An empirical analysis of Drug Reimbursement Decisions in 10 European countries*

Laia Maynou[1,2,3,4], John Cairns[1, 5]

[1]*London School of Hygiene and Tropical Medicine, London, UK.*
[2] *Research Group on Statistics, Econometrics and Health (GRECS), University of Girona, Girona, Spain*
[3] *CIBER of Epidemiology and Public Health (CIBERESP), University of Girona, Girona, Spain*
[4] *Center for Research in Health and Economics (CRES), University Pompeu Fabra, Barcelona, Spain*
[5] *CCBIO, University of Bergen, Bergen, Norway*

Abstract

Most European countries have Health Technology Assessment procedures for informing drug reimbursement decisions. Depending on the drug and the country assessing it, the decision can be Favourable, Favourable with restrictions or Non-Favourable. The main objective of this paper is to determine empirically the factors that may lead to different drug reimbursement decisions across countries. For this purpose, a taxonomy has been developed, comprising three groups of variables: system-level, product-specific and time-dependent. Our goal is achieved through modelling a Hierarchical, Random-Effects Ordered Probit. This model is run on a database containing cancer drug reimbursement decisions of ten European countries (2006-2014). The main results show that a drug-indication with a NICE favourable decision is associated with a higher probability of adoption in another country. Furthermore, the probability of reimbursement is higher when a drug is considered cost-effective by NICE/SMC, when there is a financial Managed Entry Agreement and when fewer stakeholders are involved in the process. However, the requirement of economic evaluation/budget impact, an external review of evidence, the manufacturer being the initiator of the process and the price being based on a reference pricing is associated with a lower probability of reimbursement.

Key words: Drug reimbursement, cancer drugs, Health Technology Assessment (HTA), Hierarchical model, Ordered Probit.

*This research is funded under the European 7th Framework Programme with Advance-HTA (nº 305983). The results presented reflect the author's views. The EC is not liable for any use of the information communicated.

1. Introduction

Most European countries have Health Technology Assessment (HTA) procedures for informing drug reimbursement decisions. In the last stage of the process, the final decision can be Favourable, Favourable with restrictions or Non-Favourable. Existing literature shows that depending on the drug-indication and the country assessing it, the final reimbursement decision can differ across countries (for review see [1]). Even if the European countries have common objectives for HTA systems, the processes are not homogenous. The operative procedures and the organisations work differently across these countries. Previous literature and the findings of this research show that there are differences in the final decisions across countries.

But, why do the differences in drug reimbursement decisions across European countries matter? The clinical evidence they are reviewing is largely the same, and the countries while not of equal wealth are of broadly comparable levels of economic development. As a result, we might expect countries to reach broadly similar decisions (positive or negative) on drug reimbursement. However, this is not the case. So, why do these countries reach different conclusions? The differences may, in part, reflect the different HTA objectives and procedures across countries.

In a preliminary analysis [2], we tested a number of hypotheses that could explain the differences in cancer drug reimbursement decisions across ten European countries. While, the results showed that the HTA system characteristics, the drug particularities and the socioeconomic situation can explain some of the differences between countries, a fuller explanation requires a model which determines cancer drug reimbursement decisions incorporating a wide range of health system characteristics and specific characteristics of the individual drugs.

In particular, drug reimbursement procedures have attracted attention from several authors, due to the different systems that exist. Various comparative analyses have been published recently [3-7] describing the different national models in the world. Moreover, a number of descriptive and comparative studies have specifically analysed reimbursement decisions [8-11], however, the few empirical analyses are mainly focused on the UK or include few observations [12-17].

Our aim in this paper is to contribute to the empirical literature, determining the factors that might lead to different drug reimbursement decisions on cancer drugs in ten European countries (Belgium, England, France, Germany, Netherlands, Poland, Portugal, Scotland, Spain, and Sweden). This paper extends existing research both in terms of the methodology used and in the range of countries and decisions analysed. This paper has the following structure. Section 2 describes the taxonomy. In section 3, the database is defined. Section 4 explains the econometric approach to be applied. The results are presented in section 5. The main findings are discussed in the final section.

2. Taxonomy

A detailed analysis of the drug reimbursement systems in ten European countries was conducted. These countries were selected because they each have a well-defined HTA process and publicly available information on their drug reimbursement decisions. This first analysis involved: 1) a review of policy documents and relevant literature [18] and detailed examination of the study country decision-making bodies' websites 2) discussion with experts knowledgeable regarding the process in each of the countries. In this second process, we drew heavily on Advance-HTA[1] Consortium

[1] http://www.advance-hta.eu. ADVANCE-HTA, EU-funded project. The aim is to advance and strengthen the methodological tools and practices relating to the application and implementation of HTA. It is a partnership of 13 Consortium members led by the London School of Economics - LSE Health.

members representing some of the studied countries.

This analysis was essential to identify potentially important system-wide factors. This classification describes the main characteristics of the drug reimbursement system in each country and the main features of each drug-indication. Moreover, time dependent variables are also included in the classification to capture the socioeconomic situation of each country. This taxonomy can be classified into three groups, the system-wide variables (organisational, process and method), the product-specific variables (general characteristics and country-specific) and the time dependent variables. Table I defines the taxonomy variables and Table II categorises them. This taxonomy will be used as explanatory variables in the econometric model.

3. Database

Cancer drugs have been selected owing to the high level of public interest in these reimbursement decisions. Moreover, this area provides a rich database for the analysis of reimbursement decisions since many cancer drug-indication pairs have been appraised during the last decade. The drugs selected were classified under "malignant disease and immunosuppression" on the Scottish Medicines Consortium (SMC) website. SMC was the starting point of our study because it appraises all drugs approved by either the Medicine and Healthcare Product Regulatory Agency (MHRA) or the European Medicines Agency (EMA). The SMC list was validated, checking National Institute for Health and Care Excellence (NICE) decisions for any additional cancer drug-indications. After this process, the number of drug-indications was 161.

3.1. Sample

The sample includes the technology appraisals for cancer drugs from January 2006 to November 2014 appraised in the ten selected countries. Drugs appraised from 2006 onwards are included since, by this stage, many Europeans countries had introduced formal HTA systems. Moreover, it is the period when the European Network for Health Technology Assessment (EUnetHTA) started. The dataset contains the outcome of the decision, the date when the decision was published and all the variables defined in the taxonomy section. We considered 161 drug-indications per country, however, decisions made before 2006, non-assessed drug-indications and decisions when the date was missing, are not included (since there is no possibility of linking them with the time variables). The final sample contains 158 drug-indications.

The decision outcome describes the final decision regarding the adoption of the technology: Non-Favourable, Favourable with restrictions and Favourable. To distinguish between "Favourable with restrictions" and "Favourable", the decision is considered to be restricted only when it differs from the indication detailed in the marketing authorisation (e.g. when the indication is limited to a sub-population). However, it is not considered restricted when the recommendation refers to which doctors can prescribe or if it has a Managed Entry Agreement.

In order to capture all possible decisions, the decision variable has another category: Non-submission. This category collects the decisions where the reimbursement body asked the manufacturer to make a submission and it failed to do so. Under this category, there are only decisions from NICE and SMC, as the other countries do not document this information. Even though a Non-submission is considered as a Non-Favourable decision for NICE or SMC, it is classified in a different category because this negative decision is reached through a different process. As a result, the

non-submission category is not included in the econometric model. Exclusion of this category of "decision" from the analysis potentially introduces a sample selection problem. In addition, this might not be the only cause of sample selection, as there are drug-indications that have not been assessed by all countries and as a result they do not enter into the final model. However, robustness checks are performed to validate the approach.

Table III reports the data source for each country. For some countries, all decisions were publicly available through official websites, but for others, assistance was required from the National HTA Agencies or the Health Departments. Our database is formed of 792 decisions, 59% of them were favourable, while only 15% were rejected and 21% were restricted. The non-submission category only account for 5% of the total sample. Table IV disaggregates the information by country. France has assessed most cancer drug-indications, followed by Scotland and Belgium. Belgium, Poland and Scotland have the highest rates of restricted decisions. However, there are few decisions for Germany, Portugal, Netherlands and Sweden. For some countries, these results show some data availability problems (e.g. Germany only data from 2011; "Pharmaceuticals Market Reorganisation Act", AMNOG; and Netherlands).

The dataset also records all the variables defined in the taxonomy. Table V and VI report the descriptive statistics for these variables. Moreover, the annex collects the outcome of the system-wide variables for each country. Table V shows the results of the categorical variables for 792 decisions. For some variables, such as, type of patient, disease stage, Incremental Cost-Effectiveness Ratio (ICER), initiator, decision level, transparency, Managed Entry Agreement (MEA), some categories have few observations. This table helped us determine which variables/categories to omit from the model due to a lack of variation. Table VI shows the number of observations, the mean, the standard deviation, the minimum and maximum of the continuous variables. The variables, except for incidence rate, are time dependent and range from 2006 to 2014. Even if there are some missing mortality data, mortality was included in the econometric model in order to capture differences across countries.

3.2. Data assumptions

Owing largely to data limitations some assumptions were required in order to produce comparable data. For instance, the date of decision was not always available. For Spain, for some drug-indications, there was only the commercialisation date or the authorisation date. The date of decision was then approximated by computing the time difference between the date of authorisation and the date of decision for the drug-indications where this information was available. For Spain, the result was a median difference of seven months, which was added to the authorisation date for the drug-indications without a decision date. For France, a similar approximation involved adding five months to the recommendation date of the Transparency Committee in order to approximate the decision date.

The reimbursement process for some countries changed between January 2006 and November 2014. Consequently, for these countries, the same variable was recorded differently depending on the year considered (see the annex). In addition, when a later decision changed the outcome, the latest decision for a particular drug-indication was taken for the database. This criterion was taken because some countries make revaluations of previous decisions (e.g. England and France) or they allow for a resubmission after a negative decision (e.g. Scotland).

The taxonomy was designed to capture all relevant characteristics of the drug reimbursement system. However, the meaning of decision outcome may differ across countries. For instance, in Germany, a new technology enters the market directly after the marketing authorisation approval. However, this technology has to be assessed in a period of time, in

order to make a price decision. In Germany, negative or restricted decisions do not occur. In the price decision, if the new technology is regarded as adding value, the price is set following negotiation. Conversely, if it is not deemed to add value, the price is set using a reference pricing procedure. Whereas, in the French system, if the drug is adding value (ASMR[2] I-III), the price is set using reference pricing, while, it is based on a negotiation when the new technology is not adding value (ASMR IV-V). A distinctive feature of these two countries is that the drug reimbursement assessment is used to determine the pricing decision procedure rather than to accept or reject the new technology. This is different from other countries, such us England or Sweden, where the outcome of the decision covers all aspects of the reimbursement and it is used to accept or reject the new technology. Consequently, an additional analysis takes this into account by controlling for the pricing system in each situation. Spain also has a particular situation. There is a national assessment within the Ministry of Health (MoH), which is followed by price setting by the CIPM (Comisión Interministerial de Precios de los Medicamentos) using reference pricing. Although there is a national decision, the Spanish regions have some freedom of implementation due to the decentralised nature of the health system. Our Spanish data is based on the national decision.

Another limitation of the database is that some countries do not report negative decisions. In some cases, e.g. Spain, it is not possible to know if a particular drug has been assessed but has subsequently not entered the positive list. For Spain, a negative decision fell under the non-assessed category, as it was not possible to distinguish among them. However, this particular limitation is not major because the Spanish system tends to accept nearly all new drugs. As noted above, the French system has no negative decisions. However, as the recommendation of the Transparency Commission is available, if a particular number of months after the recommendation (based on the other French reimbursement decisions), there was not a final decision, we have assumed that the decision was negative. This approach was not possible in the case of Spain. For Germany, there are only positive decisions, and it is not possible to define negative decisions, as we did for France. For the Netherlands, due to data availability, the database only has positive decisions. Despite these data challenges, these countries are retained in the initial analysis because they are considered relevant for the study. Furthermore, the focus of the analysis is cross-country and not within-country. Due to these data limitations, we also perform robustness checks to validate the approach.

A final limitation with respect to this database is that an ICER cannot be defined for each decision for each country, as this information is not always available in the decisions. As a result, the best option was to work with an approximation of this ICER taking NICE's ICER for each drug-indication (or SMC's ICER if NICE did not report it). The price of each drug would have been also a useful variable but it is usually not available.

4. Econometric Model

The objective is to determine empirically which of the taxonomy variables are associated with a higher or lower probability of reimbursement. The previous empirical studies [12-17] show evidence of a significant impact of clinical evidence and the ICER on drug reimbursement decisions.

The dataset is designed as a particular case of a hierarchical model. It considers the decisions for each drug-indication in ten European countries from January 2006 to November 2014. A panel data design is not feasible, as there is one

[2] ASMR- Amélioration du Service Médical Rendu

decision per drug-indication per country in a particular year as opposed to annual decisions. However, the year of decision is taken into account to construct the time dependent variables and time-specific effects are also considered.

4.1. Dependent variable

Our primary interest is to observe the effect of the explanatory variables on the probability of reimbursement. Although this variable is unobserved, it can be approximated through a categorical variable corresponding to the final decision: 0. Non-Favourable. 1. Favourable with restrictions. 2. Favourable. Due to this problem, the response (Y) cannot be modelled as a linear combination of explanatory variables plus an error. Instead, it will use probabilities.

The recent literature has treated the categorical dependent variable as nominal [13,15], however, it could also be considered ordinal since the "Favourable with restriction" outcome represents an intermediate point between "Favourable" and "Non-Favourable" decision. In this analysis we treat it as ordinal believing that this specification allows us to capture more information regarding the decision outcome.

The dependent variable Y can be considered as a latent variable y_i^* ,

$$y_i^* = \beta_0 + \beta_1 x_1 + \varepsilon_i$$

While the latent variable y_i^* is unobserved, y_i can be observed,

$$\begin{aligned} y_i &= 0 \quad if\ y_i^* \leq 0 \\ y_i &= 1 \quad if\ 0 < y_i^* \leq m_1 \\ y_i &= 2 \quad if\ m_1 < y_i^* \leq m_2 \end{aligned}$$

Assuming that ε_i is normally distributed (with zero mean and unit variance). In our case, Y has the form of a categorical variable, as the outcome of the variable can have three options. The category that is taken as the baseline in the model is "Non-Favourable" decision. Consequently the coefficients and the odds ratios represent deviations from a "Non-Favourable" decision and that "Favourable with restriction" decisions are not directly compared with the "Favourable" decisions.

There are two main options for specifying the model when the dependent variable is: *ordered probit* (when the latent variable is normally distributed) or *proportional odds ratio model/cumulative logit model* (when the latent variable is non-linear). Since the plot of the residuals showed normality, the model was specified as an ordered probit [19].

$$\begin{aligned} \Pr ob(y_i = 0 | x_i) &= \Phi\left(-(b_0 + b_1 x_i)\right) \\ \Pr ob(y_i = 1 | x_i) &= \Phi\left(m_{i,1} - (b_0 + b_1 x_i)\right) - \Phi\left(b_0 + b_1 x_i\right) \\ \Pr ob(y_i = 2 | x_i) &= \Phi\left(m_{i,2} - (b_0 + b_1 x_i)\right) - \Phi\left(m_{i,1} - (b_0 + b_1 x_i)\right) \end{aligned}$$

where $\Phi(.)$ denotes the cumulative normal distribution. For all the probabilities to be positive the following restriction must be fulfilled, $0<\mu_{i,1}<\mu_{i,2}<1$.

Under ordered categorical data, there is an important assumption to fulfil, i.e. the parallel lines/proportional odds assumption. This assumption states that the location parameters (slope coefficients) are the same across response

categories. If this assumption is violated the estimators are biased. As a result, it will be checked through the test of parallel lines. If our model specification violates this assumption, there are different options in order to get better estimators.

The commonly used option is to consider the dependent variable (y_i) as a nominal variable (non-ordered). The model that has been used in the literature is the multinomial logistic regression. As Dakin et al [13] defines, this estimation shows the effect of independent variables on the natural log of the odds of the outcome of a discrete choice being outcome A or B as opposed to the comparison outcome C. In this model, no ranking is assumed between outcomes.

$$\Pr ob\left(Y_i = j|x_i\right) = \frac{e^{b_0 + b_1 x_i}}{1 + \sum_{k=1}^{J} e^{b_{0k} + b_{1k} x_i}} \quad j = 1, 2$$

$$\Pr ob\left(Y_i = 0|x_i\right) = \frac{1}{1 + \sum_{k=1}^{J} e^{b_{0k} + b_{1k} x_i}}$$

The multinomial logistic regression assumes the Independence of Irrelevant Alternatives (IIA), which implies adding another alternative or changing the characteristics of the original options, does not affect the relative odds between the options considered. This hypothesis is quite often violated when the options are similar [20]. Due to our particular type of dependent variable, we strongly believe that this hypothesis is likely to be violated because, the three decisions outcomes are related; since Favourable with restrictions can be regarded as a middle solution between Favourable and Non-Favourable.

Other feasible options suggested by Liu and Agresti [21] are: a) trying different link functions, such as log-log, b) adding additional terms, such as interactions, to the linear predictor, c) generalising the model by using dispersion parameters, d) permitting separate effects for each category for some but not all predictors, i.e. *partial proportional odds* (e.g. random-effects) and e) using the ordinary model for a nominal response.

Liu and Agresti [21] also state that even if a model specification has an inadequate fit, because the parallel lines assumption is violated, this is not the only reason why the model needs to be changed. They note that "when n is large, statistical significance need not imply an inadequate fit in a practical sense, and the decrease in the bias obtained with a more complex model may be more than offset by the increased mean square error in estimating the effects caused by the large increase in the number of model parameters" (Liu and Agresti, 2005).

4.2. Explanatory variables

The main set of explanatory variables, defined in the taxonomy section, is made up of dummies and categorical variables. In addition, some continuous variables are also included to control for the principal socioeconomic characteristics of each country. This combination of variables adds some degree of complexity to our model specification.

4.3. Model specification

Our particular hierarchical model [22] can be specified as follows,

$$Y_{ijt} = a_{ijt} + bx_{ijt} + e_{ijt}$$

where α denotes intercept, β the coefficients, x a matrix of explanatory variables, e the error term and subscripts, i, j, t denote drug-indication, country and year.

The year of the decision is a time effect that needs to be considered. Temporal dependency might create some problems of autocorrelation that need to be adjusted in the model specification. With spatial data, as in this case (country), it is necessary to distinguish between two sources of extra variability [23,24]. First, the largest source is usually named 'spatial dependence', or clustering, and is a consequence of the correlation between the spatial unit and the neighbouring spatial units. The second source is independent, spatially uncorrelated extra variability, which is due to unobserved non-spatial variables that could influence the dependent variable [23,24].

In many datasets, subjects are unlike one another, that is, they are heterogeneous. In our data, these subjects are the decisions. Failure to include heterogeneous quantities in the model may introduce serious bias into the model estimators. In our case, when using a complex design with multiple levels (drug-indication and country) and dimensions (spatial and temporal), there is important heterogeneity in the initial conditions (i.e. intercept). This heterogeneity can be controlled introducing *random-effects* in the intercept (e.g. drug-indication, country-specific and time-specific).

In this particular case, to take into account the spatio-temporal extra variability, we introduce some structure into the model. Heterogeneity is captured by using the random effect associated with the intercept (α_{ij}) (varying at country level j and drug-indication i). Temporal dependency is approximated through a random walk of order 1, and is linked to the random effect associated with the intercept (α_t) (varying at a year level, t).

Some explanatory variables, mainly system-level, were removed from the final model or were regrouped, as they were irrelevant according to t-tests, or were correlated with other variables (e.g. male and female mortality rate). The best model was chosen using the Deviance Information Criterion (DIC) and the Conditional Predictive Ordinate (CPO), both, goodness of fit of the Bayesian approach. The final model specification is a *Hierarchical, Random-Effects Ordered Probit*:

$$P(Y_{ijt}=K)=a_{ijt}+b_1\left(Evidence_j\right)+b_2\left(Initiator=1_j\right)+b_3\left(Stakeholders<2_j\right)+b_4\left(EconomicEvaluation=2_j\right)+b_5\left(Budget\,\mathrm{Im}\,pact_j\right)+$$
$$b_6\left(\mathrm{Pr}\,icedecision=1_j\right)+b_7\left(\mathrm{Pr}\,icedecision=2_j\right)+b_8\left(MEA=1_{ij}\right)+b_9\left(TimediffFav_{ij}\right)+b_{10}\left(ICER=1_i\right)+b_{11}\left(Endoflife=1_i\right)+$$
$$b_{12}\left(Endoflife=2_i\right)+b_{13}\left(Incidencerate_{ij}\right)+b_{14}\left(HEcap_{jt}\right)+b_{15}\left(HE\%GDP_{jt}\right)+b_{16}\left(Mortalityratefemale_{jt}\right)+b_{17}\left(Pop>65_{jt}\right)+$$
$$b_{18}\left(Pop>14_{jt}\right)+b_{19}\left(Healthsystem_j\right)+u_{ijt}$$

K= 0, 1, 2 (decision), i=drug-indication, j=country, t=year.

Once the model was specified, we could test whether it fulfilled the parallel lines assumption. In this case, the result of the test showed that the assumption was violated. One reason for this result is that the sample is not balanced; there is not a decision for each drug-indication for each country (i.e. there are non-assessed drug-indications). As a robustness check, we tried to remove some countries from the sample (i.e. the ones with mostly favourable decisions), but the assumption was still violated, even if the test was closer to the non-rejection. The sample was still unbalanced because some of the remaining countries did not assess all drug-indications and the number of observations dropped significantly.

In order to overcome the bias of the estimates, we follow the strategy suggested by Liu and Agresti [21], outlined above. They state that rather than relying purely on testing, a sensible strategy is to fit models for the separate categories, taking into account ordinary sampling variability. As a result, instead of moving to a multinomial model, where we would face

the same problem with IIA, a good option is to introduce some random-effects in the model specification. As it has been explained above, the model has random-effects in the intercept accounting for drug-indication, country and year. This approach improves the efficiency of the model.

4.4. Inference

Our model is quite complex. The variables, both the dependent and the explanatory variables, are mainly categorical, there is a space-time dimension and there is heterogeneity. Under these circumstances, it is more suitable for our database and model specification, to use the Bayesian approach. Within the (pure) Bayesian framework, the Integrated Nested Laplace Approximation (INLA) [25] approach is followed (for review see [26,27]). All analyses are made with the free software R (version 2.15.3) [28], through the INLA library.

The Bayesian approach is considered the most suitable for accounting for model uncertainty, both in the parameters and in the specification of the models, either in cross-sectional studies [29-31] or in panel data models [32-35]. Moreover, within the Bayesian approach, it is easy to specify a hierarchical structure on the (observable) data and (unobservable) parameters where all are considered random quantities. The Bayesian approach is criticized because it is time consuming, as it uses the Markov Chain Monte Carlo (MCMC) simulations to get to the final result. However, there is a recent implementation (INLA) that avoids the simulation and gives a straightforward result. So, this improvement allows us to use Bayesian methods much more efficiently and is less time consuming.

5. Results

There is substantial evidence of different reimbursement decisions for the same drug-indication across the ten countries. However, the specification of the econometric model is able to control for the differences across the HTA systems and the socioeconomic conditions of each country, as the random-effect used to allow for country variation is not significant. In addition, the other two random-effects included in the intercept (i.e. drug-indication and time) show a very low variability. However, they are relevant in order to control for the different levels of the dataset and for the ordered probit specification.

The results of the econometric model can be classified into three main analyses: A) Ten European countries, B) Eight European countries (without France and Germany) and C) Ten European countries, controlling for the pricing decision. The results of each analysis are shown in Tables VII and VIII. These tables classifies the results according to the variables forming the taxonomy and it highlights the variables that have a statistically significant effect (95%) on the probability of reimbursement.

The results of the first model (A) are shown in Table VII. In terms of the system-wide variables, the results of the econometric model show, that less involvement of stakeholders in the process is associated with a higher probability of reimbursement. However, the requirement of economic evaluation for all drugs, an external review of the evidence, the manufacturer being the initiator of the process, price based on a reference pricing and the budget impact requirement are related to a lower probability of reimbursement.

According to the results of model A, for product-specific variables, a drug-indication with a NICE favourable decision is linked with a higher probability of this drug-indication being accepted in another country. Moreover, a drug considered

cost-effective by NICE/SMC or having a financial MEA is associated with a higher probability of reimbursement. The other product-specific variables included in the model were not statistically significant.

Finally, the results also show the effects of the time variables included in the model. In particular, the rate of female mortality and percentage of population older than 65 are positively and significantly associated with the probability of reimbursement. However, higher proportion of population below 14 years old is related with a lower probability of reimbursement. The remaining time variables included in the model did not have a significant association with the probability of reimbursement.

5.1. Additional analyses

In the second analysis (B) the model is estimated without data from France and Germany in order to overcome one of the limitations identified in Section 3.2, namely that in France and Germany drug reimbursement assessment is used to define the pricing procedure and not to accept or reject the new technology. The new model specification (B) follows the same structure as model A but some adjustments were needed due to new database.

Comparing models A and B, the main finding is that model B is more efficient in terms of the DIC and the CPO. In terms of the individual significance of the explanatory variables, model B broadly follows the results of model A. However, the main differences of this new model are that the initiator and the time dependent variables are not statistically significant compared to model A.

The third, and last, model (C) (reported in Table VIII) takes an alternative approach to address the limitation, regarding the meaning of the decision outcome. In this case, instead of removing data for both countries, it captures the difference by controlling for the pricing system in each situation. Although in models A and B there was already a variable defining the pricing decision, our hypothesis is that the effect of the explanatory variables on the final decision may differ according to the pricing system. In model C, the pricing decision variable is removed, but some interactions between the main explanatory variables and the pricing systems are included in order to observe different effects depending on the pricing system used in each decision. The first result to note is that the efficiency of models A and C, in terms of the DIC and CPO, is very similar. However, model C explains more according to the pricing system. The conclusion of this analysis is that there is a statistically significant effect when the price is set by the manufacturer (such as in NICE, SMC, Sweden, Netherlands and Poland) compared to the price based on a negotiation (France and Germany when the drug is adding value). In other words, our hypothesis can be corroborated, since the pricing system of each country has an effect on the decision outcome.

5.2. Robustness checks[3]

The models were re-run without the observations from Spain, France, Germany and the Netherlands for the different limitations arising with respect to these countries. In the case of Germany, there was only data from 2011; 21 Favourable decisions. For Spain, there were more observations, but nearly all of them were Favourable decisions (except from a few restricted decisions). Regarding the Netherlands, there were only positive decisions (45 decisions). For France, 93% of

[3] Tables with the results of the robustness checks can be obtained from authors on request.

the decisions were also favourable. They were also removed in order to test the parallel lines assumption, as explained above.

By removing these countries from the analysis (one by one and all together), the models improved slightly in terms of efficiency (DIC and CPO reduced), but in terms of individual significance, fewer variables were significant but the ones that kept the significance had the same sign of the original model. In addition, removing the countries was not solving the parallel lines problem. We believe that it is desirable to retain these countries, due to the interest in their systems and the importance of these countries inside Europe. In econometrics terms, without these countries the drop in the number of observations was substantial.

Another robustness check was performed to see whether the exclusion of non-submission decisions from the main analysis was appropriate. The results of the model changed significantly and the model lost efficiency when the non-submission decisions were treated as non-favourable. This confirms that these two types of decision are best considered separately.

In a final robustness check, the econometric model was specified as a two-part model, which followed the initial specification, i.e. Random- effects Hierarchical model, in order to test for sample selection regarding the assessment. The first part of the model defines whether the drug-indication has been assessed or not. The model is specified as a Probit with a binary variable (0. Non-assessed /1. Assessed) with a set of explanatory variables that can explain the assessment decision (i.e. Evidence, Initiator, Economic Evaluation, Health system, Budget impact, Pricing decision, NICE has assessed it, ICER, Orphan, Incidence rate and Health expenditure). The second part of the model is the Ordered Probit previously defined.

The results of the model did not show any statistical difference for the second part (i.e. efficiency and significance did not change) compared to the original analysis. In other words, the ordered model that we estimated initially is robust. However, an interesting finding for the first part is that the model shows significant differences across countries when accounting for country random-effects. As a result, it seems that country-effects are relevant in order to explain if a drug-indication has been assessed or not, but not for the final decision once controlling for system-level and socioeconomic characteristics.

6. Discussion

The main objective of this paper was to determine the factors that may lead to different drug reimbursement decisions across ten European countries, through designing a taxonomy and estimating a Hierarchical, Random-Effects Ordered Probit. There are some interesting results.

According to previous literature and our data collection, differences exist in the final reimbursement decision across the ten countries analysed. As a result, one of the main findings of our study is that it documents the differences in decisions across the ten European countries. The database shows that there are different reimbursement decisions for the same drug-indication for the selected countries. However, the specifications of the econometric models were able to control for the differences across the HTA systems and the socioeconomic conditions of each country, as the random-effect used to allow for country variation was not significant.

The results also show, contrary to expectations, that less involvement of stakeholders in the process is linked to a higher probability of acceptance. However, this may arise from the combination of two extremes Spain, which does not involve stakeholders, accepts most of the cancer drugs and NICE, which fully involves stakeholders, rejects a substantial number. Another interesting result is the introduction of economic requirements in the assessment is associated with a lower probability of reimbursement. These requisites allow for a deeper analysis and increase the strictness of the evaluation.

As with previous studies [12-15], the incremental cost-effectiveness ratio (ICER) is important for HTA decision-making. In particular, these studies and our results, show that a lower ICER (better cost-effectiveness) increases the probability of reimbursement. However, our results differ from the research of Hernandez-Villafuerte et al. [36] on the effect of NICE decisions on other countries. Our findings show that a favourable NICE decision is associated with a higher probability of reimbursement in another country. However, Hernandez-Villafuerte et al [36] state that other countries tend to follow NICE decisions when NICE restricts or rejects the drug, but not when it gives a positive recommendation. So, our results are different. This difference may be related to the small number of decisions that they are analysing, as they note.

The findings on cost-effectiveness, the effect of NICE on other countries and the Economic evaluation requirement are in line with the results from our previous research. The results of testing these three hypotheses had the same outcome. However, the health system and the Health expenditure per capita are not showing any significance. Compared to the previous study, the main contribution of this paper is that the econometric analysis gives more validity to our findings.

Another interesting finding is what we show in the second and third analysis (model B and C). Both analyses were done to overcome one of the main limitations of the database, i.e. meaning of the decision outcome. Model B was run without France and Germany. In Model C the specification tried to capture the effect of the explanatory variables on the final decision according to the pricing system. The main results of both analyses were that without France and Germany the efficiency of the model improved and that depending on the price system, the effect on the final reimbursement decision was different.

We performed a number of robustness checks to validate our results. Very interesting findings are the result of the two-part model. The first part of the model, where assessment is evaluated, shows significant differences across countries when including country random-effects. As a result, it seems that country-effects are relevant in order to explain if a drug-indication has been assessed or not, but not for the final decision once controlling for system-level and socioeconomic characteristics. This finding should be further investigated in future research.

Even if the results are satisfactory, during this research we encountered a number of limitations. Assembling the database for this analysis was both time consuming and complicated because not all countries make their decisions publicly available or provide sufficient detail. These issues were overcome by contacting national experts who helped us validating our database. Moreover, a number of assumptions were needed in order to combine all these data in a single analysis. Another important challenge was the model specification. Our dependent variable has been treated as an ordered categorical variable and this specification requires the fulfilment of the parallel lines assumption. Our model did not fulfil this assumption and basing our strategy on Liu and Agresti (2005), we kept with the ordered probit but introducing some random-effects.

As future research, it would be interesting to make a comparison across therapeutic areas. In order to explore whether type of drug is an additional determinant of the probability of reimbursement. In this study, we have only looked at cancer drugs.

Acknowledgements

This paper was developed within the research project funded by the European Commission's Research Framework Programme (FP7), Advance-HTA (nº 305983). The authors would like to thank the Advance-HTA Consortium members and the experts from the National HTA Agencies for their support on the data collection process. Moreover, the authors appreciate the good feedback received in the different conferences where this work has been presented (HESG, EvaluAES, AES, HTAi, iHEA, internal seminars at LSHTM and CRES, UPF). We would also like to thank Prof. Marc Saez. Finally, we appreciate the comments of two anonymous reviewers that, without doubt, helped us improve our work.

Conflicts of interest

There are no conflicts of interest for any of the authors. Both authors freely disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations that could inappropriately influence, or be perceived to influence, their work.

7. References

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

[2] Maynou, L., Cairns, J.: Why do some countries approve a cancer drug and others don't?. Journal of Cancer Policy 4, 21-25 (2015).

[3] Sorenson, C., Drummond, M., Kanavos, P.: Ensuring value for money in health care: the role of health technology assessment in the European Union. European Observatory on Health Systems and Policies, Copenhagen (2008).

[4] Wilsdon, T., Serota, A.: A comparative role and impact of health technology assessment. Charles River Associates, London (2011).

[5] Franken, M., Le Polain, M., Koopmanschap, M., Cleemput, I.: Similarities and differences between five European drug reimbursement systems. International Journal of Technology Assessment in Health Care 28(4), 349–357 (2012).

[6] Paris, V., Belloni, A.: Value in Pharmaceutical Pricing. OECD Health Working Papers 63 (2013).

[7] Barneih, L., Manns, B., Harris, A., Blom, M., Donaldson, C., Klarenbach, S., Husereau, D., Lorenzetti, D., Clement, F.: A synthesis of Drug Reimbursement Decision-Making Processes in Organisation for Economic Co-operation and Development countries. Value in Health 17, 98-108 (2014).

[8] Mason, AR., Drummond, MF.: Public funding of new cancer drugs: Is NICE getting nastier?. European Journal of Cancer 45(7), 1188–1192 (2009).

[9] Clement, F., Harris, A., Li, J., Yong, K., Lee, K., Manns, B.: Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia, and Canada. JAMA 302(13), 1437–1443 (2009).

[10] Kanavos, P., Nicod, E., Van den Aardweg, S., Pomedli, S.: The impact of health technology assessments: an international comparison. Euro Observer 12(4), 1–7 (2010).

[11] Nicod, E., Kanavos, P.: Commonalities and differences in HTA outcomes: A comparative analysis of five countries and implications for coverage decisions. Health Policy 108,167–177 (2012).

[12] Devlin, N., Parkin, D.: Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Economics 13(5), 437–452 (2004).

[13] Dakin, H., Devlin, N., Odeyemi, I.: ''Yes'', ''No'' or ''Yes, but''? Multinomial modelling of NICE decision-making. Health Policy 77, 352–367 (2006).

[14] Dakin, H., Devlin, N., Feng, Y., Rice, N., O'Neill, P., Parkin, D.: The Influence of Cost-effectiveness and Other Factors on NICE Decisions. Health Economics 24, 1256-71 (2015).

[15] Cerri, K., Knapp, M., Fernandez, JL.: Decision making by NICE: examining the influences of evidence, process and context. Health Economics, Policy and Law 1 – 23 (2013).

[16] Fischer, K., Rogowski, W., Leidl, R., Stollenwerk, B.: Transparency vs. closed-door policy: Do process characteristics have an impact on the outcomes of coverage decisions? A statistical analysis. Health Policy 112, 187-196 (2013).

[17] Charokopou, M., Majer, IM., Raad, J., Broekhuizen, S., Postma, M., Heeg, B.: Which Factors Enhance Positive Drug Reimbursement Recommendation in Scotland? A Retrospective Analysis 2006–2013. Value in Health 18(2), 284-91 (2015).

[18] Hutton, J., McGrath, C., Frybourg, JM., Tremblay, M., Bramley-Harker, E., Henshall, C.: Framework for describing and classifying decision-making systems using technology assessment to determine the reimbursement of health technologies (fourth hurdle systems). International Journal of Technology Assessment in Health Care 22(1), 10-8 (2006).

[19] Zavoina, R., McElvey, W.: A statistical model for the analysis of ordinal level dependent variables. Journal of Mathematical Sociology 103-120 (1975).

[20] Frees, E.: Longitudinal and Panel Data: Analysis and Applications for the Social Sciences, Cambridge University Press (2004).

[21] Liu, I. Agresti, A.: The Analysis of Ordered Categorical Data: An Overview and a Survey of Recent Developments. Test 14(1), 1-73 (2005).

[22] Garson D.: Hierarchical Linear Modelling: Guide and Applications. SAGE Publications, Inc (2013).

[23] Lawson, AB., Browne, WJ., Vidal-Rodeiro, CL. Disease mapping with WinBUGS and MLwiN'. John Wiley & Sons: Chichester (2003).

[24] Barceló, MA., Saez, M., Saurina, C.: Spatial variability in mortality inequalities, socioeconomic deprivation, and air pollution in small areas of the Barcelona Metropolitan Region, Spain. Science of the Total Environment 407(21), 5501-5523 (2009).

[25] R-INLA project. http://www.r-inla.org/. Accessed 30 September 2015.

[26] Rue, H., Martino, S., Chopin, N.: Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion). Journal of the Royal Statistical Society Series B 71, 319-392 (2009).

[27] Blangiardo, M., Cameletti, M., Baio, G., Rue, H.: Spatial and spatio-temporal models with R-INLA. Spatial and Spatio-temporal Epidemiology 4, 33-49 (2013).

[28] R Development Core Team: R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, ISBN 3,900051,07,0 (2012).

[29] Raftery A.: Bayesian model selection in social research. Sociological Methodology 25, 111-163 (1995).

[30] Fernandez, C., Ley, E., Steel, M.: Model uncertainty in cross-country growth regressions. Journal of Applied Econometrics 16, 563-576 (2001).

[31] Sala-i-Martin, X., Doppelhofer, G., Miller, R.: Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. American Economic Review 94(4), 813-835 (2004).

[32] Hsiao, C., Pesaran, MH., Tahmiscioglu, AK.: Bayes estimation of short-run coefficients in dynamic panel data models. In Hsiao, C., Lee, LF., Lahiri, K., Pesaran, MH. (eds): Analysis of Panels and Limited Dependent Variables Models, 268-206. Cambridge University Press: Cambridge (1999).

[33] Hsiao, C., Pesaran, MH.: Random coefficient panel data models', In Mátyás L, Sevestre P. (eds): The Econometrics of Panel Data, Advances Studies in Theoretical and Applied Econometrics (46), 185-213, Springer (2008).

[34] Moral-Benito, E.: Determinants of economic growth: a Bayesian panel data approach. Documento de Trabajo 1031, Madrid: Banco de España (2010).

[35] Rendon, SR.: Fixed and random effects in classical and Bayesian regression. Oxford Bulletin of Economics and Statistics 75(3), 460-476 (2012).

[36] Hernandez-Villafuerte, K., Garau, M., Devlin, N.: Do NICE decision affect decisions in other countries. *Value in Health* 17(7), A418 (2014).

Table I. Taxonomy

1. System-wide variables	
	1.1. Organisational
Evidence	Whether the assessment is produced or reviewed inside the body dealing with the drug reimbursement (agency) or, by contrast, this evidence is produced or reviewed by an independent body (outside the agency), for instance, universities, or independent committees.
Body Independence	The body in charge of the drug reimbursement is an independent scientific body or the government manages it. Moreover, if it is independent from the Ministry of Health (MoH), does it make recommendations or the final decision?
Decision level	This variable indicates if the decision and recommendation is taken at a national or regional level.
Health System	This variable collects whether the country health system is based on a Social Health Insurance (SHI) or a Tax-based system.
	1.2. Process
Initiator	In most cases the manufacturer applies for reimbursement, however, in some countries, the initiative comes from the Department of Health, from the body in charge of HTA or it is an automatic procedure. In these last cases, then, the manufacturer is asked to make a specific submission.
Stakeholders	The different systems have a diverse degree of involvement of stakeholders. In some countries, they are fully involved in the whole procedure, while in other countries their presence is just limited to some comments at the early assessment.
Transparency	This variable indicates the transparency of the system, in terms of documentation publicly available, without taking into account the information of price negotiations.
Appeal	This variable records whether or not there is a formal system to appeal the final decision taken by the decision-making body.
	1.3. Method
Economic Evaluation	This variable indicates whether an economic evaluation (cost effectiveness, cost utility or cost-benefit analysis) is required for the decision-making process. It can be that it is always needed for the assessment or that it is only required for some group of drugs (e.g. drugs which increase the therapeutic value) or non-required.
Budget Impact	This variable shows if a budget impact analysis is required for the decision-making process.
Pricing location	This variable indicates what type of institution deals with price setting (inside the Ministry of Health (MoH), external body or none of them, price set by manufacturer).
Pricing decision	The variable records how the pricing decision is taken in each of the previous cases. It can be a price negotiation, it can be based on referencing pricing or, by contrast, it can be set by the manufacturer in the submission and used for the corresponding calculations.

Source: own construction

Table I. (Continued)

2. Product-specific variables	
	2.1. General drug characteristics
Type of patient	This variable identifies whether the drug-indication is for adults, for children or for both.
Orphan	This variable indicates whether or not a drug is designated as an orphan by European Medicine Agency (EMA)[1]. A drug is qualified under orphan when it fulfils the following criteria (EMA, orphan designation): 1) it must be intended for the treatment, prevention or diagnosis of a disease that is life-threatening or chronically debilitating; 2) the prevalence of the condition in the European Union (EU) must not be more than 5 in 10,000; 3) no satisfactory method of the condition concerned can be authorised.
Incidence rate	This variable tries to collect an estimate the number of patients for whom the drug is indicated. From the technology appraisals, it is not always feasible to know for how many patients are eligible for that drug in each country. This variable is approximated through the incidence rate. The information is taken from the age-standardised incidence rate per 100,000 for each therapeutic area and country[2]. For cancer, the incidence rate is disaggregated per type of cancer and country (GLOBOCAN 2012 project).
Disease stage	This variable determines whether the drug-indication is a treatment for an early stage or late stage of the condition.
ICER	This variable indicates whether the Incremental Cost-Effectiveness Ratio (ICER) determined by NICE is above or below £30,000 per QALY. When the drug indication has not been appraised by NICE, the SMC ICER is taken to define this variable. NICE generally performs a more detailed analysis than SMC in calculating the ICER, while SMC usually accepts the ICER identified by the manufacturer. ICER variable is not a continuous variable because of two main reasons. Firstly, NICE and SMC are the bodies that always document this value (transparency). Secondly, for simplicity, it is used as an indicator of cost-effectiveness (i.e. threshold from NICE). A categorical variable is able to show the relationship between cost-effectiveness and the probability of reimbursement, while a continuous variable will not take into account the specific criteria of cost-effectiveness.
End of life	Was the drug-indication accepted by NICE as an end of life treatment? For the drug-indications assessed before 2009 (year of implementation of the criteria), it is categorised with another code.
	2.2. ***Specific drug-country characteristics***
Managed Entry Agreement (MEA)	This variable indicates the existence of a MEA (also called: "Risk sharing agreements" or "Patient access schemes") during the decision-making process. It also collects the type of MEA: financial, performance-based or a combination of both.
Alternative	This variable shows whether or not there are alternative active treatments for this drug-indication already available in the positive list of each of the countries. It is not considered to be an alternative treatment when the comparator is best supportive care, standard chemotherapy or standard care.
Time difference to NICE (Timedifffav, Timediffres, Timediffnonfav)	This variable shows if at the moment of the decision, there was already a decision made by NICE for that particular indication and also, if the decision was the same as NICE. It is consider that there is a NICE decision 4 weeks before NICE final decision (because some information is already available). NICE HTA analyses are considered among the most complete and strict. Thus, regardless of whether a country's decision precedes or follows a NICE decision, other countries will tend to say yes to drugs for which NICE make a Favourable decision. A NICE favourable decision is a proxy for the quality of the supporting evidence.

[1] European Medicines Agency, Human medicines, Orphan designations. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000029.jsp&mid=WC0b01ac05800240ce. Accessed 30 September 2015
[2] http://globocan.iarc.fr/Default.aspx. Accessed 30 September 2015

Table I. (Continued)

3. Time Dependent variables
Gross Domestic Product (GDP) growth rate
Health expenditure (HE) per capita Purshasing Parity Power (constant 2005 international $)
Health expenditure (HE) public (%GDP)
Mortality rate adult female (per 1,000 female adults)
Mortality rate adult male (per 1,000 male adults)
Population ages 0-14 (% of total)
Population ages >65 (% of total)

Source: World Bank Data

Table II. Variables categorised

SYSTEM-LEVEL VARIABLES	
Evidence	0. Internal (done inside the agency) 1. External (review body outside the agency)
Body Independence	0. Inside Ministry of Health 1. Independent body only does a recommendation 2. Independent body who decides
Decision level	0. Recommendation and decision at National level 1. Recommendation National / decision Regional 2. Recommendation Regional/decision National 3. Recommendation and decision National, freedom for implementation at Regional
Health system	0. Tax-based system 1. Social Health Insurance system
Initiator	0. Department of Health 1. Manufacturer submission 2. Body in charge of the HTA 3. Automatic 4. Both, manufacturer and Department of Health
Stakeholders	0. Non-involved 1. Only comments at an early stage 2. Involved but not in the final meeting 3. Fully involved
Transparency	0. Nothing 1. Some documents 2. Everything
Appeal	0. No 1. Yes
Economic evaluation	0. Never 1. Only for some drugs 2. Yes, for all cases
Budget impact	0. No 1. Yes
Pricing location	0. No negotiation (e.g. price set by Manufacturer) 1. External body 2. Inside Ministry of Health
Pricing decision	0. Based on a negotiation 1. Calculation based on price referencing 2. Set by the manufacturer
PRODUCT-SPECIFIC VARIABLES	
Type of patient	0. Adults 1. Children 2. Both
Orphan	0. No 1. Yes
Incidence rate	Numeric variable
Disease stage	0. Early treatment 1. Late treatment 2. Not specified
ICER	0. Above £30,000 per QALY 1. Below £30,000 per QALY 2. Non submission 3. Non data
End of life	0. No 1. Yes

	2.	Not determined (before 2009)
Managed Entry Agreement (MEA)	0. 1. 2. 3.	No Yes (financial schemes) Yes (performance-based) Yes (combination of both)
Alternative	0. 1.	No Yes
Time difference to NICE (Timedifffav, Timediffres, Timediffnonfav)	0. 1.	No Yes

TIME DEPENDENT VARIABLES
GDP growth rate
HE per capita PPP (constant 2005 international $)
HE public (%GDP)
Mortality rate adult female (per 1,000 female adults)
Mortality rate adult male (per 1,000 male adults)
Population ages 0-14 (% of total)
Population ages >65 (% of total)

Source: own construction and World Bank Data

Table III. Decision data by country (sources)

Country	Institution/Database	Data source
England	NICE	HTA decisions from the NICE website
Scotland	SMC	HTA decisions from the SMC website
Sweden	TLV / NLT	HTA decisions from the TLV/NLT website. Validation from the TLV team.
Belgium	RIZIV INAMI	HTA decisions from the INAMI database (online). Validation of the data and information on MEA from the INAMI team.
Portugal	INFARMED	HTA decisions from INFARMED database (online). Information on the MEA from the INFARMED team.
Poland	AHTAPol	Database created by AHTAPol.
Spain	BOTPLUS	Database created by EASP and UCLM from BOTPLUS. Validation of data by GENESIS.
Germany	G-BA	HTA decisions from the G-BA website. Only decisions from 2011 onwards (AMNOG)
Netherlands	ZIN/MoH	Information on decisions provided by MoH.
France	HAS/MoH	Database created by the UPEC.

NICE - National Institute for Health and Care Excellence; SMC - Scottish Medicine Consortium; TLV - The Dental and Pharmaceutical Benefits Agency, NLT - New pharmaceutical product therapies, RIZIV-INAMI - Belgium Health Insurance Agency; INFARMED - National Authority of Medicines and Health Products, IP; AHTAPol - Agency for Health Technology Assessment in Poland; EASP - Andalusian School of Public Health; UCLM - University of Castilla la Mancha; GENESIS - Spanish Society of Hospital Pharmacy; G-BA - Federal Joint Committee; AMNOG - Pharmaceuticals Market Reorganisation Act; ZIN - Dutch Health Care Insurance Board; HAS - French National Authority for Health; UPEC - University Paris-Est Créteil. *Source: own construction*

Table IV. Decision outcome per country[1]

	Scotland		England		Belgium		Sweden		France	
Non-Favourable	32	(25.2%)	34	(37%)	4	(3.6%)	5	(11.1%)	6	(4.6%)
Restricted	35	(27.6%)	20	(21.7%)	45	(40.5%)	6	(13.3%)	3	(2.3%)
Favourable	24	(18.9%)	31	(33.7%)	62	(55.9%)	34	(75.6%)	122	(93.1%)
Non-submission	36	(28.3%)	7	(7.6%)	0	(0%)	0	(0%)	0	(0%)
Total	**127**	**100%**	**92**	**100%**	**111**	**100%**	**45**	**100%**	**131**	**100%**

	Poland		Portugal		Germany		Spain		Netherlands	
Non-Favourable	27	(28.7%)	7	(15.9%)	0	(0.0%)	0	(0%)	0	(0%)
Restricted	47	(50%)	3	(6.8%)	0	(0%)	10	(12.2%)	0	(0%)
Favourable	20	(21.3%)	34	(77.3%)	21	(100%)	72	(87.8%)	45	(100%)
Non-submission	0	(0%)	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Total	**94**	**100%**	**44**	**100%**	**21**	**100%**	**82**	**100%**	**45**	**100%**

[1] We considered 161 drug-indications per country, however, decisions before 2006, non-assessed drug-indications and decisions when the date was missing, are not included (no possibility of linking them with time variables). Finally, this sample accounts for 158 drug-indications.

Source: own construction

Table V. Descriptive statistics: categorical variables

Variable	Category	N (%)
Evidence	Internal	607 (76.6%)
	External	185 (23.4%)
Body Independence	Inside MoH	171 (21.6%)
	Indep. Recom.	495 (62.5%)
	Indep. Decision	126 (15.9%)
Decision level	Recom./Decision National	538 (67.9%)
	Recom. National /Decision Regional	172 (21.7%)
	Recom./Decision National. Regional freedom	82 (10.4%)
Initiator	Department of Health (DoH)	174 (22%)
	Manufacturer	390 (49.2%)
	Body in charge of HTA	56 (7.1%)
	Automatic	127 (16.0%)
	Both, manufacturer and DoH	45 (5.7%)
Stakeholders	Non-involved	93 (11.7%)
	Only early assessment	349 (44.1%)
	Involvement, not final meeting	258 (32.6%)
	Fully involved	92 (11.6%)
Transparency	No documents available	82 (10.4%)
	Some documents available	142 (17.9%)
	Everything publicly available	568 (71.7%)
Appeal	No	150 (18.9%)
	Yes	642 (81.1%)
Economic Evaluation	No	101 (12.8%)
	Only for some drugs	268 (33.8%)
	Yes, for all cases	423 (53.4%)
Budget Impact	No	126 (15.9%)
	Yes	666 (84.1%)
Pricing location	No negotiation	253 (31.9%)
	External body	122 (15.4%)
	Inside MoH	417 (52.7%)
Health system	Tax-based system	390 (49.2%)
	Social Health Insurance system	402 (50.8%)

Variable	Category	N (%)
Pricing decision	Based on a negotiation	102 (13%)
	Calculation based on price referencing	293 (37%)
	Set by the manufacturer	392 (49.8%)
Type of patient	Adults	758 (95.7%)
	Children	10 (1.3%)
	Both	24 (3%)
Orphan	No	609 (76.9%)
	Yes	183 (23.1%)
Disease stage	Early treatment	149 (18.8%)
	Late treatment	630 (79.5%)
	Not specified	13 (1.6%)
ICER	Above £30,000 per QALY	381 (48.1%)
	Below £30,000 per QALY	253 (31.9%)
	Non-submission	119 (15%)
	No-data	39 (4.9%)
End of life treatment	No	345 (43.6%)
	Yes	145 (18.3%)
	Not determined (before 2009)	302 (38.1%)
Managed Entry Agreement	No	633 (79.9%)
	Yes, financial scheme	136 (17.2%)
	Yes, performance-based	20 (2.5%)
	Yes, combination	3 (0.4%)
Alternative	No	511 (64.5%)
	Yes	280 (35.4%)
Timedifffav	No	480 (84.4%)
	Yes	89 (15.6%)
Timediffres	No	534 (93.8%)
	Yes	35 (6.2%)
Timediffnonfav	No	518 (91%)
	Yes	51 (9%)

Source: own construction

Table VI. Descriptive statistics: Continuous variables

Variable	N	Mean (SD)	Min	Max
Incidence rate	747	31.86 (35.97)	1.60	159.1
GDP growth	704	1.01 (2.17)	-5.17	6.56
HE capita (PPP, $)	619	3340.43 (861.97)	1240.39	5384.61
HE public (% GDP)	619	7.46 (1.20)	4.71	9.93
Mortality rate (per 1000 female)	377	56.61 (6.17)	42.63	77.35
Mortality rate (per 1000 male)	377	103.63 (22)	70.81	204.98
Population 0-14 (% total pop.)	720	16.77 (1.36)	13.09	18.46
Population > 65 (% total pop.)	720	16.68 (1.42)	13.34	21.14

Source: own construction

Table VII. Results of the model

	A) CANCER 2006	*B) CANCER 2006 (no France/Germany)*
Intercept	2.785[1] [1.624,4.811][2]	3.541 [1.755,7.114]
System-wide variables		
Evidence (=1 external review)	0.548 [0.362,0.819]	0.531 [0.340,0.819]
Health system (=1 SHI)	1.170 [0.910,1.502]	0.788 [0.520,1.194]
Initiator (1= Manufacturer)	0.670 [0.504,0.884]	0.853 [0.613,1.178]
Stakeholders (<2 not involved/early stage)	1.281 [1.151,1.426]	1.298 [1.144,1.472]
Economic Evaluation (2= required for all drugs)	0.784 [0.622,0.987]	0.529 [0.358,0.783]
Budget Impact	0.554 [0.356,0.855]	0.589 [0.368,0.933]
Pricing decision (1=reference pricing)	0.638 [0.474,0.852]	0.519 [0.320,0.850]
Pricing decision (2=set by the manufacturer)	0.838 [0.665,1.053]	1.226 [0,781,1.915]
Product-specific variables		
Managed Entry Agreement (1= Financial MEA)	1.256 [1.093,1.441]	1.325 [1.143,1.532]
Timedifffav (1=Yes, NICE Fav. Before)	1.494 [1.248,1.781]	1.627 [1.328,1.982]
ICER (1=cost-effectiveness)	1.114 [1.008,1.231]	1.154 [1.028,1.296]
End of life (1=end of life criteria fulfilled)	1.099 [0.995,1.213]	1.104 [0.982,1.240]
End of life (2= before 2009, criteria not applicable)	0.982 [0.855,1.129]	0.986 [0.840,1.159]
Incidence rate	0.999 [0.997,1.001]	0.999 [0.997,1.001]
Time variables		
HE per capita ($ PPP)	1.0002 [0.999,1.001]	0.999 [0.999,1.0002]
HE public (% GDP)	0.929 [0.826,1.046]	1.100 [0.920,1.316]
Mortality rate 1000 female	1.004 [1.001,1.007]	1.003 [0.999,1.006]
Population >65 (% total population)	1.087 [1.037,1.140]	1.020 [0.914,1.139]
Population <14 (% total population)	0.914 [0.867,0.964]	0.975 [0.864,1.102]
N of observations	749	597
Study period	2006-2014	2006-2014
N of countries	10	8
N of Drug-Indications	158	157
DIC	1899.15	1481.54
Effective number of parameters	20.10	20.01
CPO	1.258	1.230

[1] Odds Ratio [2] 95% CI Statistically significant: 95% Credible Interval did not contain the unity. *Source: own construction*

Table VIII. Third analysis: results of the model

	C) CANCER 2006 (pricing effect)
Intercept	1.645[1] [1.206,2.227][2]
System-wide variables	
Evidence (=1 external) * Price set by manufacturer	0.574 [0.371,0.884]
Health system (=1 SHI system)	0.915 [0.738,1.134]
Initiator (1= manufacturer) * Price set by manufacturer	0.699 [0.442,1.099]
Stakeholders (<2 not involved/early stage) * Price referencing	1.063 [0.852,1.323]
Stakeholders (<2 not involved/early stage) * Price set by manufacturer	1.614 [1.194,2.178]
Economic Evaluation (1= for all drugs) * Price referencing	0.666 [0.431,1.026]
Economic Evaluation (1= for all drugs) * Price set by manufacturer	0.593 [0.391,0.899]
Budget Impact (1= required) * Price set by manufacturer	0.603 [0.426,0.851]
Product-specific variables	
Timedifffav (1=Yes, NICE Fav. Before)	1.507 [1.260,1.795]
MEA (=1 Financial) * Price referencing	1.321 [0.912,1.885]
MEA (=1 Financial) * Price set by manufacturer	1.536 [1.173,2.003]
ICER (1=cost-effectiveness) * Price referencing	0.966 [0.799,1.164]
ICER (1=cost-effectiveness) * Price set by manufacturer	1.430 [1.155,1.768]
End of life (1=end of life criteria fulfilled)	1.078 [0.975,1.192]
End of life (2= before 2009, criteria not applicable)	0.973 [0.849,1.116]
Incidence rate	0.999 [0.997,1.001]
Time variables	
HE per capita ($ PPP)	1.00001 [0.999,1.0003]
HE public (% GDP)	1.001 [0.879,1.138]
Mortality rate 1000 female	1.003 [1.0003,1.006]
Population >65 (% total population)	1.017 [0.974,1.059]
Population <14 (% total population)	0.982 [0.938,1.031]
N of observations	749
Study period	2006-2014
N of countries	10
N of Drug-indications	158
DIC	1899.66
Effective number of parameters	22.09
CPO	1.258

[1] Odds Ratio [2] 95% CI Statistically significant: 95% Credible Interval did not contain the unity. *Source: own construction*

8. Annex

System-level variables

Countries /variables	Agency	Evidence	Body independency	Decision level	Health system	Initiator	Stakeholders	Transparency	Appeal	Economic evaluation	Budget impact	Pricing location	Pricing decision
Belgium	RIZIV INAMI	0	1	0	1	1	2	2	1	1	1	1	1
England	NICE	1	2	0	0	0	3	2	1	2	0	0	2
France	HAS	0	1	0	1	1	1	1	1	1[~]	1	2	1
Germany	IQWiG	0	1	0	1	1	2	2	1	2[$]	1	2	0[&]
Netherlands	ZIN	0	0	0	1	2	1	2	0	1	1	2	2
Poland	AHTAPol	0	1	0	1	1[#]	1	2	0	2	1	2	2
Portugal	INFARMED	0	0	0	0	1	1	2	1	2	1	2	1
Scotland	SMC	0	1	1	0	3	1	2	1	2	1	0	2
Spain	MoH/CIPM	1	0	3	0	0	0	1	1	1	1	2	1
Sweden	TLV prescribed	0	2	1	0	1	1	2	1	2	0	0	2
Sweden	NLT hospital	1	1	1	0	2	0	1	0	2	1	1	0

France: product specific. Price decision=1 for ASMR I,II and III.
Price decision=0 for ASMR IV and V.

Germany: product specific. Price decision=0 for drugs with added benefit
Price decision=1 for drugs with no added benefit.

Time changes: [~] 0 - before 2013; [$] 1- from 2010 to 2007 / 0 – before 2007; [&] 2 – before 2011 for drugs with added benefit; [#] 4 – before 2012

Source: own construction. Validated by National experts.