HEALTH ECONOMICS
Health Econ. **18**: 629–643 (2009)
Published online 1 August 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hec.1397

TRANSFERABILITY INDICES FOR HEALTH ECONOMIC EVALUATIONS: METHODS AND APPLICATIONS

FERNANDO ANTONANZAS*, ROBERTO RODRÍGUEZ-IBEAS, CARMELO JUÁREZ, FLORENCIA HUTTER, REYES LORENTE and MARIOLA PINILLOS

Universidad de La Rioja, Economia y Empresa, Cigüena 60, Logrono 26004, Spain

SUMMARY

In this paper, we have elaborated an index in two phases to measure the degree of transferability of the results of the economic evaluation of health technologies. In the first phase, we have considered the objective factors (critical and non-critical) to derive a general transferability index, which can be used to measure this internal property of the studies of economic evaluation applied to health technologies. In the second phase, with a more specific index, we have measured the degree of applicability of the results of a given study to a different setting. Both indices have been combined (arithmetic and geometric mean) to obtain a global transferability index. We have applied the global index to a sample of 27 Spanish studies on infectious diseases. We have obtained an average value for the index of 0.54, quite far from the maximum theoretical value of 1. We also found that 11 studies lacked some critical factor and were directly deemed as not transferable. Copyright © 2008 John Wiley & Sons, Ltd.

Received 27 June 2007; Revised 2 April 2008; Accepted 6 June 2008

KEY WORDS: economic evaluation; transferability; generalisability

1. INTRODUCTION

Economic evaluation of health technologies started to be used as an informative tool to determine the efficiency of such technologies at the end of the 1970s. In the last decades, research dealing with the evaluation of health technologies has grown and, in Spain, there actually are six regional agencies plus a nationwide agency. Therefore, there exist available human resources that are familiar with economic evaluation techniques, read available research and elaborate reports on their own geographical areas. These agencies usually work after receiving a demand from their regional users, and then, they provide them with reports dealing with efficacy, safety, budget impact and efficiency issues related to the technologies of their interest. The publication of economic evaluation reports in Spain is growing at a similar rate to other surrounding countries (more than 30 published economic evaluations per year) according to the number of available references in different databases, with the exception of the UK, which dominates this research area. In spite of this, the publication rate of economic evaluation studies in the decision-making domain is not replicated.

Many economic evaluation reports are published in specialised journals. In order to guarantee that these studies have the necessary elements for decision-making and can be used to make more reliable the publication and the application of these studies, as well as the verification of their results, some guidelines of economic evaluation have been developed. These guidelines are publicly available and deal

*Correspondence to: Departamento de Economía y Empresa, Universidad de La Rioja, C/ La Cigüena 60, 26004 Logrono, Spain.
E-mail: fernando.antonanzas@unirioja.es

with different elements leading to the setting of quality standards in the field of economic evaluation. A comparative list of the available guidelines in the EC countries can be found in Graaf von der Shulenburg and Hoffmann (2000).

For the purpose of providing information on efficiency in a more comfortable and quick way, some databases with either public or private access have been created. These databases contain thousands of summaries structured according to a list of field headings that make it easier for the reader to understand the economic evaluation studies. Some examples of these databases are: the NHS Economic Evaluation Database (NHS EED), maintained by the University of York, with public access, and containing summaries of studies carried out and published in any country of the world, *Connoisances et Décision en Economie de la Santé* (CODECS), also with public access, containing studies relevant to France, and the Office of Health Economics database (OHE), with subscription access and shorter summaries covering any published study in the world. More recently, the European Network of Health Economics Evaluations Databases (EURONHEED), funded by a EU project, have been established. They contain more than 1500 summaries (2–3 pages each) of studies published in the UK, France, Switzerland, Belgium, Germany, The Netherlands, Italy, Spain, Portugal, Sweden and Denmark.

Currently, given the existence of many studies on different health technologies carried out in several countries, one of the more relevant issues concerns the capacity of transferring the results from one health context to another. As Oliva *et al.* (2001) point out, this issue constitutes a barrier in expanding the utilisation of the studies in an efficient manner. This transfer of results is performed routinely in the field of medicine. In this sense, for example, the results of a clinical trial carried out in one country are usually accepted in other countries without major problems, when efficacy or safety is assessed. However, if the aim of the study is to assess efficiency, the results obtained in one country are not directly applicable in another country and, usually require adjustments to take into account the new context (medical practice, costs, comparators, discount rate, etc.) before being incorporated within the local decision-making process. This is a crucial issue, as it constitutes an important barrier to extend the use of this informative tool.

So far, one usual way that makes easier the transference of the results of the economic evaluations from one jurisdiction to another is the design of multinational clinical trials so that different health environments are represented in the final outcomes. Also multilevel models and meta-regressions are used with the same intention; these models address both the transferability of the health outcomes and the resource use (i.e. the costs) of the assessed technologies. However, some health authorities may consider that the heavy data handling needed in these multilevel models reduces the reliability on the final results of the economic evaluations, and they would rather look for a particular and simpler adaptation to their context.

Boulenger *et al.* (2005), Mason and Mason (2006) and Welte *et al.* (2004) have analysed this issue of transferability and suggested some ways to check and improve the application of the results of economic evaluation studies to other jurisdictions. In Spain, given the existence of several health technology assessment agencies, it becomes very important to have studies on efficiency with easily transferable results. This will also guarantee the efficiency of the very process of performing economic evaluations from a global perspective.

In this paper, we report the development of a transferability index for published economic evaluations to help in the decision-making processes of different health contexts. The paper is structured as follows. In the following section, the literature on this issue is reviewed. In Section 3, we present an index in two phases to check the degree of transferability. We describe the relevant elements in both phases and define the index. In Section 4, we apply the index to 27 studies on infectious diseases and comment on the results. Finally, some conclusions and caveats of our methods and findings are presented.

DOI: 10.1002/hec

2. LITERATURE REVIEW: CONCEPTS AND MEASURES OF TRANSFERABILITY AND GENERALISABILITY

While growing activity of economic evaluation has generated extensive results on the efficiency of health technologies, it also raises the question of the applicability of their results to different contexts, as is the case with studies on efficacy. To the best of our knowledge, this topic began to be addressed by Drummond *et al*. (1992) and other authors (Barbieri *et al*. 2005; Cook *et al*., 2003; Hoffman *et al*., 2002; Manca *et al*., 2005; Manca and Willan, 2006; Nixon and Pang, 2000; Sculpher *et al*., 2004; Spath *et al*., 1999) have also expressed their concerns with the issue of transferability. The main concepts are still subject to debate. For this reason, we next review definitions of transferability and generalisability used by different authors.

2.1. Concepts

The study by Boulenger *et al*. (2005) pointed out that the glossaries of terms in health economics do not distinguish between 'generalisability, transferability and transportability' and often group them under the concept of 'external validity'. In their paper, they refer to generalisability as 'the degree to which the results of a study hold true in other settings', while transferability is defined as 'the data, methods and results of a given study are transferable if (a) potential users can assess their applicability to their setting and (b) they are applicable to that setting'. Thus, for these authors, transferability is a wider concept than generalisability. For Mason and Mason (2006) the term 'generalisability' considers three elements: technical quality (the author must be sure that the study refers to the original sequence of the research and that the primary economic analysis was performed with the best available methods to minimise errors and biases), applicability (referred to the extrapolation of the original setting of the clinical trial where the results were obtained, e. g. results derived from a given medical practice and costs were similar in both settings) and transferability (capacity to directly use the complete results of the economic evaluation in a setting different from the original one in which the technology was assessed). Welte *et al*. (2004), in contrast, do not explicitly define the meaning of transferability, although they implicitly refer to it as the capacity to use the results obtained in one setting in other health context.

In summary, these definitions refer to transferability as a property that either the studies or their results have. As with any other property, it can be achieved with a degree within a scale, as it happens, for instance, with quality. Therefore, to measure the degree of transferability, it is necessary to have a scale and some items relevant to transferability aspects that need to be checked. Regarding this issue, there have been so far two studies whose purpose was to propose different ways to assess the transferability of the results of published economic evaluations.

2.2. Two previous proposals to measure transferability

Boulenger *et al*. (2005) suggested check lists containing concepts that, according to interviewed experts, should be considered in the economic evaluation studies to guarantee transferable results (these check lists have guidelines which have currently been submitted for publication, as the authors have personally communicated to us). Since the transferability property should be checked in studies having a minimum standard of quality, their transferability lists include some standard elements related to the general quality of the studies (e. g. see the British Medical Journal check list available in its web site www.bmj.com). From this stand point, Boulenger *et al*. (2005) have produced two lists. The first one is wider and includes 42 items related to overall quality (including generalisability and transferability), while the second one is more specific and focuses on items related to transferability. They check whether a set of studies verify the elements of the two lists, grading 1, 0.5 or 0 each item if they respectively were fully, partially or not addressed in the study. Further, if the study does not contain information on a given item, it is also given a score of 0. When a concept is not applicable to the study, it is excluded from

DOI: 10.1002/hec

the calculation of the transferability index, which is defined as:

$$\frac{1}{n-x}\sum_{i=1}^{n} s_i \times 100$$

where n is the number of items in the check list, x the number of non-applicable items and s the score each item obtains. This particular approach was first utilised by Nixon and Pang (2000) in a review of economic evaluations published in Japan.

This index can be applied to several studies of the same technology and averaged so that it will summarise the state-of-the art of transferability issues for a health area, for a group of countries, etc.

However, Boulenger *et al.* (2005) do not suggest a minimum threshold to qualify the study (or group of studies) as being transferable, do not apply weighting to items of the check list according to their importance and do not select knock-out items whose null scores would render a study as 'non-transferable'.

Welte *et al.* (2004) suggested a decision chart that establishes a sequence of elements to measure transferability. They considered general and specific knock-out elements such that if a study lacked any of these critical elements, it would be deemed as non-transferable, making it not necessary to check additional items. Further, they suggested, besides analysing each item in the list, to check whether it would be possible (with or without a lot of effort) to adjust the original study to other contexts to make it transferable. Although this approach may seem a little harsh, it could be smoothed by double-checking with the authors whether the critical elements should in fact imply knocking out the study for transferability purposes.

However, Welte *et al.* (2004) do not build an index to summarise the results but use a qualitative approach. This feature makes it more difficult to aggregate the results for a group of studies related to a given technology and to establish a threshold to qualify a study as being transferable. The introduction of critical factors is interesting, however, as it allows the elimination of non-transferable studies with only little effort. Also, the assessment by experts of how difficult it is to adapt each factor to the new health setting is important in itself as the decision makers are able to appreciate in advance the necessary effort required to take advantage of the findings of previous studies.

3. A PROPOSED TRANSFERABILITY INDEX

The studies carried out by Boulenger *et al.* (2005) and Welte *et al.* (2004) can be extended to obtain a measurement of the transferability that summarises the results, weights items according to their relative importance and considers the possibility of stopping the checking process when some critical factor is identified.

We propose a numerical index based on weighted objective and subjective elements that takes into account the decisive influence of the critical factors regarding the transferability of a given study.

The capacity to transfer the results of a given study of economic evaluation of health technologies is an objective question, and therefore, it does not depend on the setting to which the results are transferred. However, the capacity to use and apply such results to a different context depends on the specific characteristics of the context itself. The assessment of this issue (the transferability of the study) requires that the analyst, with real information about the health setting to which the results may be transferable, be able to evaluate how difficult it is to make the needed adjustments to apply the results in the new setting.

We have therefore designed a new index in two phases. In the first phase, we examine the objective factors in a given study, and determine its degree of general transferability. In the second phase, we evaluate the level of difficulty that exists in applying or adopting the information in the original study to the new setting, and we obtain a second index that measures the degree of specific transferability. It is

DOI: 10.1002/hec

useful to recall that this idea perfectly fits the definition of transferability given by Boulenger *et al.* (2005): the data, methods and results of a given study are transferable if (a) the potential users can assess their applicability to their settings (first phase), and (b) the results are applicable (or adjustable) to their settings (second phase).

The index we have derived in the first phase includes some critical objective factors – such that a study can be initially excluded if it fails to meet certain standards of quality – (as suggested by Welte *et al.* 2004), and some non-critical objective factors, whose assessment will allow the classification of studies according to their scores.

The index derived in the second phase measures the degree of applicability of each study to the specific setting of the decision maker. The value taken by the index will depend on the available information on the setting as well as the ability to adapt the results of the original study to the new setting. This index is, therefore, subjective and its value will depend on the specific characteristics of both the concrete setting and the decision maker. Thus, the concept of transferability does not have a universal meaning, as the index value for a study might be different depending on the values given to the relevant factors in the second phase.

3.1. The first phase: General Transferability Index

Let us assume that there are n critical objective factors and m non-critical objective factors. Let IT_1 be the general transferability index for a given study:

$$\mathrm{IT}_1 = \beta \sum_{j=1}^{m} \alpha_j \mathrm{FO}_j$$

where

$$\beta = \begin{cases} 0 & \text{if any critical factor is applicable to the study} \\ 1 & \text{otherwise} \end{cases}$$

The variable FO_j takes the value 1 if the non-critical objective factor j is fully specified in the study, 0.5 if it is partially specified and 0 if it is not specified (similar to the proposal by Boulenger *et al.* 2005).

The coefficient α_j is the weight for the non-critical factor j. Factor relevance can be assessed by a committee of experts. Let h be the number of experts and v_{js} be the value given by expert s to factor j in a Likert scale from zero to four, $s = 1, 2, ..., h$ and $j = 1, 2, .., m$. Let $w_j = \sum_{s=1}^{h} v_{js}$. The average value given by the committee of experts to factor j is $\bar{w}_j = \frac{w_j}{h}$. Thus, the weight assigned to factor j is $\alpha_j = \bar{w}_j / \sum_{j=1}^{m} \bar{w}_j$.

Transferability studies have usually weighted each factor equally – in the sense that they do not weight them at all – although it is acknowledged that relevance differs across factors. In this paper, we intend to assign weights to the factors based on their relevance for transferability. Weights are usually based on opinions given by experts or they are fixed by administrative authorities, competent in the matter, taking into account strategic issues (authorities state the importance of each element with the weights and guide the users' actions, as it happens, for example, when the activity of universities is evaluated and summarised). In our case, as a preliminary approach to obtain the relative importance of each factor, we have sent a questionnaire to the seven Health Technology Assessment Agencies in Spain. We consider that future research should be devoted to this issue to determine the weighs more accurately. An alternative method to assign weights would be to apply the proposal suggested by Chiou *et al.* (2003) to the transferability context. Such a proposal refers to quality and is based on conjoint analysis. However, this proposal can be difficult to implement as the correlations between global assessments of transferability and the experts' partial opinions on the relevance of each factor are difficult to obtain.

DOI: 10.1002/hec

Notice that the index takes the value 0 if $\beta = 0$. In this case, the study would be considered as not transferable. Thus, it would not make sense to check specific transferability in the second stage, unless additional information could be obtained from the authors.

The critical factors are:

(1) The relevant parameters needed to calculate the ratio cost/effectiveness are not given in the study.

(2) The quality of the study is low (it contains mistakes, e.g. costs are double counted, as Welte *et al.* (2004) pointed out). Study quality should be checked together with the general objective factors (non-critical objective factors) listed later. If two or more of these factors received a score of zero, the study would be considered as not transferable. Regarding the quality of the economic evaluation studies, Chiou *et al.* (2003) assigned weights to 16 items. The following items were considered as the most relevant to assess the quality of a given study as they received 7 or more points (over 100): (a) the study objectives are presented in a clear, specific and measurable manner; (b) the variable estimates used in the analysis come from the best available source; (c) the measurement of costs is appropriate and the methodology for the estimation of quantity and unit costs is clearly described; (d) the health outcome measures are based on valid and reliable scales, when available. Otherwise, the scales used in the study must be fully justified; (e) the economic model (including its structure), study methods and components of the costs and effectiveness are presented in a clear manner and (f) the conclusions and recommendations of the study are justified and based on the study results.

Some of these items (a, c, d) will be mentioned later as general objective factors. Notice that Boulenger *et al.* (2005) also included them in their relevant factors list to assess study transferability.

The non-critical factors we propose to assess objective or general transferability are similar to the list of 16 factors considered by Boulenger *et al.* (2005). Most of the informative content in their list coincide with the information available in the template for full economic evaluation of the EURONHEED database.

1. *HT1. Is the intervention described in sufficient detail?
2. *HT2. Is the comparator described in sufficient detail?
3. *SE2. Is the country in which the economic study took place clearly specified?
4. P1. Did the authors correctly state the perspective for the economic analysis?
5. SP1. Is the target population of the health technology clearly stated or can it be inferred by reading the article?
6. SP3. Does the article provide sufficient detail about the study sample?
7. *E5. Have the principal estimates of effectiveness measures been reported?
8. E7. Are the results of a statistical analysis of the effectiveness results provided?
9. B5. Is the level of reporting of benefit data adequate (incremental analysis, statistical analyses)?
10. *C1. Are the cost components used in the analysis presented?
11. C5. Are unit price for resources given?
12. *C6. Are costs and quantities reported separately?
13. *C7. Is the price year given?
14. *C9. Is the currency unit reported?
15. *S1. Are quantitative and/or descriptive analyses conducted to explore variability from place to place?
16. *O1. Did the authors discuss the generalisability of their results?

(The codes as well as the descriptions of the items used in the index are similar to those considered by Boulenger *et al.* (2005) and will be used later in the tables of Section 4.) The 10 items with an asterisk were also chosen as most relevant to assess transferability by 16 health economists participating in the EURONHEED project.

DOI: 10.1002/hec

The value the index IT_1 takes for a non-transferable study is 0, as in that case, β is 0. Otherwise, the index takes a strictly positive value in the interval (0,1].

3.2. The second phase: Specific Transferability Index

In the second phase, we build an index to measure the degree of specific transferability of a given study. The index is also based upon a list of items whose values depend on how easy the data, methods and results of a given study can be applied to the specific health context. We use a Likert scale with five levels (from 0 to 4) to assign the values. The index, as in the previous phase, takes into account critical subjective factors as well as non-critical subjective factors. Critical factors are previously checked, and if any of them receives the value 0, the study is deemed to be not transferable. Notice that a study with a high degree of general transferability can be not transferable to a specific context as its degree of specific transferability is zero.

The specific transferability index in the second phase IT_2 is:

$$IT_2 = \delta \sum_{i=1}^{r} \frac{FS_i}{4r}$$

where FS_i is the value of the subjective factor i in the Likert scale (from 0 to 4), the denominator $4r$ is the maximum aggregate value of all subjective factors and δ is:

$$\delta = \begin{cases} 0 & \text{if any critical factor is applicable to the study} \\ 1 & \text{otherwise} \end{cases}$$

The subjective factors in the second stage are not weighted due to the specific nature of the index. Weights can obviously be included, but they will be different for each decision maker according to his preferences on the importance of each subjective factor.

The subjective critical factors, whose lack of compliance makes transferability of a given study not possible to another health context, are:

1. The evaluated technology is used in the new health context. (This factor will not be taken into account if the economic evaluation is carried out to obtain relevant information before the potential use of the technology in the new context.)
2. The comparator is available or used in the new context.
3. Treatment and comparator data, as well as relevant epidemiological parameters for the technology, are valid in the new context.
4. The study perspective coincides with that used in the new context.

Following some of the ideas by Welte *et al*. (2004), we propose a general list of non-critical subjective factors that we consider relevant to assess the final applicability of a give study to a new health context. Final users could eventually adapt this list to their specific features. The proposed non-critical subjective factors are the following:

1. Cost components correspond to the medical practice related to the evaluated technology in the original study. If medical practice differs in the new context, additional costs components must be taken into account. (For example, if in the original study all patients were explored with the CT-scanner but in the new context, both MRI and CT-scanner are used; a new cost component should be included, which can be unavailable.)
2. The model connecting variables and parameters can be adapted to the new context.
3. Life expectancy is similar in both contexts.
4. Health-status preferences are similar in both contexts. (applicable to cost/utility analyses)

DOI: 10.1002/hec

5. Productivity measures are similar in both contexts (applicable to cost minimisation and cost/benefit analyses).
6. The evolution of the disease is similar in both contexts.
7. The applied discount rate is similar in both contexts.
8. Costs and health effects data are presented in current and discounted units.

The value the index IT_2 takes for a non-transferable study is 0, as in that case, δ is 0. Otherwise, the index takes a strictly positive value in the interval (0,1].

3.3. The Global Transferability Index

The global transferability index, IT, combines the indices IT_1 and IT_2 through a mathematical formula. The index can be the simple or weighted arithmetic mean, the geometric mean or it may adopt any other combination of both indices:

$$IT = IT_1^a \, IT_2^b \tag{1}$$

$$IT = \frac{aIT_1 + bIT_2}{a+b} \tag{2}$$

$$IT = IT_1(1 + IT_2) \tag{3}$$

where $a \in (0, 1]$ and $b \in (0, 1]$ are constants and denote the importance given to the first and second phase. Both values are subjectively determined. The above formulas would be applicable as long as none of the indices are zero.

The transferability index in (1) is the geometric mean for $a = b = 0.5$. The formula (2) is the arithmetic mean for $a = b$. The index in (3) can take values in the interval (0,2]. It differs from the other two formulations, for which the maximum value the index can take is 1. However, it takes into account the relationship between the two phases as both index enter the formula in a multiplicative way.

Lacking a better criterion, we propose to use formulations (1) and (2) with $a = b = 0.5$, which is equivalent to calculating the geometric and arithmetic means of the first and second phase indices. Whatever the formulation finally adopted, we need to specify a rule to break ties when several studies obtain the same value for the transferability index. A possible solution would be to choose the study closer to the fully transferable study. The degree of approximation of a given study to the fully transferable one can be measured by the Euclidean distance between the points (1,1) and (IT_1, IT_2), where (IT_1, IT_2) are the indices of the study:

$$d = \sqrt{(1 - IT_1)^2 + (1 - IT_2)^2}$$

The lower the distance, the more transferable the study is.

However, we can mention at this point that some analysts may consider that it is better to take both indices as separate concepts due to their different informative contents. Therefore, we suggest to present both the disaggregated figures of the indices together with the global index value.

Before presenting the results of our empirical work we would like to highlight two issues dealing with the selection of the factors for the index. One refers to the little importance we have given to how uncertainty is handled in economic evaluations. The transfer of the results of an economic evaluation is easier when a sensitivity analysis is performed. However, if this analysis does not include all the multiple combinations of parameters values and variables, other jurisdictions will not be fully represented in the final results and hence they will not have all the information needed to adopt a decision based on efficiency grounds. The other issue is related to the similar treatment we give in the transferability index to the economic evaluations regardless of the type of model they are based on (decision or individual patient data models); we believe that this difference in the nature of the clinical data will scarcely affect

DOI: 10.1002/hec

the final degree of transferability, because most of the incremental cost-effectiveness ratios refers to gains of life years accruing in the future, while the studies based on patient data derived from clinical trials have a much shorter time horizon, and consequently a model would be needed to extrapolate the final outcomes.

4. EMPIRICAL RESULTS

In order for the index to be applicable, we need to have available a set of weights assigned to the non-critical factors in the first stage. As it was mentioned before, a questionnaire was sent to the seven Health Technology Assessment Agencies in Spain, who responded to it. Their answers were then averaged, and are presented in Table I. As we can see, the factor with the largest weight (0.0802) is P1 (do the authors state the study perspective?) and the factor with the lowest weight (0.0428) is C1 (are costs components specified in the study?). Notice that all weights belong to a narrow interval. There were no relevant incidents in the questionnaires administrative process (except a suggestion to include the discount rate as a non-critical objective factor, made by one expert. The discount rate is one of the non-critical specific factors).

Once the weights were determined, we applied the first index to a set of economic evaluations of health technologies referred to infectious diseases. We have used the respective abstracts listed in the EURONHEED database. We chose infectious diseases as they deal with preventive and curative technologies and include models with diverse time horizons, which allow the checking of different aspects included in the proposed index (in other health areas, studies tend to be more homogeneous). The total number of available studies in Spain is 27. The studies that have been considered appear at the bottom of Table III.

The second stage index has been derived through a simulation as we lacked an agency responsible to undertake an assessment. For all studies, the authors revised all factors and assigned, based on their experience, values to each of them. As all the studies were relevant to Spain, we have opted to apply the index from the perspective of a hypothetical regional agency in charge of evaluating health technologies and assuming the availability of information these agencies have regarding their reference context. This enabled us to check the relevant items and obtain the global transferability index.

Regarding the critical objective factors, nine studies have been deemed as not transferable as some critical items received a value of 0. Consequently, six of them were excluded as they lacked a clear and transparent model, two did not measure costs correctly and one study was excluded, as the relevant parameters required to calculate the cost-effectiveness ratio were not specified.

The non-critical objective factors are summarised in Table II. All studies include three factors (country, costs components and currency), while none considers either the possibility that the results were transferable to other contexts or how difficult it could be to undertake this. Most of the studies (range from 18 to 24, over a maximum of 27) take into account the usual methodological factors (perspective, target population, effectiveness measures, unit resource costs and reference year). However, there are six factors whose frequencies are low.

Table I. Weights for the non-critical objective factors

	HT1	HT2	SE2	P1	SP1	SP3	E5	E7
α	0.0615	0.0588	0.0615	0.0802	0.0695	0.0588	0.0722	0.0615
	B5	C1	C5	C6	C7	C9	S1	O1
α	0.0642	0.0428	0.0588	0.0695	0.0588	0.0642	0.0642	0.0535

Source: Elaborated by authors.

DOI: 10.1002/hec

Table II. Non-critical objective factors

		Number of Answers			
	NCOF	'Yes'	'Partial'	'No'	N/A
1	SE2	27	0	0	0
2	C1	27	0	0	0
3	C9	27	0	0	0
4	SP1	24	2	1	0
5	E5	23	0	4	0
6	P1	20	1	6	0
7	C7	19	0	8	0
8	C5	18	1	8	0
9	HT2	13	8	3	3
10	C6	13	3	11	0
11	HT1	12	13	2	0
12	SP3	7	5	3	12
13	B5	7	3	4	13
14	E7	6	2	19	0
15	S1	0	0	27	0
16	O1	0	0	27	0

Source: Elaborated by authors.

Table III. Non-critical subjective factors (Likert Scale 0–4)

Study Number	1	2	3	4	5	6	7	8	Sum
1	4	2	2	N/A	N/A	4	N/A	4	16
2	4	0	2	N/A	N/A	4	4	4	18
3	4	0	4	N/A	N/A	4	N/A	4	16
4	4	2	2	N/A	N/A	4	N/A	4	16
5	4	3	4	N/A	N/A	4	N/A	4	19
6	4	3	4	N/A	N/A	4	N/A	4	19
7	4	1	3	N/A	N/A	4	N/A	4	16
8	4	3	4	N/A	N/A	N/A	N/A	4	15
9	4	3	4	4	N/A	N/A	2	3	20
10	4	3	4	N/A	N/A	N/A	N/A	3	14
11	4	4	2	N/A	N/A	N/A	N/A	4	14
12	4	4	3	N/A	N/A	N/A	N/A	4	15
13	3	2	2	N/A	N/A	N/A	N/A	2	9
14	4	3	3	N/A	N/A	N/A	N/A	3	13
15	4	3	3	N/A	N/A	N/A	N/A	4	14
16	4	3	3	N/A	N/A	N/A	N/A	2	12
17	4	N/A	4	N/A	N/A	N/A	N/A	3	11
18	4	N/A	4	N/A	3	1	4	3	19
19	4	N/A	4	N/A	N/A	N/A	N/A	3	11
20	4	0	4	N/A	N/A	N/A	N/A	3	11
21	4	N/A	4	N/A	N/A	N/A	N/A	4	12
22	4	4	4	N/A	N/A	N/A	N/A	4	16
23	4	4	4	N/A	N/A	N/A	N/A	4	16
24	4	0	3	N/A	N/A	N/A	N/A	3	10
25	4	0	4	N/A	N/A	N/A	N/A	4	12
26	4	0	4	N/A	N/A	N/A	N/A	4	12
27	4	0	4	N/A	N/A	N/A	N/A	4	12

N/A: No applicable Source: Elaborated by authors. The 27 studies are: Abad *et al.* (2000), Badia *et al.* (1999), Bonal and Bravo (1999), Calvet *et al.* (2001), Catalán *et al.* (1995), Domínguez-Gil *et al.* (1988), Fernández Arjona *et al.* (1996), Fernández Muñoz *et al.* (1998), García de Villaescusa *et al.* (2003), García and Ingesta (1997), Gené *et al.* (2000), Gomollón *et al.* (2000), González *et al.* (2004), Gómez Vargas *et al.* (1998), Jiménez *et al.* (1999), Martínez Jabaloyas (1997), Plans *et al.* (1995), Plans (2002), Plaza and de los Santos (2003), Redondo and Nocea (2003), Ricart *et al.* (2000), Rubio-Terrés *et al.* (2003), Rubio-Terrés (2004), Sanz *et al.* (2005), Sábat *et al.*(1998), Sánchez García *et al* .(1998) and Soto *et al.* (2003). We have randomly assigned a number to each study. Such a number does not coincide with the position of each study in the alphabetical list above.

DOI: 10.1002/hec

With regard to critical subjective factors, two studies are not transferable, as both were assigned 0 in the factor related to effectiveness data and relevant parameters in the new context.

For the non-critical subjective factors (Table III), several of them have not been considered due to the nature of the analysed studies. For example, the factor dealing with health status preferences is not applicable to cost minimisation or cost-effectiveness studies, the measure of productivity is not relevant in studies that do not consider indirect costs and the discount rate is not used in studies with short time horizons. The remaining factors have obtained high scores, which is probably due to the fact that most of the studies refer to similar health contexts. Nevertheless, the exercise allowed us to obtain values for the index in the second phase as well as to detect how difficult is to interpret various factors. In this sense, the authors believe it is necessary to produce a guide to help understand each factor and hence its valuation in each possible situation the index is applied to.

Table IV shows the final results for the global transferability index using the arithmetic and geometric means. It should be mentioned here that, due to the lack of any of the critical factors, 11 studies were eliminated (nine in the first phase and two in the second one). Therefore, we have calculated the value of the index for the 16 remaining studies. As can be seen in Table IV, the values hardly show any variation, regardless of the formulation used. The values of the index using the geometric mean range from 0.39 to 0.65, while those calculated with the arithmetic mean range from 0.41 to 0.66. The transferability of the analysed studies is relatively low (five studies obtained values between 0.40 and 0.50, seven studies between 0.51 and 0.60 and four studies between 0.61 and 0.70, regardless the formula

Table IV. Transferability indices

Study Number	General Index IT_1	General Index IT_1 ($\beta = 1$)	Boulenger *et al.* index	Specific Index IT_2	GLOBAL INDEX GEOMETRIC MEAN ($a = b = 0{,}5$)	GLOBAL INDEX ARITHMETIC MEAN ($a = b = 0{,}5$)
1	0.6671	0.6671	0.6333	0.5000	0.5775	0.5836
2	0.7634	0.7634	0.8000	0.5625	0.6553	0.6629
3	0.5655	0.5655	0.6429	0.5000	0.5317	0.5328
4	0.7313	0.7313	0.7667	0.5000	0.6047	0.6156
5		0.6257	0.7143			
6	0.5374	0.5374	0.5667	0.5938	0.5649	0.5656
7	0.5254	0.5254	0.5313	0.5000	0.5125	0.5127
8	0.5414	0.5414	0.5667	0.4688	0.5038	0.5051
9		0.5441	0.6071			
10		0.5401	0.5313			
11		0.5414	0.5667			
12	0.8503	0.8503	0.8438	0.4688	0.6313	0.6595
13	0.5481	0.5481	0.5625	0.2813	0.3926	0.4147
14		0.5321	0.5625			
15	0.5227	0.5227	0.6333	0.4375	0.4782	0.4801
16		0.4893	0.5333			
17	0.6725	0.6725	0.7500			
18	0.7032	0.7032	0.7857	0.5938	0.6462	0.6485
19	0.5441	0.5441	0.6071	0.3438	0.4325	0.4439
20	0.4866	0.4866	0.5333	0.3438	0.4090	0.4152
21		0.7594	0.8000			
22	0.6979	0.6979	0.7333	0.5000	0.5907	0.5989
23	0.6631	0.6631	0.7000	0.5000	0.5758	0.5816
24	0.6377	0.6377	0.6667	0.3125	0.4464	0.4751
25		0.6845	0.6875			
26	0.5481	0.5481	0.6000			
27		0.5548	0.5625			
Media	0.6225	0.6103	0.6477	0.4629	0.5346	0.5435
SD	0.1012	0.0966	0.0972	0.0956	0.0847	0.0836

Source: Elaborated by authors.

DOI: 10.1002/hec

used). The mean value of the index is between 0.534–0.543. We have also applied the formula used by Boulenger *et al.* (2005) to the 27 studies. Note that these authors only consider the first phase factors (the so-called objective factors) without any weighting, and do not exclude any study due to the absence of the critical factors. We find that the mean value of the general index IT_1 calculated with this approach is 0.64. Our general transferability index (IT_1) without the application of the knock-out criterion (i.e. assuming that $\beta = 1$ for each study) is then comparable to the one by Boulenger *et al.* (2005), although we include the weighs for the items. The result for the mean value of the index is 0.61 (statistically, the difference in mean values is not significant at the usual 0.05 level). When the knock-out criterion in the first phase is taken into account the mean value for the general index is 0.62 (again not statistically significant). We think that the similarities in the results are due to the weighs assigned by the experts to the items that do not show big differences (see Table I). Further, when the knock-out criterion is applied, the mean value of the index again does not change too much, as both the numerator and the denominator are reduced (the former, because the index is not calculated for the omitted studies and the latter for the lower number of the studies considered).

As other authors that have studied transferability, we have not set a threshold value above which a given study would be deemed to be transferable. In our case, there are a technical and a political justification. As it was mentioned earlier, a study can have a high score in the first phase meaning that it has a high potential degree of transferability but however, its features may not fully match the actual requirements in another setting. Consequently, it will score low in the second phase and the global threshold will not be meaningful for this jurisdiction. Nevertheless, the same study could be meaningful in another jurisdiction; therefore, a fixed threshold value would not be universally acceptable and useful. In regard to the political reason, health economists perform economic evaluations (they adapt procedures and methods, assess specific health technologies and so on) but health authorities take the final decisions on the utilisation of the assessments and on the technologies. As health economists we consider that health authorities, in a broad sense, should select the threshold value of the index that better fits their purposes. As can be seen in Table IV, the main difference in the final results appears when the comparison is done between the first phase and the global indices (statistically significant); the reason being that only the global index takes into account the subjective second phase.

Finally, it is important to mention that the Euclidean distance has not been used due to the fact that no two studies referred to the same technology and with the same comparators had the same value for the index.

5. CONCLUSIONS

In this paper, we have developed an index in two phases to measure the degree of transferability of the results of economic evaluation studies. In the first phase, we took into account the objective factors (critical and non-critical) to obtain a general transferability index, which can be used to measure this internal property of studies of economic evaluation applied to health technologies. In the second phase, we have built a specific index to measure the applicability of the data, methods and results of the original studies to a different context. We have shown that the degree of applicability depends on the characteristics of the decision maker and his ability to use the information contained in the original study. Therefore, the index in the second phase must be tailored to the specific context where the original study is to be applied.

The index, in its two phases, was applied to 27 Spanish studies dealing with infectious diseases. These studies obtained, on average, a value of 0.54. Thus, all of them are quite far from the maximum theoretical value of 1. Importantly, it must be mentioned that 11 out of 27 studies were excluded as one or more critical factors were assigned a value of zero.

DOI: 10.1002/hec

Our analysis suggests that more attention should be paid to aspects related to the degree of transferability of the results of published economic evaluations. This will not only help to increase efficiency with regard to the utilisation of resources in the studies of economic evaluation but also make the decision-making process easier in dealing with the adoption of health technologies in different contexts.

ACKNOWLEDGEMENTS

The authors thank Professor John Nixon for his valuable comments and suggestions. This research was partially funded by the Fondo de Investigaciones Sanitarias (Spanish Health Research Fund) project number PI060282.

REFERENCES

Abad F, Calbo F, Zapater P, Rodríguez-Vilanova F, García-Pérez LE, Sacristán JA. 2000. Comparative pharmacoeconomic study of vancomycin and teicoplanin in intensive care patients. *International Journal of Antimicrobial Agents* **15**(1): 65–71.

Badia X, Brosa M, Casado A, Segú JL, Álvarez A. 1999. Análisis coste efectividad de estrategias de diagnóstico-tratamiento del ulcus péptico asociados a Helicobacter pilori en atención primaria. *Atención Primaria* **24**: 344–351.

Barbieri M, Drummond M, Willke R, Chancellor J, Jolain B, Towse A . 2005. Variability of cost-effectiveness estimates for pharmaceuticals in Western Europe: lessons for inferring generalisability. *Value and Health* **8**: 10–23.

Bonal J, Bravo P. 1999. Estudio farmacoeconómico imipenem monovial/meropenem. *Revista Española de Farmacoeconomía* **5**: 21–26.

Boulenger S, Nixon J, Drummond M, Ulmann P, Rice S, Pouvourville G. 2005. Can economic evaluations be made more transferable? *European Journal of Health Economics* **6**: 334–346.

Calvet X, Gené E, López T, Gisbert JP. 2001. What is the optimal length of proton pump inhibitor-based triple therapies for H. pylori? A cost-effectiveness analysis. *Alimentary Pharmacology and Therapeutics* **15**(7): 1067–1076.

Catalán JI, Juan J, Font I, Jiménez NV. 1995. Impacto económico de un programa farmacoterápico coordinado centrado en aminoglucósidos. *Revista Española de Farmacoeconomía* **2**: 21–26.

Chiou CF, Hay JW, Wallace JF, Bloom BS, Neumann PJ, Sullivan SD, Yu HT, Keeler EB, Henning JM, Ofman JJ. 2003. Development and validation of a grading system for the quality of cost-effectiveness studies. *Medical Care* **41**(1): 32–44.

Cook JR, Drummond M, Glick H, Heyse JF. 2003. Assessing the appropriateness of combining economic data from multinational clinical trials. *Statistics in Medicine* **22**:1855–1876.

Domínguez-Gil A, Rubio C. 1998. Análisis coste efectividad del tratamiento de las infecciones intra-abdominales con piperaciclina/tazobactam, en comparación con imipenem/cilastatina. *Revista Española de Farmacoeconomía (Marzo)*: 15–24.

Drummond M, Bloom BS, Carrin G, Hillman AL, Hutchings HC, Knill-Jones RP, de Pouvourville G, Torfs K. 1992. Issues in the cross-national assessment of health technology. *International Journal of Technology Assessment in Health Care* **8**: 671–682.

Fernández Arjona M, Herruzo R, Gómez-Sancha F, Nieto S, Rey J. 1996. Economical Saving due to prophylaxis in the prevention of surgical wound infection. *European Journal of Epidemiology* **12**(5): 455–459.

Fernández Muñoz J, López A, Zapater P, Abad F. 1998. Análisis coste-efectividad de la erradicación del Helicobacter pylori como tratamiento de la úlcera duodenal. *Anales de Medicina Interna* **15**: 515–522.

García de Villaescusa R, Barallobre J, Staginnus U. 2003. Coste-efectividad de las transfusiones de componentes plaquetarios preparados con tratamiento de inactivación de patógenos en España. *Revista Española de Economía de la Salud* **2**(3): 166–175.

García FM, Ingesta A. 1997. Análisis coste efectividad de diferentes pautas de erradicación de Helicobacter pylori en pacientes diagnosticados de úlcera doudenal en un área de salud. *Farmacia Clínica* **14**: 570–580.

DOI: 10.1002/hec

Gené E, Calvet X, Azagra R. 2000. Diagnosis of Helicobacter pylori after triple therapy in uncomplicated duodenal ulcers- a cost-effectiveness analysis. *Alimentary Pharmacology and Therapeutics* **14**: 433–442.
Gomollón F, Valdepérez J, Graus Bellido R, Fuentes J, Barrera F, Malo J, Tirado M, Simon MA. 2000. Análisis coste-efectividad de dos estrategias de erradicación de Helicobacter pylori: resultados de un estudio prospectivo y aleatorizado en atención primaria. *Medicina Clínica* **115**: 1–6.
González P, Rejas J, González G, Montesinos A. 2004. Estudio farmacoeconómico del tratamiento de la exacerbación aguda de la bronquitis crónica (EABC) durante un periodo de 6 meses en atención primaria: azitromicina frente a tratamiento estándar. *Pharmacoeconomics Spanish Research Articles* **1**(3): 109–121.
Gómez Vargas J, Gómez J, Ruiz J, Simarro E, San Miguel T, Canteras M, Valdés M. 1998. Influencia de la protocolización razonada y consensuada en el uso de ceftriaxona en un hospital general. *Revista Española de Quimioterapia* **11**(4): 327–332.
Graaf von der Schulenburg JM, Hoffmann C. 2000. Review of European guidelines for economic evaluation of medical technologies and pharmaceuticals. *HEPAC* **1**: 2–8.
Hoffman C, Stoykova BA, Nixon J, Glanville JM, Misso K, Drummond M. 2002. Do decision makers find economic evaluations useful? Result of recent focus group research in the UK. *Value Health* **5**: 71–78.
Jiménez F, Guallar-Castillón P, Rubio C, Guallar E. 1999. Cost benefit analysis of Haemophylus influenzae type B vaccination in children in Spain. *Pharmacoeconomics* **15**(1): 75–83.
Manca A, Rice N, Sculpher MJ, Briggs AH. 2005. Assessing the generalisability by location in trial-based cost-effectiveness analysis: the use of multilevel models. *Health Economics* **14**(5): 471–485.
Manca A, Willan AR. 2006. 'Lost in translation': accounting for between-country differences in the analysis of multinational cost-effectiveness data. *Pharmacoeconomics* **24**(11): 1101–1119.
Martínez Jabaloyas JM. 1997. Estudio comparativo coste eficacia ceftriaxona vs. Cefotaxima en el tratamiento de las infecciones urinarias complicadas. *Actas Urológicas Españolas* **21**(7): 668–674.
Mason JM, Mason AR. 2006. The generalisability of pharmacoeconomic studies: issues and challenges ahead. *Pharmacoeconomics* **24**(10): 937–945.
Nixon J, Pang F. 2000. Economic evaluations in Japan: a review of published studies, methodological issues and practice. In: *Probabilistic safety assessment and management 5*. Vol. 3. Kondo S, Furuta K (eds). Universal Academy Press Inc.: Tokyo, 2077–2084.
Oliva J, del Llano J, Antoñanzas F, Juárez C, Rovira J, Figueras M, Gervás J. 2001. Impacto de los estudios de evaluación económica en la toma de decisiones sanitarias en Atención Primaria. *Cuadernos de Gestión* **7**: 192–202.
Plans P. 2002. Coste-efectividad de la vacunación antineumocócica 23-valente en Cataluña. *Gaceta Sanitaria* **16**(5): 392–400.
Plans P, Garrido P, Salleras L. 1995. Coste-efectividad de la vacunación neumocócica en Cataluña. *Revista Española de Salud Pública* **69**: 409–417.
Plaza G, de los Santos G. 2003. Análisis de coste-efectividad del tratamiento de la otitis media serosa infantil. *Otorrinolaringológica Española* **54**(5): 316–324.
Redondo E, Nocea G. 2003. Análisis coste-efectividad de ertapenem (Invanz) en infecciones intraabdominales frente a piperacilina/tazobactam. *Revista Española de Economía de la Salud* **2**(6): 306–310.
Ricart E, Soriano G, Novella MT, Ortiz J, Sàbat M, Kolle L, Sola-Vera J, Miñana J, Dedéu JM, Gómez C, Barrio JL, Guarner C. 2000. Amoxiliclin–clavulanic acid versus cefotaxime in the therapy of bacterial infections in cirrhotic patients. *Journal of Hepatology* **32**(4): 596–602.
Rubio-Terrés C. 2004. Análisis farmacoeconómico del tratamiento de pacientes con exacerbación aguda de la bronquitis crónica con telitromcina o cefuroxima-axetilo. *Revista Clínica Española* **204**(11): 567–573.
Rubio-Terrés C, Cots JM, Domínguez-Gil A, Herreras A, Sánchez Gascón F, Chang J, Trilla A. 2003. Análisis farmacoeconómico del tratamiento de la neumonía adquirida en la comunidad con telitromicina o claritromicina. *Revista Española de Quimioterapia* **16**(3): 295–303.
Sanz MA, Bermúdez A, Rovira M, Besalduch J, Pascual MJ, Nocea G, Sanz-Rodríguez C. 2005. Imipenem/cilastatin versus piperacillin/tazobactam plus amikacin for empirical therapy in febrile neutropenic patients: results of the COSTINE study. *Current Medical Research and Opinion* **21**(5): 645–655.
Sábat M, Kolle L, Soriano G, Ortiz J, Pamplona J, Novella MT, Villanueva C, Sainz S, Torras J, Balanz J, Guarner C. 1998. Parenteral antibiotic prophylaxis of bacterial infections does not improve cost-efficacy of oral norfloxacin in cirrhotic patients with gastrointestinal bleeding. *American Journal of Gastroenterology* **93**: 2457–2462.
Sánchez García M, Cambronero JA, López J, Cerdá E, Rubio J, Gómez MA, Núñez A, Rogero S, Oñoro JJ, Sacristán JA. 1998. Effectiveness and cost of selective decontamination of the digestive tract in critically ill intubated patients. *American Journal of Respiratory and Critical Care Medicine* **158**: 908–916.

DOI: 10.1002/hec

Sculpher MJ, Pang FS, Manca A, Drummond M, Golder S, Urdahl H, Davies LM, Eastwood A. 2004. Generalisability in economic evaluation studies in health care: a review and case studies. *Health Technology Assessment* **8**(49): iii–iv,1–192.

Soto J, Grau S, Mateu-de Antonio J. 2003. Eficiencia de voriconazol (V-fend) en el tratamiento de la aspergilosis invasiva. *Revista Española de Economía de la Salud* **2**(6): 358–362.

Spath HM, Carrère MO, Fervers B, Phillip T. 1999. Analysis of the elegibility of published economic evaluations to a given health care system. *Health Policy* **49**: 161–177.

Welte R, Feenstra T, Jager H, Leidl R. 2004. A decision chart for assessing and improving the transferability of economic evaluation results between countries. *Pharmacoeconomics* **22**(13): 857–876.

DOI: 10.1002/hec