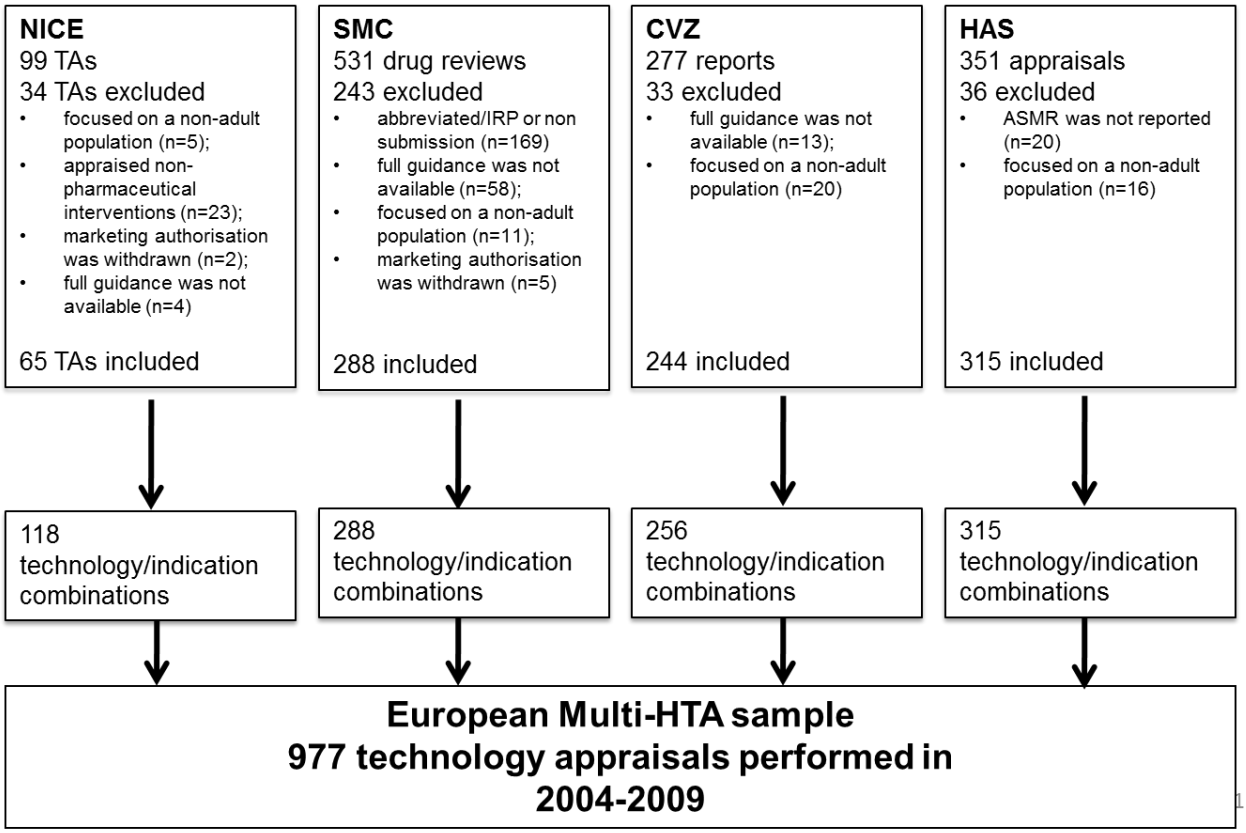
(21) Fischer, K.E. A systematic review of coverage decision-making on health technologies— Evidence from the real world. Health Policy 2012 107: 218– 230.

(22) Cerri KH, Knapp M, Fernandez JL. Decision making by NICE: examining the influences of evidence, process and context. Health Econ Policy Law 2013 May 21;1-23.

(23) Cerri KH, Knapp M, Fernandez JL. Public funding of pharmaceuticals in the Netherlands: investigating the effect of evidence, process and context on CVZ decision-making. Eur J Health Econ 2013 July 18.

Figures and tables

Figure I. Flowchart of Multi-HTA dataset including technology appraisals performed by NICE, SMC, CVZ and HAS during 2004-2009



Legend:

TA- Technology Appraisals; NICE –National Institute for Health and Care Excellence (NICE); Scottish Medicines Consortium (SMC); College voor Zorgverzekeringen (CVZ); Haute Autorité de Santé (HAS)

Figure II a. Distribution of type of HTA coverage decisions on pharmaceutical technologies by NICE, SMC, CVZ and HAS in 2004-2009 (n=977)

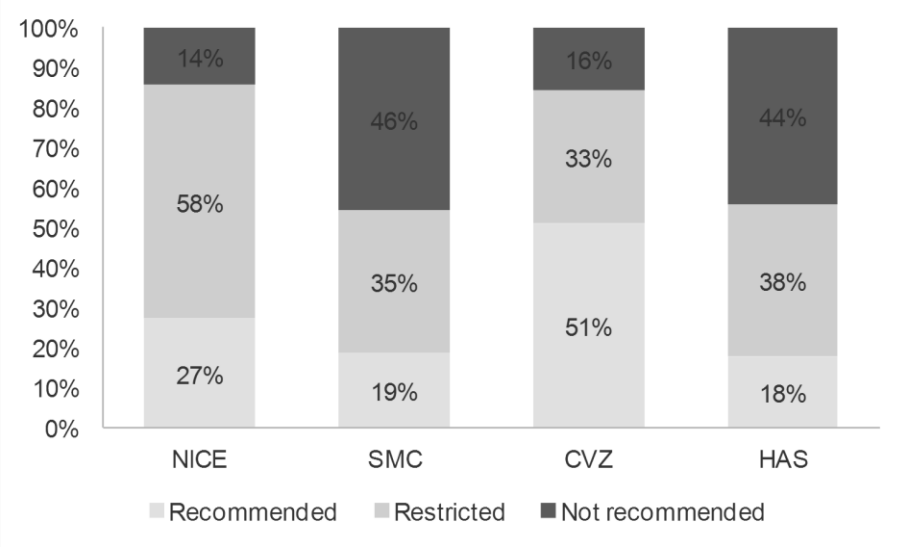
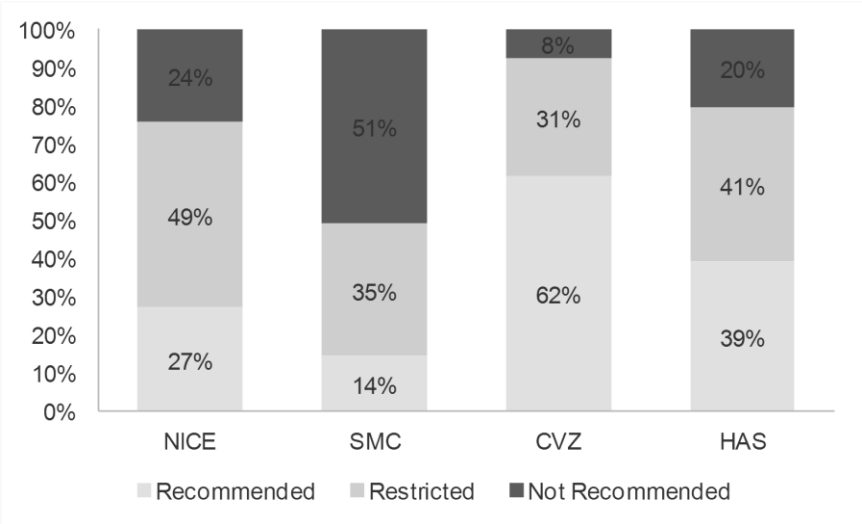


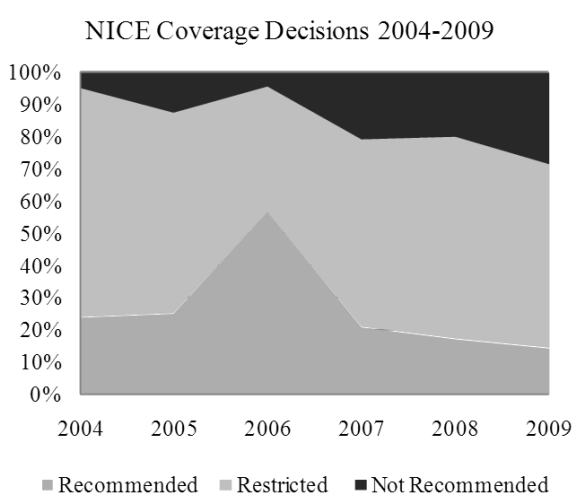
Figure II b. Distribution of type of HTA coverage decisions made in 2004-2009 on subset of pharmaceutical technologies appraised by NICE, SMC, CVZ and HAS (n=192)

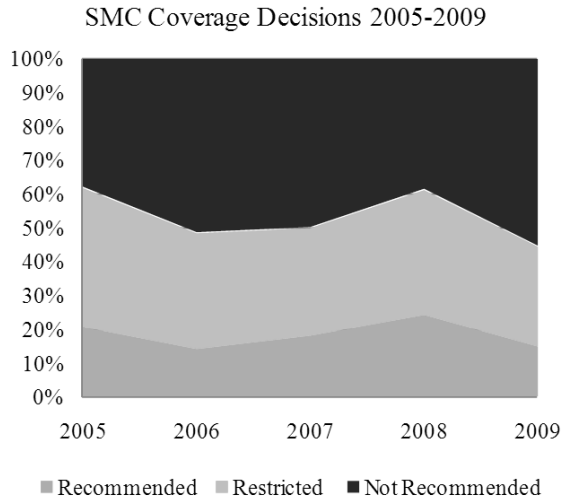


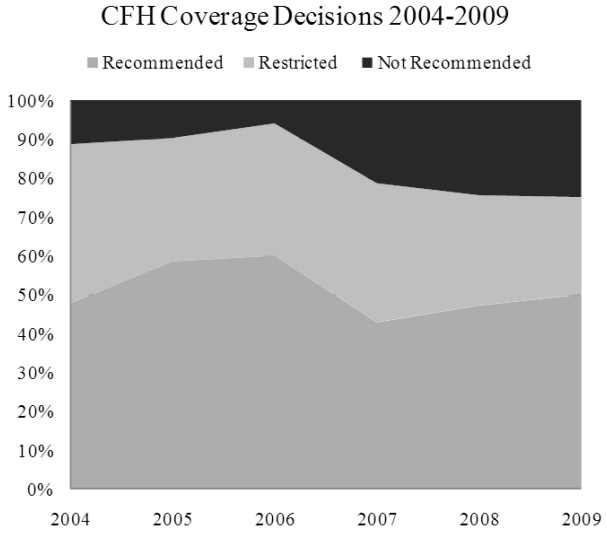
Legend:

NICE –National Institute for Health and Care Excellence (NICE); Scottish Medicines Consortium (SMC); College voor Zorgverzekeringen (CVZ); Haute Autorité de Santé (HAS). Recommended = pharmaceutical technology recommended for routine use; Restricted = pharmaceutical technology recommended for restricted use; Not Recommended = pharmaceutical technology not recommended for use

Figure III NICE, SMC, CVZ and HAS HTA decisions between 2004-2009 (June), by year (n=977)







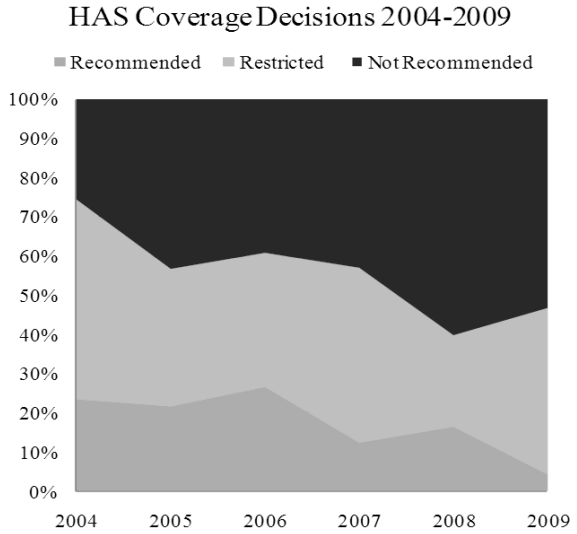


Table I Summary of Included Explanatory Variables, their Definition, and Data Sources

#	Variable	Unit of Measure	Definition	Data Sources
Evidence – clinical and disease-related variables				
1	Number of RCTs considered in decision	Count	The number of distinct randomised controlled trials (RCTs) that provide data related to the therapeutic indication under evaluation. Excluded: studies that are single-arm, have no randomization, or are non-interventional.	NICE: Technical Appraisal Report (TAR) Section 2 SMC: SMC Advice, "Summary of evidence on comparative efficacy" section SMC Advice CVZ: CFH-rapport , section 2a; Farmacotherapeutisch rapport, section 3, 4a-f HAS: Advice from the Commission de la Transparence, section 3 (3.1 – 3.3)
2	Size of population included in RCTs	Numeric	Mean number of patients per RCT.	
3	Length/extent of follow-up in RCT	Numeric	Mean number of weeks over which data are collected on patients that entered the RCTs (see variable no. 1).	
4	Statistically Significant superior results	Categorical (superior/ not superior /inconsistent results)	Presence of statistically significant superiority of technology vs. comparator for primary endpoint(s). If the technology showed statistically significant superiority in one trial, but not in another, the results were considered to be 'inconsistent'. RCTs designed as 'non-inferiority' studies were classified as not demonstrating superiority (i.e. 'no').	NICE: Technical Appraisal Report (TAR) Section 4 SMC: SMC Advice, "Summary of evidence on comparative efficacy" section SMC Advice CVZ: CFH-rapport , section 2a; Farmacotherapeutisch rapport, section 3, 4a-f HAS: Advice from the Commission de la Transparence, section 3 (3.1 – 3.3)
5	Use of active comparator	Numeric	Percentage of RCTs where active comparator was used, as opposed to placebo.	NICE: Technical Appraisal Report (TAR) Section 4 SMC: SMC Advice, "Summary of evidence on comparative efficacy" section SMC Advice; "Summary of clinical effectiveness issues" section of SMC Advice; CVZ: CFH-rapport , section 2a; Farmacotherapeutisch rapport, section 3, 4a-f HAS: Advice from the Commission de la Transparence, section 3 (3.1 – 3.3)
6	Number of observational studies considered in guidance	Count	Number of observational studies providing information to support study drug. Observational studies are defined as non-interventional (i.e. do not explicitly request the patient to take particular medication or the physician to follow particular protocol).	
7	Priority disease area	Categorical – yes/no	This variable captures whether the pharmaceutical in question is linked to a disease area that is prioritized by the Ministry of Health. Priority disease areas were identified by examining government plans/health documents that highlight national health care system focus.	NICE: Department of Health (DH) (2002), DH (2006a), DH (2006b); DH (2007) SMC: NHS Health Scotland (2004, 2005, 2006, 2007, 2008) CVZ: Ministerie van Volksgezondheid, Welzijn en Sport HAS: CNRS (2004)
8	Orphan Designated	Categorical	This variable captured information on whether or not	European Medicines Agency [accessed 2011]

		– yes/no	the technology was recognized by the European Medicines Agency (EMA) as an orphan designated medicine.	
9	Therapeutic Indication	Categorical - 12 categories	The British National Formulary (BNF) categories were used to classify each technology into the corresponding therapeutic area.	British National Formulary (2010)
10	Size of eligible population for treatment	Numeric	Reported number of patients eligible for treatment, as per the Summary Product Characteristics and indication of the medication under evaluation is indicated.	NICE: TAR section 2 and/or 5 SMC: SMC Advice, " Additional information: budget impact" section; CVZ: CFH-rapport , section 2a Farmacotherapeutisch rapport, section 3, 4a-f HAS: Advice from the Commission de la Transparence, section 4.4
11	Availability of alternative therapies in current treatment setting.	Categorical – yes/no	An alternative was considered to be available if comparators were clearly defined in the review by the HTA agency. An alternative was considered NOT available if it was stated as such in the appraisal, or if 'best supportive care' or 'palliative care' was specified as the comparator.	NICE: TAR section 2 SMC: SMC Advice " Additional information: comparators" section CVZ: CFH-rapport , section 2a; Farmacotherapeutisch rapport, section 3, 4a-f HAS: Advice from the Commission de la Transparence, sections 2.2 and 2.3, and 4.3
Evidence – economic variables				
12	Consideration of Cost Utility Analysis in guidance	Categorical – CUA performed or no CUA	Presence or absence of a cost-utility analysis.	NICE: TAR section 4/ 5 SMC: SMC Advice " Summary of comparative health economic evidence" section CVZ: Farmaco-economisch rapport and Vraagstelling doelmatigheidstoets; HAS: Secretariat General de la Commission de la Transparence (2005)
13	Incremental Cost-utility ratio of technology vs. comparator in base case	Numeric	ICER (cost per QALY) included in the report for base case as accepted by the HTA body. This is defined as the ICER that is related to the recommendation. If more than one ICER is presented (as the recommendation covers more than one population) then the ICER pertaining to the larger of the populations was reported. If technology is reported as dominant or dominated, it was recorded as such on the data extraction sheet. This variable was not applicable to the HAS.	NICE: TAR section 4/ 5 SMC: SMC Advice " Summary of comparative health economic evidence" section CVZ: Farmaco-economisch rapport and Vraagstelling doelmatigheidstoets; HAS: N/A

14 - 15	Multiple CUA/CEA models reported	Categorical - Yes/No; If yes – provide range	Variable 14 reports whether more than one cost-utility or cost-effectiveness model was considered during the appraisal (yes / no). If yes, variable 15 measures the range of base-case ICERs presented between the different models reported. The difference between the lowest and highest ICER was calculated. This variable was not applicable to the HAS.	
16	Uncertainty around the base case ICER reported in submission (probabilistic)	Numeric	Measures the percentage probability of acceptance at the threshold used by the agency. For the CVZ the probability of the medication being cost-effective was reported at a EUR €50,000 threshold; for NICE and SMC, a threshold of 20,000GBP was used. This variable was not applicable to the HAS.	
17	Uncertainty around base case ICER reported in submission (univariate)	Numeric	Measured the range of ICERs (min-max) resulting from univariate sensitivity on the base case. This variable was not applicable to the HAS.	NICE: TAR section 4/ 5 SMC: SMC Advice " Summary of comparative health economic evidence" section CVZ: Farmaco-economisch rapport and Vraagstelling doelmatigheidstoets; HAS: N/A
18	Anticipated budgetary impact of introduction of new technology in health care system	Numeric	Estimated annual budgetary impact of introducing new medication into the current treatment setting, if the pharmaceutical were to be introduced without any restriction. Drug cost only (per year). This variable was not applicable to the HAS.	NICE: TAR section 4/ 5 SMC: SMC Advice, " Additional information: budget impact" section CVZ: Kostenprognose Rapport HAS: N/A
Decision-making process variables				
19	Inclusion of patient submission	Categorical – yes/no	A patient submission was considered to have been included as part of the appraisal process if a submission from a patient group was posted on the webpage pertaining to the guidance.	NICE: NICE (2011) section describing the history of the appraisal SMC: SMC Advice, " Summary of patient and public involvement" section CVZ: College voor zorgverzekeringen: Cvz-criteria voor beoordeling therapeutische waarde, 2010 HAS: Secretariat General de la Commission de la Transparence (2005)
20	No. of decision makers accountable	Numeric	Captures the number of committee members accountable for the guidance issued, as reported in the HTA process guidelines.	NICE: TAR Appendix B of each guidance SMC: SMC Meeting Minutes, http://www.scottishmedicines.org.uk/About_SMC/Minutes/Minutes

				CVZ: College voor zorgverzekeringen: Cvz-criteria voor beoordeling therapeutische waarde, 2010 HAS: Secretariat General de la Commission de la Transparence (2005)
21	Cost-effectiveness evaluation component in process	Categorical – yes/no	Captures whether or not cost-effectiveness is a component of the decision-making process. If cost-effectiveness analysis is a formal part of the appraisal process, this variable was marked as 'yes'.	NICE: TAR / NICE (2008b) SMC: SMC Advice, SMC (2010b) CVZ: College voor zorgverzekeringen: Cvz-criteria voor beoordeling therapeutische waarde, 2010 HAS: Secretariat General de la Commission de la Transparence (2005) ; Sorenson et al. (2008)
22	Budget impact evaluation component in process	Categorical – yes/no	Captures whether budget impact considerations are part of decision-making process	
23	Price Known during appraisal	Categorical – yes/no	Captures whether the price of the technology under appraisal was known during the assessment	NICE: TAR section 3 SMC: SMC Advice, SMC (2010b) CVZ: College voor zorgverzekeringen: Cvz-criteria voor beoordeling therapeutische waarde, 2010 HAS: Secretariat General de la Commission de la Transparence (2005)
24	Number of drugs appraised in same appraisal	Count	This variable captures the number of technologies appraised simultaneously in the appraisal	NICE: TAR cover page SMC: SMC Advice, SMC (2010b) CVZ: College voor zorgverzekeringen: Cvz-criteria voor beoordeling therapeutische waarde, 2010; CFH-rapport , section 2a; HAS: Secretariat General de la Commission de la Transparence (2005)
Socio-economic and system context variables				
25	Date guidance was issued	Numeric	Year when coverage decision was issued (11)	NICE: TAR cover page SMC: Decision cover page CVZ: Letter from the minister van Volksgezondheid, Welzijn en Sport HAS: Cover Page of Avis from the Commission de la Transparence
26	Healthcare exp as % GDP	Numeric (%)	Percentage of GDP spent on healthcare, during year of decision	OECD Health Data 2009, OECD 2010

27	Healthcare expenditure on pharmaceuticals	Numeric (€)	Healthcare budget spent on pharmaceuticals per patient per year, during the same year in which the appraisal was published. (12, 13)	NICE: Association of the British Pharmaceutical Industry (2010) SMC: ISD (2009) CVZ: Centraal Bureau voor de Statistiek (CBS)Statline ; GIPdatabank , HAS: Econ-Sante` France (2010) ; IRDES (2009)
28	National Population Size	Numeric	Estimate of population size within remit of the agency performing the evaluation.	NICE: National Office of Statistics (2009) SMC: National Office of Statistics (2009) CVZ: Centraal Bureau voor de Statistiek (CBS)Statline HAS: Eurostat (2010)
29	Election year at time of decision	Categoric al – yes/no	This variable captures whether the payer decision was made within an election year. An election year was defined as a year in which either national government or regional elections took place. (13)	NICE: BBC 2005 SMC: BBC 2005, BBC 2007 CVZ: Todosijevic et al HAS: Le ministre de l'intérieur, de l'outre-mer, des collectivités territoriales et de l'immigration (2004) and (2007)

Legend: RCT=Randomised Controlled Trial; HTA= Health Technology Assessment; CUA= Cost Utility Analysis; CEA=Cost Effectiveness Analysis; NICE –National Institute for Health and Care Excellence (NICE); Scottish Medicines Consortium (SMC); College voor Zorgverzekeringen (CVZ); Haute Autorité de Santé (HAS).

Table II Descriptive Statistics of pooled sample of coverage decisions from NICE, SMC, CVZ and HAS (n= 977) by HTA coverage decision (recommended, restricted or not recommended) (n=977)

	Recommended			Restricted			Not Recommended			P value
Variable	**Mean**	**95% CI**		**Mean**	**95% CI**		**Mean**	**95% CI**		
Clinical, Disease Variables										
Number of RCTs considered in decision	2.7	2.1	3.3	3.4	2.9	3.8	2.6	2.3	2.9	<0.05
Size of population included in RCTs	967	687	1247	1251	914	1587	830	607	1053	NS
Length/extent of follow-up in RCT (weeks)	58.4	50.4	66.3	49.1	42.9	55.4	41.4	34.6	48.2	<0.01
Statistically Significant superior results										
Superior	49%	43%	55%	42%	37%	47%	39%	33%	44%	<0.05
Not superior	15%	11%	20%	16%	13%	20%	18%	14%	23%	<0.01
Inconsistent results	20%	15%	25%	32%	27%	37%	29%	24%	34%	NS
Use of active comparator	52%	46%	58%	45%	40%	49%	43%	38%	48%	NS
Number of observational studies considered in guidance	0.6	0.3	0.8	0.3	0.2	0.4	1.2	-0.3	2.7	<0.05
Priority disease area	64%	58%	70%	62%	57%	67%	63%	58%	68%	NS
Orphan Designated	11%	7%	15%	9%	6%	12%	8%	5%	11%	NS
Therapeutic Indication										
cardiovascular system	10%	6%	13%	12%	9%	15%	11%	7%	14%	NS
central nervous system	12%	8%	15%	19%	15%	23%	23%	19%	28%	<0.01
ear, nose and oropharynx	1%	0%	3%	0%	0%	0%	0%	0%	1%	<0.05
endocrine system	4%	2%	7%	7%	5%	10%	6%	4%	9%	NS
Eye	3%	1%	5%	1%	0%	2%	1%	0%	2%	<0.05
gastro-intestinal system	3%	1%	6%	2%	1%	4%	6%	3%	9%	<0.05
Infections	9%	6%	12%	12%	9%	16%	10%	7%	13%	NS
malignant disease and immunosuppression	34%	29%	40%	22%	18%	26%	21%	17%	26%	<0.01
musculoskeletal and joint diseases	10%	7%	14%	12%	9%	15%	5%	2%	7%	<0.01
nutrition and blood	6%	3%	9%	4%	2%	6%	5%	3%	8%	NS
obstetrics, gynaecology, and urinary-tract disorders	0%	0%	1%	2%	1%	4%	1%	0%	2%	<0.05
respiratory system	1%	0%	2%	1%	0%	2%	6%	3%	9%	<0.01
Skin	4%	2%	7%	5%	3%	7%	4%	2%	6%	NS
Size of eligible patient population	105,118	44,430	165,805	746,591	363,878	1,129,304	561,944	326,147	797,741	<0.05
Availability of alternative therapies in	82%	77%	86%	86%	83%	90%	85%	81%	89%	NS

Variable	Recommended Mean	95% CI		Restricted Mean	95% CI		Not Recommended Mean	95% CI		P value
current treatment setting										
Economic Variables										
Consideration of Cost Utility Analysis in guidance	31%	26%	37%	39%	34%	44%	37%	31%	42%	NS
Incremental Cost-effectiveness ratio of technology vs. comparator in base case	£20,151	£12,203	£28,099	£25,358	£18,523	£32,193	£50,253	£29,179	£71,326	<0.05
Multiple CUA/CEA models reported	12%	7%	16%	17%	12%	21%	6%	3%	10%	<0.01
Lowest reported ICER	£16,256	£4,175	£28,336	£13,141	£9,888	£16,393	£20,201	£10,459	£29,944	<0.01
Highest reported ICER	£95,652	£6,764	£184,540	£125,859	£82,185	£169,533	£69,600	£28,140	£111,059	<0.01
Uncertainty around the base case ICER reported in submission (probabilistic)	65%	51%	78%	46%	36%	57%	23%	11%	35%	<0.05
Uncertainty around base case ICER reported in submission (univariate) Low	£11,278	£6,677	£15,879	£14,881	£6,640	£23,123	£59,904	£17,958	£101,849	<0.05
Uncertainty around base case ICER reported in submission (univariate) High	£105,042	£29,268	£180,816	£42,049	£28,653	£55,445	£242,362	£75,624	£409,101	<0.05
Anticipated budgetary impact of introduction of new technology in health care system (million £)	£10.9	£6	£15.3	£274.6	£25.4	£523.8	£163.4	-£40	£366.7	<0.05
Decision-Making Process variables										
Inclusion of patient submission	17%	12%	22%	27%	22%	32%	22%	18%	27%	<0.05
Number of Decision Makers Accountable	24	23	25	27	26	27	27	27	28	<0.01
Cost-effectiveness evaluation component in process	64%	58%	70%	60%	55%	65%	54%	49%	60%	<0.05
Budget impact as a component of decision-making process	81%	76%	86%	68%	63%	73%	57%	52%	62%	<0.01
Pricing known during appraisal	49%	43%	55%	23%	18%	27%	12%	9%	16%	<0.01
Number of drugs appraised in same appraisal	1.1	1.1	1.2	1.4	1.3	1.6	1.1	1.0	1.1	<0.05
Socio-economic context variables										
Date guidance was issued	2006	2006	2006	2006	2,006	2007	2007	2007	2007	<0.01
GDP-healthcare expenditure	10%	9%	10%	10%	9%	10%	10%	10%	10%	<0.01
Healthcare expenditure on	£263	£252	£274	£280	£269	£291	£306	£293	£320	<0.01

Variable	Recommended Mean	95% CI		Restricted Mean	95% CI		Not Recommended Mean	95% CI		P value
pharmaceuticals										
Population size – Agency coverage (million)	27.50	24.80	30.20	35.30	32.70	37.90	34.10	31.10	37.10	<0.01
Election year at time of decision	28%	22%	33%	31%	26%	35%	30%	25%	35%	NS

Legend: RCT=Randomised Controlled Trial; HTA= Health Technology Assessment; CUA= Cost Utility Analysis; CEA=Cost Effectiveness Analysis; NICE –National Institute for Health and Care Excellence (NICE); Scottish Medicines Consortium (SMC); College voor Zorgverzekeringen (CVZ); Haute Autorité de Santé (HAS). Recommended = pharmaceutical technology recommended for routine use; Restricted = pharmaceutical technology recommended for restricted use; Not Recommended = pharmaceutical technology not recommended for use

Table III Base case Model 1: includes all four HTA bodies (but no ICER, and including fixed effects for each HTA body) (n=977)

	Restricted vs. Recommended				Not Recommended vs. Recommended			
Variables	Log Odds	P value	95% Conf. Interval		Log Odds	P value	95% Conf. Interval	
Number of RCTs	-0.018	0.402	-0.061	0.025	-0.060	0.058	-0.122	0.002
RCT duration of follow-up	-0.001	0.453	-0.0044	0.00198	-0.003	0.112	-0.007	0.001
Use of active comparator in RCT	-0.535	0.011	-0.949	-0.121	-0.932	<0.001	-1.383	-0.481
Clinical superiority demonstrated in RCT	-0.433	0.023	-0.807	-0.059	-0.762	<0.001	-1.176	-0.349
Therapeutic Indication								
Central nervous system	0.213	0.464	-0.357	0.784	0.381	0.209	-0.213	0.976
Eye	-1.443	0.047	-2.866	-0.021	-1.314	0.075	-2.761	0.134
Malignancy/immunosuppresion therapy	-0.532	0.031	-1.016	-0.048	-0.457	0.087	-0.981	0.067
musculoskeletal and joint diseases	-0.200	0.536	-0.833	0.433	-0.958	0.017	-1.747	-0.169
obstetrics, gynaecology, and urinary-tract disorders	2.292	0.034	0.171	4.413	1.700	0.148	-0.603	4.003
Respiratory system	0.492	0.511	-0.974	1.958	1.747	0.012	0.378	3.116
Cardiovascular disease	0.067	0.828	-0.540	0.675	-0.290	0.39	-0.952	0.372
Skin	-0.011	0.980	-0.856	0.834	-0.302	0.529	-1.242	0.639
Orphan designation status	-0.311	0.294	-0.891	0.269	-0.749	0.023	-1.397	-0.101
Size of eligible patient population	5.00E-07	0.093	-0.00000134	0.00000010	0.000	0.633	0.000	0.000
Patient submission included	0.391	0.203	-0.211	0.993	0.833	0.008	0.215	1.452
Number of technologies appraised simultaneously	0.523	0.002	0.197	0.849	0.146	0.546	-0.328	0.620
National population size	0.0000005	0.024	0.00000007	0.0000009	0.000	<0.001	<0.001	<0.0001
NICE*	-18.526	0.027	-34.957	-2.096	-37.714	<0.001	-55.265	-20.163
SMC*	6.666	0.0080	1.761	11.571	13.270	<0.001	8.034	18.506
HAS*	-22.156	0.033	-42.550	-1.762	-44.951	<0.001	-66.727	-23.175
Constant	-8.412	0.022	-15.589	-1.235	-16.638	<0.001	-24.299	-8.977

Note: Recommended technologies are the reference case. Multinomial logistic regression, pseudo R-squared: 0.13. * Please note, CVZ is referent.

Table IV. Base-case Model 2: Multivariate analysis of multi-HTA sample of NICE and SMC HTA coverage decisions 2004-2009 (n=406)

	Restricted vs. Recommended				Not Recommended vs. Recommended			
Variables	Log Odds	P value	95% Conf. Interval		Log Odds	P value	95% Conf. Interval	
Number of RCTs	-0.035	0.198	-0.089	0.018	-0.103	0.052	-0.206	0.001
RCT duration of follow-up	-0.004	0.196	-0.009	0.002	-0.009	0.011	-0.015	-0.002
Use of active comparator in RCT	-0.708	0.050	-1.415	-0.001	-0.897	0.022	-1.662	-0.131
Clinical superiority demonstrated in RCT	-0.523	0.111	-1.166	0.120	-0.352	0.323	-1.050	0.346
Therapeutic Indication								
Central nervous system	0.882	0.319	-0.852	2.616	1.356	0.151	-0.493	3.205
Malignancy/immunosuppression therapy	0.906	0.295	-0.788	2.600	1.282	0.171	-0.553	3.117
Respiratory system	1.415	0.302	-1.271	4.102	2.172	0.118	-0.554	4.899
Cardiovascular disease	0.355	0.685	-1.359	2.070	0.453	0.638	-1.431	2.336
Endocrine system	0.443	0.647	-1.453	2.339	0.263	0.802	-1.790	2.316
Gastro-intestinal disorders	0.240	0.869	-2.615	3.095	2.001	0.141	-0.661	4.664
Infections	0.363	0.678	-1.348	2.074	-0.063	0.949	-1.976	1.851
musculoskeletal and joint diseases	0.348	0.721	-1.561	2.256	0.290	0.786	-1.806	2.386
Nutrition and blood	0.224	0.835	-1.879	2.327	1.033	0.351	-1.137	3.204
Skin	0.566	0.564	-1.358	2.490	-0.249	0.828	-2.498	1.999
Size of eligible patient population	-0.00000036	0.405	-0.0000012	0.0000005	-0.00000079	0.332	-0.0000024	0.0000008
ICER	0.000025	0.011	0.0000057	0.0000448	0.000033	0.001	0.000013	0.000053
Number of technologies appraised simultaneously	0.648	0.001	0.277	1.018	0.466	0.074	-0.044	0.977
National population size	0.00000094	0.087	-0.00000014	0.0000020	0.00000174	0.029	0.00000018	0.0000033
SMC	46.499	0.083	-6.060	99.059	86.641	0.027	10.061	163.222
Constant	-51.408	0.083	-109.603	6.786	-95.335	0.027	-180.057	-10.613

Untangling the Complexity of Funding Recommendations: A Comparative Analysis of Health Technology Assessment Outcomes in Four European Countries

Supplementary information

Karin H. Cerri,[1,2] Martin Knapp,[1] and Jose-Luis Fernandez[1]

3. London School of Economics and Political Science, London, England
4. Janssen Pharmaceutica NV, Beerse, Belgium

Corresponding author: Dr. Karin H. Cerri
London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK. k.h.cerri@lse.ac.uk

Part 1. Description of HTA bodies included in the analysis

NICE

Established in 1999, the National Institute for Health and Care Excellence (NICE) is responsible for providing guidance to the National Health Service (NHS) in England and Wales on the funding of new technologies and their use, including pharmaceutical technologies (1,2). One of the key rationales for setting up NICE was to help tackle the geographic inequality in access to technology, or the phenomenon commonly referred to as 'postcode prescribing'(3). Since 2002, NICE's recommendations have been mandatory and NHS organizations have had to comply, usually within three months.

The NICE HTA process involves a panel of clinical, academic, industry stakeholders and lay members (NICE 2008a). Appraisals by NICE are governed by the use of an established standard methodology for the evaluation of clinical and economic characteristics of a technology (4,5). As outlined in its methods guide(5), a range of clinical criteria are evaluated during a NICE appraisal, with the consideration of cost-utility evidence an integral part of the process. In addition to economic and clinical criteria, the patients' (and carers') perspective and patient (and carer) evidence, as well as submissions from other people working within the NHS, are taken into consideration via the consultation and stakeholder submission processes. It should be added that NICE operates a process that emphasises the role of social values in decision-making (4), especially during appraisal of clinical effectiveness and cost-effectiveness evidence (2).

SMC

Established in 2001, the Scottish Medicines Consortium (SMC) advises NHS Scotland on the use of newly licensed technologies or newly licensed indications for existing technologies. The purpose of the SMC is "to avoid duplication of new medicines assessment by individual ADTCs [NHS Board Area Drugs and Therapeutics Committees], to avoid geographical inequity in decision making and to make the best use of expertise available across Scotland" (6) p. 4). The composition of the Consortium and the New Drugs Committee (NDC) that advises the SMC on new technologies includes clinicians, pharmacists and health economists, in addition to stakeholders from manufacturers, patient groups and the government. The SMC provides guidance to NHS Scotland based on a rapid review soon after the marketing authorisation is obtained for the technology (6,7). The assessment of the

SMC is based on evidence submitted by the manufacturer, and there is no third party assessment within its appraisal process (8). A series of clinical and economic criteria are taken into account in the SMC's assessment of pharmaceutical technologies (9). Clinical effectiveness, relative to a comparator, is a key part of the SMC appraisal, which aims to assess the relevance of the efficacy and safety outcomes of the clinical studies to treatment practice and patients in Scotland. The pharmaco-economic assessment examines the economic model and results submitted by the manufacturer, including its relevance and robustness. Sensitivity analysis around the economic model results is an important component of the evaluation. In addition to pharmaco-economic assessments, resource implications associated with the introduction of the technology are assessed by the SMC. Submissions by patient groups are a formal part of the appraisal process.

CVZ

In the Netherlands, the College Voor Zorgverzekeringen (CVZ) has an important role in supporting and maintaining the quality, accessibility and affordability of health care through its HTA activities. The Commissie Farmaceutische Hulp (CFH) is a part of the CVZ. It is tasked with the assessment of the therapeutic value of new technologies for inclusion in the Medicine's Reimbursement System (Geneesmiddelen Vergoedings Systeem – GVS), or inclusion in the various reimbursement policies of the Dutch Health Authority (Nederlandse Zorgautoriteit - NZa). Collectively, the role of the CFH and CVZ is to provide advice to the Ministry of Health on the reimbursement list appropriate for the funding of each technology – the GVS 1A or 1B list, or specific reimbursement policies (e.g. the expensive drug list, or List '2')(10-12). The therapeutic value of the technology plays an important role in the CVZ recommendation (13). The CVZ takes into account several factors, including the disease for which the technology is indicated and whether or not there is a standard of care already reimbursed for that system with the same indication, based on clinical guidelines and clinical criteria. The review also collects information about the experience with the technology, such as time on the market, to assess the risk of unknown side effects occurring over time and to increase the certainty around the therapeutic benefit of the technology. The financial implications of adoption of the technology are also assessed, similar to budget impact estimation. The importance of each factor is weighted and technologies are compared, to come to a decision on the category of therapeutic benefit that should be applied (lower benefit, comparable benefit or higher benefit). The CVZ also has a particular role, perhaps less common in other HTA bodies, of reviewing technologies for unlicensed indications, on the request of health insurance bodies. In this situation, the CVZ is asked to establish if the unlicensed indication is rare (less than 1:150,000 population), if there is a scientific basis for the efficacy of the technology in this unlicensed indication and if there is no other alternative therapy available in the Netherlands for the condition under review. Once the CVZ has provided its advice to the Ministry of Health, and the Ministry of Health gives its formal approval, the reimbursement decision for extramural medicines is published on the Pharmaco-therapeutic Compass (Farmaco-therapeutisch Kompas)(10).

HAS

Since 2004, the Haute Autorité de Santé (HAS) in France has existed to reinforce the quality of treatment for the benefit of patients (14). The role of the HAS is currently undergoing change and a new law was passed in October 2012. The description of the HAS and relevant committees below reflects the objectives and processes of the system during the period when this analysis was performed (2004-2009). Through the Transparency Commission (Commission de la Transparence), it provides advice on the therapeutic value of technologies seeking reimbursement by the healthcare system. The Transparency Commission includes 31 members (15) and is composed of physicians, pharmacists and specialists in epidemiology and research methods. There are several criteria upon which the appraisal is based(16). The recommendations of the Commission, in 2004-2009, were based on an assessment of the medical service rendered by the technology ('Service Medical Rendu'- SMR), by taking into consideration various factors including the severity of the disease, the efficacy and safety profile

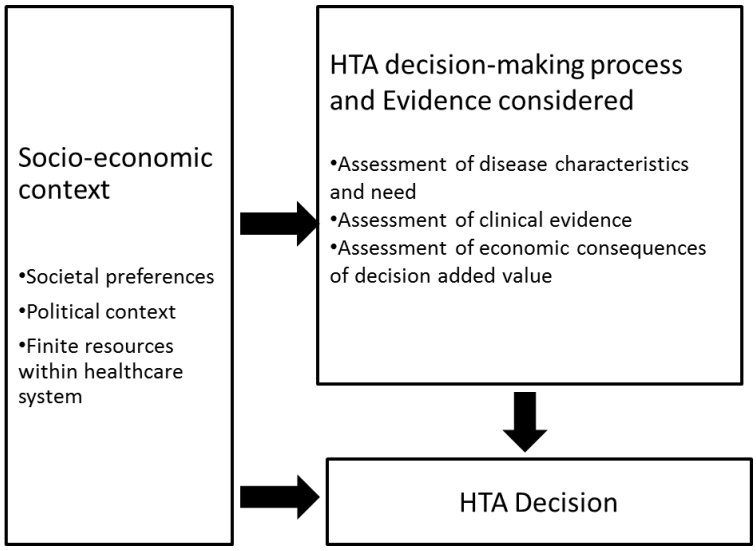
demonstrated, and the importance of the management of this disease from a public health standpoint. In addition, in 2004-2009, the Commission made a recommendation on the incremental medical service rendered (l'Amelioriation du Service Medical Rendu, -ASMR) defined by examining the efficacy and safety profile of the technology relative to a specified comparator. There are five levels of ASMR as defined in the Commission's regulations, ranging from ASMR I, which represents technologies that bring highly significant incremental medical value, to ASMR V, which represents technologies that show no incremental medical value. The Commission also considers the target population for the technology and, for example, the nature of the technology in terms of its duration of treatment and dosing. It is important to note that the cost of the technology is not considered during the appraisal process[3] and the price is not known during the appraisal by the Transparency Commission. The Commission's conclusion is transmitted to the Comité Economique des Produits de Santé (CEPS). This last committee is responsible for setting the price of medicines through negotiation with the manufacturers. These negotiations are based on several components, although they are primarily driven by the ASMR rating which represents the incremental benefit of the technology compared to the standard of care. Additional factors taken into consideration in price negotiations include the size of the target population, the manufacturer's research expenditure and advertising costs (11,17,18).

[3] In October 2012, a new law was passed (article 47 of the *loi de financement de la sécurité sociale*) which gives the HAS new competency in conducting medico-economic studies to evaluate health technologies. These medico-economic studies may be conducted for a limited number of products to inform the Comité économique des produits de santé (CEPS) in their decision-making process. In addition, cost-effectiveness analyses may be conducted once the technology has been in the market and its price and real-life utilization are known, possibly jointly with post-reimbursement studies.

Part 2. Selection of Explanatory Variables

Upon examination of the literature, three "streams" of research on factors impacting on HTA decision-making became apparent. Firstly, research on the impact of evidence on the HTA decision, which focuses on the degree to which evidence related to the medicine under review (whether clinical, economic or otherwise) can impact on HTA coverage decisions (19-23). Secondly, research on the decision-making process itself, rather than on the technology (e.g. whether economic evaluation is a component of the decision-making process or not). The literature examining the HTA appraisal process provides insights into a number of process-related factors that can potentially influence coverage decisions (23-26). Thirdly, reference in the literature is made to the impact of the overall healthcare and welfare characteristics on HTA decision-making (e.g. the impact of healthcare expenditure levels and the health policy priorities being reinforced by the Ministry of Health at the time of HTA decision-making). The literature has shown that coverage decisions are influenced by macro factors, such as healthcare spending per capita, societal willingness to pay, the structure of the healthcare system, as well as ethical and social considerations (27-31). Given these themes in the literature, it was hypothesised that HTA decisions were driven by the HTA decision-making process, the evidence considered within that process, and by the socio-economic and political context in which the decision was made (Fig. 1)

Figure 1. Health technology assessment (HTA) decision-making: hypothesised drivers



Multinomial logistic regressions were used to assess the hypothesized influence of evidence, process and contextual variables on the different coverage outcomes from the HTA bodies (recommended, restricted or not recommended). The underlying relationship between the HTA outcomes and the 3 set of influential variables was thus modelled as:

$$Log\left(\frac{P(Y=j|E_L,P_M,C_N)}{1-P(Y=j|E_L,P_M,C_N)}\right) = \alpha_j + \sum_{l=1}^{L}\varepsilon_{jl}E_l + \sum_{m=1}^{M}\pi_{jm}P_m + \sum_{n=1}^{N}\gamma_{jn}C_n$$

Which results in:

$$P(Y=j|E_L,P_M,C_N) = \frac{e^{X_j}}{1+e^{X_j}}$$

With $X_j = \alpha_j + \sum_{l=1}^{L}\varepsilon_{jl}E_l + \sum_{m=1}^{M}\pi_{jm}P_m + \sum_{n=1}^{N}\gamma_{jn}C_n$

And where:

$P(Y=j\|E_L,P_M,C_N)$	is the probability of belonging to group j given E_L,P_M,C_N
$E_1,\dots,E_L$	is the vector E of L explanatory EVIDENCE variables,
$P_1,\dots,P_M$	is the vector P of M explanatory PROCESS variables,
$C_1,\dots,C_N$	is the vector C of N explanatory CONTEXT variables,
$\varepsilon_{j1},\dots,\varepsilon_{jL}$	is the vector ε of the L coefficients corresponding to the L explanatory EVIDENCE variables
$\pi_{j1},\dots,\pi_{jM}$	is the vector π of the M coefficients corresponding to M explanatory PROCESS variables
$\gamma_{j1},\dots,\gamma_{jN}$	is the vector γ of the N coefficients corresponding to N explanatory CONTEXT variables,
α_j	is the constant.

Significant efforts were made to create a dataset that was consistent and comprehensive, and able to address the research questions. It was a specific objective of this research to ensure that a comprehensive set of common variables was collected across the HTA bodies to avoid model misspecification. From the general framework characterized by the equations presented above, overall research objectives and HTA-specific objectives were derived to reflect the understanding of the context and healthcare system within which each HTA body operates. A common set of variables, complemented by HTA–specific variables where pertinent for the research objectives, were

subsequently selected to enable the measurement of the role of evidence, process and context in HTA decision-making, illustrated by E_L , P_M and C_N above.

Part 3. Standardising the Outcome Variable

To conduct the analysis, it was necessary to standardise the way the outcome variable was defined. Each agency has its own method for defining and thinking about HTA decisions.

In order to be able to facilitate the analysis of the factors driving decision-making across HTA bodies, it was necessary to standardise the way the outcome variable was defined. Each agency has its own method for defining and thinking about coverage decisions. However, there are similarities in the types of coverage decisions made, which have been capitalised on to arrive at a series of 'rules' on how to define and classify coverage decisions by NICE, SMC, CVZ and HAS to allow for comparison and pooling (Table 1).

The analysis is based on using a three category outcome variable where the new technology can be:

- recommended for routine use
- recommended for restricted use

or

- not recommended

HAS represents a specific challenge, in that the ASMR rating reflects the incremental value associated with a technology and impacts on the willingness of the healthcare system to approve increased funding for technologies that achieve high levels of ASMR (I or II) or restrict funding for those technologies with a low ASMR (V). The ASMR does not represent the final funding decision; rather, a separate committee, the CEPS, has this responsibility as it is the entity responsible for finalising the price and volume agreements for technologies. While recognising that the ASMR represents a different form of reimbursement decision than the coverage decisions made by SMC, NICE and CVZ, to allow for pooling a classification of ASMR ratings was adopted (described in Table A). This was felt to be appropriate given that the price of the technology, and hence the funding of the technology, is primarily driven by the ASMR rating.

Recommended technologies were defined as those technologies where full coverage was granted for the totality of the licensed population. For NICE guidance, where a recommendation was made for a technology to be used in a population identical to its licensed indication, it was considered 'recommended'. For the CVZ, where the decision was to place the technology in the 'basis paket' (List 1A or 1B), or listed in the expensive drug list (Duregeneesmiddelen Beleidsregel) without any restriction or patient co-payment, the technology was considered to be recommended[4]. In France, the ASMR was used to classify outcomes. Recommended technologies in this

[4] It should be noted that category 1A within the Dutch decision-making process presented some ambiguity in terms of where it should be considered within this analysis. This is because while for products listed in 1A there is a price pre-defined based on already existing reference products, it is the manufacturer who in principle proposes in which list (1A or 1B) they would like to submit their technology, according to the therapeutic differentiation of the product –thus the manufacturer submits a proposal with a price in the anticipated range, and it is not the CVZ who asks for a price reduction to meet the reference price rule as a result of the appraisal but rather a position taken by the manufacturer at the start of the appraisal process. In addition, both 1A and 1B result in similar coverage from the patient perspective. This is why we decided to consider both 1A and 1B listings recommendations by the CVZ.

analysis were considered to be those with an ASMR I- II, representing technologies with important incremental medical value relative to standard of care.

Restricted technologies were defined as those technologies where coverage was granted for a sub-population of the licensed indication and/or with restrictions in terms of acquisition cost or utilization (e.g. monitoring or specialist use required) (Raftery 2006). For NICE and the SMC, a coverage decision was considered to be a restriction if it was recommended for use in a sub-population of its licensed indication; in a second line or higher line of therapy; required monitoring, recommended upon it bearing the lowest acquisition cost or prescription by a specialist. For the CVZ, where the decision was to place the technology in the 'basis paket', but only for use in a sub-population or with a patient co-payment, this technology was considered as restricted. With regard to HAS, technologies with ASMR III-IV representing modest or minor medical value are associated with lower price levels (and hence funding) than technologies with ASMR I-II. These were considered to be restricted technologies.

Not recommended technologies were those for which no coverage was granted. A medication was considered to be not recommended for use by NICE or SMC guidance if the words "not recommended" were stated in the guidance/report. Within CVZ decisions, the technology was considered to be not recommended when it was stated to be not recommended in the CVZ 'advies' statement and was not included in any reimbursement list. With regard to HAS, technologies with an ASMR V offer no incremental benefit versus the comparators, and as per legislation cannot be included on the reimbursement list. An ASMR V was therefore considered to be 'not recommended'. It should be noted however, that for HAS, technologies with an ASMR V can obtain reimbursement from the healthcare system, but only if associated with cost-savings. This is different from the CVZ where technologies not recommended for funding are excluded from the reimbursement list. Other research conducted on the ASMR has made a similar assumption whereby technologies with ASMR V were considered to represent non-recommendation (33). While for the purposes of this analysis it was felt appropriate to consider ASMR V as a non-recommendation, it is important to bear in mind that the implications of an ASMR V may not be the same as the implications of non recommendation from the SMC or NICE, for example.

While recognising that the ASMR represents a different form of HTA decision than the decisions made by SMC, NICE and CVZ, given that the price of the technology – and hence the funding of the technology – is primarily driven by the ASMR rating, it was considered appropriate to pursue a multi-HTA analysis by classifying HTA decisions by NICE, SMC, CVZ and HAS (Table 1).

Table 1 Classification of HTA decisions into a 3-category outcome variable: definitions per HTA body

HTA body	Recommended	Restricted	Not Recommended
NICE	Full HTA was granted for the totality of the licensed population	A sub-population of the licensed indication and/or with restrictions in terms of acquisition cost or utilization (e.g. monitoring or specialist use required)	"Not recommended" stated in section 1 of guidance
SMC	If word "recommended" used in summary statement	If word "restricted" was used in summary statement	If words "not recommended" were used in summary statement
CVZ	If technology was placed in	If technology placed in list 2, or	If the technology was not

	reimbursement lists 1A or 1B, or the expensive drug list	if patient co-payment is necessary to access medication	included in any reimbursement list
HAS	ASMR I-II	ASMR III-IV	ASMR V

ASMR = l'Amelioriation du Service Medical Rendu rating; CVZ = College voor Zorgverzekeringen; HAS = Haute Autorité de Santé; HTA= Health Technology Assessment; NICE=National Institute for Health and Care Excellence; SMC=Scottish Medicines Consortium

Part 4. List of Common Set of Technologies Appraised by All Four HTA bodies

Drugs reviewed by all 4 bodies	NICE	SMC	CVZ	HAS	Total # of Reviews
abatacept	1	1	1	1	4
adalimumab	1	4	7	6	18
bevacizumab	1	3	2	5	11
bortezomib	1	2	3	4	10
cetuximab	1	3	2	2	8
cinacalcet	1	2	1	2	6
clopidogrel	1	2	4	1	8
dabigatran	1	1	1	1	4
entecavir	1	1	1	1	4
etanercept	1	2	5	7	15
infliximab	1	3	4	4	12
lenalidomide	1	1	1	1	4
levetiracetam	1	2	2	1	6
memantine	1	1	2	1	5
natalizumab	1	2	1	1	5
pegaptanib	1	1	1	2	5
pemetrexed	1	5	2	3	11
rituximab	1	5	3	4	13
rivaroxaban	1	1	2	1	5
sunitinib	1	3	3	3	10
tacrolimus	1	1	1	3	6
telbivudine	1	1	1	1	4
teriparatide	1	1	1	4	7
topotecan	1	2	1	2	6
trastuzumab	1	1	1	2	5
Total appraisals					192

CVZ = College voor Zorgverzekeringen; HAS = Haute Autorité de Santé; HTA= Health Technology Assessment; NICE=National Institute for Health and Care Excellence; SMC=Scottish Medicines Consortium

Part 5. SUPPLEMENTARY TABLE

Table 2 Descriptive statistics - comparison of four HTA bodies included in multi-HTA analysis

	NICE (n=118)	SMC (n=288)	CVZ (n=256)	HAS (n=315)	P value	Test
	mean (95% CI)	mean (95% CI)	mean (95% CI)	mean (95% CI)		
NoRCT	6.7 (5.2,8.4)	2.2 (1.9,2.5)	2.6 (2.3,3)	2.3 (2.1,2.6)	<0.01	1
RCTsize	1249 (807,1691)	991 (689,1294)	830 (494,1165)	1154 (824,1484)	<0.01	3
RCTsignificant- yes	39% (29%,47%)	58% (48%,60%)	20% (33%,45%)	47% (33%,43%)	<0.01	2
No	16% (9%,22%)	19% (13%,22%)	46% (13%,22%)	20% (12%,20%)	<0.01	2
Inconsistent	45% (34%,52%)	23% (16%,26%)	34% (23%,34%)	33% (21%,31%)	NS	2
RCTduration	76.2 (63.5,88.9)	4490% (3840%,5140%)	3950% (3340%,4560%)	4930% (4120%,5740%)	<0.01	1
RCTcomparator	47% (39%,55%)	52% (46%,57%)	44% (38%,51%)	42% (37%,48%)	<0.01	3
ObsStudies	0.6 (0.1,1.1)	1.3 (-0.4,2.9)	0.6 (0.4,0.7)	0.3 (0.2,0.4)	<0.01	1
CUA	95% (91%,99%)	74% (69%,79%)	11% (7%,15%)	0% (0%,0%)	<0.01	2
ICER	£31,266 (£29,122,£42,410)	£34,055 (£21,630,£46,481)	£30,977 (£8,643,£53,312)	-	<0.01	3
MultipleCEA	63% (54%,72%)	1% (0%,2%)	1% (0%,2%)	-	<0.01	1
MultipleICERs – low range	£13,260 (£10,409,£16,110)	£10,399 (-£9,782,£30,580)	£85,091 (-£198,747,£368,928)	-	<0.10	3
MultipleICERs – low range	£107,421 (£66,886,£147,956)	£18,207 (-£2,672,£39,086)	£221,499 (£201,610,£241,389)	-	<0.05	3
univariateICER – low value	0.43 (0.34,0.52)	0.57 (0.42,0.73)	0.66 (0.42,0.9)	-	<0.10	3
univariateICER –high value	£25,417 (£6,412,£44,422)	£33,277 (£8,916,£57,637)	£15,187 (£4,338,£26,036)	-	<0.05	3
probabilisticICER	£167,389 (£56,865,£277,913)	£77,927 (£19,847,£43,448)	£92,826 (-£6,356,£192,008)	-	<0.01	3
NonCUA	23% (15%,30%)	30% (25%,36%)	15% (11%,20%)	-	<0.01	2
BudgetImp	£701.3 (£179.7,£1,223.0)	£1.2 (£0.9,£1.5)	£31.0 (£5.9,£56.2)	-		

Eligiblepop	2,418,119 (1,256,514,3,579,724)	11,277 (1,647,20,908)	94,543 (31,394,157,693)	511,047 (314,122,707,972)	<0.01	1
AlternativeTx	89% (83%,95%)	83% (79%,88%)	79% (74%,84%)	89% (85%,92%)	<0.01	2
PatientSub	0.87 (0.81,0.93)	0.42 (0.36,0.48)	0.04 (0.02,0.06)	0 (0%,0%)	<0.01	2
NoDecMakers	30 (28,32)	25 (24,25)	20 (20,20)	31 (31,31)	<0.01	1
CEAprocess	100% (100%,100%)	100% (100%,100%)	67% (61%,73%)	0% (0%,0%)	<0.01	2
BudgetImpProcess	100% (100%,100%)	100% (100%,100%)	100% (100%,100%)	0% (0%,0%)	<0.01	2
Price	100% (100%,100%)	100% (100%,100%)	100% (100%,100%)	0% (0%,0%)	<0.01	2
NoTech	2.8 (2.5,3.1)	1 (1,1)	1 (1,1)	1 (1,1)	<0.01	1
Accountability	0% (0%,0%)	0% (0%,0%)	0% (%,0%)	0% (0%,0%)	n/a	2
Independence	100% (100%,100%)	0% (0%,0^%)	100% (100%,100%)	100% (100%,100%)	<0.01	2
Date	2007(2006,2007)	2006(2006,2006)	2006(2006,2006)	2007(2006,2007)	<0.01	1
PopulationSize (million)	53.9(53.8,54)	5.13(5.12,5.13)	16.3(16.3,16.4)	63.4(63.3,63.5)	<0.01	3
healthcareGDP	8% (8%,8%)	8% (8%,8%)	10% (10%,10%)	11% (11%,11%)	<0.01	1
HealthExp	£175(£173,£176)	£190(£189,£190)	£249(£246,£251)	£439(£437,£441)	<0.01	3
Election	7% (2%,11%)	40% (35%,45%)	20% (15%,24%)	30% (25%,36%)	<0.01	2
Priority	56% (47%,65%)	66% (61%,71%)	55% (49%,61%)	70% (65%,75%)	<0.01	2
OrphanDesig	3% (0%,5%)	11% (8%,15%)	9% (5%,12%)	9% (6%,12%)	<0.05	2
Cardiovascular system	10% (5%,16%)	11% (7%,14%)	9% (5%,12%)	14% (10%,18%)	NS	2
Central nervous system	15% (9%,22%)	22% (16%,26%)	16% (12%,21%)	18% (14%,23%)	NS	2
Ear, nose and oropharynx	0% (0%,0%)	0% (0%,1%)	1% (0%,2%)	0% (0%,1%)	NS	2
Endocrine system	1% (-1%,3%)	9% (5%,12%)	6% (3%,9%)	6% (4%,9%)	<0.05	2
Eye	2% (-1%,4%)	1% (0%,2%)	2% (0%,3%)	2% (0%,3%)	NS	2
Gastro-intestinal system	2% (-1%,4%)	4% (2%,6%)	5% (2%,7%)	3% (1%,6%)	NS	2
Infections	12% (6%,18%)	10% (7%,14%)	9% (5%,12%)	12% (8%,15%)	NS	2
Malignant disease and immunosuppression	31% (22%,39%)	23% (20%,30%)	25% (19%,30%)	24% (19%,29%)	NS	2

Musculoskeletal and joint diseases	19% (12%,27%)	3% (1%,6%)	10% (6%,14%)	9% (6%,12%)	<0.01	2
Nutrition and blood	3% (0%,5%)	6% (3%,9%)	7% (4%,10%)	3% (1%,6%)	NS	2
Obstetrics, gynaecology, and urinary-tract disorders	0% (0%,0%)	2% (0%,3%)	3% (1%,5%)	0% (0%,1%)	<0.05	2
Respiratory system	1% (-1%,3%)	3% (2%,6%)	4% (1%,6%)	2% (1%,4%)	NS	2
Skin	5% (0.01,0.09)	3% (2%,6%)	5% (2%,7%)	4% (2%,7%)	NS	2

CEA=Cost Effectiveness Analysis; CUA= Cost Utility Analysis; CVZ = College voor Zorgverzekeringen; GDP=gross domestic product; HAS = Haute Autorité de Santé; HTA= Health Technology Assessment; ICER Incremental Cost-effectiveness Ratio; NICE=National Institute for Health and Care Excellence; QALY =Quality-Adjusted Life-Year; RCT=Randomised Controlled Trial; SMC=Scottish Medicines Consortium; TAR = Technical Appraisal Report.

References

(1) Chalkidou K. Comparative effectiveness review within the U.K.'s National Institute for Health and Clinical Excellence. Issue Brief (Commonw Fund) 2009 July;59:1-12.

(2) NICE. Social Value Judgements - Principles for the development of NICE guidance. [Second edition]. 2008.

(3) Summerhayes M, Catchpole P. Has NICE been nice to cancer? Eur J Cancer 2006 November;42(17):2881-6.

(4) Littlejohns P, Garner S, Doyle N, Macbeth F, Barnett D, Longson C. 10 years of NICE: still growing and still controversial. Lancet Oncol 2009 April;10(4):417-24.

(5) NICE. Guide to the methods of technology appraisal. 2008.

(6) Scottish Medicines Consortium. Introduction and availability of newly licensed medicines in the NHS in Scotland. 2010.

(7) Cairns J. Providing guidance to the NHS: The Scottish Medicines Consortium and the National Institute for Clinical Excellence compared. Health Policy 2006 April;76(2):134-43.

(8) Mason AR, Drummond MF. Public funding of new cancer drugs: Is NICE getting nastier? Eur J Cancer 2009 May;45(7):1188-92.

(9) Scottish Medicines Consortium. Guidance to Manufacturers for Completion of New Product Assessment Form (NPAF) . [Revised June 2010]. 2010.
(10) CVZ. Procedure beoordeling extramurale Geneesmiddelen. 2010.
(11) Sorenson C, Drummond M, Kanavos P. Ensuring value for Money in Health Care: the role of HTA in the European Union. 2008. Cornwall, UK., World Health Organization 2008, on behalf of the European Observatory on Health Systems and Policies.
(12) Stolk EA, Rutten FF. The "health benefit basket" in the Netherlands. Eur J Health Econ 2005 December;Suppl:53-7.
(13) College Voor Zorgverzekeringen. CVZ-criteria voor beoordeling therapeutische waarde. 2010.
(14) Haute Autorité de Santé. Décret no 2004-1398 du 23 décembre 2004 relatif aux médicaments remboursables par l'assurance maladie et modifiant le code de la sécurité sociale (deuxième partie : Décrets en Conseil d'Etat). 2004.
(15) Zentner A, Velasco-Garrido M, Busse R. Methods for the comparative evaluation of pharmaceuticals. GMS Health Technol Assess 2005;1:Doc09.
(16) Secretariat General de la Commission de la Transparence. Rapport d'activite de la commission de la transparence 2004. 2005.
(17) Bellanger MM, Cherilova V, Paris V. The "Health Benefit Basket" in France. Eur J Health Econ 2005 December;Suppl:24-9.
(18) Sandier S, Paris V, Polton D. Health Care Systems in Transition – France 2004. 2004. Copenhagen, Denmark, WHO Regional Office for Europe on behalf of the European Observatory on Health Systems and Policies.
(19) Clement FM, Harris A, Li JJ, Yong K, Lee KM, Manns BJ. Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia, and Canada. JAMA 2009 October 7;302(13):1437-43.
(20) Dakin HA, Devlin NJ, Odeyemi IA. "Yes", "No" or "Yes, but"? Multinomial modelling of NICE decision-making. Health Policy 2006 August;77(3):352-67.
(21) Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Econ 2004 May;13(5):437-52.
(24) Al MJ, Feenstra T, Brouwer WB. Decision makers' views on health care objectives and budget constraints: results from a pilot study. Health Policy 2004 October;70(1):33-48.
(25) Menon D, Stafinski T, Stuart G. Access to drugs for cancer: Does where you live matter? Can J Public Health 2005 November;96(6):454-8.
(26) Vuorenkoski L, Toiviainen H, Hemminki E. Decision-making in priority setting for medicines--a review of empirical studies. Health Policy 2008 April;86(1):1-9.
(27) OECD. Health Technologies and Decision Making. 2005. OECD.
(28) Packer C, Simpson S, Stevens A. International diffusion of new health technologies: a ten-country analysis of six health technologies. Int J Technol Assess Health Care 2006;22(4):419-28.
(29) Owens DK. Interpretation of cost-effectiveness analyses. J Gen Intern Med 1998 October;13(10):716-7.
(30) Buxton MJ. How much are health-care systems prepared to pay to produce a QALY? Eur J Health Econ 2005 December;6(4):285-7.
(31) Bryan S, Williams I, McIver S. Seeing the NICE side of cost-effectiveness analysis: a qualitative investigation of the use of CEA in NICE technology appraisals. Health Econ 2007 February;16(2):179-93.
(32) Ross J. The use of economic evaluation in health care: Australian decision makers' perceptions. Health Policy 1995 February;31(2):103-10.

(33) Raftery, J. Review of NICE's recommendations, 1999-2005. BMJ 332 (7552):1266-8, 2006.
(34) Kanavos P, Nicod E, van den Aardweg S, Pomedli S. The impact of health technology assessments: an international comparison. Euro Observer 4. 2010.