Instituto de Evaluación Tecnológica en Salud

Guidelines for the economic evaluation of healthcare technologies in Colombia: technical support documents

August 2014

Scientific Coordinators

Rita Faria. Centre for Health Economics. University of York.
Aurelio Mejía Mejía. Instituto de Evaluación Tecnológica en Salud.

Authors

Centre for Health Economics, University of York, York, United Kingdom.
Rita Faria.
Ana Duarte.
Claire McKenna.
Paul Revill.

Institute of Psychiatry, King's College, London, United Kingdom.
Maria Cary.
Sarah Byford.

This document should be cited as: Faria R, Mejia A (Eds.). Guidelines for the economic evaluation of healthcare technologies in Colombia: technical support documents. Instituto de Evaluación Tecnológica en Salud. Bogotá D.C.: IETS; 2014.

Contact

Instituto de Evaluación Tecnológica en Salud - IETS
Autopista Norte 118 - 30 Of. 201
Bogotá, D.C., Colombia.
www.iets.org.co
subdireccion.etes@iets.org.co

Tabla de Contenido

1. Measurement and valuation of health outcomes for economic evaluation and decision-making in healthcare

Maria Cary
Sarah Byford
Rita Faria

1.1 Overview

The measurement and valuation of health outcomes is a key component in the economic evaluation for decision making in healthcare. Economic evaluation aims to inform decisions of whether an intervention should be offered by the healthcare service by comparing its costs and health outcomes with those of its alternatives. Therefore, the method chosen to capture the health outcomes of the different interventions may have consequences in the results of the evaluation and ultimately on whether an intervention is offered by the healthcare service.

The overarching aim of economic evaluation is to maximise benefits by funding those interventions that generate the best outcomes for the resources available [1]. Therefore, three key principles have been proposed to guide the choice of outcome measure [2]. First, the measure should be appropriate to the decision problem; that is, to capture the change in health attributable to the intervention. Second, the measure should capture the impact on both the length of life and in the quality of life. Third, the measure should be generalisable across diseases and interventions.

This will allow for comparisons between different disease areas competing for a fixed budget. For these reasons, this paper will focus on the two most frequently used measures that combine length and quality of life, the quality-adjusted life year (QALY) and the disability-adjusted life year (DALY). These measures will be introduced with a brief discussion of clinical measures and their limitations. The implications for the Colombian context are discussed in the final section.

1.2 Clinical outcomes

Clinical outcome measures, used in clinical trials and clinical practice, such as blood pressure [3] or mortality [4] are relatively straightforward to use in economic evaluations. Clinical outcome

measures are included in the set of endpoints recorded in randomised controlled trials and are clinically meaningful for practitioners [5]. Some measures refer to intermediate outcomes, which are those that precede and may lead to final outcomes. For example, raised blood pressure and blood cholesterol may increase the risk of myocardial infarction. Final outcomes are those that represent the end result in health, such as myocardial infarction or death [6]. Although intermediate outcomes may have value in themselves, particularly clinical value, the measure of health outcome should relate to the impact of the intervention on health itself [5]. For example, the reduction of cholesterol is not an end in itself but rather the reduction in cardiovascular events obtained from the reduction in cholesterol. Using final outcomes may be difficult in preventative interventions, where studies to quantify the improvement in final outcomes are costly and time consuming to conduct. Intermediate outcomes may provide a practical alternative as long as one of the following conditions is met: (i) the intermediate outcome has value or clinical relevance by itself or (ii) the intermediate outcome has been linked to a final outcome by previous research and that any uncertainty in the link has been characterised appropriately [5, 7, 8].

Measuring health outcomes in terms of clinical endpoints, however, is limited by the inability to compare cost-effectiveness results across different interventions and disease areas [9]. Where in cardiovascular disease, incidence of acute myocardial infarction can be a relevant outcome, in other disease areas, such as cancer for example, it is not. As highlighted earlier, an outcome measure should be generalisable to allow for prioritisation and resource allocation among different disease areas.

Mortality is a clinical outcome measure that can be applied across many disease areas. Mortality is relevant in low income countries with lower life expectancy and high rates of communicable diseases. However, mortality still fails to capture the broader effects of a treatment on morbidity and quality of life and would be inappropriate as an outcome for disease areas where mortality is rare [9, 10]. Furthermore, there are other important health outcomes in addition to mortality. Delivery complications, for example, can have a considerable impact on quality of life of mothers and infants [11]. Moreover, the intervention itself may cause adverse effects on health that should be included in the evaluation [12]. For example, oral steroids, that help maintain asthma under control, can cause a large number of adverse effects, some of them (e.g. osteoporosis, diabetes)

are long term. For these reasons, a measure of health would ideally combine both length and health-related quality of life [5, 9, 12].

Combining length and health-related quality of life

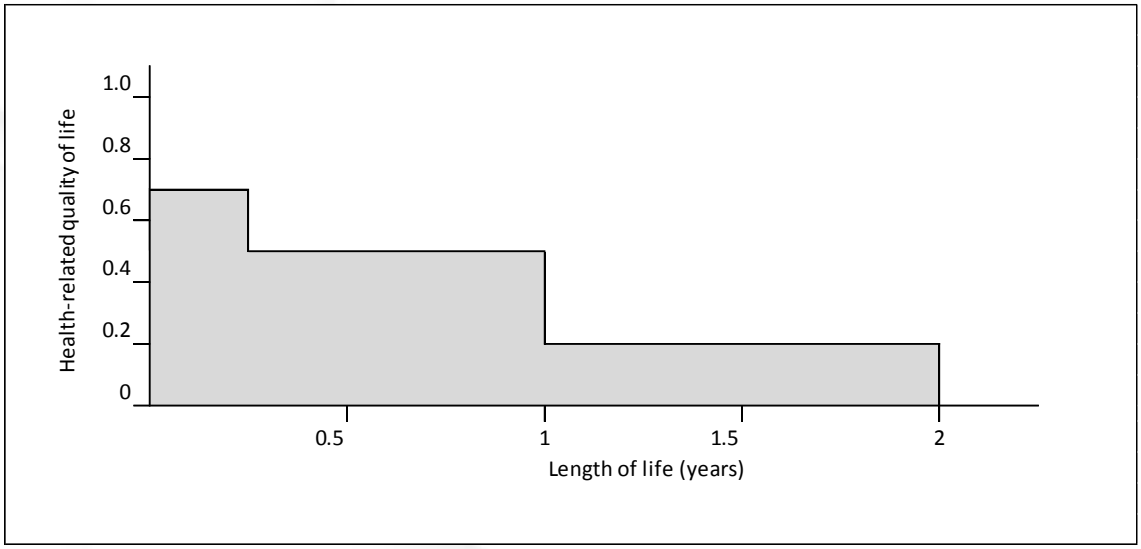
Length and health-related quality of life can be combined in a single measure of health outcome. In practice, this involves quality-adjusting the time period the individual is alive in each health state by the appropriate weight. There are a number of methods to combine length and health-related quality of life; the most common are the QALY and the DALY.

1.3 Quality-adjusted life years (QALYs)

One QALY represents one year lived in full health; one year lived in less than full health translate into QALYs lower than one. The QALY is calculated by multiplying the period of time in each health state by the score of health-related quality of life in that particular health state. The health-related quality of life of each health state is valued in a scale with an upper-bound at one, representing optimal health. Death is anchored at zero. Health states considered to be worse than dead have a negative value. The measure of health-related quality of life should have interval scale properties to allow for the aggregation of changes in QALYs. Interval scale property means that the same change represents the same impact irrespective of which part of the scale. For example, that an improvement of 0.1 from a state with a health-related quality of life of 0.2 represents the same change as in a state with a health-related quality of life of 0.8.

Figure 1 shows how to calculate a QALY. Each period of time is assigned a particular health-related quality of life score. The health-related quality of life score is used to quality-adjust the period of time. In order to obtain the score for the full length of life, each period of time is added together. For example, the first 3 months are lived in a health state with a score of 0.7; the subsequent 9 months are lived with a score of 0.5 and the last year is lived with 0.2. Therefore, the two years of life represent 0.7x0.25 + 0.5x0.75+0.2x1=0.75 QALYs.

Figure 1 Calculation of QALYs



There are a number of methods to obtain the scores of health-related quality of life: (i) professional judgment, in which a number of experts agree on the value of each health state; (ii) visual analogue scales, in which individuals are asked to rank the various health states in a line representing the health-related quality of life continuum; (iii) time trade-off, in which individuals are asked to choose between living a period of time in a given health state versus living a shorter period of time in a better health state; (iv) standard gamble, in which individuals are asked to choose between living a given health state or living in a better health state with a particular risk of death; and (v) person trade-off, in which individuals are asked on the equivalence between curing persons suffering from a disease and saving one young life [13-15].

Preference-based measurement instruments use one of the methods outlined above to obtain weights for each dimension of health included in the instrument. Patients fill in the health-related quality of life profile, which is then scored according to the appropriate tariff or value set. The most frequently used instruments are the EQ-5D-3L [16], the SF-6D (derived from the SF-36) [17] and the HUI-3 [18]. The major differences among these scales lie in their description of health states and tariff values. These scales are ready to use, cheap and require minimal analysis to translate the profiles into health-related quality of life.

The EQ-5D-3L consists of the EQ-5D descriptive system and the EQ visual analogue scale (EQ VAS). The EQ-5D descriptive system comprises of five dimensions of health: mobility, self-care, usual activities, pain or discomfort and anxiety or depression; each dimension has three levels (no problems, some problems, extreme problems) [16]. Individuals are asked to indicate their health state by ticking the box against the statement that better describes their health in each of these five dimensions. The EQ-5D-5L has the same dimensions but five levels of severity (no problems, slight problems, moderate problems, severe problems, and extreme problems) [19]. The EQ-5D-3L (or EQ-5D-5L) descriptive system is scored using the country-specific appropriate tariff. The tariff was obtained from interviews to a sample of the general population in a number of countries using time trade-off. Using the UK tariff, the most severe health state in EQ-5D-3L is valued at -0.594 [20]. There is no Colombian-specific tariff for EQ-5D-3L or EQ-5D-5L. Other potential relevant tariffs are those for individuals of Latin-America descent residing in the US [21], for Argentina [22] and Brazil [23]. The EQ-5D system has been criticised for its poor sensitivity to small changes in health-related quality of life and for changes in health states with good health-related quality of life [17].

The SF-6D is a frequently used alternative to the EQ-5D system. The SF-6D is an instrument for health-related quality of life derived from a selection of SF-36 items. It is composed of six multi-level dimensions: pain (six levels), mental health (five), physical functioning (six), social functioning (five), role limitations (four) and vitality (five) [17]. The SF-6D has been valued in a sample from the UK general population using standard gamble. The SF-6D score can be obtained from the SF-36 or the SF-12 health measures. The disadvantage of the SF-6D is that is can overestimate the value of the poorest health states and may not be sensitive to changes in conditions with poor health-related quality of life [17, 24]. The most severe health state in SF-6D is valued at 0.291 [17].

The EQ-5D and the SF-6D are generic instruments for health-related quality of life. However, condition-specific instruments have also been developed (e.g. EORTC-8D for cancer [25], AQL-5D for asthma [26], RAQoL for rheumatoid arthritis [27]). Condition-specific measures are sensitive to small changes in health outcomes related to that particular condition [28]. However, condition-specific measures may not capture all the potential impact of an intervention, particularly when the intervention may have side effects that are unrelated to the disease. In addition, it is difficult

to compare interventions in different disease areas if each evaluation uses a condition-specific measure of health-related quality of life [10, 12].

Disability-adjusted life years (DALYs)

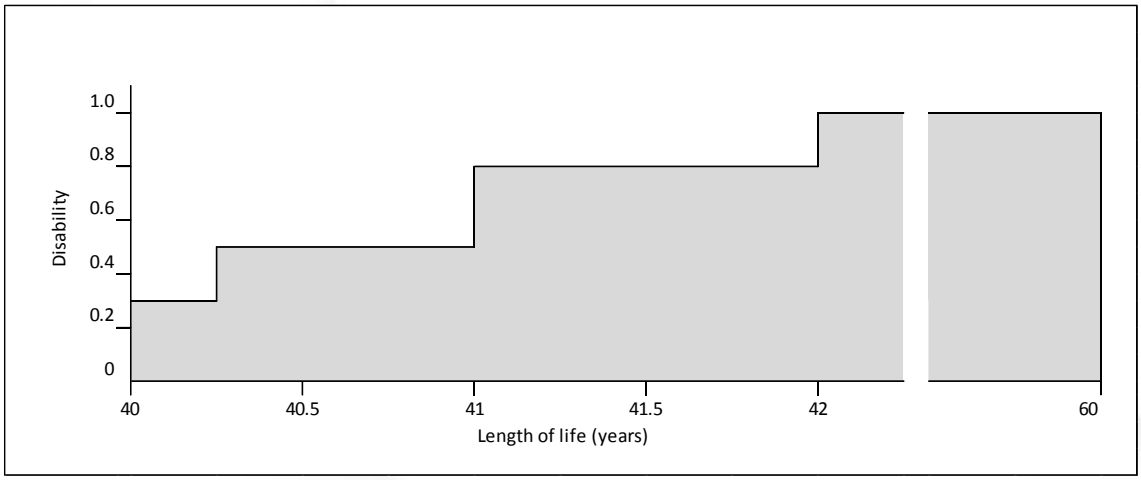
Disability-adjusted life-years (DALYs) were developed as an alternative to QALYs and are more commonly applied to research carried out in low- and middle-income countries. Similarly to the QALY, the DALY incorporates the impact of an intervention on both the length and quality of life. DALYs measure the gap between current health status and 'healthy life', where everyone lives to an advanced age free of disease and disability [29]. The standard DALY incorporates four key elements: years of life lost due to disease, disability weights that reflect the quality of life lost due to disease, age-weights which reflect the differential social value of age, and a discount rate of 3% per annum which reflects society's preference for valuing present health more than future health [30, 31]. DALYs with no discounting and no age weights ('no frills' DALY) and DALYs with no age weights but discounted at 3% are also available.

The DALY was originally developed for the evaluation of the global burden of disease study. In this analysis, the years of life lost due to premature mortality are calculated by comparing actual age of death with standard life expectancy data based on the highest national life expectancy observed, that is 82.5 for females and 80 for males [29]. For economic evaluation, the life expectancy of the country to which the analysis refers to should be used instead [30]. The disability weights are between zero and one, where zero is full health. The weights were originally allocated by a group of experts. Subsequently the weighting system was revised with a person trade-off system used to derive weights for 22 indicator conditions [32]. The age-weights were obtained from a group of public health experts and reflect the higher value placed on years of life for a young middle age person than for a child or an old person.

Figure 2 shows how to calculate the DALY using a similar example to that used for the QALY. Consider a disease with an age of onset at 40 years of age and two years of life expectancy. The local life expectancy at 40 years old is 20 years. The two years with the disease are lived with a

progressive disability, 0.3 in the first 3 months, followed by 0.5 in the subsequent 9 months and 0.8 in the last year. The 'no frills' DALY is 16.25, that is 18 years of life lost (60 years of life expectancy – 42 years age of death) plus the 1.25 years lived with disability (0.25 years x 0.3 + 0.75 years x 0.5 + 1 year x 0.8). Discounting and age-weighting can be using the approaches proposed by Fox-Rushby & Hanson [33] or Larson (2013) [34].

Figure 2 Calculation of DALYs



The key differences between DALYs and QALYs are in the measurement of life years and weights used to quality-adjust life years. DALYs measure years of life lost compared to an ideal life expectancy while QALYs measure years of life gained and, therefore, do not require knowledge of the life expectancy of the general population. DALYs use disability weights obtained from expert deliberation for specific diseases, while QALYs use weights obtained from a sample of the general population based on the desirability of particular health states. The number of DALYs saved are equivalent to the number of QALYs gained under the following conditions: (i) the health-related quality of life weight is equivalent to one minus the corresponding disability weight; (ii) both the health-related quality of life weight and the disability weight are constant throughout the disease duration; (iii) the same discount rate is used for both calculations; (iv) DALYs are not weighted according to age [35]. However, these conditions are unlikely to be met given the different methods for obtaining weights for health-related quality of life and disability and the change in health experienced by individuals over their lifetime.

DALYs have been criticized for three major reasons. First, the disability weights do not represent the preferences of the general population but the views of an expert panel [31]. Second, using age-weights implies that the value of life depends on the age of the individual. Children and older individuals are assigned weights lower than one. This means that for the same illness, young or older people will receive fewer resources, when compared to someone in the middle age-groups [36]. Third, the disability-weights assume that years lived by individuals with disabilities have lower value [37].

1.4 Implications for decision making in Colombia

DALYs and QALYs are the two main measures of health outcomes for economic evaluation. Commonly, QALYs are preferred in high income countries whilst DALYs are more routinely used in low and middle income countries. The DALY is the recommended measure for the Gates' reference case [2]. Economic evaluations using each of these measures are unlikely to return the same results given the different assumptions and methodologies. Therefore, one measure should be established as the preferred 'base-case' measure and the other used for scenario analysis and for when the preferred measure cannot be estimated. This does not preclude the development and adaptation of measures over time in order to ensure measures and values of relevance to the Colombian context are applied. However, a clear approach should be agreed from the start, upon which future methodological research can be built.

1.5 References

1. Byford, S., D. McDaid, and T. Sefton, *Because it's worth it: A practical guide to conducting economic evaluations in the social welfare field*. 2003, Joseph Rowntree Foundation.
2. Bill and Melinda Gates Foundation and NICE International, *Methods for Economic Evaluation Project*. Forthcoming.
3. Gray, A., et al., *Cost effectiveness of an intensive blood glucose control policy in patients with type 2 diabetes: economic analysis alongside randomised controlled trial (UKPDS 41).* BMJ: British Medical Journal, 2000. **320**(7246): p. 1373.
4. Tengs, T.O., et al., *Five-hundred life-saving interventions and their cost-effectiveness.* Risk analysis, 1995. **15**(3): p. 369-390.
5. Drummond, M., et al., *Methods for the Economic Evaluation of Health Care Programmes*. 2005, Oxford: Oxford Medical Publications.
6. The Physician Consortium for Performance Improvement, *Measures Development, Methodology, and Oversight Advisory Committee: Recommendations to PCPI Work Groups on Outcome Measures* 2011, American Medical Association.
7. Ramsey, S., et al., *Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report.* Value in health, 2005. **8**(5): p. 521-533.
8. Taylor, R.S. and J. Elston, *The use of surrogate outcomes in model-based cost-effectiveness analyses: a survey of UK Health Technology Assessment reports*. 2009: Prepress Projects Limited.
9. Torgerson, D. and J. Raftery, *Economics notes: measuring outcomes in economic evaluations.* BMJ: British Medical Journal, 1999. **318**(7195): p. 1413.
10. Dolan, P., *The measurement of health-related quality of life for use in resource allocation decisions in health care.* Handbook of health economics, 2000. **1**: p. 1723-1760.
11. Ashford, L., *Hidden suffering: disabilities from pregnancy and childbirth in less developed countries*. 2002: Population Reference Bureau, MEASURE Communication.
12. Gold, M.R., *Cost-effectiveness in health and medicine*. 1996: Oxford University Press.
13. Torrance, G.W., *Utility approach to measuring health-related quality of life.* Journal of chronic diseases, 1987. **40**(6): p. 593-600.
14. Torrance, G.W., *Measurement of health state utilities for economic appraisal: a review.* Journal of health economics, 1986. **5**(1): p. 1-30.
15. Nord, E., *The person-trade-off approach to valuing health care programs.* Medical decision making, 1995. **15**(3): p. 201-208.
16. Szende, A., M. Oppe, and N.J. Devlin, *EQ-5D value sets: inventory, comparative review and user guide*. Vol. 2. 2007: Springer.
17. Brazier, J., J. Roberts, and M. Deverill, *The estimation of a preference-based measure of health from the SF-36.* Journal of health economics, 2002. **21**(2): p. 271-292.
18. Furlong, W.J., et al., *The Health Utilities Index (HUI®) system for assessing health-related quality of life in clinical studies.* Annals of medicine, 2001. **33**(5): p. 375-384.
19. van Hout, B., et al., *Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets.* Value in Health, 2012. **15**(5): p. 708-715.

20. Dolan, P., et al., *A social tariff for EuroQol: results from a UK general population survey*. 1995: Centre for Health Economics.
21. Zarate, V., P. Kind, and L.-H. Chuang, *Hispanic valuation of the EQ-5D health states: a social value set for Latin Americans.* Value in Health, 2008. **11**(7): p. 1170-1177.
22. Augustovski, F.A., et al., *Argentine valuation of the EQ-5D health states.* Value in Health, 2009. **12**(4): p. 587-596.
23. Viegas Andrade, M., et al., *Societal Preferences for EQ-5D Health States from a Brazilian Population Survey.* Value in Health Regional Issues, 2013.
24. Longworth, L. and S. Bryan, *An empirical comparison of EQ-5D and SF-6D in liver transplant patients.* Health economics, 2003. **12**(12): p. 1061-1067.
25. Rowen, D., et al., *Comparison of Generic, Condition-Specific, and Mapped Health State Utility Values for Multiple Myeloma Cancer.* Value in Health, 2012. **15**(8): p. 1059-1068.
26. Yang, Y., et al., *Estimating a preference-based index for a 5-dimensional health state classification for asthma derived from the asthma quality of life questionnaire.* Medical Decision Making, 2011. **31**(2): p. 281-291.
27. Marra, C.A., et al., *A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis.* Social science & medicine, 2005. **60**(7): p. 1571-1582.
28. Donaldson, C., et al., *Should QALYs be programme-specific?* Journal of Health Economics, 1988. **7**(3): p. 239-257.
29. Murray, C.J., *Quantifying the burden of disease: the technical basis for disability-adjusted life years.* Bulletin of the World health Organization, 1994. **72**(3): p. 429.
30. Murray, C.J. and A.K. Acharya, *Understanding DALYs.* Journal of health Economics, 1997. **16**(6): p. 703-730.
31. Williams, A., *Calculating the global burden of disease: time for a strategic reappraisal?* Health economics, 1999. **8**(1): p. 1-8.
32. Clarke, P.M., J.L. Wolstenholme, and S. Wordsworth, *Applied methods of cost-effectiveness analysis in healthcare*. Vol. 3. 2010: Oxford University Press.
33. Rushby, J.F. and K. Hanson, *Calculating and presenting disability adjusted life years (DALYs) in cost-effectiveness analysis.* Health policy and planning, 2001. **16**(3): p. 326-331.
34. Larson, B.A., *Calculating disability-adjusted-life-years lost (DALYs) in discrete-time.* Cost Effectiveness and Resource Allocation, 2013. **11**(1): p. 1-6.
35. Sassi, F., *Calculating QALYs, comparing QALY and DALY calculations.* Health Policy and Planning, 2006. **21**(5): p. 402-408.
36. Anand, S. and K. Hanson, *Disability-adjusted life years: a critical review.* Journal of health economics, 1997. **16**(6): p. 685-702.
37. Arnesen, T. and E. Nord, *The value of DALY life: problems with ethics and validity of disability adjusted life years*. BMJ: British Medical Journal, 1999. **319**(7222): p. 1423.

2. Discounting in economic valuation of healthcare interventions

Ana Duarte

2.1 Introduction

Economic evaluation of healthcare interventions requires the discounting of costs and health benefits according to their occurrence in time because society (and individuals) values resources spent or saved in the present and those in the future differently.

The aim of this document is to provide an overview of the most important issues in the appropriate discounting in the context of a particular healthcare system. First, the rationale justifying the need for discounting is presented. It continues with a discussion on how to select of an appropriate discount rate for costs and health benefits. A brief overview of international recommendations and guidelines is used as a backdrop for the Colombian particular case. Finally, implications for decision making in Colombia are discussed.

2.2 The need for discounting

Discounting converts future costs and benefits into their present value; therefore, it allows the comparison of costs and benefits occurring at different points in time. The need for discounting arises from individuals and society having a positive rate of time preference. In other words, a lower weight is placed on costs and benefits occurring in the future when compared with those in the present.

A combination of factors has been identified as potentially contributing to the existence of a positive rate of time preference. The first is that individuals have a short-term view of life and place a higher value on present events versus events in the future. Secondly, the uncertainty surrounding the future is greater. Therefore, compensation needs to be offered to offset the higher uncertainty of an event that occurs later for an individual to be indifferent between two events occurring at different time points. Furthermore, historically there has been an overall tendency for positive economic growth in societies, so that individuals and societies are expected

to be wealthier in the future. As a result, one monetary unit is worth more compared with current wealth than the same unit compared with future (greater) wealth. This means that, in the absence of risk associated to a particular investment, an individual with a positive rate of time preference would always obtain a positive return on that investment, i.e. withholding current consumption and investing it at a given rate will translate in the possibility of greater consumption in the future [1]. Underlying the rationale for discounting is the assumption that wealth and health are tradable over time. This is an assumption that is usually considered to be valid in health economics, and has been at the basis of models such as the Grossman's human capital model [2].

Discounting in economic evaluation of health care can be implemented by applying the following formula:

$P = \sum_{n=1}^{t} F_n * (1 + r)^{-n}$ [1]

The present value (*P*) of a cost or benefit is estimated by applying a weight to its future value (*F*) at year *n*, where *r* is the annual discount rate and summing all weighted (discounted) values that are to be included in the analysis. The implicit assumption when applying this formula is that all costs or benefits occur at the end of each year. Alternatively, it can be assumed that costs and benefits occur at the beginning of the year, in which case discounting is applied as follows:

$P = \sum_{n=0}^{t} F_n * (1 + r)^{-n}$ [1]

2.3 Discounting of costs and health benefits – is there an argument for differential discounting?

Table 1 summarizes the appropriate discount rates to apply for costs and health benefits according to the social objective of the economic evaluation and the decision rules that are followed, as demonstrated by Claxton and colleagues [3]. The choice of discount rate depends on the objective of the healthcare system and on the type of decision rule, based on the incremental cost-effectiveness threshold (ICER) or on net benefit (more on decision rules in the Technical Document entitled 'Decision rules in economic evaluation of healthcare interventions'). Differential discounting is not appropriate if using net benefit decision rules. Here, the threshold

used to convert benefits into costs and vice versa should already account for any differential value of health versus costs over time. In contrast, differential discounting may be appropriate if using ICER decision rules. For a healthcare service with the objective of maximising health subject to a budget constraint, the discount rate is different for costs and health benefits if the threshold is expected to change over time. For an healthcare service that aims to maximise the consumption value of health, the discount rate on costs and benefits depends on changes on the threshold (if subject to a budget constraint) and on the consumption value of health. Independently of whether the objective of the optimisation problem is to maximise health or welfare, a common rate should be applied to costs and health benefits if there is a fixed health care budget and the cost-effectiveness threshold is constant over the relevant time horizon.

The discount rate to apply, $\mathbf{r_h}$ or $\mathbf{r_c}$, depends on the social objective of economic evaluation in a particular health care system. When the objective is to maximise health, $\mathbf{r_h}$ is the relevant rate. In contrast, $\mathbf{r_c}$ becomes the relevant rate when maximising the consumption value of health. In both cases, adjustments may be required depending on the decision rule that is followed, social objective, and changes in cost-effectiveness threshold and/or consumption value of health.

Table 1 Discount rates on costs and health benefits dependent on social objective and decision rules.

Social objective	Decision based on			
	Incremental cost-effectiveness ratio		Net benefit	
	DR on costs	DR on health	DR on costs	DR on health
Maximise health				
▪ Subject to budget constraint	$r_h + g_k$	r_h	r_h	r_h
Maximise consumption value of health				
▪ Subject to budget constraint	$r_c + g_k - g_v$	$r_c - g_v$	r_c	r_c
▪ Without a budget constraint	r_c	$r_c - g_v$	r_c	r_c

DR – discount rate; r_h - willingness to trade between current and future health; r_h - willingness to trade between current and future consumption; g_k - growth rate of the threshold between time periods; g_v - growth rate of the consumption value of health.

2.4 Which discount rate?

Four alternative approaches have been used to empirically estimate a discount rate: i) marginal social rate of time preference (SRTP), ii) marginal social opportunity cost of capital (SOC), iii) weighted average approach, and iv) the shadow price of capital approach. The selection of the analytical approach depends on what is perceived to be the opportunity cost of public funds [4]. A brief description of each approach is provided below. This shadow price of capital approach is not discussed because it requires the estimation of a project specific discount rate, which is unfeasible for policy making [4].

The underlying assumption of SRTP is that public projects displace current consumption. Therefore, the costs and benefits of these projects can be considered as consumption goods streams. One of the methods used to empirically estimate STRP is to apply Ramsey's formula [5]. The formula equates SRTP to the sum of a utility discount rate with the product between the elasticity of the marginal utility of consumption and the annual rate of per capita consumption growth. The utility discount rate has two components, one that reflects the individuals' pure preference for current consumption due to impatience or myopia and the other reflecting the risk of death. There are alternative approaches to estimate SRTP that assume a time declining social discount rate. However, these approaches may lead to time-inconsistent planning, and are mostly relevant when there are intergenerational equity concerns (for very long term projects, such as those in environmental projects) [4]. The SRTP approach does not account for the possibility that public investment may compete with private investment for funding. Therefore, does not adjust for the marginal social rate of return on investment in the private sector, i.e. the interest rate of investing the capital instead of applying it in the project or policy.

Ramsey's formula has been adapted by the UK treasury ([6] to define a SRTP of 3.5% to be used in the evaluation of investment projects in the public sector. This rate is composed by:

1. Pure time preference rate of 0.5%;
2. Catastrophic risk premium[1] of 1% - based on a 1% mortality rate;
3. Combined effect of elasticity of marginal utility of consumption[2] (1%) and growth in per capita consumption (2%).

The SOC approach has been suggested as an alternative to estimate social discount rate, based on the argument that public investment displaces private investment [4]. SOC can be estimated as the marginal pre-tax return of a riskless investment in the private sector. The underlying assumption in using this approach to estimate the social discount rate is that public investment will not displace any private consumption. In a perfectly competitive market SRTP and SOC would be equal. However, the existence of market distortions leads to SRTP being generally lower than SOC [4].

The third approach consist of an weighted average of SRTP, SOC and foreign borrowing where weight attributed to each component will depend on the proportion of funds obtained from each source (i.e. public, private and foreign funding)[4]. The Harberger approach can be used to estimate the social discount rate as a weighted average [7]. Although this approach attempts to overcome the limitations of the SRTP and SOC, it assumes that all benefits are immediately consumed and cannot be reinvested. This may leads to overdiscounting of benefits that can favour shorter term projects against longer term ones [4].

There has been no consensus regarding which is the best approach to estimate a social discount rate. In general, policy makers in high-income countries mostly prefer to use the SRTP approach, whereas the SOC approach has been preferred in low- and middle-income countries. The social discount rates applied in the economic evaluation of public projects are usually lower in high income countries (3-7%) than in low- and middle-income countries (8-15%), which reflects not

[1] Probability that all returns on investment of public policies, programs or projects are precluded (or at least by the occurrence of a catastrophic event).
[2] Defined as the percentage change in individuals' marginal utility by percentage change in consumption.

only the different estimating approaches but most importantly the different assumptions on how public investment affects the domestic economy [4].

2.5 Discount rates in an international context

The approaches described above are used to estimate the social discount rate for the economic evaluation of public projects in general and not specifically in health. Nevertheless, these are commonly used by policy makers to recommend a discount rate to apply in the health sector. It should be noted that the social discount rate represents $\mathbf{r_c}$, the rate at which society is willing to trade between future and current consumption. As discussed previously, for collectively funded health system where economic analysis aims to maximise health under a fixed budget, the relevant discount rate is that which reflects society's willingness to trade between current and future health (r_h). Most countries recommendations assume (if not always explicitly) that r_h and r_c are equivalent. Claxton and colleagues have criticised this assumption, which is explicit in the UK case, on the basis that it health and suggest that r_h is lower than r_c [3].

Table 2 summarises the discount rates for costs and health benefits recommended by economic evaluation in health guidelines worldwide [8]. Recommendations for the reference case are displayed where applicable. The same discount rate for costs and health benefits is recommended in the majority of the jurisdictions. Only Belgium and the Netherlands explicitly recommend differential discount rates, with costs discounted at a higher rate than health benefits. Discount rates are mostly between 3 and 5%. Nevertheless, the international context represented on the table is dominated by high-income countries which have been leading on economic evaluation in health. Importantly, it should be recognised that the economic context of each country is crucial to the selection of an appropriate rate. Discount rates could only be found for three countries in Latin America (Brazil, Cuba and Mexico). Brazil and Mexico both recommend a same discount rate of 5% on costs and health benefits. Cuban guidelines recommend discount rates for costs and health benefits as prescribed by the Finance and the Economy Ministries for other sectors in the economy, as well as 3 and 5% [9].

Table 2 Summary of discount rates for costs and health benefits in international pharmacoeconomic guidelines (adapted from the International Society for Pharmacoeconomics and Outcomes Research website [8])

	Discount rate on	
	Costs	Health benefits
Brazil	5%	5%
Cuba	Rates recommended by specific government agencies, and 3 and 5%.	As for costs.
Mexico	5%	5%
Northern America	When appropriate, adjustment for the time preference should be incorporated and should follow US PHS Panel recommendations	When appropriate, adjustment for the time preference should be incorporated and should follow US PHS Panel recommendations
Canada	5%	5%
Austria	5%	5%
Denmark	Should be done when there are effect and cost consequences over several years. Decisions on discount rate should be made in each analysis.	As for costs.
Baltic republics	5%	5%
Belgium	3%	1.5%
Finland	3 and 0%	3 and 0%
France	0, 5, and 3%*	0, 5, and 3%*
Germany	3%	3%
Hungary	5%	5%
Ireland	4%	4%
Italy	3%	3%
Norway	4%	4%
Poland	5%	3.5%
Portugal	5%	5%

	Discount rate on Costs	Health benefits
Russia Federation	5%	5%
Slovak Republic	5%	5%
Spain	3%	3%
Sweden	3%	3%
The Netherlands	4%	1.5%
England and Wales	3.5%	3.5%
Switzerland	2.5%, 5%, 10%	2.5%, 5%, 10%
China Mainland	One year interest rate.	One year interest rate.
Taiwan	3%	3 scenarios: discount costs but not health benefits; both discounted, both not discounted.
Thailand	3%	3%
Malaysia	3%	3%
Egypt	3.5%	3.5%
South Africa	5%	5%
Oceania		
Australia	5%	0% or 5%
New Zealand	3.5%	3.5%
* for international comparison purpose.		

2.6 Implications for decision-making in healthcare in Colombia

Appropriate discounting of costs and benefits is crucial to the efficient allocation of resources in economic evaluation. This document highlights the importance of discounting in economic evaluation in health, as well as the considerations to be taken into account when selecting discount rate for cost and health benefits that appropriately reflect society's willingness to trade health or the consumption value of health between time periods.

Similarly to other middle-income countries, Colombia has been working towards an expanded role for a social health insurance based health care system, with contributions linked to earnings, with some exceptions for those in unemployment or informal employment [10], and access according to need (on some basic coverage). Despite the differences in financing mechanisms and extent of coverage of the Colombian health care system when compared to the UK NHS, the social objective of economic evaluation is likely to be to maximise health subject to a fixed budget. Therefore, the selection of an appropriate discount rate will depend on i) the rate at which society is willing to trade between current and future health (r_h), and ii) changes in the cost-effectiveness threshold over time. The way in which the latter is accounted for in the selection of a discount rate for costs and health benefits will hinge on the decision rule applied (ICER or net benefit based).

Currently, the Departamento Nacional de Planeación recommends a 12% discount rate for the purpose of the socioeconomic evaluation of public investment projects [11]. A 2007 study by Rodriguez Hernandez empirically estimated a 8.5% social discount rate for Colombia using the Harberger approach and capital market data for the period between 1995 and 2005 [12]. The estimation of both 8.5 and 12% social discount rates are unlikely to correspond to r_h. The degree to which either may be a suitable proxy for r_h should also be considered by the Colombian health technology agency before making recommendations on values of appropriate discount rates for costs and benefits to be applied in economic evaluation in health.

2.7 References

1. Drummond M, et al., *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed. 2005, Oxford: Oxford University Press.
2. Grossman, M., *The demand for health: a theoretical and empirical investigation.* NBER Books, 1972.
3. Claxton, K., et al., *Discounting and decision making in the economic evaluation of health-care technologies*. Health economics, 2011. **20**(1): p. 2-15.
4. Zhuang, J., et al., *Theory and practice in the choice of social discount rate for cost-benefit analysis: a survey.* 2007.
5. Ramsey, F.P., *A mathematical theory of saving.* The Economic Journal, 1928. **38**(152): p. 543-559.
6. Treasury, H.M.s., *The green book: appraisal and evaluation in central government.* London: HM Treasury, 2003.
7. Harberger, A.C., *Project evaluation: collected papers*. 1972: University of Chicago Press.
8. International Society for Pharmacoeconomics and Outcomes Research. *COMPARATIVE TABLE (Compare selected Countries on selected Key Features)*. [cited January 2014]; Available from: http://www.ispor.org/PEguidelines/COMP1.asp.
9. Gálvez González, A.M., *Guía metodológica para la evaluación económica en salud*, M.d.S. Pública, Editor. 2003, Escuela Nacional de Salud Pública: La Habana, Cuba.
10. Wagstaff, A., *Social health insurance reexamined.* Health Economics, 2010. **19**(5): p. 503-517.
11. Departamento Nacional de Planeación. *Inversiones y Finanzas* [cited January 2014]; Available from: https://www.dnp.gov.co/PreguntasFrecuentes/InversionesyFinanzasPúblicas.aspx.
12. Hernández, R.R., *Reestimación de la tasa social de descuento en Colombia a partir del desarrollo de su mercado de capitales durante el período 1995–2005*. Equidad & Desarrollo, 2007(8): p. 55-81.

3. Uncertainty and the value of additional evidence in health care decisions

Claire McKenna

3.1 Introduction

When making recommendations about the use of health care interventions, the potential benefits and harms associated with the intervention are weighed up on the basis of what the existing evidence suggests about the intervention. However, on balance, the existing evidence might suggest that a particular intervention is expected to be the most effective (i.e. improve health outcomes with least harm to patients) but, due to uncertainty in the evidence, there is a chance that another intervention could improve health outcomes to a greater extent [1]. Therefore, it is important to consider the scale of the uncertainty in the current evidence and the consequences of that uncertainty in terms of health outcomes to patients from making a wrong decision about the intervention. The scale of the uncertainty is often indicated by the results of systematic review and meta-analysis [2]. This can be combined with information about baseline risk and incidence of the disease to estimate the expected consequences of uncertainty, expressed in terms of health outcomes, e.g. deaths averted per year. The expected consequences provides an upper bound on the health benefits that could be gained from further research which would confirm whether or not the intervention was actually more effective than the others currently available. Health outcomes can also be improved by ensuring that the accumulating findings of research are implemented into clinical practice. In fact, the potential improvements in health outcome by encouraging the implementation of what existing evidence suggests is the most effective intervention may well exceed the potential improvements in health outcomes through conducting further research [1].

When making recommendations about whether interventions represent good value for money and whether they should be reimbursed by the health care system, cost-effectiveness analysis is used to compare the costs and health outcomes of the alternative interventions under

consideration [3 4]. There are numerous sources of uncertainty in cost-effectiveness analysis, which are usually propagated through a decision-analytic model.

When decisions are made on the basis of uncertain information there is a possibility that subsequent evidence will indicate that the interventions reimbursed did not in fact represent good value for money. The difference between the realised population health gains and those that could have been achieved by reimbursing the 'optimal' set of health interventions represents the opportunity cost of uncertainty [5 6].

This document illustrates the importance of considering the scale and consequences of uncertainty in assessments of both clinical effectiveness and cost-effectiveness evidence. It begins by showing how uncertainty and its consequences are estimated in health care decisions with reference to an example on the use of corticosteroids following traumatic head injury [7 8]. The methods and presentation of results commonly used to characterise parameter uncertainty in cost-effectiveness analysis are then presented. The document finishes with a discussion on the implications for policy decisions.

3.2 Consequences of uncertainty and the value of additional evidence

A useful starting point is the distinction between uncertainty and variability. Variability refers to the natural variation between patients in their response to treatment, which cannot be reduced by additional evidence. In contrast, uncertainty refers to the fact that we cannot know with absolute certainty what the expected (mean) effects of the treatment are for the population of patients who receive the treatment. Basing decisions about a health care intervention on expected health effects will ignore the question of whether the current evidence is a sufficient basis for guiding decisions in clinical practice. It fails to address the question of whether further research is needed before making a decision that could potentially harm patients due to the consequences of the uncertainty [9].

The value of evidence or the health consequences of uncertainty are illustrated using a simple example in Table 1. The table shows the net health effects [4], which can be expressed in terms of a measure of health outcome (e.g. quality-adjusted life years (QALYs) [10], deaths averted per annum) for two treatment options A and B. Each row of the table represents a 'true' realisation of uncertainty that could occur in clinical practice. However, we don't know which of these realisations will actually occur in clinical practice; therefore, the decision maker must choose the treatment that is expected (on average) to give the greatest net health effects. The expected net health effects for treatment A and B is the average over all the possibilities, which in this example are 12 and 13 for treatments A and B, respectively (the average of the five realisations). On the basis of current evidence, the decision-maker would conclude that treatment B is the most effective and would expect to gain, on average, one additional unit of health outcome compared with treatment A. However, this decision is uncertain and treatment B is not always the best choice (only three times out of the five possibilities). Therefore, the probability that B is the most effective treatment option is 0.6 and the probability of making a wrong decision by approving treatment B based on current evidence is 0.4 (=1-0.6). Whether or not this level of uncertainty 'matters' depends on the consequences of the uncertainty, i.e. what improvement in net health effects could potentially be achieved by reducing this uncertainty. If uncertainty was removed we would always be able to make the best treatment choice; therefore, from Table 1, we could achieve 13.6 units of health outcome rather than 13 (last column of Table 1) if we could resolve the uncertainty. The difference between a decision based on perfect information and a decision based on current evidence, i.e. 0.6 units of health outcome in this example, represents the maximum value of additional evidence or the consequences of the uncertainty.

Table 3 Value of additional evidence

Realisations that could occur in clinical practice	Net health effects (e.g. QALYs, deaths averted, other measure of health outcome)			Best outcome for each realisation
	Treatment A	Treatment B	Best treatment choice	
Realisation 1	9	12	B	12

Realisations that could occur in clinical practice	Net health effects (e.g. QALYs, deaths averted, other measure of health outcome)			Best outcome for each realisation
	Treatment A	Treatment B	Best treatment choice	
Realisation 2	12	10	A	12
Realisation 3	14	17	B	17
Realisation 4	11	10	A	11
Realisation 5	14	16	B	16
Expected (mean)	12	13		13.6

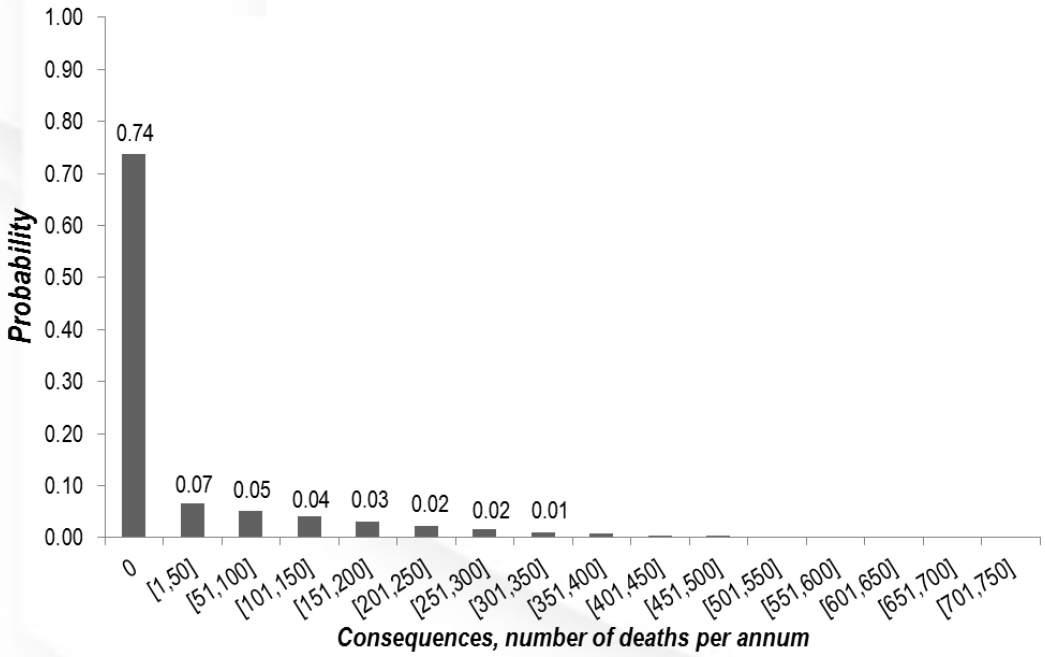
An example

Prior to the CRASH trial (Corticosteroid Randomisation After Significant Head injury) [8 11] the effect of corticosteroids (CS) on death and disability following traumatic head injury (THI) was unclear, despite 19 clinical trials conducted between 1972 and 1995 [7 12]. The CRASH trial, which first reported in 2004, was stopped early after enrolling 10,008 adults with THI. It reported a higher risk of death or severe disability associated with the use of corticosteroids (CS). As a consequence of this definitive trial clinical practice changed dramatically, resulting in many thousands of iatrogenic deaths averted around the world.

The global value of the CRASH trial appears, with hindsight, self-evident. However, the trials comparing the use of corticosteroids to placebo or no treatment in acute THI available prior to CRASH were of varying study quality, length of follow-up, steroids administered, varying doses and time to administration. A standard meta-analysis of these trials suggested substantial uncertainty about the effectiveness of CS in THI. A random effect analysis on the mortality endpoint favoured the use of CS with a mean odds ratio (OR) for death of 0.93, but the confidence interval crossed the no difference line (95% confidence interval (CI), 0.71 - 1.18) indicating that the use of CS could be harmful to patients. If a decision to use CS based on existing evidence turned out to be wrong (which was shown to be the case in 2004 following the CRASH trial) there would be consequences in terms of higher mortality in the patient population. Translating the chance that a decision about the use of CS in THI will be 'wrong' into the consequences for patient

outcomes requires an estimate of the baseline risk of mortality (without use of CS) in THI and an estimate of the size of the population who could potentially benefit from the information. In this example, using the control arms of the trials as an estimate of baseline risk (informed directly from the output of the random effect meta-analysis) suggested a mean baseline probability of death of 0.378 (95% CI, 0.248 - 0.469). The combined effect of uncertainty in relative treatment effect and baseline risk can be characterised by taking repeated random samples from their distributions, which are already estimated in the meta-analysis. Each random sample, or simulated value, can be interpreted as one possible realisation of uncertainty, i.e. one possible 'true' value of what might turn out in clinical practice given the information that is available. Assuming that the annual incidence of THI was approximately 9,000, this analysis provides a distribution of the health consequences of uncertainty in terms of number of deaths per annum, which are illustrated in Figure 1. Based on the random effect meta-analysis of evidence that favours the use of CS prior to the CRASH trial, there is a 74% chance of no adverse mortality consequences. However, there is a 26% chance that using CS in THI will not be effective and will cause iatrogenic deaths, with a greater chance of a smaller number of deaths (19% chance of greater than zero and 200 deaths per year) and a smaller chance of a large number of deaths (7% chance of greater than 200 deaths per year). The expected number of iatrogenic deaths due to this uncertainty is 40 per year (the average over the distribution illustrated in Figure 1), which is also an estimate of the expected health benefits that could potentially be gained each year if the uncertainty surrounding the use of CS could have been resolved at that time. It represents an estimate of the upper bound on the expected health benefits of additional evidence that would confirm whether or not CS reduces mortality.

Figure 3 Distribution of the consequences of uncertainty



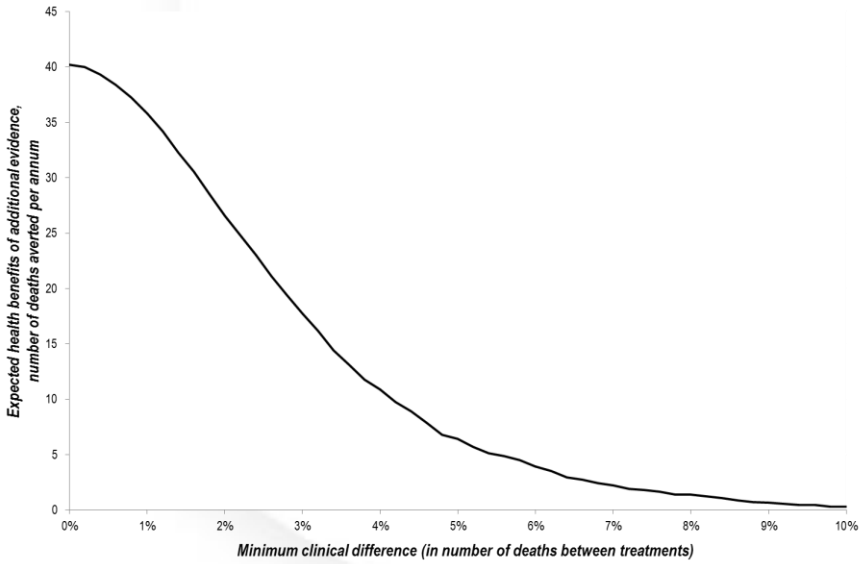
3.3 Value of implementing the findings of research

Clearly, the potential health benefits of conducting further research will only be realised (patient outcomes actually improve) if the findings of the research have an impact on clinical practice. Indeed, the potential improvements in health outcome by encouraging the implementation of what existing evidence suggests is the most effective intervention may well exceed the potential improvements in health outcomes through conducting further research [1 13]. The distinction between these two very different ways to improve patient outcomes is important because, although the results of additional research may influence clinical practice and may contribute to the implementation of research findings, it is certainly not the only, or necessarily the most effective, way to do so. Insofar as there are other mechanisms (e.g., more effective dissemination of existing evidence) or policies which fall within the remit of other bodies (e.g., incentives for implementation), then continuing to conduct research to influence implementation rather than because there is real value in acquiring additional evidence itself would seem inappropriate [1]. The research capacity could be used elsewhere to acquire additional evidence in areas where it

would offer greater potential health benefits and there are negative health effects for those patients enrolled in the less effective arm of randomised control trials.

If the impact of research on clinical practice is likely to require highly statistically significant results this will influence the design, cost and time taken for research to report. It maybe that a larger clinical difference in effectiveness would need to be demonstrated before research would have impact on practice. The concept of an effect size has been central to the design of clinical research and determines the sample size in most clinical trials. The effect size does not represent what is expected to be found by the research, but the difference in outcomes that would need to be detected for the results to be clinically significant and have an impact on clinical practice. The same concept can be used in this analysis, where estimates of the expected health benefits of additional evidence can be calculated for different minimum clinical differences (MCD) in outcomes [1]. Requiring that further research must demonstrate larger differences in effect will tend to reduce its expected potential benefits because large differences are less likely to be found than smaller ones. The expected benefits of additional evidence for a range of MCD for the example of CS are illustrated in Figure 2. When the MCD is zero the expected health benefits of additional evidence represents the upper bound on the value of evidence that would confirm whether or not CS reduces mortality. Requiring that further research must demonstrate larger difference in effect reduces these expected health benefits as larger differences are less likely to found than smaller ones; therefore, the potential benefits of further research decline as a greater MCD is required.

Figure 2 Value of additional evidence for minimum clinical difference required



3.4 Methods to characterize parameter uncertainty in cost-effectiveness analysis

As well as capturing uncertainty in the clinical effectiveness evidence, any assessment of cost-effectiveness which compares the relative value of interventions in terms of both costs and health outcomes is subject to uncertainty. This uncertainty arises from a number of sources: i) uncertainty in the estimates of parameters of the decision models commonly used to estimate costs and effects, and ii) uncertainty in the scientific value judgments and assumptions that are made when constructing a model [14]. Examples of uncertainty in decision models include uncertainty in the estimates of treatment effect, assumptions regarding extrapolation of effects over time, assumptions about the natural history of the disease, costs of treatment, health-related quality of life utility values, and assumptions about the functional form of statistical distributions used to estimate parameters.

In a decision-analytic model, statistical distributions are assigned to each of the model input parameters. Probabilistic sensitivity analysis (PSA) is then used to characterise the uncertainty in the decision associated with the cost-effectiveness assessment of the intervention [15]. This involves sampling at random (Monte Carlo simulation) from the distributions of parameter inputs

of the model. Each set of samples from all of the parameters generate a single estimate of expected costs and effects for both the intervention and the comparators. These outputs are recorded and then a new set of possible values for the parameters are sampled. This process of sampling inputs and recording the corresponding outputs (expected costs and effects) is repeated (e.g. 10,000 times) so that all the values that the parameters are likely to take are represented in the range of outputs. In doing so, PSA reflects the combined uncertainty in all the parameter inputs to quantify the implications of that uncertainty in the cost-effectiveness estimates (outputs). The total expected cost and effect for the intervention and the comparators is the average across all the corresponding PSA outputs. The output of this process also provides the proportion of times (probability) that each of the alternative interventions is expected to be cost-effective and this enables the decision maker to assess the level of uncertainty, the consequences of that uncertainty, and the maximum value of additional evidence required to reduce the uncertainty.

Other more simple methods to explore uncertainty in cost-effectiveness analysis include one-way and multi-way sensitivity analysis, extreme value analysis, and threshold analysis [16]. In these types of analysis, single mean values for each of the parameters are used to estimate outputs of expected costs and effects (known as deterministic sensitivity analysis). One-way sensitivity analysis involves varying one parameter of the model by a given amount (for example increase and decrease the mean value of the parameter by ±10%) to examine the impact of that parameter on changing the model results. This analysis can be repeated on different parameters to assess which parameters have the greatest influence on the cost-effectiveness results. While one-way sensitivity analysis is useful for demonstrating the impact of one parameter in the model, it may be necessary to examine the relationship of two or more different parameters changing simultaneously, which is known as multi-way sensitivity analysis. The interpretation of multi-way sensitivity analysis becomes increasingly difficult as the number of parameters involved increases. One method that is sometimes used to assess all the parameters simultaneously is extreme value sensitivity analysis. This involves varying all the parameters in the model to their 'best' and 'worse' case. Best and worse -case scenarios use extreme but plausibly favourable or unfavourable values for all the parameters in order to assess the implications on the cost-effectiveness results [17]. However, the probability of all parameters taking extreme but plausible values simultaneously will

be very small, which makes the interpretation of extreme value sensitivity analysis difficult. Threshold analysis may also be used to assess what value a parameter (or group of parameters) would need to take for the cost-effectiveness of an intervention to fall below a specified threshold value of cost-effectiveness.

Another source of uncertainty relates to the different scientific value judgments and assumptions that are made when constructing a model. These can be characterised using different scenario analyses, which represent alternative viewpoints that might be considered credible and plausible. In this case, there is uncertainty not only 'within' the set of model parameters used to estimate cost-effectiveness for the scenario but there is also uncertainty 'between' the alternative scenarios [17]. The uncertainty 'between' as well as 'within' scenarios can be captured by explicitly weighting the scenarios. The weighting can be done by assigning probabilities to the different scenarios to represent how credible each is believed to be [18]. This exercise can be completed using methods of expert elicitation [19].

3.5 Representing uncertainty in cost-effectiveness analysis

When there are only two alternative interventions being compared, uncertainty about incremental costs and health effects can be presented using the cost-effectiveness plane, as illustrated in Figure 3. The cost-effectiveness plane shows the joint uncertainty in the cost-effectiveness results [3]. Each of the simulated values from the PSA for costs and health effects for the interventions can be used to calculate the difference in costs (y-axis) and difference in effects (x-axis) between the interventions. The scatter plot shows the variation in results, where a large spread indicates a greater level of uncertainty. The cost-effectiveness threshold [20 21], which represents the amount in monetary terms that the decision-maker is willing to pay for an additional unit of health outcome, is shown by the solid line with slope, k = £30,000 per additional QALY. The proportion of simulated values (realisations in Table 1) which lie below the solid line provides an estimate of the probability that the intervention is considered cost-effective. The cost-effectiveness plane provides a useful visualisation of the scale of the uncertainty when there are two alternative interventions and a known cost-effectiveness threshold. However, it is difficult to visualise the

results when there are a number of alternative thresholds that might be considered and when there are more than two alternative interventions being compared. In these cases, a cost-effectiveness acceptability curve (CEAC) may provide a better representation [22 23]. CEACs are constructed by recording the number of times each alternative intervention has the highest net health effect (i.e. is considered to be cost-effective) from the simulated output of the model. The probability (% of times) that each alternative is cost-effective is plotted for a range of possible cost-effectiveness thresholds.

Figure 4 illustrates the CEACs associated with three alternative interventions. It is important to note, however, that the alternative intervention with the highest probability of being cost-effective may not always be the cost-effective alternative. This may seem counter intuitive but the intervention that is expected to be cost-effective depends on the scale of the net health effects. For example, treatment B might have the highest net health effect but the probability it is cost-effective is only 0.4, whereas treatment A might have lower net health effects but the probability it is cost-effective is 0.6. The reason is that when B is better than A it is 'much better' but when A is better than B (occurring more times) it is only a 'little bit better'.[17] Therefore, it is important to also indicate which of the alternative interventions is expected to be cost-effective as well as its probability.

The scale of the uncertainty can now be combined with information about the size of the patient population whose treatment choice can be informed by additional evidence and the time over which evidence about the interventions is expected to be useful to give the expected value of additional research. The expected costs of decision uncertainty can also be interpreted as the expected value of perfect information (EVPI) as perfect information would eliminate the possibility of making a wrong decision [14 24]. This valuation can be used as a necessary requirement for determining the potential efficiency of further research. Applying this decision rule, additional research should only be considered if the EVPI exceeds the expected cost of the research. The expected value of additional evidence will depend on the threshold of cost-effectiveness, which is illustrated in Figure 5. In addition to providing a global estimate of the total cost of uncertainty

related to all inputs in the model, EVPI can also be estimated for individual parameters (and groups of parameters) contained in the model. The objective of this analysis (termed partial EVPI or EVPPI) is to identify the model parameters for which it would be most worthwhile obtaining more precise estimates. Therefore, EVPPI provides a simple metric of the relative importance of different sources of uncertainty in contributing to the overall EVPI [14 25].

Figure 3: Cost-effectiveness plane

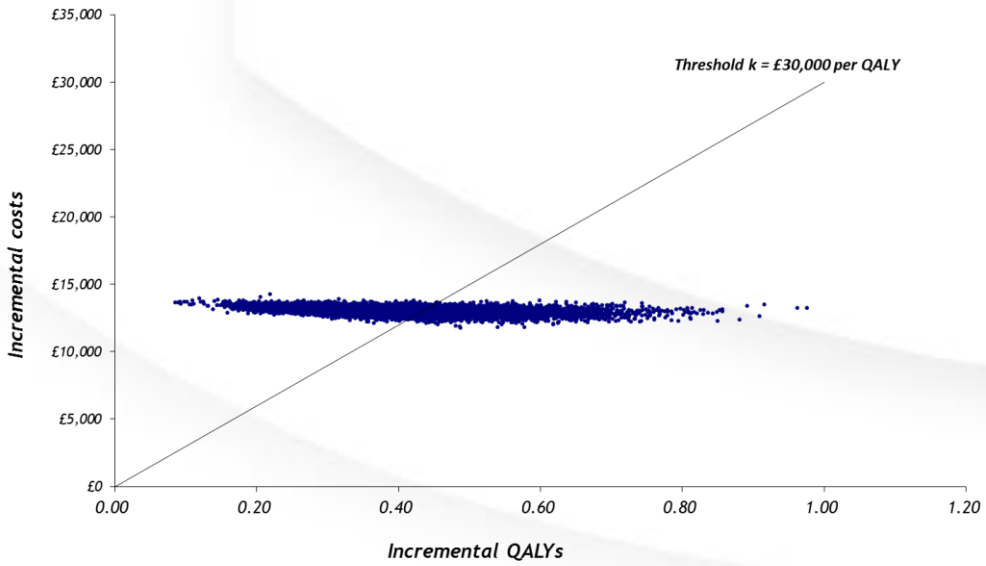


Figure 4: Cost-effectiveness acceptability curves for three alternative interventions

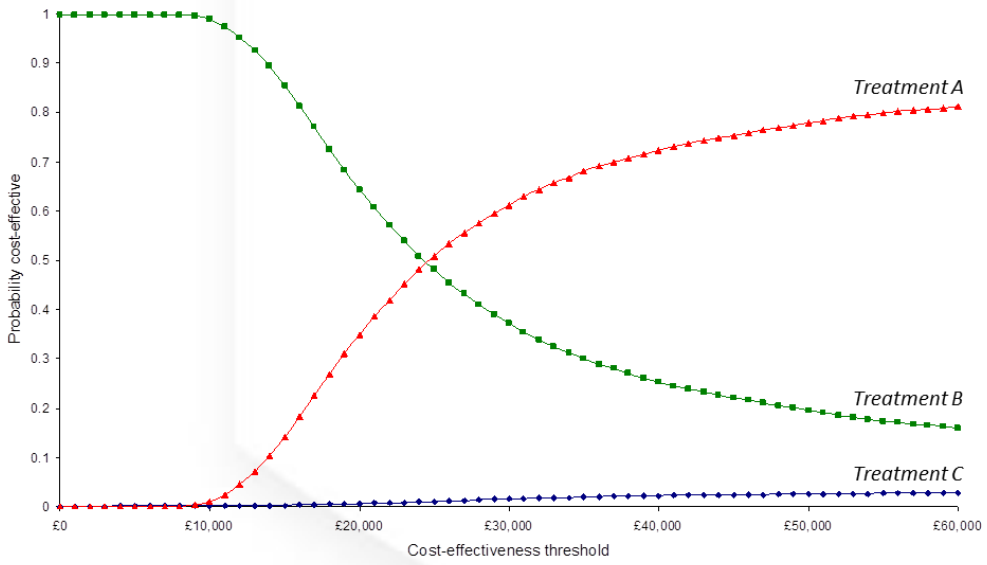
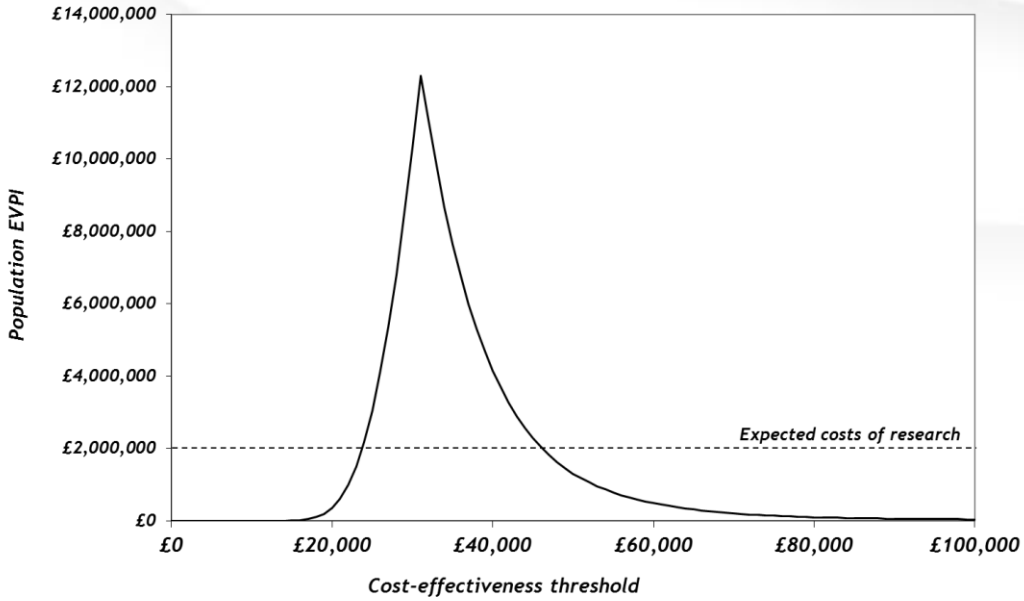


Figure 5: Population expected value of perfect information



3.6 Conclusions

Assessments regarding both the clinical effectiveness and cost-effectiveness of health care interventions are inevitably uncertain and, without sufficient and good quality evidence, decisions about the use of interventions will also be uncertain. There is a chance that the resources committed by the approval of an intervention may be wasted, or there may be harm to patients, if the expected positive net health effects are not realised in clinical practice. Similarly, in the face of uncertainty, rejecting an intervention which doesn't appear to be effective or cost-effectiveness might risk failing to provide access to a valuable intervention if the net health effects prove to be greater than expected. Therefore, if the objective of the health care system is to improve overall health for both current and future patients then the need for and the value of additional evidence is an important consideration when making decisions about the use of technologies.

This is even more critical once it is recognised that the approval of a technology for widespread use might reduce the prospects of conducting the type of research that would provide the evidence needed. In these circumstances there will be a trade-off between the net health effects for current patients from early access to a cost-effective technology and the health benefits for future patients from withholding approval until valuable research has been conducted. In addition, implementing a decision to approve an intervention may commit resources which cannot subsequently be recovered if a decision to approve or reimburse might change in the future (e.g. interventions with significant irrecoverable or sunk costs). Therefore, appropriate research and coverage decisions will depend on whether the expected benefits of research are likely to exceed the costs and whether any benefits of early approval or reimbursement are greater than withholding approval until additional research is conducted or other sources of uncertainty are resolved [18 25].

3.7 References

1. Claxton K, Griffin S, Koffijberg H, et al. Expected health benefits of additional evidence: Principles, methods and applications. Centre for Health Economics (CHE) research paper 2013;CHE RP 83
2. Uman L. Systematic reviews and meta-analyses. Journal of Canadian Academy of Child and Adolescent Psychiatry 2011;20(1):57-59
3. Drummond MF, Sculpher MJ, Torrance GW, et al. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press, 2005.
4. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Medical Decision Making 1998;18(suppl):S68-S80
5. Claxton K, Eggington S, Ginnelly L, et al. *A pilot study of value of information analysis to support research recommendations for the National Institute for Health and Clinical Excellence: Centre for Health Economics, Research paper 4*. York: Centre for Health Economics, University of York, 2005.
6. Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research. Some lessons from recent UK experience. Pharmacoeconomics 2006;24(11):1055-68
7. Alderson P, Roberts I. Corticosteroids in acute traumatic brain injury: systematic review of randomised controlled trials. British Medical Journal 1997;314:1855-59
8. CRASH trial collaborators. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury - outcomes at 6 months. Lancet 2005;365(9475):1957-9
9. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute For Clinical Excellence (NICE). Lancet 2002;360(9334):711-5
10. Bravo Vergel Y, Sculpher M. Quality-adjusted life years: How to understand them. Practical Neurology 2008;8(3):175-82
11. CRASH trial collaborators. Effect of intravenous corticosteroids on death within 14 days in 10008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial. The Lancet 2004;364:1321-28

12. Alderson P, Roberts I. *Corticosteroids for acute traumatic brain injury (Cochrane Review)*. Oxford: Update Software: The Cochrane Library, 1999.
13. Fenwick E, Claxton K, Sculpher M. The value of implementation and the value of information: combined and uneven development. Centre for Health Economics University of York 2005;CHE research paper 5.
14. Briggs A, Claxton K, Sculpher M. *Decision modelling for health economic evaluation*. Oxford: Oxford University Press, 2006.
15. Claxton K, Sculpher M, McCabe C, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Economics 2005;14(4):339-47
16. Briggs AH. Handling uncertainty in economic evaluation and presenting the results. In: Drummond M, McGuire A, eds. Economic Evaluation in Health Care. New York: Oxford University Press, 2001:172-214.
17. Claxton K. National Institute for Health and Clinical Excellence. Briefing paper for methods review. Workshop on exploring uncertainty. National Institute for Health and Care Excellence., 2007:1-63.
18. Claxton K, Palmer S, Longworth L, et al. Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. Health Technology Assessment 2012;16
19. O'Hagan A, Buck CE, Daneshkhah A, et al. *Uncertain judgements. Eliciting experts' probabilities.* Chichester: Wiley, 2006.
20. Karlsson G, Johannesson M. The decision rules of cost-effectiveness analysis. Pharmacoeconomics 1996;9(2):113-20
21. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. Pharmacoeconomics 2008;26(9):733-44
22. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. . Health Economics 2001;10:779-87
23. Fenwick E, O'Brien B, Briggs A. Cost-effectiveness acceptability curves - facts, fallacies and frequently asked questions. . Health Economics 2004;13:405-15

24. Claxton K, Ginnelly L, Sculpher MJ, et al. A pilot study on the use of decision theory and value of information analysis as part of the NHS Health Technology Assessment programme. Health Technology Assessment 2004;8(31):1-103
25. McKenna C, Claxton K. Addressing adoption and research design decisions simultaneously: the role of value of sample information analysis. Medical Decision Making 2011;31:853-65.

4. Decision rules in economic evaluation of healthcare interventions

Rita Faria
Paul Revill

4.1 Introduction

The goal of economic evaluation in healthcare is to inform decisions on whether interventions represent good value for money and should be offered (or reimbursed) by the healthcare system. An evaluation compares the costs and health outcomes of two or more alternative interventions. Health outcomes can be measured in terms of life years, quality-adjusted life years (QALYs) [1], disability-adjusted life years (DALYs) [2], physical units (e.g. blood pressure, survival) or in monetary terms by assigning a monetary value to health outcomes [3].

This document aims to provide an overview of the literature on decision rules in economic evaluation in healthcare. It starts with decision rules using the cost-effectiveness threshold with a brief mention of the net benefit framework. The use of a cost-effectiveness threshold will be exemplified with reference to its use in the international context and institutional guidelines, which are extrapolated to the Colombian context. It finishes with recommendations and implications for decision making in Colombia.

4.2 Allocation decisions across the healthcare system

Standard decision rules in economic evaluation aim to maximise health outcomes given the available budget to the healthcare system. This can be viewed as a mathematical programming problem in which the costs and health outcomes of all the possible interventions are compared in order to obtain the optimal bundle of interventions [4, 5]. Mathematical programming requires information on all costs and health outcomes of every competing intervention for every disease area. Therefore, the information requirements for a mathematical programming solution generally make it unfeasible to implement in practice.

Simplified decision rules have been proposed to overcome this issue. Birch and Gafni have argued that an intervention is cost-effective if its health outcomes are greater than the health outcomes obtained by the interventions displaced to accommodate its additional costs [6, 7]. The difficulty is identifying the interventions that will be displaced at the time of the decision. Most frequently, however, the ratio of the difference in costs by the difference in health outcomes of an intervention versus a comparator, termed the incremental cost-effectiveness ratio (ICER), is compared against a cost-effectiveness threshold [8, 9]. This is the method used in the UK by the National Institute for Health and Care Excellence (NICE) [10].

4.3 Decision rules using incremental cost-effectiveness ratios and a cost-effectiveness threshold

The ICER represents the additional cost per unit of health outcome gained from the new intervention compared with its relevant comparator. The ICER by itself does not allow for an assessment of the value of the new intervention. The ICER of the intervention versus its relevant comparator needs to be compared to the cost-effectiveness threshold. The cost-effectiveness threshold represents the opportunity costs of devoting resources to the intervention, in terms of health gains displaced or forgone as a result of those resources being unavailable to fund other alternative competing priorities [8, 9, 11]. Unless the budget is increased to account for the additional costs of the new intervention, the resources required to deliver the new intervention are found by disinvesting in other interventions elsewhere in the healthcare system. As a result, other patients, who benefited from the displaced interventions, or could have benefited from other interventions that remain unfunded, will lose out from the approval of a new intervention. The threshold represents the productivity of the healthcare system. More productive healthcare systems in 'transforming' expenditure into health will have lower cost-effectiveness thresholds as less expenditure produces one unit of health outcome.

An alternative interpretation of the cost-effectiveness threshold that is sometimes presented is that it represents the rate at which individuals are willing to forgo other forms of consumption to improve health (i.e. it represents their "willingness to pay" for health). Under this interpretation, the cost-effectiveness threshold represents the consumption value of health [12-23]. In any case,

a new intervention should be offered by the healthcare system if its health outcomes exceed the health loss from diverting funds from existing interventions; in other words, if its health outcomes exceed its opportunity cost.

The ICER of a new intervention depends on the costs and health outcomes obtained compared to a relevant comparator. In practice, the ICER of new intervention is calculated in a full incremental analysis, which may include more than one comparator. This involves ranking all the possible interventions in ascending order of costs. Interventions that are more costly and produce lower health outcomes are said to be dominated and removed from the comparison. Interventions with an ICER greater than the ICER of a more costly more effective alternative are said to be extendedly dominated and also removed from the comparison. The extendedly dominated intervention provides the same health outcomes at greater additional costs than its comparator. Ultimately, the analysis will include the least costly intervention and a set of non-dominated more costly interventions that deliver better health outcomes. The cost-effective intervention is that which delivers greatest health outcomes but with an ICER below its opportunity cost (represented by 'the threshold').

The ICER approach makes two assumptions on the nature of the decision [24]. First, that interventions are independent. This means that the costs and health outcomes of each intervention are delivered independently of the whether other interventions are included in the comparison. This assumption can be relaxed by creating composite interventions composed of combinations of the individual interventions. Second, that interventions are perfectly divisible and have constant returns to scale. In other words, that interventions can be implemented fractionally for a proportion of the costs and health outcomes. Conversely, as the intervention is expanded to a bigger population, that the costs and health outcomes increase in the same proportion. A thorough discussion on the topic can be found in references [6, 25-28].

This ICER decision rule can be reformulated in terms of the net benefit of the intervention. The net benefit is the value of the intervention over and above its additional costs (including opportunity costs, as represented by the threshold). This can be expressed in monetary terms as net monetary benefit $(\mathrm{NMB} = \Delta\mathrm{E} \cdot \lambda - \Delta\mathrm{C})$ or equivalently in health terms as the net health

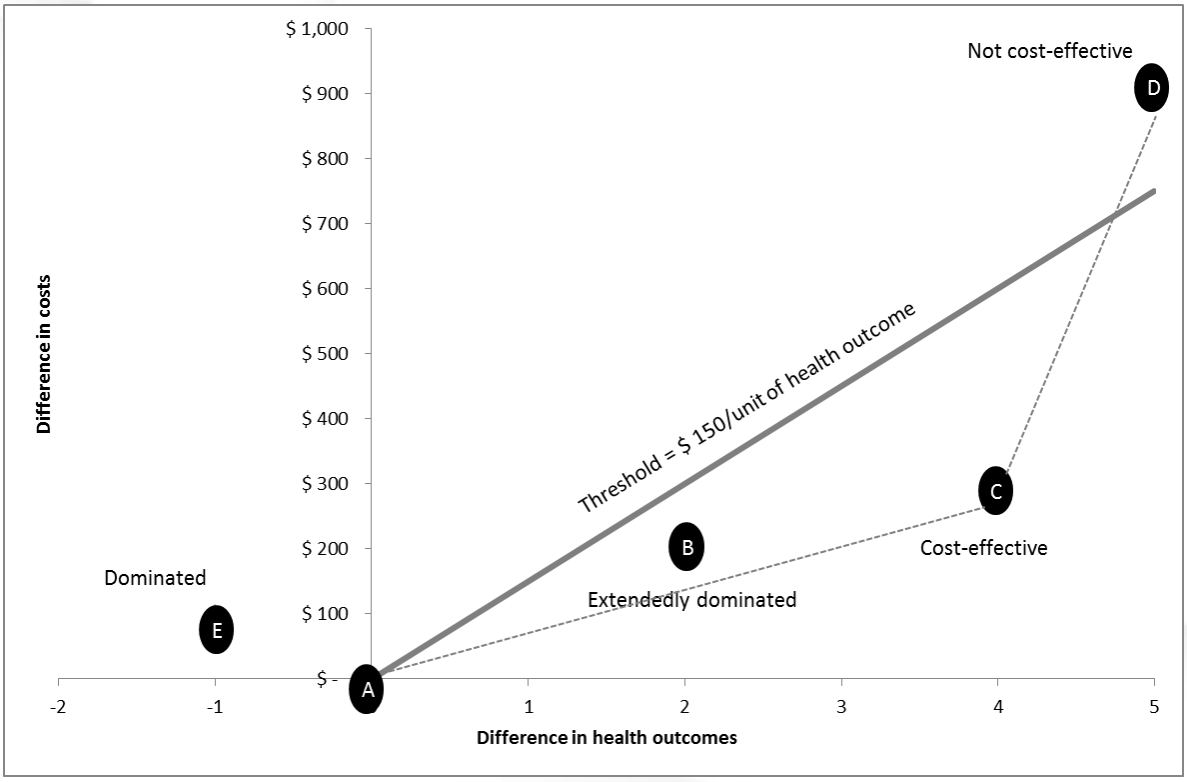
benefit ($\mathbf{NHB} = \Delta\mathbf{E} - \Delta\mathbf{C}/\boldsymbol{\lambda}$). Therefore, the cost-effective intervention is the intervention with the greatest net benefit. The net benefit approach has a number of statistical advantages over the ICER [29-31] in addition to facilitating the analysis of uncertainty [32-34] (see Technical Document entitled 'Uncertainty and the value of additional evidence in healthcare decisions').

4.4 The cost-effectiveness plane

The decision problem can be visualised in the cost-effectiveness plane, shown in Figure 1 [35]. Here the costs and health outcomes of each intervention are plotted in relation to costs and health outcomes of a less costly alternative, represented in the origin of the axis. The alternative interventions may be more costly and produce better health outcomes or be more costly and produce worse health outcomes. The decision is clear in the latter scenario as more costly interventions that produce worse health outcomes are dominated and not cost-effective (e.g. intervention E). Often it is the case that the interventions being evaluated offer better health outcomes at a greater cost compared with usual care. In this case, the cost-effectiveness of interventions depends on how their costs and health outcomes compare with each other and with the cost-effectiveness threshold, as discussed above. Intervention B is extendedly dominated; although its ICER is below the cost-effectiveness threshold at $100, it is greater than the ICER of intervention C, which produces more health. The range of interventions remaining in the analysis defines a cost-effectiveness frontier (dotted line). Any intervention lying to the northwest of the frontier is not cost-effective. The final decision of whether an intervention is cost-effective requires the cost-effectiveness threshold, represented here by grey line crossing the origin. Interventions lying to the southeast of the cost-effectiveness threshold are potentially cost-effective. The cost-effective intervention is the one which offers the most health outcomes but still lies to the south-east of the cost-effectiveness threshold line (intervention C). In this example, accepting intervention C will result in a gain of 4 units of health at an additional cost of $300 (ICER=$75 per unit of health gained). At a threshold of $150 per unit of health outcome, this means that intervention C produces 4 units of health at an additional cost of 2 units ($300 converted into health outcomes; $300/$150=2). Therefore, the net gains to the healthcare system are 2 units of health (4 units produced – 2 units displaced). Intervention D is not cost-effective. It produces 5

units of health at a cost of $900 (ICER=$180), which is equivalent to 6 units of health displaced elsewhere in the healthcare system; intervention D represents a net health loss of 1 unit of health.

Figure 4 Cost-effectiveness plane



4.5 How is the cost-effectiveness threshold expected to vary with time or for particular interventions?

The cost-effectiveness threshold tends to increase with increases in the healthcare budget and healthcare costs [36]. Greater healthcare budgets generate more health, albeit this tends to be at a diminishing rate [37, 38]. This means that increases in a small healthcare budget will have a large impact on health whereas the same increase in a large healthcare budget will have a smaller impact. Increases in the healthcare budget that are spent in areas other than healthcare interventions (e.g. increase in wages) do not increase the threshold. Conversely, decreases in the healthcare budget result in a lower threshold. Furthermore, the threshold tends to fall with increases in productivity and efficiency. The more the healthcare system can generate health from existing resources, the lower the cost-effectiveness threshold.

The cost-effectiveness threshold depends on the budget impact of the new intervention [36]. Interventions with greater budget impacts displace disproportionally more health; therefore, the threshold should be lower. This is because the threshold represents the opportunity cost of an intervention that has a marginal impact on costs. Non-marginal changes in expenditure mean that more and more valuable healthcare interventions are displaced to fund the new intervention. As a result, the opportunity cost of the offering the new intervention is greater and the threshold falls [39].

4.6 Cost-effectiveness thresholds in low- and middle-income countries

The World Health Organisation (WHO) recommended in 1996 a range of cost-effectiveness thresholds per DALY averted of $US150 as 'attractive' cost-effectiveness and $US25 as 'very attractive' cost-effectiveness for low-income countries and $US500 and $US100, respectively, for middle income countries [40]. Currently, the WHO recommends a cost-effectiveness threshold per DALY averted dependent on the gross domestic product (GDP) of each country. Technologies with ICERs below the GDP per capita are considered highly cost-effective; technologies with ICERs between one and three times the GDP per capita as cost-effective; and technologies with ICERs above three times GDP per capita as not cost-effective [41].

The WHO decision rules have been criticised for promoting unaffordable cost-effectiveness thresholds, exacerbating health inequalities and ultimately lowering population health [42]. This is because the GDP per capita is likely to be gross underestimate of the productivity of healthcare systems in low- and middle-income countries. Therefore, a threshold based on the GDP per capita will underestimate the opportunity costs associated with health care resources. Consequently, offering the new 'falsely cost-effective' intervention risks improving health outcomes of a selected few and resulting in an overall net health loss in the population as a whole.

There are a number of reasons to believe that the threshold in low- and middle-income countries is likely to be lower than the WHO's rule suggests. All things equal, the threshold should be lower in countries with smaller healthcare budgets, smaller GDP per capita, smaller share of GDP

devoted to healthcare and where unmet health needs are greatest. Smaller healthcare budgets can afford less expensive interventions, which in turn will imply a smaller threshold due to remaining unmet needs. Smaller GDP per capita and smaller share of GDP devoted to health result in smaller healthcare budgets. Countries with the same GDP per capita may have different healthcare budgets and different thresholds. Therefore, a rule that considers only GDP per capita is very unlikely to be appropriate for most countries.

4.7 The use of the cost-effectiveness threshold in high income countries

Table 1 presents the cost-effectiveness thresholds used in different countries and compares them with their GDP per capita and the share of the GDP devoted to public expenditure in healthcare. To our knowledge, only the UK uses an explicit threshold range, between £20,000 and £30,000 per QALY gained, in making decisions for the public healthcare system [10]. This threshold range is up to 1.2 times the GDP per capita of the UK. For other high-income countries, which do not use an explicit threshold, the threshold is up to 2.6 times the GDP per capita (e.g. Canada). Recent empirical research on the UK threshold estimated a lower value at $US19,600 (£12,936) per QALY gained [36]. This estimate indicates that the UK empirical threshold is 52% of its GDP per capita. The empirical threshold estimate of $US19,600 per QALY gained represents the actual productivity of the UK National Health Service (NHS). In other words, a new intervention with an ICER below $US19,600 per patient represents a net health gain to the UK NHS; whereas ICERs above $US19,600 imply that the new intervention displaces more health than what it produces. An empirical estimate of the threshold ensures that the health gains of new interventions are appropriately compared with the health displaced from their additional costs.

Table 4 Comparison of cost-effectiveness threshold with gross domestic product per capita

Country	Threshold per QALY gained	Threshold per QALY gained in PPP [43]	GDP per capita in PPP (2012) [44]	Ratio Threshold : GDP	Public health expenditure as share of GDP[45]
Explicit threshold					

UK	£20,000-£30,000[46]	$30,303-$45,455	$ 37,500	0.8-1.2	7.7%
Implicit threshold inferred from past allocation decisions					
Australia	AU$69,900[47]	$47,877	$ 43,300	1.1	6.2%
Canada	CAN$31,000-CAN$137,000[48]	$25,203-$111,382	$ 43,400	0.6-2.6	7.9%
New Zealand	NZ$20,000[49]	$13,793	$ 30,200	0.8	8.4%
Threshold proposed by individuals or institutions					
Canada	CAN$20,000-CAN$100,000[50]	$16,260-$81,301	$ 43,400	0.4-1.9	7.9%
Netherlands	€80,000[51]	$96,383	$ 42,900	2.2	10.2%
USA	$50,000[52]	$50,000	$ 50,700	1.0	8.2%
UK	£12,936[36]	$19,600	$ 37,500	0.5	7.7%
QALY – quality-adjusted life year; PPP – purchasing parity power; GDP – gross domestic product.					

4.8 Implications for decision-making in Colombia

The thresholds used in high-income countries, particularly the empirical estimate of the UK threshold, can be used to provide an indication of the upper-bound of the Colombian cost-effectiveness threshold. This upper-bound is likely to be an overestimate, given that high-income countries have greater healthcare budgets per capita, hence disproportionally greater cost-effectiveness thresholds as a result of diminishing marginal returns to healthcare expenditure. This is because as expenditure rises, returns in terms of health decrease at the margin and the threshold increases; conversely, in a low expenditure setting, the returns from expenditure are proportionally greater and the threshold should be proportionally lower.

The GDP per capita in Colombia is US$11,000 and the share of GDP devoted to public healthcare is 4.6% [44 45]. This compares with a GDP per capita in the UK of $37,500 and 7.7% of GDP devoted to public healthcare [44, 45]. In the UK, the threshold has been estimated at $19,600 (52% of GDP per capita) [36]. Assuming that the Colombian threshold is proportional to the UK threshold, given its GDP per capita and share of GDP devoted to public healthcare expenditure, the derived upper bound for the Colombian threshold is $3,434 (11,000*19,600*4.6/(37,500*7.7)).

The relationship between healthcare expenditure and appropriate thresholds across countries depends on the efficiency of a healthcare system. Healthcare systems that are relatively less efficient will have higher thresholds since more resources are required to produce the same health gain (lower productivity). However, this does not necessarily imply that more expensive interventions should be reimbursed. Instead that the healthcare system can improve population health by disinvesting in inefficient interventions and ensuring resources are committed to where they produce the most health [53].

The true cost-effectiveness threshold for Colombia is likely to be lower than a rule based on GDP per capita and share of GDP devoted to public healthcare would suggest. The rule employed here assumes that the threshold is directly proportional to public expenditure in healthcare; i.e. that the relationship between the health and expenditure increases in a constant manner. However, there is some evidence to suggest that public expenditure in healthcare in poorer countries with small expenditures results in a greater increase in health than the same increase in wealthier countries (which have large expenditures) [54]. This strengthens the argument that there are diminishing marginal returns to expenditure in health. Therefore, that the threshold is likely to disproportionally increase for greater expenditures and disproportionally decrease in the lower end of the scale. The practical implication is that even interventions with ICERs lower than this value may not represent good value for money. Using a cost-effectiveness threshold greater than the 'real' threshold (i.e. the productivity of the healthcare system in achieving health) would result in reimbursing interventions that are not cost-effective and ultimately a net loss of health for the country [42]. Although using an underestimate of the threshold can also result in losses of health, the health loss from overestimating the threshold is greater than the corresponding loss from an underestimate [36]. Therefore, it is prudent to err on the side of caution and use a lower rather than a higher value for the cost-effectiveness threshold.

4.9 Concluding remarks

Colombia is a middle-income country with an expanding economy and an increasing population with greater demands to the public healthcare system. The Colombian public healthcare system, as most healthcare systems in the world, faces increasing budgetary pressures and difficult

decisions on how to allocate funds across disease areas and technologies. Economic evaluation offers an explicit and transparent framework to help make such difficult decisions and maximise the health gains given the available budget. However, the utility of this framework is determined by the whether Colombia is able to define a set of criteria, specific to their own context and needs, to establish the value for money of technologies. These set of criteria may include other factors in addition to the health and costs of the new technology, such as equity issues, disease burden, prevalence of the disease, availability of other treatments, etc. Ultimately, the set of criteria adopted in Colombia should consider not only the patients who benefit from the new technology but also those who may lose out from its additional costs.

4.10 References

1. Torrance, G.W. and D. Feeny, *Utilities and quality-adjusted life years.* Int J Technol Assess Health Care, 1989. **5**(4): p. 559-575.
2. Rushby, J.F. and K. Hanson, *Calculating and presenting disability adjusted life years (DALYs) in cost-effectiveness analysis.* Health policy and planning, 2001. **16**(3): p. 326-331.
3. Drummond, M.F., M.J. Sculpher, and G.W. Torrance, *Methods for the economic evaluation of health care programmes*. 2005: Oxford University Press, USA.
4. Stinnett, A.A. and A.D. Paltiel, *Mathematical programming for the efficient allocation of health care resources.* Journal of Health Economics, 1996. **15**(5): p. 641-653.
5. McKenna, C. and K. Claxton, *PRM171 Optimal Shopping: An Evauation of Decision Rules in Cost-Effectiveness Analysis (CEA).* Value in Health, 2012. **15**(7): p. A491-A492.
6. Gafni, A. and S. Birch, *Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem.* CMAJ: Canadian Medical Association Journal, 1993. **148**(6): p. 913.
7. Gafni, A. and S. Birch, *Incremental cost-effectiveness ratios (ICERs): The silence of the lambda.* Social Science & Medicine, 2006. **62**(9): p. 2091-2100.
8. Karlsson, G. and M. Johannesson, *The decision rules of cost-effectiveness analysis.* Pharmacoeconomics, 1996. **9**(2): p. 113-120.
9. Weinstein, M. and R. Zeckhauser, *Critical ratios and efficient allocation.* Journal of Public Economics, 1973. **2**(2): p. 147-157.
10. Rawlins, M., D. Barnett, and A. Stevens, *Pharmacoeconomics: NICE's approach to decision-making.* British journal of clinical pharmacology, 2010. **70**(3): p. 346-349.
11. McCabe, C., K. Claxton, and A.J. Culyer, *The NICE cost-effectiveness threshold.* Pharmacoeconomics, 2008. **26**(9): p. 733-744.
12. Bobinac, A., et al., *Willingness to pay for a quality-adjusted life-year: the individual perspective.* Value Health, 2010. **13**(8): p. 1046-55.
13. Dolan, P., et al., *QALY maximisation and people's preferences: a methodological review of the literature.* Health Econ, 2005. **14**(2): p. 197-208.
14. Green, C. and K. Gerard, *Exploring the social value of health-care interventions: a stated preference discrete choice experiment.* Health Econ, 2009. **18**(8): p. 951-76.
15. Bobinac, A., et al., *GET MORE, PAY MORE? An elaborate test of construct validity of willingness to pay per QALY estimates obtained through contingent valuation.* J Health Econ, 2012. **31**(1): p. 158-68.
16. Lieu, T.A., et al., *Willingness to pay for a QALY based on community member and patient preferences for temporary health states associated with herpes zoster.* Pharmacoeconomics, 2009. **27**(12): p. 1005-16.
17. Shiroiwa, T., et al., *International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness?* Health Econ, 2010. **19**(4): p. 422-37.
18. Mason, H., R. Baker, and C. Donaldson, *Willingness to pay for a QALY: past, present and future*. Expert Rev Pharmacoecon Outcomes Res, 2008. **8**(6): p. 575-82.

19. Johnson, F.R., *Einstein on willingness to pay per QALY: is there a better way?* Med Decis Making, 2005. **25**(6): p. 607-8.
20. Gyrd-Hansen, D., *Willingness to pay for a QALY: theoretical and methodological issues.* Pharmacoeconomics, 2005. **23**(5): p. 423-32.
21. Richardson, J. and R.D. Smith, *Calculating society's willingness to pay for a QALY: key questions for discussion.* Appl Health Econ Health Policy, 2004. **3**(3): p. 125-6.
22. Gyrd-Hansen, D., *Willingness to pay for a QALY.* Health Econ, 2003. **12**(12): p. 1049-60.
23. Steurer, J., *[QALY or willingness to pay?].* Praxis (Bern 1994), 1999. **88**(8): p. 333-8.
24. Culyer, A., et al., *Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence.* Journal of health services research & policy, 2007. **12**(1): p. 56-58.
25. Buxton, M.J., *How much are health-care systems prepared to pay to produce a QALY?* The European Journal of Health Economics, 2005. **6**(4): p. 285-287.
26. Eichler, H.-G., et al., *Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge?* Value in health, 2004. **7**(5): p. 518-528.
27. Garber, A.M. and C.E. Phelps, *Economic foundations of cost-effectiveness analysis.* Journal of health economics, 1997. **16**(1): p. 1-31.
28. Elshaug, A.G., et al., *Identifying existing health care services that do not provide value for money.* Med J Aust, 2009. **190**(5): p. 269-73.
29. Briggs, A.H., C.Z. Mooney, and D.E. Wonderling, *Constructing confidence intervals for cost-effectiveness ratios: an evaluation of parametric and non-parametric techniques using Monte Carlo simulation.* Statistics in medicine, 1999. **18**(23): p. 3245-3262.
30. Polsky, D., et al., *Confidence intervals for cost-effectiveness ratios: a comparison of four methods.* Health economics, 1997. **6**(3): p. 243-252.
31. Briggs, A.H., D.E. Wonderling, and C.Z. Mooney, *Pulling cost-effectiveness analysis up by its bootstraps: A non-parametric approach to confidence interval estimation.* Health economics, 1997. **6**(4): p. 327-340.
32. Fenwick, E., K. Claxton, and M. Sculpher, *Representing uncertainty: the role of cost-effectiveness acceptability curves.* Health economics, 2001. **10**(8): p. 779-787.
33. Fenwick, E., B.J. O'Brien, and A. Briggs, *Cost-effectiveness acceptability curves–facts, fallacies and frequently asked questions.* Health economics, 2004. **13**(5): p. 405-415.
34. Van Hout, B.A., et al., *Costs, effects and c/e-ratios alongside a clinical trial.* Health economics, 1994. **3**(5): p. 309-319.
35. Black, W.C., *The CE Plane A Graphic Representation of Cost-Effectiveness.* Medical Decision Making, 1990. **10**(3): p. 212-214.
36. Claxton, K., et al., *Methods for the estimation of the NICE cost effectiveness threshold.* CHE Research Paper, 2013. **81**.
37. UC Atlas of Global Inequality. *Health Care Spending*. n.d. 28/11/2013]; Available from: http://ucatlas.ucsc.edu/spend.php.
38. Bryant, J.H. and P.F. Harrison, *Global Health in Transition: A Synthesis:Perspectives from International Organizations*. 1996: The National Academies Press.
39. Claxton, K., et al., *Appropriate perspectives for health care decisions*, in *CHE Research Paper*. 2010, Centre for Health Economics: York.

40. World Health Organisation, *Investing in health research and development. Report of the Ad Hoc committee on health research relating to future intervention options.* . 1996, World Health Organisation.: Geneva.
41. Commission on Macroeconomics and Health, *Macroeconomics and health: investing in health for economic development. Report of the Commission on Macroeconomics and Health.* 2001, Geneva: World Health Organisation.
42. Revill, P. and M. Sculpher, *Cost effectiveness of interventions to tackle non-communicable diseases.* BMJ 2012;344: d7883.
43. The World Bank. *PPP conversion factor, GDP (LCU per international $)*. 2013 27/11/2013]; Available from: http://data.worldbank.org/indicator/PA.NUS.PPP.
44. Central Intelligence Agency. *The World Factbook - GDP per capita in PPP*. 2013 2/11/2013]; Available from: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2004rank.html.
45. the World Bank. *Health expenditure, public (% of GDP)*. 2013 28/11/2013]; Available from: http://data.worldbank.org/indicator/SH.XPD.PUBL.ZS.
46. National Institute for Health and Care Excellence, *Guide to the methods of technology appraisal 2013*. 2013, National Institute for Health and Clinical Excellence: London.
47. Henry, D.A., S.R. Hill, and A. Harris, *Drug prices and value for money.* JAMA: the journal of the American Medical Association, 2005. **294**(20): p. 2630-2632.
48. Rocchi, A., et al., *The role of economic evidence in Canadian oncology reimbursement decision-making: to lambda and beyond*. Value in Health, 2008. **11**(4): p. 771-783.
49. Towse, A., C. Pritchard, and N.J. Devlin, *Cost-effectiveness thresholds: economic and ethical issues*. 2002: King's Fund.
50. Laupacis, A., et al., *How attractive does a new technology have to be to warrant adoption and utilization? tentative guidelines for using clinical and economic evaluations.* CMAJ: Canadian Medical Association Journal, 1992. **146**(4): p. 473.
51. Raad voor de Volksgezondheid en Zorg, *Zinnige en duurzame zorg.* 2006.
52. Weinstein, M.C., *How much are Americans willing to pay for a quality-adjusted life year?* Medical care, 2008. **46**(4): p. 343-345.
53. Paulden, M., McCabe, C., & Karnon, J. *Achieving Allocative Efficiency in Healthcare: Nice in Theory, not so NICE in Practice?*. PharmacoEconomics, 2014: 1-4 10.1007/s40273-014-0146-x.
54. Moreno-Serra, R., & Smith, P. C. *Broader health coverage is good for the nation's health: evidence from country level panel data*. Journal of the Royal Statistical Society: Series A (Statistics in Society), 2014. DOI: 10.1111/rssa.12048

Instituto de Evaluación
Tecnológica en Salud

Autopista Norte #118-30, oficina 201
Bogotá D.C.

contacto@iets.org.co

www.iets.org.co

ietscolombia

ietscolombia@blogspot.com

@ietscolombia